

## Genetics of CD33 in Alzheimer's Disease and Acute Myeloid Leukemia

**Manasi Malik<sup>1</sup>, Joe Chiles III<sup>1</sup>, Hualin S. Xi<sup>2</sup>, Christopher Medway<sup>4</sup>, James Simpson<sup>1</sup>, Shobha Potluri<sup>3</sup>, Dianna Howard<sup>5</sup>, Ying Liang<sup>5</sup>, Christian M. Paumi<sup>6</sup>, Shubhabrata Mukherjee<sup>7</sup>, Paul Crane<sup>7</sup>, Steven Younkin<sup>4</sup>, David W. Fardo<sup>8</sup> and Steven Estus<sup>1,\*</sup>**

<sup>1</sup>Department of Physiology, Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536 USA

<sup>2</sup>Computational Sciences Center of Emphasis, Pfizer, Inc., Cambridge, MA 02140, USA

<sup>3</sup>Rinat-Pfizer, South San Francisco CA 94080, USA

<sup>4</sup>Department of Neuroscience, Mayo Clinic Jacksonville, Jacksonville, FL 32224 USA

<sup>5</sup>Department of Internal Medicine, University of Kentucky, Lexington, KY 40536 USA

<sup>6</sup>Department of Toxicology, University of Kentucky, Lexington, KY 40536 USA

<sup>7</sup>Department of Medicine, University of Washington, Seattle, WA 98195 USA

<sup>8</sup>Department of Biostatistics, Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536 USA

\*Corresponding Author: Steven Estus, Ph.D. Sanders-Brown Center on Aging, 800 S.

Limestone St. Lexington, KY 40536, Steve.Estus@uky.edu, Phone: (859) 218-3858, Fax:

(859) 323-2866

## Abstract

The CD33 single nucleotide polymorphism (SNP) rs3865444 has been associated with the risk of Alzheimer's disease (AD). Rs3865444 is in linkage disequilibrium with rs12459419 which has been associated with efficacy of an acute myeloid leukemia (AML) chemotherapeutic agent based on a CD33 antibody. We seek to evaluate the extent to which CD33 genetics in AD and AML can inform one another and advance human disease therapy. We have previously shown that these SNPs are associated with skipping of CD33 exon 2 in brain mRNA. Here, we report that these CD33 SNPs are associated with exon 2 skipping in leukocytes from AML patients and with a novel CD33 splice variant that retains CD33 intron 1. Each copy of the minor rs12459419T allele decreases prototypic full-length CD33 expression by about 25% and decreases the AD odds ratio by about 0.10. These results suggest that CD33 antagonists may be useful in reducing AD risk. CD33 inhibitors may include humanized CD33 antibodies such as Lintuzumab which was safe but ineffective in AML clinical trials. Here, we report that Lintuzumab downregulates cell surface CD33 by 80% in phorbol-ester differentiated U937 cells, at concentrations as low as 10 ng/ml. Overall, we propose a model wherein a modest effect on RNA splicing is sufficient to mediate the CD33 association with AD risk and suggest the potential for an anti-CD33 antibody as an AD-relevant pharmacologic agent.

## Introduction

Genetic polymorphisms in the myeloid cell surface receptor CD33 have been implicated in Alzheimer's disease (AD) risk and acute myeloid leukemia (AML) treatment efficacy (1-6). More specifically, rs3865444 in the CD33 promoter has been associated with AD risk while rs12459419 within CD33 exon 2 has been associated with gemtuzumab ozogamycin (GO) efficacy in AML (1-6). We recently reported that these two SNPs are in linkage disequilibrium and associated with exon 2 splicing efficiency in human brain *in vivo* (7). We supported these *in vivo* data with *in vitro* data that rs12459419 is a functional SNP, modulating exon 2 splicing in a minigene splicing model. This association between the minor rs12459419T allele and increased CD33 exon 2 skipping was subsequently confirmed by others (8). Since exon 2 encodes the IgV domain which mediates sialic acid binding (9, 10), CD33 lacking exon 2 is likely to have reduced function. Consistent with this possibility, CD33 inhibits A $\beta$  phagocytosis in microglial cells but CD33 lacking the IgV-domain has no effect on phagocytosis (11). The domain encoded by exon 2 is also critical to the chemotherapeutic actions of GO because this agent depends upon the monoclonal antibody hP67.6, which recognizes an exon 2-encoded epitope (12). Since CD33 genetics contribute to both AD risk and cancer chemotherapy efficacy, we suggest that an exchange between these two disciplines may be enlightening. In particular, we hypothesize that rs12459419 acts on both AD risk and response to AML chemotherapeutics primarily through its effects on CD33 splicing.

To investigate this hypothesis, we have compared CD33 splicing in brain and AML. We identify a novel CD33 splice variant that retains CD33 intron 1, show that this variant is associated with rs12459419 in both brain and AML, and show that exon 2

splicing in AML cells is also associated with rs12459419. We then compare the CD33 SNP allelic dose response on splicing with the dose response on AD risk, finding that a moderate effect on RNA splicing correlates with significant reduction in AD risk. Lastly, we consider whether a CD33-based biological drug from AML may impact AD research; we report that Lintuzumab, a humanized anti-CD33 monoclonal antibody that was safe but ineffective in AML (reviewed in (13, 14)), reduces cell surface CD33 in a robust fashion, suggesting the potential for CD33 antibodies in AD pharmacology.

## Results

To elucidate the mechanism underlying the association between CD33 genetics and response to GO treatment in AML patients, we evaluated CD33 splicing in AML cells. The rationale for this study included that rs12459419 is associated with CD33 exon 2 splicing in brain (7, 8). To assess whether exon 2 shows variable splicing in leukocytes from AML patients, we performed PCR from exons 1 to 3 on cDNA from these cells. The resultant PCR products were separated on polyacrylamide gels and visualized by fluorescent labeling (Figure 1A). This analysis revealed that AML cells express the same CD33 isoforms we detected in human brain, including an isoform lacking exon 2 (D2-CD33) as well as an isoform that retains intron 1 (R1-CD33) (7). CD33 translation is initiated from an ATG within exon 1 and the 351 bp exon 2 encodes the sialic acid-binding IgV domain. Hence, the D2-CD33 isoform encodes a CD33 protein that lacks the sialic acid-binding IgV domain and appears inactive in suppressing microglial activation (Figure 1B) (10). Intron 1 is 62 base pairs in length; consequently, intron 1

retention leads to a frameshift such that the R1-CD33 isoform encodes a prematurely truncated peptide that includes only the signal peptide from CD33 (Figure 1B).

We proceeded to evaluate the extent to which rs12459419 was associated with CD33 splicing in two cohorts of cells from AML patients. In a 26 sample cohort from the University of Kentucky, we quantified D2-CD33 expression by qPCR by using a forward primer at the junction of exon 1-3 and a reverse primer in exon 3 (Figure 1B). Total CD33 was quantified by using primers in exons 4 and 5. Inspection of the relationship between D2-CD33 and total CD33 suggests that expression of D2-CD33 increases in parallel with total CD33 expression and that individuals carrying the minor rs12459419T allele have increased D2-CD33 expression (Figure 2A). This impression is confirmed by analyzing the percentage of CD33 expressed as D2-CD33, noting that D2-CD33 increases from  $10.9 \pm 3.3$  (n=13) in individuals homozygous for the major rs12459419C allele to  $24.4 \pm 8.4\%$  (n=13) in rs12459419C/T heterozygous individuals (mean  $\pm$  SD,  $p=1.6 \times 10^{-5}$ , two-tailed t-test) (Figure 2A). We confirmed these findings by analyzing expression of CD33 isoforms in RNA sequencing data from 107 AML patients available from The Cancer Genome Atlas (TCGA). We found a robust association between rs12459419 and D2-CD33 that was similar to that observed in our smaller cohort ( $p=4.58 \times 10^{-9}$ , one-way ANOVA, Figure 2B). These findings are overall similar to our previous study in human brain that the proportion of CD33 expressed as D2-CD33 increased 10.1 percentage points per rs12459419T allele (7). These results are also consistent with a recent report by Raj et al. who used exon arrays to show that the rs12459419T allele is associated with increased exon 2 skipping in purified human monocytes (8).

We hypothesized that an increase in the proportion of CD33 expressed as D2-CD33 might decrease the efficacy of AML chemotherapeutics based on humanized CD33 monoclonal antibodies such as GO and Lintuzumab. We therefore investigated the binding of Lintuzumab to HEK293T cells transfected with D2-CD33 or full length CD33. Cells were treated with Lintuzumab and the CD33 antibody PWS44, which recognizes an epitope within the IgC2 domain encoded by exons 3-4. As expected, PWS44 labeled the surface of cells transfected with either D2-CD33 or full length CD33. Lintuzumab, however, only labeled cells transfected with full-length CD33 (Figure 3), suggesting that Lintuzumab does not bind D2-CD33 and may have decreased efficacy in individuals that express a higher proportion of CD33 as D2-CD33.

