Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2918608 A1 2014/02/06

(21) 2 918 608

(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

(13) **A1**

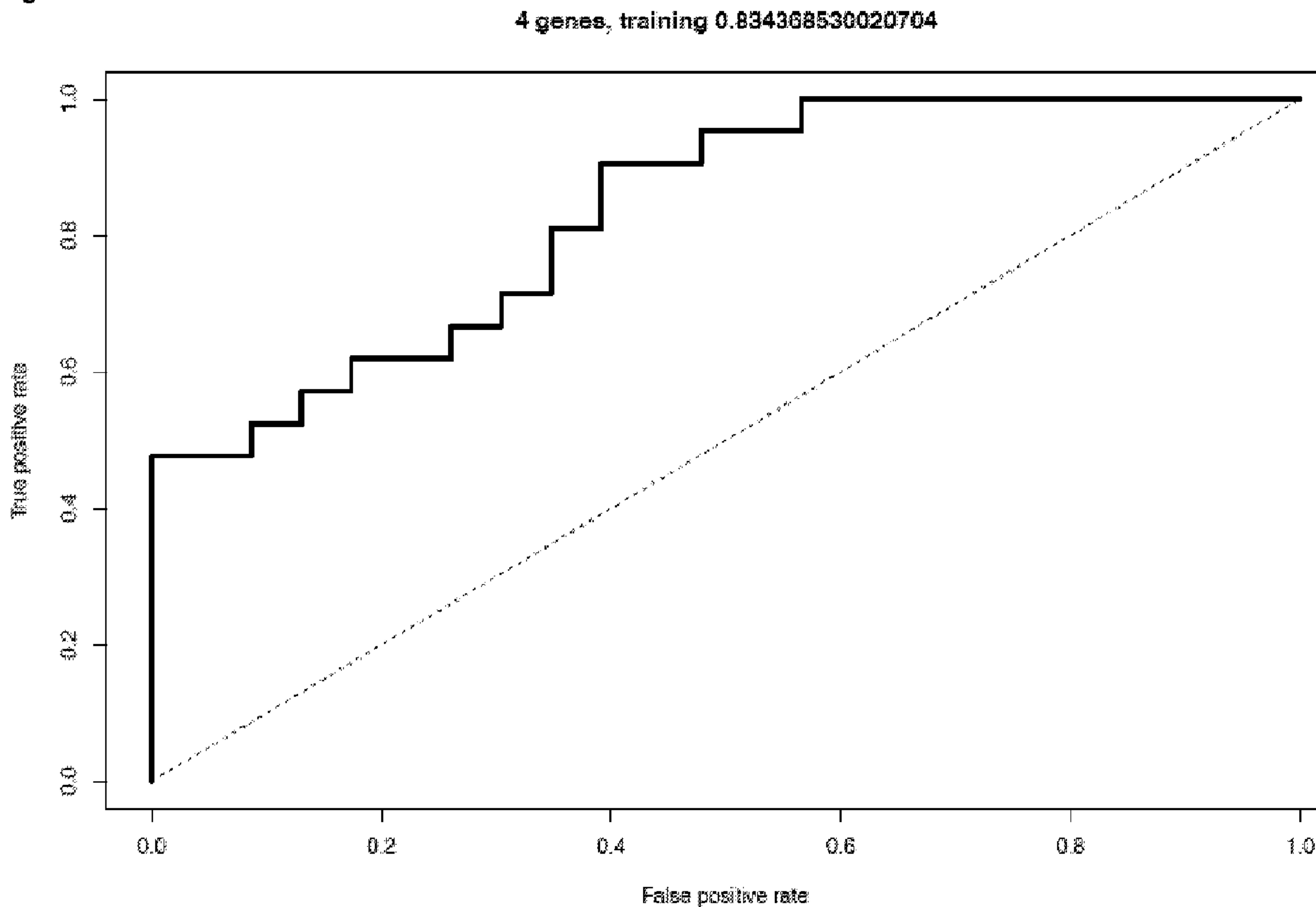
(86) Date de dépôt PCT/PCT Filing Date: 2013/08/02
(87) Date publication PCT/PCT Publication Date: 2014/02/06
(85) Entrée phase nationale/National Entry: 2016/01/18
(86) N° demande PCT/PCT Application No.: US 2013/053511
(87) N° publication PCT/PCT Publication No.: 2014/022826
(30) Priorité/Priority: 2012/08/03 (US61/679,394)
(51) Cl.Int./Int.Cl. *C12Q 1/68* (2006.01)

(71) Demandeur/Applicant:
ICAHN SCHOOL OF MEDICINE AT MOUNT SINAI, US

(72) Inventeurs/Inventors:
SAENGER, YVONNE, US;
SIVENDRAN, SHANTHI, US;
CHANG, RUI, US;
DIFEO, ANALISA, US

(74) Agent: BENOIT & COTE INC.

(54) Titre : BIOMARQUEUR ASSOCIE AU RISQUE DE RECURRENCE DU MELANOME
(54) Title: BIOMARKER ASSOCIATED WITH RISK OF MELANOMA REOCCURRENCE



(57) Abrégé/Abstract:
The present invention provides a method of predicting the risk of reoccurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following: a. obtaining a RNA-containing sample of the previously removed melanoma tissue containing RNA from the patient; b. treating the sample to determine from the RNA contained in the sample the

(57) Abrégé(suite)/Abstract(continued):
level of expression of a plurality of preselected genes; and c. comparing the level of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has a reduced risk of reoccurrence of melanoma.

CA 02918608 2016-01-18

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 February 2014 (06.02.2014)

(10) International Publication Number
WO 2014/022826 A3

(51) International Patent Classification:
G01N 33/574 (2006.01)

(21) International Application Number:
PCT/US2013/053511

(22) International Filing Date:
2 August 2013 (02.08.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/679,394 3 August 2012 (03.08.2012) US

(71) Applicant *(for all designated States except US)*: **ICAHN SCHOOL OF MEDICINE AT MOUNT SINAI** [US/US]; One Gustave L. Levy Place, New York, NY 10029 (US).

(72) Inventors; and
(71) Applicants *(for US only)*: **SAENGER, Yvonne** [US/US]; 730 Columbus Ave., #14f, New York, NY 10025 (US). **SIVENDRAN, Shanthi** [US/US]; 88 Lake View Drive, Gettysburg, PA 17325 (US). **CHANG, Rui** [CN/US]; 215 East 95th Street, New York, NY 10128 (US). **DIFEO, Analisa** [US/US]; 18 Ivy Way, Port Washington, NY 11050 (US).

(74) Agent: **WHITE, John, P.**; Cooper & Dunham LLP, 30 Rockefeller Plaza, New York, NY 10112 (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, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, 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).

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

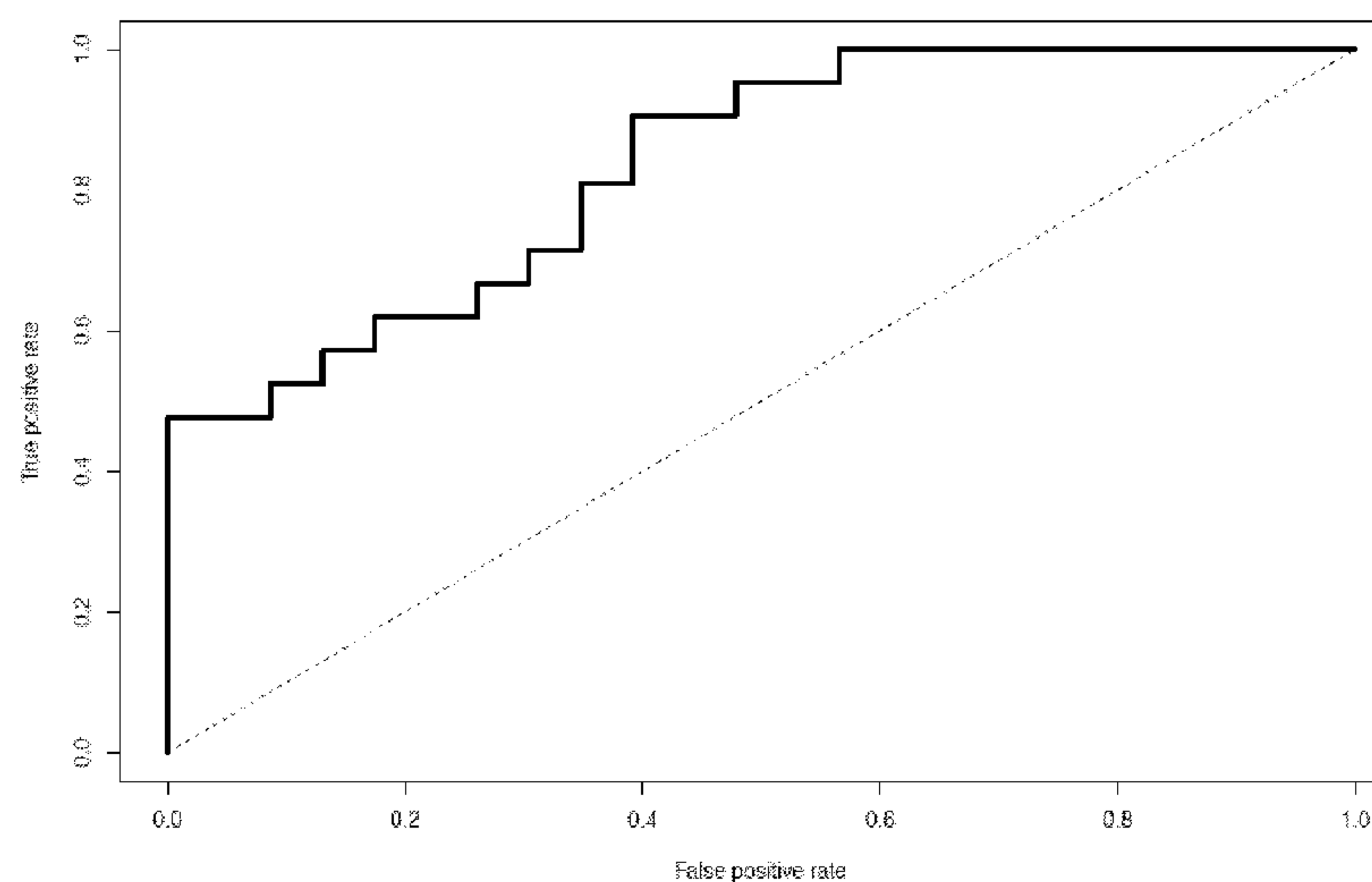
— *with sequence listing part of description (Rule 5.2(a))*

[Continued on next page]

(54) Title: BIOMARKER ASSOCIATED WITH RISK OF MELANOMA REOCCURRENCE

Figure 8

4 genes, training 0.834368530020704



(57) Abstract: The present invention provides a method of predicting the risk of reoccurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following: a. obtaining a RNA-containing sample of the previously removed melanoma tissue containing RNA from the patient; b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and c. comparing the level of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; wherein a higher level of expression of the plurality of pre- selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has a reduced risk of reoccurrence of melanoma.

WO 2014/022826 A3

CA 02918608 2016-01-18

WO 2014/022826 A3

(88) Date of publication of the international search report:
27 March 2014

BIOMARKER ASSOCIATED WITH RISK OF MELANOMA REOCCURRENCE

This application claims priority of U.S. Provisional Application No. 61/679,394, filed August 3, 2012, the entire contents of which are hereby incorporated herein by reference.

Throughout this application, various publications are referenced. Full citations for these publications may be found at the end of the specification immediately preceding the claims. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

Background of Invention

Melanoma is an aggressive malignancy associated with five-year survival rates under 5% in patients with metastatic disease.[1] Despite successful excision of the primary lesion, a five-year survival of only 68% is expected in cases of primary melanoma greater than 2mm in thickness.[5, 6] Increasing depth of the primary tumor and the presence of high risk histopathology are predictive of recurrence across populations, but do not accurately assess risk in individual patients.[7] Sentinel lymph node biopsy (SLNB) is an invasive procedure that offers limited prognostic information and has no proven survival benefit.[8] Improved biomarkers are needed to identify patients at high risk for recurrence and death.

Expression profiling has never been systematically performed in formalin-fixed, paraffin-embedded (FFPE) primary melanoma.[9] In contrast, Oncotype DX measures the expression of a 21-gene panel and offers prognostic information for patients with breast cancer.[10] The development of similar biomarkers in melanoma has been limited due in part to the clinical standard of entire tumor fixation in formalin which leads to low yields of extractable RNA and limits the quality of RNA available for molecular studies. As a result, studies of primary melanoma have

relied on cell lines, limited supplies of frozen tissue, or focused on profiling microRNA, which is less subject to degradation.[11-17] Even in rare cases where frozen tissue is available, RNA extraction is difficult due to the fibrous nature of cutaneous tissues.[9, 12]

Thus, there is a need for suitable methods and markers for providing prognostic information related to melanoma.

Summary of the Invention

The present invention provides a method of predicting the risk of reoccurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a RNA-containing sample of the previously removed melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the level of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has a reduced risk of reoccurrence of melanoma, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by a method of the invention, comprising administering an immunotherapy to the patient.

The present invention also provides a method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of

reoccurrence of melanoma by a method of the invention, comprising testing the human patient for recurrence of melanoma more frequently than a corresponding patient who was determined to have a reduced risk of reoccurrence of melanoma
5 would be tested for recurrence.

The present invention also provides a method of treating a patient afflicted with melanoma which comprises the following:

10 a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;
15 c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and
d. administering a therapy to the patient if there is a higher level of expression of the plurality of pre-
20 selected genes in the sample as compared with the predetermined reference level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the
25 following genes:

a. CD2;
b. KLRK1;
c. ITK; and
30 d. HLAE.

The present invention also provides a method of treating a patient afflicted with melanoma which comprises the following:

35 a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;

b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;

c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and

d. administering a therapy to the patient if there is a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;

b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;

c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and

d. administering a therapy to the patient if the level of expression of the plurality of pre-selected genes in the sample is i) lower as compared with the predetermined reference upper level of expression of such genes and ii) higher as compared with the predetermined reference lower level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

5 a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

10 The present invention also provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

15 a. obtaining a RNA-containing sample of melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
20 c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-
25 selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

30 wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
35 c. ITK; and
d. HLAE.

The present invention also provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;

b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a level of expression of the plurality of pre-selected genes in the sample is i) lower as compared with the predetermined reference upper level of expression of such genes and ii) higher as compared with the predetermined reference lower level of expression of such genes, indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of determining whether a therapy is effective for treating patients afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from at least one patient afflicted with melanoma, which at least one patient was administered the therapy for treatment of the melanoma previous to collection of the sample;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to the levels of

expression of the plurality of pre-selected genes to the expression level of each such gene in a corresponding at least one patient not administered the therapy;

5 wherein a higher level of expression of the plurality of pre-selected genes in the sample of step a) as compared with the level of expression of such genes in a corresponding at least one patient not administered the treatment indicates that the therapy is effective for treating patients afflicted with 10 melanoma, and

wherein the plurality of pre-selected genes comprises the following genes:

15 a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

20 The present invention also provides a method of determining whether a patient afflicted with melanoma and which patient was administered a therapy has exhibited a positive clinical response to the therapy which comprises the following:

25 a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient, which RNA-containing sample was removed from the patient who was administered the therapy;
b. treating the sample to determine from the RNA contained 30 in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

35

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the

predetermined reference level of expression of such genes indicates that the patient has exhibited a positive clinical response to the therapy, and

5 wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
10 c. ITK; and
d. HLAE.

The present invention also provides a method of determining whether a treatment should be administered to patients
15 afflicted with melanoma as an adjuvant or a neoadjuvant therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from at least one patient afflicted with
20 melanoma;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the
25 plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the
30 predetermined reference level of expression of such genes indicates that the treatment should be administered to patients afflicted with melanoma as a neoadjuvant therapy, and a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined
35 reference level of expression of such genes indicates that the treatment should be administered to patients afflicted with melanoma as an adjuvant therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

5 a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

10 The present invention provides a method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma
15 tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene
20 or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a higher level of expression product of the gene or
25 each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

30 The present invention provides a method of treating a patient afflicted with melanoma which comprises the following:
a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes
35 in the sample;
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined

reference level of the expression product for each such gene; and

d. administering a therapy to the patient if there is a higher level of expression of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene.

The present invention provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample;
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene; and
d. administering a therapy to the patient if there is a lower level of expression of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene.

The present invention provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or the each of two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

The present invention provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of the gene or each of the two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a lower level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

The present invention provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;

b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a level of of the expression product of the gene or each of the two or more genes that is i) lower as compared with a predetermined reference upper level of expression for each such gene and ii) higher as compared with a predetermined reference lower level of expression for each such gene, indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

The present invention provides a method of determining whether a therapy is effective for treating patients afflicted with melanoma which comprises the following:

a. obtaining a sample of melanoma tissue from at least one patient afflicted with melanoma, which at least one patient was administered the therapy;
b. treating the sample to determine the level of an expression product of the gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to the level of the expression product of the gene or each of the two or more genes in a corresponding at least one patient not administered the treatment,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample of step (a) as compared with the level of expression product of the gene or each of the two or more genes in the corresponding at least one patient not administered the treatment indicates that the

therapy is effective for treating patients afflicted with melanoma.

The present invention provides a method of determining whether 5 a patient afflicted with melanoma and which patient was administered a therapy has exhibited a positive clinical response to the therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
10 b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined 15 reference level for each such gene,

wherein a higher level of expression the expression product of the gene or each of the two or more genes as compared to the predetermined reference level of the expression product of 20 each such gene indicates that the patient has exhibited a positive clinical response to the therapy.

The present invention provides a method method of determining whether a treatment should be administered to patients 25 afflicted with melanoma as an adjuvant or a neoadjuvant therapy which comprises the following:

a. obtaining a sample of melanoma tissue from at least one patient afflicted with melanoma;
30 b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined 35 reference level for each such gene,

wherein a higher level of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of expression for each such gene indicates that the treatment should be administered to patients afflicted with melanoma as a neoadjuvant therapy, and a lower level of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the treatment should be administered to patients afflicted with melanoma as an adjuvant therapy.

The present invention provides a method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma tissue from the patient;
b. treating the sample to determine the level of the expression product of the CD2 gene in the melanoma tissue sample; and
c. comparing the level of the expression product of the CD2 gene to a predetermined reference level of the expression product of the CD2 gene,

wherein a higher level of expression product of the CD2 gene in the sample as compared with the predetermined reference level of the expression product of the CD2 gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

The present invention provides a method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma tissue from the patient;

b. treating the sample to determine the level of the expression product of the X gene in the melanoma tissue sample; and

c. comparing the level of the expression product of the X gene to a predetermined reference level of the expression product of the X gene,

wherein a higher level of expression product of the X gene in the sample as compared with the predetermined reference level of the expression product of the X gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

Brief Description of the Drawings

Figure 1. RNA was extracted from FFPE primary melanomas and analyzed using NanoString Technology. Relative levels of mRNA expression for each sample are depicted according to the color scale shown, with each column representing a patient sample and each row representing a gene. Genes are arranged in the heatmap and listed in this order on the right from most differentially expressed (top) to least (bottom). Panel A shows the relative expression of 92 inflammatory genes found to be differentially expressed between 21 recurrent (light grey) and 23 non-recurrent (dark grey) patient samples. Of these 92 genes, 90 were up-regulated in melanomas which did not recur (*see, e.g.*, Table 4 and 5 for fold change). In panel B, validation of these findings in 37 melanomas is shown. 41 genes out of 63 tested were differentially expressed between recurrent and non-recurrent melanomas, all of which were up-regulated in the non-recurrent group.

Figure 2. A 21-gene signature able to predict melanoma recurrence is proposed. Panel A shows a bar graph representing the number of times each gene was selected in a leave-4-out cross-validation module. From this cross validation, a compact list of genes was selected using a linear regression model to compose a 21-gene signature, listed in the inset. In panel B, the receiver operating characteristic (ROC) curves for the statistical model using these 21 genes is shown in the training (left) and validation (right) cohorts. By definition, random classification of a sample as recurrent or non-recurrent provides an AUC of 50% (dotted line). The AUC for the proposed gene signature was 0.983 and 0.794 in the training and validation cohorts respectively. In panel C, the coefficient of determination (R^2) was calculated using a linear model for each characteristic with and without the gene signature for both training and validations sets. When combined with the gene signature, the R^2 value drastically increases for each characteristic,

indicating improved ability to predict recurrence. Values are provided in the Table 7.

Figure 3. Immunohistochemistry (IHC) using anti-CD2 monoclonal antibody was performed to assess risk of disease recurrence. In panel A, photographs of non-recurrent (left) and recurrent (right) patient specimens are shown. In the non-recurrent patient sample, a brisk peritumoral infiltrate is seen at 4x magnification on H&E (top left) corresponding with cells staining positive for CD2 at 10x and 40x power (bottom left). In the recurrent patient sample, no such infiltrate is seen at 4x power on H&E (top right) with few corresponding cells positive for CD2 at 10x and 40x magnification (bottom right). Panel B displays the average number of cells counted at 40x magnification in 8 random HPFs for patients in the training and validation groups.* Significantly more CD2 positive cells were found in non-recurrent, as compared to recurrent, melanomas in the training (left, $p=0.0172$) and validation (right, $p=0.0032$) cohorts. In Panel C, linear regression is used to assess correlation of NanoString with IHC for CD2 in the training (left) and validation (right) groups. The two methods correlate with one another with r values of 0.847 and 0.538 for the training ($p<0.0001$) and validation ($p=0.0026$) cohorts respectively.

* Inadequate tissue in FFPE block to make slides for IHC in 6 recurrent patients and 2 non-recurrent patients in the validation cohort.

Figure 4. Kaplan-Meier curves of overall survival are shown. In panel A, patients in the training and validation cohorts were classified as either signature + (red) or signature - (black) based on expression of the 21-gene panel defined in Figure 2. Patients with a positive gene signature had a higher overall survival compared to those with a negative signature ($p<0.0001$). In panel B, patients were classified according to depth as either <4mm or >4mm. Depth trends

towards but does not reach statistical significance in predicting survival for the same cohort ($p=0.0509$).

Figure 5. The distribution of the AUC values for the 900 cross validation iterations, using a leave-4-out approach, are shown in the training (A) and validation (B) sets. Distribution of AUC values in training and validation sets during cross-validation.

Figure 6. Normalized mRNA count of the 21 gene signature in training and validation sets. Analysis of a gene signature for predicting melanoma recurrence using NanoString is shown. Panels A displays the differential in mRNA counts between recurrent and non-recurrent patients for 21 genes composing the proposed signature. Values represent the number of mRNA transcripts counted by the nCounter® Analysis System of NanoString Technology. For every gene, a marked upregulation in mRNA expression is demonstrated in patients with non-recurrent melanoma when compared with patients with recurrent disease. These 21 genes were validated by NanoString in an independent set of samples displayed in Panel B. With the exception of IFNAR1, all genes were upregulated in the non-recurrent group with 14 out of 21 genes found to be differentially expressed to a statistically significant degree ($p < 0.05$) in the validation cohort.

Figure 7. Percent survival for signature positive and negative groups in training and validation cohorts. Kaplan-Meier curves of survival based on a 21-gene signature are shown for the training and validation cohorts. Patients were classified as either signature + (red) or signature - (black). In both the training (A) and validation (B) cohorts, patients with a positive gene signature had a higher overall survival compared to those with a negative signature, which conferred a poor prognosis (Training: $p < 0.0001$, Validation: $p = 0.0151$).

Figure 8. The receiver operating characteristic (ROC) curves for the statistical model using the 4 core genes is shown in the training cohort. By definition, random classification of a sample as recurrent or non-recurrent provides an AUC of 50% (dotted line). The AUC for the proposed gene signature was 0.834.

Figure 9. The receiver operating characteristic (ROC) curves for the statistical model using the 4 core genes is shown in the validation cohort. By definition, random classification of a sample as recurrent or non-recurrent provides an AUC of 50% (dotted line). The AUC for the proposed gene signature was 0.782.

Figure 10. The receiver operating characteristic (ROC) curves for the statistical model using the 15 core genes is shown in the training cohort. By definition, random classification of a sample as recurrent or non-recurrent provides an AUC of 50% (dotted line). The AUC for the proposed gene signature was 0.961.

Figure 11. The receiver operating characteristic (ROC) curves for the statistical model using the 15 core genes is shown in the validation cohort. By definition, random classification of a sample as recurrent or non-recurrent provides an AUC of 50% (dotted line). The AUC for the proposed gene signature was 0.785.

Figure 12. RNA was extracted from 40 FFPE stage II-III primary melanoma specimens and analyzed using NanoString Technology and 53 genes predictive of melanoma progression were identified using elastic net and random forest classifiers. In (**A**), a bar graph depicts the number of times each of the 53 genes was selected using a leave-4-out cross-validation. ROC curve to predict melanoma progression is shown in (**B**), AUC=1.000, **p<0.001.** In (**C**), relative levels of mRNA expression for each sample are depicted according to the color scale

shown, with each column representing a different patient sample and each row representing one of the 53 genes. Unsupervised hierarchical clustering was performed on both genes and samples. Patient who progressed are labeled in dark
5 gray and patients who did not progress are labeled in dark gray.

Figure 13. RNA was extracted from 48 FFPE stage II-III primary melanoma specimens and analyzed using NanoString Technology.
10 ROC curve to predict melanoma progression is shown in (**A**), AUC =0.787, **p<0.001.** In (**B**), distribution of AUC values using a leave-4-out cross-validation test is shown. In (**C**), relative levels of mRNA expression for each sample are depicted according to the color scale shown, with each column
15 representing a different patient sample and each row representing one of the 53 genes. Unsupervised hierarchical clustering was performed on both genes and samples. Patient who progressed are labeled in dark gray and patients who did not progress are labeled in light gray.
20

Figure 14. IHC using anti-CD2 monoclonal antibody was performed to assess risk of disease progression. (A) Photographs of a tumor expressing low levels of CD2 which did progress (left) and a tumor with high levels of CD2 from a
25 patient who remained disease free are shown (right). A brisk peri-tumoral in!ltrate is seen at 4x magnification on H&E in the tumor that did not progress. (B) The average number of CD2 positive cells counted at 40x magnification in 8 random HPFs for patients in the validation test set is shown. A linear
30 regression model is used to assess correlation in Nanostring with IHC for CD2 (C), CD4 (D), and CD5 (E) in the training set.

Figure 15. Kaplan-Meier curves of survival based on a 21-gene
35 signature and ulceration using a log rank Mantel-Cox test are shown for the training (**A**,**C**,**E**) and validation (**B**,**D**,**F**) populations. In both the training (**A**) and validation (**B**),

patients were classified as either signature + (gray) or signature - (black) based on median score. In the training set, a negative gene signature conferred inferior survival (**p<0.001**) while there was a strong trend for an ulcerated tumor (p=0.050), while patients with both negative indicators also had decreased survival (**p<0.001**). In the validation population, there was a trend towards shortened survival in patients with a negative gene signature (p=0.091) or an ulcerated tumor (p=0.101), while patients with both features had significantly diminished survival (**p=0.044**).

Figure 16. (**A**) Gene-protein interaction network surrounding the 53-gene panel. The 53-gene panel (pink) forms a denser network of gene-protein or protein-protein interactions (green) than the network surrounding the original 446-gene panel tested as shown in (**B**). (**C**) Network attributes of the 53-gene panel and 446-gene panel networks. Av. local CC and local cc SE represents the mean local clustering coefficient and standard error. Global CC denotes the global clustering coefficient. Local CC Average P-value and Global CC P-value represents the p-values associated with Local CC Average P-value and Global CC.

Figure 17. Co-expression network using WGCNA27 on 46 gene expression profiles in primary melanoma patients (GEO accession ID: GSE15605). The yellow dots compose a 758-gene module within the entire gene genome (pink). Red lines denote interactions between nodes involving nodes within the module.

Figure 18. ROC curves were generated for the refined 9 gene signatures on the training (A) and validation data set (B).

Figure 19. Cutaneous melanoma. CD2 Immunohistochemistry and disease recurrence Immunohistochemistry (IHC) using anti-CD2 monoclonal antibody was performed to examine differential expression of this protein in recurrent and non-recurrent patients. In panel A, photographs of non-recurrent (left) and

recurrent (right) patient melanoma tumors are shown. In the non-recurrent patient sample, a representative field at 40x magnification stained with H&E is shown (*top left*). The tumor contained many CD2-positive cells (*bottom left*), shown at 40x and 100x power (*inset*). The recurrent patient sample is shown at 40x power stained with H&E (*top right*), with few corresponding cells staining positive for CD2 (*bottom right*) at 40x and 100x magnification (*inset*). Panel B displays the average number of cells counted at 400x magnification in 8 random HPFs for 90 patients with primary melanoma. Significantly more CD2 positive cells were counted in patients with non-recurrent disease in comparison to patients who experienced recurrent disease ($p = 0.0003$). Panel C shows the Kaplan-Meier survival curves for patients with high- and low-CD2 counts. Patients with a higher number of CD2 positive cells in their primary tumor showed superior overall survival ($p=0.004$).

Figure 20. Cutaneous melanoma. Cell surface expression of CD2 A. Charged slides from six melanoma tumors were co-stained using immunoflourescence with an anti-CD3 antibody, a pan T-cell marker, and an anti-CD2 antibody. Patterns of staining, shown at 100X, were similar for CD3 and CD2. B. Serially-sectioned charged slides from three tumors were stained with an anti-CD16 antibody, expressed on Natural Killer cells, and an anti-CD2 antibody. Patterns of staining, shown at 100X, were dissimilar with a small amount of overlap.

Figure 21. Cutaneous melanoma. T-cell subtypes and CD2 expression

A. Tumors were co-stained with either anti-CD2 and anti-CD4 or anti-CD2 and anti-CD8 using immunoflourescence. Co-expression of CD2 and CD4, as well as CD2 and CD8, is shown. Images at 200X magnification. B. Immunohistochemistry (IHC) using anti-CD2, anti-CD4 and anti-CD8 monoclonal antibodies was performed to classify the T-cell subtype expressing CD2 in primary melanoma tumors. In panel A (*top*), a tumor found to contain

high levels of CD2-positive cells was found to contain both CD4-positive (*top left*) and CD8-positive (*top right*) T-cells. CD2 staining is seen to overlap with CD4 and CD8 staining. Images shown at 40x magnification. A tumor found to contain
5 low levels of CD2-positive cells (*bottom*) was shown to contain low levels of CD4-positive (*bottom left*) and CD8-positive cells (*bottom right*). Images shown at 40x magnification. In Panel C, a comparison of the number of CD4-positive cells to the number of CD8-positive cells within the same tumor is
10 shown for four tumors expressing high levels of CD2 and four tumors expressing low levels of CD2. The distribution of CD4 and CD8-positive cells is not different between the high CD2 and low CD2 groups (p=0.5152).

15 **Figure 22.** Cutaneous melanoma. CD2 Count and TIL Characterization

A subset of slides available for analysis were examined by the Department of Dermatopathology (RGP) and tumor-infiltrating lymphocytes were characterized as absent, non-brisk, or brisk
20 using established criteria. Panel A demonstrates representative images of (*from left to right*) absent, non-brisk and brisk tumor-infiltrating lymphocytes stained with H&E at 88x magnification. Panel B demonstrates the number of CD2 positive cells by IHC seen in tumors with absent, non-
25 brisk and brisk tumor-infiltrating lymphocytes. CD2 count increases from absent to non-brisk to brisk infiltrates (p=0.0004). In Panel C, the number of CD2 positive cells in recurrent (*right*) and non-recurrent (*left*) tumors is shown for tumors containing non-brisk infiltrates. Among this large
30 subset of patients, CD2 count remains significantly elevated in non-recurrent patients, compared to recurrent patients (p=0.0006), and a high CD2 count correlates with improved overall survival (Panel D; p=0.0318).

35 **Figure 23.** Cutaneous melanoma. CD2 and CD3 immunohistochemistry, disease recurrence and overall survival.

Immunohistochemistry (IHC) using anti-CD2 and anti-CD3 monoclonal antibodies was performed to examine differential expression of these proteins in a subset of recurrent and non-recurrent patients (n=21). Significantly more CD2 positive cells were counted in patients with non-recurrent disease in comparison to patients who experienced recurrent disease (Panel A *(left)*, p = 0.041). CD2 also significantly correlated with improved overall survival (Panel A *(right)*, p=0.0123). Panel B demonstrates the relationship between CD3 immunohistochemistry, disease recurrence and overall survival. *(Left)* CD3 immunohistochemistry did not significantly differ between recurrent and non-recurrent patient populations (p=0.0514). *(Right)* CD3 immunohistochemistry did not significantly correlate with improved overall survival (p=0.0873).

Figure 24. Cutaneous melanoma. Tumor-infiltrating lymphocyte topography and intensity.
Fifty-five primary tumor specimens were determined to have non-brisk tumor-infiltrating lymphocytes. A. Subclassification of non-brisk TILs in terms of topography (central, peripheral or both) did not distinguish recurrent and non-recurrent patients in this group. B. Subclassification of non-brisk TILs in terms of intensity (focal, multifocal, or segmental) did not distinguish recurrent from non-recurrent patients in this group.

Detailed Description of the Invention

The present invention provides a method of predicting the risk of reoccurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a RNA-containing sample of the previously removed melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the level of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has a reduced risk of reoccurrence of melanoma, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

In some embodiments, the method further comprises the step of creating a report summarizing said prediction.

In some embodiments, the expression level of each gene of the plurality of pre-selected genes is normalized relative to the expression level of one or more reference genes.

In some embodiments, the expression level of each such gene is normalized relative to the expression level of the following

genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

In some embodiments, the expression level of each such gene is normalized relative to the expression level of one or more of, or each of the following genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

In some embodiments, the predetermined reference level of expression is the expression level of the one or more reference genes.

The present invention also provides a method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by a method of the invention, comprising administering an immunotherapy to the patient.

The present invention also provides a method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by a method of the invention, comprising testing the human patient for recurrence of melanoma more frequently than a corresponding patient who was determined to have a reduced risk of reoccurrence of melanoma would be tested for recurrence.

The present invention also provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;

c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and
d. administering a therapy to the patient if there is a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and
d. administering a therapy to the patient if there is a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and
d. administering a therapy to the patient if the level of expression of the plurality of pre-selected genes in the sample is i) lower as compared with the predetermined reference upper level of expression of such genes and ii) higher as compared with the predetermined reference lower level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a RNA-containing sample of melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a level of expression of the plurality of pre-selected genes in the sample is i) lower as compared with the predetermined reference upper level of expression of such genes and ii) higher as compared with the predetermined reference lower level of expression of such genes, indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
5 b. KLRK1;
c. ITK; and
d. HLAE.

The present invention also provides a method of determining
10 whether a therapy is effective for treating patients afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from at least one patient afflicted with
15 melanoma, which at least one patient was administered the therapy for treatment of the melanoma previous to collection of the sample;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of
20 preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to the levels of expression of the plurality of pre-selected genes to the expression level of each such gene in a corresponding at
25 least one patient not administered the therapy;

wherein a higher level of expression of the plurality of pre-selected genes in the sample of step a) as compared with the level of expression of such genes in a corresponding at least
30 one patient not administered the treatment indicates that the therapy is effective for treating patients afflicted with melanoma, and

wherein the plurality of pre-selected genes comprises the
35 following genes:

a. CD2;

b. KLRK1;
c. ITK; and
d. HLAE.

5 The present invention also provides a method of determining whether a patient afflicted with melanoma and which patient was administered a therapy has exhibited a positive clinical response to the therapy which comprises the following:

10 a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient, which RNA-containing sample was removed from the patient who was administered the therapy;
b. treating the sample to determine from the RNA contained
15 in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

20

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has exhibited a positive clinical
25 response to the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

30 a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

35 The present invention also provides a method of determining whether a treatment should be administered to patients

afflicted with melanoma as an adjuvant or a neoadjuvant therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from at least one patient afflicted with melanoma;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the treatment should be administered to patients afflicted with melanoma as a neoadjuvant therapy, and a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the treatment should be administered to patients afflicted with melanoma as an adjuvant therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

In some embodiments, the plurality of pre-selected genes further comprises at least one of the following genes:
a. IFNAR1;
b. LCK;
c. CD4;

d. LGMN; and
e. IFI27.

In some embodiments, the plurality of pre-selected genes
5 comprises the following genes:
a. CD2;
b. KLRK1;
c. ITK;
d. HLAE;
10 e. IFNAR1;
f. LCK;
g. CD4;
h. LGMN; and
i. IFI27.
15

In some embodiments, the plurality of pre-selected genes
further comprises at least one of the following genes:
a. CCL27;
b. SYK;
20 c. CD68;
d. IL18; and
f. IL1F7.

In some embodiments, the plurality of pre-selected genes
25 comprises the following genes:
a. CD2;
b. KLRK1;
c. ITK;
d. HLAE;
30 e. LCK;
f. IFNAR1;
g. CD48;
h. CD4;
i. CTSS;
35 j. CCR4;
k. HLA-DQB1;
l. TAP2;

m. LGMN;
n. CSF2RA;
o. IFNGR1;
p. CCL27;
5 q. SYK;
r. CD68;
s. IL18;
t. IFI27; and
u. IL1F7.
10

In some embodiments, the plurality of pre-selected genes further comprises at least one of the following genes:
a. LCK;
b. IFNAR1;
15 c. CD48;
d. CD4;
e. CTSS;
f. CCR4;
g. HLA-DQB1;
20 h. TAP2;
i. LGMN;
j. CSF2RA; and
k. IFNGR1.

25 In some embodiments, the plurality of pre-selected genes further comprises at least two, three, four, five, six, seven, eight, nine, ten, or eleven of the following genes:
a. LCK;
b. IFNAR1;
30 c. CD48;
d. CD4;
e. CTSS;
f. CCR4;
g. HLA-DQB1;
35 h. TAP2;
i. LGMN;
j. CSF2RA; and

k. IFNGR1.

In some embodiments, the plurality of pre-selected genes comprises the following genes:

5 a. CD2;
b. KLRK1;
c. ITK;
d. HLAE;
e. LCK;
10 f. IFNAR1;
g. CD48;
h. CD4;
i. CTSS;
j. CCR4;
15 k. HLA-DQB1;
l. TAP2;
m. LGMN;
n. CSF2RA; and
o. IFNGR1.

20

In some embodiments, the plurality of pre-selected genes further comprises at least one of the following genes:

a. CCL27;
b. SYK;
25 c. CD68;
d. IL18;
e. IFI27; and
f. IL1F7.

30 In some embodiments, the plurality of pre-selected genes further comprises at least two, three, four or five of the following genes:

a. CCL27;
b. SYK;
35 c. CD68;
d. IL18;
e. IFI27; and

f. IL1F7.

In some embodiments, the plurality of pre-selected genes comprises the following genes:

5 a. CD2;
b. KLRK1;
c. ITK;
d. HLAE;
e. LCK;
10 f. IFNAR1;
g. CD48;
h. CD4;
i. CTSS;
j. CCR4;
15 k. HLA-DQB1;
l. TAP2;
m. LGMN;
n. CSF2RA;
o. IFNGR1;
20 p. CCL27;
q. SYK;
r. CD68;
s. IL18;
t. IFI27; and
25 u. IL1F7.

In some embodiments, the plurality of pre-selected genes consists of less than about 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 1000, or 10,000 genes.

30

In some embodiments, the expression level is assayed by NanoString gene expression analysis.

In some embodiments, the RNA transcripts of the plurality of pre-selected genes in the sample are fragmented.
35

The present invention provides a method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

5 a. obtaining a sample of the previously removed melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
10 c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

15 wherein a higher level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

20

In some embodiments, the method further comprises the step of creating a report summarizing said prediction.

In some embodiments, the expression level of the expression
25 product of the gene or the each of two or more genes is normalized relative to the expression level of the expression product of one or more of the following genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

30

In some embodiments, the predetermined reference level of expression of the expression product is the expression level of the expression product the one or more reference genes.

35 The present invention provides a method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of

reoccurrence of melanoma by the method of the invention, comprising administering an immunotherapy to the patient.

The present invention provides a method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by the method of the invention, comprising testing the human patient for recurrence of melanoma more frequently than a corresponding patient who was determined to have a reduced risk of reoccurrence of melanoma would be tested for recurrence.

The present invention provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample;
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene; and
d. administering a therapy to the patient if there is a higher level of expression of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene.

The present invention provides a method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample;
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined

reference level of the expression product for each such gene; and

d. administering a therapy to the patient if there is a lower level of expression of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene.

The present invention provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or the each of two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

The present invention provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;

b. treating the sample to determine the level of an expression product of the gene or each of the two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a lower level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

The present invention provides a method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a level of of the expression product of the gene or each of the two or more genes that is i) lower as compared with a predetermined reference upper level of expression for each such gene and ii) higher as compared with a predetermined reference lower level of expression for each such gene, indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

The present invention provides a method of determining whether a therapy is effective for treating patients afflicted with melanoma which comprises the following:

5 a. obtaining a sample of melanoma tissue from at least one patient afflicted with melanoma, which at least one patient was administered the therapy;
b. treating the sample to determine the level of an expression product of the gene or each of two or more
10 genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to the level of the expression product of the gene or each of the two or more genes in a corresponding at least one patient not
15 administered the treatment,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample of step (a) as compared with the level of expression product of the gene or
20 each of the two or more genes in the corresponding at least one patient not administered the treatment indicates that the therapy is effective for treating patients afflicted with melanoma.

25 The present invention provides a method of determining whether a patient afflicted with melanoma and which patient was administered a therapy has exhibited a positive clinical response to the therapy which comprises the following:

30 a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene
35 or each of the two or more genes to a predetermined reference level for each such gene,

wherein a higher level of expression the expression product of the gene or each of the two or more genes as compared to the predetermined reference level of the expression product of each such gene indicates that the patient has exhibited a positive clinical response to the therapy.

The present invention provides a method method of determining whether a treatment should be administered to patients afflicted with melanoma as an adjuvant or a neoadjuvant therapy which comprises the following:

a. obtaining a sample of melanoma tissue from at least one patient afflicted with melanoma;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level for each such gene,

wherein a higher level of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of expression for each such gene indicates that the treatment should be administered to patients afflicted with melanoma as a neoadjuvant therapy, and a lower level of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the treatment should be administered to patients afflicted with melanoma as an adjuvant therapy.

The present invention provides a method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma tissue from the patient;
b. treating the sample to determine the level of the expression product of the CD2 gene in the melanoma tissue sample; and
c. comparing the level of the expression product of the CD2 gene to a predetermined reference level of the expression product of the CD2 gene,

wherein a higher level of expression product of the CD2 gene in the sample as compared with the predetermined reference level of the expression product of the CD2 gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

The present invention provides a method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma tissue from the patient;
b. treating the sample to determine the level of the expression product of the X gene in the melanoma tissue sample; and
c. comparing the level of the expression product of the X gene to a predetermined reference level of the expression product of the X gene,

wherein a higher level of expression product of the X gene in the sample as compared with the predetermined reference level of the expression product of the X gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

In some embodiments, the sample was previously removed from the patient.

In some embodiments, the sample is a fixed, wax-embedded tissue specimen.

In some embodiments, the sample is at least one week old.

In some embodiments, the sample is at least one month old.

In some embodiments, the sample is at least six months old.

In some embodiments, the sample is at least one year old.

In some embodiments, the sample is at least ten years old.

In some embodiments, a method of the invention further comprises identifying a treatment option for the patient based on the expression level of the plurality of pre-selected genes.

In some embodiments, the expression level is determined by immunohistochemistry or proteomics technology.

In some embodiments, the therapy is chemotherapy.

In some embodiments, the chemotherapy comprises dacarbazine, temozolomide, paclitaxel, cisplatin, carmustine, fotemustine, vindesine, vincristine, and bleomycin, or vemurafenib.

In some embodiments, the therapy is radiation therapy.

In some embodiments, the therapy is immunotherapy.

In some embodiments, the immunotherapy comprises an interferon (IFN).

In some embodiments, the immunotherapy comprises IFN-α.

In some embodiments, the IFN-α is IFN-α2b.

In some embodiments, the IFN-α2b is PEGylated IFN-α2b.

In some embodiments, the PEGylated IFN-α2b is Sylatron.

In some embodiments, the immunotherapy comprises an interleukin.

5

In some embodiments, the interleukin is IL-2.

In some embodiments, the IL-2 is aldesleukin.

10 In some embodiments, the immunotherapy comprises an antibody.

In some embodiments, the antibody is a monoclonal antibody.

In some embodiments, the monoclonal antibody is a humanized
15 monoclonal antibody.

In some embodiments, the monoclonal antibody is a fully human monoclonal antibody.

20 In some embodiments, the antibody is an anti-PD-1 antibody.

In some embodiments, the anti-PD-1 antibody is BMS-936558.

In some embodiments, the antibody is an anti-CTLA-4 antibody.

25

In some embodiments, the antibody is ipilimumab.

In some embodiments, the immunotherapy is an oncolytic immunotherapy.

30

In some embodiments, the oncolytic immunotherapy comprises a virus.

In some embodiments, the virus is derived from HSV-1.

35

In some embodiments, the oncolytic immunotherapy is a vaccine.

In some embodiments, the vaccine is talimogene laherparepvec (T-VEC).

In some embodiments, the level of an expression product is
5 determined for one gene.

In some embodiments, the one gene is CD2.

In some embodiments, each of the two or more genes comprises
10 two, three, four, five, six, seven, nine, ten or more of CD2, KLRK1, ITK, HLAE, LCK, IFNAR1, CD48, CD4, CTSS, CCR4, HLA-DQB1, TAP2, LGMN, CSF2RA, IFNGR1, CCL27, SYK, CD68, IL18, IFI27 or IL1F7 gene.

15 In some embodiments, each of the two or more genes consists of less than about 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 1000, or 10,000 genes.

In some embodiments, the expression product is a protein
20 encoded by the gene or is the expression product of each of the two or more genes.

In some embodiments, the expression product of each gene is a protein encoded by the gene.
25

In some embodiments, the X gene is the CD2, KLRK1, ITK, HLAE, LCK, IFNAR1, CD48, CD4, CTSS, CCR4, HLA-DQB1, TAP2, LGMN, CSF2RA, IFNGR1, CCL27, SYK, CD68, IL18, IFI27 or IL1F7 gene.

30 In some embodiments, the sample is treated to determine the level of the expression product of each of two, three, four, five, six, seven, eight nine, ten or more genes in the melanoma tissue sample.

35 In some embodiments, the two, three, four, five, six, seven, eight nine, ten or more genes are two, three, four, five, six, seven, eight nine, ten or more of CD2, KLRK1, ITK, HLAE, LCK,

IFNAR1, CD48, CD4, CTSS, CCR4, HLA-DQB1, TAP2, LGMN, CSF2RA, IFNGR1, CCL27, SYK, CD68, IL18, IFI27 or IL1F7 gene.

In some embodiments, the patient is afflicted with sentinel lymph node positive melanoma when the melanoma tissue was removed.

In some embodiments, the patient was afflicted a melanoma deeper than 2mm or deeper than 1mm and ulcerated when the melanoma tissue was removed.

In some embodiments, the melanoma is sentinel lymph node positive melanoma.

In some embodiments, the melanoma is deeper than 2mm or deeper than 1mm and ulcerated.

In some embodiments, the method further comprises the step of creating a report summarizing said prediction.

In some embodiments, the expression product is a protein encoded by the X gene.

In some embodiments, the level of the expression product of each gene is determined by immunohistochemistry or proteomics technology.

In some embodiments, the level of the expression product of each gene is determined by immunohistochemistry.

In some embodiments, the melanoma tissue is stage II or III primary melanoma tissue.

In some embodiments, the melanoma is stage II or III melanoma.

In some embodiments, the expression level of each gene of the plurality of pre-selected genes is normalized relative to the expression level of one or more reference genes.

5 In some embodiments, the expression level of each such gene is normalized relative to the expression level of the following genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

10 In some embodiments, the predetermined reference level of expression is the expression level of one or more reference genes.

In some embodiments, the predetermined reference level of
15 expression is
i) the expression level of each such gene in normal tissue; or
ii) the expression level of one or more of the following reference genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH,
20 GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

In some embodiments, the predetermined reference level of the expression product is
25 i) the level of the expression product of each such gene in normal tissue; or
ii) the level of the expression product of one or more of the following reference genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19,
30 RPLP0, SDHA, TBP and TUBB.

In some embodiments, the normal tissue is normal skin tissue.

In some embodiments of the present invention the patient is a
35 human patient.

Each embodiment disclosed herein is contemplated as being applicable to each of the other disclosed embodiments. Thus, all combinations of the various elements described herein are within the scope of the invention.

5

It is understood that where a parameter range is provided, all integers within that range. For example, “0.2-5 mg/kg/day” is a disclosure of 0.2 mg/kg/day, 0.3 mg/kg/day, 0.4 mg/kg/day, 0.5 mg/kg/day, 0.6 mg/kg/day etc. up to 5.0 mg/kg/day.

10

Terms

“About” in the context of a numerical value or range means ±10% of the numerical value or range recited or claimed, unless the context requires a more limited range.

15

The term “melanoma” is used in the broadest sense and refers to all stages and all forms of cancer arising from melanocytes. Melanoma is typically a malignant tumor associated with skin cancer.

20

The term "prediction" is used herein to refer to the likelihood that a patient will have a particular clinical outcome, whether positive or negative. For instance, the term “prediction” may refer to the likelihood that a patient will respond either
25 favorably or unfavorably to a drug or set of drugs, and also the extent of those responses, or that a patient will survive, following surgical removal of the primary tumor and/or therapy for a certain period of time without cancer recurrence. The predictive methods of the present invention can be used
30 clinically to make treatment decisions by choosing the most appropriate treatment modalities for any particular patient. The predictive methods of the present invention are valuable tools in predicting if a patient is likely to respond favorably to a treatment regimen, such as surgical intervention, therapy with a
35 given drug or drug combination, and/or radiation therapy, or whether long-term survival of the patient, following surgery and/or termination of therapy is likely. The predictive methods

of the present invention can be used clinically to make treatment decisions by choosing the most appropriate treatment modalities for any particular patient. The predictive methods of the present invention are valuable tools in predicting if a 5 patient is likely to respond favorably to a treatment regimen, such as a chemotherapy, an immunotherapy, or radiation.

The term "positive clinical response" means an improvement in any measure of patient status, including but not limited to 10 those measures ordinarily used in the art, such as an increase in the duration of Recurrence-Free Interval (RFI), an increase in the time of Overall Survival (OS), an increase in the time of Disease-Free Survival (DFS), an increase in the duration of Distant Recurrence-Free Interval (DRFI), and the like. In a non-15 limiting example, an increase in the likelihood of positive clinical response corresponds to a decrease in the likelihood of cancer recurrence.

The term "Recurrence-Free Interval (RFI)" is used herein to 20 refer to time to first melanoma cancer recurrence censoring for second primary cancer as a first event or death without evidence of recurrence. The time may be in months or years. For instance, in some embodiments, the time may be at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or more, or at least 1, 2, 3, 4, 5, 6, 7, 8, 25 9, 10 or more years.

The term "Overall Survival (OS)" is used herein to refer to time from surgery to death from any cause. For instance, in some embodiments, the time may be at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 30 10, 11, 12 or more, or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more years.

The term "Disease-Free Survival (DFS)" is used herein to refer to time to melanoma recurrence or death from any cause. For 35 instance, in some embodiments, the time may be at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or more, or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more years.

The term "Distant Recurrence-Free Interval (DRFI)" is used herein to refer to the time from surgery to the first anatomically distant cancer recurrence. For instance, in some embodiments, the time may be at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or more, or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more years.

Normalization of Data

Aspects of the present invention relate to the use of the measured expression of certain genes by melanoma tissue to provide predictive information. In some embodiments, it is necessary to correct for (normalize away) differences in the amount of RNA assayed and/or variability in the quality of the RNA used. Therefore, assays and methods of the invention may measure and incorporate the expression of certain normalizing genes, including well known housekeeping genes. Non-limiting examples of normalizing genes include ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP, and/or TUBB. In some emodiments, a combination of two or more normalizing genes may be used. In some embodiments, normalization can be based on the mean or median signal (Ct) of all of the assayed genes or a large subset thereof (global normalization approach).

In some embodiments, sample-specific normalization factors are used to normalize raw mRNA counts in order to account for slight differences in assay efficiency such as hybridization, purification, and binding. In some embodiments, normalization for sample RNA quantity and quality differences are applied to spike-normalized values using sample-specific normalization factors calculated from the geometric mean of the counts from reporters targeting the reference genes, including but not limited to any one of or all of the following reference genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP, and TUBB. The resulting normalized counts may be used in downstream analyses.

General Description and Non-Limiting Examples of mRNA Isolation, Purification and Amplification

The steps of a representative protocol for profiling gene expression using fixed, paraffin-embedded tissues as the RNA source, including mRNA isolation, purification, primer extension and amplification are provided in various published journal articles (for example: T. E. Godfrey et al., J. Malec. Diagnostics 2: 84-91 (2000); K. Specht et al., Am. J. Pathol. 158: 419-29 (2001)). Briefly, a representative process starts with cutting about 10 or 20µm thick sections of paraffin-embedded tumor tissue samples. The RNA is then extracted, and protein and DNA are removed. Another representative method for extracting RNA from a sample, such as from a fixed, paraffin-embedded tissue sample includes the Ambion® RecoverAll Total Nucleic Acid Isolation Kit (Life Technologies, Carlsbad, CA).

Gene Expression Profiling

NanoString Gene Expression Analysis

In embodiments of the invention, gene expression may be determined using melanoma samples by NanoString gene expression analysis. As a clinical standard in melanoma all of the tumor has to be formalin fixed to preserve it for morphology assessment by the pathologist. The terms "NanoString gene expression analysis" and "NanoString" are used interchangeably herein, and refer to the nCounter® Analysis System of NanoString Technology (Seattle, Washington, USA). NanoString does not require amplification of RNA, has low sample requirements and is effective for evaluating the level of gene expression in FFPE samples, such as melanoma FFPE samples. Furthermore, NanoString is a multiplexed method for detecting gene expression and provides a method for direct measurement of mRNAs without the use of transcription or amplification. The RNA extracted from formalin fixed melanoma specimens may be of very poor quality and until recently no such analysis was possible. NanoString, however, allows for analysis of these specimens. With a sensitivity of 500

attomolar NanoString can detect as little as one copy of RNA per cell using 100 nanograms of total RNA as input.

NanoString and aspects thereof are described in Geiss et al., "Direct multiplexed measurement of gene expression with color-coded probe pairs" Nature Biotechnology 26, 317 - 325 (2008); in U.S. Patent Nos. 7,473,767, 7,941,279 and 7,919,237, and in U.S. Patent Application Publication No. 2010/0112710, the entire contents of each of which are hereby incorporated by reference. NanoString is also discussed in: Payton et al., "High throughput digital quantification of mRNA abundance in primary human acute myeloid leukemia samples" The Journal of Clinical Investigation 119(6): 1714–1726 (2009); and Vladislav et al. "Multiplexed measurements of gene signatures in different analytes using the NanoString nCounter Assay System" BMC Research Notes 2: 80 (2009), the entire contents of each of which are hereby incorporated by reference.

Immunohistochemistry

Immunohistochemistry methods are also suitable for detecting the expression levels of the prognostic markers of the present invention. Thus, antibodies or antisera, preferably polyclonal antisera, and most preferably monoclonal antibodies specific for each marker are used to detect expression. The antibodies can be detected by direct labeling of the antibodies themselves, for example, with radioactive labels, fluorescent labels, hapten labels such as, biotin, or an enzyme such as horse radish peroxidase or alkaline phosphatase.

Alternatively, unlabeled primary antibody is used in conjunction with a labeled secondary antibody, comprising antisera, polyclonal antisera or a monoclonal antibody specific for the primary antibody. Immunohistochemistry protocols and kits are well known in the art and are commercially available.

Methods of gene expression profiling include methods based on hybridization analysis of polynucleotides, methods based on

sequencing of polynucleotides, and proteomics based methods. The commonly used methods known in the art for the quantification of mRNA expression in a sample include NanoString (Geiss et al., Nature Biotechnology 26, 317 - 325 (2008)), northern blotting and in situ hybridization (Parker & Barnes, Methods in Molecular Biology 106:247-283 (1999)); RNAse protection assays (Hod, Biotechniques 13:852-854 (1992)); and PCR-based methods, such as reverse transcription polymerase chain reaction (RT-PCR) (Weis et al., Trends in Genetics 8:263-264 (1992)). Alternatively, antibodies may be employed that can recognize sequence-specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. Representative methods for sequencing-based gene expression analysis include Serial Analysis of Gene Expression (SAGE), and gene expression analysis by massively parallel signature sequencing (MPSS). Methods of Gene Expression Profiling, including SAGE; MPSS; proteomics based methods; RT-PCT and other PCR based methods; microarray analysis; and Promoter Methylation Analysis are discussed in U.S. Patent No. 8,067,178 and 8,034,565. The entire contents of each of which are hereby incorporated herein in their entireties.

Antibodies

As used herein, "monoclonal antibody" means an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations that typically include different antibodies directed against different determinants, each monoclonal antibody is directed against a single determinant on the antigen. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance

with the present invention may be made by the hybridoma method first described by Kohler and Milstein, *Nature* 256:495-97 (1975), or may be made by recombinant DNA methods (see, e.g., U.S. Pat. No. 4,816,567). The monoclonal antibodies may also be isolated from phage display libraries using the techniques described, for example, in Clackson et al., *Nature* 352:624-28 (1991) and Marks et al., *J. Mol. Biol.* 222(3):581-97 (1991).

