High expression of the stem cell marker GPR56 at diagnosis identifies acute myeloid leukemia patients at higher relapse risk after allogeneic stem cell transplantation with the CD34^{+}/CD38^{-} population

Madlen Jentzsch, Marius Bill, Juliane Grimm, Julia Schulz, Luba Schuhmann, Dominic Brauer, Karoline Goldmann, Franziska Wilke, Georg-Nikolaus Franke, Gerhard Behre, Wolfram Pönisch, Vladan Vucinic, Dietger Niederwieser, Uwe Platzbecker and Sebastian Schwind

Medical Clinic and Policlinic 1, Hematology and Cellular Therapy, Leipzig University Hospital, Leipzig, Germany

Correspondence: SEBASTIAN SCHWIND - sebastian.schwind@medizin.uni-leipzig.de

doi:10.3324/haematol.2019.229260

Supplementary Material to Jentzsch *et al.*: High expression of the stem cell marker *GPR56* at diagnosis identifies acute myeloid leukemia patients at higher relapse risk after allogeneic stem cell transplantation in context with the CD34+/CD38- population

Difference to the published patient set used to analyze the prognostic impact of the CD34+/CD38- cell burden at diagnosis

Previously, we showed a high CD34+/CD38- bone marrow (BM) cell burden (≥6%) at diagnosis to associate with inferior overall (OS) and disease-free survival in acute myeloid leukemia (AML) patients receiving an allogeneic hematopoietic stem cell transplantation (HSCT) in complete remission (CR).[1] For 123 of these patients, pretreatment BM material for *GPR56* analysis at diagnosis was available. Additionally, this set was expanded by 57 patients receiving HSCT in CR or CR with incomplete peripheral recovery (CRi) with information on the CD34+/CD38- cell burden and pretreatment BM material available as well as 33 patients with pretreatment BM material at diagnosis available but without information on the CD34+/CD38- cell burden at diagnosis.

Further genetic and HSCT related information

Additional morphologic information, surface antigen expression and allogeneic HSCT related information is shown in *Supplementary Table S1*.

Treatment protocols

In the subgroup of AML patients diagnosed at the age of 60 years or younger 106 received chemotherapy according to the AML 2002 study (OSHO #061),[2] two patients received chemotherapy within the PKC412 protocol[3] one patient was diagnosed with AML as a child and treated within the AML BFM-2014 study.[4] Among AML patients older than 60 years at diagnosis, 95 were treated within the AML 2004 (#069, ClinicalTrials.gov Identifier: NCT01497002) study,[5] and five patients were treated within the OSHO #083 protocol.

Allogeneic hematopoietic stem cell transplantation

For consolidation therapy, all patients received an allogeneic HSCT after myeloablative (MAC) or non-myeloablative (NMA) conditioning in CR or CRi at the University of Leipzig. NMA conditioning was administered to 154 patients and consisted of three days of 30 mg/m^2 Fludarabine and a 2 Gy total body irradiation (TBI) on the day of HSCT. MAC HSCT was administered to 59 patients and consisted of two days of 60 mg/kg body weight (BW) Cyclophosphamide followed by or prior to 12 Gy TBI in two daily doses over three days. All patients received granulocyte-colony stimulating factor stimulated peripheral blood stem cells. Reasons for NMA as opposed to MAC-HSCT were age over 50 years in patients receiving unrelated HSCT or age over 55 years in patients receiving related HSCT (n=106), previous autologous HSCT (n=3) or active systemic infections (n=4). Patients harboring a core binding factor AML received an allogeneic HSCT in second CR or according to patients' choice in first CR if a suitable donor was available.

Prevention of graft-versus-host disease

Prevention of graft-versus-host disease (GvHD) was different according to the two conditioning regimes used. All patients receiving MAC-HSCT were treated with cyclosporine A (CyA), beginning intravenously with 5 mg/kg BW in two daily doses from day -1. Blood levels of CyA were measured from day 0 and doses were adjusted for target levels of 200 ng/ml. Patients also received methotrexate 15 mg intravenously on days +1, +3, +6 and +11 after HSCT. Furthermore, patients with an unrelated donor additionally received *in vivo* T-cell depletion with thymoglobulin 2 mg/kg BW per day on days -3 to -1. All patients receiving NMA-HSCT were treated with the same dose CyA as MAC patients and additionally received mycophenolate mofetil (MMF) 3 g per day in three daily doses if receiving unrelated HSCT or 2 g per day in two daily doses if receiving related HSCT. CyA was reduced starting on day +84 or day +180 following related or unrelated HSCT, respectively, and MMF was stopped at day +28 following related HSCT and tapered from days +40 to +96 following unrelated HSCT.[6] Immunosuppression was prolonged or extended with systemic steroids in cases of

GvHD (grade > 2 according to Glucksberg grading system). [7] Patients were evaluated for incidence of acute GvHD (aGvHD) and chronic GvHD (cGvHD), using established criteria of the Glucksberg grading system.[7] Requirement for aGvHD was engraftment while requirement for cGvHD was engraftment and survival for at least 100 days after HSCT.

Cytogenetic Analysis

Pretreatment BM cytogenetic analyses were performed centrally in our institution using standard banding techniques. In cases were no metaphases could be obtained (n=13), fluorescence in-situ hybridization (FISH) was used to screen for recurrent abnormalities (i.e. deletion [del]5/5q, del7/7q, trisomy 8, abn(p53) and abn11q23).

Flow Cytometry

For 180 patients EDTA-anticoagulated fresh BM aspirates were analyzed at AML diagnosis using our institution's standard myeloid panel. Surface antigen expression was measured with the FACSCalibur flow cytometer and analyzed using the CELLQUEST software (Becton Dickinson). Samples were stored at room temperature and processed within 12 hours. 100µl of BM aspirates were incubated for 15 minutes with Fluorescein isothiocyanate (FITC-; CD7, CD15, CD61, CD65), R-Phycoerythrin (PE-; CD13, CD64, Glykophorin A) phycoerythrin-Cy7 (PE-Cy7-; CD2, CD45, CD38, CD56, CD117) and allophycocyanin (APC-; CD11b, CD14, CD33, CD34) labeled monoclonal antibodies. After erythrocyte lysis, samples were washed in 1ml phosphate buffered saline solution (PBS) acid, fixed in 500µl PBS acid with 1% formaldehyde and 10,000 cells per sample were analyzed. The BM blast population was determined by low side scatter properties as well as a low expression of CD45.

