

University of Groningen

A proteomics and transcriptomics approach to identify leukemic stem cell (LSC) markers

Bonardi, Francesco; Fusetti, Fabrizia; Deelen, Patrick; van Gosliga, Djoke; Vellenga, Edo; Schuringa, Jan Jacob

Published in:
Molecular & Cellular Proteomics

DOI:
[10.1074/mcp.M112.021931](https://doi.org/10.1074/mcp.M112.021931)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2013

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bonardi, F., Fusetti, F., Deelen, P., van Gosliga, D., Vellenga, E., & Schuringa, J. J. (2013). A proteomics and transcriptomics approach to identify leukemic stem cell (LSC) markers. *Molecular & Cellular Proteomics*, 12(3), 626-637. <https://doi.org/10.1074/mcp.M112.021931>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A Proteomics and Transcriptomics Approach to Identify Leukemic Stem Cell (LSC) Markers*[§]

Francesco Bonardi[‡], Fabrizia Fusetti[§], Patrick Deelen[‡], Djoke van Gosliga[‡], Edo Vellenga[‡], and Jan Jacob Schuringa[‡]¶

Interactions between hematopoietic stem cells and their niche are mediated by proteins within the plasma membrane (PM) and changes in these interactions might alter hematopoietic stem cell fate and ultimately result in acute myeloid leukemia (AML). Here, using nano-LC/MS/MS, we set out to analyze the PM profile of two leukemia patient samples. We identified 867 and 610 unique CD34⁺ PM (-associated) proteins in these AML samples respectively, including previously described proteins such as CD47, CD44, CD135, CD96, and ITGA5, but also novel ones like CD82, CD97, CD99, PTH2R, ESAM, MET, and ITGA6. Further validation by flow cytometry and functional studies indicated that long-term self-renewing leukemic stem cells reside within the CD34⁺/ITGA6⁺ fraction, at least in a subset of AML cases. Furthermore, we combined proteomics with transcriptomics approaches using a large panel of AML CD34⁺ ($n = 60$) and normal bone marrow CD34⁺ ($n = 40$) samples. Thus, we identified eight subgroups of AML patients based on their specific PM expression profile. GSEA analysis revealed that these eight subgroups are enriched for specific cellular processes. *Molecular & Cellular Proteomics* 12: 10.1074/mcp.M112.021931, 626–637, 2013.

Acute myeloid leukemia (AML)¹ is a disease characterized by an increase of immature myeloid blasts in the bone marrow as a consequence of the loss of normal differentiation and proliferation of hematopoietic progenitor cells (1, 2). The cancer stem cell (CSC) model (3–6) suggests that AML is maintained by a rare population of leukemic stem cells that are thought to be relatively quiescent, therapy resistant, and frequently the cause of relapse of disease. The interaction with

the surrounding microenvironment in the bone marrow is very important for the regulation of hematopoietic stem cell fate, and probably also of leukemic stem cells (LSCs) (7). Consequently, differential expression of proteins at the plasma membrane level could account for the specific interactions of leukemic cells with their niche. Therefore, the characterization of the plasma membrane proteome of LSCs is fundamental to further unravel the biology of leukemia development. In addition, a better understanding of the membrane proteome features could contribute to improved identification, isolation, and targeting of LSCs.

It is unclear whether there is a common plasma membrane protein signature that generally defines AML, or whether subtypes of leukemia can be identified based on the expression of specific plasma membrane proteins. From a cytogenetic standpoint AML is a very heterogeneous disease with different levels of classification (8). Leukemic cells often carry several recurring mutations, either as point mutations, insertions, deletions, gene rearrangements, and/or chromosomal translocations (8, 9). Deep sequencing technology has revealed, and will most likely continue to reveal, the occurrence of many more mutations in AML (10, 11). This diversity challenges even further the search for diagnostic factors. It has been recently shown that gene expression profiling is a valid approach in determining AML signatures and prognostic factors (12, 13), especially when it is performed on the CD34⁺ cell population (14) or on LSC-containing cell populations as defined by engraftment in xenograft models (15). Distinct subgroups could indeed be identified based on transcriptome data. However, it will still be necessary to verify whether these transcriptome changes are also translated to changes at the protein level, and whether unique plasma membrane proteins exist that might aid in the identification of distinct subgroups of AML.

Over the last two decades, the advances in mass-spectrometry-based technologies have allowed the identification and characterization of diagnostic markers in complex biological samples (16–18). In our study we used liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS) to analyze the plasma membrane proteome of two different AML samples, separated into leukemic stem-cell enriched CD34⁺ and leukemic stem cell-depleted CD34⁻ fractions (19), to identify specific plasma membrane-associated signatures.

From the [‡]Department of Experimental Hematology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9700 RB, Groningen, The Netherlands, [§]Department of Biochemistry and Netherlands Proteomics Centre, Groningen Biological Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Received June 29, 2012, and in revised form, November 19, 2012
Published, MCP Papers in Press, December 11, 2012, DOI 10.1074/mcp.M112.021931

¹ The abbreviations used are: AML, acute myeloid leukemia; LC-MS/MS, liquid chromatography-tandem MS; LSC, leukemic stem cells; CSC, cancer stem cells; GSEA, gene set enrichment analysis; FACS, fluorescence-activated cell sorting; CV, column volume; IG, information gain; NBM, normal bone marrow; PM, plasma membrane.

Following this approach a CD34⁺-specific plasma membrane protein profile was identified, which included putative AML markers such as CD47, ITGα6, CD44, CD82, and CD135. We then correlated the proteomics results with gene expression profiles of a large cohort of AML CD34⁺ and normal CD34⁺ samples, which resulted in the classification of eight AML subgroups, associated to a specific PM expression profile. Subsequent gene set enrichment analysis (GSEA) revealed that each of the identified subgroups was characterized by specific cellular processes and prognosis.

EXPERIMENTAL PROCEDURES

Isolation of AML CD34⁺ and CD34⁻ cells, MS5 Cocultures, and FACS Analysis—AML blasts from peripheral blood cells or bone marrow cells from untreated patients with AML were studied after informed consent was obtained in accordance with the Declaration of Helsinki, and the protocol was approved by the Medical Ethical Committee. AML mononuclear cells were isolated by density gradient centrifugation, and CD34⁺ cells were stained using CD34-PE antibody (BD Biosciences, San Jose, CA, USA) and selected by sorting on a MoFlo (DakoCytomation, Carpinteria, CA, USA). AML cocultures were performed on MS5 stromal cells as described previously (19, 20). All fluorescence-activated cell sorter (FACS) analyses were performed on a FACScalibur (Becton-Dickinson [BD], Alphen a/d Rijn, the Netherlands) and the data were analyzed using WinList 3D (Verity Software House, Topsham, USA) or FlowJo (Tree Star, Oregon, USA) software. Cells were incubated with antibodies at 4 °C for 30 min. Antibodies against CD34, CD38, CD135, CD47, and ITGA6 were obtained from BD Biosciences (Breda, The Netherlands), antibodies against CD96 and PTH2R were obtained from Santa Cruz (Santa Cruz Biotechnology, Santa Cruz, CA, USA) and second goat-anti-rabbit-FITC antibodies (used for PTH2R stains) were obtained from Invitrogen (Breda, The Netherlands).

Membrane Protein Purification—AMLCD34⁺ and AMLCD34⁻ cell populations were sorted with a MoFlo-XDP sorter from Beckman-Coulter. The cell suspension was centrifuged and the pellet was frozen in liquid N₂. The cells were quickly thawed and mechanically lysed with six passes through a 301/2 gauge needle and diluted twofold in lysis buffer (50 mM Tris/HCl pH 8, 250 mM sucrose, 2 mM EDTA, 0.2 mM MgCl₂, and protease inhibitor mixture). The total cell lysate was depleted of the nuclear fraction by a low spin centrifugation step (1000 × g for 10 min at 4 °C), and the supernatant was layered on top of a 60% sucrose cushion and centrifuged for 2 h at 100,000 × g at 4 °C with a TLA100.1 rotor. The top layer was sixfold diluted with 50 mM Tris/HCl, pH 8, and centrifuged for 1 h at 80,000 × g at 4 °C. The pellet was resuspended in 100 mM Na₂CO₃ pH 8.5, 0.1% SDS and subsequently incubated with TCEP (tris(2-carboxyethyl)phosphine) for 1 h at 60 °C, followed by the addition of methyl methanethiosulfonate for 10 min at room temperature, to reduce and modify cysteine residues. One microgram of Trypsin Gold (mass spectrometry grade, Promega) was added and the reaction incubated overnight at 37 °C. The sample was then treated with 1 Unit of PNGase F (SIGMA) for 2 h at 37 °C followed by a second trypsin digestion overnight at 37 °C. The tryptic peptides were acidified with 5% formic acid and cleaned with C18 TopTip (Glygen) according to the manufacturer's instructions eluting with 80% methanol in 5% formic acid.

