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

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
12 October 2017 (12.10.2017)

(10) **International Publication Number**
WO 2017/177011 A1

(51) **International Patent Classification**:
A61K 39/395 (2006.01) *C12Q 1/68* (2006.01)
A61P 35/00 (2006.01)

(21) **International Application Number**:
PCT/US2017/026369

(22) **International Filing Date**:
6 April 2017 (06.04.2017)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
62/319,284 6 April 2016 (06.04.2016) US
62/320,306 8 April 2016 (08.04.2016) US

(71) **Applicant**: **UNIVERSITY OF FLORIDA RESEARCH FOUNDATAION, INC.** [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US).

(72) **Inventor**: **LAMBA, Jatinder, Kaur**; 14239 Nw 26 Avenue, Gainesville, FL 32606-5381 (US).

(74) **Agent**: **GATES, Edward, R.**; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

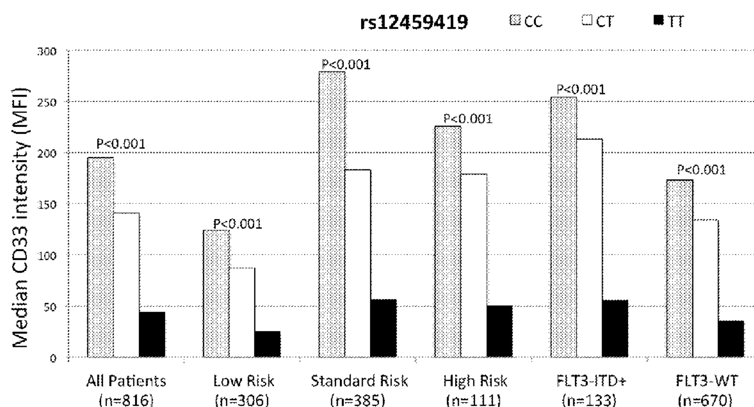
(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR,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

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

(54) **Title**: BIOMARKERS FOR ANTI-LEUKEMIC THERAPY

Figure 5



(57) **Abstract**: Aspects of the invention relate to detecting single nucleotide polymorphisms for CD33 and determining whether a subject is likely to respond to treatment with agents that selectively bind to CD33 based on the genotype of the subject for one or more CD33 single nucleotide polymorphisms.

WO 2017/177011 A1

BIOMARKERS FOR ANTI-LEUKEMIC THERAPY

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. §119 of United States provisional applications 62/319,284, filed April 6, 2016, and 62/320,306, filed April 8, 2016, the entire contents of each of which is incorporated herein by reference.

FEDERALLY SPONSORED RESEARCH

This invention was made with government support under Grant No. CA155524 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD OF THE INVENTION

The disclosure relates, at least in part, to single nucleotide polymorphisms (SNPs) that can be used to predict whether or not a subject with cancer may benefit from a particular treatment.

BACKGROUND OF THE INVENTION

Acute myeloid leukemia (AML) is a heterogeneous disease with dismal outcome. While intensive chemotherapy and hematopoietic stem cell transplantation remains the mainstay of current AML therapy, targeted therapies such as monoclonal antibodies and small molecule inhibitors have emerged as promising approaches. Gemtuzumab Ozogamicin (GO) is a humanized anti-CD33 antibody linked with cytotoxin calicheamicin, which targets AML cells, a majority of which express CD33 antigen. GO was approved in 2000 for treatment of relapsed AML in older patients and since then multiple clinical trials have investigated GO in AML. GO was withdrawn from the market due to lack of benefit and high early mortality observed in the SWOG S0106 study. Reassessment of the results from that trial as well as several other follow-up studies showed improved outcome in subset of patients.

SUMMARY OF THE INVENTION

Described herein are methods to determine CD33 single nucleotide polymorphism (SNP) genotype(s) in a subject with a cancer expressing CD33 and to determine whether the

subject is likely or unlikely to benefit from a particular treatment. Methods of treatment are also described.

In some aspects, the disclosure relates to a method of treating a subject with a cancer expressing CD33 comprising: performing an assay to detect the genotype of the subject for
5 the CD33 single-nucleotide polymorphism rs12459419, wherein the genotype may be CC, TC, or TT; and administering a therapeutically effective amount of an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide polymorphism rs12459419.

In some embodiments, the agent that selectively binds to CD33 comprises an antibody
10 that selectively binds CD33, or an antigen binding fragment thereof. In some embodiments, the antibody that selectively binds CD33 is a humanized antibody. In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof, conjugated to a toxin. In some embodiments, the agent that selectively binds to CD33 selectively binds to amino acids encoded by exon 2 of CD33.
15 In some embodiments, the agent that selectively binds to CD33 is gemtuzumab ozogamicin (GO), hP67.7, or SGN-33A.

In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject. In some embodiments, the subject is treated with chemotherapy within thirty days of the administration of the antibody.

20 In some embodiments, the assay is performed by DNA sequencing analysis. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL). In some embodiments, the assay is a hybridization assay. In some embodiments, the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis; anemia;
25 leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

In some aspects, the disclosure provides a method for determining whether a subject with a cancer expressing CD33 is likely to benefit from treatment with an agent that selectively binds to CD33 comprising: providing tissue from a subject who has been
30 diagnosed with the cancer; performing an assay on the tissue, or on a derivative of the tissue, to detect the genotype of the subject for the CD33 single-nucleotide polymorphism

rs12459419, wherein the genotype may be CC, TC, or TT; wherein the subject is likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide polymorphism rs12459419 and the subject is not likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a TC or TT genotype for the CD33 single-nucleotide polymorphism rs12459419.

In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof. In some embodiments, the agent that selectively binds to CD33 selectively binds to amino acids encoded by exon 2 of CD33. In some embodiments, the agent that selectively binds to CD33 is gemtuzumab ozogamicin (GO), hP67.7, or SGN-33A. In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject. In some embodiments, the assay comprises performing DNA sequencing analysis. In some embodiments, the assay comprises contacting a derivative of the tissue with a nucleic acid probe. In some embodiments, the assay is a hybridization assay.

In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL). In some embodiments, the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis; anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

In some aspects, the disclosure provides a method for detecting a polymorphism, comprising obtaining a biological sample of a subject that has a cancer expressing CD33, performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism (SNP) rs12459419 controlling expression of exon 2, wherein the genotype may be CC, TC, or TT and wherein the presence of the CC genotype in rs12459419 indicates expression of exon 2 of CD33.

In some embodiments, the assay is a hybridization assay comprising a probe that hybridizes specifically to the CC genotype but not the TC or the TT genotypes. In some embodiments, the hybridization assay further comprises a probe that hybridizes specifically to the TT genotype but not the CC or the TC genotypes and a probe that hybridizes specifically to the TC genotype but not the CC or the TT genotypes. In some embodiments, the method comprises detecting specific hybridization of the probes that binds specifically the CD33 single-nucleotide polymorphism rs12459419 to their respective genotype.

In some embodiments, the hybridization assay comprises detecting hybridization of a probe that binds to the a nucleic acid from the biological sample, and detecting a variant nucleic acid of CD33 single-nucleotide polymorphism rs12459419 in the sample when hybridization is detected. In some embodiments, the method further comprises performing a hybridization assay with the probes and a control genotype.

In some embodiments, the assay is a genomic sequencing assay. In some embodiments, the assay is a DNA sequencing, RNA sequencing, primer extension, enzyme-based, restriction fragment length polymorphism, PCR-based, PCR-RFLP, allele-specific PCR, flap endonuclease, 5'- nuclease, oligonucleotide ligation, SNPlex, surveyor nuclease, dynamic allele-specific hybridization, molecular beacons, or SNP microarray assay. In some embodiments, the genomic assay comprises direct sequencing of a nucleic acid containing polymorphism rs12459419, and detecting the presence of the CC, TC or TT genotype. In some embodiments, the nucleic acid is DNA, genomic DNA, RNA, cDNA, hnRNA or mRNA.

In some embodiments, the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis; anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

In some embodiments, the probe comprises a nucleotide sequence complementary to a sequence listed within Table 1. The probe can comprise additional nucleotides. In some embodiments, the probe comprises a nucleotide sequence complementary to nucleotides of SEQ ID NO:1. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

In some aspects, the disclosure provides a method for detecting a polymorphism, comprising obtaining a biological sample of a subject that has a cancer expressing CD33, performing an assay to detect the presence of amino acids encoded by exon 2 of CD33, wherein the presence of amino acids encoded by exon 2 of CD33 indicates expression of exon 2 of CD33 and presence of a CC genotype in single-nucleotide polymorphism rs12459419 of CD33. In some embodiments, the assay is an immunoassay. In some embodiments, the assay is a protein sequencing assay.

In some aspects, the disclosure provides a kit comprising an agent that selectively binds to CD33 and instructions indicating the use of the agent to treat a subject if the genotype of the subject for the CD33 single-nucleotide polymorphism rs12459419 is CC.

In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof. In some embodiments, the antibody that selectively binds CD33 is a humanized antibody. In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof, conjugated to a toxin. In some embodiments, the agent that selectively binds to CD33 selectively binds to exon 2 of CD33. In some embodiments, the agent that selectively binds to CD33 is gemtuzumab ozogamicin. In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject.

In some aspects, the disclosure provides a method for determining a CD33SNP_Score for a subject comprising: determining genotype scores of the subject for the CD33 single-nucleotide polymorphisms (SNPs) rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069, wherein the genotype score for a single nucleotide polymorphism (SNP) with two wild-type alleles is 0, the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with one wild-type allele and one variant allele is -1, the genotype score for the SNP rs2455069 with one wild-type allele and one variant allele is 1, the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with two variant alleles is -2, and the genotype score for the SNP rs2455069 with two variant alleles is 2, and adding the genotype scores to yield the CD33SNP_Score.

In some embodiments, the genotype of each SNP is determined by an assay. In some embodiments,the assay is performed by DNA sequencing analysis. In some embodiments, the assay is a hybridization assay. In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject.

In some aspects, the disclosure provides a method for determining a CD33SNP_Score for a subject comprising adding genotype scores of the subject for the CD33 single-nucleotide polymorphisms rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069 to yield the CD33SNP_Score, wherein the genotype score for a single nucleotide polymorphism (SNP) with two wild-type alleles is 0, the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with one wild-type allele and one variant allele is -1, the genotype score for the SNP rs2455069 with one wild-type

allele and one variant allele is 1, the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with two variant alleles is -2, and the genotype score for the SNP rs2455069 with two variant alleles is 2.

In some embodiments, the genotype of each SNP is determined by an assay. In some embodiments, the assay is performed by DNA sequencing analysis. In some embodiments, the assay is a hybridization assay. In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject.

In some aspects, the disclosure provides a method for determining whether a subject with cancer is likely to benefit from treatment with an agent that selectively binds to CD33 comprising: determining a CD33SNP_Score for the subject, wherein tissue is provided from the subject who has been diagnosed with the cancer; an assay is performed on the tissue, or on a derivative of the tissue, to detect the genotype of the subject for the CD33 single-nucleotide polymorphisms rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069, wherein the wild-type, single variant, or double variant genotype, respectively may be: CC, TC, or TT for rs12459419, CC, CA, or AA for rs3865444, GG, CG, or CC for rs1803254, GG, AG, or AA for rs35112940, or AA, AG, or GG for rs2455069, and wherein the subject is likely to benefit from treatment with an agent that selectively binds amino acids encoded by exon 2 of CD33 if the CD33SNP_Score for the subject is greater than or equal to zero.

In some embodiments, the assay is performed by DNA sequencing analysis. In some embodiments, the assay is a hybridization assay. In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL). In some embodiments, the tissue comprises CD33 expressing cells, *e.g.*, blast cells comprising CD33.

In some aspects, the disclosure provides a method for determining whether a subject with cancer expressing CD33 is likely to benefit from treatment with an agent that selectively binds to CD33 comprising: providing tissue from a subject who has been diagnosed with the cancer; performing an assay on the tissue, or on a derivative of the tissue, to detect the CD33 single-nucleotide polymorphism genotype of the subject for the CD33 single-nucleotide polymorphism rs12459419 or the CD33 single-nucleotide polymorphism rs3865444, and determining the single-nucleotide polymorphism genotype score, wherein the genotype score

may be 0, -1, or -2, wherein a score of 0 indicates that the subject is likely to benefit from treatment.

In some embodiments, the assay is performed by DNA sequencing analysis. In some embodiments, the assay is a hybridization assay. In some embodiments, the subject is a
5 pediatric subject. In some embodiments, the subject is an adult subject. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL). In some embodiments, the tissue comprises CD33 expressing cells, *e.g.*, blast cells comprising CD33.

In some aspects, the disclosure provides a method of treating a subject with a cancer
10 expressing CD33 comprising performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism rs3865444, wherein the genotype may be CC, CA, or AA; and administering a therapeutically effective amount of an agent that selectively binds to CD33 if the subject exhibits a GG genotype for the CD33 single-nucleotide polymorphism rs3865444.

15 In some aspects, the disclosure provides a method of treating a subject with a cancer expressing CD33 comprising performing an assay to detect the genotype of the subject for any one of the CD33 single-nucleotide polymorphisms rs1354106, rs3852865, and rs12985029, wherein the genotype may be wild-type, heterozygous variant, or homozygous variant; and administering a therapeutically effective amount of an agent that selectively
20 binds to CD33 if the subject exhibits a wild-type genotype for the CD33 single-nucleotide polymorphism rs1354106, rs3852865, or rs12985029.

In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof. In some embodiments, the antibody that selectively binds CD33 is a humanized antibody. In some embodiments, the
25 agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof , conjugated to a toxin. In some embodiments, the agent that selectively binds to CD33 selectively binds to amino acids encoded by exon 2 of CD33. In some embodiments, the agent that selectively binds to CD33 is gemtuzumab ozogamicin, hP67.7, or SGN-33A.

30 In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject. In some embodiments, the subject is treated with chemotherapy

within thirty days of the administration of the antibody. In some embodiments, the assay is performed by DNA sequencing analysis. In some embodiments, the assay is a hybridization assay. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

5 In some aspects, the disclosure provides a method for determining whether a subject with a cancer expressing CD33 is likely to benefit from treatment with an agent that selectively binds to CD33 comprising providing tissue from a subject who has been diagnosed with the cancer; performing an assay on the tissue, or on a derivative of the tissue, to detect the genotype of the subject for the CD33 single-nucleotide polymorphism 10 rs3865444, wherein the genotype may be CC, CA or AA; wherein the subject is likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide polymorphism rs3865444 and the subject is not likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a CA or AA genotype for the CD33 single-nucleotide polymorphism rs3865444.

15 In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof. In some embodiments, the agent that selectively binds to CD33 selectively binds to exon 2 of CD33. In some embodiments, the agent that selectively binds to CD33 is gemtuzumab ozogamicin, hP67.7, or SGN-33A. In some embodiments, the subject is a pediatric subject. In some 20 embodiments, the subject is an adult subject. In some embodiments, the assay comprises performing DNA sequencing analysis. In some embodiments, the assay comprises contacting a derivative of the tissue with a nucleic acid probe. In some embodiments, the assay is a hybridization assay. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

25 In some aspects, the disclosure provides a method for detecting a polymorphism, comprising obtaining a biological sample of a subject that has a cancer expressing CD33 and performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism (SNP) rs3865444 controlling expression of exon 2, wherein the genotype may be CC, CA, or AA and wherein the presence of the CC genotype in rs3865444 indicates 30 expression of exon 2 of CD33. In some embodiments, the assay is a hybridization assay comprising a probe that hybridizes specifically to the CC genotype but not the CA or the AA genotypes. In some embodiments, the hybridization assay further comprises a probe that

hybridizes specifically to the AA genotype but not the CC or the CA genotypes and a probe that hybridizes specifically to the CA genotype but not the CC or the AA genotypes. In some embodiments, the method comprises detecting specific hybridization of the probes that bind specifically the CD33 single-nucleotide polymorphism rs3865444 to their respective genotype.

In some embodiments, the hybridization assay comprises detecting hybridization of a probe that binds to the a nucleic acid from the biological sample, and detecting a variant nucleic acid of CD33 single-nucleotide polymorphism rs3865444 in the sample when hybridization is detected. In some embodiments, the method further comprises performing a hybridization assay with the probes and a control genotype.

In some embodiments, the assay is a genomic sequencing assay. In some embodiments, the assay is a DNA sequencing, RNA sequencing, primer extension, enzyme-based, restriction fragment length polymorphism, PCR-based, PCR-RFLP, allele-specific PCR, flap endonuclease, 5'- nuclease, oligonucleotide ligation, SNPlex, surveyor nuclease, dynamic allele-specific hybridization, molecular beacons, or SNP microarray assay.

In some embodiments, the genomic assay comprises direct sequencing of a nucleic acid containing polymorphism rs3865444, and detecting the presence of the CC, CA or AA genotype. In some embodiments, the nucleic acid is DNA, genomic DNA, RNA, cDNA, hnRNA or mRNA.

In some embodiments, the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis; anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma. In some embodiments, the probe comprises a nucleotide sequence complementary to a sequence listed within Table 1. The probe can comprise additional nucleotides. In some embodiments, the probe comprises a nucleotide sequence complementary to nucleotides of SEQ ID NO:2. In some embodiments, the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

In some aspects, the disclosure provides a kit comprising an agent that selectively binds to CD33 and instructions indicating the use of the agent to treat a subject if the genotype of the subject for the CD33 single-nucleotide polymorphism rs3865444 is CC. In some embodiments, the agent that selectively binds to CD33 comprises an antibody that

selectively binds CD33, or an antigen binding fragment thereof. In some embodiments, the antibody that selectively binds CD33 is a humanized antibody. In some embodiments, the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof, conjugated to a toxin. In some embodiments, the agent 5 that selectively binds to CD33 selectively binds to exon 2 of CD33. In some embodiments, the agent that selectively binds to CD33 is gemtuzumab ozogamicin (GO). In some embodiments, the subject is a pediatric subject. In some embodiments, the subject is an adult subject.

Each of the limitations of the invention can encompass various embodiments of the 10 invention. It is, therefore, anticipated that each of the limitations of the invention involving any one element or combinations of elements can be included in each aspect of the invention. This invention is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practiced or of being 15 carried out in various ways.

BRIEF DESCRIPTION OF DRAWINGS

The accompanying drawings are not intended to be drawn to scale. The figures are illustrative only and are not required for enablement of the disclosure. For purposes of clarity, not every component may be labeled in every drawing. In the drawings:

20 Figure 1 shows differences in risk of relapse (RR) and in disease free survival (DFS) from end of course 1 between GO vs No-GO arm associated with rs12459419 genotypes.

Figure 2 shows risk of relapse from end of course 1 in GO vs. No-GO arm based on different genotype groups for each of rs1803254, rs35112940, and rs2455069.

Figure 3 shows association of CD33 SNPs with outcome within GO or No-GO arm. 25 For SNP rs35112940, presence of AG or AA genotype was associated with higher risk of relapse in GO arm but not in No-GO arm.

Figure 4 shows association of rs12459419, rs1803254, rs35112940, and rs2455069 with CD33 expression determined as mean fluorescence intensity in diagnostic leukemic blasts obtained from de novo AML patients from AAML0531 clinical trial. CD33 levels 30 were determined in the diagnostic leukemic blast by multiparameter flow cytometry. Y-axis

represents Log10 CD33 mean fluorescence intensity (MFI) and X-axis represents SNP genotype. Plots show medians as a line between boxes representing the first and third quartiles; the whiskers represent the range after excluding the outliers. The outliers are defined as data points that fall outside of the first and third quartiles by more than 1.5 times
5 the interquartile range. Circles that are outside of the whiskers represent outliers.

Figure 5 shows association of rs12459419 SNP with CD33 intensity in different risk groups as well as FLT-3 status in AML patients.

Figure 6A shows that rs12459419 C>T alters binding to an exonic splicing enhancer (ESE) protein.

10 Figure 6B shows real-time PCR results from assay using isoform specific primers of CD33-D2 spliced isoform and rs12459419 genotype.

Figure 6C shows RNA-seq results of CD33-D2 spliced isoform and rs12459419.

Figure 7 shows that the rs12459419 T allele is associated with lower expression of exon 2 as reflected by exon2/exon 4 ratio.

15 Figure 8A shows the association of CD33SNP_Score with CD33 intensity. CD33SNP_Score was created for each patient using genotype information from four SNPs: rs12459419, rs1803254, rs35112940, and rs2455069.

Figure 8B shows a dichotomized CD33SNP_Score with CD33 intensity. Dichotomized CD33SNP_Score can be ≥ 0 or < 0.

20 Figure 9 shows association of dichotomized CD33SNP_Scores, ≥ 0 and < 0 with response by arm for following clinical outcomes: disease-free survival (DFS) from end of induction 1 by arm in different CD33SNP_Score groups; risk of relapse (RR) from end of induction 1 by arm in different CD33SNP_Score groups.

DETAILED DESCRIPTION

25 The following detailed description is made by way of illustration of certain aspects of the disclosure. It is to be understood that other aspects are contemplated and may be made without departing from the scope or spirit of the present disclosure. The following detailed description, therefore, is not to be taken in a limiting sense. Scientific and technical terms used herein have meanings commonly used in the art unless otherwise specified. The

definitions provided herein are to facilitate understanding of certain terms used frequently herein and are not meant to limit the scope of the present disclosure. The singular forms "a", "an", and "the" encompass the plural, unless the content clearly dictates otherwise. The term "or" is generally employed in its sense including "and/or" unless the content clearly dictates
5 otherwise.

Aspects of the disclosure relate to treating cancers that express CD33. For example, a subject with acute myeloid leukemia, in which the blast cells of acute myeloid leukemia express CD33, may be treated as disclosed herein.

CD33 and cancers expressing CD33. CD33 or sialic acid binding Ig-like lectin 3
10 (SIGLEC-3, Siglec-3) is a transmembrane receptor expressed on cells of myeloid lineage that binds sialic acids. It is a member of the SIGLEC family of lectins. The extracellular portion of the CD33 receptor contains two immunoglobulin domains - one IgV and one IgC2 domain and the intracellular portion contains immunoreceptor tyrosine-based inhibitory motifs (ITIMs) that are implicated in inhibition of cellular activity. CD33 is found on cells of
15 myeloid lineage and can also be found on some lymphoid cells. CD33 is expressed on blast cells of acute myeloid leukemia (AML) and is detected on blasts of 85-90 percent of subjects presenting with AML.

Embodiments of the disclosure relate to treating a subject having a cancer expressing CD33. Cancers that express CD33 include hematopoietic cancers, *e.g.*, acute myeloid
20 leukemia (AML), acute lymphoblastic leukemia (ALL), and acute promyelocytic leukemia (APL). In some embodiments, the subject has AML, ALL, or APL.

Agents that selectively bind to or target CD33. Aspects of the disclosure relate to administering agents that selectively bind CD33. An agent that selectively binds to CD33 without limitation can be, *e.g.*, an antibody or an antigen-binding fragment thereof, a protein
25 or peptide, a small molecule, or a nucleic acid. An agent that selectively binds to CD33 can bind to nucleic acids or amino acids of the CD33 sequence. An agent that selectively binds CD33 can bind to any region of CD33. In some embodiments, an agent that selectively binds CD33 can bind to exon 2 of CD33. In some embodiments, an agent that selectively binds CD33 can bind to amino acids encoded by exon 2 of CD33. In some embodiments, the agent
30 is an antibody that selectively binds human myeloid lineage cells that express CD33, *e.g.*, human blast cells. In some embodiments, the agent, such as an antibody, selectively binds to the IgV domain of CD33. The IgV domain of CD33 is recognized by the antibody-drug

conjugate, gemtuzumab ozogamicin (GO; Mylotarg®; Pfizer/Wyeth-Ayerst Laboratories) and by the hP67.7 antibody.

