(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 November 2016 (10.11.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/179469 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *A61K 39/395* (2006.01)
C12N 15/12 (2006.01) *A61K 35/38* (2015.01)
C12N 15/52 (2006.01) *G01N 33/53* (2006.01)

(21) International Application Number:
PCT/US2016/031145

(22) International Filing Date:
6 May 2016 (06.05.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/158,419 7 May 2015 (07.05.2015) US

(71) Applicant: **ABBVIE INC.** [US/US]; 1 North Waukegan Road, North Chicago, IL 60064 (US).

(72) Inventors: **LI, Li**; 14 Swan Circle, Shrewsbury, MA 01545 (US). **WANG, Lichun**; 177 Old Westboro Road, North Grafton, MA 01536 (US). **MCRAE, Bradford, L.**; 84 Pleasant Street, Northborough, MA 01532 (US). **DEVANARAYAN, Viswanath**; 32 Knights Crest Court, Souderton, PA 18964 (US).

(74) Agents: **COWLES, Cristin H.** et al.; McCarter & English, LLP, 265 Franklin Street, Boston, MA 02110 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING INFLAMMATORY BOWEL DISEASE

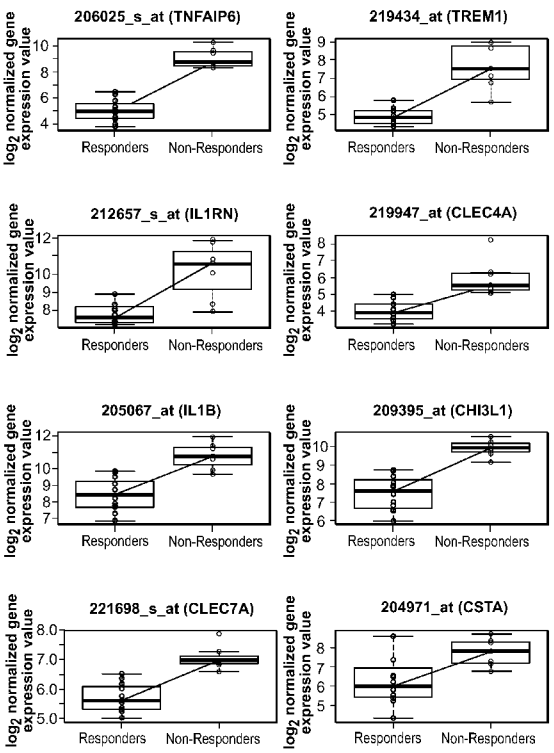

FIG. 23

(57) Abstract: The invention provides methods and compositions for treating a subject or predicting the responsiveness of a subject having inflammatory bowel disease (IBD) to treatment with a TNFα inhibitor, such as an anti-TNFα antibody.

WO 2016/179469 A1

DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING INFLAMMATORY BOWEL DISEASE

RELATED APPLICATION

This application claims the benefit of priority to U.S. Provisional Patent Application. No. 62/158,419, filed on May 7, 2015. The entire contents of the foregoing application are incorporated by reference herein.

BACKGROUND OF THE INVENTION

Inflammatory bowel disease (IBD), such as Crohn's disease and ulcerative colitis, are multifactorial autoimmune diseases of unknown etiology, characterized by chronic inflammation of the gastrointestinal tract. While therapeutic options have increased over the past decade for treating Crohn's disease and ulcerative colitis, the ability to classify subtypes of patients, predict disease progression and behavior, and target newer biologic therapies to specific subgroups of patients has lagged behind. This has led to an empiric step-up approach to therapy, in which increasingly more potent agents are offered until an effective regimen is identified.

Anti-TNFα antibodies, such as adalimumab and infliximab, have proven efficacious in treating IBD, including ulcerative colitis and Crohn's disease. For example, for more than a decade, studies have shown that antibodies against TNFα can induce and maintain remission in patients with Crohn's disease, with about 60% of patients responding. (Schreiber *et al.* (2010) *Am J Gastroenterol* 105:1574). The percentage of patients who respond to a second anti-TNFα agent, however, is lower than with the first agent. (Hanauer *et al.* (2006) *Gastroenterology* 1309:323; and Danborn *et al.* (2007) *Ann Intern Med* 146:829).

It would therefore be highly advantageous to have diagnostic methods, including molecular-based diagnostic methods, that can be used to objectively identify the presence of and/or classify an IBD in a patient, define pathophysiologic aspects of IBD, clinical activity, response to therapy, including response to treatment with various IBD therapeutic agents, prognosis, and/or risk of developing IBD. Thus, there is a continuing need to identify new molecular biomarkers associated with IBD. Such associations would also benefit the identification of pathophysiologic aspects of IBD, clinical activity, response to therapy, or

prognosis. In addition, statistically and biologically significant and reproducible information regarding such associations could be utilized as an integral component in efforts to identify specific subsets of patients who would be expected to significantly benefit from treatment with a particular therapeutic agent, for example where the therapeutic agent is or has been shown in clinical studies to be of therapeutic benefit in such specific IBD patient subpopulation.

SUMMARY OF THE INVENTION

The identification of a genetic marker or genetic markers that would help to predict or assess the effectiveness of a given treatment for IBD remains a challenge. The present invention provides methods of treating subjects having IBD, such as Crohn's disease and ulcerative colitis, as well as methods for identifying subpopulations of IBD patients, who will be responsive to treatment with a TNFα inhibitor. The present invention is based, at least in part, on the identification of gene signatures that can be used to assess the responsiveness of a subject to treatment with a TNFα inhibitor, *e.g.,* treatment with a human anti-TNFα antibodies, or antigen binding portions thereof. Thus, the invention provides more effective ways of treating IBD, such as Crohn's or ulcerative colitis, in subjects having the disease.

In certain embodiments, the invention identifies four sets of genetic markers (gene signatures) that may be used alone, or in combination with one another, to determine whether a subject having IBD will be responsive to treatment with a TNFα inhibitor.

In one embodiment, the invention includes a method for determining the responsiveness of a subject having an inflammatory bowel disease (IBD) to treatment with a TNFα inhibitor, the method comprising determining a gene signature from a biological sample from a subject having an IBD, wherein the gene signature is obtained by determining an expression level(s) of a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6; and comparing the gene signature with a responder gene signature comprising a corresponding gene or gene group of (i) to (iv), wherein the subject will be responsive to treatment with the TNFα inhibitor if the gene signature correlates with the responder gene signature.

In another embodiment, the invention includes a method for determining the responsiveness of a subject having an IBD to treatment with a TNFα inhibitor, the method comprising determining a gene signature from a biological sample from a subject having an IBD, wherein the gene signature is obtained by determining an expression level(s) of a gene or

gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6; and comparing the gene signature with a non-responder gene signature comprising a corresponding gene or gene group of (i) to (iv), wherein the subject will not be responsive to treatment with the TNFα inhibitor if the gene signature correlates with the non-responder gene signature.

In certain embodiments, the invention provides a method for treating a subject having an IBD, said method comprising determining a gene signature from a biological sample from a subject having an IBD, wherein the gene signature is obtained by determining an expression level(s) of a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and
TREM1; and (iv) TNFAIP6, and administering a TNFα inhibitor to the subject having IBD provided that the gene signature correlates with a responder gene signature comprising a corresponding gene or gene group of (i) to (iv).

In one embodiment, the invention provides a method for treating a subject having an IBD, said method comprising selecting a subject having an IBD who is identified as having a responder gene signature, wherein the responder gene signature comprises a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6; and administering a TNFα inhibitor to the subject having IBD.

In another embodiment, the invention includes a method for determining a personalized medical treatment for a subject diagnosed with an IBD comprising the steps of: measuring a gene signature of a biological sample from a subject diagnosed with an IBD comprising gene expression analysis for a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6; accessing a computer database to compare the results of the gene expression analysis in step a) to a responder or non-responder gene signature to determine whether the gene signature correlates to the responder or non-responder gene signature, such that a personalized medical treatment is determined.

In one embodiment, the invention includes a kit for predicting the responsiveness of a subject having IBD to treatment with a TNFα inhibitor, the kit comprising a means for determining the expression level(s) of a gene signature comprising a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6, in a sample obtained from the subject having IBD; and instructions for recommended treatment of the subject having IBD based on the expression level(s) of the gene signature.

In another embodiment, the invention includes a kit for identifying a responder to a TNFα inhibitor for treating IBD, the kit comprising a means for determining the expression level(s) of a gene signature comprising a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6, in a patient sample obtained from the subject having IBD; and a control sample from a responder or a non-responder which can be used to determine the expression level(s) of the gene signature determined for the patient sample.

In one embodiment, the kit comprises a means for determining the presence of the gene signature (i) CSF1R, CSF3R, CCL18, and IL1B in a sample obtained from the subject having IBD. In one embodiment,the means for determining the presence of the gene signature (i) CSF1R, CSF3R, CCL18, and IL1B comprises either nucleic acids that hybridize to the nucleic acid molecules encoding each of genes CSF1R, CSF3R, CCL18, and IL1B; or antibodies which specifically bind to the proteins corresponding to each of genes CSF1R, CSF3R, CCL18, and IL1B.

In one embodiment, the kit comprises a means for determining the presence of the gene signature (ii) CSF1R, CSF3R, CD36, and IL1B in a sample obtained from the subject having Crohn's disease. In one embodiment, the means for determining the presence of the gene signature (ii) CSF1R, CSF3R, CD36, and IL1B comprises either nucleic acids that hybridize to the nucleic acid molecules encoding each of genes CSF1R, CSF3R, CD36, and IL1B; or antibodies which specifically bind to the proteins corresponding to each of genes CSF1R, CSF3R, CD36, and IL1B.

In one embodiment, the kit comprises a means for determining the presence of the gene signature (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6,

and TREM1 in a sample obtained from the subject having Crohn's disease. In one embodiment, the means for determining the presence of the gene signature (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1 comprises either: nucleic acids that hybridize to the nucleic acid molecules encoding each of genes CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; or antibodies which specifically bind to the proteins corresponding to each of genes CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1.

In one embodiment, the kit comprises a means for determining the presence of the gene signature (iv) TNFAIP6 in a sample obtained from the subject having IBD. In one embodiment, the means for determining the presence of the gene signature (iv) TNFAIP6 comprises either a nucleic acid that hybridizes to the nucleic acid molecule encoding the gene TNFAIP6; or an antibody which specifically binds to the protein corresponding to the gene TNFAIP6.

In one embodiment, the biological sample is a tissue sample. Examples of tissue samples useful in the invention include a tissue sample is from an intestine of the subject.

In one embodiment, the biological sample is a blood sample or a stool sample.

In one embodiment, the biological sample comprises peripheral blood mononuclear cells (PBMCs).

In one embodiment, the gene signature comprises the gene group CSF1R, CSF3R, CCL18, and IL1B.

In one embodiment,the gene signature comprises the gene group CSF1R, CSF3R, CD36, and IL1B.

In one embodiment,the gene signature comprises the gene group CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1.

In one embodiment, the gene signature comprises TNFAIP6.

In one embodiment, the expression level is determined using an microarray analysis and/or polymerase chain reaction (PCR) technology. In one embodiment, the PCR technology is quantitative real time PCR (qRT-PCR).

In one embodiment, the expression level is determined using an immunoassay, *e.g.,* ELISA.

In one embodiment, the TNFα inhibitor is an anti-TNFα antibody, or antigen-binding fragment thereof. In one embodiment, the anti-TNFα antibody, or antigen-

binding portion thereof, is selected from the group consisting of a human antibody, a chimeric antibody, and a humanized antibody. In one embodiment, the anti-TNFα antibody is infliximab, or a biosimilar thereof.

In one embodiment, the anti-TNFα antibody, or antigen-binding portion thereof, is an isolated human antibody that dissociates from human TNFα with a K_d of 1 x 10^{-8} M or less and a k_{off} rate constant of 1 x 10^{-3} s^{-1} or less, both determined by surface plasmon resonance, and inhibits human TNFα-induced expression of ELAM-1 in human umbilical vein endothelial cells.

In one embodiment, the anti-TNFα antibody, or antigen-binding portion thereof, comprises a light chain variable region comprising a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 27, a CDR2 domain comprising the amino acid sequence of SEQ ID NO: 29, and a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 31, and a heavy chain variable region comprising a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 28, a CDR2 domain comprising the amino acid sequence of SEQ ID NO: 30, and a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 32.

In one embodiment, the anti-TNFα antibody, or antigen-binding portion thereof, comprises a light chain variable region (LCVR) comprising the amino acid sequence of SEQ ID NO: 25 and a heavy chain variable region (HCVR) comprising the amino acid sequence of SEQ ID NO: 26. In one embodiment, the light chain is a kappa light chain.

In one embodiment, the anti-TNFα antibody, or antigen-binding portion thereof, is an IgG isotype, *e.g.,* an IgG1 or an IgG4.

In one embodiment, the anti-TNFα antibody is adalimumab, or a biosimilar thereof.

In one embodiment, the anti-TNFα antibody is golimumab, or a biosimilar thereof.

In one embodiment, the anti-TNFα antibody fragment is certolizumab pegol, or a biosimilar thereof.

In one embodiment, a first 160 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject, a second 80 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject two weeks after the first dose, and a third 40 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject two weeks after the second dose.

In one embodiment, a 40 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject every other week after the third dose.

In one embodiment, the IBD is Crohn's disease. In one embodiment, the Crohn's disease is moderately to severely active Crohn's disease. In one embodiment, the IBD is ulcerative colitis.

In one embodiment, the subject is a human subject.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: shows a boxplot of the $\log_2$ normalized gene expression values for the before treatment Crohn's disease training data (GEO GSE16879, Arijs *et al.* (2009) *PLoS One*, Vol. 4(11): e7984) comparing non-responders and responders for the probesets of gene signature #1. A diagonal line connects the median expression values.

Figure 2: shows a boxplot of the composite score for gene signature #1 for the Crohn's disease training data set, (GSE16879, Arijs 2009), comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 3: shows the ROC plot of the composite score of gene signature #1 applied to Test Set #1, (GSE52746, Leal 2015) comparing responders versus non-responders. An ROC AUC of 1 was obtained.

Figure 4: shows the canonical Principal Component Analysis (PCA) plot of the normalized gene expressions of gene signature #1 applied to Test Set #1 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 5: shows a boxplot of the composite score for gene signature #1 for Test Set #1, comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 6: shows the ROC plot of the composite score of gene signature #1 applied to Test Set #2, Crohn's disease dataset provided by UMASS, comparing responders versus non-responders. An ROC AUC of 0.79 was obtained

Figure 7: shows the canonical plot of the normalized gene expressions of gene signature #1 applied to Test Set #2 (UMASS dataset) comparing responders versus non-responders. Individual responders are shown as circles and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 8: shows a boxplot of the composite score for gene signature #1 for Test Set #2: Private Dataset (UMASS) comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 9: shows the ROC plot of the composite score of gene signature #1 applied to Test Set #3 UC disease dataset, (GSE16879, Arijs 2009) comparing responders versus non-responders. An ROC AUC of 0.86 was obtained.

Figure 10: shows the canonical PCA plot of the normalized gene expressions of gene signature #1 applied to Test Set #3 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 11: shows a boxplot of the composite score for gene signature #1 for Test Set #3 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 12: shows boxplots of the $\log_2$ normalized gene expression values for the before treatment Crohn's disease training data (GEO GSE16879, Arijs 2009) comparing non-responders and responders for the probesets of gene signature #2. A diagonal line connects the median expression values.

Figure 13: shows a boxplot of the composite score for gene signature #2 for the Crohn's disease training data set, (GSE16879, Arijs 2009), comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups

Figure 14: shows the ROC plot of the composite score of gene signature #2 applied to Test Set #1 , (GSE52746, Leal *et al.* (2014) *Gut* 0:1-10) comparing responders versus non-responders. An ROC AUC of 0.93 was obtained.

Figure 15: shows the canonical PCA plot of the normalized gene expressions of gene signature #2 applied to Test Set #1 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 16: shows a boxplot of the composite score for gene signature #2 for Test Set #1 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 17: shows the ROC plot of the composite score of gene signature #2 applied to Test Set #2, Crohn's disease dataset provided by UMASS, comparing responders versus non-responders. An ROC AUC of 0.75 was obtained.

Figure 18: shows the canonical PCA plot of the normalized gene expressions of gene signature #2 applied to Test Set #2 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 19: shows a boxplot of the composite score for gene signature #2 for Test Set #2 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 20: shows the ROC plot of the composite score of gene signature #2 applied to Test Set #3 UC disease dataset, (GSE16879, Arijs 2009) comparing responders versus non-responders. An ROC AUC of 0.87 was obtained.

Figure 21: shows the canonical PCA plot of the normalized gene expressions of gene signature #2 applied to Test Set #3 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 22: shows a boxplot of the composite score for gene signature #2 for Test Set #3 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 23: show the $\log_2$ normalized gene expression values for the before treatment Crohn's disease training data (GEO GSE16879, Arijs 2009) comparing non-responders and responders for the probesets of gene signature #3. A diagonal line connects the median expression values.

Figure 24: shows a boxplot of the composite score for gene signature #3 for the Crohn's disease training data set, (GSE16879, Arijs 2009), comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 25: shows the ROC plot of the composite score of gene signature #3 applied to Test Set #1 , (GSE52746, Leal 2014) comparing responders versus non-responders. An ROC AUC of 0.93 was obtained.

Figure 26: shows the canonical PCA plot of the normalized gene expressions of gene signature #3 applied to Test Set #1 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 27: shows a boxplot of the composite score for gene signature #3 for Test Set #1 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 28: shows the ROC plot of the composite score of gene signature #3 applied to Test Set #2, Crohn's disease dataset provided by UMASS, comparing responders versus non-responders. An ROC AUC of 0.84 was obtained.

Figure 29: shows the canonical PCA plot of the normalized gene expressions of gene signature #3 applied to Test Set #2 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 30: The above plot shows a boxplot of the composite score for gene signature #3 for Test Set #2 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 31: shows the ROC plot of the composite score of gene signature #3 applied to Test Set #3 UC disease dataset, (GSE16879, Arijs 2009)comparing responders versus non-responders. An ROC AUC of 0.90 was obtained.

Figure 32: shows the canonical PCA plot of the normalized gene expressions of gene signature #3 applied to Test Set #3 comparing responders versus non-responders. Individual responders are shown as points and individual non-responders are shown as stars. Open circles show the 95% confidence region for the true mean of each group.

Figure 33: shows a boxplot of the composite score for gene signature #3 for Test Set #3 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 34: show the $\log_2$ normalized gene expression values for the before treatment Crohn's disease training data (GEO GSE16879, Arijs 2009) comparing non-responders and responders for the probesets of gene signature #4. A diagonal line connects the median expression values.

Figure 35: shows a boxplot of the composite score for gene signature #4 for the Crohn's disease training data set, (GEO GSE16879, Arijs 2009),comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 36: shows the ROC plot of the composite score of gene signature #4 applied to Test Set #1 , (GSE52746, Leal 2014) comparing responders versus non-responders. We obtained a ROC AUC of 0.87.

Figure 37: shows a boxplot of the composite score for gene signature #4 for Test Set #1 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 38: shows the ROC plot of the composite score of gene signature #4 applied to Test Set #2, Crohn's disease dataset provided by UMASS, comparing responders versus non-responders. An ROC AUC of 0.84 was obtained.

Figure 39: shows a boxplot of the composite score for gene signature #4 for Test Set #2 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

Figure 40: shows the ROC plot of the composite score of gene signature #4 applied to Test Set #3 UC disease dataset, (GSE16879, Arijs 2009) comparing responders versus non-responders. An ROC AUC of 0.88 was obtained.

Figure 41: shows a boxplot of the composite score for gene signature #4 for Test Set #3 comparing responders versus non-responders. Individual values are shown as circles for the responders and the non-responders. A diagonal line connects the median across groups.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

In order that the present invention may be more readily understood, certain terms are first defined.

As used herein, the term “inflammatory bowel disease” or “IBD”, used interchangeably herein, refers to inflammatory conditions of the large and small intestine. Examples of inflammatory bowel disease include, but are not limited to, Crohn's disease (also referred to herein as “CD”) and ulcerative colitis.

The term “gene signature” as used herein refers to a gene or group of genes whose expression (*e.g.* expression level or activity) is uniquely characteristic of a biological phenotype or medical condition. In a preferred embodiment of the invention, a gene signature refers to a gene or a group of genes whose expression indicates whether a subject who has inflammatory bowel disease (IBD), *e.g.,* Crohn’s disease or ulcerative colitis, will be responsive to treatment with a TNFα inhibitor, such as, but not limited to, adalimumab.

As used herein, the term "expression" or “gene expression”, refers to transcription of a gene or to translation of a protein. To quantitate expression refers to the act of determining the level of the transcript of the gene or the protein of the gene. Detecting and/or quantitating expression can include determining whether gene expression is upregulated as compared to a control level, downregulated as compared to a control level, or substantially unchanged as compared to a control level. Therefore, the step of quantitating and/or detecting expression does not require that expression of the gene is actually is upregulated or downregulated, but rather, can also include detecting no expression of the gene or detecting that the expression of the gene has not changed or is not different (i.e., detecting no significant expression of the gene or no significant change in expression of TNFα as compared to a control). In one embodiment, the term “gene expression” refers to transcription of a gene. In one embodiment, to detect gene expression refers to the act of actively determining whether the mRNA of the gene is expressed or not, *e.g.,* using real time or quantitative PCR (RT PCR or Q PCR).

The term "level" or "amount" as used herein refers to the measurable quantity of gene expression (transcript or protein). The amount may be either (a) an absolute amount as measured in an appropriate unit, *e.g.,* number of cells, fluorescence intensity, molecules, moles or weight per unit volume or cell or (b) a relative amount. The level of expression of a gene can be considered “high”, “low”, “increased” or “decreased” relative to a control level of expression or relative to the level of expression of the gene in a “responder”, relative to either the level of expression of the gene in a “non-responder”, or, in another embodiment, the level of expression of a subject who does not have an IBD. In one embodiment, the “level of expression” refers to the transcription level of expression of a gene (*e.g.,* a gene from one of the gene signatures described herein) in a sample from a subject having IBD.

The term "responder gene signature" as used herein, refers to a gene signature of a subject or group of subjects having a disease or disorder, such as IBD, (*e.g.,* Crohn's disease or ulcerative colitis), that correlates with treatment when a subject having the disease is administered a certain therapeutic agent. Thus, in one embodiment, a responder gene signature indicates that a subject having that gene signature and IBD will show improvement when given a specific therapeutic agent for the treatment of a certain disorder, such as IBD. In contrast, the term "non-responder gene signature" as used herein, refers to a gene signature of a subject or group of subjects having a disease or disorder, such as IBD, (*e.g.,* Crohn's disease or ulcerative colitis), that correlates with no significant improvement or a deficient response when the subject having the disease is treated with a certain therapeutic agent. In a preferred embodiment of the invention, a non-responder gene signature is a gene signature which indicates that a subject having IBD, (*e.g.,* Crohn's disease or ulcerative colitis), who is treated with a TNFα inhibitor, such as, but not limited to, adalimumab, will not be responsive to or have a deficient response to the TNFα inhibitor treatment for IBD.

The term "correlate" or "correlating", as used herein, is intended to mean that two gene signatures substantially match in their gene expression patterns. For example, if a subject having IBD (*e.g.,* Crohn's disease or ulcerative colitis) has a gene signature that correlates to a responder gene signature, then the subject would be expected to respond to TNFα inhibitor treatment. In contrast, if a subject having IBD (*e.g.,* Crohn's disease or ulcerative colitis) has a gene signature that correlates to a non-responder gene signature, then the subject would *not* be expected to respond to TNFα inhibitor treatment for IBD (*e.g.,* Crohn's disease or ulcerative colitis). In an alternative, the gene signature of a subject having IBD (*e.g.,* Crohn's disease or ulcerative colitis) could be compared to both a responder gene signature and a non-responder gene signature to determine which gene signature the subject's own gene signature correlates more closely with. In one embodiment, two gene signatures correlate if the two gene expression profiles are substantially similar. In one embodiment, gene expression levels of a gene signature are statistically higher in a non-responder versus a responder.

The term "substantially similar" or "substantially the same," as used herein, denotes a sufficiently high degree of similarity between two or more numeric values (for example, one associated with gene expression levels of a gene signature and the other associated with a reference/comparator gene signature), such that one of skill in the art would consider the difference between the two values to be of little or no biological and/or statistical significance within the context of the biological characteristic measured by said values (*e.g.,* gene

expression levels). In one embodiment, the difference between said two values may be, for example, less than about 50%, less than about 40%, less than about 30%, less than about 20%, and/or less than about 10% as a function of the reference/comparator value.

The term "biological sample" as used herein, includes, but is not limited to, any quantity of a substance from a living thing or formerly living thing. A biological sample may be obtained from an individual, body fluid, body tissue, cell line, tissue culture, or other source. Body fluids are, *e.g.,* lymph, sera, whole fresh blood, peripheral blood mononuclear cells, frozen whole blood, plasma (including fresh or frozen), urine, saliva, semen, synovial fluid, and spinal fluid. Biological samples also include synovial tissue, skin, hair follicle, and bone marrow. Methods for obtaining tissue biopsies and body fluids from mammals are well known in the art. If the term "sample" is used alone, it shall still mean that the "sample" is a "biological sample", *i.e.*, the terms are used interchangeably. In the context of the invention, a biological sample may contain genetic material, such as nucleic acids (DNA and/or RNA), or proteins.

The term "polynucleotide" or "nucleic acid," as used interchangeably herein, refers to polymers of nucleotides of any length, and include DNA and RNA. The nucleotides can be deoxyribonucleotides, ribonucleotides, modified nucleotides or bases, and/or their analogs, or any substrate that can be incorporated into a polymer by DNA or RNA polymerase.

The term "primer" refers to a single stranded polynucleotide that is capable of hybridizing to a nucleic acid and allowing the polymerization of a complementary nucleic acid, generally by providing a free 3'-OH group.

The term "array" or "microarray" refers to an ordered arrangement of hybridizable array elements, preferably polynucleotide probes on a substrate. The substrate can be a solid substrate, such as a glass slide, or a semi-solid substrate, such as nitrocellulose membrane.

The term "amplification" refers to the process of producing one or more copies of a reference nucleic acid sequence or its complement. Amplification may be linear or exponential (*e.g.,* PCR).

The term "detection" includes any means of detecting, including direct and indirect detection. As used herein, detection is understood to refer to an assay performed for identification of a specific gene or genes in a sample. In one embodiment, expression of a gene or group of genes is detected, *e.g.,* by PCR (*e.g.,* real time or quantitiative real time PCR (qRT-PCR) and/or microarray analysis.

The term "molecular phenotype" refers to a subpopulation of subjects having IBD, *e.g.,* Crohn's disease or ulcerative colitis, where the subpopulation is characterized by a certain gene signature, *e.g.,* a responder gene signature or a non-responder gene signature.

The term "prediction" is used herein to refer to the likelihood that a subject having a disease or disorder will respond either favorably or unfavorably to a therapeutic agent. In one embodiment, the invention provides a method for predicting whether a subject having IBD, *e.g.,* Crohn's disease or ulcerative colitis, will respond to treatment with a TNFα inhibitor.

The term "human TNFα" (abbreviated herein as hTNFα, or simply hTNF or hTNFalpha), as used herein, is intended to refer to a human cytokine that exists as a 17 kD secreted form and a 26 kD membrane associated form, the biologically active form of which is composed of a trimer of noncovalently bound 17 kD molecules. The structure of hTNFα is described further in, for example, Pennica, D., *et al.* (1984) *Nature* 312:724-729; Davis, J.M., *et al.* (1987) *Biochemistry* 26:1322-1326; and Jones, E.Y., *et al.* (1989) *Nature* 338:225-228. The term human TNFα is intended to include, in one embodiment, recombinant human TNFα (rhTNFα), which can be prepared by standard recombinant expression methods or purchased commercially (R & D Systems, Catalog No. 210-TA, Minneapolis, MN).

The term "TNFα inhibitor" or "TNF inhibitor" includes agents which interfere with TNFα activity. In a preferred embodiment, the TNFα inhibitor is an agent which interferes with human TNFα activity. The term also includes each of the anti-TNFα human antibodies and antibody portions described herein as well as those described in U.S. Patent Nos. 6,090,382; 6,258,562; 6,509,015, and in U.S. Patent Application Serial Nos. 09/801185 and 10/302356. In one embodiment, the TNFα inhibitor used in the invention is an anti-TNFα antibody, or a fragment thereof, including infliximab (REMICADE®, Janssen; described in U.S. Patent No. 5,656,272, incorporated by reference herein), CDP571 (a humanized monoclonal anti-TNF-alpha IgG4 antibody), CDP 870 (a humanized monoclonal anti-TNFα antibody fragment; certolizumab pegol or CIMZIA®; UCB Group), an anti-TNF dAb (Peptech), CNTO 148 (golimumab; Medarex and Janssen, see WO 02/12502), and adalimumab (HUMIRA®, AbbVie, a human anti-TNFα mAb, described in US 6,090,382 as D2E7). Additional anti-TNFα antibodies which may be used in the invention are described in U.S. Patent Nos. 6,593,458; 6,498,237; 6,451,983; and 6,448,380, each of which is incorporated by reference herein. In another embodiment, the TNFα inhibitor is a TNF fusion protein, *e.g.,* etanercept (ENBREL®, Amgen; described in WO 91/03553 and WO 09/406476, incorporated

by reference herein). In another embodiment, the TNFα inhibitor is a recombinant TNF binding protein (r-TBP-I) (Serono).

The term "antibody", as used herein, is intended to refer to immunoglobulin molecules comprised of four polypeptide chains, two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. Each heavy chain is comprised of a heavy chain variable region (abbreviated herein as HCVR or VH) and a heavy chain constant region. The heavy chain constant region is comprised of three domains, CH1, CH2 and CH3. Each light chain is comprised of a light chain variable region (abbreviated herein as LCVR or VL) and a light chain constant region. The light chain constant region is comprised of one domain, CL. The VH and VL regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDR), interspersed with regions that are more conserved, termed framework regions (FR). Each VH and VL is composed of three CDRs and four FRs, arranged from amino-terminus to carboxyl-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The antibodies of the invention are described in further detail in U.S. Patent Nos. 6,090,382; 6,258,562; and 6,509,015, each of which is incorporated herein by reference in its entirety.

The term "antigen-binding portion" or "antigen-binding fragment" of an antibody (or simply "antibody portion"), as used herein, refers to one or more fragments of an antibody that retain the ability to specifically bind to an antigen (*e.g.*, hTNFα). It has been shown that the antigen-binding function of an antibody can be performed by fragments of a full-length antibody. Binding fragments include Fab, Fab', $F(ab')_2$, Fabc, Fv, single chains, and single-chain antibodies. Examples of binding fragments encompassed within the term "antigen-binding portion" of an antibody include (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a $F(ab')_2$ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment (Ward *et al.* (1989) *Nature* 341:544-546), which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Furthermore, although the two domains of the Fv fragment, VL and VH, are coded for by separate genes, they can be joined, using recombinant methods, by a synthetic linker that enables them to be made as a single protein chain in which the VL and VH regions pair to form monovalent molecules (known as single chain Fv (scFv); see *e.g.*, Bird *et al.* (1988) *Science* 242:423-426; and Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5879-

5883). Such single chain antibodies are also intended to be encompassed within the term "antigen-binding portion" of an antibody. Other forms of single chain antibodies, such as diabodies are also encompassed. Diabodies are bivalent, bispecific antibodies in which VH and VL domains are expressed on a single polypeptide chain, but using a linker that is too short to allow for pairing between the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen binding sites (see *e.g.,* Holliger *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:6444-6448; Poljak *et al.* (1994) *Structure* 2:1121-1123). The antibody portions of the invention are described in further detail in U.S. Patent Nos. 6,090,382, 6,258,562, 6,509,015, each of which is incorporated herein by reference in its entirety.