The term "hybridoma" or "hybridoma cell line" refers to a cell line derived by cell fusion, or somatic cell hybridization, between a normal lymphocyte and an immortalized lymphocyte tumor line. In particular, B cell hybridomas are created by fusion of normal B cells of defined antigen specificity with a myeloma cell line, to yield immortal cell lines that produce monoclonal antibodies. In general, techniques for producing human B cell hybridomas, are well known in the art [Kozbor et al., *Immunol. Today* 4:72 (1983); Cole et al., in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. 77-96 (1985)].

The term "epitope" refers to a portion of a molecule (the antigen) that is capable of being bound by a binding agent, e.g., an antibody, at one or more of the binding agent's antigen binding regions. Epitopes usually consist of specific three-dimensional structural characteristics, as well as specific charge characteristics.

As used herein, "fully human antibody" is an antibody that is completely human. Fully human antibodies may be generated by, e.g., phage display, or in animals (such as mice) which have been genetically engineered to produce human antibodies. Exemplary methods of producing fully human antibodies are described in U.S. Patent Nos. 7,414,170; 7,803,981; in U.S. Patent Application No. 2008/0248531, and in McCafferty et al., "Phage antibodies: filamentous phage displaying antibody variable domains" Nature (1990) 348 (6301): 552–554; Osbourn JK, "Proximity-guided (ProxiMol) antibody selection" Methods Mol. Biol. (2002) 178: 201–5; and Lonberg et al., "Human antibodies

from transgenic mice" Int. Rev. Immunol. (1995) 13(1):65-93, the contents of each of which are hereby incorporated by reference in their entireties.

5 "Humanized antibodies" means antibodies that contain minimal sequence derived from non-human immunoglobulin sequences. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hyper variable region of the recipient are replaced by residues from a 10 hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit or nonhuman primate having the desired specificity, affinity, and capacity. See, for example, U.S. Pat. Nos. 5,225,539; 5,585,089; 5,693,761; 5,693,762; 5,859,205, each herein incorporated by reference. In some 15 instances, framework residues of the human immunoglobulin are replaced by corresponding non-human residues (see, for example, U.S. Pat. Nos. 5,585,089; 5,693,761; 5,693,762, each herein incorporated by reference). Furthermore, humanized antibodies may comprise residues that are not found in the recipient 20 antibody or in the donor antibody. These modifications are made to further refine antibody performance (e.g., to obtain desired affinity). In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable 25 regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. 30 For further details see Jones et al., Nature 331:522-25 (1986); Riechmann et al., *Nature* 332:323-27 (1988); and Presta, *Curro Opin. Struct. Biol.* 2:593-96 (1992), each of which is incorporated herein by reference.

35 Antibodies of the invention also include antibodies produced in a non-human mammalian host, more particularly a transgenic mouse, characterized by inactivated endogenous immunoglobulin (Ig) loci.

In such transgenic animals, competent endogenous genes for the expression of light and heavy subunits of host immunoglobulins are rendered non-functional and substituted with the analogous human immunoglobulin loci. These transgenic animals produce human antibodies in the substantial absence of light or heavy host immunoglobulin subunits. See, for example, U.S. Pat. No. 5,939,598, the entire contents of which are incorporated herein by reference.

Those skilled in the art will be aware of how to produce antibody molecules of the present invention. For example, polyclonal antisera or monoclonal antibodies can be made using standard methods. A mammal, (e.g., a mouse, hamster, or rabbit) can be immunized with an immunogenic form of the protein which elicits an antibody response in the mammal. For instance, a mammal can be immunized with irradiated cells that were transfected with a nucleic acid encoding the protein such that high levels of the protein were expressed on the cell surface. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassay can be used with the immunogen as antigen to assess the levels of antibodies. Following immunization, antisera can be obtained, and, if desired IgG molecules corresponding to the polyclonal antibodies may be isolated from the sera.

To produce monoclonal antibodies, antibody producing cells (lymphocytes) can be harvested from an immunized animal and fused with myeloma cells by standard somatic cell fusion procedures thus immortalizing these cells and yielding hybridoma cells. Such techniques are well known in the art. Hybridoma cells can be screened immunochemically for production of antibodies which are specifically reactive with the oligopeptide, and monoclonal antibodies isolated.

Immunotherapy

As used herein, "immunotherapy" is a treatment that induces or enhances the immune system to reject cancer in a patient. In

some embodiments of the invention an immunotherapy comprises at least one cytokine. In some embodiments, an immunotherapy comprises a vaccine that causes a patient's immune system to reject a cancer. In some embodiments, an immunotherapy comprises an antibody that binds to an antigen on the surface of cancer cells. Non-limiting examples of immunotherapies include high-dose interleukin-2 (IL-2); PEGylated IL-2 and interferon-α2b (IFN-α2b); ipilimumab; BMS-936558; and vaccines such as talimogene laherparepvec (T-VEC).

Recombinant interleukin-2 (IL-2) is sold under the trade name Proleukin, and is also known as aldesleukin. Aldesleukin is available from Prometheus, Inc. (San Diego, CA, USA)

PEGylated IFN-α2b is also known as Peginterferon alfa-2b and Sylatron, and is available from Merck (Whitehouse Station, NJ, USA).

Ipilimumab is a fully human anti-CTLA-4 antibody which is marketed as Yervoy. Yervoy is available from Bristol-Myers Squibb (New York, NY, USA).

BMS-936558 is a fully human anti-PD-1 andibody which is also known as MDX 1106. BMS-936558 is available from Bristol-Myers Squibb (New York, NY, USA).

Talimogene laherparepvec (T-VEC) is an oncolytic immunotherapy derived from HSV-1, and is also known as Oncovex. T-VEC is available from Amgen Inc. (Thousand Oaks, CA, USA).

Chemotherapy

Chemotherapies of the invention include but are not limited to any agent which specifically kills or reduces the proliferation of cancer cells in a patient afflicted with cancer. Non-limiting examples of chemotherapeutic agents which may be used in aspects of the invention are dacarbazine (DTIC-Dome), temozolomide (Temodar, Temodal), paclitaxel (Taxol), cisplatin (Paraplatin),

carmustine (BCNU), fotemustine, vindesine (Eldisine, Fildesin), vincristine (Oncovin, Vincasar), and bleomycin (Blenoxane).

Another non-limiting example of a chemotherapeutic agent for the treatment of melanoma is vemurafenib, which is useful for the treatment of patients with a V600E BRAF mutation. The CAS number for vemurafenib is 1029872-54-5. Vemurafenib is also known as Zelboraf, PLX4032, RG7204, RO5185426, has the formula: $C_{23}H_{18}C_{1}F_{2}N_{3}O_{3}S$, and is available from Plexxikon, Inc. (Berkeley, CA, USA).

In some embodiments, chemotherapy may be combined with an immunotherapy and/or radiation. In some embodiments, vemurafenib is combined with an immunotherapy.

Administration

"Administering" the therapies described herein can be effected or performed using any of the various methods and delivery systems known to those skilled in the art. The administering can be, for example, intravenous, oral, intramuscular, intravascular, intra-arterial, intracoronary, intramyocardial, intraperitoneal, and subcutaneous. Other non-limiting examples include topical administration, or coating of a device to be placed within the subject. In embodiments, administration is effected by injection or via a catheter.

Injectable drug delivery systems may be employed in the methods described herein include solutions, suspensions, gels. Oral delivery systems include tablets and capsules. These can contain excipients such as binders (e.g., hydroxypropylmethylcellulose, polyvinyl pyrilodone, other cellulosic materials and starch), diluents (e.g., lactose and other sugars, starch, dicalcium phosphate and cellulosic materials), disintegrating agents (e.g., starch polymers and cellulosic materials) and lubricating agents (e.g., stearates and talc). Solutions, suspensions and powders for reconstitutable delivery systems include vehicles such as suspending agents (e.g., gums, zanthans, cellulosics and sugars),

humectants (e.g., sorbitol), solubilizers (e.g., ethanol, water, PEG and propylene glycol), surfactants (e.g., sodium lauryl sulfate, Spans, Tweens, and cetyl pyridine), preservatives and antioxidants (e.g., parabens, vitamins E and C, and ascorbic 5 acid), anti-caking agents, coating agents, and chelating agents (e.g., EDTA).

The administration of therapies and compounds described herein may be by way of compositions containing one of the antagonists 10 and a pharmaceutically acceptable carrier. As used herein, a "pharmaceutical acceptable carrier" is a pharmaceutically acceptable solvent, suspending agent or vehicle, for delivering an active compound to a mammal, including humans. The carrier may be liquid, aerosol, gel or solid and is selected with the 15 planned manner of administration in mind. In some embodiments, the pharmaceutical carrier is a sterile pharmaceutically acceptable solvent suitable for intravenous administration. In an embodiment, the pharmaceutical carrier is a pharmaceutically acceptable solid suitable for oral administration.

20

As used herein, the term "effective amount" refers to the quantity of a component that is sufficient to treat a subject without undue adverse side effects (such as toxicity, irritation, or allergic response) commensurate with a reasonable 25 benefit/risk ratio when used in the manner of this invention, i.e. a therapeutically effective amount. The specific effective amount will vary with such factors as the particular condition being treated, the physical condition of the patient, the type of mammal being treated, the duration of the treatment, the 30 nature of concurrent therapy (if any), and the specific formulations employed and the structure of the compounds or its derivatives.

By treating the patient there are multiple possible outcomes. 35 For instance, treating a subject may comprise substantially reducing, slowing, stopping, preventing or reversing the progression of a disease, particularly melanoma. Additionally,

treating a patient may comprise substantially reducing, slowing, stopping, preventing or reversing a symptom of a disease. In some embodiments, an outcome of treating the patient is substantially reducing, slowing, stopping, preventing, or reversing metastasis, wherein the patient has, or has been treated for, a solid tumor. In some embodiments, treating the patient comprises reducing the likelihood of metastasis in the patient. In some embodiments the patient is treated after melanoma tissue has been removed from the patient. In some embodiments, a therapy is used for prevention and treatment of melanoma metastasis or recurrence. In the most favorable case, reduction is equivalent to prevention.

The methods provided by the present invention may also be automated in whole or in part.

The following abbreviated gene names are used herein:
IFI27: interferon, alpha-inducible protein 27
HLA-DPB1: major histocompatibility complex, class II, DP beta 1
STAT1: signal transducer and activator of transcription 1, 91kDa
MRC1: mannose receptor, C type 1
B2M: beta-2-microglobulin
IL18: interleukin 18 (interferon-gamma-inducing factor)
IFNGR1: interferon gamma receptor 1
CXCL11: chemokine (C-X-C motif) ligand 11
TAP2: transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)
CXCL9: chemokine (C-X-C motif) ligand 9
CLEC2A: C-type lectin domain family 2, member A
HLA-DPA1: major histocompatibility complex, class II, DP alpha 1
XCL2: chemokine (C motif) ligand 2
CTSS: cathepsin S
CCL27: chemokine (C-C motif) ligand 27
ICOS: inducible T-cell co-stimulator
IRF8: interferon regulatory factor 8

IFITM1: interferon induced transmembrane protein 1
HLAE: major histocompatibility complex, class I, E
GATA3: GATA binding protein 3
TAP1: transporter 1, ATP-binding cassette, sub-family B
5 (MDR/TAP)
CD2: CD2 molecule
CD37: CD37 molecule
KLRK1: killer cell lectin-like receptor subfamily K, member 1
CD5: CD5 molecule
10 LY9: lymphocyte antigen 9
CXCR3: chemokine (C-X-C motif) receptor 3
CD3E: CD3e molecule, epsilon (CD3-TCR complex)
TNFSF13B: tumor necrosis factor (ligand) superfamily, member 13b
15 LCK: lymphocyte-specific protein tyrosine kinase
IL1F7: interleukin 37
XCR1: chemokine (C motif) receptor 1
C3: complement component 3
CD4: CD4 molecule
20 CD48: CD48 molecule
LGMN: legumain
TNFRSF18: tumor necrosis factor receptor superfamily, member 18
IRF9: interferon regulatory factor 9
SKAP1: src kinase associated phosphoprotein 1
25 TARP: TCR gamma alternate reading frame protein
GZMK: granzyme K (granzyme 3; tryptase II)
ITK: IL2-inducible T-cell kinase
CSF2RA: colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)
30 PGK1: phosphoglycerate kinase 1
HLA-DQB1: major histocompatibility complex, class II, DQ beta 1
CD40: CD40 molecule, TNF receptor superfamily member 5
CYBB: cytochrome b-245, beta polypeptide
CCL5: chemokine (C-C motif) ligand 5
35 PTPRC: protein tyrosine phosphatase, receptor type, C
ITGAL: integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide)

IRF2: interferon regulatory factor 2
CD68: CD68 molecule
TLR7: toll-like receptor 7
CD53: CD53 molecule
5 SDHA: succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
CD8A: CD8a molecule
POLR1B: polymerase (RNA) I polypeptide B, 128kDa
IKZF1: IKAROS family zinc finger 1 (Ikaros)
10 ITGB2: integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)
ACTB: actin, beta
CLTC: clathrin, heavy polypeptide (Hc)
CCR4: chemokine (C-C motif) receptor 4
15 IFNAR1: interferon (alpha, beta and omega) receptor 1
SYK: spleen tyrosine kinase
G6PD: glucose-6-phosphate dehydrogenase
IRF5: interferon regulatory factor 5
RPLP0: ribosomal protein, large, P0
20 LDHA: lactate dehydrogenase A
CCR5: chemokine (C-C motif) receptor 5
CD27: CD27 molecule
GAPDH: glyceraldehyde-3-phosphate dehydrogenase
TUBB: tubulin, beta class I
25 TBP: TATA box binding protein
RPL19: ribosomal protein L19
HPRT1: hypoxanthine phosphoribosyltransferase 1
ALAS1: aminolevulinate, delta-, synthase 1
POLR2A: polymerase (RNA) II (DNA directed) polypeptide A,
30 220kDa
GUSB: glucuronidase, beta
ABCF1: ATP-binding cassette, sub-family F (GCN20), member 1
LY64: Lymphocyte Antigen-64
LY64 is also known as: CD180 molecule (CD180).
35

All publications and other references mentioned herein are incorporated by reference in their entirety, as if each

individual publication or reference were specifically and individually indicated to be incorporated by reference. Publications and references cited herein are not admitted to be prior art.

5

This invention will be better understood by reference to the Experimental Details which follow, but those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention as described more fully
10 in the claims which follow thereafter.

15

20

25

30

35

Experimental Details

EXAMPLE 1.

METHODS

5 **Patients and Samples**

The training set included FFPE primary melanomas from 44 consecutive patients with tumors either ≥ 2mm or ≥1 mm with one of the following high-risk features: ulceration, satellite lesions, and/or a positive sentinel lymph node. Thirty-four
10 specimens were obtained from Geisinger Health System (Danville, PA) and 10 from Mount Sinai School of Medicine and affiliates (New York, NY). The validation set included 30 FFPE melanomas obtained from NYU Medical Center (New York, NY), 6 from Mount Sinai School of Medicine and affiliates, and 1 from Geisinger
15 Health System. Recurrence was defined as metastatic melanoma beyond the local lymph node basin (stage IV) or the development of unresectable stage III disease. Non-recurrence was defined as no further evidence of melanoma following excision of the primary lesion. Minimal follow up for non-
20 recurrent patients in the validation cohort was 2 years. All samples and clinical information were obtained following approval by local institutional review boards (IRBs).

Dermatopathology

25 Primary and recurrent melanomas were selected from the Tamtron PowerPath database at the Mount Sinai Medical Center (New York, NY), the Cerner CoPathPlus database at Geisinger Heath System (Danville, PA), the Oracle Clinical RDC 4i database at NYU Medical Center (New York, NY), as well as the NovoPath
30 database at Englewood Hospital and Medical Center (Englewood, NJ). Combined, the databases consisted of over 2,500,000 specimens collected since 1985 and derived from surgical pathology, dermatopathology, neuropathology and cytology. Criteria for inclusion were primary melanoma specimens
35 measuring ≥ 2mm without ulceration or ≥ 1mm with ulceration. Selected slides and paraffin blocks were reviewed by two of the study participants (RGP and SS). Each sample was evaluated

for histogenetic type, extent and type of inflammatory infiltrate, thickness, and ulceration.

Analysis of gene expression

RNA was extracted using the Ambion® RecoverAll Total Nucleic Acid Isolation Kit (Life Technologies, Carlsbad, CA). 446 genes were selected based on a PubMed literature review (Table 6). The nCounter platform (NanoString Technologies, Seattle, WA) was used to quantify relative gene expression (described below).[31]

NanoString

Gene Expression Analysis

Using the nCounter® platform (NanoString Technologies, Seattle, WA), relative gene expression was quantified in a multiplex reaction. A custom CodeSet, designated MtSinai0511, was synthesized by NanoString for the 446 selected genes as well as 17 housekeeping genes and 14 controls in a 477-plex reaction (listed in the supplemental reference file). Hybridizations were carried out according to the supplier protocols.[40] In a total reaction volume of 30µl, 100ng of each RNA sample in 5µl H_2O was mixed with 10µl nCounter Reporter probes, 10µl hybridization buffer (1x hybridization buffer = 5x SSPE, 0.1% Tween-20), and 5µl of nCounter Capture probes. Hybridizations were incubated at 65°C for approximately 16-20 hours. Following hybridization, the samples were processed in a PrepStation and counted in a DigitalAnalyzer (Nanostring Technologies) according to standard protocol recommended by NanoString Technologies.

Normalization of Data

Calculated from the sum counts of reporters of 6 positive control RNA spikes, sample-specific normalization factors were used to normalize raw mRNA counts in order to account for slight differences in assay efficiency such as hybridization, purification, and binding. Concentrations of the control RNA spikes range from 0.125-128fM. Normalization for sample RNA

quantity and quality differences were applied to the spike-normalized values using sample-specific normalization factors calculated from the geometric mean of the counts from reporters targeting the following reference genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP, and TUBB. The resulting normalized counts were used in downstream analyses.

RNA Extraction

FFPE tissue blocks were cut into four 20µm sections and treated with 100% xylene (Fisher Scientific, Pittsburgh, PA) to deparaffinize. Samples were washed twice with 100% ethanol (Absolute Ethanol Molecular Biology Grade 200 proof, Fisher Scientific, Pittsburgh, PA) and dried via vacuum centrifugation at 40°C. Tissue was then incubated in Digestion Buffer and Protease (Ambion® RecoverAll, Life Technologies, Carlsbad, CA) at 50°C for 3 hours, followed by a 15-minute incubation at 80°C. RNA was separated using an Isolation Additive/Ethanol mixture (Ambion® RecoverAll, Life Technologies, Carlsbad, CA), and filtered by centrifugation at 10,000rpm. The sample was rinsed with Wash 1 and Wash 2 (Ambion® RecoverAll, Life Technologies, Carlsbad, CA), before and after incubation with DNase for 30 minutes at RT. RNA was eluted with 60µl of Elution Solution (Ambion® RecoverAll, Life Technologies, Carlsbad, CA) at RT.

Immunohistochemistry

IHC was performed on 5-µm charged slides using anti-CD2 monoclonal antibody (MRQ-11, Ventana Medical Systems, Tucson, AZ). Sections were deparaffinized and stained using a Ventana BenchMark XT immunostainer. Slides were evaluated by two of the study authors (SGB & MMM) in a blinded manner in 8 random High Powered Fields (HPFs) using an ocular micrometer with a 1 mm^2 grid (Nikon Eclipse E400®).

Five µm sections of the same paraffin-embedded tissue samples analyzed by NanoString were prepared for immunophenotypic

analysis. Immunohistochemistry (IHC) was performed using primary, pre-diluted anti-CD2 (MRQ-11, mouse anti-human, Ventana Medical Systems, Tucson, AZ). Sections were deparaffinized, stained according to standard protocol using a Ventana BenchMark XT immunostainer and manually counterstained.[40] The immunohistochemical slides were evaluated and interpreted by two of the study authors (SGB & MMM) in a blinded manner without knowledge of corresponding clinical data. For each sample, cells with circumferential membrane staining were counted and averaged in 8 random HPFs using an ocular micrometer with a 1 mm^2 grid (Nikon Eclipse E400®).

Statistics

Ensemble classification/regression method and ROC curves Classification was performed using two standard linear regression classifiers: random forest and elastic net. The 446 genes from the training experiment were ranked based on prediction of melanoma recurrence in the training cohort and sequentially reduced using a nested cross-validation procedure. Selected genes were then further reduced to eliminate redundant genes with α=0.2 and an optimal value of λ selected by an internal leave-one-out cross-validation yielding a final list of 21 genes. ROC curves were generated and the area under the curve (AUC) with class labels was defined for each sample to maximize prediction accuracy.

Survival and Demographic Analysis Kaplan-Meier analysis and Log-Rank (Mantel Cox) tests were performed. Mann-Whitney U tests generated p values for age and depth. Other non-continuous characteristics were analyzed using a two-tailed Fisher’s exact test. Graphpad Prism version 5.0 was used (San Diego California USA) and statistical significance was defined as $p<0.05$. Detailed description of the above methods is included below.

Cross-validation

900 iterations of an 11-fold cross-validation on the training dataset were simulated with random sample reordering in each iteration to strengthen the robustness of the final classifier model disclosed herein. Going from the top of the list to the bottom of the training cohort, every 4 samples were removed. These sample sets were then used as training data to fit a statistical model. 10,000 model training tasks were performed. The trained model and gene predictors selected were recorded in each task yielding 10,000 models and 10,000 lists of gene predictors based on randomly sub-sampled training samples. For each model, a classification was performed for the entire training (44 samples) and validation datasets (37 samples). To derive a robust list of gene hits by these models, the 10,000 gene lists were pooled and the statistical count of each gene (out of 446 genes) was selected by these models using the training cohort only. A higher count value for a given gene indicates that it is frequently selected as a predictor during the cross-validation process. Finally, all genes selected at least once in the 10,000 cross-validation were put into a final model training task to yield an optimal, compact predictor gene list of 24 genes. Three genes (IFNG, TNFSF18, and CREB1) were excluded from the signature because the p value did not meet the cutoff in the preliminary analysis of the training data and levels were therefore not tested in the validation set.

Ensemble classification/regression method

A two-step sequential ensemble classification scheme, that sequentially concatenated two widely applied classifiers: random forest and elastic net, was employed. Random forest itself is an ensemble classifier consisting of many decision trees that generates the mode of individual classes yielded by independent trees. A powerful variable selector and classification/regression method, elastic net integrates a linear regression model with Lasso and Ridge regularization. Elastic net is particularly useful when there are many more

predictors than samples, serving to further exclude genes that are only correlated with, but not most predictive, of recurrence. This two-step ensemble classification scheme was applied to the cross-validated training data for the outer 5 loop of the 900, 11-fold training cross-validation data points.

Random Forest

With 446 genes as an initial set of features and 40 samples from the cross-validation used as training data, a random 10 forest model was fit. Initially, random forest was run without feature selection to determine the importance of all 446 genes based on various metrics in the RandomForest R package.[41] Next, an independent run was started that incorporated feature selection into random forest by sequentially reducing a 15 certain number of predictors, ranked by variable importance, by employing a nested cross-validation procedure. In the simulation, a leave-one-out strategy was used. In each internal cross-validation, (step=30%) the least important genes/features, ranked by variable importance, was removed 20 from the last cross-validation iteration. Next, (Ntree=50k) bootstrap samples from the original data (40 samples) were drawn. For each of the bootstrap samples, an untrimmed classification/regression tree with randomly selected (mtry=22) genes was generated from the pool of genes leftover following 25 removal. Following cross-validation, the number of genes that resulted in the lowest error rate among all the cross-validation runs was selected. This number represents the number of genes (N_{RF}) selected by random forest after cross-validation. Next, the top N_{RF} genes based on the averaged gene 30 rank were selected from the initial run without feature selection, yielding the final gene selections by random forest. The selected genes G_{RF} were used as input for an elastic net model in order to identify the constituents of a gene signature predictive of melanoma recurrence.

35

Elastic Net

A powerful variable selector and classification/regression method, elastic net integrates a linear regression model with Lasso and Ridge regularization.[42] Elastic net is particularly useful when there are many more predictors than samples, serving to further exclude genes that are only correlated with, but not most predictive of recurrence. In each round of 11-fold cross-validation on the training data, there were 40 training samples. The number of genes selected by random forest varied from 50 to 446 depending on both the leave-one-out training data and randomized feature selection used during decision tree growth. Elastic net extends the basic form of linear/logistic regression via L1 and L2-regularization. λ controls the model complexity with higher values resulting in a less complex model (less number of genes). α controls the balance between two types of model complexity penalties, including the ridge-regression penalty (α =0) and the lasso penalty (α=1). The Ridge penalty is particularly useful when there are more genes and fewer samples. Ridge regression is known to shrink the coefficients of correlated predictors towards each other. In contrast, lasso tends to pick one out of an entire set and ignore the rest. In this study, we set α=0.2 and used an internal leave-one-out cross-validation to select an optimal value of λ[4]. The input training data was a subset of the original training data based on the gene lists G_{RF} determined by random forest. The output gene lists by elastic net with non-zero coefficients is our final gene list G_{EN} for the cross-validation run.

Final Gene List Generation

Employing the two-step ensemble classification method outlined above, the gene list G_{EN} was recorded for each of the 11-fold data cross-validation runs. After more than 900 runs, 10,000 lists of final genes were selected from the cross-validation training data. The number of times each gene was selected among the 10,000 lists was counted and the p-value for the count distribution against otherwise random selection was

calculated. Since this combined gene list compressed 10k lists in the cross-validation based on different subsampled training data, it may contain correlated genes from different runs. Therefore, to filter these out and obtain our final gene signature, elastic net was used again with the same parameter configurations outlined previously (α=0.2, λ retrained based on the cross-validation of the 56-gene subset training data) for all 44 training samples. This yielded the gene signature composed of 21 genes described herein.

RESULTS

Patient demographics

The populations used are representative of patients with high-risk primary melanomas in the United States (Table 1). Patients in the training set had melanomas either ≥2mm thick or ≥1mm thick with ulceration, satellite lesions, and/or a positive sentinel node. With the exception of 4 patients in the training cohort who received adjuvant interferon alpha, patients were not treated unless they recurred.

No significant correlation between known prognostic factors and disease recurrence was observed in the training cohort, but a significant correlation with lesion depth was found in the validation cohort ($p=0.044$). This is consistent with prior observations that prognostic factors may not necessarily predict risk in smaller groups of melanoma patients. There were no statistically significant differences between demographics of the training and validation cohorts with the exception of immune infiltrate, which closely correlated with the institution where the pathology was interpreted (Table 2).

Immune gene expression is increased in non-recurrent patients.

To test the hypothesis that the immune system limits progression of early-stage melanoma, mRNA transcripts for 446 genes were measured using NanoString technology. RNA of sufficient quality for NanoString analysis was obtained in 44 out of 59 samples. Ninety-two of these 446 genes were

differentially ($p < 0.05$) expressed between recurrent and non-recurrent groups. Of these 92 genes, 90 were upregulated in the non-recurrent group (Table 4). A heat map clustered according to expression of these 92 genes (Figure 1A) demonstrates that non-recurrent samples cluster at higher extremes of expression while recurrent samples cluster at lower extremes. These findings establish that immune gene expression is predictive of non-recurrence in the training cohort.

Identification of an immune gene expression signature protective against melanoma recurrence

Next, a panel of genes for clinical application was defined. 53 genes selected at least once using a leave-4-out cross-validation on the training data set (Figure 2A) were refined by linear regression to select a 21-gene panel (Figure 2A, inset). This panel predicted recurrence with an area under the curve (AUC) of 0.983 in the training cohort (Figure 2B top). The best correlation with melanoma recurrence was found using the gene signature in combination with clinical predictors (R^2=0.794). Alone, these clinical predictors yielded a correlation of 0.318. Thus, expression of 21 immune-associated genes showed a strong correlation with recurrence in the training cohort.

Identification of CD2 as an immunohistochemical marker of favorable prognosis

CD2 was the most differentially expressed gene between the recurrent and non-recurrent groups in the training cohort ($p<0.001$). In order to validate NanoString findings at the protein level, tumors were stained for CD2 (Figure 3A). Low number of CD2 positive staining by IHC correlated with melanoma recurrence ($p=0.017$). These findings were concordant with NanoString results as determined by linear regression ($r=0.847$, $p<0.001$) (Figure 3B and 3C left).

Validation of an immune gene signature protective against disease recurrence

Next, the above findings were validated in an independent set of tissue samples. RNA of sufficient quality for NanoString analysis was obtained in 37 out of 51 samples. Based on preliminary analysis, the 63 immune-associated genes most differentially expressed by p value between recurrent and non-recurrent samples in the training cohort were measured in the validation cohort (Figure 1B). Of these genes, 41 were upregulated in the non-recurrent group (Table 5). A 21-gene signature was able to predict recurrence in the validation cohort with an AUC of 0.794 ($p=0.002$, Figure 2B). When combined with clinical factors, the gene signature correlated with recurrence with a coefficient of determination of 0.947 (Figure 3C). Cross-validation using a leave-4-out approach demonstrated that this signature was statistically robust (Table 7). Corresponding slides were then stained for CD2 in the validation cohort and expression of this protein was found to correlate with recurrence ($p=0.003$) (Figure 3B, right). These data show that immune gene expression profiling predicts recurrence in primary melanoma with an accuracy exceeding those of standard clinical predictors.

Correlation of immune gene expression signature with survival

To investigate correlation between the proposed 21-gene signature and survival, a Kaplan-Meier analysis was performed, yielding a precise prediction of overall survival ($p<0.001$, Figure 4A). By comparison, the American Joint Committee on Cancer (AJCC) standard of 4mm was less accurate in predicting survival ($p=0.084$, Figure 4B). The 21-gene signature was derived from the training cohort, and therefore may only reflect survival benefit in this group. To exclude this possibility, the 21-gene signature was applied to the validation cohort alone and found patients with a positive gene signature survived longer ($p=0.015$, Table 2).

Table 1. Clinical Characteristics of Patients with Primary Melanoma in the Training and Validation Cohorts

Characteristic	Training			Validation		
	Non-recurrent (N=23)	Recurrent (N=21)	P Value	Non-recurrent (N=17)	Recurrent (N=20)	P Value
Sex			1			0.193
Male -- no.(%)	16 (70)	15 (71)		8 (47)	14 (70)	
Female -- no. (%)	7 (30)	6 (29)		9 (53)	6 (30)	
Age			1			1
Median -- no.	66	69		68	73.5	
Clinical Stage of Disease			0.225			0.288
Stage I/II -- no. (%)	15 (65)	9 (43)		14 (82)	13 (65)	
Stage III -- no. (%)	8 (35)	12 (57)		3 (18)	7 (35)	
Location of Tumor			0.136			0.31
Trunk -- no. (%)	11 (48)	15 (71)		8 (47)	10 (50)	
Extremity -- no. (%)	12 (52)	6 (29)		9 (53)	10 (50)	
Pathological characteristics						
Depth (mm) -- avg ± CI	2.81 ± 0.62	4.72 ± 1.51	0.074	4.26 ± 2.54	6.94 ± 2.87	0.044
Ulceration			0.227			0.512
Absent -- no. (%)	14 (61)	8 (38)		7 (41)	6 (30)	
Present -- no. (%)	9 (39)	13 (62)		10 (59)	14 (70)	
Inflammatory Infiltrate			1			0.169
Absent -- no. (%)	17 (74)	15 (74)		3 (17)	8 (40)	
Present -- no. (%)	6 (26)	6 (26)		14 (82)	12 (60)	
Patient Outcome (Months)						
Time to recurrence -- avg ± CI	-	16.8 ± 5.5		-	28.6 ± 12.0	
Death from melanoma -- no. (%)	0 (0)	16 (76)		0 (0)	13 (65)	
Time to death -- avg ± CI	-	26.6 ± 9.8		-	48.6 ± 19.1	
Time to follow-up -- avg ± CI*	55.7 ± 13.5	60.4 ± 34.6		52.1 ± 8.0	79 ± 39.1	

*Data for living patients only

Table 2. Clinical Characteristics of Patients in the Training and Validation Cohorts.

Characteristic	Training (N=44)	Validation (N=37)	P Value
Sex			0.487
Male -- no.(%)	31 (70)	22 (60)	
Female -- no. (%)	14 (31)	15 (40)	
Age			0.3775
Median -- no.	67	69	
Clinical Stage of Disease			0.109
Stage I/II -- no. (%)	24 (55)	27 (73)	
Stage III -- no. (%)	20 (45)	10 (27)	
Location of Tumor			0.378
Trunk -- no. (%)	26 (60)	18 (49)	
Extremity -- no. (%)	18 (41)	19 (51)	
Pathological characteristics			
Depth (mm) -- avg ± CI	3.72 ± 0.83	5.71 ± 1.94	0.0975
Ulceration			0.26
Absent -- no. (%)	22 (50)	13 (35)	
Present -- no. (%)	22 (50)	24 (65)	
Inflammatory Infiltrate			0.0001
Absent -- no. (%)	32 (73)	11 (30)	
Present -- no. (%)	12 (27)	26 (70)	
Patient Outcome (Months)			
Time to recurrence -- avg ± CI	16.8 ± 5.45	28.6 ± 12.02	
Death from melanoma -- no. (%)	16 (36)	13 (35)	
Time to death -- avg ± CI	26.6 ± 9.77	48.6 ± 19.06	
Time to follow-up -- avg ± CI*	56.5 ± 12.7	60.3 ± 13.3	

5

10

Table 3. Cross Validation on Clinical Characteristics of Patients in the Training Cohort*

Characteristic -- no.[range]	Training		
	Non-recurrent (N=23)	Recurrent (N=21)	P Value
Sex			1
Male	15 [12-16]	14[11-15]	
Female	6 [3-7]	6 [3-6]	
Age			
Median	66 [57-72]	69 [67-72]	
Clinical Stage of Disease			0.203
Stage I/II	14 [11-15]	8 [5-9]	
Stage III	7 [4-8]	11 [8-12]	
Location of Tumor			0.208
Trunk	10 [7-11]	14 [11-15]	
Extremity	11 [8-12]	6 [2-6]	
Pathological characteristics			
Depth (mm)	2.25 [2.2-2.5]	3 [2.8-3.75]	0.035
Ulceration			
Absent	13 [10-14]	7 [4-8]	
Present	8 [5-9]	12 [9-13]	
Inflammatory Infiltrate			1
Absent	15 [13-17]	14 [11-15]	
Present	6 [2-6]	6 [3-6]	
Number of Samples			
Median	21 [19-23]	19 [17-21]	

*Using a leave-4-out approach

5

Table 4. Gene Function of Statistically Significant Genes in Training Set

5

Official Symbol	**Fold Change (Non-reccurent/ Recurrent)**	**P-value**	**Gene Function**
CD2	2.015246	0.000296	CD2 is a surface antigen of the human T-lymphocyte lineage that is expressed on all peripheral blood T cells. It is one of the earliest T-cell markers, being present on more than 95% of thymocytes; it is also found on some natural killer cells but not on B lymphocytes.
KLRK1	1.907905	0.000752	KLRK1 encodes for a member of the NKG2 family which is located within the NK complex, a region that contains several C-type lectin genes preferentially expressed in NK cells. It binds to a diverse family of ligands that can result in the activation of NK and T cells. This protein and its ligands are therapeutic targets for the treatment of immune diseases and cancers.
ITK	1.841831	0.000823	ITK encodes an intracellular tyrosine kinase expressed in T-cells. It is thought to play a role in T-cell proliferation and differentiation.
HLAE	1.545037	0.001147	HLA-E binds a restricted subset of peptides derived

			from the leader peptides of other class I molecules.
LCK	1.998149	0.001462	LCK is a member of the Src family of protein tyrosine kinases (PTKs). The encoded protein localizes to the plasma membrane where it binds to cell surface receptors, including CD4 and CD8, and is a key signaling molecule in the selection and maturation of developing T-cells.
CREB1	1.267037	0.001947	CREB1 encodes for a transcription factor that is phosphorylated by several protein kinases, and induces transcription of genes in response to hormonal stimulation of the cAMP pathway.
IFNAR1	1.248822	0.002154	The protein encoded by IFNAR1 forms one of the two chains of a receptor for interferons alpha and beta. Binding and activation of the receptor stimulates Janus protein kinases, which in turn phosphorylate several proteins, including STAT1 and STAT2. The encoded protein also functions as an antiviral factor.
CD48	1.774847	0.002713	CD48 encodes a member of the CD2 subfamily of immunoglobulin-like receptors which includes SLAM (signaling lymphocyte activation molecules) proteins. The

			encoded protein is found on the surface of lymphocytes and other immune cells, dendritic cells and endothelial cells, participating in activation and differentiation pathways in these cells.
CXCR3	1.794847	0.002995	This gene encodes a G protein-coupled receptor with selectivity for three chemokines, termed CXCL9/Mig, CXCL10/IP10, and CXCL11/I-TAC. Binding of chemokines to this protein induces cellular responses that are involved in leukocyte traffic, most notably integrin activation, cytoskeletal changes and chemotactic migration.
CD4	1.590889	0.003089	CD4 encodes a membrane glycoprotein of T lymphocytes that interacts with major histocompatibility complex class II antigens. CD4 is expressed in T lymphocytes, B cells, macrophages, and granulocytes. The protein functions to initiate or augment the early phase of T-cell activation.
IFNG	1.620499	0.003532	IFNG encodes a member of the type II interferon family. The protein encoded is a soluble cytokine with antiviral, immunoregulatory and anti-tumor properties and is a potent activator of macrophages.

CTSS	2.012265	0.003876	The protein encoded by CTSS is a member of the peptidase C1 family and a lysosomal cysteine proteinase that may participate in the degradation of antigenic proteins to peptides for presentation on MHC class II molecules.
CCR4	1.372039	0.004007	The protein encoded by CCR4 is a receptor for the CC chemokine - MIP-1, RANTES, TARC and MCP-1.
HLA-DQB1	2.09166	0.004592	HLA-DQB1 plays a central role in the immune system by presenting peptides derived from extracellular proteins. Class II molecules are expressed in antigen presenting cells such as B lymphocytes, dendritic cells, macrophages.
TAP2	1.444613	0.004638	The membrane-associated protein encoded by TAP2 is a member of the superfamily of ATP-binding cassette (ABC) transporters. The protein encoded by this gene is involved in antigen presentation.
CD37	1.681894	0.005394	The protein encoded by CD37 is a member of the tetraspanin family, most of which are cell-surface proteins that mediate signal transduction events that play a role in the regulation of cell development, activation, growth and motility. It may

			play a role in T-cell-B-cell interactions.
IRF2	1.247296	0.005432	IRF2 encodes interferon regulatory factor 2, a member of the interferon regulatory transcription factor (IRF) family. It competitively inhibits the IRF1-mediated transcriptional activation of interferons alpha and beta, and presumably other genes that employ IRF1 for transcription activation.
TNFSF18	1.604035	0.00557	The protein encoded by TNFSF18 is a cytokine that belongs to the tumor necrosis factor (TNF) ligand family. It has been shown to modulate T lymphocyte survival in peripheral tissues. This cytokine is also found to be expressed in endothelial cells, and is thought to be important for interaction between T lymphocytes and endothelial cells.
LGMN	1.652632	0.006518	LGMN encodes a cysteine protease that may be involved in the processing of bacterial peptides and endogenous proteins for MHC class II presentation in the lysosomal/endosomal systems. Overexpression of this gene may be associated with the majority of solid tumor types.
CCL5	1.943322	0.006747	CCL5 is one of several CC cytokine genes that functions

			as a chemoattractant for blood monocytes, memory T helper cells and eosinophils. It causes the release of histamine from basophils and activates eosinophils.
CSF2RA	1.523143	0.006915	The protein encoded by CSF2RA is the alpha subunit of the heterodimeric receptor for colony stimulating factor 2, a cytokine which controls the production, differentiation, and function of granulocytes and macrophages.
ITGAL	1.675467	0.006919	ITGAL encodes the integrin alpha L chain. This I-domain containing alpha integrin combines with the beta 2 chain (ITGB2) to form the integrin lymphocyte function-associated antigen-1 (LFA-1), which is expressed on all leukocytes. LFA-1 plays a central role in leukocyte intercellular adhesion.
BTK	1.389055	0.007196	BTK plays a crucial role in B-cell development. Mutations in this gene cause X-linked agammaglobulinemia type 1.
CD53	1.512738	0.007369	The protein encoded by CD53 is a member of the tetraspanin family, most of which serve to mediate signal transduction events that play a role in the regulation of cell development, activation, growth and motility. It contributes to the

			transduction of CD2-generated signals in T cells and natural killer cells and has been suggested to play a role in growth regulation.
IRF5	1.403363	0.007497	IRF5 encodes a member of the interferon regulatory factor (IRF) family, a group of transcription factors with diverse roles, including modulation of cell growth, differentiation, apoptosis, and immune system activity.
IL17RA	1.292772	0.00885	The protein encoded by IL17RA binds with low affinity to interleukin 17A which is a pro-inflammatory cytokine secreted by activated T-lymphocytes. It is a potent inducer of the maturation of CD34-positive hematopoietic precursors into neutrophils.
HLA-DPB1	1.725111	0.009152	HLA-DPB belongs to the HLA class II beta chain paralogues that plays a central role in the immune system by presenting peptides derived from extracellular proteins.
CCL27	2.459453	0.009421	The protein encoded by CCL27 is chemotactic for skin-associated memory T lymphocytes and may also play a role in mediating the homing of lymphocytes to cutaneous sites.
IFNGR1	1.288055	0.009535	IFNGR1 encodes the ligand-binding chain (alpha) of the gamma interferon receptor.

SYK	1.401605	0.01008	SYC encodes a member of the family of non-receptor type Tyr protein kinases that is widely expressed in hematopoietic cells and is involved in coupling activated immunoreceptors to downstream signaling events that mediate diverse cellular responses, including proliferation, differentiation, and phagocytosis. It is thought to be a modulator of epithelial cell growth and a potential tumor suppressor in human breast carcinomas.
CD180	1.435879	0.010121	CD180 is a cell surface molecule whose interactions serve to control B cell recognition and signaling of lipopolysaccharide (LPS).
CD68	1.605002	0.01036	CD68 encodes a transmembrane glycoprotein that is highly expressed by human monocytes and tissue macrophages. The protein is a member of the scavenger receptor family. that typically functions to clear cellular debris, promote phagocytosis, and mediate the recruitment and activation of macrophages.
B2M	1.66008	0.011068	B2M encodes a serum protein found in association with the major histocompatibility complex (MHC) class I heavy chain on the surface of nearly all nucleated cells.

IRF9	1.310939	0.011403	IRF9 encodes for interferon regulatory factor 9.
CD27	1.836914	0.011584	The protein encoded by CD27 is a member of the TNF-receptor superfamily and is required for generation and long-term maintenance of T cell immunity. It plays a key role in regulating B-cell activation and immunoglobulin synthesis.
KLRD1	1.614182	0.011947	KLRD1 is an antigen preferentially expressed on NK cells which are a distinct lineage of lymphocytes that mediate cytotoxic activity and secrete cytokines upon immune stimulation.
CD40	1.340179	0.012885	The protein encoded by CD40 is a member of the TNF-receptor superfamily and has been found to be essential in mediating a broad variety of immune and inflammatory responses including T cell-dependent immunoglobulin class switching, memory B cell development, and germinal center formation.
PTPRC	1.518673	0.012959	The protein encoded by PTPRC is a member of the protein tyrosine phosphatase (PTP) family which are signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic

			transformation.
NFATC3	1.284201	0.013039	The product of NFATC3 plays a role in the regulation of gene expression in T cells and immature thymocytes.
C3	2.335095	0.013304	Complement component C3 plays a central role in the activation of complement system. Its activation is required for both classical and alternative complement activation pathways.
CD8A	1.847047	0.013679	The CD8 antigen is a cell surface glycoprotein found on most cytotoxic T lymphocytes that mediates efficient cell-cell interactions within the immune system.
IKZF1	1.493992	0.013683	IKZF1 encodes a transcription factor associated with chromatin remodeling that functions as a regulator of lymphocyte differentiation. Overexpression of some dominant-negative isoforms have been associated with B-cell malignancies, such as acute lymphoblastic leukemia (ALL).
IL18	1.537873	0.014344	The protein encoded by IL18 is a proinflammatory cytokine that augments natural killer cell activity in spleen cells and stimulates interferon gamma production in T-helper type I cells.
HLA-DPA1	1.605975	0.014775	HLA-DPA1 belongs to the HLA class II alpha chain

			paralogues that plays a central role in the immune system by presenting peptides derived from extracellular proteins.
TAP1	1.420407	0.014927	The protein encoded by TAP1 is involved in the pumping of degraded cytosolic peptides across the endoplasmic reticulum into the membrane-bound compartment where class I molecules assemble.
ITGB2	1.557211	0.016195	The product of ITGB2 belongs to the integrin beta chain family of proteins which are known to participate in cell adhesion as well as cell-surface mediated signalling.
IFI27	1.679386	0.016642	Encodes for the protein interferon alpha-inducible protein 27.
STAT1	1.601101	0.018416	The protein encoded by STAT1 can be activated by various ligands such as interferon-alpha, interferon-gamma, EGF, PDGF, and IL6. It mediates the expression of a variety of genes and is thought to be important for cell viability in response to different cell stimuli and pathogens.
CD3E	1.558632	0.01849	The protein encoded by CD3E plays an important role in coupling antigen recognition to several intracellular signal-transduction pathways. The epsilon polypeptide plays an essential role in T-cell

			development.
TBX21	1.504635	0.018557	Expression of TBX21 has been shown to correlate with IFNG expression in Th1 and natural killer cells, suggesting a role for this gene in initiating Th1 lineage development from naive Th precursor cells.
CD5	1.478656	0.019625	CD5 is a cluster of differentiation found on a subset of IgM-secreting B cells known as B-1 cells as well as T cells.
PLCG2	1.346729	0.020213	The protein encoded by PLCG2 is a transmembrane signaling enzyme that plays an important role in the synthesis of IP3 and DAG which are important for transmitting signals from growth factor receptors and immune system receptors across the cell membrane.
TNFSF13B	1.563995	0.021075	The protein encoded by TNFSF13B belongs to the tumor necrosis factor (TNF) ligand family that has been shown to play an important role in the proliferation and differentiation of B cells.
LAMP1	1.21012	0.022709	The protein encoded by LAMP1 is a glycoprotein that provides selectins with carbohydrate ligands. It may also play a role in tumor cell metastasis.
IL37	1.753289	0.022796	The protein encoded by IL37 is a member of the interleukin 1

			cytokine family that can bind to, and may be a ligand for interleukin 18 receptor (IL18R1/IL-1Rrp) as well as interleukin 18 binding protein (IL18BP).
GATA3	1.495402	0.023877	GATA3 encodes a protein that is an important regulator of T-cell development and plays an important role in endothelial cell biology.
LTA	1.722909	0.02389	LTA encodes for a cytokine produced by lymphocytes that is highly inducible, secreted, and forms heterotrimers with lymphotoxin-beta which anchor lymphotoxin-alpha to the cell surface. This protein also mediates a large variety of inflammatory and immunostimulatory responses and plays a role in apoptosis.
CLEC2A	1.945567	0.026066	CLEC2A belongs to the CLEC2 family of activation-induced, natural killer gene complex-encoded C-type lectin-like receptors.
IKZF5	1.239951	0.026122	IKZF5 is expressed in lymphocytes and is implicated in the control of lymphoid development.
XCL2	1.459574	0.026267	XCL2 is a cytokine related to XCL1 that is predominantly expressed in activated T cells which induces chemotaxis of cells expressing the chemokine receptor XCR1.
ZAP70	1.631185	0.026582	ZAP70 encodes an enzyme that

			plays a role in T-cell development and lymphocyte activation. This enzyme functions in the initial step of TCR-mediated signal transduction in combination with the Src family kinases, Lck and Fyn.
PILRA	1.367106	0.026666	PILRA encodes for an ITIM-bearing member and serves an inhibitory function that is central to the regulation of several cell signaling pathways.
GZMK	2.111418	0.027669	GZMK is a member of a group of related serine proteases from the cytoplasmic granules of cytotoxic lymphocytes.
IRF8	1.543724	0.0279	The IRF family proteins bind to the IFN-stimulated response element (ISRE) and regulate expression of genes stimulated by type I IFNs, namely IFN-alpha and IFN-beta.
TNFRSF18	1.444177	0.02993	TNFRSF18 is thought to play a key role in dominant immunological self-tolerance maintained by CD25(+)CD4(+) regulatory T cells. Knockout studies in mice also suggest the role of this receptor is in the regulation of CD3-driven T-cell activation and programmed cell death.
CLECL1	1.461028	0.030971	CLECL1 encodes a C-type lectin-like protein that is highly expressed on dendritic and B cells. It may act as a

			T-cell costimulatory molecule that enhances interleukin-4 production, and maybe involved in the regulation of the immune response.
MST1R	1.522389	0.032201	MST1R encodes a cell surface receptor for macrophage-stimulating protein (MSP) with tyrosine kinase activity. It is expressed on the ciliated epithelia of the mucociliary transport apparatus of the lung, and together with MSP, thought to be involved in host defense.
TARP	1.446441	0.032526	In some non-lymphoid tissues, the unrearranged T cell receptor gamma (TRG@) locus is expressed. The resulting transcript contains a subset of the TRG@ gene segments and is shorter than TRG@ transcripts expressed in lymphoid tissues.
IFITM1	1.609253	0.032627	IFITM1 codes for an intrinsic membrane protein that is induced by interferon and is part of the interferon signaling pathway.
MFGE8	1.422697	0.033878	Mfge9 contains a phosphatidylserine (PS) binding domain that allows it to bind to PS on the surface of apoptotic cells. This helps facilitate opsonization of apoptotic cells.
CYBB	1.455351	0.034018	CYBB is the beta chain of Cytochrome b which has been

			proposed as a primary component of the microbicidal oxidase system of phagocytes.
NFKB1	1.250284	0.034122	NFKB1 encodes a protein that can undergo cotranslational processing that forms a DNA binding subunit of the NF-kappa-B (NFKB) protein complex. Inappropriate activation of NFKB has been associated with a number of inflammatory diseases while persistent inhibition of NFKB leads to inappropriate immune cell development or delayed cell growth.
LY9	1.422617	0.035493	LY9 belongs to the SLAM family of immunomodulatory receptors and interacts with the adaptor molecule SAP.
STAT2	1.284081	0.035692	The protein encoded by STAT2 is phosphorylated In response to cytokines and growth factors. In response to interferon (IFN), it forms a complex with STAT1 and IFN regulatory factor family protein p48 (ISGF3G).
XCR1	1.403745	0.037724	The protein encoded by XCR1 is a chemokine receptor most closely related to RBS11 and the MIP1-alpha/RANTES receptor. It transduces a signal by increasing the intracellular calcium ions level.
MRC1	1.45079	0.03779	The protein encoded by MRC1 mediates the endocytosis of

			glycoproteins by macrophages.
CCR5	1.479252	0.038168	CCR5 is expressed by T cells and macrophages. Its expression has also been detected in a promyeloblastic cell line, suggesting that this protein may play a role in granulocyte lineage proliferation and differentiation.
CXCL9	1.893082	0.038221	While the exact function of CXCL9 has not been specifically defined, it is thought to be involved in T cell trafficking.
SKAP1	1.776735	0.038853	SKAP1 encodes a protein that plays a critical role in inside-out signaling by coupling T-cell antigen receptor stimulation to the activation of integrins.
CD1C	1.43729	0.039292	CD1C encodes a protein that mediates the presentation of primarily lipid and glycolipid antigens of self or microbial origin to T cells.
SP110	1.257066	0.039624	SP110 encodes a leukocyte-specific nuclear body component that can function as an activator of gene transcription and may play a role in ribosome biogenesis as well as the induction of myeloid cell differentiation.
IFNGR2	1.170558	0.040199	IFNGR2 encodes the non-ligand-binding beta chain of the gamma interferon receptor which is a heterodimer of

			IFNGR1 and IFNGR2.
MAP3K7	1.213028	0.040203	The protein encoded by MAP3K7 mediates the signaling transduction induced by TGF beta and morphogenetic protein (BMP). It controls a variety of cell functions including transcription regulation and apoptosis.
CEBPA	1.545808	0.041975	The protein encoded by CEBPA can bind as a homodimer to certain promoters and enhancers. It can interact with CDK2 and CDK4, thereby inhibiting these kinases and causing growth arrest in cultured cells.
NLRC5	1.439237	0.044939	NLRC5 plays a role in cytokine response and antiviral immunity through its inhibition of NF-kappa-B activation and negative regulation of type I interferon signaling pathways.
CXCL11	1.584466	0.04521	CXCL11 encodes a protein that induces a chemotactic response in activated T-cells and is the dominant ligand for CXC receptor-3. IFN-gamma is a potent inducer of transcription of this gene.
ICOS	1.415572	0.045259	The protein encoded by ICOS belongs to the CD28 and CTLA-4 cell-surface receptor family and plays an important role in cell-cell signaling, immune responses, and regulation of cell proliferation.

CTLA4	1.644172	0.04539	CTLA4 encodes a protein that transmits an inhibitory signal to T cells.
BIRC5	0.67022	0.045418	BIRC5 is a member of the inhibitor of apoptosis (IAP) gene family which encodes negative regulatory proteins that prevent apoptotic cell death. Gene expression is high during fetal development and in most tumors, yet low in adult tissues.
TLR6	1.561918	0.04578	The protein encoded by TLR6 is a member of the Toll-like receptor (TLR) family and plays a fundamental role in pathogen recognition and activation of innate immunity.
IL10RA	1.408962	0.048235	The protein encoded by IL10RA has been shown to mediate the immunosuppressive signal of interleukin 10, thus serving to inhibit the synthesis of proinflammatory cytokines. This receptor is reported to promote survival of progenitor myeloid cells through the insulin receptor substrate-2/PI 3-kinase/AKT pathway.
CXCL6	0.559425	0.048912	CXCL6 encodes a protein that serves as a chemoattractant for neutrophilic granulocytes by interacting with the chemokine receptors CXCR1 and CXCR2.