Analysis of gene expression levels

For gene expression analysis, total RNA was extracted using TRIzol (Invitrogen, Carlsbad, CA) following the manufacturer's protocols. cDNA was synthesized using Superscript III (Invitrogen, Carlsbad, CA) which was used for a real time polymerase chain reaction (RT-

qPCR). RT-qPCR of *GPR56* (Hs00173754_m1), *BAALC* (Hs00227249_m1), *MN1* (Hs00159202_m1), *EVI1* (Hs00602795_m1), *ABL1* (Hs01104728_m1) and *18S* (Hs99999901_s1) was performed with Taqman gene expression assays (Applied Biosystems, Carlsbad, CA) following the manufacturer's protocols. *GPR56*, *BAALC,* and *MN1* expression was normalized to *ABL1* as internal control. The comparative cycle threshold ($\Delta\Delta C_T$) method was used for relative quantification of gene expression as previously described.[8] The to *18S* normalized expression of *EVI1* in the cell line SKOV3 was used to define *EVI1* positive expressers, *i.e.* patients with an expression higher than 0.1 relative to the *EVI1* expression of SKOV3 were labeled *EVI1* positive.[9]

Analysis of recurrent gene mutations in AML

The presence of internal tandem duplication in the *FLT3* gene (*FLT3*-ITD), mutations in the *FLT3* tyrosine kinase domain (*FLT3*-TKD) and in the *NPM1* and *CEBPA* genes were determined as previously described.[10,11] Insertion and deletion mutations at codon 646 in the gene *ASXL1* were analyzed by Sanger sequencing using a proofreading polymerase as previously reported.[12] Additionally, in 83 patients genomic DNA at diagnosis was available for targeted amplicon sequencing of 54 recurrently mutated genes in myeloid malignancies on the MiSeq platform (Illumina, San Diego, CA, USA) using the TruSight Myeloid Sequencing Panel (Illumina) as previously described.[13] Analyzed genes were *ABL1*, *ATRX*, *ASXL1*, *BCOR*, *BCORL1*, *BRAF*, *CALR*, *CBL*, *CBLB, CBLC*, *CDKN2A*, *CSF3R*, *CUX1*, *DNMT3A*, *ETV6*/*TEL*, *EZH2*, *FBXW7*, *GATA1*, *GATA2*, *GNAS*, *HRAS*, *IDH1*, *IDH2*, *IKZF1*, *JAK2*, *JAK3*, *KDM6A*, *KIT*, *KRAS*, *MLL*/*KMT2A*, *MPL*, *MYD88*, *NOTCH1*, *NRAS*, *PDGFRA*, *PHF6*, *PTEN*, *PTPN11*, *RAD21*, *RUNX1*, *SETBP1*, *SF3B1*, *SMC1A*, *SMC3*, *SRSF2*, *STAG2*, *TET2*, *TP53*, *U2AF1*, *WT1*, *ZRSR2*. Patients were grouped according to the European LeukemiaNet (ELN) 2017 classification.[14]

Evaluation of the leukemia initiating cell burden

The median normalized *GPR56* expression was used to define high (n=107) and low (n=106) *GPR56* expressers. For the CD34+/CD38- cell burden at diagnosis, the previously used 6% cut-off defined patients with a high (n=26) and a low (n=154) CD34+/CD38- cell burden.[1]

Definition of complete remission

CR was defined as the presence of <5% of blasts in BM, neutrophils >1.0 x 10^9/L, platelets >100 x 10^9/L, absence of blasts with Auer rods, independence of blood transfusion and no extramedullary disease.[15] CRi was defined as CR with platelets <100 x 10^9/L or neutrophils <1.0 x 10^9/L. The presence of CR or CRi was confirmed within 28 days prior to HSCT by BM and peripheral blood analysis.

Definition of clinical end points and statistical analysis

All statistical analyses were performed using the R statistical software platform (version 3.4.3).[16] Cumulative incidence of relapse (CIR) was calculated from HSCT to relapse and leukemia free survival (LFS) was calculated from HSCT to death from any cause. For LFS, survival estimates were calculated using the Kaplan-Meier method and groups were compared with the log-rank test. CIR was calculated considering the competing risk (non-relapse mortality) using the Fine and Gray model. Associations of *GPR56* expression levels with baseline clinical, demographic, and molecular features were compared using the Kruskal-Wallis test and Fisher's exact test for continuous and categorical variables, respectively.

Multivariable analyses

We constructed a multivariable proportional hazard model for CIR and LFS to evaluate the impact of the leukemia initiating cell burden defined by the *GPR56* expression and the CD34+/CD38- cell burden at diagnosis by forward adjusting for other variables.[17] Additionally to the *GPR56* expression (high *vs.* low, median cut), the following variables were considered

for multivariable analyses: sex, hemoglobin count, platelet count, white blood cell count, blast percentages in peripheral blood and BM at diagnosis, age at diagnosis, AML type (*de novo vs.* secondary), ELN 2010 Genetic Groups, *FLT3*-TKD mutation status at diagnosis, *BAALC* expression at diagnosis, *MN1* expression at diagnosis, *EVI1* expression status at diagnosis, disease status at HSCT (CRi *vs.* CR), cytomegalovirus (CMV) status of recipient and donor (high risk [+/-] *vs.* all others), HLA match (antigen match *vs.* mismatch), HLA donor type (related *vs.* unrelated) and sex of the donor (female into male *vs.* all others). Variables significant at α=.20 in univariable analyses were considered for multivariable analyses. Hazard/odds ratios with their corresponding 95% confidence intervals (CI) were obtained using forward selection analysis and indicated for every significant prognostic factor.

GPR56 in patients with a low CD34+/CD38- cell burden

Because of the prognostic impact of *GPR56* expression levels in patients with a low CD34+/CD38- cell burden at diagnosis, associations were also analyzed in this patients' subset (n=154). Except a lacking association with *JAK2* and only a trend for *RUNX1* mutation status – which are likely due to restricted patient numbers - similar clinical, cytogenetic and molecular associations were observed as in the entire set of patients (*Supplementary Table S2*).

Additional associations of a high *GPR56* expression

We observed a distinct immunophenotype at diagnosis in patients with a high *GPR56* expression (*Supplementary Table S1*). Analyzing BM mononuclear cells at diagnosis, patients with a high *GPR56* expression had a significantly higher expression of the immature antigen CD34 (*P*<.001) as well as a higher expression of antigens indicating T-cellular differentiation (*i.e.* CD2, *P*=.006; CD7, *P*=.002) and antigens indicating thrombocytic or erythrocytic differentiation (CD61, *P*=.002 and Glykophorin A, *P*=.02, respectively) but a lower expression of antigens indicating myeloid differentiation (*i.e.* CD15, *P*<.001; CD33, *P*<.001; CD38, *P*=.03; CD64, *P*=.002 and CD65, *P*<.001). All analyzed HSCT-related factors

did not significantly differ between patients with a high or a low *GPR56* expression (*Supplementary Table S1*).