Strong Cation Exchange Fractionation—Off-line peptide pre-fractionation by strong cation-exchange (SCX) was performed on a silica-based polysulfoethyl aspartamide column (200 × 2.1 mm, 200 Å, Cat.: 202SE0502 PolyLC Inc., Columbia USA) mounted on an Ettan-MDLC system (Amersham Biosciences AB, Uppsala, Sweden), and

run at a flow rate of 200 μl/min. The pH of the sample was adjusted to 3.0 with phosphoric acid before separation. Gradient solutions A: 10 mM triethylammonium phosphate, pH 2.7, 25% acetonitrile; B: 10 mM triethylammonium phosphate, pH 2.7, 25% acetonitrile, 1 M KCl. Gradient conditions: column equilibration with five column volumes (CV) (1 CV = 0.7 ml) of 100% A. After sample loading, the column was washed with 10 CV at 100% A. Peptides were eluted stepwise in: (1) 0 to 5% B in 5 CV; (2) followed by 12 to 30% B in 10 CV; and (3) 24–60% B in 5 CV. Elution fractions were collected every 120 s in a 96-well plate and subsequently dried in a vacuum centrifuge. Eluted peptides were concentrated to ~40 μl in a vacuum centrifuge and diluted 1:2 with 0.2% trifluoroacetic acid. Depending on the complexity, either separate fractions or pools of two fractions were analyzed by RP-LC MS/MS.

Reverse Phase Liquid Chromatography and ESI-MS—SCX fractions resuspended in 5% formic acid were separated on a capillary column (C18 PepMap 300, 75 μm × 250 mm, 3-μm particle size, Dionex, Amsterdam, The Netherlands) mounted in line with a precolumn (EASY-Column C18, 100 μm × 20 mm, 5-μm particle size, Thermo Scientific, Bremen, Germany) on a Proxeon Easy-LC system (Proxeon Biosystems, Odense, Denmark). Solutions of 0.1% formic acid in water and a 0.1% formic acid in 100% acetonitrile were used as the mobile phases. A gradient from 2 to 35% acetonitrile was performed in 140 min at a flow rate of 200 nl/min. Eluted peptides were analyzed using a linear ion trap-Orbitrap hybrid mass spectrometer (LTQ-Orbitrap, Thermo Scientific). The LTQ was operated in data dependent mode in which one full MS scan was followed by an MS/MS scan with dynamic exclusion set to: 1 repeat count, 30 s exclusion duration, and 500 exclusion list size. MS scans were acquired in the Orbitrap in the range from 250 to 2000 m/z, with a resolution of 60,000 (full-width at half-maximum). The seven most intense ions per scan were submitted to MS/MS fragmentation (35% Normalized Collision Energy™) and detected in the linear ion trap. The capillary temperature was set at 200 °C, spray voltage was 1.7 kV, and capillary voltage was 39.96 V. Each fraction was analyzed in triplicate with exclusion lists.

Database Searching—The MS raw data were submitted to Mascot (Version 2.1, Matrix Science, London, UK) and Sequest using the Proteome Discoverer 1.1 analysis platform (Thermo Scientific) and searched against ipi.HUMAN.v3.83 database proteome covering 91,464 entries. Peptide tolerance was set at 20 ppm and 2.0 Da for intact peptides and fragment ions respectively, using semitrypsin as protease specificity and allowing for up to two missed cleavages. Oxidation of methionine residues, deamidation of asparagine and glutamine, and MMTS modification of cysteines were specified as variable modifications. The MS/MS-based peptide and protein identifications were further validated with the program Scaffold (Version Scaffold_3.1.4, Proteome Software Inc., Portland, OR). Protein identifications were accepted if they could be established at greater than 50% and contained at least one identified peptide. X! Tandem (thegpm.org; version 2007.01.01.1) was set up to search a subset of the ipi.HUMAN.v3.83 database also assuming trypsin.

The data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche using the following hash: `ixv+b6vfwuUwT9OCMTxjkJNjE+0aF4klO8y/fkQ0sR4A1j1telRkrH38zfjGKGoJ/u0Vlg/s/0hb0FZSw3wCYU6it20AAAAAACZg = =`

Principal Components Analysis—Gene expression profiling of in total 147 samples was performed previously (14) using the Illumina HumanHT-12 Expression BeadChips. These 147 samples were divided as follows: 40 CD34⁺ NBM, 60 AML CD34⁺, and 47 AML CD34⁻ samples (see Supplemental File S2). All samples were corrected for background using Illumina GenomeStudio and then jointly forced to positive values, normalized, and transformed using the R packages Bioconductor (21) and Lumi (22). Probes with a detection p

value larger than 0.01 in all samples, as provided by GenomeStudio, were deleted. Log₂ transformation and quantile normalization were applied. As a measure of quality control we performed a principal component analysis on the correlation matrix of all 147 samples (23). The first component was removed from the data (24). To ensure reliability and reproducibility of the results we used multivariate permutations to determine the significance of our results using a false discovery rate of 0.4.

Information Gain—Information gain (IG) is a measurement of the expected reduction in information entropy in the presence of a known variable (25). More specifically, in our case we have used two sets of samples, AML CD34⁺ or NBM CD34⁺. This cell-type annotation contains certain information entropy, but without prior knowledge about these samples it would not be possible to classify them correctly. By making use of additional available information about these samples, in our case gene expression profiles, one can reduce the information entropy of the cell-type annotation, which aids in the classification of the cell-type of a sample. The IG is the reduction in information entropy when using the extra knowledge about these samples. In our case, we calculated the IG using the gene expression levels of PM proteins that were upregulated in AML CD34⁺ cells. The IG allowed us to prioritize genes based on their predictive value of the cell type.

RESULTS

Identification of the Plasma Membrane Proteome of Leukemic Stem Cell-Enriched Fractions of Primary Leukemia Patient Samples—To investigate the plasma membrane signature of leukemic stem cells, we selected cells from two patients: a poor risk AML patient (FAB M1, FLT3 ITD, NPM1wt; Inv(3q), -7, -10; designated AML1) and a myeloid blast crisis patient sample (CML patient that progressed to AML; FLT3wt, NPM1wt, Inv(16) and t(9;22); designated AML2). Both samples were sorted into CD34⁺ and CD34⁻ populations to obtain stem cell-enriched and stem cell-depleted fractions (3, 14, 19, 26). About five million cells could be sorted per sample and because of this relatively small amount of cells it was chosen to simplify membrane purification procedures, thus minimizing the loss of membrane proteins. Therefore, after cell lysis and a low centrifugation step, the sample preparation included just a sucrose cushion centrifugation to separate the membrane component from the nuclear and cytoplasmic fractions, as summarized in the scheme of Fig. 1A. Because of this technical approach, the resulting membrane enriched fractions were still a very complex mixture of cellular and subcellular membrane proteins. To deal with the high level of sample complexity and still obtain a comprehensive inventory of the plasma membrane proteome, we applied a multidimensional protein identification technology (MuDPIT) approach (27) combined with a high resolution LCMS/MS workflow. After protein digestion, the resulting tryptic peptide mixtures were eluted combining off-line a SCX chromatography with a Reverse Phase (RP) chromatography column directly coupled to an LTQ-Orbitrap mass spectrometer. The elution of each sample was carried out with a 4-hour gradient to increase the separation of the peptide peaks. Each SCX fraction was analyzed in three technical replicates with the use of incremental exclusion lists, which have been shown to increase the num-

ber of identified peptides in label-free systems (28) for highly complex samples. The MS/MS spectra were then searched against the ipi-HUMAN database using Mascot, Sequest, and X!Tandem to increase the confidence of identification probability. The results were further validated with the program Scaffold. The total number of proteins identified with at least one unique peptide was: 3504 for AML1 CD34⁺, 1162 for AML1 CD34⁻, 2580 for AML2 CD34⁺, and 4058 for AML2 CD34⁻ (Fig. 1B and Supplemental File S1). Gene ontology annotation indicated that 32% of the total amount of identified proteins was classified as “membrane,” of which 61% were plasma membrane proteins (Fig. 1C). In total 867 and 610 unique AML CD34⁺ PM-associated proteins (PM+) could be detected in AML1 and AML2, respectively, of which 619 and 386 were annotated as specific PM proteins (Fig. 1B and Supplemental File S1). Less PM proteins were detected in the CD34⁻ fraction of AML1 compared with AML2, although we do not know whether differences in heterogeneity in the CD34⁻ compartments or technical issues underlie these observations. As shown in the VENN diagram, some but limited overlap in plasma membrane (-associated) proteins was observed between the CD34⁺ and CD34⁻ fractions within each AML patient sample (Fig. 1B). Apparently, the leukemic stem cell-enriched CD34⁺ population is quite distinct from the leukemic stem cell-depleted CD34⁻ fraction in terms of its plasma membrane proteome composition. Moreover, some but limited overlap in CD34⁺ plasma membrane transcriptome was also observed between the two AML samples (Fig. 1D). This indicates that, as expected, there is considerable heterogeneity between the plasma membrane proteome of individual patients as well. Gene ontology annotation for biological processes, obtained by using the combined list of all identified AML CD34⁺ PM proteins revealed enrichment for processes such as cell adhesion, ion transport, cell migration, and cytoskeleton organization (Fig. 1E). A short list of identified AML CD34⁺ plasma membrane proteins is shown in Fig. 1F.