Table 5. Statistically Significant Genes in Validation Set

Official Symbol	**Fold Change (Non-reccurent/Recurrent)**	**P-value**
IFI27	1.977386	0.000887
HLA-DPB1	1.649143	0.00097
STAT1	1.805251	0.000972
MRC1	1.563158	0.001282
B2M	1.584912	0.002027
IL18	1.864116	0.002164
IFNGR1	1.350452	0.002413
CXCL11	2.850672	0.002644
TAP2	1.489447	0.003244
CXCL9	3.114418	0.003475
CLEC2A	3.20643	0.003833
HLA-DPA1	1.603543	0.005165
XCL2	1.431471	0.005597
CTSS	1.59295	0.005937
CCL27	3.513776	0.006878
ICOS	1.399578	0.011073
IRF8	1.435517	0.011883
IFITM1	1.391794	0.012035
HLAE	1.356181	0.012934
GATA3	1.6531	0.014346
TAP1	1.707536	0.014598
CD2	1.487405	0.015139
CD37	1.347055	0.016238
KLRK1	1.612816	0.018172
CD5	1.300416	0.018254
LY9	1.282556	0.018897
CXCR3	1.334443	0.020895
CD3E	1.400876	0.02122
TNFSF13B	1.42329	0.021607
LCK	1.325899	0.021821
IL37	1.935279	0.027492

XCR1	1.29547	0.032381
C3	1.392606	0.035138
CD4	1.19348	0.037494
CD48	1.224441	0.038406
LGMN	1.186092	0.039788
TNFRSF18	1.323574	0.04165
IRF9	1.220289	0.043863
SKAP1	1.221685	0.044579
TARP	1.49682	0.045214
GZMK	1.412444	0.046383

5

10

15

20

25

30

35

Table 6. Official names of all 446 genes selected for NanoString analysis in training cohort with housekeeping genes (17) highlighted

A2M	CCR8	CSF1	HLAA	IKBKG	IRF1	MRC1	TIA1
ABCF1	CCR9	CSF1R	HLAB	IKZF1	IRF2	MSR1	TICAM1
ACTB	CCRL1	CSF2	HLAC	IKZF5	IRF3	MST1R	TICAM2
ALAS1	CCRL2	CSF2RA	HLA-DMA	IL10	IRF4	MX1	TIMP3
ALCAM	CD101	CSF2RB	HLA-DOB	IL10RA	IRF5	MYADM	TIRAP
ALOX5	CD14	CSF3	HLA-DPA1	IL11	IRF6	MYD88	TLR1
AMICA1	CD163	CSF3R	HLA-DPB1	IL12A	IRF7	NFAM1	TLR10
ANGPTL4	CD164	CTLA4	HLA-DQA1	IL12B	IRF8	NFATC3	TLR2
ANXA1	CD180	CTSS	HLA-DQA2	IL12RB1	IRF9	NFKB1	TLR3
ANXA11	CD19	CX3CL1	HLA-DQB1	IL12RB2	ISG15	NFKB2	TLR4
B2M	CD1A	CX3CR1	HLA-DRB1	IL13	ISG20	NFKBIA	TLR5
BCL10	CD1B	CXCL1	HLA-DRB3	IL13RA1	ITGA1	NFKBIZ	TLR6
BCL2A1	CD1C	CXCL10	HLA-DRB4	IL13RA2	ITGA2	NLRC3	TLR7
BCL3	CD1D	CXCL11	HLAE	IL15	ITGA4	NLRC5	TLR8
BCL6	CD2	CXCL12	HLAF	IL15RA	ITGA5	NOS2A	TLR9
BDCA3	CD20	CXCL13	HLAG	IL17D	ITGA6	OAS1	TNF
BIK	CD200	CXCL14	HMGB1	IL17F	ITGA9	OPTN	TNFAIP3
BIRC5	CD207	CXCL16	HPRT1	IL17RA	ITGAL	OSM	TNFAIP6
CXCR5	CD209	CXCL2	ICAM1	IL17RB	ITGAM	PDCD1	TNFRSF10B
BTK	CD24	CXCL3	ICOS	IL18	ITGB1	PDL1	TNFRSF11A
C1QA	CD27	CXCL5	ICOSLG	IL18RAP	ITGB2	PGK1	TNFRSF11B
C3	CD36	CXCL6	IFI27	IL19	ITGB3	PILRA	TNFRSF12A
C3AR1	CD37	CXCL7	IFI35	IL1A	ITGB4	PILRB	TNFRSF13B
CASP1	CD38	CXCL9	IFI44	IL1B	ITK	PLCG2	TNFRSF13C
CCBP2	CD3E	CXCR3	IFI6	IL1F10	JAK1	POLR1B	TNFRSF14
CCL1	CD4	CXCR4	IFIH1	IL1R1	KCNIP2	POLR2A	TNFRSF17
CCL11	CD40	CXCR6	IFIT1	IL1R2	KLF6	PPARG	TNFRSF18

CCL13	CD40LG	CYBB	IFIT2	IL1RAP	KLRD1	PRG1	TNFRSF19L
CCL14	CD47	CYFIP2	IFITM1	IL1RAPL2	KLRK1	PTGS2	TNFRSF1A
CCL15	CD48	DUSP1	IFITM2	IL1RL1	LAMP1	PTPRC	TNFRSF21
CCL16	CD5	DUSP5	IFNA1	IL1RL2	LAMP2	REL	TNFRSF25
CCL17	CD53	EHD1	IFNA14	IL1RN	LAMP3	RELA	TNFRSF4
CCL18	CD55	ENG	IFNA2	IL2	LAT2	RELB	TNFRSF8
CCL19	CD58	F13A1	IFNA21	IL21	LAX1	RIPK2	TNFRSF9
CCL2	CD63	FAS	IFNA4	IL22RA1	LCK	RPL19	TNFSF10
CCL20	CD68	FCAMR	IFNA5	IL23A	LDHA	RPLP0	TNFSF11
CCL21	CD70	FCER1A	IFNA6	IL23R	LEPR	Runx1	TNFSF12
CCL22	CD74	FCER2	IFNA8	IL24	LGMN	S100A12	TNFSF13
CCL23	CD79A	FCGR1A	IFNAR1	IL25	LRP1	SAA1	TNFSF13B
CCL24	CD79B	FCGR2A	IFNAR2	IL27	LSP1	SDHA	TNFSF14
CCL25	CD80	FCGR3A	IFNE1	IL28A	LTA	SERPINB2	TNFSF18
CCL26	CD83	FCGR3B	IFNG	IL28RA	LTB	SIGIRR	TNFSF4
CCL27	CD86	FCGRT	IFNGR1	IL2R	LTBR	SIGLEC1	TNFSF9
CCL28	CD8A	FLT3	IFNGR2	IL33	LY9	SKAP1	TRAF1
CCL3	CDC42	FN1	IFNK	IL34	LY96	SOCS1	TRAF2
CCL3L1	CEACAM1	Foxp3	IFRG28	IL37	LYVE1	SP110	TRAF3
CCL4	CEBPA	FPR1	IGCL2	IL3RA	MAL2	SPP1	TRAF6
CCL5	CHST4	FYN	IGF1R	IL4	MALT1	STAT1	TRAT1
CCL7	CISH3	G6PD	IGHA1	IL4R	MAP3K7	STAT2	TSLP
CCL8	CKLF	GAPDH	IGHG1	IL5	MAPK1	STAT3	TUBB
CCR1	CLEC2A	GATA3	IGHG2	IL6	MCAM	SYK	TXK
CCR10	CLEC4C	GBP1	IGHG3	IL6R	MDK	TAP1	VCAM1
CCR2	CLECL1	GBP2	IGHG4	IL7	MERTK	TAP2	VEGFC
CCR3	CLTC	GHR	IGHM	IL8	MFGE8	TARP	XCL1
CCR4	CMKLR1	GPR44	IGKC	CXCR2	MGLL	TBP	XCL2
CCR5	COLEC12	GUSB	IGLL1	INHBA	MIF	TBX21	XCR1

CCR6	CREB1	GZMK	IGSF4	IRAK1	MITF	TCL1A	ZAP70
CCR7	CRP	HIF1A	IGSF9	IRAK2	MMP9	THBS1	

Table 7. Coefficient of Determination With and Without Gene Signature

5

Characteristic	Training Set		Validation Set	
	Characteristic Alone	+ Gene signature	Characteristic Alone	+ Gene signature
	coefficient of determination R^2			
Stage	0.061	0.710	0.040	0.678
Depth	0.118	0.719	0.049	0.679
Lymph node status	0.073	0.710	0.073	0.677
Inflammatory infiltrate	0.001	0.710	0.059	0.684
Location of tumor	0.057	0.723	0.001	0.688
Ulceration	0.052	0.707	0.014	0.678
Age	0.097	0.713	0.001	0.853
Gender	0.000	0.719	0.054	0.700
All clinical factors	0.318	0.794	0.224	0.947
Gene signature	0.706	-	0.676	-

10

15

Table 8. Core Gene Picks

Total Input Genes	Training AUC	Validation AUC	Total AUC	Total misclassification in training	Total misclassification in validation	Total misclassification	Final gene number																					
29	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
30	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
31	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
32	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
33	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
34	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
35	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
36	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
37	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
38	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
39	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
40	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
41	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
42	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
43	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
44	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
45	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
46	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
47	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
48	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
49	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
50	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
51	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
52	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
53	0.983437	0.794117647	1.7775545	4	8	12	21	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IL18	IFI27	IL1F7
27	0.979296	0.791176471	1.77047254	4	8	12	20	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IFI27	IL1F7	
28	0.979296	0.791176471	1.77047254	4	8	12	20	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	SYK	CD68	IFI27	IL1F7	
26	0.966874	0.794117647	1.76099135	5	9	14	19	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68	IFI27	IL1F7		
20	0.971014	0.782352941	1.75336743	5	10	15	17	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68				
21	0.971014	0.776470588	1.74748508	6	10	16	18	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68	IL1F7			
22	0.971014	0.776470588	1.74748508	6	10	16	18	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68	IL1F7			
23	0.971014	0.776470588	1.74748508	6	10	16	18	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68	IL1F7			
24	0.971014	0.776470588	1.74748508	6	10	16	18	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68	IL1F7			
25	0.971014	0.776470588	1.74748508	6	10	16	18	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	CCL27	IFNGR1	CD68	IL1F7			
16	0.960663	0.785294118	1.74595664	5	10	15	15	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	IFNGR1						
19	0.973085	0.770588235	1.74367312	5	10	15	16	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	IFNGR1	CD68					
17	0.94617	0.782352941	1.72852271	7	10	17	15	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	IRF5	IFNGR1						
18	0.94617	0.782352941	1.72852271	7	10	17	15	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA	IRF5	IFNGR1						
14	0.954451	0.720588235	1.67503958	7	12	19	14	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA							
15	0.954451	0.720588235	1.67503958	7	12	19	14	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	LGMN	CSF2RA							
13	0.929607	0.720588235	1.65019486	9	12	21	13	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2	CSF2RA								
12	0.929607	0.714705882	1.64431251	9	12	21	12	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD48	CD4	CTSS	CCR4	HLA-DQB1	TAP2									
2	0.84058	0.794117647	1.63469736	14	10	24	2	CD2	HLAE																			
3	0.836439	0.782352941	1.61879186	13	9	22	3	CD2	KLRK1	HLAE																		
4	0.834369	0.782352941	1.61672147	13	9	22	4	CD2	KLRK1	ITK	HLAE																	
11	0.904762	0.694117647	1.59887955	10	12	22	11	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4	CTSS	CCR4	HLA-DQB1	TAP2										
10	0.904762	0.691176471	1.59593838	10	12	22	10	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4	CTSS	CCR4	TAP2											
8	0.886128	0.708823529	1.59495189	9	12	21	8	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4	CCR4													
9	0.89234	0.691176471	1.58351602	8	14	22	9	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4	CTSS	CCR4												
7	0.869565	0.694117647	1.56368286	9	13	22	7	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1	CD4														
5	0.861284	0.688235294	1.54951894	9	13	22	5	CD2	KLRK1	ITK	HLAE	IFNAR1																
6	0.859213	0.688235294	1.54744854	9	13	22	6	CD2	KLRK1	ITK	HLAE	LCK	IFNAR1															

Table 9. Gene Accession Numbers

Gene Name	Accession #	Gene Name	Accession #	Gene Name	Accession #	Gene Name	Accession #
IFI27	NM_005532.3 (SEQ ID NO:1)	TAP1	NM_000593.5 (SEQ ID NO:21)	GZMK	NM_002104.2 (SEQ ID NO:41)	CLTC	NM_004859.2 (SEQ ID NO:61)
HLA-DPB1	NM_002121.4 (SEQ ID NO:2)	CD2	NM_001767.2 (SEQ ID NO:22)	ITK	NM_005546.3 (SEQ ID NO:42)	CCR4	NM_005508.4 (SEQ ID NO:62)
STAT1	NM_007315.2 (SEQ ID NO:3)	CD37	NM_001774.2 (SEQ ID NO:23)	CSF2RA	NM_006140.3 (SEQ ID NO:43)	IFNAR1	NM_000629.2 (SEQ ID NO:63)
MRC1	NM_002438.2 (SEQ ID NO:4)	KLRK1	NM_007360.3 (SEQ ID NO:24)	PGK1	NM_000291.2 (SEQ ID NO:44)	SYK	NM_003177.3 (SEQ ID NO:64)
B2M	NM_004048.2 (SEQ ID NO:5)	CD5	NM_014207.2 (SEQ ID NO:25)	HLA-DQB1	NM_002123.2 (SEQ ID NO:45)	G6PD	NM_000402.2 (SEQ ID NO:65)
IL18	NM_001562.2 (SEQ ID NO:6)	LY9	NM_001033667.1 (SEQ ID NO:26)	CD40	NM_001250.4 (SEQ ID NO:46)	IRF5	NM_002200.3 (SEQ ID NO:66)
IFNGR1	NM_000416.1 (SEQ ID NO:7)	CXCR3	NM_001504.1 (SEQ ID NO:27)	CYBB	NM_000397.3 (SEQ ID NO:47)	RPLP0	NM_001002.3 (SEQ ID NO:67)
CXCL11	NM_005409.3 (SEQ ID NO:8)	CD3E	NM_000733.2 (SEQ ID NO:28)	CCL5	NM_002985.2 (SEQ ID NO:48)	LDHA	NM_005566.1 (SEQ ID NO:68)
TAP2	NM_000544.3 (SEQ ID NO:9)	TNFSF13B	NM_006573.3 (SEQ ID NO:29)	PTPRC	NM_002838.2 (SEQ ID NO:49)	CCR5	NM_000579.1 (SEQ ID NO:69)
CXCL9	NM_002416.1 (SEQ ID NO:10)	LCK	NM_005356.2 (SEQ ID NO:30)	ITGAL	NM_002209.2 (SEQ ID NO:50)	CD27	NM_001242.4 (SEQ ID NO:70)
CLEC2A	NM_207375.1 (SEQ ID NO:11)	IL1F7	NM_014439.3 (SEQ ID NO:31)	IRF2	NM_002199.2 (SEQ ID NO:51)	GAPDH	NM_002046.3 (SEQ ID NO:71)
HLA-DPA1	NM_033554.2 (SEQ ID NO:12)	XCR1	NM_005283.2 (SEQ ID NO:32)	CD68	NM_001251.2 (SEQ ID NO:52)	TUBB	NM_178014.2 (SEQ ID NO:72)
XCL2	NM_003175.3 (SEQ ID NO:13)	C3	NM_000064.2 (SEQ ID NO:33)	TLR7	NM_016562.3 (SEQ ID NO:53)	TBP	NM_003194.3 (SEQ ID NO:73)
CTSS	NM_004079.3 (SEQ ID NO:14)	CD4	NM_000616.3 (SEQ ID NO:34)	CD53	NM_000560.3 (SEQ ID NO:54)	RPL19	NM_000981.3 (SEQ ID NO:74)
CCL27	NM_006664.2 (SEQ ID NO:15)	CD48	NM_001778.2 (SEQ ID NO:35)	SDHA	NM_004168.1 (SEQ ID NO:55)	HPRT1	NM_000194.1 (SEQ ID NO:75)
ICOS	NM_012092.2 (SEQ ID NO:16)	LGMN	NM_001008530.1 (SEQ ID NO:36)	CD8A	NM_001768.5 (SEQ ID NO:56)	ALAS1	NM_000688.4 (SEQ ID NO:76)
IRF8	NM_002163.2 (SEQ ID NO:17)	TNFRSF18	NM_004195.2 (SEQ ID NO:37)	POLR1B	NM_019014.3 (SEQ ID NO:57)	POLR2A	NM_000937.2 (SEQ ID NO:77)
IFITM1	NM_003641.3 (SEQ ID NO:18)	IRF9	NM_006084.4 (SEQ ID NO:38)	IKZF1	NM_006060.3 (SEQ ID NO:58)	GUSB	NM_000181.1 (SEQ ID NO:78)
HLAE	NM_005516.4 (SEQ ID NO:19)	SKAP1	NM_003726.3 (SEQ ID NO:39)	ITGB2	NM_000211.2 (SEQ ID NO:59)	ABCF1	NM_001090.2 (SEQ ID NO:79)
GATA3	NM_001002295.1 (SEQ ID NO:20)	TARP	NM_001003799.1 (SEQ ID NO:40)	ACTB	NM_001101.2 (SEQ ID NO:60)	LY64	NM_005582.2 (SEQ ID NO:80)

The entire sequence of each accession number provided in Table 9
5 above is incorporated herein by reference in its entirety.

EXAMPLE 2.

Abstract

10 Melanoma is a devastating form of skin cancer that is rarely curative in advanced stages of the disease. Even early stage melanomas can metastasize and accurate diagnosis and clinical staging is vitally important. Evidence shows that the host immune system plays a determinative role in clinical outcomes in
15 cancer. Studies in liver cancer, lung cancer, prostate cancer, and advanced melanoma have revealed that the expression of inflammatory genes by "normal" host cells correlates with survival. Prognostic information in cancer can therefore be

learned from study of the host immunologic milieu. Characterization of the immune signature of melanoma is clearly an important step in disease staging, prognostication, and therapeutics. Primary melanomas have been inadequately studied to date because clinical standards require that the entire specimen be fixed in formalin to preserve morphology for pathology diagnosis. This process is damaging to RNA, hindering analysis of gene expression signatures. Novel technologies recently developed, however, allow for analysis of partially degraded RNA derived from formalin fixed paraffin embedded (FFPE) tissue blocks. Pure RNA extracted from these FFPE tissue blocks can be analyzed using an nCounter system (NanoString) which is capable of detecting as little as 0.5 fM of a specific mRNA. This has opened a new avenue for high throughput research in this area. In this study such cutting-edge technologies are used to create a gene signature for recurrent melanoma. Specifically this study has three aims: screen the dermatopathology database at Mount Sinai and identify melanoma specimens from patients who subsequently recurred and matched specimens from patients who did not recur; establish a protocol for extraction of RNA from paraffin embedded primary melanoma tissues; and establish an inflammatory signature for early stage melanoma at high risk of recurrence. Characterization of the immune signature of melanoma is clearly an important step in disease staging, prognostication, and therapeutics of this devastating disease.

Introduction

Melanoma is the deadliest form of skin cancer. Over 160,000 new cases of melanoma are diagnosed annually worldwide with 40,800 deaths per year from advanced (stage IV) melanoma.[69] The median survival time in stage IV melanoma is six months and few effective therapies exist to treat this disease. A meta-analysis of all phase II cooperative group trials in stage IV melanoma demonstrated a median survival time of 6.2 months with 25.5% of patients alive at one year.[70] A review of 35 immunotherapy trials involving 765 patients demonstrated an overall response rate of 3.3%.[71] Clearly, a better way is

needed to identify patients at risk or not at risk for progression to metastatic disease so as to guide therapeutic decisions and enable one to tailor therapeutic approaches. There are several important examples in cancer of how gene signatures have affected management and the course of disease. In breast cancer, a prognostic tool called Oncotype DX was developed to identify patients whose prognosis was sufficiently good that they may not require chemotherapy in addition to hormonal therapy.[72] HER2 (tyrosine kinase receptor) amplification is a predictive signature for those patients who will benefit from trastuzumab.[72] The KRAS mutation in colorectal cancer is a predictive signature for patients who will benefit from the EGFR inhibitors and the EGFR mutation in non small lung cancer is a predictive signature for patients who would benefit from first line treatment with erlotinib.[73] Systematic, well designed trials have shown that gene signatures are effective for prediction or prognostication in a variety of malignancies.

Gene signature studies in melanoma have been hampered by several factors including the need for fresh tissue to perform RNA extractions, restrictions to studying metastatic lesions in order to preserve primary biopsies for clinical use, and the descriptive nature of most studies with few clinical endpoints being utilized. Additionally, melanoma studies are particularly difficult given how small the specimens are and the need to preserve some tissue for clinical use. Two key studies have evaluated gene expression in primary melanomas. Kauffman et al evaluated 60 primary fresh frozen melanoma lesions (17 in the validation set) with four years of follow-up. They specifically reviewed only DNA repair/replication genes and found a 48 gene signature associated with metastatic progression.[74] Winnepenninckx et al evaluated 234 primary melanoma lesions with at least 4 years of follow up and found a 254 gene signature associated with distant metastases free survival. These genes were mainly involved in DNA replication;

a subsequent studies have not shown this gene signature to be of utility in the clinical setting.[75]

Melanoma has classically been regarded as an immune mediated disease. Pathological studies have demonstrated that skin melanomas are characterized by lymphoid infiltrates to various degrees with brisk infiltrates being a good prognostic indicator. Regression of primary melanoma due to heavy lymphoid infiltrates is a common clinicopathologic feature. Tumor infiltrating cells in melanomas contain T cells, B cells, dendritic cells, and macrophages.[75] Several gene signature studies have demonstrated immune gene upregulation in several settings. Metastatic melanoma in comparison to benign nevi had increased expression of immune related genes such as HLA-B, HLA-H, and STAT1.[76] Bogunovic et al. found that in metastatic melanoma, upregulation of a gene signature profile reflective of immune activation was associated with improved survival while cell proliferation genes negatively impacted survival.[77] Evidence suggests that the host immune system plays a determinative role in clinical outcomes in melanoma. Therefore, prognostic information in melanoma can possibly be informed by examining the expression of genes involved in the host immunologic response milieu.

This study establishes an immune signature for Stage II cutaneous melanoma at high risk of recurrence. Such an immune signature helps identify which subset of stage II melanoma patients benefits from earlier therapeutic interventions. Although, research in this field has been hampered by the need for cell lines or fresh tissue to conduct functional genomics newer technologies are used that allow RNA extraction from formalin fixed paraffinized melanoma samples in conjunction with microarray assays designed specifically to assay to degraded RNA have opened a new avenue for large volume research in this area.

Specific Aims:

Specific Aim 1: To screen the dermatopathology database at Mount Sinai and identify melanoma specimens from patients who recurred and specimens from patients who did not recur; and to create a database of clinical and pathologic characteristics (obtained from electronic and paper medical records as well as patient surveys) for each patient.

A. Evaluate stage II melanomas from the dermatopathology database and examine the clinical record to identify 24 patients who recurred

B. Identify lesions from 24 patients who did not recur matched for stage, age, gender and ethnicity.

Specific Aim 2: To extract RNA from paraffin embedded primary melanoma tissues using the Ambion RecoverAll Total Nucleic Acid Isolation kit.

A. To optimize a standard protocol for extraction of RNA (**Ambion** RecoverAll Total Nucleic Acid Isolation kit) from paraffinized tissue, for use with small amounts of previously archived tissue samples.

Specific Aim 3: To establish an inflammatory immune signature for early stage melanoma at high risk of recurrence

A. To identify a panel of 500 inflammatory and cancer genes based on a systematic search of the literature for genes that are significant in both melanoma and other malignancies

B. To employ NanoString technology to screen 24 melanomas which recurred and 24 melanomas (discovery set) that did not recur for expression of a panel of 500 inflammatory markers to identify candidate genes. A gene signature is created from these genes

C. To validate a gene signature from the discovery set in a training set of recurrent and non recurrent melanomas and determine whether differential expression of inflammatory genes provides prognostic information beyond pathological markers of prognostication

Methods

Patient selection: Stage II melanoma is defined as a tumor greater than 2mm in depth or a tumor between 1-2mm with ulceration. Stage II melanomas are chosen as they are the melanomas at highest risk of recurrence or progression to metastatic disease. The Mount Sinai Hospital pathology department has all melanoma related pathology stored from 1999-2009. An electronic database is queried to identify all patients diagnosed with melanoma and then each pathology report is reviewed by the investigators to determine if they fit the criteria of either stage II melanoma. A 10 year time period is examined as the FFPE extraction kit has been validated for samples up to 10 years old. As melanoma is most likely to recur in the first two years after diagnosis, the patient must have at least two years of follow up to be included in this study. Each pathology specimen is reviewed with Dr. Robert Phelps, the head of dermatopathology. As part of this study a database that stores both pathologic and clinical characteristics of every specimen in the study is created. Pathologic characteristics (as reviewed by Dr. Phelps) including depth of lesion, ulceration, immune infiltration, number of mitoses, degree of sun damage, pathologic subtype, Clarks level, satellite lesions, blood vessel invasion, and lymphatic invasion, and lymph node involvement are recorded.

As the dermatopathology database does not contain any clinical information about patient outcomes, it is created. A systematic review of the electronic medical records (2 existing databases at Mount Sinai) and paper charts (Main medical records, Dermatology records, Oncology records) is conducted on every patient in the study. Clinical characteristics that are recorded include recurrence status, gender, ethnicity, alive/deceased, cause of death if applicable, metastatic sites, treatments (chemotherapy, surgery, radiation), other skin tumors, other cancers, a family history of melanoma, a history of immune disease, site

of recurrence, how many months since primary melanoma did the recurrence happen, and what treatment did the patient receive for the recurrence.

5 Additionally, in cases where the medical record is incomplete, a patient phone questionnaire is administered (Table 10). Patients are mailed a cover letter explaining the project as well as a consent form. A follow up phone call by one of the investigators is conducted to administer the questionnaire to 10 consenting patients. The phone survey attempts to gather the following information: does the patient receive regular dermatology follow up, how many moles do they have, what is there hair and skin color, do they have freckles, when they are exposed to the sun how frequently do the burn versus tan, 15 how many blistering sunburns have they had in the past, what is there occupation, how many hours do they spend outside for their occupation, how many hours do they spend outside for their leisure activities, do they wear sunscreen of protective clothing on a daily basis.

20

RNA extraction: For extraction of RNA from FFPE tissues, several commercially available kits were tested. The best yield of total RNA came from the RecoverAll Total Nucleic Acid Isolation Kit which is optimized for and can only be used in 25 FFPE samples. Historically, the chemicals used in preserving tissue in paraffin have made the samples unusable in molecular analysis. The paraffinization process made the RNA from these samples too fragmented to be compatible with molecular techniques. The RecoverAll Total Nucleic Acid Isolation kit 30 uses a protease digestion process that releases the maximal amount of RNA, of all sizes, as possible. RNA from frozen tissue are extracted using a four step protocol that involves: 1. Phase separation (uses Trizol based reagents, homogenization); 2. RNA precipitation and incubation; 3. RNA 35 wash; 4. Redisolving the RNA in Rnase free water. The microarray requires no more than 5uL of sample with a

concentration of 20ng/ul. More than 33-50% of the samples should be greater than 300 base pairs.

Immune assay: NanoString is a gene expression assay that directly captures and counts individual mRNA transcripts. It is uniquely suited for measuring partially degraded RNA as found in FFPE tissues. Total RNA is mixed with a pool of probes bound to strings of fluorophores. The color sequence encoded by each nanostring is specific to a given probe. Experimentally, 100ng of total RNA is mixed with a mixture of up to 550 unique DNA/fluorophore and a hybridization step follows. The reporter probe is a 50mer oligonucleotide. As a result partially fragmented samples can be detected using this technology without affecting the quality of the results. After hybridization the excess reporter probes are washed off. The transcripts present in the total RNA sample are identified by binding the hybridized RNA/probe to a substrate and scanning the substrate with a laser device. The surface is imaged by a CCD camera and the signal processed by software which determines total counts for each reporter probe. With a sensitivity of 500 attomolar this assay can detect as little as one copy of RNA per cell using 100 nanograms of total RNA as input. 500 genes can be evaluated using the NanoString assay. To identify these genes a PubMed literature search for gene expression profiling in melanoma, gene signatures in melanoma, immune/inflammatory genes in melanoma, and immune signatures in other malignancies (eg. prostate, breast, liver, lung) was conducted. Additionally, commercially available inflammatory panels were screened for possible candidate genes.

A discovery set of at least 24 recurrent and at least 24 non recurrent melanomas is evaluated to identify genes for the immune signature. The gene signature from the discovery set is then be applied to the validation set of samples (obtained from the dermatopathlogy database) to estimate prediction accuracy. To demonstrate that the new signature is significant

it is compared to standard prognostic factors (depth, ulceration).

Sample Size: The number of samples necessary for the identification of a robust biomarker signature is variable. Sample size depends on the amplitude of the difference between and variability within study groups. Little consensus exists for the calculation of sample size for microarray experiments.[78-79] Best practices utilize independent sets of samples for the purpose of validating candidate signatures. Thus the robustness of the signature relies on a statistically significant association between the predicted and true phenotypic class in the sample sets. In the discovery set, this is indicated by the Fisher's Exact Test result, as well as the estimates for sensitivity and specificity and their corresponding exact 95% confidence intervals. In this study, results obtained from the discovery set are used for power calculations for the validation set.

Statistical Analysis: Continuous variables (e.g. depth) are described by their frequency of observations, mean, median, standard deviation, minimum, and maximum values. *Categorical variables* (e.g. recurrence) are described by their frequency and percentage. In addition to the previously mentioned microarray analysis techniques, other inferential statistics are used to assess the association between transcripts based variables and clinical outcomes. Appropriate methods are chosen depending on the specific outcome's level of measurement and whether or not observations are independent. For *continuous* variables with *independent observations* comparisons of central tendency are made using ANOVA or Kruskal-Wallis test. For *dependent observations* (e.g. clustered or longitudinal data), linear mixed model analyses is used. For *categorical* variables with *independent observations* likelihood-ratio chi-square tests are used to univariately test for differences among groups. For *dependent observations* McNemar's or Cochran's Q (for tables larger than

2 by 2) test is used. For multivariate analyses of binary outcomes generalized linear mixed models (assuming a binomially distributed outcome and using the logit link function) are used to account for correlated observations. For
5 time to event (e.g. survival and disease progression) analysis, the Cox proportional hazards model isused. The Benjamini and Hochberg method for controlling the false discovery rate are used to account for multiple testing. Descriptive statistics for clinical and demographic variables are given overall and
10 by appropriate classifications (e.g. disease stage).

Preliminary results:

Clinical Database: To date 70 patients with stage II melanoma
15 have been identified for the database. Obtaining the corresponding clinical information and recurrence status is ongoing and currently 12 recurrent and 12 non recurrent patients have been identified that can be used for the discovery set. This information was initially collected
20 utilizing excel and is currently being converted to an access database.

Phone Questionnaire to ascertain additional clinical information: Ten patients were consented for administration of
25 the questionnaire as seen in Table 10. The questionnaire is used to supplement clinical information that is often not found in the patients clinical record but is relevant to their dermatology and oncology history. The clinical information that is abstracted from the patient's medical record includes:
30 recurrence status, gender, ethnicity, alive/deceased, cause of death if applicable, metastatic sites, treatments (chemotherapy, surgery, radiation), other skin tumors, other cancers, a family history of melanoma, a history of immune disease, site of recurrence, how months since primary melanoma
35 did the recurrence happen, what treatment did the patient receive for the recurrence, does the patient receive regular dermatology follow up, how many moles do they have, what is

there hair and skin color, do they have freckles, when they are exposed to the sun how frequently do they burn versus tan, how many blistering sunburns have they had in the past, what is there occupation, how many hours do they spend outside for their occupation, how many hours do they spend outside for their leisure activities, and do they wear sunscreen of protective clothing on a daily basis.

RNA extraction and custom immune gene assay: RNA has been successfully and repeatedly been extracted from FFPE specimens using the Ambion RecoverAll Total Nucleic Acid Isolation kit. This commercial protocol was optimized in the Saenger Lab for extraction of RNA from skin tissue. Over several months this commercial protocol was adjusted to increase RNA yield from cutaneous tissue. The commercial protocol involves four major steps – deparaffinization, protease digestion, nucleic acid isolation, and nuclease digestion and final nucleic acid purification. Two major modifications were made. During the deparaffinization step four 20micron sections are deparaffinized in one tube which is the maximum number of sections the protocol allows for. It was found that given the small size of the melanoma samples, using less than this amount gave RNA yields unsuitable for nanostring analysis. Additionally, when incubating the samples in 100% xylene, extending the incubation period to a maximum of 30 minutes at 50 degrees Celsius rather than 3 minutes resulted in higher RNA yields, presumably secondary to more complete deparaffinization. The second major step of the protocol requires digestion of the melanoma tissue with digestion buffer and protease. This step allows for release of the RNA from the deparaffinized melanoma tissue. The commercial protocol recommends incubation at 50 degrees for 15 minutes followed by 15 minutes at 80 degrees Celsius. It was found that this incubation time did not produce yields high enough to utilize for nanostring analysis. In the literature, RNA extraction from fresh skin tissue has been historically difficult because the tissue is tough, hard to homogenize, and

contains many RNAases. Some of these issues may be similar in extraction from paraffin. It was found that extending the digestion time to three hours at 50 degrees Celsius followed by 15 minutes at 80 degrees Celsius increased the yield of RNA to sufficient amounts suitable for nanostring analysis. This is the maximum recommended time per the commercial protocol. This has been demonstrated by using an Agilent Bioanalyzer for total RNA. A 500 gene immune panel has been assembled for the nanostring assay. A PubMed literature using the following key phrases were used to identify relevant genes: gene expression profiling in melanoma; gene signatures in melanoma; immune/inflammatory genes in melanoma; and immune signatures in other malignancies (eg. prostate, breast, liver, lung). Additionally, commercially available inflammatory panels were screened for possible candidate genes. Genes from the following functional categories were chosen – macrophages, neutrophils, natural killer cells, dendritic cells, cytokines, chemokines, adhesion molecules, toll like receptors, complement, t cells, b cells, cell death, cell signaling, major histocompatability complex I and II, immunoglobulins, NF kappa B, and the JAK-STAT pathway. These genes were mainly chosen because of their relevance to cancer surveillance or progression in melanoma or other malignancies as found in our literature search. This gene set that was created has not been used, in any other studies. The gene set being used can be viewed in Table 6.

Ongoing work

12 recurrent and 12 non recurrent patients were identified that can be used for the discovery set and have extracted RNA on these specimens. These have been sent to NanoString Technologies for analysis on a custom immune gene assay that created with NanoString Technologies. These results are obtained and a gene signature from the discovery set is applied to the validation set of samples (obtained from the dermatopathlogy database) to estimate prediction accuracy. To maximize the sample number, Tammie Ferringer M.D., a

dermatopathologist in the Geisinger Health Network is collaborated with.

Summary of Novel Findings to Date:

5
1. Establishment of a new database linking clinical and pathological information on patients with early stage melanoma with recurrent verses non recurrent disease
2. Adaption, modification, and optimization of an RNA extraction procedure for extraction of RNA from a small sample of archived tissue material previously embedded in paraffin.
10
3. Development of a new phone survey to capture important clinical information
4. Development & utilization of a customized & novel 500 immune gene panel (Nanostring immune gene set).
15

Table 10: Institutional Review Board (IRB) approved phone questionnaire administered to study subjects with incomplete clinical records

1) Have you been diagnosed with melanoma? When and how many times?
2) Do you have any melanoma in your body currently as far as you know?
3) Did any of your melanomas spread beyond the skin?
4) Did any of your melanomas reoccur again in the skin after it was removed? How long after?
5) Have you had any other melanomas that were diagnosed by a doctor not associated with Mount Sinai Hospital? If so, where was the melanoma located, when were you diagnosed, and did you receive any treatment for this melanoma?
6) Do you get regular follow up for your melanoma? How often do you visit your dermatologist?
7) Do you have more than 10 moles?
8) What would you describe as your race?
9) Do you have red hair? Freckles?
10) Does anyone in your family have melanoma
11) Have you ever been diagnosed with a cancer other than melanoma, if so what cancer, when, and are you receiving active treatment?
12) When you are exposed to the sun do you always burn? Burn sometimes? Always tan? Tan sometimes? Never tan? Never burn?
13) As a child/young adult how many blistering sunburns did you get?
14) What type of work do you do and how many hours a day are your exposed to the sun in your job?
15) What kind of leisure activities do you do and how many hours a day to those activities expose you to the sun?
16) How often do you use sunscreen and/or where sun-protective clothing such as hats, long sleeves, and/or long pants?
17) What is your ethnicity?
18) Have you been diagnosed with any other skin tumors

EXAMPLE 3.

1) There are 70,000 cases of melanoma a year in the US of which approximately 25,000 are deep primary melanomas. Early stage 3 melanomas (sentinel lymph node positive) and late stafe 2 melanomas (deeper than 2mm or deeper than 1mm and ulcerated)

are included. These melanomas are at high risk of recurrence, causing advanced disease and death Depth provides some prognostic information but generally only allows the estimation of mortality risk to between 25-50%. Therefore there is a need for better information to guide patient and physician choices.

In order to define a better biomarker for melanoma recurrence, RNA was isolated from primary melanoma tumors and measured expression of inflammatory genes. This has not been done yet because primary melanomas are preserved in paraffin and this affects RNA quality, making the RNA difficult to analyze. A specialized technology, NanoString, was used to analyze the degraded RNA. Excitingly, of the 33 genes that were significantly different between recurrent and non-recurrent groups, all of them were up-regulated in the non-recurrent patients, suggesting that inflammation is protective.

More importantly, a 10 gene signature was determined, similar to Oncotype Dx for breast cancer, which allows for determination of risk of recurrence for breast cancer. This signature, in our test sample, allows for detection of recurrence risk with 90% specificity and 80% sensitivity. Genes included in this signature make biologic sense as they correlate with markers of T cell infiltration. Current morphologic assays of T cell infiltration are crude and do not allow for any phenotypic differentiation between lymphocyte population so our genetic screen would add a great deal of information to current clinical parameters.

Notably this approach may have application beyond the primary melanoma setting. Inflammatory markers may be predictive of prognosis as well as response to immunotherapy in the metastatic setting. The same genes permitting the tumor to escape the immune system early on in disease may also be operative at more advanced stages. This technology

could therefore be applied to patients with metastatic disease to predict survival and also to predict response to immunotherapy such as ipilimumab, anti-PD1, oncovex treatments, or potentially conventional or targeted therapies resulting in antigen release and potential immune response. NanoString would have great application here because, although it is theoretically possible to preserve frozen tumor specimens, most samples available in clinical practice are paraffin embedded. An identical or similar panel of genes may have utility here.

The 10 relevant genes are: HLAE, CD2, ITK, KLRK1, CCR4, LCK, CD48, CD4, CXCR3, CD53

CD2 is particularly intriguing because CD53 and CD48 are both implicated as having interactions with CD2. CD2 is a co-stimulatory marker on T cells and also implicated in NK cell and dendritic cell function. These genes are associated with T cell responses and with the recruitment of inflammatory cells to the skin.

2) Patients can have the test done to determine their risk of recurrence. This helps define monitoring as far as whether they need regular imaging tests and also help patients to assess their own risk and decide whether to take adjuvant therapies which can be very toxic. Patients want to know this information.

3) Oncotype Dx

There is currently no biomarker for recurrence of primary melanoma based on gene expression and there is no biomarker for cancer recurrence based exclusively on inflammatory gene expression, and no biomarker related specifically to T cell genes and interferon response genes or to any of the genes listed above.

EXAMPLE 4.

ABSTRACT

Improved biomarkers are needed for patients with resected stage II-III melanoma. Clinico-pathologic features such depth, ulceration, and sentinel lymph node status, while essential to clinical practice, often fail to predict progression in individual patients. Biomarker development has been hindered by clinical standards dictating that the entire specimen be formalin fixed and paraffin embedded (FFPE) for morphology evaluation, a process damaging to RNA. To define a biomarker for melanoma progression, mRNA copy number of 446 genes was measured in completely resected stage II-III FFPE primary melanoma using NanoString, a hybridization assay suited for analysis of partially degraded RNA. A 53-gene biomarker of progression was defined using receiver operating characteristic (ROC) curves in a test population (N=40). Prediction power of this panel was tested in a second independent population (N=48, AUC=0.787, p<0.001). Protein levels of the most differentially expressed gene, CD2, associated with non-progression by immunohistochemistry. In the validation population, multivariable analysis identified gene signature score as an independent predictors of progression (p<0.001) and survival (p=0.03 Analysis of publicly available expression data in primary melanoma identified a co-expression network and a module enriched for the 53-gene panel and immune response.). Signaling pathway analysis revealed the 53 genes to form a dense network enriched in T and NK cell signaling pathways. mRNA levels of 53 genes with immune-surveillance function are co-regulated in primary FFPE melanoma, predict non-progression, and should be evaluated in larger studies as a biomarker.

INTRODUCTION

Metastatic melanoma is a devastating illness, taking the lives of over 48,000 people worldwide per year.[106] Newer immune therapies are bringing hope to patients with advanced disease. Nonetheless, mortality rates remain very high for patients

with stage IV melanoma where the estimated survival rate at 5 years is less than 20%.[107] Surgery, for decades, has been the only reliably curative therapy for this cancer, and, unfortunately, despite significant advances, medical treatments remain non-curative at the present time for the majority of patients.[108]

Patients who have had a stage II or stage III melanoma surgically removed remain at high risk for progression and death because micro-metastasis may have spread to other body sites prior to resection. No highly effective therapy is available to prevent progression. While interferon is FDA approved in patients with stage IIB-III melanoma, it has limited benefit and a difficult toxicity profile, and therefore is inconsistently prescribed in oncology practices across the United States.[109-111]

Critical prognostic features in the pathology report describing a newly resected primary melanoma are depth and ulceration, and these are incorporated into the AJCC melanoma staging system, with stage II melanoma defined as a lesion 2mm or greater, or 1mm or greater with ulceration.[112,113] The best test available to further estimate risk is the sentinel lymph node biopsy procedure, and stage III disease is defined by a positive sentinel lymph node.[114,115] Stage III disease, however is highly heterogeneous. Five year survival ranges from 87% for stage III patients with one nodal micro-metastasis and a primary lesion less than 2mm down to 36% for stage III patients with four or more involved nodes.[116]

Meanwhile, patients with a clean sentinel lymph node (Stage II) are not safe from progression either as patients with IIC disease (primary lesion 4mm or greater, or 2mm with ulceration and a negative node) have a five year survival of only 48%.[112] Thus, a primary melanoma greater than 4mm in depth confers a worse prognosis than a microscopic focus of melanoma in the sentinel node, likely due to hematogenous spread.[112] There is a

clear need for accurately, broadly applicable prognostic tools for patients with resectable stage II-III melanoma, both for clinical care, and because improved prognostication would greatly enhance stratification for study of adjuvant therapies.

Evidence is growing that the phenomenon of immunosurveillance, originally defined in mice, plays a key role in human solid tumors.[117-119] Thus, the immunoscore, has recently been developed as a biomarker for cancer progression.[120] In melanoma, it has long been known that tumor infiltrating lymphocytes (TILs) can confer a more favorable prognosis, and this has recently been validated in patients with stage I-III melanoma.[121,122] Two factors, however, limit the widespread clinical application of TIL quantification. First, TIL quantification is subjective and subject to observer variability.[123] Second, the majority of patients have "non-brisk" TILs, an intermediate category which offers little further clarification of the prognosis.[122] More objective, molecular immune markers are needed.

A major barrier to the development of molecular markers in primary melanoma tumors in particular is the fact that most clinical treatment centers require that the entire specimen be formalin fixed and paraffin embedded (FFPE). This is because the tumors are very small and key features including depth and ulceration can be accurately determined only in FFPE specimens.[124] Thus only those markers which can be assayed in FFPE tissues are applicable to the vast majority of stage I-III melanoma patients in the United States. Genomic markers of inflammation have shown promise in more advanced cases where frozen tissue can be obtained from larger metastatic lesions.[125-127] In melanoma, however, most of the uncertainty exists in the clinical setting before these large metastatic lesions develop. Furthermore, the immune-surveillance hypothesis suggests that it is precisely at the earlier stages of tumor growth, that the determinative balance between tumor and immune system is established.[128]

In order to address the need for an FFPE based immune biomarker in primary melanoma, we tested the ability of NanoString, a technology developed to quantify mRNA transcripts in partially degraded samples, to distinguish patients with a good prognosis from patients with a poor one.[129] We find that expression levels within the original biopsy specimen of an 53 gene panel comprised of genes implicated in immune surveillance predicts clinically non-progression and prolonged survival in two independent sets of patients with resectable melanoma. Herein we present, to our knowledge, the first genomic based immune biomarker based on analysis of FFPE primary melanoma. Large scale prospective studies should be initiated to define the role of mRNA quantification of genes with immune function using NanoString in primary tumors of patients with resectable melanoma.

MATERIALS AND METHODS

Patients and Samples

The training set included FFPE primary melanoma tumors from 40 patients with completely resected stage II/III melanoma identified by screening dermatopathology databases between January 2001 and January of 2011 at Geisinger Medical Center (GMC, Danville PA, 32 patients) and Mount Sinai School of Medicine (MSSM. New York, NY, 8 patients). Following approval by the local institutional review board (IRB), authorized personnel obtained clinical information at each institution. Progression was defined as biopsy proven melanoma which had spread beyond the local lymph node basin (stage IV) or was no longer amenable to surgical resection. Non-progression was defined as no further evidence of melanoma following excision of the primary lesion with a minimum follow up of 24 months. Patients with incomplete clinical follow-up were contacted by mail and telephone under an IRB-approved protocol. The validation set included additional patients from GHS (15) and MSSM (7) as well as 25 patients meeting criteria defined above from New York University Medical Center (New York, NY). A complete review of all patient records was performed on

December 31, 2011 for the training set and December 31 2012 for the validation set and living patients were censured.

Analysis of gene expression

5 RNA was extracted from primary melanoma specimens using the Ambion® RecoverAll Total Nucleic Acid Isolation Kit (Life Technologies, Carlsbad, CA). 446 genes were selected based on a PubMed literature review (Table 13). The nCounter platform (NanoString Technologies, Seattle, WA) was used to quantify
10 relative mRNA copy number.[130]

Immunohistochemistry

IHC was performed on 5-µm charged slides using anti-CD2 monoclonal antibody (MRQ-11, Ventana Medical Systems, Tucson,
15 AZ). Sections were deparaffinized and stained using a Ventana BenchMark XT immunostainer. Slides were evaluated by two of the study authors (SGB & MMM) in a blinded manner in 8 random High Powered Fields (HPFs) using an ocular micrometer with a 1 mm^2 grid (Nikon Eclipse E400®).
20

Statistics

Ensemble classification/regression method and ROC curves

Classification was performed using an ensemble feature selection method encapsulating two standard classifiers:
25 random forest and elastic net, both embedded in data bootstrapping to boost the robustness of the finalgene panel. The starting 446 genes from the training experiment were ranked and filtered based on prediction power of melanoma progression in the training cohort and a subset of 53 genes
30 was selected as final gene panel. ROC curves were generated and the area under the curve (AUC) was calculated on both training and test datasets. Detailed methods are included in the below.

35 *Demographic, survival and multivariable analysis*

Two tailed student T tests generated p values for continuous variables including age, depth, and mitotic rate. Other non-

continuous characteristics were analyzed using a two-tailed Fisher's exact test or, in the case of TILs, a chi square test. For survival analysis, Kaplan-Meier analysis and Log-Rank (Mantel Cox) tests were performed. Graphpad Prism version 5.0 was used (San Diego California USA) and statistical significance was defined as $p<0.05$ without correction for multiple comparisons. Standard multivariable logistic and Cox propotional hazards analysis were performed using XLSTAT (Addinsoft) software.

Co-expression network analysis

From the NIH GEO database, 46 samples of gene expression data identified based on origin in primary melanoma tissue and expression platform (Table 14) were collected (GEO accession ID: GSE15605)[131]. Co-expression network analysis was performed using Weighted Gene Co-expression Network Analysis (WGCNA)[132] to identify highly correlated gene modules among whole-genome genes in early stage melanoma patients. Let N denote the total number of genes in the whole-genome. For an overlap of m genes between a module of size M and a panel of genes of size n, an enrichment fold was computed using the ratio of the proportion of panel genes contained in the module (m/n) to the proportion of whole genome genes contained in the module (M/N). That is, enrichment fold = $(m/n)/(M/N)$. The p-value of this enrichment fold is calculated by using Fisher exact test.

Physical interaction network analysis

To analyze the enrichment of the 53-gene panel from both a network perspective and a functional perspective, a gene network was constructed using the gene network tool VisAnt 4.0.[133,134] A reference network was similarly constructed using the original 446-gene panel. For a detailed description of the network construction, see supplemental appendix methods. Density, clustering coefficients, and other network statistics were compared across networks. Furthermore P-values associated with clustering coefficients on each network were generated by randomizing networks of the same size and density.

Pathway and gene ontology enrichment

Gene panels were annotated using the functional database and tool DAVID.[135,136] The default list of whole-genome was chosen as the background gene set, and each network gene list was
5 tested for enrichment of KEGG pathways or GO term biological process (BP) or GO term molecular function (MF).

RNA Extraction

FFPE tissue blocks were cut into four 20μm sections and
10 treated with 100% xylene (Fisher Scientific, Pittsburgh, PA) to deparaffinize. Samples were washed twice with 100% ethanol (Absolute Ethanol Molecular Biology Grade 200 proof, Fisher Scientific, Pittsburgh, PA) and dried via vacuum centrifugation at 40°C. Tissue was then incubated in Digestion
15 Buffer and Protease (Ambion® RecoverAll, Life Technologies, Carlsbad, CA) at 50°C for 3 hours, followed by a 15-minute incubation at 80°C. RNA was separated using an Isolation Additive/Ethanol mixture (Ambion® RecoverAll, Life Technologies, Carlsbad, CA), and filtered by centrifugation at
20 10,000rpm. The sample was rinsed with Wash 1 and Wash 2 (Ambion® RecoverAll, Life Technologies, Carlsbad, CA), before and after incubation with DNase for 30 minutes at RT. RNA was eluted with 60μl of Elution Solution (Ambion® RecoverAll, Life Technologies, Carlsbad, CA) at RT.

25

Dermatopathology

Primary and recurrent melanomas were selected from the Tamtron PowerPath database at the Mount Sinai Medical Center (New York, NY), the Cerner CoPathPlus database at Geisinger Heath System
30 (Danville, PA), the Oracle Clinical RDC 4i database at NYU Medical Center (New York, NY), as well as the NovoPath database at Englewood Hospital and Medical Center (Englewood, NJ). Combined, the databases contained more than 2,500,000 specimens collected since 1985 and derived from surgical
35 pathology, dermatopathology, neuropathology and cytology. Criteria for inclusion were: completely resected stage I-III melanoma, adequate clinical follow-up for all variables listed

in the demographic table, and availability of tissue of sufficient quality to extract RNA. Selected slides and paraffin blocks were reviewed by two of the study authors (RGP and SS). Each sample was evaluated for histogenetic type, 5 extent and type of inflammatory infiltrate, thickness, and ulceration.

Tumor infiltrating lymphocytes (TILs) were defined as ones that percolated between and around tumor cells, as previously 10 described by Rao et al.[1] High magnification images of H&E stained tumor specimens were evaluated for TILs by a dermatopathologist. Brisk refers to lymphocytes present throughout the substance or infiltrating the entire base of the vertical growth phase; non-brisk refers to lymphocytes in 15 one focus or more of the vertical growth phase, either dispersed throughout or situated focally in the periphery; and absent if there were no lymphocytes or if they were present but did not infiltrate the melanoma.[107]

20 *Immunohistochemistry* Five micron sections of the same paraffin-embedded tissue samples analyzed by Nanostring were prepared for immunophenotypic analysis. Immunohistochemistry (IHC) was performed using primary, pre-diluted anti-CD2 (MRQ-11, mouse anti-human, Ventana Medical Systems, Tucson, AZ). 25 Sections were deparaffinized, stained according to standard protocol using a Ventana BenchMark XT immunostainer and manually counterstained.[108] The immunohistochemical slides were evaluated and interpreted by two of the study authors (SGB & MMM) in a blinded manner without knowledge of corresponding 30 clinical data. For each sample, cells with circumferential membrane staining were counted and averaged in 8 random HPFs using an ocular micrometer with a 1 mm^2 grid (Nikon Eclipse E400®).

35 **Statistical Analysis**

Cross-validation We simulated 900 iterations of a 11-fold cross-validation on the training dataset with random sample

reordering in each iteration to strengthen the robustness of our final classifier model. 4 samples were removed at random. These sample sets were then used as training data to fit a statistical model. 10,000 model training tasks were performed. The trained model and gene predictors selected were recorded in each task yielding 10,000 models and 10,000 lists of gene predictors based on randomly sub-sampled training samples. For each model, we performed a classification for the entire training (44 samples) and validation datasets (37 samples). To derive a robust list of gene hits by these models, the 10,000 gene lists were pooled and the statistical count of each gene (out of 446 genes) was selected by these models using the training cohort only. A higher count value for a given gene indicates that it is frequently selected as a predictor during the cross-validation process. Three genes (IFNG, TNFSF18, and CREB1) were excluded from the signature because the p value did not meet the cutoff in the preliminary analysis of the training data and levels were therefore not tested in the validation set. Finally, all genes selected at least once in the 10,000 cross-validation were put into a final model training task to yield an optimal, compact predictor gene list of 53 genes.

Ensemble classification/regression method We employed a two-step sequential ensemble classification scheme that sequentially concatenated two widely applied classifiers: random forest and elastic net. Random forest itself is an ensemble classifier consisting of many decision trees that generates the mode of individual classes yielded by independent trees. A powerful variable selector and classification/regression method, elastic net integrates a linear regression model with Lasso and Ridge regularization. Elastic net is particularly useful when there are many more predictors than samples, serving to further exclude genes that are only correlated with, but not most predictive, of recurrence. We applied this two-step ensemble classification

scheme to the cross-validated training data for the outer loop of the 900, 11-fold training cross-validation data points.

Random Forest With 446 genes as an initial set of features and 40 samples from the cross-validation used as training data, a random forest model was fit. Initially, random forest was run without feature selection to determine the importance of all 446 genes based on various metrics in the RandomForest R package.[109] Next, an independent run was started that incorporated feature selection into random forest by sequentially reducing a certain number of predictors, ranked by variable importance, by employing a nested cross-validation procedure. In our simulation, a leave-one-out strategy was used. In each internal cross-validation, we removed (step=30%) the least important genes/features, ranked by variable importance, from the last cross-validation iteration. Next, we drew (Ntree=50k) bootstrap samples from the original data (40 samples). For each of the bootstrap samples, we generated an untrimmed classification/regression tree with randomly selected (mtry=22) genes from the pool of genes leftover following removal. Following cross-validation, we selected the number of genes that resulted in the lowest error rate among all the cross-validation runs. This number represents the number of genes (N_{RF}) selected by random forest after cross-validation. Next, we selected the top N_{RF} genes based on the averaged gene rank from the initial run without feature selection, yielding our final gene selections by random forest. The selected genes G_{RF} were used as input for an elastic net model in order to identify the constituents of a gene signature predictive of melanoma recurrence.

Elastic Net A powerful variable selector and classification/regression method, elastic net integrates a linear regression model with Lasso and Ridge regularization.[110] Elastic net is particularly useful when there are many more predictors than samples, serving to further exclude genes that are only correlated with, but not most predictive of

recurrence. In each round of 11-fold cross-validation on the training data, there were 40 training samples. The number of genes selected by random forest varied from 50 to 446 depending on both the leave-one-out training data and randomized feature selection used during decision tree growth. Elastic net extends the basic form of linear/logistic regression via L1 and L2-regularization. λ controls the model complexity with higher values resulting in a less complex model (less number of genes). α controls the balance between two types of model complexity penalties, including the ridge-regression penalty (α =0) and the lasso penalty (α=1). The Ridge penalty is particularly useful when there are more genes and fewer samples. Ridge regression is known to shrink the coefficients of correlated predictors towards each other. In contrast, lasso tends to pick one out of an entire set and ignore the rest. In our study, we set α=0.2 and used an internal leave-one-out cross-validation to select an optimal value of λ.[111] Our input training data was a subset of the original training data based on the gene lists G_{RF} determined by random forest. The output gene lists by elastic net with non-zero coefficients is our final gene list G_{EN} for the cross-validation run.

NanoString

Gene Expression Analysis 446 candidate genes were selected based on a PubMed literature search using the reference terms: melanoma, biomarker, immune, and gene signature [112-135]. The nCounter® platform (NanoString Technologies, Seattle, WA), was used to quantify relative gene expression in a multiplex reaction. A custom CodeSet, designated MtSinai0511, was synthesized by NanoString for the 446 selected genes as well as 17 housekeeping genes and 14 controls in a 477-plex reaction (listed in the supplemental reference file). Hybridizations were carried out according to the supplier protocols.[108] In a total reaction volume of 30µl, 100ng of each RNA sample in 5µl H_2O was mixed with 10µl nCounter Reporter probes, 10µl hybridization buffer (1x hybridization buffer =

5x SSPE, 0.1% Tween-20), and 5µl of nCounter Capture probes. Hybridizations were incubated at 65°C for approximately 16-20 hours. Following hybridization, the samples were processed in a PrepStation and counted in a DigitalAnalyzer (Nanostring 5 Technologies) according to standard protocol recommended by NanoString Technologies.

Normalization of Data Calculated from the sum counts of reporters of 6 positive control RNA spikes, sample-specific 10 normalization factors were used to normalize raw mRNA counts in order to account for slight differences in assay efficiency such as hybridization, purification, and binding. Concentrations of the control RNA spikes range from 0.125-128fM. Normalization for sample RNA quantity and quality 15 differences were applied to the spike-normalized values using sample-specific normalization factors calculated from the geometric mean of the counts from reporters targeting reference genes: ABCF1, ACTB, ALAS1, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPLP0, SDHA, and TUBB. The resulting 20 normalized counts were used in downstream analyses. For the validation set data was run in two batches, with some common samples between them (technical replicates). In order to account for batch effect, the ratio of probe counts between each of the replicate samples tested in both codesets 25 (excluding the 22nd replicate (M87) which was removed at the beginning of this process due to normalization flag). Standard ratios were then calculated for each probe and then used to adjust the two data sets so they could be analyzed together.

30 **RESULTS**

Patient populations

To test the hypothesis that the immune system limits progression of completely resected localized melanoma, mRNA transcripts for immune genes were measured using NanoString 35 technology in melanoma tissues of patients with documented clinical follow up. All tumor tissue was from the initial primary lesion and no patient received any treatment for

melanoma prior to tissue harvesting. Patients were scored as "progressors" if they presented with unresectable and/or systemic (stage IV) disease. Patients were scored as "non-progressors," if they remained free of melanoma with a minimum follow up of 24 months. Based on these criteria, an initial test set of 47 patients with completely resected stage II-III primary melanoma was identified for whom sufficient tissue and clinical follow up was available. RNA of sufficient quality for NanoString analysis was obtained in 40 of these cases (85%). A validation test set of 57 patients was identified using identical criteria and RNA was successfully extracted from 48 melanomas (84%). Thus out of a total of 104 patients identified, RNA could be extracted from 88 for an overall success rate of 85%. The 16 patients from whom RNA could not be extracted had significantly thinner melanomas (p= 0.024) and a lower mitotic rate (p=0.0067) and were not significantly different in terms of any other clinical characteristics (Table 15).

Clinical characteristics of the two test populations are shown in Table 11. There was no statistically significant difference between demographics of the training and validation populations with the exception of mitosis that were higher in the training population (p=0.002, and all others p>0.05 without correcting for multiple comparisons). 52% of patients in the training cohort and 44% in the validation cohort progressed. Univariate logistic regression showed that ulceration (p=0.003), depth (p=0.005), and age (p=0.016) associated significantly with progression. TILs, mitotic rate, location of the primary tumor and stage of disease (II A-C vs III A-C) did not significantly associate with progression.[112] Death rates were 43% and 36% in each test population, generally consistent with expected death rates based on AJCC staging over the follow up time (median 61 months in test set 1 and 45 months in test set 2).[112]

Table 11. Clinical Characteristics of Patients with Primary Melanoma in Test Set 1 and Test Set 2.