First validation set: patients treated with an allogeneic HSCT

To further validate the prognostic significance of the *GPR56* expression levels in the context of the CD34+/CD38- cell burden at AML diagnosis we analysed a validation set of 96 AML patients that received allogeneic HSCT at our institution and were not included in the first set of patients described above. Conditioning regimens for this cohort were either MAC (n=20, as defined above), of reduced intensity (FLAMSA-based conditioning, n=17; Busulfan/Fludarabine, n= 5; or Fludarabine/Melphalan, n=2) or NMA (n=52, as defined above). Also in this cohort, a high CD34+/CD38- cell burden (6% cut) associated with higher CIR (*P*=.01, *Supplementary Figure S3A*) and by trend shorter LFS (*P*=.06, *Supplementary Figure S3B*). A high *GPR56* expression (median cut) associated with by trend higher CIR (*P*=.06, *Supplementary Figure S3C*) and shorter LFS (*P*=.05, *Supplementary Figure S3D*). Finally, there was a stepwise higher CIR (*P*=.04, *Supplementary Figure S3E*) and by trend shorter LFS (*P*=.08, *Supplementary Figure S3F*) for patients with low expression of both markers, patients with a low CD34+/CD38- cell burden, but high *GPR56* expression and patients with a high CD34+/CD38- cell burden.

Second validation set: patients treated with chemotherapy

We analysed a second validation set of 105 AML patients that received chemotherapy-based consolidation at our institution to validate the previously described impact of *GPR56* expression on outcome in chemotherapy-based consolidated AML patients. Similar to the results from our HSCT-treated cohort – although restricted due to the lower number of patients - we observed a separation of the survival curves according to the CD34+/CD38- cell burden (6% cut) for incidence of CR/CRi achievement (*P*=.14, *Supplementary Figure S4A*) and LFS (*P*=.20, *Supplementary Figure S4B*). Furthermore, a high *GPR56* expression (median cut) associated with a higher incidence of CR/CRi achievement (*P*=.006,

Supplementary Figure S4C) and a trend for shorter LFS (*P*=.08, *Supplementary Figure S4D*). Finally, we also observed a separation of the survival curves according to the CD34+/CD38- cell burden and the *GPR56* expression at diagnosis in patients consolidated with chemotherapy for both incidence of CR/CRi achievement (*P*=.05, *Supplementary Figure S4E*) and LFS (*P*=.08, *Supplementary Figure S4F*). Due to restricted patient numbers, no CIR analysis could be performed for this cohort.

Third validation set: TCGA dataset

We evaluated the TCGA dataset (from http://cancergenome.nih.gov) for the mRNA expression levels of *GPR56* as well as *CD34* and *CD38* (n=172). Consistent with our patient sets, a high *GPR56* expression (median cut) associated with a significantly shorter LFS (*P*=.03, *Supplementary Figure S5C*) and by trend shorter OS (*P*=.06, *Supplementary Figure S5D)* after diagnosis. Using RNA data for *CD34* and *CD38* expression levels, we divided the cohort into quartiles for both genes and patients with the highest expression of *CD34* (4th quartile) and the lowest expression of *CD38* (1st quartile) were defined as having high *CD34*^{high}/*CD38*^{low} expression (n=9). Also for these patients, we observed a trend for shorter LFS (*P*=.08, *Supplementary Figure S5A)* and significantly shorter OS (*P*=.01*, Supplementary Figure S5B).* Finally, combining all three parameters we divided the cohort into patients with a high *CD34*^{high}/*CD38*^{low} expression, patients with a low *CD34*^{high}/*CD38*^{low} expression but a high *GPR56* expression and patients with a low *CD34*^{high}/*CD38*^{low} expression and low *GPR56* expression. Again, a stepwise shorter LFS (*P*=.01, *Supplementary Figure S5E*) and OS (*P*=.02, *Supplementary Figure S5F*) was observed for patients with a high expression of one or both markers.

Correlation of *GPR56* RT-qPCR results with GPR56 analysis by flow cytometry

In six patients, surface antigen GPR56 expression levels on mononuclear cells at diagnosis were evaluated using flow cytometry on the FACSCalibur flow cytometer and analyzed using the CELLQUEST software (Becton Dickinson). For labeling the GPR56 antibody (Santa Cruz

Biotechnology, Dallas, Texas, USA) was used. Results were correlated with mRNA *GPR56/ABL1* expression levels by RT-qPCR. GPR56 surface antigen expression levels and mRNA expression correlated well (R^2=.80, *Supplementary Figure S1*).

References

1. Jentzsch M, Bill M, Nicolet D, et al. Prognostic impact of the bone marrow CD34+/CD38- cell burden at diagnosis in acute myeloid leukemia patients undergoing allogeneic stem cell transplantation. *Am J Hematol*. 2017;**92**(4):388-396.

2. Büchner T, Schlenk RF, Schaich M, et al. Acute Myeloid Leukemia (AML): different treatment strategies versus a common standard arm-combined prospective analysis by the German AML Intergroup. *J Clin Oncol*. 2012;**30**(29):3604-3610.

3. Stone RM, Mandrekar SJ, Sanford BL, et al. Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a *FLT3* Mutation. *N Engl J Med*. 2017;**377**(5):454-464.

4. Creutzig U, Zimmermann M, Bourquin JP, et al. Randomized trial comparing liposomal daunorubicin with idarubicin as induction for pediatric acute myeloid leukemia: results from Study AML-BFM 2004. *Blood*. 2013;**122**(1):37-43.

5. Niederwieser D, Hoffmann VS, Pfirrmann M, et al. Comparison of Treatment Strategies in Patients over 60 Years with AML: Final Analysis of a Prospective Randomized German AML Intergroup Study. [abstract]. *Blood*. 2016;128:1066. Abstract 1066.

6. Niederwieser D, Maris M, Shizuru JA, et al. Low-dose total body irradiation (TBI) and fludarabine followed by hematopoietic cell transplantation (HCT) from HLA-matched or mismatched unrelated donors and postgrafting immunosuppression with cyclosporine and mycophenolate mofetil (MMF) can induce durable complete chimerism and sustained remissions in patients with hematological diseases. *Blood.* 2003;**101**(4):1620-1629.

7. Glucksberg H, Storb R, Fefer A, et al. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HLA matched sibling donors. *Transplantation*. 1974;**18**(4): 295-304.