Since we also detected a CD33 isoform that retains intron 1 which is contiguous with exon 2, we hypothesized that rs12459419 may also associate with intron 1 retention. To evaluate this hypothesis, we first quantified R1-CD33 expression by qPCR in our initial cohort of 26 AML samples. We found that R1-CD33 expression ranged from 3.9% to 32.0% of total expression (mean = 20.3%) and showed a modest increase with rs12459419T that was not significant ( $p=0.681$ , two-tailed t-test) (Figure 4A). Since this cohort of 26 individuals offers limited statistical power to detect an association between R1-CD33 expression and rs12459419, we proceeded to analyze expression of the R1-CD33 expression in the TCGA cohort. We found that R1-CD33 expression ranged from 3.3% to 35.1% of total CD33 expression (mean = 13.3%). The percentage of CD33 expressed as R1-CD33 increased from  $10.3 \pm 4.4\%$  ( $n=55$ ) for rs12459419CC individuals to  $14.5 \pm 5.3\%$  ( $n=42$ ) for heterozygotes to  $25.1 \pm 6.8\%$  ( $n=10$ ) for rs12459419TT individuals (mean  $\pm$  SD Figure 4B). The association between R1-CD33 and rs12459419

was statistically significant ( $p=2.17\times 10^{-13}$ ) by one-way ANOVA. We attribute the results discrepancy between the 26 AML sample cohort and the 107 TCGA cohorts to the increased statistical power present in the larger TCGA cohort.

We hypothesized that R1-CD33 may undergo nonsense mediated decay (NMD) because retention of this intron is predicted to lead to a CD33 frameshift and premature translation termination; NMD commonly occurs when a ribosome encounters a termination codon upstream of an exon junction complex (15). NMD can be detected by comparing mRNA levels in the presence and absence of a translation inhibitor. To evaluate the possibility of NMD in *CD33* isoforms, we compared total CD33, D2-CD33, and R1-CD33 levels in K562 cells treated with the transcription inhibitor actinomycin D, with or without the translation inhibitor cycloheximide. This paradigm was shown to be an effective model for NMD as cycloheximide treatment stabilized the NMD-susceptible d7 splice isoform of cyclin T1 (D7-CCNT), as previously reported (16) (Figure 5C-D). However, cycloheximide did not affect the levels of total CD33, D2-CD33 or R1-CD33 indicating that NMD likely does not influence CD33 isoforms (Figure 5A-B). The lack of NMD may be explained by recent findings that mRNA transcripts with AUG-proximal premature termination codons commonly escape NMD due to the interaction of the poly(A)-binding protein 1 (PABP) with the eukaryotic translation initiation factors eIF4G and eIF3, which block binding of the NMD-activating UPF1 to the translation complex (17-19). In summary, R1-CD33 does not appear to undergo NMD.

We proceeded to evaluate the R1-CD33 isoform in the brain. We found that R1-CD33 increased in parallel with the expression of total CD33 (Figure 6A) as well as other microglial marker genes (data not shown); the percentage of CD33 expressed as R1-

CD33 increased in a genotype-dependent manner from  $7.0 \pm 2.9$  (mean  $\pm$  SD) to  $10.2 \pm 4.0$  to  $10.8 \pm 4.6$  for rs12459419CC (n=25), rs12459419CT (n=22), and rs12459419TT (n=4) genotypes, respectively (Figure 6A-B). This was an average of  $2.5 \pm 0.8$  percentage point increase per rs12459419T allele in the percentage of total CD33 expressed as R1-CD33 (ANOVA,  $p=0.003$ ). In summary, the proportion of CD33 expressed as R1-CD33 was associated with rs12459419 genotype in both brain and AML.

To quantify the overall impact of rs3865444 and its proxy functional SNP, rs12459419, on full length CD33 expression in human brain, we subtracted expression of the two atypical isoforms, D2-CD33 and R1-CD33, from total CD33 expression for each brain sample. Using a main-effects ANOVA model accounting for age, microglial marker expression, sex, AD status, and rs12459419 genotype, we found that normalized full length CD33 expression decreased in a genotype-dependent manner from  $0.007580 \pm 0.000373$  (estimated marginal mean  $\pm$  SE, n=25) to  $0.005666 \pm 0.000395$  (n=21) to  $0.004058 \pm 0.000929$  (n=4) for the rs12459419CC, rs12459419 CT, and rs12459419 TT genotypes, respectively ( $p=0.001$ , Figure 6C-D). This represents a 25.2% decrease in full length CD33 expression from the rs12459419 CC genotype to the rs12459419 CT genotypes, and a 46.4% decrease in full length CD33 expression from the CC genotype to the TT genotype.

Since rs3865444 and its proxy rs12459419 show an allelic dose dependence for CD33 splicing, we hypothesized that rs3865444 shows an allelic dose-dependence with AD risk. Previous reports that associated rs3865444 with AD risk used an additive



model, which is the standard for GWAS (4, 5). Here, to evaluate the effects of one and two copies of the rs3865444 minor allele on AD risk, we used a co-dominant model. We performed this analysis first with data on 9,259 AD and 8,361 non-AD DNA samples from the AD Genetics Consortium (ADGC) (4). We found that the SNP showed a dose-dependent association with AD odds (Table 1). This pattern was replicated in 3,455 AD and 5,006 non-AD individuals from the Mayo Clinic cohort (Table 1). A meta-analysis of these overall data shows that rs3865444CA and rs3865444AA confer AD odds ratios of 0.87 and 0.82, respectively. Hence, these data suggest a dose dependent model of rs3865444 in AD and are consistent with the additive action of rs3865444 and its functional proxy, rs12459419, in modulating CD33 splicing.

Since rs3865444A acts through the functional allele rs12459419T to reduce the amount of cell surface CD33 that contains exon 2, pharmacologic agents that act similarly may also reduce AD risk. Antibody-induced cell surface receptor downregulation can be a robust pharmacologic approach (20); others have shown that CD33 is internalized following antibody treatment (21). CD33 antibodies have been developed as possible AML treatment strategies, with the antibody-toxin conjugate GO in use from 2000 to 2010 (reviewed in (13)). The humanized monoclonal antibody Lintuzumab was not toxin conjugated and was found to be safe but ineffective in AML (reviewed in (13, 14)). Additionally, Lintuzumab recognizes an epitope encoded within exon 2 and hence may preferentially decrease CD33 isoforms that include exon 2 (Figure 3). To evaluate the efficacy and potency of agents such as Lintuzumab to induce cell surface CD33 downregulation, we evaluated the dose response and time course for

Lintuzumab actions in U937 cells. We first studied Lintuzumab effects in rapidly dividing, non-differentiated U937 cells. For this assay, cells were treated with either Lintuzumab or human IgG control antibody. Subsequent cell surface CD33 was detected by flow cytometry with an antibody, HIM3-4, which recognizes an epitope within the IgC<sub>2</sub> domain encoded by exons 3-4 (22). Lintuzumab promoted CD33 internalization in a time and concentration dependent fashion (Figure 7A). The maximal Lintuzumab efficacy was a 50% reduction in cell surface CD33 at 70 ng/ml of antibody; higher Lintuzumab concentrations were not more effective. We proceeded to evaluate the concentration-dependent actions of Lintuzumab in U937 cells that were differentiated into a “microglial” phenotype by treatment with 10 or 50 ng/ml phorbol-12-myristate-13-acetate (PMA) (23). Cells were treated with Lintuzumab for 24 hours to model conditions of chronic Lintuzumab treatment. In this study, we found that Lintuzumab was effective at concentrations of 10 ng/ml and above, and that Lintuzumab reduced cell surface CD33 by up to 80% (Figure 7B). For both studies, at least a portion of CD33 remaining on the cell surface likely reflects D2-CD33 because this isoform is recognized by HIM 3-4 but not by Lintuzumab (22). Consistent with this possibility, qPCR studies indicate that  $18.5 \pm 1.0$  % (mean  $\pm$  SD) of CD33 is expressed as D2-CD33 in U937 cells. In summary, humanized monoclonal antibodies such as Lintuzumab offer the possibility of robustly decreasing cell surface CD33 in a fashion that mimics and amplifies the actions of the AD-protective rs3865444A allele.

## Discussion

The primary impact of this study is quantitation of the CD33 genetic relationship with CD33 splicing and human disease coupled with recognition that a CD33 antibody derived from AML pharmacology may be useful in an AD context. More specifically, primary findings include (i) CD33 exon 2 splicing is associated with the linked SNPs rs12459419 and rs3865444 in AML leukocytes, (ii) CD33 intron 1 splicing is associated with rs12459419 in brain and AML, (iii) the rs12459419 T allele results in a dose-dependent decrease in full length CD33 mRNA expression, (iv) the rs12459419 proxy SNP rs3865444 shows allele-dependent association with AD risk and (v) the CD33 antibody Lintuzumab robustly decreases cell surface CD33. Overall, we interpret these results as suggesting that (i) genotype-dependent differences in exon 2 splicing may modulate the efficacy of AML treatments that target exon-2 encoded epitopes, (ii) modest decreases in CD33 splicing may reduce the odds ratio for AD, and (iii) CD33 antibodies may offer the means to pharmacologically replicate and potentially amplify the protective action of rs3865444A on AD risk.