Characteristic	Test Set 1 (N=40)	Test Set 2 (N=48)	P Value
Sex			
Male -- no.(%)	28 (70)	26 (54)	0.187
Female -- no.(%)	12 (30)	22 (46)	
Age			
Median (range) --no.	67 (29-87)	65 (27-90)	0.531
Location of Tumor			
Trunk -- no.(%)	24 (60)	25 (52)	0.521
Extremity -- no.(%)	16 (40)	23 (48)	
Pathological characteristics			
Depth (mm) -- median (range)	2.65 (1.2-13)	3.47 (1-30)	0.179
Ulceration			
Absent -- no.(%)	21 (52)	20 (42)	0.392
Present -- no.(%)	19 (48)	28 (58)	
Tumor-infiltrating lymphocytes[+]			
Absent -- no.(%)	7 (17)	0 (0)	0.071
Non-brisk -- no. (%)	29 (73)	24 (89)	
Brisk -- no.(%)	4 (10)	3 (11)	
Mitoses -- median (range)	6.5 (0-26)	3 (0-20)	0.00204
Stage			
II -- no. (%)	12 (30)	25 (52)	0.051
III -- no. (%)	28 (70)	23 (48)	
Patient Outcome (months)			
Disease Progression			
Yes -- no.(%)	21 (52)	22 (46)	0.669
No -- no.(%)	19 (48)	26 (54)	
Time to Recurrence -- median(range)	14 (2-72)	20 (2-130)	0.885
Died from Melanoma -- no.(%)	17 (43)	18 (38)	0.667
Time to death -- median (range)	19 (6-82)	42 (25-160)	0.036
Time to censoring -- median(range)	61 (27-130)	47 (31-160)	0.159

[+]Tumor-infiltrating lymphocytes assessed for 27 set 2 patients

Definition of a 53 immune gene panel predictive of melanoma progression based on the training population

446 genes of interest were identified based on a pubmed search of the literature using the search terms “gene signature,” “inflammatory,” “immune,” “melanoma,” and “biomarker.” The list of 446 genes is provided in the supplement, along with a list of these genes. The starting 446 genes from the training experiment were ranked and filtered based on prediction power of melanoma progression in the training cohort using two standard classifiers, random forest and elastic net. A subset of 53 genes was selected as final gene panel (Fig 12A). ROC curves were generated and the area under the curve (AUC) was calculated on the training data (Fig 13B). A heat map clustered according to expression of these 53 genes (Fig. 12C) shows that these genes differentiate between patients who progress and those who do not. Furthermore, all 53 genes were up-regulated in the non-progressors as shown, a distribution which was significantly not random (p<0.0001).

Next, the ability of the gene signature to predict progression was evaluated in the context of known clinic-pathologic predictors. Within the training population depth (p=0.022) and age (p-0.014) significantly correlated with progression by logistic regression, while there was a strong trend for ulceration (p=0.053). Mitotic rate, TILs, gender, stage, and location of the primary tumor did not significantly correlate with progression. Multivariable logistic regression showed that gene signature score alone was the best predictive mocel of progression (p<0.001) and that clinico-pathologic features did not enhance the gene signature.

Survival analysis was then performed on the training set. Cox proportional hazards showed that the gene signature also correlated with prolonged survival (p<0.001). Multivariable cox proportional hazards analysis showed that the best model to predict survival included age and gene signature (p<0.001). Thus, the immune signature correlated strongly with

progression and survival in the context of clinic-pathologic predictors in the training patient population.

Validation of the immune gene signature of melanoma progression in a second independent cohort.

Next, experiments were conducted that sought to validate the above findings in an independent set of tissue samples. Demographics for this group are shown in Table 11. To test whether immune-associated genes in the panel were reproducibly up-regulated in tumor that did not progress, the 53 genes included in the signature were measured (Fig 12A). Notably, the proposed 53-gene signature was able to predict progression in the validation cohort with an AUC of 0.787 ($p<0.001$, Figure 13B). Cross-validation using a leave-4-out approach to rule out the possibility that individual samples were biasing the final result demonstrated that this signature was statistically robust (Fig 12B). Heat map of expression of these 53 genes in the training set confirms that these genes discriminate between progressing and non-progressing patients, with notable higher expression levels seen again in non-progressors (Fig 12C).

When the gene signature was evaluated in the validation test set in the context of clinic-pathologic predictors, it was noted that, within the validation population univariate logistic analysis showed that depth ($p=0.044$) and ulceration ($p=0.013$) correlated with progression. Multivariable logistic regression showed that the best model predictive of progression included gene signature and ulceration ($p<0.0001$).

The gene signature was then examined in terms of survival in the validation cohort. The gene signature correlated with survival by cox proportional hazards ($p=0.037$). Multivariable analysis showed that the best model to predict survival within the validation included gene signature and ulceration ($p=0.028$). Ulceration and an unfavorable immune signature identified a population at high risk of death with median

survival of 49 months as compared to 139 months in patients with one or none of these risk factors (Fig 15, p=0.044). Thus, the immune gene signature enhances the ability of established clinical-pathologic features to predict progression and survival in a second independent patient population.

Validation of expression data at the protein level and identification of CD2 as an immunohistochemical (IHC) marker of favorable prognosis

In order to validate mRNA data obtained by NanoString, CD2 staining, IHC was performed for top genes for which antibodies in clinical use were readily available. Results were concordant with NanoString results as determined by linear regression for CD2, the most differentially expressed gene between the patients who progressed and those who did not. (r=0.799; Fig. 14C). Tumors from the training cohort were also stained for CD4 and CD5, and findings correlated with the NanoString data, validating expression of these genes at the protein level (r=0.543 and r=0.666; Fig. 14D and 14E respectively). Thus, immunohistochemistry correlated with the mRNA results from NanoString.

CD2 was the most differentially expressed gene between the tumors that progressed and those that did not within the training cohort (p=0.002). Low number of CD2 positive staining by IHC correlated with melanoma progression in the second independent population (p<0.001; Fig. 14B). Thus, the NanoString analysis allowed for the identification of a novel IHC stain that may be predictive of progression in patients with completely resected stage II/III melanoma.

Physical interaction network analysis results.

Next, experiments were conducted that sought to determine whether there were any factors distinguishing the final 53 genes from the original 446 candidates (Table 13). To analyze the density of physical interactions among the 53-gene panel

relative to the original 446-gene panel, gene/protein physical interaction networks were constructed using VisAnt[133,134] (see methods). Fig. 16 is a visualization (using the software Cytoscape[137]) of the gene networks induced by the 53-gene panel (16A), and the the original 446 genes (16B). Descriptive statistics across each network (e.g. size, density, average local clustering coefficient, global clustering coefficient) are listed below (16C). Interestingly, the density of the networks is higher with the smaller 53-gene panel network, indicating that the induced subgraph (from the 53-gene panel) is proportionally more connected. That is, genes within the 53-gene panel network are interacting at a greater level than the genes in the broader 446-gene panel network. There is a 4.81 density fold change of the 53-gene panel network to the 446-gene panel network. Importantly, the P-values associated with average local CC was significant for the 53-gene panel network but not for the 446-gene panel network (Figure 5). Therefore, a significant difference in the connectivity was observed when the 446-gene panel was refined to the 53-gene panel of predictive genes.

Table 12a Top 10 enriched KEGG and GO terms (using DAVID) in the 53-gene module relative to the whole genome.

Category	Term	P Value	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006955~immune response	2.34E-57	5.987716412	5.76E-54	5.76E-54	4.12E-54
GOTERM_BP_FAT	GO:0002684~positive regulation of immune system process	1.48E-42	9.524596866	3.65E-39	1.82E-39	2.61E-39
GOTERM_BP_FAT	GO:0048584~positive regulation of response to stimulus	9.87E-33	8.210994045	2.43E-29	8.11E-30	1.74E-29
GOTERM_BP_FAT	GO:0045321~leukocyte activation	4.12E-31	7.856332365	1.02E-27	2.54E-28	7.26E-28
GOTERM_BP_FAT	GO:0050778~positive regulation of immune response	1.39E-30	10.59041938	3.43E-27	6.86E-28	2.45E-27
GOTERM_BP_FAT	GO:0001775~cell activation	2.75E-30	7.006686129	6.79E-27	1.13E-27	4.85E-27
GOTERM_BP_FAT	GO:0046649~lymphocyte activation	6.60E-29	8.451555073	1.63E-25	2.32E-26	1.16E-25
KEGG_PATHWAY	hsa04650:Natural killer cell mediated cytotoxicity	1.68E-28	7.031371532	1.78E-26	1.78E-26	1.88E-25
GOTERM_BP_FAT	GO:0042110~T cell activation	2.44E-28	11.02668383	6.01E-25	7.51E-26	4.29E-25
KEGG_PATHWAY	hsa04660:T cell receptor signaling pathway	1.91E-27	7.757024266	2.02E-25	1.01E-25	2.14E-24

Table 12b. Top 10 enriched KEGG and GO terms (using DAVID) in the 446-gene module relative to the whole genome.

Module	Category	Term	P Value	Fold Enrichment	Bonferroni	Benjamini	FDR
446 gene	GOTERM_BP_FAT	GO:0006955~immune response	2.92E-104	3.130862049	1.44E-100	1.44E-100	5.55E-101
	GOTERM_BP_FAT	GO:0010941~regulation of cell death	3.14E-85	2.726186103	1.55E-81	7.73E-82	5.96E-82
	GOTERM_BP_FAT	GO:0043067~regulation of programmed cell death	4.09E-85	2.728678557	2.01E-81	6.71E-82	7.77E-82
	KEGG_PATHWAY	hsa04060:Cytokine-cytokine receptor interaction	1.05E-84	3.192464986	1.72E-82	1.72E-82	1.28E-81
	GOTERM_BP_FAT	GO:0042981~regulation of apoptosis	3.02E-83	2.717554176	1.49E-79	3.72E-80	5.74E-80
	GOTERM_BP_FAT	GO:0006952~defense response	1.35E-74	2.892202076	6.63E-71	1.33E-71	2.56E-71
	GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	1.87E-73	2.542309068	9.19E-70	1.53E-70	3.55E-70
	GOTERM_BP_FAT	GO:0009611~response to wounding	1.49E-69	2.984442117	7.36E-66	1.05E-66	2.84E-66
	GOTERM_BP_FAT	GO:0007243~protein kinase cascade	4.28E-67	3.393394486	2.11E-63	2.63E-64	8.13E-64
	GOTERM_BP_FAT	GO:0002684~positive regulation of immune system process	1.50E-66	4.06002401	7.40E-63	8.22E-64	2.86E-63

Table 12c. Top 10 enriched KEGG and GO terms (using DAVID) in the 758 gene module relative to the whole genome.

Category	Term	P Value	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006955~immune response	2.92E-104	3.130862049	1.44E-100	1.44E-100	5.55E-101
GOTERM_BP_FAT	GO:0010941~regulation of cell death	3.14E-85	2.726186103	1.55E-81	7.73E-82	5.96E-82
GOTERM_BP_FAT	GO:0043067~regulation of programmed cell death	4.09E-85	2.728678557	2.01E-81	6.71E-82	7.77E-82
KEGG_PATHWAY	hsa04060:Cytokine-cytokine receptor interaction	1.05E-84	3.192464986	1.72E-82	1.72E-82	1.28E-81
GOTERM_BP_FAT	GO:0042981~regulation of apoptosis	3.02E-83	2.717554176	1.49E-79	3.72E-80	5.74E-80
GOTERM_BP_FAT	GO:0006952~defense response	1.35E-74	2.892202076	6.63E-71	1.33E-71	2.56E-71
GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	1.87E-73	2.542309068	9.19E-70	1.53E-70	3.55E-70
GOTERM_BP_FAT	GO:0009611~response to wounding	1.49E-69	2.984442117	7.36E-66	1.05E-66	2.84E-66
GOTERM_BP_FAT	GO:0007243~protein kinase cascade	4.28E-67	3.393394486	2.11E-63	2.63E-64	8.13E-64
GOTERM_BP_FAT	GO:0002684~positive regulation of immune system process	1.50324E-66	4.06002401	7.40198E-63	8.22442E-64	2.85615E-63

5 **Co-expression network analysis**

In order to further assess the applicability of the findings herein to patients diagnosed with primary melanoma, a co-expression network, consisting of 16,745 genes, (Fig. 17) was reconstructed using the 46 samples of gene expression data in
10 primary melanoma patients (GEO accession ID: GSE15605)[26]. A 758-gene module (highlighted in yellow in Fig. 17) was found to be the most enriched for both the 53-gene panel and 446-gene panel. For the 53-gene panel, there was an enrichment fold of 13.75 with a p-value of 1.985e-31. An enrichment fold

was similarly computed for the same module against the 446-gene panel, yielding an enrichment fold of 7.03 with p-value of 3.99e-80. The enrichment fold increased almost two times in the more refined set of genes, which indicates a higher correlation among the selected 53 genes than the original 446 genes. This data shows that the 53 gene panel is closely related to a module of genes with immune function discovered though unbiased network analysis of publicly available data from primary melanoma tumor samples.

Physical interaction network and co-expression network pathway enrichment analyses

Next, experiments were conducted that sought to determine which functional pathways were enriched in our 53-gene panel. The gene lists generated by both the 53 and 446 gene networks were annotated with Pathway and GO molecular function. The top 10 most significant enriched pathways or GO terms are shown in Tables 12a and 12b, for the 446-gene panel network genes and the 53-gene panel network genes, respectively. Interestingly, the smaller network surrounding the 53 genes shows a higher enrichment of biological processes that characterize lymphocyte fundction and immune-surveillance. Moreover, the enrichment fold change (Table 12a) in the top enriched terms for the 53- gene panel network ranges from 5 to 11 fold whereas the enrichment fold change of the top 10 terms corresponding to the 446-gene panel network (Table 12b) ranges from just 2 to 4 fold. Therefore, there a higher functional enrichment was observed in the network induced by the 53-gene panel.

Finally, experiments were conducted that sought to determine whether the module identified in publicly available samples on GEO, correlated well functionally with our proposed 53-gene signature. The functional pathways enriched by the yellow module derived from the GEO model (Fig. 17) are listed in Table 12c. The top 10 terms are listed in Table 12c. Immune processes enriched for include T cell and NK cell related

functions. These findings show that a module enriched for immune processes known to be implicated in immune surveillance is identified both in two independent melanoma patient populations of matched stage and also in publicly available
5 primary melanoma data from GEO. These experiments find that, in the two populations for which clinical follow up is available, including a training set and a test set, higher expression of this immune surveillance module associates with non-progression.
10

Table 13. Official names of all 446 genes selected for NanoString analysis in training cohort with housekeeping genes (17) highlighted in grey

A2M	CCR8	CSF1	HLAA	IKBKG	IRF1	MRC1	TIA1
ABCF1	CCR9	CSF1R	HLAB	IKZF1	IRF2	MSR1	TICAM1
ACTB	CCRL1	CSF2	HLAC	IKZF5	IRF3	MST1R	TICAM2
ALAS1	CCRL2	CSF2RA	HLA-DMA	IL10	IRF4	MX1	TIMP3
ALCAM	CD101	CSF2RB	HLA-DOB	IL10RA	IRF5	MYADM	TIRAP
ALOX5	CD14	CSF3	HLA-DPA1	IL11	IRF6	MYD88	TLR1
AMICA1	CD163	CSF3R	HLA-DPB1	IL12A	IRF7	NFAM1	TLR10
ANGPTL4	CD164	CTLA4	HLA-DQA1	IL12B	IRF8	NFATC3	TLR2
ANXA1	CD180	CTSS	HLA-DQA2	IL12RB1	IRF9	NFKB1	TLR3
ANXA11	CD19	CX3CL1	HLA-DQB1	IL12RB2	ISG15	NFKB2	TLR4
B2M	CD1A	CX3CR1	HLA-DRB1	IL13	ISG20	NFKBIA	TLR5
BCL10	CD1B	CXCL1	HLA-DRB3	IL13RA1	ITGA1	NFKBIZ	TLR6
BCL2A1	CD1C	CXCL10	HLA-DRB4	IL13RA2	ITGA2	NLRC3	TLR7
BCL3	CD1D	CXCL11	HLAE	IL15	ITGA4	NLRC5	TLR8
BCL6	CD2	CXCL12	HLAF	IL15RA	ITGA5	NOS2A	TLR9
BDCA3	CD20	CXCL13	HLAG	IL17D	ITGA6	OAS1	TNF
BIK	CD200	CXCL14	HMGB1	IL17F	ITGA9	OPTN	TNFAIP3
BIRC5	CD207	CXCL16	HPRT1	IL17RA	ITGAL	OSM	TNFAIP6
CXCR5	CD209	CXCL2	ICAM1	IL17RB	ITGAM	PDCD1	TNFRSF10B

BTK	CD24	CXCL3	ICOS	IL18	ITGB1	PDL1	TNFRSF11A
C1QA	CD27	CXCL5	ICOSLG	IL18RAP	ITGB2	PGK1	TNFRSF11B
C3	CD36	CXCL6	IFI27	IL19	ITGB3	PILRA	TNFRSF12A
C3AR1	CD37	CXCL7	IFI35	IL1A	ITGB4	PILRB	TNFRSF13B
CASP1	CD38	CXCL9	IFI44	IL1B	ITK	PLCG2	TNFRSF13C
CCBP2	CD3E	CXCR3	IFI6	IL1F10	JAK1	POLR1B	TNFRSF14
CCL1	CD4	CXCR4	IFIH1	IL1R1	KCNIP2	POLR2A	TNFRSF17
CCL11	CD40	CXCR6	IFIT1	IL1R2	KLF6	PPARG	TNFRSF18
CCL13	CD40LG	CYBB	IFIT2	IL1RAP	KLRD1	PRG1	TNFRSF19L
CCL14	CD47	CYFIP2	IFITM1	IL1RAPL2	KLRK1	PTGS2	TNFRSF1A
CCL15	CD48	DUSP1	IFITM2	IL1RL1	LAMP1	PTPRC	TNFRSF21
CCL16	CD5	DUSP5	IFNA1	IL1RL2	LAMP2	REL	TNFRSF25
CCL17	CD53	EHD1	IFNA14	IL1RN	LAMP3	RELA	TNFRSF4
CCL18	CD55	ENG	IFNA2	IL2	LAT2	RELB	TNFRSF8
CCL19	CD58	F13A1	IFNA21	IL21	LAX1	RIPK2	TNFRSF9
CCL2	CD63	FAS	IFNA4	IL22RA1	LCK	RPL19	TNFSF10
CCL20	CD68	FCAMR	IFNA5	IL23A	LDHA	RPLP0	TNFSF11
CCL21	CD70	FCER1A	IFNA6	IL23R	LEPR	Runx1	TNFSF12
CCL22	CD74	FCER2	IFNA8	IL24	LGMN	S100A12	TNFSF13
CCL23	CD79A	FCGR1A	IFNAR1	IL25	LRP1	SAA1	TNFSF13B
CCL24	CD79B	FCGR2A	IFNAR2	IL27	LSP1	SDHA	TNFSF14
CCL25	CD80	FCGR3A	IFNE1	IL28A	LTA	SERPINB2	TNFSF18
CCL26	CD83	FCGR3B	IFNG	IL28RA	LTB	SIGIRR	TNFSF4
CCL27	CD86	FCGRT	IFNGR1	IL2R	LTBR	SIGLEC1	TNFSF9
CCL28	CD8A	FLT3	IFNGR2	IL33	LY9	SKAP1	TRAF1
CCL3	CDC42	FN1	IFNK	IL34	LY96	SOCS1	TRAF2
CCL3L1	CEACAM1	Foxp3	IFRG28	IL37	LYVE1	SP110	TRAF3
CCL4	CEBPA	FPR1	IGCL2	IL3RA	MAL2	SPP1	TRAF6
CCL5	CHST4	FYN	IGF1R	IL4	MALT1	STAT1	TRAT1

CCL7	CISH3	G6PD	IGHA1	IL4R	MAP3K7	STAT2	TSLP
CCL8	CKLF	GAPDH	IGHG1	IL5	MAPK1	STAT3	TUBB
CCR1	CLEC2A	GATA3	IGHG2	IL6	MCAM	SYK	TXK
CCR10	CLEC4C	GBP1	IGHG3	IL6R	MDK	TAP1	VCAM1
CCR2	CLECL1	GBP2	IGHG4	IL7	MERTK	TAP2	VEGFC
CCR3	CLTC	GHR	IGHM	IL8	MFGE8	TARP	XCL1
CCR4	CMKLR1	GPR44	IGKC	CXCR2	MGLL	TBP	XCL2
CCR5	COLEC12	GUSB	IGLL1	INHBA	MIF	TBX21	XCR1
CCR6	CREB1	GZMK	IGSF4	IRAK1	MITF	TCL1A	ZAP70
CCR7	CRP	HIF1A	IGSF9	IRAK2	MMP9	THBS1	

Table 14. **Gene expression samples in primary melanoma patients from GEO (GSE15605)**

GEO sample	Phenotype
GSM390224	Primary_melanoma MEL101
GSM390225	Primary_melanoma MEL128
GSM390226	Primary_melanoma MEL131
GSM390227	Primary_melanoma MEL135
GSM390228	Primary_melanoma MEL142
GSM390229	Primary_melanoma MEL145
GSM390230	Primary_melanoma MEL157
GSM390231	Primary_melanoma MEL173
GSM390232	Primary_melanoma MEL176
GSM390233	Primary_melanoma MEL185
GSM390234	Primary_melanoma MEL190
GSM390235	Primary_melanoma MEL197
GSM390236	Primary_melanoma MEL209
GSM390237	Primary_melanoma MEL213
GSM390238	Primary_melanoma MEL233
GSM390239	Primary_melanoma MEL236
GSM390240	Primary_melanoma MEL243
GSM390241	Primary_melanoma MEL244
GSM390242	Primary_melanoma MEL250
GSM390243	Primary_melanoma MEL257
GSM390244	Primary_melanoma MEL258
GSM390245	Primary_melanoma MEL272
GSM390246	Primary_melanoma MEL275
GSM390247	Primary_melanoma MEL276
GSM390248	Primary_melanoma MEL280
GSM390249	Primary_melanoma MEL282

GSM390250	Primary_melanoma MEL283
GSM390251	Primary_melanoma MEL287
GSM390252	Primary_melanoma MEL290
GSM390253	Primary_melanoma MEL294
GSM390254	Primary_melanoma MEL298
GSM390255	Primary_melanoma MEL307
GSM390256	Primary_melanoma MEL310
GSM390257	Primary_melanoma MEL326
GSM390258	Primary_melanoma MEL339
GSM390259	Primary_melanoma MEL340
GSM390260	Primary_melanoma MEL356
GSM390261	Primary_melanoma MEL362
GSM390262	Primary_melanoma MEL364
GSM390263	Primary_melanoma MEL375
GSM390264	Primary_melanoma MEL380
GSM390265	Primary_melanoma MEL385
GSM390266	Primary_melanoma MEL395
GSM390267	Primary_melanoma MEL420
GSM390268	Primary_melanoma MEL429
GSM390269	Primary_melanoma MEL430

Table 15. Clinical Characteristics of Patients with Primary Melanoma (Extracted Vs. Non-Extracted)

Characteristic	Extracted (N=88)	Unextracted (N=16)	P Value
Sex			
Male -- no.(%)	54 (61)	11 (69)	0.78
Female -- no.(%)	34 (39	5 (31)	
Age *			
Median (range) --no.	66 (27-90)	71.5 (46-77)	0.645
Location of Tumor			
Trunk -- no.(%)	49 (56)	12 (75)	0.177
Extremity -- no.(%)	39 (44)	4 (25)	
Pathological characteristics **			
Depth (mm) -- median (range)	3 (1-30)	2.45 (1-8)	0.024
Ulceration			
Absent -- no.(%)	41 (47)	10 (62)	0.285
Present -- no.(%)	47 (53)	6 (38)	
Mitoses -- median (range)	4 (0-26)	1 (0-10)	0.0067
Stage			
II -- no. (%)	37 (42)	9 (56)	0.413
III -- no. (%)	51 (58)	7 (44)	
Patient Outcome			
Development of Recurrent Disease			
Yes -- no.(%)	43 (49)	4 (25)	0.103
No -- no.(%)	45 (51)	12 (75)	

* Age assessed for 8 patients in non-extracted set

** Mitosis assessed for 86 patients in extracted set and 14 patients in non-extracted set

EXAMPLE 5.

DEFINING A KEY 9 GENE SUBNETWORK PREDICTIVE OF MELANOMA PROGRESSION

In order to define key genes predictive of melanoma progression the 53 gene network was further refined to a smaller network. This smaller network was then validated in a larger validation set including patients with completely resected stage I-III melanoma and known progression status. Thus, the difference between this validation set and the earlier one for the 53 gene set is that patients with completely resected stage I disease which were provided to us by NYU and Geisinger Health Systems were also tested to test whether the more compact panel would also be applicable to stage I patients. The 9 genes resulting from this algorithm include: CD2, KLRK1, IFNAR1, HLAE, ITK, LCK, CD4, LGMN, IFI27.

ROC curves showing the predictive accuracy of this panel in the training and validation set are shown in Figure 18 for the training set (A) and the validation set (B).

Statistical Methods:

In order to refine the 53 gene panel a cross-validation procedure was performed. We simulated 900 iterations of an 11-fold cross-validation on the training dataset with random sample reordering in each iteration to strengthen the robustness of our final classifier model. Going from the top of the list to the bottom of the training cohort, every 4 samples were removed. These sample sets were then used as training data to fit a statistical model. 10,000 model training tasks were performed. The trained model and gene predictors selected were recorded in each task yielding 10,000 models and 10,000 lists of gene predictors based on randomly sub-sampled training samples. For each model, a classification was performed for the entire training (40 samples) and validation datasets (70 samples).

To derive a robust list of gene hits by these models, the 10,000 gene lists were pooled and the statistical count of each gene (out of 446 genes) was selected by these models using the training cohort only. A higher count value for a given gene indicates that it is frequently selected as a predictor during the cross-validation process. Finally, all genes selected at least once in the 10,000 cross-validation were put into a final model training task to yield an optimal, compact predictor gene list of 9 genes.

Ensemble classification/regression method

A two-step sequential ensemble classification scheme that sequentially concatenated two widely applied classifiers: random forest and elastic net, was employed. Random forest itself is an ensemble classifier consisting of many decision trees that generates the mode of individual classes yielded by independent trees. A powerful variable selector and classification/regression method, elastic net integrates a linear regression model with Lasso and Ridge regularization. Elastic net is particularly useful when there are many more predictors than samples, serving to further exclude genes that are only correlated with, but not most predictive, of recurrence. This two-step ensemble classification scheme was applied to the cross-validated training data for the outer loop of the 900, 11-fold training cross-validation data points.

Random Forest

With 446 genes as an initial set of features and 40 samples from the cross-validation used as training data, a random forest model was fit. Initially, random forest was run without feature selection to determine the importance of all 446 genes based on various metrics in the RandomForest R package.[2] Next, an independent run was started that incorporated feature selection into random forest by sequentially reducing a certain number of predictors, ranked by variable importance, by employing a nested cross-validation procedure. In the simulation, a leave-one-out strategy was used. In each

internal cross-validation, we removed (step=30%) the least important genes/features, ranked by variable importance, from the last cross-validation iteration. Next, we drew (Ntree=50k) bootstrap samples from the original data (40 samples). For each of the bootstrap samples, an untrimmed classification/regression tree with randomly selected (mtry=22) genes was generated from the pool of genes leftover following removal. Following cross-validation, the number of genes that resulted in the lowest error rate among all the cross-validation runs was selected. This number represents the number of genes (N_{RF}) selected by random forest after cross-validation. Next, we selected the top N_{RF} genes based on the averaged gene rank from the initial run without feature selection, yielding the final gene selections by random forest. The selected genes G_{RF} were used as input for an elastic net model in order to identify the constituents of a gene signature predictive of melanoma recurrence.

Elastic Net

A powerful variable selector and classification/regression method, elastic net integrates a linear regression model with Lasso and Ridge regularization.[3] Elastic net is particularly useful when there are many more predictors than samples, serving to further exclude genes that are only correlated with, but not most predictive of recurrence. In each round of 11-fold cross-validation on the training data, there were 40 training samples. The number of genes selected by random forest varied from 50 to 446 depending on both the leave-one-out training data and randomized feature selection used during decision tree growth. Elastic net extends the basic form of linear/logistic regression via L1 and L2-regularization. λ controls the model complexity with higher values resulting in a less complex model (less number of genes). α controls the balance between two types of model complexity penalties, including the ridge-regression penalty (α =0) and the lasso penalty (α=1). The Ridge penalty is particularly useful when there are more genes and fewer samples. Ridge regression is

known to shrink the coefficients of correlated predictors towards each other. In contrast, lasso tends to pick one out of an entire set and ignore the rest. In our study, we set $\alpha=0.2$ and used an internal leave-one-out cross-validation to select an optimal value of λ^4. The input training data was a subset of the original training data based on the gene lists G_{RF} determined by random forest. The output gene lists by elastic net with non-zero coefficients is our final gene list G_{EN} for the cross-validation run.

Final Gene List Generation

Employing the two-step ensemble classification method outlined above, the gene list G_{EN} was recorded for each of the 11-fold data cross-validation runs. After more than 900 runs, we collected 10,000 lists of final genes were cfrom the cross-validation training data. We counted the number of times each gene was selected among the 10,000 lists and calculated the p-value for the count distribution against otherwise random selection. Since this combined gene list compressed 10k lists in the cross-validation based on different subsampled training data, it may contain correlated genes from different runs. Therefore, to filter these out and obtain the final gene signature, elastic net was used again with the same parameter configurations outlined previously ($\alpha=0.2$, λ retrained based on the cross-validation of the 53-gene subset training data) for all 44 training samples. This yielded the 9 genes composing our gene signature.

DEFINING A KEY 4 GENE SUBNETWORK PREDICTIVE OF MELANOMA PROGRESSION

A 4 gene subnetwork with predictive value was identified in the two independent populations. These genes are: CD2, KLRK1, HLAE, and ITK and AUC curves are shown in Figures 18 C and D for the training and validation sets respectively.

EXAMPLE 6.

DEFINING THE ROLE OF CD2 IN DISEASE PROGRESSION AND OVERALL SURVIVAL AMONG PATIENTS WITH COMPLETELY RESECTED STAGE II-III CUTANEOUS MELANOMA

5

ABSTRACT

Background: Accurate assessment of prognosis remains clinically challenging in stage II-III cutaneous melanoma. Studies have implicated CD2 in immune surveillance, T-cell activation and
10 anti-tumor immunity, but its role in melanoma progression warrants further investigation.

Objective: To investigate the prognostic role of CD2 in primary cutaneous melanoma

15

Methods: Patients with American Joint Committee on Cancer Stage II and III cutaneous melanoma were identified by retrospective review of dermatopathology databases from 2001-2010 at Mount Sinai Medical Center and Geisinger Medical Center. Additional
20 patients were provided by New York University Medical Center based on tissue availability. Immunohistochemistry was performed on tumors from 90 patients with known recurrence status and documented follow-up.

25 Results: Primary tumors from patients who developed recurrent disease had fewer CD2-positive cells (p=0.0003). In multivariable analyses including standard clinicopathologic predictors, CD2 was an independent predictor of disease recurrence (p=0.008) and overall survival (p=0.007). CD2 count
30 correlated with characterization of tumor infiltrating lymphocytes (TILs) (p=0.0004). Among the intermediate prognosis group of patients with non-brisk TILs, CD2 count was predictive of disease recurrence (p=0.0006) and overall survival (p=0.0318).

35

Limitations: The retrospective design of this study may have resulted in incomplete representation of patients lacking documented follow-up.

5 Conclusions: CD2 may be an independent predictor of disease recurrence and overall survival among patients with primary cutaneous melanoma.

ABBREVIATIONS AND ACRONYMS:

10 AJCC – American Joint Committee on Cancer
IHC – Immunohistochemistry
TILs- Tumor-infiltrating lymphocytes
Ig – Immunoglobulin
NK – Natural killer
15 IRB – institutional review board
HPF – high-powered fields

CAPSULE SUMMARY:

Melanoma is the deadliest form of skin cancer and
20 accurate prognostication remains clinically challenging
CD2 immunohistochemistry correlates with non-recurrence and improved overall survival among patients with Stage II-III cutaneous melanoma
CD2 immunohistochemistry may inform clinician and patient
25 decision-making regarding treatment and follow-up

INTRODUCTION

Melanoma is a devastating disease that is steadily increasing in incidence. An estimated 1 million survivors are living in
30 the United States, and an additional 76,250 individuals will be diagnosed in 2012 alone[80]. The current American Joint Committee on Cancer (AJCC) staging system, revised in 2009, emphasizes the importance of tumor thickness, ulceration and mitotic rate in predicting prognosis. Further, the presence of
35 a single melanoma cell by immunohistochemistry (IHC) within a lymph node is considered stage III disease[81]. While survival generally correlates with AJCC staging, the subgroup of stage

II and stage III patients has a variable prognosis, with 5-year survival rates ranging from 24%-70%[81,82]. Further, while surgical resection is often curative, patients who develop recurrent disease are at a high-risk for unfavorable outcomes. Though rates vary widely with depth of the primary lesion, some studies report that up to one-third of all patients treated for primary cutaneous melanoma will experience disease recurrence[83-85]. Thus there is a need for additional biomarkers capable of enhancing prognostication and guiding clinical follow-up for these high-risk patients.

For many years, studies have demonstrated the critical role of the host immune system in the development or progression of melanoma. Interestingly, the evolving interaction between the immune system and tumor cells can result in elimination of malignant cells – a phenomenon known as immune-surveillance[86]. When elimination is incomplete, the tumor and immune cell microenvironment enter a state of equilibrium, where tumor growth is controlled by the immune system. Ultimately, however, continuous selection pressure by infiltrating immune cells can lead to an escape phase, where tumor cells are unrecognized by the immune system or induce an immunosuppressive microenvironment[86]. The presence of tumor-infiltrating lymphocytes has been associated with decreased lymphatic spread and improved survival in patients with cutaneous melanoma though studies have demonstrated inconsistencies in TIL characterization, limiting its use in clinical practice[87,88]. Additional immune based markers might add to the predictive value of TILs.

CD2 is a member of the immunoglobulin (Ig) superfamily shown to be present on T-cells, thymocytes and natural-killer (NK) cells. Various studies have implicated CD2 in immune surveillance, T-cell activation and anti-tumor immunity[89-91]. CD2 is expressed at much higher levels on activated and memory T cells than on naïve T cells, and it binds LFA3 expressed by antigen presenting cells[97,98]. The interaction between CD2 and

LFA3 enhances IL-2 production in response to antigen stimulation[92,93]. CD2 polymorphisms have been associated with systemic auto-immunity, and, more recently gene expression studies in stage II-IV melanoma have implicated CD2, among other immune genes, in prognostic molecular signatures[94-96]. CD2 has also been implicated in the adhesion of T cells to their target including tumor cells[97,98]. Examples herein examine the role of CD2 as a biomarker in cutaneous melanoma and its ability to enhance the predictive power of infiltrating lymphocyte characterization.

RESULTS

Patients

Unstained charged slides were obtained from 90 patients with Stage II or Stage III primary melanoma with documented follow-up and available tissue. These included 25 patients from MSMC, 39 patients from GMS and 26 patients from NYU. Patient demographics are shown in Table 16. The majority of patients were male (59 patients) and elderly, with a median age of 69 years (range 27-90). The median depth of primary melanoma was 3mm (range 1.2-30). 59 patients presented with American Joint Committee on Cancer (AJCC) Stage II disease, while 31 presented with AJCC stage III disease. Patient and tumor characteristics were consistent across the three contributing institutions with the exception of greater depth in the NYU cases and shorter follow-up in the non-recurrent MSMC cases (Table 19).

Table 16. Clinicopathologic Characteristics of 90 Patients with Cutaneous Melanoma

Characteristic	Value
Age (years), median (range)	69 (27-90)
Gender--no. (%)	
Male	59 (66)
Female	31 (34)
Site of Primary Lesion--no. (%)	
Axial	52 (58)
Extremity	38 (42)
Thickness (mm)	
Mean	4.19
Median (range)	3 (1.2-30)
Ulceration--no. (%)	
Yes	51 (57)
No	39 (43)
Stage, no. (%)	
Stage IIA	27 (30)
Stage IIB	18 (20)
Stage IIC	14 (16)
Stage III	31 (34)
Sentinel Lymph Node (SLN) Status--no. (%)	
Positive	31 (34)
Negative	59 (66)
Number of mitoses--median (range)	4 (0-30)
Development of Recurrent Disease--no. (%)	
Yes	45 (50)
No	45 (50)

5 Half of all patients experienced disease recurrence (45 patients). Known clinicopathologic predictors of poor prognosis, including tumor depth, patient age and AJCC Stage, significantly differed between recurrent and non-recurrent patients (Table 17). Among patients developing recurrent disease, the median time to 10 recurrence was 13.5 months (range 2-72 months). Non-recurrent patients remaining disease-free had a median follow-up of 54 months (range 26-132 months).

Table 17. Comparison of Clincopathologic Characteristics of Recurrent and Non-recurrent Patients

Characteristic	Recurrent (n=45)	Non-recurrent (n=45)	P-value
Age (years)--median (range)	71 (28-87)	64.5 (27-90)	0.0281
Gender--no. (%)			1.0000
Male	30 (67)	29 (64)	
Female	15 (33)	16 (36)	
Site of Primary Lesion--no. (%)			0.8312
Axial	27 (60)	25 (56)	
Extremity	18 (40)	20 (44)	
Thickness (mm)			0.0085
Mean	5.25	3.12	
Median (range)	3.5 (1.4-30)	2.6 (1.2-8.5)	
Ulceration-- no. (%)			0.0882
Yes	30 (67)	21 (47)	
No	15 (33)	24 (53)	
Stage--no. (%)			0.0246
Stage IIA	7 (16)	20 (44)	
Stage IIB	10 (22)	8 (18)	
Stage IIC	8 (18)	6 (13)	
Stage III	20 (44)	11 (24)	
Sentinel Lymph Node Status--no. (%)			0.0751
Positive	20 (44)	11 (24)	
Negative	25 (56)	34 (76)	
Number of mitoses--median (range)	4 (0-30)	4 (0-17)	
Time to recurrence (months)--median (range)	13.5 (2-72)	------	
Follow-up for non-recurrence (months)--median (range)	------	54(26-132)	

5

CD2 Immunohistochemistry

Quantification of the number of CD2 positive cells per HPF was reproducible with good inter rater reliability (r=0.880, p<0.0001). The number of CD2 positive cells was significantly increased in primary melanomas from non-recurrent patients compared with recurrent patients (Figure 19A+B, p=0.0003). Among recurrent patients, the median number of CD2 positive cells was 75.6 cells per HPF, compared with 37.5 cells per HPF for patients who subsequently developed recurrent disease. The number of CD2 positive cells per HPF was inversely correlated with SLN positivity, and positively correlated with improved overall survival (p=0.004 and p=0.003, respectively). CD2 did not significantly vary within Stage II substages (p=0.569), however there was a decrease in CD2 expression in stage III patients compared with stage II patients (Figure 19C, p=0.0039). Additionally, using the median CD2 count to stratify the cohort into two groups, the high-CD2 group demonstrated superior overall survival (p=0.0065; Figure 19D). After the inclusion of clinicopathologic predictors, the CD2 count was found to be an independent predictor of disease recurrence and overall survival (p=0.008 and p=0.007, respectively; Table 18). In a stepwise Cox Proportional Hazards analysis, the CD2 count significantly improved the ability of clinicopathologic variables to predict overall survival (p=0.004).

Table 18. Predictors of Disease Recurrence and Overall Survival

	Predictors of Disease recurrence		Predictors of Overall Survival	
	Univariate	Multivariate	Univariate	Multivariate
Gender	0.767	0.954	0.567	0.939
Age	0.028	0.020	0.014	0.011
Site of Disease	0.762	0.912	0.097	0.204
Depth	0.007	0.020	0.037	0.391
Ulceration	0.048	0.274	0.086	0.286
Number of mitoses	0.141	0.164	0.024	0.036
SLN	0.036	0.128	0.505	0.993
CD2 count	0.001	0.008	0.003	0.007

CD2 Expression on T-cells and Natural Killer (NK)-cells

Serial sections stained with antibodies for CD2 and CD3, a pan T-cell marker, demonstrated similar patterns of membrane staining within primary melanoma tumors (Figure 20A). However, when quantified in the same manner, CD2 immunohistochemistry was a superior predictor of overall survival compared with CD3 immunohistochemistry (Figure 23). Further, sections stained with antibodies against CD2 and CD16, a marker expressed by NK-cells, demonstrated dissimilar patterns of membrane staining but with slight overlap (Figure 20B).

Immunoflourescent co-staining with antibodies against CD4, present on the subset of helper T-cells, and CD8, present of the subset of cytotoxic T-cells, demonstrated CD2 expression by both cell types (Figure 21A). Further, serial sections demonstrated similar patterns of CD4, CD8 and CD2 staining by immunohistochemistry (Figure 21B). To assess whether CD2 surface expression correlated with a change in the distribution of T-cell subtype, the ratio of CD4 to CD8 staining was compared in a subset of tumors with high CD2 expression and a subset with low CD2 expression. While ratios within individual tumors varied, overall the ratio of CD4 to CD8-positive T-cells was similar within tumors demonstrating high and low levels of CD2 staining ($p=0.5167$; Figure 21B).

CD2 and Tumor Infiltrating Lymphocytes (TILs)

The number of CD2-positive cells significantly correlated with the characterization of TILs as absent, non-brisk and brisk ($p=0.0004$, Figure 22B). The majority of the 70 cases evaluated (55 patients, 79%) were classified as having non-brisk TILs. Among this large group with non-brisk TILs, further classification in terms of topography (central, peripheral, both) and intensity (focal, multifocal, segmental) failed to distinguish recurrent from non-recurrent patients (Figure 24). The CD2 count did not correlate with topography or intensity of TILs. However, the CD2 count remained predictive of disease

recurrence and overall survival among patients with non-brisk TILs (p=0.0006 and p=0.0318, respectively; Figure 22C-D).

5

10

15

20

25

30

Table 19. Comparison of Clincopathologic Characteristics Among Contributing Institutions

Characteristic	GMS (n=39)	MSSM (n=25)	NYU (n=26)	P-value
Age (years), median (range)	66.5 (29-86)	71 (27-87)	67 (28-90)	0.3230
Gender--no. (%)				
Male	27 (69)	16 (64)	16 (62)	0.8002
Female	12 (31)	9 (36)	10 (38)	
Site of Primary Lesion--no. (%)				0.4217
Axial	25 (64)	12 (48)	14 (54)	
Extremity	14 (36)	13 (52)	12 (46)	
Thickness (mm)				0.0006*
Mean	3.07	4.16	5.89*	
Median (range)	2.6 (1.2-13)	3.25 (1.7-11)	4.15 (1.4-30)	
Ulceration--no. (%)				0.0894
Yes	17 (44)	17 (68)	17 (65)	
No	22 (56)	8 (32)	9 (35)	
Number of mitoses, median (range)	6 (0-26)*	3 (0-30)	3 (0-14)	0.0260*
Stage--no. (%)				0.1622
Stage II	21 (54)	20 (80)	17 (65)	
Stage III	18 (46)	5 (20)	9 (35)	
Time to recurrence (months), median (range)	12 (3-72)	9 (2-36)	15 (7-35)	0.3232
Follow-up for non-recurrence (months), median (range)	70 (30-132)	32 (27-71)*	55 (26-110)	0.0006*
CD2 Count (cells per HPF), median (range)	46.3 (4.13-161.9)	70.9 (7.1-171.5)	52.3 (0.63-118.5)	0.1977

MATERIALS AND METHODS

Patient Selection

A retrospective review of dermatopathology database records from 2001-2010 at The Mount Sinai Medical Center (MSMC) and Geisinger Medical Center (GMC) was conducted between July 2010 and July 2011. Patients with American Joint Committee on Cancer (AJCC) Stage II or Stage III primary melanoma were selected for possible inclusion[81]. Disease recurrence was defined as local, regionally advanced or systemic. Local recurrences were those occurring within the scar of the primary resection. Cutaneous lesions beyond the resection scar, as well as clinically palpable lymph nodes found to contain malignant melanoma were classified as regionally advanced recurrences. Non-recurrence was defined as no further evidence of melanoma following excision of the primary lesion. A minimum follow-up of two years was required for all non-recurrent patients. Additional patient samples meeting clinical criteria were contributed by New York University Medical Center in March 2012 based on tissue availability. All living patients were censored on March 31, 2012 or on the last date of documented clinical follow-up if they were lost to follow-up prior to that date.

Patient demographics, tumor histopathologic features and clinical follow-up were extracted from electronic medical records by authorized personnel at each institution following approval by the institutional review board (IRB). Due to the retrospective nature of the study, treatment and monitoring following the diagnosis of primary melanoma were dictated by each patients' dermatologist and/or oncologist. Information was obtained from physician records of these visits. MSMC patients with incomplete clinical records were contacted by mail and telephone under an IRB approved protocol by authorized personnel to obtain clinical follow-up.

Immunohistochemistry

Immunohistochemistry (IHC) was performed on 5-micron charged slides obtained from eligible patients with known clinical follow-up. Cut sections from each tumor specimen were deparaffinized in xylene, rehydrated in ethanol and stained with an anti-CD2 monoclonal antibody (pre-diluted, Ventana Medical Systems, Tucson, AZ) using the Ventana BenchMark XT immunostainer. Staining was visualized using the i-View DAB kit solutions (Ventana Medical Systems, Tucson, AZ). Each slide set included a negative control without the addition of primary antibody to confirm specificity of the stain. Once stained, each slide was evaluated twice independently by blinded investigators using an ocular micrometer with a 1mm^2 130 grid (Nikon Eclipse E400®) and the number of CD2 positive cells in 8 high-powered fields (HPF) per slide was counted. Scores for each slide were averaged to yield a single score for use in subsequent analyses. A subset of tumors with available tissue were stained and quantified in an identical manner using anti-CD3 monoclonal antibody (pre-diluted, Ventana Medical Systems, Tucson, AZ).

IHC was also performed on serial 5-micron sections cut from a subset of tumors. Serially sectioned slides were stained with anti-CD4 (pre-diluted, Ventana Medical Systems, Tucson, AZ) and anti-CD8 (pre-diluted, Ventana Medical Systems, Tucson, AZ) as described above. One adjacent section was stained with anti-CD16 monoclonal antibody (2H7, 1:20 dilution, Thermo Scientific Lab Vision, Kalamazoo, MI). Slides were deparaffinized in xylene and rehydrated in ethanol. Following antigen retrieval in 10mmol/L citrate solution, endogenous peroxidase activity was blocked with 3% hydrogen peroxide in phosphate buffered saline. Following blocking for non-specific binding, slides were incubated with anti-CD16 antibody overnight at 4°C. Staining was detected using diaminobenzine (Sigma-Aldrich, Saint Louis, MO) and slideswere counterstained with Mayer's hematoxylin. As CD16 146 is also expressed on the

surface of macrophages, membrane staining was evaluated by a dermatopathologist to ensure its expression on NK-cells.

Immunofluorescence

Charged slides from a subset of tumors were deparaffinized in xylene, rehydrated in ethanol and heated in ethylenediaminetetraacetic acid (EDTA) pH 9.0 for antigen retrieval. Slides were then co-stained with anti-CD2 (1:40 dilution, Dako, Glostrup, Denmark) and anti-CD3 (pre-diluted, Ventana Medical Systems, Tucson, AZ) monoclonal antibodies. Staining was visualized using fluorochrome-conjugated secondary antibodies. Slides were sealed using mounting media with 4',6-diamidino-2-phenylindole (DAPI). Additional slides were co-stained in the same manner using anti-CD2 (1:40 dilution) and anti-CD4 (1:100 dilution, Abcam, Cambridge, MA, USA) or anti-CD2 and anti-CD8 (1:100 dilution, Abcam, Cambridge, MA, USA).

Tumor-infiltrating lymphocyte characterization

Additional slides were stained with hematoxylin and eosin and reviewed by Dr. Robert G. Phelps, Director of the Department of Dermatopathology at MSMC. TILs were characterized as brisk, non-brisk and absent according to published criteria[99]. Brisk lymphocytic infiltration was used to describe lymphocyte density greater than 20 lymphocytes per high-powered field throughout the lesion. Non-brisk was used to describe collections of a few lymphoctyes per high-powered field. Non-brisk tumor-infiltrating lymphocytes were further characterized in terms of their topography and intensity within the tumor. TIL topography was classified as central, peripheral or both and intensity was classified as focal, multifocal or segmental. Focal infiltration was used to describe a single collection of a few lymphocytes. Multifocal describes multiple collections of a few lymphocytes. Segmental describes a large collection 170 of lymphocytes occupying a substantial portion of the vertical growth phase[100].

Statistical Analysis

Patient demographics were analyzed using the Student's t-test for continuous variables and the Fisher's exact test for categorical variables. CD2 count was analyzed as a continuous variable using the Student's t-test. Number of CD2 positive cells was correlated with disease recurrence and overall survival using univariable logistic regression and Cox proportional hazards models, respectively, in SPSS Statistical Analysis Software Package Version 20. The predictive power of CD2 was also examined in the context of known clinical predictors using multivariable logistic regression and Cox proportional hazards models. The relationship between CD2 and overall survival was also examined using the Kaplan-Meier method and log-rank test in GraphPad Prism Version 5.0. The relationship between the number of CD2 positive cells and the number of CD4 and CD8 positive cells was analyzed using Spearman correlation in GraphPad Prism Version 5.0. The relationship between the CD2 count and various TILs characterizations was analyzed using the Kruskal-Wallis test in GraphPad Prism Version 5.0.

Discussion

Patients with history of melanoma are at risk of recurrence and death. Current predictors used in staging do not accurately assess prognosis for individuals.[1] Although the host immune system may modulate melanoma progression, no evidence-based immune biomarkers are in clinical use.[2]

Technological developments now allow for the analysis of partially degraded RNA, facilitating analysis of FFPE melanoma. Through the use of molecular "barcodes," NanoString technology quantifies gene expression based on individual mRNA transcripts with a high level of precision and sensitivity.[3, 4]

The role of the immune system in melanoma progression is complex. Although studies of intransit metastasis (advanced stage III) have suggested that inflammation plays a protective role in disease progression, it is well established that melanoma is able to co-opt the immune system by recruiting regulatory T cells as well as myeloid cell types that promote angiogenesis.[18-22] Tumor infiltrating lymphocytes (TILs) have been reported to favorably impact prognosis in primary melanoma, particularly if the deepest portion of the tumor was studied.[23-26] TIL quantification, however, is not routinely performed due to variability in pathologic interpretation as well as a lack of consistent correlation with outcome.[27-29]

Although the immune system has been proposed to limit melanoma progression, the exact mechanisms and clinical relevance of immune activation remain to be elucidated.[30] The studies described herein sought to determine whether expression patterns of immune-associated genes correlate with tumor recurrence in patients with previously excised localized melanoma.

In order to define a biomarker for melanoma recurrence, the expression of immune-related genes from formalin-fixed, paraffin-embedded (FFPE) primary melanoma was measured using NanoString, a hybridization assay uniquely suited for

quantifying gene expression in samples with partially degraded RNA.[3, 4]

This work represents the first time an immune biomarker based on gene expression measured in FFPE tissues has been proposed for early-stage malignancy in humans. FFPE tissues are easier to obtain than frozen specimens, which require careful intra-operative processing. Using NanoString technology, a 21-immune gene signature was established which was predictive of recurrence with greater accuracy than any currently established predictor.

FFPE tissue and corresponding clinical information were obtained in 44 consecutive patients who had complete resection of primary melanoma. mRNA transcripts of 446 genes were measured. Immunohistochemistry (IHC) was used to assess protein expression of the most differentially expressed gene, CD2. Findings were validated in an independent cohort of 37 patients.

Expression profiling yielded a panel of 21 immune genes predictive of melanoma recurrence using receiver operating characteristic (ROC) curves. This panel was validated in an independent patient cohort (AUC=0.794) and correlated with improved overall survival ($p<0.001$). CD2 expression correlated with non-recurrence ($p=0.017$).

The role of the immune system in tumor development is complex, with evidence to support both protective and harmful roles for a variety of immune cell types.[32, 33] Nonetheless, there is convincing pre-clinical and clinical data that tumors undergo immune-editing whereby immunogenic clones are controlled and/or eliminated.[34, 35] The role of tumor immunosurveillance in humans is best established in colon cancer where lymphocytic infiltrates and expression of genes implicated in adaptive immunity correlate closely with prognosis.[36, 37] In melanoma, evidence for immunosurveillance was found in a previous study of cutaneous metastasis in patients with stage III disease.[20]

In the study of primary tumors described herein, also led to the discovery that higher expression of immune genes is associated with lower risk of recurrence and, although genes known to have regulatory functions were tested, these genes did not confer a poor prognosis. The data discussed herein strongly implicates a protective role for the immune system during melanoma development.

Immune gene expression showed higher correlation with recurrence and death than standard clinical predictors. Notably, pathological features characterizing melanoma are subject to observer bias.[29] In the studies described herein, the presence or absence of immune infiltrate was not predictive of prognosis in either patient cohort, but more closely correlated with the institution where the melanoma was examined. In contrast, the proposed gene signature provides an objective, more accurate indicator for risk of disease recurrence than available clinical and pathologic predictors.

Due to the small volume of specimen available from primary melanoma, it was not feasible to micro-dissect stroma from tumor. Therefore, in thestudies described herein, genes may be expressed by the tumor itself and/or by stromal cells. However, CD2 is expressed by T, NK, and NKT lymphocytes, and increased CD2 levels likely indicate increased anti-melanoma immunity.[38, 39] Genes identified in the panel are implicated in cutaneous T cell trafficking and activation, interferon signaling, antigen presentation, and natural killer activity. Thus, without wishing to be bound by any scientific theory, it is speculated that the observed patterns of gene expression between patients with recurrent and non-recurrent disease are the result of an interaction between the tumor and host immune system that plays a determinative role in tumor progression. Findings presented herein lead to the development of a biomarker that informs clinicians as to which patients warrant close observation and follow up, and allow for improved stratification of patients in clinical trials.

A 21-immune gene signature is predictive of recurrence and mortality in early-stage melanoma. These data provide evidence that the immune system limits progression of early-stage melanoma and demonstrate that expression analysis of FFPE specimens yield prognostic information. A novel way to identify early-stage melanoma patients at high risk of recurrence and death is established herein.

Defining the role of CD2 in disease progression and overall survival among patients with completely resected stage II-III cutaneous melanoma.

The patients in Example 6 represent a clinically high-risk primary melanoma population. Half of the population experienced some form of disease recurrence, which is consistent with estimates for patients with AJCC Stage III disease.[81,101] Standard predictors, including depth and patient age, correlated with disease recurrence in our population as has been previously shown.

The data of Example 6 demonstrated that CD2 expression levels within the primary tumor associate with non-recurrence and prolonged survival in 90 patients with stage II-III cutaneous melanoma. CD2 has been implicated in melanoma specific anti-tumor immunity in the past. Altomante et al suggested that the interaction between immune cells and malignant cells of the melanocytic lineage was mediated in part by CD2 expressing lymphocytes[102]. More recently, Bogunovic et al described an immune response gene expression signature, including CD2, predictive of improved survival in metastatic melanoma15. Our study expands upon these findings by demonstrating that increased CD2 staining in primary melanoma tissue sections correlates with a lower recurrence rate and improved overall survival.

Characterization of TILs is complicated by heterogeneity in the composition and location of these cells. Tumor-infiltrating lymphocytes may be "helper" T-cells (CD4 positive), "cytotoxic" T cells (CD8 positive), NK-cells, B-cells or the counterproductive regulatory T-cell, which limits the host immune response[103]. The results in Example 6 demonstrate CD2 expression predominantly on T-cells, both CD4-positive and CD8-positive, but also on NK-cells. Thus, CD2 may be a specific indicator of activated or anti-tumor infiltrating lymphocytes. Additionally, TIL characterization is inconsistently reported and subject to inter-rater variability. Monshizadeh et al examined concordance between referring pathologists and pathologists reviewing cases for the Western Australia Melanoma Advisory Service (WAMAS). They found that TILs were not reported in 51% of cases referred to WAMAS, and further found only slight agreement in those that did report TILs (52.4%, $\kappa = 0.12$)[88]. Similarly, a multi-site review of histopathology reports found that only 54.2% included information about lymphocytic infiltration[104]. CD2 immunohistochemistry may be a widely accessible, more objective method for TIL characterization.

Finally, while a brisk lymphocytic infiltrate has been shown to be strongly protective, this designation only applies to a minority of patients[87]. In the population of Example 6, tumor-infiltrating lymphocytes were characterized by the Department of Dermatopathology (RGP). More than 75% of patients were classified as having non-brisk TILs, an intermediate category that adds little to the clinician's estimation of risk. The Examples herein have demonstrated that CD2 immunohistochemistry may be useful to estimate prognosis in this large subpopulation.

Limitations

Due to the retrospective nature of the study in Example 6, reporting standards and follow-up guidelines were heterogeneous across the population. Further, due to the

requirement of a minimum of 24-months of clinical follow-up, the study did not capture patients who had not recurred but for whom limited follow up was available. The prognostic role of CD2 should be validated prospectively in an independent cohort to address these limitations.

Conclusion

Accurate assessment of prognosis at an individual patient level in Stage II and Stage III primary melanoma represents a clinical challenge. Recurrence and survival estimates range widely, and treatment options span from observation alone to enrollment in clinical trials[105]. The results provided herein demonstrate that CD2 quantification by immunohistochemistry associates with non-recurrence and prolonged survival within this group of patients. Prospective studies may define a role for CD2 immunohistochemistry that helps patients and clinicians make informed decisions regarding treatment and follow-up, and may enhance stratification for studies of adjuvant therapies.

