ARTICLE

Reduced hematopoietic stem cell frequency predicts outcome in acute myeloid leukemia

Wenwen Wang,[1,8] Thomas Stiehl,[2] Simon Raffel,[1,3,4] Van T. Hoang,[1,9] Isabel Hoffmann,[1] Laura Poisa-Beiro,[1] Borhan R. Saeed,[1] Rachel Blume,[1] Linda Manta,[1] Volker Eckstein,[1] Tilmann Bochtler,[1,5] Patrick Wuchter,[1,10] Marieke Essers,[3,4] Anna Jauch,[6] Andreas Trumpp,[3,4,7] Anna Marciniak-Czochra,[2] Anthony D. Ho[1] and Christoph Lutz[1,7]

[1]Department of Medicine V, Heidelberg University, Germany; [2]Institute of Applied Mathematics, Interdisciplinary Center for Scientific Computing (IWR), BIOQUANT, Heidelberg University, Germany; [3]Division of Stem Cells and Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany; [4]Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM gGmbH), Germany; [5]Clinical Cooperation Unit Molecular Hematology/Oncology, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany; [6]Institute of Human Genetics, Heidelberg University, Germany; [7]German Cancer Consortium (DKTK), Heidelberg, Germany; [8]Present Address: Department of Oncology, Nanjing Medical University Affiliated Wuxi Second Hospital, Jiangsu, China; [9]Present Address: Georg-Speyer-Haus, Institute for Tumor Biology and Experimental Therapy, Frankfurt am Main, Germany and [10]Present Address: Institute of Transfusion Medicine and Immunology, Medical Faculty Mannheim, Heidelberg University; German Red Cross Blood Service Baden-Württemberg – Hessen, Germany

Haematologica 2017
Volume 102(9):1567-1577

ABSTRACT

In patients with acute myeloid leukemia and low percentages of aldehyde-dehydrogenase-positive cells, non-leukemic hematopoietic stem cells can be separated from leukemic cells. By relating hematopoietic stem cell frequencies to outcome we detected poor overall- and disease-free survival of patients with low hematopoietic stem cell frequencies. Serial analysis of matched diagnostic and follow-up samples further demonstrated that hematopoietic stem cells increased after chemotherapy in patients who achieved durable remissions. However, in patients who eventually relapsed, hematopoietic stem cell numbers decreased dramatically at the time of molecular relapse demonstrating that hematopoietic stem cell levels represent an indirect marker of minimal residual disease, which heralds leukemic relapse. Upon transplantation in immune-deficient mice cases with low percentages of hematopoietic stem cells of our cohort gave rise to leukemic or no engraftment, whereas cases with normal hematopoietic stem cell levels mostly resulted in multi-lineage engraftment. Based on our experimental data, we propose that leukemic stem cells have increased niche affinity in cases with low percentages of hematopoietic stem cells. To validate this hypothesis, we developed new mathematical models describing the dynamics of healthy and leukemic cells under different regulatory scenarios. These models suggest that the mechanism leading to decreases in hematopoietic stem cell frequencies before leukemic relapse must be based on expansion of leukemic stem cells with high niche affinity and the ability to dislodge hematopoietic stem cells. Thus, our data suggest that decreasing numbers of hematopoietic stem cells indicate leukemic stem cell persistence and the emergence of leukemic relapse.

Introduction

Acute myeloid leukemia (AML) is a malignant disease and affected people have poor overall survival (OS) rates.[1-3] Although most AML patients achieve complete remissions after standard chemotherapy, the majority subsequently relapse with

Correspondence:
christoph.lutz@med.uni-heidelberg.de

Received: January 5, 2017.
Accepted: May 17, 2017.
Pre-published: May 26, 2017.