Identification of Leukemic Stem Cell Markers Using a Transcriptomics Approach—We determined the gene expression profile of AML and NBM samples using Illumina Bead Arrays. Transcriptomes of NBM CD34⁺ samples ($n = 40$) were compared with 60 AML CD34⁺ stem cell enriched, and 47-paired leukemic stem cell-depleted CD34⁻ samples (an overview of patient characteristics is provided in Supplemental File S2) (14). Genes that were more highly expressed in the AML CD34⁺ fraction were identified using a single sided Kruskal-Wallis U test, and significance was determined by multivariate permutation MP (29), which robustly limits the false discovery rate (30) of the performed analysis. Multivariate permutation exploits the correlation structure of the data and combines the low false positive rate of the Bonferroni correction with the high true positive rate of the Benjamini and Hochberg FDR correction. Within these AML CD34⁺-specific transcriptomes, GO annotation was used to select all proteins associated with

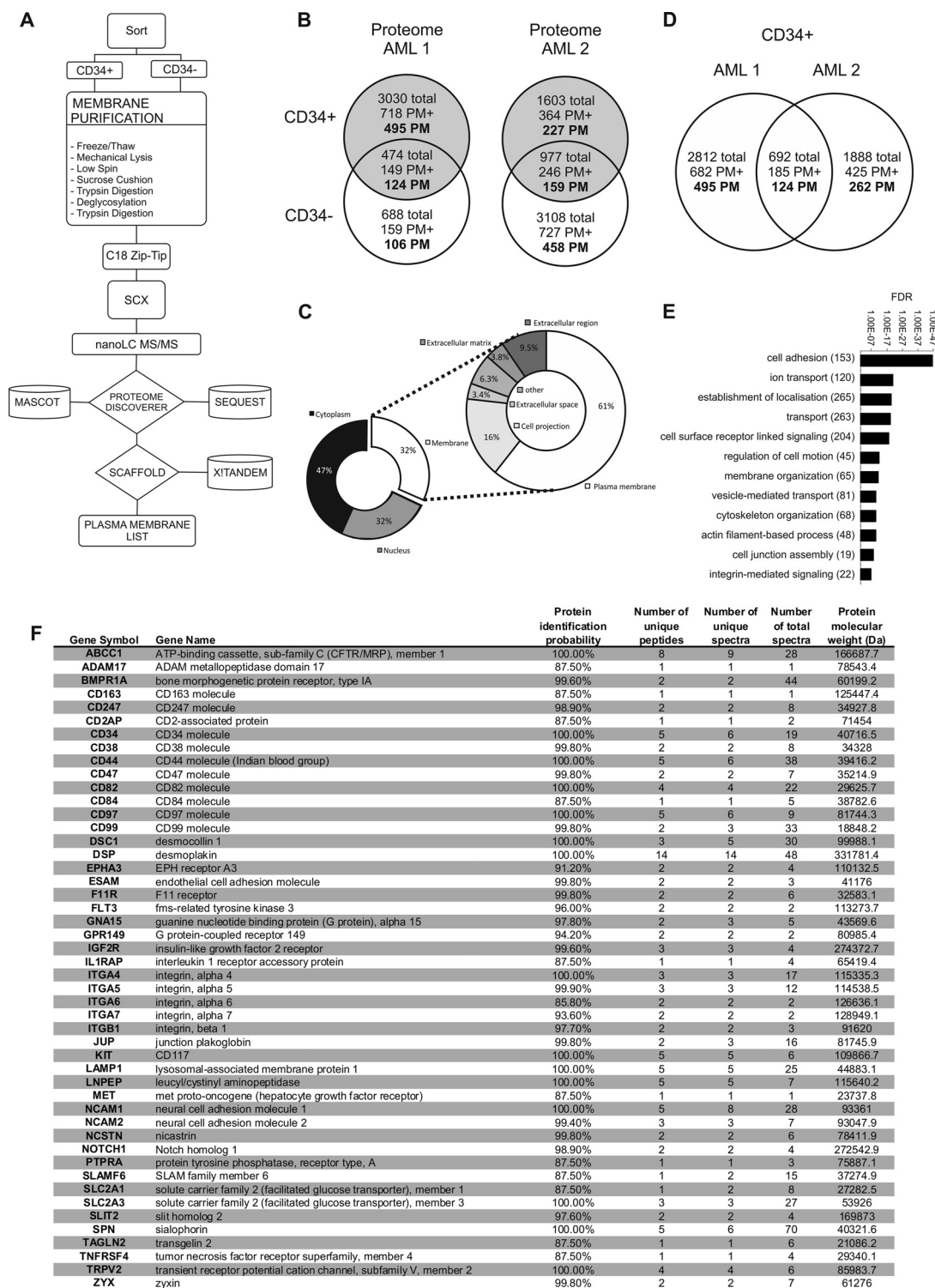


FIG. 1. Identification of the plasma membrane proteome of leukemic stem cell-enriched fractions of primary leukemia patient samples. *A*, Schematic workflow of the proteome strategy adopted for the AML samples analysis. *B*, Venn diagram showing the number of proteins identified for the two AML samples. *Total* refers to the total amount of identified proteins. *PM+* refers to a subgroup of selected proteins included in the following GO terms: plasma membrane, extracellular region, cell projection, extrinsic to membrane, extracellular matrix. *PM* indicated the number of proteins annotated as plasma membrane only. *C*, Pie chart indicating the composition of the samples after the purification procedure indicated in panel *A*. *D*, Venn diagram showing the number of total proteins, *PM+* proteins, and *PM* proteins identified in the CD34⁺ fractions of the two AML samples. *E*, Gene ontology annotation for biological processes using the combined list of all identified AML CD34⁺ *PM* proteins. *F*, Subset of *PM* proteins identified in AML CD34⁺ fractions.

the GO terms plasma membrane (GO:0005886), external side of plasma membrane (GO:0009897), integral to plasma membrane (GO:0005887), and cell surface (GO:0009986). Thus, 238 AML CD34⁺-specific up-regulated probe sets encoding 200 unique genes were identified (Fig. 2A, Supplemental File S3). In Fig. 2B, a supervised cluster analysis of these differentially expressed probesets in NBM CD34⁺ samples versus AML CD34⁺ samples is shown, and the top 20 up-regulated genes are indicated (Fig. 2B).

Comparison of the Plasma Membrane Proteome and Transcriptome of AML CD34⁺ Populations—Next, data sets obtained from our proteomics and transcriptomics approaches were compared. Among the 200 up-regulated genes in AML CD34⁺ cells at the transcriptome level, 59 were also present in at the proteome level (Fig. 2C, Supplemental File S3). GO annotation for the term molecular function showed enrichment of signal transducer activity, receptor activity, kinase activity, integrin binding, cytokine binding, receptor binding, and calcium ion binding (Fig. 2C). Clearly, these 59 genes belong to a potentially larger set of proteins that would define differences in the plasma membrane proteome between AML versus NBM CD34⁺ cells. However, because the expression of these 59 genes was confirmed at both the transcriptome as well as proteome level, we consider that this list contains putative leukemic stem cell markers that can be used to further understand the molecular biology of AML.

Verification and Functional Characterization of a Number of Putative Leukemic Stem Cell Markers—First, we set out to determine the stem cell frequencies within the CD34⁺/CD38⁻ and CD34⁺/CD38⁺ compartments in 10 primary AML patient samples by long-term culture-initiating cell assays (LTC-IC) in limiting dilution. As shown in Fig. 3A, it is clear that there was a high level of heterogeneity in the percentage of CD34⁺/CD38⁻ and CD34⁺/CD38⁺ populations. Also, we observed that the AML LTC-IC frequency was not uniquely present in the CD34⁺/CD38⁻ fraction, but that stem cell activity was observed within the CD34⁺/CD38⁺ compartment as well, in line with previously published data (15, 31). Therefore, we focused in our further studies on the AML CD34⁺ compartment.