References

1. Wolchok JD, Saenger YM. Current topics in melanoma. Curr Opin Oncol 2007;19:116-20.
2. Piris A, Mihm MC, Jr. Progress in melanoma histopathology and diagnosis. Hematol Oncol Clin North Am 2009;23:467-80, viii.
3. Fortina P, Surrey S. Digital mRNA profiling. Nat Biotechnol 2008;26:293-4.
4. Payton JE, Grieselhuber NR, Chang LW, et al. High throughput digital quantification of mRNA abundance in primary human acute myeloid leukemia samples. J Clin Invest 2009;119:1714-26.
5. Jeffrey E, Gershenwald CMB, Seng-Jaw Soung, John F Thompson. Prognostic factors and natural history of melanoma. In: Chalres M Balch ANH, Arthur J Sober, Seng-jaw Soong, Michael B Atkins, John F Thompson, ed. Cutaneous Melanoma. Fifth ed. St Louis, Missouri: Quality Medical Publishing; 2009:35-64.
6. Balch CM, Soong SJ, Gershenwald JE, et al. Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. J Clin Oncol 2001;19:3622-34.
7. Green AC, Baade P, Coory M, Aitken JF, Smithers M. Population-based 20-year survival among people diagnosed with thin melanomas in queensland, australia. J Clin Oncol 2012;30:1462-7.
8. Morton DL, Thompson JF, Cochran AJ, et al. Sentinel-node biopsy or nodal observation in melanoma. N Engl J Med 2006;355:1307-17.
9. Timar J, Gyorffy B, Raso E. Gene signature of the metastatic potential of cutaneous melanoma: too much for too little? Clin Exp Metastasis 2010;27:371-87.
10. Oakman C, Santarpia L, Di Leo A. Breast cancer assessment tools and optimizing adjuvant therapy. Nat Rev Clin Oncol 2010;7:725-32.
11. Conway C, Mitra A, Jewell R, et al. Gene expression profiling of paraffin-embedded primary melanoma using the

DASL assay identifies increased osteopontin expression as predictive of reduced relapse-free survival. Clin Cancer Res 2009;15:6939-46.
12. Winnepenninckx V, Lazar V, Michiels S, et al. Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl Cancer Inst 2006;98:472-82.
13. Smith AP, Hoek K, Becker D. Whole-genome expression profiling of the melanoma progression pathway reveals marked molecular differences between nevi/melanoma in situ and advanced-stage melanomas. Cancer Biol Ther 2005;4:1018-29.
14. Koh SS, Opel ML, Wei JP, et al. Molecular classification of melanomas and nevi using gene expression microarray signatures and formalin-fixed and paraffin-embedded tissue. Mod Pathol 2009;22:538-46.
15. Kabbarah O, Nogueira C, Feng B, et al. Integrative genome comparison of primary and metastatic melanomas. PLoS One 2010;5:e10770.
16. Segura MF, Belitskaya-Levy I, Rose AE, et al. Melanoma MicroRNA signature predicts post-recurrence survival. Clin Cancer Res 2010;16:1577-86.
17. Bittner M, Meltzer P, Chen Y, et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 2000;406:536-40.
18. Taube JM, Anders RA, Young GD, et al. Colocalization of inflammatory response with B7-h1 expression in human melanocytic lesions supports an adaptive resistance mechanism of immune escape. Sci Transl Med 2012;4:127ra37.
19. Wang E, Miller LD, Ohnmacht GA, et al. Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. Cancer Res 2002;62:3581-6.
20. Bogunovic D, O'Neill DW, Belitskaya-Levy I, et al. Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc Natl Acad Sci U S A 2009;106:20429-34.

21. Gajewski TF, Fuertes M, Spaapen R, Zheng Y, Kline J. Molecular profiling to identify relevant immune resistance mechanisms in the tumor microenvironment. Curr Opin Immunol 2011;23:286-92.
22. Drake CG, Jaffee E, Pardoll DM. Mechanisms of immune evasion by tumors. Adv Immunol 2006;90:51-81.
23. Tuthill RJ, Unger JM, Liu PY, Flaherty LE, Sondak VK. Risk assessment in localized primary cutaneous melanoma: a Southwest Oncology Group study evaluating nine factors and a test of the Clark logistic regression prediction model. Am J Clin Pathol 2002;118:504-11.
24. Piras F, Colombari R, Minerba L, et al. The predictive value of CD8, CD4, CD68, and human leukocyte antigen-D-related cells in the prognosis of cutaneous malignant melanoma with vertical growth phase. Cancer 2005;104:1246-54.
25. Day CL, Jr., Mihm MC, Jr., Sober AJ, et al. Prognostic factors for melanoma patients with lesions 0.76 - 1.69 mm in thickness. An appraisal of "thin" level IV lesions. Ann Surg 1982;195:30-4.
26. Clemente CG, Mihm MC, Jr., Bufalino R, Zurrida S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. Cancer 1996;77:1303-10.
27. Chao C, Wong SL, Ross MI, et al. Patterns of early recurrence after sentinel lymph node biopsy for melanoma. Am J Surg 2002;184:520-4; discussion 5.
28. Scheri RP, Essner R, Turner RR, Ye X, Morton DL. Isolated tumor cells in the sentinel node affect long-term prognosis of patients with melanoma. Ann Surg Oncol 2007;14:2861-6.
29. Busam KJ, Antonescu CR, Marghoob AA, et al. Histologic classification of tumor-infiltrating lymphocytes in primary cutaneous malignant melanoma. A study of interobserver agreement. Am J Clin Pathol 2001;115:856-60.
30. Parmiani G. Tumor-infiltrating T cells--friend or foe of neoplastic cells? N Engl J Med 2005;353:2640-1.

31. Geiss GK, Bumgarner RE, Birditt B, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol 2008;26:317-25.
32. Colotta F, Allavena P, Sica A, Garlanda C, Mantovani A. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. Carcinogenesis 2009;30:1073-81.
33. Allavena P, Sica A, Solinas G, Porta C, Mantovani A. The inflammatory micro-environment in tumor progression: the role of tumor-associated macrophages. Crit Rev Oncol Hematol 2008;66:1-9.
34. Chow MT, Moller A, Smyth MJ. Inflammation and immune surveillance in cancer. Semin Cancer Biol 2012;22:23-32.
35. Koebel CM, Vermi W, Swann JB, et al. Adaptive immunity maintains occult cancer in an equilibrium state. Nature 2007;450:903-7.
36. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 2006;313:1960-4.
37. Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 2012;12:298-306.
38. Altomonte M, Gloghini A, Bertola G, et al. Differential expression of cell adhesion molecules CD54/CD11a and CD58/CD2 by human melanoma cells and functional role in their interaction with cytotoxic cells. Cancer Res 1993;53:3343-8.
39. Li Y, Hellstrom KE, Newby SA, Chen L. Costimulation by CD48 and B7-1 induces immunity against poorly immunogenic tumors. J Exp Med 1996;183:639-44.
40. CD2 (MRQ-11) [package insert]. Tucson, AZ: Ventana Medical Systems; 2010.
41. Liaw A, Wiener M. Classification and regression by randomForest. R news 2002;2(3):18-22.
42. Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society 2005;SeriesB:301-320.

43. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software 2010;33:1-22.
44. Bogunovic D, O'Neill DW, Belitskaya-Levy I, et al. Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc Natl Acad Sci U S A 2009;106:20429-34.
45. Budhu A, Forgues M, Ye QH, et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell 2006;10:99-111.
46. Conway C, Mitra A, Jewell R, et al. Gene expression profiling of paraffin-embedded primary melanoma using the DASL assay identifies increased osteopontin expression as predictive of reduced relapse-free survival. Clin Cancer Res 2009;15:6939-46.
47. Desai KV, Michalowska AM, Kondaiah P, Ward JM, Shih JH, Green JE. Gene expression profiling identifies a unique androgen-mediated inflammatory/immune signature and a PTEN (phosphatase and tensin homolog deleted on chromosome 10)-mediated apoptotic response specific to the rat ventral prostate. Mol Endocrinol 2004;18:2895-907.
48. Haass NK, Smalley KS. Melanoma biomarkers: current status and utility in diagnosis, prognosis, and response to therapy. Mol Diagn Ther 2009;13:283-96.
49. Harlin H, Meng Y, Peterson AC, et al. Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment. Cancer Res 2009;69:3077-85.
50. Hoek KS. Melanoma progression, gene expression and DNA microarrays. G Ital Dermatol Venereol 2009;144:39-49.
51. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995-2004.
52. Jaeger J, Koczan D, Thiesen HJ, et al. Gene expression signatures for tumor progression, tumor subtype, and tumor thickness in laser-microdissected melanoma tissues. Clin Cancer Res 2007;13:806-15.

53. Jonsson G, Busch C, Knappskog S, et al. Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. Clin Cancer Res 2010;16:3356-67.
54. Kannengiesser C, Spatz A, Michiels S, et al. Gene expression signature associated with BRAF mutations in human primary cutaneous melanomas. Mol Oncol 2008;1:425-30.
55. Kashani-Sabet M, Venna S, Nosrati M, et al. A multimarker prognostic assay for primary cutaneous melanoma. Clin Cancer Res 2009;15:6987-92.
56. Koh SS, Opel ML, Wei JP, et al. Molecular classification of melanomas and nevi using gene expression microarray signatures and formalin-fixed and paraffin-embedded tissue. Mod Pathol 2009;22:538-46.
57. Mandruzzato S, Callegaro A, Turcatel G, et al. A gene expression signature associated with survival in metastatic melanoma. J Transl Med 2006;4:50.
58. Riker AI, Enkemann SA, Fodstad O, et al. The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. BMC Med Genomics 2008;1:13.
59. Roepman P, Jassem J, Smit EF, et al. An immune response enriched 72-gene prognostic profile for early-stage non-small-cell lung cancer. Clin Cancer Res 2009;15:284-90.
60. Ryu B, Kim DS, Deluca AM, Alani RM. Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. PLoS One 2007;2:e594.
61. Schwartzentruber DJ, Hom SS, Dadmarz R, et al. In vitro predictors of therapeutic response in melanoma patients receiving tumor-infiltrating lymphocytes and interleukin-2. J Clin Oncol 1994;12:1475-83.
62. Smith AP, Hoek K, Becker D. Whole-genome expression profiling of the melanoma progression pathway reveals marked molecular differences between nevi/melanoma in

situ and advanced-stage melanomas. Cancer Biol Ther 2005;4:1018-29.

63. Teschendorff AE, Caldas C. A robust classifier of high predictive value to identify good prognosis patients in ER-negative breast cancer. Breast Cancer Res 2008;10:R73.
64. Timar J, Gyorffy B, Raso E. Gene signature of the metastatic potential of cutaneous melanoma: too much for too little? Clin Exp Metastasis 2010;27:371-87.
65. Ugurel S, Utikal J, Becker JC. Tumor biomarkers in melanoma. Cancer Control 2009;16:219-24.
66. Wang E, Miller LD, Ohnmacht GA, et al. Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. Cancer Res 2002;62:3581-6.
67. Winnepenninckx V, Lazar V, Michiels S, et al. Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl Cancer Inst 2006;98:472-82.
68. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree S, George RD, Grogan T, James JJ, Maysuria M, Mitton JD, Oliveri P, Osborn JL, Peng T, Ratcliffe AL, Webster PJ, Davidson EH, Hood L, Dimitrov K. Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol. 2008 Mar;26(3):317-25. Epub 2008 Feb 17.
69. Balch CM. Cutaneous Melanoma. 5th edition. Quality Medical Publishing, S.L., MO; 2009.
70. Korn, E.L., et al., Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. J Clin Oncol, 2008. 26(4): p. 527-34.
71. Rosenberg, S.A., J.C. Yang, and N.P. Restifo, Cancer immunotherapy: moving beyond current vaccines. Nat Med, 2004. 10(9): p. 909-15.
72. Subramanian, J. and R. Simon, What should physicians look for in evaluating prognostic gene-expression signatures? Nat Rev Clin Oncol, 2010. 7(6): p. 327-34.

73. Kauffmann, A., et al., High expression of DNA repair pathways is associated with metastasis in melanoma patients. Oncogene, 2008. 27(5): p. 565-73.
74. Winnepenninckx, V., et al., Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl Cancer Inst, 2006. 98(7): p. 472-82.
75. Timar, J., B. Gyorffy, and E. Raso, Gene signature of the metastatic potential of cutaneous melanoma: too much for too little? Clin Exp Metastasis, 2010. 27(6): p. 371-87.
76. Koh, S.S., et al., Molecular classification of melanomas and nevi using gene expression microarray signatures and formalin-fixed and paraffin-embedded tissue. Mod Pathol, 2009. 22(4): p. 538-46.
77. Bogunovic, D., et al., Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc Natl Acad Sci U S A, 2009. 106(48): p. 20429-34.
78. Dobbin, K. and R. Simon, Sample size determination in microarray experiments for class comparison and prognostic classification. Biostatistics, 2005. 6(1): p. 27-38.
79. Lee, W.C., Detecting Differentially Expressed Genes: Minimizing Burden of Testing and Maximizing Number of Discoveries. Ann Epidemiol, 2010.
80. Siegel R, DeSantis C, Virgo K, et al. Cancer treatment and survivorship statistics. CA Cancer J Clin 2012;62:220-41.
81. Balch CM, Gershenwald JE, Soong SJ, et al. Final version of 2009 AJCC melanoma staging and classification. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009;27:6199-206.
82. Ben-Porat L, Panageas KS, Hanlon C, et al. Estimates of stage-specific survival are altered by changes in the 2002 American Joint Committee on Cancer staging system for melanoma. Cancer 2006;106:163-71.
83. Soong SJ, Shaw HM, Balch CM, McCarthy WH, Urist MM, Lee JY. Predicting survival and recurrence in localized

melanoma: a multivariate approach. World J Surg 1992;16:191-5.

84. Soong SJ, Harrison RA, McCarthy WH, Urist MM, Balch CM. Factors affecting survival following local, regional, or distant recurrence from localized melanoma. J Surg Oncol 1998;67:228-33.

85. Green AC, Baade P, Coory M, Aitken JF, Smithers M. Population-based 20-year survival among people diagnosed with thin melanomas in Queensland, Australia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:1462-7.

86. Aris M, Barrio MM, Mordoh J. Lessons from cancer immunoediting in cutaneous melanoma. Clin Dev Immunol 2012;2012:192719.

87. Azimi F, Scolyer RA, Rumcheva P, et al. Tumor-infiltrating lymphocyte grade is an independent predictor of sentinel lymph node status and survival in patients with cutaneous melanoma. Journal of clinical oncology: official journal of the American Society of Clinical Oncology 2012;30:2678-83.

88. Monshizadeh L, Hanikeri M, Beer TW, Heenan PJ. A critical review of melanoma pathology reports for patients referred to the Western Australian Melanoma Advisory Service. Pathology 2012;44:441-7.

89. Crawford K, Stark A, Kitchens B, et al. CD2 engagement induces dendritic cell activation: implications for immune surveillance and T-cell activation. Blood 2003;102:1745-52.

90. Tibaldi EV, Salgia R, Reinherz EL. CD2 molecules redistribute to the uropod during T cell scanning: implications for cellular activation and immune surveillance. Proc Natl Acad Sci U S A 2002;99:7582-7.

91. Sabry M, Tsirogianni M, Bakhsh IA, et al. Leukemic priming of resting NK cells is killer Ig-like receptor independent but requires CD15-mediated CD2 ligation and natural cytotoxicity receptors. J Immunol 2011;187:6227-34.

92. Bierer BE, Peterson A, Gorga JC, Herrmann SH, Burakoff SJ. Synergistic T cell activation via the physiological ligands for CD2 and the T cell receptor. J Exp Med 1988;168:1145-56.
93. Moingeon P, Chang HC, Wallner BP, Stebbins C, Frey AZ, Reinherz EL. CD2-mediated adhesion facilitates T lymphocyte antigen recognition function. Nature 1989;339:312-4.
94. Bogunovic D, O'Neill DW, 477 Belitskaya-Levy I, et al. Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc Natl Acad Sci U S A 2009;106:20429-34.
95. Mann GJ, Pupo GM, Campain AE, et al. BRAF Mutation, NRAS Mutation, and the Absence of an Immune-Related Expressed Gene Profile Predict Poor Outcome in Patients with Stage III Melanoma. The Journal of investigative dermatology 2013;133:509-17.
96. Sivendran S CR, Harcharik S, Hall L, Bernardo S, Moskalenko M, et al. Immune gene expression in primary melanomas to predict lower risk of recurrence and death [abstract]. In: 2013 ASCO Annual Meeting. Poster Discussion Session: Developmental Therapeutics - Immunotherapy; 2013 May 31- June 4. Accepted.2013.
97. Patarroyo M, Prieto J, Rincon J, et al. Leukocyte-cell adhesion: a molecular process fundamental in leukocyte physiology. Immunological reviews 1990;114:67-108.
98. Wang A, Batteux F, Wakeland EK. The role of SLAM/CD2 polymorphisms in systemic autoimmunity. Current opinion in immunology 2010;22:706-14.
99. Mihm MC, Jr., Clemente CG, Cascinelli N. Tumor infiltrating lymphocytes in lymph node melanoma metastases: a histopathologic prognostic indicator and an expression of local immune response. Lab Invest 1996;74:43-7.
100. Rao UN, Lee SJ, Luo W, Mihm MC, Jr., Kirkwood JM. Presence of tumor infiltrating lymphocytes and a dominant nodule within primary melanoma are prognostic factors for

relapse-free survival of patients with thick (t4) primary melanoma: pathologic analysis of the e1690 and e1694 intergroup trials. Am J Clin Pathol;133:646-53.

101. Leiter U, Buettner PG, Eigentler TK, et al. Hazard rates for recurrent and secondary cutaneous melanoma: an analysis of 33,384 patients in the German Central Malignant Melanoma Registry. J Am Acad Dermatol 2012;66:37-45.

102. Altomonte M, Gloghini A, Bertola G, et al. Differential expression of cell adhesion molecules CD54/CD11a and CD58/CD2 by human melanoma cells and functional role in their interaction with cytotoxic cells. Cancer Res 1993;53:3343-8.

103. Oble DA, Loewe R, Yu P, Mihm MC, Jr. Focus on TILs: prognostic significance of tumor infiltrating lymphocytes in human melanoma. Cancer Immun 2009;9:3.

104. Thompson B, Austin R, Coory M, et al. Completeness of histopathology reporting of melanoma in a high-incidence geographical region. Dermatology 2009;218:7-14.

105. Fox MC, Lao CD, Schwartz JL, Frohm ML, Bichakjian CK, Johnson TM. Management options for metastatic melanoma in the era of novel therapies: a primer for the practicing dermatologist: part I: Management of stage III disease. J Am Acad Dermatol 2013;68:1 e-9; quiz 10-2.

106. Manolio TA, Chisholm RL, Ozenberger B, et al. Implementing genomic medicine in the clinic: the future is here. Genetics in medicine : official journal of the American College of Medical Genetics 2013;15:258-67.

107. Balch CM, Gershenwald JE, Soong SJ, Thompson JF. Update on the melanoma staging system: the importance of sentinel node staging and primary tumor mitotic rate. Journal of surgical oncology 2011;104:379-85.

108. Pennock GK, Waterfield W, Wolchok JD. Patient responses to ipilimumab, a novel immunopotentiator for metastatic melanoma: how different are these from conventional treatment responses? American journal of clinical oncology 2012;35:606-11.

109. Agarwala SS. An update on pegylated IFN-alpha2b for the adjuvant treatment of melanoma. Expert review of anticancer therapy 2012;12:1449-59.
110. Eggermont AM, Suciu S, Testori A, et al. Long-term results of the randomized phase III trial EORTC 18991 of adjuvant therapy with pegylated interferon alfa-2b versus observation in resected stage III melanoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:3810-8.
111. Sondak VK, Kudchadkar R. Pegylated interferon for the adjuvant treatment of melanoma: FDA approved, but what is its role? The oncologist 2012;17:1223-4.
112. Balch CM, Gershenwald JE, Soong SJ, et al. Final version of 2009 AJCC melanoma staging and classification. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009;27:6199-206.
113. Balch CM, Soong SJ, Gershenwald JE, et al. Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2001;19:3622-34.
114. Wong SL, Balch CM, Hurley P, et al. Sentinel lymph node biopsy for melanoma: American Society of Clinical Oncology and Society of Surgical Oncology joint clinical practice guideline. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:2912-8.
115. Morton DL, Thompson JF, Cochran AJ, et al. Sentinel-node biopsy or nodal observation in melanoma. The New England journal of medicine 2006;355:1307-17.
116. Balch CM, Gershenwald JE, Soong SJ, et al. Multivariate analysis of prognostic factors among 2,313 patients with stage III melanoma: comparison of nodal micrometastases versus macrometastases. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2010;28:2452-9.

117. Bindea G, Mlecnik B, Fridman WH, Galon J. The prognostic impact of anti-cancer immune response: a novel classification of cancer patients. Semin Immunopathol 2011;33:335-40.

118. Ott PA, Carvajal RD, Pandit-Taskar N, et al. Phase I/II study of pegylated arginine deiminase (ADI-PEG 20) in patients with advanced melanoma. Investigational new drugs 2013;31:425-34.

119. Fridman WH, Galon J, Dieu-Nosjean MC, et al. Immune infiltration in human cancer: prognostic significance and disease control. Curr Top Microbiol Immunol 2011;344:1-24.

120. Ascierto PA, Capone M, Urba WJ, et al. The additional facet of immunoscore: immunoprofiling as a possible predictive tool for cancer treatment. Journal of translational medicine 2013;11:54.

121. Clemente CG, Mihm MC, Jr., Bufalino R, Zurrida S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. Cancer 1996;77:1303-10.

122. Azimi F, Scolyer RA, Rumcheva P, et al. Tumor-infiltrating lymphocyte grade is an independent predictor of sentinel lymph node status and survival in patients with cutaneous melanoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:2678-83.

123. Busam KJ, Antonescu CR, Marghoob AA, et al. Histologic classification of tumor-infiltrating lymphocytes in primary cutaneous malignant melanoma. A study of interobserver agreement. Am J Clin Pathol 2001;115:856-60.

124. Frishberg DP, Balch C, Balzer BL, et al. Protocol for the examination of specimens from patients with melanoma of the skin. Arch Pathol Lab Med 2009;133:1560-7.

125. Bogunovic D, O'Neill DW, Belitskaya-Levy I, et al. Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc Natl Acad Sci U S A 2009;106:20429-34.

126. Mann GJ, Pupo GM, Campain AE, et al. BRAF mutation, NRAS mutation, and the absence of an immune-related expressed gene profile predict poor outcome in patients with stage III melanoma. The Journal of investigative dermatology 2013;133:509-17.

127. Messina JL, Fenstermacher DA, Eschrich S, et al. 12-Chemokine gene signature identifies lymph node-like structures in melanoma: potential for patient selection for immunotherapy? Scientific reports 2012;2:765.

128. Chow MT, Moller A, Smyth MJ. Inflammation and immune surveillance in cancer. Seminars in cancer biology 2012;22:23-32.

129. Fortina P, Surrey S. Digital mRNA profiling. Nature biotechnology 2008;26:293-4.

130. Geiss GK, Bumgarner RE, Birditt B, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. Nature biotechnology 2008;26:317-

131. Raskin L, Fullen DR, Giordano TJ, Thomas DG et al. Transcriptome Profiling Identifies HMGA2 as a Biomarker of Melanoma Progression and Prognosis. J. Invest Dermatol 2013; doi: 10.1038/jid.2013.197

132. Zhang B, and Horvath S. A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical Applications in Genetics and Molecular Biology. 2005;4:(1) Article 17

133. Hu Z, et al. VisANT 4.0: Integrative network platform to connect genes, drugs, diseases and therapies. Nucl. Acids Res. 2013;41:W225-31.

134. Hu Z, Mellor J, Wu J and DeLisi C. VisANT: an online visualization and analysis tool for biological interaction data. BMC Bioinformatics. 2004;5:17.

135. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009;4(1):44-57.

136. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive

functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

137. Christmas R, Avila-Campillo I, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Am Assoc Cancer Res Educ Book. 2005;12-16.

Claims

What is claimed is:

1. A method of predicting the risk of reoccurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

    a. obtaining a RNA-containing sample of the previously removed melanoma tissue containing RNA from the patient;
    b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
    c. comparing the level of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

    wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has a reduced risk of reoccurrence of melanoma, and

    wherein the plurality of pre-selected genes comprises the following genes:

    a. CD2;
    b. KLRK1;
    c. ITK; and
    d. HLAE.

2. The method of claim 1, further comprising the step of creating a report summarizing said prediction.

3. The method of claim 1 or 2, wherein the expression level of each gene of the plurality of pre-selected genes is

normalized relative to the expression level of one or more reference genes.

4. The method of claim 3, wherein the expression level of each such gene is normalized relative to the expression level of the following genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

5. The method of claim 3 or 4, wherein the predetermined reference level of expression is the expression level of the one or more reference genes.

6. A method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by the method of any one of claims 1-5, comprising administering an immunotherapy to the patient.

7. A method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by the method of any one of claims 1-5, comprising testing the human patient for recurrence of melanoma more frequently than a corresponding patient who was determined to have a reduced risk of reoccurrence of melanoma would be tested for recurrence.

8. A method of treating a patient afflicted with melanoma which comprises the following:

   a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
   b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;

c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and

d. administering a therapy to the patient if there is a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;

b. KLRK1;

c. ITK; and

d. HLAE.

9. A method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;

b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;

c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and

d. administering a therapy to the patient if there is a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

10. A method of treating a patient afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes;
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene; and
d. administering a therapy to the patient if the level of expression of the plurality of pre-selected genes in the sample is i) lower as compared with the predetermined reference upper level of expression of such genes and ii) higher as compared with the predetermined reference lower level of expression of such genes, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

11. A method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

    a. obtaining a RNA-containing sample of melanoma tissue containing RNA from the patient;
    b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
    c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

    wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

    wherein the plurality of pre-selected genes comprises the following genes:

    a. CD2;
    b. KLRK1;
    c. ITK; and
    d. HLAE.

12. A method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

    a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;

b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

13. A method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a level of expression of the plurality of pre-selected genes in the sample is i) lower as compared with the predetermined reference upper level of expression of such genes and ii) higher as compared with the predetermined reference lower level of expression of such genes, indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

14. A method of determining whether a therapy is effective for treating patients afflicted with melanoma which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from at least one patient afflicted with melanoma, which at least one patient was administered the therapy for treatment of the melanoma previous to collection of the sample;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to the levels of expression of the plurality of pre-selected genes to the expression level of each such gene in a corresponding at least one patient not administered the therapy;

wherein a higher level of expression of the plurality of pre-selected genes in the sample of step a) as compared with the level of expression of such genes in a corresponding at least one patient not administered the treatment indicates that the therapy is effective for treating patients afflicted with melanoma, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

15. A method of determining whether a patient afflicted with melanoma and which patient was administered a therapy has exhibited a positive clinical response to the therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from the patient, which RNA-containing sample was removed from the patient who was administered the therapy;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the patient has exhibited a positive clinical response to the therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

16. A method of determining whether a treatment should be administered to patients afflicted with melanoma as an adjuvant or a neoadjuvant therapy which comprises the following:

a. obtaining a RNA-containing sample of the melanoma tissue containing RNA from at least one patient afflicted with melanoma;
b. treating the sample to determine from the RNA contained in the sample the level of expression of a plurality of preselected genes; and
c. comparing the levels of expression of each gene of the plurality of pre-selected genes to a predetermined reference level of expression for each such gene;

wherein a higher level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the treatment should be administered to patients afflicted with melanoma as a neoadjuvant therapy, and a lower level of expression of the plurality of pre-selected genes in the sample as compared with the predetermined reference level of expression of such genes indicates that the treatment should be administered to patients afflicted with melanoma as an adjuvant therapy, and

wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK; and
d. HLAE.

17. The method of any one of claims 1-16, wherein the plurality of pre-selected genes further comprises at least one of the following genes:
a. IFNAR1;
b. LCK;
c. CD4;
d. LGMN; and
e. IFI27.

18. The method of any one of claims 1-16, wherein the plurality of pre-selected genes comprises the following genes:
a. CD2;
b. KLRK1;
c. ITK;
d. HLAE;
e. IFNAR1;
f. LCK;
g. CD4;
h. LGMN; and
i. IFI27.

19. The method of claim 17 or 18, wherein the plurality of pre-selected genes further comprises at least one of the following genes:
a. CCL27;
b. SYK;
c. CD68;
d. IL18; and

f. IL1F7.

20. The method of claim 17 or 18, wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;
b. KLRK1;
c. ITK;
d. HLAE;
e. LCK;
f. IFNAR1;
g. CD48;
h. CD4;
i. CTSS;
j. CCR4;
k. HLA-DQB1;
l. TAP2;
m. LGMN;
n. CSF2RA;
o. IFNGR1;
p. CCL27;
q. SYK;
r. CD68;
s. IL18;
t. IFI27; and
u. IL1F7.

21. The method of any one of claims 1-16, wherein the plurality of pre-selected genes further comprises at least one of the following genes:

a. LCK;
b. IFNAR1;
c. CD48;
d. CD4;
e. CTSS;
f. CCR4;
g. HLA-DQB1;
h. TAP2;

i. LGMN;

j. CSF2RA; and

k. IFNGR1.

22. The method of any one of claims 1-16, wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;

b. KLRK1;

c. ITK;

d. HLAE;

e. LCK;

f. IFNAR1;

g. CD48;

h. CD4;

i. CTSS;

j. CCR4;

k. HLA-DQB1;

l. TAP2;

m. LGMN;

n. CSF2RA; and

o. IFNGR1.

23. The method of claim 21 or 22, wherein the plurality of pre-selected genes further comprises at least one of the following genes:

a. CCL27;

b. SYK;

c. CD68;

d. IL18;

e. IFI27; and

f. IL1F7.

24. The method of claim 21 or 22, wherein the plurality of pre-selected genes comprises the following genes:

a. CD2;

b. KLRK1;

c. ITK;
d. HLAE;
e. LCK;
f. IFNAR1;
g. CD48;
h. CD4;
i. CTSS;
j. CCR4;
k. HLA-DQB1;
l. TAP2;
m. LGMN;
n. CSF2RA;
o. IFNGR1;
p. CCL27;
q. SYK;
r. CD68;
s. IL18;
t. IFI27; and
u. IL1F7.

25. The method of any one of claims 1-24, wherein the plurality of pre-selected genes consists of less than about 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 1000, or 10,000 genes.

26. The method of any one of claims 8-25, wherein the sample was previously removed from the patient.

27. The method of any one of claims 1-26, wherein the sample is a fixed, wax-embedded tissue specimen.

28. The method of any one of claims 1-27, wherein the expression level is assayed by NanoString gene expression analysis.

29. The method of any one of claims 1-7 or 26-28, wherein the sample is at least one week old.

30. The method of any one of claims claim 26-29, wherein the sample is at least one month old.

31. The method of claim 30, wherein the sample is at least six months old.

32. The method of claim 31, wherein the sample is at least one year old.

33. The method of claim 32, wherein the sample is at least ten years old.

34. The method of any one of claims 1-33, wherein the RNA transcripts of the plurality of pre-selected genes in the sample are fragmented.

35. The method of any one of claims 1-34, further comprising identifying a treatment option for the patient based on the expression level of the plurality of pre-selected genes.

36. The method of any one of claims 1-35, wherein the expression level is determined by immunohistochemistry or proteomics technology.

37. The method of any one of claims 8-36, wherein the therapy is chemotherapy.

38. The method of claim 37, wherein the chemotherapy comprises dacarbazine, temozolomide, paclitaxel, cisplatin, carmustine, fotemustine, vindesine, vincristine, and bleomycin, or vemurafenib.

39. The method of any one of claims 8-36, wherein the therapy is radiation therapy.

40. The method of any one of claims 8-36, wherein the therapy is immunotherapy.

41. The method of claim 6 or 40, wherein the immunotherapy comprises an interferon (IFN).

42. The method of claim 41, wherein the immunotherapy comprises IFN-α.

43. The method of claim 41, wherein the IFN-α is IFN-α2b.

44. The method of claim 43, wherein the IFN-α2b is PEGylated IFN-α2b.

45. The method of claim 44, wherein the PEGylated IFN-α2b is Sylatron.

46. The method of claim 6 or 40, wherein the immunotherapy comprises an interleukin.

47. The method of claim 46, wherein the interleukin is IL-2.

48. The method of claim 47, wherein the IL-2 is aldesleukin.

49. The method of claim 6 or 40, wherein the immunotherapy comprises an antibody.

50. The method of claim 49, wherein the antibody is a monoclonal antibody.

51. The method of claim 50, wherein the monoclonal antibody is a humanized monoclonal antibody.

52. The method of claim 50, wherein the monoclonal antibody is a fully human monoclonal antibody.

53. The method of any one of claims 49-52, wherein the antibody is an anti-PD-1 antibody.

54. The method of claim 53, wherein the anti-PD-1 antibody is BMS-936558.

55. The method of any one of claims 50-53, wherein the antibody is an anti-CTLA-4 antibody.

56. The method of claim 55, wherein the antibody is ipilimumab.

57. The method of claim 6 or 40, wherein the immunotherapy is an oncolytic immunotherapy.

58. The method of claim 57, wherein the oncolytic immunotherapy comprises a virus.

59. The method of claim 58, wherein the virus is derived from HSV-1.

60. The method of any one of claims 57-59, wherein the oncolytic immunotherapy is a vaccine.

61. The method of claim 60, wherein the vaccine is talimogene laherparepvec (T-VEC).

62. A method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

    a. obtaining a sample of the previously removed melanoma tissue from the patient;
    b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and

c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

63. The method of claim 62, further comprising the step of creating a report summarizing said prediction.

64. The method of claim 62 or 63, wherein the expression level of the expression product of the gene or the each of two or more genes is normalized relative to the expression level of the expression product of one or more of the following genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

65. The method of claim 63 or 64, wherein the predetermined reference level of expression of the expression product is the expression level of the expression product the one or more reference genes.

66. A method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by the method of any one of claims 62-65, comprising administering an immunotherapy to the patient.

67. A method of treating a patient from whom melanoma tissue was previously removed, and which patient was determined to not have a reduced risk of reoccurrence of melanoma by the method of any one of claims 62-65, comprising testing

the human patient for recurrence of melanoma more frequently than a corresponding patient who was determined to have a reduced risk of reoccurrence of melanoma would be tested for recurrence.

68. A method of treating a patient afflicted with melanoma which comprises the following:
   a. obtaining a sample of melanoma tissue from the patient;
   b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample;
   c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene; and
   d. administering a therapy to the patient if there is a higher level of expression of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene.

69. A method of treating a patient afflicted with melanoma which comprises the following:
   a. obtaining a sample of melanoma tissue from the patient;
   b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample;
   c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene; and
   d. administering a therapy to the patient if there is a lower level of expression of the expression product of the gene or each of the two or more genes in the

sample as compared with the predetermined reference level of the expression product for each such gene.

70. A method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or the each of two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

71. A method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of the gene or each of the two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a

predetermined reference level of the expression product for each such gene,

wherein a lower level of expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

72. A method of predicting whether a patient afflicted with melanoma is likely to exhibit a positive clinical response to treatment with a therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level of the expression product for each such gene,

wherein a level of of the expression product of the gene or each of the two or more genes that is i) lower as compared with a predetermined reference upper level of expression for each such gene and ii) higher as compared with a predetermined reference lower level of expression for each such gene, indicates that the patient is likely to exhibit a positive clinical response to treatment with the therapy.

73. A method of determining whether a therapy is effective for treating patients afflicted with melanoma which comprises the following:

a. obtaining a sample of melanoma tissue from at least one patient afflicted with melanoma, which at least one patient was administered the therapy;
b. treating the sample to determine the level of an expression product of the gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to the level of the expression product of the gene or each of the two or more genes in a corresponding at least one patient not administered the treatment,

wherein a higher level of expression product of the gene or each of the two or more genes in the sample of step (a) as compared with the level of expression product of the gene or each of the two or more genes in the corresponding at least one patient not administered the treatment indicates that the therapy is effective for treating patients afflicted with melanoma.

74. A method of determining whether a patient afflicted with melanoma and which patient was administered a therapy has exhibited a positive clinical response to the therapy which comprises the following:

a. obtaining a sample of melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level for each such gene,

wherein a higher level of expression the expression product of the gene or each of the two or more genes as

compared to the predetermined reference level of the expression product of each such gene indicates that the patient has exhibited a positive clinical response to the therapy.

75. A method of determining whether a treatment should be administered to patients afflicted with melanoma as an adjuvant or a neoadjuvant therapy which comprises the following:

   a. obtaining a sample of melanoma tissue from at least one patient afflicted with melanoma;
   b. treating the sample to determine the level of an expression product of a gene or each of two or more genes in the sample; and
   c. comparing the level of the expression product of the gene or each of the two or more genes to a predetermined reference level for each such gene,

   wherein a higher level of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of expression for each such gene indicates that the treatment should be administered to patients afflicted with melanoma as a neoadjuvant therapy, and a lower level of the expression product of the gene or each of the two or more genes in the sample as compared with the predetermined reference level of the expression product for each such gene indicates that the treatment should be administered to patients afflicted with melanoma as an adjuvant therapy.

76. The method of any one of claims 62-75, wherein the level of an expression product is determined for one gene.

77. The method of claim 76, wherein the one gene is CD2.

78. The method of any one of claims 62-75, wherein each of the two or more genes comprises two, three, four, five, six, seven, nine, ten or more of CD2, KLRK1, ITK, HLAE, LCK, IFNAR1, CD48, CD4, CTSS, CCR4, HLA-DQB1, TAP2, LGMN, CSF2RA, IFNGR1, CCL27, SYK, CD68, IL18, IFI27 or IL1F7 gene.

79. The method of any one of claims 62-75 or 78, wherein the two or more genes consists of less than about 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 1000, or 10,000 genes.

80. The method of any one of claims 62-79, wherein the expression product is a protein encoded by the gene or is the expression product of each of the two or more genes.

81. A method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of a X gene in the melanoma tissue sample; and
c. comparing the level of the expression product of the X gene to a predetermined reference level of the expression product of the X gene,

wherein a higher level of expression product of the X gene in the sample as compared with the predetermined reference level of the expression product of the X gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

82. The method of claim 81, wherein the X gene is the CD2, KLRK1, ITK, HLAE, LCK, IFNAR1, CD48, CD4, CTSS, CCR4,

HLA-DQB1, TAP2, LGMN, CSF2RA, IFNGR1, CCL27, SYK, CD68, IL18, IFI27 or IL1F7 gene.

83. A method of predicting the risk of reocurrence of melanoma in a patient from whom melanoma tissue was previously removed which comprises the following:

a. obtaining a sample of the previously removed melanoma tissue from the patient;
b. treating the sample to determine the level of an expression product of the CD2 gene in the melanoma tissue sample; and
c. comparing the level of the expression product of the CD2 gene to a predetermined reference level of the expression product of the CD2 gene,

wherein a higher level of expression product of the CD2 gene in the sample as compared with the predetermined reference level of the expression product of the CD2 gene indicates that the patient has a reduced risk of reoccurrence of melanoma.

84. The method of any one of claims 1-7 or 25-83, wherein the patient was afflicted with sentinel lymph node positive melanoma when the melanoma tissue was removed.

85. The method of any one of claims 1-7 or 25-84, wherein the patient was afflicted a melanoma deeper than 2mm or deeper than 1mm and ulcerated when the melanoma tissue was removed.

86. The method of any one of claims 8-84, wherein the melanoma is sentinel lymph node positive melanoma.

87. The method of any one of claims 8-86, wherein the melanoma is deeper than 2mm or deeper than 1mm and ulcerated.

88. The method of any one of claims 81-87, wherein the expression product is a protein encoded by the X gene.

89. The method of any one of claims 1-88, wherein the level of the expression product of each gene is determined by immunohistochemistry or proteomics technology.

90. The method of claim 89, wherein the level of the expression product of each gene is determined by immunohistochemistry.

91. The method of any one of claims 1-90, wherein the melanoma tissue is stage II or III primary melanoma tissue.

92. The method of any one of claims 1-93, wherein the melanoma is stage II or III melanoma.

93. The method of any one of claims 1-61, wherein the predetermined reference level of expression is
    i) the expression level of each such gene in normal tissue; or
    ii) the expression level of one or more of the following reference genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

94. The method of any one of claims 62-92, wherein the predetermined reference level of the expression product is
    i) the level of the expression product of each such gene in normal tissue; or
    ii) the level of the expression product of one or more of the following reference genes: ABCF1, ACTB, ALAS1, CLTC, G6PD, GAPDH, GUSB, HPRT1, LDHA, PGK1, POLR1B, POLR2A, RPL19, RPLP0, SDHA, TBP and TUBB.