Our finding that rs12459419 is associated with CD33 exon 2 splicing efficiency in leukocytes from AML patients may have significant implications for CD33-based AML therapies. Chemotherapeutic drugs based upon antibodies against CD33 have been used to target AML cells because CD33 is overexpressed in 90% of AML cases (24). Two of these biological drugs, GO and Lintuzumab, have been used extensively in humans; GO was approved for patient use from 2000 to 2010. Since both drugs rely on antibodies against the domain encoded by exon 2 ((22) and Figure 3), these drugs will not recognize D2-CD33, which comprises 2% - 40% of total CD33 in the TCGA cohort. Individuals

homozygous for the major allele of rs12459419 expressed 7.2% of their CD33 as D2-CD33; this portion increased to 17.4% in homozygous minor individuals. We obtained similar results from our smaller cohort of AML patients from University of Kentucky. Long-term studies involving CD33-based therapy have not yet analyzed the effect of CD33 genotype on efficacy. In pilot studies such as the St. Jude's AML02 clinical trial for childhood AML, individuals carrying the minor rs12459419T allele responded less well to a chemotherapy course that included GO (1). An association with rs12459419 was not seen in patients from the AML02 trial who did not receive GO (2). An analysis of the subsequent Children's Oncology Group AAML03P1 trial did not replicate the association of rs12459419 with clinical outcome in patients treated once or twice with GO (2). Our results regarding CD33 splicing in AML cells suggest that individuals carrying the minor allele of the SNP may be less responsive to treatment in part because their cells produce more of the CD33 variant that is not recognized by the GO antibody. These patients, who constitute about 42% of the population, might be more responsive to a modified treatment using an antibody against a constitutively present CD33 epitope, e.g., HIM3-4 which recognizes the IgC<sub>2</sub> epitope encoded by exons 3-4 (22). Alternatively, a patient's rs12459419 genotype may be useful in determining their optimal treatment.

CD33 effects on microglial activation may be critical to CD33 actions in AD. Griciuc et al. reported that ectopic CD33 overexpression in murine BV-2 cells reduces A $\beta$  uptake while CD33 deletion decreases A $\beta$  levels in murine models of AD (11). This action may be mediated by CD33 interaction with CD14, which appears to be an A $\beta$  receptor (25-27), by CD33 modulation of immune activation (reviewed in (28, 29)), or by both mechanisms. Our studies in human brain cDNA indicate that rs12459419T is

associated with increased production of the atypical CD33 splice variants R1-CD33 and D2-CD33, both of which have unclear function with respect to AD mechanisms. R1-CD33 encodes an 18 amino acid CD33 signal peptide followed by a 31 amino acid peptide and premature stop codon. Since this secreted peptide has no known homology to existing proteins, its function, if any, is unclear. D2-CD33 encodes a CD33 protein lacking the IgV domain that is expressed at the cell surface (Figure 3 and (22)). However, in contrast to CD33, this protein does not appear to be functional in A $\beta$  phagocytosis as ectopic expression of CD33 but not D2-CD33 in BV-2 cells reduces A $\beta$  uptake (11). A similar assay performed by Bradshaw et al. suggests that peripheral monocytes from individuals with the minor rs3865444 allele, which produce a high proportion of D2-CD33, exhibit enhanced A $\beta$  phagocytosis relative to individuals who produce more full length CD33 (30). Hence, we predict that an increase in the proportion of CD33 expressed as R1-CD33 and D2-CD33 represents a decrease in normal CD33 function.

The finding that retention of CD33 intron 1 increases with the AD-protective minor allele of rs3865444 (and its proxy rs12459419) has implications for a model of CD33 splicing as the primary mechanism of rs3865444's modulation of AD risk. Previously, we reported that rs12459419T is subtly associated with full length CD33 expression and strongly associated with splicing of CD33 exon 2 (7). In the present study, we report a smaller association between rs12459419 and intron 1 retention. These results combine to produce a model wherein one copy of rs12459419T decreases the production of full length CD33 mRNA by 25.2% while two copies of rs12459419T decrease the production of full length CD33 mRNA by 46.4%. This dose-dependent reduction in

CD33 functionality per copy of rs12459419T corresponds with a dose-dependent decrease in AD risk. While a modest decrease in functional CD33 enables a modest reduction in AD risk, a more robust knockdown of CD33 function by pharmacological agents may enable a more complete alleviation of AD risk.

CD33 antibodies may offer the means to target cell surface CD33 with high specificity and efficacy. The humanized monoclonal antibody Lintuzumab downregulated cell surface CD33 up to 50% in non-differentiated U937 cells and up to 80% in PMA-differentiated U937 cells. We speculate that this difference in efficacy with differentiation reflects that cell surface proteins are replenished more efficiently in rapidly dividing, PMA-naïve cells than in PMA-differentiated, non-dividing cells. In PMA-differentiated cells, Lintuzumab effectively downregulated cell-surface CD33 at 10 ng/mL. This concentration is about 0.1% of the plasma concentration of AML patients treated with Lintuzumab (31). Recognizing that antibody concentrations in the brain are about 0.1% of those in the plasma (32), peripheral infusion of Lintuzumab at doses similar to those used in AML trials may be sufficient to impact CD33 in the brain. While the utility of Lintuzumab in AD will require extended *in vitro* and *in vivo* analysis, AML trials have shown the antibody to be safe ((31, 33), reviewed in (13, 14)). Overall, our results, combined with the strong safety profile, support further evaluation of this antibody in AD research.

This study has several limitations. First, our ability to quantify CD33 at the protein level in genetically diverse human samples is limited by the low CD33 expression in brain and by our limited access to primary cell samples from AML patients. Griciuc et

al. reported that the AD-protective rs3865444A allele was associated with a 30-50% decrease in full length CD33 protein expression in brain (11). In monocytes, leukemic blasts, and PBMCs, rs3865444 has been associated with variable decreases in CD33 expression (8, 11, 30). The quantitative mRNA analysis described here is more consistent with the Grieciuc et al. study. The major contribution of this study is to further explain the mechanism of the association between genotype and total CD33 expression by demonstrating the SNP's effect on CD33 splicing. The Raj et al. study was similar to our finding in reporting a decrease in CD33 exon 2-containing transcripts with the AD-protective rs3865444A allele (8). Second, this study is underpowered to evaluate a potential link between total CD33 expression, SNP genotype and AML risk group, which has been previously evaluated by others (2, 34-36). Third, the effects of chronic Lintuzumab treatment on CD33 function and, ultimately AD risk, are unclear and may depend upon the differentiation state of the target cell. In differentiated cells such as microglia, the predominant action of Lintuzumab may be to act as a CD33 inhibitor: in these cells, Lintuzumab is more efficacious in downregulating cell surface CD33. These lower CD33 levels at the cell surface may result in reduced CD33 signaling. In contrast, in rapidly dividing cells such as cell lines *in vitro*, the predominant action of chronic Lintuzumab treatment may be to increase CD33 activation, as shown by Kung-Sutherland in leukemic cell lines (37); in these cells, cell surface CD33 may be replenished sufficiently rapidly that moderate levels of CD33 are continuously stimulated by Lintuzumab. Extended studies to evaluate among these possibilities are underway.

In summary, we interpret our results as showing that the AD-associated CD33 polymorphism rs3865444 and its proxy rs124549419 are associated with altered exon 2

and intron 1 splicing in human brain tissue and exon 2 splicing in leukocytes from AML patients. In particular, the minor allele consistently promotes increased exon 2 deletion. The allelic dose-dependent effects of rs12459419 on CD33 splicing are consistent with the allelic dose response for AD risk. Lastly, antibodies such as Lintuzumab may represent the means to translate these genetic findings into a pharmacologic agent. Further studies are necessary to elucidate the actions of chronic Lintuzumab treatment on microglial function to understand their potential relevance to AD.

## Materials and Methods

**Human tissue samples.** Human samples were obtained with appropriate institutional review board approval. Brain RNA and DNA were prepared from de-identified human brain specimens provided by the University of Kentucky AD Center Neuropathology Core and have been previously described (38, 39). Samples were from 31 men (16 AD and 15 non-AD) and 26 women (12 AD and 14 non-AD). The age at death for AD individuals was  $81.7 \pm 6.3$  (mean  $\pm$  SD) while the age at death of non-AD cases was  $82.3 \pm 8.6$ . AD diagnoses were made on the basis of dementia and neuropathology (amyloid plaques and neurofibrillary tangles) as previously described (40). Leukocytes from 24 AML and two chronic myeloid leukemia patients were obtained from the University of Kentucky Markey Cancer Center, prepared by Ficoll gradients, and frozen at  $-80^{\circ}\text{C}$ . RNA was extracted from these cells by using the Trizol extraction method. Although these samples represent both leukemic and non-leukemic leukocytes, CD33 expression has been shown to be largely restricted to blasts and myeloid progenitor cells



(41). Similar results were obtained among the two CML and 24 AML cases and hence these samples are referred to as AML for simplicity.

**CD33 mRNA stability.** K562 cells (ATCC) were maintained in Iscove's Modified Dulbecco's Medium with 10% fetal bovine serum, supplemented with non-essential amino acids, 50 U/ml penicillin, and 50 µg/ml streptomycin in a humidified 5% carbon dioxide atmosphere.  $10^6$  cells were plated in 0.9 mL media in a 24-well plate. Cells were treated with actinomycin D (5 µg/mL final concentration) with or without cycloheximide (100 µg/mL). Cells were harvested after 1, 2, 5, and 8 hours of treatment. Triplicate wells were treated for each time point. Cell suspensions were centrifuged at 300xg for 5 minutes, and RNA was extracted from cell pellets using RNEasy kits according to the manufacturer's instructions (Life Technologies).

**Genotyping.** Rs12459419 and rs3865444 genotypes were determined by using TaqMan-based assays (Life Technologies). These two SNPs were in perfect LD in all samples. We generally refer to rs12459419 in the quantitative studies as we have previously shown that this is the functional SNP in exon 2 splicing (7).

**Analysis of gene expression in cDNA.** cDNA was prepared from 1 µg total RNA using SuperScript III Reverse Transcriptase with random primers according to the manufacturer's instructions (Life Technologies). CD33 splice variants in AML patients were initially characterized by performing 30 cycles of PCR on pooled cDNA from five individuals, along with cDNA prepared from the U937 cell line. Amplification was

performed from exons 1 to 3 using forward primer 5'-CTCAGACATGCCGCTGCT and reverse primer 5'-GCACCGAGGAGTGAGTAGTCC. PCR products were separated by polyacrylamide gel electrophoresis and visualized by SYBR-Gold fluorescence.