Still further, an antibody or antigen-binding portion thereof may be part of a larger immunoadhesion molecules, formed by covalent or noncovalent association of the antibody or antibody portion with one or more other proteins or peptides. Examples of such immunoadhesion molecules include use of the streptavidin core region to make a tetrameric scFv molecule (Kipriyanov, S.M., *et al.* (1995) *Human Antibodies and Hybridomas* 6:93-101) and use of a cysteine residue, a marker peptide and a C-terminal polyhistidine tag to make bivalent and biotinylated scFv molecules (Kipriyanov, S.M., *et al.* (1994) *Mol. Immunol.* 31:1047-1058). Antibody portions, such as Fab and $F(ab')_2$ fragments, can be prepared from whole antibodies using conventional techniques, such as papain or pepsin digestion, respectively, of whole antibodies. Moreover, antibodies, antibody portions and immunoadhesion molecules can be obtained using standard recombinant DNA techniques, as described herein.

A "conservative amino acid substitution", as used herein, is one in which one amino acid residue is replaced with another amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, including basic side chains (*e.g.,* lysine, arginine, histidine), acidic side chains (*e.g.,* aspartic acid, glutamic acid), uncharged polar side chains (*e.g.,* glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.,* alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.,* threonine, valine, isoleucine) and aromatic side chains (*e.g.,* tyrosine, phenylalanine, tryptophan, histidine).

"Chimeric antibodies" refers to antibodies wherein one portion of each of the amino acid sequences of heavy and light chains is homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular class, while the

remaining segment of the chains is homologous to corresponding sequences from another species. In one embodiment, the invention features a chimeric antibody or antigen-binding fragment, in which the variable regions of both light and heavy chains mimics the variable regions of antibodies derived from one species of mammals, while the constant portions are homologous to the sequences in antibodies derived from another species. In a preferred embodiment of the invention, chimeric antibodies are made by grafting CDRs from a mouse antibody onto the framework regions of a human antibody.

"Humanized antibodies" refer to antibodies which comprise at least one chain comprising variable region framework residues substantially from a human antibody chain (referred to as the acceptor immunoglobulin or antibody) and at least one complementarity determining region (CDR) substantially from a non-human-antibody (*e.g.,* mouse). In addition to the grafting of the CDRs, humanized antibodies typically undergo further alterations in order to improve affinity and/or immmunogenicity.

The term "multivalent antibody" refers to an antibody comprising more than one antigen recognition site. For example, a "bivalent" antibody has two antigen recognition sites, whereas a "tetravalent" antibody has four antigen recognition sites. The terms "monospecific", "bispecific", "trispecific", "tetraspecific", *etc.* refer to the number of different antigen recognition site specificities (as opposed to the number of antigen recognition sites) present in a multivalent antibody. For example, a "monospecific" antibody's antigen recognition sites all bind the same epitope. A "bispecific" or "dual specific" antibody has at least one antigen recognition site that binds a first epitope and at least one antigen recognition site that binds a second epitope that is different from the first epitope. A "multivalent monospecific" antibody has multiple antigen recognition sites that all bind the same epitope. A "multivalent bispecific" antibody has multiple antigen recognition sites, some number of which bind a first epitope and some number of which bind a second epitope that is different from the first epitope

The term "human antibody", as used herein, is intended to include recombinant antibodies having variable and constant regions derived from human germline immunoglobulin sequences. The human antibodies of the invention may include amino acid residues not encoded by human germline immunoglobulin sequences (*e.g.,* mutations introduced by random or site-specific mutagenesis *in vitro* or by somatic mutation *in vivo*), for example in the CDRs and in particular CDR3. However, the term "human antibody", as used herein, is not intended to include antibodies in which CDR sequences derived from the germline of another mammalian species, such as a mouse, have been grafted onto human framework sequences.

The term "isolated protein" or "isolated polypeptide" is a protein or polypeptide that by virtue of its origin or source of derivation is not associated with naturally associated components that accompany it in its native state; is substantially free of other proteins from the same species; is expressed by a cell from a different species; or does not occur in nature. Thus, a polypeptide that is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be "isolated" from its naturally associated components. A protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art.

A "neutralizing antibody", as used herein (or an "antibody that neutralized hTNFα activity"), is intended to refer to an antibody whose binding to hTNFα results in inhibition of the biological activity of hTNFα. This inhibition of the biological activity of hTNFα can be assessed by measuring one or more indicators of hTNFα biological activity, such as hTNFα-induced cytotoxicity (either *in vitro* or *in vivo*), hTNFα-induced cellular activation and hTNFα binding to hTNFα receptors. These indicators of hTNFα biological activity can be assessed by one or more of several standard *in vitro* or *in vivo* assays known in the art (see U.S. Patent No. 6,090,382). Preferably, the ability of an antibody to neutralize hTNFα activity is assessed by inhibition of hTNFα-induced cytotoxicity of L929 cells. As an additional or alternative parameter of hTNFα activity, the ability of an antibody to inhibit hTNFα-induced expression of ELAM-1 on HUVEC, as a measure of hTNFα-induced cellular activation, can be assessed.

The term "surface plasmon resonance", as used herein, refers to an optical phenomenon that allows for the analysis of real-time biospecific interactions by detection of alterations in protein concentrations within a biosensor matrix, for example using the BIAcore system (Pharmacia Biosensor AB, Uppsala, Sweden and Piscataway, NJ). For further descriptions, see Example 1 of U.S. Patent 6,258,562 and Jönsson *et al.* (1993) *Ann. Biol. Clin.* 51:19; Jönsson *et al.* (1991) *Biotechniques* 11:620-627; Johnsson *et al.* (1995) *J. Mol. Recognit.* 8:125; and Johnnson *et al.* (1991) *Anal.Biochem.*198:268.

The term "k_{off}," as used herein, is intended to refer to the off rate constant for dissociation of an antibody from the antibody/antigen complex.

The term "K_d", as used herein, is intended to refer to the dissociation constant of a particular antibody-antigen interaction.

The term "IC_{50}" as used herein, is intended to refer to the concentration of the inhibitor required to inhibit the biological endpoint of interest, *e.g.,* neutralize cytotoxicity activity.

The term "dose," as used herein, refers to an amount of TNFα inhibitor which is administered to a subject.

The term "dosing", as used herein, refers to the administration of a substance (*e.g.,* an anti-TNFα antibody) to achieve a therapeutic objective (*e.g.,* treatment of rheumatoid arthritis).

A "dosing regimen" describes a treatment schedule for a TNFα inhibitor, *e.g.,* a treatment schedule over a prolonged period of time and/or throughout the course of treatment.

The terms "biweekly dosing regimen", "biweekly dosing", and "biweekly administration", as used herein, refer to the time course of administering a substance (*e.g.,* an anti-TNFα antibody) to a subject to achieve a therapeutic objective, *e.g.,* throughout the course of treatment. The biweekly dosing regimen is not intended to include a weekly dosing regimen. In one embodiment, the substance is administered every 9-19 days, every 11-17 days, every 13-15 days, and every 14 days.

The term "combination" as in the phrase "a first agent in combination with a second agent" includes co-administration of a first agent and a second agent, which for example may be dissolved or intermixed in the same pharmaceutically acceptable carrier, or administration of a first agent, followed by the second agent, or administration of the second agent, followed by the first agent. The present invention, therefore, includes methods of combination therapeutic treatment and combination pharmaceutical compositions.

The term "concomitant" as in the phrase "concomitant therapeutic treatment" includes administering an agent in the presence of a second agent. A concomitant therapeutic treatment method includes methods in which the first, second, third, or additional agents are co-administered. A concomitant therapeutic treatment method also includes methods in which the first or additional agents are administered in the presence of a second or additional agents, wherein the second or additional agents, for example, may have been previously administered. A concomitant therapeutic treatment method may be executed step-wise by different actors. For example, one actor may administer to a subject a first agent and a second actor may to administer to the subject a second agent, and the administering steps may be executed at the same time, or nearly the same time, or at distant times, so long as the first agent (and additional agents) are after administration in the presence of the second agent (and additional agents). The actor and the subject may be the same entity (*e.g.,* human).

The term "combination therapy", as used herein, refers to the administration of two or more therapeutic substances, *e.g.,* an anti-TNFα antibody and another drug. The other drug(s)

may be administered concomitant with, prior to, or following the administration of an anti-TNFα antibody.

The term "subject" or "patient," as used herein interchangeably, refers to either a human or non-human animal. In one embodiment, the subject is a human. In one embodiment, the subject is a human having IBD, *e.g.,* Crohn's disease, who has had an inadequate response to conventional therapy. In one embodiment, the subject is a human having IBD, *e.g.,* Crohn's disease who has had an inadequate response to corticosteroids or immunomodulators such as azathioprine, 6-mercaptopurine, or methotrexate. In one embodiment, the subject has ulcerative colitis and has had an inadequate response to immunosuppressants such as corticosteroids, azathioprine or 6-mercaptopurine (6-MP).

A "kit" is any article of manufacture (*e.g.,* a package or container) comprising at least one reagent, *e.g.,* a medicament for treatment of IBD, *e.g.*, Crohn's disease or ulcerative colitis, or a probe for specifically detecting a gene or protein for use in the invention. The article of manufacture is preferably promoted, distributed, or sold as a unit for performing the methods of the present invention.

II. Gene Signatures of Invention

Gene Signatures

An unmet need in the treatment of IBD is to establish predictive biomarkers for therapeutic responders in order to avoid exposure of non-responders to anti-TNFα therapy, thus decreasing morbidity in patients with a low likelihood of response and enhancing safety and cost effective use of this treatment.

The invention provides gene signatures that have been identified as being useful for determining whether a subject having IBD, *e.g.,* Crohn's disease or ulcerative colitis, will be responsive to treatment with a TNFα inhibitor, *e.g.,* adalimumab. Specifically, a panel of gene signatures showing statistically significant differences between responders and non-responders of Crohn's disease and ulcerative colitis patients by performing global gene expression analyses in a training cohort (GSE16879) have been identified, and were cross-validated in two independent cohorts. The sensitive methods disclosed in this application yielded several genes relevant to IBD, including Crohn's disease and ulcerative colitis, that can be used, for example, for determining an appropriate treatment for a subject having IBD.

The invention provides the following four sets of genetic markers (gene signatures):

i) CSF1R, CSF3R, CCL18, and IL1B;

ii) CSF1R, CSF3R, CD36, and IL1B;

iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

iv) TNFAIP6.

The gene signatures described herein represent genes or groups of genes whose expression patterns may be used for determining whether a subject having IBD (*e.g.,* Crohn's disease or ulcerative colitis) will be responsive to treatment with a TNFα inhibitor, including, for example, a human anti-TNFα antibody, such as adalimumab. The gene signatures described herein may also be useful in the treatment of subjects having IBD and for compositions, such as kits, relating to the determination of a treatment for a subject having IBD. The members of each gene signature are described in more detail below:

Gene signature #1

The first gene signature that may be used in the methods and compositions described herein comprises the genes CSF1R, CSF3R, CCL18, and IL1B.

Colony stimulating factor 1 receptor (CSF1R) is a protein tyrosine-kinase transmembrane receptor for CSF1 and IL34. The receptor controls the production, differentiation, and function of macrophages. The nucleic acid sequence encoding CSFIR is described in SEQ ID NO: 1, and the amino acid sequence of the CSFIR protein is provided in SEQ ID NO: 2. In one embodiment, a nucleic acid encoding the CSFIR protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CSFIR gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CSFIR protein may be determined in accordance with the invention.

Colony stimulating factor 3 receptor (CSF3R) is a receptor for granulocyte colony-stimulating factor (CSF3), a cytokine that controls the production, differentiation, and function of granulocytes. In addition CSF3R may function in some adhesion or recognition events at the cell surface. The nucleic acid sequence encoding CSF3R is described in SEQ ID NO: 3, and the amino acid sequence of the CSF3R protein is provided in SEQ ID NO: 4. In one

embodiment, a nucleic acid encoding the CSF3R protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CSF3R gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CSFIR protein may be determined in accordance with the invention.

Chemokine (C-C motif) ligand 18 (CCL18) is a chemotactic factor that attracts lymphocytes but not monocytes or granulocytes. This chemokine may be involved in B-cell migration into B-cell follicles in lymph nodes, and also has chemotactic activity for naive T-cells, CD4+ and CD8+ T-cells and thus may play a role in both humoral and cell-mediated immunity responses. The nucleic acid sequence encoding CCL18 is described in SEQ ID NO: 5, and the amino acid sequence of the CCL18 protein is provided in SEQ ID NO: 6. In one embodiment, a nucleic acid encoding the CCL18 protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CCL18 gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CCL18 protein may be determined in accordance with the invention.

Interleukin 1, beta (IL1B) is produced by activated macrophages and other non-lymphoid cells, such as epithelial cells. IL-1 stimulates thymocyte proliferation by inducing IL-2 release, B-cell maturation and proliferation, and fibroblast growth factor activity. IL-1 proteins are involved in the inflammatory response, being identified as endogenous pyrogens, and are reported to stimulate the release of prostaglandin and collagenase from synovial cells. The nucleic acid sequence encoding the IL1B protein is described in SEQ ID NO: 7, and the amino acid sequence of IL1B protein is provided in SEQ ID NO: 8. In one embodiment, a nucleic acid encoding the IL1B protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the IL1B gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the IL1B protein may be determined in accordance with the invention.

Thus, expression levels from the gene group comprising CSF1R, CSF3R, CCL18, and IL1B may be detected in a sample from a subject having IBD and used in the methods and compositions described herein for the treatment of the subject. In particular, expression levels of CSF1R, CSF3R, CCL18, and IL1B in a sample from a subject having IBD may be use to determine whether that subject will be responsive to treatment with a TNFα inhibitor.

Gene signature #2

The second signature that may be used in the methods and composition described herein comprises the genes CSF1R, CSF3R, CD36, and IL1B. Sequences and descriptions of CSF1R, CSF3R, and IL1B are provided above in the first gene signature.

CD36 is believed to have numerous potential physiological functions. It binds to collagen, thrombospondin, anionic phospholipids and oxidized LDL. CD36 may function as a cell adhesion molecule. The nucleic acid sequence encoding CD36 is described in SEQ ID NO: 9, and the amino acid sequence of the CD36 protein is provided in SEQ ID NO: 10. In one embodiment, a nucleic acid encoding the CD36 protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CD36 gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CD36 protein may be determined in accordance with the invention.

Thus, expression levels from the gene group comprising CSF1R, CSF3R, CCL18, and CD36 may be detected in a sample from a subject having IBD and used in the methods and compositions described herein for the treatment of the subject. In particular, expression levels of CSF1R, CSF3R, CCL18, and CD36 in a sample from a subject having IBD may be use to determine whether that subject will be responsive to treatment with a TNFα inhibitor.

Gene signature #3

The third gene signature that may be used in the methods and composition described herein comprises the genes CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1. Sequences and a description of IL1B are provided above in the second gene signature.

C-type lectin domain family 4, member A (CLEC4A) may be involved in regulating immune reactivity. It may play a role in modulating dendritic cells (DC) differentiation and/or maturation. CLEC4A may be involved via its ITIM motif (immunoreceptor tyrosine-based inhibitory motifs) in the inhibition of B-cell-receptor-mediated calcium mobilization and protein tyrosine phosphorylation. The nucleic acid sequence encoding CLEC4A is described in SEQ ID NO: 11, and the amino acid sequence of the CLEC4A protein is provided in SEQ ID NO: 12. In one embodiment, a nucleic acid encoding the CLEC4A protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CLEC4A gene in accordance with the methods and

compositions of the invention. In another embodiment, levels of the CLEC4A protein may be determined in accordance with the invention.

C-type lectin domain family 7, member A (CLEC7A) is a type of lectin that functions as pattern receptor specific for beta-1,3-linked and beta-1,6-linked glucans, such as cell wall constituents from pathogenic bacteria and fungi. It is necessary for the TLR2-mediated inflammatory response and for TLR2-mediated activation of NF-kappa-B. CLEC7A enhances cytokine production in macrophages and dendritic cells, mediates production of reactive oxygen species in the cell, and binds T-cells in a way that does not involve their surface glycans and plays a role in T-cell activation. The nucleic acid sequence encoding CLEC7A is described in SEQ ID NO: 13, and the amino acid sequence of the CLEC7A protein is provided in SEQ ID NO: 14. In one embodiment, a nucleic acid encoding the CLEC7A protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CLEC7A gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CLEC7A protein may be determined in accordance with the invention.

Cystatin A (stefin A) (CSTA) is an intracellular thiol proteinase inhibitor, may play a role in controlling microorganism invasion. The nucleic acid sequence encoding CSTA is described in SEQ ID NO: 15, and the amino acid sequence of the CSTA protein is provided in SEQ ID NO: 16. In one embodiment, a nucleic acid encoding the CSTA protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CSTA gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CSTA protein may be determined in accordance with the invention.

Chitinase 3-like 1 (cartilage glycoprotein-39) (CHI3L1) is carbohydrate-binding lectin with a preference for chitin. It may play a role in defense against pathogens, or in tissue remodeling. The nucleic acid sequence encoding CHI3LI is described in SEQ ID NO: 17, and the amino acid sequence of the CHI3LI protein is provided in SEQ ID NO: 18. In one embodiment, a nucleic acid encoding the CHI3LI protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the CHI3LI gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the CHI3LI protein may be determined in accordance with the invention.

Interleukin 1 receptor antagonist (IL1RN) inhibits the activity of IL-1 by binding to its receptor. It has no IL-1 like activity and may play a role in regulating cytokine IL1 function. The nucleic acid sequence encoding IL1RN is described in SEQ ID NO: 19, and the amino acid sequence of the IL1RN protein is provided in SEQ ID NO: 20. In one embodiment, a nucleic acid encoding the IL1RN protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the IL1RN gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the IL1RN protein may be determined in accordance with the invention.

Tumor necrosis factor, alpha-induced protein 6 (TNFAIP6) is involved in cell-cell and cell-matrix interactions during inflammation and tumorigenesis. TNFAIP6 is upregulated in response to proinflammaotry cytokines (*e.g.*, IL-1 and TNFα), and it is detected in many inflammatory disease. The nucleic acid sequence encoding TNFAIP6 is described in SEQ ID NO: 21, and the amino acid sequence of the TNFAIP6 protein is provided in SEQ ID NO: 22. In one embodiment, a nucleic acid encoding the TNFAIP6 protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the TNFAIP6 gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the TNFAIP6 protein may be determined in accordance with the invention.

Triggering receptor expressed on myeloid cells 1 (TREM1) stimulates neutrophil and monocyte-mediated inflammatory responses through an unknown ligand. It triggers release of pro-inflammatory chemokines and cytokines, as well as increased surface expression of cell activation markers. The nucleic acid sequence encoding TREM1 is described in SEQ ID NO: 23, and the amino acid sequence of the TREM1 protein is provided in SEQ ID NO: 24. In one embodiment, a nucleic acid encoding the TREM1 protein (mRNA or cDNA when reverse transcribed in a PCR reaction) (or probes or primers thereto) may be used in determining expression levels of the TREM1 gene in accordance with the methods and compositions of the invention. In another embodiment, levels of the TREM1 protein may be determined in accordance with the invention.

Gene signature #4

The fourth predictive gene signature for determining or predicting responsiveness of a subject to treatment with TNFα inhibitor contains one gene set: TNFAIP6. The sequences of TNFAIP6 are described above.

A sequence listing is being filed herewith to provide sequences of the markers provided herein. The gene names, associated accession numbers, and corresponding nucleotide and amino acid SEQ ID NOs are provided in the summary table below. All of the associated accession numbers are incorporated herein by reference.

Gene Name	Associated Gene (Corresponding Protein) NCBI Accession Numbers	Nucleotide SEQ ID NO:	Protein SEQ ID No:
CSF1R colony stimulating factor 1 receptor; Also known as FMS; CSFR; FIM2; HDLS; C-FMS; Crohn's disease115; CSF-1R; M-CSF-R	NM_005211.3 (NP_005202.2)	1	2
CSF3R colony stimulating factor 3 receptor (granulocyte) Also known as Crohn's disease114; GCSFR	NM_000760.3 (NP_000751.1)	3	4
CCL18 chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated) Also known as CKb7; PARC; AMAC1; DCCK1; MIP-4; AMAC-1; DC-CK1; SCYA18	NM_002988.3 (NP_002979.1)	5	6
IL1B interleukin 1, beta Also known as IL-1; IL1F2; IL1-BETA	NM_000576.2 (NP_000567.1)	7	8

Gene Name	Associated Gene (Corresponding Protein) NCBI Accession Numbers	Nucleotide SEQ ID NO:	Protein SEQ ID No:
CD36 CD36 molecule (thrombospondin receptor) Also known as FAT; GP4; GP3B; GPIV; CHDS7; PASIV; SCARB3; BDPLT10	NM_001001548.2 (NP_001001548.1)	9	10
CLEC4A C-type lectin domain family 4, member Also known as DCIR; LLIR; DDB27; CLECSF6; HDCGC13P	transcript variant 1 NM_016184.3 (NP_057268.1)	11	12
CLEC7A C-type lectin domain family 7, member Aprovided Also known as BGR; CANDF4; DECTIN1; CLECSF12	transcript variant 1 NM_197947.2 (NP_922938.1)	13	14
CSTA cystatin A (stefin A) Also known as AREI; STF1; STFA	NM_005213.3 (NP_005204.1)	15	16

Gene Name	Associated Gene (Corresponding Protein) NCBI Accession Numbers	Nucleotide SEQ ID NO:	Protein SEQ ID No:
CHI3L1 chitinase 3-like 1 (cartilage glycoprotein-39) Also known as GP39; ASRT7; GP-39; YKL40; CGP-39; YKL-40; YYL-40; HC-gp39; HCGP-3P; hCGP-39	NM_001276.2 (NP_001267.2)	17	18
IL1RN interleukin 1 receptor antagonist Also known as DIRA; IRAP; IL1F3; IL1RA; MVCrohn's disease4; IL-1RN; IL-1ra; IL-1ra3; ICIL-1RA	t NM_173842.2 (NP_776214.1)	19	20
TNFAIP6 tumor necrosis factor, alpha-induced protein 6 Also known as TSG6; TSG-6	NM_007115.3 (NP_009046.2)	21	22
TREM1 triggering receptor expressed on myeloid cells 1 Also known as Crohn's disease354; TREM-1	NM_018643.3 (NP_061113.1)	23	24

The foregoing sequences are representative of the genes identified as useful in the gene signatures of the invention. Variants of the aforementioned gene sequences are also included in the invention.

In one aspect, the present invention provides methods for predicting or assessing responsiveness of a subject having IBD, *e.g.,* Crohn's disease or ulcerative colitis, to a TNFα inhibitor. The methods generally include determining the expression level(s) of at least one of the following gene signatures: CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; or TNFAIP6 in a biological sample obtained from the subject, wherein the expression level(s) of the marker(s) in the sample is an indication that the subject will respond to treatment with the TNFα inhibitor when compared to a responder gene signature or a non-responder gene signature.

In certain embodiments of the invention, the gene signatures provide a method for determining a personalized medical treatment for a subject diagnosed with an IBD. Such a method includes measuring a gene signature of a biological sample from a subject diagnosed with an IBD comprising gene expression analysis for a gene or gene group selected from the group consisting of (i) CSF1R, CSF3R, CCL18, and IL1B; (ii) CSF1R, CSF3R, CD36, and IL1B; (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and (iv) TNFAIP6. Further the method comprises accessing a computer database to compare the results of the gene expression analysis of the gene signature(s) to a responder or non-responder gene signature to determine whether the gene signature correlates to the responder or non-responder gene signature.

Using the methods described herein and known in the art, a sample from a subject may be tested for the expression level(s) of at least one of the above-mentioned gene signatures.

Methods for detecting the genes of the gene signatures described herein include protocols that examine the presence and/or expression of the gene(s) of the gene signature in a sample from a subject. Determining the presence or absence, and/or expression level(s) of the gene(s) of the gene signature in the biological sample may be accomplished using any well known techniques in the art, such microarrays or a polymerase chain reaction (PCR) amplification reaction, such as reverse-transcriptase PCR (RT-PCR) analysis. Various methods for determining expression of mRNA or protein include, but are not limited to, gene expression profiling, polymerase chain reaction (PCR) including quantitative real time PCR (qRT-PCR), microarray analysis, serial analysis of gene expression (SAGE), MassARRAY, Gene

Expression Analysis by Massively Parallel Signature Sequencing (MPSS), proteomics, immunohistochemistry (IHC), etc. Preferably mRNA is quantified. Such mRNA analysis is preferably performed using the technique of polymerase chain reaction (PCR), or by microarray analysis. In one embodiment, PCR is quantitative real time PCR (qRT-PCR).

Samples, including tissue or cell samples, can be conveniently assayed for expression, *e.g.,* mRNAs, using, for example, a Northern blot method, a Southern blot method, a dot-blot, PCR analysis, array hybridization, RNase protection assay, a FISH method, a CGH method, an RNA chip method, or a DNA chip method, such as a DNA chip microarray (*e.g.,* Affymetrix's microarray system or Illumina's BeadArray Technology). DNA chip microarrays are commercially available, including DNA microarray snapshots. In one embodiment, the methods of the invention are performed using a genechip or DNA microarray comprising nucleic acid probes specific for CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6.

In one embodiment, an mRNA sample may be obtained from the subject (*e.g.,* isolated from peripheral blood mononuclear cells, by standard methods) and expression of mRNA(s) encoding CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 in the mRNA sample may be detected using standard molecular biology techniques, such as PCR analysis. A preferred method of PCR analysis is reverse transcriptase-polymerase chain reaction (RT-PCR) or quantitative real time PCR (qRT-PCR). Other suitable systems for mRNA sample analysis include microarray analysis (*e.g.,* using Affymetrix's microarray system or Illumina's BeadArray Technology).

In certain embodiments of the invention, real-time PCR (RT-PCR) assays such as quantitative PCR assays may be used to detect the expression level(s) of the gene(s) described herein, and such methods are well known in the art. In an illustrative embodiment of the invention, a method for detecting CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 mRNA in a biological sample comprises producing cDNA from the sample by reverse transcription using at least one primer; amplifying the cDNA so produced using CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 polynucleotides as sense and antisense primers to amplify CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R,

CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 cDNAs therein; and detecting the presence of the amplified CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 cDNA. In addition, such methods can include one or more steps that allow one to determine the levels of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 mRNA in a biological sample (*e.g.,* by simultaneously examining the levels of a comparative control mRNA sequence of a "housekeeping" gene such as an actin family member). Optionally, the sequence of the amplified CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 cDNA can be determined.

In certain embodiments of the invention, the method used for determining the expression level(s) of genes in the gene signatures described herein, *i.e.,* CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6, is serotyping of biological samples from a subject using, *e.g.,* commercially available antibodies for CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 in an ELISA assay. For example, samples may be assayed for the expression of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 at the protein level, using a detection reagent that detects the protein product encoded by the mRNA of the marker. For example, if an antibody reagent is available that binds specifically to the protein products of the CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 to be detected, and not to other proteins, then such an antibody reagent can be used to detect the expression of the CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 in a sample from the subject, or a preparation derived from the sample, using standard antibody-based techniques known in the art, such as FACS analysis, ELISA and the like.

Subcombinations within each of the gene signatures described herein, *e.g.,* CLEC4A and CLEC7A, are also contemplated in the methods and compositions of the invention for use in identifying responders and non-responders.

Included in the techniques listed above for determining gene expression levels of the gene signatures described herein, is PCR. PCR can be used to compare mRNA levels in different sample populations, *e.g.,* in responder and/or non-responder samples, to characterize patterns of gene expression of the identified gene signatures.

The first step is the isolation of mRNA from a target sample. As described below, the starting material may be a blood sample from a subject having IBD. General methods for mRNA extraction are well known in the art and are disclosed in standard textbooks of molecular biology, including Ausubel et al., *Current Protocols of Molecular Biology,* John Wiley and Sons (1997). Methods for RNA extraction from paraffin embedded tissues are disclosed, for example, in Rupp and Locker, Lab Invest. 56:A67 (1987), and De Andres et al., BioTechniques 18:42044 (1995). In particular, RNA isolation can be performed using purification kit, buffer set and protease from commercial manufacturers, such as Qiagen, according to the manufacturer's instructions. For example, total RNA from cells in culture can be isolated using Qiagen RNeasy mini-columns. Other commercially available RNA isolation kits include MASTERPURE™. Complete DNA and RNA Purification Kit (EPICENTRE™, Madison, Wis.), and Paraffin Block RNA Isolation Kit (Ambion, Inc.). Total RNA from tissue samples can be isolated using RNA Stat-60 (Tel-Test). RNA prepared from a biological sample can be isolated, for example, by cesium chloride density gradient centrifugation.

In one embodiment, quantitative real time PCR (qRT-PCR) is used to determine the gene expression patterns of the gene signatures described herein. qRT-PCR is well known in the art and measures PCR product accumulation through a dual-labeled fluorigenic probe (i.e., TAQMAN™ probe). Real time PCR is compatible both with quantitative competitive PCR, where internal competitor for each target sequence is used for normalization, and with quantitative comparative PCR using a normalization gene contained within the sample, or a housekeeping gene for PCR. For further details see, *e.g.* Held *et al.,* Genome Research 6:986-994 (1996).

For guidelines for PCR primer and probe design see, *e.g.* Dieffenbach et al., "General Concepts for PCR Primer Design" in PCR Primer, A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 133-155; Innis and Gelfand, "Optimization of PCRs" in PCR Protocols, A Guide to Methods and Applications, CRC Press, London, 1994, pp. 5-11;

and Plasterer, T. N. Primerselect: Primer and probe design. Methods Mol. Biol. 70:520-527 (1997), the entire disclosures of which are hereby expressly incorporated by reference. Primers may be designed using known techniques specific to the gene(s) described herein as part of the four gene signatures of invention. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 1, or a portion thereof, and can be used to detect gene expression levels of CSF1R. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 3, or a portion thereof, and can be used to detect gene expression levels of CSF3R. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 5, or a portion thereof, and can be used to detect gene expression levels of CCL18. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 7, or a portion thereof, and can be used to detect gene expression levels of IL1B. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 9, or a portion thereof, and can be used to detect gene expression levels of CD36. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 11, or a portion thereof, and can be used to detect gene expression levels of CLEC4A. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 13, or a portion thereof, and can be used to detect gene expression levels of CLEC7A. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 15, or a portion thereof, and can be used to detect gene expression levels of CSTA. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 17, or a portion thereof, and can be used to detect gene expression levels of CHI3L1. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 19, or a portion thereof, and can be used to detect gene expression levels of IL1RN. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 21, or a portion thereof, and can be used to detect gene expression levels of TNFAIP6. In one embodiment, a primer(s) used in the methods and compositions described herein is specific to SEQ ID NO: 23, or a portion thereof, and can be used to detect gene expression levels of TREM1. Notably, the methods and compositions described herein (including kits) may include one or any combination of the foregoing primers specific to the gene signatures identified herein.

Expression level(s) of gene(s) from a gene signature described in gene signatures (i) to (iv), are compared to control responder gene signatures and/or non-responder gene signatures. In certain embodiments, a responder, whose gene expression levels defines a responder gene signature, is a subject having IBD who has shown a clinical response when treated with a TNFα inhibitor. Thus, a responder (or a gene signature from a responder) may include, but is not limited to, a subject with IBD who has improved clinical disease status following treatment with a TNFα inhibitor (*e.g.,* reduction in CDAI score or reduction in use of corticosteroids). In one embodiment, a responder is a subject having IBD who achieves a reduction of 100 points or more in their Crohn's Disease Activity Index (CDAI) score following treatment with a TNFα inhibitor. In one embodiment, a responder is a subject having IBD who achieves a reduction of 100 points or more in their Crohn's Disease Activity Index (CDAI) score in a specific time frame following treatment with a TNFα inhibitor. As used herein, "non-responder" includes, but is not limited to, a subject with IBD who has no, or limited improvement in their clinical disease status following treatment with a TNFα inhibitor (*e.g.,* lack of reduction in CDAI score, lack of reduction in use of corticosteroids). In one embodiment, a non-responder is a subject having IBD who fails to achieve a reduction of 100 points or more in their Crohn's Disease Activity Index (CDAI) score following treatment with a TNFα inhibitor. In one embodiment, a non-responder is a subject having IBD who fails to achieve a reduction of 100 points or more in their Crohn's Disease Activity Index (CDAI) score in a specific time frame following treatment with a TNFα inhibitor.

In certain embodiments, the methods and compositions described herein may be used to identify responders and non-responders prior to treatment. Thus, in certain embodiments, a responder or non-responder is a subject having IBD who has not yet received treatment with a TNFα inhibitor.