95. The method of claim 93 or 94, wherein the normal tissue is normal skin tissue.

Figure 1

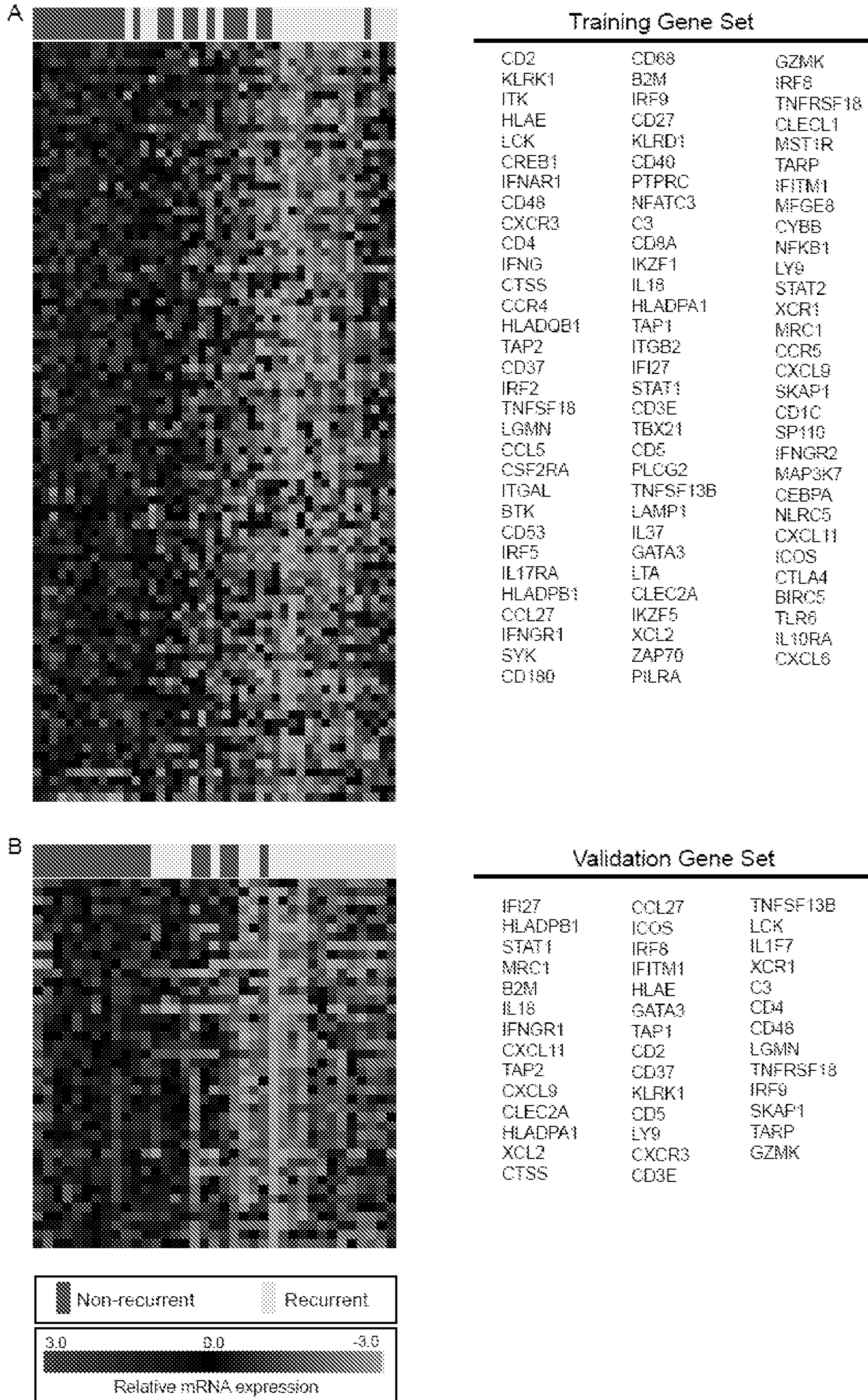


Figure 2

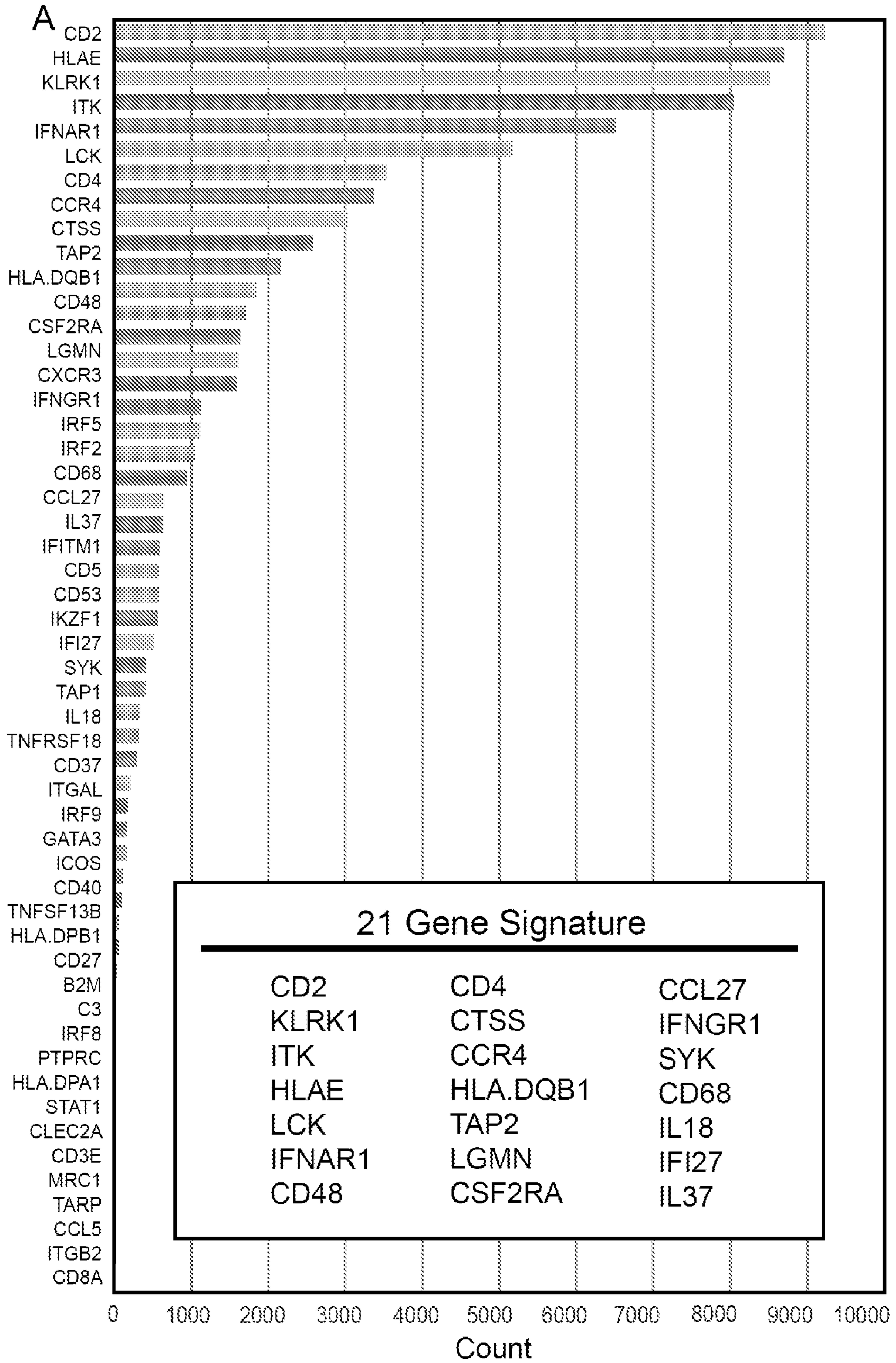


Figure 2, continued

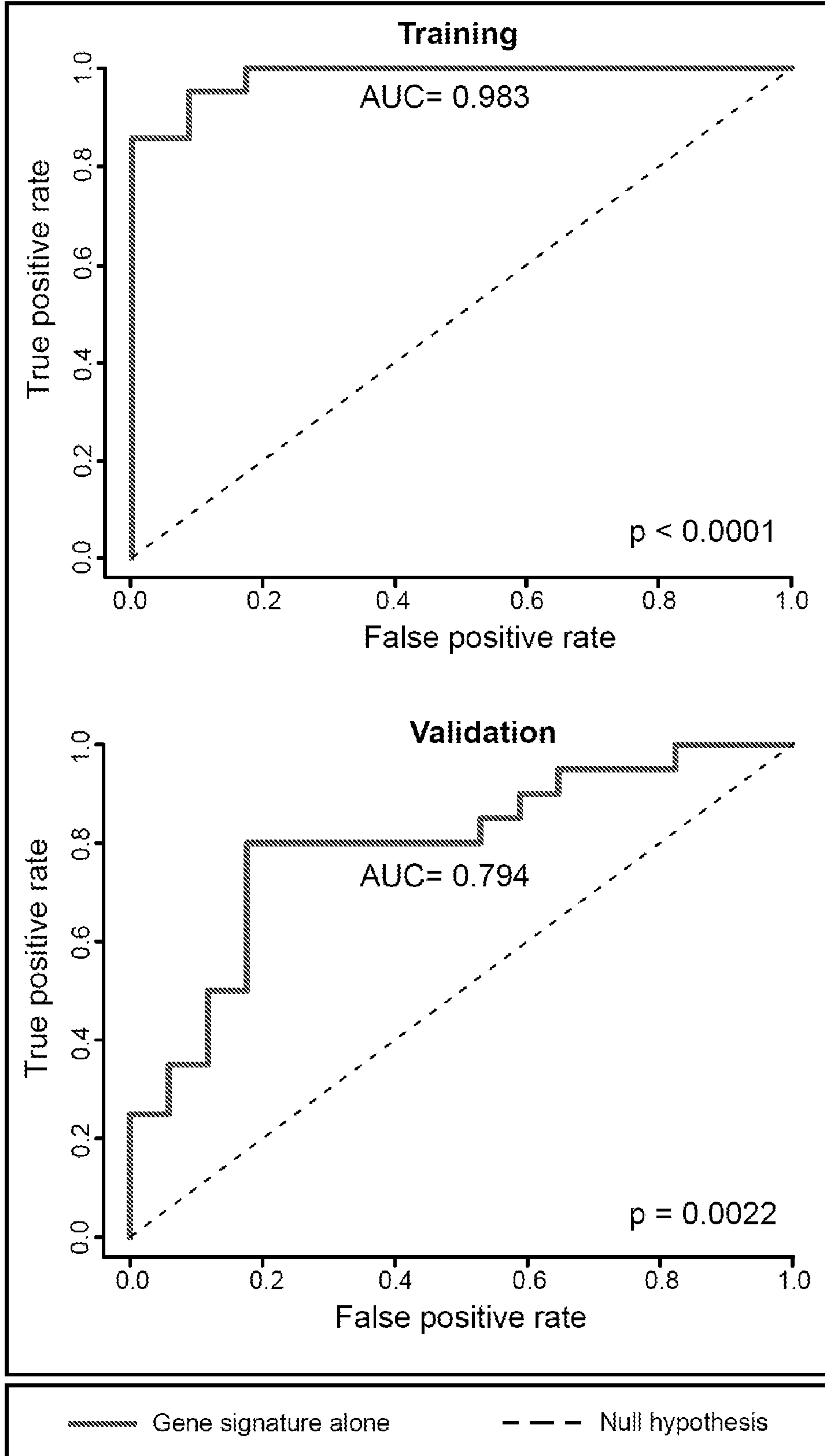


Figure 2, continued

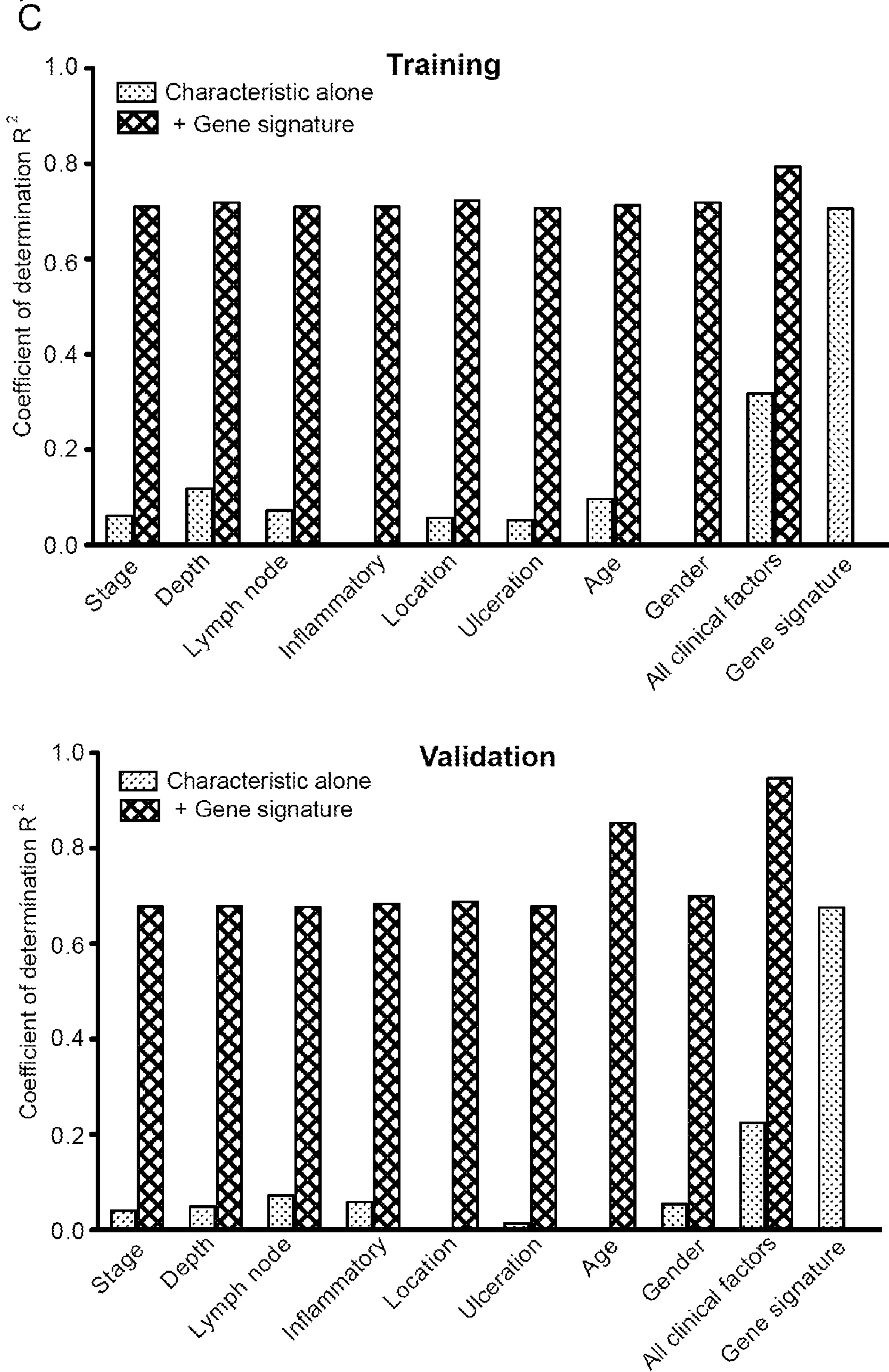


Figure 3

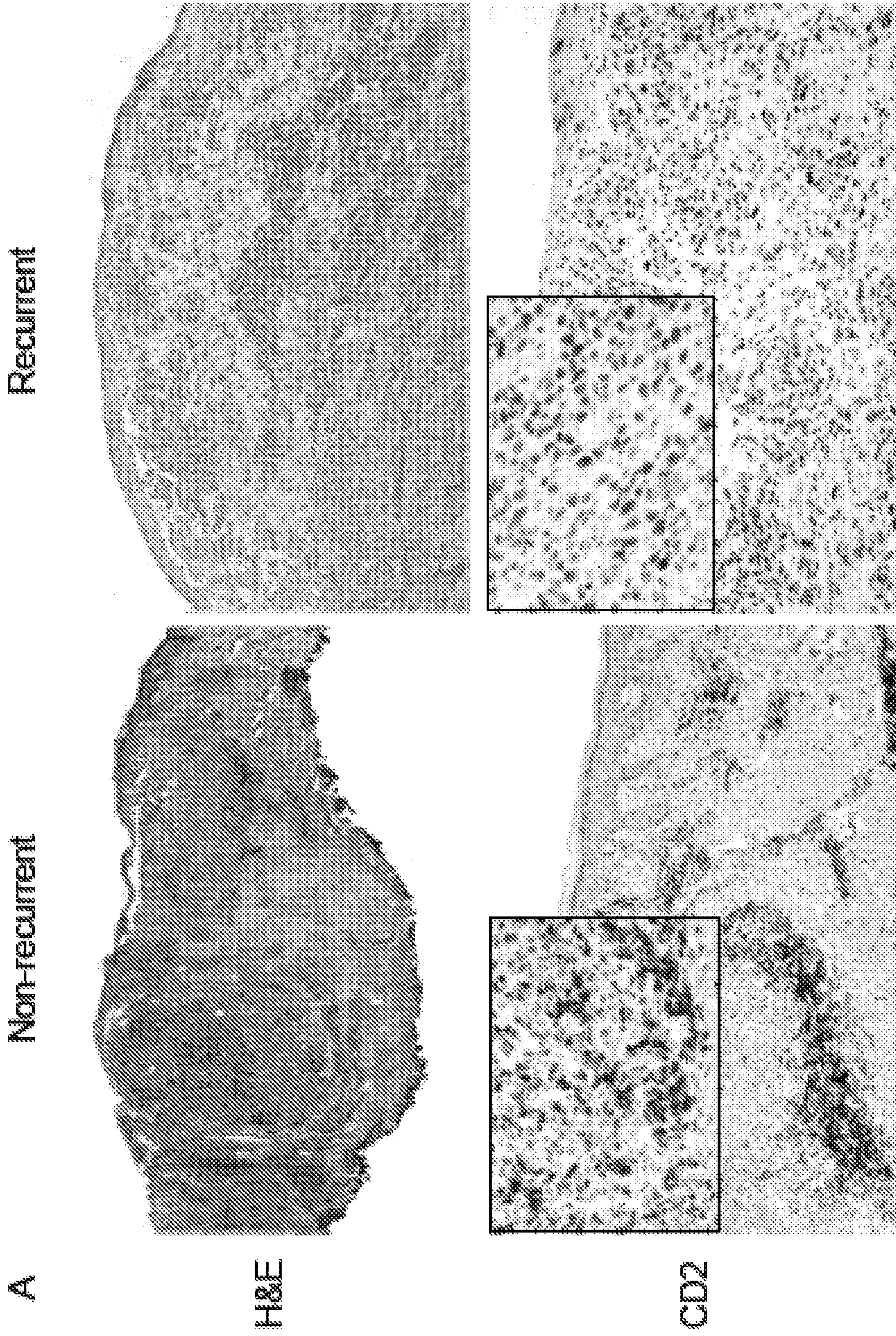


Figure 3, continued

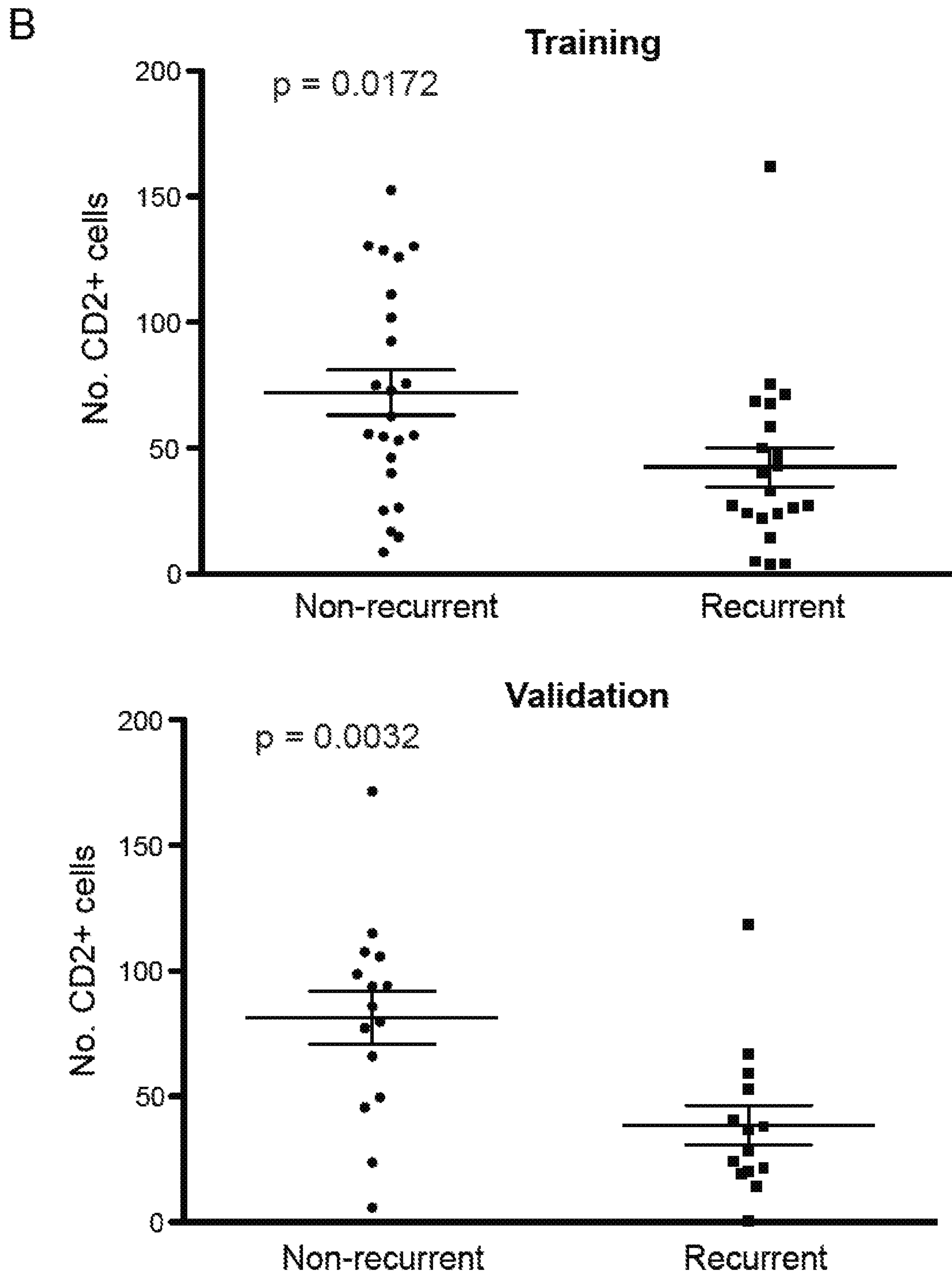


Figure 3, continued

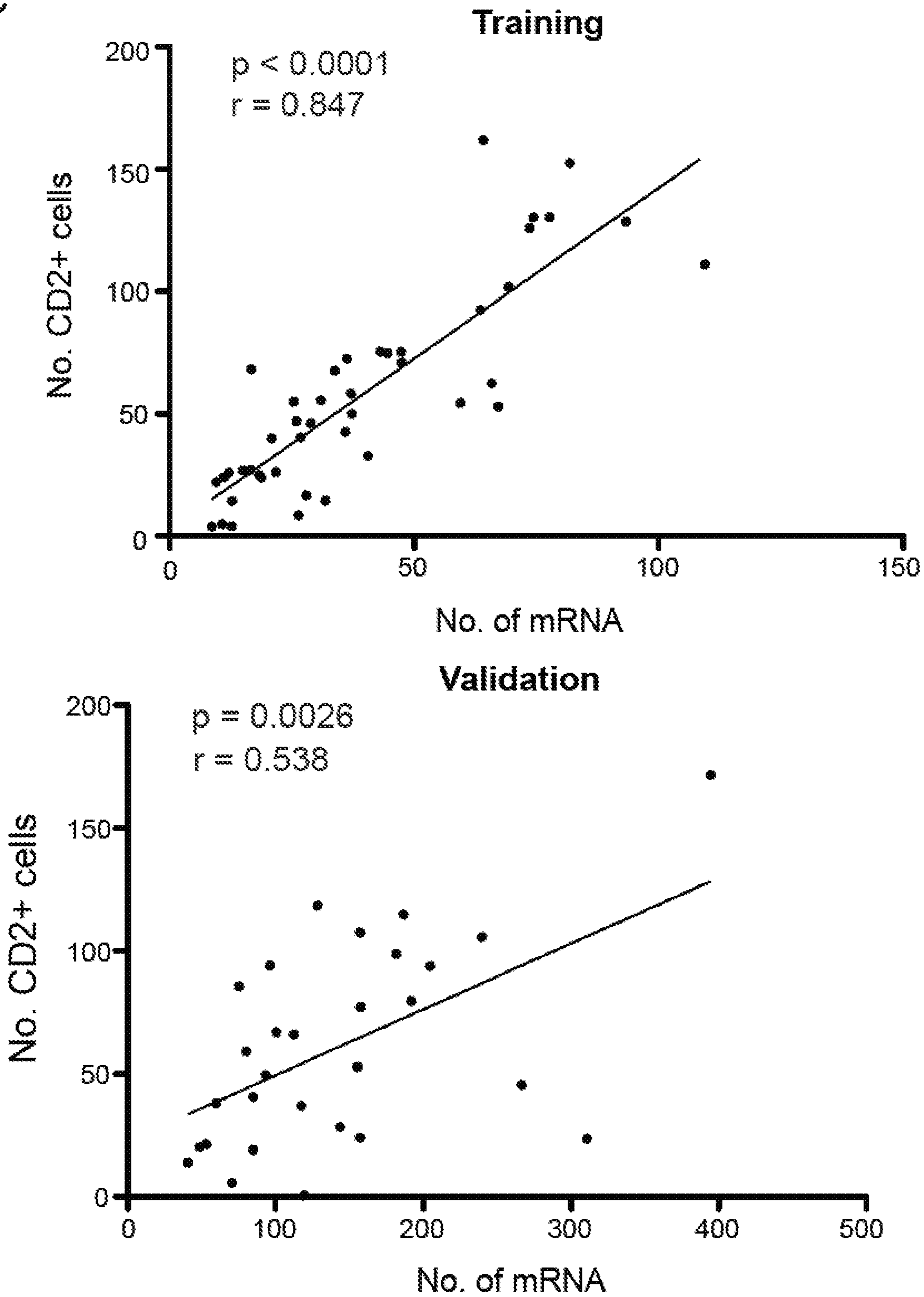
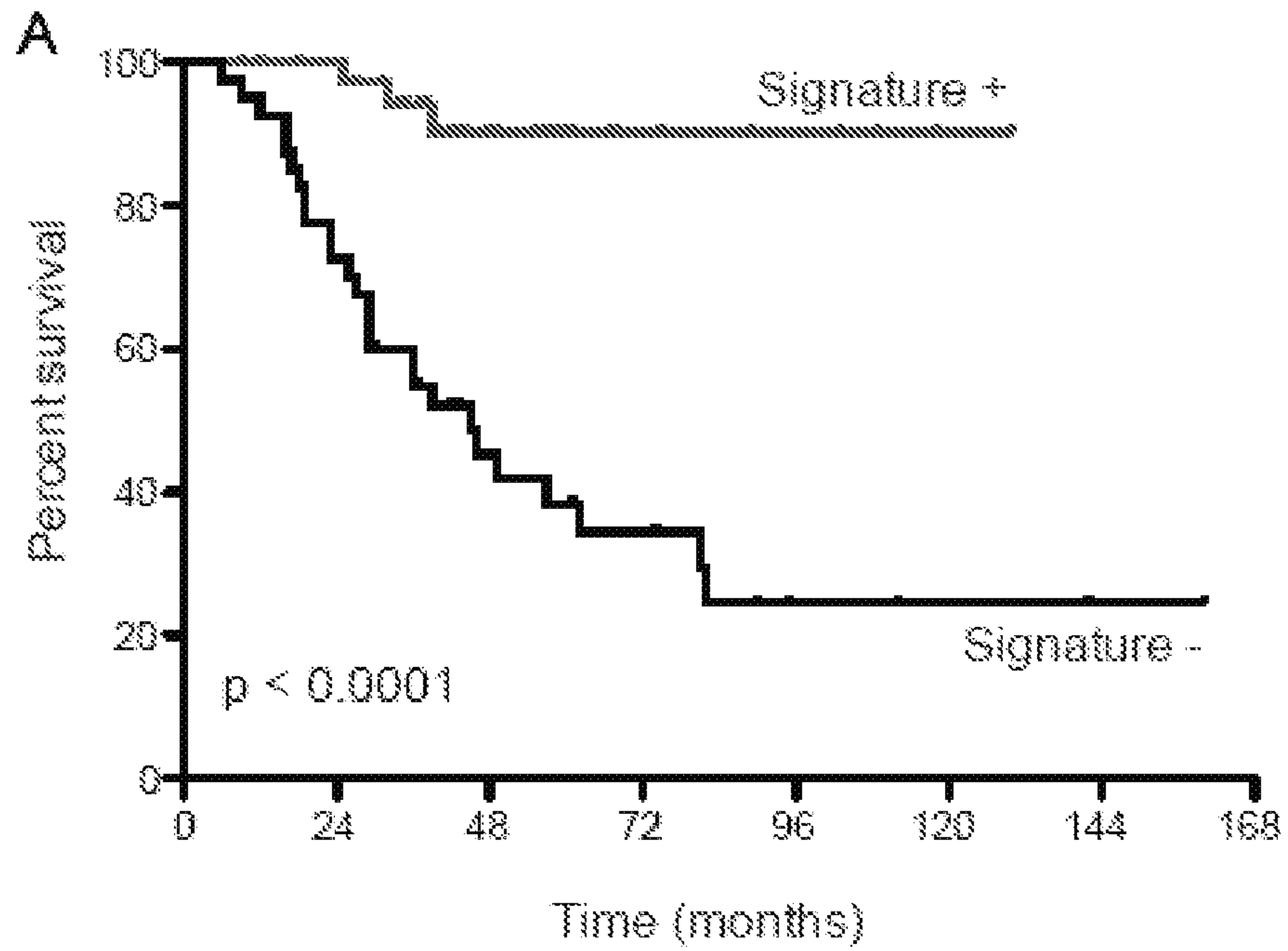
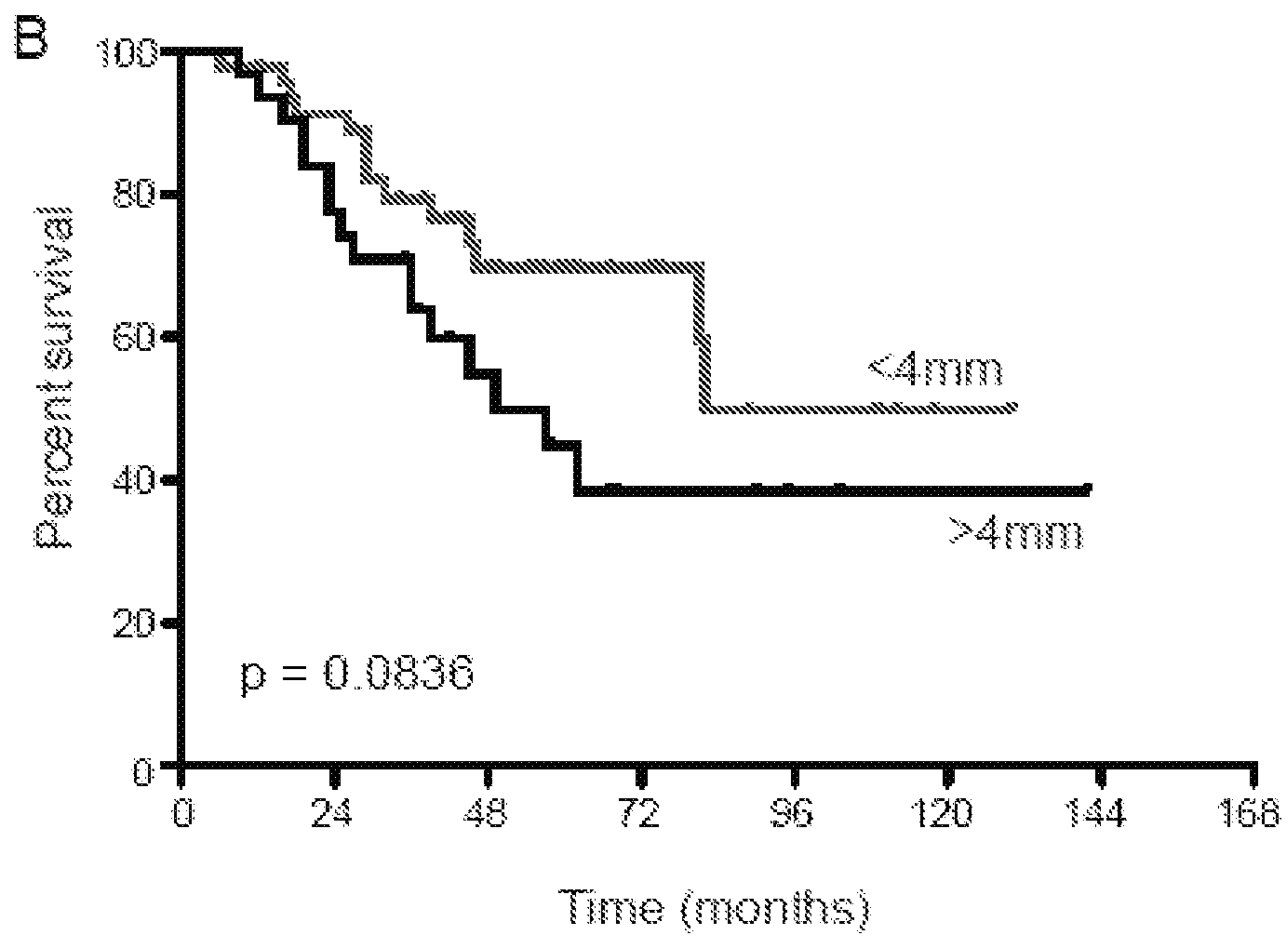


Figure 4



No. at Risk								
Signature +	39	36	18	6	4	1	0	0
Signature -	42	30	14	9	4	3	2	0



No. at Risk								
<4mm	48	40	19	9	4	1	0	0
>4mm	31	24	11	4	2	1	0	0

Figure 5

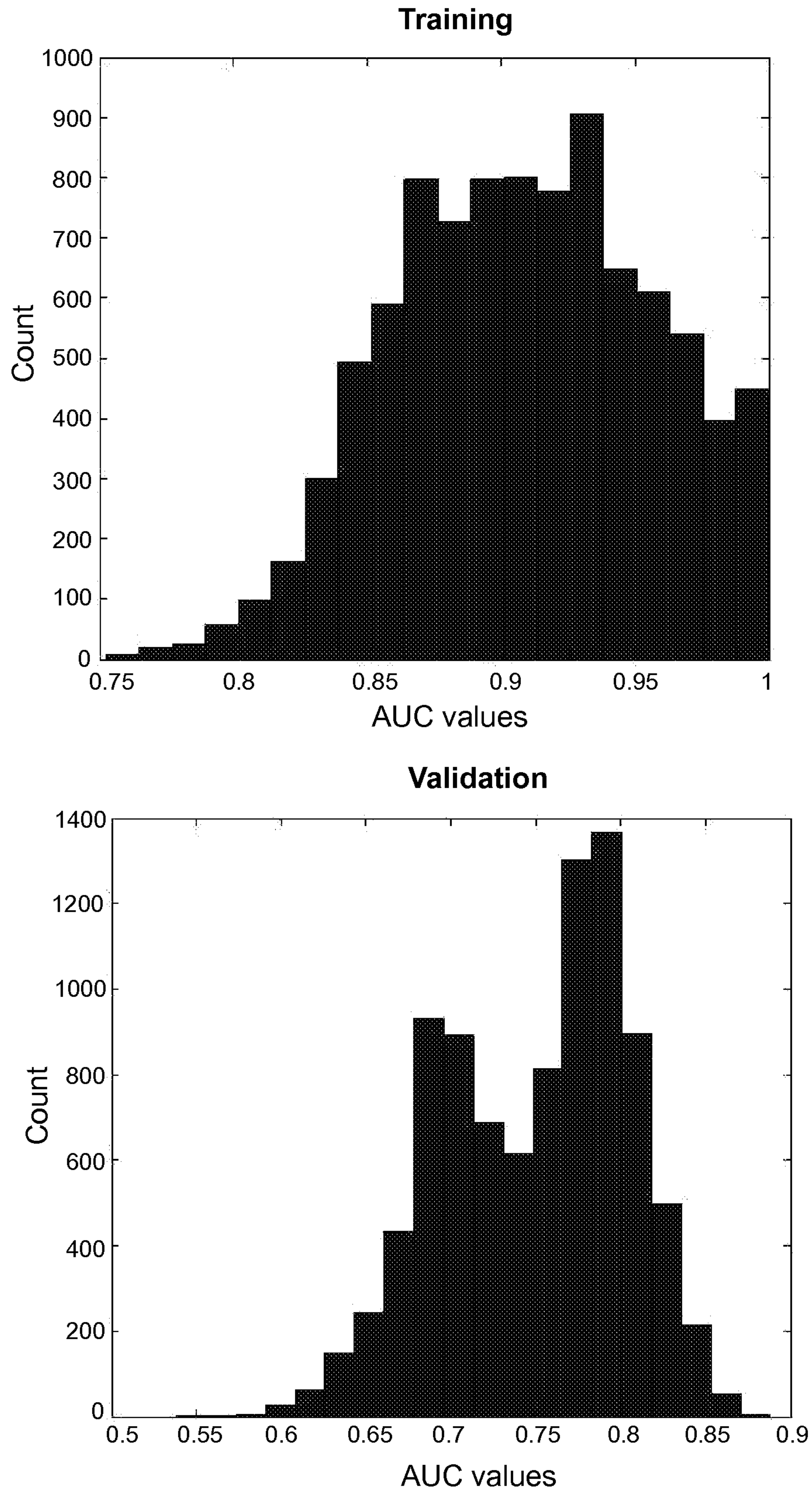


Figure 6 Panel A

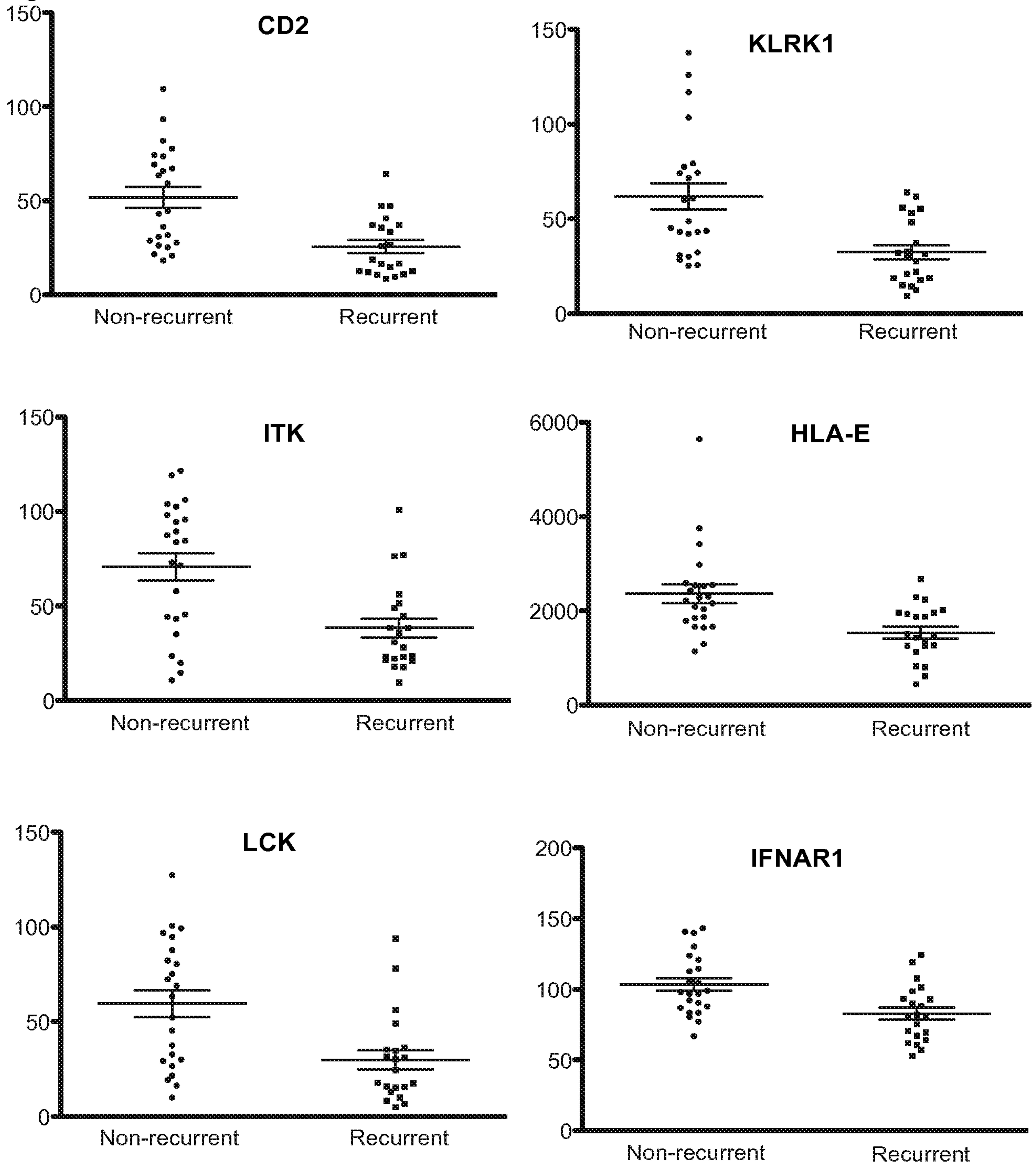


Figure 6 Panel A, continued

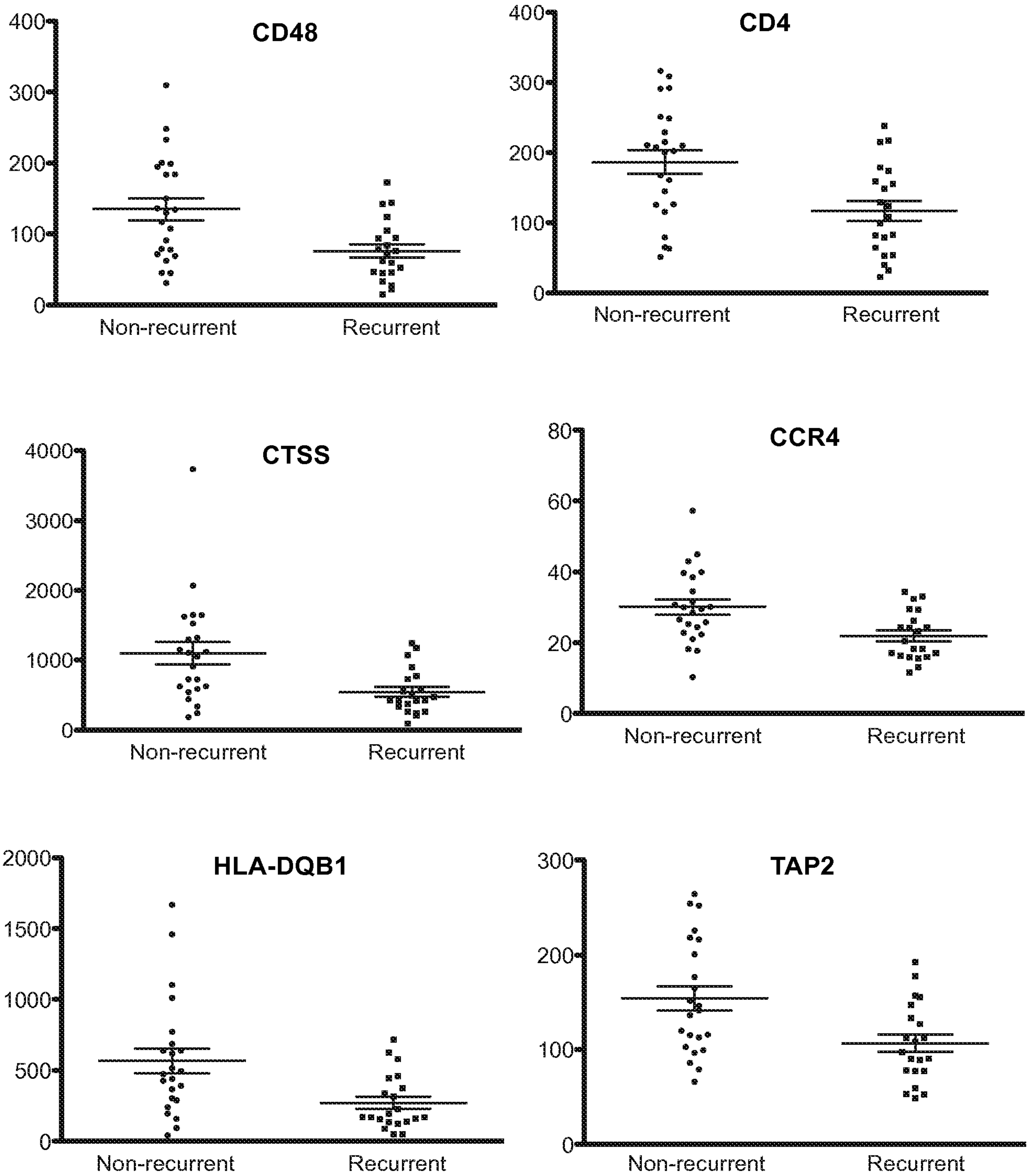
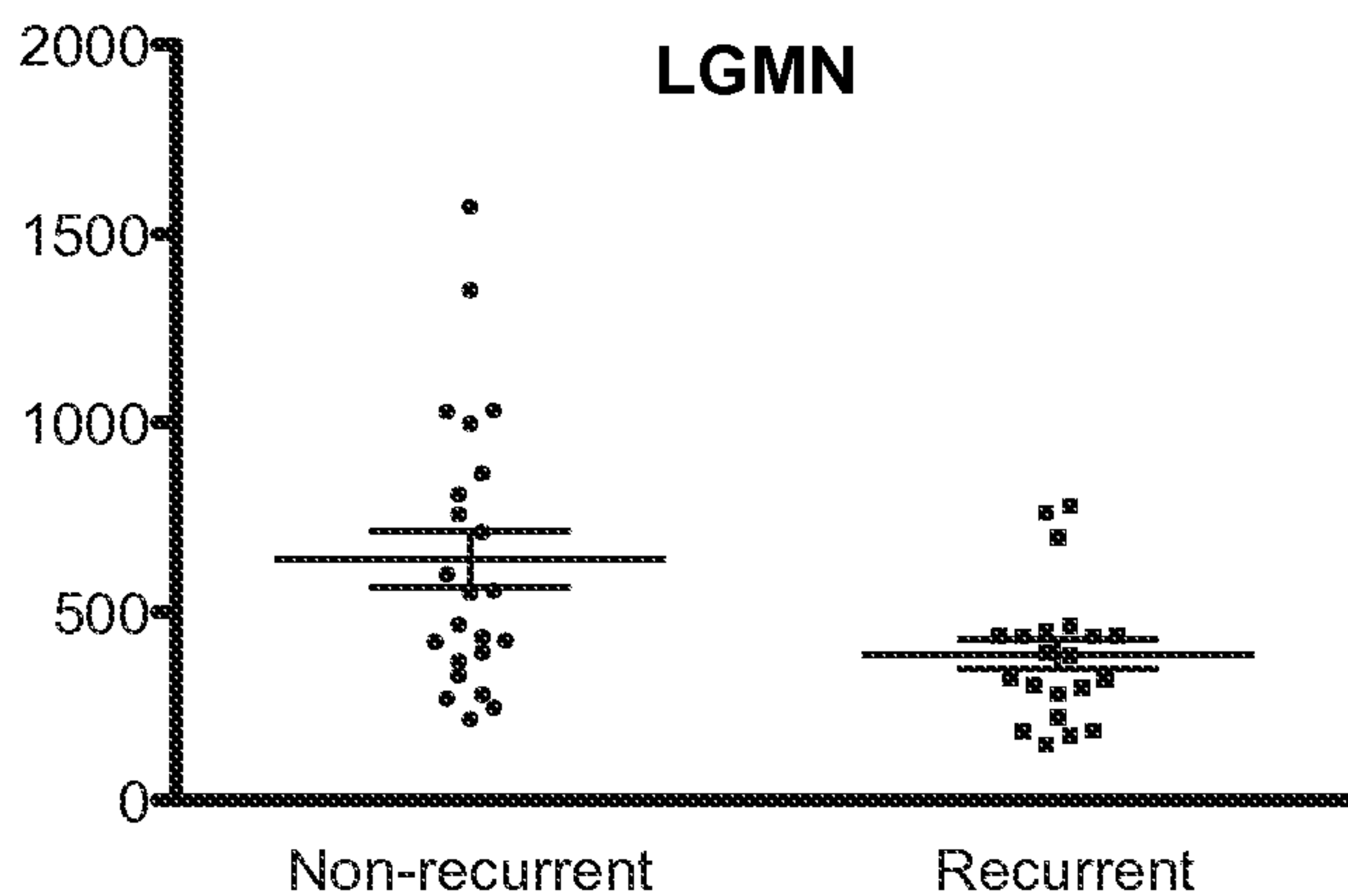
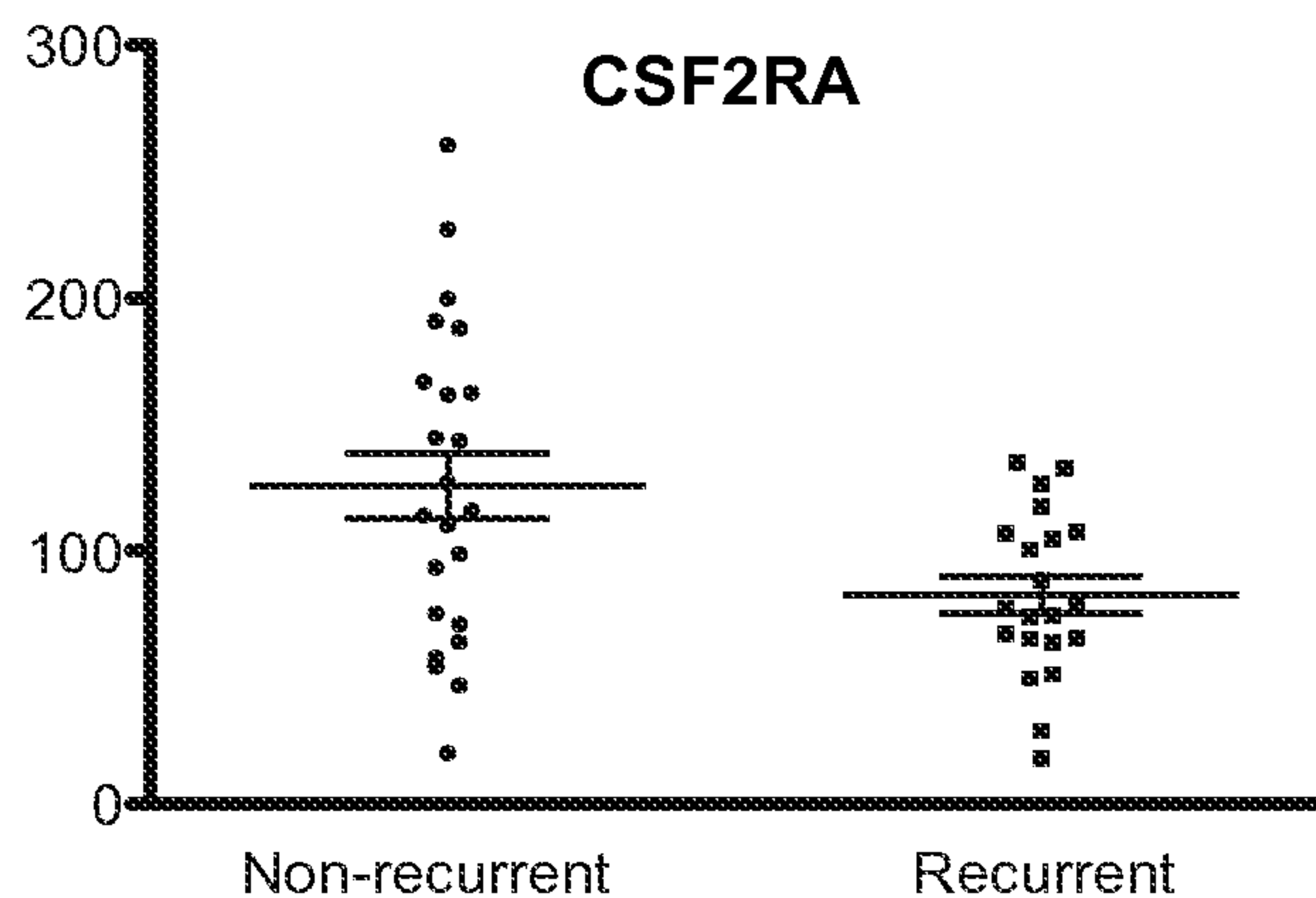
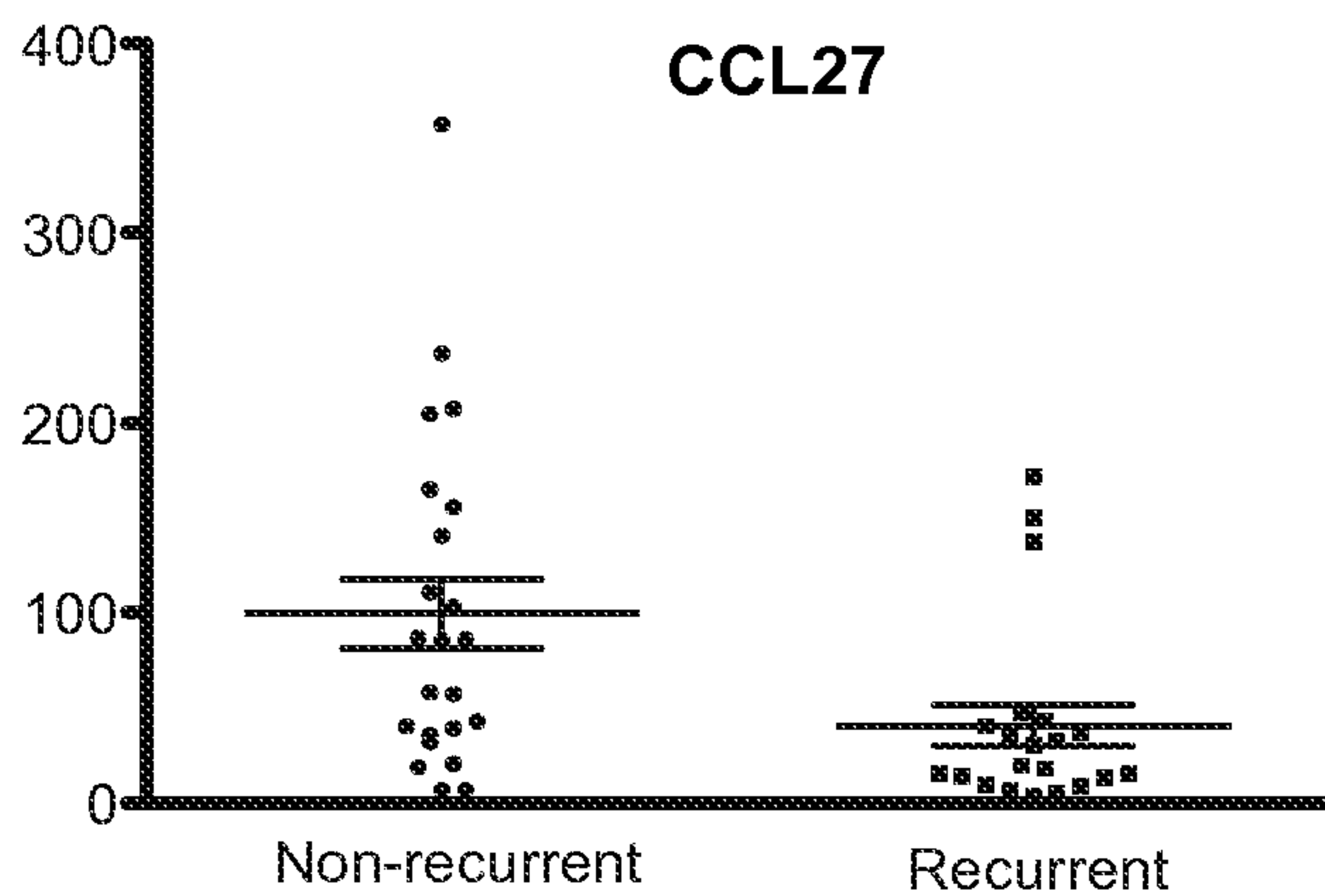
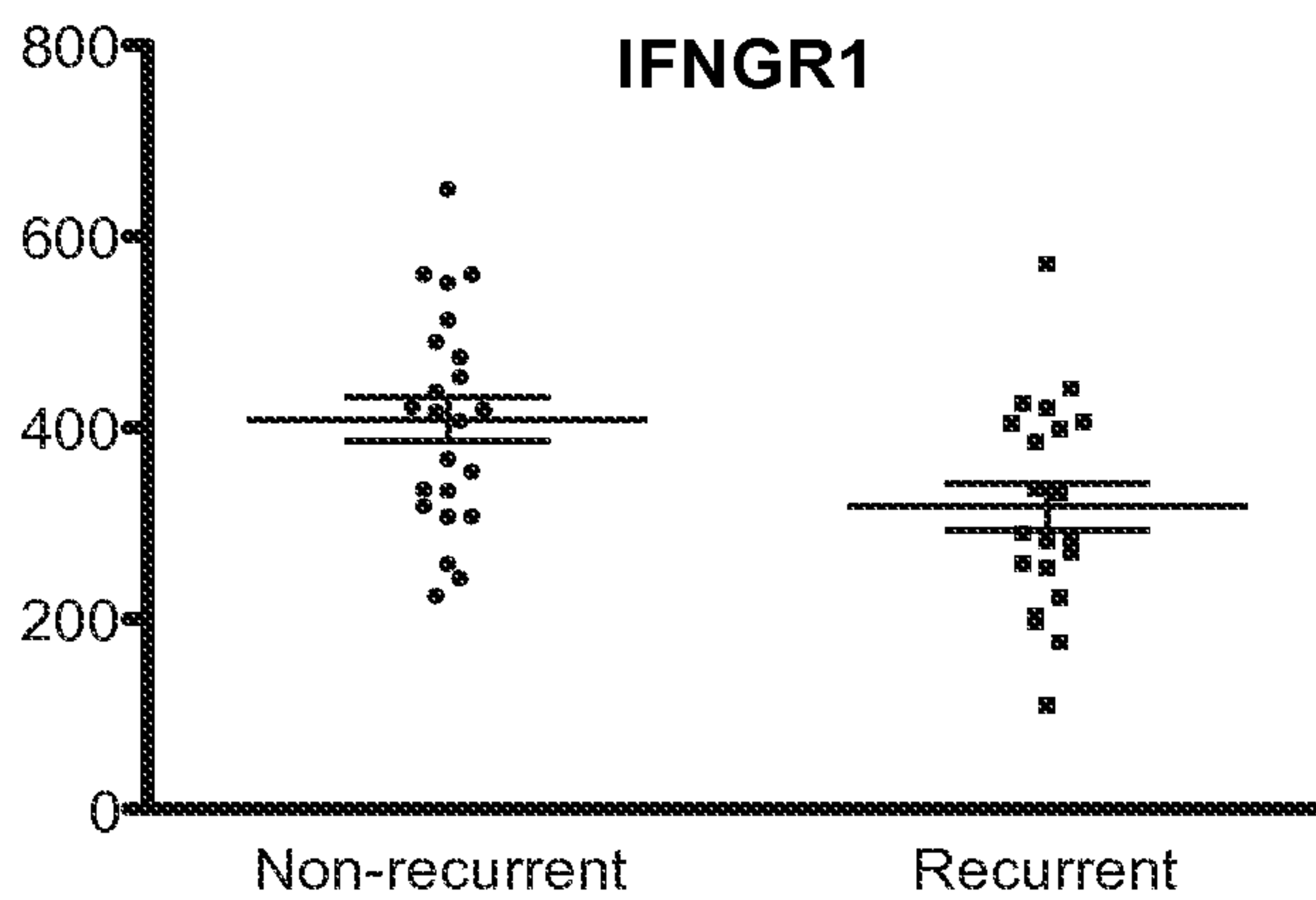


