SUPPLEMENTARY METHODS

Parental and engineered human acute leukemia cell lines. Human myeloid ML-1 cells were grown in Iscove's Modified Dulbecco's Medium (IMDM), OCI-AML3 cells were maintained in Minimal Essential Medium (MEM) alpha, and U937, THP-1, Kasumi and KG-1a were cultured in Roswell Park Memorial Institute (RPMI)-1640 medium, all with 10% fetal bovine serum (FBS, Hyclone, ThermoFisher Scientific, Waltham, MA, USA). NB4 cells were grown in RPMI-1640 + 5% bovine calf serum (BCS), HL-60 in RPMI-1640 + 10% BCS, TF-1 in RPMI + 10% BCS + 5 ng/mL granulocyte-macrophage colony-stimulating factor (GM-CSF, PeproTech, Rocky Hill, NJ, USA), and KG-1 and ME-1 cells in RPMI-1640 + 20% FBS. Human lymphoid REH and RCH-ACV cells were maintained in RPMI-1640 with 10% FBS while RS4;11 cells were maintained in MEM alpha with 10% FBS. All cultured cells were grown in the presence of penicillin and streptomycin (ThermoFisher Scientific, Waltham, MA, USA) and were routinely tested for mycoplasma contamination using the MycoAlert™ Mycoplasma Detection Kit (Lonza, Basel, Switzerland). Sublines of cells overexpressing human $CD33^{FL}$ or human $CD33^{\Delta E2}$ were generated through transduction with pRRLsin.cPPT.MSCV lentivirus constructs containing an IRES-Enhanced Green Fluorescent Protein (EGFP) cassette at a multiplicity of infection (MOI) of 0.25-25 as previously described.[1-4] EGFP-positive cells were isolated by flow cytometry and re-cultured for further analysis. All cell lines were confirmed to be mycoplasma-negative and authenticated using standard STR CODIS typing.

Genetic deletion of CD33 in AML cell lines. Clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9-editing was carried out as described before[5] by electroporation of purified Cas9 protein (TrueCut Cas9 V2; ThermoFisher Scientific) complexed with synthetic guide RNA (sgRNA) targeting exon 1 of CD33 (sequence 5'- CTGCTGCCCCTGCTGTGGGC-3'). sgRNA was modified at the 5′ and 3′ ends with 2′-*O*-methyl-3′-phosphorothiate (Synthego, Redwood City, CA) and cell lines electroporated using the ECM 380 Square Wave Electroporation system (Harvard Apparatus, Cambridge, MA). Subclones with complete knockout of CD33 were identified by flow cytometry (see below) and CD33-negative single cells isolated via fluorescence activated cell sorting using a BD FACSAria II (BD Biosciences Pharmingen, San Diego, CA, USA). Genomic DNA from individual clones were analyzed by PCR amplification of exon 1 and Sanger DNA sequence analysis to confirm disruption and frame-shift mutation at both CD33 loci. The ML-1 clone E subline harboring a CRISPR/Cas9 deletion of CD33 exon-2 has been described previously.[5]

Primary AML patient and healthy donor biospecimens. For initial biochemical characterization studies, frozen aliquots of Ficoll-isolated mononuclear cells from peripheral blood or bone marrow specimens from 3 adults with AML and 4 healthy donors were obtained from institutional cell repositories. Subjects provided written informed consent for the collection and use of their specimens for research purposes under protocols approved by the Fred Hutchinson Cancer Research Center Institutional Review Board (IRB). For subsequent immunophenotypic characterization studies, fresh left-over peripheral blood and/or bone marrow material was used under IRB-approved protocols from specimens from AML patients submitted for analysis to the University of Washington/Seattle Cancer Care Alliance Hematopathology Laboratory. Clinical data were de-identified in compliance with the Health Insurance Portability and Accountability Act.

Generation of murine antibodies against human CD33$^{\Delta E2}$. Peptide immunogens consisting of the mouse Fc domain and either the entire extracellular domain of human CD33^{FL} or the entire extracellular domain of human CD33$^{\Delta E2}$ were generated, expressed in Freestyle^{TM} 293-F cells, and purified using nickel IMAC resin on an ÄKTA FPLC system as previously described.[6] Both immunogens were characterized by SDS-PAGE and size-exclusion chromatography (SEC) and confirmed to be of very high biochemical quality. BALB/c, CD1 and F1 mice were immunized with a mixture of both immunogens. Resulting hybridomas were screened by flow cytometry with beads coupled to the peptide immunogens as well as parental human lymphoid cells and sublines overexpressing human CD33^{FL} or human CD33$^{\Delta E2}$. Hybridomas showing reactivity with CD33$^{\Delta E2}$ but not CD33^{FL} were selected and subcloned. Antibodies were then isotyped and variable regions sequenced following 5' RACE cloning with isotype-specific PCR primers.

Expression and purification of recombinant murine anti-human CD33$^{\Delta E2}$ antibodies. Protein sequences were reverse-translated using human codons and cloned into a modified pCVL lentiviral vector as described.[6] Lentivirus was produced by transient co-transfection with psPAX2 and pMD2.G of 293T cells and used to transduce FreestyleTM 293-F cells. Cultures were expanded to 2 liters and supernatants were harvested by centrifugation. Secreted protein was extracted from the conditioned media using immobilized metal-affinity chromatography (Ni-NTA) and subsequently polished using size exclusion chromatography on an ÄKTA pure instrument (GE Healthcare Life Sciences, Pittsburgh, PA, USA) equipped with a Superdex 200 Increase 10/300GL running at 0.75 mL/min in PBS. Fractions corresponding to the monomeric

proteins were pooled, quantitated using a Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) and further analyzed by SDS-PAGE under non-reducing and reducing conditions.