8. Mims A, Walker AR, Huang X, et al. Increased anti-leukemic activity of decitabine via AR-42-induced upregulation of miR-29b: a novel epigenetic-targeting approach in acute myeloid leukemia. *Leukemia*. 2013;**27**(4):871-878.

9. Gröschel S, Lugthart S, Schlenk RF, et al. High *EVI1* Expression Predicts Outcome in Younger Adult Patients With Acute Myeloid Leukemia and Is Associated With Distinct Cytogenetic Abnormalities. *J Clin Oncol*. 2010;**28**(12):2101-2107.

10. Bill, M, Jentzsch, M, Grimm, J, et al. Prognostic impact of the European LeukemiaNet standardized reporting system in older AML patients receiving stem cell transplantation after non-myeloablative conditioning. *Bone Marrow Transplantation*. 2017;**52**(6):932-935.

11. Benthaus T, Schneider F, Mellert G, et al. Rapid and sensitive screening for *CEBPA* mutations in acute myeloid leukaemia. *Br J Haematol*. 2008;**143**(2):230–239.

12. Metzeler KH, Becker H, Maharry K, et al. *ASXL1* mutations identify a high-risk subgroup of older patients with primary cytogenetically normal AML within the ELN favorable genetic category. *Blood*. 2011;**118**(26):6920–6929.

13. Grimm J, Bill M, Jentzsch M, et al. Clinical impact of clonal hematopoiesis in acute myeloid leukemia patients receiving allogeneic transplantation. *Bone Marrow Transplant.* 2018 Nov 30. doi: 10.1038/s41409-018-0413-0.

14. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;**129**(4):424-447.

15. Döhner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood.* 2010;**115**(3):453-474.

16. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2017. URL https://www.R-project.org/.

17. Scrucca L, Santucci A, Aversa F. Regression modeling of competing risk using R: an in depth guide for clinicians. Bone Marrow Transplantation. 2010;**45**(9):1388–1395.

Supplementary Tables

Supplementary Table S1. Additional characteristics of HSCT treated AML patients according to the *GPR56* expression at diagnosis, median cut, n=213.

Characteristic	All patients (n=213)	low *GPR56* expression (n = 106)	high *GPR56* expression (n = 107)	*P*
Morphologic information				
FAB classification, n (%)				
M0	7	1 (1)	6 (7)	0.12
M1	18	5 (6)	13 (15)	0.08
M2	92	44 (52)	48 (56)	0.65
M4	14	8 (9)	6 (7)	0.59
M4eo	13	9 (11)	4 (5)	0.16
M5	9	15 (18)	3 (4)	0.003
M6	2	2 (2)	2 (2)	1
M7	5	1 (1)	4 (5)	0.37
Flow Cytometry				
BM CD2 expression, %				0.006
Median	13	11	18	
Range	1-93	1-92	2-93	
BM CD7 expression, %				0.002
Median	16	11	21	
Range	1-94	2-94	1-90	
BM CD11b expression, %				0.78
Median	15	14	17	
Range	0.5-93	0.5-93	1-80	
BM CD13 expression, %				0.95
Median	61	61	62	
Range	0.5-96	0.5-95	7-96	
BM CD14 expression, %				0.86
Median	2	2	2	
Range	0.5-74	0.5-74	0.5-44	
BM CD15 expression, %				<0.001
Median	30	47	24	
Range	2-94	2-94	2-91	
BM CD33 expression, %				<0.001
Median	67	79	57	
Range	1-98	1-98	2-98	
BM CD34 expression, %				<0.001
Median	35	18.0	42	
Range	0-97	0-97	0-93	
BM CD38 expression, %				0.03
Median	76	80	72	
Range	0.5-98	0.5-98	4-98	
BM CD45 expression, %				0.11
Median	94	95	93	
Range	6-100	13-100	6-99	
BM CD56 expression, %				0.42
Median	8	8	8	
Range	0-97	0-93	0.5-97	
BM CD61 expression, %				0.002
Median	4	3	6	
Range	0.5-39	0.5-30	0.5-39	
BM CD64 expression, %				0.002
Median	15	22	13	
Range	0-98	0-98	0.5-85	
BM CD65 expression, %				<0.001
Median	18	25	14	
Range	0.5-93	0.5-93	0.5-81	
BM CD117 expression, %				0.86
Median	40	41	38	
Range	0.5-96	0.5-95	3-96	

BM Glykophorin A expression, %				0.02
Median	8	6	10	
Range	0.5-90	0.5-90	0.5-61	
HSCT related information				
Age at HSCT, years				0.17
Median	60.6	59.2	61.5	
Range	19.6-76.2	20.6-75.3	19.6-76.2	
Conditioning, n (%)				0.29
NMA	154	73 (69)	81 (76)	
MAC	59	33 (31)	26 (24)	
Remission at HSCT, n (%)				0.70
CR1	151	77 (73)	74 (69)	
CR2	35	15 (14)	20 (19)	
CRi	28	14 (13)	14 (13)	
HLA antigen match, n (%)				0.50
Matched	170	83 (78)	87 (82)	
Mismatched	43	24 (22)	19 (18)	
Donor, n (%)				0.74
Related	45	21 (20)	24 (22)	
Unrelated	168	85 (80)	83 (78)	
Female into male, n (%)				1
Absent	181	90 (86)	91 (87)	
Present	29	15 (14)	14 (13)	
aGvHD, n (%)				0.41
Absent	137	70 (77)	67 (71)	
Present	48	21 (23)	27 (29)	
cGvHD, n (%)				0.40
Absent	48	21 (29)	27 (33)	
Limited	27	10 (14)	17 (20)	
Extended	80	41 (56)	39 (47)	

Abbreviations: aGvHD, acute graft-versus-host disease; BM, bone marrow; CD, cluster of differentiation; cGvHD, chronic graft-versus-host disease; CR, complete remission; CRi, complete remission with incomplete peripheral recovery; FAB, French American British classification; HSCT, hematopoietic stem cell transplantation; HLA, human leukocyte antigen; MAC, myeloablative conditioning; NMA, non-myeloablative conditioning.

Supplementary Table S2. Clinical and genetic characteristics of AML patients with a low CD34+/CD38- cell burden at diagnosis treated with HSCT according to the *GPR56* expression at diagnosis, median cut, n=154.