In some embodiments, the agent that selectively binds CD33 is gemtuzumab ozogamicin (GO). GO is a recombinant, humanized anti-CD33 monoclonal antibody (IgG4 κ antibody hP67.6) linked with (covalently attached to) the cytotoxic antitumor antibiotic calicheamicin (N-acetyl-γ-calicheamicin) via a bifunctional linker (4-(4-acetylphenoxy) butanoic acid). GO targets AML blast cells, the majority of which express a CD33 antigen. In some embodiments, GO is used to treat the subject. In some embodiments, GO is used in a therapeutically effective amount. In some embodiments, a subject is also treated with chemotherapy.

CD33 is also a target of the anti-CD33 immunotoxin Vadastuximab talirine (SGN-CD33A) (Seattle Genetics). SGN-CD33A is an antibody-drug conjugate that may reduce multidrug resistance observed in response to treatment with GO. In some embodiments, SGN-CD33A is used to treat the subject. In some embodiments, GO and SGN-CD33A are used in combination to treat the subject. In some embodiments, the subject is also treated with chemotherapy.

In some embodiments, one or more other antibodies that selectively bind CD33, or antigen binding fragments thereof, may be used to treat the subject.

In some embodiments, an antibody or an antigen binding fragment thereof that selectively binds to CD33 is linked to a toxin to target CD33 expressing cancer cells in a subject. Any antibody that selectively binds CD33 may be used.

Isolated. In some embodiments, the antibodies and other therapeutic molecules used herein are isolated. Isolated means, in the context of an antibody or other biologic, the antibody or other biologic has been removed from its natural milieu or has been altered from its natural state. As such, isolated does not necessarily reflect the extent to which the molecule has been removed from its natural milieu or has been altered from its natural state. However, it will be understood that an antibody or other biologic that has been purified to some degree and to an extent to which it can be used for its intended therapeutic purpose is "isolated".

Antibody. In some embodiments, the methods herein employ antibodies. The term antibody is used in the broadest sense and specifically includes, for example, single

monoclonal antibodies, antibody compositions with polyepitopic specificity, single chain antibodies, and antigen-binding fragments of antibodies. An antibody may include an immunoglobulin constant domain from any immunoglobulin, such as IgG1, IgG2, IgG3, or IgG4 subtypes, IgA (including IgA1 and IgA2), IgE, IgD, or IgM.

5 In some embodiments, the methods herein employ antigen-binding fragments. As used herein, an antigen-binding fragment refers to a portion of an intact antibody that binds antigen. Examples of antibody fragments include Fab, Fab', F (ab')$_2$, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8 (10): 1057-1062 [1995]); and single-chain antibody molecules. Fv is the minimum antibody fragment containing a 10 complete antigen-recognition binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. In this configuration the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. The Fab fragment also contains the constant domain of the light chain and the first 15 constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. F(ab')$_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known. In some embodiments, the 20 antibody is a full length antibody (i.e., contains an Fc region, which can be IgG4 for example).

Humanized. In some embodiments, the antibodies used herein are humanized. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins (including full length immunoglobulins), immunoglobulin chains or fragments thereof (such 25 as Fv, Fab, Fab', F(ab')2, scFv or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from the non-human immunoglobulin. Humanized antibodies typically include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the 30 desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. 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 CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

A composite antibody is an antibody that contains sequence segments from different antibodies. Humanized antibodies can be formed of a composite of overlapping human sequences, with one segment of the CDR found in one human sequence and another segment of the same CDR found in another human sequence, each of the two sequences having common sequences at an overlapping region where the segments meet. In an embodiment, the composite human sequence is free of known T cell epitopes. In some embodiments, the composite human sequence does not elicit an immune response in humans. In any of the embodiments, the subject can be human and the antibody can be a humanized antibody. In any of the embodiments, the antibody can be a composite antibody. In any of the embodiments, the subject can be human and the antibody can be a fully human antibody. A fully human antibody is an antibody consisting only of human amino acid sequences.

Further details respecting antibodies and general methods of making antibodies can be found in U.S. patent application publication number 2013/ 0136735, the entire disclosure of which is incorporated herein by reference.

The antibodies selectively bind their targets, such as CD33 on blast cells. An antibody that selectively binds its target cell(s) means it has the ability to be used *in vitro* or *in vivo* to bind to and distinguish such target bearing tissue from other tissue types of the species, including other closely related cell types under the conditions in which the antibody is used, such as under physiologic conditions. In some embodiments, the antibody selectively binds human blast cells that express CD33. In some embodiments, the antibody selectively binds to any region of CD33. In some embodiments, the antibody selectively binds to the IgV domain of CD33. In some embodiments, the antibody is GO. In some embodiments, the antibody is SGN-CD33A. In some embodiments, the antibody is hP67.7. In some embodiments, the antibody is hP67.7 linked to a toxin. The antibody can be any antibody or antigen binding fragment thereof that selectively binds CD33 and is linked to a toxin.

Aspects of the invention relate to treatment with an antibody drug conjugate (ADC), such as an antibody or antigen binding fragment thereof that selectively binds to CD33, which is directly linked to a toxin or linked to a toxin through a linker.

ADCs and Toxins. Antibodies or antigen binding fragments thereof of the disclosure may be conjugated (covalently or non-covalently linked) to a toxin or they may be linked to a toxin through a linker. The toxin may be any toxin that can elicit a therapeutic effect. The toxin may be an enzymatically active toxin of bacterial, fungal, plant or animal origin or a synthetic toxin, or fragments thereof.

The use of antibody-drug conjugates (ADCs), *e.g.*, immunoconjugates, for the local delivery of cytotoxic or cytostatic agents to kill or inhibit tumor cells in the treatment of cancer (Syrigos and Epenetos (1999) Anticancer Research 19:605-614; Niculescu-Duvaz and Springer (1997) Adv. Drg Del. Rev. 26:151-172; U.S. Pat. No. 4,975,278) theoretically allows targeted delivery of the drug moiety to tumors, and intracellular accumulation therein, where systemic administration of these unconjugated drug agents may result in unacceptable levels of toxicity to normal cells as well as the tumor cells sought to be eliminated (Baldwin et al., Lancet pp., 1986: 603-05; Thorpe, (1985) "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review," in Monoclonal Antibodies '84: Biological And Clinical Applications, A. Pinchera et al. (ed.s), pp. 475-506). Efforts to design and refine ADC have focused on the selectivity of monoclonal antibodies (mAbs) as well as drug-linking and drug-releasing properties. Both polyclonal antibodies and monoclonal antibodies have been reported as useful in these strategies (Rowland et al., (1986) Cancer Immunol. Immunother., 21:183-87). Drugs used in these methods include daunomycin, doxorubicin, methotrexate, and vindesine (Rowland et al., (1986) supra).

Toxins useful as therapeutics are known to those skilled in the art. Toxins used in antibody-toxin conjugates include bacterial toxins such as diphtheria toxin, plant toxins such as ricin, small molecule toxins such as geldanamycin (Mandler et al (2000) Jour. of the Nat. Cancer Inst. 92(19):1573-1581; Mandler et al (2000) Bioorganic & Med. Chem. Letters 10:1025-1028; Mandler et al (2002) Bioconjugate Chem. 13:786-791), maytansinoids (US 20050169933 A1; EP 1391213; Liu et al., (1996) Proc. Natl. Acad. Sci. USA 93:8618-8623), and calicheamicin (Lode et al (1998) Cancer Res. 58:2928; Hinman et al (1993) Cancer Res. 53:3336-3342). Other toxins include plant and bacterial toxins, such as, abrin, alpha toxin, exotoxin, gelonin, pokeweed antiviral protein, and saporin. Toxins can effect their cytotoxic

and cytostatic effects by mechanisms including tubulin binding, DNA binding, or topoisomerase inhibition.

In some embodiments, a toxin is linked to an antibody or an antigen binding fragment thereof, through a linker.

Linkers. Linkers of the disclosure can be any chemical linker. The linker can be a peptide linker. In some embodiments, the peptide linker ranges from about 2 to about 25 amino acids in length. In some embodiments, the peptide linker is 20 amino acids in length. In some embodiments, the peptide linker ranges from about 4 to about 16 amino acids in length. In some embodiments, the peptide linker is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids in length. In some embodiments, the peptide linker is longer than 25 amino acids in length. Linkers of the disclosure can be any cleavable or non-cleavable linker. In some embodiments, a peptide linker provides a protease-dependent cleavable site. Examples of protease-cleavable peptide linkers include, without limitation, the MMP sensitive linker and the factor Xa-sensitive linker IEGR. The art is familiar with a variety of cleavable sequences that may be employed for the methods provided herein, for example those disclosed in Chen et al., Adv. Drug Deliv. Rev. (2013), 65(10): 1357-69).

In some embodiments of the present invention, a flexible peptide linker can be used. A flexible peptide linker is preferably about 25 or fewer amino acids in length. In some embodiments, a flexible peptide linker is 20 amino acids in length. In some embodiments, a peptide linker contains about 20 or fewer amino acid residues, *e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20. In some embodiments, a peptide linker contains about 12 or fewer amino acid residues, e.g., 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12. In some cases, a peptide linker comprises two or more of the following amino acids: glycine, serine, alanine, and threonine. In some embodiments, the flexible peptide linker is a glycine-serine linker.

In some embodiments, the glycine-serine linker is represented by the formula $(GS)_n$, wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12. In some embodiments, the glycine-serine linker is represented by the formula $(GGGGS)_n$, wherein n is 1, 2, 3, 4, or 5.

Alternatively, a linking molecule may be a non-peptide linker. As used herein, a "non-peptide linker" refers to a biocompatible polymer including two or more repeating units linked to each other. Examples of the non-peptide polymer include but are not limited to:

polyethylene glycol (PEG) , polypropylene glycol (PPG), co-poly (ethylene/propylene) glycol, polyoxyethylene (POE), polyurethane, polyphosphazene, polysaccharides, dextran, polyvinyl alcohol, polyvinylpyrrolidones, polyvinyl ethyl ether, polyacryl amide, polyacrylate, polycyanoacrylates, lipid polymers, chitins, hyaluronic acid, and heparin. For more detailed descriptions of non-peptide linkers useful for Fc fusion molecules, see, for example, WO/2006/107124, which is incorporated by reference herein. Typically such linkers will have a range of molecular weight of from about 1 kDa to 50 kDa. For example, a typical PEG has a molecular weight of about 1 to 5 kDa, and polyethylene glycol has a molecular weight of about 5 kDa to 50 kDa, and more preferably about 10 kDa to 40 kDa.

The present disclosure shows that a CD33 protein isoform produced from an alternatively spliced transcript of CD33 lacking exon 2 lacks the IgV domain. Absence of the IgV domain of CD33, which can be indicated by the absence of exon 2 in CD33 mRNA or of the absence of amino acids encoded by exon 2 in the CD33 protein prevents binding of GO or other immunoconjugates targeting this region of CD33.

Aspects of the invention relate to determining whether a subject is likely to benefit from treatment with an agent that selectively binds to CD33. In some embodiments, an assay is performed on a biological sample from a subject with a cancer expressing CD33 to determine whether the subject is likely to benefit from treatments disclosed herein. One of ordinary skill in the art would appreciate that a variety of biological samples could be compatible with aspects of the invention. In some embodiments, assays are performed on tissue from a subject or a derivative of the tissue. In some embodiments, the subject has been diagnosed with a cancer expressing CD33. In some embodiments, an assay is performed to detect the presence of one or more CD33 SNPs. Any suitable assay for detection of a SNP may be compatible with aspects of the invention.

Single Nucleotide Polymorphism (SNP). Single nucleotide polymorphisms (SNPs) are single-nucleotide substitutions of one base for another that occur in more than one percent of the general population. SNPs occur throughout DNA in the human genome, at about one in every 300 nucleotides. To be classified as a SNP, two or more versions of a sequence must each be present in at least one percent of the general population.

In the present disclosure, six common SNPs: rs12459419-Ala14Val (rs12459419); rs2455069-Arg60Gly (rs2455069); rs35112940-Arg304Gly (rs35112940); rs61736475-Ser305Pro (rs61736475); promoterSNP-rs3865444 (rs3865444); and 3’UTRSNP-rs1803254

(rs1803254) in CD33 were genotyped and screened for association with clinical outcome, mRNA levels, and cell surface CD33 expression levels in 942 *de novo* AML patients enrolled in a COGAAML0531 trial. Nucleic acid sequences flanking each of the SNPs are shown in Table 1.

5 SNPs can be identified using various assays including, *e.g.*, without limitation, DNA sequencing, RNA sequencing, primer extension, enzyme-based methods, restriction fragment length polymorphism, PCR-based methods, PCR-RFLP, allele-specific PCR, flap endonuclease, 5'- nuclease, oligonucleotide ligation assay, other post-amplification methods based on physical properties of DNA, single strand conformation polymorphism, temperature 10 gradient gel electrophoresis, denaturing high performance liquid chromatography, high-resolution melting of the entire amplicon, use of DNA mismatch-binding proteins, SNPlex, surveyor nuclease assay, and hybridization-based methods including dynamic allele-specific hybridization, molecular beacons, and SNP microarrays. Other assays known to one of ordinary skill in the art may also be used to identify or genotype SNPs.

15 As described in the Examples, the presence of a variant single-nucleotide polymorphism (SNP) allele for the SNP rs12459419 (genotype TC or TT) in CD33 increases production of alternatively spliced transcript lacking exon 2 (p=1.41e-10).

In some embodiments, a subject is considered likely to benefit from treatment with an agent that selectively binds to CD33, such as GO if the subject exhibits a genotype of CC for 20 the CD33 SNP rs12459419. Accordingly, in some embodiments, an agent that selectively binds to CD33, such as GO is administered to a subject who exhibits a genotype of CC for the CD33 SNP rs12459419. In some embodiments, a subject is not considered likely to benefit from treatment with an agent that selectively binds to CD33, such as GO if the subject exhibits a genotype of TC or TT genotype for the CD33 SNP rs12459419. Accordingly, in 25 some embodiments, an agent that selectively binds to CD33, such as GO is not administered to a subject who exhibits a genotype of TC or TT for the CD33 SNP rs12459419.

In some embodiments, a subject is considered likely to benefit from treatment with an agent that selectively binds to CD33, such as GO if the subject exhibits a genotype of CC for the CD33 SNP rs3865444. Accordingly, in some embodiments, an agent that selectively 30 binds to CD33, such as GO is administered to a subject who exhibits a genotype of CC for the CD33 SNP rs3865444. In some embodiments, a subject is not considered likely to benefit from treatment with an agent that selectively binds to CD33, such as GO if the subject

exhibits a genotype of CA or AA for the CD33 SNP rs3865444. Accordingly, in some embodiments, an agent that selectively binds to CD33, such as GO is not administered to a subject who exhibits a genotype of CA or AA for the CD33 SNP rs3865444.

In some embodiments, an immunoassay is used to detect presence of a CD33 antigen. In some embodiments, an immunoassay can be used to detect presence of amino acids encoded by exon 2 of CD33. In some embodiments, binding of GO to the CD33 protein indicates presence of the CC genotype for rs12459419. In some embodiments, lack of binding of GO to the CD33 protein indicates presence of the TC or TT genotype for rs12459419. In some embodiments, detection of amino acids encoded by exon 2 of CD33 can indicate whether the genotype for SNP rs12459419 is wild-type (genotype CC) or variant (genotype TC or TT).

The antibodies or antigen binding fragments of the disclosure are, for example, suited for use in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. Examples of immunoassays which can utilize the antibody or antigen binding fragments are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the Enzyme Linked Immunoassay (ELISA), radioimmunoassay (RIA), the sandwich (immunometric assay), flow cytometry, the western blot assay, immunoprecipitation assays, immunohistochemistry, immuno-microscopy, lateral flow immuno-chromatographic assays, and proteomics arrays. The antigens and antibodies or antigen binding fragments can be bound to many different solid supports (*e.g.*, carriers, membrane, columns, proteomics array, etc.).

In some embodiments, an assay used to detect the presence or absence of amino acids encoded by exon 2 of CD33 can be a protein sequencing assay, *e.g.*, mass spectrometry or Edman degradation.

Table 1. CD33 SNPs identified by sequencing and genotyped in the present study

CD33 SNP	rs#	Location	Flanking sequence	SEQ ID NO:
Ala14Val	rs12459419	Exon 2	CCCACAGGGG-C-CCTGGCTATG	1
promoterSNP	rs3865444	Promoter	CTAAACACCC-C-ATGGATCTAG	2
3'UTR	rs1803254	3'UTR	CATTTGCTAA-G-TGTATGATGT	3
Arg304Gly	rs35112940	Exon 6	CCCTACCACA-G-GGTCAGCCTC	4
Arg69Gly	rs2455069	Exon 2	CCATTATATCC-A-GGGACTCTCC	5
Ser305Pro	rs61736475	Exon 6	TACCACAGGG-T-CAGCCTCCCC	6

Hybridization. The nucleic acid sequence of the flanking sequences of the SNPs of the disclosure are provided in SEQ ID NOs: 1 to 6, in Table 1. The amino acid sequence of the variant SNPs encoded by variant SNPs of SEQ ID NO:1 (variant SNP genotypes TC and TT of rs12459419) excludes amino acids encoded by exon 2 of CD33. The details of each SNP and its variant genotypes are useful for designing nucleic acid probes or primers which may be used to detect the mutant forms of the genes in individuals. For example, a nucleic acid probe may be designed which binds specifically to a sequence of CD33 which includes one or more of the variant SNPs (i.e. it either binds specifically to the variant SNP and not to the wild-type SNP, or it binds preferentially to the variant SNP as opposed to the wild-type). Typically, the probe will have a nucleic acid sequence that is complementary to a sequence of CD33 which includes one or more of the SNPs identified in Table 1. Suitably, the probe comprises a nucleic acid sequence which hybridizes under conditions of suitable stringency to at least 7, at least 14, at least 25, 50, 75, 100, 150, 200, 250, 300, 350, or 400 consecutive nucleotides of the sequence of the CD33 gene (SEQ ID NO: 7) which includes at least one of the SNPs rs12459419, rs3865444, rs1803254, rs35112940, and rs2455069.

Thus, in a further embodiment, there is provided an oligonucleotide which is complementary to a sequence of the CD33 gene which includes at least one of the SNPs indicated in Table 1. Typically, the oligonucleotide is a probe or a primer. Ideally, the primer is a primer for PCR nucleic acid amplification, ideally RT-PCR amplification. In one embodiment, the oligonucleotide consists of at least 7, at least 14, at least 25, 50, 75, 100, 150, 200, 250, 300, 350, or 400 consecutive nucleotides. A sample of myeloid lineage or blast cells from a subject with a cancer expressing CD33 may be isolated, the DNA extracted, and then assayed using a probe of the disclosure for the presence of one of the variant SNPs indicated in Table 1. Likewise, an oligonucleotide primer of the disclosure may be used to perform RT-PCR on the DNA sample. As the primer is designed to bind with the target DNA only when a desired variant SNP or desired wild-type SNP is present, amplification will only take place when the variant or wild-type SNP, respectively is present. The design of primers and probes of the disclosure, and their use in determining the presence of any wild-type or variant SNPs of the disclosure will be well known to a person skilled in the art.

CD33SNP_Score. In some embodiments, a genotype score is determined for one or more of the SNPs of the disclosure. The genotype score can be used to determine whether a subject is likely to respond to treatment with an agent that selectively binds to CD33. The genotype scores can be -2, -1, 0, 1, or 2. For each of the SNPs rs12459419/ rs3865444,

rs1803254 and rs35112940, patient genotypes are given a genotype score of 0 (wt/wt), -1 (wt/var) or -2 (var/var) based on the number of low expression allele. Since rs2455069 was associated with high expression genotype, genotype scores were 0 (wt/wt), 1 (wt/var) and 2 (var/var). A CD33SNP _Score can be calculated by adding the directional genotype scores of
5 the SNPs: rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069. Therefore the CD33SNP_Score = Σ (genotype scores of rs12459419 (or rs3865444), rs1803254, rs2455069 and rs35112940).

In some embodiments, a CD33SNP_Score greater than or equal to zero indicates that the subject is likely to benefit from treatment. In some embodiments, a genotype score equal
10 to zero for SNPs rs12459419 and/or rs3865444 indicates that the subject is likely to benefit from treatment. In some embodiments, a CD33SNP_Score greater than or equal to zero indicates that the subject is likely to benefit from treatment. In some embodiments, a genotype score of less than zero for SNPs rs12459419 and/or rs3865444 indicates that the subject is not likely to benefit from treatment. In some embodiments, the subject is likely to
15 benefit from treatment with an agent that selectively binds amino acids encoded by exon 2 of CD33 if the CD33SNP_Score for the subject is greater than or equal to zero. In some embodiments, the subject is not likely to benefit from treatment with an agent that selectively binds amino acids encoded by exon 2 of CD33 if the CD33SNP_Score for the subject is less than zero. In some embodiments, the subject is likely to benefit from treatment with an agent
20 that selectively binds amino acids encoded by exon 2 of CD33 if the genotype score for SNPs rs12459419 and/or rs3865444 is equal to zero for the subject. In some embodiments, the subject is not likely to benefit from treatment with an agent that selectively binds amino acids encoded by exon 2 of CD33 if the genotype score for SNPs rs12459419 and/or rs3865444 is less than zero for the subject.

25 *Subject.* "Subject" means a mammal, such as a human, a nonhuman primate, a dog, a cat, a sheep, a horse, a cow, a pig or a goat. In an important embodiment, the mammal is a human. The subject as used herein can be an adult subject or a pediatric subject.

Biological sample. A "biological sample" from a subject can include any cellular, tissue, bone marrow, or blood sample from the subject. Any type of biological sample
30 appropriate for conducting assays described herein can be compatible with aspects of the invention, as would be understood by one of ordinary skill in the art.

Treatment. "Treat", "treating" and "treatment" encompass an action that occurs while a subject is suffering from a condition which reduces the severity of the condition (or a symptom associated with the condition) or retards or slows the progression of the condition (or a symptom associated with the condition). This is therapeutic treatment.

5 Subjects are treated with effective amounts of the agents of the disclosure. An "effective amount" of an agent generally refers to an amount sufficient to elicit the desired biological response, *i.e.*, treat the condition. As will be appreciated by those of ordinary skill in this art, the effective amount of an agent described herein may vary depending on such factors as the condition being treated, the mode of administration, the therapy, if any, with 10 which it is combined, and the age and health of the subject.

For therapeutic treatment, a therapeutically effective amount is an amount sufficient to provide a therapeutic benefit in the treatment of a condition or to reduce or eliminate one or more symptoms associated with the condition. This may encompass an amount that improves overall therapy, reduces or avoids symptoms or causes of the condition, or 15 enhances the therapeutic efficacy of another therapeutic agent.

In some embodiments, the subject is treated with chemotherapy and/or radiation in addition to treatment with an agent that selectively binds to CD33, such as GO. In some embodiments, the subject is treated with chemotherapy within 30 days of administration of the agent that selectively binds CD33. In some embodiments, the subject is treated with 20 radiation therapy within 30 days of administration of the agent that selectively binds CD33.

In general, effective amounts are administered to reduce the number of or kill cancer cells. In connection with a specific disease or condition, a therapeutically effective amount can halt the development of, inhibit the progression of, reverse the development of, or otherwise reduce or ameliorate one or more symptoms of the disease or condition, for 25 example, one or more symptoms of cancer.