CD33 single nucleotide polymorphism (SNP) rs12459419 genotyping. CD33 rs12459419 genotyping was done as previously described.[7] Specifically, gDNA was purified using Qiagen AllPrep DNA/RNA mini kits (Qiagen, Valencia, CA) and amplified with the following primers: For 5′-CTGGAAGCTGCTTCCTCAGACATG-3′ and Rev 5′-GAACCAGTAACCATGAACTGGGGAG-3′. The resulting 266 bp amplicon was digested overnight with HaeIII (New England Biolabs, Ipswich, MA, USA) and separated on a 2% agarose gel.

Quantification of CD33 expression. Expression of CD33 variants on human leukemia cell lines and primary AML cells was quantified by flow cytometry either using a directly labeled CD33 antibody (clone P67.6; BD Biosciences)[1,2] or unlabeled 13E11 or 11D5 followed by APC-conjugated goat anti-mouse Ig (Multiple Adsorption, (ThermoFisher Scientific). Fluorochrome labeling of anti-CD33 antibodies was performed using a DyLight 650 Conjugation kit (Abcam, Cambridge, UK). To identify non-viable cells, samples were stained with 4',6-diamidino-2-phenylindole (DAPI). 10,000 events were acquired on a BD FACSCanto II flow cytometer (BD Biosciences), and DAPI-negative cells analyzed using FlowJo (Tree Star, Ashland, OR).

Multiparameter flow cytometric analysis of $CD33^{FL}$ and $CD33^{\Delta E2}$ expression in clinical AML patient specimens. Flow cytometry was performed on 21 residual clinical samples positive for AML or myelodysplastic syndrome (MDS) using the anti-CD33v PE-labeled antibody in a backbone of CD15 FITC, CD19 PE-CF594, CD117 PE-Cy5, CD13 PE-Cy7, HLA-DR PB, CD34 APC, CD71 APC-A700 and CD45 APC-H7. Positivity was assessed using a PE fluorescence minus one (FMO) control for background and signal obtained on the gated CD34+ progenitor population compared with that obtained on the mature lymphocytes as an internal control.

Immunoprecipitation and western blotting. 2 x 10^7 cells were washed twice in ice-cold Dulbecco's Phosphate Buffered Saline (PBS; ThermoFisher Scientific) and then lysed in 1% Triton-X (TX), consisting of 10 mM HEPES, 150 mM NaCl, 2mM EDTA, 50 mM NaF, 100nM sodium orthovanadate with protease inhibitor cocktail (Roche Diagnostics GmbH, Mannheim, Germany). After incubation on ice for 45 min, lysates were precleared by centrifugation. For immunoprecipitations, precleared lysates were incubated with 1 µg of antibody and Protein A/G

beads (Santa Cruz Biotechnology, Dallas, TX) overnight at 4°C. Beads were then washed 2 times in ice-cold PBS and boiled in 1x Laemmli sample buffer (BioRad, Hercules, CA). Samples were then separated for western blotting using min protean TGX sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-10%-PAGE; BioRad), transferred to nitrocellulose, placed in blocking solution consisting of 5% non-fat dry milk in Tris-buffered saline (TBS; 50 mM Tris [pH 7.5], 150 mM NaCl) containing 0.1% Tween 20 (TBS-T) for 30 minutes at room temperature, and then incubated with primary antibodies overnight in blocking solution at 4°C. After washing, blots were incubated for 180 minutes at RT with horseradish peroxidase-conjugated sheep anti-mouse secondary antibody in blocking solution, washed, and immunoreactive signals visualized with enhanced chemiluminescence (ECL) Plus western blotting detection system (GE Healthcare, Amersham Biosciences, Buckinghamshire, England) and Blue Devil Autoradiography film (Genesee Scientific, El Cajon, CA). Mouse monoclonal CD33 antibodies against the c-terminus of CD33, B-9 and E-6 (sc-374450, and sc-376184, respectively, SantaCruz Biotechnology, Dallas, TX) were used for immunoprecipitation and CD33 western blot detection.

Immunofluorescence microscopy. 30,000 cells were washed in PBS, resuspended in PBS with 1% Bovine Serum Albumin (BSA; Sigma, Massachusetts, USA), adhered to 1 mm thick White Epoxy Ink frosted glass slides (ThermoFisher Scientific) using a Shandon Cytospin 3 (Thermo Shandon, Cheshire, UK). Cells were fixed with 4% PFA (Paraformaldehyde, Electron Microscopy Sciences, Pennsylvania, USA) for 20 minutes at 4°C. Cells were washed three times in PBS for 5 minutes. Cells were permeabilized with 0.1% Triton X-100 (LabChem, Pennsylvania, USA) in PBS at room temperature for 5 minutes. Cells were washed three times with PBS for 5 minutes. Cells were blocked in 2% BSA/5% Normal Goat Serum (Jackson ImmunoResearch, Pennsylvania, USA) in PBS for 30 minutes. 5 µg/ml of primary antibody was added to cells and incubated overnight at 4°C. After 24 hours, slides were washed 3 times with PBS. 5 µg/ml goat anti-mouse IgG1 Alexa Fluor 568 (ThermoFisher Scientific) was applied to cells for 45 min at room temperature. Secondary antibody was aspirated off cells, and DAPI solution was added for 10 minutes. Cells were washed three times in PBS. ProLong Gold Antifade Mountant (ThermoFisher Scientific) was added to cells, 12 mm thick Platinum Line coverslip (Electron Microscopy Sciences, Pennsylvania, USA) was applied to slide, and allowed to dry overnight before imaging. Cells were examined on a Nikon E800 microscope (Nikon, New York, USA) using a Zeiss Axiocam Mrm fluorescent camera (Zeiss, Oberkochen, Germany) and

Nikon 20x Plan Apo objective (Nikon, New York, USA). Images were analyzed using ImageJ software (National Institutes of Health, Maryland, USA).