Next, we validated the expression of CD135 (FLT3), CD47, CD96, PTH2R, and CD49f (ITGA6) by FACS in the CD34⁺ and CD34⁻ compartments of normal peripheral blood/bone marrow and AML patients samples (*n* = 9–12). Enhanced expression of FLT3 (Fig. 3B), CD47 (Fig. 3C), ITGA6 (Fig. 3D), CD96, and PTH2R (Supplemental Fig. S1) was observed in AML samples compared with normal controls, and protein expression nicely correlated with RNA data obtained by Illumina BeadArray analyses (Figs. 3B–3C, data not shown). ITGA6 was further characterized functionally in long-term MS5 stromal coculture assays. Two AML samples were sorted into CD34⁺/ITGA6⁺ and CD34⁺/ITGA6⁻ populations and plated onto MS5 cocultures, and as depicted in Fig. 3D, cells that

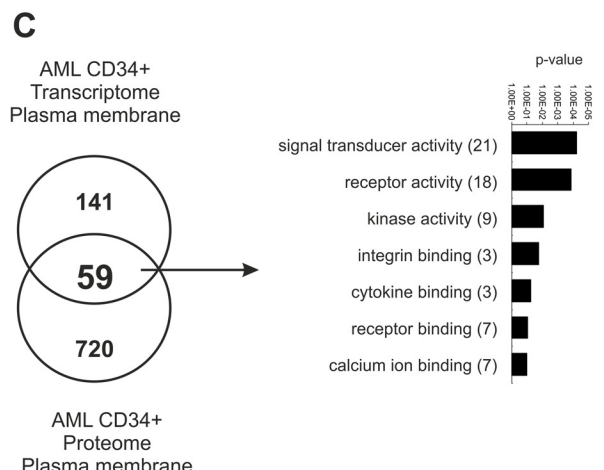
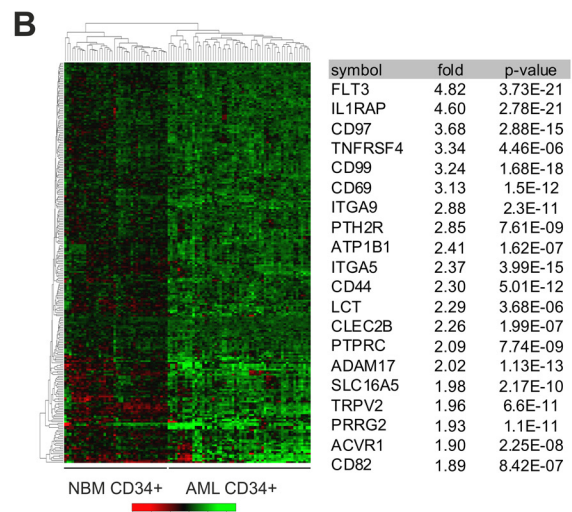
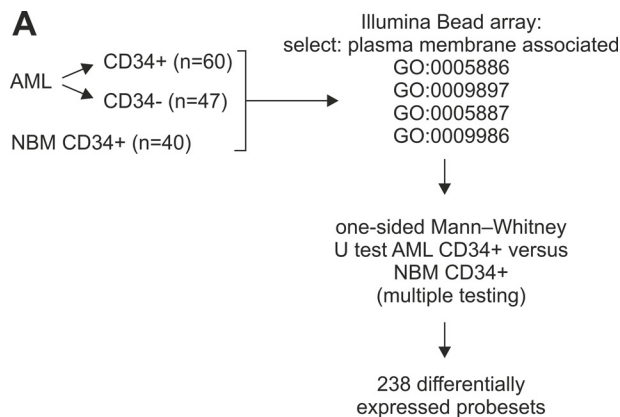


Fig. 2. Identification of leukemic stem cell markers using a transcriptomics approach. A, Scheme representing the transcriptome strategy adopted to analyze the AML and NBM samples. B, Heat map showing how the supervised clustering of the 238 differentially expressed genes obtained with the procedures indicated in panel A. C, Venn diagram showing the overlap between the plasma membrane proteins identified with the proteome approach and those identified with the transcriptome approach. The 59 overlapping proteins are further GO-annotated for molecular function.

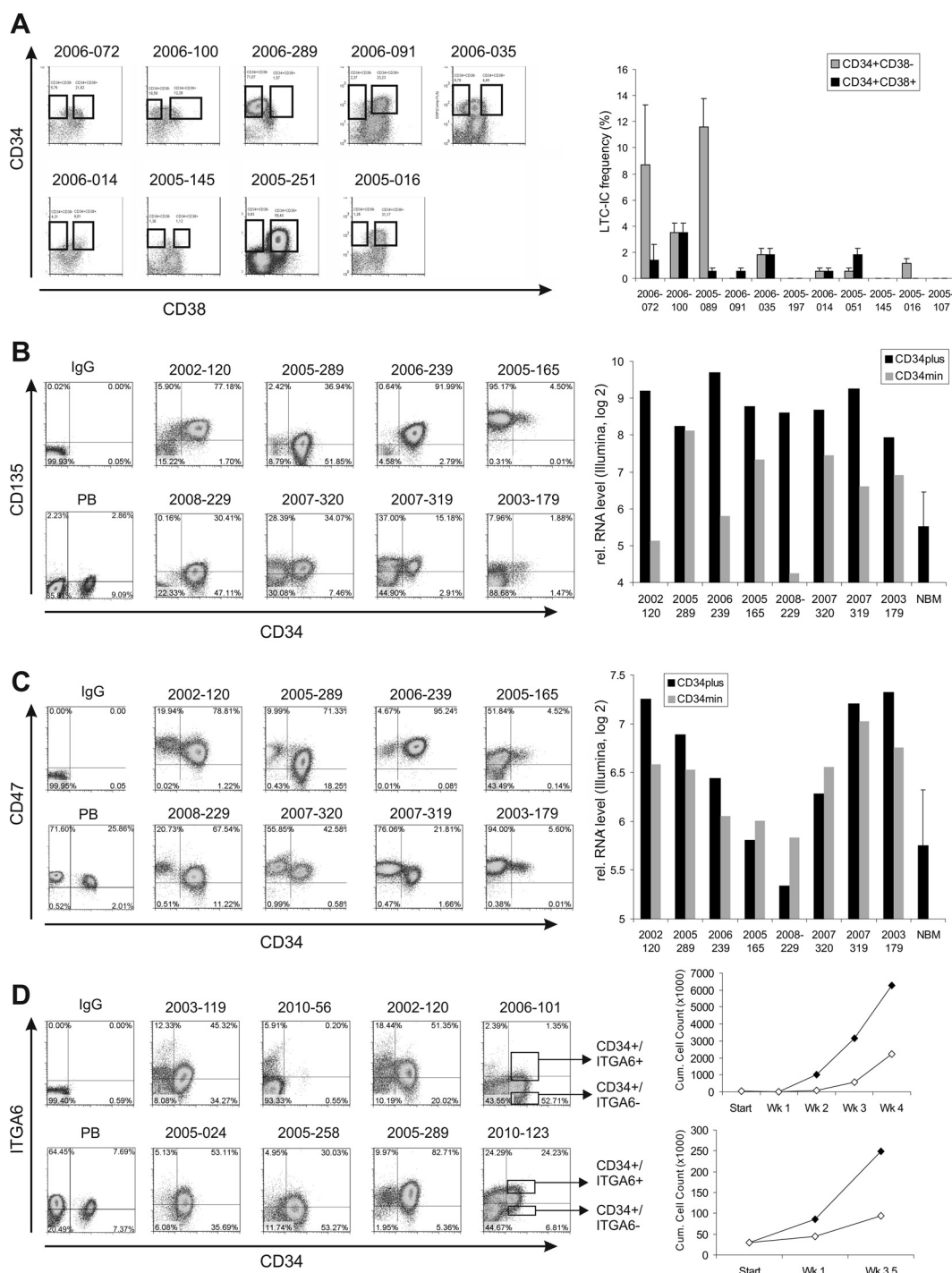


FIG. 3. Verification and functional characterization of a number of putative leukemic stem cell markers. A, Expression of CD34 and CD38 among 10 primary AML patients (left) and relative stem cell frequencies as determined by long-term culture-initiating cell assays (right). B, Validation of the expression of CD135, C, CD47, and (D) ITGA6, by FACS analysis (left panels) by FACS analysis. In (B) and (C), comparisons with mRNA levels obtained with Illumina BeadArray are also shown (right graphs). In two cases, CD34⁺/ITGA6⁺ and CD34⁺/ITGA6⁻ populations were sorted and long-term expansion was analyzed in MS5 cocultures (D, right panels).

could initiate long-term *in vitro* expansion resided predominantly within the CD34⁺/ITGA6⁺ fraction in both cases.