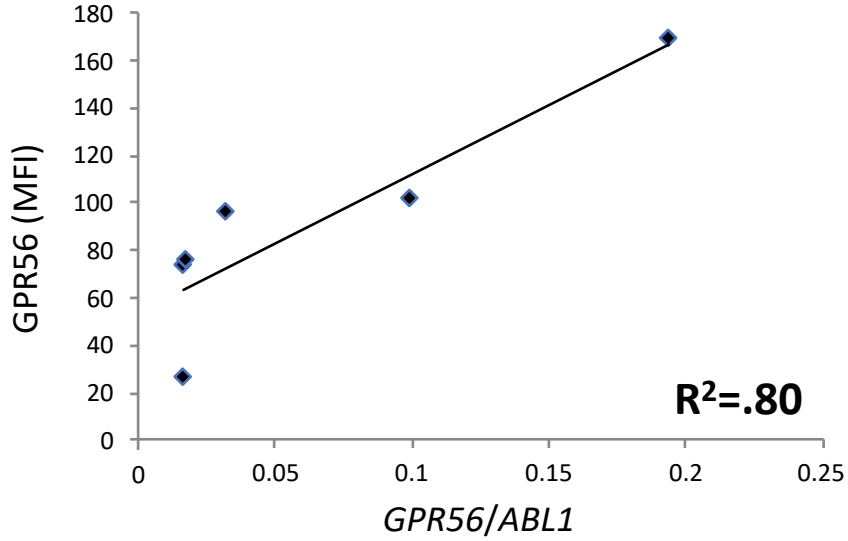
Characteristic	All patients (n=154)	Low *GPR56* expression (n=84)	High *GPR56* expression (n=70)	*P*
Clinical Characteristics				
Age at diagnosis, years				0.28
Median	59.9	59.0	62.0	
Range	14.5-74.8	14.5-74.8	19.5-73.7	
Sex, n (%)				0.87
Male	73	39 (46)	34 (49)	
Female	81	45 (54)	36 (51)	
Disease origin, n (%)				0.06
De novo	117	69 (82)	48 (69)	
Secondary	37	15 (18)	22 (31)	
Hemoglobin, g/dL				0.15
Median	8.9	9.1	8.6	
Range	4.5-15.7	4.5-15.7	4.5-14.9	
Platelet count, x 10^9/L				0.35
Median	65	65	79	
Range	1-327	1-327	3-287	
WBC count, x 10^9/L				0.002
Median	7.2	11	4	
Range	0.7-385	0.9-385	0.7-177	
Percentage of blood blasts, %				0.18
Median	23	26	17	
Range	0-98	0-97	0-98	
Percentage of BM blasts, %				0.96
Median	57	60	55	
Range	3-95	20-95	3-95	
BM CD34+/CD38- burden, %				0.008
Median	0.4	0.1	0.5	
Range	0-5	0-5	0-5	
Genetic Characteristics				
Karyotype, n (%)				0.25
Abnormal	78	38 (49)	40 (59)	
Normal	68	40 (51)	28 (41)	
Monosomal karyotype, n (%)				0.01
Absent	136	77 (99)	59 (88)	
Present	9	1 (1)	8 (12)	
Complex karyotype, n (%)				0.07
Absent	129	73 (94)	56 (84)	
Present	16	5 (6)	11 (16)	
del5/del(5q), n (%)				0.09
Absent	145	81 (99)	64 (93)	
Present	6	1 (1)	5 (7)	
del7/del(7q), n (%)				0.02
Absent	134	78 (95)	56 (82)	
Present	16	4 (5)	12 (18)	
CBF AML, n (%)				0.02
Absent	129	64 (82)	65 (96)	
Present	17	14 (18)	3 (4)	
Trisomy 8, n (%)				0.80
Absent	133	73 (89)	60 (87)	
Present	18	9 (11)	9 (13)	
ELN 2017 Genetic Group, n (%)				<0.001
Favorable	48	41 (69)	7 (18)	
Intermediate	12	3 (5)	9 (23)	
Adverse	38	15 (25)	23 (59)	
***NPM1*, n (%)**				<0.001
Wild-type	113	52 (62)	61 (87)	
Mutated	41	32 (38)	9 (13)	

***FLT3*-ITD, n (%)**				0.48
Absent	124	70 (83)	54 (77)	
Present	30	14 (17)	16 (23)	
***NPM1* mut/no *FLT3*-ITD, n (%)**				<0.001
Absent	128	62 (74)	66 (94)	
Present	26	22 (26)	4 (6)	
***CEBPA*, n (%)**				0.19
Wild-type	126	65 (88)	61 (94)	
Single mutated	9	5 (7)	4 (6)	
Double mutated	4	4 (5)	0 (0)	
***DNMT3A*, n (%)**				0.24
Wild-type	44	21 (68)	22 (79)	
Mutated	15	10 (32)	6 (21)	
***FLT3-TKD*, n (%)**				0.48
Wild-type	128	73 (88)	55 (83)	
Mutated	21	10 (12)	11 (17)	
***IDH1*, n (%)**				0.49
Wild-type	49	27 (87)	22 (79)	
Mutated	10	4 (13)	6 (21)	
***IDH2*, n (%)**				1
Wild-type	50	26 (84)	24 (86)	
Mutated	9	5 (16)	4 (14)	
***JAK2*, n (%)**				0.22
Wild-type	57	31 (100)	26 (93)	
Mutated	2	0 (0)	2 (7)	
***RUNX1*, n (%)**				0.07
Wild-type	50	29 (94)	21 (75)	
Mutated	9	2 (6)	7 (25)	
***TP53*, n (%)**				1
Wild-type	54	28 (90)	26 (93)	
Mutated	5	3 (10)	2 (7)	
***ASXL1*, n (%)**				0.48
Wild-type	50	25 (81)	25 (89)	
Mutated	9	6 (19)	3 (11)	
***EVI1* expression status, n (%)**				0.005
Negative	92	55 (95)	37 (76)	
Positive	15	3 (5)	12 (25)	
***MN1/ABL1* expression**				<0.001
Median	0.11	0.02	0.23	
Range	0.00-75.15	0.00-1.63	0.01-59.67	
***BAALC/ABL1* expression**				<0.001
Median	0.06	0.01	0.21	
Range	0.00-56.31	0.00-2.41	0.00-56.31	

Abbreviations: ABL1, Abelson murine leukemia viral oncogene homolog 1 gene; AML, acute myeloid leukemia; BAALC, brain and acute leukemia cytogenetic gene; CBF, core-binding factor; CEBPA, CCAAT/enhancer-binding protein alpha gene; del, deletion; DNMT3A, DNA-methyltransferase 3A gene; ELN, European LeukemiaNet; EVI1, ecotropic viral integration site gene; FLT3-ITD, internal tandem duplication of the FLT3 gene; FLT3-TKD, tyrosine kinase domain of the FLT3 gene; HSCT, hematopoietic stem cell transplantation; IDH1, isocitrat dehydrogenase 1 gene; IDH2, isocitrat dehydrogenase 2 gene; JAK2, janus kinase 2 gene; MN1, meningioma-1 gene; NPM1, nucleophosmin-1 gene; RUNX1, Runt-related transcription factor 1 gene; TP53, tumor protein 53; WBC, white blood cell.