doi:10.3324/haematol.2016.163584

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/102/9/1567

more aggressive and resistant disease demonstrating the necessity to improve therapeutic strategies.[4-6] As relapse is the main cause of death, it is important to effectively stratify AML patients according to their individual risk of relapse and to identify patients who require more aggressive treatment protocols, such as allogeneic hematopoietic stem cell transplantation (HSCT). Cytogenetic analysis is so far the most important and reliable risk stratification tool with specific abnormalities and the degree of mutations indicating cases with good and poor prognosis. However, up to 50% of AML patients have normal karyotypes and can be classified in the cytogenetic intermediate-risk group that combines patients with highly variable clinical outcomes.[1,5] The systematic analysis and identification of AML-specific mutations, such as Fms-related tyrosine kinase 3 (*FLT3*)-internal tandem duplication (ITD), CCAAT/enhancer-binding protein-α (*CEBPA*) and nucleophosmin 1 (*NPM1*), further improved AML risk stratification, increasing the number of patients who can be stratified.[7,8] Another important improvement in AML management and risk stratification is the ability to track minimal residual disease (MRD) in follow-up samples of patients undergoing therapy; MRD can be assessed by flow cytometry or molecular analysis of existing mutations, such as *NPM1*.[9]

So far, tracking of molecular MRD by polymerase chain reaction analysis represents the most sensitive approach, with detection levels ranging from 0.1-0.001%. This approach is, however, limited by the fact that only 60-70% of patients have suitable markers that can be monitored on a molecular basis.[10,11] In contrast flow cytometry-based MRD analysis is applicable in over 90% of patients but its level of sensitivity can be low (0.1-0.01%) and optimal cut off values are debatable.[11-13]

A completely different approach for AML risk stratification has been suggested by our recent study in which we have shown that functionally normal and non-leukemic hematopoietic stem cells (HSC) can be isolated from a subgroup of patients using the $CD34^+(CD38^-)ALDH^+$ (aldehyde dehydrogenase) phenotype.[14] Importantly, patients suitable for this distinction can be prospectively identified by their low frequency of total $ALDH^+$ cells [<1.9% of total mononuclear cells (MNC)] - named ALDH-rare AML.[6,14,15] Patients in whom non-leukemic HSC (nl-HSC) can be separated represent a cohort with favorable outcome compared to the rest of AML patients.[14,16,17] However, within ALDH-rare AML, nl-HSC frequencies vary significantly and even patients without any detectable $ALDH^+$ cells exist (*unpublished observation*).[18] Studies analyzing HSC in AML have mostly focused on the cells' functional and genetic properties. In contrast to previous studies, we have analyzed how nl-HSC relate to survival and behave during the course of disease in relation to disease burden. We have also analyzed the functional properties of nl-HSC *in vitro* and *in vivo* in correlation to their frequency within the bone marrow compartment which may serve as an indicator of niche changes or competition both facilitated by leukemic infiltration.

Our experimental results show that nl-HSC frequencies predict outcome and correlate to MRD in follow-up samples. We hypothesized that the correlation of nl-HSC frequencies, MRD status and patients' survival can be explained by an ongoing competition between leukemic stem cells (LSC) and nl-HSC in the bone marrow niche. To support this hypothesis, we proposed a novel mathematical model. Mathematical modeling has been shown to be a useful tool, improving our understanding of the hematopoietic system and its diseases, as it allows the study of processes that cannot be observed in conventional experiments. In this context, our model enables linkage of clinical data to unobservable dynamic processes in the human stem cell niche.[19-23] Based on our model simulations we conclude that cell competition within the niche is required to explain the decline of nl-HSC before overt relapse.

Methods

Sample collection

Bone marrow aspirates derived from 61 AML patients and 11 healthy donors were collected after informed consent between October 2011 and August 2015. All experiments were approved by the Ethics Committee of the Medical Faculty of Heidelberg University. The patients' characteristics are shown in *Online Supplementary Table S1*. The exact time-points of follow-up sample collection are indicated in the captions of the relevant figures.