Evaluating Heterogeneity in Plasma Membrane Markers in AML—Leukemia is not a single type of disease, but in fact a

number of different leukemia subtypes exist. Such different subtypes of leukemia might also be reflected in how leukemic cells interact with and respond to their environment, and thus by differences in their plasma membrane composition. There-

fore, we set out to determine whether a subdivision of leukemia subtypes would be possible on the basis of differential expression of PM proteins based on our transcriptome data. To select the best discriminating uncorrelated markers we designed the following algorithm. First we calculated the information gain (25) for all the genes. Then the gene with the highest information gain was selected and all the genes that were moderately correlated (Pearson's $r > = 0.1$) to this gene were removed. This process was repeated until we obtained the best possible list of uncorrelated genes that were candidates for leukemic stem cell markers. This whole process is depicted by Algorithm 1 (see [Supplemental materials and Methods](#)). Thus, eight plasma membrane markers were identified that were almost completely uncorrelated and could significantly discriminate eight subgroups of AML within our cohort of 60 samples: being FLT3, GPR114, ITGA5, CD44, TNFRSF10B, PTH2R, FCGR1A, and TMEM5, ranked in their decreasing order of IG. Expression for the majority of these could also be confirmed at the proteome level ([Supplemental File S1](#)). Supervised cluster analysis shown in Fig. 4A clearly indicates that AML CD34⁺ samples can be separated from NBM CD34⁺ samples on the basis of the expression of these eight markers. The eight identified subgroups were not associated with a certain karyotype, risk group, or FLT3-ITD or NPM mutation status. Finally, we questioned whether these eight subgroups would be characterized by specific cell processes. Expression of all genes was ranked according to their Pearson correlation coefficient in relation to the eight uncorrelated plasma membrane markers. Thus, we generated eight individual lists, headed by our eight identified plasma membrane proteins. Genes for which the expression strongly correlated with the plasma membrane marker would reside at the top of these lists and correlation would decrease toward the bottom. To evaluate whether these lists would be useful to further understand biological aspects of these subgroups, we performed GSEA analyses and the data is summarized in Fig. 4B and [Supplemental File S4](#). Clearly, these eight subgroups were enriched for very specific GSEA terms. The FLT3 group, characterized by the highest IG, was enriched for genes associated with doxorubicin resistance, MYC signaling, glucose metabolism, and stem cell signatures. Even though this group was not significantly enriched for AMLs carrying FLT3-ITDs, we did observe enrichment for FLT3-ITD gene expression programs, suggesting that high FLT3 expressing AMLs might use similar signaling pathways. The association with stem cell and MYC signatures was further confirmed by performing GSEA directly with gene sets obtained from Wong and colleagues (ESC-like module, (32)) and Neff and colleagues (MYC signature, (33)). Enrichment for MYC and stem cell signatures was also observed for the PTH2R and TMEM5 groups. The ITGA5 group was enriched for gene sets associated with adhesion, actin cytoskeleton, CXCR4, and integrin signaling, whereas enrichment for MET signaling was observed for the CD44 and TNFRSF10B groups. Finally, we

analyzed whether any of the subgroups was enriched for good or poor prognosis gene sets as described by Yagi and colleagues (34). A strong and significant correlation with poor prognosis signature was observed for the GPR114 and TMEM5 groups. Associations with good prognosis signature were observed in the following groups: FLT3, CD44, PTH2R, ITGA5, and FCGR1A. In the last two groups, there was a concomitant negative correlation with poor prognosis signature (Fig. 4B).

DISCUSSION

Hematopoietic stem cells reside within specialized niches within the bone marrow, with which they interact via proteins within the PM. Changes in these interactions might alter hematopoietic stem cell fate and ultimately result in hematological malignancies including AML. AML is still difficult to treat, often because of relapse of disease caused by therapy-resistant LSCs. Thus, identification of markers to recognize and ultimately target LSCs is warranted.

The main aim of the current study was to characterize the plasma membrane composition of primary leukemic stem and progenitor cells from leukemia patients in detail. We have used proteome and transcriptome approaches, both of which have advantages and disadvantages. The transcriptome can easily be quantified in a large series of samples with low cell numbers, but the presence of mRNA does not always correlate directly with protein expression. In contrast, the proteomics approach will provide insight into whether certain proteins reside in the plasma membrane, but the drawback there is that the proteome is much more difficult to quantify, particularly when only limited amounts of cells can be obtained as is the case for primary leukemia stem and progenitor cells isolated from patient samples. Furthermore, the absolute number of plasma membrane identifications is strongly dependent on the number of cells analyzed. In particular, the availability of low amounts of starting material limits the precise quantification of the least abundant membrane proteins. We therefore set out to combine proteome and transcriptome approaches to gain further insight into the PM proteome composition of primary hematopoietic stem cells isolated from AML patients.

Isolation of plasma membrane proteins for mass spectrometry from embryonic stem cells, murine hematopoietic stem and progenitor cells, and carcinoma cell lines for proteomic analysis has been recently reported (35–38). We have now adapted these methods to gain further insight into the plasma membrane proteome of primary AML patient cells. We were able to analyze sorted CD34⁺ stem cell-enriched and CD34⁻ stem cell-depleted AML populations by shotgun proteomics. Among all identified proteins, 32% could be annotated as membrane proteins, of which 61% were plasma membrane proteins. Thus, 619 and 386 unique plasma membrane proteins were identified in the CD34⁺ compartments of AML1 and AML2, respectively. These lists included novel markers like CD82, CD97, CD99, PTH2R, ESAM, MET, and ITGA6, as

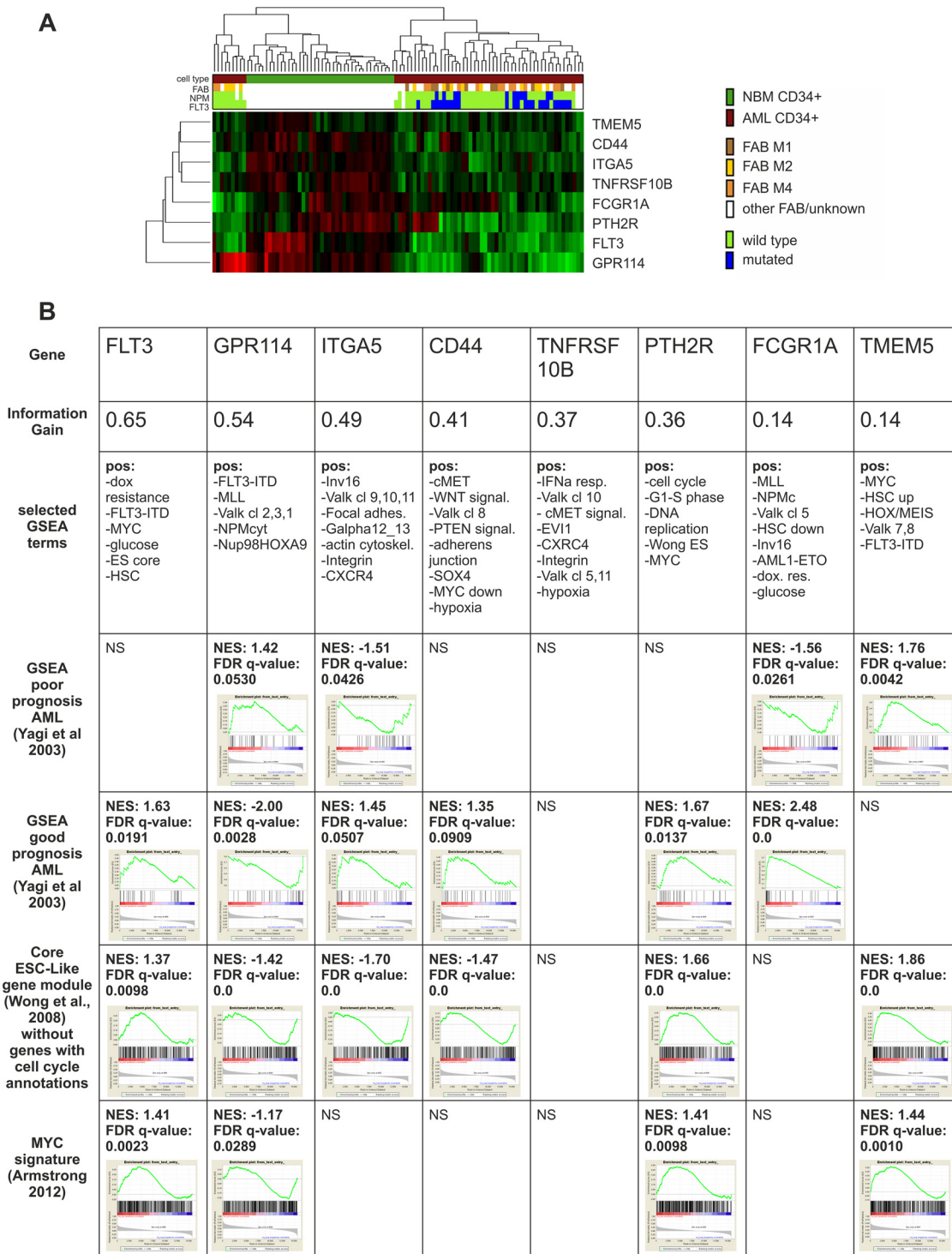


FIG. 4. Evaluating heterogeneity in plasma membrane markers in AML. A, By identifying the best discriminating uncorrelated markers using an information gain approach (see “Results” section and Materials and Methods sections for details) we were able to identify eight plasma membrane markers that were almost completely uncorrelated. Supervised cluster analysis of expression of these eight markers in AML CD34⁺ and NBM CD34⁺ samples is shown. B, An overview of the eight uncorrelated markers including information gain is shown. Gene set enrichment analysis (GSEA) of the eight plasma membrane markers indicates that the identified subgroups associate with specific gene signatures. NS denotes Not Significant.

well as previously described ones such as CD44, CD47, CD135, CD96, and ITGA5.