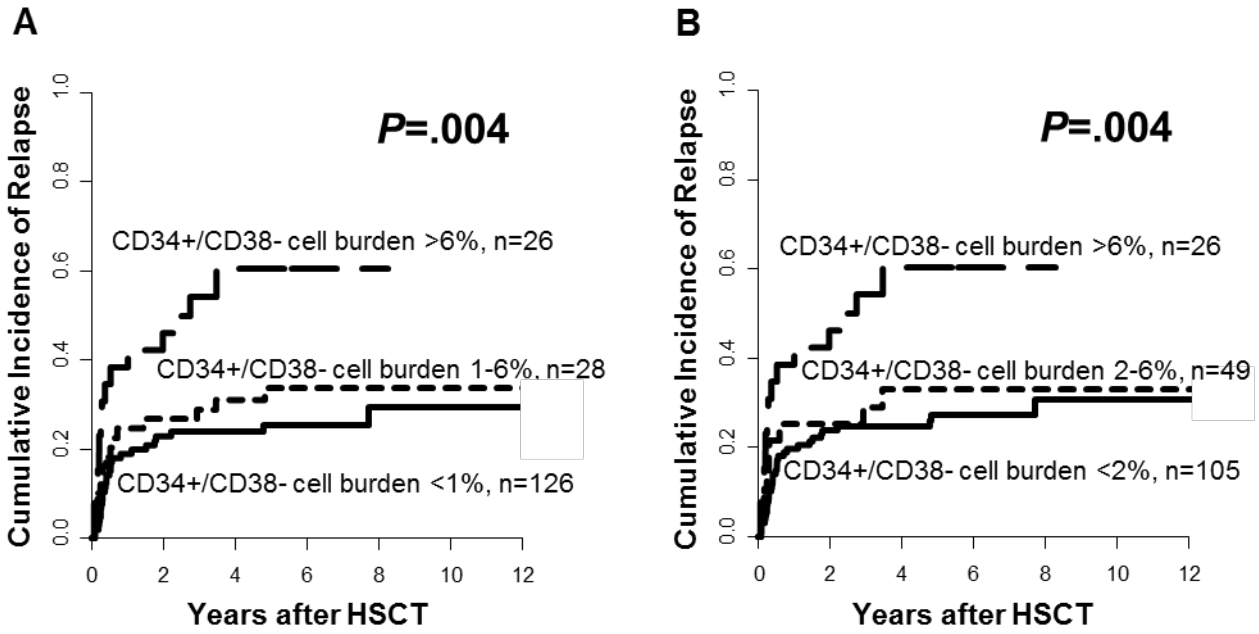
Supplementary Figures

Supplementary Figure S1



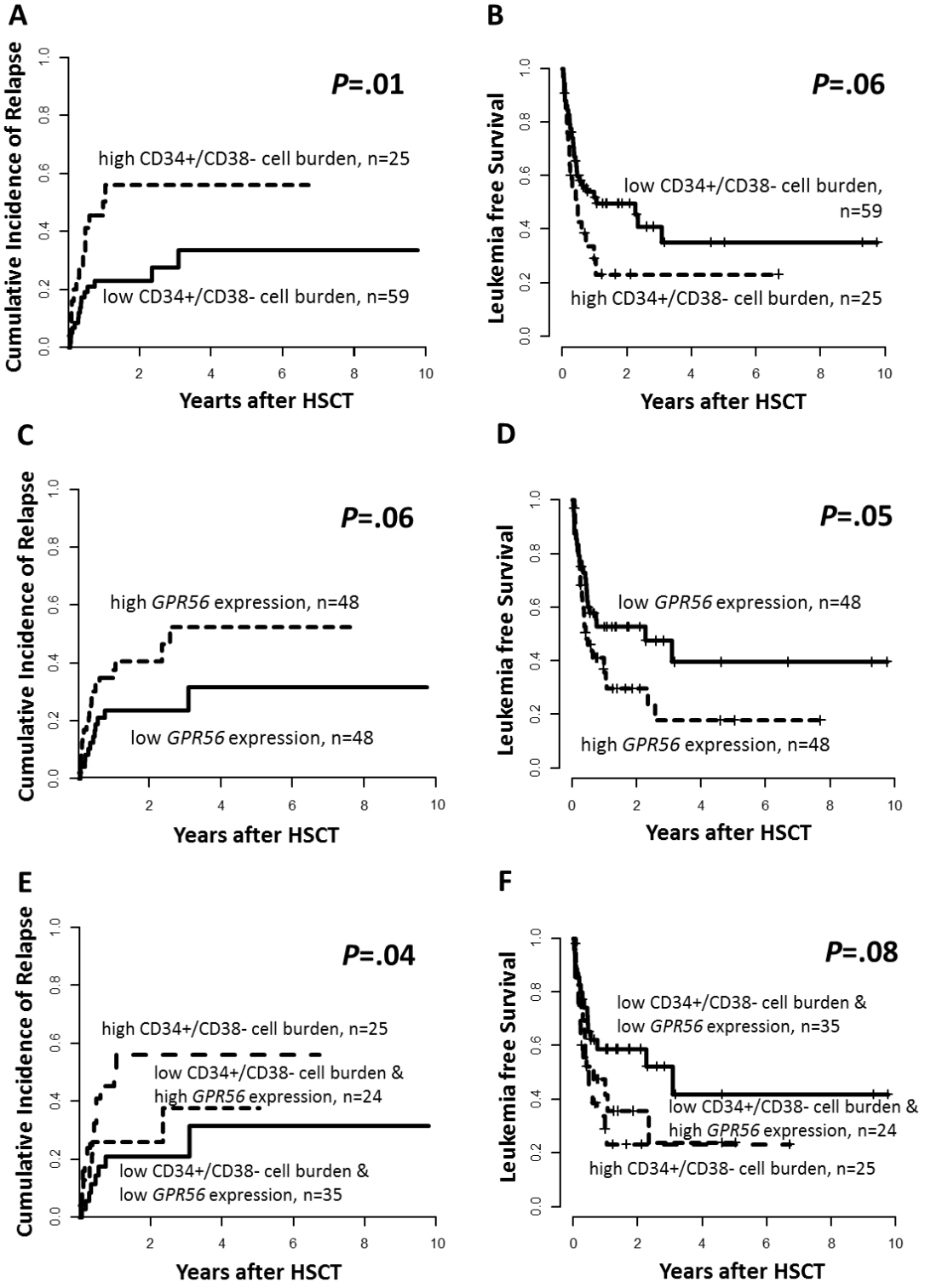
Supplementary Figure S1: Correlation of the GPR56 protein surface expression (assessed by flow cytometry) and *GPR56* mRNA expression (assessed by RT-qPCR) in diagnostic bone marrow samples from six AML patients. MFI, mean fluorescence intensity.