Pharmaceutical compositions. Agents that selectively bind to CD33, such as humanized antibodies, biologics and other molecules can be administered for the treatment of various disorders in the form of pharmaceutical compositions. Such compositions include the therapeutic(s) and one or more other pharmaceutically acceptable components. See 30 Remington's Pharmaceutical Science (15th ed., Mack Publishing Company, Easton, Pa. (1980)). The preferred form depends on the intended mode of administration and therapeutic

application. The compositions can also include, depending on the formulation desired, pharmaceutically-acceptable, non-toxic carriers or diluents, which are defined as vehicles commonly used to formulate pharmaceutical compositions for animal or human administration. The diluent is selected so as not to adversely affect the biological activity of
5 the antibody. Examples of such diluents are distilled water, physiological phosphate-buffered saline, Ringer's solutions, dextrose solution, and Hank's solution. In addition, the pharmaceutical composition or formulation may also include other carriers or nontoxic, nontherapeutic, nonimmunogenic stabilizers and the like.

Pharmaceutical compositions can also include large, slowly metabolized
10 macromolecules such as proteins, polysaccharides such as chitosan, polylactic acids, polyglycolic acids and copolymers (such as latex functionalized SEPHAROSE™ (GE Healthcare Bio-Sciences Ltd.), agarose, cellulose, and the like), polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes).

Pharmaceutical compositions may be injectable compositions. Injectable
15 compositions include solutions, suspensions, dispersions, and the like. Injectable solutions, suspensions, dispersions, and the like may be formulated according to techniques well-known in the art (see, for example, Remington's Pharmaceutical Sciences, Chapter 43, 14th Ed., Mack Publishing Co., Easton, Pa.), using suitable dispersing or wetting and suspending agents, such as sterile oils, including synthetic mono- or diglycerides, and fatty acids,
20 including oleic acid.

Injectable compositions that include an agent useful in the invention may be prepared in water, saline, isotonic saline, phosphate-buffered saline, citrate-buffered saline, and the like and may optionally mixed with a nontoxic surfactant. Dispersions may also be prepared in glycerol, liquid polyethylene, glycols, DNA, vegetable oils, triacetin, and the like and
25 mixtures thereof. Under ordinary conditions of storage and use, these preparations may contain a preservative to prevent the growth of microorganisms. Pharmaceutical dosage forms suitable for injection or infusion include sterile, aqueous solutions or dispersions or sterile powders comprising an active ingredient which powders are adapted for the extemporaneous preparation of sterile injectable or infusible solutions or dispersions.
30 Preferably, the ultimate dosage form is a sterile fluid and stable under the conditions of manufacture and storage. A liquid carrier or vehicle of the solution, suspension or dispersion may be a solvent or liquid dispersion medium comprising, for example, water, ethanol, a

polyol such as glycerol, propylene glycol, or liquid polyethylene glycols and the like, vegetable oils, nontoxic glyceryl esters, and suitable mixtures thereof. Proper fluidity of solutions, suspensions or dispersions may be maintained, for example, by the formation of liposomes, by the maintenance of the desired particle size, in the case of dispersion, or by the
5 use of nontoxic surfactants. The prevention of the action of microorganisms can be accomplished by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. Isotonic agents such as sugars, buffers, or sodium chloride may be included. Prolonged absorption of the injectable compositions can be brought about by the inclusion in the composition of agents delaying
10 absorption--for example, aluminum monostearate hydrogels and gelatin. Solubility enhancers may be added.

Sterile injectable compositions may be prepared by incorporating the therapeutic in the desired amount in the appropriate solvent with various other ingredients, e.g. as enumerated above, and followed by sterilization, as desired, by, for example filter
15 sterilization. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation include vacuum drying and freeze-drying techniques, which yield a powder of the active ingredient plus any additional desired ingredient present in a previously sterile-filtered solution. Any suitable sterilization process may be employed, such as filter sterilization, e.g. 0.22 micron filter or nanofiltration, gamma or electron beam sterilization, or
20 pulsed white light. Other suitable sterilization processes include UtiSter (Pegasus Biologics, Irvine Calif.) and those described in, e.g., U.S. Pat. No. 6,946,098 and U.S. Pat. No. 5,730,933.

In various embodiments, the final solution typically is adjusted to have a pH between about 4 and about 9, between about 5 and about 7, between about 5.5 and about 6.5, or about
25 6. The pH of the composition may be adjusted with a pharmacologically acceptable acid, base or buffer. Hydrochloric acid is an example of a suitable acid, and sodium hydroxide is an example of a suitable base. The hydrochloric acid or sodium hydroxide may be in any suitable form, such as a 1N solution

A resultant injectable solution preferably contains an amount of one or more
30 therapeutics effective to treat a disease. In various embodiments, a therapeutic such as an antibody is present in an injectable composition at a concentration between about 0.0001

mg/ml and about 50 mg/ml. In various embodiments, an antibody is present in an injectable composition at a concentration between about 0.01 mg/mL and about 10 mg/mL.

Agents, such as antibodies, also may be administered via other modes of administration known in the art. Such modes of administration include inhalation, ingestion 5 and topical application. Oral administration is also possible for therapeutics, although this form of administration is more challenging for certain biologics such as antibodies.

Kits

In certain aspects, the disclosure provides kits. Kits can include an agent that selectively binds to CD33. In some embodiments, the agent is conjugated to a toxin. In some 10 embodiments, the agent is linked to a toxin through a linker. In some embodiments, the agent is in sterile container(s). In some embodiments, the kit comprises instructions for administration of the kit components. In some embodiments, the instructions indicate the use of the agent to treat a subject if the genotype of the subject for the CD33 SNP rs12459419 is CC and/or if the genotype of the subject for the CD33 SNP rs3865444 is CC. In some 15 embodiments, the kit includes a pharmaceutical preparation vial, a pharmaceutical preparation diluent vial, and the agent. The diluent vial contains a diluent such as physiological saline for diluting what could be a concentrated solution or lyophilized powder of a composition of the disclosure. In some embodiments, the instructions include instructions for mixing a particular amount of the diluent with a particular amount of a 20 concentrated pharmaceutical composition, whereby a final formulation for injection or infusion is prepared. In some embodiments, the instructions include instructions for use in a syringe or other administration device. In some embodiments, the instructions include instructions for treating a patient with an effective amount of a composition of the disclosure. It also will be understood that the containers containing the preparations, whether the 25 container is a bottle, a vial with a septum, an ampoule with a septum, an infusion bag, and the like, may contain indicia such as conventional markings which change color when the preparation has been autoclaved or otherwise sterilized.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than 30 routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless required by context, singular terms shall include pluralities and plural terms shall include the singular. The methods and techniques of the present disclosure are generally performed according to conventional methods well-known in the art. Generally, nomenclatures used in connection with, and techniques of biochemistry, enzymology, molecular and cellular biology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art. The methods and techniques of the present disclosure are generally performed according to conventional methods known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated.

Exemplary embodiments of the disclosure will be described in more detail by the following examples. These embodiments are exemplary of the disclosure, which one skilled in art will recognize is in no way limited to the exemplary embodiments. The entire contents of all of the references (including literature references, issued patents, published patent applications, and co-pending patent applications) cited throughout this application are hereby expressly incorporated by reference, in particular for the teaching that is referenced hereinabove. However, the citation of any reference is not intended to be an admission that the reference is prior art.

EXAMPLES

Example 1: CD33 genetic variation predicts clinical response to gemtuzumab ozogamicin vs. standard therapy in de novo AML

Gemtuzumab Ozogamicin (GO) is a CD33-targeting agent that holds promise in both adult and pediatric AML populations. Two Meta-analysis reports including 11 and 5 randomized control trials suggest GO provided significant survival benefit in patients with favorable cytogenetics (9,10). Within pediatric AML, two trials from Children's Oncology Group COG- AAML03P1, (all patients received GO based chemotherapy) and more recently AAML0531 (patients were randomized to receive either standard five-course chemotherapy with or without addition of 2 doses of GO) demonstrated that addition of GO improved event free survival (EFS) and overall survival (OS) in AAML03P1 (11) and EFS and reduction in

relapse risk (RR) within AAML0531(12). These results indicate that GO has a potential to improve outcome in AML and was prematurely with drawn from market (13,14).

In AML, wide inter-patient variation in leukemic CD33 cell surface intensity as well as association with negative prognostic factors such as FLT3-ITD has been reported (15). Recent investigation from AAML0531 shows benefit from GO addition in patients with high leukemic cell surface CD33 levels and which was not seen in patient with low levels CD33 (16). We have previously shown association of CD33 SNPs with leukemic cell surface CD33 intensity and clinical outcome in patients within the AAML03P1 study (17). As the AAML03P1 clinical trial did not have a randomized control arm, it was not possible to investigate whether CD33 SNPs demonstrated differences in clinical outcome in patients in GO vs. No-GO arm. Here, we report results of clinical significance of 6 common CD33 SNPs by GO vs. No-GO treatment arms in a large cohort (n=1022) of de novo AML patients treated under COG AAML0531 trial. These results demonstrate significant clinical impact of a splicing and coding SNP rs12459419 on clinical outcome and provide rationale for selection of patients based on the genotype for GO based chemotherapy. Presence of the SNP results in exon 2 skipping and thus deletion of the epitope targeted by GO, thereby compromising response to GO.

Methods

Patients and treatment

Pediatric patients enrolled in the COG trial AAML0531 were included for this study. Details of the study design, treatment regimen and clinical outcome have been described in detail previously (Gamis et al., *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 2014;32(27):3021-32). Overall, 1022 patients between ages 0-29 years with *de novo* AML were randomly assigned to either standard five course chemotherapy (No-GO, n=511) or to the same chemotherapy with addition of two doses of 3 mg/m2 of GO (GO, n=511) during induction course I and intensification II. Low risk (LR) features were presence of t(8;21), inv(16) or t(16;16); high risk (HR) features included presence of monosomy 7, monosomy 5/5q deletion, or persistent disease at end of induction 1. All high-risk group patients received allogeneic stem cell transplant. Patients not assigned to low or high-risk groups were classified as intermediate risk (IR) patients and these patients received SCT if available. For this study, in addition to features listed above, presence of FLT-3/ITD mutations was considered high-risk and presence of NPM1 or CEBP mutation

with absence of FLT3/ITD was considered low-risk. All AAML0531 patient specimens with consent for biology studies were used in this study. The institutional review boards of all participating institutions approved the clinical protocol and the COG Myeloid Disease Biology Committee approved this research.

5 ***Genotyping of CD33 SNPs***

Six SNPs in CD33, including: a promoter SNP - rs3865444; 4 coding SNPs: rs12459419-Ala14Val, rs2455069-Arg60Gly, rs35112940-Arg304Gly and rs61736475-Ser305Pro; and a 3'UTR SNP - rs1803254 were genotyped using Sequenome platform at Biomedical Genomics Center at University of Minnesota (Minneapolis, MN). All the 6 SNPs 10 had a call rate of >0.98.

CD33 expression levels on leukemic cells

CD33 expression levels as determined by mean fluorescent intensity (MFI) of myeloid progenitor cells, as defined by CD45 low and side scatter, was determined by flow cytometry using the P67.7 antibody that specifically recognizes the IgV domain of CD33 as 15 described previously (Pollard et al., *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 2016;34(7):747-55; Mortland et al., *Clinical cancer research: an official journal of the American Association for Cancer Research*, 2013;19(6):1620-7; Pollard et al., *Blood*, 2012;119(16):3705-11; and Wells, *Leukemia research*, 2008;32(6).).

20 ***mRNA levels of CD33 and its isoforms***

CD33-WT and alternatively spliced transcript levels, as well as exon levels were obtained from RNA-seq data from 80 patients. Additionally CD33-WT and D2 splice variant lacking exon 2 was quantitated in 30 samples (10 with each genotype) with real time PCR using isoform specific primers.

25 *RNA seq:* RNA-seq data was available from 80 patients. CD33 mRNA transcript levels as well as levels of each exon was obtained. Additionally, the expression of a CD33-D2 isoform with deletion of exon 2 expressed as the "percent-spliced-in" (PSI, or ψ) value was obtained. PSI represents percentage of all mRNAs that correspond to the isoform with exclusion of exon 2.

Real-Time Quantitative PCR Analysis: mRNA expression levels of CD33 WT and D2-splice variant isoforms were quantitated in 30 patients selected based on genotype for rs12459419 (10 in each genotype) from the COG-AML0531 study. Briefly, total RNA was extracted and cDNA was synthesized using High Capacity cDNA reverse transcription kit (Life Technologies, USA). Quantitative real-time PCR was carried out using ABI7900-HT Real-time PCR system and QuantiTect SYBR Green PCR Kit. To quantify expression of total CD33: Forward, 5-TGTTCCACAGAACC CAACAA-3 and reverse, 5-GGCTGTAACACCAGCTC CTC-3 primers corresponding to sequences within exons 4 and 5, respectively were used. To quantify expression of D2-CD33: Forward, 5 -CCCTGCTGTGGGCAGACTTG-3; and reverse, 5 - GCACCGAGGAGTGAGTAGTCC-3 primers corresponding to sequences at the exon 1–3 junction and exon 3, respectively were used. The relative levels of CD33 isoforms were determined by $2^{-\Delta\Delta CT}$ method after normalization to GAPDH levels.

CD33 SNPs are not associated with patient characteristics

6 SNPs in the CD33 gene in 942 patients treated in the COG AAML0531 clinical trial were genotyped. Table 8 provides CD33 SNP status with respect to patient and disease characteristics. CD33 SNPs differed in frequency by race; however, rs12459419 (LD SNP rs3865444), rs1803254, rs2455069 and rs35112940 did not show any significant difference in representation by treatment arm, cytogenetics, risk group and FLT3-ITD status, or whether patients received SCT in the study or not. rs61736475, a less commonly occurring SNP, differed significantly among cytogenetic features, risk group, FLI3-ITD status and whether or not patients received SCT.

Table 8: Patient and disease characteristics by CD33 SNPs.

	Population	rs12459419						
		Genotype Frequency						
		CC		TT		TC		Genotypes
Characteristics	**N**	**N**	**%**	**N**	**%**	**N**	**%**	**p**
Gender								
Male	414	209	50%	41	48%	163	52%	0.852
Female	403	206	50%	44	52%	153	48%	
Race								
White	599	283	74%	69	99%	247	86%	<0.001
Black	93	68	18%	1	1%	24	8%	<0.001
Other	44	29	8%	0	0%	15	5%	0.037
Unknown	81	35		15		30		
Treatment Arm								
Arm A	408	204	49%	46	54%	158	50%	0.707
Arm B	409	211	51%	39	46%	158	50%	
Cytogenetics								
Normal	180	104	26%	18	22%	58	19%	0.083
t(8:21)	117	61	15%	10	12%	46	15%	0.761
inv(16)	96	41	10%	13	16%	41	13%	0.247
Abn 11	162	80	20%	17	20%	65	21%	0.925
t(6:9)	11	7	2%	2	2%	2	1%	0.328
monosomy 7	16	6	1%	1	1%	9	3%	0.348
del (7q)	11	6	1%	2	2%	3	1%	0.591
-5/5q	11	8	2%	0	0%	3	1%	0.270
+8	51	23	6%	2	2%	26	8%	0.099
Other	139	66	16%	18	22%	55	18%	0.507
Unknown	23							
FLT3/ITD								
No	671	341	84%	68	82%	261	83%	0.917
Yes	133	66	16%	15	18%	52	17%	
Unknown	8							
CEBPA status								
CEBPA mutant	49	30	7%	4	5%	15	5%	0.327
CEBPA WT	758	380	93%	79	95%	298	95%	
Unknown	7							
NPM status								
NPM mutant	64	36	12%	5	6%	23	7%	0.622
NPM WT	742	374	88%	78	94%	289	93%	
Risk group								
Intermediate	385	192	47%	44	53%	149	48%	0.626
Low	307	159	39%	28	34%	119	38%	0.649
High	111	55	14%	11	13%	45	14%	0.938
MRD at End of Induction 1								
No	472	229	55%	53	62%	189	60%	0.620
Yes	212	111	27%	24	28%	77	24%	
Received SCT on study								
Yes	128	57	14%	15	18%	56	18%	0.297
No	689	358	86%	70	82%	260	82%	

	Population	rs1803254						
		Genotype Frequency						
		CC		CG		GG		Genotypes
Characteristics	**N**	**N**	**%**	**N**	**%**	**N**	**%**	**p**
Gender								
Male	414	21	53%	111	50%	281	51%	0.916
Female	403	19	48%	113	50%	271	49%	
Race								
White	599	28	82%	133	69%	438	86%	<0.001
Black	93	6	18%	47	24%	40	8%	<0.001
Other	44	0	0%	13	7%	31	6%	0.306
Unknow n	81	6		31		43		
Treatment Arm								
Arm A	408	24	60%	119	53%	265	48%	0.187
Arm B	409	16	40%	105	47%	287	52%	
Cytogenetics								
Normal	180	6	15%	49	23%	125	23%	0.483
t(8:21)	117	6	15%	38	18%	73	14%	0.375
inv(16)	96	7	18%	24	11%	64	12%	0.519
Abn 11	162	7	18%	38	18%	117	22%	0.389
t(6:9)	11	0	0%	3	1%	8	1%	0.739
monosomy 7	16	0	0%	6	3%	10	2%	0.468
del (7q)	11	1	3%	2	1%	8	1%	0.691
-5/5q	11	0	0%	2	1%	9	2%	0.542
+8	51	1	3%	13	6%	37	7%	0.528
Other	139	12	30%	41	19%	86	16%	0.065
Unknow n	23							
FLT3/ITD								
No	671	37	93%	180	81%	453	84%	0.221
Yes	133	3	8%	41	19%	89	16%	
Unknow n	8							
CEBPA status								
CEBPA mutant	49	3	8%	10	5%	36	7%	0.501
CEBPA WT	758	37	93%	212	95%	508	93%	
Unknow n	7							
NPM status								
NPM mutant	64	2	5%	15	7%	47	9%	0.539
NPM WT	742	38	95%	206	93%	497	91%	
Risk group								
Intermediate	385	21	53%	104	47%	260	48%	0.817
Low	307	16	40%	81	37%	209	39%	0.852
High	111	3	8%	36	16%	72	13%	0.274
MRD at End of Induction 1								
No	472	25	63%	126	56%	320	58%	0.572
Yes	212	11	28%	65	29%	136	25%	
Received SCT on study								
Yes	128	8	20%	37	17%	83	15%	0.652
No	689	32	80%	187	83%	469	85%	

	Population	rs2455069						
		Genotype Frequency						
		AA		AG		GG		Genotypes
Characteristics	N	N	%	N	%	N	%	p
Gender								
Male	414	148	47%	194	52%	65	54%	0.309
Female	403	164	53%	176	48%	55	46%	
Race								
White	599	201	75%	294	85%	95	87%	0.001
Black	93	31	12%	45	13%	14	13%	0.853
Other	44	36	13%	6	2%	0	0%	<0.001
Unknown	81	44		25		11		
Treatment Arm								
Arm A	408	165	53%	177	48%	60	50%	0.422
Arm B	409	147	47%	193	52%	60	50%	
Cytogenetics								
Normal	180	76	25%	69	19%	34	30%	0.036
t(8:21)	117	41	14%	57	16%	16	14%	0.695
inv(16)	96	43	14%	37	10%	15	13%	0.289
Abn 11	162	49	16%	88	24%	22	19%	0.031
t(6:9)	11	5	2%	3	1%	2	2%	0.579
monosomy 7	16	8	3%	8	2%	0	0%	0.226
del (7q)	11	2	1%	7	2%	1	1%	0.315
-5/5q	11	2	1%	6	2%	2	2%	0.466
+8	51	17	6%	24	7%	8	7%	0.818
Other	139	60	20%	62	17%	15	13%	0.259
Unknown	23							
FLT3/ITD								
No	671	253	82%	316	87%	89	77%	0.034
Yes	133	55	18%	49	13%	27	23%	
Unknown	8							
CEBPA status								
CEBPA mutant	49	14	5%	25	7%	9	8%	0.344
CEBPA WT	758	294	95%	341	93%	109	92%	
Unknown	7							
NPM status								
NPM mutant	64	28	9%	21	6%	14	12%	0.067
NPM WT	742	280	91%	344	94%	104	88%	
Risk group								
Intermediate	385	146	48%	184	51%	47	40%	0.128
Low	307	115	38%	135	37%	51	43%	0.472
High	111	45	15%	45	12%	20	17%	0.408
MRD at End of Induction 1								
No	472	170	54%	229	62%	86	72%	0.178
Yes	212	89	29%	86	23%	31	26%	
Received SCT on study								
Yes	128	51	16%	55	15%	19	16%	0.866
No	689	261	84%	315	85%	101	84%	

	Population	rs35112940						
		Genotype Frequency						
		AA		AG		GG		Genotypes
Characteristics	N	N	%	N	%	N	%	p
Gender								
Male	414	9	41%	82	47%	322	52%	0.348
Female	403	13	59%	92	53%	298	48%	
Race								
White	599	20	95%	140	89%	439	79%	0.003
Black	93	1	5%	12	8%	80	14%	0.045
Other	44	0	0%	5	3%	39	7%	0.104
Unknown	81	1		17		62		
Treatment Arm								
Arm A	408	12	55%	91	52%	305	49%	0.701
Arm B	409	10	45%	83	48%	315	51%	
Cytogenetics								
Normal	180	4	19%	32	19%	144	24%	0.345
t(8:21)	117	2	10%	22	13%	93	15%	0.568
inv(16)	96	5	24%	21	12%	69	11%	0.227
Abn 11	162	4	19%	42	25%	116	19%	0.296
t(6:9)	11	0	0%	3	2%	8	1%	0.784
monosomy 7	16	1	5%	4	2%	11	2%	0.604
del (7q)	11	0	0%	3	2%	8	1%	0.784
-5/5q	11	0	0%	2	1%	9	1%	0.818
+8	51	0	0%	11	6%	40	7%	0.475
Other	139	5	24%	30	18%	104	17%	0.740
Unknown	23							
FLT3/ITD								
No	671	16	80%	144	83%	510	84%	0.910
Yes	133	4	20%	29	17%	100	16%	
Unknown	8							
CEBPA status								
CEBPA mutant	49	0	0%	10	6%	39	6%	0.498
CEBPA WT	758	20	100%	162	94%	575	94%	
Unknown	7							
NPM status								
NPM mutant	64	0	0%	13	8%	51	8%	0.391
NPM WT	742	20	100%	159	92%	562	92%	
Risk group								
Intermediate	385	12	57%	86	50%	287	47%	0.526
Low	307	7	33%	61	36%	238	39%	0.655
High	111	2	10%	24	14%	85	14%	0.845
MRD at End of Induction 1								
No	472	15	68%	109	63%	347	56%	0.211
Yes	212	4	18%	39	22%	169	27%	
Received SCT on study								
Yes	128	2	9%	25	14%	101	16%	0.570
No	689	20	91%	149	86%	519	84%	

	Population	rs61736475 Genotype Frequency						
		CC		TT		TC		Genotypes
Characteristics	N	N	%	N	%	N	%	p
Gender								
Male	414	2	33%	379	52%	32	42%	0.155
Female	403	4	67%	351	48%	45	58%	
Race								
White	599	0	0%	565	86%	31	44%	<0.001
Black	93	6	100%	48	7%	38	54%	<0.001
Other	44	0	0%	43	7%	1	1%	0.189
Unknown	81	0		74		7		
Treatment Arm								
Arm A	408	3	50%	368	50%	36	47%	0.830
Arm B	409	3	50%	362	50%	41	53%	
Cytogenetics								
Normal	180	1	17%	163	23%	15	20%	0.761
t(8:21)	117	2	33%	91	13%	23	30%	<0.001
inv(16)	96	1	17%	88	12%	7	9%	0.677
Abn 11	162	1	17%	147	21%	14	18%	0.867
t(6:9)	11	0	0%	10	1%	0	0%	0.556
monosomy 7	16	0	0%	15	2%	1	1%	0.840
del (7q)	11	0	0%	10	1%	1	1%	0.956
-5/5q	11	0	0%	10	1%	1	1%	0.956
+8	51	0	0%	48	7%	3	4%	0.515
Other	139	1	17%	126	18%	11	14%	0.768
Unknown	23							
FLT3/ITD								
No	671	6	100%	591	82%	71	93%	0.025
Yes	133	0	0%	127	18%	5	7%	
Unknown	8							
CEBPA status								
CEBPA mutant	49	0	0%	43	6%	5	7%	0.806
CEBPA WT	758	6	100%	678	94%	71	93%	
Unknown	7							
NPM status								
NPM mutant	64	0	0%	56	8%	7	9%	0.701
NPM WT	742	6	100%	664	92%	69	91%	
Risk group								
Intermediate	385	3	50%	350	49%	31	41%	0.410
Low	307	3	50%	261	36%	41	54%	0.010
High	111	0	0%	106	15%	4	5%	0.045
MRD at End of Induction 1								
No	472	4	67%	423	58%	44	57%	0.649
Yes	212	1	17%	194	27%	16	21%	
Received SCT on study								
Yes	128	0	0%	120	16%	7	9%	0.137
No	689	6	100%	610	84%	70	91%	

CD33 SNPs are strong predictors of clinical response by treatment arm

Survival outcome and risk of relapse (RR) analysis indicated that patients with rs12459419 CC genotype had significantly lower RR at 26±7% with addition of GO as
5 compared to 49±9% in No-GO arm (p<0.001) Figure 1. Similarly, DFS and EFS demonstrated significant improvement in GO vs. No-GO arm (DFS: 65±7% vs. 46±9%;,

p=0.004; EFS: 54±7% vs. 41±7%; p=0.04). This difference in outcome between GO and No-GO arms was not seen in patients with at least one variant -T allele for rs12459419 (TC and TT genotypes), Table 2.

