

**A NETWORK-BASED FRAMEWORK TO DISCOVER TREATMENT-RESPONSE-
PREDICTING BIOMARKERS FOR COMPLEX DISEASES****CROSS-REFERENCE TO RELATED APPLICATIONS**

[0001] This application claims the benefit of U.S. Provisional Application No. 63/518,783, filed August 10, 2023, and U.S. Provisional Application No. 63/639,488, filed April 26, 2024, each of which is incorporated by reference herein in its entirety.

BACKGROUND

[0002] Precision medicine's potential to transform complex inflammatory and immunology ("I&I") disease treatment can be challenged by, e.g., limited data and inadequate sample size compared to the number of molecular features found in high-throughput, multi-omics datasets. Addressing at least this issue, described herein is a novel framework, a Predictive Response Biomarkers using Network medicine (PRoBeNet) framework, which can operate with limited data and inadequate sample size. The PRoBeNet framework herein recognizes that the therapeutic effect of a drug directed to I&I indications can propagate through a protein-protein interaction network to reverse disease states. The PRoBeNet framework herein can prioritize biomarkers by considering at least (1) therapy-targeted proteins, (2) disease-specific molecular signatures, and (3) an underlying network of interactions among cellular components (e.g., a human interactome). Using the PRoBeNet framework herein, biomarkers were discovered, which can predict patient responses to an established autoimmune therapy (e.g., infliximab) and an investigational compound (e.g., an ERK1/2 inhibitor). The predictive utility of PRoBeNet biomarkers herein was validated with retrospective gene-expression data from ulcerative colitis (UC) and rheumatoid arthritis (RA) patients and prospective data from UC and Crohn's disease (CD) patient-derived tissues.

[0003] Machine-learning models using PRoBeNet biomarkers herein significantly outperformed other models using either all genes or randomly selected genes, especially when data was limited (e.g., fewer than 20 samples). These results illustrate the utility of the PRoBeNet framework herein for reducing features and for constructing robust machine-learning models when limited data is available. The PRoBeNet framework herein can be used to develop companion and complementary diagnostic assays for complex I&I disease therapies. Such use can help to stratify suitable patient subgroups in clinical trials, approve new drugs, and improve patient outcomes.

SUMMARY

[0004] Additional aspects and advantages of the present disclosure will become readily apparent from the following detailed description, wherein only illustrative embodiments of the present

disclosure are shown and described. The present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and descriptions are to be regarded as illustrative in nature and not as restrictive.

[0005] A novel network-based system biology framework, a Predictive Response Biomarkers using Network medicine (PRoBeNet) framework, is described herein to predict treatment response biomarkers in complex I&I diseases without relying on, e.g., extensive patient datasets. The PRoBeNet framework herein can successfully determine and validate response-predicting biomarkers retrospectively for, e.g., infliximab response and prospectively for, e.g., **extracellular signal-regulated kinase (ERK1/2) inhibitors**. Nodes in a human interactome (HI) can be ranked by the PRoBeNet framework using a dual PageRank score, which considers proteins important to the therapy's mechanism of action and disease pathogenesis. The PRoBeNet framework's ability to determine response biomarkers can be a useful and valuable tool for, e.g., the development of companion diagnostic tests, amplifying drug efficacies, and optimizing personalized treatment strategies.

[0006] In an aspect, disclosed herein is a method of treating a subject suffering from a disease or disorder, the method comprising: administering to the subject a therapy, wherein the therapy has been predicted to generate a treatment response in the subject based at least in part on a trained machine learning (ML) classifier that classifies the subject as responsive or non-responsive to the therapy, wherein the trained ML classifier predicts the treatment response based at least in part on analyzing in the subject a set of cytokines or a set of transcripts. In some embodiments, the method further comprising: identifying one or more source nodes associated with a set of proteins targeted by the therapy; identifying one or more response nodes associated with the set of cytokines or the set of transcripts indicative of the treatment response; determining, by a network-based diffusion process, one or more intermediate nodes associated with the one or more source nodes or the one or more response nodes, wherein the one or more intermediate nodes are indicative of the treatment response; generating a ranked set of the one or more intermediate nodes, wherein the ranked set is based at least in part on one or more interactions of the intermediate nodes with the one or more source nodes or the one or more response nodes; and processing a subset of the ranked set of intermediate nodes to determine a set of biomarkers that are predictive of the treatment response to the therapy. In some embodiments, the ranked set of intermediate nodes comprises a ranked set of proteins, a ranked set of cytokines, or a ranked set of transcripts. In some embodiments, the ranked set of proteins, the ranked set of cytokines, or the ranked set of transcripts is used as features to obtain the trained ML classifier. In some embodiments, the set of biomarkers is predictive of the

treatment response to the therapy. In some embodiments, (i) the ranked set of cytokines is different from at least one cytokine of the set of cytokines or (ii) the ranked set of cytokines is the same as the set of cytokines. In some embodiments, (i) the ranked set of transcripts is different from at least one transcript of the set of transcripts or (ii) the ranked set of transcripts is the same as the set of transcripts. In some embodiments, the ranked set of cytokines or the ranked set of transcripts is analyzed to determine statistically significant changes after administering to the subject the therapy. In some embodiments, generating the ranked set of the one or more intermediate nodes comprises: determining a first rank between the one or more source nodes and the one or more intermediate nodes; determining a second rank between the one or more response nodes and the one or more intermediate nodes; and determining a rank product of the first rank and the second rank. In some embodiments, the method further comprising: constructing a subset of a protein-protein network associated with the disease or disorder, wherein the constructing comprises: (i) determining the one or more source nodes in the protein-protein network, based at least in part on the set of proteins targeted by the therapy; (ii) determining the one or more response nodes in the protein-protein network, based on least in part on the set of cytokines or the set of transcripts indicative of the treatment response; and (iii) determining one or more connections between the one or more source nodes and the one or more response nodes, based at least in part on one or more interactions between the one or more source nodes and the one or more response nodes. In some embodiments, the protein-protein network comprises a human interactome (HI). In some embodiments, the disease or disorder comprises an autoimmune disease, an immunology disease, or a cancer. In some embodiments, the immunology disease or the autoimmune disease comprises **rheumatoid arthritis (RA), ulcerative colitis (UC), Crohn's disease (CD), inflammatory bowel disease, psoriatic arthritis, ankylosing spondylitis, chronic psoriasis, hidradenitis suppurativa, multiple sclerosis, juvenile idiopathic arthritis, uveitis, systemic lupus, Type 1 diabetes, Graves' Disease, Hashimoto's Thyroiditis, celiac disease, Sjögren's Syndrome, or an immune-mediated disease.** In some embodiments, the immunology disease or the autoimmune disease comprises RA. In some embodiments, the immunology disease or the autoimmune disease comprises UC. In some embodiments, the immunology disease or the autoimmune disease comprises CD. In some embodiments, the therapy for the immunology disease or the autoimmune disease comprises an anti-TNF therapy. In some embodiments, the anti-TNF therapy comprises infliximab, adalimumab, etanercept, certolizumab pegol, or golimumab, or a biosimilar thereof. In some embodiments, the therapy for the immunology disease or the autoimmune disease comprises an alternative to anti-TNF therapy. In some embodiments, the alternative to anti-TNF therapy comprises anti-CD20, JAK, anti-IL6, rituximab, sarilumab, tofacitinib citrate, leflunomide, vedolizumab, tocilizumab,

anakinra, abatacept, TL1A, co-stimulation blockade, or a biosimilar thereof. In some embodiments, the cancer comprises bladder cancer, breast cancer, colon cancer, rectal cancer, endometrial cancer, kidney cancer, leukemia, liver cancer, lung cancer, melanoma, Non-Hodgkin's lymphoma, pancreatic cancer, prostate cancer, or thyroid cancer. In some embodiments, the therapy for the cancer comprises a targeted therapy, chemotherapy, hormone therapy, immunotherapy, photodynamic therapy, radiation therapy, stem cell transplant therapy, or any combination thereof. In some embodiments, the trained ML classifier predicts the treatment response to the therapy using a non-linear relationship between (i) measurements of the ranked set of intermediate nodes and (ii) the treatment response indicated by statistically significant changes in the ranked set of intermediate nodes after administering to the subject the therapy. In some embodiments, the trained ML classifier is trained using measurements of the one or more intermediate nodes in (i) a first set of subjects with the disease or disorder determined to be responsive to the therapy and (ii) a second set of subjects with the disease or disorder determined to be non-responsive to the therapy. In some embodiments, the trained ML classifier comprises a neural network, deep learning, perceptron, random forest, Bayes, Markov, Gaussian process, clustering algorithm, support vector machine, generative model, or kernel. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to an immunology disease or an autoimmune disease comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL-12A/B, IL12B, IL13, IL15, IL16, IL17A, IL18, IL1B, IL2, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, IL8, MIF, TGFB1, TGFB2, TGFB3, or TNF. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL13, IL17A, IL1B, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, TGFB1, or TNF. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to rheumatoid arthritis (RA) comprises: CSF2, CXCL8, IFN γ , IL10, IL12A, IL12B, IL15, IL16, IL17A, IL18, IL1B, IL4, IL6, MIF, TGFB1, TGFB2, TGFB3, TNF. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to Crohn's disease (CD) comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B. In some embodiments, the trained ML classifier predicts the treatment response with an area under receiver operating characteristic curve (AUC) of at least about 0.8. In some embodiments, the set of cytokines or the ranked set of cytokines is measured using an immunoassay. In some embodiments, the set of transcripts or the ranked set of transcripts is measured using reverse transcription quantitative real-

time polymerase chain reaction (RT-qPCR), microarrays, or RNA-sequencing (RNA-seq). In some embodiments, the therapy comprises an experimental therapy in a clinical trial. In some embodiments, the experimental therapy comprises KO-947.

[0007] In another aspect, disclosed herein is a method for determining biomarkers predictive of treatment response to a therapy for a subject suffering from a disease or disorder, the method comprising: identifying one or more source nodes associated with a set of proteins targeted by the therapy; determining, by a network-based diffusion process, one or more intermediate nodes associated with the one or more source nodes or the one or more response nodes, wherein the one or more intermediate nodes are indicative of the treatment response; generating a ranked set of the one or more intermediate nodes, wherein the ranked set is based at least in part on one or more interactions of the intermediate nodes with the one or more source nodes or the one or more response nodes; and processing a subset of the ranked set of intermediate nodes to determine a set of biomarkers that are predictive of the treatment response to the therapy. In some embodiments, the ranked set of intermediate nodes comprises a ranked set of proteins, a ranked set of cytokines, or a ranked set of transcripts. In some embodiments, the ranked set of proteins, the ranked set of cytokines, or the ranked set of transcripts is used as features to obtain a trained ML classifier. In some embodiments, the set of biomarkers is predictive of the treatment response to the therapy. In some embodiments, (i) the ranked set of cytokines is different from at least one cytokine of the set of cytokines or (ii) the ranked set of cytokines is the same as the set of cytokines. In some embodiments, (i) the ranked set of transcripts is different from at least one transcript of the set of transcripts or (ii) the ranked set of transcripts is the same as the set of transcripts. In some embodiments, the ranked set of cytokines or the ranked set of transcripts is analyzed to determine statistically significant changes after administering to the subject the therapy. In some embodiments, generating the ranked set of the one or more intermediate nodes comprises: determining a first rank between the one or more source nodes and the one or more intermediate nodes; determining a second rank between the one or more response nodes and the one or more intermediate nodes; and determining a rank product of the first rank and the second rank. In some embodiments, the method further comprising: constructing a subset of a protein-protein network associated with the disease or disorder, wherein the constructing comprises: (i) determining the one or more source nodes in the protein-protein network, based at least in part on the set of proteins targeted by the therapy; (ii) determining the one or more response nodes in the protein-protein network, based on least in part on the set of cytokines or the set of transcripts indicative of the treatment response; and (iii) determining one or more connections between the one or more source nodes and the one or more response nodes, based at least in part on one or more interactions

between the one or more source nodes and the one or more response nodes. In some embodiments, the protein-protein network comprises a human interactome (HI). In some embodiments, the disease or disorder comprises an autoimmune disease, an immunology disease, or a cancer. In some embodiments, the immunology disease or the autoimmune disease comprises rheumatoid arthritis (RA), ulcerative colitis (UC), Crohn's disease (CD), inflammatory bowel disease, psoriatic arthritis, ankylosing spondylitis, chronic psoriasis, hidradenitis suppurativa, multiple sclerosis, juvenile idiopathic arthritis, uveitis, systemic lupus, Type 1 diabetes, Graves' Disease, Hashimoto's Thyroiditis, celiac disease, Sjögren's Syndrome, or an immune-mediated disease. In some embodiments, the immunology disease or the autoimmune disease comprises RA. In some embodiments, the immunology disease or the autoimmune disease comprises UC. In some embodiments, the immunology disease or the autoimmune disease comprises CD. In some embodiments, the therapy for the immunology disease or the autoimmune disease comprises an anti-TNF therapy. In some embodiments, the anti-TNF therapy comprises infliximab, adalimumab, etanercept, certolizumab pegol, or golimumab, or a biosimilar thereof. In some embodiments, the therapy for the immunology disease or the autoimmune disease comprises an alternative to anti-TNF therapy. In some embodiments, the alternative to anti-TNF therapy comprises anti-CD20, JAK, anti-IL6, rituximab, sarilumab, tofacitinib citrate, leflunomide, vedolizumab, tocilizumab, anakinra, abatacept, TL1A, co-stimulation blockade, or a biosimilar thereof. In some embodiments, the cancer comprises bladder cancer, breast cancer, colon cancer, rectal cancer, endometrial cancer, kidney cancer, leukemia, liver cancer, lung cancer, melanoma, Non-Hodgkin's lymphoma, pancreatic cancer, prostate cancer, or thyroid cancer. In some embodiments, the therapy for the cancer comprises a targeted therapy, chemotherapy, hormone therapy, immunotherapy, photodynamic therapy, radiation therapy, stem cell transplant therapy, or any combination thereof. In some embodiments, the trained ML classifier predicts the treatment response to the therapy using a non-linear relationship between (i) measurements of the ranked set of intermediate nodes and (ii) the treatment response indicated by statistically significant changes in the ranked set of intermediate nodes after administering to the subject the therapy. In some embodiments, the trained ML classifier is trained using measurements of the one or more intermediate nodes in (i) a first set of subjects with the disease or disorder determined to be responsive to the therapy and (ii) a second set of subjects with the disease or disorder determined to be non-responsive to the therapy. In some embodiments, the trained ML classifier comprises a neural network, deep learning, perceptron, random forest, Bayes, Markov, Gaussian process, clustering algorithm, support vector machine, generative model, or kernel. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to an immunology disease or an autoimmune disease

comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL-12A/B, IL12B, IL13, IL15, IL16, IL17A, IL18, IL1B, IL2, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, IL8, MIF, TGFB1, TGFB2, TGFB3, or TNF. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL13, IL17A, IL1B, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, TGFB1, or TNF. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to rheumatoid arthritis (RA) comprises: CSF2, CXCL8, IFN γ , IL10, IL12A, IL12B, IL15, IL16, IL17A, IL18, IL1B, IL4, IL6, MIF, TGFB1, TGFB2, TGFB3, TNF. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to Crohn's disease (CD) comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B. In some embodiments, the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B. In some embodiments, the trained ML classifier predicts the treatment response with an area under receiver operating characteristic curve (AUC) of at least about 0.8. In some embodiments, the set of cytokines or the ranked set of cytokines is measured using an immunoassay. In some embodiments, the set of transcripts or the ranked set of transcripts is measured using reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR), microarrays, or RNA-sequencing (RNA-seq). In some embodiments, the therapy comprises an experimental therapy in a clinical trial.