Total CD33 was quantified by using qPCR and primers corresponding to sequences within exon 4 and exon 5. D2-CD33 expression was quantified by using a forward primer corresponding to the junction of exons 1 and 3 and a reverse primer within exon 3, as described previously (7). R1-CD33 expression was quantified by using forward primer 5'-CGAGCTGACCCTGTTTC corresponding to sequence within intron 1 and reverse primer 5'-GCCTGTGGGTCAAGTCTGTC corresponding to sequence at the junction of exons 2 and 3. Expression of CD33 transcripts encoding an IgV domain was calculated by subtracting D2-CD33 expression and R1-CD33 expression from total CD33 expression. Although we recognize that these transcripts may exhibit splicing variations in exons 6 and 7, we refer to these transcripts as full length CD33 for simplicity.

qPCR was performed by using a Chromo4 thermal cycler (MJ Research) with PerfeCTa SYBR Green SuperMix (Quanta). Each 20  $\mu$ L sample, containing 20 ng of cDNA and 1  $\mu$ M of each primer, underwent an initial denaturation at 95°C for 3 minutes followed by 40 cycles of 15 second denaturation at 95°C, 40 second annealing at 60°C, and 15 second extension at 72°C. A melting curve was performed after each qPCR run to ensure specific amplification. Samples were run in parallel with standard curves to generate accurate copy numbers from C(t)s. Samples of cDNA with fewer than 5 copies of D2-CD33 or R1-CD33 were excluded from subsequent analysis. For brain cDNAs, copy numbers were normalized to the geometric mean of ribosomal protein L32 (RPL32) and eukaryotic initiation factor 4H (EIF4H). RPL32 was quantified using primers described

previously (42). EIF4H was quantified using forward primer 5' - TCTCAGCATAAGGAGTGTACGG in exon 2 with reverse primer 5' - GGGAATCCACTTCATCGAAT in exon 3. Expression of microglial markers CD11b and AIF was quantified as previously described (7). The geometric mean of normalized CD11b and AIF expression was used as an approximation of microglial marker expression.

To validate the nonsense mediated RNA decay experiments, we analysed cyclin T1 splice variants which were previously demonstrated to be differentially susceptible to NMD; cyclin T1 variants were PCR amplified between exons 6 and 9 as previously described (16). PCR products were separated by polyacrylamide gel electrophoresis and stained with SYBR-Gold for 30 minutes. Fluorescence intensity was quantified (Fuji FLA-2000) to calculate relative expression of each isoform.

**RNA sequencing analysis of TCGA samples.** RNA sequencing data from 123 acute myeloid leukemia bone marrow samples was generated as previously described (43). D2-CD33 was quantified by counting sequencing reads that contained the junction of exons 1 and 3. R1-CD33 was quantified by averaging the count of sequencing reads that contained the junction of exon 1 and intron 1 with the count of sequencing reads that contained the junction of intron 1 and exon 2. Total expression was quantified by adding the count of sequencing reads that contained the junction of exons 2 and 3 to the count of sequencing reads that contained the junction of exons 1 and 3. Samples with less than 3 counts were excluded from downstream analysis, resulting in a final sample size of 107. The genotypes at rs12459419 (a coding SNP) for the TCGA AML samples were also

determined from the RNA-seq data. For each sample, the numbers of sequencing reads mapped to the two rs12459419 alleles were counted, and then genotypes were called based on the proportion of reads mapped to the rs12459419 allele (CC: proportion  $\leq$  0.2, CT:  $0.2 < \text{proportion} \leq 0.8$  and TT: proportion  $> 0.8$ ).

**Statistical analysis.** In the cohort of 60 brain samples, two outliers were excluded from subsequent analysis that had normalized CD33 expression 15.5-fold higher than the median or percent R1-CD33 expression 7.7-fold higher than the median. Both of these samples came from individuals heterozygous for rs3865444 and rs12459419 and were classified as outliers according to the Grubb's test for outliers ( $p < 0.05$ ) (44). The associations between normalized CD33 variant expression and rs12459419 genotype were analyzed by linear regression (SPSS) with total CD33, AD status, and rs12459419 as independent variables. The dependence of the proportion of CD33 expressed as R1-CD33 on rs12459419 genotype was calculated using the relevant  $\beta$  coefficient and standard error from a linear regression of the relevant isoform(s) with SNP genotype, AD status, sex, and age as independent variables. Estimated marginal means for full-length CD33 expression for each genotype were calculated using a main-effects ANOVA model with AD status, rs12459419 genotype, and sex as fixed factors and microglial marker expression and age as covariates. In the TCGA cohort, one sample was excluded from subsequent analysis that had percent D2-CD33 expression 8-fold higher than the median. This sample was homozygous for rs12459419CC and was classified as an outlier according to Grubb's test for outliers reference (44).

**Lintuzumab binding to CD33 and D2-CD33.** HEK293T cells (ATCC) were maintained in DMEM with 10% fetal bovine serum, supplemented with 50 U/ml penicillin and 50 µg/ml streptomycin in a humidified 5% carbon dioxide atmosphere. Cells were plated on eight-well LabTek chambered coverglass plates and transfected with TOPO 3.1 expression vectors encoding full-length CD33 (exons 1-7), D2-CD33, or, as a negative control, “empty” 3.1 vector, using Lipofectamine 3000 (Life Technologies) according to the manufacturer’s instructions. Twenty-four hours after transfection, cells were washed with PBS and fixed using cold methanol. Cells were blocked with 5% goat serum and 0.1% Tween in PBS for 1 hour, then incubated overnight at 4°C with either Lintuzumab (200 ng/mL) or PWS44 antibody (1:100 dilution, Leica Biosystems) in goat serum block. Cells were washed three times with PBS-Tween, then incubated with Alexa 488 goat anti-human IgG and Alexa 516 goat anti-mouse IgG at 1:200 dilution in goat serum block for two hours. Cells were washed with PBST, incubated with Hoechst nuclear stain (0.2 µg/mL in PBS) for five minutes, then washed again and maintained in PBS for fluorescent microscopy.

**CD33 allelic dose dependence for AD risk.** The Alzheimer’s Disease Genetics Consortium (ADGC) and Mayo Clinic datasets have been described previously (45) (4). Briefly, the ADGC dataset comprises subjects from 15 cohorts using either Affymetrix or Illumina genotyping arrays (4). Extensive quality control filtering was conducted and included MACH imputation to HapMap phase 2 (release 22) to combine SNPs across genotyping platforms (46). Because some subjects in the Mayo dataset participated in the ADGC study, that cohort was removed from the ADGC dataset for the current analysis.

Of the remaining 14 cohorts, rs3865444 was directly genotyped in 11 of them; results using only those 11 were consistent with those incorporating imputed genotypes.

Therefore, we report combined data from all cohorts.

The Mayo Clinic dataset contained 3455 cases and 5006 controls collected from six centers from the US and Europe as previously described (4). Direct genotyping of rs3865444 was performed using a TaqMan<sup>®</sup> SNP genotyping assay in an ABI PRISM<sup>®</sup> 7900HT Sequence Detection System with 384-well block module from Applied Biosystems (California, USA). First-pass genotype cluster calling was analyzed using the SDS software version 2.2.3 (Applied Biosystems, California, USA).

Association testing was carried out in PLINK (47) using additive, dominant, recessive, and co-dominant logistic regression models that corrected for appropriate covariates; diagnosis age, *APOE*  $\epsilon$ 4 allele dose, *APOE*  $\epsilon$ 2 allele dose, sex and contributing center.

Rs3865444 was significantly associated with AD in each model, with negligible differences in Akaike's Information Criterion between models. To assess the dose dependence of rs3865444 on disease outcome, a co-dominant model was selected to calculate ORs conferred by each genotype of rs3865444 (AA and CA individuals relative to CC individuals.)

**CD33 internalization by Lintuzumab.** U937 cells (ATCC) were maintained in RPMI with 10% fetal bovine serum, supplemented with 50 U/ml penicillin and 50  $\mu$ g/ml streptomycin in a humidified 5% carbon dioxide atmosphere.  $7.5 \times 10^5$  cells were plated in 2 mL of media in six-well plates. For experiments with undifferentiated monocytes, cells were treated with 100 ng/mL human IgG or with the specified concentration of

Lintuzumab. For experiments with differentiated monocytes cells were treated with 10 ng/mL or 50 ng/mL phorbol 12-myristate 13-acetate (ATCC) for 24 hours prior to Lintuzumab or hIgG treatment. Cells were harvested, blocked for 10 minutes on ice with PBS containing 1% BSA and 0.1% sodium azide, and then incubated with either FITC-labeled HIM 3-4 antibody (Beckton-Dickinson) or the relevant isotype control (FITC-labeled mouse IgG $\kappa$ 1, Beckton-Dickinson) per the manufacturer's instructions for at least 30 minutes on ice. Samples were then washed twice with cold PBS, fixed with 1% formaldehyde for 5 minutes at room temperature, then washed again and resuspended in PBS. Flow cytometry was performed using the 488 nm laser on a FACSCalibur (Becton-Dickinson), with at least 10,000 gated events collected per sample. Specific CD33 surface labeling was quantified by subtracting the geometric mean of fluorescence for the isotype-labeled sample from the geometric mean of fluorescence from the HIM 3-4 labeled samples. Data are presented as the percent CD33 remaining after Lintuzumab treatment, relative to the hIgG-treated control for each time point or PMA dose. Independent experiments confirmed that pre-incubation of cells with Lintuzumab at 4°C does not decrease the binding efficiency of the HIM 3-4 antibody (data not shown).