Preparation of cells

MNC were isolated with Biocoll separation solution (Biochrom, Berlin, Germany) by gradient centrifugation and used fresh or frozen in liquid nitrogen with fetal calf serum/12.5% dimethyl sulfoxide for future studies. Numbers of nl-HSC were calculated as described in the *Online Supplement*. Mesenchymal stromal cells from healthy donors were isolated and cultured as described previously.[24]

Flow cytometry and fluorescence-activated cell sorting

MNC were stained with Aldefluor reagent (Stem Cell Technologies, Vancouver, BC, Canada), CD34-APC, CD45-APC-H7, propidium iodide (PI) or 7-AAD, CD38-PE or CD38-PE-Cy7 (clone HB7) (BD Bioscience, Heidelberg, Germany) as described previously.[14] Cells were analyzed using a BD LSR II flow cytometer and subpopulations were sorted using a FACSAria II sorter (BD Bioscience). Guided by the percentages of HSC detected in the bone marrow of healthy donors, 0.01% ($CD34^+CD38^-ALDH^+$/MNC) was chosen as a cut-off value to distinguish between nl-$HSC^{-/low}$ and nl-HSC^+ AML (Figure 2C). This corresponds to the 5% percentile of HSC frequencies in healthy individuals.

In vitro colony assays

To compare the long-term colony forming abilities of different subpopulations, we performed the limiting dilution long-term culture-initiating cell (LTC-IC) assay with HSC-CFU lite with Epo (Miltenyi Biotec, Bergisch Gladbach, Germany) as described previously.[14] To compare progenitor potential of different fractions, short-term colony-forming cell (CFC) assays were performed using HSC-CFU complete with Epo (Miltenyi Biotec) according to the manufacturer's instructions. LTC-IC frequencies were determined by L-Calc Limiting Dilution Software (Stem Cell Technologies, Vancouver, BC, Canada).

Mathematical modeling

To uncover mechanisms leading to the dynamics observed in the experimental data, we used computer simulations of mathematical models reflecting different plausible interactions of healthy and leukemic cells. In particular, we developed a novel mathematical model describing dynamics in the bone marrow niche. The model is an extension of our previously published

model on LSC dynamics in acute leukemias.[23] The new model includes competition of healthy and leukemic cells for niche spaces and dislodgement of healthy cells from the niche by leukemic cells. The model is based on a system of non-linear ordinary differential equations, describing proliferation, self-renewal, differentiation, death and various possible interactions of healthy and leukemic cells.

For additional details of the study methods see the *Online Supplementary Methods* section.

Results

CD34$^+$CD38$^-$ALDH$^+$ cells are enriched for hematopoietic stem cell potential

Our previous analyses using *in vivo* and *in vitro* assays showed that nl-HSC can be isolated from leukemic cells in a subgroup of AML patients using the CD34$^+$ALDH$^+$ marker combination (Figure 1A).[14] In an attempt to further enrich nl-HSC within this compartment we analyzed CD34$^+$CD38$^-$ALDH$^+$ cells for their *in vitro* colony-forming potential in comparison to that of other subcompartments including CD34$^+$CD38$^+$ALDH$^+$, CD34$^-$ALDH$^+$, ALDH$^-$ and CD34$^+$CD38$^-$ALDH$^-$ cells. Limiting dilution LTC-IC assays quantifying the long-term colony-forming ability of tested subpopulations were performed for 24 diagnostic ALDH-rare AML cases. LTC-IC frequencies were highest for CD34$^+$CD38$^-$ALDH$^+$ cells (1 in 2.5 cells) confirming that these cells are enriched for HSC potential (Figure 1B). Normal bone marrow samples served as healthy controls and showed comparable LTC-IC frequencies for CD34$^+$CD38$^-$ALDH$^+$ cells (Figure 1C). Progenitor potential was also determined by plating the same subpopulations in methylcellulose. The results obtained confirmed the LTC-IC data