[0008] In another aspect, disclosed herein is a method for administering an experimental therapy to a subject in a clinical trial, comprising: administering to the subject the experimental therapy, wherein the experimental therapy has been predicted to generate a treatment response in the subject based at least in part on a trained machine learning (ML) classifier that classifies the subject as responsive or non-responsive to the experimental therapy, wherein the trained ML classifier predicts the treatment response based at least in part on analyzing in the subject a set of cytokines or a set of transcripts.

[0009] In another aspect, disclosed herein is a method for selecting a subject to receive an experimental therapy in a clinical trial, comprising: selecting to the subject to receive the experimental therapy, wherein the experimental therapy has been predicted to generate a treatment response in the subject based at least in part on a trained machine learning (ML) classifier that classifies the subject as responsive or non-responsive to the therapy, wherein the trained ML classifier predicts the treatment response based at least in part on analyzing in the subject a set of cytokines or a set of transcripts.

[0010] In another aspect, disclosed herein is a computer program product for determining biomarkers predictive of treatment response to a therapy for a subject suffering from a disease or disorder, the computer program product comprising at least one non-transitory computer-readable medium having computer-readable program code portions embodied therein, the computer-readable program code portions comprising: an executable portion configured to identify one or more source nodes associated with a set of proteins targeted by the therapy; an executable portion configured to determine, by a network-based diffusion process, one or more intermediate nodes associated with the one or more source nodes or the one or more response nodes, wherein the one or more intermediate nodes are indicative of the treatment response; an executable portion configured to generate a ranked set of the one or more intermediate nodes, wherein the ranked set is based at least in part on one or more interactions of the intermediate nodes with the one or more source nodes or the one or more response nodes; and an executable portion configured to process a subset of the ranked set of intermediate nodes to determine a set of biomarkers that are predictive of the treatment response to the therapy.

INCORPORATION BY REFERENCE

[0011] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0013] **FIGs. 1A-1H** depict an overview, implementation, and validation of the biomarker-discovery framework. **FIG. 1A** depicts a schematic representation of a human interactome (HI) after running the personalized PageRank (PPR) algorithm herein starting from the drug target as a source node. **FIG. 1B** depicts a schematic representation of the HI after running the personalized PageRank algorithm starting from the diseases-signature cytokines as readout nodes. **FIG. 1C** depicts a schematic representation of the HI reflecting a final node ranking based on two rankings in **FIGs. 1A-1B**. Schematic representations of framework implementation: **FIG. 1D** depicts rheumatoid (RA) cytokines and tumor necrosis factor (TNF) (e.g., infliximab target) mapped on

the whole-blood network. **FIG. 1E** depicts ulcerative colitis (UC) cytokines and TNF mapped on the colon network. **FIG. 1F** depicts UC cytokines and extracellular signal-regulated kinase (ERK1/2) (KO-947 targets) mapped on the colon network. Validation of framework using patient-derived gene-expression data: **FIG. 1G** depicts supervised-classification methods for reasonable sample size and with repeated k-fold cross-validation (CV), and **FIG. 1H** depicts unsupervised clustering for limited sample size. Abbreviations: CV, cross-validation; ERK, extracellular signal-regulated kinase; rheumatoid arthritis, RA; TNF, tumor necrosis factor; UC, ulcerative colitis;

[0014] **FIGs. 2A-2C** depict retrospective validation of response predictive biomarkers in UC-patient gene-expression data. **FIG. 2A** depicts average results for six performance metrics for 50 stratified splits on three feature sets. The models fed with the top 100 Predictive Response Biomarkers using Network medicine (PRoBeNet) biomarkers using methods herein outperform models fed with all genes and 20 sets of randomly selected 100 genes in all performance metrics. Asterisks reflect whether the outperformance was significant (p-value < 0.05). **FIG. 2B** depicts that sample-size analysis using the area under the curve (AUC) of the receiver operating characteristic (ROC) (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 30% of data is used. **FIG. 2C** depicts the effect of varying the number of biomarkers (top k) used for model training and validation. Abbreviations: AUCROC, the area under the curve of the receiver operating characteristic; AUPRC, the area under the precision-recall curve; TNFi, tumor necrosis factor inhibitor. Abbreviations: as in **FIGs. 1A-1H**;

[0015] **FIGs. 3A-3C** depict retrospective validation of response predictive biomarkers in RA-patient whole-blood gene-expression data. **FIG. 3A** depicts average results for six performance metrics for 50 stratified splits on three feature sets. The model fed with the top 100 Predictive Response Biomarkers using Network medicine (PRoBeNet) biomarkers using methods herein outperforms models fed with all genes and 20 sets of randomly selected 100 genes in all performance metrics. Asterisks reflect whether the outperformance was significant (p-value < 0.05). **FIG. 3B** depicts that sample-size analysis using the area under the curve (AUC) of the receiver operating characteristic (ROC) (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. **FIG. 3C** depicts the effect of varying the number of biomarkers (top k) used for model training and validation. Abbreviations: as in **FIGs. 1A-1H** and **2A-2C**;

[0016] **FIGs. 4A-4C** depict preclinical validation of KO-947 (an extracellular signal-regulated kinase (ERK1/ERK2) inhibitor) response predictive biomarkers. **FIG. 4A** depicts an example

study design using methods herein. **FIG. 4B** depicts a uniform manifold approximation and projection (UMAP) generated using all genes as inputs. **FIG. 4C** depicts a UMAP generated from the top 100 predicted biomarkers; responders are well separated from non-responders. Abbreviations: Crohn's disease, CD; extracellular signal-regulated kinases, ERK; Staphylococcus enterotoxin B, SEB; uniform manifold approximation and projection, UMAP;

[0017] **FIGs. 5A-5B** depict rank-based correlations as a heat map. Predictive Response Biomarkers using Network medicine (PRoBeNet) biomarkers are obtained using the rank product of source-specific PageRank and readout-specific PageRank. The PRoBeNet rankings do not show a correlation with rankings from most common network centrality measures such as degree centrality, betweenness centrality, and generic PageRank. Generic PageRank is run without personalization. **FIG. 5A** depicts ulcerative colitis (UC). **FIG. 5B** depicts rheumatoid arthritis (RA);

[0018] **FIGs. 6A-6C** depict a cohort of rheumatoid arthritis (RA) patients with extreme responses. **FIG. 6A** depicts that the cohort consists of 67 extreme non-responder patients who achieved less than 20% improvement as assessed by six-month American College of Rheumatology criteria (e.g., ACR20). Fourteen patients achieved greater than 70% improvement in RA symptoms as assessed by six-month ACR70. **FIG. 6B** depicts the performance of different feature sets compared using the area under the curve (AUC) of the receiver operating characteristic (ROC) (AUCROC). **FIG. 6C** depicts that sample size analysis shows Predictive Response Biomarkers using Network medicine (PRoBeNet) outperforming ALL and RAND models;

[0019] **FIGs. 7A-7C** depict pathway enrichment analysis performed using Predictive Response Biomarkers using Network medicine (PRoBeNet) biomarkers. **FIG. 7A** depicts a circus plot showing the intersection of PRoBeNet biomarkers between rheumatoid arthritis (RA) and ulcerative colitis (UC), where the interior lines link identical genes. **FIG. 7B** depicts functional similarity in terms of Gene Ontology between RA-specific, UC-specific, and common PRoBeNet biomarkers with the interior lines linking genes within the same ontology term. **FIG. 7C** depicts enriched pathways identified by the PRoBeNet biomarkers for RA and UC;

[0020] **FIG. 8** depicts an example list of cytokines related to ulcerative colitis (UC), rheumatoid arthritis (RA), and Crohn's Disease (CD) disease signatures;

[0021] **FIGs. 9A-9B** depict examples of hierarchical clustering analysis (HCA). **FIG. 9A** depicts an example dendrogram obtained after performing HCA with Predictive Response Biomarkers using Network medicine (PRoBeNet) biomarkers. Three of the 12 samples are misclassified. **FIG. 9B** depicts an example dendrogram obtained with all genes while performing HCA, and 5 of 12 samples are misclassified;

[0022] **FIG. 10** depicts an example histogram showing a distribution of adjusted rand index (ARI) scores for 10,000 sets of 100 randomly selected genes. Predictive Response Biomarkers using Network medicine (PRoBeNet) biomarkers yielded an ARI of 0.1773 (p-value = 0.0109), whereas ALL genes yielded an ARI of -0.0433 (p-value = 0.7161);

[0023] **FIGs. 11A-11D** depict the comprehensive network integration achieved by Predictive Response Biomarkers using Network medicine (PRoBeNet). **FIG. 11A** depicts seventy-eight differentially expressed (DE) genes between responders and non-responders to infliximab treatment obtained from a cohort, e.g., GSE12251. **FIG. 11B** depicts that no significant DE genes were seen in a cohort, e.g., GSE14580. **FIG. 11C** depicts DE genes mapped to a human interactome (HI). DE genes are scattered on a human interactome and show a few interactions. **FIG. 11D** depicts PRoBeNet biomarkers and DE genes mapped to a HI. Most of the DE genes appear as neighbors of one or more PRoBeNet biomarkers; and

[0024] **FIG. 12** depicts an example computing device configured to perform methods herein.

DETAILED DESCRIPTION

[0025] While various embodiments of the present disclosure have been shown and described herein, such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur without departing from the present disclosure. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed.

[0026] The Predictive Response Biomarkers using Network medicine (PRoBeNet) framework of the present disclosure provides at least a technical solution as a feature-reduction tool, which may be particularly useful for constructing more robust machine-learning models when limited patient data may be available. The PRoBeNet framework herein may also be used to develop companion and complementary diagnostic tests for under-development drugs (e.g., experimental drugs in clinical trials) and developed drugs (e.g., established or approved drugs) for inflammatory and immunology (I&I) diseases. Thus, the PRoBeNet framework herein can increase the likelihood of new drugs being approved by stratifying suitable patient subgroups in clinical trials, which can potentially improve patient outcomes, e.g., response to therapy.

Introduction

[0027] A key promise of precision medicine can include matching patient subgroups with the most appropriate treatments (1). This can be **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 (ML) models that unveil complex, generalizable patterns from large molecular and clinical datasets, usually comprising data from hundreds to thousands of patients. For example, analyzing extensive molecular and clinical

datasets from cancer patients, machine-learning models have 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 healthcare payers (9,10).

[0028] Biomarker discovery and validation using typical machine-learning approaches can require relatively large and relevant data sets consisting of both high-quality molecular data collected from patients before initiation of therapy and matched clinical outcome data indicating which patients ultimately responded to treatment and which did not. Typical machine-learning models may not be generalizable when insufficient data is available. A technical challenge in modeling insufficient data is the curse of dimensionality where the amount of data sufficient to generalize from machine-learning results can increase exponentially as the number of features increases. In such cases, machine-learning models can tend to overfit the training data. For example, the models may find biomarkers that seem to be predictive for patients in the training cohort but may not be predictive for patients in other cohorts (11-15). An important example is chronic I&I diseases such as rheumatoid arthritis (RA) and ulcerative colitis (UC), each of which is an immune-mediated inflammatory disease (IMID) (11,12,16-22). The volume of publicly available molecular data from complex I&I disease patients is less than that of cancer patients. For example, as of June 2023, the number of “cancer” datasets in the Gene Expression Omnibus was over one million whereas the number of “autoimmune-disease” datasets was only 27,000 (23). This can occur partly because I&I diseases cause fewer deaths each year than cancer (24). As a result, research pertaining to I&I diseases can garner less attention and funding than cancer research. For example, in fiscal year 2022, the National Institutes of Health (NIH) invested approximately \$3.4 billion in funding for cancer research and approximately \$380 million in I&I disease research. This data can be obtained directly from the NIH RePORTER website by searching for “funding for cancer research” and “funding for autoimmune disease research.” Thus, employing typical machine-learning approaches to find generalizable treatment-response patterns that predict biomarkers for existing and novel I&I therapies can be technically challenging (25).

[0029] A network-based framework, Predictive Response Biomarkers using Network medicine (PRoBeNet), that can usefully reduce the pool of candidate biomarkers (e.g., features) is described herein. The PRoBeNet framework herein can enable machine-learning models to be successfully trained on cohorts with relatively few samples. The PRoBeNet framework recognizes that the effect of a treatment (e.g., a drug targeting a specific protein) can ripple through a cascade of protein-protein interactions, reach the disease signature (e.g., a panel of proteins characteristically dysregulated in that disease), and revert them to normal (e.g., in responding patients). The

PRoBeNet framework, therefore, can use a treatment-targeted protein, defined by a drug's mechanism of action, as input, and the disease-signature proteins, which can be defined by the biological effects of the disease. For example, to find biomarkers predicting RA-patient responses to infliximab, the PRoBeNet framework can consider as input the infliximab-targeted protein (e.g., tumor necrosis factor-alpha; TNF α) together with the appropriate disease-signature cytokine proteins, such as interleukin IL-4, IL-6, and IL-10. Abnormal expression of these cytokines is detectable in RA patient serum years before onset of symptoms. They become increasingly prominent as the disease progresses (26, 27) and are usually targeted by RA treatments (28).

[0030] The PRoBeNet framework herein can successfully identify response-predicting biomarkers for both approved autoimmune disease treatments or therapies and new investigational compounds. The PRoBeNet framework herein can also be applied to other complex diseases, e.g., cancers, when the treatment-targeted proteins and appropriate disease-signature molecules are determined.

Results

Overview of the PRoBeNet biomarker-discovery framework

[0031] Described herein is a novel network-based framework, Predictive Response Biomarkers using Network medicine (PRoBeNet). The PRoBeNet framework can be useful for discovering treatment-response biomarkers for complex diseases. In some cases, the framework may rely on protein interactions integrated by a human interactome (HI). In some cases, the HI can be a consolidated network of experimentally validated physical interactions among cellular components. In the HI, nodes (e.g., N = 18,627 nodes) can represent proteins, and edges (e.g., N = 439,260 edges) can represent experimentally validated binding interactions among these proteins. The HI is typically comprised mostly of proteins but can also include genes. For simplicity, proteins (and their physical interactions) are described here. To find biomarkers using methods herein for an autoimmune disease (e.g., RA) and IMIDs (e.g., UC, CD), a subgraph of the HI specific to each disease can be constructed. Each subgraph can include protein products of genes expressed in specific tissues where data were available (e.g., whole blood for RA, colon tissue for UC and CD) (42).