Among other CD33 SNPs: rs1803254, rs35112940 and rs2455609, specific genotype groups demonstrated differences in survival outcome by GO vs. No-GO treatment arms. Table 2 and Figure 2 provide results for the association analysis between CD33 SNPs and differences in clinical response in patients on GO and No-GO arm. None of the SNPs were associated with overall survival (OS).

Table 2. CD33 SNPs and clinical outcome analysis by treatment arm No-GO vs. GO in patients treated under COG-AAML0531.

rs12459419	**CC (N=415)**			**TC (N=316)**			**TT (N=85)**		
	No-GO (N=204)	**GO (N=211)**	**No-GO vs. GO**	**No-GO (N=158)**	**GO (N=158)**	**No-GO vs. GO**	**No-GO (N=46)**	**GO (N=39)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**
5yr OS from study entry	58 ± 7%	66 ± 7%	0.159	69 ± 7%	64 ± 8%	0.658	70 ± 14%	69 ± 15%	0.824
5yr EFS from study entry	41 ± 7%	54 ± 7%	**0.041**	53 ± 8%	51 ± 8%	0.749	44 ± 15%	49 ± 17%	0.461
	No-GO (N=145)	**GO (N=154)**	**p**	**No-GO (N=101)**	**GO (N=125)**	**p**	**No-GO (N=34)**	**GO (N=30)**	**p**
5yr DFS from End of Course 1	46 ± 9%	65 ± 7%	**0.004**	60 ± 10%	56 ± 9%	0.821	54 ± 18%	51 ± 20%	0.972
5yr RR from End of Course 1	49 ± 9%	26 ± 7%	**<0.001**	37 ± 10%	38 ± 9%	0.975	46 ± 18%	46 ± 20%	0.798
5yr TRM from End of Course 1	4 ± 3%	9 ± 5%	0.135	3 ± 3%	5 ± 4%	0.496	--	3 ± 7%	0.290
rs3865444	**CC(N=412)**			**CA (N=313)**			**AA (N=86)**		
	No-GO (N=203)	**GO (N=209)**	**No-GO vs. GO**	**No-GO (N=156)**	**GO (N=157)**	**No-GO vs. GO**	**No-GO (N=47)**	**GO (N=39)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**
5yr OS from study entry	58 ± 7%	67 ± 7%	0.134	69 ± 7%	66 ± 8%	0.773	69 ± 14%	63 ± 20%	0.768
5yr EFS from study entry	41 ± 7%	54 ± 7%	**0.034**	53 ± 8%	52 ± 8%	0.618	43 ± 15%	47 ± 17%	0.558
	No-GO (N=144)	**GO (N=151)**	**p**	**No-GO (N=100)**	**GO (N=124)**	**p**	**No-GO (N=34)**	**GO (N=30)**	**p**
5yr DFS from End of Course 1	46 ± 9%	65 ± 8%	**0.003**	60 ± 10%	58 ± 9%	0.954	54 ± 18%	47 ± 20%	0.792
5yr RR from End of Course 1	50 ± 9%	26 ± 7%	**<0.001**	37 ± 10%	37 ± 9%	0.750	46 ± 18%	49 ± 20%	0.964
5yr TRM from End of Course 1	4 ± 3%	9 ± 5%	0.128	3 ± 3%	5 ± 4%	0.500	--	3 ± 7%	0.290
rs1803254	**GG (N=552)**			**CG (N=224)**			**CC (N=40)**		
	No-GO (N=268)	**GO (N=287)**	**No-GO vs. GO**	**No-GO (N=119)**	**GO (N=219)**	**No-GO vs. GO**	**No-GO (N=24)**	**GO (N=16)**	**No-GO vs. GO**
	% ± 2SE%	**% ± 2SE%**	**p**	**% ± 2SE%**	**% ± 2SE%**	**p**	**% ± 2SE%**	**% ± 2SE%**	**p**
5yr OS from study entry	65 ± 6%	67 ± 6%	0.640	58 ± 10%	62 ± 10%	0.560	71 ± 19%	66 ± 26%	0.810
5yr EFS from study entry	47 ± 6%	55 ± 6%	0.093	39 ± 9%	46 ± 10%	0.190	67 ± 19%	56 ± 25%	0.462
	No-GO (N=182)	**GO (N=287)**	**p**	**No-GO (N=81)**	**GO (N=79)**	**p**	**No-GO (N=17)**	**GO (N=11)**	**p**

5yr DFS from End of Course 1	53 ± 7%	62 ± 7%	0.097	44 ± 11%	53 ± 12%	0.165	76 ± 20%	73 ± 27%	0.754
5yr RR from End of Course 1	43 ± 8%	32 ± 6%	**0.022**	52 ± 12%	38 ± 11%	**0.040**	24 ± 14%	27 ± 12%	0.756
5yr TRM from End of Course 1	3 ± 3%	6 ± 3%	0.220	4 ± 4%	9 ± 6%	0.190	--	--	--
rs35112940	**GG (N=620)**			**AG (N=174)**			**AA (N=22)**		
	No-GO (N=305)	**GO (N=315)**	**No-GO vs. GO**	**No-GO (N=91)**	**GO (N=83)**	**No-GO vs. GO**	**No-GO (N=12)**	**GO (N=10)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	p	% ± 2SE%	% ± 2SE%	p	% ± 2SE%	% ± 2SE%	p
5yr OS from study entry	61 ± 6%	66 ± 6%	0.360	70 ± 10%	65 ± 13%	0.820	75 ± 25%	70 ± 30%	0.850
5yr EFS from study entry	45 ± 6%	53 ± 5%	0.091	50 ± 11%	52 ± 11%	0.671	--	45 ± 35%	0.172
	No-GO (N=211)	**GO (N=234)**	p	**No-GO (N=61)**	**GO (N=67)**	p	**No-GO (N=8)**	**GO (N=8)**	p
5yr DFS from End of Course 1	51 ± 7%	63 ± 6%	**0.019**	59 ± 13%	52 ± 12%	0.536	--	47 ± 37%	0.228
5yr RR from End of Course 1	45 ± 7%	28 ± 6%	**<0.001**	38 ± 13%	47 ± 12%	0.383	75 ± 36%	53 ± 42%	0.236
5yr TRM from End of Course 1	3 ± 2%	8 ± 4%	**0.033**	3 ± 5%	1 ± 3%	0.501	--	--	--
rs2455069	**AA (N=312)**			**AG (N=370)**			**GG (N=120)**		
	No-GO (N=165)	**GO (N=147)**	**No-GO vs. GO**	**No-GO (N=177)**	**GO (N=193)**	**No-GO vs. GO**	**No-GO (N=60)**	**GO (N=60)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	p	% ± 2SE%	% ± 2SE%	p	% ± 2SE%	% ± 2SE%	p
5yr OS from study entry	69 ± 4%	67 ± 8%	0.700	59 ± 8%	63 ± 8%	0.244	61 ± 13%	74 ± 12%	0.247
5yr EFS from study entry	50 ± 8%	49 ± 8%	0.876	43 ± 8%	54 ± 7%	**0.015**	42 ± 13%	58 ± 13%	0.216
	No-GO (N=116)	**GO (N=115)**	p	**No-GO (N=113)**	**GO (N=143)**	p	**No-GO (N=45)**	**GO (N=49)**	p
5yr DFS from End of Course 1	54 ± 10%	55 ± 10%	0.608	82 ± 10%	62 ± 8%	0.084	45 ± 15%	65 ± 14%	0.105
5yr RR from End of Course 1	44 ± 10%	39 ± 9%	0.294	43 ± 10%	32 ± 8%	0.090	53 ± 16%	20 ± 12%	**0.002**
5yr TRM from End of Course 1	2 ± 2%	5 ± 4%	0.153	5 ± 4%	6 ± 4%	0.907	2 ± 4%	14 ± 10%	**0.041**
rs61736475	**TT (N=730)**			**TC (N=77)**			**CC (N=6)**		
	No-GO (N=368)	**GO (N=362)**	**No-GO vs. GO**	**No-GO (N=36)**	**GO (N=41)**	**No-GO vs. GO**	No-GO (N=3)	**GO (N=3)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	p	% ± 2SE%	% ± 2SE%	p	% ± 2SE%	% ± 2SE%	p
5yr OS from study entry	65 ± 5%	66 ± 5%	0.662	52 ± 17%	61 ± 16%	0.297	67 ± 5%	67 ± 5%	0.886
5yr EFS from study entry	46 ± 5%	53 ± 5%	0.056	41 ± 17%	55 ± 16%	0.365	67 ± 54%	0 ± 0%	0.110
	No-GO (N=252)	**GO (N=275)**	p	**No-GO (N=26)**	**GO (N=30)**	p	**No-GO (N=2)**	**GO (N=2)**	p
5yr DFS from End of Course 1	52 ± 6%	60 ± 6%	0.054	57 ± 20%	69 ± 17%	0.428	--	0 ± 0%	0.433
5yr RR from End of Course 1	46 ± 6%	34 ± 6%	**0.004**	32 ± 19%	21 ± 16%	0.415	50 ± 100%	100 ± 100%	0.434
5yr TRM from End of Course 1	2 ± 2%	6 ± 3%	**0.035**	12 ± 13%	10 ± 11%	0.864	--	--	--

Among other CD33 SNPs, rs1803254 C allele was associated with lower RR in GO vs. No-GO treatment arms (GG: 32±6% vs. 43±8%, p=0.02; CG: 38±11% vs. 52±12%, p=0.04). This difference was not observed in patients homozygous for the low expressing CC genotype (p=0.76). For coding SNP rs35112940, patients homozygous with GG genotype

had better outcome by treatment arm (GO vs. No-GO: DFS 63±6% vs. 51.7%, p=0.019; RR 28±6% vs. 45±7%, p<0.001) as compared to patients with at least one low expressing variant allele (AA or AG; p>0.05). Patients homozygous for variant allele (GG) for rs2455609 had significant improvement with RR in GO vs.No-GO arm (20±12 vs. 53±16% p=0.002). Only one SNP –rs35112940 was associated with treatment related mortality (TRM), with GG genotype experiencing higher TRM in GO vs. No-GO arm (14±10% vs. 2±4%, p=0.041). Table 2 provides results for the association analysis between CD33 SNPs and differences in clinical response in patients on GO and No-GO arm. None of the SNPs were associated with OS.

Association of CD33 SNPs with clinical outcome within each arm

Previous results from AAML03P1 showed rs35112940 GG genotype to be associated with reduced RR and better OS in White patients. Results presented herein are consistent with previous results, in that rs35112940 was associated with lower incidence of RR in GG genotype as compared to patients at least one A allele (GG vs. AG vs. AA; 28±6% vs. 47±12% vs. 53±42% p=0.012, Table 4 and Figure 3) in GO arm, but not in No-GO arm (p=0.093). Additionally for rs3865444 improvement in RR in patients in GO but not in No-GO arm was observed (GO arm -GG vs. GT vs. TT: 26±7% vs.37±9% vs. 49±20%, p =0.041, No-GO arm - GG vs. GT vs. TT: 50±9% vs.37±10% vs. 46±18%, p =0.299).

Table 4. Association of CD33 SNPs with clinical outcome within each arm

CD33 SNP	NO-GO ARM							GO-ARM						
rs35112940	AA		AG		GG		Genotypes	AA		AG		GG		Genotypes
	N	% ± 2SE%	N	% ± 2SE%	N	% ± 2SE%	p	N	% ± 2SE%	N	% ± 2SE%	N	% ± 2SE%	p
5yr RR from End of Course 1	8	75 ± 36%	61	38 ± 13%	211	45 ± 7%	0.093	8	53 ± 42%	67	47 ± 12%	234	28 ± 6%	**0.012**
rs3865444	GG		GT		TT		Genotypes	GG		GT		TT		Genotypes
	N	% ± 2SE%	N	% ± 2SE%	N	% ± 2SE%	p	N	% ± 2SE%	N	% ± 2SE%	N	% ± 2SE%	p
5yr RR from End of Course 1	144	50 ± 9%	100	37 ± 10%	34	46 ± 18%	0.299	151	26 ± 7%	124	37 ± 9%	30	49 ± 20%	**0.041**
rs12459419	CC		TT		TC		Genotypes	CC		TT		TC		Genotypes
	N	% ± 2SE%	N	% ± 2SE%	N	% ± 2SE%	p	N	% ± 2SE%	N	% ± 2SE%	N	% ± 2SE%	p
5yr RR from End of Course 1	145	49 ± 9%	34	46 ± 18%	101	37 ± 10%	0.290	154	26 ± 7%	30	46 ± 20%	125	38 ± 9%	**0.059**

CD33 SNP rs12459419 is the strongest predictor of CD33 cell surface intensity in patients from COG-AAML0531

Both genotype and CD33 cell surface intensity (mean fluorescence intensity-MFI) were available in 816 patients (n=408 in No-GO arm and n=408 in GO arm). Presence of variant allele for rs12459419 (and linked promoter SNP rs3865444) was significantly

associated with lower CD33 MFI, p=1.9E-29 (Table 3 and Figure 4). Variant alleles for rs1803254 and rs35112940 were also significantly associated with lower CD33 intensity on leukemic cells (p=7.3E-10 and p=3.9E-14). Coding missense SNP rs2455069 SNP was associated with higher CD33 expression (p=4.5E-8).

Table 3. Association of CD33 SNPs with CD33 cell surface intensity on Leukemic cells

CD33 SNP	SNP type	Allele Change	MAF Whites/ Blacks	P value (Kruskal Wallis)	Direction of Association of variant allele	Genotype distribution	CD33 Expression Median (Range)
rs12459419 (LD with promoter SNP rs3865444)	Missense Ala14Val	C>T (C>A)	0.32/0.13	**p=1.93e-29**	Low	CC (n=415) CT (n=316) TT (n=85)	195 (2.65 - 1351) 140.84 (3 - 748.47) 44.5 (6 - 813)
rs1803254	3'UTR	G>C	0.15/0.32	**p =7.32e-10**	Low	GG (n=552) CG (n=224) CC (n=40)	50.76 (5 - 584) 121.67 (6 - 1066) 168.17 (2.66 - 1351)
rs2455069	Missense Arg60Gly	A>G	0.41/0.40	**p= 3.93e-14**	High	AA (n=312) AG (n=370) GG (n=120)	96.9 (3-1130.6) 172 (3 - 2.68-1351) 204.15 (6.12 - 1225.8)
rs35112940	Missens Arg304Gly	G>A	0.15/0.07	**p=4.54e-8**	Low	GG (n=620) AG (n=174) AA (n=22)	157.57 (2.68 - 1351) 116.05 (3 - 797.2) 52.69 (8-813)
rs61736475	Missens Ser305Pro	T>C	0.028/0.27	p=0.129	NA	TT (n=730) TC (n=77) CC (n=6)	146,97 (2.68 - 1351) 111 (5 - 1025,07) 381,24 (29.39 - 1119.5)

Analysis within each arm, risk group as well as by FLT3-ITD status was consistent with results from the total study cohort (Table 5, and Figure 5). Since frequency of the selected CD33 SNPs within risk groups or by FLT3-ITD status was not different, the impact of CD33 SNPs on its cell surface intensity appears to be independent of risk group characteristics. Previous studies (Pollard JA et al., *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 2016;34(7):747-55; Pollard et al., *Blood* 2012;119(16):3705-11) have utilized quartile system to classify patients based on CD33 cell surface expression; consistent results were obtained here with association of SNPs with quartiles (Table 6).

Table 5. Association of CD33 SNPs expression within different subgroups of AML patients

rs12459419	Total	C (N=415)			T (N=85)			TC (N=316)			Genotype Comparison
SNPs	**N**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**p**
CD33 Expression, NoGo	408	204	205.15	(4.27 - 1077.7)	46	33.4	(6 - 813)	158	126.47	(3 - 748.47)	**<0.001**
CD33 Expression, Go	408	211	175	(2.68 - 1351)	39	58.62	(6.84 - 134.47)	158	158.77	(4 - 713)	**<0.001**
CD33 Expression, Standard Risk	385	192	279.26	(2.68 - 1351)	44	56.49	(6 - 813)	149	183	(3 -713)	**<0.001**
CD33 Expression, Low Risk	306	159	124	(11.23 - 876)	28	25.34	(6.84 - 91)	119	87.31	(5 - 545.03)	**<0.001**
CD33 Expression, High Risk	111	55	225.74	(4.27 - 1225.87)	11	50.55	(12 - 139.47)	45	178.73	(9.83 - 748.47)	**<0.001**
CD33 Expression, FLT3+	133	66	254.26	(4.27 - 1225.87)	15	55.9	(12 - 165.1)	52	213.28	(25 - 706)	**<0.001**
CD33 Expression, FLT3-	670	341	173	(2.68 - 1351)	68	35.48	(6 - 165.14)	261	133.85	(3 - 748.47)	**<0.001**

rs3865444	Total	G (N=412)			GT (N=313)			T (N=86)			Genotype Comparison
SNPs	**N**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**p**
CD33 Expression, All	811	412	193.16	(2.68 - 1351)	313	141	(3 - 748.47)	86	46.25	(6 - 813)	**<0.001**
CD33 Expression, NoGo	406	203	203.29	(4.27 - 1077.72)	156	126.83	(3 - 748.47)	47	34.22	(6 - 813)	**<0.001**
CD33 Expression, Go	405	209	173	(2.68 - 1351)	157	156	(4 - 713)	39	58.62	(6.84 - 439)	**<0.001**
CD33 Expression, Standard Risk	383	191	273.92	(2.68 - 1351)	147	183	(3 - 713)	45	58.62	(6 - 813)	**<0.001**
CD33 Expression, Low Risk	303	157	123	(11.23 - 876)	118	89.86	(5 - 545.03)	28	25.34	(6.84 - 91)	**<0.001**
CD33 Expression, High Risk	111	55	225.74	(4.27 - 1225.87)	45	178.73	(9.83 - 748.47)	11	50.55	(12 - 139.47)	**<0.001**
CD33 Expression, FLT3+	133	66	254.26	(4.27 - 1225.87)	52	213.275	(25 - 706)	15	55.9	(12 - 165.1)	**<0.001**
CD33 Expression, FLT3-	666	338	171	(2.68 - 1351)	259	133.85	(3 - 748.47)	69	35.95	(6 - 439)	**<0.001**

rs1803254	Total	C (N=40)			CG (N=224)			G (N=552)			Genotype Comparison
SNPs	**N**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**p**
CD33 Expression, NoGo	408	24	50.76	(20 - 584)	119	117	(6 - 748.47)	265	163	(3 - 1077.72)	**<0.001**
CD33 Expression, Go	408	16	56.26	(5 - 318.8)	105	121.89	(10 - 1086)	287	171	(2.68 - 1351)	**<0.001**
CD33 Expression, Standard Risk	385	21	80	(7.6 - 584)	104	174.19	(6 - 1086)	260	226.5	(2.68 - 1351)	**<0.001**
CD33 Expression, Low Risk	306	16	42.98	(5 - 153.88)	81	87.31	(10 - 455.24)	209	113.26	(9.34 - 876)	**<0.001**
CD33 Expression, High Risk	111	3	50.55	(32.58 - 55.90)	36	128.24	(9.83 - 748.47)	72	215.31	(4.27 - 1225.87)	**0.005**
CD33 Expression, FLT3+	133	3	50.55	(32.58 - 55.9)	41	146.01	(10.51 - 478.64)	89	239.25	(4.27 - 1225.87)	**0.001**
CD33 Expression, FLT3-	670	37	50.96	(5 - 584)	180	119.28	(6 - 1086)	453	146.94	(2.68 - 1351)	**<0.001**

rs2455069	Total	A (N=312)			AG (N=370)			G (N=120)			Genotype Comparison
SNPs	**N**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**p**
CD33 Expression, NoGo	402	165	37	(3 - 911)	177	164	(4.27 - 1077.72)	60	220.26	(6.12 - 900)	**<0.001**
CD33 Expression, Go	400	147	117	(4 - 1130.6)	193	174.12	(2.68 - 1351)	60	186.32	12.25 - 1225.87	**<0.001**
CD33 Expression, Standard Risk	377	146	143.5	(3 - 1130.6)	184	225	(2.68 - 1351)	47	324.09	(6.12 - 1086)	**<0.001**
CD33 Expression, Low Risk	301	115	63.42	(5 - 545.03)	135	99.75	(11.23 - 876)	51	147.11	(19.22 - 698.08)	**<0.001**
CD33 Expression, High Risk	110	45	114.11	(12 - 628.93)	45	207	(4.27 - 860.09)	20	358.05	12.25 - 1225.87	**0.003**
CD33 Expression, FLT3+	131	55	144.17	(12 - 628.93)	49	207.24	(4.27 - 711.68)	27	362.73	12.25 - 1225.87	**<0.001**
CD33 Expression, FLT3-	658	253	89.44	(3 - 1130.6)	316	160.77	(2.68 - 1351)	89	171	(6.12 - 1086)	**<0.001**

rs35112940	Total	A (N=22)			AG (N=174)			G (N=620)			Genotype Comparison
SNPs	**N**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**p**
CD33 Expression, NoGo	408	12	22.5	(8 - 813)	91	85	(3 - 706)	305	158.38	(4.27 - 1077.72)	**<0.001**
CD33 Expression, Go	408	10	62.21	(16 - 129.45)	83	148	(10 - 797.2)	315	157.14	(2.68 - 1351)	**0.011**
CD33 Expression, Standard Risk	385	12	56.705	(8 - 813)	86	156.94	(3 - 797.2)	287	218.45	(2.68 - 1351)	**<0.001**
CD33 Expression, Low Risk	306	7	23	(9.34 - 82.38)	61	77	(10 - 355.74)	238	107.12	(5 - 876)	**<0.001**
CD33 Expression, High Risk	111	2	83.135	(52.97 - 113.3)	24	147.33	(12 - 459.02)	85	207.01	(4.27 - 1225.87)	0.199
CD33 Expression, FLT3+	133	4	71.24	(40 - 107.65)	29	178.73	(12 - 797.2)	100	235.77	(4.27 - 1225.87)	0.068
CD33 Expression, FLT3-	670	16	27.01	(8 - 129.45)	144	97.08	(3 - 713)	510	147.06	(2.68 - 1351)	**<0.001**

rs61736475	Total	C (N=6)			T (N=730)			TC (N=77)			Genotype Comparison
SNPs	**N**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**N**	**Median**	**Range**	**p**
CD33 Expression, NoGo	407	3	86	(29.39 - 318.47)	368	142	(3 - 1077.72)	36	138.69	13.49 - 1025.07	0.866
CD33 Expression, Go	406	3	1086	(444 - 1119.5)	362	155	(2.68 - 1351)	41	94.13	(5 - 543.81)	**0.004**
CD33 Expression, Standard Risk	384	3	1086	(444 - 1119.5)	350	199	(2.68 - 1351)	31	186.87	13.49 - 1025.07	**0.027**
CD33 Expression, Low Risk	305	3	86	(29.39 - 318.47)	261	99.75	(6.84 - 698.08)	41	85	(5 - 876)	0.575
CD33 Expression, High Risk	110	0	--	--	106	182.915	(4.27 - 1225.87)	4	217.125	(67 - 436.73)	0.737
CD33 Expression, FLT3+	132	0	--	--	127	207.24	(4.27 - 1225.87)	5	347	(85 - 436.73)	0.478
CD33 Expression, FLT3-	668	6	381.235	(29.39 - 1119.5)	591	136.84	(2.68 - 1351)	71	100.7	(5 - 1025.07)	0.119