Permeabilization assay. Cells were washed in ice-cold PBS, and permeabilized with a Transcription Factor Buffer Set (BD Biosciences) according to manufacturer's instructions. Cell surface and intracellular staining were performed with fluorochrome-labeled 11D5 and FITC-labeled anti-human myeloperoxidase (MPO) antibody (BioLegend, San Diego, CA, USA) as a positive control. 10,000 events per sample were acquired on a BD FACSymphony flow cytometer (BD Biosciences) and analyzed with FlowJo.

Signal peptide prediction and mass spectrometry. Publicly available signal peptide prediction algorithms, SignalP 4.1, Phobius, and WoLF PSORT were used to predict CD33 protein cleavage sites. For mass spectrometry, purified peptide immunogen was diluted with 100mM ammonium bicarbonate buffer/8M urea and reduced with 10mM dithiothreitol (DTT) at 56ºC for 45 minutes. The solution was removed from heat and alkylated (in the dark) using iodoacetamide 30 minutes at room temperature. The urea concentration was then reduced to 1M with further addition of 100mM ammonium bicarbonate buffer and trypsin was added to the sample in a 1:25 enzyme:substrate ratio. The sample was digested overnight at 37ºC. The peptide mixture was removed from heat the following morning and the sample was acidified using TFA to a final concentration of 0.1%. The peptides were desalted on a C18 ZipTip (Millipore, Burlington, MA) using a protocol based on the manufacturer's recommendation and eluted from the tip in 60% acetonitrile:40% water (0.1% TFA). Peptides were analyzed by Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC-ESI MS/MS) with a Thermo Scientific Easy-nLC II nano HPLC system coupled to a hybrid Orbitrap Elite (both ThermoFisher Scientific) mass spectrometer. The peptide sample was reconstituted in a 20µL aliquot of water (0.1% formic acid):acetonitrile (98:2 volume:volume) and a 2µL injection was loaded onto a reversed-phase trap column (100 µm × 20 mm) packed with Magic C_{18}AQ (5-µm 200Å resin; Michrom Bioresources, Auburn, CA). The sample was eluted off of a reversed-phase column (75 µm × 250 mm) packed with Magic C_{18}AQ (5-µm 100Å resin; Michrom Bioresources) using a 60-minute gradient from 5% to 35% acetonitrile in 0.1% formic acid at a flow rate of 400 nL/minute. The heated capillary temperature was set to 300 C and a spray voltage of 2750 V was applied to the electrospray tip. The mass spectrometer was operated in the data-dependent mode, switching automatically between MS survey scans in the Orbitrap (using an AGC target value of 1,000,000, resolution of 120,000, and injection time of 250

milliseconds) with MS/MS spectra acquisition in the dual linear ion trap. The top 20 most intense ions from the Fourier-transform (FT) full scan were selected for fragmentation by collision induced dissociation (CID) with a normalized collision energy of 35%. Following fragmentation, the selected ions were placed on the dynamic exclusion list for 10 seconds with a list size of 500 and exclusion mass width of +/- 10 ppm. X!TANDEM 2013.06.15 software was used to perform spectra search. Precursor mass tolerances were -1.5 to 2.5 Daltons with variable modification on Methionine and static modification on Cysteine. Tryptic enzyme restriction was used with up to two mis-cleavage. E-value 0.005 cut off was used to identify matching sequences.

RNA analysis of CD33 isoforms. RNA analysis was performed as described.[4] Briefly, approximately 25 ng of RNA was reverse transcribed for each PCR reaction and cDNA amplified with primers specific for $CD33^{FL}$ and $CD33^{\Delta E2}$. Amplicons were electrophoretically separated using 1x Tris/acetic acid/EDTA (TAE) buffered 0.9% agarose gel electrophoresis and visualized with ethidium bromide (AlphaImager HP; ProteinSimple, San Jose, CA, USA).

References

1. Walter RB, Raden BW, Kamikura DM, Cooper JA, Bernstein ID. Influence of CD33 expression levels and ITIM-dependent internalization on gemtuzumab ozogamicin-induced cytotoxicity. *Blood* 2005; **105**(3)**:** 1295-1302.

2. Laszlo GS, Gudgeon CJ, Harrington KH, Dell'Aringa J, Newhall KJ, Means GD*, et al.* Cellular determinants for preclinical activity of a novel CD33/CD3 bispecific T-cell engager (BiTE) antibody, AMG 330, against human AML. *Blood* 2014; **123**(4)**:** 554-561.

3. Laszlo GS, Gudgeon CJ, Harrington KH, Walter RB. T-cell ligands modulate the cytolytic activity of the CD33/CD3 BiTE antibody construct, AMG 330. *Blood Cancer J* 2015; **5:** e340.