Levels of expression of the gene signature of a subject having IBD are compared to levels of expression of the same gene signature from a responder gene signature and/or a non-responder gene signature. Given the responder and non-responder expression profiles, one of ordinary skill would be able to determine which gene signature *i.e.,* responder or non-responder, the subject correlates to. Measurement of gene signature expression levels, for example, may be performed by using a software program executed by a suitable processor. Suitable software and processors are well known in the art and are commercially available. The program may be embodied in software stored on a tangible medium such as a CD-ROM, a floppy disk, a hard drive, a DVD, or a memory associated with the processor, but persons of

ordinary skill in the art will readily appreciate that the entire program or parts thereof could alternatively be executed by a device other than a processor, and/or embodied in firmware and/or dedicated hardware in a well known manner.

Following the measurement of the expression levels or presence of the genes identified herein, or their expression products, and the determination that a subject is likely or not likely to respond to treatment with a TNFα inhibitor, the assay results, findings, diagnoses, predictions, and/or treatment recommendations may be recorded and communicated to technicians, physicians, and/or patients, for example. In certain embodiments, computers will be used to communicate such information to interested parties, such as patients and/or the attending physicians. In some embodiments, the assays will be performed or the assay results analyzed in a country or jurisdiction that differs from the country or jurisdiction to which the results or diagnoses are communicated.

In certain embodiments of the invention, a diagnosis, prediction, and/or treatment recommendation based on the expression level of genes in a gene signature described herein, in a subject having IBD is communicated to the subject as soon as possible after the assay is completed and the diagnosis and/or prediction is generated. The results and/or related information may be communicated to the subject by the subject's treating physician. Alternatively, the results may be communicated directly to a test subject by any means of communication, including writing, electronic forms of communication, such as e-mail, or telephone. Communication may be facilitated by use of a computer, such as in the case of e-mail communications. In certain embodiments, the communication containing results of a diagnostic test and/or conclusions drawn from and/or treatment recommendations based on the test may be generated and delivered automatically to the subject using a combination of computer hardware and software that will be familiar to artisans skilled in telecommunications. One example of a healthcare-oriented communications system is described in U.S. Pat. No. 6,283,761; however, the present invention is not limited to methods that utilize this particular communications system. In certain embodiments of the methods of the invention, all or some of the method steps, including the assaying of samples, diagnosing of diseases, and communicating of assay results or diagnoses, may be carried out in diverse (*e.g.,* foreign) jurisdictions.

Any sample obtained from a subject having IBD, *e.g.,* Crohn's disease or ulcerative colitis, may be used to determine the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN,

TNFAIP6, AND TREM1; and/or TNFAIP6. For example, the sample may be any fluid or sub-component thereof, *e.g.,* blood fluids, vomit, intra-articular fluid, saliva, lymph, cystic fluid, urine, feces, fluids collected by bronchial lavage, fluids collected by peritoneal rinsing, or gynecological fluids, obtained from the subject. In a typical situation, the fluid may be a blood sample, or a component thereof, obtained from the subject. The sample may also be any tissue or fragment or sub-component thereof, *e.g.,* bone, connective tissue, cartilage, lung, liver, kidney, muscle tissue, heart, pancreas, and skin, obtained from the subject.

Techniques or methods for obtaining samples from a subject are well known in the art and include, for example, obtaining samples by a mouth swab or a mouth wash; drawing blood; or obtaining a biopsy. Isolating sub-components of fluid or tissue samples (*e.g.,* cells or RNA or DNA) may be accomplished using well known techniques in the art.

Methods of Treatment

The invention also provides methods for treating a subject having IBD. Examples of inflammatory bowel diseases that may be treated by the methods described herein include, but are not limited to, Crohn's disease and ulcerative colitis.

Crohn's disease represents one of the major entities of inflammatory bowel diseases, and is characterized by a chronic relapsing inflammation of the intestinal mucosa (Strober et al. *J Clin Invest* 117, 514-521 (2007) and Danese, S. New therapies for inflammatory bowel disease: from the bench to the bedside. *Gut* 61, 918-932 (2012)). Patients with this incurable disease suffer from chronic diarrhea, rectal bleeding, abdominal cramping and fistula formation and many patients require surgical intervention over time. It is general consensus that inappropriate activation of the mucosal immune system leading to augmented cytokine production contributes to disease pathogenesis (Neurath et al. *Immunity* 31, 357-361 (2009) and Atreya, et al. *Nat Med* 6, 583-588 (2000)). In this context, the pro-inflammatory cytokine tumor necrosis factor-α (TNFα) plays a pivotal role in Crohn's disease immunopathogenesis.

Ulcerative colitis is a chronic inflammatory disease of unknown etiology afflicting the large intestine. The course of the disease may be continuous or relapsing, mild or severe. The earliest lesion is an inflammatory infiltration with abscess formation at the base of the crypts of Lieberkuhn. Coalescence of these distended and ruptured crypts tends to separate the overlying mucosa from its blood supply, leading to ulceration. Signs and symptoms of the disease include cramping, lower abdominal pain, rectal bleeding, and frequent, loose discharges consisting

mainly of blood, pus, and mucus with scanty fecal particles. A total colectomy may be required for acute, severe or chronic, unremitting ulcerative colitis.

In one embodiment, the gene signatures described herein may be used in a method of selecting a patient having IBD, *e.g.,* Crohn's disease or ulcerative colitis, who will be responsive to treatment with a TNFα inhibitor, *e.g.,* a human anti-TNFα antibody, or antigen binding portion thereof, such as, but not limited to, adalimumab. The method includes determining the expression levels of a gene signature, *i.e.,* CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6, in the subject, selecting a treatment regimen with an TNFα inhibitor based upon the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 in the subject, and administering the TNFα inhibitor according to the treatment regimen such that the subject is treated for the Crohn's disease.

In certain embodiments of the invention, the gene signatures described herein may be used in a method for treating a subject having an IBD comprising determining a gene signature from a biological sample from a subject having IBD, wherein the gene signature is obtained by determining an expression level(s) of any one of gene signatures (i) to (iv) described above. Following the determination of the gene signature, a TNFα inhibitor may be administered to the subject having IBD provided that the gene signature correlates with a responder gene signature comprising the same gene or gene group of (i) to (iv) tested. Should the gene signature of the subject correlate to a non-responder gene signature, however, an alternative therapy other than a TNFα inhibitor may need to be considered. Thus, the use of the gene signatures described herein may be incorporated into the treatment process, whereby a determination is first made as to whether or not a TNFα inhibitor is an appropriate therapy for a subject having IBD, *e.g.,* Crohn's disease or ulcerative colitis.

The pharmaceutical composition used in the methods of the invention may include a "therapeutically effective amount" or a "prophylactically effective amount" of an antibody or antibody portion of the invention. A "therapeutically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired therapeutic result. A therapeutically effective amount of the antibody, antibody portion, or other TNFα inhibitor may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of the antibody, antibody portion, other TNFα inhibitor to elicit a desired

response in the individual. A therapeutically effective amount is also one in which any toxic or detrimental effects of the antibody, antibody portion, or other TNFα inhibitor are outweighed by the therapeutically beneficial effects. A "prophylactically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired prophylactic result. Typically, since a prophylactic dose is used in subjects prior to or at an earlier stage of disease, the prophylactically effective amount will be less than the therapeutically effective amount.

The treatment regimen that is selected typically includes at least one of the following parameters and may include many or all of the following parameters: the type of agent chosen for administration, the dosage, the formulation, the route of administration and/or the frequency of administration.

In one embodiment, the subject having IBD who is identified as a responder to TNFα inhibitor therapy according to the methods described herein, is treated with a human anti-TNFα antibody, or antigen-binding portion thereof, according to a multiple variable dose regimen. Multiple-variable dose regimens are described in US Publication No. 20060009385, which is incorporated by reference herein in its entirety. In one embodiment, a subject identified as a responder is subcutaneously administered a loading or induction dose (s) followed by subsequent treatment or maintenance doses. In one embodiment, the subject is subcutaneously administered a first dose of 160 mg, a second dose of 80 mg, and a dose of 40 mg. In a further embodiment a dose of 80 mg is administered subcutaneously followed by a dose of 40 mg for treatment of IBD, including Crohn's disease and ulcerative colitis in a subject identified as a responder.

In one embodiment, the subject is treated with a TNFα inhibitor in accordance with a biweekly dosing regimen. Biweekly dosing regimens are further described in US Appln. No. 10/163657 (US 20030235585), incorporated by reference herein. In one embodiment, the subject is subcutaneously administered 40 mg of a human TNFα antibody, or antigen binding portion thereof, every other week for the treatment of Crohn's disease or ulcerative colitis.

In certain embodiments of the invention, the methods described herein are used for reducing signs and symptoms and inducing and maintaining clinical remission in adult patients with moderately to severely active Crohn's disease. In certain embodiments of the invention, the methods described herein are used for reducing signs and symptoms and inducing clinical remission in patients having Crohn's disease if they have also lost response to or are intolerant

to infliximab. In certain embodiments of the invention, the methods described herein are used for reducing signs and symptoms and inducing and maintaining clinical remission in patients 6 years of age and older with moderately to severely active Crohn's disease who have had an inadequate response to corticosteroids or immunomodulators such as azathioprine, 6-mercaptopurine, or methotrexate. In certain embodiments of the invention, the methods described herein are used to for inducing and sustaining clinical remission in adult patients with moderately to severely active ulcerative colitis.

III. TNFα Inhibitors for Use in the Methods and Compositions of the Invention

The invention provides a method for determining whether a subject having IBD will respond to treatment with a TNFα inhibitor based on the identified gene signatures (i) to (iv), as well as methods and kits based on the use of the gene signatures described herein.

In one embodiment, the TNFα inhibitor used in the methods and compositions of the invention is an anti-TNFα antibody, or antigen-binding portion thereof, such as, but not limited to, a human antibody, a chimeric antibody, and a humanized antibody. An example of a chimeric antibody that may be used is infliximab, or a biosimilar thereof. In one embodiment, the anti-TNFα antibody fragment is certolizumab pegol, or a biosimilar thereof.

In one embodiment, the invention features uses and composition for predicting or determining the responsiveness of a subject having an IBD to treatment with a TNFα inhibitor, or for treating said subject, wherein the TNFα antibody is an isolated human antibody, or antigen-binding portion thereof, that binds to human TNFα with high affinity and a low off rate, and also has a high neutralizing capacity. Examples of such antibodies include adalimumab or golimumab. Preferably, the human antibodies used in the invention are recombinant, neutralizing human anti-hTNFα antibodies. The most preferred recombinant, neutralizing antibody of the invention is referred to herein as adalimumab, also referred to as HUMIRA® or D2E7 (the amino acid sequence of the adalimumab VL region is shown in SEQ ID NO: 25; the amino acid sequence of the adalimumab VH region is shown in SEQ ID NO: 26). The properties (and sequences) of D2E7 (adalimumab / HUMIRA®) have been described in Salfeld *et al.*, U.S. Patent Nos. 6,090,382, 6,258,562, and 6,509,015, which are each incorporated by reference herein.

In one embodiment, the TNFα inhibitor for use in the invention is a fully human TNFα antibody which is a biosimilar to adalimumab. In one embodiment, the TNFα inhibitor is

highly similar to adalimumab, and may, for example, include minor differences in clinically inactive components. In one embodiment, the TNFα inhibitor is interchangeable with adalimumab, and is, for example, able to produce the same clinical result as adalimumab in any given patient. In one embodiment, the adalimumab biosimilar is Exemptia (Zydus Cadila), ABP 501 (Amgen), ONS-3010 (Oncobiologics), and BI 695501 (Boehringer Ingelheim).

In one embodiment, the method of the invention includes determining the responsiveness of a subject to treatment of IBD with adalimumab or other human antibodies and antibody portions with equivalent properties to adalimumab, such as high affinity binding to hTNFα with low dissociation kinetics and high neutralizing capacity, for the treatment of an IBD, *e.g.,* Crohn's disease or ulcerative colitis. In one embodiment, the invention provides treatment with an isolated human antibody, or an antigen-binding portion thereof, that dissociates from human TNFα with a K_d of 1 x 10^{-8} M or less and a k_{off} rate constant of 1 x 10^{-3} s^{-1} or less, both determined by surface plasmon resonance, and neutralizes human TNFα cytotoxicity in a standard *in vitro* L929 assay with an IC_{50} of 1 x 10^{-7} M or less. More preferably, the isolated human antibody, or antigen-binding portion thereof, dissociates from human TNFα with a k_{off} of 5 x 10^{-4} s^{-1} or less, or even more preferably, with a k_{off} of 1 x 10^{-4} s^{-1} or less. More preferably, the isolated human antibody, or antigen-binding portion thereof, neutralizes human TNFα cytotoxicity in a standard *in vitro* L929 assay with an IC_{50} of 1 x 10^{-8} M or less, even more preferably with an IC_{50} of 1 x 10^{-9} M or less and still more preferably with an IC_{50} of 1 x 10^{-10} M or less.

It is well known in the art that antibody heavy and light chain CDR3 domains play an important role in the binding specificity/affinity of an antibody for an antigen. Accordingly, in another aspect, the invention pertains to treating an IBD, *e.g.,* Crohn's disease or ulcerative colitis, by administering human antibodies that have slow dissociation kinetics for association with hTNFα and that have light and heavy chain CDR3 domains that structurally are identical to or related to those of adalimumab. Position 9 of the adalimumab VL CDR3 can be occupied by Ala or Thr without substantially affecting the k_{off}. Accordingly, a consensus motif for the adalimumab VL CDR3 comprises the amino acid sequence: Q-R-Y-N-R-A-P-Y-(T/A) (SEQ ID NO: 27). Additionally, position 12 of the adalimumab VH CDR3 can be occupied by Tyr or Asn, without substantially affecting the k_{off}. Accordingly, a consensus motif for the adalimumab VH CDR3 comprises the amino acid sequence: V-S-Y-L-S-T-A-S-S-L-D-(Y/N) (SEQ ID NO: 28). Moreover, as demonstrated in Example 2 of U.S. Patent No. 6,090,382, the CDR3 domain of the adalimumab heavy and light chains is amenable to substitution with a

single alanine residue (at position 1, 4, 5, 7 or 8 within the VL CDR3 or at position 2, 3, 4, 5, 6, 8, 9, 10 or 11 within the VH CDR3) without substantially affecting the k_{off}. Still further, the skilled artisan will appreciate that, given the amenability of the adalimumab VL and VH CDR3 domains to substitutions by alanine, substitution of other amino acids within the CDR3 domains may be possible while still retaining the low off rate constant of the antibody, in particular substitutions with conservative amino acids. Preferably, no more than one to five conservative amino acid substitutions are made within the adalimumab VL and/or VH CDR3 domains. More preferably, no more than one to three conservative amino acid substitutions are made within the adalimumab VL and/or VH CDR3 domains. Additionally, conservative amino acid substitutions should not be made at amino acid positions critical for binding to hTNFα. Positions 2 and 5 of the adalimumab VL CDR3 and positions 1 and 7 of the adalimumab VH CDR3 are critical for interaction with hTNFα and thus, conservative amino acid substitutions preferably are not made at these positions (although an alanine substitution at position 5 of the adalimumab VL CDR3 is acceptable, as described above) (*see* U.S. Patent No. 6,090,382).

Accordingly, in another embodiment, the antibody or antigen-binding portion thereof preferably dissociates from human TNFα with a k_{off} rate constant of 1 x 10^{-3} s^{-1} or less, as determined by surface plasmon resonance; has a light chain CDR3 domain comprising the amino acid sequence of SEQ ID NO: 27, or modified from SEQ ID NO: 27 by a single alanine substitution at position 1, 4, 5, 7 or 8 or by one to five conservative amino acid substitutions at positions 1, 3, 4, 6, 7, 8 and/or 9; and has a heavy chain CDR3 domain comprising the amino acid sequence of SEQ ID NO: 28, or modified from SEQ ID NO: 28 by a single alanine substitution at position 2, 3, 4, 5, 6, 8, 9, 10 or 11 or by one to five conservative amino acid substitutions at positions 2, 3, 4, 5, 6, 8, 9, 10, 11 and/or 12.

More preferably, the antibody, or antigen-binding portion thereof, dissociates from human TNFα with a k_{off} of 5 x 10^{-4} s^{-1} or less. Even more preferably, the antibody, or antigen-binding portion thereof, dissociates from human TNFα with a k_{off} of 1 x 10^{-4} s^{-1} or less.

In yet another embodiment, the antibody or antigen-binding portion thereof preferably contains a light chain variable region (LCVR) having a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 27, or modified from SEQ ID NO: 27 by a single alanine substitution at position 1, 4, 5, 7 or 8, and with a heavy chain variable region (HCVR) having a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 28, or modified from SEQ ID NO: 28 by a single alanine substitution at position 2, 3, 4, 5, 6, 8, 9, 10 or 11. In one embodiment, the LCVR further has a CDR2 domain comprising the amino acid sequence of

SEQ ID NO: 29 (*i.e.*, the adalimumab VL CDR2) and the HCVR further has a CDR2 domain comprising the amino acid sequence of SEQ ID NO: 30 (*i.e.*, the adalimumab VH CDR2). In one embodiment, the LCVR further has CDR1 domain comprising the amino acid sequence of SEQ ID NO: 31 (*i.e.*, the adalimumab VL CDR1) and the HCVR has a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 32 (*i.e.*, the adalimumab VH CDR1). The framework regions for VL preferably are from the V_KI human germline family, more preferably from the A20 human germline Vk gene and most preferably from the adalimumab VL framework sequences shown in Figures 1A and 1B of U.S. Patent No. 6,090,382. The framework regions for VH preferably are from the V_H3 human germline family, more preferably from the DP-31 human germline VH gene and most preferably from the adalimumab VH framework sequences shown in Figures 2A and 2B of U.S. Patent No. 6,090,382.

Accordingly, in another embodiment, the antibody or antigen-binding portion thereof preferably contains a light chain variable region (LCVR) comprising the amino acid sequence of SEQ ID NO: 25 (*i.e.*, the adalimumab VL) and a heavy chain variable region (HCVR) comprising the amino acid sequence of SEQ ID NO: 26 (*i.e.*, the adalimumab VH). In certain embodiments, the antibody comprises a heavy chain constant region, such as an IgG1, IgG2, IgG3, IgG4, IgA, IgE, IgM or IgD constant region. Preferably, the heavy chain constant region is an IgG1 heavy chain constant region or an IgG4 heavy chain constant region. Furthermore, the antibody can comprise a light chain constant region, either a kappa light chain constant region or a lambda light chain constant region. Preferably, the antibody comprises a kappa light chain constant region. Alternatively, the antibody portion can be, for example, a Fab fragment or a single chain Fv fragment. In one embodiment, the antibody or antigen-binding portion thereof, comprises a light chain variable region (LCVR) comprising CDRs as described by the amino acid sequence of SEQ ID NO: 25 (*i.e.*, the adalimumab VL) and a heavy chain variable region (HCVR) comprising CDRs as described in the amino acid sequence of SEQ ID NO: 26 (*i.e.*, the adalimumab VH).

The TNFα antibody used in the methods and compositions of the invention may be modified for improved treatment of an IBD, *e.g.,* Crohn's disease or ulcerative colitis. In some embodiments, the TNFα antibody or antigen binding fragments thereof, is chemically modified to provide a desired effect. For example, pegylation of antibodies and antibody fragments of the invention may be carried out by any of the pegylation reactions known in the art, as described, for example, in the following references: *Focus on Growth Factors* 3:4-10 (1992);

EP 0 154 316; and EP 0 401 384 (each of which is incorporated by reference herein in its entirety). Preferably, the pegylation is carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer). A preferred water-soluble polymer for pegylation of the antibodies and antibody fragments of the invention is polyethylene glycol (PEG). As used herein, "polyethylene glycol" is meant to encompass any of the forms of PEG that have been used to derivatize other proteins, such as mono (Cl-ClO) alkoxy- or aryloxy-polyethylene glycol.

Methods for preparing pegylated antibodies and antibody fragments of the invention will generally comprise the steps of (a) reacting the antibody or antibody fragment with polyethylene glycol, such as a reactive ester or aldehyde derivative of PEG, under conditions whereby the antibody or antibody fragment becomes attached to one or more PEG groups, and (b) obtaining the reaction products. It will be apparent to one of ordinary skill in the art to select the optimal reaction conditions or the acylation reactions based on known parameters and the desired result.

Pegylated antibodies and antibody fragments may generally be used to treat IBD by administration of the TNFα antibodies and antibody fragments described herein. Generally the pegylated antibodies and antibody fragments have increased half-life, as compared to the nonpegylated antibodies and antibody fragments. The pegylated antibodies and antibody fragments may be employed alone, together, or in combination with other pharmaceutical compositions.

In yet another embodiment of the invention, TNFα antibodies or fragments thereof can be altered wherein the constant region of the antibody is modified to reduce at least one constant region-mediated biological effector function relative to an unmodified antibody. To modify an antibody of the invention such that it exhibits reduced binding to the Fc receptor, the immunoglobulin constant region segment of the antibody can be mutated at particular regions necessary for Fc receptor (FcR) interactions (see *e.g.,* Canfield, S.M. and S.L. Morrison (1991) *J. Exp. Med.* 173:1483-1491; and Lund, J. *et al.* (1991) *J. of Immunol.* 147:2657-2662). Reduction in FcR binding ability of the antibody may also reduce other effector functions which rely on FcR interactions, such as opsonization and phagocytosis and antigen-dependent cellular cytotoxicity.

An antibody or antibody portion used in the methods of the invention can be derivatized or linked to another functional molecule (*e.g.,* another peptide or protein). Accordingly, the antibodies and antibody portions of the invention are intended to include derivatized and

otherwise modified forms of the human anti-hTNFα antibodies described herein, including immunoadhesion molecules. For example, an antibody or antibody portion of the invention can be functionally linked (by chemical coupling, genetic fusion, noncovalent association or otherwise) to one or more other molecular entities, such as another antibody (*e.g.,* a bispecific antibody or a diabody), a detectable agent, a cytotoxic agent, a pharmaceutical agent, and/or a protein or peptide that can mediate associate of the antibody or antibody portion with another molecule (such as a streptavidin core region or a polyhistidine tag).

One type of derivatized antibody is produced by cross-linking two or more antibodies (of the same type or of different types, *e.g.,* to create bispecific antibodies). Suitable cross-linkers include those that are heterobifunctional, having two distinctly reactive groups separated by an appropriate spacer (*e.g.,* m-maleimidobenzoyl-N-hydroxysuccinimide ester) or homobifunctional (*e.g.,* disuccinimidyl suberate). Such linkers are available from Pierce Chemical Company, Rockford, IL.

An antibody, or antibody portion, used in the methods and compositions of the invention, can be prepared by recombinant expression of immunoglobulin light and heavy chain genes in a host cell. To express an antibody recombinantly, a host cell is transfected with one or more recombinant expression vectors carrying DNA fragments encoding the immunoglobulin light and heavy chains of the antibody such that the light and heavy chains are expressed in the host cell and, preferably, secreted into the medium in which the host cells are cultured, from which medium the antibodies can be recovered. Standard recombinant DNA methodologies are used to obtain antibody heavy and light chain genes, incorporate these genes into recombinant expression vectors and introduce the vectors into host cells, such as those described in Sambrook, Fritsch and Maniatis (eds), *Molecular Cloning; A Laboratory Manual, Second Edition*, Cold Spring Harbor, N.Y., (1989), Ausubel, F.M. *et al.* (eds.) *Current Protocols in Molecular Biology*, Greene Publishing Associates, (1989) and in U.S. Patent No. 4,816,397 by Boss *et al.*

To express an anti-TNFα antibody, such as adalimumab or an adalimumab-related antibody (*e.g.,* an antibody have at least 95% identity in sequence to the amino acid sequence set forth in SEQ ID NOs: 25 and/or 26), DNA fragments encoding the light and heavy chain variable regions are first obtained. These DNAs can be obtained by amplification and modification of germline light and heavy chain variable sequences using the polymerase chain reaction (PCR). Germline DNA sequences for human heavy and light chain variable region genes are known in the art (see *e.g.,* the "Vbase" human germline sequence database; see also

Kabat, E.A., *et al.* (1991) *Sequences of Proteins of Immunological Interest, Fifth Edition*, U.S. Department of Health and Human Services, NIH Publication No. 91-3242; Tomlinson, I.M., *et al.* (1992) "The Repertoire of Human Germline V_H Sequences Reveals about Fifty Groups of V_H Segments with Different Hypervariable Loops" *J. Mol. Biol.* 227:776-798; and Cox, J.P.L. *et al.* (1994) "A Directory of Human Germ-line V_{78} Segments Reveals a Strong Bias in their Usage" *Eur. J. Immunol.* 24:827-836; the contents of each of which are expressly incorporated herein by reference). To obtain a DNA fragment encoding the heavy chain variable region of adalimumab, a member of the V_H3 family of human germline VH genes is amplified by standard PCR. Most preferably, the DP-31 VH germline sequence is amplified. To obtain a DNA fragment encoding the light chain variable region of adalimumab, or an adalimumab-related antibody, a member of the V_KI family of human germline VL genes is amplified by standard PCR. Most preferably, the A20 VL germline sequence is amplified. PCR primers suitable for use in amplifying the DP-31 germline VH and A20 germline VL sequences can be designed based on the nucleotide sequences disclosed in the references cited *supra*, using standard methods.

Once the germline VH and VL fragments are obtained, these sequences can be mutated to encode the adalimumab, or an adalimumab-related amino acid sequences disclosed herein. The amino acid sequences encoded by the germline VH and VL DNA sequences are first compared to the adalimumab, or an adalimumab-related VH and VL amino acid sequences to identify amino acid residues in the adalimumab, or an adalimumab-related sequence that differ from germline. Then, the appropriate nucleotides of the germline DNA sequences are mutated such that the mutated germline sequence encodes the adalimumab, or an adalimumab-related amino acid sequence, using the genetic code to determine which nucleotide changes should be made. Mutagenesis of the germline sequences is carried out by standard methods, such as PCR-mediated mutagenesis (in which the mutated nucleotides are incorporated into the PCR primers such that the PCR product contains the mutations) or site-directed mutagenesis.

Moreover, it should be noted that if the "germline" sequences obtained by PCR amplification encode amino acid differences in the framework regions from the true germline configuration (*i.e.*, differences in the amplified sequence as compared to the true germline sequence, for example as a result of somatic mutation), it may be desirable to change these amino acid differences back to the true germline sequences (*i.e.*, "backmutation" of framework residues to the germline configuration).

Once DNA fragments encoding the anti-TNFα antibody, *e.g.*, adalimumab, VH and VL segments are obtained (by amplification and mutagenesis of germline VH and VL genes, as described above), these DNA fragments can be further manipulated by standard recombinant DNA techniques, for example to convert the variable region genes to full-length antibody chain genes, to Fab fragment genes or to a scFv gene. In these manipulations, a VL- or VH-encoding DNA fragment is operatively linked to another DNA fragment encoding another protein, such as an antibody constant region or a flexible linker. The term "operatively linked", as used in this context, is intended to mean that the two DNA fragments are joined such that the amino acid sequences encoded by the two DNA fragments remain in-frame.

The isolated DNA encoding the VH region can be converted to a full-length heavy chain gene by operatively linking the VH-encoding DNA to another DNA molecule encoding heavy chain constant regions (CH1, CH2 and CH3). The sequences of human heavy chain constant region genes are known in the art (see *e.g.*, Kabat, E.A., *et al.* (1991) *Sequences of Proteins of Immunological Interest, Fifth Edition*, U.S. Department of Health and Human Services, NIH Publication No. 91-3242) and DNA fragments encompassing these regions can be obtained by standard PCR amplification. The heavy chain constant region can be an IgG1, IgG2, IgG3, IgG4, IgA, IgE, IgM or IgD constant region, but most preferably is an IgG1 or IgG4 constant region. For a Fab fragment heavy chain gene, the VH-encoding DNA can be operatively linked to another DNA molecule encoding only the heavy chain CH1 constant region.

The isolated DNA encoding the VL region can be converted to a full-length light chain gene (as well as a Fab light chain gene) by operatively linking the VL-encoding DNA to another DNA molecule encoding the light chain constant region, CL. The sequences of human light chain constant region genes are known in the art (see *e.g.*, Kabat, E.A., *et al.* (1991) *Sequences of Proteins of Immunological Interest, Fifth Edition*, U.S. Department of Health and Human Services, NIH Publication No. 91-3242) and DNA fragments encompassing these regions can be obtained by standard PCR amplification. The light chain constant region can be a kappa or lambda constant region, but most preferably is a kappa constant region.

To create a scFv gene, the VH- and VL-encoding DNA fragments are operatively linked to another fragment encoding a flexible linker, *e.g.*, encoding the amino acid sequence $(Gly_4\text{-}Ser)_3$ (SEQ ID NO: 33) such that the VH and VL sequences can be expressed as a contiguous single-chain protein, with the VL and VH regions joined by the flexible linker (see

e.g., Bird *et al.* (1988) *Science* 242:423-426; Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5879-5883; McCafferty *et al.*, *Nature* (1990) 348:552-554).

To express the antibodies, or antibody portions used in the invention, DNAs encoding partial or full-length light and heavy chains, obtained as described above, are inserted into expression vectors such that the genes are operatively linked to transcriptional and translational control sequences. In this context, the term "operatively linked" is intended to mean that an antibody gene is ligated into a vector such that transcriptional and translational control sequences within the vector serve their intended function of regulating the transcription and translation of the antibody gene. The expression vector and expression control sequences are chosen to be compatible with the expression host cell used. The antibody light chain gene and the antibody heavy chain gene can be inserted into separate vector or, more typically, both genes are inserted into the same expression vector. The antibody genes are inserted into the expression vector by standard methods (*e.g.,* ligation of complementary restriction sites on the antibody gene fragment and vector, or blunt end ligation if no restriction sites are present). Prior to insertion of the adalimumab, or an adalimumab-related light or heavy chain sequences, the expression vector may already carry antibody constant region sequences. For example, one approach to converting the adalimumab, or an adalimumab-related VH and VL sequences to full-length antibody genes is to insert them into expression vectors already encoding heavy chain constant and light chain constant regions, respectively, such that the VH segment is operatively linked to the CH segment(s) within the vector and the VL segment is operatively linked to the CL segment within the vector. Additionally or alternatively, the recombinant expression vector can encode a signal peptide that facilitates secretion of the antibody chain from a host cell. The antibody chain gene can be cloned into the vector such that the signal peptide is linked in-frame to the amino terminus of the antibody chain gene. The signal peptide can be an immunoglobulin signal peptide or a heterologous signal peptide (*i.e.*, a signal peptide from a non-immunoglobulin protein).

In addition to the antibody chain genes, the recombinant expression vectors of the invention carry regulatory sequences that control the expression of the antibody chain genes in a host cell. The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.,* polyadenylation signals) that control the transcription or translation of the antibody chain genes. Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). It will be appreciated by those skilled in the art that the design of

the expression vector, including the selection of regulatory sequences may depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, *etc.* Preferred regulatory sequences for mammalian host cell expression include viral elements that direct high levels of protein expression in mammalian cells, such as promoters and/or enhancers derived from cytomegalovirus (CMV) (such as the CMV promoter/enhancer), Simian Virus 40 (SV40) (such as the SV40 promoter/enhancer), adenovirus, (*e.g.,* the adenovirus major late promoter (AdMLP)) and polyoma. For further description of viral regulatory elements, and sequences thereof, see *e.g.,* U.S. Patent No. 5,168,062 by Stinski, U.S. Patent No. 4,510,245 by Bell *et al.* and U.S. Patent No. 4,968,615 by Schaffner *et al.*

In addition to the antibody chain genes and regulatory sequences, the recombinant expression vectors used in the invention may carry additional sequences, such as sequences that regulate replication of the vector in host cells (*e.g.,* origins of replication) and selectable marker genes. The selectable marker gene facilitates selection of host cells into which the vector has been introduced (see *e.g.,* U.S. Patents Nos. 4,399,216, 4,634,665 and 5,179,017, all by Axel *et al.*). For example, typically the selectable marker gene confers resistance to drugs, such as G418, hygromycin or methotrexate, on a host cell into which the vector has been introduced. Preferred selectable marker genes include the dihydrofolate reductase (DHFR) gene (for use in $dhfr^-$ host cells with methotrexate selection/amplification) and the *neo* gene (for G418 selection).