Table 6. Association of CD33 SNPs with CD33 expression Quartiles

rs12459419	Total	CD33 Q1 (N=207)		CD33 Q2 (N=204)		CD33 Q3 (N=203)		CD33 Q4 (N=203)		Q1 vs. Q2 vs. Q3 vs. Q4	Q1 vs. Q2-4
	N	N	%	N	%	N	%	N	%	p	p
CC	415	62	30%	103	50%	105	52%	145	71%	<0.001	<0.001
TT	85	59	29%	23	11%	2	1%	1	0%	<0.001	<0.001
TC	316	85	41%	78	38%	96	47%	57	28%	0.001	0.415
rs3865444	**Total**	**CD33 Q1 (N=205)**		**CD33 Q2 (N=204)**		**CD33 Q3 (N=201)**		**CD33 Q4 (N=201)**		**Q1 vs. Q2 vs. Q3 vs. Q4**	**Q1 vs. Q2-4**
	N	N	%	N	%	N	%	N	%	p	p
GG	412	63	31%	103	50%	103	51%	143	71%	<0.001	<0.001
GT	313	83	40%	78	38%	96	48%	56	28%	0.001	0.519
TT	86	59	29%	23	11%	2	1%	2	1%	<0.001	<0.001
rs1803254	**Total**	**CD33 Q1 (N=206)**		**CD33 Q2 (N=204)**		**CD33 Q3 (N=203)**		**CD33 Q4 (N=203)**		**Q1 vs. Q2 vs. Q3 vs. Q4**	**Q1 vs. Q2-4**
SNPs	N	N	%	N	%	N	%	N	%	p	p
CC	40	23	11%	10	5%	4	2%	3	1%	<0.001	<0.001
CG	224	64	31%	61	30%	66	33%	33	16%	0.001	0.179
GG	552	119	58%	133	65%	133	66%	167	82%	<0.001	0.001
rs2455069	**Total**	**CD33 Q1 (N=203)**		**CD33 Q2 (N=200)**		**CD33 Q3 (N=198)**		**CD33 Q4 (N=201)**		**Q1 vs. Q2 vs. Q3 vs. Q4**	**Q1 vs. Q2-4**
	N	N	%	N	%	N	%	N	%	p	p
AA	312	117	58%	80	40%	66	33%	49	24%	<0.001	<0.001
AG	370	75	37%	89	45%	102	52%	104	52%	0.008	0.002
GG	120	11	5%	31	16%	30	15%	48	24%	<0.001	<0.001
rs35112940	**Total**	**CD33 Q1 (N=206)**		**CD33 Q2 (N=204)**		**CD33 Q3 (N=203)**		**CD33 Q4 (N=203)**		**Q1 vs. Q2 vs. Q3 vs. Q4**	**Q1 vs. Q2-4**
	N	N	%	N	%	N	%	N	%	p	p
AA	22	15	7%	6	3%	0	0%	1	0%	<0.001	<0.001
AG	174	57	28%	44	22%	42	21%	31	15%	0.024	0.010
GG	620	134	65%	154	75%	161	79%	171	84%	<0.001	<0.001
rs61736475	**Total**	**CD33 Q1 (N=207)**		**CD33 Q2 (N=203)**		**CD33 Q3 (N=201)**		**CD33 Q4 (N=202)**		**Q1 vs. Q2 vs. Q3 vs. Q4**	**Q1 vs. Q2-4**
	N	N	%	N	%	N	%	N	%	p	p
CC	6	1	0%	1	0%	0	0%	4	2%	0.107	1.000
TT	730	182	88%	183	90%	183	91%	182	90%	0.757	0.304
TC	77	24	12%	19	9%	18	9%	16	8%	0.633	0.227

rs12459419 C>T is associated with higher levels of exon 2 skipped spliced isoform and production of shorter CD33 isoform lacking IgV domain

Although mechanisms underlying the functional consequences of CD33 coding and regulatory SNPs are yet unknown, an alternatively spliced isoform of CD33 skipping exon 2 (D2 isoform) and thus lacking IgV domain has been reported previously(20-22). hP67.6-CD33 antibody as well as GO recognizes IgV region of CD33; thus, loss of this domain can not only interfere with detection of total CD33 (only detects full length but not short form lacking IgV domain (Perez-Oliva et al., *Glycobiology*, 2011;21(6):757-70), but can influence cellular sensitivity to GO, Figure 6A. CD33 SNP rs12459419 in exon 2 is present within 4 bp

of exonic junction, and C>T results in Ala>Val change at codon 14. This SNP also resides within and alters exonic splicing enhancer binding site for SRSF2 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home) and has been recently associated with exon 2 deletion (Raj et al., *Hum Mol Genet*, 2014;23(10):2729-36; et al., *Hum Mol Genet*, 2015;24(12):3557-70; Malik et al., *J Neurosci,* 2013;33(33):13320-5). This alternatively spliced, shorter isoform of CD33 missing exon-2 (D2 isoform) lacks the IgV domain. The exon-2 encoding IgV domain is known to be the epitope for hP67.6-CD33 antibody, which is used for diagnostic immunophenotyping as well as for the antibody that is conjugated to calicheamicin in GO. As a result, loss of this domain would not only interfere with detection of total CD33, but can impact binding, internalization and thus clinical efficacy of GO, Figure 6A.

We inquired whether minor allele is associated with alternate splicing. We used transcriptome sequencing data available from 80 patients from AAML0531 to correlate rs12459419 genotype with CD33 isoforms. We demonstrate that rs12459419 variant T allele is significantly associated with increased levels of alternatively spliced CD33-D2 isoform lacking exon2 (p=1.4E-10; Figure 6B).

Quantitative RT-PCR of RNA from diagnostic leukemic blasts from a completely different set of 30 AAML0531 samples (CC=9, CT=11, and TT=10) using isoform specific primers shown in Figure 1 showed significant association of T allele with increased levels isoform lacking exon 2 (Figure 6C; p= 0.02). No association of total CD33 transcript levels with rs12459419 was observed within this cohort (p>0.05) for both the sets investigated above (RNA-seq and RT-PCR).

Using transcriptome sequencing data we also show that rs12459419 T allele is associated with lower levels of exon 2 relative to levels of exon 4 which is constitutive as reflected by lower exon 2:exon 4 ratio in patients with CT or TT genotype in Figure 7.These results are consistent with lower cell surface intensity of CD33 in leukemic cells from patients with TT genotype as detected by current method that uses hP67.6.

CD33SNP_SCORE score is predictive of outcome by arm

Since 5 CD33 SNPs: rs12459419, rs3865444, rs1803254, rs2455069 and rs35112940 were associated with clinical outcome by arm and CD33 cell surface expression, and other than rs12459419 and rs3865444, none of the SNPs were in LD, a scoring system referred to

herein as "CD33SNP_Score" (ranging from 2 to -4) was created for each patient by adding the directional genotype scores of each SNP. Further CD33SNP_Scores were combined to generate a dichotomized score: Scores of ≥ 0 and <0 and demonstrated association with CD33 cell surface expression (Figure 8) . Patients with scores of ≥ **0** demonstrated better outcome with DFS being 49 ± 9% vs. 67 ± 8% (p=0.004); and RR being 24 ± 7% vs. 48 ± 9% (p <0.001) in NO-GO vs. GO arm, Table 7 and Figure 9. Analysis within each risk group showed significant improvement in outcome in Low-Risk group patients with score ≥ 0, DFS of 82±10% vs. 60±12% (p=0.005) and RR, 10±7% vs. 36±17% (p=0.001) in GO vs. No-GO arms. Similarly in standard risk group patients with score ≥ 0, RR was significantly lower in GO vs. No-Go arm (36 ±12% vs. 57±13%; p=0.030); in High-Risk group we observed a trend towards lower RR 38±25% vs. 70±32% in GO arm vs. No-GO (p=0.073); Table 7. The results in high-risk group category might be confounded by sample size. Conversely patients with score of <0 showed no benefit from addition of GO with DFS and RR being similar in both arms (p>0.05).

Rs12459419 was the primary driver of the CD33_SNP score with ~ 95% patients with score of ≥ 0 being CC genotype (no one with TT genotype) and around 94% of patients with score <0 having at least one T allele (with CT or TT genotype, indicating that single CD33 SNP-rs12459419 could be used as a biomarker to determine patients' response to CD33 targeted agents.

Table 7. Association of dichotomized CD33-3SNP score with clinical outcome by GO or No-GO arm in all patients and by risk group

	Score < 0 (N=405)					Score ≥ 0 (N=412)				
All Patients	**NoGO (N=210)**		**GO (N=195)**		**NoGO vs. GO**	**NoGO (N=198)**		**GO (N=214)**		**NoGO vs. GO**
	N		N		p	N		N		p
5yr OS from study entry	210	67 ± 7%	195	63 ± 7%	0.696	198	60 ± 7%	214	69 ± 7%	0.174
5yr EFS from study entry	210	49 ± 7%	195	49 ± 7%	0.579	198	43 ± 7%	214	56 ± 7%	**0.030**
5yr DFS from End of Course 1	137	56 ± 9%	153	53 ± 8%	0.909	143	49 ± 9%	156	67 ± 8%	**0.004**
5yr RR from End of Course 1	137	41 ± 9%	153	42 ± 8%	0.887	143	48 ± 9%	156	24 ± 7%	**<0.001**

Low Risk	**Score < 0 (N=143)**					**Score ≥ 0 (N=164)**				
	NoGO (N=76)		**GO (N=67)**		**NoGO vs. GO**	**NoGO (N=77)**		**GO (N=87)**		**NoGO vs. GO**
	N		N		p	N		N		p
5yr OS from study entry	76	87 ± 8%	67	76 ± 11%	0.123	77	75 ± 10%	87	89 ± 7%	**0.036**
5yr EFS from study entry	76	69 ± 11%	67	66 ± 12%	0.574	77	57 ± 12%	87	79 ± 9%	**0.003**
5yr DFS from End of Course 1	58	67 ± 13%	57	68 ± 12%	0.981	67	60 ± 12%	71	82 ± 10%	**0.005**
5yr RR from End of Course 1	58	30 ± 12%	57	25 ± 12%	0.620	67	36 ± 12%	71	10 ± 7%	**0.001**
Standard Risk	**Score < 0 (N=200)**					**Score ≥ 0 (N=185)**				
	NoGO (N=101)		**GO (N=99)**		**NoGO vs. GO**	**NoGO (N=94)**		**GO (N=91)**		**NoGO vs. GO**
	N		N		p	N		N		p
5yr OS from study entry	101	59 ± 10%	99	57 ± 11%	0.675	94	52 ± 11%	91	60 ± 10%	0.620
5yr EFS from study entry	101	39 ± 10%	99	41 ± 10%	0.319	94	35 ± 10%	91	47 ± 11%	0.278
5yr DFS from End of Course 1	58	49 ± 13%	77	42 ± 12%	0.840	64	40 ± 12%	65	58 ± 12%	0.073
5yr RR from End of Course 1	58	49 ± 13%	77	55 ± 12%	0.898	64	57 ± 13%	65	36 ± 12%	**0.030**
High Risk	**Score < 0 (N=57)**					**Score ≥ 0 (N=54)**				
	NoGO (N=30)		**GO (N=27)**		**NoGO vs. GO**	**NoGO (N=24)**		**GO (N=30)**		**NoGO vs. GO**
	N		N		p	N		N		p
5yr OS from study entry	30	38 ± 19%	27	54 ± 20%	0.251	24	46 ± 23%	30	45 ± 19%	0.993
5yr EFS from study entry	30	28 ± 17%	27	41 ± 19%	0.197	24	29 ± 19%	30	26 ± 16%	0.449
5yr DFS from End of Course 1	18	40 ± 25%	18	56 ± 23%	0.232	10	30 ± 29%	16	44 ± 25%	0.496
5yr RR from End of Course 1	18	54 ± 26%	18	39 ± 24%	0.293	10	70 ± 32%	16	38 ± 25%	0.090

Discussion

5 Here, common CD33 SNPs are reported that are significant predictors of clinical response to GO-based therapy in pediatric AML patients. GO is an anti-CD33 targeted antibody linked to cytotoxin calicheamicin and has shown promising results in AML. Results from meta-analysis studies further suggest GO provides survival benefits in patients with favorable cytogenetics(9,10). In pediatric AML, results from two Children's Oncology group 10 trials: AAML03P1 and AAML0531 demonstrated improved outcome with addition of GO to chemotherapy (11,12).

GO targets CD33 and higher CD33 expression levels have been previously correlated with in vitro GO chemosensitivity, indicating the significance of CD33 expression levels on GO response (25-27).

Results from COG AAML03P1 study demonstrated wide inter-patient variation (~ 2 log fold) in CD33 intensity on leukemic cells with higher expression correlating with negative prognostic factors such as FLT3-ITD (15). A follow-up study in a larger AAML0531 cohort randomly assigned to GO or standard chemotherapy arm (No-GO) reported that addition of GO is more likely to benefit patients with higher leukemic CD33 intensity, as compared to patients with low CD33 expression where no benefit of adding GO was observed (16).

Coding and regulatory polymorphisms in CD33 have previously been sequenced and identified by us (17,28). In the AAML031 study, rs12459419, rs2455069 and rs1803254 were associated with leukemic cell surface CD33 expression and rs35112940 was predictive of EFS and RR in patients who received GO (17). Since that AAML031 clinical trial did not have a randomized control arm, it was not possible to compare the impact of CD33 SNPs in patients who did or did not receive GO-based chemotherapy. AAML0531, being a randomized study, allowed investigation of the influence of CD33 SNPs on treatment outcome by comparing patients who did or did not receive GO.

CD33 SNPs, particularly rs12459419, were a significant determinant of whether patients would or would not benefit from addition of GO to conventional chemotherapy. Patients with CC genotype for rs12459419, which is also associated with higher cell surface expression of CD33 WT isoform had almost 50 % reduction in the risk of relapse and experience superior DFS with addition of GO to standard chemotherapy. Conversely for patients with at least one T allele, which results in alternate splicing with deletion of exon 2 and hence loss of IgV, no difference in outcome was observed with or without addition of GO. Since multiple SNPs were associated with cell surface expression and outcome, a CD33 SNP score was created. Patients with a CD33SNP score of ≥ 0 (around 50 % of patients) have survival benefit with better EFS, DFS and lower risk of relapse when treated with GO containing chemotherapy as compared to standard arm. In patients with score <0, there was no benefit from addition of GO. This relationship of CD33SNP score was consistent in low and standard risk group patients and demonstrated a similar trend in high-risk group patients. However, this score was primarily driven by rs12459419 (or linked rs3865444) SNP. In fact

for rs12459419 also, risk group analysis showed that low and standard risk group patients had significantly better outcome in CC genotype and similar trend in high-risk group patients. In fact in Low risk group patients overall survival (OS) from study entry was 90 ± 7% when patients had CC genotype and were given GO based therapy, in contrast CC patients in No-GO arm had OS 73 ± 11% (Table 9).

Table 9: CD33 SNPrs12459419 and clinical outcome in all patients and within each risk group by treatment arm NO-GO vs. GO in patients treated under COG-AAML0531 study.

rs12459419									
Low Risk	**CC (N=159)**			**CT (N=119)**			**TT (N=28)**		
	No-GO (N=76)	**GO (N=83)**	**No-GO vs. GO**	**No-GO (N=61)**	**GO (N=58)**	**No-GO vs. GO**	**No-GO (N=16)**	**GO (N=12)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**
5yr OS from study entry	73 ± 11%	90 ± 7%	**0.014**	88 ± 8%	74 ± 12%	0.058	88 ± 17%	83 ± 22%	0.761
5yr EFS from study entry	55 ± 12%	82 ± 9%	**0.001**	73 ± 11%	64 ± 13%	0.276	60 ± 26%	58 ± 28%	0.929
	No-GO (N=65)	**GO (N=69)**	**No-GO vs. GO**	**No-GO (N=46)**	**GO (N=48)**	**No-GO vs. GO**	**No-GO (N=14)**	**GO (N=11)**	**No-GO vs. GO**
5yr DFS from End of Course 1	57 ± 13%	84 ± 9%	**0.001**	72 ± 13%	70 ± 14%	0.794	61 ± 28%	55 ± 30%	0.683
5yr RR from End of Course 1	37 ± 13%	10 ± 8%	**<0.001**	26 ± 13%	19 ± 11%	0.483	39 ± 30%	45 ± 32%	0.683
5yr TRM from End of Course 1	6 ± 6%	6 ± 6%	0.944	2 ± 4%	11 ± 10%	0.107	0 ± 0%	0 ± 0%	1
Standard Risk	**CC (N=192)**			**CT (N=149)**			**TT (N=44)**		
	No-GO (N=101)	**GO (N=91)**	**No-GO vs. GO**	**No-GO (N=72)**	**GO (N=77)**	**No-GO vs. GO**	**No-GO (N=22)**	**GO (N=22)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**
5yr OS from study entry	49 ± 11%	56 ± 11%	0.599	64 ± 12%	60 ± 12%	0.936	59 ± 21%	64 ± 21%	0.684
5yr EFS from study entry	33 ± 10%	42 ± 10%	0.397	44 ± 12%	44 ± 12%	0.532	31 ± 20%	47 ± 23%	0.197
	No-GO (N=68)	**GO (N=67)**	**No-GO vs. GO**	**No-GO (N=40)**	**GO (N=60)**	**No-GO vs. GO**	**No-GO (N=14)**	**GO (N=15)**	**No-GO vs. GO**
5yr DFS from End of Course 1	38 ± 12%	51 ± 12%	0.181	55 ± 16%	46 ± 14%	0.721	43 ± 26%	50 ± 28%	0.439
5yr RR from End of Course 1	59 ± 12%	41 ± 12%	**0.056**	43 ± 16%	52 ± 14%	0.635	57 ± 28%	50 ± 30%	0.441
5yr TRM from End of Course 1	3 ± 4%	8 ± 7%	0.248	3 ± 5%	2 ± 3%	0.766	0 ± 0%	0 ± 0%	1
High Risk	**CC (N=55)**			**CT (N=45)**			**TT (N=11)**		
	No-GO (N=24)	**GO (N=31)**	**No-GO vs. GO**	**No-GO (N=24)**	**GO (N=21)**	**No-GO vs. GO**	**No-GO (N=6)**	**GO (N=5)**	**No-GO vs. GO**
	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**	% ± 2SE%	% ± 2SE%	**p**
5yr OS from study entry	46 ± 23%	40 ± 19%	0.767	32 ± 20%	62 ± 21%	0.056	56 ± 50%	--	0.639

5yr EFS from study entry	29 ± 19%	23 ± 15%	0.222	21 ± 18%	48 ± 22%	**0.037**	--	--	0.981
	No-GO (N=10)	**GO (N=14)**	**No-GO vs. GO**	**No-GO (N=14)**	**GO (N=16)**	**No-GO vs. GO**	**No-GO (N=4)**	**GO (N=4)**	**No-GO vs. GO**
5yr DFS from End of Course 1	30 ± 29%	43 ± 26%	0.5	28 ± 26%	56 ± 25%	0.112	--	--	0.62
5yr RR from End of Course 1	70 ± 32%	36 ± 27%	**0.073**	65 ± 31%	44 ± 26%	0.261	25 ± 50%	25 ± 50%	0.919
5yr TRM from End of Course 1	0 ± 0%	21 ± 23%	0.129	7 ± 14%	0 ± 0%	0.286	0 ± 0%	25 ± 53%	0.33

Although functional consequences of CD33 SNPs are still under investigation, the promoter SNP rs3865444 A>G was identified in a genome-wide association analysis to be associated with late-onset of Alzheimer's (29) and later confirmed by other studies (30-33).
5 The variant G allele was also shown to be associated with low CD33 expression levels in the microglial cells (34,35), which is consistent with our results in AML patients. Rs3865444 occurs in LD with rs12459419 (Ala14Val), which has recently been associated with transcript lacking exon 2 (22-24). CD33 alternately spliced transcripts characterized by skipping of exon 2 lacks IgV domain (21). As shown in Figure 6A, the shorter CD33 variant lacks the
10 IgV domain and thus would not be detected by hP67.7 antibody, as the epitope for hP67.7 lies in the IgV domain. Results on RNA transcript levels in diagnostic samples from AML patients presented herein confirm that the presence of T allele is associated with higher levels of alternatively spliced D2 isoform. GO targets the IgV domain, in consensus with the observation of lack of improvement in outcome with addition of GO in patients with variant
15 T allele. By contrast, the HIM-34 antibody targets the IgC2 domain and hence would detect both CD33-WT and shorter CD33-D2 isoform (exon 2 skipped) (21).

Although in depth mechanistic study of rs12459419 and other SNPs in CD33 is warranted, our results show that CD33 rs12459419 genotype allows for identification of patients expressing CD33 isoform lacking the antibody binding site for GO and holds
20 promise as a marker to select patients likely to benefit from addition GO to chemotherapy (CC genotype) regardless of clinical risk or surface CD33 expression level. This opens up opportunities to utilize patient genotypes for selection of CD33 targeted therapies. Additionally the fact that presence of splicing SNP results in shorter CD33 isoform lacking IgV domain warrants further investigation in developing CD33 immunoconjugates with
25 epitope targeted to regions not influenced by splicing and/or CD33 genotypes, thus being effective in larger patient cohort. The results of our study would also be relevant to other CD33 targeted therapies such as SGN-33A, lintuzumab etc. Since genotype calls are stable as

compared to use of CD33 mRNA or cell surface levels which is sensitive to- type of antibody utilized, storage conditions of specimens, heterogeneity of cell populations if not sorted or enriched; use of genotype as a marker for identification of patients will enhance clinical translation of CD33 genetics and will be adaptable in clinical practice.

5 ***References***

1. Walter RB, Appelbaum FR, Estey EH, Bernstein ID. Acute myeloid leukemia stem cells and CD33-targeted immunotherapy. Blood 2012;119(26):6198-208.