[0032] Alternatively, to using a disease-specific HI network, the present framework can determine, for each of one or more biomarkers, a significance of correlation with response outcome (e.g., response or non-response) to an anti-TNF therapy or an alternative to anti-TNF therapy among the responders as compared to the non-responders. In some cases, anti-TNF therapies can include infliximab, adalimumab, etanercept, certolizumab pegol, or golimumab, or a biosimilar thereof. In some cases, alternative to anti-TNF therapies can include anti-CD20, JAK, anti-IL6, rituximab,

sarilumab, tofacitinib citrate, leflunomide, vedolizumab, tocilizumab, anakinra, abatacept, TL1A co-stimulation blockade, or a biosimilar thereof.

[0033] In some cases, biomarkers can include expression levels of genes or proteins (e.g., cytokines) associated with response or non-response in RA. For example, expression levels of genes or transcripts associated with RA (e.g., source nodes, response nodes, or intermediate nodes) can include genes or transcripts comprising: ALPL, BCL6, CDK11A, CFLAR, IL1B, JAK3, LIMK2, NOD2, NOTCH1, RGL4, RNF141, RNF175, RPL38, RYBP, SELENBP1, SIGLEC14, SIGLEC5, SLC1A5, SLC25A29, SNAPC5, SOCS3, SORD, STAT1, STAT2, STOML2, STS, SULT1A1, SULT1A3, TAP2, TAS2R43, TDRD7, THBS1, THOC3, TIGD3, TIMM10, TLDC2, TLR1, TLR4, TNFSF10, TNFSF13B, TOR1B, TRAPPC5, TRIM22, TRIM24, TRIM25, TRIM5, TSTA3, UBA52, USP15, VSTM1, WIP1, XAF1, ZBP1, ZC3HAV1, ZCCHC2, ZDHHC17, ZDHHC2, ZFP36, ZNF814, ZNFX1, or ZYG11B. For example, expression levels of proteins or cytokines (e.g., source nodes, response nodes, or intermediate nodes) associated with RA can include CSF2, CXCL8, IFN γ (or IFNG), IL10, IL12A, IL12B, IL15, IL16, IL17A, IL18, IL1B, IL4, IL6, MIF, TGFB1, TGFB2, TGFB3, TNF.

[0034] For example, expression levels of genes or transcripts associated with UC (e.g., source nodes, response nodes, or intermediate nodes) can include genes or transcripts comprising: CEBPB, KLHL12, AMIGO2, MS4A7, CXCL2, MMP12, DRAM1, CXCL1, IGFBP5, MAP3K20, NR3C1, MEIS1, or CXCL6. For example, expression levels of proteins or cytokines (e.g., source nodes, response nodes, or intermediate nodes) associated with UC can include CCL2, CSF2, CXCL8, IFN γ (or IFNG), IL10, IL12A, IL13, IL17A, IL1B, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, TGFB1, TNF. For example, expression levels of proteins or cytokines (e.g., source nodes, response nodes, or intermediate nodes) associated with both Crohn's disease (CD) and UC can include IL10, IFN γ (or IFNG), IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, IL-12A/B.

[0035] Next, two key protein sets in each HI subgraph can be considered: (i) protein(s) targeted by the therapeutic drug (e.g., TNF α targeted by infliximab), called source node(s) and (ii) the repertoire of disease-signature cytokines, called readout nodes (26). Readout nodes can be objectively selected by criteria, e.g., (i) being generally important for inflammation or (ii) being characteristically upregulated or downregulated in RA or UC patients (see **FIG. 8**) (21,26,27,43,44). For example, cytokine readout nodes associated with RA can include CSF2, CXCL8, IFN γ (or IFNG), IL10, IL12A, IL12B, IL15, IL16, IL17A, IL18, IL1B, IL4, IL6, MIF, TGFB1, TGFB2, TGFB3, or TNF. Cytokine readout nodes associated with UC can include CCL2, CSF2, CXCL8, IFN γ (IFNG), IL10, IL12A, IL13, IL17A, IL1B, IL21, IL22, IL23A, IL27, IL4,

IL5, IL6, IL7, TGFB1, or TNF. Cytokine readout nodes common to both CD and UC can include IL10, IFN γ (or IFNG), IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, IL-12A/B.

[0036] Recognized herein by the present framework is that a drug can first affect its target protein(s); then affect intermediary or intermediate proteins; and ultimately affect the disease-signature cytokines. The intermediary or intermediate proteins important for relaying these effects can include the most important response-predicting biomarkers. The framework herein can be agnostic to the amount of patient data available, since it can operate using three inputs: (1) source node(s), (2) readout node(s), and (3) a tissue-specific HI subgraph.

[0037] Some or all nodes of the HI subgraph can be ranked with the personalized PageRank (PPR) algorithm described herein to measure their random-walk-based scores relative to both source nodes and readout nodes (**FIGs. 1A-1B**) (45). Two scores can be determined 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 PageRank algorithm can be useful for integrating the impact of the network topology due to its robustness against noise and uncertainties in the network structure. Furthermore, the personalization aspect of the PageRank algorithm can allow computations to be specific to drug targets and cytokines relevant to the disease being studied.

[0038] The two scores can yield two rankings for each node: (v): $r_{source}(v)$ corresponds to node v 's ranking personalized to drug-target proteins (**FIG. 1A**), and $r_{readout}(v)$ corresponds to node v 's ranking personalized to disease-signature cytokines (**FIG. 1B**). Source and readout rankings can be combined using a rank product (46) (**FIG. 1C**) for node v as:

$$r(v) = \sqrt{r_{source}(v) \cdot r_{readout}(v)} \quad (\text{Equation 1}).$$

[0039] The rank product can be useful (as opposed to other options, like summation) because the ranking distribution follows a Fat-tailed distribution, which may be indicative of multiplicative processes. Importantly, the rank product can also be robust to variations in data quality and scale, making it particularly suitable for integrating two ranks. By using a rank product, high-ranking nodes may be more central than low-ranking ones on the paths between source and readout nodes. In some cases, the top 100 ranked nodes can be selected as candidates for treatment-response-predicting biomarkers and for further validation. These biomarker rankings can be compared with rankings from the most common network-centrality measures, such as degree centrality, betweenness centrality, and generic PageRank. In some cases, the rankings may not correlate, which demonstrates that the PROBeNet framework herein can determine metrics beyond node degree and global centrality (**FIGs. 5A-5B**).

Examples

[0040] While various examples of the present disclosure have been shown and described herein, such examples are provided by way of example only. Numerous variations, changes, or substitutions may occur without departing from the present disclosure. It should be understood that various alternatives to the examples described herein may be employed.

Example 1: Retrospective validation of response biomarkers using methods herein

[0041] The predictive utility of the PRoBeNet framework for discovering biomarkers for complex diseases was retrospectively validated in two cases and was prospectively validated with an investigational drug (Table 1, **FIGs. 1D-1F**). For example, in UC and RA, one-third to one-half of patients respond to the anti-TNF α monoclonal antibody infliximab (47). To determine biomarkers of response, models herein were built and validated with real-world data from UC patients (e.g., N = 47) and RA patients (e.g., N = 107). For these two cases, clinical responses were predicted herein using retrospective gene-expression data collected from patients before treatment initiation. For prospective validation, biomarkers predicting response to the investigational compound KO-947, which inhibits ERK1/2 (a potential target for treating UC and CD patients), were prospectively evaluated with data from four UC and eight CD tissue-donor patients (48). 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. For each donor, cytokine (e.g., N = 15) levels from samples collected before and after treatment were used to identify patients as either responders or non-responders. Gene expression levels from patient samples collected before treatment were used to predict response labels.

[0042] In some cases, repeated cross-validation was used to train and test models predicting response to infliximab (**FIG. 1G**). Because of the small sample size used for the ERK1/2-inhibitor-response model, unsupervised clustering was used to assess that model's predictive utility (**FIG. 1H**).

Table 1. Three validation cases of treatment-response-predicting biomarkers identified by the network-based framework herein in distinct tissue sample types

Indication	n	Treatment (target) ^a	Sample Type ^b	Platform	Response definition ^c
UC	47	Infliximab (TNF α)	Colon biopsy	Microarray	Endoscopic healing at 4–8 weeks
RA	107	Infliximab (TNF α)	Whole blood	RNAseq	6-month ACR50 outcome
UC, CD	12	KO-947 (ERK1/2)	Intestinal biopsy	RNAseq	Change in cytokine expression ^d

^a The first two cases are dedicated to an established drug (e.g., infliximab). The third case is dedicated to an investigational compound (e.g., KO-947).

^b Sample source for data used in model development

^c ACR50 score is calculated as described herein elsewhere.

^d A response was defined as $\geq 20\%$ change in levels of expression.

Example 2: Discovery and validation of biomarkers predicting response to infliximab in UC patients using methods herein

[0043] To discover treatment-response-predicting biomarkers of infliximab in UC, TNF α (the source node), the UC-signature cytokines (the readout nodes; **FIG. 9**), and the colon-specific human interactome (HI) subgraph were used to rank biomarkers of response. Ranked P_{RoBeNet} biomarkers were then validated using pre-treatment gene-expression data from colonic mucosal biopsies from two UC-patient cohorts (49). The cohorts were combined, and batch correction was performed. In this combined cohort, patient-derived samples (e.g., N = 47 comprising 20 responders and 27 non-responders) were treated with infliximab. The combined cohort and the P_{RoBeNet} 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%. Model training and validation were repeated across 50 splits to ensure the results were robust and consistent (**FIG. 1G**). A nested cross-validation scheme was used to avoid overfitting the model to the training data, and an L1 regularization was used to eliminate non-informative features.

[0044] Model performances were then compared across three feature sets used to train the model: (1) RAND, 20 sets of 100 randomly selected genes (a random control); (2) ALL, a set comprising all genes; and (3) P_{RoBeNet} herein, a set comprising the top 100 ranked genes. The use of randomly selected genes served as a baseline comparison. The RAND model had the lowest prediction accuracy (using area under the curve (AUC) of a receiver operating characteristic (ROC) curve (AUCROC) of 0.65; **FIG. 2A**). Noise may be introduced in the ALL model because of inclusion of many irrelevant or less informative features 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 (P_{RoBeNet} biomarkers herein) is often more useful. Indeed, the model relying on the P_{RoBeNet} biomarkers herein (e.g., 100 top-ranked) significantly outperformed (p-value < 0.05) the other two models (RAND and ALL) as assessed by several performance metrics (e.g., AUCROC; area under the precision recall curve (AUPRC); average precision; classification accuracy; balanced accuracy; and F1 score; **FIG. 2A**). An improvement in the accuracy was observed, increasing from 0.70 (ALL) to 0.75 (P_{RoBeNet} herein). This improvement in patient response prediction becomes predominant in the context of clinical decision-making. For example, when extrapolated to a large cohort of 10,000 patients, this improvement in accuracy can lead to better patient stratification and treatment decisions for an additional 500 patients. This improvement can offer substantial real-world benefits in improving healthcare outcomes in patients, e.g., optimizing patient care, causing fewer adverse

effects, and improving survival rates. Such improvements demonstrate the utility of the PProBeNet framework in the context of IMIDs.

[0045] The performance of PProBeNet biomarkers herein, when patient data was limited, was evaluated by artificially reducing the number of patient data (e.g., through 10 different random samplings) used to train and validate the model. This analysis allowed both an assessment of the improved utility of PProBeNet biomarkers under limited-data circumstances and an estimate of the minimum sample size required by methods herein. For example, using as few as 14 samples, constituting one-third of the data, the performance of PProBeNet model (e.g., AUCROC, > 0.8) consistently surpassed performance of the RAND and ALL models (e.g., AUCROCs, 0.65-0.75; **FIG. 2B**). The performance of the ALL model linearly correlated with sample size, illustrating that typical machine-learning methods need large cohorts to accurately determine and filter biomarkers.

[0046] Next, the effect of number of top-ranked PProBeNet biomarkers herein on model's performance was assessed. Model performances were compared after varying the number of top PProBeNet biomarkers used to train the model (e.g., from 50 to 500, in increments of 50). In all variations, the PProBeNet model outperformed the RAND and ALL models (**FIG. 2C**). Overall, these results confirm that the biomarkers found by the PProBeNet framework can be more predictive than biomarkers determined by the RAND and ALL models. These results demonstrate how the PProBeNet framework herein can be practically applied to limited-size datasets to build predictive models.

Example 3: Discovery and validation of biomarkers predicting response to infliximab in RA patients using methods herein

[0047] The PProBeNet framework herein can detect and identify response biomarkers in other complex diseases by modifying source nodes, readout nodes, and HI subgraphs. To show the practical application of this approach in a different context, the PProBeNet framework was used to find infliximab-response-predicting biomarkers for RA, an autoimmune disease. For example, infliximab-targeted TNF α protein was retained as the source node, but RA-signature cytokines rather than UC-signature cytokines were used as readout nodes (**FIG. 8**). Also, a whole-blood-specific HI subgraph was constructed, since whole-blood RA gene-expression data were now used for validation. With these three inputs (source node, readout nodes, and whole-blood HI 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 RA patients (e.g., N = 107) before treatment, was used to assess the predictive utility of the biomarkers herein (31). Patient clinical data needed to calculate American College of Rheumatology (ACR) treatment-response scores were collected at baseline and at six-month follow-up visits. Patients

were defined as either responders or non-responders to infliximab treatment by using six-month ACR50 scores.

[0048] Predicted biomarkers were validated for the RA-patient cohort using the same methods as that used to validate UC-patient-cohort biomarkers. The RA-patient cohort and the PProBeNet biomarkers herein were used to train and validate the model (**FIG. 1G**). Using multiple metrics, the model trained on PProBeNet biomarkers significantly outperformed the models trained on either ALL or RAND genes (**FIG. 3A**). The model trained on ALL performed worse than the model trained on RAND genes (e.g., 100 genes/set). Using many features of the ALL set, machine learning models can overfit to the noise in training data.

[0049] The sample size analysis was repeated for the RA-patient cohort to emulate the real-world scenario of data paucity or scarcity. For example, using as little as one-half of the data (e.g., $N = 53$), the PProBeNet ranked-biomarker model consistently outperformed the ALL and RAND models (**FIG. 3B**). Thus, even with limited patient data, the PProBeNet biomarkers herein can predict responders and non-responders better than do other models. In some cases, model performance peaked when using about 100 PProBeNet biomarkers (**FIG. 3C**). Thus, the PProBeNet framework can determine the biomarkers with the greatest predictive utility.

[0050] 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, methods herein were performed by using extreme-responder RA patients (e.g., patients who achieved $> 70\%$ improvement in RA symptoms, as assessed by six-month ACR70) and extreme-non-responder patients (e.g., patients who achieved $< 20\%$ improvement, as assessed by six-month ACR20). When using these patients, the PProBeNet ranked-biomarker model (e.g., AUCROC, 0.64) further outperformed the other models (e.g., AUCROCs, RAND = 0.46 and ALL = 0.5; **FIGs. 6A-6C**). For insight into the biology of determining whether RA and UC patients will respond to TNFi, pathways overrepresented by PProBeNet biomarkers (e.g., top PProBeNet biomarkers for RA and for UC) were determined using Metascape[®] gene annotation and analysis. Several canonical pathways were enriched in both the RA and UC biomarkers that are related to TNF α and cytokine signaling. Comparing the enriched pathways for each disease revealed several UC-specific pathways, such as cadherin and endothelial cell adhesion, thereby gaining further disease specific insights from the predicted biomarkers herein (50) (**FIGs. 7A-7C**).