For expression of the light and heavy chains, the expression vector(s) encoding the heavy and light chains is transfected into a host cell by standard techniques. The various forms of the term "transfection" are intended to encompass a wide variety of techniques commonly used for the introduction of exogenous DNA into a prokaryotic or eukaryotic host cell, *e.g.,* electroporation, calcium-phosphate precipitation, DEAE-dextran transfection and the like. Although it is theoretically possible to express the antibodies of the invention in either prokaryotic or eukaryotic host cells, expression of antibodies in eukaryotic cells, and most preferably mammalian host cells, is the most preferred because such eukaryotic cells, and in particular mammalian cells, are more likely than prokaryotic cells to assemble and secrete a properly folded and immunologically active antibody. Prokaryotic expression of antibody genes has been reported to be ineffective for production of high yields of active antibody (Boss, M.A. and Wood, C. R. (1985) *Immunology Today* 6:12-13).

Preferred mammalian host cells for expressing the recombinant antibodies of the invention include Chinese Hamster Ovary (CHO cells) (including dhfr- CHO cells, described

in Urlaub and Chasin, (1980) *Proc. Natl. Acad. Sci. USA* 77:4216-4220, used with a DHFR selectable marker, *e.g.,* as described in R.J. Kaufman and P.A. Sharp (1982) *Mol. Biol.* 159:601-621), NS0 myeloma cells, COS cells and SP2 cells. When recombinant expression vectors encoding antibody genes are introduced into mammalian host cells, the antibodies are produced by culturing the host cells for a period of time sufficient to allow for expression of the antibody in the host cells or, more preferably, secretion of the antibody into the culture medium in which the host cells are grown. Antibodies can be recovered from the culture medium using standard protein purification methods.

Host cells can also be used to produce portions of intact antibodies, such as Fab fragments or scFv molecules. It is understood that variations on the above procedure are within the scope of the present invention. For example, it may be desirable to transfect a host cell with DNA encoding either the light chain or the heavy chain (but not both) of an antibody of this invention. Recombinant DNA technology may also be used to remove some or all of the DNA encoding either or both of the light and heavy chains that is not necessary for binding to hTNFα. The molecules expressed from such truncated DNA molecules are also encompassed by the antibodies of the invention. In addition, bifunctional antibodies may be produced in which one heavy and one light chain are an antibody of the invention and the other heavy and light chain are specific for an antigen other than hTNFα by crosslinking an antibody of the invention to a second antibody by standard chemical crosslinking methods.

In a preferred system for recombinant expression of an antibody, or antigen-binding portion thereof, of the invention, a recombinant expression vector encoding both the antibody heavy chain and the antibody light chain is introduced into dhfr-CHO cells by calcium phosphate-mediated transfection. Within the recombinant expression vector, the antibody heavy and light chain genes are each operatively linked to CMV enhancer/AdMLP promoter regulatory elements to drive high levels of transcription of the genes. The recombinant expression vector also carries a DHFR gene, which allows for selection of CHO cells that have been transfected with the vector using methotrexate selection/amplification. The selected transformant host cells are culture to allow for expression of the antibody heavy and light chains and intact antibody is recovered from the culture medium. Standard molecular biology techniques are used to prepare the recombinant expression vector, transfect the host cells, select for transformants, culture the host cells and recover the antibody from the culture medium.

In view of the foregoing, nucleic acid, vector and host cell compositions that can be used for recombinant expression of the antibodies and antibody portions used in the invention

include nucleic acids, and vectors comprising said nucleic acids, comprising the human TNFα antibody adalimumab (D2E7). The nucleotide sequences encoding the adalimumab light and heavy chain variable regions are known in the art, *e.g.,* see U.S. Patent No. 6,090,382. It will be appreciated by the skilled artisan that nucleotide sequences encoding adalimumab-related antibodies, or portions thereof (*e.g.,* a CDR domain, such as a CDR3 domain), can be derived from the nucleotide sequences encoding the adalimumab LCVR and HCVR using the genetic code and standard molecular biology techniques.

Methods of isolating human neutralizing antibodies with high affinity and a low off rate constant for hTNFα are described in U.S. Patent Nos. 6,090,382, 6,258,562, and 6,509,015, each of which is incorporated by reference herein.

Antibodies, antibody-portions, and other TNFα inhibitors for use in the methods of the invention, can be incorporated into pharmaceutical compositions suitable for administration to a subject. Typically, the pharmaceutical composition comprises an antibody, antibody portion, or other TNFα inhibitor, and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like that are physiologically compatible. Examples of pharmaceutically acceptable carriers include one or more of water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. In many cases, it is preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Pharmaceutically acceptable carriers may further comprise minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives or buffers, which enhance the shelf life or effectiveness of the antibody, antibody portion, or other TNFα inhibitor.

Therapeutic compositions typically must be sterile and stable under the conditions of manufacture and storage. The composition can be formulated as a solution, microemulsion, dispersion, liposome, or other ordered structure suitable to high drug concentration. Sterile injectable solutions can be prepared by incorporating the active compound (*i.e.*, antibody, antibody portion, or other TNFα inhibitor) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable

solutions, the preferred methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof. The proper fluidity of a solution can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prolonged absorption of injectable compositions can be brought about by including in the composition an agent that delays absorption, for example, monostearate salts and gelatin.

In one embodiment, the invention includes pharmaceutical compositions comprising an effective TNFα inhibitor and a pharmaceutically acceptable carrier, wherein the effective TNFα inhibitor may be used to treat IBD.

IV. Kits of Invention

The invention also provides kits for assessing a subject's responsiveness to a TNFα inhibitor for the treatment of IBD, *e.g.,* Crohn's disease or ulcerative colitis, as well as kits for treating a subject having an IBD, *e.g.,* Crohn's disease or ulcerative colitis. Thus, kits of the invention can be used to determine if a subject with IBD, *e.g.,* Crohn's disease or ulcerative colitis, will be effectively responsive to a TNFα inhibitor. These kits may comprise a carrier means being compartmentalized to receive in close confinement one or more container means such as vials, tubes, and the like, each of the container means comprising one of the separate elements to be used in the method. For example, one of the container means may comprise a probe that is or can be detectably labeled directed to one of the genes in the gene signature described herein. Such probe may be an antibody or polynucleotide specific for a protein or gene transcript, respectively. Where the kit utilizes nucleic acid hybridization to detect the target nucleic acid, the kit may also have containers containing nucleotide(s) for amplification of the target nucleic acid sequence and/or a container comprising a reporter-means, such as a biotin-binding protein, *e.g.,* avidin or streptavidin, bound to a reporter molecule, such as an enzymatic, florescent, or radioisotope label.

In certain embodiments of the invention, the kit includes a means for determining the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 and instructions for use of the kit. Means for determining expression levels may include nucleic acid molecules or antibodies, as described in more detail below.

Kits of the invention can be used to determine if a subject with Crohn's disease will be effectively responsive to a TNFα inhibitor. These kits may comprise a carrier means being compartmentalized to receive in close confinement one or more container means such as vials, tubes, and the like, each of the container means comprising one of the separate elements to be used in the method. For example, one of the container means may comprise a probe that is or can be detectably labeled. Such probe may be an antibody or polynucleotide specific for a protein or a biomarker (CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6) gene or message, respectively. Where the kit utilizes nucleic acid hybridization to detect the target nucleic acid, the kit may also have containers containing nucleotide(s) for amplification of the target nucleic acid sequence and/or a container comprising a reporter-means, such as a biotin-binding protein, *e.g.,* avidin or streptavidin, bound to a reporter molecule, such as an enzymatic, florescent, or radioisotope label.

Such kit will typically comprise the container described above and one or more other containers comprising materials desirable from a commercial and user standpoint, including buffers, diluents, filters, needles, syringes, and package inserts with instructions for use. A label may be present on the container to indicate that the composition is used for a specific application, and may also indicate directions for either in vivo or in vitro use, such as those described above.

The kits of the invention have a number of embodiments. A typical embodiment is a kit comprising a container, a label on the container, and a composition contained within the container, wherein the composition includes one or more polynucleotides that hybridize to a complement of the CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 under stringent conditions, and the label on the container indicates that the composition can be used to evaluate the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 in a sample, and wherein the kit includes instructions for using the polynucleotide(s) for evaluating the expression level(s) of the CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 RNA or DNA in a particular sample type.

Another aspect is a kit comprising a container, a label on the container, and a composition contained within the container, wherein the composition includes a primary antibody that binds to a protein or autoantibody biomarker, and the label on the container indicates that the composition can be used to evaluate the presence of such proteins or antibodies in a sample, and wherein the kit includes instructions for using the antibody for evaluating the presence of biomarker proteins in a particular sample type. The kit can further comprise a set of instructions and materials for preparing a sample and applying antibody to the sample. The kit may include both a primary and secondary antibody, wherein the secondary antibody is conjugated to a label, *e.g.,* an enzymatic label.

Other optional components of the kit include one or more buffers (*e.g.,* block buffer, wash buffer, substrate buffer, etc.), other reagents such as substrate (*e.g.,* chromogen) that is chemically altered by an enzymatic label, epitope retrieval solution, control samples (positive and/or negative controls), control slide(s), etc. Kits can also include instructions for interpreting the results obtained using the kit.

In further specific embodiments, for antibody-based kits, the kit can comprise, for example: (1) a first antibody (*e.g.,* attached to a solid support) that binds to a biomarker protein; and, optionally, (2) a second, different antibody that binds to either the protein or the first antibody and is conjugated to a detectable label.

For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, *e.g.,* a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a biomarker protein or (2) a pair of primers useful for amplifying a biomarker nucleic acid molecule. The kit can also comprise, *e.g.,* a buffering agent, a preservative, or a protein-stabilizing agent. The kit can further comprise components necessary for detecting the detectable label (*e.g.,* an enzyme or a substrate). The kit can also contain a control sample or a series of control samples that can be assayed and compared to the test sample. Each component of the kit can be enclosed within an individual container, and all of the various containers can be included within a single package, along with instructions for interpreting the results of the assays performed using the kit.

The kits of the invention may optionally comprise additional components useful for performing the methods of the invention. By way of example, the kits may comprise means for obtaining a biological sample from a subject, a control sample, *e.g.,* a sample from a subject, one or more sample compartments, an instructional material which describes performance of a method of the invention and specific controls/standards.

The instructions can be, for example, printed instructions for performing the assay for evaluating the results.

The means for isolating a biological sample from a subject can comprise one or more reagents that can be used to obtain a fluid or tissue from a subject. The means for obtaining a biological sample from a subject may also comprise means for isolating peripheral blood mononuclear cells from a blood sample, for example by positive selection of the monocytes or by negative selection in which all other cell types other than monocytes are removed.

The kits of the invention may further a TNFα inhibitor.

Preferably, the kit is designed for use with a human subject.

Also provided by the invention are articles of manufacture containing materials useful for the treatment of Crohn's disease (Crohn's disease). The article of manufacture comprises a container and a label or package insert on or associated with the container. In this aspect, the package insert is on or associated with the container. Suitable containers include, for example, bottles, vials, syringes, etc. The containers may be formed from a variety of materials such as glass or plastic. The container holds or contains the antagonist that is effective for treating Crohn's disease and may have a sterile access port (for example, the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). At least one active agent in the composition is the B-cell antagonist. The label or package insert indicates that the composition is used for treating IBD, *e.g.,* Crohn's disease or ulcerative colitis, in a subject eligible for treatment with specific guidance regarding dosing amounts and intervals of antagonist and any other medicament being provided.

The kits and articles of manufacture herein also include information, for example in the form of a package insert or label, indicating that the composition is used for treating IBD, *e.g.,* Crohn's disease or ulcerative colitis, where the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6 herein are detected in a genetic sample from the patient with the disease. The insert or label may take any form, such as paper or electronic media, for example, a magnetically recorded medium (*e.g.,* floppy disk) or a CD-ROM. The label or insert may also include other information concerning the pharmaceutical compositions and dosage forms in the kit or article of manufacture.

Generally, such information aids patients and physicians in using the enclosed pharmaceutical compositions and dosage forms effectively and safely. For example, the following information regarding the antagonist may be supplied in the insert:

pharmacokinetics, pharmacodynamics, clinical studies, efficacy parameters, indications and usage, contraindications, warnings, precautions, adverse reactions, overdosage, proper dosage and administration, how supplied, proper storage conditions, references, and patent information.

In a specific embodiment of the invention, an article of manufacture is provided comprising, packaged together, a pharmaceutical composition comprising a TNFα inhibitor and a pharmaceutically acceptable carrier and a label stating that the inhibitor or pharmaceutical composition is indicated for treating patients with IBD, *e.g.,* Crohn's disease or ulcerative colitis, from which a genetic sample has been obtained showing the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6. This can be shown by assessing genetic expression as a biomarker of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6.

Also the invention provides a method for manufacturing a TNFα inhibitor or a pharmaceutical composition thereof comprising combining in a package the TNFα inhibitor or pharmaceutical composition and a label stating that the TNFα inhibitor or pharmaceutical composition is indicated for treating patients with IBD, *e.g.,* Crohn's disease or ulcerative colitis, from which a genetic sample has been obtained showing the expression level(s) of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6. The label may further state that this can be shown by assessing genetic expression as a biomarker of CSF1R, CSF3R, CCL18, and IL1B; CSF1R, CSF3R, CD36, and IL1B; CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, AND TREM1; and/or TNFAIP6. Notably, each of the genetic markers identified in the invention may be described individually or in combination with one another.

Any compositions or methods provided herein can be combined with one or more of any of the other compositions and methods provided herein.

Ranges provided herein are understood to be shorthand for all of the values within the range. For example, a range of 1 to 50 is understood to include any number, combination of numbers, or sub-range from the group consisting 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50.

The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference

This invention is further illustrated by the following example, which should not be construed as limiting.

EXAMPLES

The following examples describe the identification of gene signatures that may be used, for example, to predict whether a patient having IBD (*e.g.*, a Crohn's disease patient) will respond or not respond to treatment with a TNFα inhibitor, *e.g.,* adalimumab.

To identify a gene signature(s) separating responders and non-responders to anti-TNFα treatment in IBD patients, various pathways, microarray analysis, and statistical/machine-learning were collectively considered. The gene signatures described in the Examples are based on the biology of myeloid cells in IBD, and provide predictive accuracy which was confirmed by at least two independent cohorts of studies. .As described in the Examples below, four different gene signatures were identified that showed significant differences between responders and non-responders of IBD patients to anti-TNF therapy. These gene signatures were identified by performing global gene expression analyses in a training cohort (GSE16879), and were then cross-validated in two independent cohorts (Test Sets #1 and #2). The results described in Examples 1 - 5 provide accurate and effective predictors to identify TNF responders and non-responders. The four identified gene signatures can be used as markers to identify patients having IBD (including Crohn's disease or ulcerative colitis) who will be responsive to treatment with a TNFα inhibitor (such as adalimumab).

MATERIALS AND METHODS

The materials and methodologies used in Examples 1-5 are described below.

Patient data sets

In Examples 1 - 5, microarray data were derived from the following three independent cohorts of Crohn's disease (CD) and ulcerative colitis (UC): GSE16879 (accessible through GEO series accession number GSE16879 at NCBI's Gene Expression Omnibus; dataset for CD and UC; see Arijs *et al.* (2009)), GSE52746 (accessible through GEO series accession number GSE52746 at NCBI's Gene Expression Omnibus), and a test data set provided by UMASS.

For the training dataset GSE16879, nineteen patients with active CD who were refractory to corticosteroids and/or immunosuppression and 6 control patients were included

for this study. The clinical data for the CD and UC patients provided by the literature is included in Arijs *et al.* (2009).

For CD, inclusion criteria for this study are CD patients were examined by endoscopy with biopsies from diseased colon. All patients underwent endoscopy with median CD Activity Index =267 (240 – 409.5), and the presence of median C-reactive protein level = 10.2 (4.3-35) mg/dl at first drug infusion. In addition, simplified histology score were determined in all patients, in which histology scores were graded 0, 1, 2 and 3 if they were normal, mild, moderate and severe, respectively. Patients were intravenously injected 5 mg infliximab per kg body weight, followed by a second endoscopy with biopsies 4 weeks after the first infliximab infusion for a single infusion or at 6 weeks for a series loading doses of infliximab at weeks 0, 2 and 6. The response to infliximab was assessed 4 and 6 weeks after the first infliximab treatment. CD responders to infliximab were defined as a complete mucosal healing with a decrease of at least 3 points on the histological score for CD. CD patients did not achieve this complete mucosal healing but with some endoscopic and/or histologic improvement were defined as non-responders. Biopsy samples were taken at sites of active, then immediately snap-frozen in liquid nitrogen until needed for RNA isolation and/or immunohistochemistry.

For the test dataset GSE52746, CD patients diagnosed with 4 months' disease duration were included in this study and controls with non-inflammatory bowel disease (non-IBD). The clinical data for the CD patients provided by the literature is included in Leal *et al.* (2014). Non-CD controls were selected for mild gastrointestinal symptoms or a screening colonoscopy for colorectal cancer. All patients were assessed according to the Montreal classification with CD Endoscopic Index of Severity (CDEIS) of ≥5 (5.4-20.5), CD Activity Index >140 (142-189), and the presence of large ulcers (>0.5 cm diameter) in at least one of the explored segments. The exclusion categories were: 1. patients with concomitant infections; 2. Pharmacokinetics of anti-TNF biologics with low levels of drug (infliximab ≤2 μg/mL; adalimumab ≤4 μg/mL) or existence of any levels of anti-drug antibodies. Adalimumab was given subcutaneously 160/80 mg at weeks 0 and 2 for induction and 40 mg every other week for maintenance. Infliximab was given intravenously (5 mg/kg) at weeks 0, 2 and 6 for induction and every 8 weeks for maintenance. Response to anti-TNFα therapy was assessed by endoscopy at 12 weeks after first drug administration. Responders to anti-TNF therapy were defined as CDEIS <5 and an absence of ulcers, whereas patients who did not achieve such improvement were defined as non-responders. Biopsy samples were immediately snap-frozen

in RNA later-RNA Stabilization Reagent (QIAGEN) and kept in liquid nitrogen until needed for RNA isolation and/or immunohistochemistry.

An additional test data set was a dataset provided by UMASS Medical School, Worcester, MA, and was obtained from CD patients. Crohn's disease samples were a mix of treatment naïve and patients who were on the CD standard of care who then showed diagnostic improvement or failed anti-TNF treatment.

RNA Preparation from FFPE tissue blocks

The following RNA preparation protocol was used to obtain RNA from samples for the bioinformatics analysis described in Examples 2-4 for the UMASS dataset.

Archival Formalin–Fixed, Paraffin-Embedded (FFPE) tissue samples represent a potentially invaluable source for gene expression analysis data, but with a potential caveat due to cross-linked RNA leading to ineffectively serve as a substrate for subsequent enzyme reactions. Thus, high quality RNA was extracted from FFPE tissue sections using High Pure RNA Paraffin Kit (Cat. No. 03270289001, Roche, USA) following the manufacture's protocol. 5µm FFPE tissue section was treated with xylene. One hundred percent ethanol was then added to extract any residual xylene from the sample. Next, samples were incubated in an optimized lysis buffer, which contained proteinase K, to release RNA from the sections. The DNase treatment that followed was designed to eliminate all genomic DNA, including very small fragments that could have been present in FFPE samples. Appropriate binding conditions for RNA were created by adding ethanol. The sample was then transferred to a filter tube and spin column to elute RNA with RNase-free water, contaminants being efficiently washed away.

Total RNA concentration was measured using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, DE) and RNA purity and integrity was verified using la-on-chip technology (Agilent 2100 Bioanalyzer, Santa Clara, CA, USA). RNA was kept at -80 ^{o}C until needed for microarray analysis.

Gene expression data generation

Gene expression data generation was performed in Examples 2-4 for the UMASS dataset , as described below.

Tissue samples prepared according to the above FFPE protocol were analyzed using microarray analysis to obtain gene expression profiles. The Standard Whole-Genome DASL (cDNA-mediated Annealing, Selection, extension, and Ligation) HT Assay label protocol

(WGDASL HT Assay Guide (15018211 D) (Illumina, Inc.) was carried out using the WG-DASL HT Assay Profiling Reagent Kit. In accordance with the Whole-Genome DASL assay, total RNA was converted into biotinylated cDNA through a reverse transcription reaction using biotinylated oligo (dT) and random nonamer primers. The biotinylated cDNA was then annealed with assay-specific oligonucleotides (ASO) specially designed for a single contiguous 50 nucleotide sequence on each cDNA (provided with the Illumina Gene Chip kit). The resulting products were PCR amplified and labeled with a universal fluorescently labeled primer.

The resulting PCR products were then hybridized to the single-stranded labeled products were then hybridized on the complementary gene specific sequence bead to Illumina Whole-Genome Gene Expression Human HT-12 v4 BeadChips and scanned with the iScan™ Reader. The bead chips scan generated intensity data files (.idat files) for each sample, each file containing raw intensity data values for every bead in the scanned image. Bead summary intensities were log2-transformed and normalized using quantile normalization imbedded in Illumina BeadStudio v3.1.3 with GeneExpression plugin v3.4.0 and GenomeStudio 2010.1 with GeneExpression plugin v1.0 were used to generate reference data (Illumina, San Diego, CA, USA).

Bioinformatics analysis of whole tissue gene expression profiling

Genome-wide transcriptome analysis on Crohn's disease patients was derived from three independent experiments. For the public datasets analyzed, all external dataset GSE16879 and GSE52746 were downloaded as compressed *.CEL file packages from the GEO database, and normalized and background-corrected using Robust Multi-array Average (*RMA*) algorithm. The data were analyzed for differential gene expression using an empirical Bayes moderated t-test, implemented in the Bioconductor package Linear Models for Microarray Data LIMMA (www.bioconductor.org). The results were sorted by the adjusted p-value and fold change of gene expression.

EXAMPLE 1: Bioinformatics analysis of anti-TNFα non-responder using immune cell lineage gene signature

Datasets were obtained based on the differential gene expression profiles among multiple human immune cell lineages using a customized gene chip. The generated datasets contained 136 human immune cell samples for 82 cell lineage gene expression datasets. Using

this human immune cell database, multiple human immune cell lineage gene signatures, including T, B, NK, myeloid cells, were obtained using a multivariate comparison test implemented in the Bioconductor package Linear Models for Microarray Data LIMMA. Three fold change cutoff was chosen to select T, B and NK cell lineage signatures while 5-fold change cutoff to generate myeloid cell lineage signature. The final myeloid cell signature contained the most overexpressed 64 genes compared to all other immune cell gene expression.

Changes in transcript levels in the anti-TNFα CD responder and non-responder populations could reflect alterations in the hematopoietic cell composition of the leukocytic infiltrate and/or modifications in gene expression programs. To bioinformatically address this issue, a cell distribution profile was generated using GSE16879 dataset, plotting for each of the most differentially over-expressed genes in each human immune cell lineages using its normalized expression value in a panel of cell-types. The set of genes whose transcript levels were most significantly upregulated in anti-TNFα non-responder were characteristic of cells of the myeloid lineage, as opposed to lymphoid cell lineages, such as T and B cell lineages.

In order to define the genome wide phenotype of anti-TNF CD non-responders with the goal to stratify CD patients who would not respond to anti-TNF therapy, a manually curated gene signature was selected from myeloid cell lineage signature using hierarchical clustering with Euclidean distance metric and selecting the genes with most significant fold changes and statistical difference between non-responders and responders. The final 9 representative signature genes (CLEC4A: C-type lectin domain family 4, member A; CHI3L1: chitinase 3-like 1 (cartilage glycoprotein-39); MMP7: matrix metallopeptidase 7 (matrilysin, uterine); CLEC7A: C-type lectin domain family 7, member A; IL1B: interleukin 1, beta; CSTA: cystatin A (stefin A); IL1RN: interleukin 1 receptor antagonist; TNFAIP6: tumor necrosis factor, alpha-induced protein 6; TREM1: triggering receptor expressed on myeloid cells 1) were most significantly differentially upregulated in GSE16879 training dataset between anti-TNF CD non-responders and responders.

Examples 2 to 4 describe multivariate data analysis based on gene set enrichment analysis (GSEA) of the 64 genes identified in the examination of human myeloid cells in CD patients. Gene Set Enrichment Analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states (*e.g.*, IBD TNFα inhibitor responder phenotype).

EXAMPLE 2: Multivariate Data Analysis and Identification of Gene Signature #1

The following Example describes the identification of a gene signature (Gene Signature #1) that can separate responders from non-responders to treatment with a TNF inhibitor in IBD patients, specifically Crohn's and ulcerative colitis patients. The approach considered Gene Set Enrichment Analysis (GSEA) pathway enrichment analysis (Subramanian 2005) to the myeloid 62 gene cell signature that was identified based on human immune cell subsets, and described in Example 1. The top pathways were ranked by p-value, followed by application of a Monte-Carlo based logistic regression with lasso to the 2nd most enriched pathway consisting of 9 input genes to the Crohn's disease before treatment (BT) data (training set, GEO GSE16879, Arijs *et al* 2009), as described in Example 1. This analysis resulted in a signature of 4 genes: CSF1R, CSF3R, CCL18, and IL1B.

In this algorithm, the lasso penalty parameter estimate from the logistic-lasso model (Friedman *et al* 2010) that controls the feature selection process was refined by a Monte-Carlo resampling procedure. The gene expressions were first normalized by forming the $\log_2$ RMA normalized expression values of the probe sets. This procedure entails generating 50 bootstrap replications of the normalized dataset and refitting the logistic-lasso model to estimate the lasso penalty parameter. The median of all the lasso penalty parameter estimates across the 50 Monte-Carlo replications was then used to finalize the lasso penalty parameter estimate in the logistic regression model, which in turn finalizes the number of features to be used in the signature.

The logistic regression coefficients of the final list of features from the above algorithm were then used to create a linear composite score. The linear composite score is a weighted sum of the normalized expressions of CSF1R, CSF3R, CCL18, and IL1B. Specifically, our composite score (LC1) can be calculated for Affymetrix Human Genome U133 Plus 2.0 Arrays as: *0.18 * 203104_at (CSF1R) +*
*0.68 * 1553297_a_at (CSF3R) + 0.36 * 32128_at (CCL18) + 0.53 * 205067_at (IL1B).*

The performance of this 4 gene signature was assessed on 3 independent datasets (test sets): a public dataset on Crohn's disease patients (Test Set #1, GEO GSE52746; Affymetrix), a dataset from UMASS on Crohn's disease patients (Illumina; Test Set #2), and a dataset consisting of ulcerative colitis patients (GEO Affymetrix GSE16879, Arijs 2009; Test Set #3).

The area under the receiver operator curve (AUC), which is a commonly used metric to report predictive performance (Hanley 1982, Zou 2007, and Swets 1988), was used to measure the performance of our signature on these datasets. The value of AUC is reflective of the specificity and sensitivity of the marker, with values of 0.7-0.8, 0.8-0.9, and 0.9-1.0 representing "acceptable", "excellent", and "outstanding" discriminatory potential, respectively.

An AUC of 1 was obtained for the Test Set #1, an AUC of 0.79 was obtained for Test Set #2, and an AUC of 0.86 was obtained for the ulcerative colitis dataset (Test Set #3) was obtained. In view of the AUC scores, the gene signature comprising CSF1R, CSF3R, CCL18, and IL1B was determined to be predictive of response to treatment.

Figures 1 and 2 illustrate the separation of responders versus non-responders for the training set for normalized gene expressions and for the composite score respectively.

Figures 3-11 illustrate the performance in the three independent Test-Sets 1 to 3 (public dataset, private dataset: UMASS, and an ulcerative colitis dataset, respectively).

EXAMPLE 3: Multivariate Data Analysis and Identification of Gene Signature #2

As in Example 2, the same resampling-based logistic regression with lasso was applied to the most enriched GSEA pathway consisting of 6 input genes to the Crohn's disease before-treatment data (training set, GEO GSE16879, Arijs 2009) to obtain a second signature of 4 genes: CSF1R, CSF3R, CD36, and IL1B.

The linear composite score is a weighted sum of the normalized expressions of these four genes: CSF1R, CSF3R, CD36, and IL1B. Specifically, the composite score (LC2) can be calculated for Affymetrix Human Genome U133 Plus 2.0 Arrays as:*2.31 * 203104_at (CSF1R) + 3.05 * 1553297_a_at (CSF3R) + 0.39 * 206488_s_at (CD36) + 1.18 * 205067_at (IL1B).* Upon testing the performance of this composite score (Signature 2) in the three independent test-sets, an AUC of 0.93 for the public dataset (Test Set #1; GSE52746), an AUC of 0.75 for the UMASS dataset (Test Set #2), and an AUC of 0.87 for the ulcerative colitis dataset (Test Set #3; GSE16879) was obtained. In view of the AUC scores, the gene signature comprising CSF1R, CSF3R, CD36, and IL1B was determined to be predictive of response to treatment.

Figures 12 and 13 illustrate the univariate separation of responders versus non-responders for the training set for normalized gene expressions and for the composite score respectively.

Figures 14-22 illustrate the performance in the three independent Test-Sets 1 to 3 (public dataset, UMASS dataset, and an ulcerative colitis dataset, respectively).

EXAMPLE 4: Multivariate Data Analysis and Identification of Gene Signature #3

The elastic-net method from the R package glmnet (function glmnet) was applied to Crohn's disease before-treatment data (training set, GEO GSE16879, Arijs 2009) focusing on an initial manually curated list of 9 candidate genes: CHI3L1, CLEC4A, CLEC7A, CSTA, IL1B, IL1RN, MMP7, TNFAIP6, and TREM1. The elastic net mixing parameter was set to 1/4 and the penalty coefficient, lambda, was selected by BIC (the Bayesian information criterion). A subset of 8 genes (TNFAIP6, TREM1, IL1RN, CLEC4A, IL1B, CHI3L1, CLEC7A, and CSTA) was selected for model construction.

The final signature is a weighted sum of the normalized expressions of the 8 genes. Specifically, the composite score (LC_3) can be calculated for Affymetrix Human Genome U133 Plus 2.0 Arrays as:*1.489 * 206025_s_at (TNFAIP6) + 0.949 * 219434_at (TREM1) + 0.536 * 212657_s_at (IL1RN) + 1.145 * 219947_at (CLEC4A) + 0.943 * 205067_at (IL1B) + 1.400 * 209395_at (CHI3L1) + 1.537 * 221698_s_at (CLEC7A) + 0.382 * 204971_at (CSTA).* Upon testing the performance of this composite score (Signature 3) in the three independent test-sets, an AUC of 0.93 for the public dataset (GSE52746), an AUC of 0.84 for the dataset provided by UMASS, and an AUC of 0.90 for the ulcerative colitis dataset was obtained (GSE16879). In view of the AUC scores, the gene signature comprising TNFAIP6, TREM1, IL1RN, CLEC4A, IL1B, CHI3L1, CLEC7A, and CSTA was determined to be predictive of response to treatment.

Figures 23-24 illustrate the univariate separation of responders versus non-responders for the training set for normalized gene expressions and for the composite score respectively.

Figures 25-33 illustrate the performance in the three independent test sets 1 to 3 (public dataset, private dataset: UMASS, and an ulcerative colitis dataset, respectively).

In sum, for the eight gene signature, an AUC of 1 was reached in a training cohort, which means that this test correctly classifies 100% of 7 patients of randomly drawn pairs.

Based on these data, the panel of genes signatures was chosen to predict non-response to anti-TNFα inhibitors with a specificity of around 84% to 90% based on the data sets tested.

EXAMPLE 5: Multivariate Data Analysis and Identification of TNFAIP6 as Predictive Marker

The same Monte-Carlo based logistic regression with lasso was applied to the Crohn's disease before-treatment data (training set, GEO GSE16879, Arijs 2009) focusing on an initial list of 9 candidate genes: CHI3L1, CLEC4A, CLEC7A, CSTA, IL1B, IL1RN, MMP7, TNFAIP6, and TREM1 to obtain a gene signature of a single gene, TNFAIP6. Specifically, the composite score (LC_4) is simply the normalized expression value of 206025_s_at. An AUC of 0.87 for the public dataset (GSE52746), an AUC of 0.84 for the UMASS dataset, and an AUC of 0.88 for the ulcerative colitis dataset was obtained (GSE16879). In view of the AUC scores, TNFAIP6 was determined to be predictive of response to treatment.