2. Bross PF, Beitz J, Chen G, Chen XH, Duffy E, Kieffer L, et al. Approval summary: gemtuzumab ozogamicin in relapsed acute myeloid leukemia. Clinical cancer research
10 : an official journal of the American Association for Cancer Research 2001;7(6):1490-6.

3. Petersdorf SH, Kopecky KJ, Slovak M, Willman C, Nevill T, Brandwein J, et al. A phase 3 study of gemtuzumab ozogamicin during induction and postconsolidation therapy in younger patients with acute myeloid leukemia. Blood 2013;121(24):4854-
15 60.

4. Burnett AK, Hills RK, Hunter AE, Milligan D, Kell WJ, Wheatley K, et al. The addition of gemtuzumab ozogamicin to low-dose Ara-C improves remission rate but does not significantly prolong survival in older patients with acute myeloid leukaemia: results from the LRF AML14 and NCRI AML16 pick-a-winner
20 comparison. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK 2013;27(1):75-81.

5. Burnett AK, Hills RK, Milligan D, Kjeldsen L, Kell J, Russell NH, et al. Identification of patients with acute myeloblastic leukemia who benefit from the addition of gemtuzumab ozogamicin: results of the MRC AML15 trial. Journal of
25 clinical oncology : official journal of the American Society of Clinical Oncology 2011;29(4):369-77.

6. Burnett AK, Russell NH, Hills RK, Kell J, Freeman S, Kjeldsen L, et al. Addition of gemtuzumab ozogamicin to induction chemotherapy improves survival in older patients with acute myeloid leukemia. Journal of clinical oncology : official journal of
30 the American Society of Clinical Oncology 2012;30(32):3924-31.

7. Castaigne S, Pautas C, Terre C, Raffoux E, Bordessoule D, Bastie JN, et al. Effect of gemtuzumab ozogamicin on survival of adult patients with de-novo acute myeloid leukaemia (ALFA-0701): a randomised, open-label, phase 3 study. Lancet 2012.

8. Thol F, Schlenk RF. Gemtuzumab ozogamicin in acute myeloid leukemia revisited. Expert opinion on biological therapy 2014;14(8):1185-95.

9. Hills RK, Castaigne S, Appelbaum FR, Delaunay J, Petersdorf S, Othus M, et al. Addition of gemtuzumab ozogamicin to induction chemotherapy in adult patients with acute myeloid leukaemia: a meta-analysis of individual patient data from randomised controlled trials. The lancet oncology 2014;15(9):986-96.

10. Loke J, Khan JN, Wilson JS, Craddock C, Wheatley K. Mylotarg has potent anti-leukaemic effect: a systematic review and meta-analysis of anti-CD33 antibody treatment in acute myeloid leukaemia. Annals of hematology 2015;94(3):361-73.

11. Cooper TM, Franklin J, Gerbing RB, Alonzo TA, Hurwitz C, Raimondi SC, et al. AAML03P1, a pilot study of the safety of gemtuzumab ozogamicin in combination with chemotherapy for newly diagnosed childhood acute myeloid leukemia: a report from the Children's Oncology Group. Cancer 2012;118(3):761-9.

12. Gamis AS, Alonzo TA, Meshinchi S, Sung L, Gerbing RB, Raimondi SC, et al. Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves event-free survival by reducing relapse risk: results from the randomized phase III Children's Oncology Group trial AAML0531. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2014;32(27):3021-32.

13. Clarke WT, Marks PW. Gemtuzumab ozogamicin: is there room for salvage? Blood 2010;116(14):2618-9.

14. Ravandi F, Estey EH, Appelbaum FR, Lo-Coco F, Schiffer CA, Larson RA, et al. Gemtuzumab ozogamicin: time to resurrect? Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30(32):3921-3.

15. Pollard JA, Alonzo TA, Loken M, Gerbing RB, Ho PA, Bernstein ID, et al. Correlation of CD33 expression level with disease characteristics and response to gemtuzumab ozogamicin containing chemotherapy in childhood AML. Blood 2012.

16. Pollard JA, Loken M, Gerbing RB, Raimondi SC, Hirsch BA, Aplenc R, et al. CD33 Expression and Its Association With Gemtuzumab Ozogamicin Response: Results From the Randomized Phase III Children's Oncology Group Trial AAML0531. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2016;34(7):747-55.

17. Mortland L, Alonzo TA, Walter RB, Gerbing RB, Mitra AK, Pollard JA, et al. Clinical significance of CD33 nonsynonymous single-nucleotide polymorphisms in pediatric patients with acute myeloid leukemia treated with gemtuzumab-ozogamicin-containing chemotherapy. Clinical cancer research : an official journal of the American Association for Cancer Research 2013;19(6):1620-7.

18. Pollard JA, Alonzo TA, Loken M, Gerbing RB, Ho PA, Bernstein ID, et al. Correlation of CD33 expression level with disease characteristics and response to gemtuzumab ozogamicin containing chemotherapy in childhood AML. Blood 2012;119(16):3705-11.

19. Wells DAL, M.R. Flow cytometric mean fluorescence intensity: The biophysics behind the number. Leukemia research 2008;32(6).

20. Hernandez-Caselles T, Martinez-Esparza M, Perez-Oliva AB, Quintanilla-Cecconi AM, Garcia-Alonso A, Alvarez-Lopez DM, et al. A study of CD33 (SIGLEC-3) antigen expression and function on activated human T and NK cells: two isoforms of CD33 are generated by alternative splicing. J Leukoc Biol 2006;79(1):46-58.

21. Perez-Oliva AB, Martinez-Esparza M, Vicente-Fernandez JJ, Corral-San Miguel R, Garcia-Penarrubia P, Hernandez-Caselles T. Epitope mapping, expression and post-translational modifications of two isoforms of CD33 (CD33M and CD33m) on lymphoid and myeloid human cells. Glycobiology 2011;21(6):757-70.

22. Raj T, Ryan KJ, Replogle JM, Chibnik LB, Rosenkrantz L, Tang A, et al. CD33: increased inclusion of exon 2 implicates the Ig V-set domain in Alzheimer's disease susceptibility. Hum Mol Genet 2014;23(10):2729-36.

23. Malik M, Chiles J, 3rd, Xi HS, Medway C, Simpson J, Potluri S, et al. Genetics of CD33 in Alzheimer's disease and acute myeloid leukemia. Hum Mol Genet 2015;24(12):3557-70.

24. Malik M, Simpson JF, Parikh I, Wilfred BR, Fardo DW, Nelson PT, et al. CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. J Neurosci 2013;33(33):13320-5.

25. Jager E, van der Velden VH, te Marvelde JG, Walter RB, Agur Z, Vainstein V. Targeted drug delivery by gemtuzumab ozogamicin: mechanism-based mathematical model for treatment strategy improvement and therapy individualization. PloS one 2011;6(9):e24265.

26. Jedema I, Barge RM, van der Velden VH, Nijmeijer BA, van Dongen JJ, Willemze R, et al. Internalization and cell cycle-dependent killing of leukemic cells by Gemtuzumab Ozogamicin: rationale for efficacy in CD33-negative malignancies with endocytic capacity. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK 2004;18(2):316-25.

27. Walter RB, Gooley TA, van der Velden VH, Loken MR, van Dongen JJ, Flowers DA, et al. CD33 expression and P-glycoprotein-mediated drug efflux inversely correlate and predict clinical outcome in patients with acute myeloid leukemia treated with gemtuzumab ozogamicin monotherapy. Blood 2007;109(10):4168-70.

28. Lamba JK, Pounds S, Cao X, Downing JR, Campana D, Ribeiro RC, et al. Coding polymorphisms in CD33 and response to gemtuzumab ozogamicin in pediatric patients with AML: a pilot study. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK 2009;23(2):402-4.

29. Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet 2011;43(5):436-41.

30. Bao J, Wang XJ, Mao ZF. Associations Between Genetic Variants in 19p13 and 19q13 Regions and Susceptibility to Alzheimer Disease: A Meta-Analysis. Med Sci Monit 2016;22:234-43.

31. Omoumi A, Fok A, Greenwood T, Sadovnick AD, Feldman HH, Hsiung GY. Evaluation of late-onset Alzheimer disease genetic susceptibility risks in a Canadian population. Neurobiol Aging 2014;35(4):936 e5-12.

32. Tan L, Yu JT, Zhang W, Wu ZC, Zhang Q, Liu QY, et al. Association of GWAS-linked loci with late-onset Alzheimer's disease in a northern Han Chinese population. Alzheimers Dement 2013;9(5):546-53.

33. Carrasquillo MM, Belbin O, Hunter TA, Ma L, Bisceglio GD, Zou F, et al. Replication of EPHA1 and CD33 associations with late-onset Alzheimer's disease: a multi-centre case-control study. Mol Neurodegener 2011;6(1):54.

34. Bradshaw EM, Chibnik LB, Keenan BT, Ottoboni L, Raj T, Tang A, et al. CD33 Alzheimer's disease locus: altered monocyte function and amyloid biology. Nat Neurosci 2013;16(7):848-50.

35. Griciuc A, Serrano-Pozo A, Parrado AR, Lesinski AN, Asselin CN, Mullin K, et al. Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. Neuron 2013;78(4):631-43.

CLAIMS

What is claimed is:

1. A method of treating a subject with a cancer expressing CD33 comprising:

performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism rs12459419, wherein the genotype may be CC, TC, or TT; and

administering a therapeutically effective amount of an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide polymorphism rs12459419.

2. The method of claim 1, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof.

3. The method of claim 1, wherein the antibody that selectively binds CD33 is a humanized antibody.

4. The method of claim 1, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof, conjugated to a toxin.

5. The method of any one of claims 1-4, wherein the agent that selectively binds to CD33 selectively binds to amino acids encoded by exon 2 of CD33.

6. The method of claim 1, wherein the agent that selectively binds to CD33 is gemtuzumab ozogamicin, hP67.7, or SGN-33A.

7. The method of any one of claims 1-6, wherein the subject is a pediatric subject.

8. The method of any one of claims 1-7 wherein the subject is an adult subject.

9. The method of any one of claims 1-8, wherein the subject is treated with chemotherapy within thirty days of the administration of the antibody.

10. The method of any one of claims 1-9, wherein the assay is performed by DNA sequencing analysis.

11. The method of any one of claims 1-10, wherein the cancer is any one of acute lymphoblastic leukemia (ALL), acute promyelocytic leukemia (APL), and acute myeloid leukemia (AML).

12. The method of any one of claims 1-11, wherein the assay is a hybridization assay.

5 13. The method of claim 11, wherein the cancer is AML.

14. The method of any one of claims 1-13, wherein the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis; anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

10 15. A method for determining whether a subject with a cancer expressing CD33 is likely to benefit from treatment with an agent that selectively binds to CD33 comprising:

providing tissue from a subject who has been diagnosed with the cancer;

performing an assay on the tissue, or on a derivative of the tissue, to detect the genotype of the subject for the CD33 single-nucleotide polymorphism rs12459419, wherein
15 the genotype may be CC, TC, or TT;

wherein the subject is likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide polymorphism rs12459419 and the subject is not likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a TC or TT genotype for the CD33
20 single-nucleotide polymorphism rs12459419.

16. The method of claim 15, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof.

17. The method of any one of claims 16, wherein the agent that selectively binds to CD33 selectively binds to amino acids encoded by exon 2 of CD33.

25 18. The method of claim 15, wherein the agent that selectively binds to CD33 is gemtuzumab ozogamicin, hP67.7, or SGN-33A.

19. The method of any one of claims 15-18, wherein the subject is a pediatric subject.

20. The method of any one of claims 15-19, wherein the subject is an adult subject.

21. The method of any one of claims 15-20, wherein the assay comprises performing DNA sequencing analysis.

22. The method of any one of claims 15-20, wherein the assay comprises contacting a derivative of the tissue with a nucleic acid probe.

5 23. The method of any one of claims 15-20 and 22, wherein the assay is a hybridization assay.

24. The method of any one of claims 15-23, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

25. The method of any one of claims 15-24, wherein the subject has one or more of: the 10 presence of blast cells that express CD33 within the hematopoietic system; leukostasis; anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

26. A method for detecting a polymorphism, comprising:

obtaining a biological sample of a subject that has a cancer expressing CD33, and

15 performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism (SNP) rs12459419 controlling expression of exon 2, wherein the genotype may be CC, TC, or TT and wherein the presence of the CC genotype in rs12459419 indicates expression of exon 2 of CD33.

27. The method of 26, wherein the assay is a hybridization assay comprising a probe that 20 hybridizes specifically to the CC genotype but not the TC or the TT genotypes.

28. The method of claim 27, wherein the hybridization assay further comprises a probe that hybridizes specifically to the TT genotype but not the CC or the TC genotypes and a probe that hybridizes specifically to the TC genotype but not the CC or the TT genotypes.

29. The method claim 28, comprising detecting specific hybridization of the probes that 25 bind specifically the CD33 single-nucleotide polymorphism rs12459419 to their respective genotype.

30. The method of claim 27, wherein the hybridization assay comprises

(a) detecting hybridization of a probe that binds to the a nucleic acid from the biological sample, and

(b) detecting a variant nucleic acid of CD33 single-nucleotide polymorphism rs12459419 in the sample when hybridization is detected.

5 31. The method claim 28, further comprising performing a hybridization assay with the probes and a control genotype.

32. The method of 26, wherein the assay is a genomic sequencing assay.

33. The method of 26, wherein the assay is a DNA sequencing, RNA sequencing, primer extension, enzyme-based, restriction fragment length polymorphism, PCR-based, PCR-
10 RFLP, allele-specific PCR, flap endonuclease, 5'- nuclease, oligonucleotide ligation, SNPlex, surveyor nuclease, dynamic allele-specific hybridization, molecular beacons, or SNP microarray assay.

34. The method of claim 30, wherein the genomic assay comprises:

direct sequencing of a nucleic acid containing polymorphism rs12459419, and

15 detecting the presence of the CC, TC or TT genotype.

35. The method of claim 34, wherein the nucleic acid is DNA, genomic DNA, RNA, cDNA, hnRNA or mRNA.

36. The method of any one of claims 15-35, wherein the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis;
20 anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

37. The method of any one of claims 26-36, wherein the probe comprises a nucleotide sequence complementary to a sequence listed within Table 1.

38. The method of any one of claims 26-37, wherein the probe comprises a nucleotide
25 sequence complementary to nucleotides of SEQ ID NO:1.

39. The method of any one of claims 26-37, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

40. A method for detecting a polymorphism, comprising:

obtaining a biological sample of a subject that has a cancer expressing CD33,

performing an assay to detect the presence of amino acids encoded by exon 2 of CD33, wherein the presence of amino acids encoded by exon 2 of CD33 indicates expression
5 of exon 2 of CD33 and presence of a CC genotype in single-nucleotide polymorphism rs12459419 of CD33.

41. The method of claim 40, wherein the assay is an immunoassay.

42. The method of claim 40, wherein the assay is a protein sequencing assay.

43. A kit comprising:

10 an agent that selectively binds to CD33, and

instructions indicating the use of the agent to treat a subject if the genotype of the subject for the CD33 single-nucleotide polymorphism rs12459419 is CC.

44. The kit of claim 43, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof.

15 45. The kit of claim 43, wherein the antibody that selectively binds CD33 is a humanized antibody.

46. The kit of claim 43, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof, conjugated to a toxin.

20 47. The kit of any one of claims 43-46, wherein the agent that selectively binds to CD33 selectively binds to exon 2 of CD33.

48. The kit of claim 43, wherein the agent that selectively binds to CD33 is gemtuzumab ozogamicin.

49. The kit of any one of claims 43-48, wherein the subject is a pediatric subject.

25 50. The kit of any one of claims 43-49, wherein the subject is an adult subject.

51. A method for determining a CD33SNP_Score for a subject comprising:

determining genotype scores of the subject for the CD33 single-nucleotide polymorphisms (SNPs) rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069, wherein

the genotype score for a single nucleotide polymorphism (SNP) with two wild-type alleles is 0,

the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with one wild-type allele and one variant allele is -1,

the genotype score for the SNP rs2455069 with one wild-type allele and one variant allele is 1,

the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with two variant alleles is -2, and

the genotype score for the SNP rs2455069 with two variant alleles is 2, and

adding the genotype scores to yield the CD33SNP_Score.

52. The method of claim 51, wherein the genotype of each SNP is determined by an assay.

53. The method of claim 52, wherein the assay is performed by DNA sequencing analysis.

54. The method of claim 52, wherein the assay is a hybridization assay.

55. The method of any one of claims 51-54, wherein the subject is a pediatric subject.

56. The method of any one of claims 51-54 wherein the subject is an adult subject.

57. A method for determining a CD33SNP_Score for a subject comprising adding genotype scores of the subject for the CD33 single-nucleotide polymorphisms rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069 to yield the CD33SNP_Score, wherein

the genotype score for a single nucleotide polymorphism (SNP) with two wild-type alleles is 0,

the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with one wild-type allele and one variant allele is -1,

the genotype score for the SNP rs2455069 with one wild-type allele and one variant allele is 1,

the genotype score separately for each of the SNPs rs12459419, rs3865444, rs1803254, and rs35112940 with two variant alleles is -2, and

5 the genotype score for the SNP rs2455069 with two variant alleles is 2.

58. The method of claim 57, wherein the genotype of each SNP is determined by an assay.

59. The method of claim 58, wherein the assay is performed by DNA sequencing analysis.

10 60. The method of claim 58, wherein the assay is a hybridization assay.

61. The method of any one of claims 57-60, wherein the subject is a pediatric subject.

62. The method of any one of claims 57-60, wherein the subject is an adult subject.

63. A method for determining whether a subject with cancer is likely to benefit from treatment with an agent that selectively binds to CD33 comprising:

15 determining a CD33SNP_Score for the subject, wherein

tissue is provided from the subject who has been diagnosed with the cancer;

an assay is performed on the tissue, or on a derivative of the tissue, to detect the genotype of the subject for the CD33 single-nucleotide polymorphisms rs12459419 or rs3865444, rs1803254, rs35112940, and rs2455069, wherein the wild-type, single variant, or
20 double variant genotype, respectively may be:

CC, TC, or TT for rs12459419,

CC, CA, or AA for rs3865444,

GG, CG, or CC for rs1803254,

GG, AG, or AA for rs35112940, or

25 AA, AG, or GG for rs2455069, and

wherein the subject is likely to benefit from treatment with an agent that selectively binds amino acids encoded by exon 2 of CD33 if the CD33SNP_Score for the subject is greater than or equal to zero.

64. The method of claim 63, wherein the assay is performed by DNA sequencing analysis.

65. The method of claim 63, wherein the assay is a hybridization assay.

66. The method of any one of claims 63-65, wherein the subject is a pediatric subject.

67. The method of any one of claims 63-65, wherein the subject is an adult subject.

68. The method of any one of claims 63-67, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

69. The method of any one of claims 63-68, wherein the tissue comprises CD33 expressing cells, e.g., blast cells comprising CD33.

70. A method for determining whether a subject with cancer expressing CD33 is likely to benefit from treatment with an agent that selectively binds to CD33 comprising:

providing tissue from a subject who has been diagnosed with the cancer;

performing an assay on the tissue, or on a derivative of the tissue, to detect the CD33 single-nucleotide polymorphism genotype of the subject for the CD33 single-nucleotide polymorphism rs12459419 or the CD33 single-nucleotide polymorphism rs3865444, and

determining the single-nucleotide polymorphism genotype score, wherein the genotype score may be 0, -1, or -2, wherein a score of 0 indicates that the subject is likely to benefit from treatment.

71. The method of claim 70, wherein the assay is performed by DNA sequencing analysis.

72. The method of claim 70, wherein the assay is a hybridization assay.

73. The method of any one of claims 70-72, wherein the subject is a pediatric subject.

74. The method of any one of claims 70-72, wherein the subject is an adult subject.

75. The method of any one of claims 70-74, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

76. The method of any one of claims 70-75, wherein the tissue comprises CD33 expressing cells, e.g., blast cells comprising CD33.

5 77. A method of treating a subject with a cancer expressing CD33 comprising:

performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism rs3865444, wherein the genotype may be CC, CA, or AA; and

administering a therapeutically effective amount of an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide polymorphism 10 rs3865444.

78. A method of treating a subject with a cancer expressing CD33 comprising:

performing an assay to detect the genotype of the subject for any one of the CD33 single-nucleotide polymorphisms rs1354106, rs3852865, and rs12985029, wherein the genotype may be wild-type, heterozygous variant, or homozygous variant; and

15 administering a therapeutically effective amount of an agent that selectively binds to CD33 if the subject exhibits a wild-type genotype for the CD33 single-nucleotide polymorphism rs1354106, rs3852865, or rs12985029.

79. The method of claims 77 or 78, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof.

20 80. The method of claim 77 or 78, wherein the antibody that selectively binds CD33 is a humanized antibody.

81. The method of claim 77 or 78, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof , conjugated to a toxin.

25 82. The method of any one of claims 77 or 78, wherein the agent that selectively binds to CD33 selectively binds to amino acids encoded by exon 2 of CD33.

83. The method of claim 77 or 78, wherein the agent that selectively binds to CD33 is gemtuzumab ozogamicin, hP67.7, or SGN-33A.

84. The method of any one of claims 77-83, wherein the subject is a pediatric subject.

85. The method of any one of claims 77-84, wherein the subject is an adult subject.

86. The method of any one of claims 77-85, wherein the subject is treated with chemotherapy within thirty days of the administration of the antibody.

5 87. The method of any one of claims 77-86, wherein the assay is performed by DNA sequencing analysis.

88. The method of any one of claims 77-87, wherein the cancer is any one of ALL, APL, and AML.

89. The method of any one of claims 77-88, wherein the assay is a hybridization assay.

10 90. The method of any one of claims 77-89, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

91. A method for determining whether a subject with a cancer expressing CD33 is likely to benefit from treatment with an agent that selectively binds to CD33 comprising:

providing tissue from a subject who has been diagnosed with the cancer;

15 performing an assay on the tissue, or on a derivative of the tissue, to detect the genotype of the subject for the CD33 single-nucleotide polymorphism rs3865444, wherein the genotype may be CC, CA, or AA;

wherein the subject is likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a CC genotype for the CD33 single-nucleotide
20 polymorphism rs3865444 and the subject is not likely to benefit from treatment with an agent that selectively binds to CD33 if the subject exhibits a CA or AA genotype for the CD33 single-nucleotide polymorphism rs3865444.

92. The method of claim 91, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof.

25 93. The method of any one of claims 91 and 92, wherein the agent that selectively binds to CD33 selectively binds to exon 2 of CD33.

94. The method of claim 91, wherein the agent that selectively binds to CD33 is gemtuzumab ozogamicin, hP67.7, or SGN-33A.

95. The method of any one of claims 91-94, wherein the subject is a pediatric subject.

96. The method of any one of claims 91-95 wherein the subject is an adult subject.

5 97. The method of any one of claims 91-96, wherein the assay comprises performing DNA sequencing analysis.

98. The method of any one of claims 91-96, wherein the assay comprises contacting a derivative of the tissue with a nucleic acid probe.

99. The method of any one of claims 91-96 and 98, wherein the assay is a hybridization
10 assay.

100. The method of any one of claims 91-99, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

101. A method for detecting a polymorphism, comprising:

obtaining a biological sample of a subject that has a cancer expressing CD33, and

15 performing an assay to detect the genotype of the subject for the CD33 single-nucleotide polymorphism (SNP) rs3865444 controlling expression of exon 2, wherein the genotype may be CC, CA, or AA and wherein the presence of the CC genotype in rs3865444 indicates expression of exon 2 of CD33.