[0051] Overall, these results demonstrate that the PProBeNet framework herein can determine response-predicting biomarkers for different complex diseases, even when using limited data.

Example 4: Discovering biomarkers predicting response to ERK1/ERK2 inhibitor using methods herein

[0052] Precision medicine has the potential to improve drug-response rates by discovering biomarkers of therapeutic response and identifying patients carrying those biomarkers. This may also translate into U.S. Food and Drug Administration (FDA) approval for drugs that are administered with an accompanying diagnostic tool, e.g., companion and complementary diagnostics. Using methods herein, preclinical validation was performed to assess the performance of PRoBeNet biomarkers herein for an investigational compound (e.g., KO-947), whose mechanism of action can differ from that of infliximab. KO-947 inhibits extracellular signal-regulated kinase 1 and 2 (ERK1/ERK2), which can modulate inflammatory cytokine signaling. Thus, KO-947 may be efficacious for treating I&I diseases and IMIDs (51). Molecular and outcome data were prospectively generated using methods herein in a preclinical setting (52). PRoBeNet biomarkers herein were identified using ERK1 and ERK2 as source nodes, fifteen cytokines (e.g., both pro-inflammatory and anti-inflammatory) whose levels were measured as readout nodes (**FIG. 8**), and the colon-specific HI subgraph. Despite the differences in UC and CD etiologies, ERK1/2 inhibitors work in a similar manner in the tissues derived from UC and CD patients (e.g., cytokine reduction). Colon or ileum samples were collected from four UC and eight CD patients undergoing therapeutic resection (**FIG. 4A**). Tissue samples were cultured and treated either with Staphylococcus enterotoxin and vehicle (e.g., baseline samples) or Staphylococcus enterotoxin and KO-947 (e.g., treated samples). Disease-signature cytokines secreted into media and gene expression in tissue were measured for all samples. Response scores for patients were calculated using changes in cytokine-expression levels between baseline and treated samples and across replicates. The response scores for UC and CD samples differed among donor types (Table 2), confirming this ex vivo approach was useful for discovering response biomarkers in a preclinical setting.

[0053] Using the few samples, the utility of the PRoBeNet biomarkers herein to classify responders and non-responders 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 (UMAP) algorithm. Three-fourths of samples (e.g., 9 of 12) were properly classified by treatment-response levels (**FIG. 4B**). To compare the predictive utility of the PRoBeNet biomarkers over other approaches, UMAP was implemented on data from ALL. Here, the samples did not clearly cluster (**FIG. 4C**). The hierarchical clustering analysis (HCA) performed using the PRoBeNet biomarkers had fewer misclassifications (e.g., 3 of 12) than did HCA performed using

ALL genes (e.g., 5 of 12) (**FIGs. 9A-9B**). The comparison of these results was quantitatively measured using an adjusted rand index (ARI) score, a clustering similarity metric that ranges from -1 to 1, where a higher score indicates better clustering and a score of 0 indicates random labeling. Clustering using the PProBeNet biomarkers yielded an ARI score of 0.1773 (p-value = 0.0109; **FIG. 10**), demonstrating statistically improved accuracy in representing true labels than that obtained using ALL, which yielded an ARI score of -0.0433.

[0054] These results demonstrate the PProBeNet framework herein can be employed for finding response-predicting biomarkers for both approved drugs as well as investigational compounds, even with few samples.

Table 2. Ex vivo responses of UC- and CD-patient tissue samples to KO-947 treatment^a

Donor ID	Cytokine-response change	Binary response ^b	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

^a Abbreviations: non-responder, NR; responder, R.

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

Discussion

[0055] The PProBeNet framework herein demonstrates a practical application for discovering treatment response-predicting biomarkers. The PProBeNet framework can be applied to leveraging precision-medicine strategies to treat complex diseases because the framework does not require an extensive patient dataset. Instead, the PProBeNet framework can take advantage of network medicine methods demonstrating the predictive utility of tissue-specific HI networks. Recognized herein is that a therapy can propagate through a cascade of protein-protein interactions to bring expression levels of disease-specific signatures (e.g., cytokines abnormally expressed in each disease) closer to a normal range. In some cases, the PProBeNet framework may prioritize intermediary or intermediate nodes in the HI that play important roles in transmitting the effects of therapies (e.g., infliximab or KO-947) from source-node target proteins (e.g., TNF or ERK1/2) to readout-node disease-signature cytokines (e.g., RA, UC, or CD). For example, predictive biomarkers herein can be selected from prioritized intermediary or intermediate nodes to determine patient responses to each treatment. The PProBeNet biomarkers can more accurately predict

treatment responses and can be less sensitive to training-set sizes compared to biomarkers found by other approaches. The P_{Ro}BeNet framework can also reduce or optimize potential biomarkers to a more manageable number, which can mitigate a common problem of typical machine-learning approaches when dealing with small cohorts and many features (e.g., the curse of dimensionality). **[0056]** A distinctive characteristic of the P_{Ro}BeNet framework herein can be the dual use of independent PageRank scores. The first score, recognizing proteins important for source-node-derived effects, can determine proteins critical to the therapy's mechanism of action. The second score, recognizing proteins connected to readout nodes, can determine proteins that strongly influence disease pathogenesis and cytokine dysregulation. Using methods herein to integrate both scores via the rank product (Equation 1), the P_{Ro}BeNet framework can evaluate each biomarker's relevance to both source nodes and readout nodes, which can ensure that high-scoring biomarkers are not only associated with the therapy but also important for transmitting effects to cytokines. Importantly, this dual-score approach can help to better understand the molecular mechanisms underlying a successful response. Thus, the dual-score approach can also complement typical, purely data-driven biomarker-discovery methods. For example, typical approaches may rely on machine learning or statistical models to identify biomarkers based on their correlation with response outcomes. Unlike other machine learning approaches, the P_{Ro}BeNet framework can consider molecular mechanisms of response, so it may reveal useful biomarkers that other approaches fail to identify. Also, differentially expressed (DE) genes between responders and non-responders of therapy may not be generalizable. For example, DE genes when mapped onto the HI tend to be scattered with sparse interactions. In contrast, P_{Ro}BeNet biomarkers can consistently form a connected network component. Notably, many of these DE genes may be situated as neighbors to one or more P_{Ro}BeNet biomarkers herein, which can demonstrate the comprehensive network integration achieved by the P_{Ro}BeNet framework (**FIGs. 11A-11C**).

[0057] Pathway analysis of TNFi biomarkers in UC and RA demonstrates that the P_{Ro}BeNet framework herein can also explain crosstalk among pathways driving resistance to established therapies. Thus, the P_{Ro}BeNet framework can identify new therapies to overcome drug resistance. For example, infliximab response-predicting biomarkers were significantly enriched with genes belonging to glucocorticoid-receptor-dependent gene-regulatory networks (**FIGs. 7A-7C**). Glucocorticoids (e.g., prednisone) can rapidly reduce inflammation in RA and UC patients (53,54). However, TNF α therapy can induce glucocorticoid resistance because the TNF α -signaling pathway is significantly interconnected with and coregulated with the glucocorticoid-receptor pathway (55,56). These results demonstrate that drug resistance may be overcome by using drugs conjugated to antibodies that target cell-surface TNF α and glucocorticoid receptor (57). Similar

drug-antibody conjugates can be used to treat cancer patients who inadequately respond to monotherapies but are yet to be used to treat patients with autoimmune disease and IMID (58). Thus, methods herein correctly indicated a key mechanism of glucocorticoid resistance and supported a distinct combination therapy for overcoming this resistance.

[0058] Although predictive biomarkers were discovered and successfully validated with the PRoBeNet framework herein, some challenges can exist. A challenge when finding RA biomarkers was a low signal-to-noise ratio (SNR) in patients' whole-blood gene-expression data. While RA biomarkers were found by using data from whole blood, which may not directly manifest RA disease, UC and CD biomarkers were found using data from colon tissue, which can manifest UC and CD. SNRs in data from tissues that do not actively manifest a disease can be lower than SNRs in data from tissues that do. For example, the SNR in the RA data was much lower than it was in the UC data. Despite this challenge, the PRoBeNet framework identified response-predicting biomarkers for RA, thus demonstrating the successful application of methods herein in datasets with poor SNR. These results show a practical application of the PRoBeNet framework in identifying relevant biomarkers to stratify or classify patients as responders or non-responders.

[0059] In conclusion, the PRoBeNet framework herein can be used for predicting a reduced pool of therapy-response biomarkers to build generalizable models for diseases with little data available (e.g., as few as 14 samples). The PRoBeNet framework can be applied to developing companion and complementary diagnostic tests. For example, using methods herein, such applications can determine appropriate patient subpopulations and direct the subpopulations to specific therapies thereby improving drug efficacies. Also, the PRoBeNet framework can improve clinical development programs by determining or predicting suitable patient populations in phase 2/3 studies. Further, the PRoBeNet framework can be generalized to other complex diseases (e.g., cancers) by using different sets of molecular phenotype markers, e.g., differentially expressed genes or proteins, as readout nodes. Additionally, the PRoBeNet framework can be used to reduce the number of features when searching for biomarkers in high dimensional data. The PRoBeNet framework can bring the scientific community closer to realizing the full potential of precision medicine for treating I&I diseases and other complex conditions by improving personalized treatment strategies.

Methods

Human interactome

[0060] 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. For each disease, a tissue-specific subgraph was constructed by extracting a largest connected component of the HI. Each tissue-specific subgraph had 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

[0061] The Personalized PageRank (PPR) algorithm (e.g., a network-based diffusion process) is a variant of the 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:

$$PR(v) = \frac{1-d}{N} + d \cdot \sum \frac{PR(u)}{L(u)} \cdot \delta(u, v) \quad (\text{Equation 2})$$

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.

[0062] The personalization probabilities are equally distributed among all personalization nodes (e.g., $1/|T|$ 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).

[0063] The NetworkX library ver. 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 based on the two PageRank scores, and these two ranks are merged into a single combined rank by using rank product.

Logistic regression and validation

[0064] 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 create sparse models by encouraging some feature weights to be zero. 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. 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 was used for training and hyperparameter optimization; and 20%, for evaluating model performance. In the inner loop, a five-fold stratified split with grid search was used to select the best hyperparameters. This process was repeated 50 times, and the average of the six performance metrics was reported.

UC cohort

[0065] Two publicly available datasets associated with UC (GSE14580 and GSE12251) were collected from Gene Expression Omnibus. In the first cohort (GSE14580), 24 patients with active UC, who were refractory to corticosteroids and/or immunosuppression, underwent colonoscopy with biopsies from diseased colon tissues at least one week before receiving their first intravenous infusion of infliximab (5 mg/kg body weight). Response to infliximab was defined as endoscopic and histologic healing four to six weeks after the first infusion. In the second cohort (GSE12251), 22 patients underwent colonoscopy with biopsy before infliximab infusion (two samples were derived from the same patient). Response to infliximab was defined as endoscopic and histologic healing after eight weeks. For both cohorts, total RNA was isolated from colonic mucosal biopsies, labeled, and hybridized to Affymetrix Human Genome U133 Plus 2.0 Arrays. Batch effects among studies were corrected by using ComBat ver. 3.44.0. Gender and other demographic information were not available for both cohorts.

RA cohort

[0066] Patients with a clinical diagnosis of RA were included in the Molecular Signature Response Classifier (MSRC) test arm (31). Patients were at least 18 years of age and had never received TNFi treatment (e.g., N = 107). Patient data were collected in the Study to Accelerate Information of Molecular Signatures (AIMS) between August 2020 and August 2022. All patients in this study had moderate-to-high baseline clinical disease activity scores (e.g., > 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 six-month follow-up visits ($\pm$ four 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 RA patients 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 serological and blood-marker levels (e.g., C-reactive protein, erythrocyte sedimentation rate) (29,32,33). 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 six-month ACR50 scores, the RA cohort (e.g., N = 107) comprised more non-responders (e.g., N = 82) than responders (e.g., N = 25), consistent with insufficient infliximab response being common among RA patients. Most patients were women (e.g., 78%, 84 out of 107), as in the general RA population.

Statistical analyses

[0067] A two-tailed t-test from stats library in scipy package ver. 1.5.4 (41) was performed to assess the significance of differences in model performances (for each of six performance metrics) between the 100 top-ranked biomarkers, all genes, and 100 randomly selected gene sets. Differences were considered significant for p-value < 0.05.

Tissue sample collection

[0068] Colon or ileum tissue sections were obtained from four UC and eight CD patients undergoing therapeutic resection surgery. Tissues were maintained in tissue solution until use. The cohort consisted of seven women and five men.

Ex vivo cultures

[0069] Biopsy tissues (e.g., 5 mm²) were cultured in triplicate in 12-well plates containing a modified Connaught Medical Research Laboratory (CMRL)-based cell-culture media. Tissues were positioned such that the apical mucosal surfaces faced up at the liquid-air interface. Culture plates were incubated (e.g., at 37 °C, 5% CO₂) in media with either test compounds or vehicle. Tissues were treated with Staphylococcus enterotoxin B to stimulate inflammatory cytokine secretion. Cell-culture supernatants (e.g., about 1 milliliter (mL)) were collected and flash frozen 18 hours after treatment. Two samples from each condition were stabilized in RNeasy Lysis Buffer (Qiagen) and frozen.

Cytokine analysis

[0070] Disease-signature cytokines were selected for analyses to determine if their dysregulation is associated with either RA or UC or if they are generally important for inflammation (34-37). For UC and CD patient derived tissues, supernatant cytokine levels were analyzed with the Luminex 200 platform by using Miliplex[®] MAP kits (Merck Millipore[®]), according to manufacturer's instructions. These 15 cytokines were measured: GM-CSF, IFN γ (or IFNG), IL-10, IL-12p70, IL-13, IL-17, IL-1b, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, MCP-1, and TNF α . All measurements were made in duplicate.

RNA extraction and sequencing

[0071] 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 nanograms per picoliter (ng/pL) (> 0.4 picograms (pg) yield in > 20 pL), and RNA integrity numbers were confirmed to be > 6.8 (with flat baselines).

[0072] Paired-end reads (e.g., 150 nucleotides long) were mapped to the GRCh38 human genome with STAR[®] alignment software (38,39). Per gene abundance in fragments per kilobase of transcript per million mapped reads was calculated with the RSEM[®] software package (40).

Cytokine-response score

[0073] The cytokine-response score for a donor d was defined as the following weighted sum:

$$S(d) = \frac{1}{n_c} \sum_c p(c) \delta_c \quad (\text{Equation 3})$$

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

$$p = \frac{\bar{l}_{treatment}}{\bar{l}_{control}} - 1 \quad (\text{Equation 4})$$

where n_c is the number of cytokines for which this percentage can be computed (e.g., 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 pro-inflammatory (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 pre-defined threshold, S_{cutOff} (20%).

Computing systems

[0074] Referring to FIG. 12, a block diagram is shown depicting an exemplary machine that includes a computer system 1200 (e.g., a processing or computing system) within which a set of instructions can execute for causing a device to perform or execute any one or more of the aspects and/or methodologies herein. The components in FIG. 12 are examples only and do not limit the scope of use or functionality of any hardware, software, embedded logic component, or a combination of two or more such components implementing particular embodiments.