### **Acknowledgments**

The authors would like to express their gratitude to study participants and their families. This research was sponsored and funded by AbbVie Inc. in that Abbvie provided the Lintuzumab used in this study. The RNA-seq results shown here are based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>.

## Alzheimer's Disease Genetics Consortium

Biological samples and associated phenotypic data used in primary data analysis were stored at the Principal Investigator's institutions, and at the National Cell Repository for Alzheimer's Disease (NCRAD), at the NIA Genetics of Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania, and the NIA Alzheimer's Disease Genetics Consortium Data Storage Site at the University of Pennsylvania.

The members of the Alzheimer's Disease Genetics Consortium are: Marilyn S. Albert<sup>1</sup>, Roger L. Albin<sup>2-4</sup>, Liana G. Apostolova<sup>5</sup>, Steven E. Arnold<sup>6</sup>, Clinton T. Baldwin<sup>7</sup>, Robert Barber<sup>8</sup>, Michael M. Barmada<sup>9</sup>, Lisa L. Barnes<sup>10, 11</sup>, Thomas G. Beach<sup>12</sup>, Gary W. Beecham<sup>13, 14</sup>, Duane Beekly<sup>15</sup>, David A. Bennett<sup>10, 16</sup>, Eileen H. Bigio<sup>17</sup>, Thomas D. Bird<sup>18</sup>, Deborah Blacker<sup>19,20</sup>, Bradley F. Boeve<sup>21</sup>, James D. Bowen<sup>22</sup>, Adam Boxer<sup>23</sup>, James R. Burke<sup>24</sup>, Joseph D. Buxbaum<sup>25, 26, 27</sup>, Nigel J. Cairns<sup>28</sup>, Laura B. Cantwell<sup>29</sup>, Chuanhai Cao<sup>30</sup>, Chris S. Carlson<sup>31</sup>, Regina M. Carney<sup>13</sup>, Minerva M. Carrasquillo<sup>33</sup>, Steven L. Carroll<sup>34</sup>, Helena C. Chui<sup>35</sup>, David G. Clark<sup>36</sup>, Jason Corneveaux<sup>37</sup>, Paul K. Crane<sup>38</sup>, David H. Cribbs<sup>39</sup>, Elizabeth A. Crocco<sup>40</sup>, Carlos Cruchaga<sup>41</sup>, Philip L. De Jager<sup>42,43</sup>, Charles DeCarli<sup>44</sup>, Steven T. DeKosky<sup>45</sup>, F. Yesim Demirci<sup>9</sup>, Malcolm Dick<sup>46</sup>, Dennis W. Dickson<sup>33</sup>, Ranjan Duara<sup>47</sup>, Nilufer Ertekin-Taner<sup>33,48</sup>, Denis Evans<sup>49</sup>, Kelley M. Faber<sup>50</sup>, Kenneth B. Fallon<sup>34</sup>, Martin R. Farlow<sup>51</sup>, Lindsay A Farrer<sup>7,52,76,77,83</sup>, Steven Ferris<sup>53</sup>, Tatiana M. Foroud<sup>50</sup>, Matthew P. Frosch<sup>54</sup>, Douglas R. Galasko<sup>55</sup>, Mary Ganguli<sup>56</sup>, Marla Gearing<sup>57,58</sup>, Daniel H. Geschwind<sup>59</sup>, Bernardino Ghetti<sup>60</sup>, John R. Gilbert<sup>13,14</sup>, Sid Gilman<sup>2</sup>, Jonathan D. Glass<sup>61</sup>, Alison M. Goate<sup>41</sup>, Neill R. Graff Radford<sup>33,48</sup>, Robert C. Green<sup>62</sup>, John H. Growdon<sup>63</sup>, Jonathan L. Haines<sup>64, 65</sup>, Hakon



Hakonarson<sup>66</sup>, Kara L. Hamilton-Nelson<sup>13</sup>, Ronald L. Hamilton<sup>67</sup>, John Hardy<sup>68</sup>, Lindy E. Harrell<sup>36</sup>, Elizabeth Head<sup>69</sup>, Lawrence S. Honig<sup>70</sup>, Matthew J. Huentelman<sup>37</sup>, Christine M. Hulette<sup>71</sup>, Bradley T. Hyman<sup>63</sup>, Gail P. Jarvik<sup>72,73</sup>, Gregory A. Jicha<sup>74</sup>, Lee Way Jin<sup>75</sup>, Gyungah Jun<sup>7,76,77</sup>, M. Ilyas Kamboh<sup>9,78</sup>, Anna Karydas<sup>23</sup>, John S.K. Kauwe<sup>79</sup>, Jeffrey A. Kaye<sup>80,81</sup>, Ronald Kim<sup>82</sup>, Edward H. Koo<sup>55</sup>, Neil W. Kowall<sup>83,84</sup>, Joel H. Kramer<sup>85</sup>, Patricia Kramer<sup>80,86</sup>, Walter A. Kukull<sup>87</sup>, Frank M. LaFerla<sup>88</sup>, James J. Lah<sup>61</sup>, Eric B. Larson<sup>38,89</sup>, James B. Leverenz<sup>90</sup>, Allan I. Levey<sup>61</sup>, Ge Li<sup>91</sup>, Andrew P. Lieberman<sup>92</sup>, Chiao Feng Lin<sup>29</sup>, Oscar L. Lopez<sup>78</sup>, Kathryn L. Lunetta<sup>76</sup>, Constantine G. Lyketsos<sup>93</sup>, Wendy J. Mack<sup>94</sup>, Daniel C. Marson<sup>36</sup>, Eden R. Martin<sup>13,14</sup>, Frank Martiniuk<sup>95</sup>, Deborah C. Mash<sup>96</sup>, Eliezer Masliah<sup>55,97</sup>, Richard Mayeux<sup>70,109,110</sup>, Wayne C. McCormick<sup>38</sup>, Susan M. McCurry<sup>98</sup>, Andrew N. McDavid<sup>31</sup>, Ann C. McKee<sup>83,84</sup>, Marsel Mesulam<sup>99</sup>, Bruce L. Miller<sup>23</sup>, Carol A. Miller<sup>100</sup>, Joshua W. Miller<sup>75</sup>, Thomas J. Montine<sup>90</sup>, John C. Morris<sup>28,101</sup>, Jill R. Murrell<sup>50,60</sup>, Amanda J. Myers<sup>40</sup>, Adam C. Naj<sup>13</sup>, John M. Olichney<sup>44</sup>, Vernon S. Pankratz<sup>102</sup>, Joseph E. Parisi<sup>103,104</sup>, Margaret A. Pericak-Vance<sup>13,14</sup>, Elaine Peskind<sup>91</sup>, Ronald C. Petersen<sup>21</sup>, Aimee Pierce<sup>39</sup>, Wayne W. Poon<sup>46</sup>, Huntington Potter<sup>30</sup>, Joseph F. Quinn<sup>80</sup>, Ashok Raj<sup>30</sup>, Murray Raskind<sup>91</sup>, Eric M. Reiman<sup>37,105-107</sup>, Barry Reisberg<sup>53,108</sup>, Christiane Reitz<sup>70,109,110</sup>, John M. Ringman<sup>5</sup>, Erik D. Roberson<sup>36</sup>, Ekaterina Rogaeva<sup>111</sup>, Howard J. Rosen<sup>23</sup>, Roger N. Rosenberg<sup>112</sup>, Mary Sano<sup>26</sup>, Andrew J. Saykin<sup>50,113</sup>, Gerard D. Schellenberg<sup>29</sup>, Julie A. Schneider<sup>10,114</sup>, Lon S. Schneider<sup>35,115</sup>, William W. Seeley<sup>23</sup>, Amanda G. Smith<sup>30</sup>, Joshua A. Sonnen<sup>90</sup>, Salvatore Spina<sup>60</sup>, Peter St George-Hyslop<sup>111,116</sup>, Robert A. Stern<sup>83</sup>, Rudolph E. Tanzi<sup>63</sup>, John Q. Trojanowski<sup>29</sup>, Juan C. Troncoso<sup>117</sup>, Debby W. Tsuang<sup>91</sup>, Otto Valladares<sup>29</sup>, Vivianna M. Van Deerlin<sup>29</sup>, Linda J. Van Eldik<sup>118</sup>, Badri N. Vardarajan<sup>7</sup>, Harry V. Vinters<sup>5,119</sup>, Jean Paul Vonsattel<sup>120</sup>,

Li San Wang<sup>29</sup>, Sandra Weintraub<sup>99</sup>, Kathleen A. Welsh Bohmer<sup>24, 121</sup>, Jennifer Williamson<sup>70</sup>, Randall L. Woltjer<sup>122</sup>, Clinton B. Wright<sup>123</sup>, Steven G. Younkin<sup>33</sup>, Chang En Yu<sup>38</sup>, Lei Yu<sup>10</sup>