102. The method of 101, wherein the assay is a hybridization assay comprises a probe that
20 hybridizes specifically to the CC genotype but not the CA or the AA genotypes.

103. The method of claim 102, wherein the hybridization assay further comprises a probe that hybridizes specifically to the AA genotype but not the CC or the CA genotypes and a probe that hybridizes specifically to the CA genotype but not the CC or the AA genotypes.

104. The method claim 103, comprising detecting specific hybridization of the probes that
25 bind specifically the CD33 single-nucleotide polymorphism rs3865444 to their respective genotype.

105. The method of claim 102, wherein the hybridization assay comprises:

(a) detecting hybridization of a probe that binds to the a nucleic acid from the biological sample, and

(b) detecting a variant nucleic acid of CD33 single-nucleotide polymorphism rs3865444 in the sample when hybridization is detected.

5 106. The method claim 103, further comprising performing a hybridization assay with the probes and a control genotype.

107. The method of 101, wherein the assay is a genomic sequencing assay.

108. The method of 101, wherein the assay is a DNA sequencing, RNA sequencing, primer extension, enzyme-based, restriction fragment length polymorphism, PCR-based, PCR-
10 RFLP, allele-specific PCR, flap endonuclease, 5’- nuclease, oligonucleotide ligation, SNPlex, surveyor nuclease, dynamic allele-specific hybridization, molecular beacons, or SNP microarray assay.

109. The method of claim 105, wherein the genomic assay comprises:

direct sequencing of a nucleic acid containing polymorphism rs3865444, and

15 detecting the presence of the CC, CA or AA genotype.

110. The method of claim 109, wherein the nucleic acid is DNA, genomic DNA, RNA, cDNA, hnRNA or mRNA.

111. The method of any one of claims 51-110, wherein the subject has one or more of: the presence of blast cells that express CD33 within the hematopoietic system; leukostasis;
20 anemia; leukopenia; neutropenia; thrombocytopenia; chloroma; granulocytic sarcoma; and myeloid sarcoma.

112. The method of any one of claims 101-111, wherein the probe comprises a nucleotide sequence complementary to a sequence listed within Table 1.

113. The method of any one of claims 101-112, wherein the probe comprises a nucleotide
25 sequence complementary to nucleotides of SEQ ID NO:2.

114. The method of any one of claims 101-112, wherein the cancer is acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), or acute promyelocytic leukemia (APL).

115. A kit comprising:

an agent that selectively binds to CD33, and

instructions indicating the use of the agent to treat a subject if the genotype of the subject for the CD33 single-nucleotide polymorphism rs3865444 is CC.

5 116. The kit of claim 115, wherein the agent that selectively binds to CD33 comprises an antibody that selectively binds CD33, or an antigen binding fragment thereof.

117. The kit of claim 115, wherein the antibody that selectively binds CD33 is a humanized antibody.

118. The kit of claim 115, wherein the agent that selectively binds to CD33 comprises an

10 antibody that selectively binds CD33, or an antigen binding fragment thereof , conjugated to a toxin.

119. The kit of any one of claims 115-118, wherein the agent that selectively binds to CD33 selectively binds to exon 2 of CD33.

120. The kit of claim 115, wherein the agent that selectively binds to CD33 is gemtuzumab

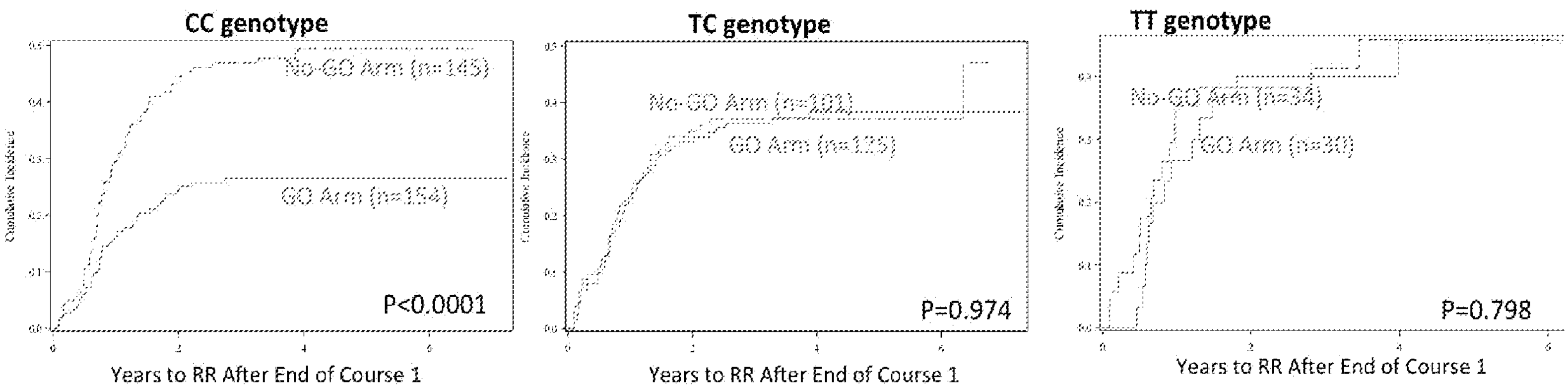
15 ozogamicin.

121. The kit of any one of claims 115-120, wherein the subject is a pediatric subject.

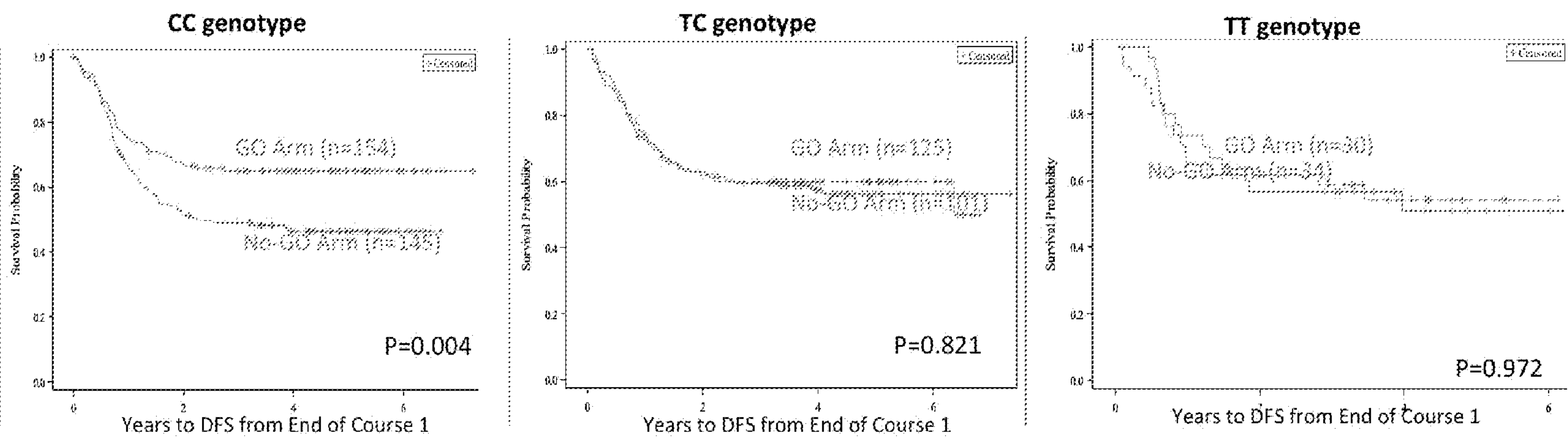
122. The kit of any one of claims 115-121, wherein the subject is an adult subject.

Figure 1

Risk of Relapse from end of Course 1- rs12459419



DFS from End of Course 1- rs12459419



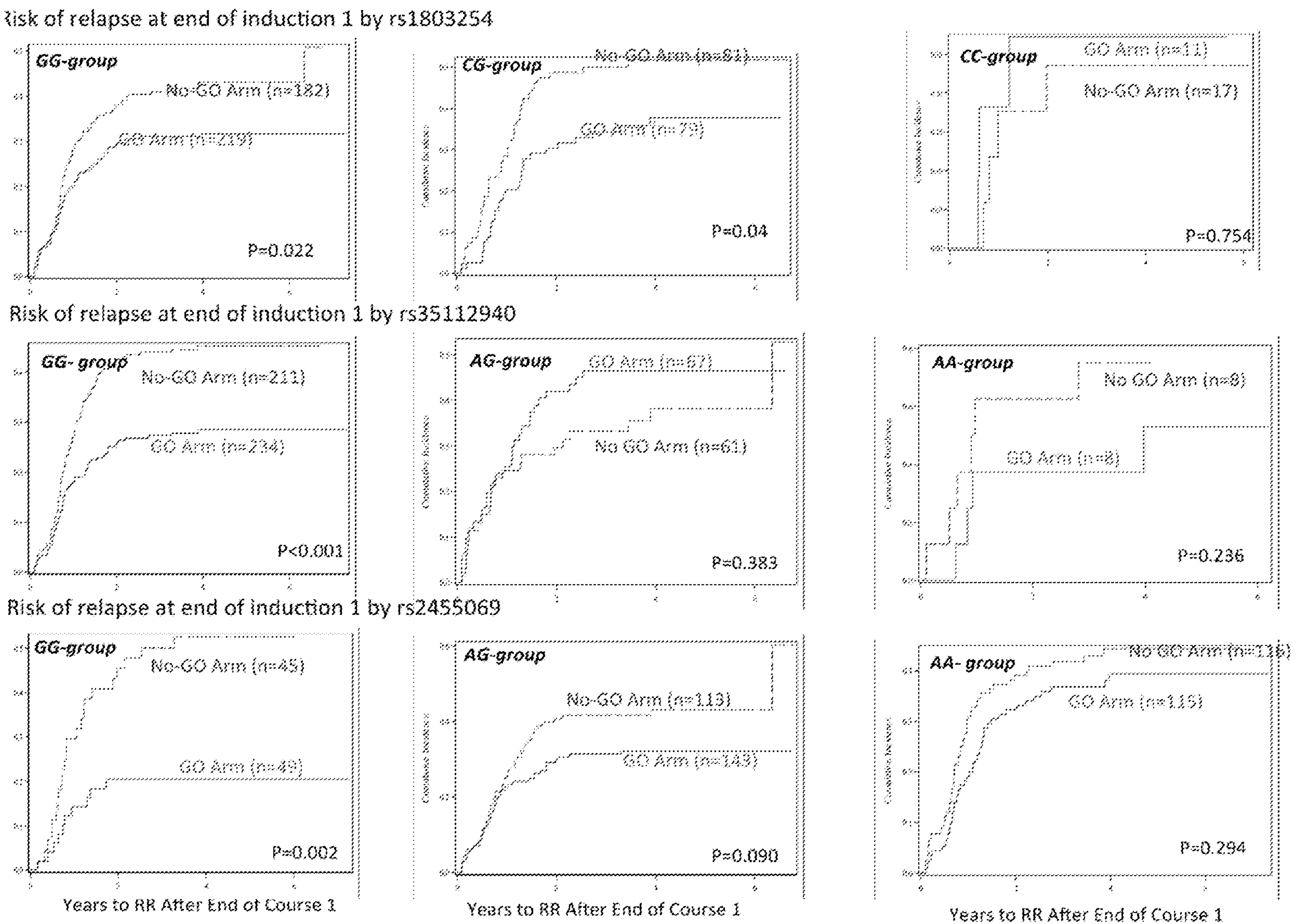


Figure 2
Risk of relapse at end of induction 1 by rs1803254
GG-group
No-GO Arm (n=182)
GO Arm (n=219)
P=0.022
CG-group
No-GO Arm (n=81)
GO Arm (n=79)
P=0.04
CC-group
GO Arm (n=11)
No-GO Arm (n=17)
P=0.754
Risk of relapse at end of induction 1 by rs35112940
GG- group
No-GO Arm (n=211)
GO Arm (n=234)
P<0.001
AG-group
GO Arm (n=67)
No GO Arm (n=61)
P=0.383
AA-group
No GO Arm (n=8)
GO Arm (n=8)
P=0.236
Risk of relapse at end of induction 1 by rs2455069
GG-group
No-GO Arm (n=45)
GO Arm (n=49)
P=0.002
AG-group
No-GO Arm (n=113)
GO Arm (n=143)
P=0.090
AA- group
No GO Arm (n=116)
GO Arm (n=115)
P=0.294
Years to RR After End of Course 1
Years to RR After End of Course 1
Years to RR After End of Course 1

Figure 3

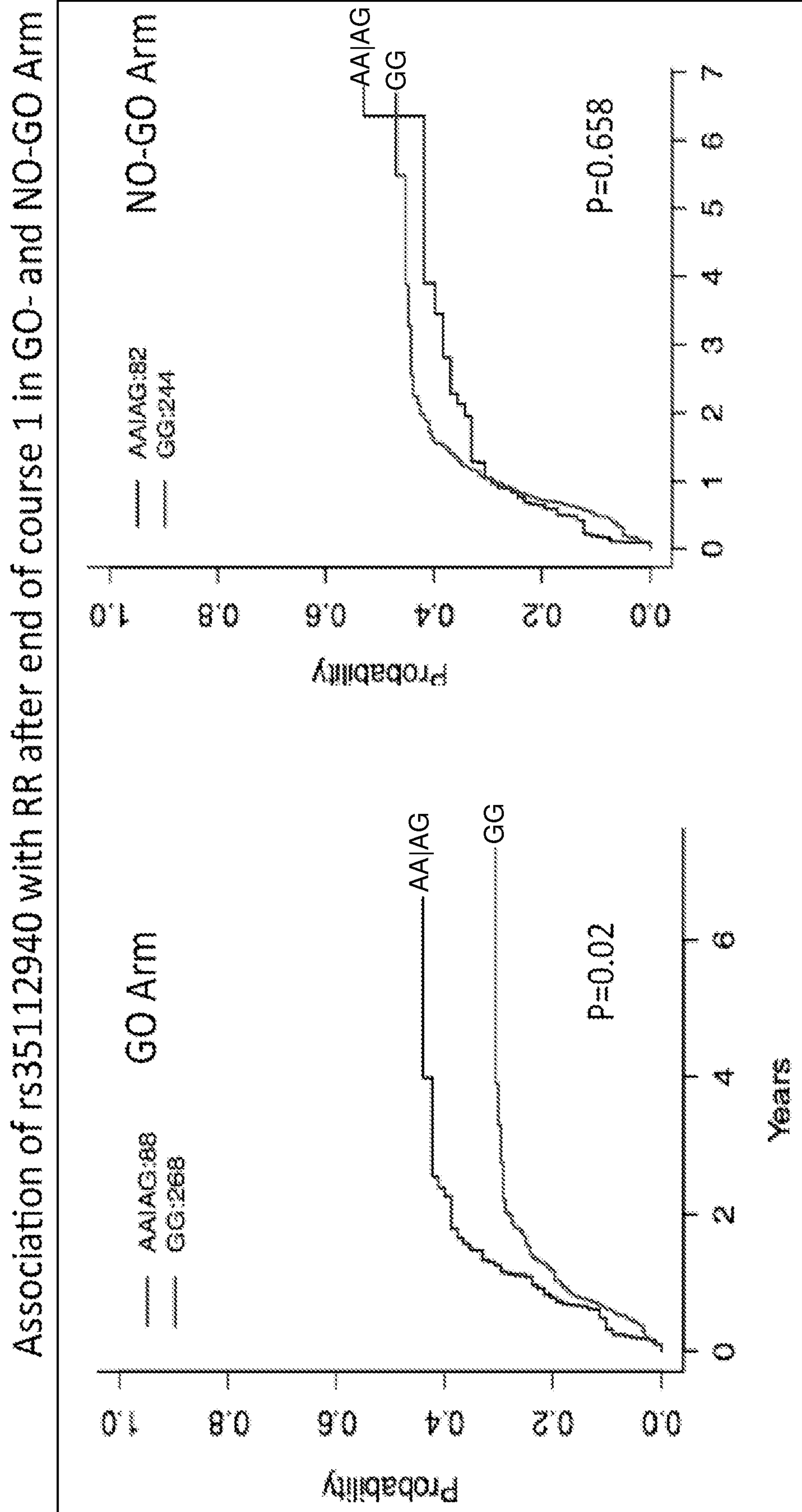


Figure 4

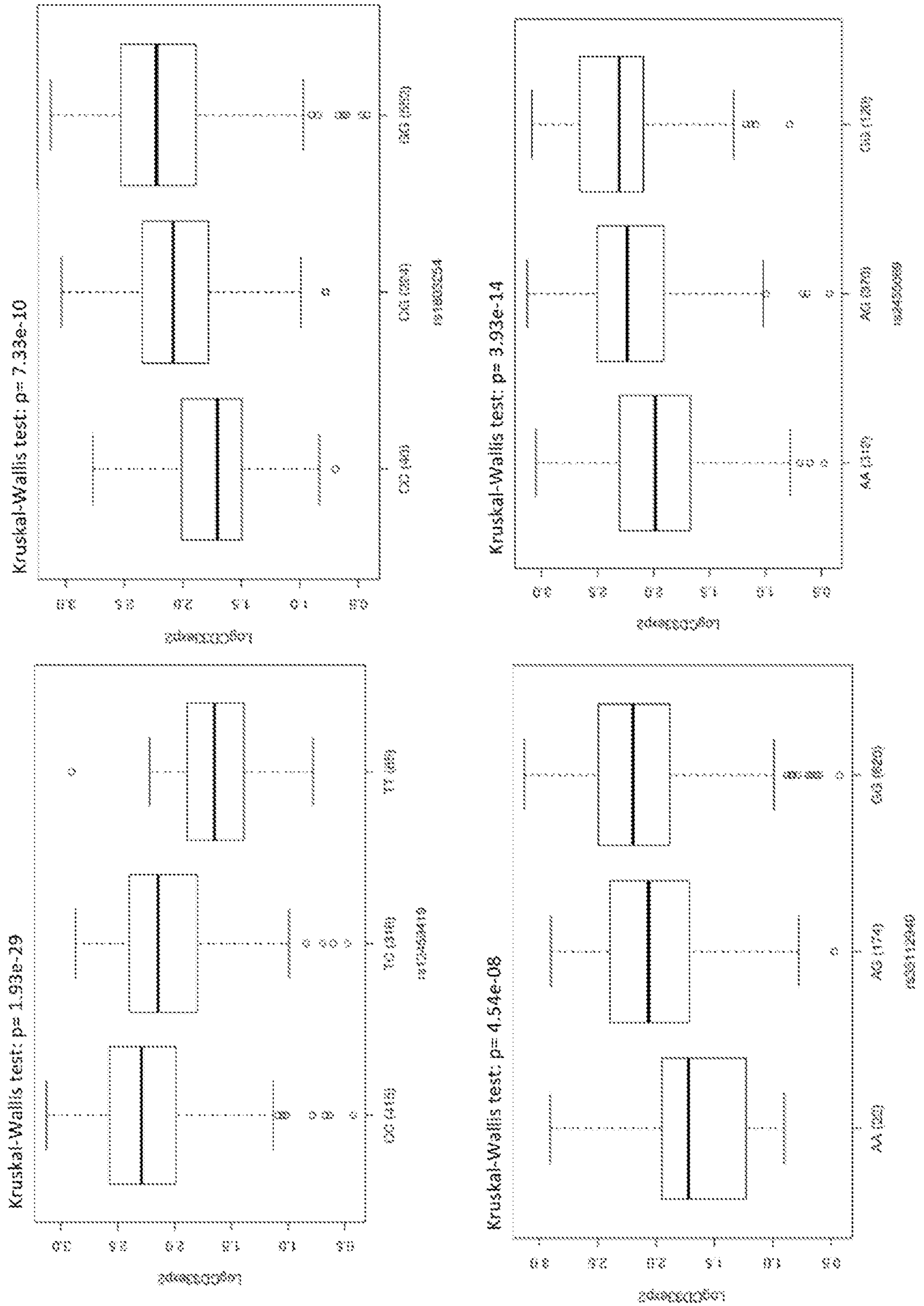
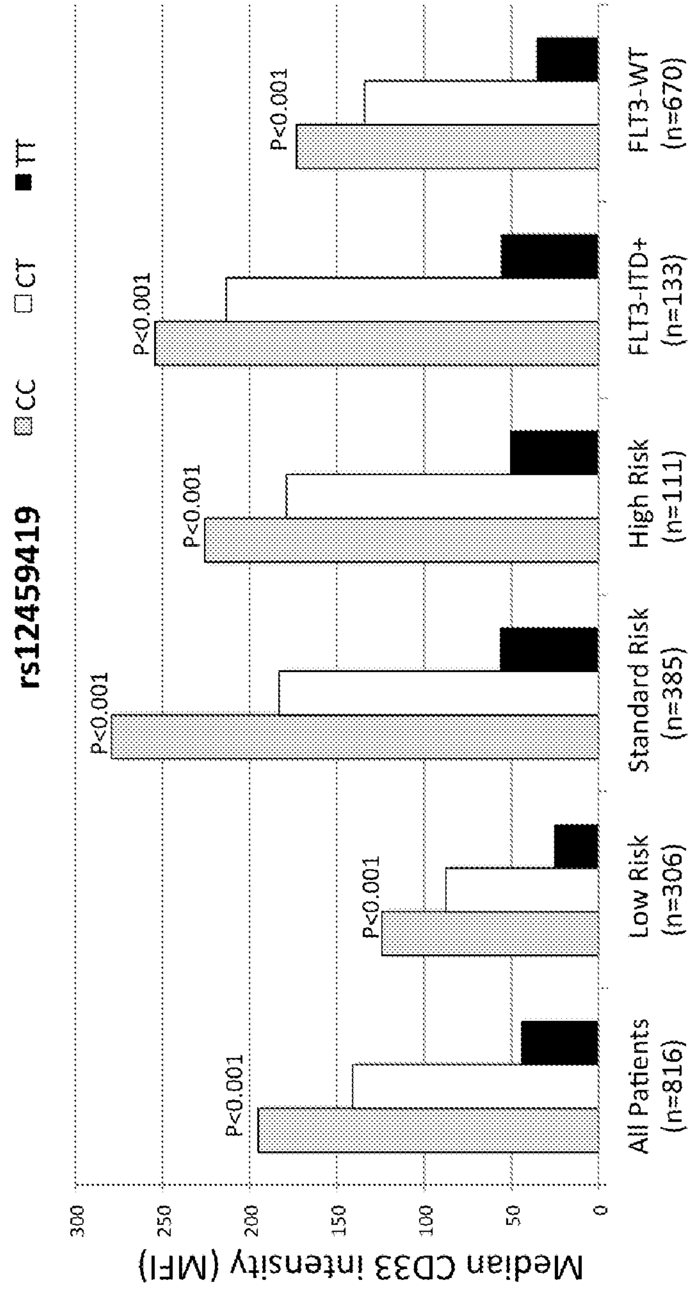