[0075] Computer system 1200 may include one or more processors 1201, a memory 1203, and a storage 1208 that communicate with each other, and with other components, via a bus 1240. The bus 1240 may also link a display 1232, one or more input devices 1233 (which may, for example,

include a keypad, a keyboard, a mouse, a stylus, etc.), one or more output devices 1234, one or more storage devices 1235, and various tangible storage media 1236. All of these elements may interface directly or via one or more interfaces or adaptors to the bus 1240. For instance, the various tangible storage media 1236 can interface with the bus 1240 via storage medium interface 1226. Computer system 1200 may have any suitable physical form, including but not limited to one or more integrated circuits (ICs), printed circuit boards (PCBs), mobile handheld devices (such as mobile telephones or PDAs), laptop or notebook computers, distributed computer systems, computing grids, or servers.

[0076] Computer system 1200 includes one or more processor(s) 1201 (e.g., central processing units (CPUs) or general purpose graphics processing units (GPUs)) that carry out functions. Processor(s) 1201 optionally contains a cache memory unit 1202 for temporary local storage of instructions, data, or computer addresses. Processor(s) 1201 are configured to assist in execution of computer-readable instructions. Computer system 1200 may provide functionality for the components depicted in **FIG. 12** as a result of the processor(s) 1201 executing non-transitory, processor-executable instructions embodied in one or more tangible computer-readable storage media, such as memory 1203, storage 1208, storage devices 1235, and/or storage medium 1236. The computer-readable media may store software that implements particular embodiments, and processor(s) 1201 may execute the software. Memory 1203 may read the software from one or more other computer-readable media (such as mass storage device(s) 1235, 1236) or from one or more other sources through a suitable interface, such as network interface 1220. The software may cause processor(s) 1201 to carry out one or more processes or one or more steps of one or more processes described or illustrated herein. Carrying out such processes or steps may include defining data structures stored in memory 1203 and modifying the data structures as directed by the software.

[0077] The memory 1203 may include various components (e.g., machine-readable media) including, but not limited to, a random access memory component (e.g., RAM 1204) (e.g., static RAM (SRAM), dynamic RAM (DRAM), ferroelectric random access memory (FRAM), phase-change random access memory (PRAM), etc.), a read-only memory component (e.g., ROM 1205), and any combinations thereof. ROM 1205 may act to communicate data and instructions unidirectionally to processor(s) 1201, and RAM 1204 may act to communicate data and instructions bidirectionally with processor(s) 1201. ROM 1205 and RAM 1204 may include any suitable tangible computer-readable media described below. In one example, a basic input/output system 1206 (BIOS), including basic routines that help to transfer information between elements within computer system 1200, such as during start-up, may be stored in memory 1203.

[0078] Fixed storage 1208 is connected bidirectionally to processor(s) 1201, optionally through storage control unit 1207. Fixed storage 1208 provides additional data storage capacity and may also include any suitable tangible computer-readable media described herein. Storage 1208 may be used to store operating system 1209, executable(s) 1210, data 1211, applications 1212 (application programs), and the like. Storage 1208 can also include an optical disk drive, a solid-state memory device (e.g., flash-based systems), or a combination of any of the above. Information in storage 1208 may, in appropriate cases, be incorporated as virtual memory in memory 1203.

[0079] In one example, storage device(s) 1235 may be removably interfaced with computer system 1200 (e.g., via an external port connector (not shown)) via a storage device interface 1225. Particularly, storage device(s) 1235 and an associated machine-readable medium may provide non-volatile and/or volatile storage of machine-readable instructions, data structures, program modules, and/or other data for the computer system 1200. In one example, software may reside, completely or partially, within a machine-readable medium on storage device(s) 1235. In another example, software may reside, completely or partially, within processor(s) 1201.

[0080] Bus 1240 connects a wide variety of subsystems. Herein, reference to a bus may encompass one or more digital signal lines serving a common function, where appropriate. Bus 1240 may be any of several types of bus structures including, but not limited to, a memory bus, a memory controller, a peripheral bus, a local bus, and any combinations thereof, using any of a variety of bus architectures. As an example and not by way of limitation, such architectures include an Industry Standard Architecture (ISA) bus, an Enhanced ISA (EISA) bus, a Micro Channel Architecture (MCA) bus, a Video Electronics Standards Association local bus (VLB), a Peripheral Component Interconnect (PCI) bus, a PCI-Express (PCI-X) bus, an Accelerated Graphics Port (AGP) bus, HyperTransport (HTX) bus, serial advanced technology attachment (SATA) bus, and any combinations thereof.

[0081] Computer system 1200 may also include an input device 1233. In one example, a user of computer system 1200 may enter commands and/or other information into computer system 1200 via input device(s) 1233. Examples of an input device(s) 1233 include but are not limited to, an alpha-numeric input device (e.g., a keyboard), a pointing device (e.g., a mouse or touchpad), a touchpad, a touch screen, a multi-touch screen, a joystick, a stylus, a gamepad, an audio input device (e.g., a microphone, a voice response system, etc.), an optical scanner, a video or still image capture device (e.g., a camera), and any combinations thereof. In some embodiments, the input device is a Kinect[®], Leap Motion[®], or the like. Input device(s) 1233 may be interfaced to bus 1240 via any of a variety of input interfaces 1223 (e.g., input interface 1223) including, but not limited

to, serial, parallel, game port, USB, FIREWIRE, THUNDERBOLT, or any combination of the above.

[0082] In particular embodiments, when computer system 1200 is connected to network 1230, computer system 1200 may communicate with other devices, specifically mobile devices and enterprise systems, distributed computing systems, cloud storage systems, cloud computing systems, and the like, connected to network 1230. Communications to and from computer system 1200 may be sent through network interface 1220. For example, network interface 1220 may receive incoming communications (such as requests or responses from other devices) in the form of one or more packets (such as Internet Protocol (IP) packets) from network 1230, and computer system 1200 may store the incoming communications in memory 1203 for processing. Computer system 1200 may similarly store outgoing communications (such as requests or responses to other devices) in the form of one or more packets in memory 1203 and communicated to network 1230 from network interface 1220. Processor(s) 1201 may access these communication packets stored in memory 1203 for processing.

[0083] Examples of the network interface 1220 include but are not limited to, a network interface card, a modem, and any combination thereof. Examples of a network 1230 or network segment 1230 include, but are not limited to, a distributed computing system, a cloud computing system, a wide area network (WAN) (e.g., the Internet, an enterprise network), a local area network (LAN) (e.g., a network associated with an office, a building, a campus or other relatively small geographic space), a telephone network, a direct connection between two computing devices, a peer-to-peer network, and any combinations thereof. A network, such as network 1230, may employ a wired and/or wireless mode of communication. In general, any network topology may be used.

[0084] Information and data can be displayed through a display 1232. Examples of a display 1232 include, but are not limited to, a cathode ray tube (CRT), a liquid crystal display (LCD), a thin film transistor liquid crystal display (TFT-LCD), an organic liquid crystal display (OLED) such as a passive-matrix OLED (PMOLED) or active-matrix OLED (AMOLED) display, a plasma display, and any combinations thereof. The display 1232 can interface to the processor(s) 1201, memory 1203, and fixed storage 1208, as well as other devices, such as input device(s) 1233, via the bus 1240. The display 1232 is linked to the bus 1240 via a video interface 1222, and transport of data between the display 1232 and the bus 1240 can be controlled via the graphics control 1221. In some embodiments, the display is a video projector. In some embodiments, the display is a head-mounted display (HMD) such as a VR headset. In further embodiments, suitable VR headsets include, by way of non-limiting examples, HTC Vive®, Oculus Rift®, Samsung Gear VR®, Microsoft HoloLens®, Razer OSVR®, FOVE VR®, Zeiss VR One®, Avegant Glyph®, Freefly VR®

headset, and the like. In still further embodiments, the display is a combination of devices such as those disclosed herein.

[0085] In addition to a display 1232, computer system 1200 may include one or more other peripheral output devices 1234 including, but not limited to, an audio speaker, a printer, a storage device, and any combinations thereof. Such peripheral output devices may be connected to the bus 1240 via an output interface 1224. Examples of an output interface 1224 include, but are not limited to, a serial port, a parallel connection, a USB port, a FIREWIRE port, a THUNDERBOLT port, and any combinations thereof.

[0086] In addition or as an alternative, computer system 1200 may provide functionality as a result of logic hardwired or otherwise embodied in a circuit, which may operate in place of or together with software to execute one or more processes or one or more steps of one or more processes described or illustrated herein. Reference to software in this disclosure may encompass logic, and reference to logic may encompass software. Moreover, reference to a computer-readable medium may encompass a circuit (such as an IC) storing software for execution, a circuit embodying logic for execution, or both, where appropriate. The present disclosure encompasses any suitable combination of hardware, software, or both.

[0087] Various illustrative logical blocks, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, computer software, or combinations of both. To clearly illustrate this interchangeability of hardware and software, various illustrative components, blocks, modules, circuits, and steps have been described above generally in terms of their functionality.

[0088] The various illustrative logical blocks, modules, and circuits described in connection with the embodiments disclosed herein may be implemented or performed with a general purpose processor, a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array (FPGA) or other programmable logic device, discrete gate or transistor logic, discrete hardware components, or any combination thereof designed to perform the functions described herein. A general-purpose processor may be a microprocessor, but in the alternative, the processor may be any processor, controller, microcontroller, or state machine. A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality of microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

[0089] The steps of a method or algorithm described in connection with the embodiments disclosed herein may be embodied directly in hardware, in a software module executed by one or more processor(s), or in a combination of the two. A software module may reside in RAM memory, flash

memory, ROM memory, EPROM memory, EEPROM memory, registers, hard disk, a removable disk, CD-ROM, or any other form of storage medium. An exemplary storage medium is coupled to the processor such the processor can read information from and write information to, the storage medium. In the alternative, the storage medium may be integral to the processor. The processor and the storage medium may reside in an ASIC. The ASIC may reside in a user terminal. In the alternative, the processor and the storage medium may reside as discrete components in a user terminal.

[0090] In accordance with the description herein, suitable computing devices include, by way of non-limiting examples, server computers, desktop computers, laptop computers, notebook computers, sub-notebook computers, netbook computers, netpad computers, set-top computers, media streaming devices, handheld computers, Internet appliances, mobile smartphones, tablet computers, personal digital assistants, video game consoles, and vehicles. Select televisions, video players, and digital music players with optional computer network connectivity are suitable for use in the system described herein. Suitable tablet computers, in various embodiments, include those with booklet, slate, and convertible configurations.

[0091] In some embodiments, the computing device includes an operating system configured to perform executable instructions. The operating system is, for example, software, including **programs and data, which manages the device's hardware and provides services for execution of applications.** Suitable server operating systems include, by way of non-limiting examples, FreeBSD®, OpenBSD®, NetBSD®, Linux®, Apple® Mac OS X Server®, Oracle Solaris®, Windows Server®, and Novell NetWare®. Suitable personal computer operating systems include, by way of non-limiting examples, Microsoft Windows®, Apple Mac® OS X, UNIX®, and UNIX-like operating systems such as GNU/Linux®. In some embodiments, the operating system is provided by cloud computing. Suitable mobile smartphone operating systems include, by way of non-limiting examples, Nokia Symbian® OS, Apple® iOS, Research In Motion BlackBerry® OS, Google® Android®, Microsoft® Windows Phone® OS, Microsoft® Windows Mobile OS, Linux®, and Palm® WebOS. Suitable media streaming device operating systems include, by way of non-limiting examples, Apple TV®, Roku®, Boxee®, Google TV®, Google Chromecast®, Amazon Fire®, and Samsung® HomeSync®. Suitable video game console operating systems include, by way of non-limiting examples, Sony® PS3®, Sony® PS4®, Microsoft® Xbox 360®, Microsoft Xbox One®, Nintendo Wii®, Nintendo Wii U®, and Ouya®. Suitable virtual reality headset systems include, by way of non-limiting example, Meta Oculus®.

Non-transitory computer readable storage mediums

[0092] In some embodiments, the platforms, systems, media, and methods disclosed herein include one or more non-transitory computer readable storage media encoded with a program including instructions executable by the operating system of an optionally networked computing device. In further embodiments, a computer readable storage medium is a tangible component of a computing device. In still further embodiments, a computer readable storage medium is optionally removable from a computing device. In some embodiments, a computer readable storage medium includes, by way of non-limiting examples, CD-ROMs, DVDs, flash memory devices, solid state memory, magnetic disk drives, magnetic tape drives, optical disk drives, distributed computing systems including cloud computing systems and services, and the like. In some cases, the program and instructions are permanently, substantially permanently, semi-permanently, or non-transitorily encoded on the media.

Computer programs

[0093] In some embodiments, the platforms, systems, media, and methods disclosed herein include at least one computer program, or use of the same. A computer program includes a sequence of instructions, executable by one or more processor(s) of the computing device's CPU, written to perform a specified task. Computer readable instructions may be implemented as program modules, such as functions, objects, Application Programming Interfaces (APIs), computing data structures, and the like, that perform particular tasks or implement particular abstract data types. In light of the disclosure provided herein, a computer program may be written in various versions of various languages.

[0094] The functionality of the computer readable instructions may be combined or distributed as desired in various environments. In some embodiments, a computer program comprises one sequence of instructions. In some embodiments, a computer program comprises a plurality of sequences of instructions. In some embodiments, a computer program is provided from one location. In other embodiments, a computer program is provided from a plurality of locations. In various embodiments, a computer program includes one or more software modules. In various embodiments, a computer program includes, in part or in whole, one or more web applications, one or more mobile applications, one or more standalone applications, one or more web browser plug-ins, extensions, add-ins, or add-ons, or combinations thereof.

Targeted therapies

[0095] Table 3 lists examples of targeted therapy drugs for different types of cancers that may be used with methods herein.