Figures 34-35 illustrate the univariate separation of responders versus non-responders for the training set for normalized gene expressions and for the composite score respectively.

Figures 36-41 illustrate the performance in three independent test sets 1 to 3 (public dataset, UMASS dataset, and an ulcerative colitis dataset, respectively).

Table 1: Key to probes mentioned in above examples and figures

Probe	Gene
203104_at	CSF1R
1553297_a_at	CSF3R
32128_at	CCL18
205067_at	IL1B
206488_s_at	CD36
206025_s_at	TNFAIP6
219434_at	TREM1
212657_s_at	IL1RN

Probe	Gene
219947_at	CLEC4A
Probe	Gene
209395_at	CHI3L1
221698_s_at	CLEC7A
204971_at	CSTA

References

Arijs I, De Hertogh G, Lemaire K, Quintens R et al, "Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment", PLoS One, Vol. 4(11): e7984, 1-10, Nov 24, 2009.

Leal, R.F. *et al.*, "Identification of inflammatory mediators in patients with Crohn's disease unresponsible to anti-TNFa therapy", Gut, 0:1-10, 2014 (published online first on April 3, 2014 as 10.1136/gutjnl-2013-306518).

Friedman J., Hastie T., and Tibshirani, R., *"Regularization Paths for Generalized Linear Models via Coordinate Descent",* Journal of Statistical Software, Vol. 33(1), 1-22, Feb 2010.

Subramanian A., Tamayo P. et al, "Gene set enrichment analysis: A knowledge based approach for interpreting genome wide expression profiles," PNAS, Vol 102 (43), 15545-15550, 2005.

Friedman J., Hastie T., and Tibshirani, R., *"Regularization Paths for Generalized Linear Models via Coordinate Descent", Journal of Statistical Software,* Vol. 33(1), 1-22 Feb 2010.

Hanley JA and McNeil BJ, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," Radiology, 143:29–36, 1982.

Subramanian A., Tamayo P. et al, "Gene set enrichment analysis: A knowledge based approach for interpreting genome wide expression profiles," PNAS, Vol 102 (43), 15545-15550, 2005.

Swets JA, "Measuring the accuracy of diagnostic systems," Science, 240: 1285–93, 1988.

Zou KH, O'Malley AJ, Mauri L, "Receiver operating characteristic analysis for evaluation diagnostic tests and predictive models," Circulation, 115: 654–57, 2007.

INCORPORATION BY REFERENCE

The contents of all references, patents, pending patent applications and published patents, cited throughout this application are hereby expressly incorporated by reference.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

CLAIMS

What is claimed is:

1. A method for determining the responsiveness of a subject having an inflammatory bowel disease (IBD) to treatment with a TNFα inhibitor, the method comprising

determining a gene signature from a biological sample from a subject having an IBD, wherein the gene signature is obtained by determining an expression level(s) of a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6; and

comparing the gene signature with a responder gene signature comprising a corresponding gene or gene group of (i) to (iv), wherein the subject will be responsive to treatment with the TNFα inhibitor if the gene signature correlates with the responder gene signature.

2. A method for determining the responsiveness of a subject having an IBD to treatment with a TNFα inhibitor, the method comprising

determining a gene signature from a biological sample from a subject having an IBD, wherein the gene signature is obtained by determining an expression level(s) of a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6; and

comparing the gene signature with a non-responder gene signature comprising a corresponding gene or gene group of (i) to (iv), wherein the subject will not be responsive to treatment with the TNFα inhibitor if the gene signature correlates with the non-responder gene signature.

3. A method for treating a subject having an IBD, said method comprising

determining a gene signature from a biological sample from a subject having an IBD, wherein the gene signature is obtained by determining an expression level(s) of a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6, and

administering a TNFα inhibitor to the subject having IBD provided that the gene signature correlates with a responder gene signature comprising a corresponding gene or gene group of (i) to (iv).

4. The method of any one of claims 1-3, wherein the biological sample is a tissue sample.

5. The method claim 4, wherein the tissue sample is from an intestine of the subject.

6. The method of any one of claims 1-3, wherein the biological sample is a blood sample or a stool sample.

7. The method of any one of claims 1-3, wherein the biological sample comprises peripheral blood mononuclear cells (PBMCs).

8. The method of any one of claims 1-7, wherein the gene signature comprises the gene group CSF1R, CSF3R, CCL18, and IL1B.

9. The method of any one of claims 1-7, wherein the gene signature comprises the gene group CSF1R, CSF3R, CD36, and IL1B.

10. The method of any one of claims 1-7, wherein the gene signature comprises the ' gene group CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1.

11. The method of any one of claims 1-7, wherein the gene signature comprises TNFAIP6.

12. The method of any one of claims 1-11, wherein the expression level is determined using an microarray analysis and/or polymerase chain reaction (PCR) technology.

13. The method of claim 12, wherein the PCR technology is quantitative real time PCR (qRT-PCR).

14. The method of any one of claims 1-13, wherein the expression level is determined using an immunoassay.

15. The method of claim 14, wherein the immunoassay is an ELISA.

16. The method of any one of claims 1-15, wherein the TNFα inhibitor is an anti-TNFα antibody, or antigen-binding fragment thereof.

17. The method of claim 16, wherein the anti-TNFα antibody, or antigen-binding portion thereof, is selected from the group consisting of a human antibody, a chimeric antibody, and a humanized antibody.

18. The method of claim 16, wherein the anti-TNFα antibody is infliximab, or a biosimilar thereof.

19. The method of claim 16, wherein the anti-TNFα antibody, or antigen-binding portion thereof, is an isolated human antibody that dissociates from human TNFα with a K_d of 1 x 10^{-8} M or less and a k_{off} rate constant of 1 x 10^{-3} s^{-1} or less, both determined by surface plasmon

resonance, and inhibits human TNFα-induced expression of ELAM-1 in human umbilical vein endothelial cells.

20. The method of claim 16, wherein the anti-TNFα antibody, or antigen-binding portion thereof, comprises

a light chain variable region comprising a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 27, a CDR2 domain comprising the amino acid sequence of SEQ ID NO: 29, and a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 31, and

a heavy chain variable region comprising a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 28, a CDR2 domain comprising the amino acid sequence of SEQ ID NO: 30, and a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 32.

21. The method of claim 16, wherein the anti-TNFα antibody, or antigen-binding portion thereof, comprises a light chain variable region (LCVR) comprising the amino acid sequence of SEQ ID NO: 25 and a heavy chain variable region (HCVR) comprising the amino acid sequence of SEQ ID NO: 26.

22. The method of claim 20 or 21, wherein the light chain is a kappa light chain.

23. The method of any one of claims 19-22, wherein the anti-TNFα antibody, or antigen-binding portion thereof, is an IgG isotype.

24. The method of claim 23, wherein the IgG isotype is an IgG1 or an IgG4.

25. The method of claim 16, wherein the anti-TNFα antibody is adalimumab, or a biosimilar thereof.

26. The method of claim 16, wherein the anti-TNFα antibody is golimumab, or a biosimilar thereof.

27. The method of claim 16, wherein the anti-TNFα antibody fragment is certolizumab pegol, or a biosimilar thereof.

28. A method for treating a subject having an IBD, said method comprising

selecting a subject having an IBD who is identified as having a responder gene signature, wherein the responder gene signature comprises a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6; and

administering a TNFα inhibitor to the subject having IBD.

29. The method of claim 28, wherein the TNFα inhibitor is an anti-TNFα antibody, or antigen-binding fragment thereof.

30. The method of claim 29, wherein the anti-TNFα antibody, or antigen-binding portion thereof, is selected from the group consisting of a human antibody, a chimeric antibody, and a humanized antibody.

31. The method of claim 29, wherein the anti-TNFα antibody is infliximab or golimumab, or a biosimilar thereof; or wherein the anti-TNFα antibody fragment is certolizumab pegol, or a biosimilar thereof.

32. The method of claim 29, wherein the anti-TNFα antibody, or antigen-binding portion thereof, is an isolated human antibody that dissociates from human TNFα with a K_d of 1 x 10^{-8} M or less and a k_{off} rate constant of 1 x 10^{-3} s^{-1} or less, both determined by surface plasmon resonance, and inhibits human TNFα-induced expression of ELAM-1 in human umbilical vein endothelial cells.

33. The method of claim 29, wherein the anti-TNFα antibody, or antigen-binding portion thereof, comprises

a light chain variable region comprising a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 27, a CDR2 domain comprising the amino acid sequence of

SEQ ID NO: 29, and a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 31, and

a heavy chain variable region comprising a CDR3 domain comprising the amino acid sequence of SEQ ID NO: 28, a CDR2 domain comprising the amino acid sequence of SEQ ID NO: 30, and a CDR1 domain comprising the amino acid sequence of SEQ ID NO: 32.

34. The method of claim 29, wherein the anti-TNFα antibody, or antigen-binding portion thereof, comprises a light chain variable region (LCVR) comprising the amino acid sequence of SEQ ID NO: 25 and a heavy chain variable region (HCVR) comprising the amino acid sequence of SEQ ID NO: 26.

35. The method of claim 33 or 34, wherein the light chain is a kappa light chain.

36. The method of any one of claims 32-35, wherein the anti-TNFα antibody, or antigen-binding portion thereof, is an IgG isotype.

37. The method of claim 36, wherein the IgG isotype is an IgG1 or an IgG4.

38. The method of claim 32, wherein the anti-TNFα antibody is adalimumab, or a biosimilar thereof.

39. The method of any one of claims 32-38, wherein a first 160 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject, a second 80 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject two weeks after the first dose, and a third 40 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject two weeks after the second dose.

40. The method of claim 39, wherein a 40 mg dose of the anti-TNFα antibody, or antigen-binding portion thereof, is administered to the subject every other week after the third dose.

41. A method for determining a personalized medical treatment for a subject diagnosed with an IBD comprising the steps of:

a) measuring a gene signature of a biological sample from a subject diagnosed with an IBD comprising gene expression analysis for a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6,

b) accessing a computer database to compare the results of the gene expression analysis in step a) to a responder or non-responder gene signature to determine whether the gene signature correlates to the responder or non-responder gene signature, such that a personalized medical treatment is determined.

42. The method of any one of claims 1-41, wherein the IBD is Crohn's disease.

43. The method of claim 42, wherein the Crohn's disease is moderately to severely active Crohn's disease.

44. The method of any one of claims 1-41, wherein the IBD is ulcerative colitis.

45. The method of any one of claims 1-44, wherein the subject is a human subject.

46. A kit for predicting the responsiveness of a subject having IBD to treatment with a TNFα inhibitor, the kit comprising

a means for determining the expression level(s) of a gene signature comprising a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6,

in a sample obtained from the subject having IBD; and

instructions for recommended treatment of the subject having IBD based on the expression level(s) of the gene signature.

47. A kit for identifying a responder to a TNFα inhibitor for treating IBD, the kit comprising

a means for determining the expression level(s) of a gene signature comprising a gene or gene group selected from the group consisting of

(i) CSF1R, CSF3R, CCL18, and IL1B;

(ii) CSF1R, CSF3R, CD36, and IL1B;

(iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; and

(iv) TNFAIP6,

in a patient sample obtained from the subject having IBD; and

a control sample from a responder or a non-responder which can be used to determine the expression level(s) of the gene signature determined for the patient sample.

48. The kit of claim 46 or 47, wherein the kit comprises a means for determining the presence of the gene signature (i) CSF1R, CSF3R, CCL18, and IL1B in a sample obtained from the subject having IBD.

49. The kit of claim 48, wherein the means for determining the presence of the gene signature (i) CSF1R, CSF3R, CCL18, and IL1B comprises either:

nucleic acids that hybridize to the nucleic acid molecules encoding each of genes CSF1R, CSF3R, CCL18, and IL1B; or

antibodies which specifically bind to the proteins corresponding to each of genes CSF1R, CSF3R, CCL18, and IL1B.

50. The kit of claim 46 or 47, wherein the kit comprises a means for determining the presence of the gene signature (ii) CSF1R, CSF3R, CD36, and IL1B in a sample obtained from the subject having Crohn's disease.

51. The kit of claim 50, wherein the means for determining the presence of the gene signature (ii) CSF1R, CSF3R, CD36, and IL1B comprises either:

nucleic acids that hybridize to the nucleic acid molecules encoding each of genes CSF1R, CSF3R, CD36, and IL1B; or

antibodies which specifically bind to the proteins corresponding to each of genes CSF1R, CSF3R, CD36, and IL1B.

52. The kit of claim 46 or 47, wherein the kit comprises a means for determining the presence of the gene signature (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1 in a sample obtained from the subject having Crohn's disease.

53. The kit of claim 52, wherein the means for determining the presence of the gene signature (iii) CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1 comprises either:

nucleic acids that hybridize to the nucleic acid molecules encoding each of genes CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1; or

antibodies which specifically bind to the proteins corresponding to each of genes CLEC4A, CLEC7A, CSTA, CHI3L1, IL1B, IL1RN, TNFAIP6, and TREM1.

54. The kit of claim 46 or 47, wherein the kit comprises a means for determining the presence of the gene signature (iv) TNFAIP6 in a sample obtained from the subject having IBD.

55. The kit of claim 54, wherein the means for determining the presence of the gene signature (iv) TNFAIP6 comprises either:

a nucleic acid that hybridizes to the nucleic acid molecule encoding the gene TNFAIP6; or

an antibody which specifically binds to the protein corresponding to the gene TNFAIP6.

56. The kit of any one of claims 46-55, wherein the subject has Crohn's disease or ulcerative colitis.

57. The kit of claim 56, wherein the subject has moderate to severe Crohn's disease.

58. The kit of any one of claims 46-57, wherein the subject is a human.

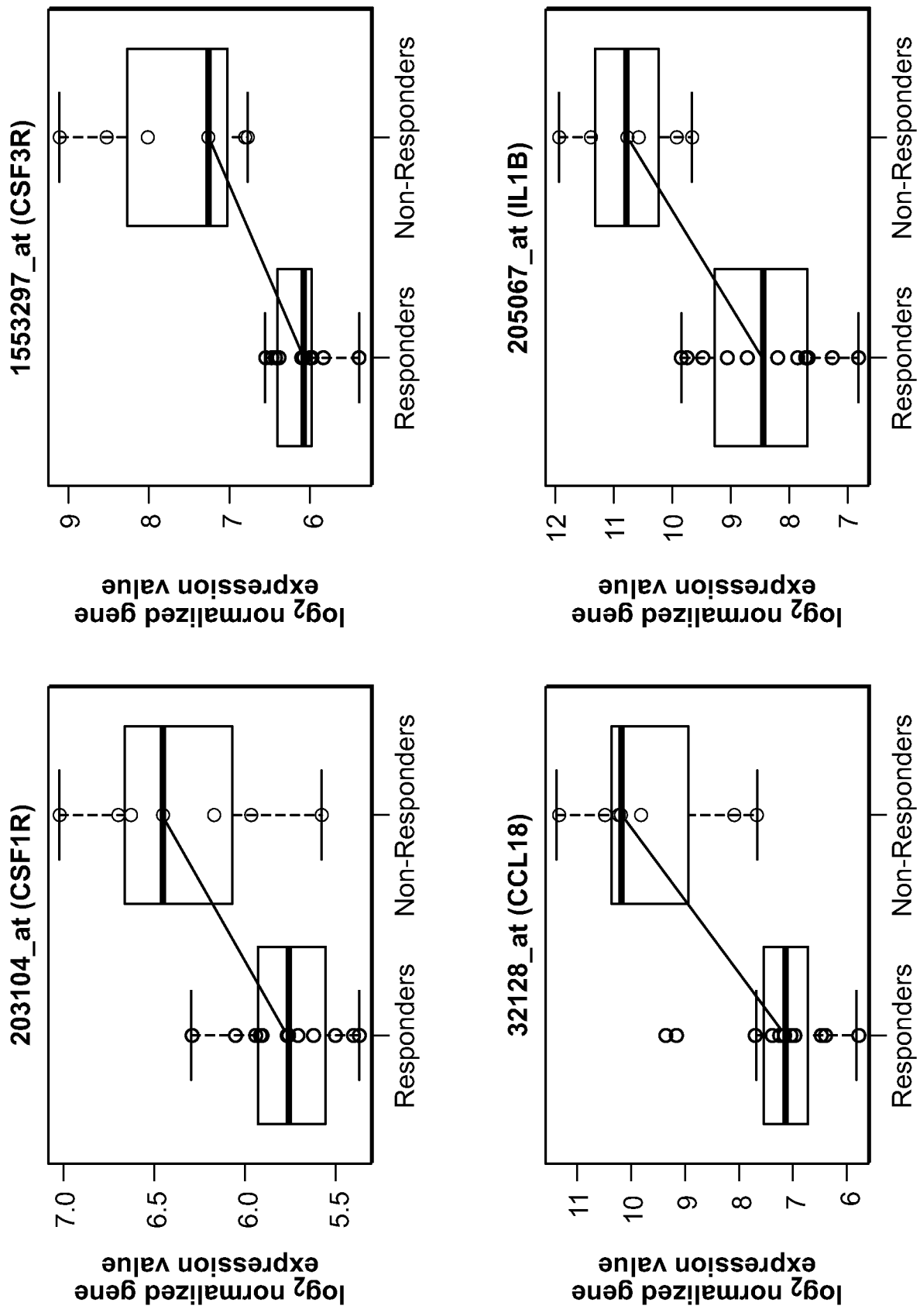
1553297_at (CSF3R)
log2 normalized gene expression value
Responders
Non-Responders
205067_at (IL1B)
log2 normalized gene expression value
Responders
Non-Responders
203104_at (CSF1R)
log2 normalized gene expression value
Responders
Non-Responders
32128_at (CCL18)
log2 normalized gene expression value
Responders
Non-Responders

FIG. 1

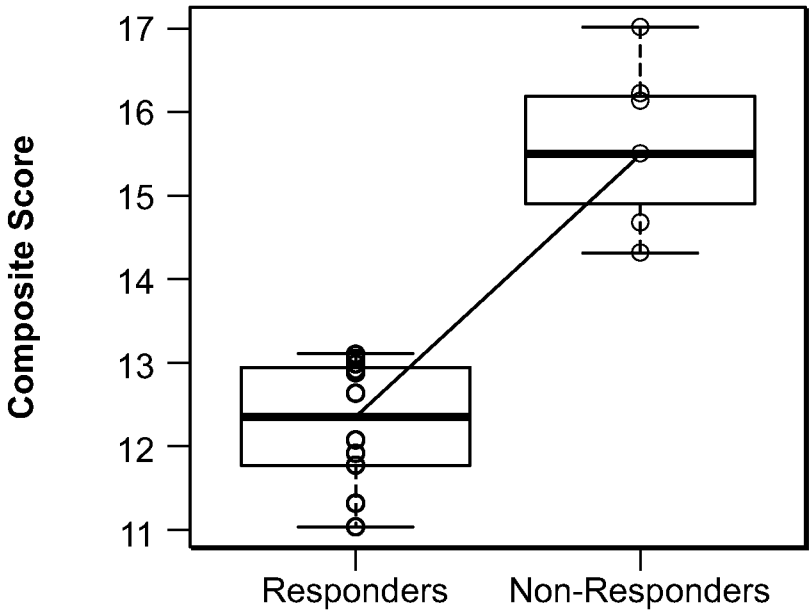
Composite Score
17
16
15
14
13
12
11
Responders
Non-Responders

FIG. 2

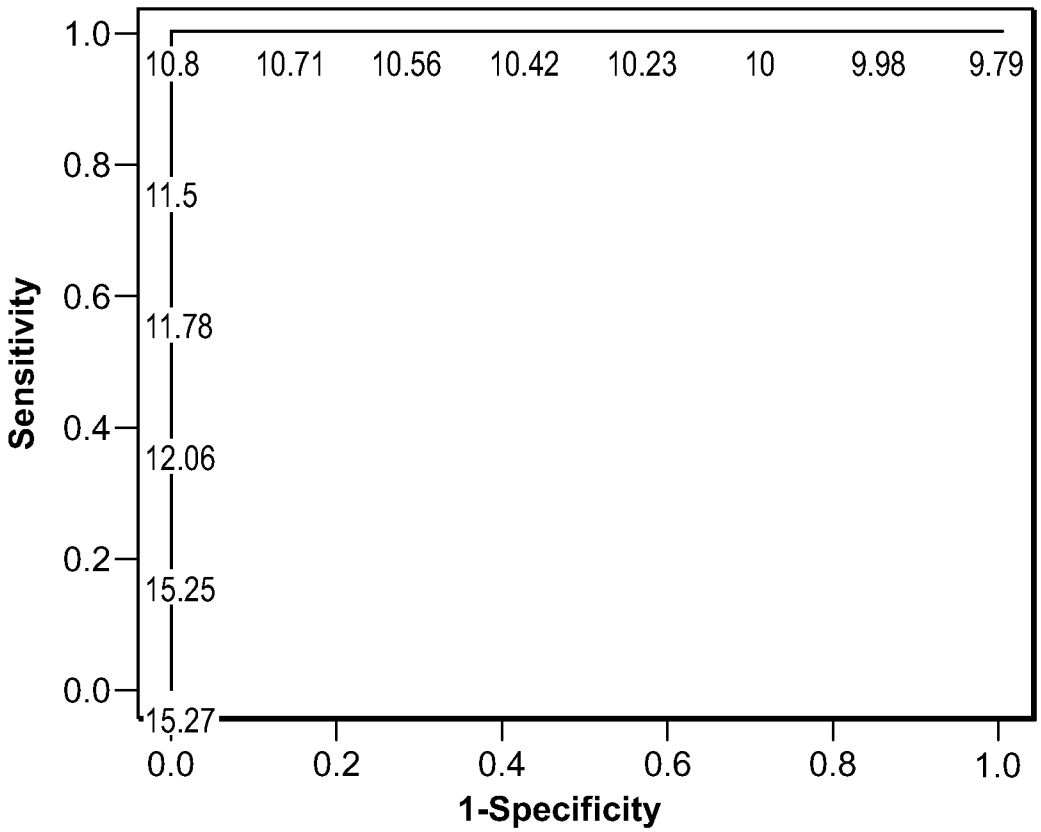
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
10.8
10.71
10.56
10.42
10.23
10
9.98
9.79
11.5
11.78
12.06
15.25
15.27
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity

FIG. 3

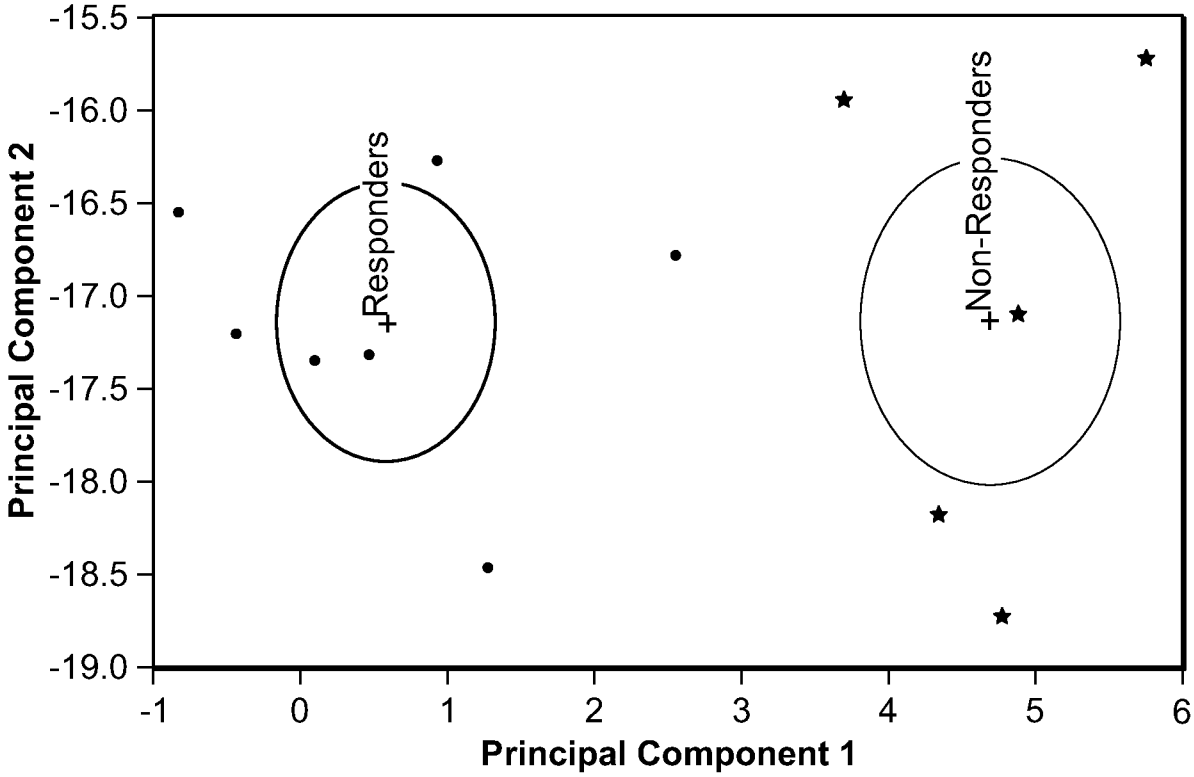
Principal Component 2
-15.5
-16.0
-16.5
-17.0
-17.5
-18.0
-18.5
-19.0
-1
0
1
2
3
4
5
6
Principal Component 1
Responders
Non-Responders

FIG. 4

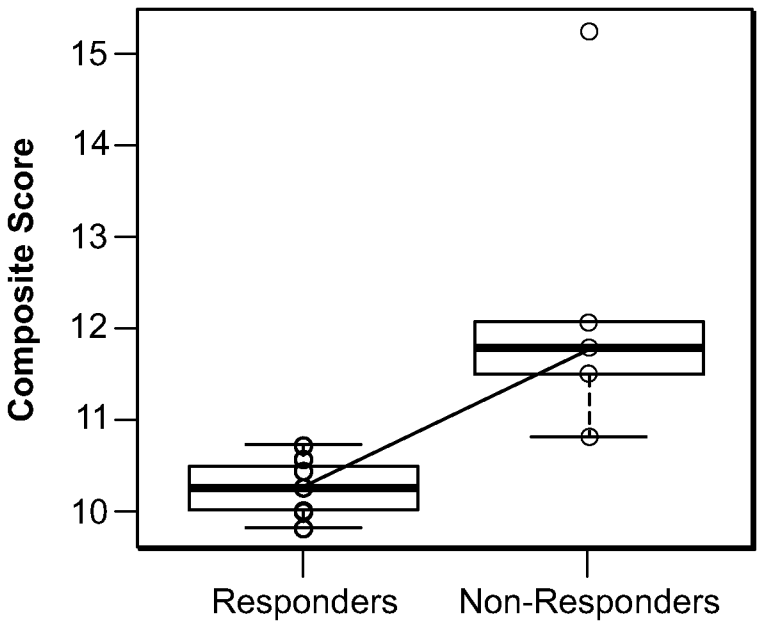
Composite Score
15
14
13
12
11
10
Responders
Non-Responders

FIG. 5

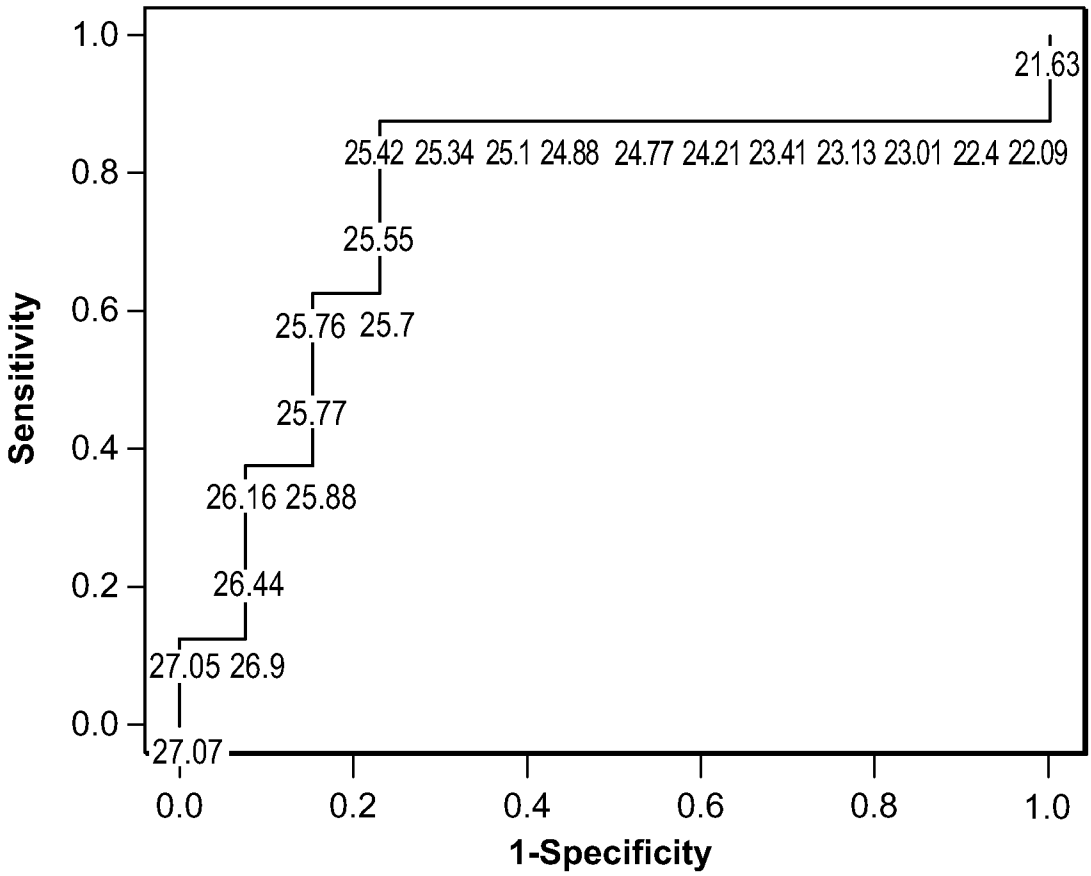
1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
21.63
25.42 25.34 25.1 24.88 24.77 24.21 23.41 23.13 23.01 22.4 22.09
25.55
25.76 25.7
25.77
26.16 25.88
26.44
27.05 26.9
27.07
0.0 0.2 0.4 0.6 0.8 1.0
1-Specificity