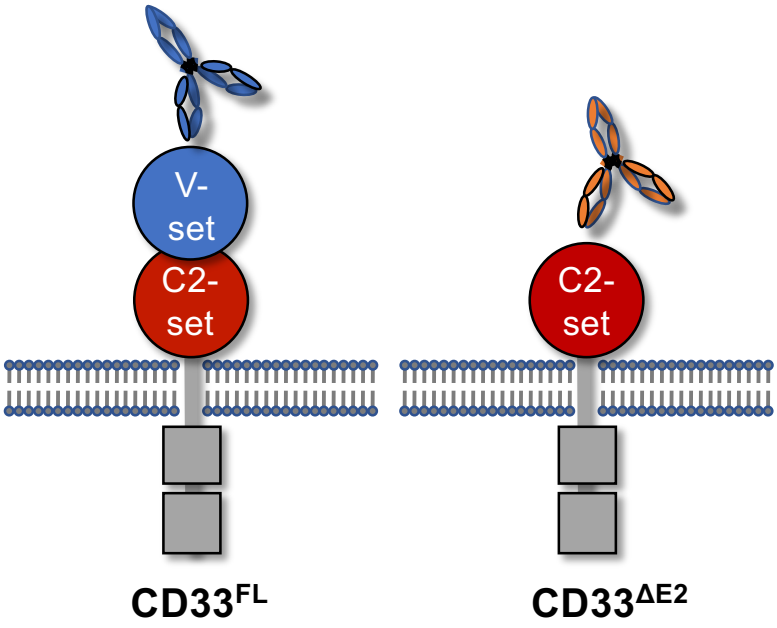
4. Laszlo GS, Harrington KH, Gudgeon CJ, Beddoe ME, Fitzgibbon MP, Ries RE*, et al.* Expression and functional characterization of CD33 transcript variants in human acute myeloid leukemia. *Oncotarget* 2016; **7**(28)**:** 43281-43294.

5. Humbert O, Laszlo GS, Sichel S, Ironside C, Haworth KG, Bates OM*, et al.* Engineering resistance to CD33-targeted immunotherapy in normal hematopoiesis by CRISPR/Cas9-deletion of CD33 exon 2. *Leukemia* 2019; **33**(3)**:** 762-808.

6. Bandaranayake AD, Correnti C, Ryu BY, Brault M, Strong RK, Rawlings DJ. Daedalus: a robust, turnkey platform for rapid production of decigram quantities of active recombinant proteins in human cell lines using novel lentiviral vectors. *Nucleic Acids Res* 2011; **39**(21)**:** e143.

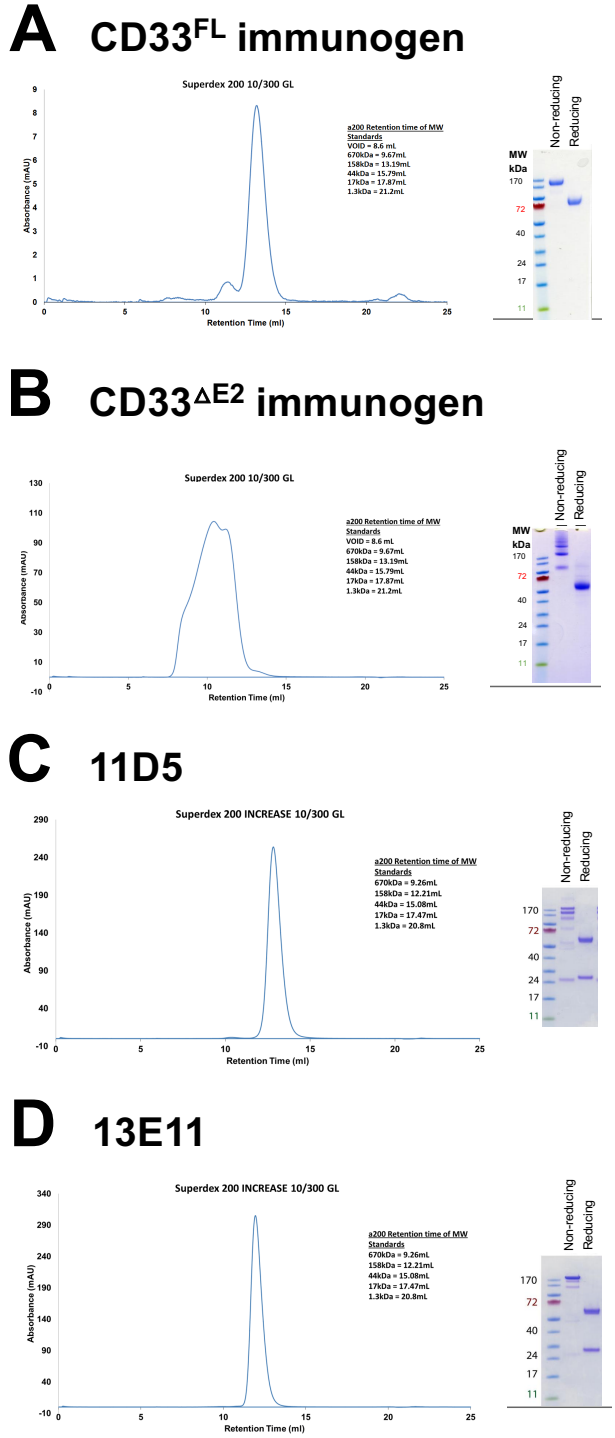
7. Laszlo GS, Beddoe ME, Godwin CD, Bates OM, Gudgeon CJ, Harrington KH*, et al.* Relationship between CD33 expression, splicing polymorphism, and in vitro cytotoxicity of gemtuzumab ozogamicin and the CD33/CD3 BiTE(R) AMG 330. *Haematologica* 2019; **104**(2)**:** e59-e62.