Table 3

Type of Cancer	Drug (Example Brand Name)
bladder cancer	atezolizumab (Tecentriq [®]), enfortumab vedotin-ejfv (Padcev [®]), erdafitinib (Balversa [®]), nivolumab (Opdivo [®]), avelumab (Bavencio [®]), pembrolizumab (Keytruda [®]), sacituzumab govitecan-hziy (Trodelvy [®])
brain cancer	abemaciclib (Verzenio [®]), ado-trastuzumab emtansine (Kadcyla [®]), alpelisib (Piqray [®]), anastrozole (Arimidex [®]), capivasertib (Truqap [®]), elacestrant dihydrochloride (Orserdu [®]), everolimus (Afinitor [®]), exemestane (Aromasin [®]), fam-trastuzumab deruxtecan-nxki (Enhertu [®]), fulvestrant (Faslodex [®]), lapatinib ditosylate (Tykerb [®]), letrozole (Femara [®]), margetuximab-cmkb (Margenza [®]), neratinib maleate (Nerlynx [®]), olaparib (Lynparza [®]), palbociclib (Ibrance [®]), pembrolizumab (Keytruda [®]), pertuzumab (Perjeta [®]), pertuzumab, trastuzumab, and hyaluronidase-zzxf (Phesgo [®]), ribociclib (Kisqali [®]), sacituzumab govitecan-hziy (Trodelvy [®]), talazoparib tosylate (Talzenna [®]), tamoxifen citrate (Soltamox [®]), toremifene (Fareston [®]), trastuzumab (Herceptin [®]), tucatinib (Tukysa [®])
cervical cancer	bevacizumab (Avastin [®]), pembrolizumab (Keytruda [®]), tisotumab vedotin-tftv (Tivdak [®])
colorectal cancer	bevacizumab (Avastin [®]), cetuximab (Erbix [®]), encorafenib (Braftovi [®]), fruquintinib (Fruzalqa [®]), ipilimumab (Yervoy [®]), nivolumab (Opdivo [®]), panitumumab (Vectibix [®]), pembrolizumab (Keytruda [®]), ramucirumab (Cyramza [®]), regorafenib (Stivarga [®]), tucatinib (Tukysa [®]), ziv-aflibercept (Zaltrap [®])
dermatofibrosarcoma protuberans	imatinib mesylate (Gleevec [®])
endocrine and neuroendocrine tumors	avelumab (Bavencio [®]), iobenguane I 131 (Azedra [®]), lanreotide acetate (Somatuline Depot [®]), lutetium Lu 177-dotatate (Lutathera [®])
endometrial cancer	bevacizumab (Avastin [®]), pembrolizumab (Keytruda [®]), tisotumab vedotin-tftv (Tivdak [®]),

Type of Cancer	Drug (Example Brand Name)
esophageal cancer	fam-trastuzumab deruxtecan-nxki (Enhertu [®]), ipilimumab (Yervoy [®]), nivolumab (Opdivo [®]), pembrolizumab (Keytruda [®]), ramucirumab (Cyramza [®]), tislelizumab-jsgr (Tevimbra [®]), trastuzumab (Herceptin [®])
head and neck cancer	cetuximab (Erbix [®]), nivolumab (Opdivo [®]), pembrolizumab (Keytruda [®]), toripalimab-tpzi (Loqtorzi [®]),
gastrointestinal stromal tumor	avapritinib (Ayvakit [®]), imatinib mesylate (Gleevec [®]), regorafenib (Stivarga [®]), ripretinib (Qinlock [®]), sunitinib malate (Sutent [®])
giant cell tumor	denosumab (Xgeva [®]), pexidartinib hydrochloride (Turalio [®])
kidney cancer	avelumab (Bavencio [®]), axitinib (Inlyta [®]), belzutifan (Welireg [®]), bevacizumab (Avastin [®]), cabozantinib-s-malate (Cabometyx [®]), everolimus (Afinitor [®]), ipilimumab (Yervoy [®]), lenvatinib mesylate (Lenvima [®]), nivolumab (Opdivo [®]), pazopanib hydrochloride (Votrient [®]), pembrolizumab (Keytruda [®]), sorafenib tosylate (Nexavar [®]), sunitinib malate (Sutent [®]), temsirolimus (Torisel [®]), tivozanib hydrochloride (Fotivda [®]),
leukemia	acalabrutinib (Calquence [®]), alemtuzumab (Campath [®]), asciminib hydrochloride (Scemblix [®]), avapritinib (Ayvakit [®]), blinatumomab (Blinicyto [®]), bosutinib (Bosulif [®]), brexucabtagene autoleucel (Tecartus [®]), dasatinib (Sprycel [®]), duvelisib (Copiktra [®]), enasidenib mesylate (Idhifa [®]), gemtuzumab ozogamicin (Mylotarg [®]), gilteritinib fumarate (Xospata [®]), glasdegib maleate (Daurismo [®]), ibrutinib (Imbruvica [®]), idelalisib (Zydelig [®]), imatinib mesylate (Gleevec [®]), inotuzumab ozogamicin (Besponsa [®]), ivosidenib (Tibsovo [®]), lisocabtagene maraleucel (Breyanzi [®]), midostaurin (Rydapt [®]), moxetumomab pasudotox-tdfk (Lumoxiti [®]), nilotinib (Tasigna [®]), obinutuzumab (Gazyva [®]), ofatumumab (Arzerra [®]), olutasidenib (Rezlidhia [®]),
liver and bile duct cancer	atezolizumab (Tecentriq [®]), bevacizumab (Avastin [®]), cabozantinib-s-malate (Cabometyx [®]), durvalumab (Imfinzi [®]), futibatinib (Lytgobi [®]), ipilimumab (Yervoy [®]), ivosidenib (Tibsovo [®]), lenvatinib mesylate (Lenvima [®]), nivolumab

Type of Cancer	Drug (Example Brand Name)
	(Opdivo [®]), pembrolizumab (Keytruda [®]), pemigatinib (Pemazyre [®]), ramucirumab (Cyramza [®]), regorafenib (Stivarga [®]), sorafenib tosylate (Nexavar [®]), tremelimumab-actl (Imjudo [®])
lung cancer	adagrasib (Krazati [®]), afatinib dimaleate (Gilotrif [®]), alectinib (Alecensa [®]), amivantamab-vmjw (Rybrevant [®]), atezolizumab (Tecentriq [®]), bevacizumab (Avastin [®]), binimetinib (Mektovi [®]), brigatinib (Alunbrig [®]), capmatinib hydrochloride (Tabrecta [®]), cemiplimab-rwlc (Libtayo [®]), ceritinib (Zykadia [®]), crizotinib (Xalkori [®]), dabrafenib mesylate (Tafinlar [®]), dacomitinib (Vizimpro [®]), durvalumab (Imfinzi [®]), encorafenib (Braftovi [®]), entrectinib (Rozlytrek [®]), erlotinib hydrochloride (Tarceva [®]), fam-trastuzumab deruxtecan-nxki (Enhertu [®]), gefitinib (Iressa [®]), ipilimumab (Yervoy [®]), lorlatinib (Lorbrena [®]), necitumumab (Portrazza [®]), nivolumab (Opdivo [®]), osimertinib mesylate (Tagrisso [®]), pembrolizumab (Keytruda [®]), pralsetinib (Gavreto [®]), ramucirumab (Cyramza [®]), repotrectinib (Augtyro [®]), selpercatinib (Retevmo [®]), sotorasib (Lumakras [®]), tepotinib hydrochloride (Tepmetko [®]), trametinib dimethyl sulfoxide (Mekinist [®]), tremelimumab-actl (Imjudo [®])
lymphoma	acalabrutinib (Calquence [®]), axicabtagene ciloleucel (Yescarta [®]), belinostat (Beleodaq [®]), bexarotene (Targretin [®]), bortezomib (Velcade [®]), brentuximab vedotin (Adcetris [®]), brexucabtagene autoleucel (Tecartus [®]), crizotinib (Xalkori [®]), denileukin diftitox (Ontak [®]), duvelisib (Copiktra [®]), epcoritamab-bysp (Epkincy [®]), glofitamab-gxbm (Columvi [®]), ibritumomab tiuxetan (Zevalin [®]), ibrutinib (Imbruvica [®]), lisocabtagene maraleucel (Breyanzi [®]), loncastuximab tesirine-lpyl (Zynlonta [®]), mogamulizumab-kpkc (Poteligeo [®]), mosunetuzumab-axgb (Lunsumio [®]), nivolumab (Opdivo [®]), obinutuzumab (Gazyva [®]), pembrolizumab (Keytruda [®]), pemigatinib (Pemazyre [®]), pirtobrutinib (Jaypirca [®]), polatuzumab vedotin-piiq (Polivy [®]), pralatrexate (Foloty [®]), rituximab (Rituxan [®]), rituximab and hyaluronidase human

Type of Cancer	Drug (Example Brand Name)
	(Rituxan Hycela [®]), romidepsin (Istodax [®]), selinexor (Xpovio [®]), siltuximab (Sylvant [®]), tafasitamab-cxix (Monjuvi [®]), tazemetostat hydrobromide (Tazverik [®]), tisagenlecleucel (Kymriah [®]), venetoclax (Venclexta [®]), vorinostat (Zolinza [®]), zanubrutinib (Brukinsa [®])
malignant mesothelioma	ipilimumab (Yervoy [®]), nivolumab (Opdivo [®])
multiple myeloma	bortezomib (Velcade [®]), carfilzomib (Kyprolis [®]), ciltacabtagene autoleucel (Carvykti [®]), daratumumab (Darzalex [®]), daratumumab and hyaluronidase-fihj (Darzalex Faspro [®]), elranatamab-bcmm (Elrexio [®]), elotuzumab (Empliciti [®]), idecabtagene vicleucel (Abecma [®]), isatuximab-irfc (Sarclisa [®]), ixazomib citrate (Ninlaro [®]), talquetamab-tgvs (Talvey [®]), selinexor (Xpovio [®]), teclistamab-cqyv (Tecvayli [®])
myelodysplastic and myeloproliferative disorders	fedratinib hydrochloride (Inrebic [®]), imatinib mesylate (Gleevec [®]), ivosidenib (Tibsovo [®]), momelotinib dihydrochloride monohydrate (Ojjaara [®]), pacritinib citrate (Vonjo [®]), pemigatinib (Pemazyre [®]), ruxolitinib phosphate (Jakafi [®])
neuroblastoma	dinutuximab (Unituxin [®]), naxitamab-gqgk (Danyelza [®])
ovarian epithelial, fallopian tube, and primary peritoneal cancers	bevacizumab (Avastin [®]), mirvetuximab soravtansine-gynx (Elahere [®]), niraparib tosylate monohydrate (Zejula [®]), olaparib (Lynparza [®]), rucaparib camsylate (Rubraca [®])
pancreatic cancer	belzutifan (Welireg [®]), erlotinib hydrochloride (Tarceva [®]), everolimus (Afinitor [®]), olaparib (Lynparza [®]), sunitinib malate (Sutent [®])
plexiform neurofibroma	selumetinib sulfate (Koselugo [®])
prostate cancer	abiraterone acetate (Zytiga [®]), apalutamide (Erleada [®]), cabazitaxel (Jevtana [®]), darolutamide (Nubeqa [®]), enzalutamide (Xtandi [®]), lutetium Lu 177 vipivotide tetraxetan (Pluvicto [®]), niraparib tosylate monohydrate and abiraterone acetate (Akeega [®]), olaparib (Lynparza [®]), talazoparib tosylate (Talzenna [®]), radium 223 dichloride (Xofigo [®]), rucaparib camsylate (Rubraca [®])

Type of Cancer	Drug (Example Brand Name)
skin cancer	alitretinoin (Panretin [®]), atezolizumab (Tecentriq [®]), avelumab (Bavencio [®]), binimetinib (Mektovi [®]), cemiplimab-rwlc (Libtayo [®]), cobimetinib fumarate (Cotellic [®]), dabrafenib mesylate (Tafinlar [®]), encorafenib (Braftovi [®]), ipilimumab (Yervoy [®]), nivolumab (Opdivo [®]), nivolumab and relatlimab-rmbw (Opdualag [®]), pembrolizumab (Keytruda [®]), retifanlimab-dlwr (Zynyz [®]), sonidegib (Odomzo [®]), tebentafusp-tebn (Kimmtrak [®]), trametinib dimethyl sulfoxide (Mekinist [®]), vismodegib (Erivedge [®]), vemurafenib (Zelboraf [®])
soft tissue sarcoma	alitretinoin (Panretin [®]), atezolizumab (Tecentriq [®]), crizotinib (Xalkori [®]), nirogacestat hydrobromide (Ogsiveo [®]), pazopanib hydrochloride (Votrient [®]), sirolimus protein-bound particles (Fyarro [®]), tazemetostat hydrobromide (Tazverik [®])
solid tumors	dabrafenib mesylate (Tafinlar [®]), dostarlimab-gxly (Jemperli [®]), entrectinib (Rozlytrek [®]), larotrectinib sulfate (Vitrakvi [®]), pembrolizumab (Keytruda [®]), selpercatinib (Retevmo [®]), trametinib dimethyl sulfoxide (Mekinist [®])
stomach (gastric) cancer	fam-trastuzumab deruxtecan-nxki (Enhertu [®]), nivolumab (Opdivo [®]), pembrolizumab (Keytruda [®]), ramucirumab (Cyramza [®]), trastuzumab (Herceptin [®])
systemic mastocytosis	avapritinib (Ayvakit [®]), imatinib mesylate (Gleevec [®]), midostaurin (Rydapt [®])
thyroid cancer	cabozantinib-s-malate (Cometriq [®]), dabrafenib mesylate (Tafinlar [®]), lenvatinib mesylate (Lenvima [®]), pralsetinib (Gavreto [®]), selpercatinib (Retevmo [®]), sorafenib tosylate (Nexavar [®]), trametinib dimethyl sulfoxide (Mekinist [®]), vandetanib (Caprelsa [®])

References

[0096] 1. Conte F, Fiscon G, Licursi V, Bizzarri D, D'Anto T, Farina L, et al. A paradigm shift in medicine: A comprehensive review of network-based approaches. Biochimica et Biophysica Acta

(BBA) - Gene Regulatory Mechanisms. 2020 Jun 1;1863(6):194416, which is incorporated by reference herein in its entirety.

[0097] 2. Vafaee F, Diakos C, Kirschner MB, Reid G, Michael MZ, Horvath LG, et al. A data-driven, knowledge-based approach to biomarker discovery: application to circulating microRNA markers of colorectal cancer prognosis. *npj Syst Biol Appl*. 2018 Jun 1;4(1):1-12, which is incorporated by reference herein in its entirety.

[0098] 3. Cheng F, Kovacs IA, Barabasi AL. Network-based prediction of drug combinations. *Nat Commun*. 2019 Mar 13;10:1197, which is incorporated by reference herein in its entirety.

[0099] 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 May;39(3):1205-27, which is incorporated by reference herein in its entirety.

[0100] 5. Tabernero J, Lenz HJ, Siena S, Sobrero A, Falcone A, Ychou M, et al. 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 Aug;16(8):937-48, which is incorporated by reference herein in its entirety.

[0101] 6. Verdaguer H, Saun T, Macarulla T. Predictive and prognostic biomarkers in personalized gastrointestinal cancer treatment. *J Gastrointest Oncol*. 2017 Jun;8(3):405-17, which is incorporated by reference herein in its entirety.

[0102] 7. Tejpar S, Bertagnolli M, Bosman F, Lenz HJ, Garraway L, Waldman F, et al. Prognostic and Predictive Biomarkers in Resected Colon Cancer: Current Status and Future Perspectives for Integrating Genomics into Biomarker Discovery. *Oncologist*. 2010 Apr;15(4):390-404, which is incorporated by reference herein in its entirety.

[0103] 8. Li J, Lei K, Wu Z, Li W, Liu G, Liu J, et al. Network-based identification of microRNAs as potential pharmacogenomic biomarkers for anticancer drugs. *Oncotarget* [Internet]. 2016 [cited 2023 Jun 5]; Available from: <https://dash.harvard.edu/handle/1/30371077>, which is incorporated by reference herein in its entirety.

[0104] 9. Kong J, Ha D, Lee J, Kim I, Park M, Im SH, et al. Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nat Commun*. 2022 Jun 28;13:3703, which is incorporated by reference herein in its entirety.

[0105] 10. Amin S, Bathe OF. Response biomarkers: re-envisioning the approach to tailoring drug therapy for cancer. *BMC Cancer*. 2016 Nov 5;16:850, which is incorporated by reference herein in its entirety.