FIG. 6

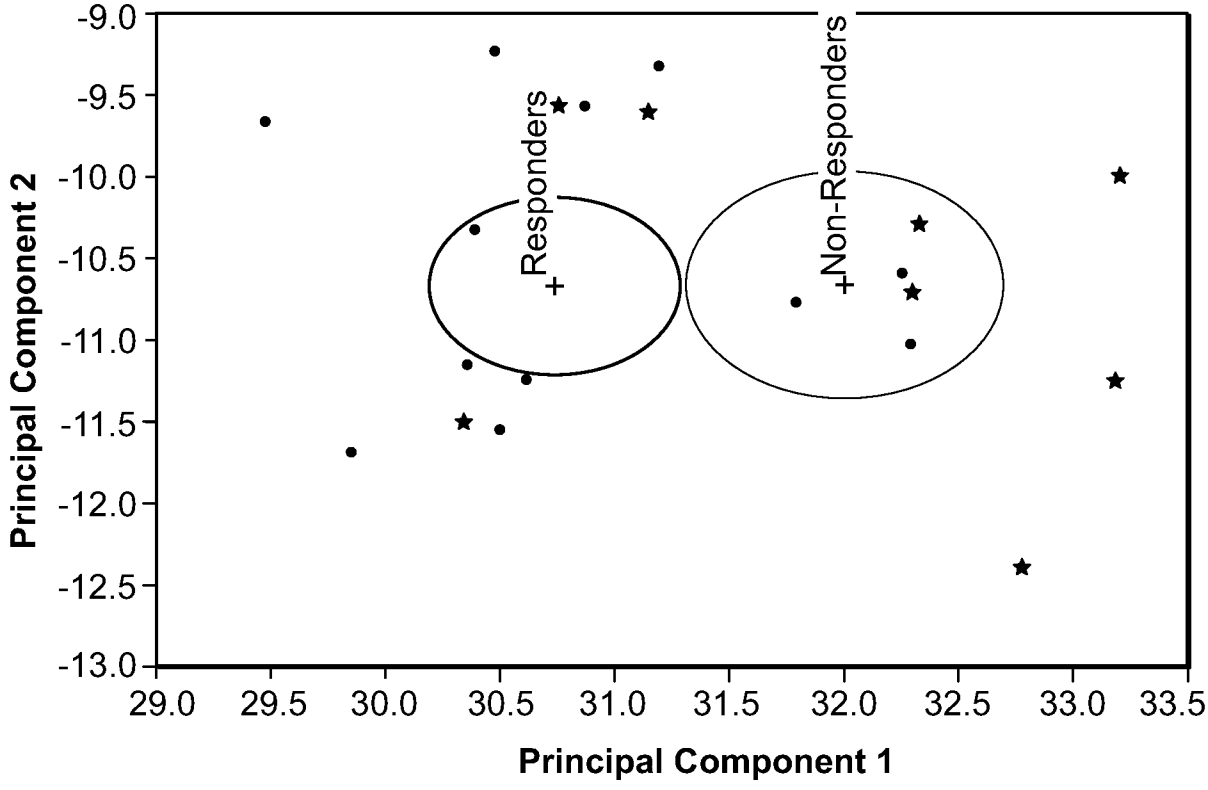
-9.0
-9.5
-10.0
-10.5
-11.0
-11.5
-12.0
-12.5
-13.0
Principal Component 2
Responders
Non-Responders
29.0 29.5 30.0 30.5 31.0 31.5 32.0 32.5 33.0 33.5
Principal Component 1

FIG. 7

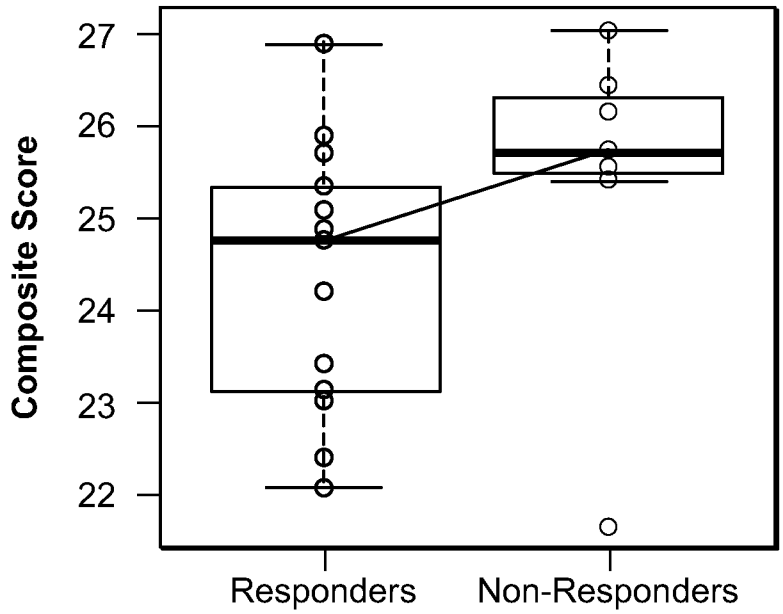
Composite Score
27
26
25
24
23
22
Responders
Non-Responders

FIG. 8

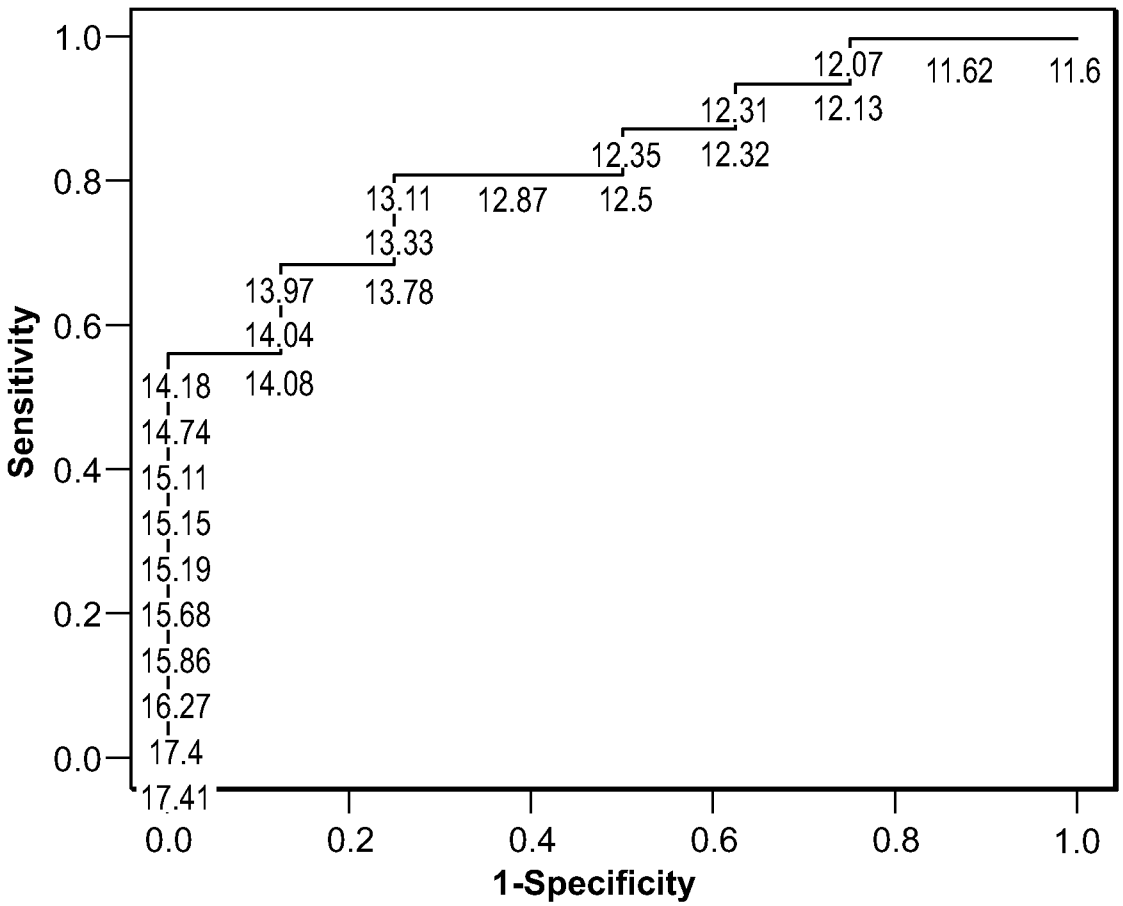
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity
12.07
11.62
11.6
12.31
12.13
12.35
12.32
13.11
12.87
12.5
13.33
13.97
13.78
14.04
14.18
14.08
14.74
15.11
15.15
15.19
15.68
15.86
16.27
17.4
17.41

FIG. 9

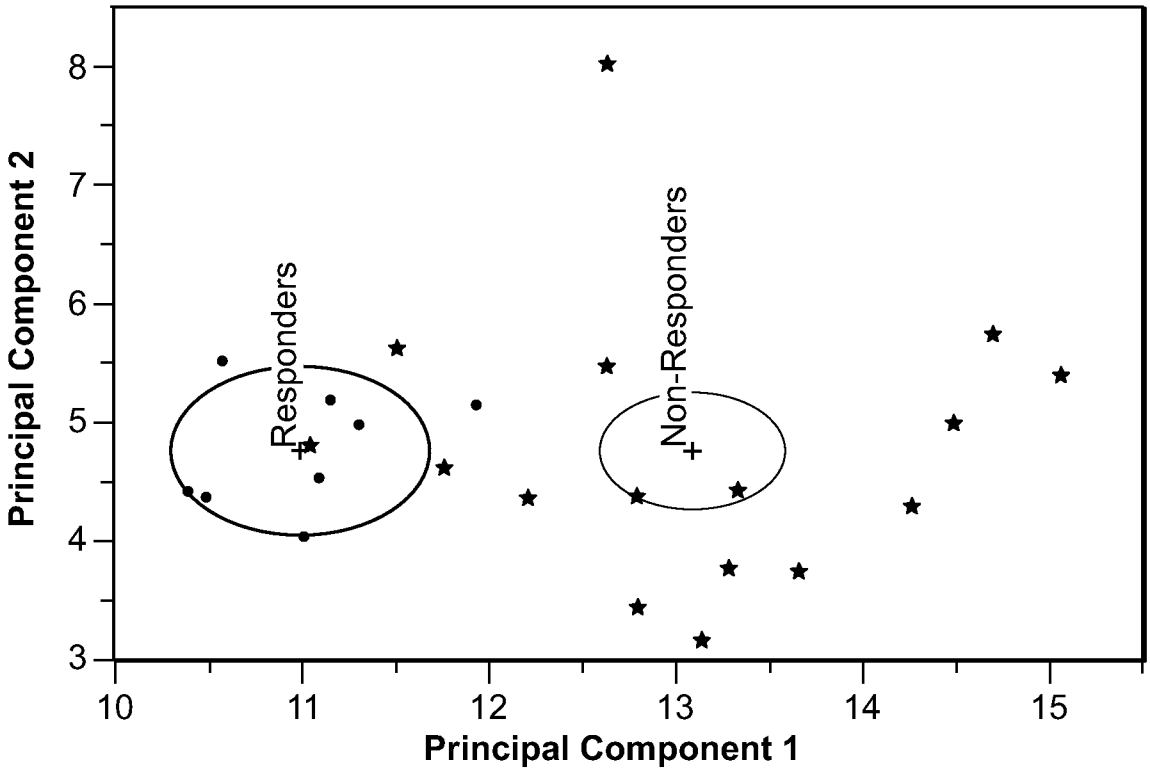
Principal Component 2
Principal Component 1
Responders
Non-Responders
8
7
6
5
4
3
10
11
12
13
14
15

FIG. 10

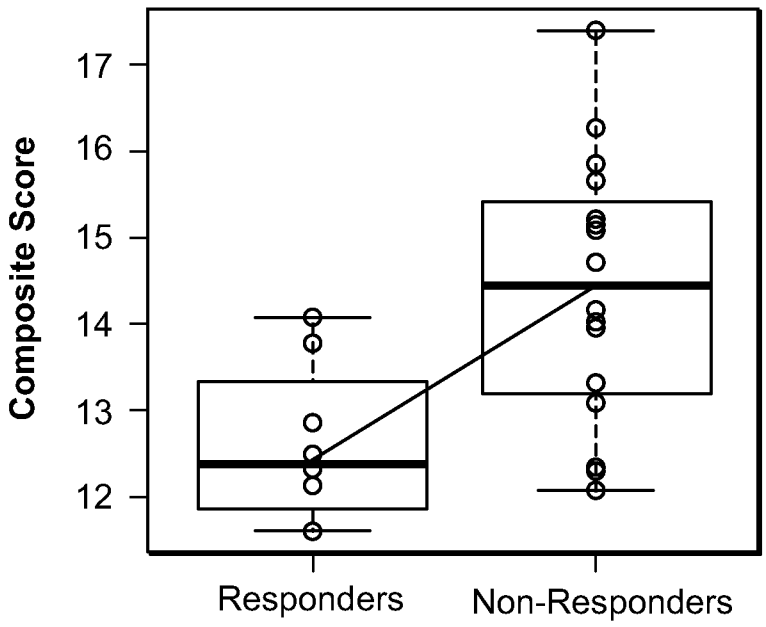
Composite Score
17
16
15
14
13
12
Responders
Non-Responders

FIG. 11

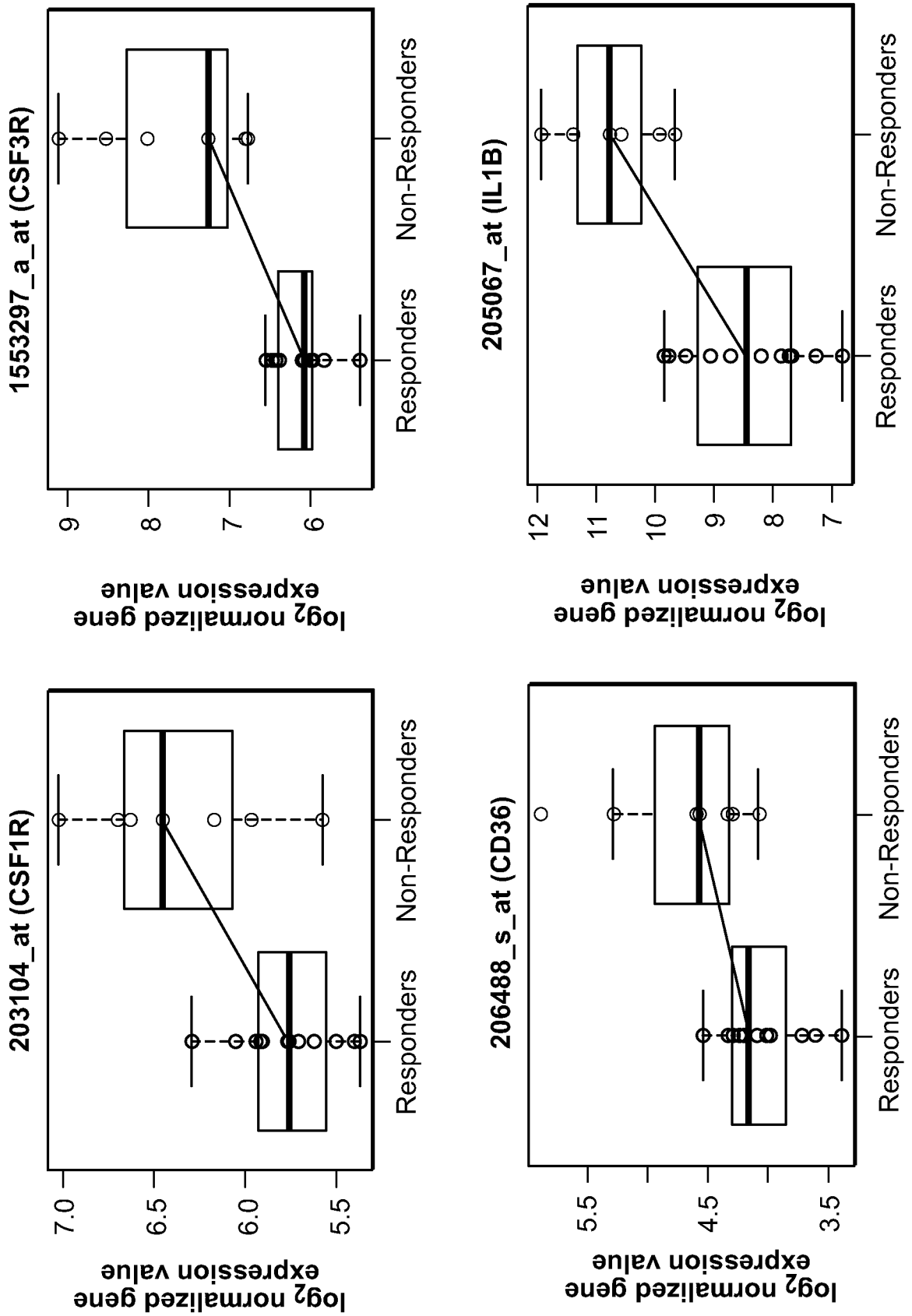
1553297_a_at (CSF3R)
log2 normalized gene expression value
9
8
7
6
Responders
Non-Responders
205067_at (IL1B)
log2 normalized gene expression value
12
11
10
9
8
7
Responders
Non-Responders
203104_at (CSF1R)
log2 normalized gene expression value
7.0
6.5
6.0
5.5
Responders
Non-Responders
206488_s_at (CD36)
log2 normalized gene expression value
5.5
4.5
3.5
Responders
Non-Responders

FIG. 12

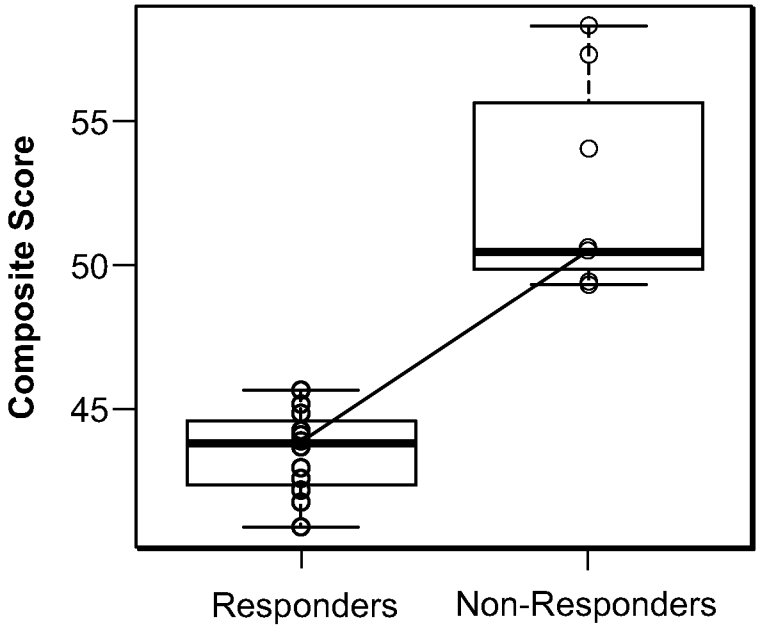
Composite Score
55
50
45
Responders
Non-Responders

FIG. 13

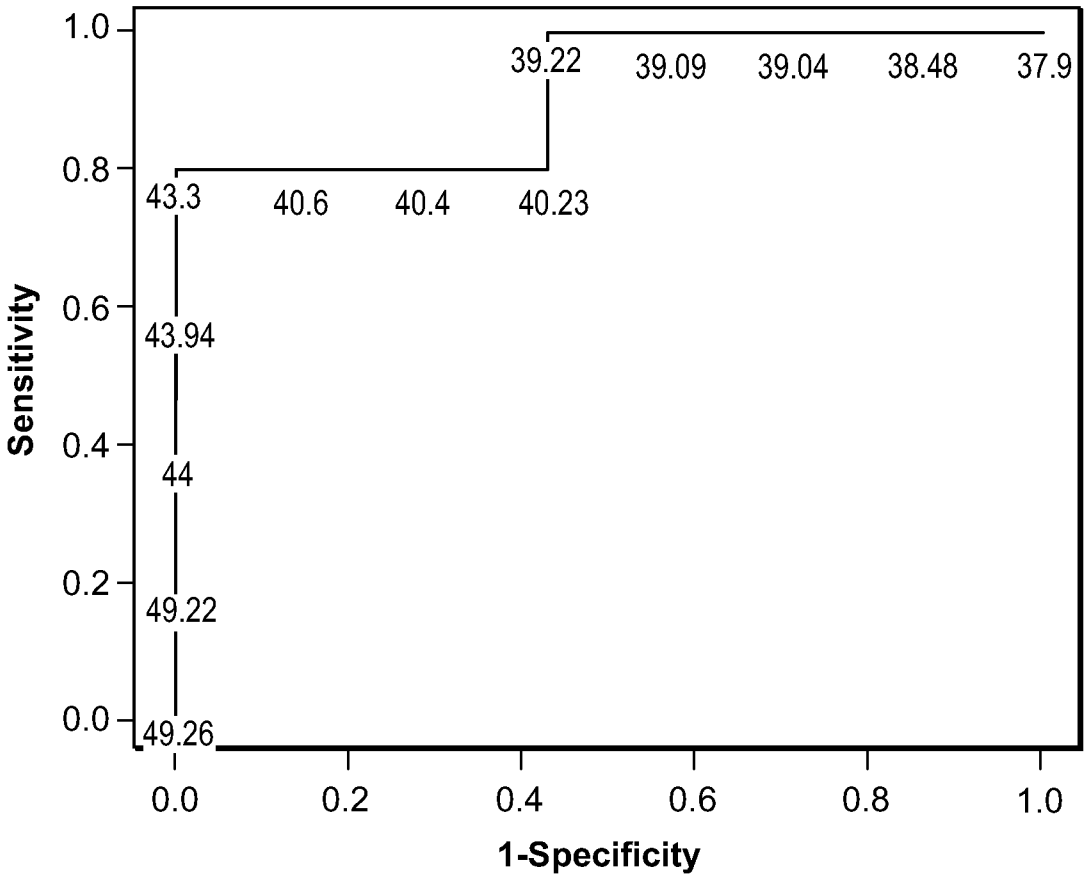
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
39.22
39.09
39.04
38.48
37.9
43.3
40.6
40.4
40.23
43.94
44
49.22
49.26
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity

FIG. 14

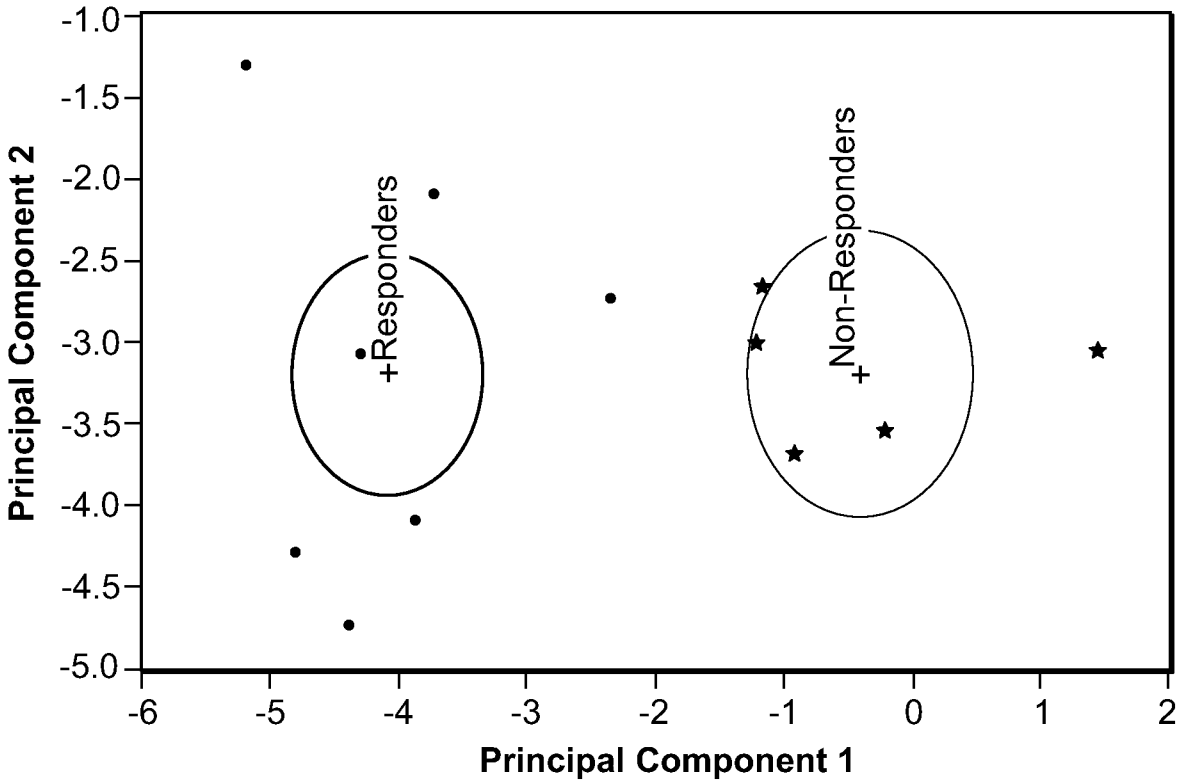
Principal Component 2
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-4.5
-5.0
-6
-5
-4
-3
-2
-1
0
1
2
Principal Component 1
+Responders
+Non-Responders

FIG. 15

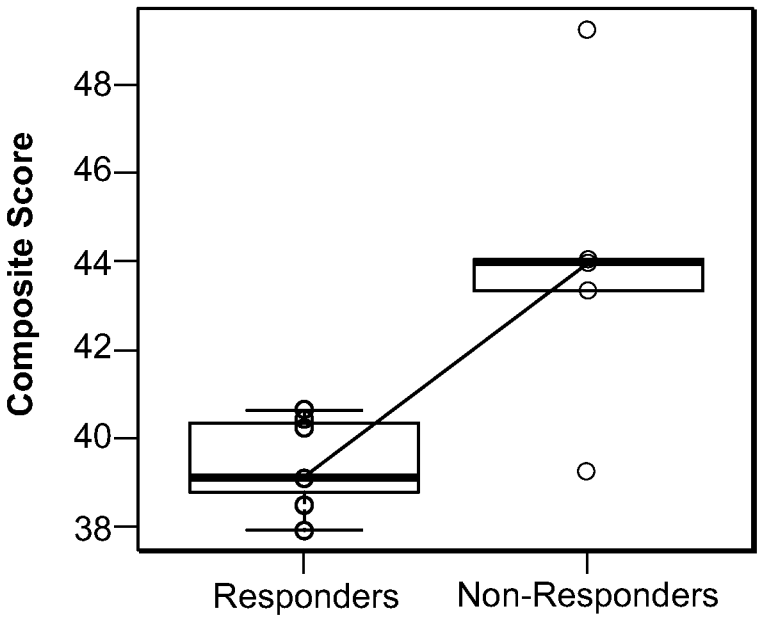
Composite Score
48
46
44
42
40
38
Responders
Non-Responders

FIG. 16

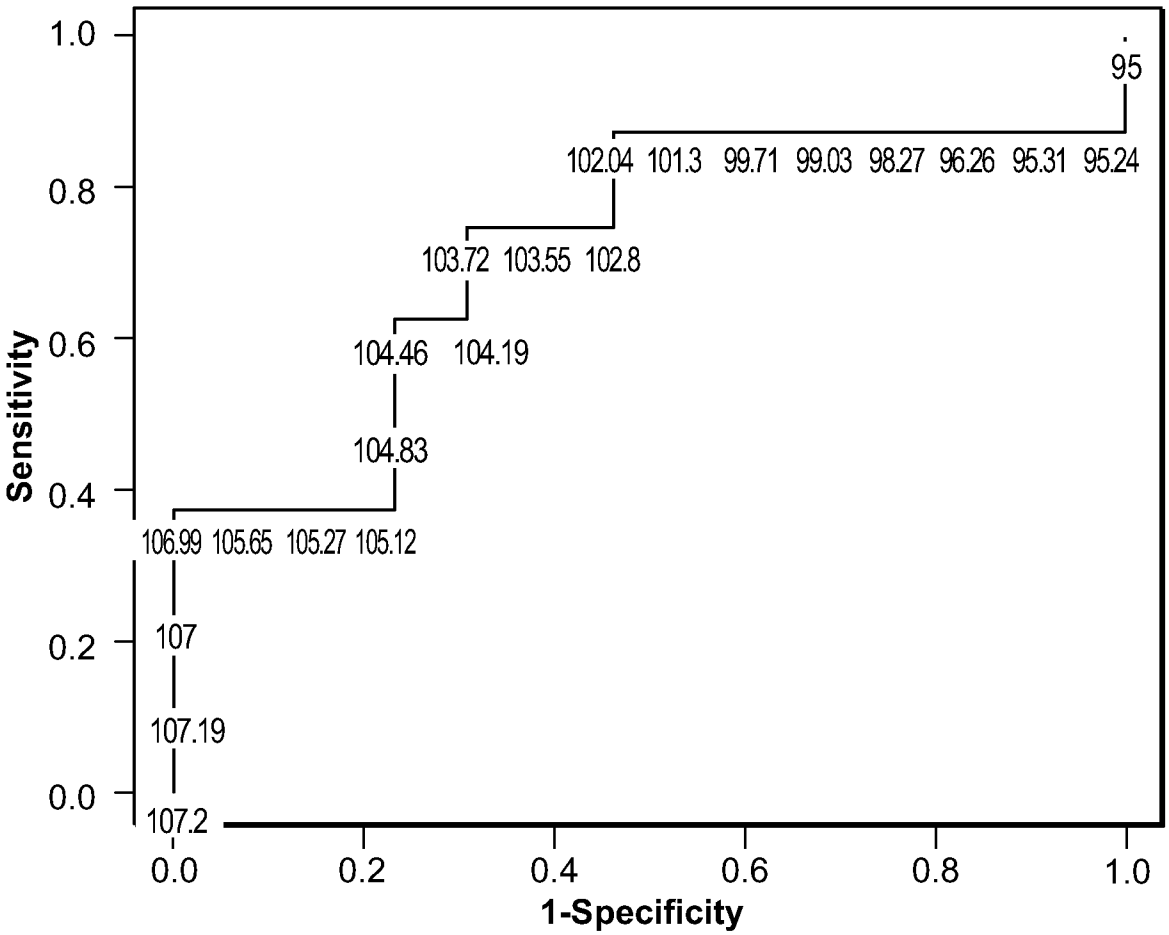
Sensitivity
1-Specificity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
95
102.04 101.3 99.71 99.03 98.27 96.26 95.31 95.24
103.72 103.55 102.8
104.46 104.19
104.83
106.99 105.65 105.27 105.12
107
107.19
107.2

FIG. 17

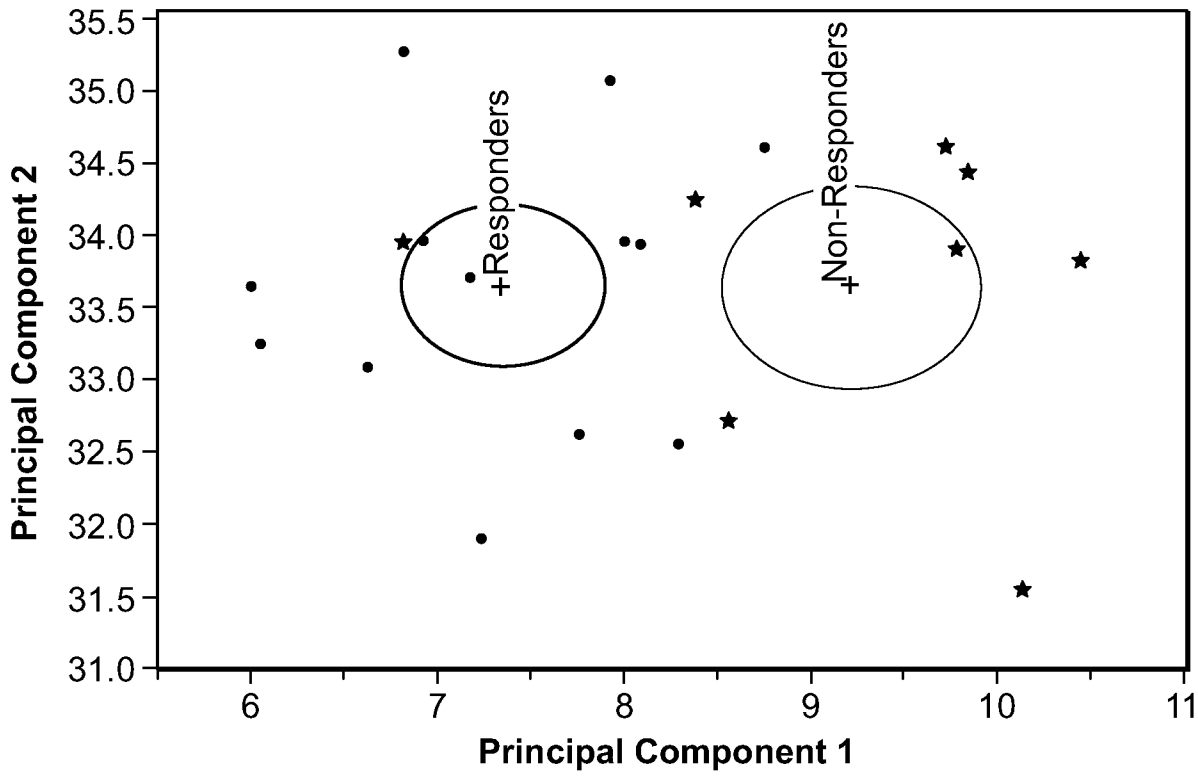
Principal Component 2
Principal Component 1
35.5
35.0
34.5
34.0
33.5
33.0
32.5
32.0
31.5
31.0
6
7
8
9
10
11
+Responders
+Non-Responders