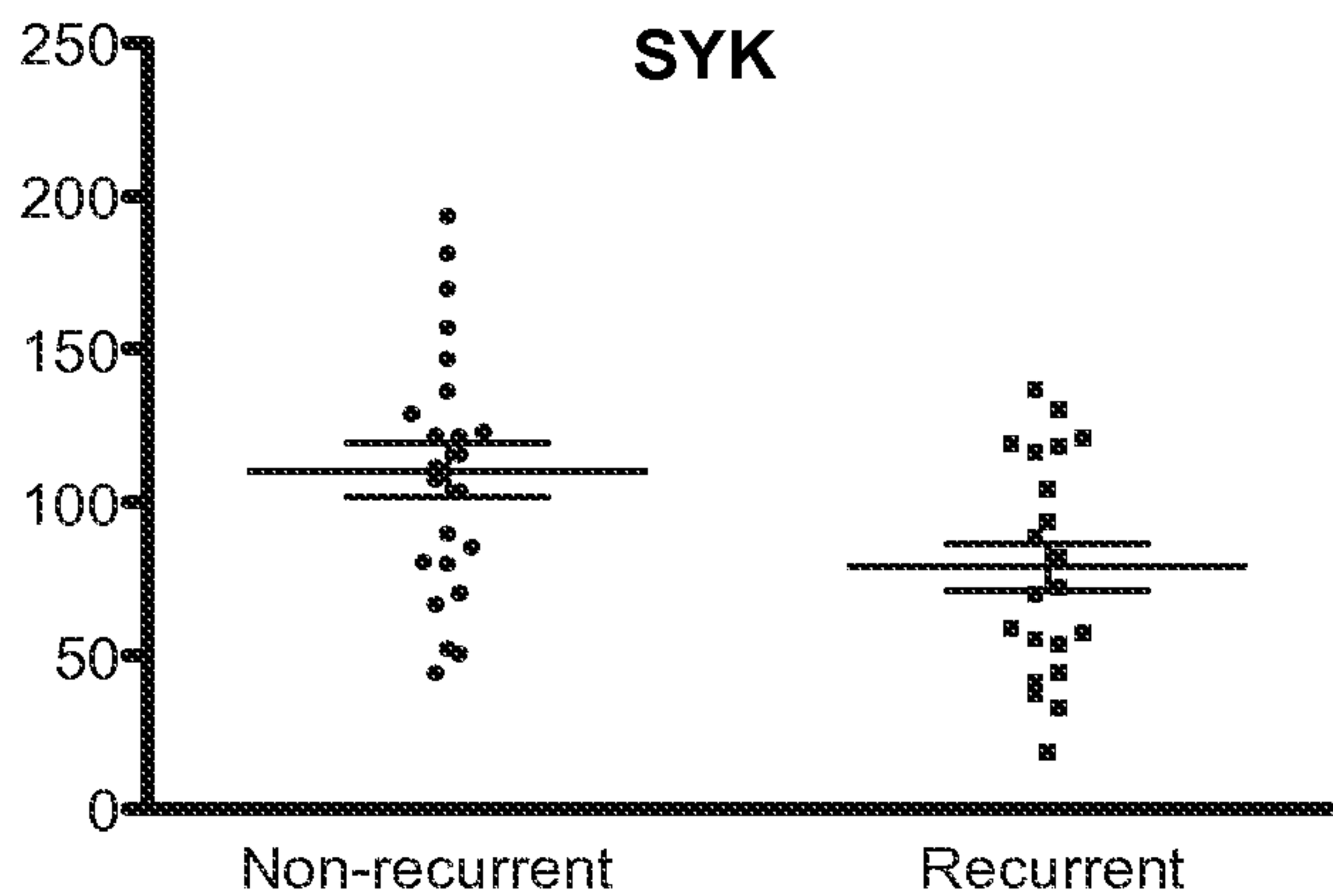
Figure 6 Panel A, continued











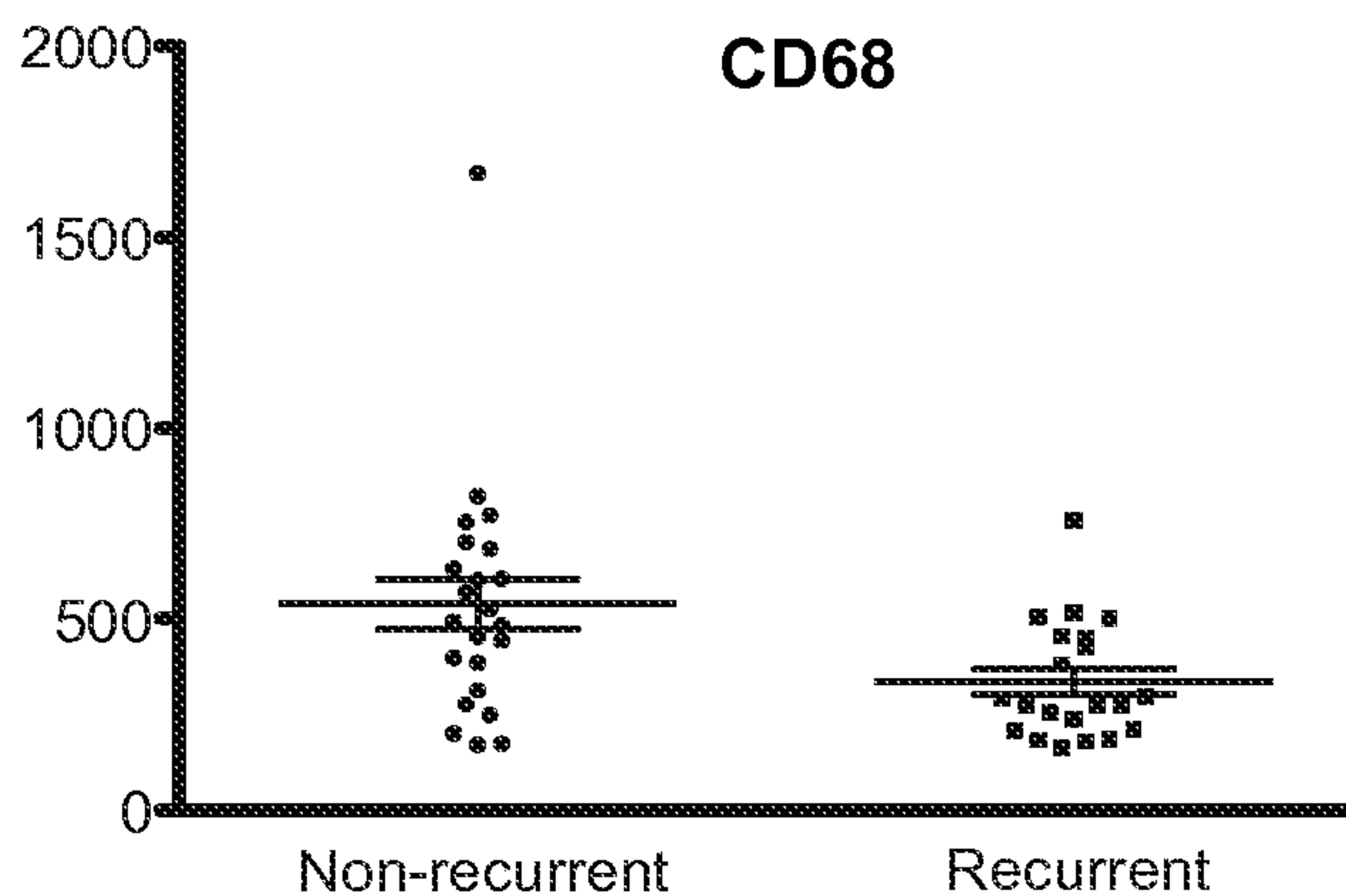


Figure 6 Panel A, continued

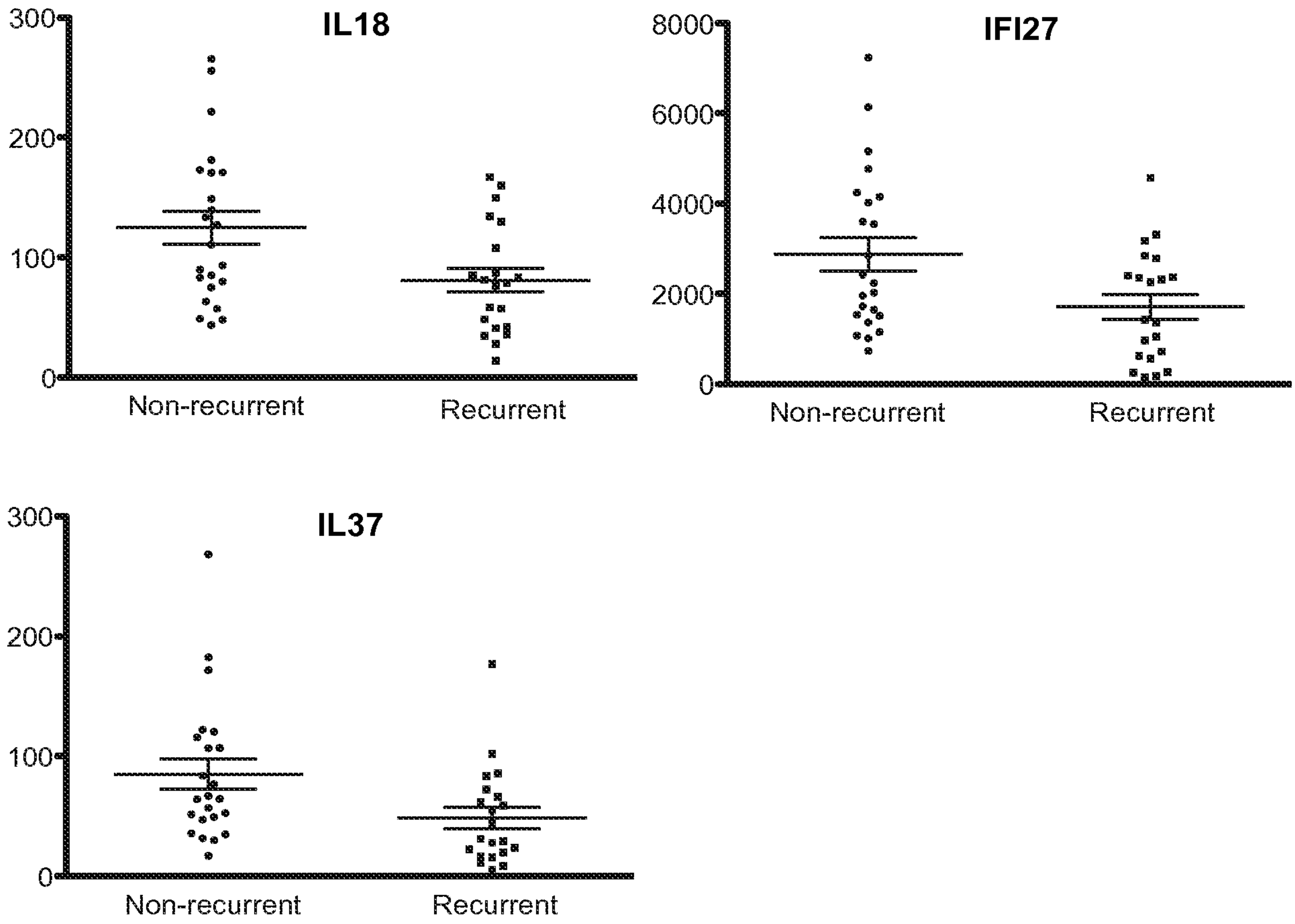


Figure 6 Panel B

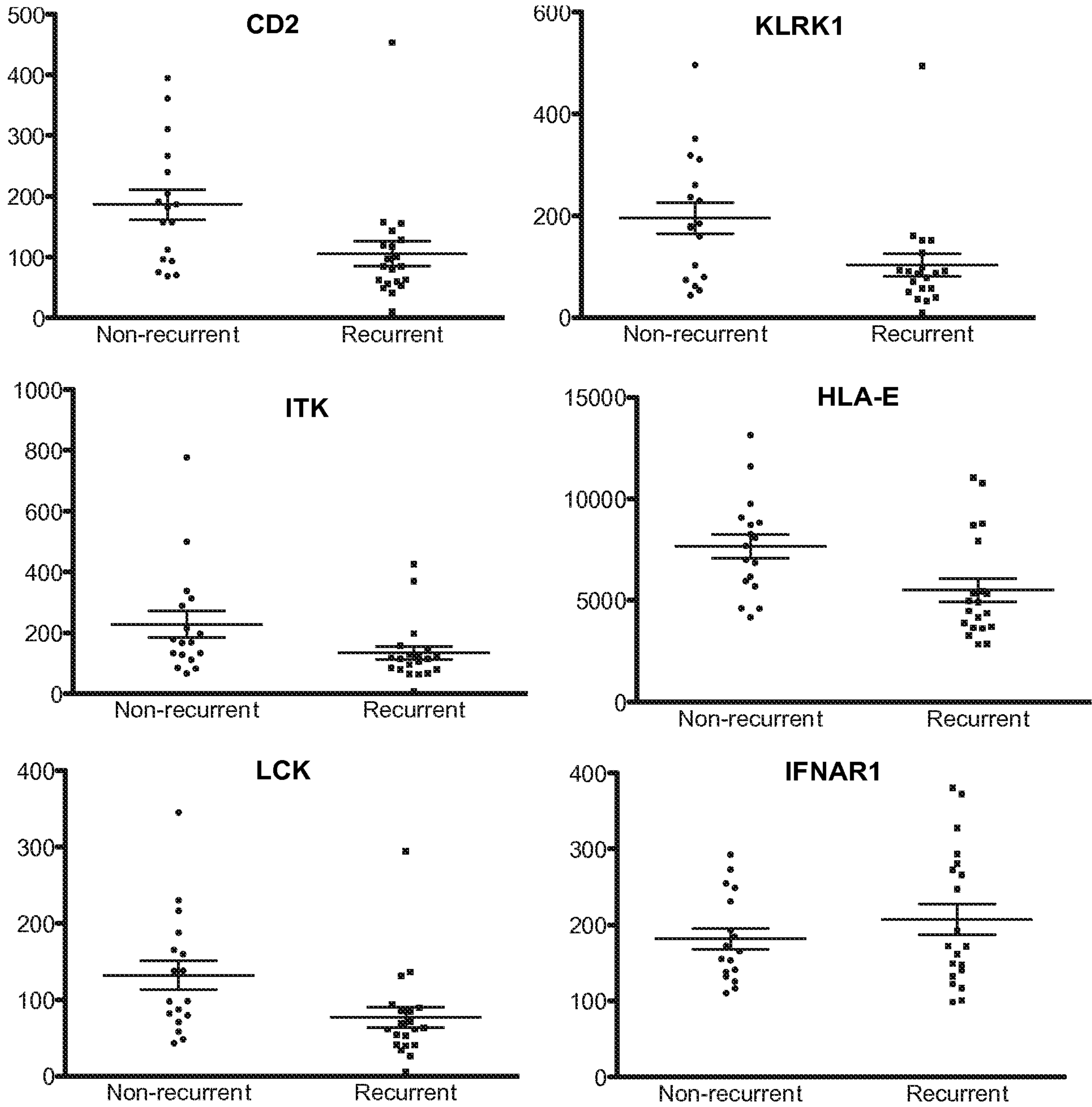


Figure 6 Panel B, continued

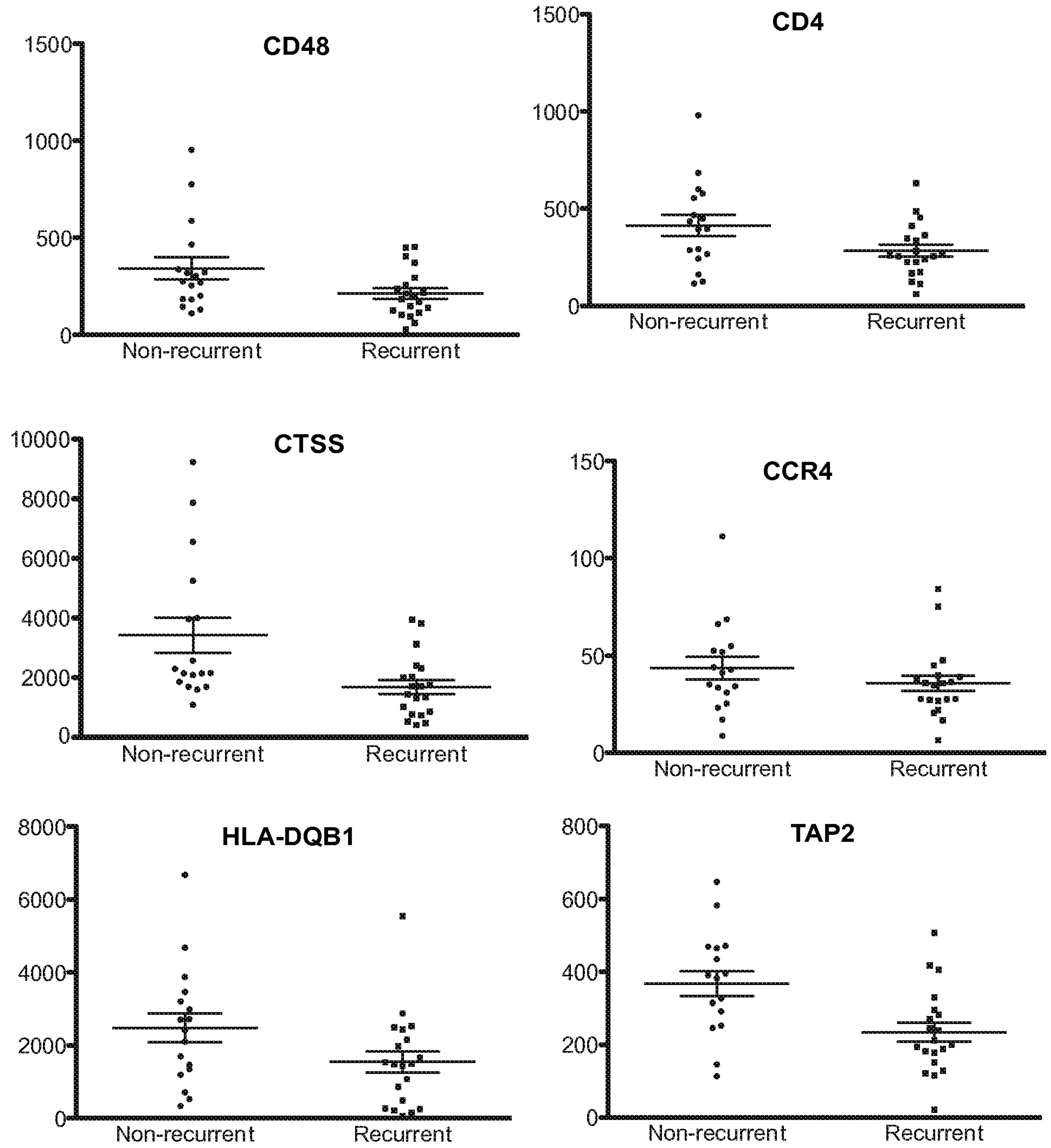


Figure 6 Panel B, continued

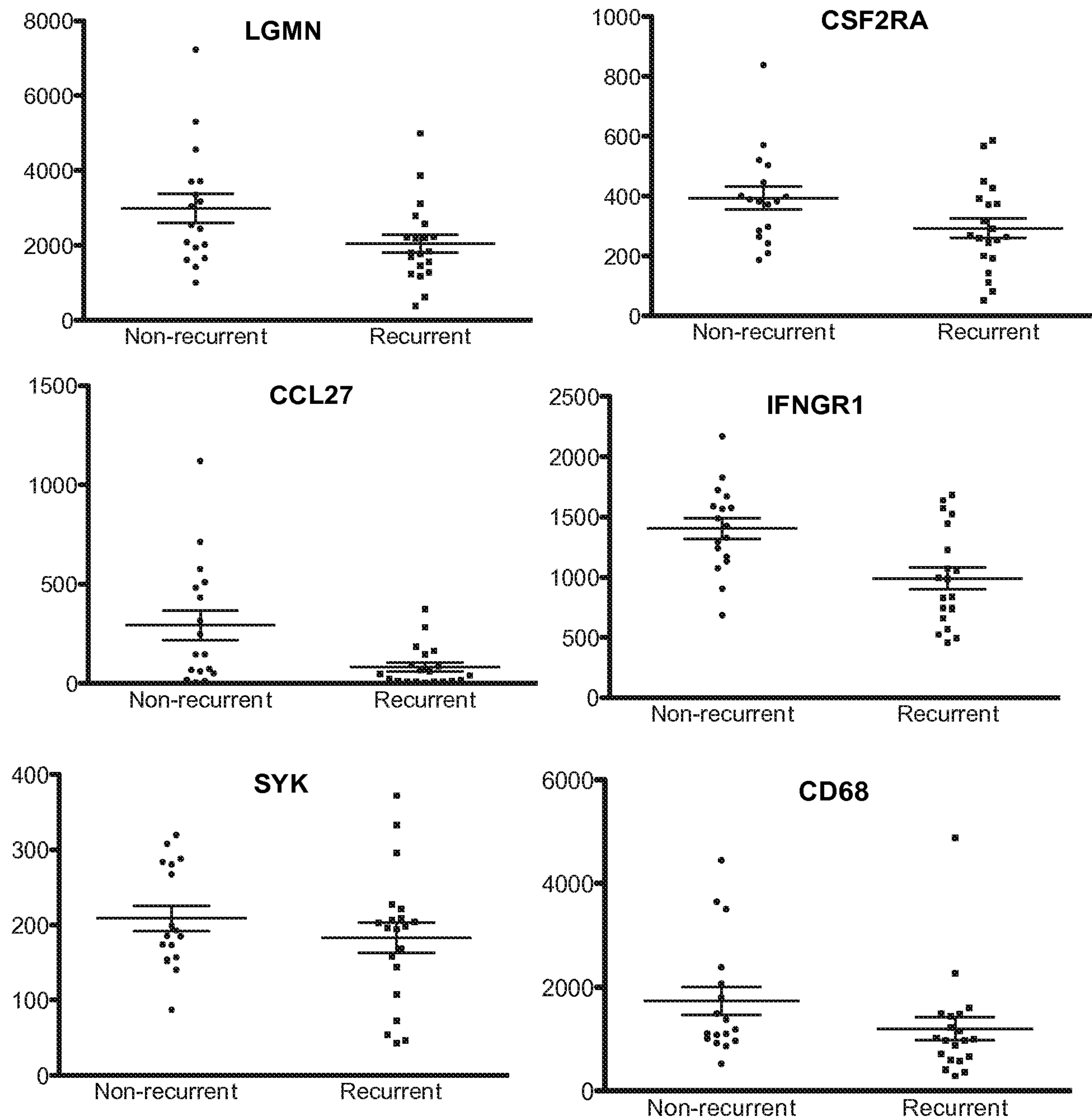


Figure 6 Panel B, continued

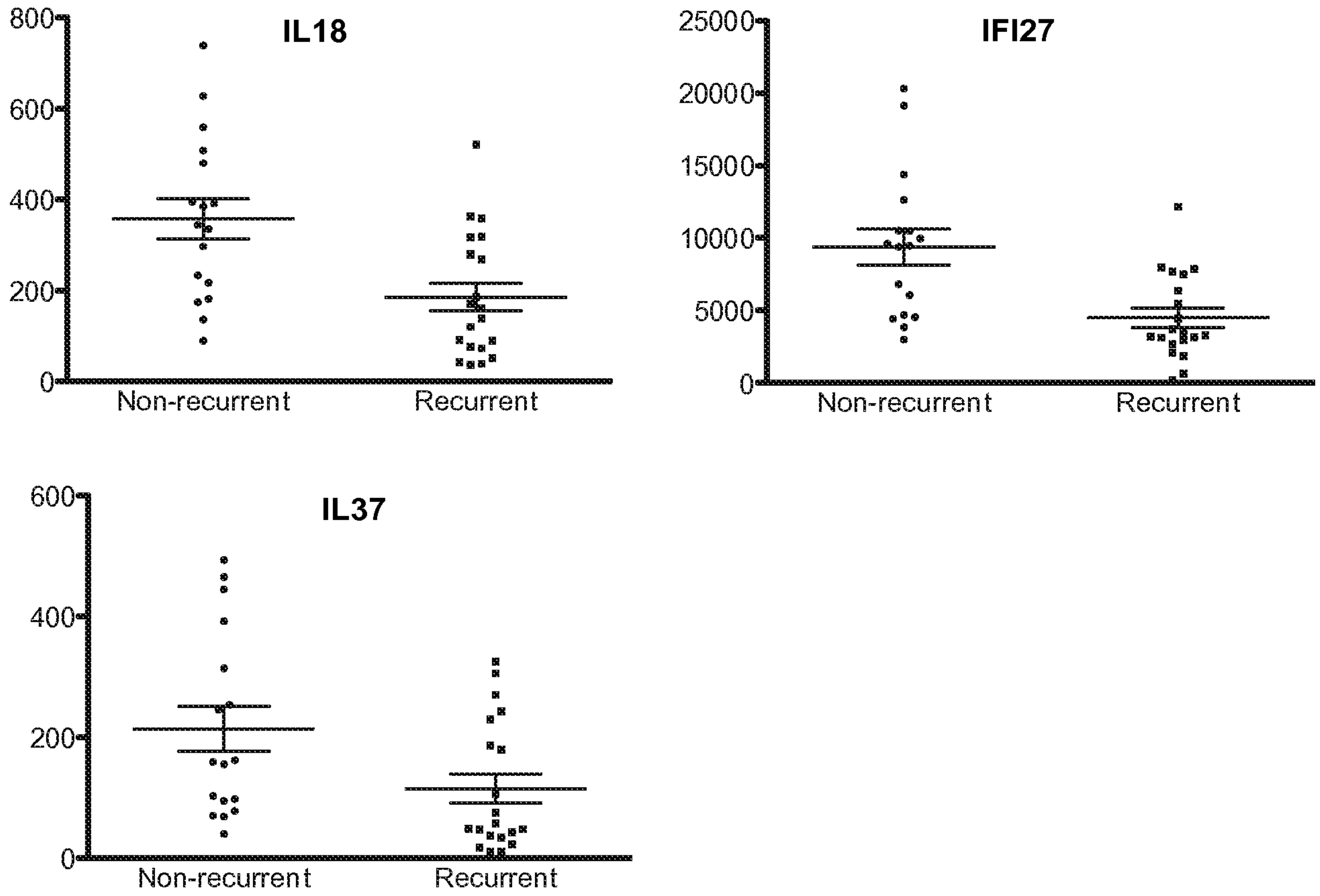


Figure 7

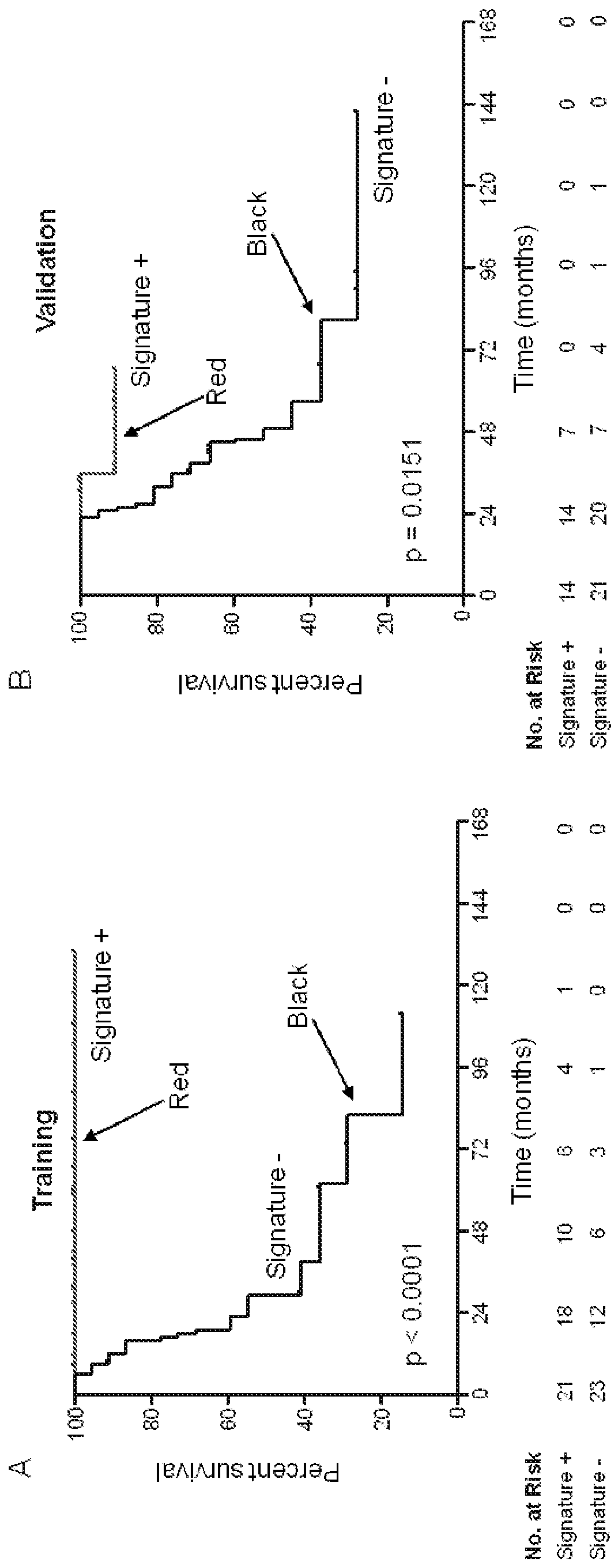


Figure 8

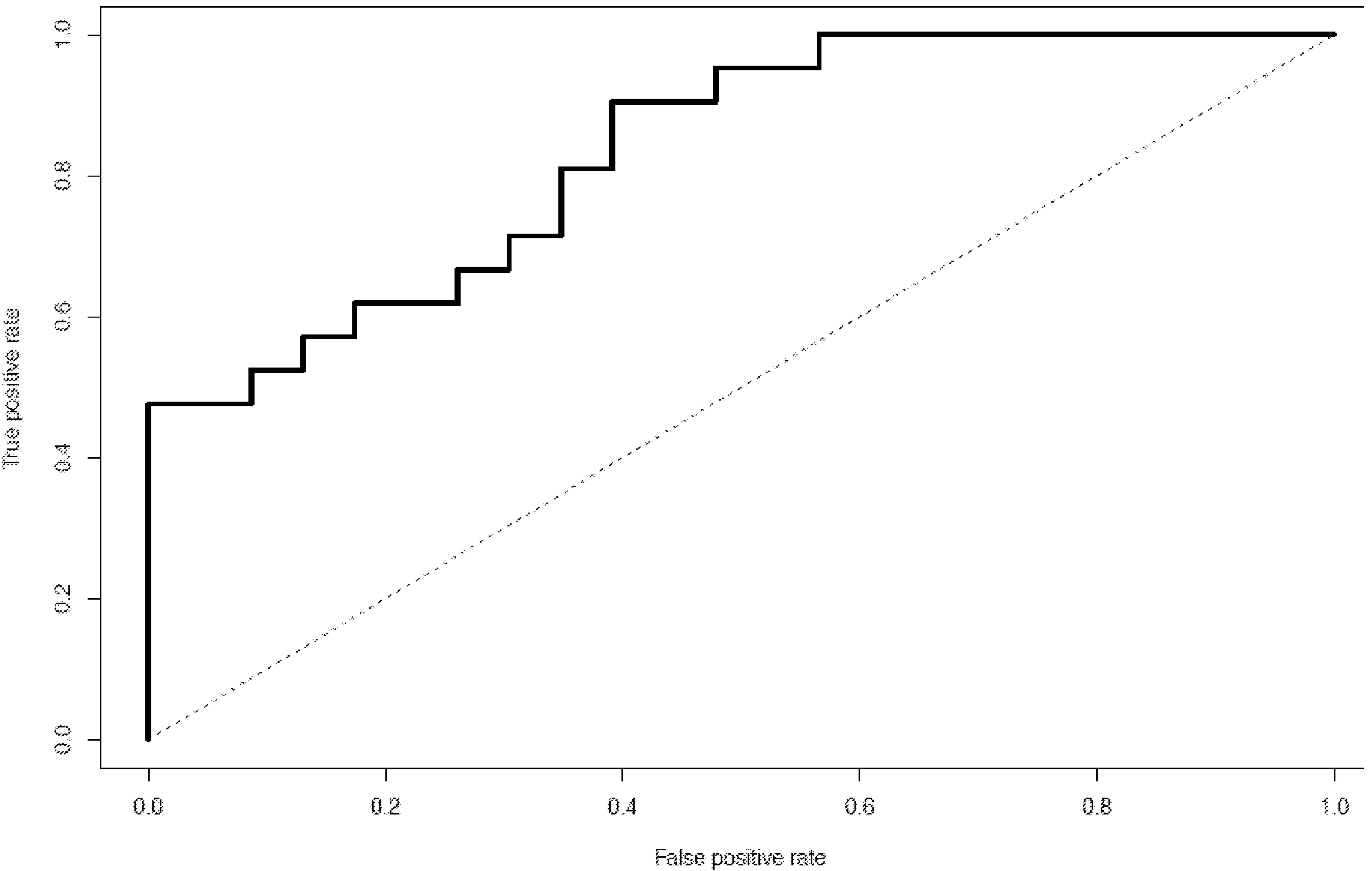


Figure 9

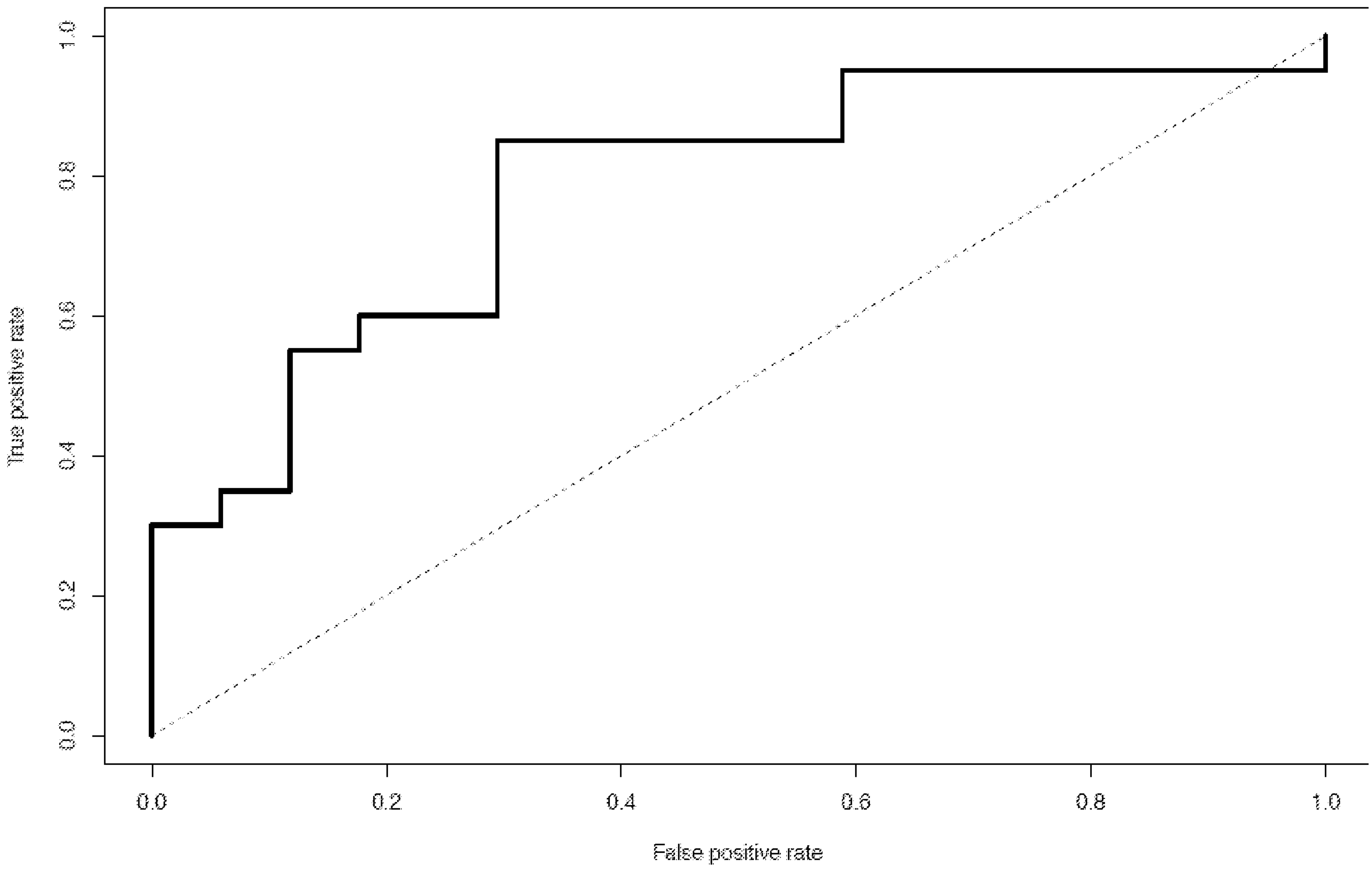


Figure 10

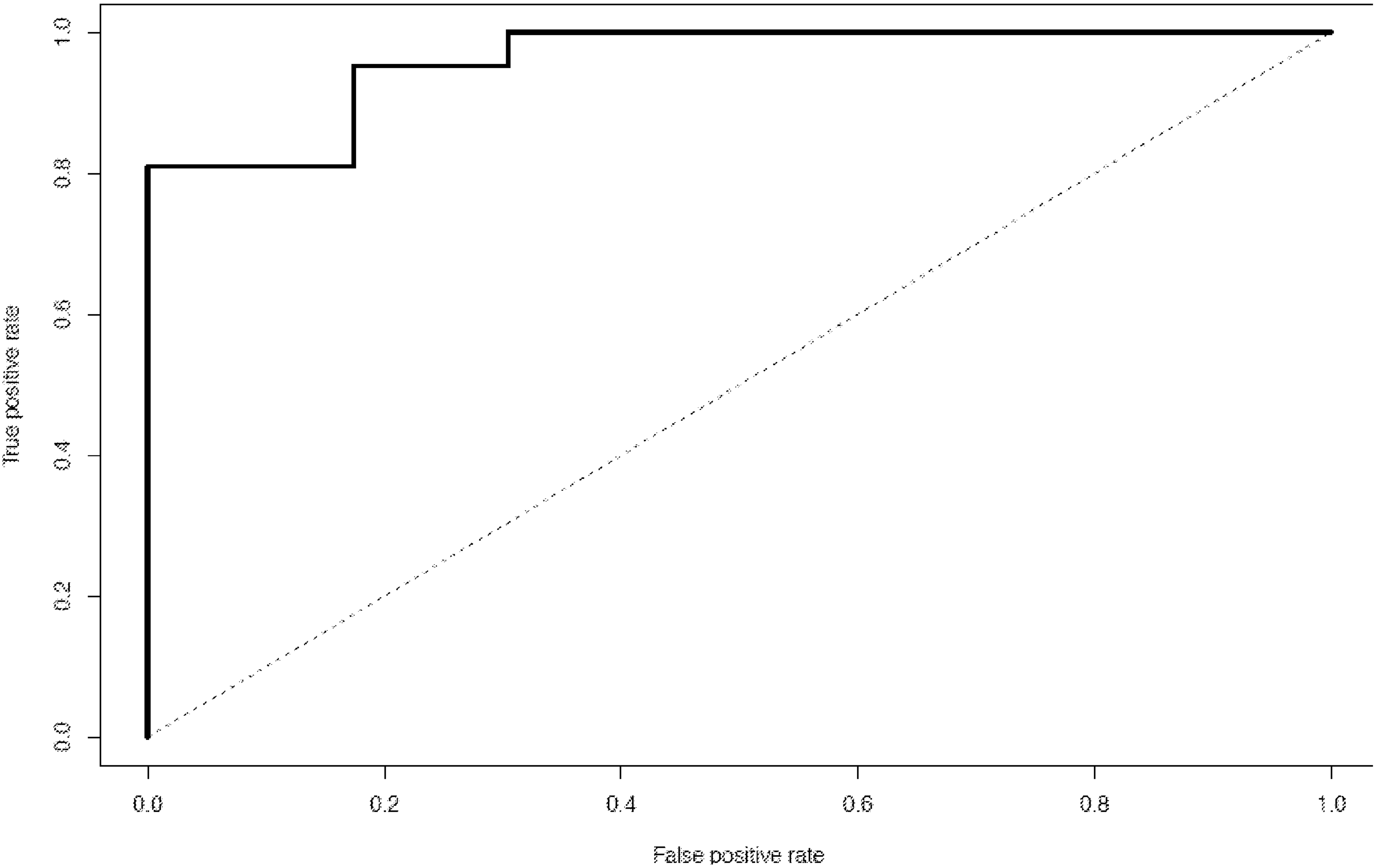


Figure 11

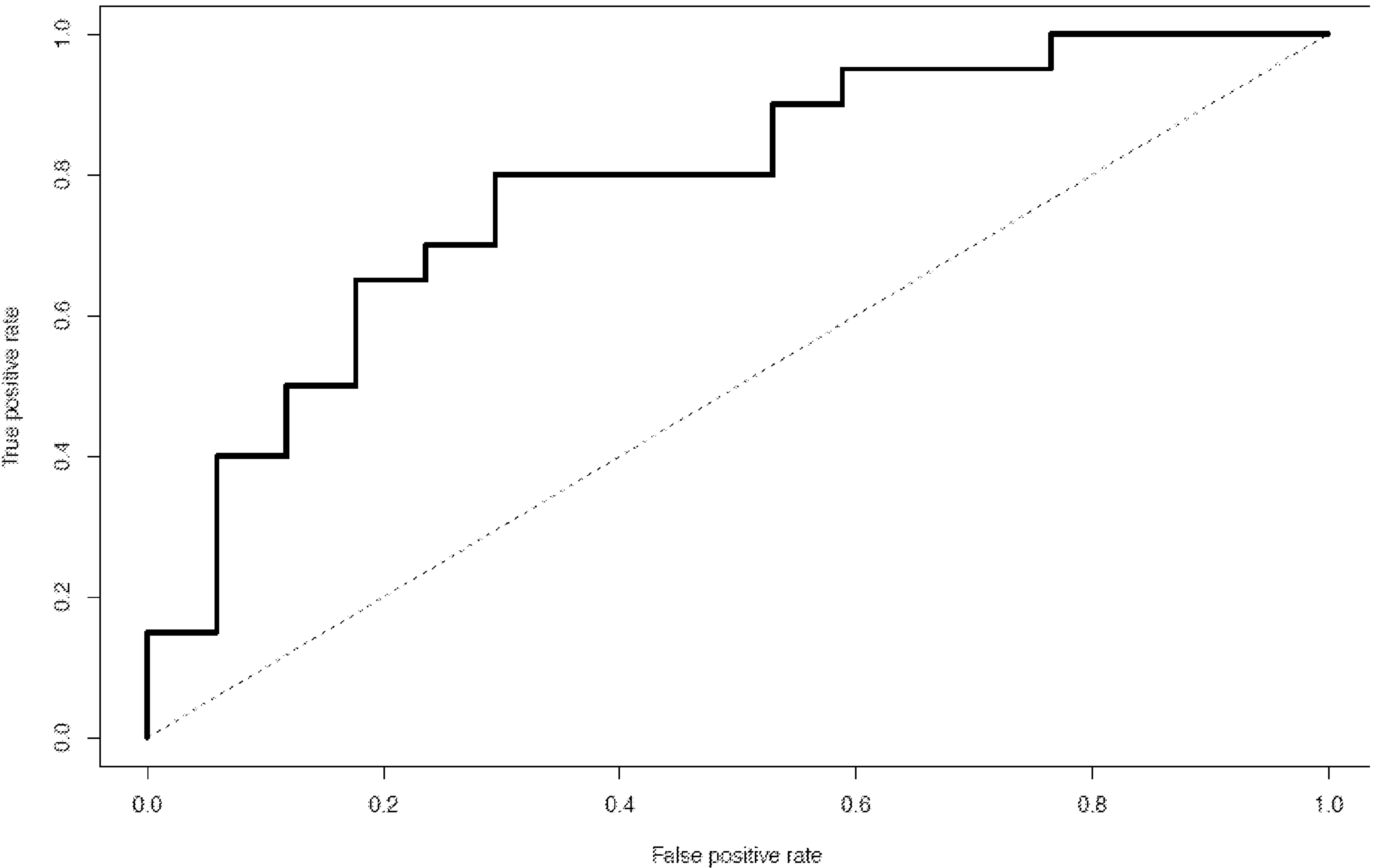


Figure 12

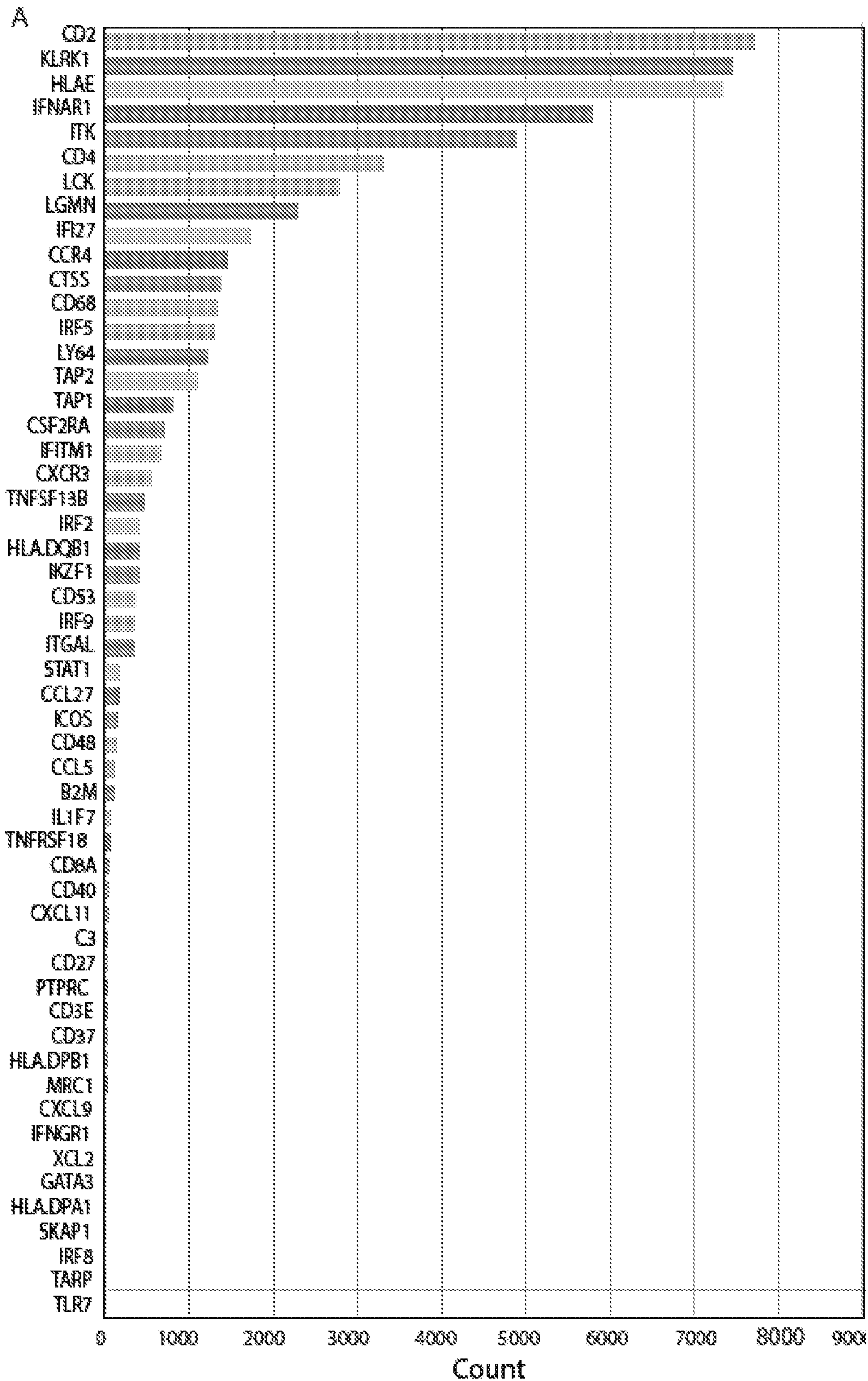
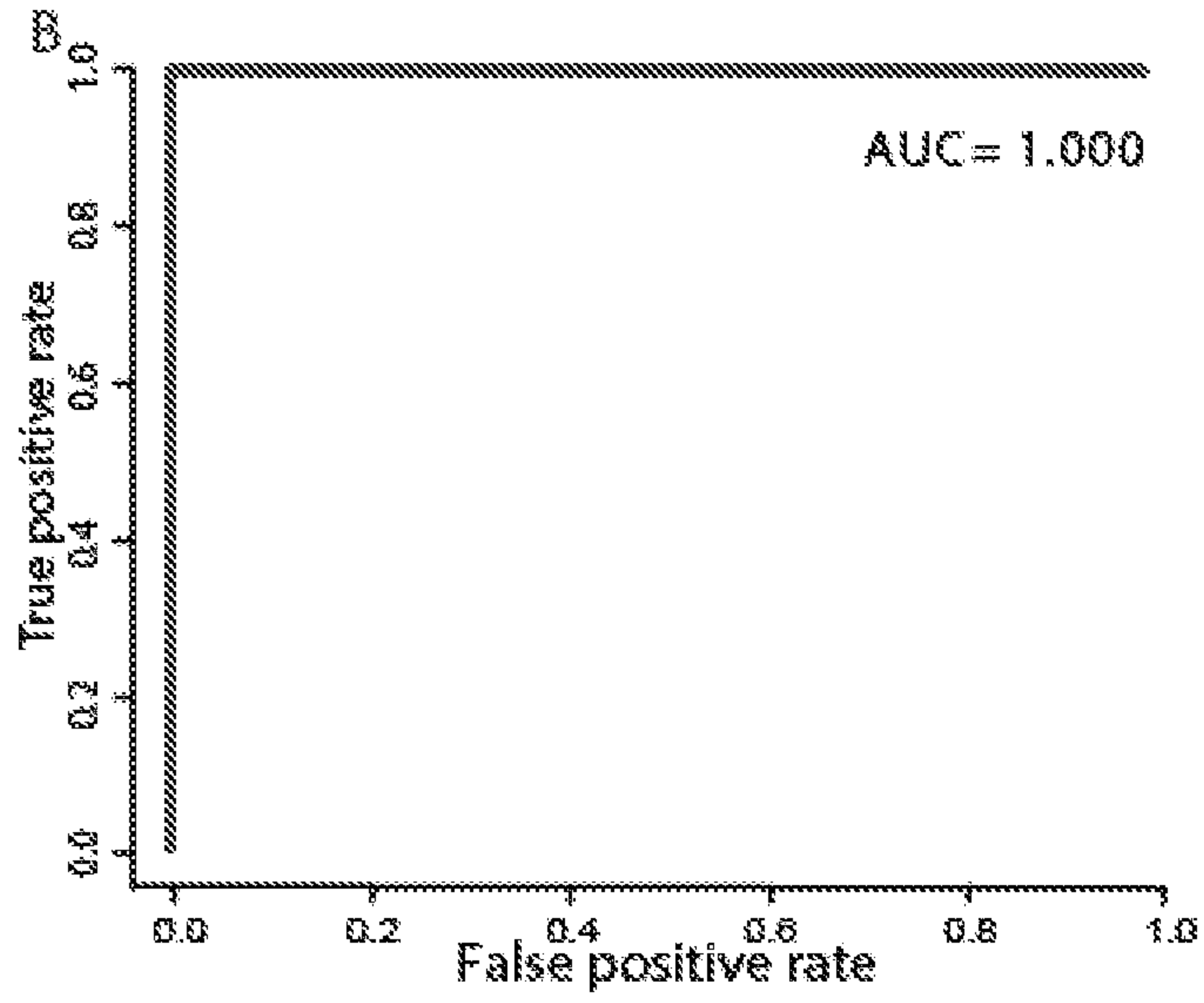


Figure 12, continued



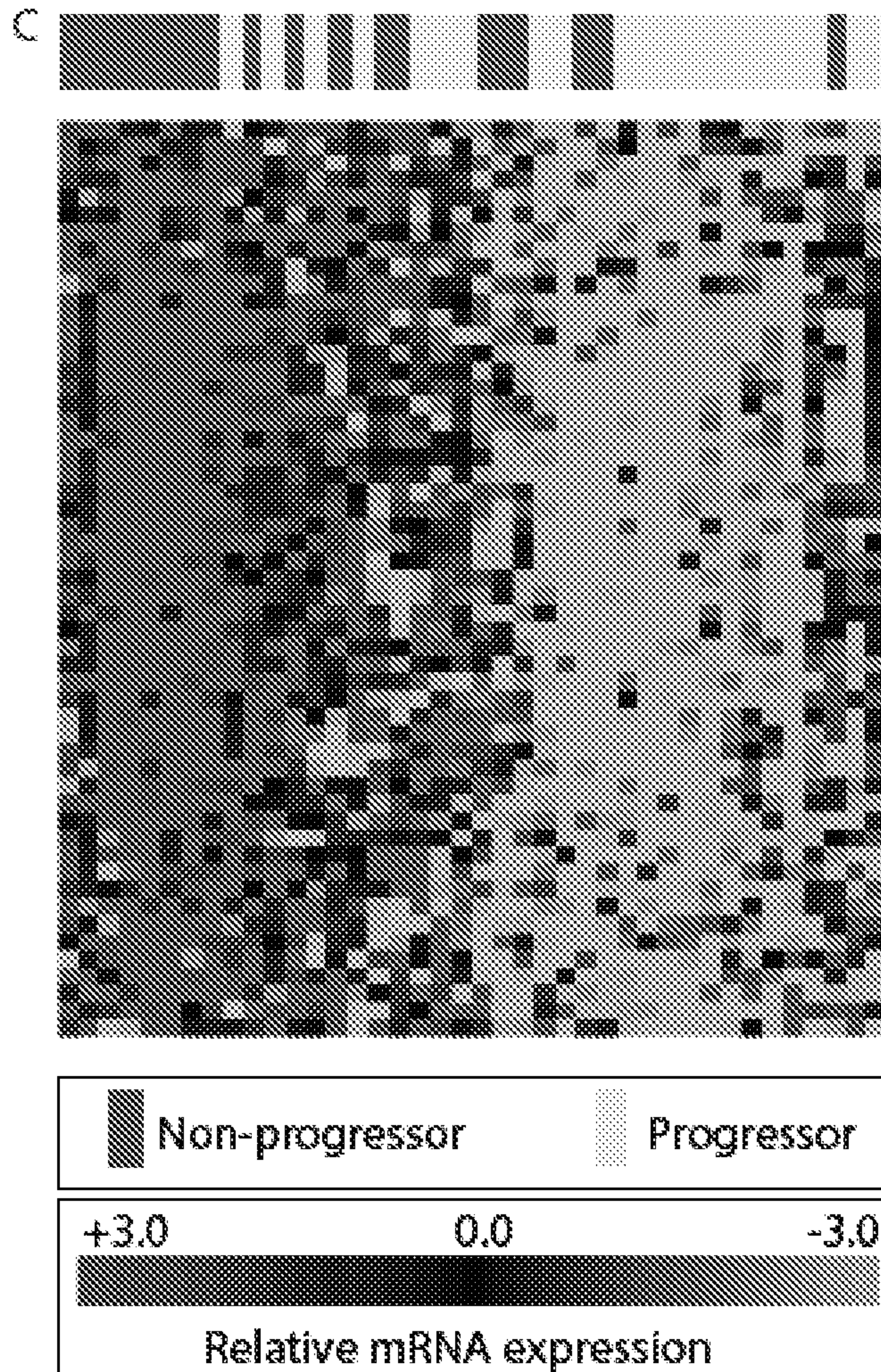


Figure 13

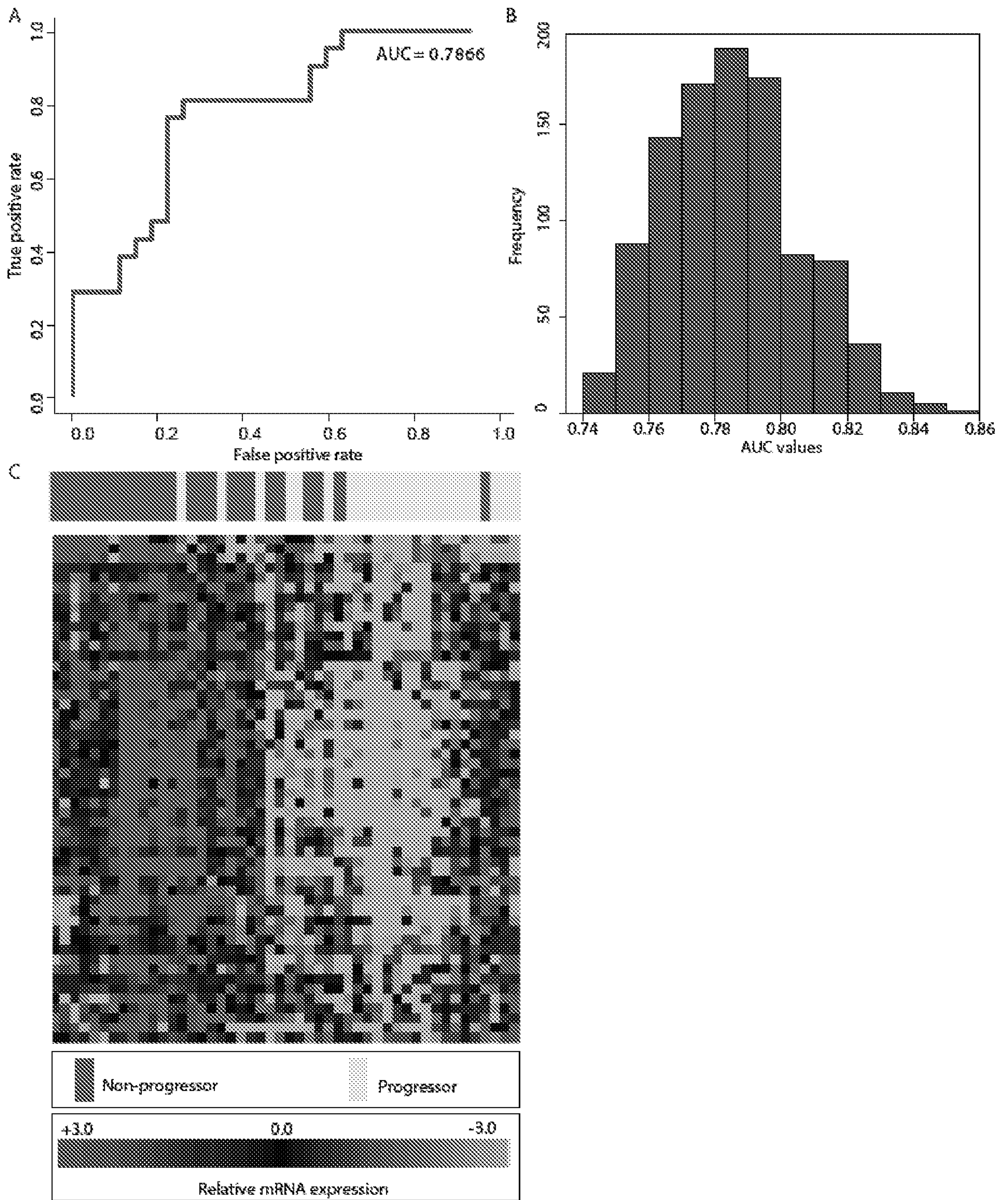


Figure 14

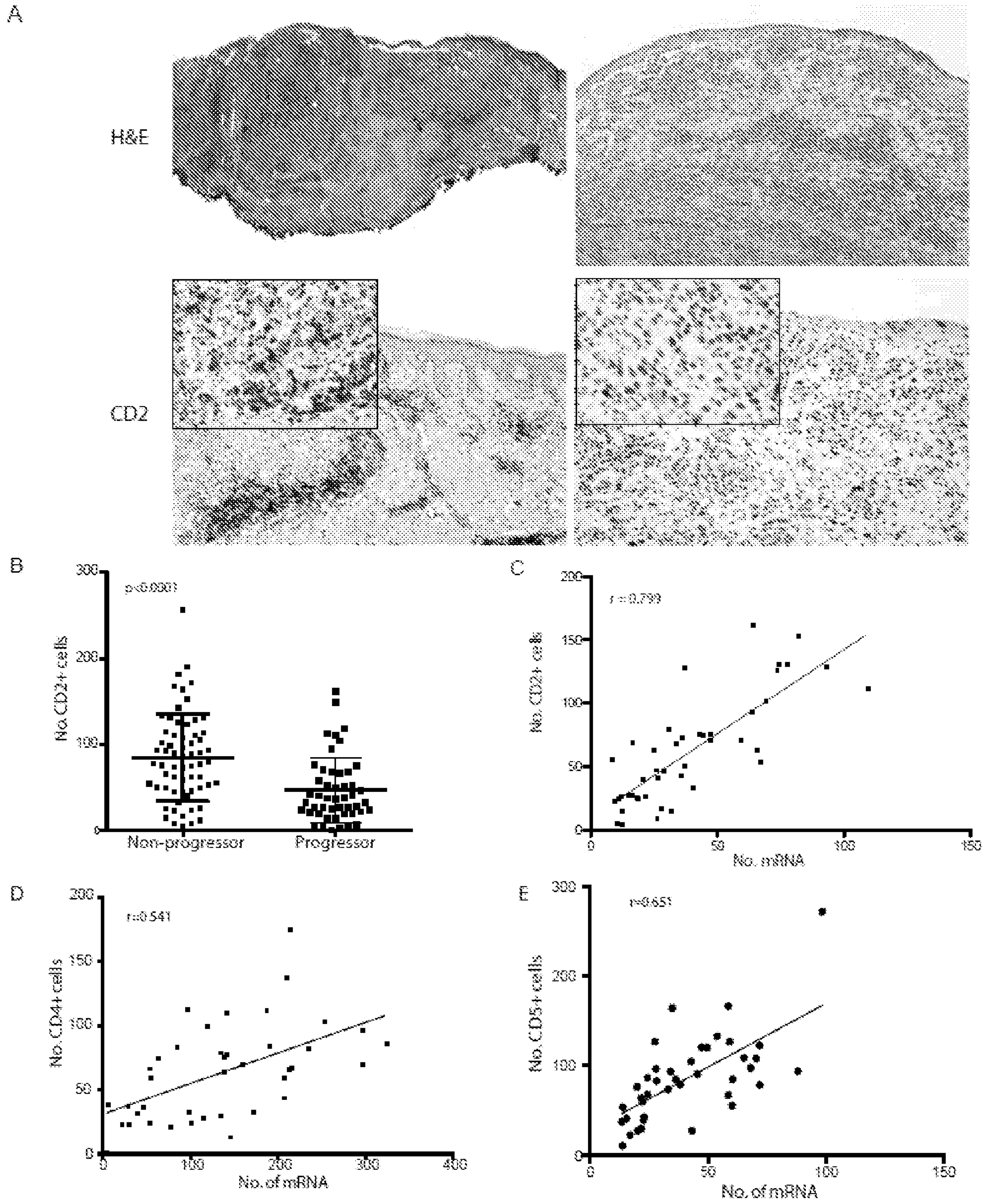


Figure 15

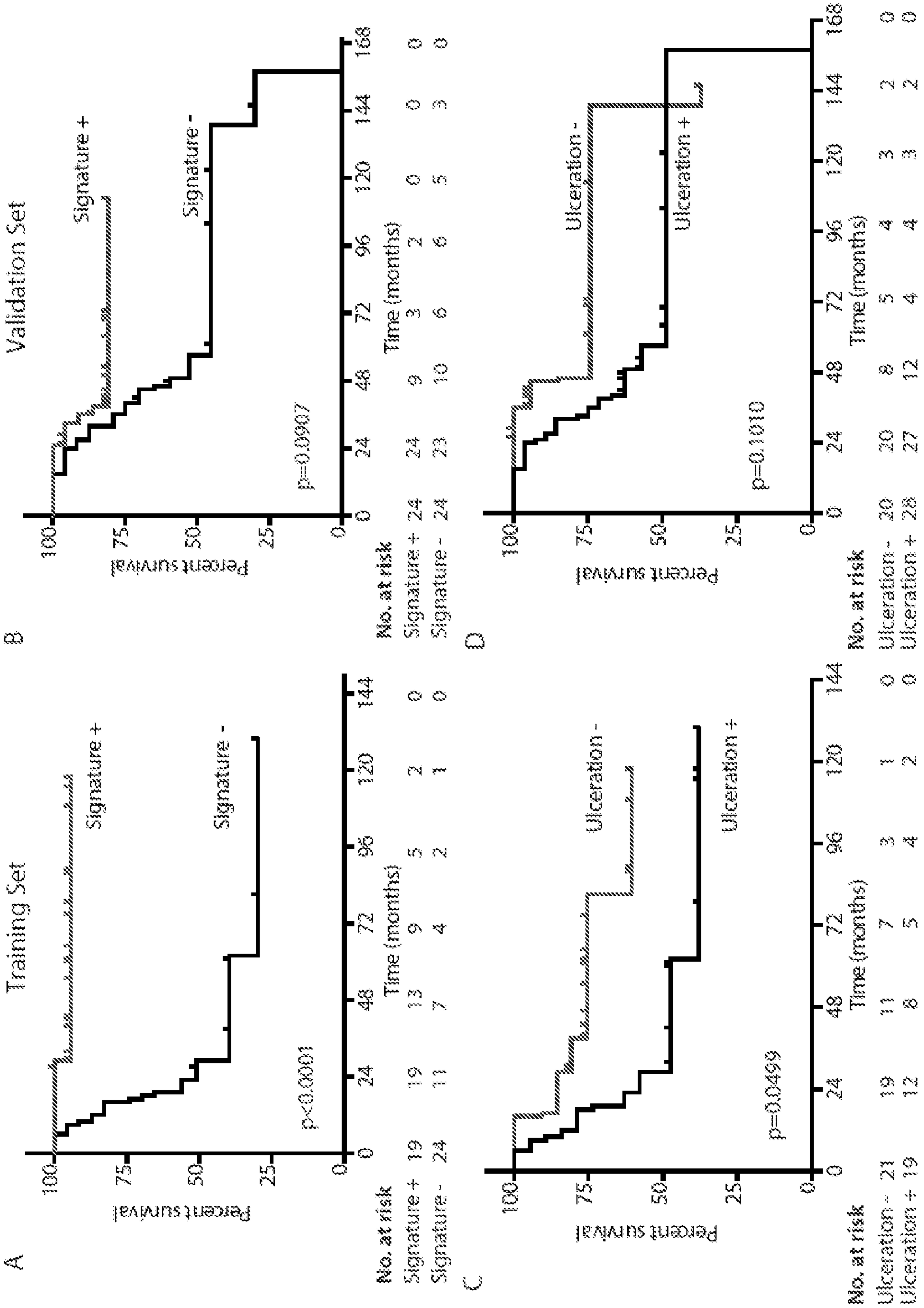


Figure 15, continued

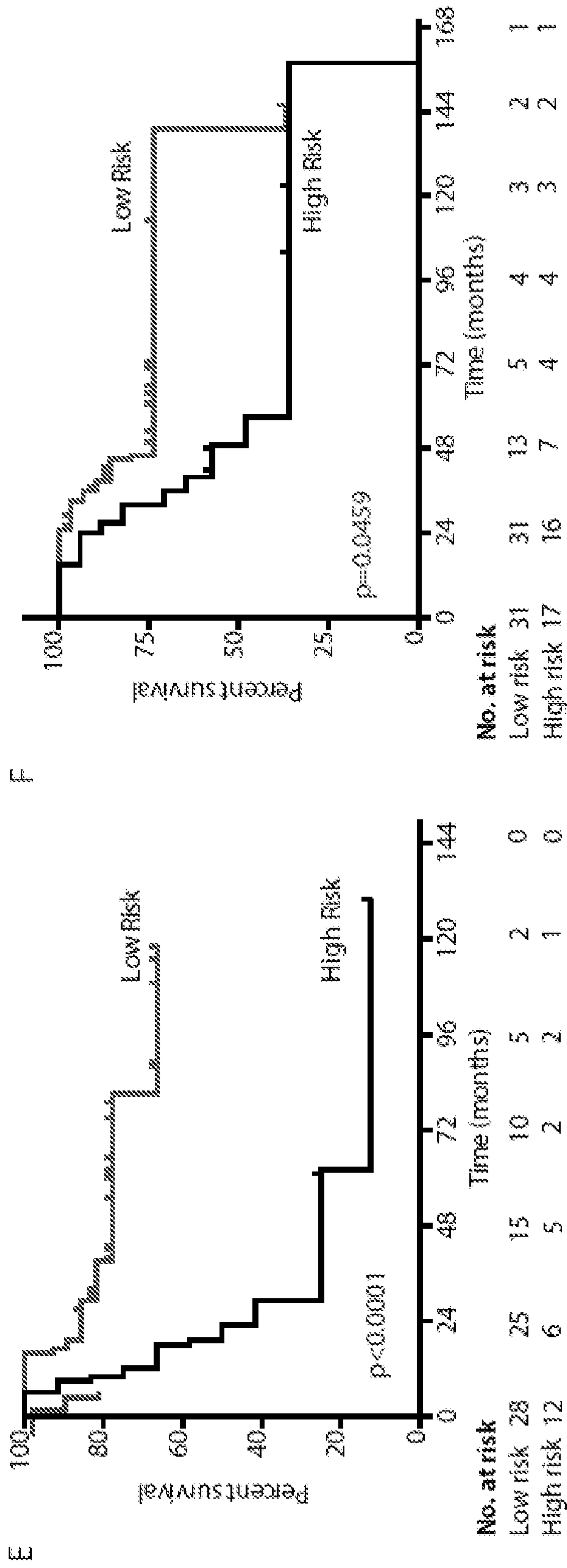
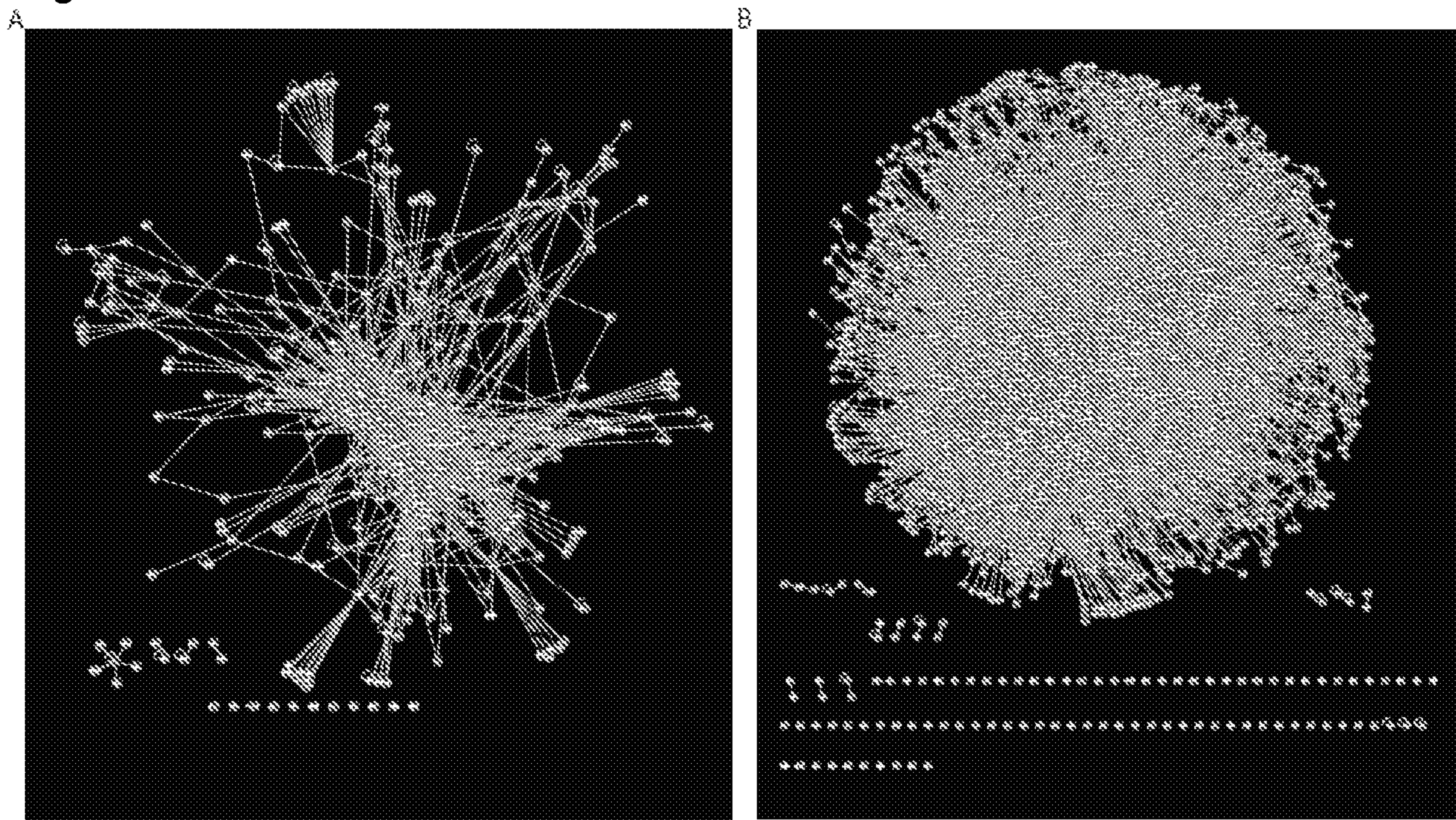