SUPPLEMENTARY FIGURE 1



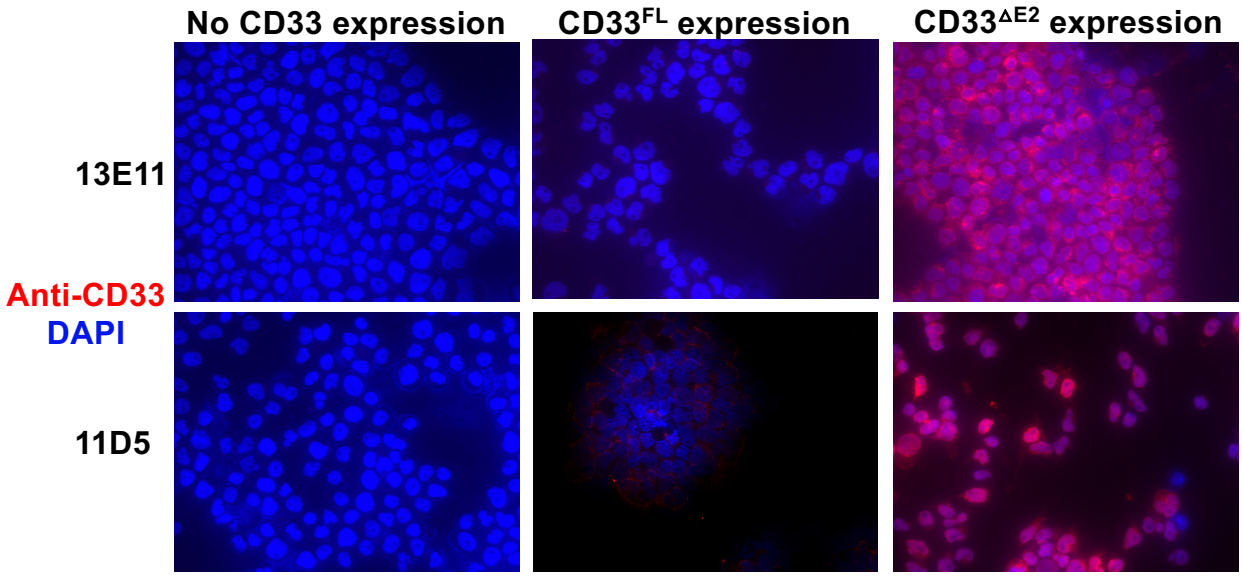
Schematic of diverse antibody binding specificities to CD33 isoforms. Current antibody-based therapies target the V-set domain of CD33 (**blue antibody**), which is present in the full length CD33 isoform but not CD33$^{\Delta E2}$. Antibodies binding to the C2-set domain only when the V-set domain is not present (**orange antibody**) are CD33$^{\Delta E2}$-specific antibodies.

SUPPLEMENTARY FIGURE 2



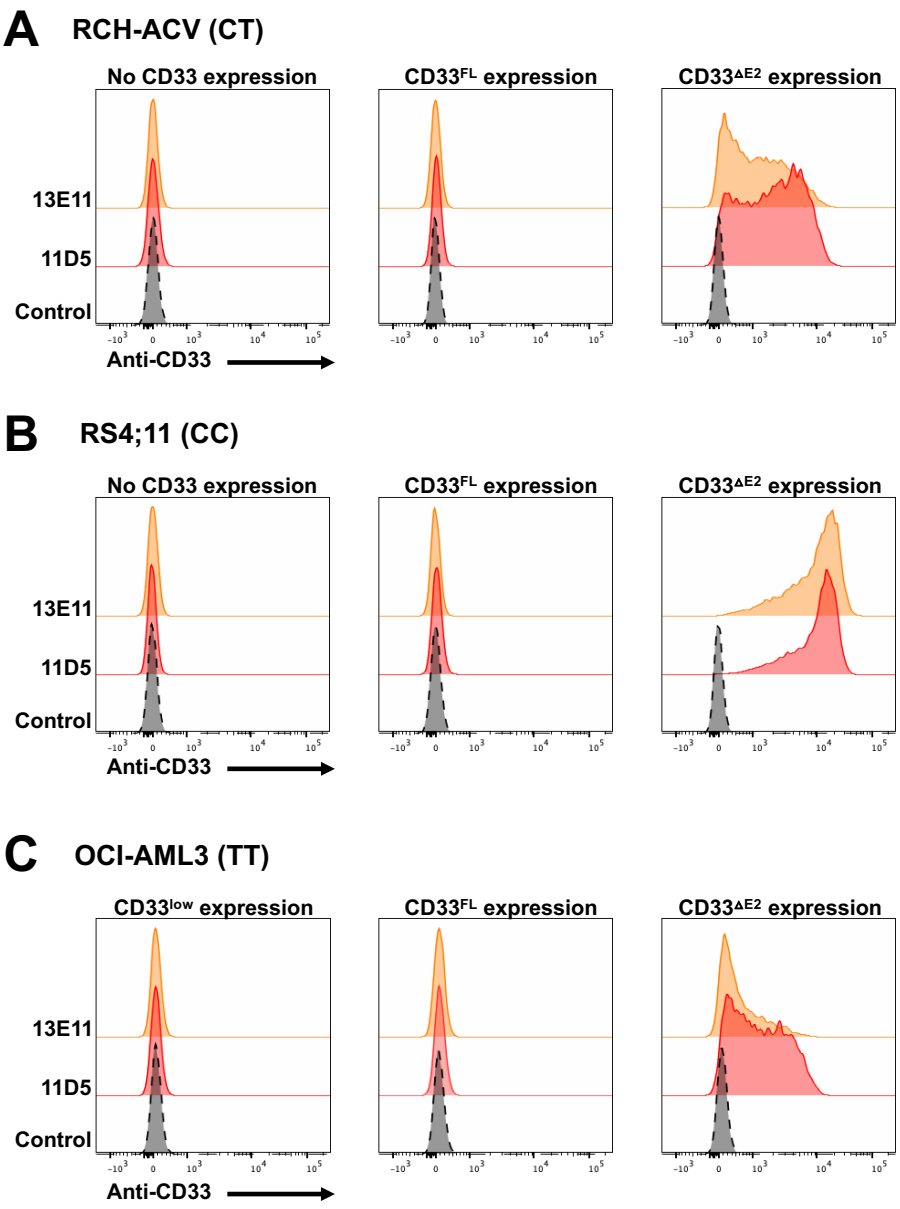
Quality control of CD33 immunogens and anti- CD33$^{\Delta E2}$ monoclonal antibodies. Size exclusion chromatography and SDS-PAGE analysis (under non-reducing and reducing conditions) of **(A)** CD33^{FL} peptide immunogen, **(B)** CD33$^{\Delta E2}$ peptide immunogen, **(C)** 11D5 monoclonal antibody, and **(D)** 13E11 monoclonal antibody.