- [0106] 11. Glaab E. Using prior knowledge from cellular pathways and molecular networks for diagnostic specimen classification. *Brief Bioinform.* 2016 May;17(3):440-52, which is incorporated by reference herein in its entirety.
- [0107] 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 Dec;17(12):710-30, which is incorporated by reference herein in its entirety.
- [0108] 13. Taylor JMG, Ankerst DP, Andridge RR. Validation of Biomarker-based risk prediction models. *Clin Cancer Res.* 2008 Oct 1;14(19):5977-83, which is incorporated by reference herein in its entirety.
- [0109] 14. Archer R, Hock E, Hamilton J, Stevens J, Essat M, Poku E, et al. Assessing prognosis and prediction of treatment response in early rheumatoid arthritis: systematic reviews. *Health Technol Assess.* 2018 Nov;22(66):1-294, which is incorporated by reference herein in its entirety.
- [0110] 15. Diaz-Uriarte R, Gomez de Lope E, Giugno R, Frohlich H, Nazarov PV, Nepomuceno-Chamorro IA, et al. Ten quick tips for biomarker discovery and validation analyses using machine learning. *PLoS Comput Biol.* 2022 Aug 11;18(8):e1010357, which is incorporated by reference herein in its entirety.
- [0111] 16. Choobdar S, Ahsen ME, Crawford J, Tomasoni M, Fang T, Lamparter D, et al. Assessment of network module identification across complex diseases. *Nat Methods.* 2019 Sep;16(9):843-52, which is incorporated by reference herein in its entirety.
- [0112] 17. Strunz S, Wolkenhauer O, de la Fuente A. Network-Assisted Disease Classification and Biomarker Discovery. *Methods Mol Biol.* 2016;1386:353-74, which is incorporated by reference herein in its entirety.
- [0113] 18. Toro-Dominguez D, Alarcon-Riquelme ME. Precision medicine in autoimmune diseases: fact or fiction. *Rheumatology.* 2021 Sep 1;60(9):3977-85, which is incorporated by reference herein in its entirety.
- [0114] 19. Menche J, Sharma A, Kitsak M, Ghiassian S, Vidal M, Loscalzo J, et al. Uncovering disease-disease relationships through the incomplete human interactome. *Science.* 2015 Feb 20;347(6224):1257601, which is incorporated by reference herein in its entirety.
- [0115] 20. Guthridge JM, Wagner CA, James JA. The promise of precision medicine in rheumatology. *Nat Med.* 2022 Jul;28(7):1363-71, which is incorporated by reference herein in its entirety.
- [0116] 21. Lin CMA, Cooles FAH, Isaacs JD. Precision medicine: the precision gap in rheumatic disease. *Nat Rev Rheumatol.* 2022 Dec;18(12):725-33, which is incorporated by reference herein in its entirety.

- [0117] 22. Zhao SS, Moots RJ. Biomarkers for Treatment Response in Rheumatoid Arthritis: Where are they? *Rheumatol Immunol Res.* 2020 Dec 1;1(1):1-3, which is incorporated by reference herein in its entirety.
- [0118] 23. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002 Jan 1;30(1):207-10, which is incorporated by reference herein in its entirety.
- [0119] 24. Xu J, Murphy SL, Kochanek KD, Arias E. Mortality in the United States, 2021. *NCHS Data Brief.* 2022 Dec;(456):1-8, which is incorporated by reference herein in its entirety.
- [0120] 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 Aug;246:78-86, which is incorporated by reference herein in its entirety.
- [0121] 26. Kokkonen H, Soderstrom I, Rocklov J, Hallmans G, Lejon K, Rantapaa Dahlqvist S. Up-regulation of cytokines and chemokines predates the onset of rheumatoid arthritis. *Arthritis & Rheumatism.* 2010;62(2):383-91, which is incorporated by reference herein in its entirety.
- [0122] 27. Brink M, Lundquist A, Alexeyenko A, Lejon K, Rantapaa-Dahlqvist S. Protein profiling and network enrichment analysis in individuals before and after the onset of rheumatoid arthritis. *Arthritis Res Ther.* 2019;21:288, which is incorporated by reference herein in its entirety.
- [0123] 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 Dec;76(2):527-36, which is incorporated by reference herein in its entirety.
- [0124] 29. Mellors T, Withers JB, Ameli A, Jones A, Wang M, Zhang L, et al. Clinical Validation of a Blood-Based Predictive Test for Stratification of Response to Tumor Necrosis Factor Inhibitor Therapies in Rheumatoid Arthritis Patients. *Network and Systems Medicine.* 2020 Aug;3(1):91-104, which is incorporated by reference herein in its entirety.
- [0125] 30. Koh K, Kim SJ, Boyd S. A Method for Large-Scale L1-Regularized Logistic Regression. *AAAI.* 2007 Jul 22;565-71, which is incorporated by reference herein in its entirety.
- [0126] 31. Curtis JR, Strand V, Golombek S, Zhang L, Wong A, Zielinski MC, et al. Patient outcomes improve when a molecular signature test guides treatment decision-making in rheumatoid arthritis. *Expert Rev Mol Diagn.* 2022 Nov 3;22:1-10, which is incorporated by reference herein in its entirety.
- [0127] 32. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet.* 2016 Oct 22;388(10055):2023-38, which is incorporated by reference herein in its entirety.

- [0128] 33. Cohen S, Wells AF, Curtis JR, Dhar R, Mellors T, Zhang L, et al. A Molecular Signature Response Classifier to Predict Inadequate Response to Tumor Necrosis Factor- α Inhibitors: The NETWORK-004 Prospective Observational Study. *Rheumatol Ther*. 2021 Jun 19;8(3):1159-76, which is incorporated by reference herein in its entirety.
- [0129] 34. Nakase H, Sato N, Mizuno N, Ikawa Y. The influence of cytokines on the complex pathology of ulcerative colitis. *Autoimmun Rev*. 2022 Mar;21(3):103017, which is incorporated by reference herein in its entirety.
- [0130] 35. Friedrich M, Pohin M, Powrie F. Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease. *Immunity*. 2019 Apr 16;50(4):992-1006, which is incorporated by reference herein in its entirety.
- [0131] 36. Kondo N, Kuroda T, Kobayashi D. Cytokine Networks in the Pathogenesis of Rheumatoid Arthritis. *Int J Mol Sci*. 2021 Oct 10;22(20):10922, which is incorporated by reference herein in its entirety.
- [0132] 37. 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 Apr 1;455:161-71, which is incorporated by reference herein in its entirety.
- [0133] 38. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013 Jan 1;29(1):15-21, which is incorporated by reference herein in its entirety.
- [0134] 39. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics*. 2015 Sep 3;51:11.14.1-11.14.19, which is incorporated by reference herein in its entirety.
- [0135] 40. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011 Aug 4;12:323, which is incorporated by reference herein in its entirety.
- [0136] 41. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020 Mar;17(3):261-72, which is incorporated by reference herein in its entirety.
- [0137] 42. Kitsak M, Sharma A, Menche J, Guney E, Ghiassian SD, Loscalzo J, et al. Tissue Specificity of Human Disease Module. *Sci Rep*. 2016 Oct 17;6:35241, which is incorporated by reference herein in its entirety.
- [0138] 43. Prasad B, McGeough C, Eakin A, Ahmed T, Small D, Gardiner P, et al. ATRPred: A machine learning based tool for clinical decision making of anti-TNF treatment in rheumatoid

arthritis patients. PLoS Comput Biol. 2022 Jul 5;18(7):e1010204, which is incorporated by reference herein in its entirety.

[0139] 44. Lerner A, Neidhofer 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 Aug;32(4):550-62, which is incorporated by reference herein in its entirety.

[0140] 45. Ivan G, Grolmusz V. When the Web meets the cell: using personalized PageRank for analyzing protein interaction networks. Bioinformatics. 2011 Feb 1;27(3):405-7, which is incorporated by reference herein in its entirety.

[0141] 46. 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 Aug 27;573(1-3):83-92, which is incorporated by reference herein in its entirety.

[0142] 47. Roda G, Jharap B, Neeraj N, Colombel JF. Loss of Response to Anti-TNFs: Definition, Epidemiology, and Management. Clin Transl Gastroenterol. 2016 Jan 7;7(1):e135, which is incorporated by reference herein in its entirety.

[0143] 48. Voitalov I, Zhang L, Kilpatrick C, Withers JB, Saleh A, Akmaev VR, et al. The module triad: a novel network biology approach to utilize patients' multi-omics data for target discovery in ulcerative colitis. Sci Rep. 2022 Dec 15;12(1):21685, which is incorporated by reference herein in its entirety.

[0144] 49. Arijs I, Li K, Toedter G, Quintens R, Van Lommel L, Van Steen K, et al. Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. Gut. 2009 Dec;58(12):1612-9, which is incorporated by reference herein in its entirety.

[0145] 50. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019 Apr 3;10(1):1523, which is incorporated by reference herein in its entirety.

[0146] 51. Lucas RM, Luo L, Stow JL. ERK1/2 in immune signalling. Biochem Soc Trans. 2022 Oct 31;50(5):1341-52, which is incorporated by reference herein in its entirety.

[0147] 52. 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 Jul 16;25(7):1637-47, which is incorporated by reference herein in its entirety.

[0148] 53. Hua C, Buttgerit F, Combe B. Glucocorticoids in rheumatoid arthritis: current status and future studies. RMD Open. 2020 Jan;6(1):e000536, which is incorporated by reference herein in its entirety.

[0149] 54. Bruscoli S, Febo M, Riccardi C, Migliorati G. Glucocorticoid Therapy in Inflammatory Bowel Disease: Mechanisms and Clinical Practice. *Front Immunol.* 2021;12:691480, which is incorporated by reference herein in its entirety.

[0150] 55. Van Bogaert T, De Bosscher K, Libert C. Crosstalk between TNF and glucocorticoid receptor signaling pathways. *Cytokine Growth Factor Rev.* 2010 Aug;21(4):275-86, which is incorporated by reference herein in its entirety.

[0151] 56. Dendoncker K, Timmermans S, Vandewalle J, Eggermont M, Lempiainen J, Paakinaho V, et al. TNF- α inhibits glucocorticoid receptor-induced gene expression by reshaping the GR nuclear cofactor profile. *Proc Natl Acad Sci U S A.* 2019 Jun 25;116(26):12942-51, which is incorporated by reference herein in its entirety.

[0152] 57. Hobson AD, McPherson MJ, Hayes ME, Goess C, Li X, Zhou J, et al. Discovery of ABBV-3373, an Anti-TNF Glucocorticoid Receptor Modulator Immunology Antibody Drug Conjugate. *J Med Chem.* 2022 Dec 8;65(23):15893-934, which is incorporated by reference herein in its entirety.

[0153] 58. Bayat Mokhtari R, Homayouni TS, Baluch N, Morgatskaya E, Kumar S, Das B, et al. Combination therapy in combating cancer. *Oncotarget.* 2017 Jun 6;8(23):38022-43, which is incorporated by reference herein in its entirety.

Terms and Definitions

[0154] Unless otherwise defined, all technical terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this present disclosure belongs.

[0155] As used herein, the singular forms “a,” “an,” and “the” include plural references unless the context clearly dictates otherwise. Any reference to “or” herein is intended to encompass “and/or” unless otherwise stated.

[0156] As used herein, the term “about” in some cases refers to an amount that is approximately the stated amount.

[0157] As used herein, the term “about” refers to an amount that is near the stated amount by 10%, 5%, or 1%, including increments therein.

[0158] As used herein, the term “about” in reference to a percentage refers to an amount that is greater or less the stated percentage by 10%, 5%, or 1%, including increments therein.

[0159] As used herein, the phrases “at least one”, “one or more”, and “and/or” are open-ended expressions that are both conjunctive and disjunctive in operation. For example, each of the expressions “at least one of A, B and C”, “at least one of A, B, or C”, “one or more of A, B, and C”, “one or more of A, B, or C” and “A, B, and/or C” means A alone, B alone, C alone, A and B together, A and C together, B and C together, or A, B and C together.

[0160] As used herein, in some cases, the terms “immunology disease”, “autoimmune disease”, and “inflammatory disease” can be used interchangeably. In some cases, the terms “immunology disease”, “autoimmune disease”, and “inflammatory disease” can generally refer to different diseases based on how the disease affects the body’s immune system. For example, autoimmune diseases can occur when the body’s immune system wrongly attacks normal, healthy cells and mistakes them for harmful invaders. Immunology diseases, sometimes also called immunodeficiencies, can occur when the body’s immune system is not strong enough and cannot adequately protect the body against infection.

Non-exhaustive list of abbreviations

- [0161] ACR: American College of Rheumatology
- [0162] AD: autoimmune disorders
- [0163] AIMS: Accelerate Information of Molecular Signatures
- [0164] AUCROC: area under the curve (AUC) of receiver operating characteristic (ROC)
- [0165] AUPRC: area under the precision recall curve
- [0166] CD: Crohn’s disease
- [0167] CMRL: Connaught Medical Research Laboratory
- [0168] CV: cross-validation
- [0169] DE: differentially expressed
- [0170] ERK: extracellular signal-regulated kinase 1 and 2
- [0171] GM-CSF: granulocyte-macrophage colony-stimulating factor
- [0172] HI: human interactome
- [0173] IFN γ (or IFNG): interferon gamma
- [0174] IL: interleukin
- [0175] IMID: immune mediated inflammatory disease
- [0176] MSRC: Molecular Signature Response Classifier
- [0177] NR: non-responder or nonresponder
- [0178] R: responder
- [0179] RA: rheumatoid arthritis
- [0180] SEB: Staphylococcus enterotoxin B
- [0181] SNR: signal-to-noise ratio
- [0182] TNF α : tumor necrosis factor alpha
- [0183] TNFi: tumor necrosis factor inhibitor
- [0184] UC: ulcerative colitis
- [0185] UMAP: uniform manifold approximation and projection

[0186] While preferred embodiments of the present disclosure have been shown and described herein, such embodiments are provided by way of example only. It is not intended that the present disclosure be limited by the specific examples provided within the specification. While the present disclosure has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions may occur without departing from the present disclosure. Furthermore, it shall be understood that all aspects of the present disclosure are not limited to the specific depictions, configurations, or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the present disclosure described herein may be employed in practicing the present disclosure. It is therefore contemplated that the present disclosure shall also cover any such alternatives, modifications, variations, or equivalents. It is intended that the following claims define the scope of the present disclosure and that systems, methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS**WHAT IS CLAIMED IS:**

1. A method of treating a subject suffering from a disease or disorder, the method comprising:
administering to the subject a therapy,
wherein the therapy has been predicted to generate a treatment response in the subject based at least in part on a trained machine learning (ML) classifier that classifies the subject as responsive or non-responsive to the therapy,
wherein the trained ML classifier predicts the treatment response based at least in part on analyzing in the subject a set of cytokines or a set of transcripts.
2. The method of claim 1, further comprising:
identifying one or more source nodes associated with a set of proteins targeted by the therapy;
identifying one or more response nodes associated with the set of cytokines or the set of transcripts indicative of the treatment response;
determining, by a network-based diffusion process, one or more intermediate nodes associated with the one or more source nodes or the one or more response nodes, wherein the one or more intermediate nodes are indicative of the treatment response;
generating a ranked set of the one or more intermediate nodes, wherein the ranked set is based at least in part on one or more interactions of the intermediate nodes with the one or more source nodes or the one or more response nodes; and
processing a subset of the ranked set of intermediate nodes to determine a set of biomarkers that are predictive of the treatment response to the therapy.
3. The method of claim 2, wherein the ranked set of intermediate nodes comprises a ranked set of proteins, a ranked set of cytokines, or a ranked set of transcripts.
4. The method of claim 3, wherein the ranked set of proteins, the ranked set of cytokines, or the ranked set of transcripts is used as features to obtain the trained ML classifier.
5. The method of claim 2, wherein the set of biomarkers is predictive of the treatment response to the therapy.
6. The method of claim 3, wherein (i) the ranked set of cytokines is different from at least one cytokine of the set of cytokines or (ii) the ranked set of cytokines is the same as the set of cytokines.