Supplementary Figure S2



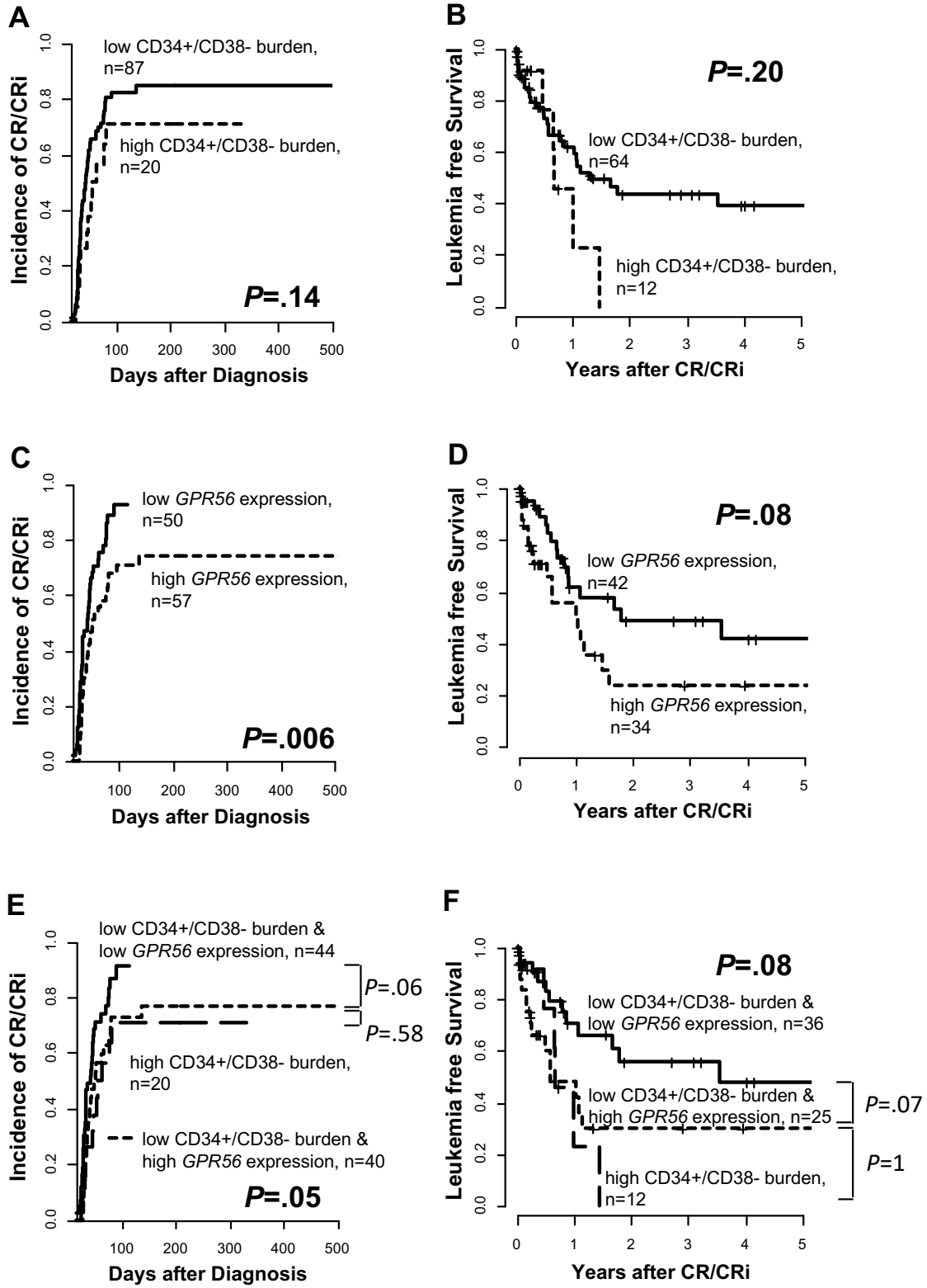
Supplementary Figure S2: Cumulative incidence of relapse according to the CD34+/CD38- cell burden at diagnosis for patients in the initial set (n=180). (**A**) 1% and 6% cut and (**B**) 2% and 6% cut.