SUPPLEMENTARY FIGURE 3



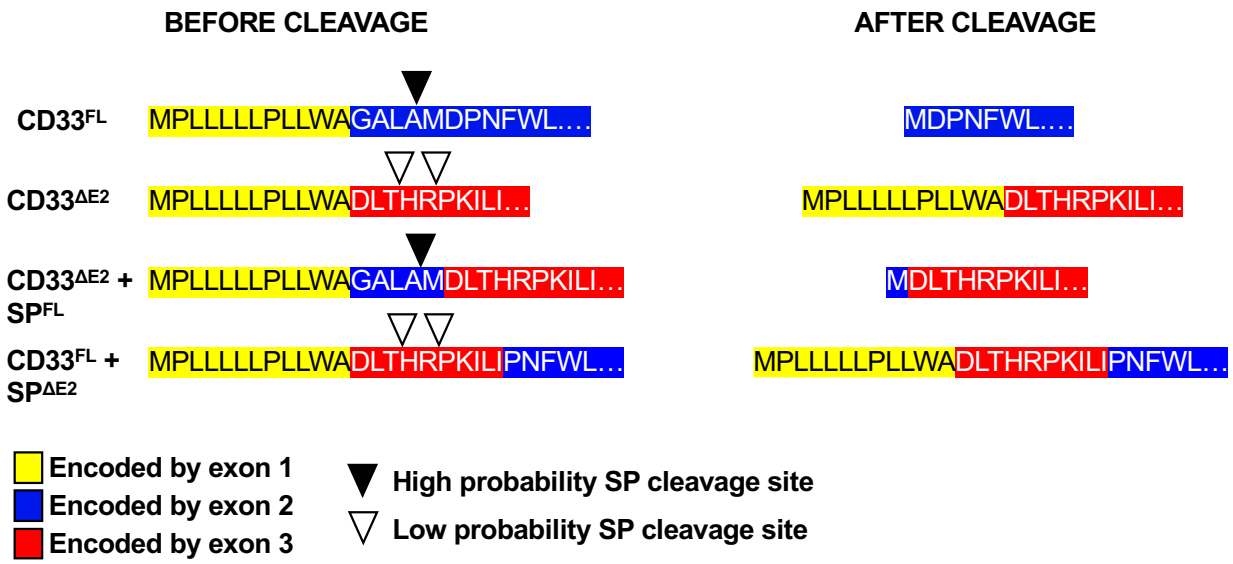
Immunofluorescence microscopy studies of CD33-engineered cell line with 11D5 and 13E11. Anti-$CD33^{\Delta E2}$ antibody clones 11D5 and 13E11 were tested against parental REH cells (human acute lymphoblastic leukemia cells, endogenously $CD33^{neg}$) and sublines engineered to overexpress $CD33^{FL}$ or $CD33^{\Delta E2}$. Data are representative of three separate experiments. $CD33^{FL}$, full length CD33 isoform; $CD33^{\Delta E2}$, CD33 isoform arising from transcript missing exon 2.

SUPPLEMENTARY FIGURE 4



Validation of binding specificity of anti-CD33$^{\Delta E2}$ antibody clones 11D5 and 13E11. Anti-CD33$^{\Delta E2}$ antibody clones 11D5 and 13E11 were tested flow cytometrically against parental RCH-ACV **(A),** RS4;11 cells (**B**; both human acute lymphoblastic cell lines, endogenously CD33^{neg}) as well as OCI-AML3 cells (**C**, human acute myeloid cell line, endogenously CD33^{low}) and sublines engineered to overexpress CD33^{FL} and CD33$^{\Delta E2}$. No primary antibody control was included. rs12459419 genotype is shown in parentheses. CD33^{FL}, full length CD33 isoform; CD33$^{\Delta E2}$, CD33 isoform arising from transcript missing exon 2.

SUPPLEMENTARY FIGURE 5



Signal peptides and putative signal peptide cleavage sites. Schematic of N-terminal amino acids, putative signal peptide cleavage sites, and anticipated amino acid sequences following signal peptide cleavage in endogenous ($CD33^{FL}$, $CD33^{\Delta E2}$) and artificial ($CD33^{\Delta E2}$ + SP^{FL}, $CD33^{FL}$ + $SP^{\Delta E2}$) CD33 isoforms.