Figure 16



C

	Genes (Nodes)	Interactions (Edges)	Density	Average Local CC	Local CC SE	Average Local CC P-value	Global CC	Global CC P-value
53-gene panel network	406	1927	0.023	0.286	0.014	0	0.192	0
446-gene panel network	2474	14866	0.005	0.206	0.005	0.511	0.092	0
Fold Change			4.815	1.39			2.08	

Figure 17

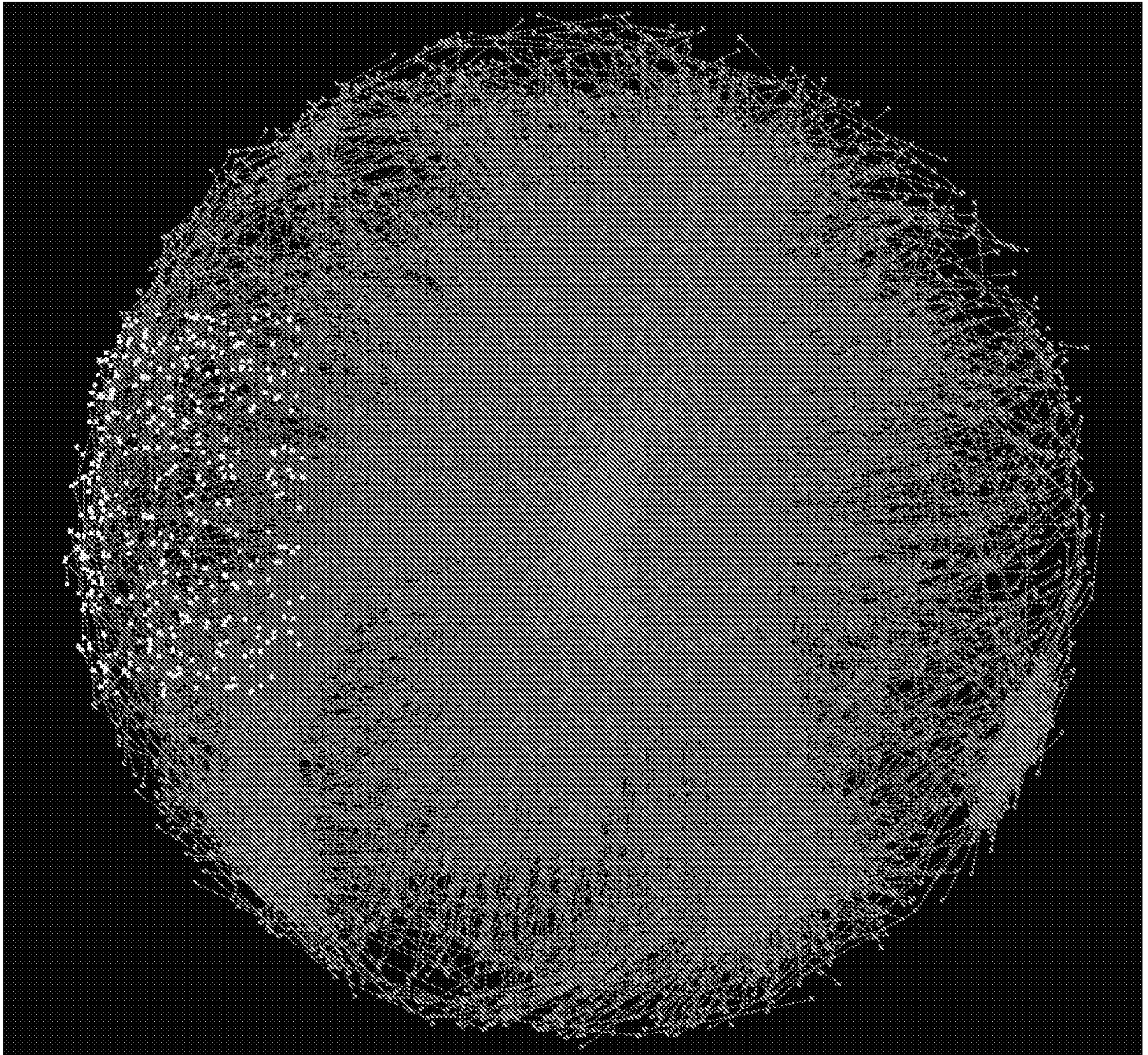


Figure 18

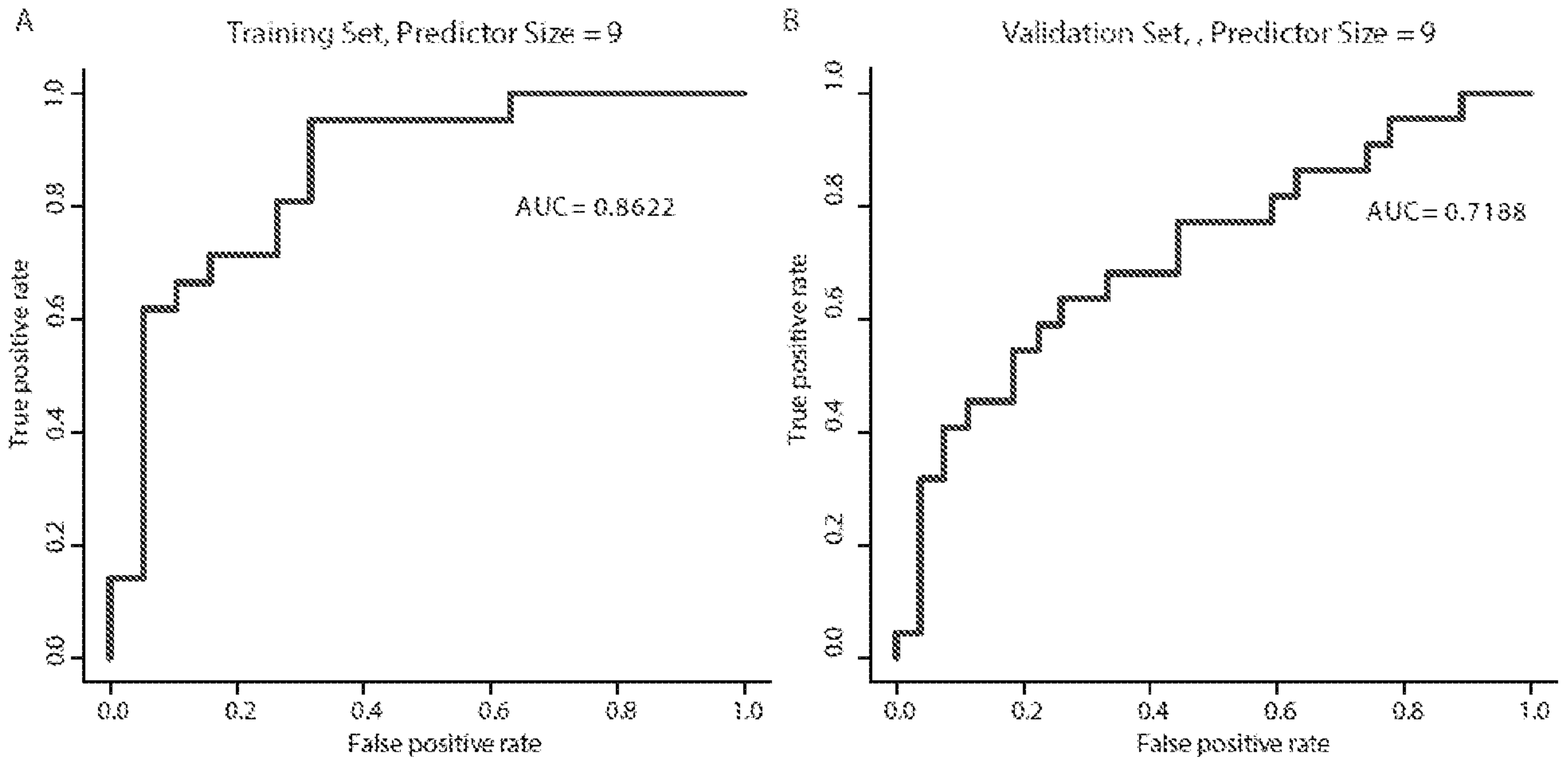


Figure 19

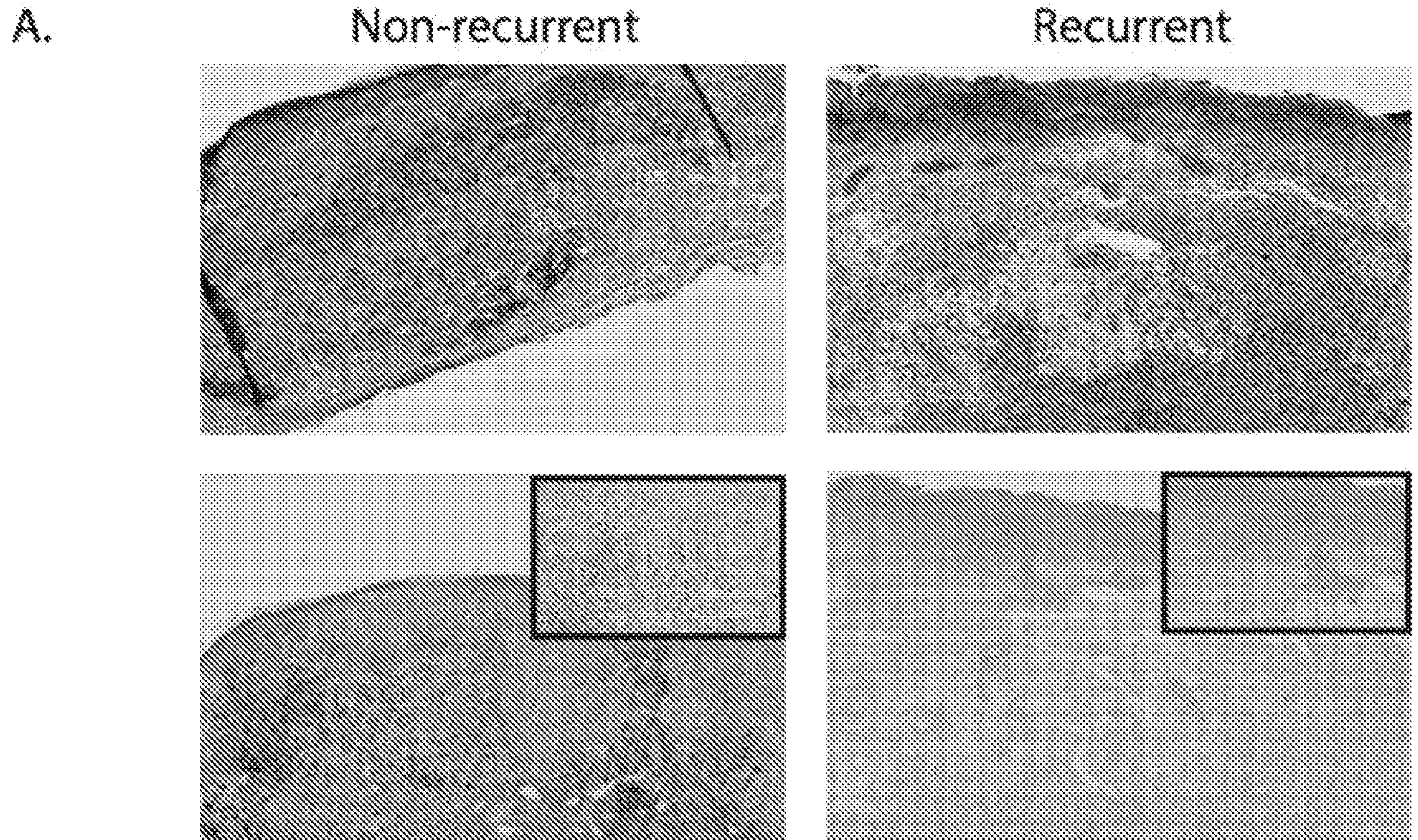


Figure 19, continued

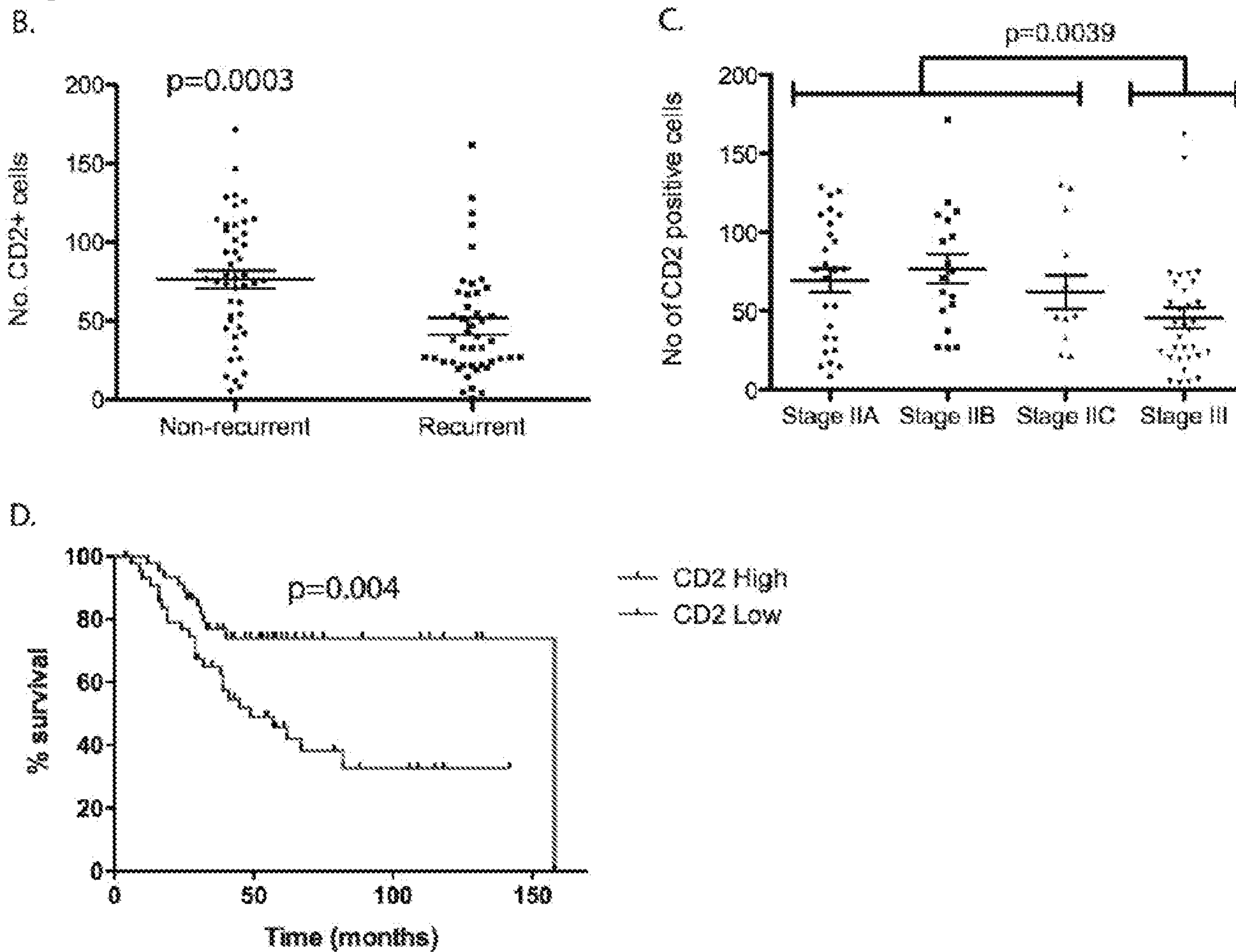


Figure 20

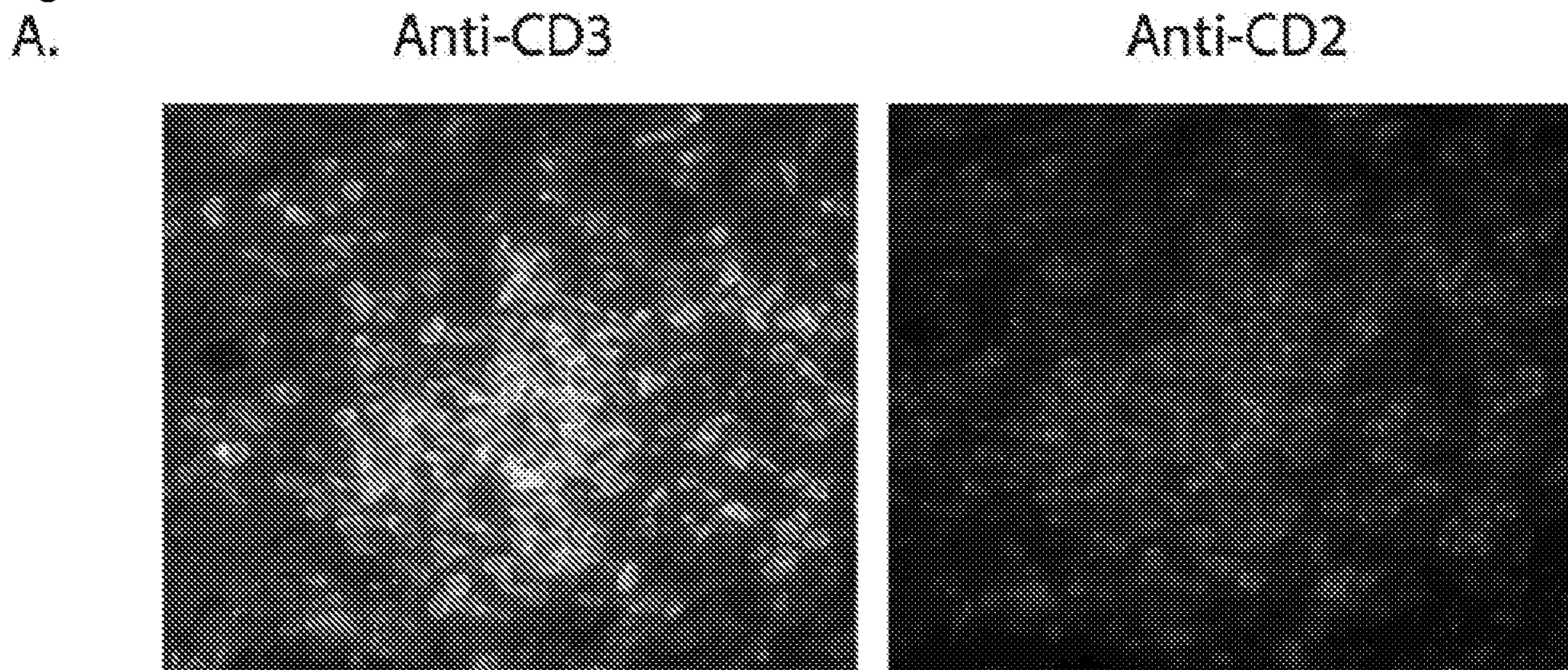


Figure 20, continued

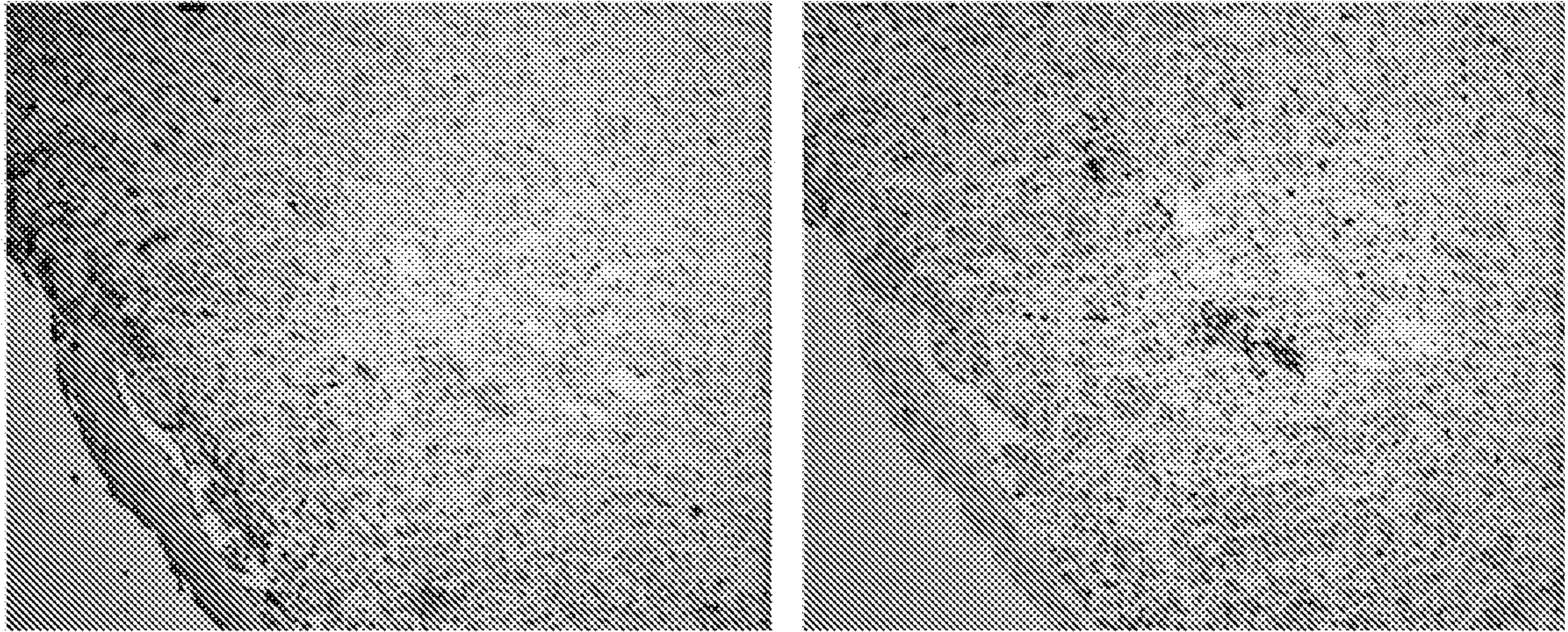


Figure 21

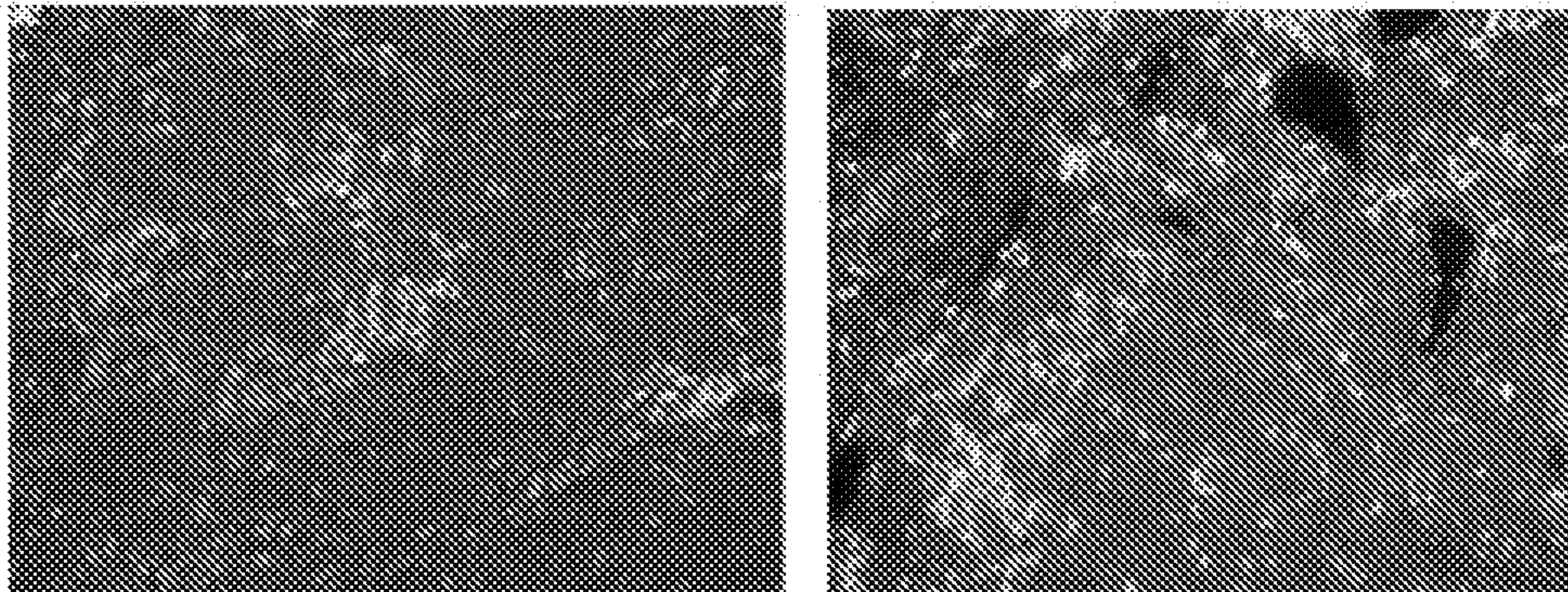
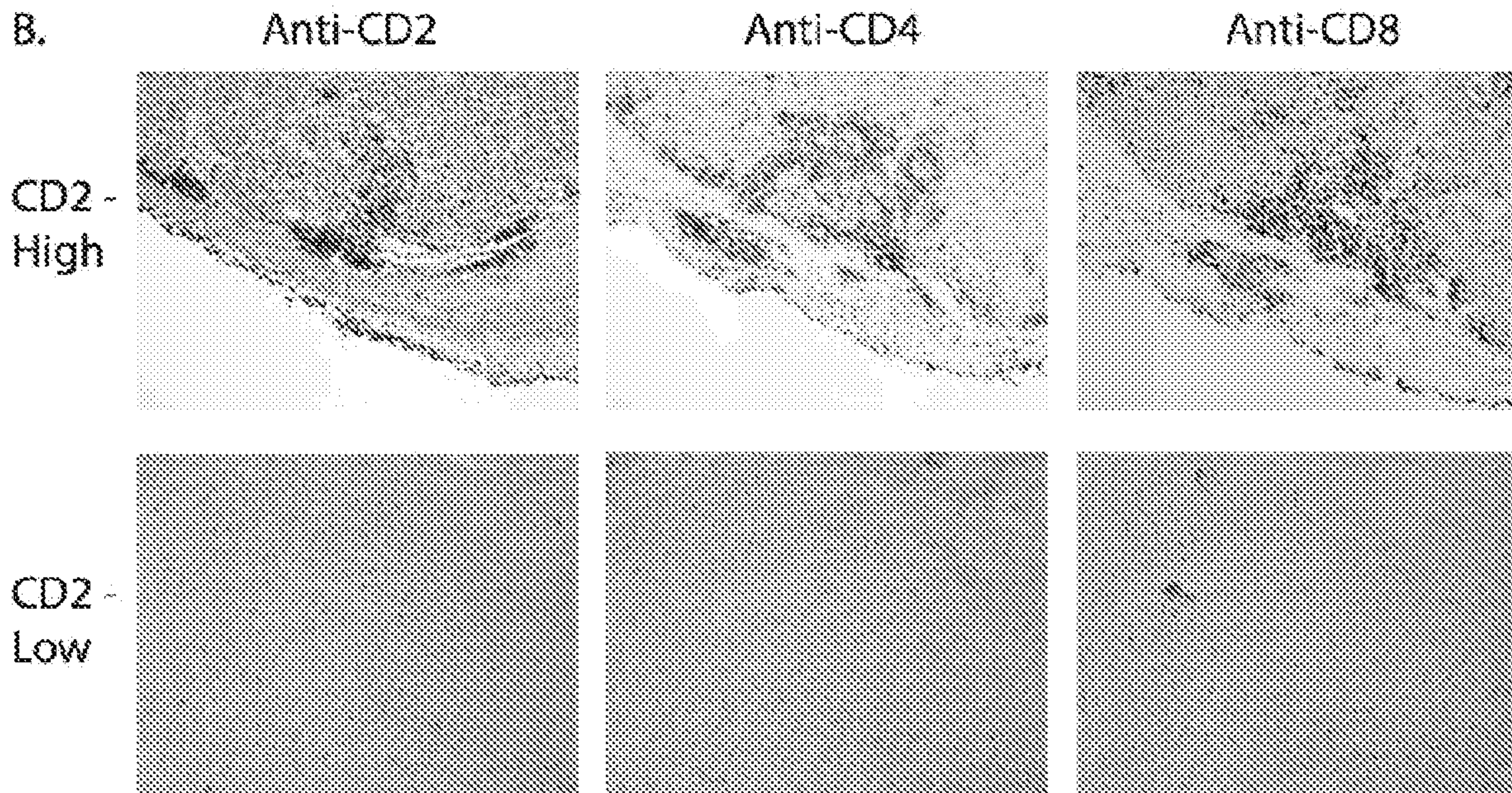


Figure 21 (continued)



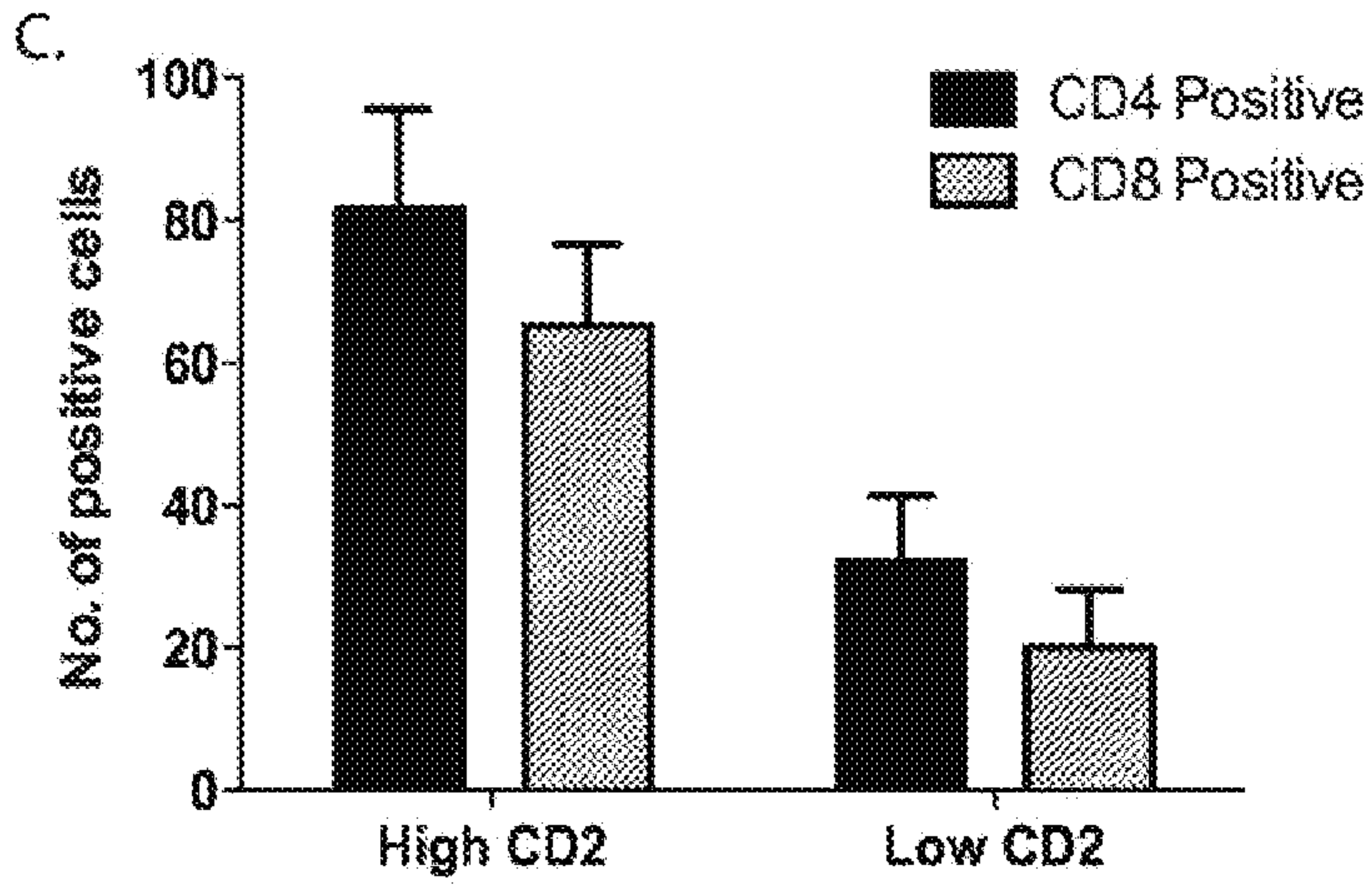


Figure 22

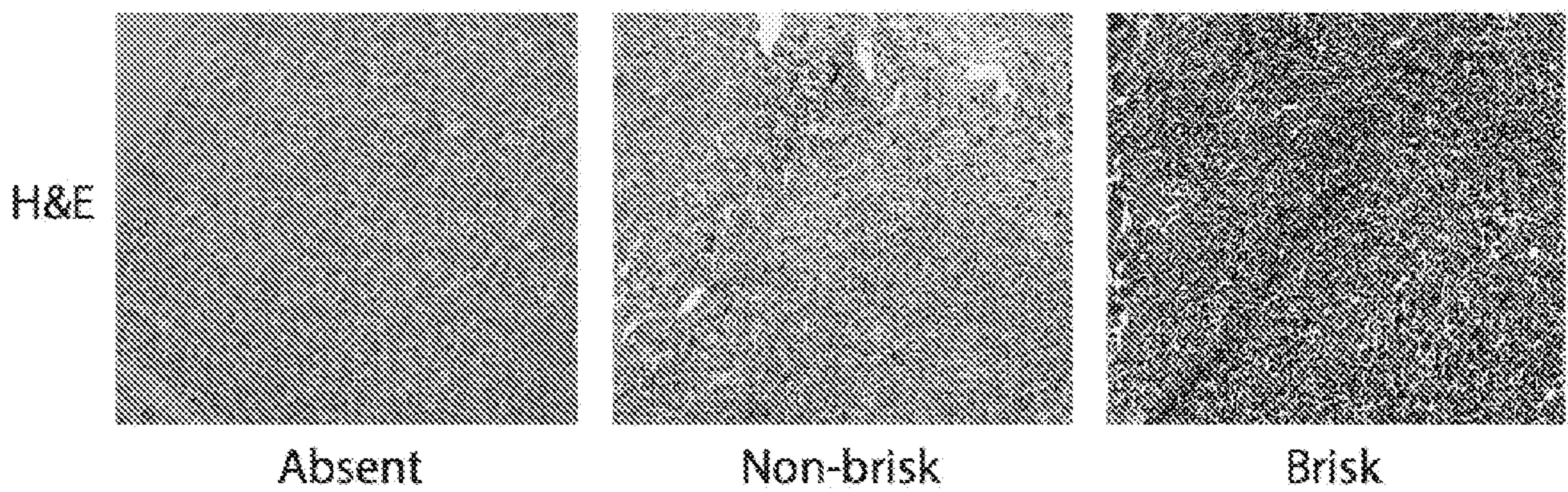


Figure 22 (continued)

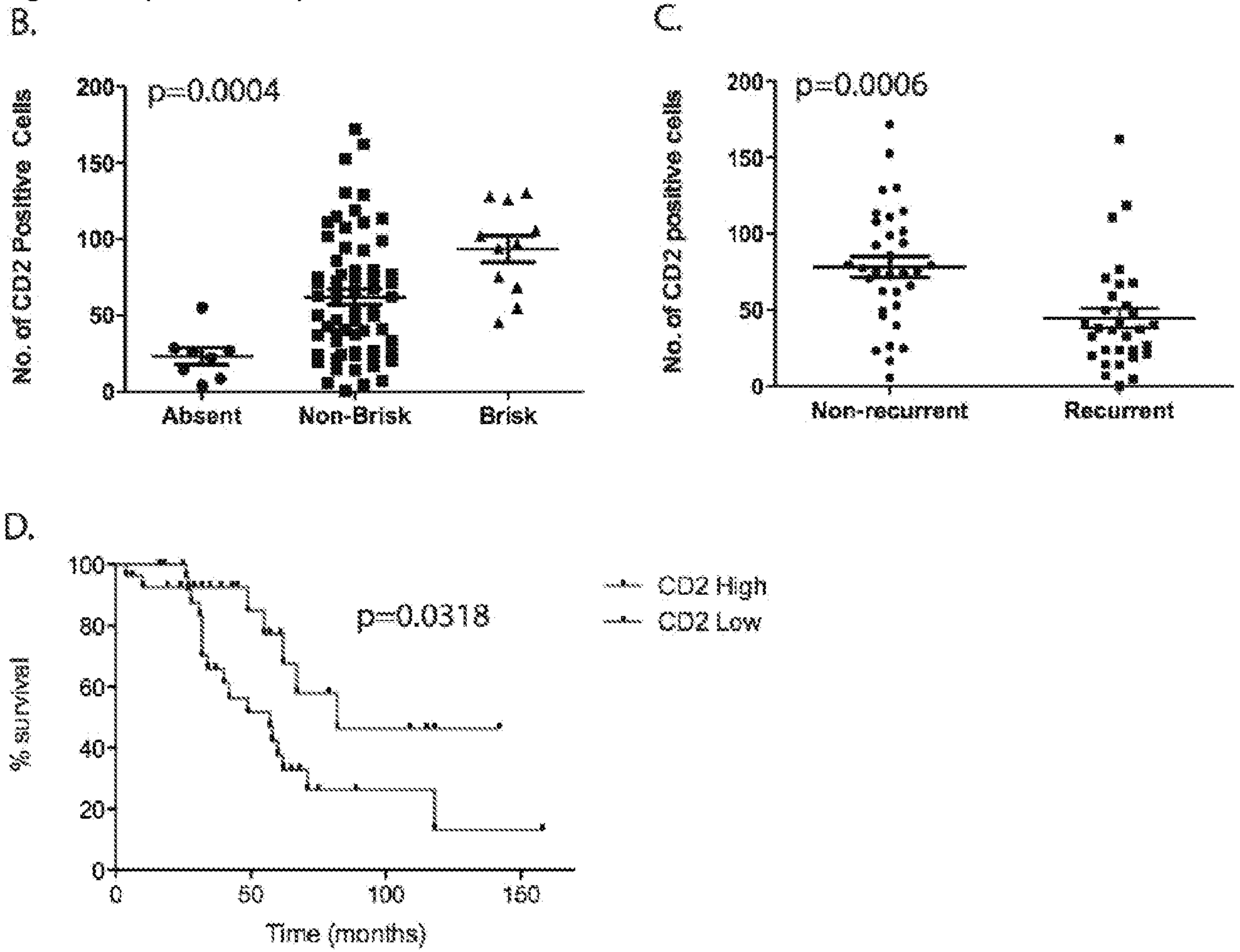


Figure 23

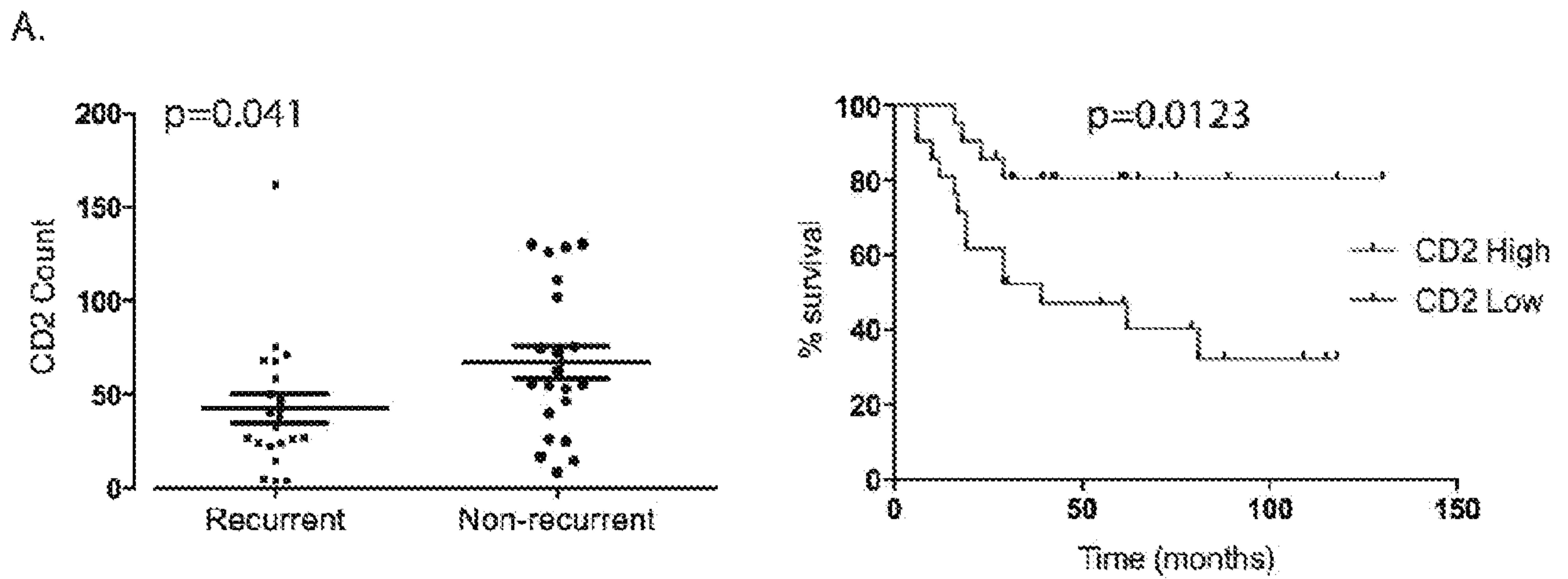
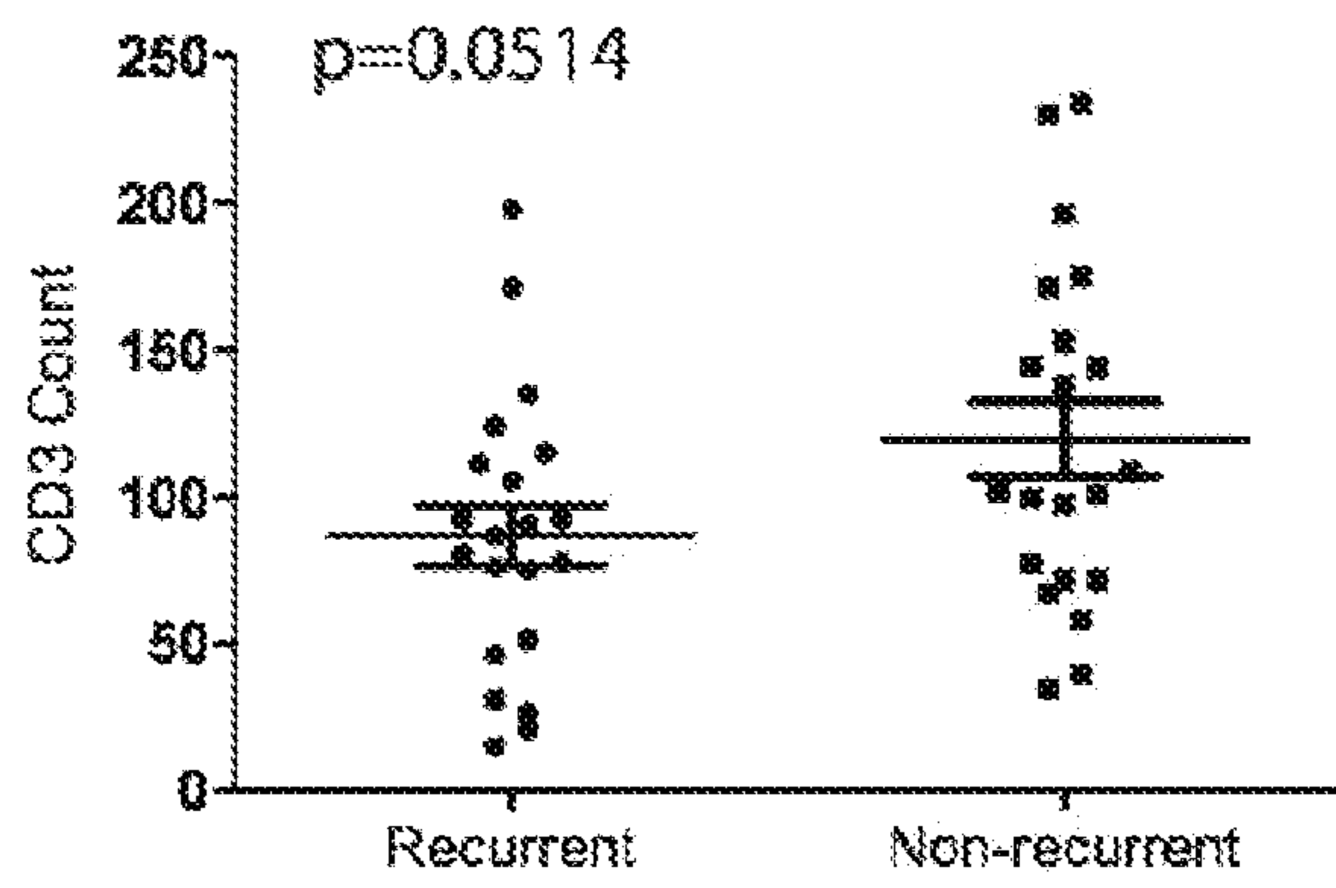


Figure 23 (continued)

B.



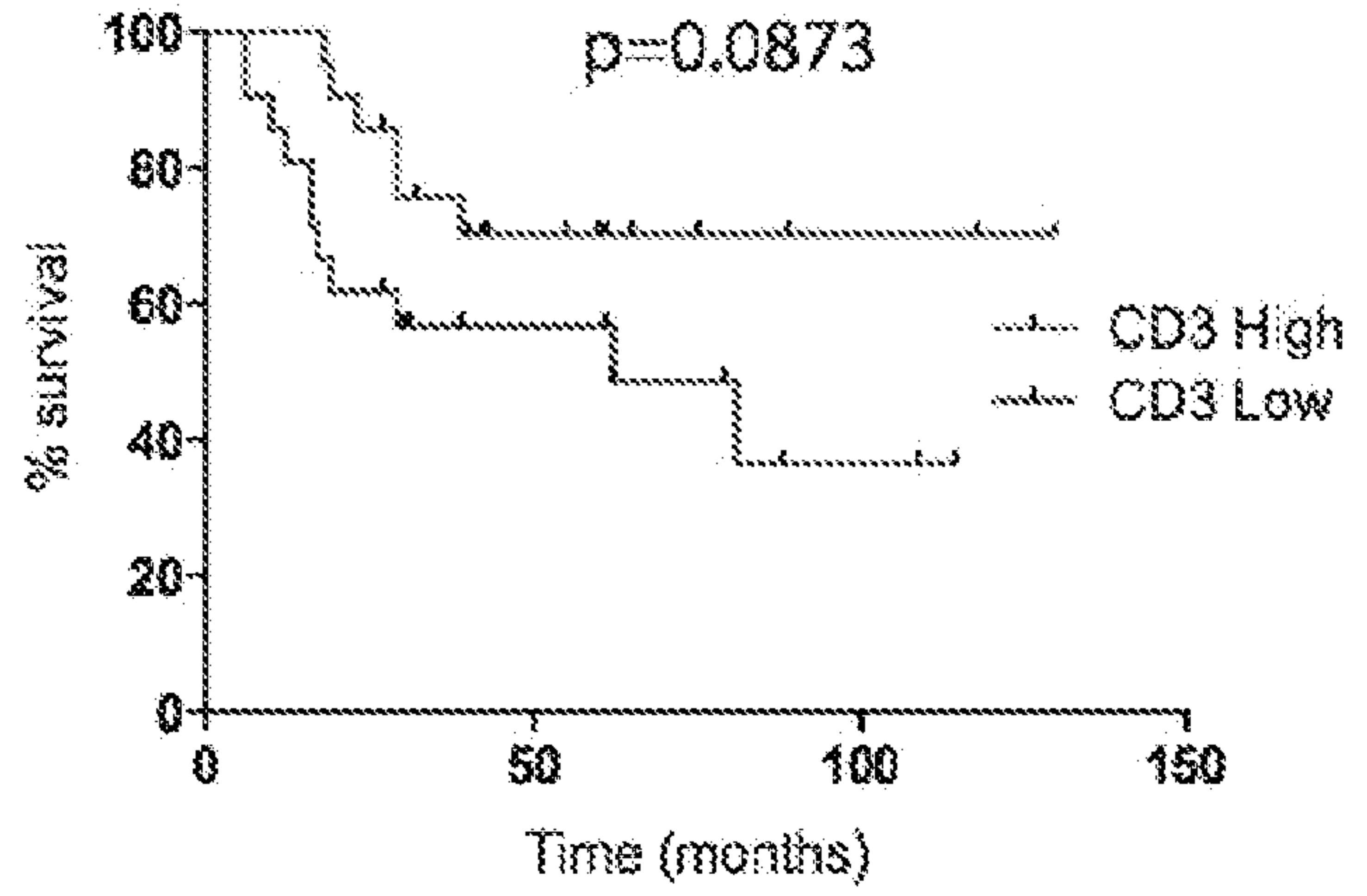
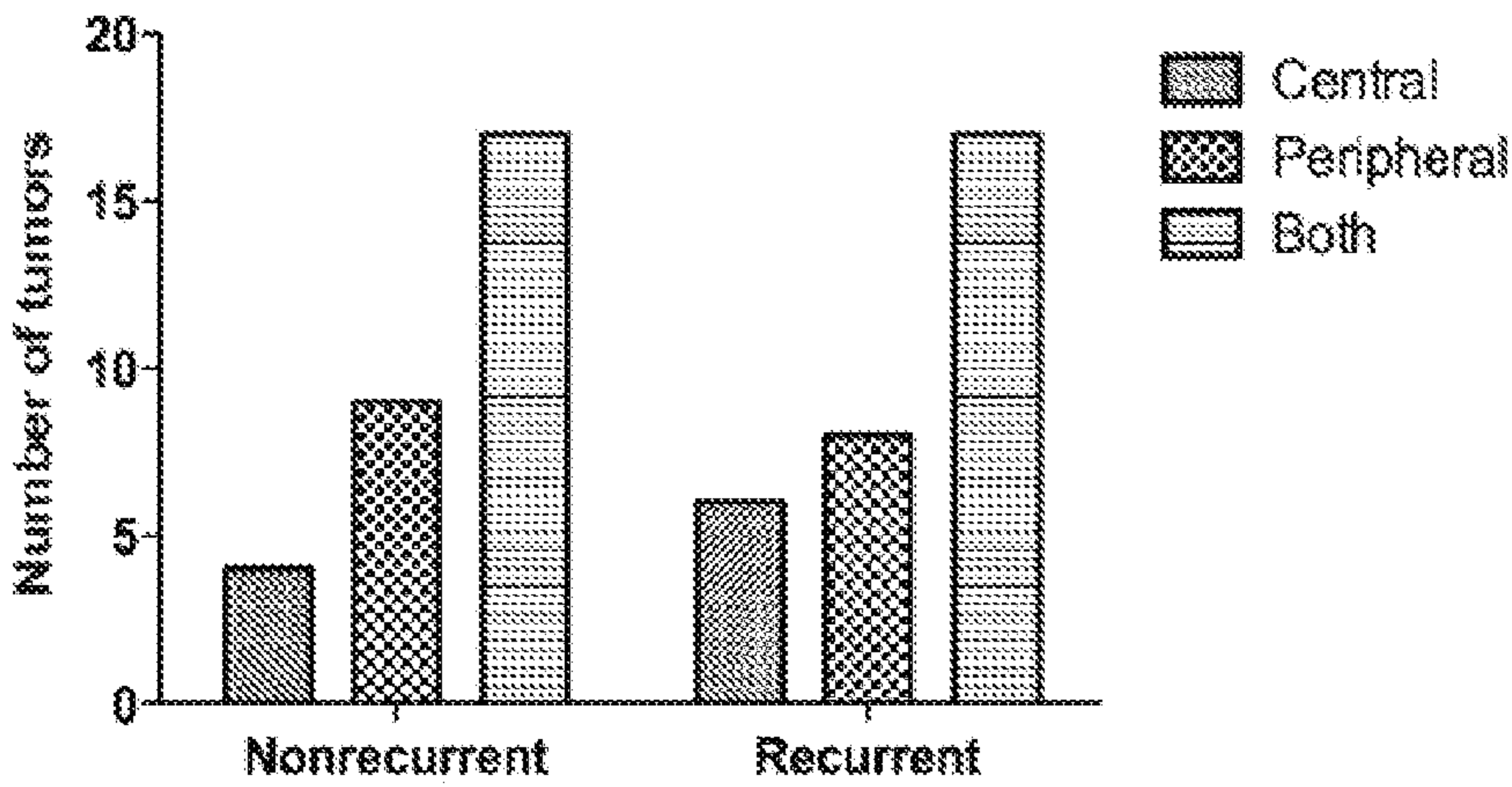


Figure 24

A.



B.

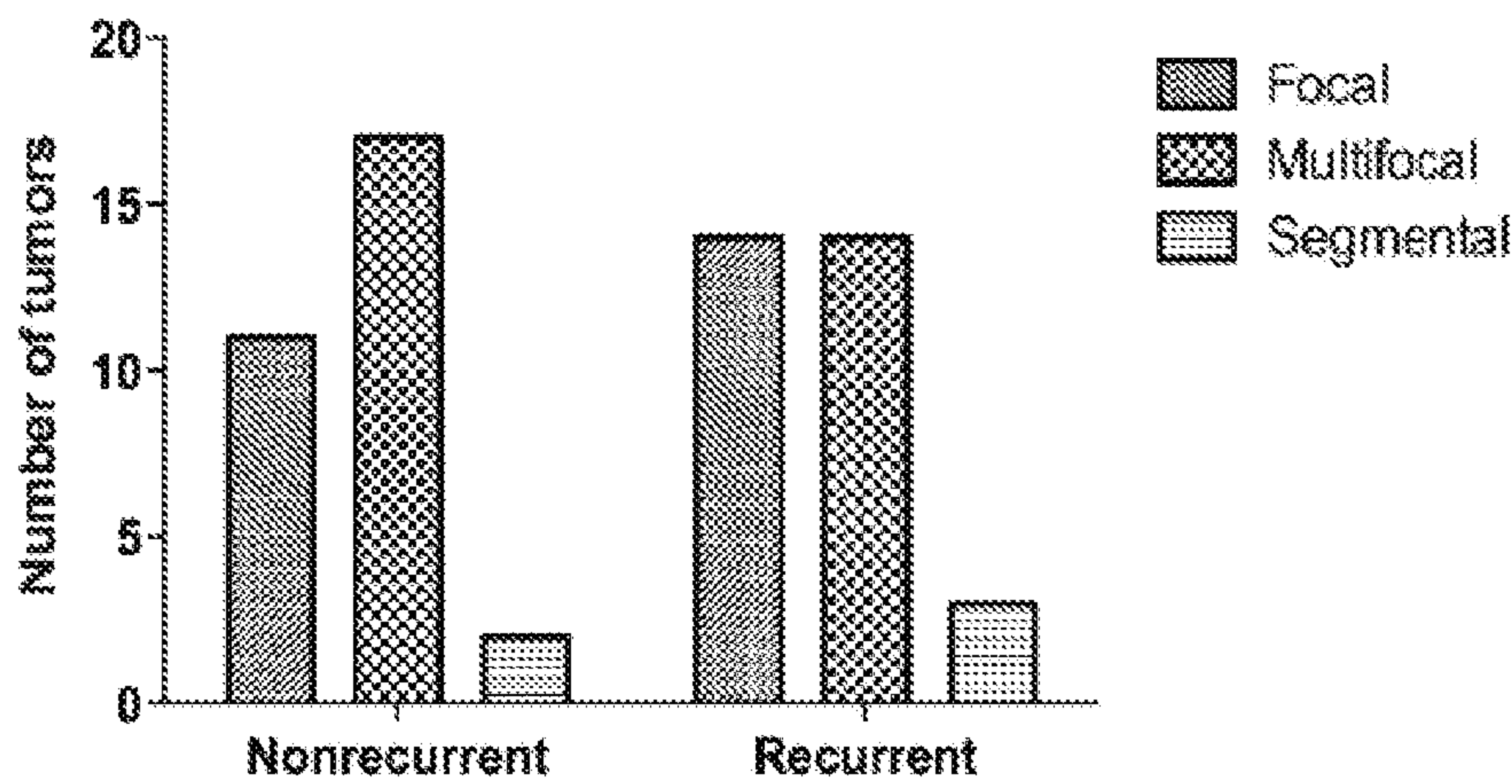


Figure 8