Although these novel plasma membrane proteomes will help to design focused future studies to further unravel the biology of leukemias, we realize that this study was based on only two patients. Also, our proteome approach did not allow a quantitative evaluation to determine which of these plasma membrane proteins would be higher expressed in leukemic stem and progenitor cells as compared with normal CD34⁺ stem and progenitor cells. Therefore, we continued with a transcriptome approach in which 60 AML patient samples were sorted into CD34⁺ stem cell-enriched and CD34⁻ leukemic stem cell-depleted fractions (of which 47 could be analyzed). As a comparison, 40 NBM CD34⁺ samples were also included in the analyses. Thus, 238 probe sets encoding 200 unique plasma-membrane associated genes were identified that were significantly up-regulated in the AML CD34⁺ fraction. Out of these 200, we found that 59 were indeed expressed at the protein plasma membrane level based on our proteome studies. It is currently unclear why the additional 141 were not identified in our proteome analysis. It is possible that (some of) these transcripts are not translated into protein, or that technical issues associated with the limited number of cells available for analysis, and/or to the presence of large hydrophobic membrane spanning regions make proteins less easy to detect. Nevertheless, our proteome and transcriptome data together clearly indicate that these 59 plasma membrane proteins are truly overexpressed in leukemic stem cell-enriched CD34⁺ cells. Increased expression of a number of these, including CD135, CD47, ITGA6, CD96, and PTH2R, was further confirmed at the protein level in an independent cohort of AML patients by FACS analyses.

In our analysis we find that CD135 is the strongest and most significantly up-regulated plasma membrane protein encoding gene in AML CD34⁺ cells compared with normal BM CD34⁺ cells at the RNA level, and increased expression was also confirmed at the protein level. CD135 (or FLT3) is a membrane receptor that is expressed in the majority of AML cases (39, 40). Activating mutations such as internal tandem duplications (FLT3-ITD) and point mutations in the tyrosine kinase domain (FLT3-TDK) are present in about 30% of AML patients (41, 42). Overexpression of FLT3 was found to be an unfavorable prognostic factor for overall survival in AML cases without FLT3/ITD (43). Indeed, very high expression of CD135 was also observed in the absence of FLT3-ITDs, suggesting that targeting this pathway might also be beneficial in patients that do not carry mutations in this receptor.

Similarly, we observed that CD47 expression is increased in AML CD34⁺ cells, both at the RNA and protein level, although there was a considerable heterogeneity among different samples. CD47 overexpression has been shown to be associated with a decreased overall survival in human AML (44). Its interaction with the signal regulatory protein alpha (SIRP α) is involved in cell-to-cell communication by prevention of phag-

ocytosis of red blood cells or platelets by macrophages (45, 46). In AML, it has been shown that disruption of the CD47-SIRP α interaction, using monoclonal antibodies, leads to phagocytosis of AML-LSC and inhibition of engraftment (44). Other previously described leukemic stem cell markers for which we provide additional evidence include CD96 (47), ITGA5 (48), CD44 (49–51), and IL3RA (52, 53). Aberrant expression of CD97 and CD99 has been observed particularly in lymphoid malignancies (54), however, in line with observations by Akashi and colleagues (55), we observe up-regulation in AML as well.

Interesting new potential leukemic stem cells markers identified in our analysis include CD82, PTH2R, ESAM, MET, and ITGA6. Recently, CD82 was shown to mediate homing and engraftment of human stem and progenitor cells (56). Parathyroid hormone receptors are typically associated with formation of the stem cell niche (57), but we find them to be up-regulated in AML CD34⁺ cells as well. ESAM was recently reported as a marker that identifies actively cycling human stem cells that do retain long-term reconstitution activity (58), and it will be interesting to determine its role in myeloid leukemias. The hepatocyte growth factor receptor MET has been shown to play an essential role in numerous cancers (59), and it was recently shown that autocrine activation of MET is frequently observed in AML (60). Recently, Notta and colleagues described that integrin alpha 6 (ITGA6, CD49F) is expressed on human hematopoietic stem cells and can be used to isolate single cells that can provide long-term reconstitution in mice (61). Interestingly, we find that ITGA6 is strongly up-regulated in AML CD34⁺ cells. Based on these results we have initiated preliminary functional studies, revealing that cells capable of long-term *in vitro* expansion reside indeed within the CD34⁺/ITGA6⁺ compartment. Ongoing studies include *in vivo* experiments in which long-term engraftment in xenograft models is evaluated, to further validate these findings.

Leukemia is clearly not a single type of disease. Typically, patients are grouped into risk categories based on their prognosis, in which the karyotype, mutation status of proteins such as FLT3, NPMc, and CEBP α , or expression levels of proteins such as EVI1 and BAALC play a dominant role (8). Because biological differences between different types of leukemia might also be initiated by interactions of leukemic cells with their environment, possibly based on differences in the composition of their plasma membrane proteome, we set out to determine whether the classification of leukemia subtypes could be associated to differential expression of PM proteins. By identifying the best discriminating uncorrelated genes, using an iterative approach that applies information gain (see “Results” section and Materials and Methods sections for details), we were able to identify eight plasma membrane markers that were almost completely uncorrelated within our cohort of 60 AML samples. These were FLT3, GPR114, ITGA5, CD44, TNFRSF10B, PTH2R, FCGR1A, and TMEM5,

ranked in their decreasing order of IG. Although the limited size of our sample group of 60 patients did not allow a thorough statistical analysis of whether some of our identified subgroups were associated with a certain karyotype, mutation status, or risk group, we did observe in an unsupervised cluster analysis based on the expression of the eight markers that a number of different clusters appeared, some of which were enriched for mutated NPMcvt and FLT3-ITDs. Furthermore, the expression of all genes was ranked according to their Pearson correlation coefficient in relation to the eight uncorrelated plasma membrane markers. Thus, we generated eight individual lists in which the expression of genes was correlated to the plasma membrane protein, ranked according to their correlation coefficient. Next, we used GSEA to evaluate whether correlations existed with previously published transcriptome sets. First we analyzed whether enrichment would be observed with gene sets that had been associated with poor or good prognosis (Yagi *et al.* 2003). Two of the identified subgroups positively correlated with a good prognosis signature, and negatively correlated with a poor prognosis signature. Reversely, in 1 subgroup we identified a strong positive correlation with poor prognosis and a negative correlation with good prognosis. Although these are interesting observations, without direct links to clinical outcome in patients, further studies are required to confirm whether these findings indeed have important clinical relevance. Also, GSEA analyses indicated that the eight subgroups could be characterized by specific cellular processes pinpointing to their possible biological relevance. For instance, a strong positive correlation with MYC signatures was identified in three out of eight of the identified subgroups. In one out of eight, a significant negative correlation was observed, and in four out of eight subgroups no significant correlation with MYC signatures was observed. Also, it is extensively debated in the field whether certain oncogenes would enforce gene expression profiles that are similar to self-renewing embryonic stem cells, and we find that in three out of eight subgroups strong positive correlations exist with embryonic-like signatures defined by Wong *et al.* (32). Clearly, one cannot make definitive statements based on GSEA approaches alone, and furthermore, future studies will be required to functionally validate these findings. However, these data will be useful for the setup of such experiments and it will be interesting to further characterize these differences in detail in the future and analyze whether the differences in plasma membrane transcriptome that we have described here allow a deeper understanding of the biology of the various subtypes of human myeloid leukemias.

Acknowledgments—We thank Rudolf Fehrmann (Department of Genetics, UMCG, Groningen, The Netherlands) for help with PCA.

* This work was supported by a grant from The Netherlands Organisation for Scientific Research (NWO-VIDI 91796312) to J.J.S.

§ This article contains supplemental Fig. S1 and Files S1 to S4.
¶ To whom correspondence should be addressed: Tel.: +31-50-3619391; Fax: +31-50-3614862; E-mail: j.j.schuringa@umcg.nl.