FIG. 18

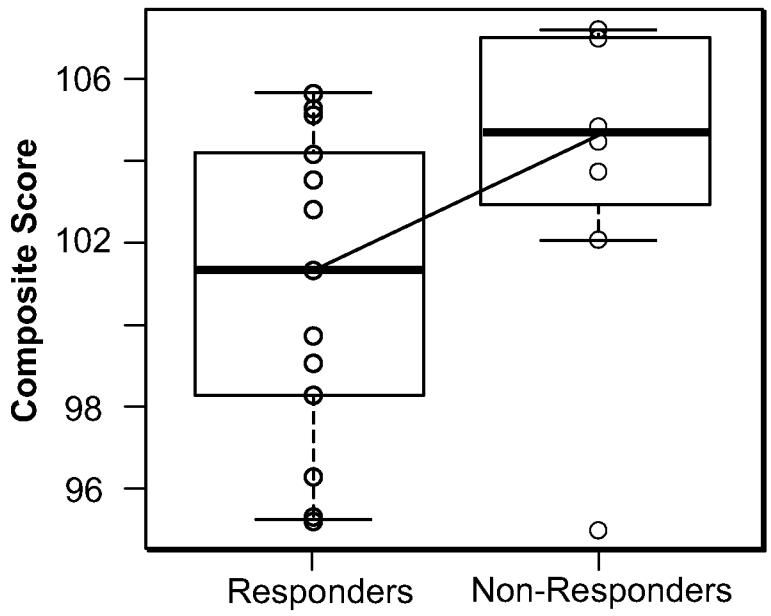
Composite Score
106
102
98
96
Responders
Non-Responders

FIG. 19

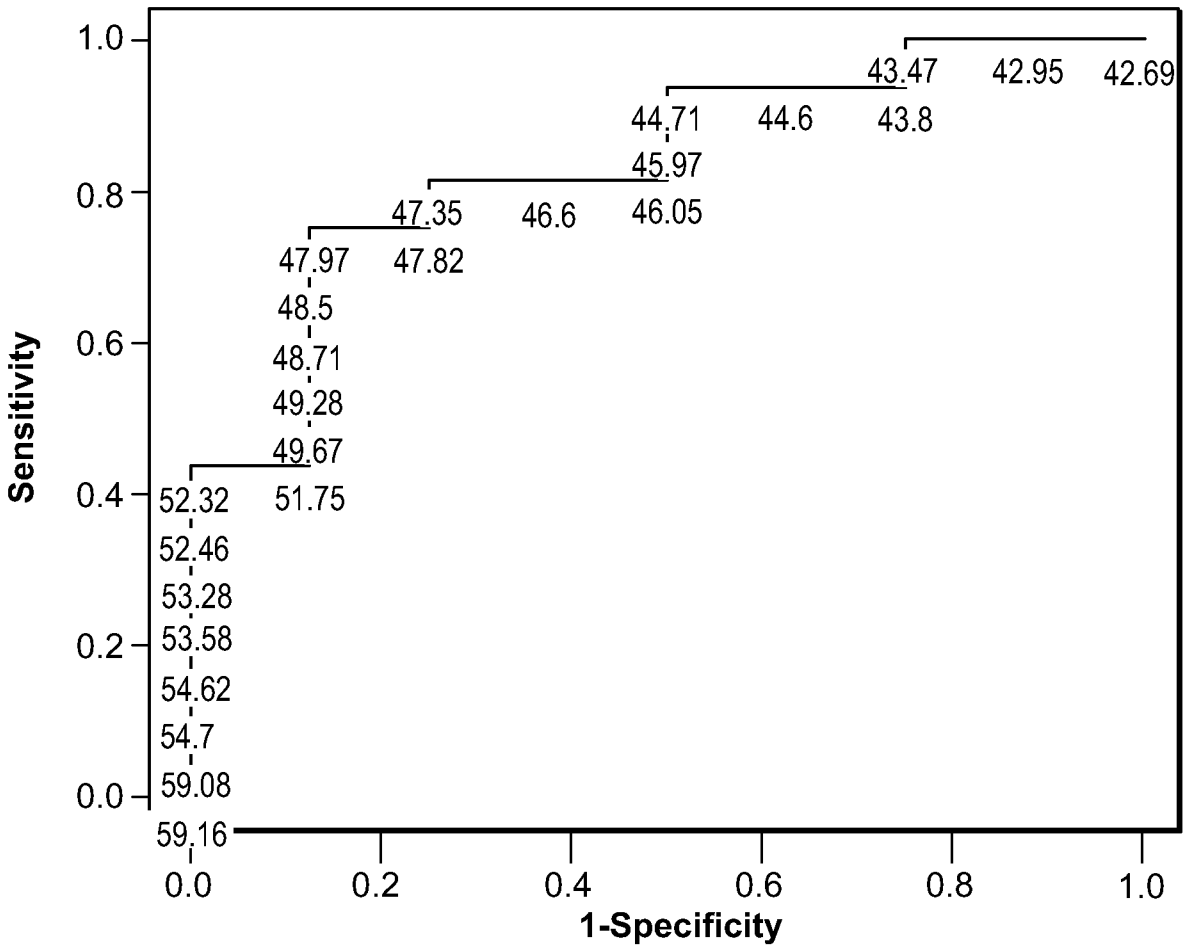
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
43.47
42.95
42.69
44.71
44.6
43.8
45.97
47.35
46.6
46.05
47.97
47.82
48.5
48.71
49.28
49.67
52.32
51.75
52.46
53.28
53.58
54.62
54.7
59.08
59.16

FIG. 20

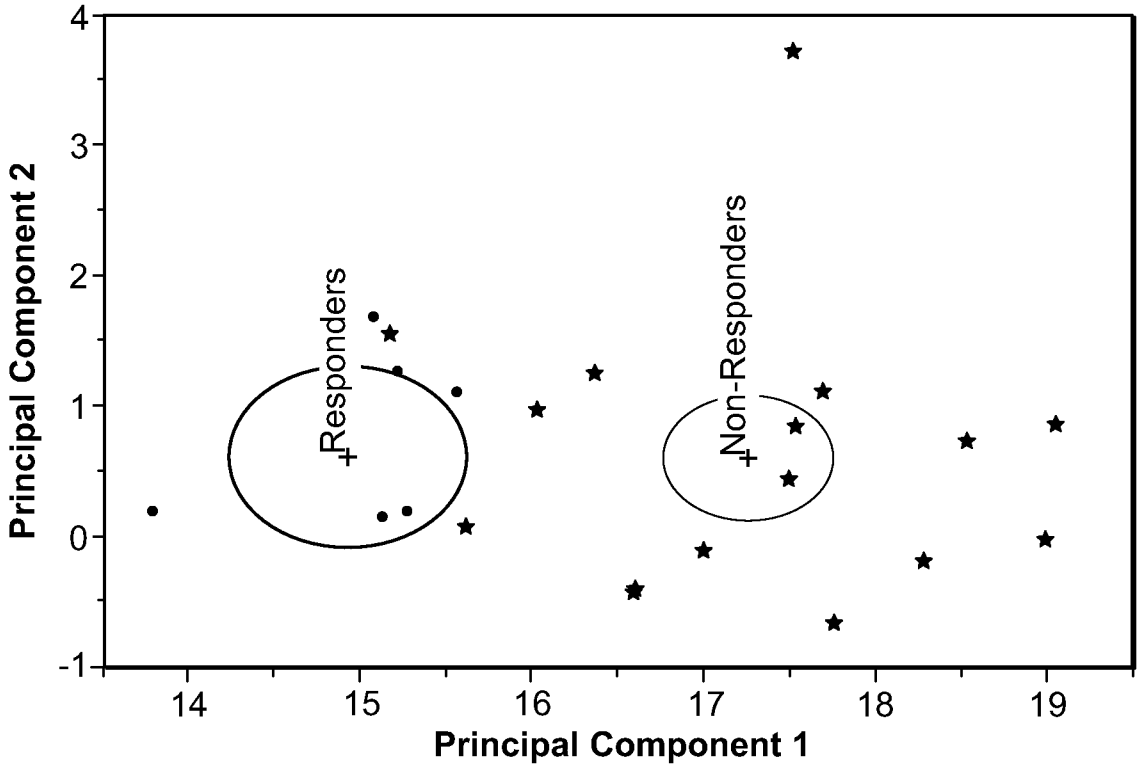
Principal Component 2
Principal Component 1
Responders
Non-Responders
4
3
2
1
0
-1
14
15
16
17
18
19

FIG. 21

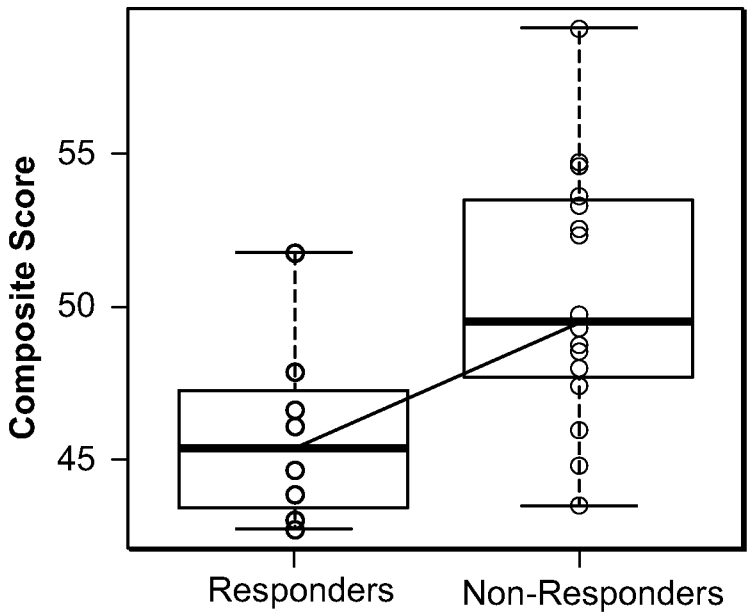
Composite Score
55
50
45
Responders
Non-Responders

FIG. 22

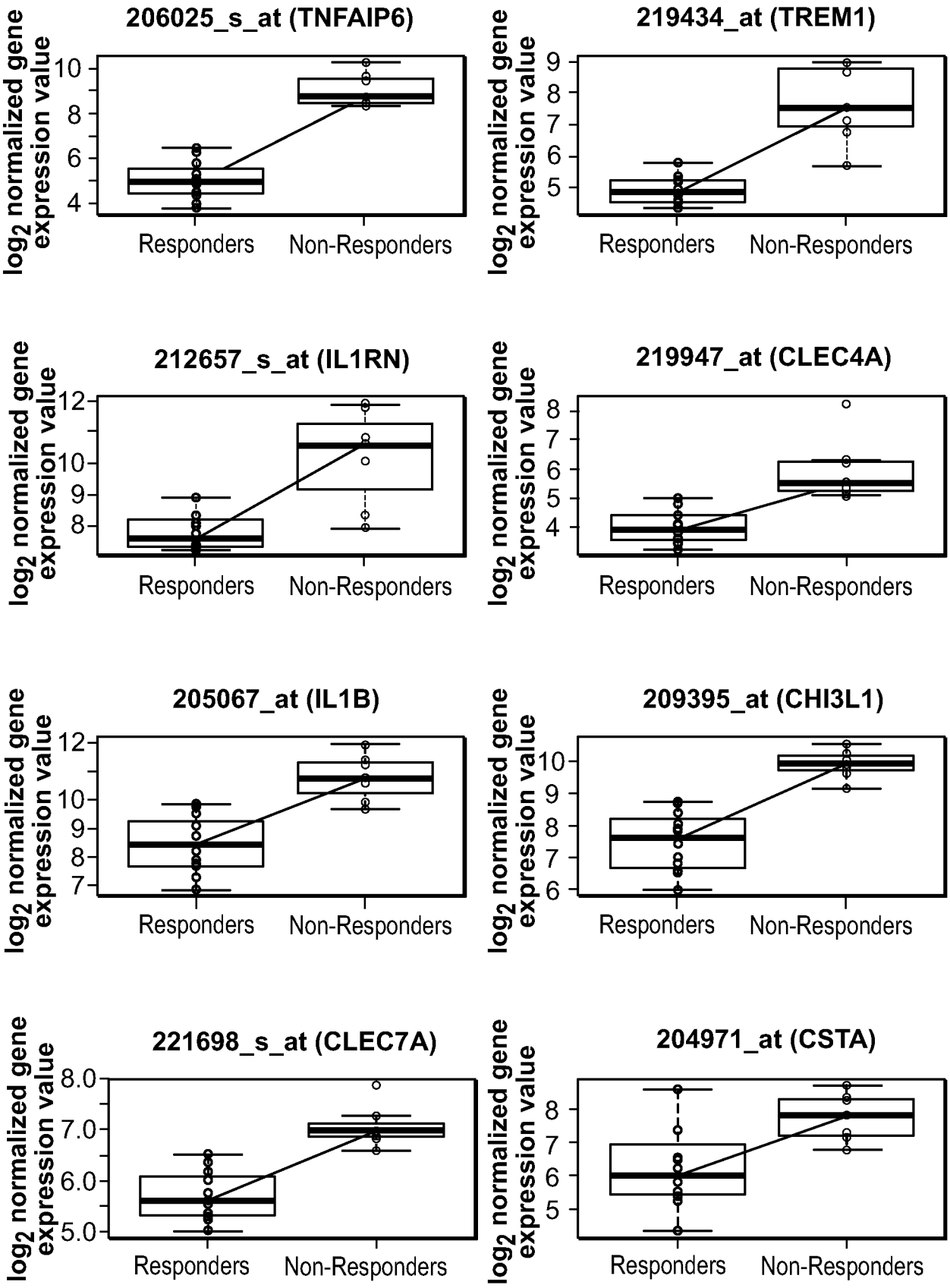
206025_s_at (TNFAIP6)
219434_at (TREM1)
212657_s_at (IL1RN)
219947_at (CLEC4A)
205067_at (IL1B)
209395_at (CHI3L1)
221698_s_at (CLEC7A)
204971_at (CSTA)
log2 normalized gene expression value
Responders
Non-Responders

FIG. 23

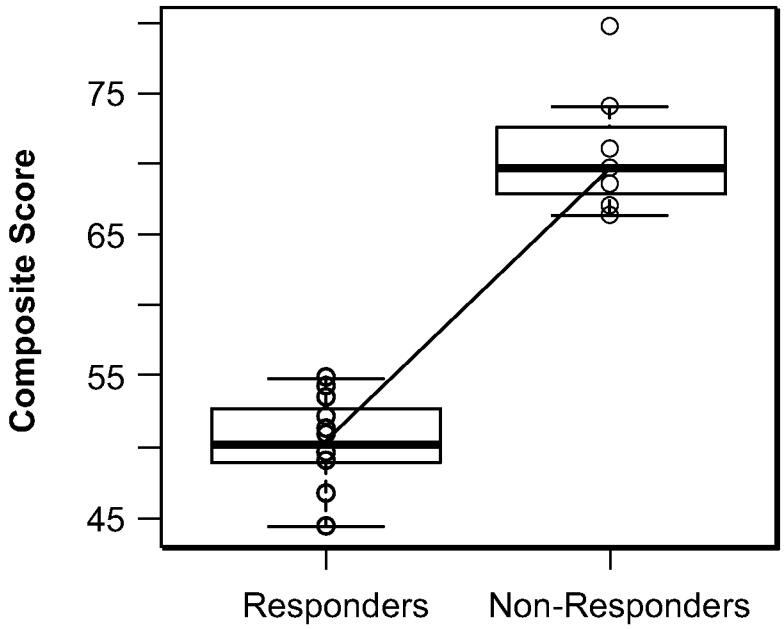
Composite Score
75
65
55
45
Responders
Non-Responders

FIG. 24

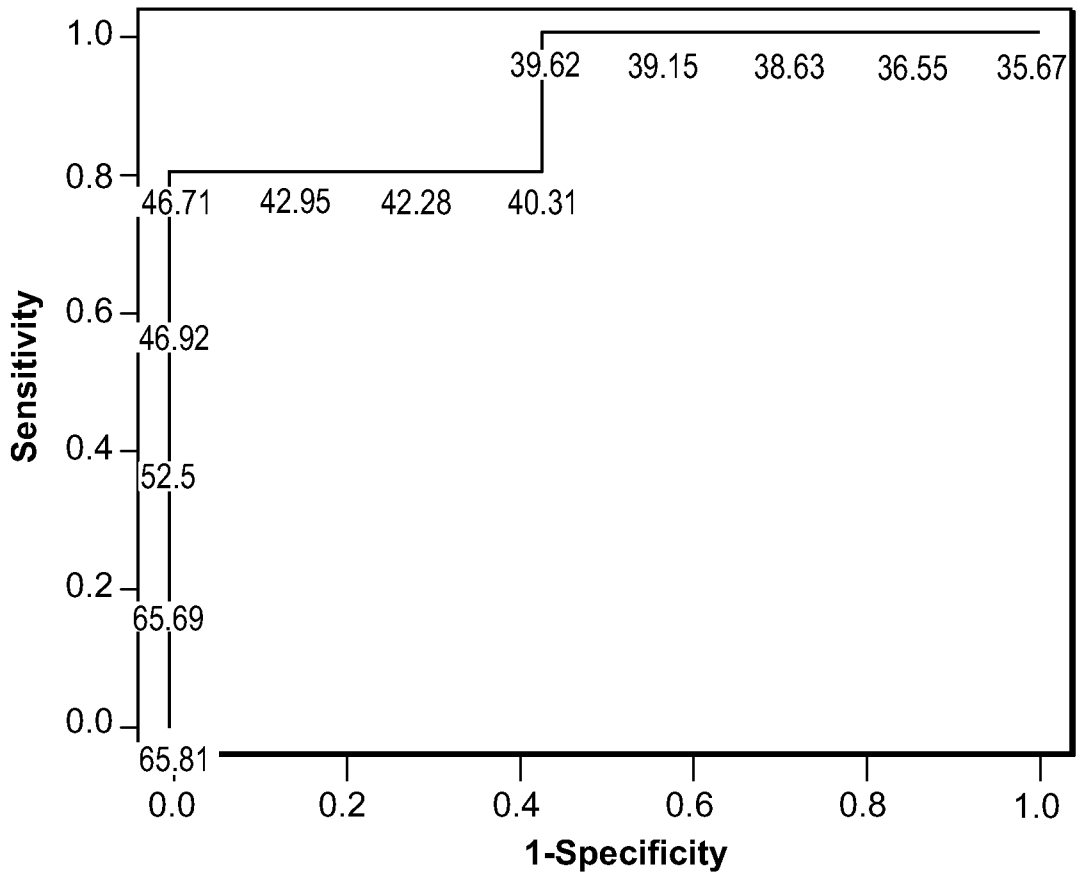
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
39.62
39.15
38.63
36.55
35.67
46.71
42.95
42.28
40.31
46.92
52.5
65.69
65.81
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity

FIG. 25

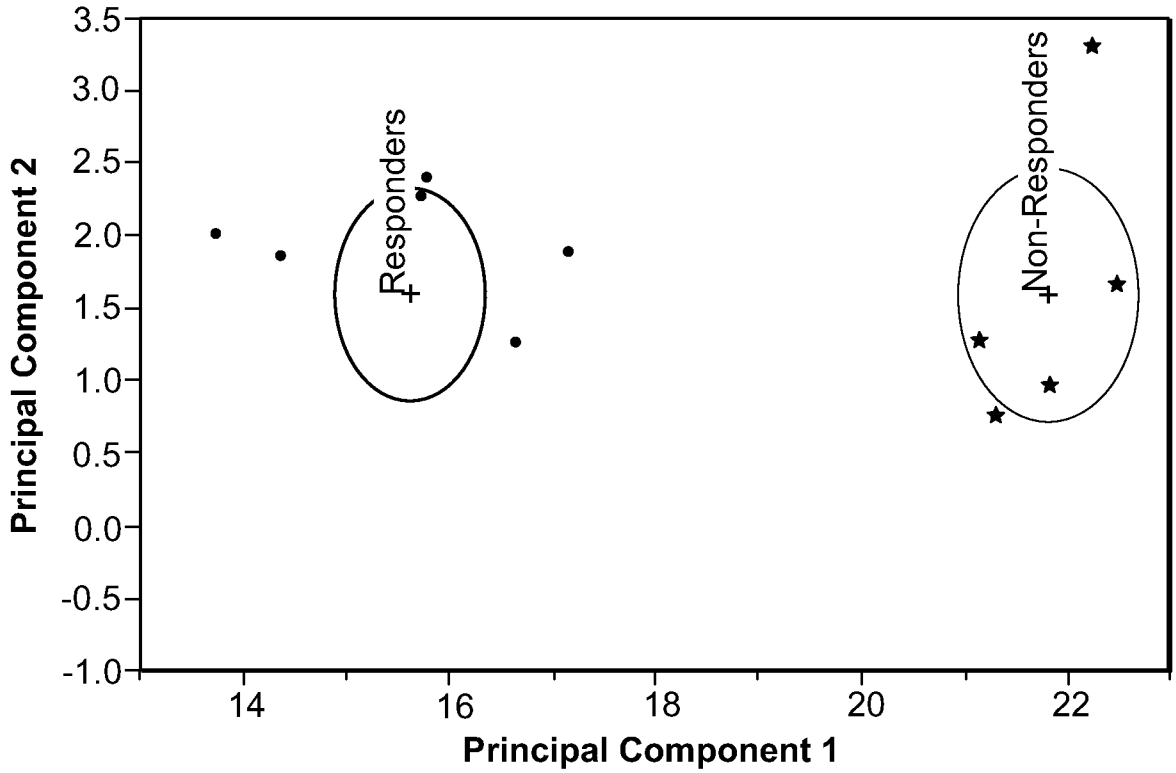
Principal Component 2
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
14
16
18
20
22
Principal Component 1
Responders
Non-Responders

FIG. 26

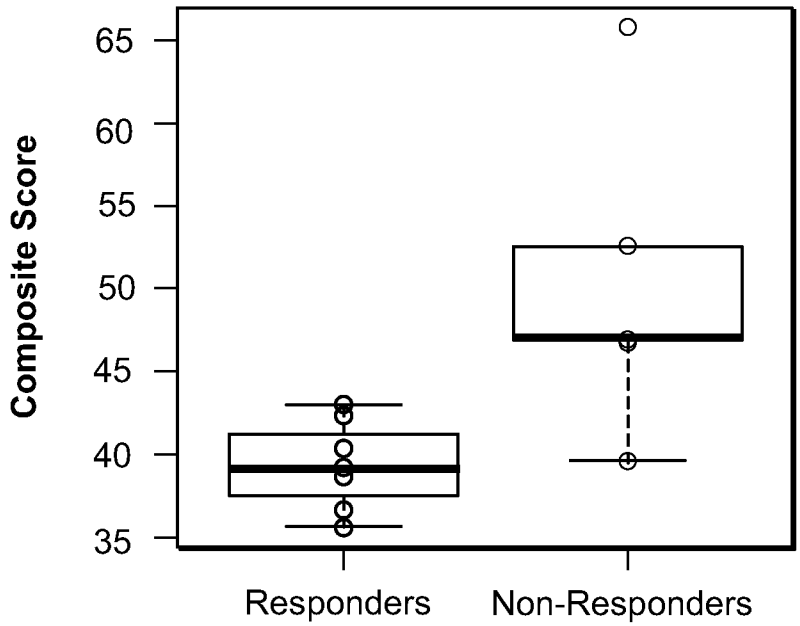
Composite Score
65
60
55
50
45
40
35
Responders
Non-Responders

FIG. 27

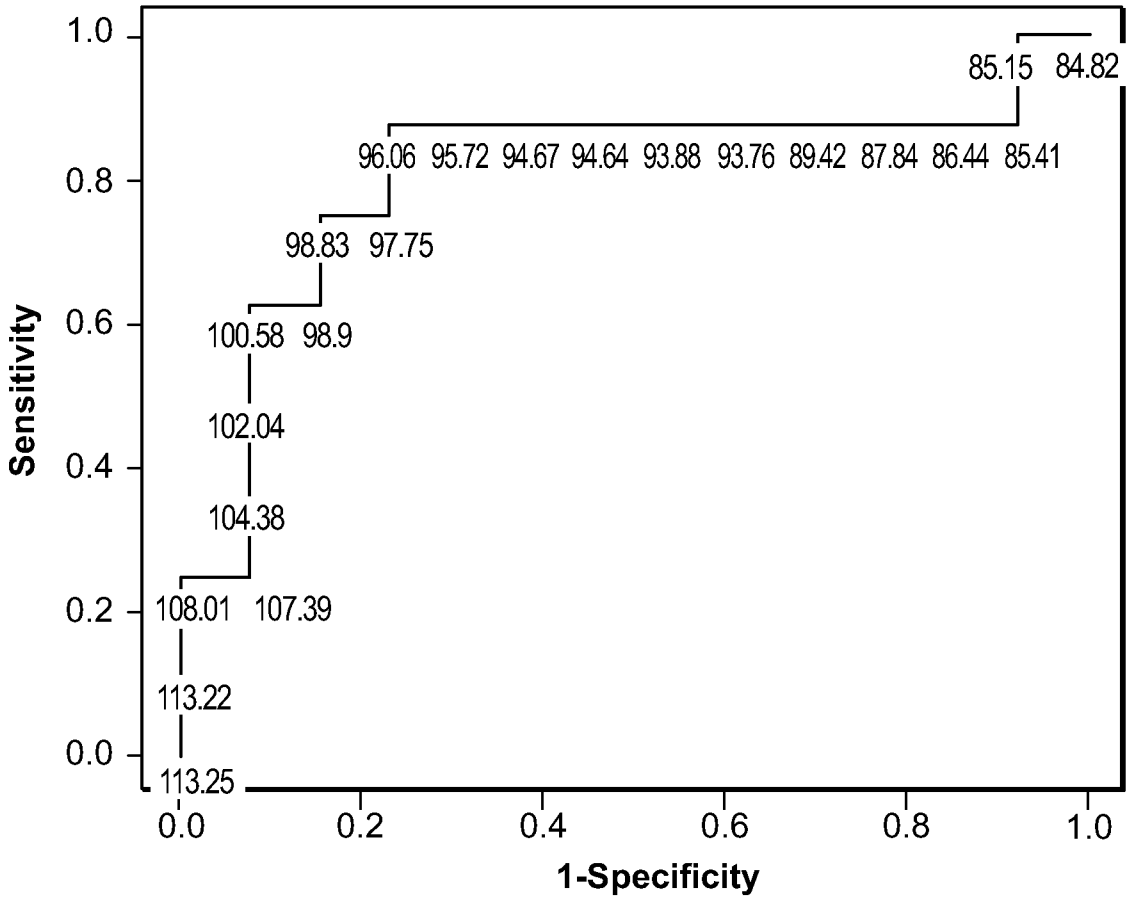
1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
85.15 84.82
96.06 95.72 94.67 94.64 93.88 93.76 89.42 87.84 86.44 85.41
98.83 97.75
100.58 98.9
102.04
104.38
108.01 107.39
113.22
113.25
0.0 0.2 0.4 0.6 0.8 1.0
1-Specificity

FIG. 28

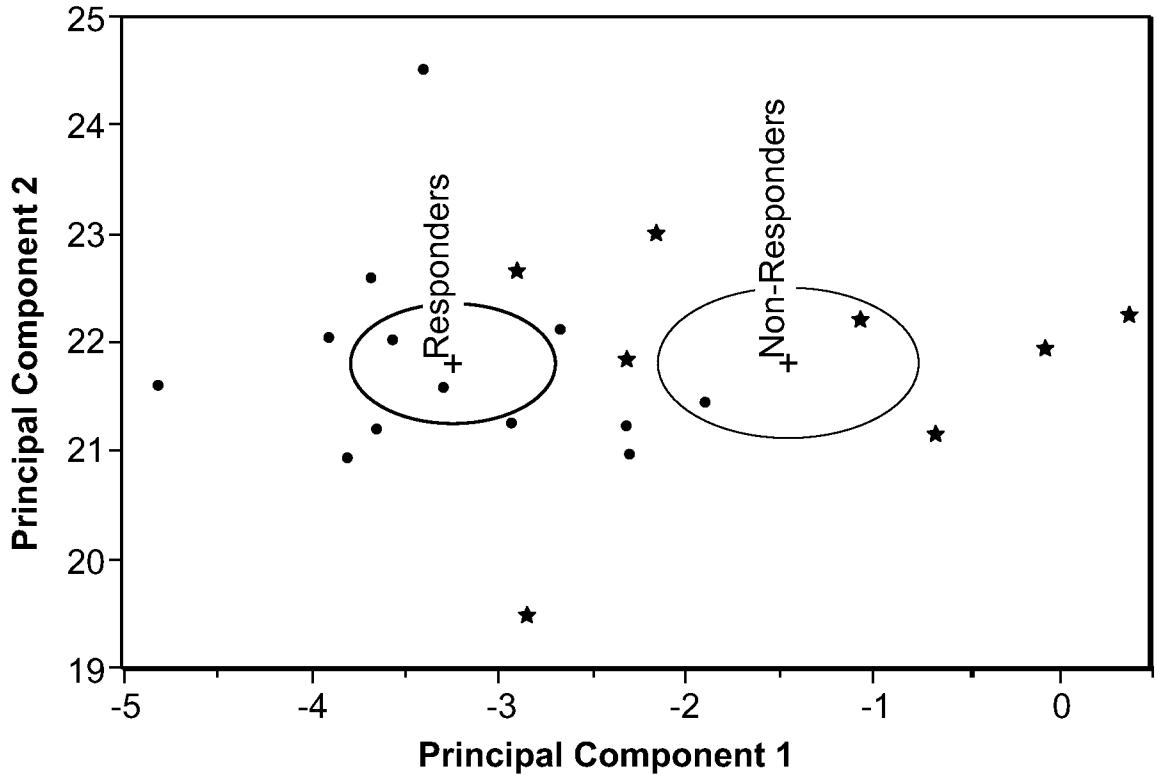
25
24
23
22
21
20
19
Principal Component 2
Responders
Non-Responders
-5 -4 -3 -2 -1 0
Principal Component 1

FIG. 29

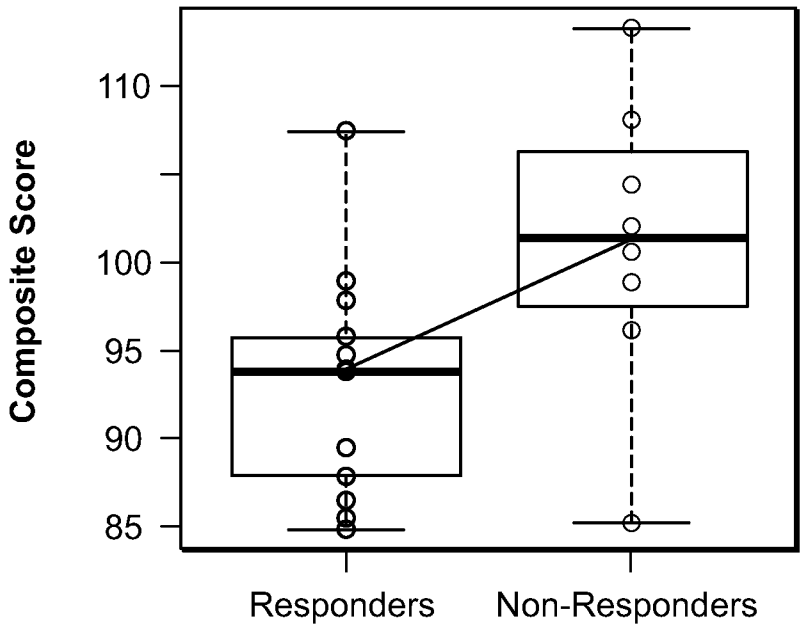
Composite Score
110
100
95
90
85
Responders
Non-Responders