SUPPLEMENTARY FIGURE 6

Amino terminal peptides identified by MS

MPLLLLLPLLWADLTHRPK (E-value: 2E-3)

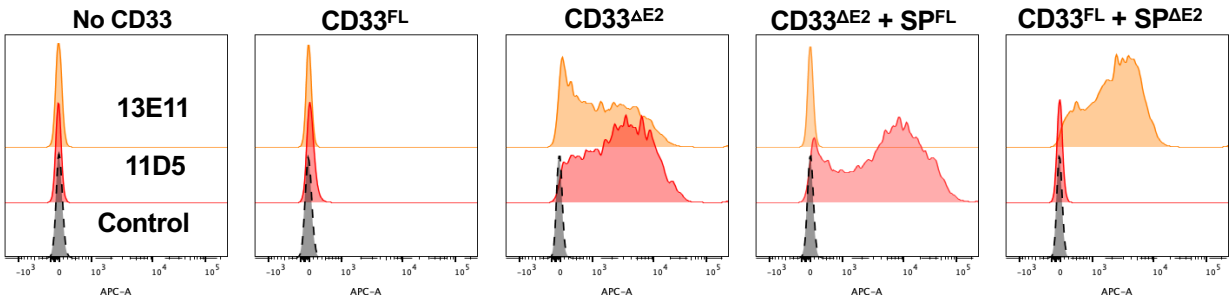
M^{oxy}PLLLLLPLLWADLTHRPK (E-value: 2E-3)

LLLLLPLLWADLTHRPK (E-value: 4.9E-4)

PLLLLLPLLWADLTHRPK (E-value: 1.2E-4)

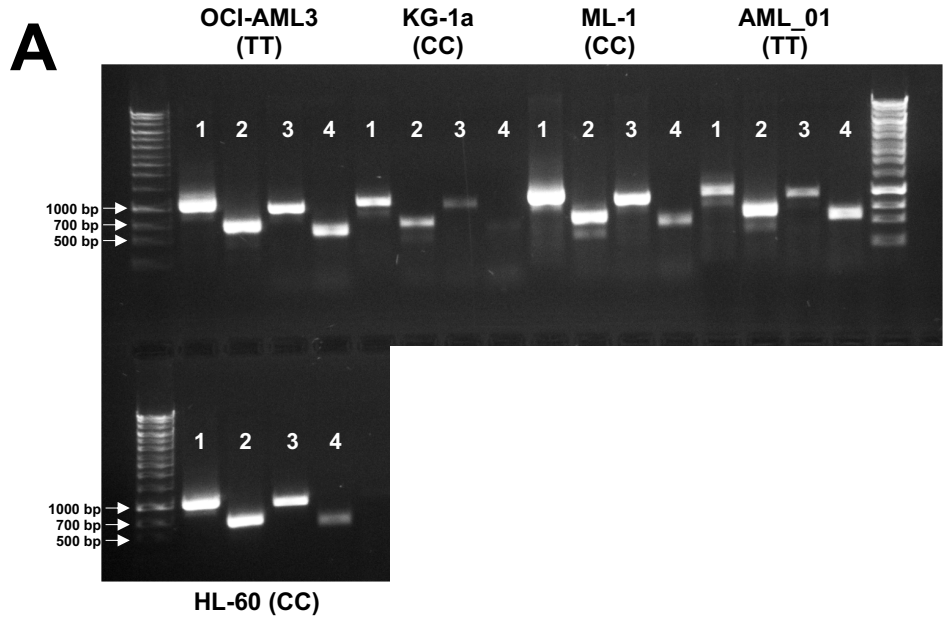
CD33$^{\Delta E2}$ immunogen retains exon 1 amino acid sequences. Liquid chromatography/tandem mass spectrometry was performed on the CD33$^{\Delta E2}$ peptide immunogen used in our antibody campaign. N-terminal amino acids sequences corresponding to those predicted to be encoded by exon 1 were identified.

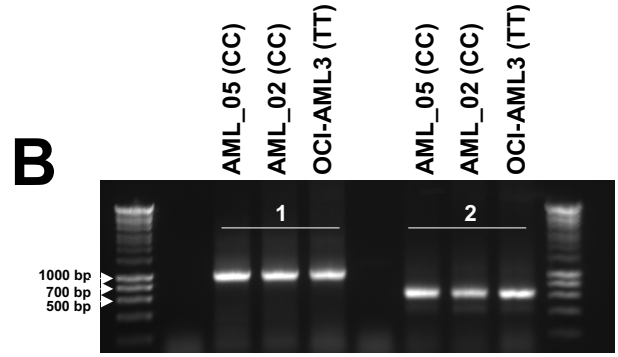
SUPPLEMENTARY FIGURE 7



Binding of 11D5 and 13E11 to various CD33 proteins. Flow cytometric analysis of binding of anti-CD33$^{\Delta E2}$ antibody clones 11D5 and 13E11 to parental REH cells (human acute lymphoblastic leukemia cell line, endogenously CD33^{neg}) and sublines engineered to overexpress endogenous CD33 isoforms (CD33^{FL}, CD33$^{\Delta E2}$) and artificial CD33 isoforms (CD33$^{\Delta E2}$ + SP^{FL}, CD33^{FL} + SP$^{\Delta E2}$). CD33^{FL}, full length CD33 isoform; CD33$^{\Delta E2}$, CD33 isoform arising from transcript missing exon 2; SP, signal peptide; CD33$^{\Delta E2}$ + SP^{FL}, CD33$^{\Delta E2}$ with CD33^{FL} signal peptide; CD33^{FL} + SP$^{\Delta E2}$, CD33^{FL} with CD33$^{\Delta E2}$ signal peptide.