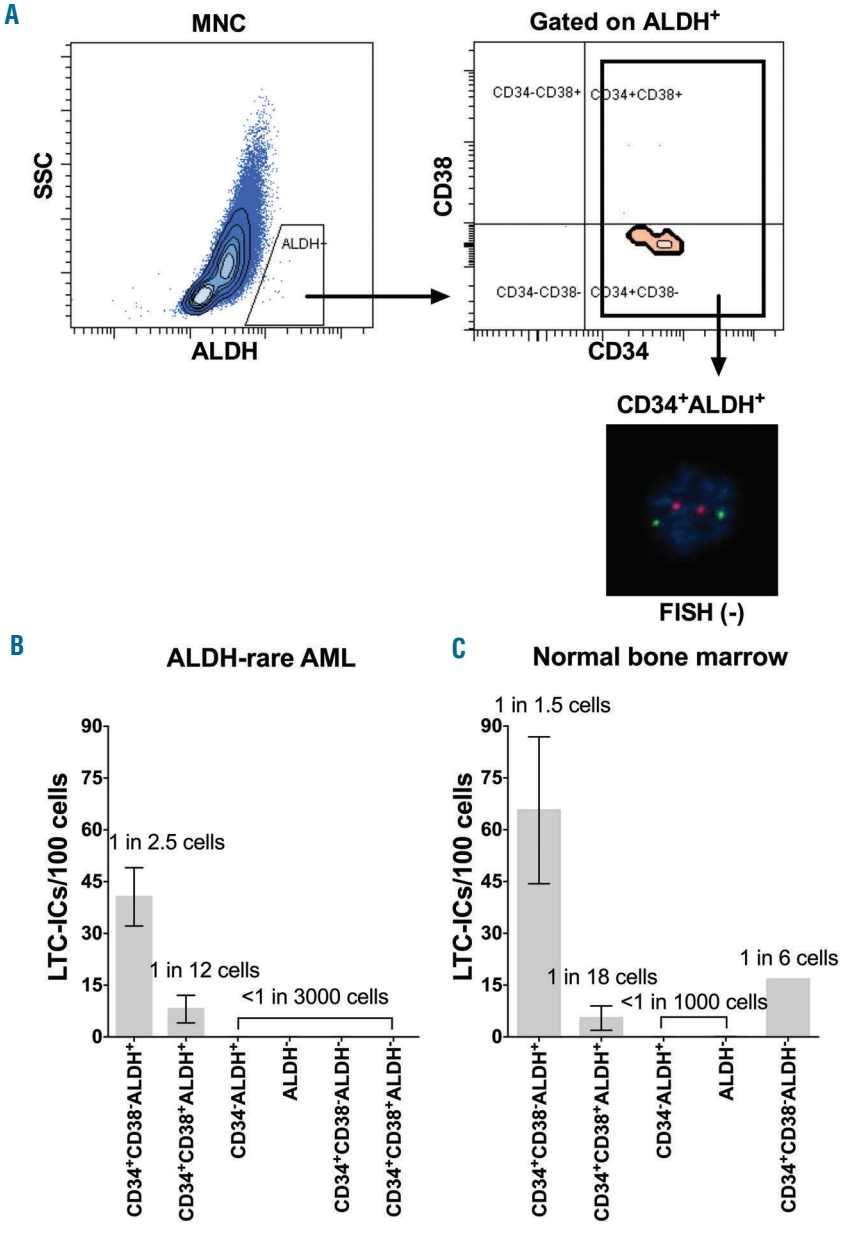

Figure 1. CD34$^+$CD38$^-$ALDH$^+$ cells are enriched for hematopoietic stem cell potential. (A) Example of sorting and analysis strategy of CD34$^+$ALDH$^+$ cells. (B) LTC-IC frequencies of sorted subpopulations of 24 ALDH-rare AML samples. Results are shown in LTC-IC per 100 plated cells. (C) LTC-IC frequencies of sorted subpopulations of four normal bone marrow samples. Results are shown in LTC-IC per 100 plated cells. Data are shown as mean ± SEM.