<sup>1</sup>Department of Neurology, Johns Hopkins University, Baltimore, Maryland, <sup>2</sup>Department of Neurology, University of Michigan, Ann Arbor, Michigan, <sup>3</sup>Geriatric Research, Education and Clinical Center (GRECC), VA Ann Arbor Healthcare System (VAAHS), Ann Arbor, Michigan, <sup>4</sup>Michigan Alzheimer Disease Center, Ann Arbor, Michigan, <sup>5</sup>Department of Neurology, University of California Los Angeles, Los Angeles, California, <sup>6</sup>Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, <sup>7</sup>Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts, <sup>8</sup>Department of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, Texas, <sup>9</sup>Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, <sup>10</sup>Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, <sup>11</sup>Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, <sup>12</sup>Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona, <sup>13</sup>The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida, <sup>14</sup>Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, <sup>15</sup>National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington, <sup>16</sup>Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, <sup>17</sup>Department of Pathology, Northwestern University, Chicago, Illinois, <sup>18</sup>Department of Neurology, University of Washington, Seattle,

Washington,<sup>19</sup>Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts,<sup>20</sup>Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts,<sup>21</sup>Department of Neurology, Mayo Clinic, Rochester, Minnesota,<sup>22</sup>Swedish Medical Center, Seattle, Washington,<sup>23</sup>Department of Neurology, University of California San Francisco, San Francisco, California,<sup>24</sup>Department of Medicine, Duke University, Durham, North Carolina,<sup>25</sup>Department of Neuroscience, Mount Sinai School of Medicine, New York, New York,<sup>26</sup>Department of Psychiatry, Mount Sinai School of Medicine, New York, New York,<sup>27</sup>Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York,<sup>28</sup>Department of Pathology and Immunology, Washington University, St. Louis, Missouri,<sup>29</sup>Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania,<sup>30</sup>USF Health Byrd Alzheimer's Institute, University of South Florida, Tampa, Florida,<sup>31</sup>Fred Hutchinson Cancer Research Center, Seattle, Washington,<sup>32</sup>Department of Psychiatry, Vanderbilt University, Nashville, Tennessee,<sup>33</sup>Department of Neuroscience, Mayo Clinic, Jacksonville, Florida,<sup>34</sup>Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama,<sup>35</sup>Department of Neurology, University of Southern California, Los Angeles, California,<sup>36</sup>Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama,<sup>37</sup>Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona,<sup>38</sup>Department of Medicine, University of Washington, Seattle, Washington,<sup>39</sup>Department of Neurology, University of California Irvine, Irvine, California,<sup>40</sup>Department of Psychiatry and Behavioral Sciences, Miller School of

Medicine, University of Miami, Miami, Florida,<sup>41</sup>Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri,<sup>42</sup>Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology & Psychiatry, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts,<sup>43</sup>Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts,<sup>44</sup>Department of Neurology, University of California Davis, Sacramento, California,<sup>45</sup>University of Virginia School of Medicine, Charlottesville, Virginia,<sup>46</sup>Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California,<sup>47</sup>Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida,<sup>48</sup>Department of Neurology, Mayo Clinic, Jacksonville, Florida,<sup>49</sup>Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois,<sup>50</sup>Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana,<sup>51</sup>Department of Neurology, Indiana University, Indianapolis, Indiana,<sup>52</sup>Department of Epidemiology, Boston University, Boston, Massachusetts,<sup>53</sup>Department of Psychiatry, New York University, New York, New York,<sup>54</sup>C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts,<sup>55</sup>Department of Neurosciences, University of California San Diego, La Jolla, California,<sup>56</sup>Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania,<sup>57</sup>Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia,<sup>58</sup>Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia,<sup>59</sup>Neurogenetics Program, University of California Los Angeles, Los

Angeles, California, <sup>60</sup>Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana, <sup>61</sup>Department of Neurology, Emory University, Atlanta, Georgia, <sup>62</sup>Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, <sup>63</sup>Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, <sup>64</sup>Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee, <sup>65</sup>Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, <sup>66</sup>Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, <sup>67</sup>Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania, <sup>68</sup>Institute of Neurology, University College London, Queen Square, London, <sup>69</sup>Sanders Brown Center on Aging, Department of Molecular and Biomedical Pharmacology, University of Kentucky, Lexington, Kentucky, <sup>70</sup>Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York, <sup>71</sup>Department of Pathology, Duke University, Durham, North Carolina, <sup>72</sup>Department of Genome Sciences, University of Washington, Seattle, Washington, <sup>73</sup>Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington, <sup>74</sup>Sanders Brown Center on Aging, Department Neurology, University of Kentucky, Lexington, Kentucky, <sup>75</sup>Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California, <sup>76</sup>Department of Biostatistics, Boston University, Boston, Massachusetts, <sup>77</sup>Department of Ophthalmology, Boston University, Boston, Massachusetts, <sup>78</sup>University of Pittsburgh Alzheimer's Disease Research Center,

Pittsburgh, Pennsylvania,<sup>79</sup>Department of Biology, Brigham Young University, Provo, Utah,<sup>80</sup>Department of Neurology, Oregon Health & Science University, Portland, Oregon,<sup>81</sup>Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon,<sup>82</sup>Department of Pathology and Laboratory Medicine, University of California Irvine, Irvine, California,<sup>83</sup>Department of Neurology, Boston University, Boston, Massachusetts,<sup>84</sup>Department of Pathology, Boston University, Boston, Massachusetts,<sup>85</sup>Department of Neuropsychology, University of California San Francisco, San Francisco, California,<sup>86</sup>Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon,<sup>87</sup>Department of Epidemiology, University of Washington, Seattle, Washington,<sup>88</sup>Department of Neurobiology and Behavior, University of California Irvine, Irvine, California,<sup>89</sup>Group Health Research Institute, Group Health, Seattle, Washington,<sup>90</sup>Department of Pathology, University of Washington, Seattle, Washington,<sup>91</sup>Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington,<sup>92</sup>Department of Pathology, University of Michigan, Ann Arbor, Michigan,<sup>93</sup>Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland,<sup>94</sup>Department of Preventive Medicine, University of Southern California, Los Angeles, California,<sup>95</sup>Department of Medicine - Pulmonary, New York University, New York, New York,<sup>96</sup>Department of Neurology, University of Miami, Miami, Florida,<sup>97</sup>Department of Pathology, University of California San Diego, La Jolla, California,<sup>98</sup>School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington,<sup>99</sup>Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Chicago, Illinois,<sup>100</sup>Department of Pathology, University of Southern California, Los Angeles, California,<sup>101</sup>Department of Neurology,

Washington University, St. Louis, Missouri, <sup>102</sup>Department of Biostatistics, Mayo Clinic, Rochester, Minnesota, <sup>103</sup>Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota, <sup>104</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, <sup>105</sup>Arizona Alzheimer's Consortium, Phoenix, Arizona, <sup>106</sup>Department of Psychiatry, University of Arizona, Phoenix, Arizona, <sup>107</sup>Banner Alzheimer's Institute, Phoenix, Arizona, <sup>108</sup>Alzheimer's Disease Center, New York University, New York, New York, <sup>109</sup>Gertrude H. Sergievsky Center, Columbia University, New York, New York, <sup>110</sup>Department of Neurology, Columbia University, New York, New York, <sup>111</sup>Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario, <sup>112</sup>Department of Neurology, University of Texas Southwestern, Dallas, Texas, <sup>113</sup>Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana, <sup>114</sup>Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois, <sup>115</sup>Department of Psychiatry, University of Southern California, Los Angeles, California, <sup>116</sup>Cambridge Institute for Medical Research and Department of Clinical Neurosciences, University of Cambridge, Cambridge, <sup>117</sup>Department of Pathology, Johns Hopkins University, Baltimore, Maryland, <sup>118</sup>Sanders Brown Center on Aging, Department of Anatomy and Neurobiology, University of Kentucky, Lexington, Kentucky, <sup>119</sup>Department of Pathology & Laboratory Medicine, University of California Los Angeles, Los Angeles, California, <sup>120</sup>Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York, <sup>121</sup>Department of Psychiatry & Behavioral Sciences, Duke University, Durham, North Carolina, <sup>122</sup>Department of Pathology, Oregon Health & Science University, Portland, Oregon, <sup>123</sup>Evelyn F.

McKnight Brain Institute, Department of Neurology, Miller School of Medicine,  
University of Miami, Miami, Florida

### Conflict of Interest Statement

The University of Kentucky has a patent pending on the use of CD33 inhibitors relative to AD.

### Funding

This work is funded by National Institutes of Health [P01-AGO30128 and R01-AG045775 (SE), P30-AG028383, R25GM093044 and K25-AG043546 (DWF)] and the University of Kentucky Bucks for Brains program (MM). The ADGC is also NIH funded (U01 AG032984).

### References

1. Lamba, J.K., Pounds, S., Cao, X., Downing, J.R., Campana, D., Ribeiro, R.C., Pui, C.H. and Rubnitz, J.E. (2009) Coding polymorphisms in CD33 and response to gemtuzumab ozogamicin in pediatric patients with AML: A pilot study. *Leukemia*, **23**, 402-404.
2. Mortland, L., Alonzo, T.A., Walter, R.B., Gerbing, R.B., Mitra, A.K., Pollard, J.A., Loken, M.R., Hirsch, B., Raimondi, S., Franklin, J. *et al.* (2013) Clinical significance of CD33 nonsynonymous single-nucleotide polymorphisms in pediatric patients with acute myeloid leukemia treated with gemtuzumab-ozogamicin-containing chemotherapy. *Clin Cancer Res*, **19**, 1620-1627.
3. Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., Jun, G., Destefano, A.L., Bis, J.C., Beecham, G.W. *et al.* (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*, **45**, 1452-1458.
4. Naj, A.C., Jun, G., Beecham, G.W., Wang, L.S., Vardarajan, B.N., Buross, J., Gallins, P.J., Buxbaum, J.D., Jarvik, G.P., Crane, P.K. *et al.* (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet*, **43**, 436-441.



5. Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J.C., Carrasquillo, M.M., Abraham, R., Hamshere, M.L., Pahwa, J.S., Moskvina, V. *et al.* (2011) Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet*, **43**, 429-435.
6. Carrasquillo, M.M., Belbin, O., Hunter, T.A., Ma, L., Bisceglia, G.D., Zou, F., Crook, J.E., Pankratz, V.S., Sando, S.B., Aasly, J.O. *et al.* (2011) Replication of EPHA1 and CD33 associations with late-onset Alzheimer's disease: A multi-centre case-control study. *Mol Neurodegener*, **6**, 54.
7. Malik, M., Simpson, J.F., Parikh, I., Wilfred, B.R., Fardo, D.W., Nelson, P.T. and Estus, S. (2013) CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. *J Neurosci*, **33**, 13320-13325.
8. Raj, T., Ryan, K.J., Replogle, J.M., Chibnik, L.B., Rosenkrantz, L., Tang, A., Rothamel, K., Stranger, B.E., Bennett, D.A., Evans, D.A. *et al.* (2014) CD33: Increased inclusion of exon 2 implicates the IgV-set domain in Alzheimer's disease susceptibility. *Hum Mol Genet*. **23**, 2729-2736.
9. Crocker, P.R., Vinson, M., Kelm, S. and Drickamer, K. (1999) Molecular analysis of sialoside binding to sialoadhesin by NMR and site-directed mutagenesis. *Biochem J*, **341** ( Pt 2), 355-361.
10. Hernandez-Caselles, T., Martinez-Esparza, M., Perez-Oliva, A.B., Quintanilla-Cecconi, A.M., Garcia-Alonso, A., Alvarez-Lopez, D.M. and Garcia-Penarrubia, P. (2006) 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*, **79**, 46-58.
11. Griciuc, A., Serrano-Pozo, A., Parrado, A.R., Lesinski, A.N., Asselin, C.N., Mullin, K., Hooli, B., Choi, S.H., Hyman, B.T. and Tanzi, R.E. (2013) Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. *Neuron*, **78**, 631-643.
12. Dowell, J.A., Korth-Bradley, J., Liu, H., King, S.P. and Berger, M.S. (2001) Pharmacokinetics of gemtuzumab ozogamicin, an antibody-targeted chemotherapy agent for the treatment of patients with acute myeloid leukemia in first relapse. *J Clin Pharmacol*, **41**, 1206-1214.
13. Jurcic, J.G. (2012) What happened to anti-CD33 therapy for acute myeloid leukemia? *Curr Hematol Malig Rep*, **7**, 65-73.
14. Laszlo, G.S., Estey, E.H. and Walter, R.B. (2014) The past and future of CD33 as therapeutic target in acute myeloid leukemia. *Blood Rev*, **28**, 143-153.
15. Silva, A.L. and Romao, L. (2009) The mammalian nonsense-mediated mRNA decay pathway: To decay or not to decay! Which players make the decision? *FEBS Lett*, **583**, 499-505.
16. Urano, E., Miyauchi, K., Ichikawa, R., Futahashi, Y. and Komano, J. (2012) Regulation of cyclin T1 expression and function by an alternative splice variant that skips exon 7 and contains a premature termination codon. *Gene*, **505**, 1-8.
17. Inacio, A., Silva, A.L., Pinto, J., Ji, X., Morgado, A., Almeida, F., Faustino, P., Lavinha, J., Liebhaber, S.A. and Romao, L. (2004) Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay. *J Biol Chem*, **279**, 32170-32180.
18. Peixeiro, I., Inacio, A., Barbosa, C., Silva, A.L., Liebhaber, S.A. and Romao, L. (2012) Interaction of PABPC1 with the translation initiation complex is critical to the

- NMD resistance of AUG-proximal nonsense mutations. *Nucleic Acids Res*, **40**, 1160-1173.
19. Neu-Yilik, G., Amthor, B., Gehring, N.H., Bahri, S., Paidassi, H., Hentze, M.W. and Kulozik, A.E. (2011) Mechanism of escape from nonsense-mediated mRNA decay of human beta-globin transcripts with nonsense mutations in the first exon. *RNA*, **17**, 843-854.
  20. Albanell, J., Codony, J., Rovira, A., Mellado, B. and Gascon, P. (2003) Mechanism of action of anti-HER2 monoclonal antibodies: Scientific update on trastuzumab and 2c4. *Adv Exp Med Biol*, **532**, 253-268.
  21. Walter, R.B., Raden, B.W., Zeng, R., Hausermann, P., Bernstein, I.D. and Cooper, J.A. (2008) ITIM-dependent endocytosis of CD33-related siglecs: Role of intracellular domain, tyrosine phosphorylation, and the tyrosine phosphatases, SHP1 and SHP2. *Journal of leukocyte biology*, **83**, 200-211.
  22. Perez-Oliva, A.B., Martinez-Esparza, M., Vicente-Fernandez, J.J., Corral-San Miguel, R., Garcia-Penarrubia, P. and Hernandez-Caselles, T. (2011) Epitope mapping, expression and post-translational modifications of two isoforms of CD33 (CD33M and CD33m) on lymphoid and myeloid human cells. *Glycobiology*, **21**, 757-770.
  23. Liu, M.Y. and Wu, M.C. (1992) Induction of human monocyte cell line U937 differentiation and CSF-1 production by phorbol ester. *Exp Hematol*, **20**, 974-979.
  24. Legrand, O., Perrot, J.Y., Baudard, M., Cordier, A., Lautier, R., Simonin, G., Zittoun, R., Casadevall, N. and Marie, J.P. (2000) The immunophenotype of 177 adults with acute myeloid leukemia: Proposal of a prognostic score. *Blood*, **96**, 870-877.
  25. Liu, Y., Walter, S., Stagi, M., Cherny, D., Letiembre, M., Schulz-Schaeffer, W., Heine, H., Penke, B., Neumann, H. and Fassbender, K. (2005) LPS receptor (CD14): A receptor for phagocytosis of Alzheimer's amyloid peptide. *Brain*, **128**, 1778-1789.
  26. Reed-Geaghan, E.G., Savage, J.C., Hise, A.G. and Landreth, G.E. (2009) CD14 and Toll-like receptors 2 and 4 are required for fibrillar A{beta}-stimulated microglial activation. *J Neurosci*, **29**, 11982-11992.
  27. Ishida, A., Akita, K., Mori, Y., Tanida, S., Toda, M., Inoue, M. and Nakada, H. (2014) Negative regulation of Toll-like receptor-4 signaling through the binding of glycosylphosphatidylinositol-anchored glycoprotein, CD14, with the sialic acid-binding lectin, CD33. *J Biol Chem*, **289**, 25341-25350.
  28. Linnartz, B. and Neumann, H. (2013) Microglial activatory (immunoreceptor tyrosine-based activation motif)- and inhibitory (immunoreceptor tyrosine-based inhibition motif)-signaling receptors for recognition of the neuronal glycocalyx. *Glia*, **61**, 37-46.
  29. Linnartz, B., Wang, Y. and Neumann, H. (2010) Microglial immunoreceptor tyrosine-based activation and inhibition motif signaling in neuroinflammation. *Int J Alzheimers Dis*, **2010**.
  30. Bradshaw, E.M., Chibnik, L.B., Keenan, B.T., Ottoboni, L., Raj, T., Tang, A., Rosenkrantz, L.L., Imboywa, S., Lee, M., Von Korff, A. *et al.* (2013) CD33 Alzheimer's disease locus: Altered monocyte function and amyloid biology. *Nat Neurosci*, **16**, 848-850.
  31. Caron, P.C., Dumont, L. and Scheinberg, D.A. (1998) Supersaturating infusional humanized anti-CD33 monoclonal antibody HUM195 in myelogenous leukemia. *Clin Cancer Res*, **4**, 1421-1428.