Author contribution: F.B. performed experiments, interpreted data, wrote the manuscript; F.F. performed and interpreted proteome experiments; P.D. Analyzed transcriptome data; D.v.G. performed FACS analysis and leukemic stem cell assays; E.V. discussed data; and J.J.S. designed experiments, interpreted data, wrote the manuscript.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

1. Estey, E., and Döhner, H. (2006) Acute myeloid leukaemia. *Lancet* **368**, 1894–1907
2. Löwenberg, B., Downing, J. R., and Burnett, A. (1999) Acute myeloid leukemia. *N. Engl. J. Med.* **341**, 1051–1062
3. Bonnet, D., and Dick, J. E. (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737
4. Dalerba, P., Cho, R. W., and Clarke, M. F. (2007) Cancer stem cells: models and concepts. *Annu. Rev. Med.* **58**, 267–284
5. Doulatov, S., Notta, F., Laurenti, E., and Dick, J. E. (2012) Hematopoiesis: a human perspective. *Cell Stem Cell* **10**, 120–136
6. Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A., and Dick, J. E. (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645–648
7. Rizo, A., Vellenga, E., de Haan, G., and Schuringa, J. J. (2006) Signaling pathways in self-renewing hematopoietic and leukemic stem cells: do all stem cells need a niche? *Hum. Mol. Genet.* **15**, R210–R219
8. Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., Harris, N. L., Le Beau, M. M., Hellström-Lindberg, E., Tefferi, A., and Bloomfield, C. D. (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* **114**, 937–951
9. Kondo, M., Wagers, A. J., Manz, M. G., Prohaska, S. S., Scherer, D. C., Beilhack, G. F., Shizuru, J. A., and Weissman, I. L. (2003) Biology of hematopoietic stem cells and progenitors: implications for clinical application. *Annu. Rev. Immunol.* **21**, 759–806
10. Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., Ritchey, J. K., Young, M. A., Lamprecht, T., McLellan, M. D., McMichael, J. F., Wallis, J. W., Lu, C., Shen, D., Harris, C. C., Dooling, D. J., Fulton, R. S., Fulton, L. L., Chen, K., Schmidt, H., Kalicki-Veizer, J., Magrini, V. J., Cook, L., McGrath, S. D., Vickery, T. L., Wendt, M. C., Heath, S., Watson, M. A., Link, D. C., Tomasson, M. H., Shannon, W. D., Payton, J. E., Kulkarni, S., Westervelt, P., Walter, M. J., Graubert, T. A., Mardis, E. R., Wilson, R. K., and DiPersio, J. F. (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510
11. Walter, M. J., Shen, D., Ding, L., Shao, J., Koboldt, D. C., Chen, K., Larson, D. E., McLellan, M. D., Dooling, D., Abbott, R., Fulton, R., Magrini, V., Schmidt, H., Kalicki-Veizer, J., O’Laughlin, M., Fan, X., Grillo, M., Witowski, S., Heath, S., Frater, J. L., Eades, W., Tomasson, M., Westervelt, P., DiPersio, J. F., Link, D. C., Mardis, E. R., Ley, T. J., Wilson, R. K., and Graubert, T. A. (2012) Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1090–1098
12. Valk, P. J., Verhaak, R. G., Beijten, M. A., Erpelinck, C. A., Barjesteh van Waalwijk van Doorn-Khosrovani, S., Boer, J. M., Beverloo, H. B., Moorhouse, M. J., van der Spek, P. J., Löwenberg, B., and Delwel, R. (2004) Prognostically useful gene-expression profiles in acute myeloid leukemia. *N. Engl. J. Med.* **350**, 1617–1628
13. Metzeler, K. H., Hummel, M., Bloomfield, C. D., Spiekermann, K., Braess, J., Sauerland, M. C., Heinecke, A., Radmacher, M., Marcucci, G., Whitman, S. P., Maharry, K., Paschka, P., Larson, R. A., Berdel, W. E., Buchner, T., Wormann, B., Mansmann, U., Hiddemann, W., Bohlander, S. K., and Buske, C. (2008) An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood* **112**, 4193–4201
14. de Jonge, H. J., Woolthuis, C. M., Vos, A. Z., Mulder, A., van den Berg, E., Kluijn, P. M., van der Weide, K., de Bont, E. S., Huls, G., Vellenga, E., and