Supplementary Figure S3



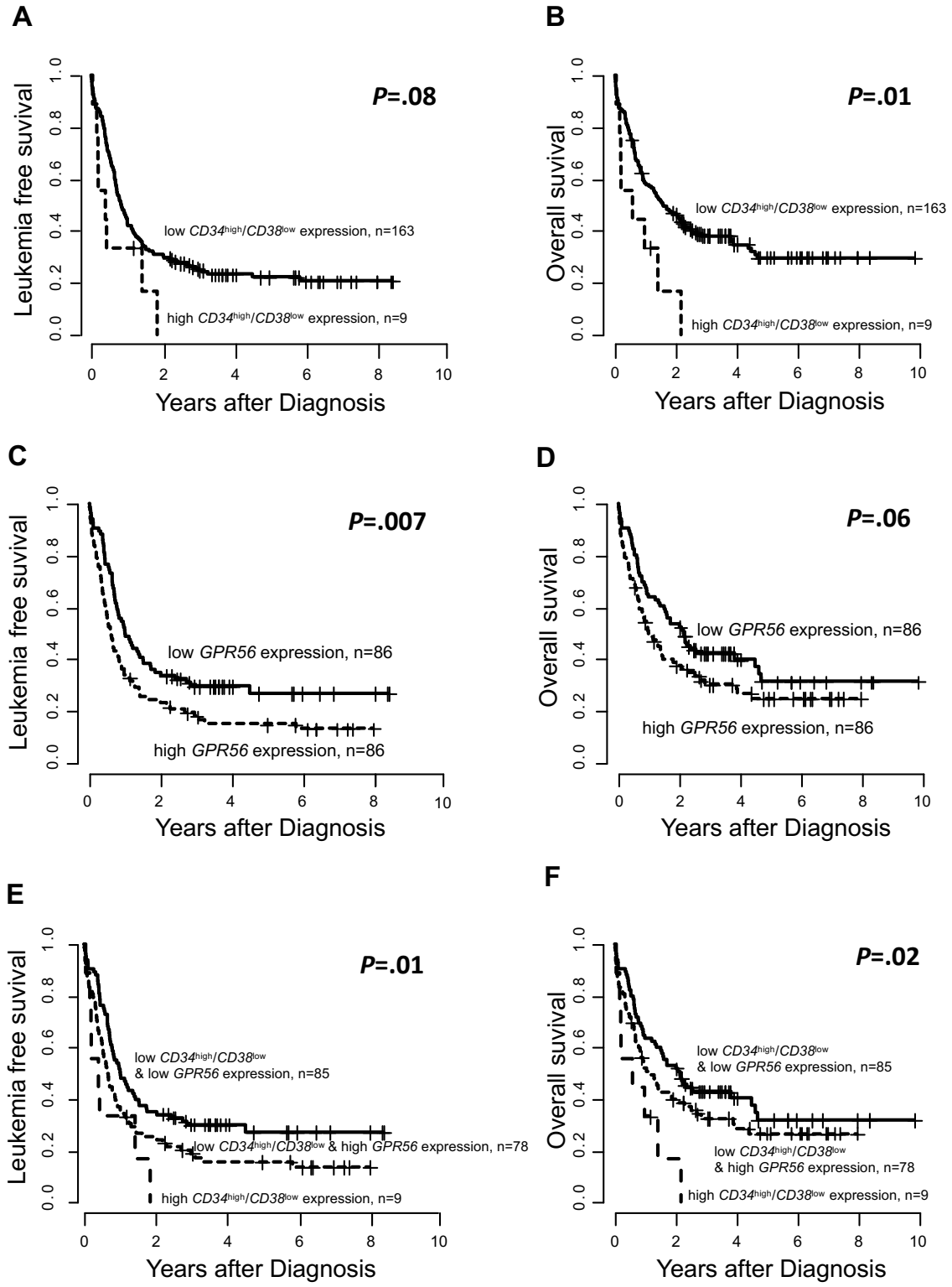
Supplementary Figure S3: First validation set of AML patients treated with HSCT (n=96). **(A)** Cumulative Incidence of relapse and **(B)** Leukemia free survival according to the CD34+/CD38- cell burden, high *vs.* low, 6% cut, **(C)** Cumulative incidence of relapse and **(D)** Leukemia free survival according to the *GPR56* expression, high *vs.* low, median cut and **(E)** Cumulative incidence of relapse and **(F)** Leukemia free survival according to the CD34+/CD38- cell burden and the *GPR56* expression.

Supplementary Figure S4



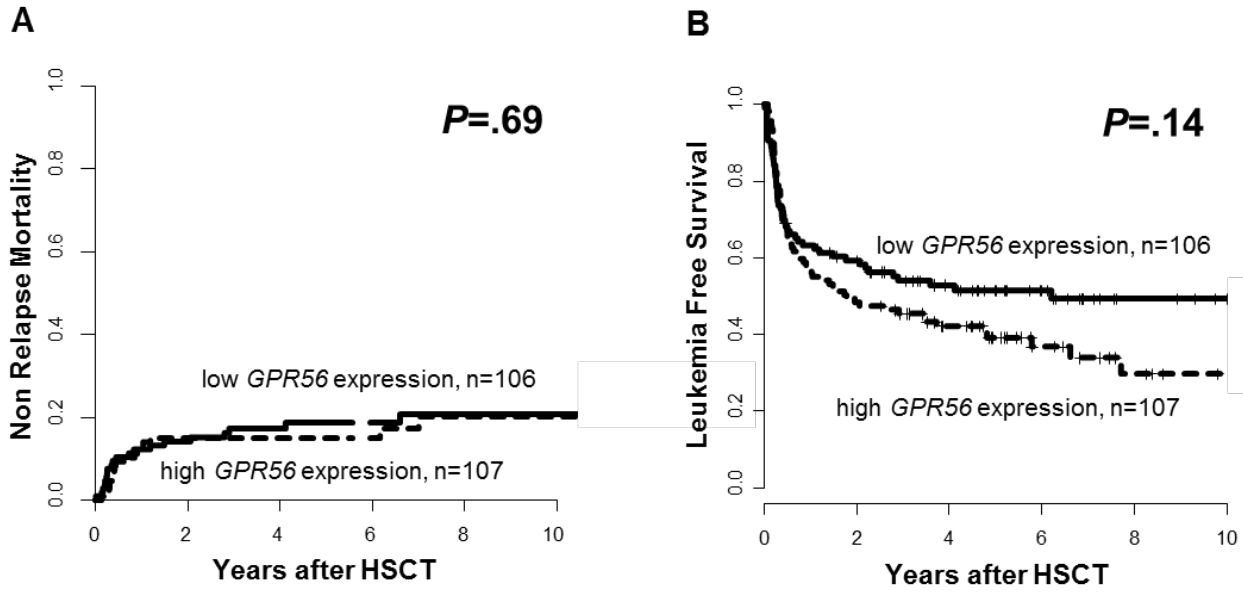
Supplementary Figure S4: Second validation set of AML patients treated with chemotherapy (n=105). **(A)** Cumulative incidence of CR/CRi achievement and **(B)** Leukemia free survival according to the CD34+/CD38- cell burden, high *vs.* low, 6% cut, **(C)** Cumulative incidence of CR/CRi achievement and **(D)** Leukemia free survival according to the *GPR56* expression, high *vs.* low, median cut and **(E)** Cumulative incidence of CR/CRi achievement and **(F)** Leukemia free survival according to the CD34+/CD38- cell burden and the *GPR56* expression.

Supplementary Figure S5



Supplementary Figure S5: Third validation set of AML patients using the TCGA dataset (n=172). **(A)** Leukemia free survival and **(B)** Overall survival according to the *CD34* and *CD38* mRNA expression (patients with *CD34* expression in the 4th quartile and *CD38* expression in the 1st quartile *vs.* all others) **(C)** Leukemia free survival and **(D)** Overall survival according to the *GPR56* expression (high *vs.* low, median cut) and **(E)** Leukemia free survival and **(F)** Overall survival according to the *CD34*^{high}/*CD38*^{low} and the *GPR56* expression.

Supplementary Figure S6



Supplementary Figure S6: Non-relapse mortality **(A)** and Leukemia free survival **(B)** according to the *GPR56* expression at diagnosis in the initial patient set, high *vs.* low, median cut (n=213).