32. Levites, Y., Smithson, L.A., Price, R.W., Dakin, R.S., Yuan, B., Sierks, M.R., Kim, J., McGowan, E., Reed, D.K., Rosenberry, T.L. *et al.* (2006) Insights into the mechanisms of action of anti-Abeta antibodies in Alzheimer's disease mouse models. *FASEB J*, **20**, 2576-2578.
33. Caron, P.C., Jurcic, J.G., Scott, A.M., Finn, R.D., Divgi, C.R., Graham, M.C., Jureidini, I.M., Sgouros, G., Tyson, D., Old, L.J. *et al.* (1994) A Phase 1B trial of humanized monoclonal antibody M195 (anti-CD33) in myeloid leukemia: Specific targeting without immunogenicity. *Blood*, **83**, 1760-1768.
34. Pollard, J.A., Alonzo, T.A., Loken, M., Gerbing, R.B., Ho, P.A., Bernstein, I.D., Raimondi, S.C., Hirsch, B., Franklin, J., Walter, R.B. *et al.* (2012) Correlation of CD33 expression level with disease characteristics and response to gemtuzumab ozogamicin containing chemotherapy in childhood AML. *Blood*, **119**, 3705-3711.
35. Mingari, M.C., Vitale, C., Romagnani, C., Falco, M. and Moretta, L. (2001) Regulation of myeloid cell proliferation and survival by p75/AIRM1 and CD33 surface receptors. *Adv Exp Med Biol*, **495**, 55-61.
36. Mingari, M.C., Vitale, C., Romagnani, C., Falco, M. and Moretta, L. (2001) P75/AIRM1 and CD33, two sialoadhesin receptors that regulate the proliferation or the survival of normal and leukemic myeloid cells. *Immunol Rev*, **181**, 260-268.
37. Sutherland, M.K., Yu, C., Lewis, T.S., Miyamoto, J.B., Morris-Tilden, C.A., Jonas, M., Sutherland, J., Nesterova, A., Gerber, H.P., Sievers, E.L. *et al.* (2009) Anti-leukemic activity of lintuzumab (SGN-33) in preclinical models of acute myeloid leukemia. *MABs*, **1**, 481-490.
38. Ling, I.F., Bhongsatiern, J., Simpson, J.F., Fardo, D.W. and Estus, S. (2012) Genetics of clusterin isoform expression and Alzheimer's disease risk. *PLoS One*, **7**, e33923.
39. Zou, F., Gopalraj, R.K., Lok, J., Zhu, H., Ling, I.F., Simpson, J.F., Tucker, H.M., Kelly, J.F., Younkin, S.G., Dickson, D.W. *et al.* (2008) Sex-dependent association of a common low-density lipoprotein receptor polymorphism with RNA splicing efficiency in the brain and Alzheimer's disease. *Hum Mol Genet*, **17**, 929-935.
40. Nelson, P.T., Braak, H. and Markesbery, W.R. (2009) Neuropathology and cognitive impairment in Alzheimer disease: A complex but coherent relationship. *J Neuropathol Exp Neurol*, **68**, 1-14.
41. Griffin, J.D., Linch, D., Sabbath, K., Larcom, P. and Schlossman, S.F. (1984) A monoclonal antibody reactive with normal and leukemic human myeloid progenitor cells. *Leuk Res*, **8**, 521-534.
42. Zhang, X., Ding, L. and Sandford, A.J. (2005) Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Mol Biol*, **6**, 4.
43. Cancer Genome Atlas Research, N. (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*, **368**, 2059-2074.
44. Grubbs, F.E. (1950) Sample criteria for testing outlying observations. *Ann. Math. Statist.*, **21**, 27-58.
45. Ridge, P.G., Mukherjee, S., Crane, P.K., Kauwe, J.S. and Alzheimer's Disease Genetics, C. (2013) Alzheimer's disease: Analyzing the missing heritability. *PLoS One*, **8**, e79771.

46. Li, Y., Willer, C.J., Ding, J., Scheet, P. and Abecasis, G.R. (2010) Mach: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol*, **34**, 816-834.
47. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, **81**, 559-575.

## Legends to Figures

### Figure 1. *CD33* Splicing in AML Leukocytes

*CD33* splice variants identified in cDNA from five AML patients (Pool) and cDNA prepared from the U937 cell line after PCR of exons 1 to 3 are shown in (A). These are similar to those identified previously in human brain cDNA, and include full length *CD33*, D2-*CD33*, and R1-*CD33*. The gene and protein structure of *CD33* are depicted (B), including translation start site (Met), signal peptide (SP), sialic-acid binding immunoglobulin-like variable domain (IgV), immunoglobulin-like structural domain (IgC), transmembrane domain (TMD), and putative immunoreceptor tyrosine-based inhibitory motif (ITIM). The D2-*CD33* variant lacks the IgV domain encoded by exon 2, while R1-*CD33* is prematurely terminated at the beginning of exon 2 and consequently encodes only a signal peptide. The arrows refer to the locations of the primers used for qPCR.

### Figure 2. *CD33* Exon 2 Splicing in AML

The percentage of *CD33* expressed as D2-*CD33* increases with the presence of the minor rs12459419T allele in the cohort of 26 AML leukocytes ( $p=1.6\times 10^{-5}$ , two-tailed t-test; power=0.99; A). The association between rs12459419 and exon 2 splicing was confirmed by RNA-sequencing analysis of 107 bone marrow aspirates from AML patients ( $p=4.58\times 10^{-9}$ , one-way ANOVA; power=1.00; B).

### Figure 3. Lintuzumab binds CD33 but not D2-CD33

HEK293T cells were transiently transfected with either CD33 (A-D) or D2-CD33 (E-H). Cells were then labeled with the CD33 antibody PWS44 which recognizes an IgC<sub>2</sub> epitope (A, E) or Lintuzumab (B, F). Cellular nuclei are visualized by Hoechst fluorescence (C and G). Overlays of all three fluorescent labels (D, H) show that PWS44 and Lintuzumab both recognize CD33 (D) but only PWS44 labels D2-CD33 (H).

### Figure 4. CD33 Intron 1 Splicing in AML Leukocytes

In the 26 AML sample University of Kentucky cohort, the percent of CD33 expressed as R1-CD33 showed a modest trend towards an increase with the rs12459419T allele that was not significant ( $p=0.681$ , two-tailed t-test; power $<0.1$ ; A). However, in the larger 107 TCGA sample cohort, the percentage of CD33 expressed as R1-CD33 was found to increase significantly with the rs12459419T allele ( $p=1.54\times 10^{-13}$ , one-way ANOVA; power=1.00; B).

### Figure 5. R1-CD33 does not undergo NMD

The proportion of CD33 expressed as R1-CD33 (A) and D2-CD33 (B) remained constant between samples treated with actinomycin D alone (white bars) and samples treated with both actinomycin D and cycloheximide (grey bars). NMD was present in these samples as discerned by the positive control, i.e., the ratio of D7-CCNT to FL-CCNT (C). Error bars show standard deviation in triplicate samples. To compare the decay rates of total CD33, D2-CD33, and R1-CD33, K562 cells were treated with actinomycin D and either cycloheximide (CHX) or solvent control for the indicated time and each CD33 isoform

quantified by qPCR. CCNT isoforms were quantified using polyacrylamide gel electrophoresis followed by SYBR-gold fluorescent detection and analysis of gel images. A representative gel image from triplicates treated for 5 hours is shown (D).

**Figure 6. R1-CD33 and full length CD33 expression in brain are associated with rs12459419 genotype**

R1-CD33 expression is associated with total CD33 expression in human brain as well as rs12459419 genotype (A). Linear regression analysis of R1-CD33 expression revealed a significant model (adjusted  $r^2=0.513$ ) wherein R1-CD33 expression was associated with total CD33 expression ( $p=9.4\times 10^{-9}$ , standardized  $\beta$  coefficient=0.738) as well as rs12459419 genotype ( $p=0.0034$ , standardized  $\beta$  coefficient=0.314). R1-CD33 expression was not associated with AD status ( $p=0.78$ , standardized  $\beta$  coefficient=0.029). (A,  $r^2=0.602$ , 0.433, and 0.799 for the CC, CT, and TT genotypes, respectively). The percent of CD33 expressed as R1-CD33 increases as a function of rs12459419 in human brain ( $p=4.53\times 10^{-3}$ ), with a  $2.5\pm 0.8$  percentage point increase per copy of rs12459419T (B). Full length CD33 expression (the result of subtracting D2-CD33 and R1-CD33 expression from total CD33 expression) is shown relative to microglial marker expression (the geometric mean of CD11b and AIF expression) and rs12459419 genotype (C). Linear regression analysis of full length CD33 expression reveals a significant model (adjusted  $r^2=0.726$ ) wherein mRNA encoding full length CD33 was associated with microglial marker expression ( $p=3.32\times 10^{-11}$ ), standardized  $\beta$  coefficient=0.668), with rs12459419 genotype ( $p=1.58\times 10^{-5}$ , standardized  $\beta$  coefficient=-0.376), and with AD status ( $p=2.45\times 10^{-4}$ , standardized  $\beta$  coefficient=0.310). Full length CD33 expression

is divided by microglial marker expression to account for variations in cell type composition of brain samples (D). Full length CD33 expression normalized to microglial content decreases from  $0.0948 \pm 0.0048$  (mean  $\pm$  SE, n=25) to  $0.0718 \pm 0.0043$  (n=22) to  $0.0478 \pm 0.0035$  (n=4) for the rs12459419 CC, CT, and TT genotypes respectively.

**Figure 7. Lintuzumab decreases cell surface CD33.**

Total cell surface CD33 was quantified by using flow cytometry with HIM3-4, a CD33 antibody that recognizes an epitope encoded by exons 3-4 (22). U937 were treated with Lintuzumab at the indicated concentrations and times, placed on ice, and then exposed to HIM3-4. Lintuzumab reduced cell surface CD33 up to 50% in PMA-naïve cells (A). Lintuzumab was more effective in reducing CD33 in PMA-treated cells, reaching a maximum of 80% efficacy (B). The data for PMA (0 ng/ml) in B is reproduced from A to allow for direct comparison.



Cohort (n)	Rs3865444	AD Odds Ratio	p value
Mayo Clinic (8,461)	Rs3865444CA	0.85 (0.76- 0.95)	0.0049
	Rs3865444AA	0.79 (0.65-0.96)	0.0148
ADGC (17,620)	Rs3865444CA	0.88 (0.82-0.94)	0.0002
	Rs3865444AA	0.83 (0.73-0.94)	0.0021
Combined (26,081)	Rs3865444CA	0.87 (0.82-0.92)	$3.88 \times 10^{-6}$
	Rs3865444AA	0.82 (0.74-0.91)	$9.76 \times 10^{-5}$

**Table 1. Rs3865444 allelic dose dependence for AD risk.**

These results were adjusted for PCs, cohort, sex, age and APOE genotypes. The rs3865444CA and AA genotypes are compared to rs3865444CC major allele homozygotes.