- Schuringa, J. J. (2011) Gene expression profiling in the leukemic stem cell-enriched CD34+ fraction identifies target genes that predict prognosis in normal karyotype AML. *Leukemia* **25**, 1825–1833
15. Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K. H., Poepl, A., Ling, V., Beyene, J., Canty, A. J., Danska, J. S., Bohlander, S. K., Buske, C., Minden, M. D., Golub, T. R., Jurisica, I., Ebert, B. L., and Dick, J. E. (2011) Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* **17**, 1086–1093
 16. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207
 17. Cravatt, B. F., Simon, G. M., and Yates, J. R. 3rd (2007) The biological impact of mass-spectrometry-based proteomics. *Nature* **450**, 991–1000
 18. Patterson, S. D., and Aebersold, R. H. (2003) Proteomics: the first decade and beyond. *Nat. Genet.* **33**, 311–323
 19. van Goslign, D., Schepers, H., Rizo, A., van der Kolk, D., Vellenga, E., and Schuringa, J. J. (2007) Establishing long-term cultures with self-renewing acute myeloid leukemia stem/progenitor cells. *Exp. Hematol.* **35**, 1538–1549
 20. Schuringa, J. J., and Schepers, H. (2009) Ex vivo assays to study self-renewal and long-term expansion of genetically modified primary human acute myeloid leukemia stem cells. *Methods Mol. Biol.* **538**, 287–300
 21. Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smyth, G., Tierney, L., Yang, J. Y., and Zhang, J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome. Biol.* **5**, R80
 22. Du, P., Kibbe, W. A., and Lin, S. M. (2008) lumi: a pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547–1548
 23. Crijns, A. P., Fehrmann, R. S., de, J. S., Gerbens, F., Meersma, G. J., Klip, H. G., Hollema, H., Hofstra, R. M., te Meerman, G. J., de Vries, E. G., and van der Zee, A. G. (2009) Survival-related profile, pathways, and transcription factors in ovarian cancer. *PLoS. Med.* **6**, e24
 24. Alter, O., Brown, P. O., and Botstein, D. (2000) Singular value decomposition for genome-wide expression data processing and modeling. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10101–10106
 25. Mitchell, T. M. (1997) Machine Learning. *New York, McGraw-Hill*
 26. Hope, K. J., Jin, L., and Dick, J. E. (2004) Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. *Nat. Immunol.* **5**, 738–743
 27. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**, 242–247
 28. Muntel, J., Hecker, M., and Becher, D. (2012) An exclusion list based label-free proteome quantification approach using an LTQ Orbitrap. *Rapid Commun. Mass. Spectrom.* **26**, 701–709
 29. Korn, E. L., Troendle, J. F., McShane, L. M., and Simon, R. (2004) Controlling the number of false discoveries: application to high-dimensional genomic data. *J. Stat. Plan. Inference.* **124**, 379–398
 30. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B. Meth.* **57**, 289–300
 31. Taussig, D. C., Miraki-Moud, F., Anjos-Afonso, F., Pearce, D. J., Allen, K., Ridler, C., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S. G., Lister, T. A., Gribben, J. G., and Bonnet, D. (2008) Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood* **112**, 568–575
 32. Wong, D. J., Liu, H., Ridky, T. W., Cassarino, D., Segal, E., and Chang, H. Y. (2008) Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell* **2**, 333–344
 33. Neff, T., Sinha, A. U., Kluk, M. J., Zhu, N., Khattab, M. H., Stein, L., Xie, H., Orkin, S. H., and Armstrong, S. A. (2012) Polycomb repressive complex 2 is required for MLL-AF9 leukemia. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5028–5033
 34. Yagi, T., Morimoto, A., Eguchi, M., Hibi, S., Sako, M., Ishii, E., Mizutani, S., Imashuku, S., Ohki, M., and Ichikawa, H. (2003) Identification of a gene expression signature associated with pediatric AML prognosis. *Blood* **102**, 1849–1856
 35. Bruce, S. J., Jonsson, P., Antti, H., Cloarec, O., Trygg, J., Marklund, S. L., and Moritz, T. (2008) Evaluation of a protocol for metabolic profiling studies on human blood plasma by combined ultra-performance liquid chromatography/mass spectrometry: From extraction to data analysis. *Anal. Biochem.* **372**, 237–249
 36. Dormeyer, W., van Hoof, D., Braam, S. R., Heck, A. J., Mummery, C. L., and Krijgsvel, J. (2008) Plasma membrane proteomics of human embryonic stem cells and human embryonal carcinoma cells. *J. Proteome. Res.* **7**, 2936–2951
 37. Lund, R., Leth-Larsen, R., Jensen, O. N., and Ditzel, H. J. (2009) Efficient isolation and quantitative proteomic analysis of cancer cell plasma membrane proteins for identification of metastasis-associated cell surface markers. *J. Proteome Res.* **8**, 3078–3090
 38. Klimm, D., Hansson, J., Raffel, S., Vakhrushev, S. Y., Trumpp, A., and Krijgsvel, J. (2012) Proteomic cornerstones of hematopoietic stem cell differentiation: distinct signatures of multipotent progenitors and myeloid committed cells. *Mol. Cell Proteomics*. Epub ahead of print
 39. Rosnet, O., Bühring, H. J., Marchetto, S., Rappold, I., Lavagna, C., Sainy, D., Arnoulet, C., Chabannon, C., Kanz, L., Hannum, C., and Birnbaum, D. (1996) Human FLT3/FLK2 receptor tyrosine kinase is expressed at the surface of normal and malignant hematopoietic cells. *Leukemia* **10**, 238–248
 40. Carow, C. E., Levenstein, M., Kaufmann, S. H., Chen, J., Amin, S., Rockwell, P., Witte, L., Borowitz, M. J., Civin, C. I., and Small, D. (1996) Expression of the hematopoietic growth factor receptor FLT3 (STK-1/Flk2) in human leukemias. *Blood* **87**, 1089–1096
 41. Nakao, M., Yokota, S., Iwai, T., Kaneko, H., Horiike, S., Kashima, K., Sonoda, Y., Fujimoto, T., and Misawa, S. (1996) Internal tandem duplication of the flt3 gene found in acute myeloid leukemia. *Leukemia* **10**, 1911–1918
 42. Yamamoto, Y., Kiyoi, H., Nakano, Y., Suzuki, R., Kodera, Y., Miyawaki, S., Asou, N., Kuriyama, K., Yagasaki, F., Shimazaki, C., Akiyama, H., Saito, K., Nishimura, M., Motoji, T., Shinagawa, K., Takeshita, A., Saito, H., Ueda, R., Ohno, R., and Naoe, T. (2001) Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood* **97**, 2434–2439
 43. Ozeki, K., Kiyoi, H., Hirose, Y., Iwai, M., Ninomiya, M., Kodera, Y., Miyawaki, S., Kuriyama, K., Shimazaki, C., Akiyama, H., Nishimura, M., Motoji, T., Shinagawa, K., Takeshita, A., Ueda, R., Ohno, R., Emi, N., and Naoe, T. (2004) Biologic and clinical significance of the FLT3 transcript level in acute myeloid leukemia. *Blood* **103**, 1901–1908
 44. Majeti, R., Chao, M. P., Alizadeh, A. A., Pang, W. W., Jaiswal, S., Gibbs, K. D., Jr., van Rooijen, N., and Weissman, I. L. (2009) CD47 is an adverse prognostic factor and therapeutic antibody target on human acute myeloid leukemia stem cells. *Cell* **138**, 286–299
 45. Oldenburg, P. A., Zheleznyak, A., Fang, Y. F., Lagenaur, C. F., Gresham, H. D., and Lindberg, F. P. (2000) Role of CD47 as a marker of self on red blood cells. *Science* **288**, 2051–2054
 46. Yamao, T., Noguchi, T., Takeuchi, O., Nishiyama, U., Morita, H., Hagiwara, T., Akahori, H., Kato, T., Inagaki, K., Okazawa, H., Hayashi, Y., Matozaki, T., Takeda, K., Akira, S., and Kasuga, M. (2002) Negative regulation of platelet clearance and of the macrophage phagocytic response by the transmembrane glycoprotein SHPS-1. *J. Biol. Chem.* **277**, 39833–39839
 47. Hosen, N., Park, C. Y., Tatsumi, N., Oji, Y., Sugiyama, H., Gramatzki, M., Krensky, A. M., and Weissman, I. L. (2007) CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11008–11013
 48. Bendall, L. J., Kortlepel, K., and Gottlieb, D. J. (1993) Human acute myeloid leukemia cells bind to bone marrow stroma via a combination of beta-1 and beta-2 integrin mechanisms. *Blood* **82**, 3125–3132
 49. Jin, L., Hope, K. J., Zhai, Q., Smadja-Joffe, F., and Dick, J. E. (2006) Targeting of CD44 eradicates human acute myeloid leukemic stem cells. *Nat. Med.* **12**, 1167–1174
 50. Quéré, R., Andradottir, S., Brun, A. C., Zubarev, R. A., Karlsson, G., Olsson, K., Magnusson, M., Cammenga, J., and Karlsson, S. (2011) High levels of the adhesion molecule CD44 on leukemic cells generate acute myeloid leukemia relapse after withdrawal of the initial transforming event. *Leukemia* **25**, 515–526
 51. Charrad, R. S., Li, Y., Delpech, B., Balitrand, N., Clay, D., Jasmin, C., Chomienne, C., and Smadja-Joffe, F. (1999) Ligation of the CD44 adhesion molecule reverses blockage of differentiation in human acute myeloid leukemia. *Nat. Med.* **5**, 669–676
 52. Jordan, C. T., Upchurch, D., Szilvassy, S. J., Guzman, M. L., Howard, D. S.,

- Pettigrew, A. L., Meyerrose, T., Rossi, R., Grimes, B., Rizzieri, D. A., Luger, S. M., and Phillips, G. L. (2000) The interleukin-3 receptor alpha chain is a unique marker for human acute myelogenous leukemia stem cells. *Leukemia* **14**, 1777–1784
53. Jin, L., Lee, E. M., Ramshaw, H. S., Busfield, S. J., Peoppl, A. G., Wilkinson, L., Guthridge, M. A., Thomas, D., Barry, E. F., Boyd, A., Gearing, D. P., Vairo, G., Lopez, A. F., Dick, J. E., and Lock, R. B. (2009) Monoclonal antibody-mediated targeting of CD123, IL-3 receptor alpha chain, eliminates human acute myeloid leukemic stem cells. *Cell Stem Cell* **5**, 31–42
54. Coustan-Smith, E., Song, G., Clark, C., Key, L., Liu, P., Mehrpooya, M., Stow, P., Su, X., Shurtleff, S., Pui, C. H., Downing, J. R., and Campana, D. (2011) New markers for minimal residual disease detection in acute lymphoblastic leukemia. *Blood* **117**, 6267–6276
55. Kikushige, Y., Shima, T., Takayanagi, S., Urata, S., Miyamoto, T., Iwasaki, H., Takenaka, K., Teshima, T., Tanaka, T., Inagaki, Y., and Akashi, K. (2010) TIM-3 is a promising target to selectively kill acute myeloid leukemia stem cells. *Cell Stem Cell* **7**, 708–717
56. Larochelle, A., Gillette, J. M., Desmond, R., Ichwan, B., Cantilena, A., Cerf, A., Barrett, A. J., Wayne, A. S., Lippincott-Schwartz, J., and Dunbar, C. E. (2012) Bone marrow homing and engraftment of human hematopoietic stem and progenitor cells is mediated by a polarized membrane domain. *Blood* **119**, 1848–1855
57. Calvi, L. M., Adams, G. B., Weibrecht, K. W., Weber, J. M., Olson, D. P., Knight, M. C., Martin, R. P., Schipani, E., Divieti, P., Bringhurst, F. R., Milner, L. A., Kronenberg, H. M., and Scadden, D. T. (2003) Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425**, 841–846
58. Sudo, T., Yokota, T., Oritani, K., Satoh, Y., Sugiyama, T., Ishida, T., Shibayama, H., Ezoe, S., Fujita, N., Tanaka, H., Maeda, T., Nagasawa, T., and Kanakura, Y. (2012) The endothelial antigen ESAM monitors hematopoietic stem cell status between quiescence and self-renewal. *J. Immunol.* **189**, 200–210
59. Gherardi, E., Birchmeier, W., Birchmeier, C., and Vande, Woude, G. (2012) Targeting MET in cancer: rationale and progress. *Nat. Rev. Cancer* **12**, 89–103
60. Kentsis, A., Reed, C., Rice, K. L., Sanda, T., Rodig, S. J., Tholouli, E., Christie, A., Valk, P. J., Delwel, R., Ngo, V., Kutok, J. L., Dahlberg, S. E., Moreau, L. A., Byers, R. J., Christensen, J. G., Woude, G. V., Licht, J. D., Kung, A. L., Staudt, L. M., and Look, A. T. (10–6-2012) Autocrine activation of the MET receptor tyrosine kinase in acute myeloid leukemia. *Nat. Med.* Epub ahead of print
61. Notta, F., Doulatov, S., Laurenti, E., Poepl, A., Jurisica, I., and Dick, J. E. (2011) Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* **333**, 218–221