FIG. 30

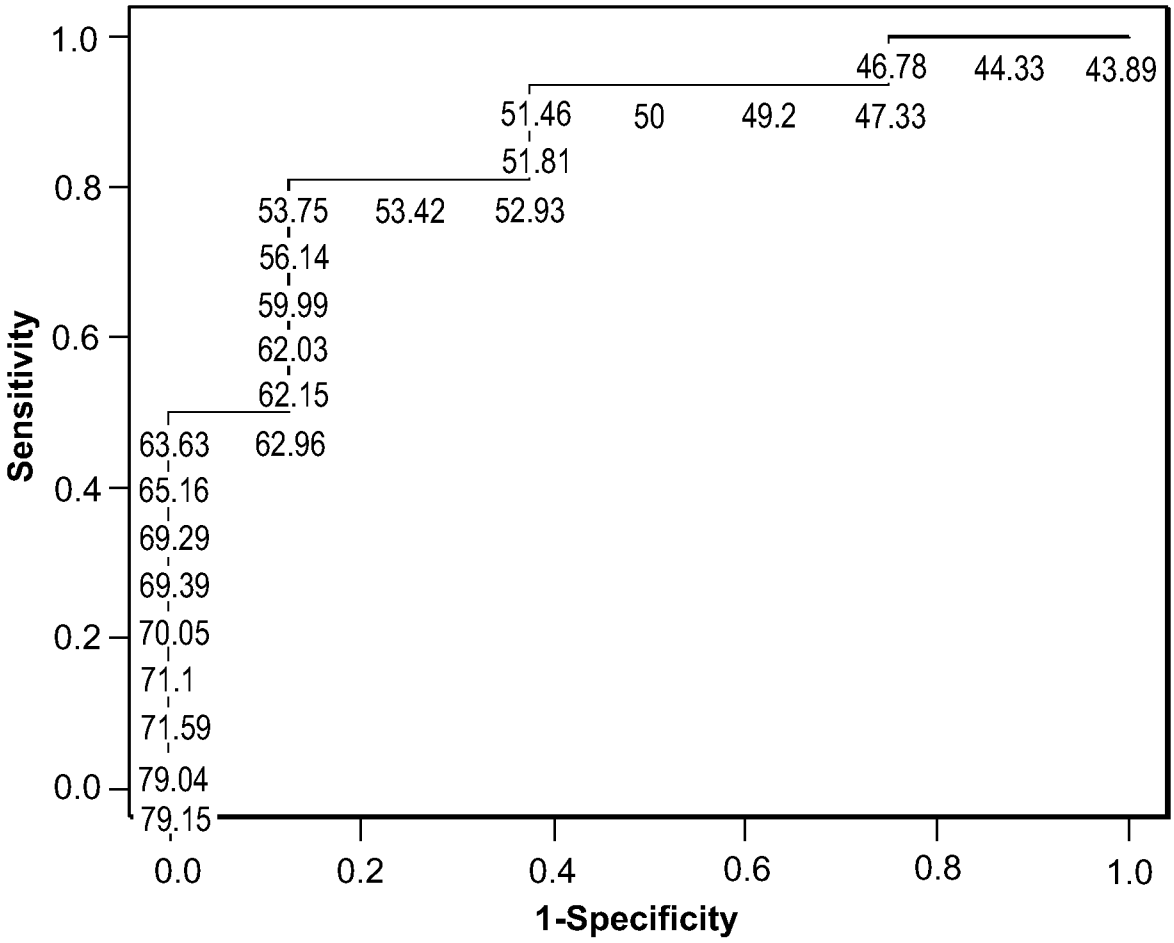
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
46.78
44.33
43.89
51.46
50
49.2
47.33
51.81
53.75
53.42
52.93
56.14
59.99
62.03
62.15
63.63
62.96
65.16
69.29
69.39
70.05
71.1
71.59
79.04
79.15
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity

FIG. 31

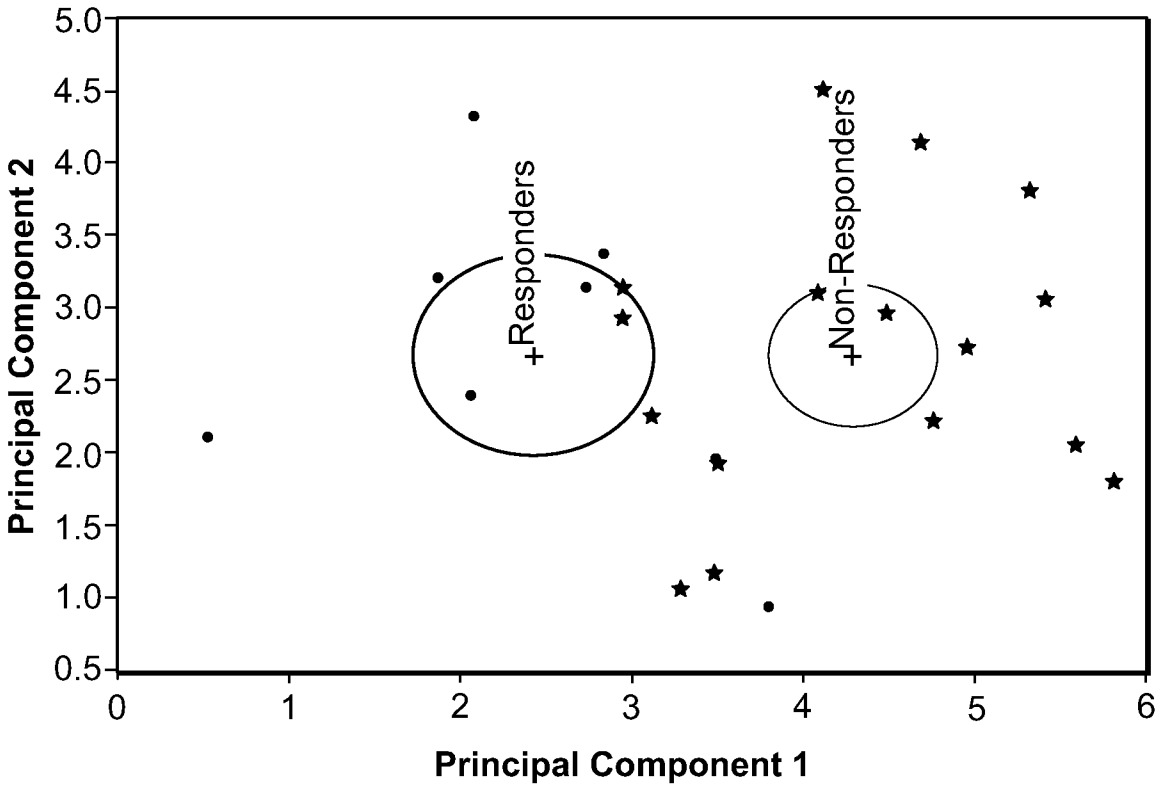
Principal Component 2
Principal Component 1
Responders
Non-Responders
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4
5
6

FIG. 32

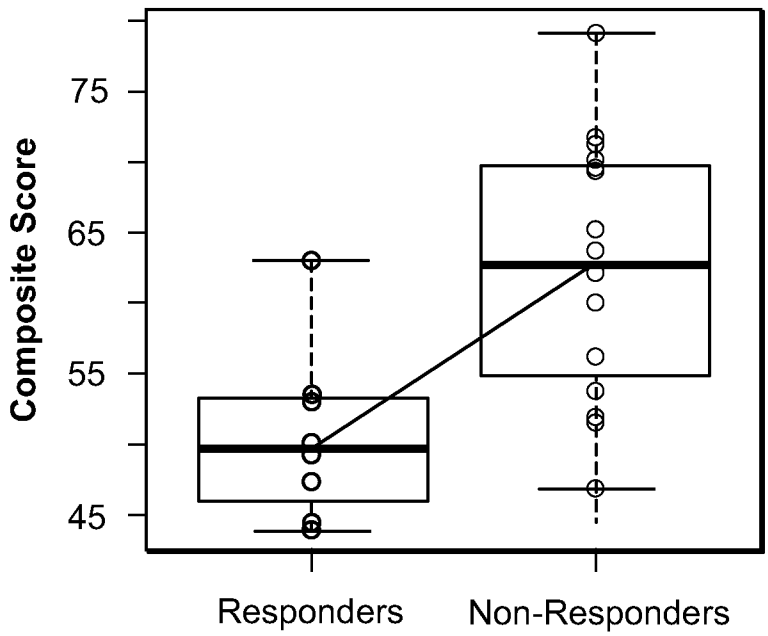
Composite Score
75
65
55
45
Responders
Non-Responders

FIG. 33

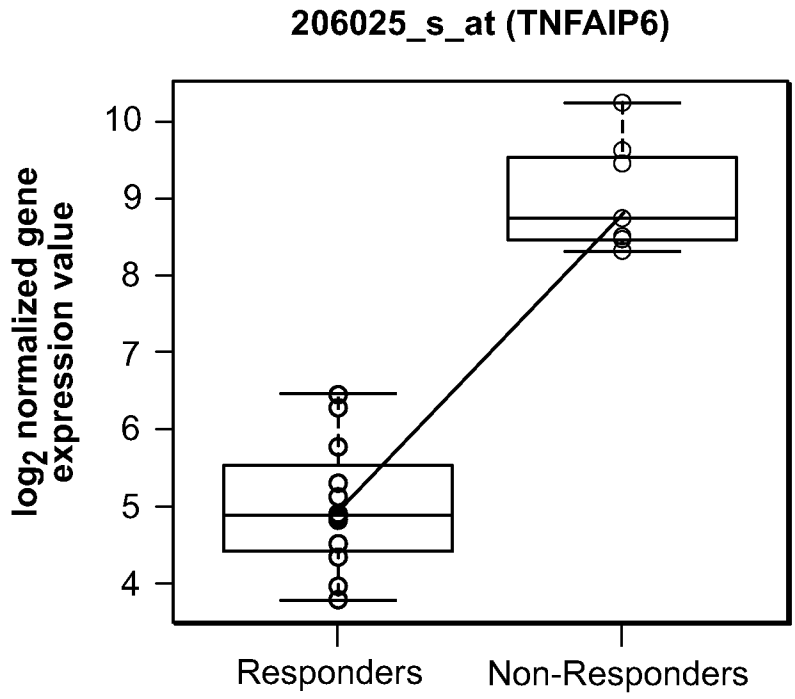
206025_s_at (TNFAIP6)
log2 normalized gene expression value
10
9
8
7
6
5
4
Responders
Non-Responders

FIG. 34

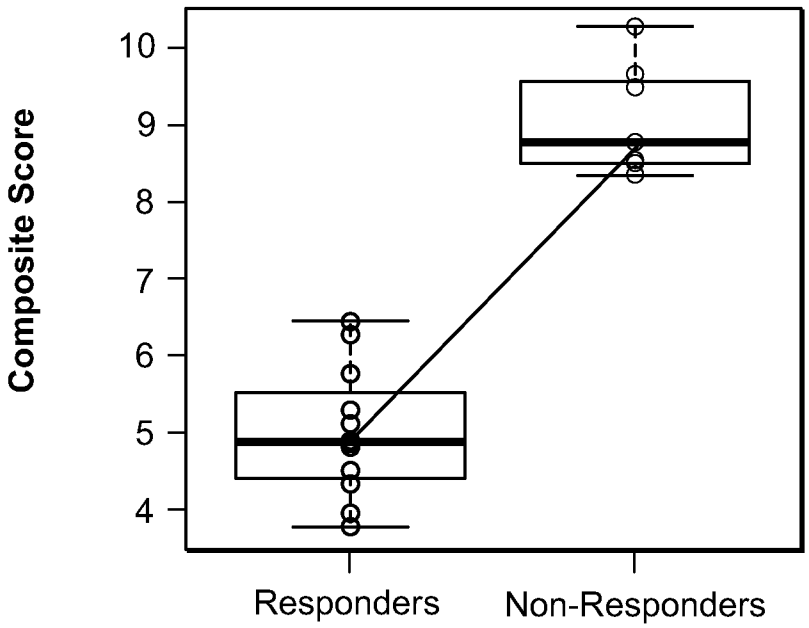
Composite Score
10
9
8
7
6
5
4
Responders
Non-Responders

FIG. 35

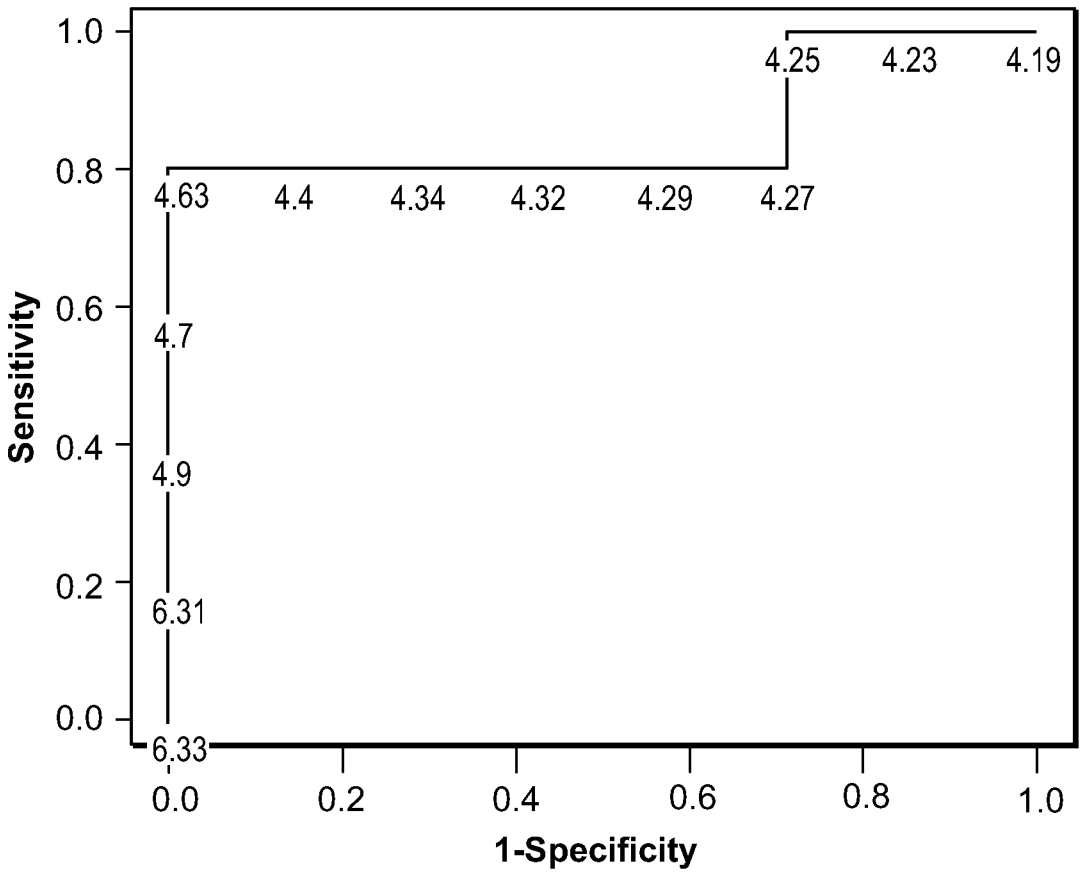
1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
4.25
4.23
4.19
4.63
4.4
4.34
4.32
4.29
4.27
4.7
4.9
6.31
6.33
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity

FIG. 36

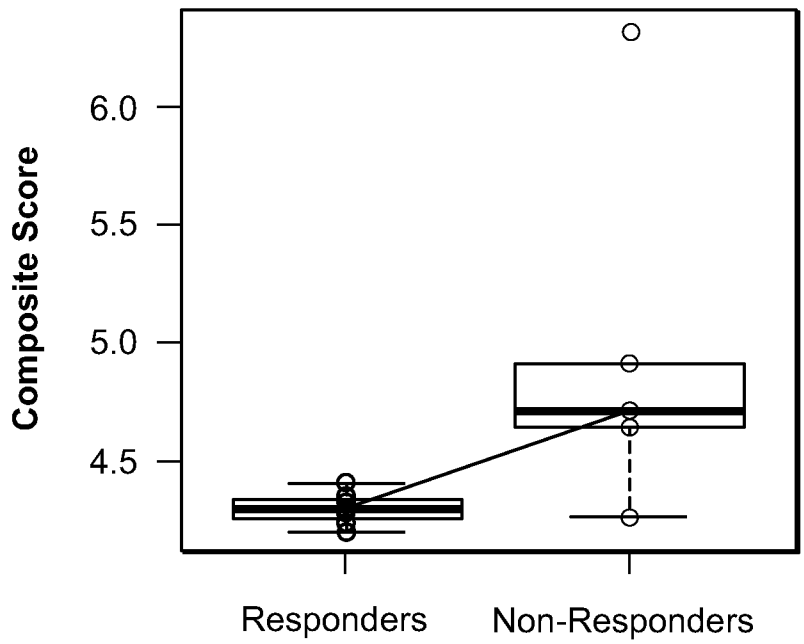
6.0
5.5
5.0
4.5
Composite Score
Responders
Non-Responders

FIG. 37

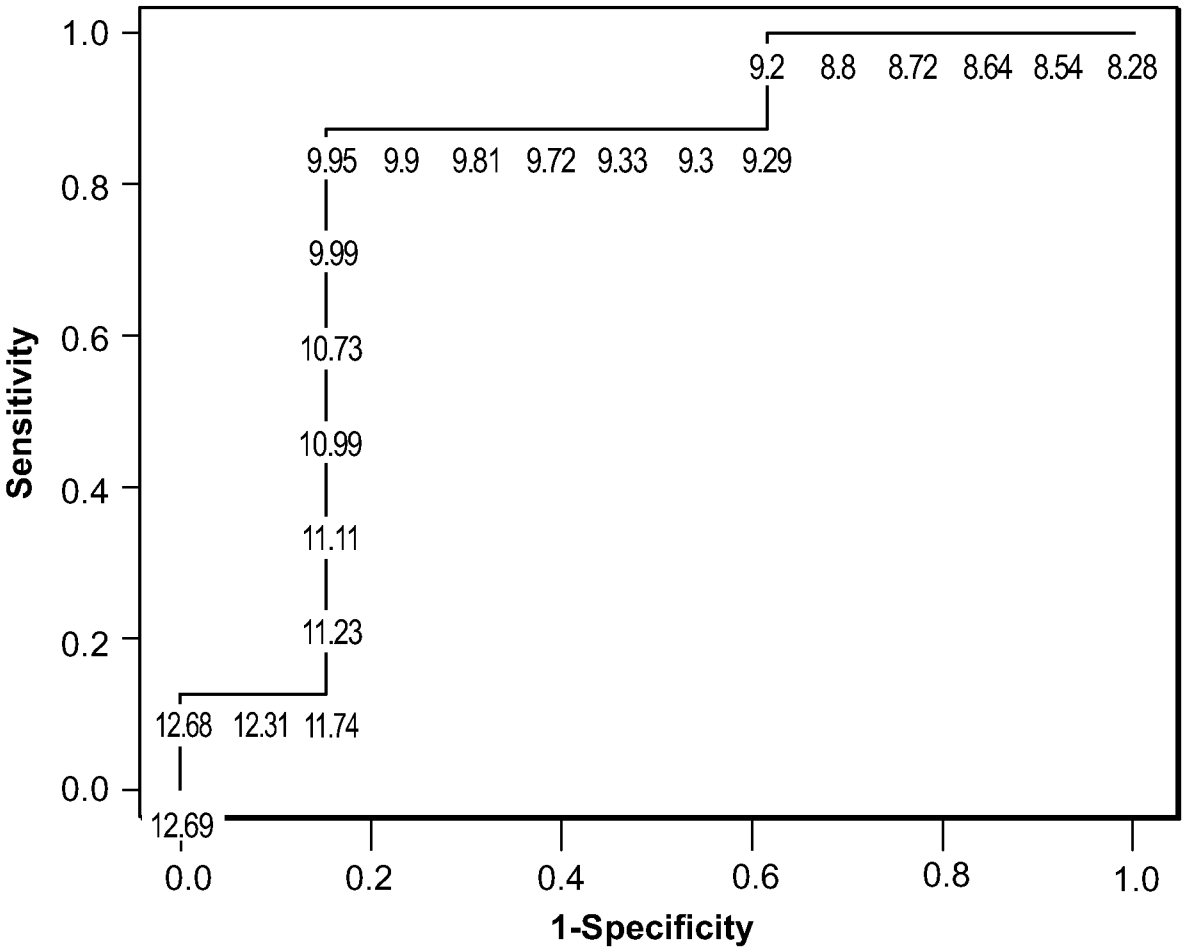
1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
9.2 8.8 8.72 8.64 8.54 8.28
9.95 9.9 9.81 9.72 9.33 9.3 9.29
9.99
10.73
10.99
11.11
11.23
12.68 12.31 11.74
12.69
0.0 0.2 0.4 0.6 0.8 1.0
1-Specificity

FIG. 38

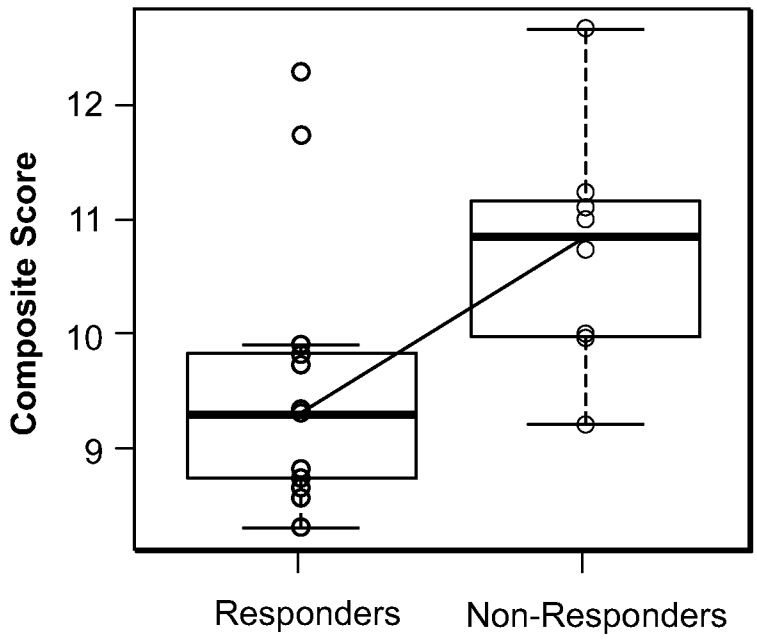
12
11
10
9
Composite Score
Responders
Non-Responders

FIG. 39

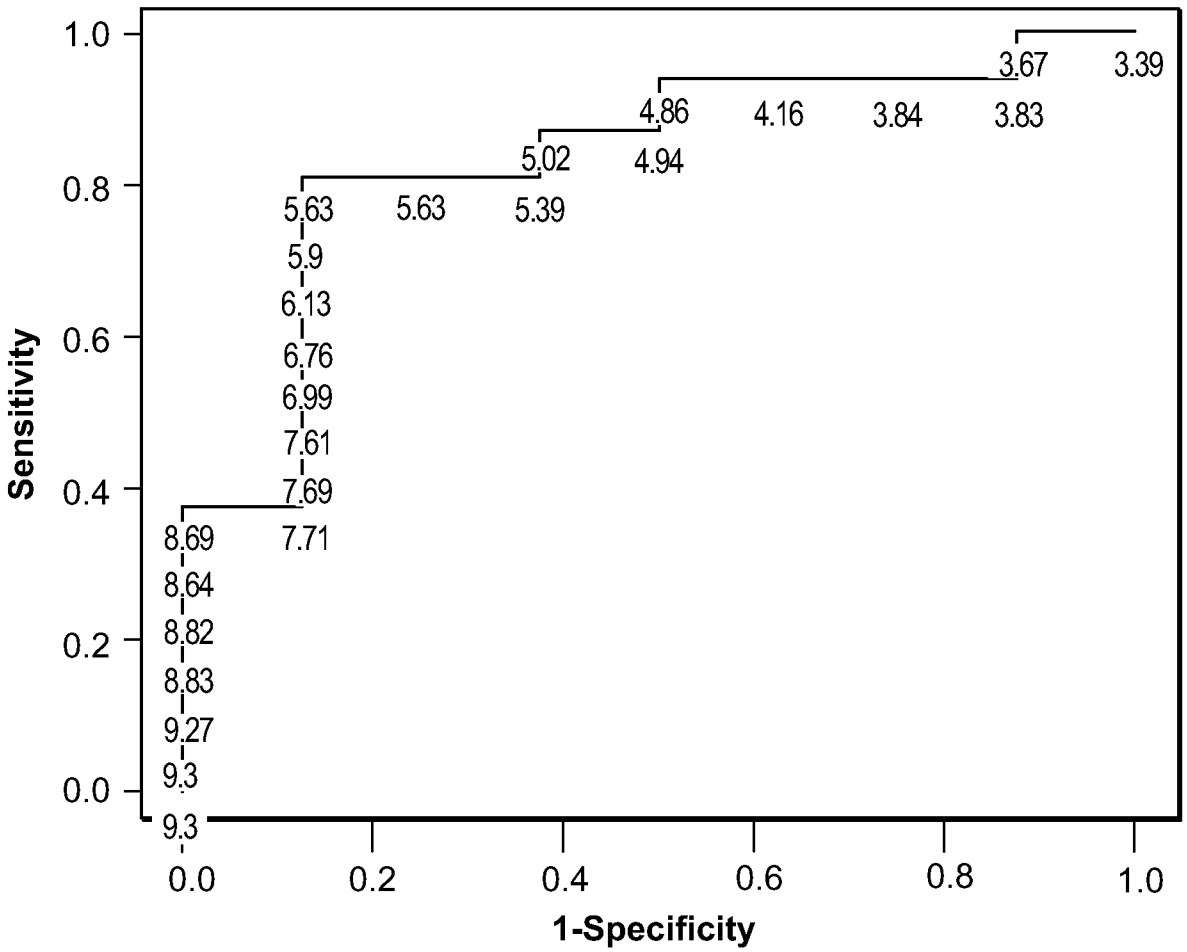
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
3.67
3.39
4.86
4.16
3.84
3.83
5.02
4.94
5.63
5.63
5.39
5.9
6.13
6.76
6.99
7.61
7.69
8.69
7.71
8.64
8.82
8.83
9.27
9.3
9.3

FIG. 40

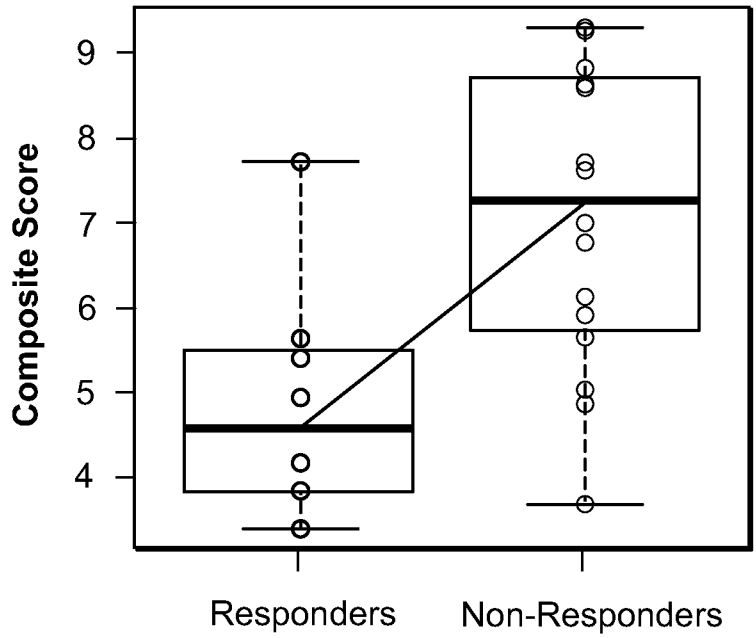
Composite Score
4
5
6
7
8
9
Responders
Non-Responders

FIG. 41

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/31145

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12Q 1/68; C12N 15/12, 15/52; A61K 39/395, 35/38; G01N 33/53 (2016.01)
CPC - C12Q 1/6844, 1/6876, 1/6883; C12N 15/52; A61K 39/39583, 35/38; G01N 33/53
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8): C12Q 1/68; C12N 15/12,15/52, 1/19,1/21, 5/63, 5/10; A61K 39/395, 35/38; A61P 35/00, 19/02; C07K 16/24, 16/28; G01N 33/53 (2016.01) CPC: C12Q 1/6844, 1/6876, 1/6883; C12N 15/52; A61K 39/39583, 35/38, 2039/505, 39/3955, 45/06; G01N 33/53, 2800/065

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); Google Scholar; Pubmed; EBSCO
Keywords: inflammatory bowel disease, Crohn's, colitis, gene, nucleic acid, protein, antibody*, hybridiz*, cystatin A, overexpress*, upregulat*, TNF*, adalimumab, infliximab

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	WO 2014/060785 A2 (EGIS PHARMACEUTICALS PUBLIC LIMITED COMPANY) 24 April 2014; abstract; page 2, lines 9, 26; page 3, lines 4, 25-30; page 7, lines 1-8; page 8, line 30 – page 9, line 6; page 15, lines 13-15; figures 3, 5; claims 9, 12	1-3, 4/1-3, 5/4/1-3, 6/1-3, 7/1-3, 28-31, 41 ---------- 32-34, 35/33-34, 38, 46-47, 54/46-47, 55/54/46-47
Y	US 2013/0330357 A1 (ABBVIE BIOTECHNOLOGY LTD,) 12 December 2013; abstract; paragraphs [0011], [0041]; SEQ ID NO.s 1-8 (supplemental document with the sequences 1-8)	32-34, 35/33-34, 38
Y	US 2014/0017174 A1 (ATREYA, R et al) 16 January 2014; paragraphs [0012], [0014], [0016], [0029], [0056], [0083], [0092]-[0093], [0155], [0213]-[0214]; claim 34	46-47, 54/46-47, 55/54/46-47
A	US 2007/0042364 A1 (MANNICK, EE et al) 22 February 2007; abstract; paragraph [0014]	48/46-47, 49/48/46-47, 50/46-47, 51/50/46-47
A	US 2012/0165220 A1 (COLGAN, SP et al) 28 June 2012; paragraph [0028]; figure 1B	48/46-47, 49/48/46-47, 50/46-47, 51/50/46-47
A	(WADDELL, A et al) Intestinal CCL11 and Eosinophilic Inflammation Is Regulated by Myeloid Cell-Specific Re1A/p65 in Mice. The Journal of Immunology. 05 April 2013; Vol. 190, No. 9; pages 1-14; page 3, column 1, paragraph 2; page 6, column 1, paragraph 2; page 7, column 1, paragraph 1; page 9, column 2, paragraph 2; DOI: 10.4049/jimmunol.1200057.	48/46-47, 49/48/46-47, 50/46-47, 51/50/46-47, 52/46-47, 53/52/46-47

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 July 2016 (27 July 2016)	22 AUG 2016
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/31145

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/31145

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 8-27, 36-37, 39-40, 42-45, 56-58
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/31145

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2002/0593367 A2 (BOARD OF SUPERVISORS OF LOUISIANA STATE UNIVERSITY AND AGRICULTURAL AND MECHANICAL COLLEGE) 01 August 2002; page 3, paragraph 3 – page 4, paragraph 1	50/46-47, 51/50/46-47,
A	(HUTTER, J et al) Role of the C-Type Lectin Receptors MCL and DCIR in Experimental Colitis. PLoS One. 28 July 2014; Vol. 9, No. 7; pages 1-11; abstract; page 2, column 1, paragraph 1; page 2, column 2, paragraph 1; DOI: 10.1371/journal.pone.0103281	52/46-47, 53/52/46-47
A	(KAMBA, A et al) Potential Association Between TLR4 and Chitinase 3-like 1 (CHI3L1/YKL-40) Signaling on Colonic Epithelial Cells in Inflammatory Bowel Disease and Colitis-Associated Cancer. Current Molecular Medicine. August 2013; Vol. 13, No. 7; pages 1-22; abstract; page 2, paragraph 3	52/46-47, 53/52/46-47
A	(WITKIN, SS et al) Influence of Interleukin-1 Receptor Antagonist Gene Polymorphism on Disease. Clinical Infectious Diseases. 07 December 2001; Vol. 34, No. 2; pages 204-209; abstract; page 205, column 1, paragraph 2	52/46-47, 53/52/46-47
A	(GENUA, M et al) The Triggering Receptor Expressed on Myeloid Cells (TREM) in Inflammatory Bowel Disease Pathogenesis. Journal of Translational Medicine. 28 October 2014; Vol. 12, No. 293; pages 1-12; abstract; page 3, column 1, paragraph 3; DOI: 10.1186/s12967-014-0293-z	52/46-47, 53/52/46-47