with comparable CFC potential for CD34$^+$CD38$^-$ALDH$^+$ cells from leukemic and normal bone marrow samples (*Online Supplementary Figure S1A,B*). These colonies also contained erythroid colonies which have been shown to indicate the non-leukemic origin of the tested cells.[25] Fluorescence *in situ* hybridization analyses of CD34$^+$ALDH$^+$ cells were negative for leukemia-specific mutations (*Online Supplementary Figure S2A,B*). These data are in line with results from our previous study, in which we showed, on a larger scale by polymerase chain reaction analysis and fluorescence *in situ* hybridization, that CD34$^+$ALDH$^+$ cells are functionally normal, non-leukemic and capable of multi-lineage engraftment in NSG mice.[14]

Hematopoietic stem cell frequencies vary in the cohort of ALDH-rare acute myeloid leukemias

We next systematically determined nl-HSC frequencies of 61 ALDH-rare AML diagnostic samples and found highly variable percentages (range, 0% to 0.564%) (Figure 2A). These patients also included two cases without any detectable CD34$^+$CD38$^-$ALDH$^+$ cells. These two patients had extremely poor survival and both died of early relapse (Figure 2A and *Online Supplementary Figure S3*). Guided by these observations and the percentages of HSC detected in the bone marrow of healthy donors, 0.01% (CD34$^+$CD38$^-$ALDH$^+$/MNC) was chosen as a cut-off value to stratify ALDH-rare AML patients in order to assess whether nl-HSC frequencies correlate with disease outcome (Figure 2B). The majority of all 61 studied patients (45 out of 61) had relatively high HSC proportions (nl-HSC$^+$ AML; ≥0.01% of total MNC) while patients with nl-HSC$^{-/low}$ AML (<0.01% of total MNC) were less frequent. Since nl-HSC proportions do not necessarily reflect absolute nl-HSC numbers, we calculated absolute numbers according to bone marrow sample volumes, which were available for 47 AML patients. Results of nl-HSC proportions and nl-HSC numbers (cells/mL) were consistent and showed low absolute numbers for patients with nl-HSC$^{-/low}$ AML compared to healthy subjects and patients with nl-HSC$^+$ AML (Figure 2C).

Non-leukemic hematopoietic stem cell frequency in patients' bone-marrow predicts acute myeloid leukemia outcome