Figure 5



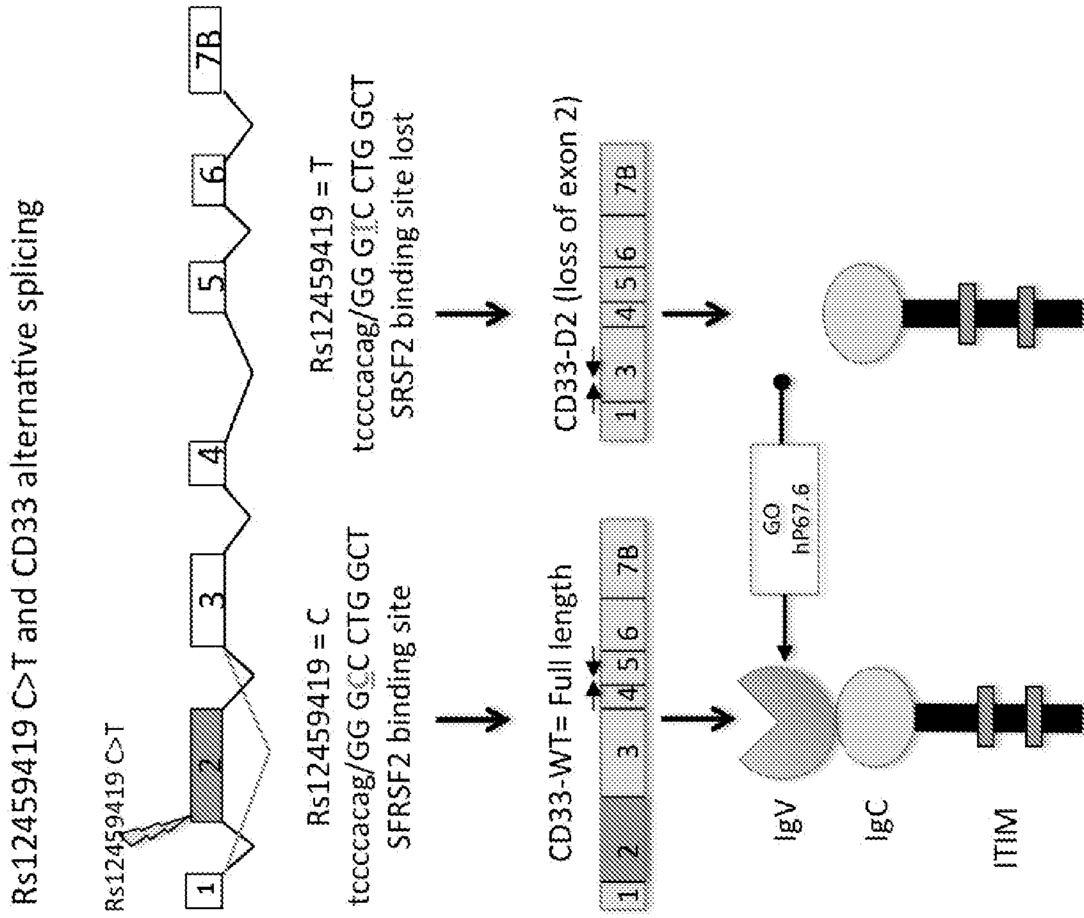


Figure 6A
Rs12459419 C>T and CD33 alternative splicing
Rs12459419 C>T
1
2
3
4
5
6
7B
Rs12459419 = C
tccccacag/GG G C CTG GCT
SFRSF2 binding site
Rs12459419 = T
tccccacag/GG G C CTG GCT
SRSF2 binding site lost
CD33-WT= Full length
1
2
3
4
5
6
7B
CD33-D2 (loss of exon 2)
1
3
4
5
6
7B
GO
hP67.6
IgV
IgC
ITIM

Figure 6B

CD33-D2 spliced isoform and rs12459419 : RT-PCR using isoform specific primers

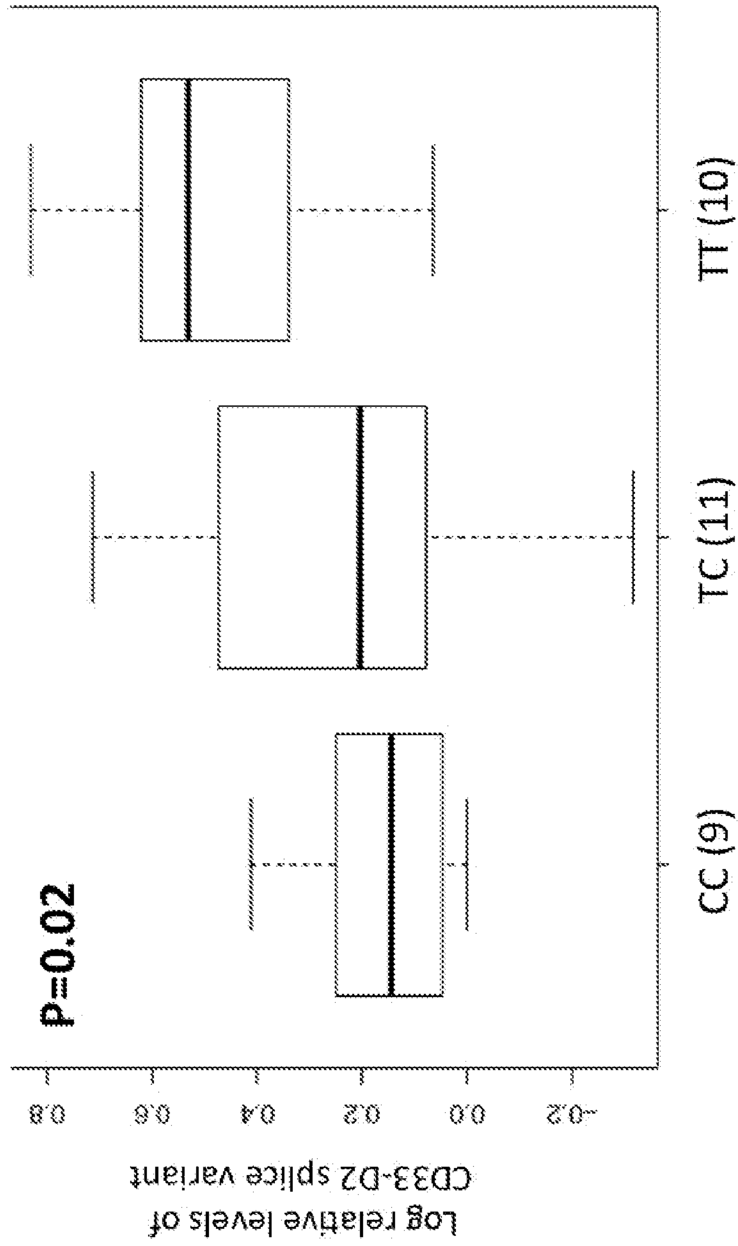


Figure 6C

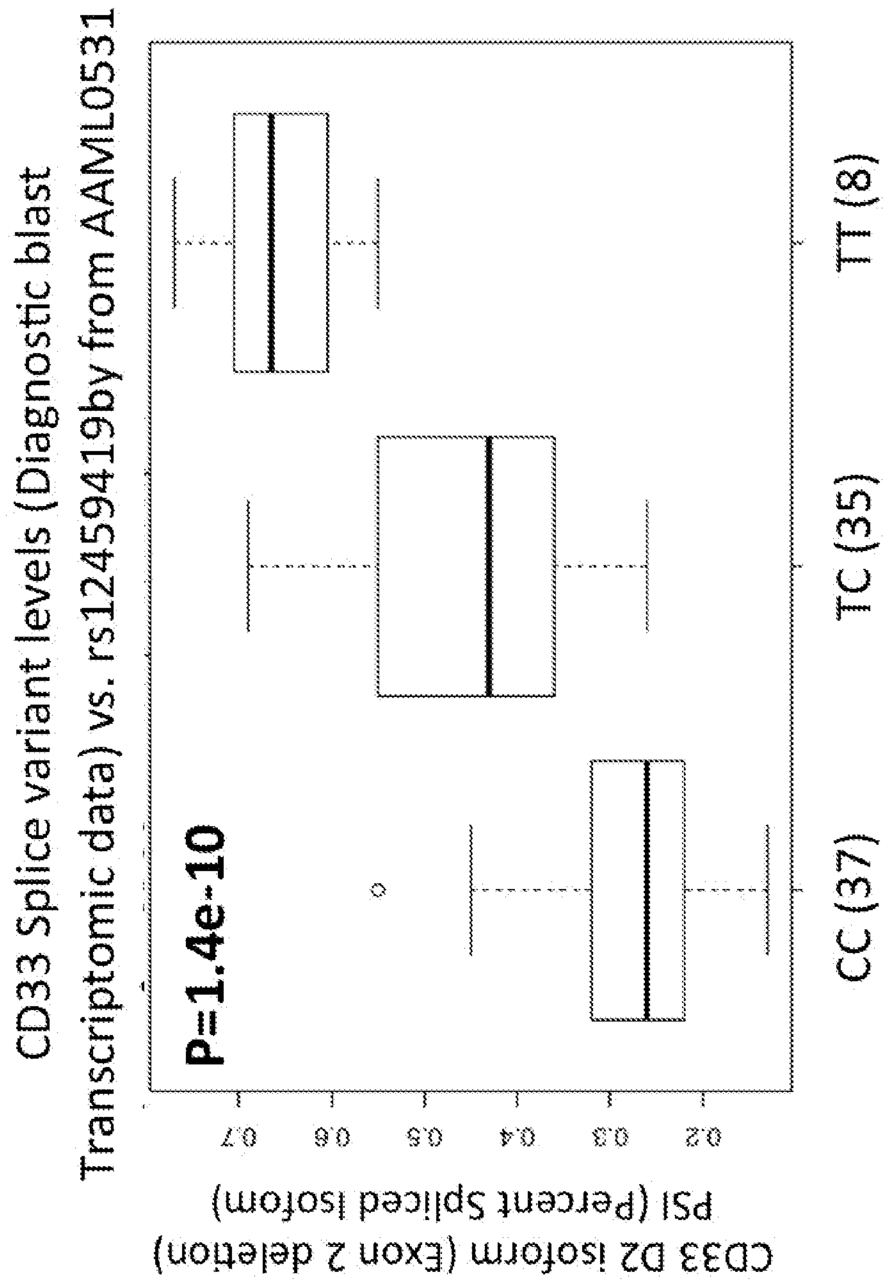


Figure 7

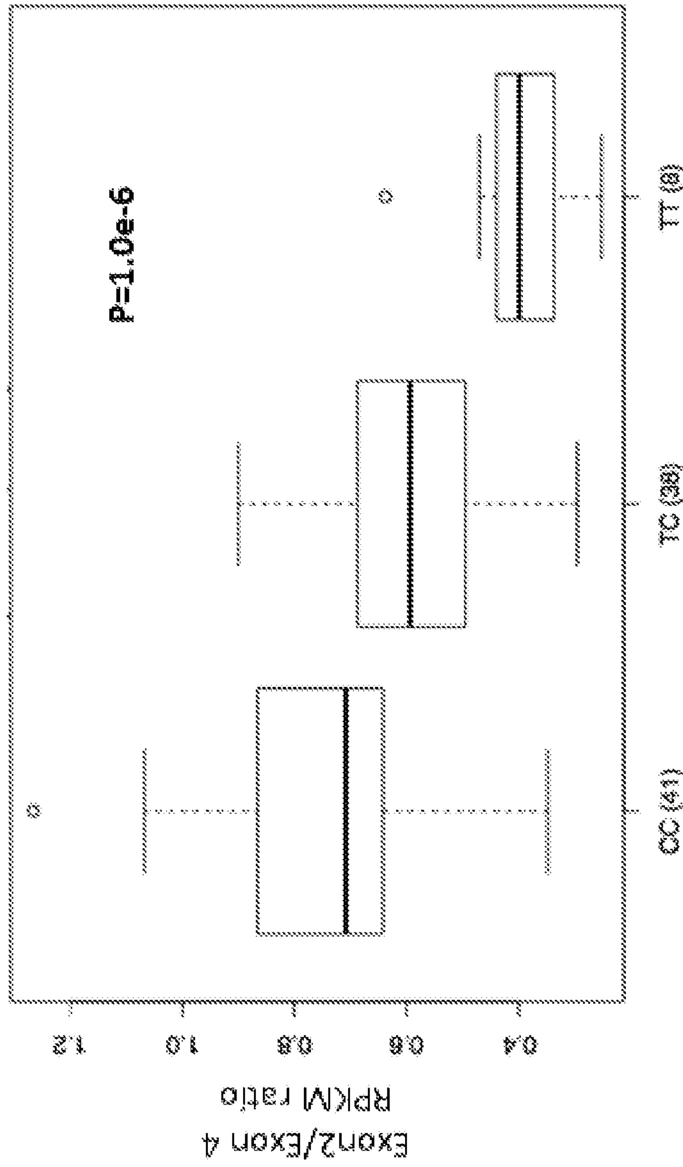


Figure 8A

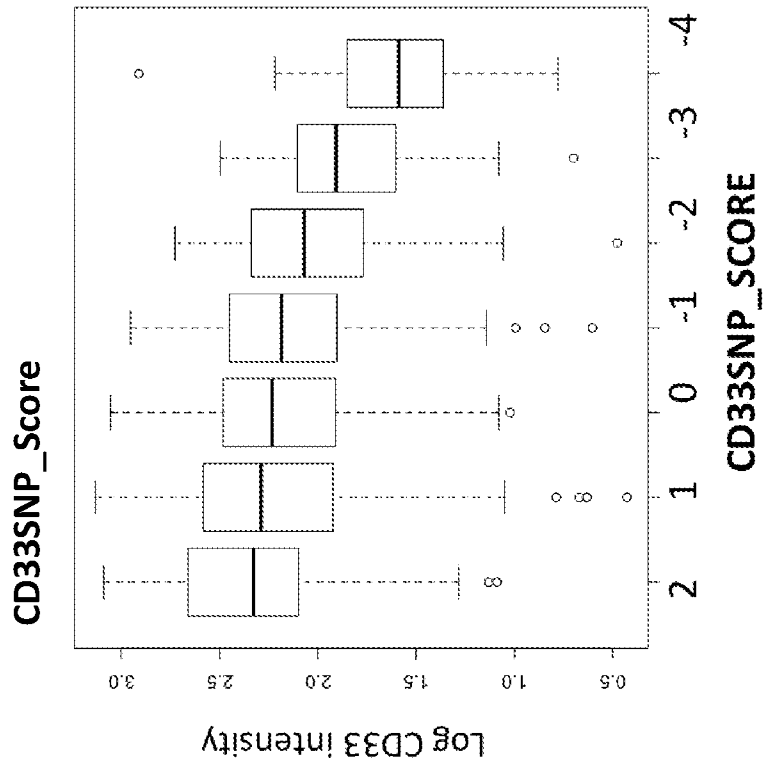
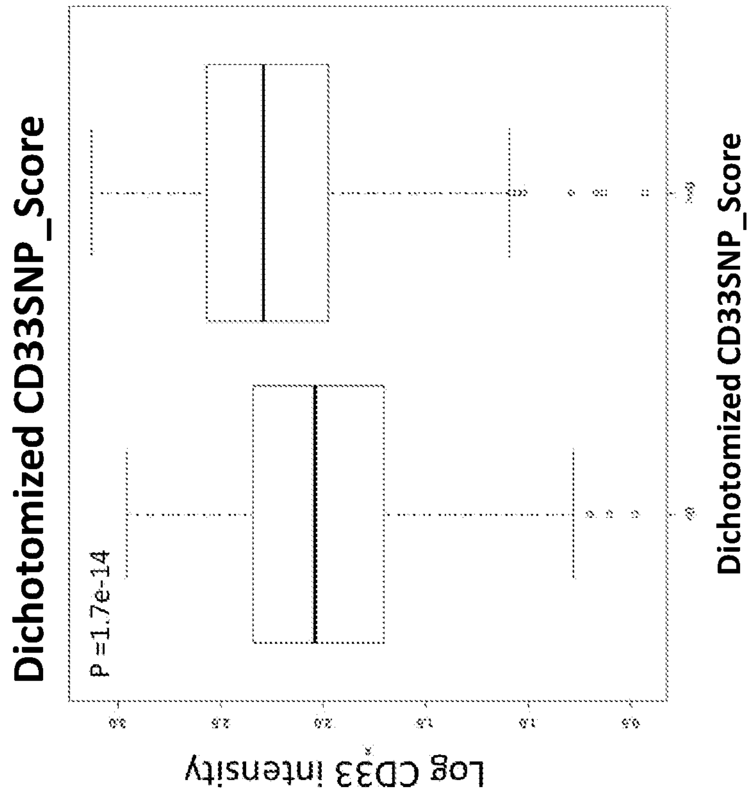


Figure 8B



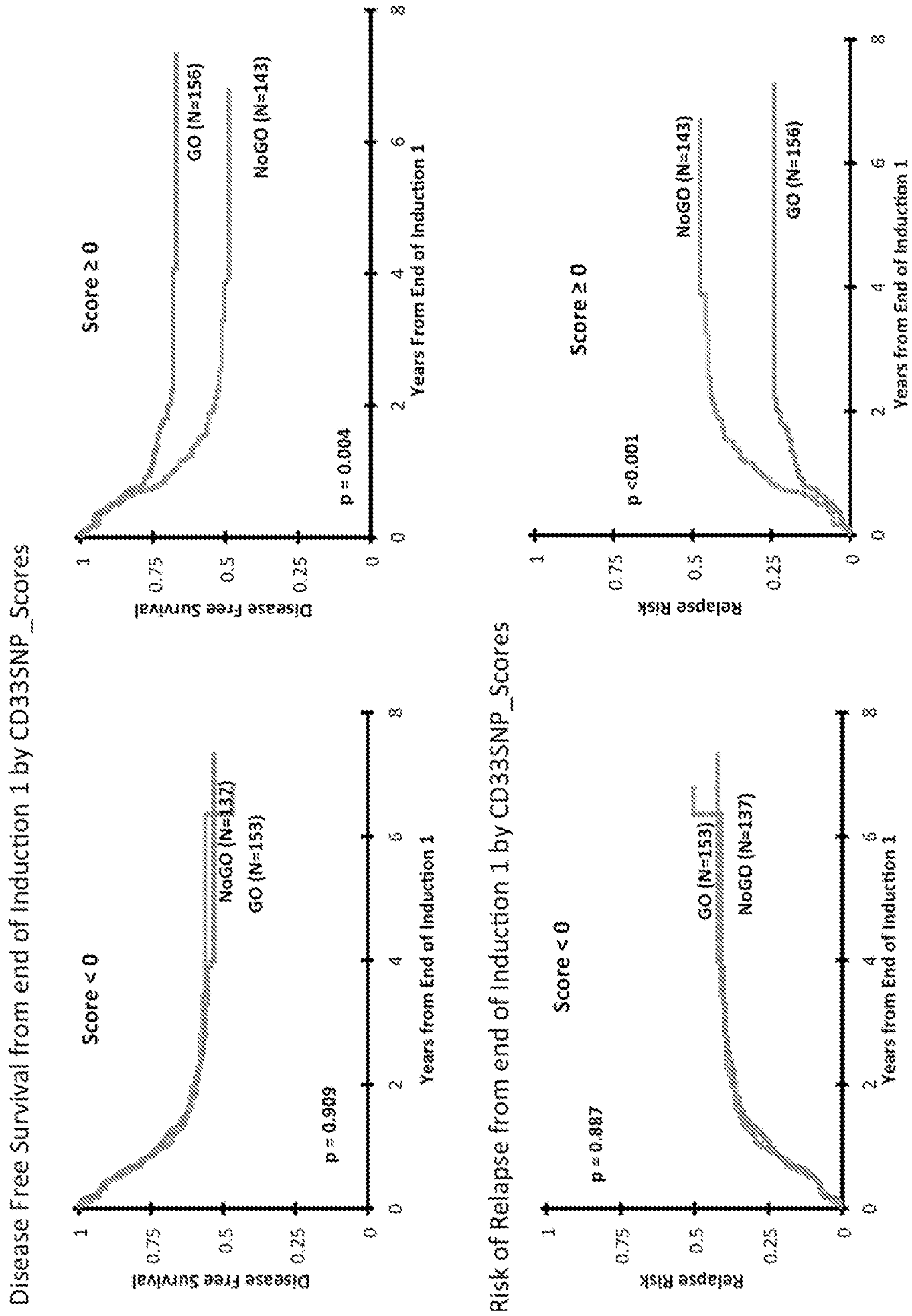


Figure 9
Disease Free Survival from end of Induction 1 by CD33SNP_Scores
Score < 0
p = 0.909
NoGO (N=137)
GO (N=153)
Disease Free Survival
Years from End of Induction 1
Score ≥ 0
p = 0.004
GO (N=156)
NoGO (N=143)
Disease Free Survival
Years From End of Induction 1
Risk of Relapse from end of Induction 1 by CD33SNP_Scores
Score < 0
p = 0.887
GO (N=153)
NoGO (N=137)
Relapse Risk
Years from End of Induction 1
Score ≥ 0
p <0.001
NoGO (N=143)
GO (N=156)
Relapse Risk
Years from End of Induction 1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/026369

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 39/395; A61P 35/00; C12Q 1/68 (2017.01)

CPC - C07K 16/2803; C12Q 1/6883; C12Q 2600/156 (2017.02)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

USPC - 435/343; 435/7.23; 514/908 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	MORTLAND et al. "Clinical Significance of CD33 Non-Synonymous Single Nucleotide Polymorphisms (SNPs) in Pediatric Patients with Acute Myeloid Leukemia Treated with Gemtuzumab-Ozogamicin-Containing Chemotherapy", Clin Cancer Res, 15 March 2013 (15.03.2013), Vol. 19, Pgs. 1620-1627. entire document	1-4, 6, 15, 16, 18, 19, 26, 32, 33, 40 -------------------- 5, 17, 27-31, 34, 35, 41, 42, 77, 79-83, 91-94, 101-110
X -- Y	US 9,066,928 B1 (ESTUS et al) 30 June 2015 (30.06.2015) entire document	43-46, 48, 115-118, 120 -------------------- 47, 77, 79-83, 91-94, 101-110, 119
Y	WALKER et al. "Association of CD33 Polymorphism rs3865444 with Alzheimer's Disease Pathology and CD33 Expression in Human Cerebral Cortex," Neurobiol Aging, 02 October 2014 (02.10.2014) Vol. 36, Pgs. 571-582. entire document	5, 17, 47, 82, 93, 119
Y	WO 2014/074942 A1 (ILLUMINA, INC.) 15 May 2014 (15.05.2014) entire document	27-31, 34, 35, 41, 42, 102-106, 109, 110
A	LOURDUSAMY et al. "Identification of cis-regulatory variation influencing protein abundance levels in human plasma," Human Molecular Genetics, 16 May 2012 (16.05.2012), Vol. 21, Pgs. 3719-3726. entire document	1-6, 15-19, 26-35, 40-48, 51-67, 70-74, 77-83, 91-94, 101-110, 115-120

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

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 June 2017	20 JUL 2017

Name and mailing address of the ISA/US	Authorized officer
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/026369

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

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

   a. ☐ forming part of the international application as filed:
      - ☐ in the form of an Annex C/ST.25 text file.
      - ☐ on paper or in the form of an image file.

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

   c. ☒ furnished subsequent to the international filing date for the purposes of international search only:
      - ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
      - ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/026369

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

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

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

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

3. ☒ Claims Nos.: 7-14, 20-25, 36-39, 49, 50, 68, 69, 75, 76, 84-90, 95-100, 111-114, 121, 122
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

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

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

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

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

Remark on Protest

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

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

☐ No protest accompanied the payment of additional search fees.

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SHERVA et al. "Genome Wide Association Study of the Rate of Cognitive Decline in Alzheimer's Disease," Alzheimers Dement, 25 March 2013 (25.03.2013), Vol. 10, Pgs. 45-52. entire document	1-6, 15-19, 26-35, 40-48, 51-67, 70-74, 77-83, 91-94, 101-110, 115-120
A	SLEEGERS et al. "A 22-single nucleotide polymorphism Alzheimer's disease risk score correlates with family history, onset age, and cerebrospinal fluid Ab42," Alzheimers Dement, 15 June 2015 (15.06.2015), Vol. 11, Pgs. 1452-1460. entire document	1-6, 15-19, 26-35, 40-48, 51-67, 70-74, 77-83, 91-94, 101-110, 115-120