7. The method of claim 3, wherein (i) the ranked set of transcripts is different from at least one transcript of the set of transcripts or (ii) the ranked set of transcripts is the same as the set of transcripts.
8. The method of claim 3, wherein the ranked set of cytokines or the ranked set of transcripts is analyzed to determine statistically significant changes after administering to the subject the therapy.
9. The method of claim 3, wherein generating the ranked set of the one or more intermediate nodes comprises:
 - determining a first rank between the one or more source nodes and the one or more intermediate nodes;
 - determining a second rank between the one or more response nodes and the one or more intermediate nodes; and
 - determining a rank product of the first rank and the second rank.
10. The method of claim 2, further comprising:
 - constructing a subset of a protein-protein network associated with the disease or disorder, wherein the constructing comprises:
 - (i) determining the one or more source nodes in the protein-protein network, based at least in part on the set of proteins targeted by the therapy;
 - (ii) determining the one or more response nodes in the protein-protein network, based on least in part on the set of cytokines or the set of transcripts indicative of the treatment response; and
 - (iii) determining one or more connections between the one or more source nodes and the one or more response nodes, based at least in part on one or more interactions between the one or more source nodes and the one or more response nodes.
11. The method of claim 10, wherein the protein-protein network comprises a human interactome (HI).
12. The method of claim 1, wherein the disease or disorder comprises an autoimmune disease, an immunology disease, or a cancer.
13. The method of claim 12, wherein the immunology disease or the autoimmune disease comprises **rheumatoid arthritis (RA), ulcerative colitis (UC), Crohn's disease (CD),** inflammatory bowel disease, psoriatic arthritis, ankylosing spondylitis, chronic psoriasis, hidradenitis suppurativa, multiple sclerosis, juvenile idiopathic arthritis, uveitis, systemic lupus, Type 1 diabetes, Graves' Disease, Hashimoto's Thyroiditis, celiac disease, **Sjögren's Syndrome**, or an immune-mediated disease.

14. The method of claim 13, wherein the immunology disease or the autoimmune disease comprises RA.
15. The method of claim 13, wherein the immunology disease or the autoimmune disease comprises UC.
16. The method of claim 13, wherein the immunology disease or the autoimmune disease comprises CD.
17. The method of claim 12, wherein the therapy for the immunology disease or the autoimmune disease comprises an anti-TNF therapy.
18. The method of claim 17, wherein the anti-TNF therapy comprises infliximab, adalimumab, etanercept, certolizumab pegol, or golimumab, or a biosimilar thereof.
19. The method of claim 12, wherein the therapy for the immunology disease or the autoimmune disease comprises an alternative to anti-TNF therapy.
20. The method of claim 19, wherein the alternative to anti-TNF therapy comprises anti-CD20, JAK, anti-IL6, rituximab, sarilumab, tofacitinib citrate, leflunomide, vedolizumab, tocilizumab, anakinra, abatacept, TL1A, co-stimulation blockade, or a biosimilar thereof.
21. The method of claim 12, wherein the cancer comprises bladder cancer, breast cancer, colon cancer, rectal cancer, endometrial cancer, kidney cancer, leukemia, liver cancer, lung cancer, melanoma, Non-Hodgkin's lymphoma, pancreatic cancer, prostate cancer, or thyroid cancer.
22. The method of claim 21, wherein the therapy for the cancer comprises a targeted therapy, chemotherapy, hormone therapy, immunotherapy, photodynamic therapy, radiation therapy, stem cell transplant therapy, or any combination thereof.
23. The method of claim 2, wherein the trained ML classifier predicts the treatment response to the therapy using a non-linear relationship between (i) measurements of the ranked set of intermediate nodes and (ii) the treatment response indicated by statistically significant changes in the ranked set of intermediate nodes after administering to the subject the therapy.
24. The method of claim 2, wherein the trained ML classifier is trained using measurements of the one or more intermediate nodes in (i) a first set of subjects with the disease or disorder determined to be responsive to the therapy and (ii) a second set of subjects with the disease or disorder determined to be non-responsive to the therapy.
25. The method of claim 1, wherein the trained ML classifier comprises a neural network, deep learning, perceptron, random forest, Bayes, Markov, Gaussian process, clustering algorithm, support vector machine, generative model, or kernel.
26. The method of claim 3, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to an immunology disease or an autoimmune disease comprises:

CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL-12A/B, IL12B, IL13, IL15, IL16, IL17A, IL18, IL1B, IL2, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, IL8, MIF, TGFB1, TGFB2, TGFB3, or TNF.

27. The method of claim 3, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL13, IL17A, IL1B, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, TGFB1, or TNF.
28. The method of claim 3, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to rheumatoid arthritis (RA) comprises: CSF2, CXCL8, IFN γ , IL10, IL12A, IL12B, IL15, IL16, IL17A, IL18, IL1B, IL4, IL6, MIF, TGFB1, TGFB2, TGFB3, TNF.
29. The method of claim 3, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to **Crohn's disease (CD)** comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B.
30. The method of claim 3, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B.
31. The method of claim 1, wherein the trained ML classifier predicts the treatment response with an area under receiver operating characteristic curve (AUC) of at least about 0.8.
32. The method of claim 3, wherein the set of cytokines or the ranked set of cytokines is measured using an immunoassay.
33. The method of claim 3, wherein the set of transcripts or the ranked set of transcripts is measured using reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR), microarrays, or RNA-sequencing (RNA-seq).
34. The method of claim 1, wherein the therapy comprises an experimental therapy in a clinical trial.
35. The method of claim 34, wherein the experimental therapy comprises KO-947.
36. A method for determining biomarkers predictive of treatment response to a therapy for a subject suffering from a disease or disorder, the method comprising:
 - identifying one or more source nodes associated with a set of proteins targeted by the therapy;
 - determining, by a network-based diffusion process, one or more intermediate nodes associated with the one or more source nodes or the one or more response nodes, wherein the one or more intermediate nodes are indicative of the treatment response;

generating a ranked set of the one or more intermediate nodes, wherein the ranked set is based at least in part on one or more interactions of the intermediate nodes with the one or more source nodes or the one or more response nodes; and

processing a subset of the ranked set of intermediate nodes to determine a set of biomarkers that are predictive of the treatment response to the therapy.

37. The method of claim 36, wherein the ranked set of intermediate nodes comprises a ranked set of proteins, a ranked set of cytokines, or a ranked set of transcripts.
38. The method of claim 37, wherein the ranked set of proteins, the ranked set of cytokines, or the ranked set of transcripts is used as features to obtain a trained ML classifier.
39. The method of claim 36, wherein the set of biomarkers is predictive of the treatment response to the therapy.
40. The method of claim 37, wherein (i) the ranked set of cytokines is different from at least one cytokine of the set of cytokines or (ii) the ranked set of cytokines is the same as the set of cytokines.
41. The method of claim 37, wherein (i) the ranked set of transcripts is different from at least one transcript of the set of transcripts or (ii) the ranked set of transcripts is the same as the set of transcripts.
42. The method of claim 37, wherein the ranked set of cytokines or the ranked set of transcripts is analyzed to determine statistically significant changes after administering to the subject the therapy.
43. The method of claim 37, wherein generating the ranked set of the one or more intermediate nodes comprises:
 - determining a first rank between the one or more source nodes and the one or more intermediate nodes;
 - determining a second rank between the one or more response nodes and the one or more intermediate nodes; and
 - determining a rank product of the first rank and the second rank.
44. The method of claim 36, further comprising:
 - constructing a subset of a protein-protein network associated with the disease or disorder, wherein the constructing comprises:
 - (i) determining the one or more source nodes in the protein-protein network, based at least in part on the set of proteins targeted by the therapy;

- (ii) determining the one or more response nodes in the protein-protein network, based on least in part on the set of cytokines or the set of transcripts indicative of the treatment response; and
- (iii) determining one or more connections between the one or more source nodes and the one or more response nodes, based at least in part on one or more interactions between the one or more source nodes and the one or more response nodes.
45. The method of claim 44, wherein the protein-protein network comprises a human interactome (HI).
46. The method of claim 36, wherein the disease or disorder comprises an autoimmune disease, an immunology disease, or a cancer.
47. The method of claim 46, wherein the immunology disease or the autoimmune disease comprises **rheumatoid arthritis (RA), ulcerative colitis (UC), Crohn's disease (CD)**, inflammatory bowel disease, psoriatic arthritis, ankylosing spondylitis, chronic psoriasis, hidradenitis suppurativa, multiple sclerosis, juvenile idiopathic arthritis, uveitis, systemic lupus, Type 1 diabetes, Graves' Disease, Hashimoto's Thyroiditis, celiac disease, **Sjögren's Syndrome**, or an immune-mediated disease.
48. The method of claim 47, wherein the immunology disease or the autoimmune disease comprises RA.
49. The method of claim 47, wherein the immunology disease or the autoimmune disease comprises UC.
50. The method of claim 47, wherein the immunology disease or the autoimmune disease comprises CD.
51. The method of claim 46, wherein the therapy for the immunology disease or the autoimmune disease comprises an anti-TNF therapy.
52. The method of claim 51, wherein the anti-TNF therapy comprises infliximab, adalimumab, etanercept, certolizumab pegol, or golimumab, or a biosimilar thereof.
53. The method of claim 46, wherein the therapy for the immunology disease or the autoimmune disease comprises an alternative to anti-TNF therapy.
54. The method of claim 53, wherein the alternative to anti-TNF therapy comprises anti-CD20, JAK, anti-IL6, rituximab, sarilumab, tofacitinib citrate, leflunomide, vedolizumab, tocilizumab, anakinra, abatacept, TL1A, co-stimulation blockade, or a biosimilar thereof.
55. The method of claim 46, wherein the cancer comprises bladder cancer, breast cancer, colon cancer, rectal cancer, endometrial cancer, kidney cancer, leukemia, liver cancer, lung cancer, melanoma, Non-Hodgkin's lymphoma, pancreatic cancer, prostate cancer, or thyroid cancer.

56. The method of claim 55, wherein the therapy for the cancer comprises a targeted therapy, chemotherapy, hormone therapy, immunotherapy, photodynamic therapy, radiation therapy, stem cell transplant therapy, or any combination thereof.
57. The method of claim 38, wherein the trained ML classifier predicts the treatment response to the therapy using a non-linear relationship between (i) measurements of the ranked set of intermediate nodes and (ii) the treatment response indicated by statistically significant changes in the ranked set of intermediate nodes after administering to the subject the therapy.
58. The method of claim 38, wherein the trained ML classifier is trained using measurements of the one or more intermediate nodes in (i) a first set of subjects with the disease or disorder determined to be responsive to the therapy and (ii) a second set of subjects with the disease or disorder determined to be non-responsive to the therapy.
59. The method of claim 38, wherein the trained ML classifier comprises a neural network, deep learning, perceptron, random forest, Bayes, Markov, Gaussian process, clustering algorithm, support vector machine, generative model, or kernel.
60. The method of claim 37, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to an immunology disease or an autoimmune disease comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL-12A/B, IL12B, IL13, IL15, IL16, IL17A, IL18, IL1B, IL2, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, IL8, MIF, TGFB1, TGFB2, TGFB3, or TNF.
61. The method of claim 37, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: CCL2, CSF2, CXCL8, IFN γ , IL10, IL12A, IL13, IL17A, IL1B, IL21, IL22, IL23A, IL27, IL4, IL5, IL6, IL7, TGFB1, or TNF.
62. The method of claim 37, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to rheumatoid arthritis (RA) comprises: CSF2, CXCL8, IFN γ , IL10, IL12A, IL12B, IL15, IL16, IL17A, IL18, IL1B, IL4, IL6, MIF, TGFB1, TGFB2, TGFB3, TNF.
63. The method of claim 37, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to **Crohn's disease (CD)** comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B.
64. The method of claim 37, wherein the set of cytokines or the ranked set of cytokines indicative of the treatment response to ulcerative colitis (UC) comprises: IL10, IFN γ , IL13, IL17A, IL1B, IL2, IL4, IL5, IL6, IL7, IL8, CCL2, TNF, or IL-12A/B.

65. The method of claim 38, wherein the trained ML classifier predicts the treatment response with an area under receiver operating characteristic curve (AUC) of at least about 0.8.
66. The method of claim 37, wherein the set of cytokines or the ranked set of cytokines is measured using an immunoassay.
67. The method of claim 37, wherein the set of transcripts or the ranked set of transcripts is measured using reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR), microarrays, or RNA-sequencing (RNA-seq).
68. The method of claim 36, wherein the therapy comprises an experimental therapy in a clinical trial.
69. A computer program product for determining biomarkers predictive of treatment response to a therapy for a subject suffering from a disease or disorder, the computer program product comprising at least one non-transitory computer-readable medium having computer-readable program code portions embodied therein, the computer-readable program code portions comprising:
- an executable portion configured to identify one or more source nodes associated with a set of proteins targeted by the therapy;
 - an executable portion configured to determine, by a network-based diffusion process, one or more intermediate nodes associated with the one or more source nodes or the one or more response nodes, wherein the one or more intermediate nodes are indicative of the treatment response;
 - an executable portion configured to generate a ranked set of the one or more intermediate nodes, wherein the ranked set is based at least in part on one or more interactions of the intermediate nodes with the one or more source nodes or the one or more response nodes; and
 - an executable portion configured to process a subset of the ranked set of intermediate nodes to determine a set of biomarkers that are predictive of the treatment response to the therapy.
70. A method for administering an experimental therapy to a subject in a clinical trial, comprising:
- administering to the subject the experimental therapy,
 - wherein the experimental therapy has been predicted to generate a treatment response in the subject based at least in part on a trained machine learning (ML) classifier that classifies the subject as responsive or non-responsive to the experimental therapy,
 - wherein the trained ML classifier predicts the treatment response based at least in part on analyzing in the subject a set of cytokines or a set of transcripts.

71. A method for selecting a subject to receive an experimental therapy in a clinical trial, comprising:
- selecting to the subject to receive the experimental therapy,
 - wherein the experimental therapy has been predicted to generate a treatment response in the subject based at least in part on a trained machine learning (ML) classifier that classifies the subject as responsive or non-responsive to the therapy,
 - wherein the trained ML classifier predicts the treatment response based at least in part on analyzing in the subject a set of cytokines or a set of transcripts.