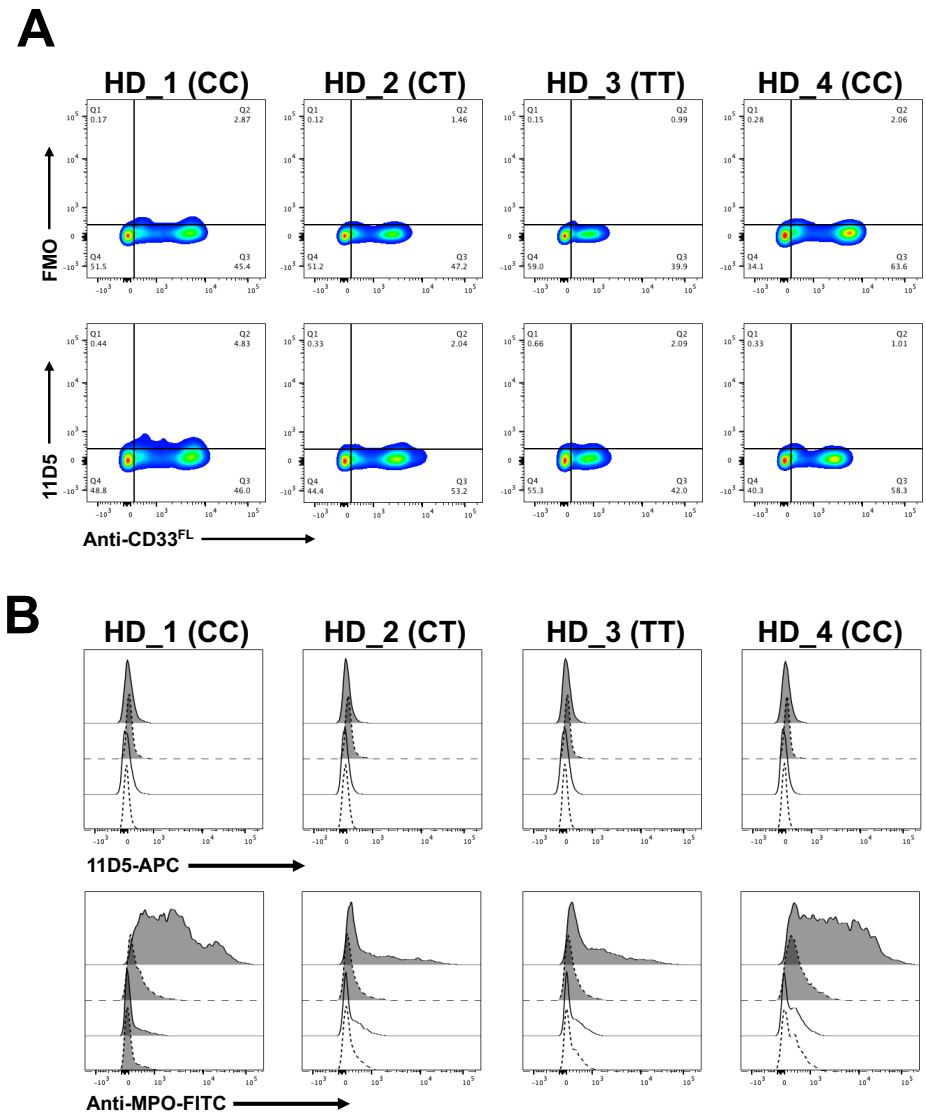
SUPPLEMENTARY FIGURE 8





Expression of *CD33*$^{\Delta E2}$ transcript in AML cell lines and patient samples. **A.** In the indicated cell lines and patient sample (AML_01), RT-PCR was used to identify expression of CD33 isoform transcripts corresponding to 1) CD33^{FL}; 2) CD33$^{\Delta E2}$; 3) CD33^{FL} with truncated exon 7; and 4) CD33$^{\Delta E2}$ with truncated exon 7. **B.** In two additional AML patient samples and a CD33-positive control cell line (OCI-AML3), RT-PCR was used to identify 1) CD33^{FL} and 2) CD33$^{\Delta E2}$ as in *A*. RT-PCR, reverse-transcriptase polymerase chain reaction. Brightness of gel in *A* was increased for clarity. rs12459419 genotype is shown in parentheses.

SUPPLEMENTARY FIGURE 9



Cell surface and intracellular staining of healthy donor bone marrow specimen with anti-CD33$^{\Delta E2}$ antibodies. A. The anti-CD33$^{\Delta E2}$ antibody 11D5 and standard anti-CD33^{FL} antibody (clone p67.6) were used to co-stain four frozen/thawed bone marrow specimens from healthy donors (HD_01, HD_02, HD_03, HD_04). Both full stain and fluorescence minus one (FMO, -11D5) are shown; rs12459419 genotype is shown in parentheses. **B.** The specimens described in *A* were stained with 11D5-APC with (gray histograms) or without (white histograms) permeabilization to detect intracellular antigens. Anti-myeloperoxidase (MPO) antibody was used as a positive control for intracellular antigen detection. Dotted lines represent isotype control antibody. rs12459419 genotype is shown in parentheses.