Survival analyses of 58 ALDH-rare AML patients for

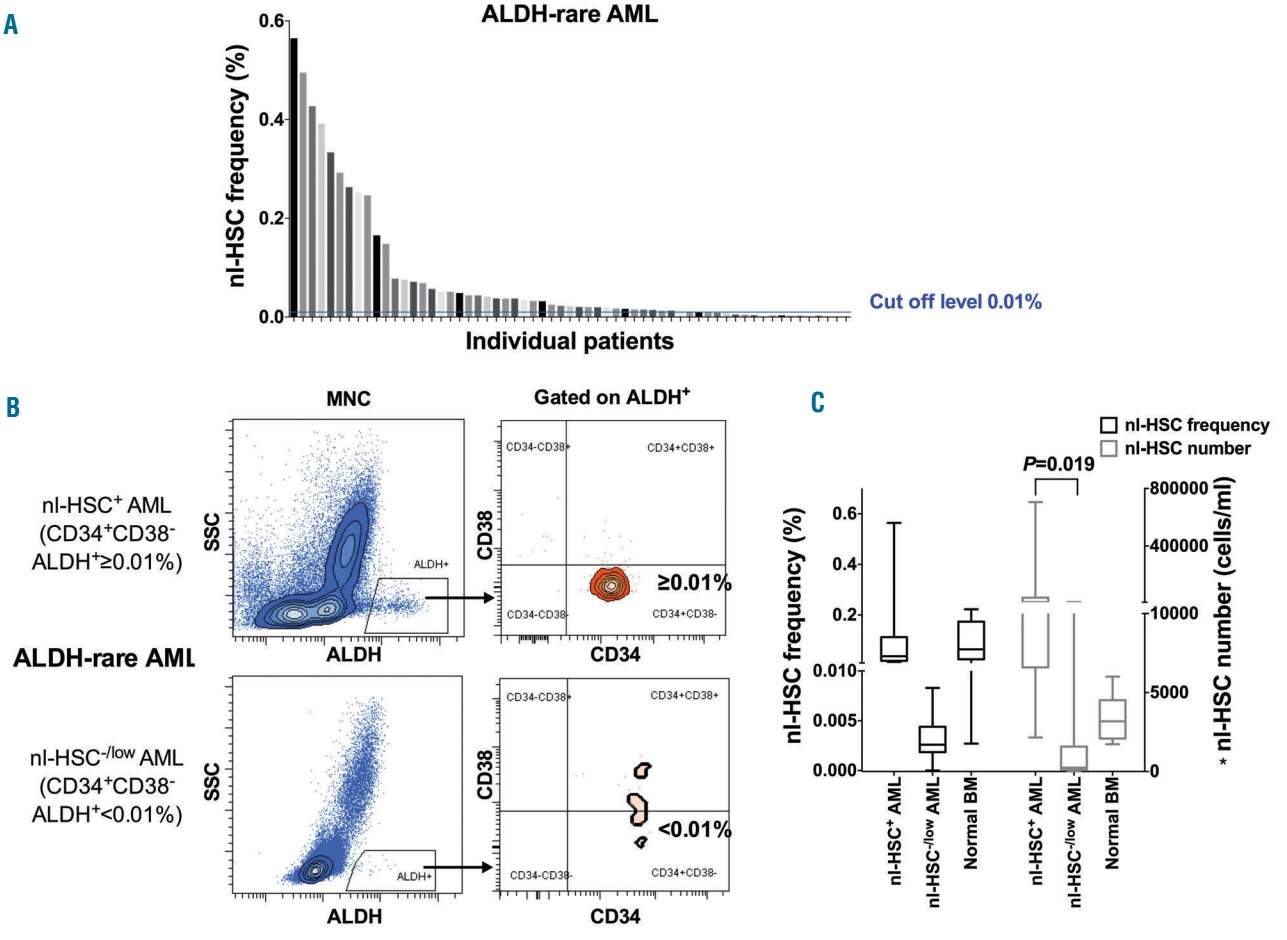

Figure 2. Frequencies of non-leukemic hematopoietic stem cells vary in ALDH-rare acute myeloid leukemia. (A) nl-HSC frequencies of all 61 studied patients [HSC% of total bone marrow (BM) MNC]. (B) ALDH-rare AML patients were stratified according to the frequencies of CD34$^+$CD38$^-$ALDH$^+$ cells into nl-HSC$^+$ AML (≥0.01% of total) and nl-HSC$^{-/low}$ AML (<0.01% of total). Representative FACS plots of each subgroup are shown. (C) CD34$^+$CD38$^-$ALDH$^+$ cells/mL frequencies displayed as percentages of total MNC and cell numbers as cells/mL of 45 nl-HSC$^+$ AML (median: 0.037%), 16 nl-HSC$^{-/low}$ AML (median: 0.0026%) and 11 healthy bone marrow controls (median: 0.064%). nl-HSC numbers of nl-HSC+ AML (median: 10623) were significantly higher compared to nl-HSC$^{-/low}$ AML (median: 240) (P<0.05). * Only patients with available nl-HSC numbers were included [35 nl-HSC$^+$ AML, 12 nl-HSC$^{-/low}$ AML and 6 healthy bone marrow controls (median: 3176)].

whom we had survival data showed a correlation between nl-HSC frequencies at the time of diagnosis and survival. HSC$^{-/low}$ AML patients had extremely poor OS with a median OS of 383 days and median disease-free survival (DFS) of 197 days, while the median OS and DFS have not been reached for patients with nl-HSC^{+} AML. Mean values for OS and DFS were substantially longer for nl-HSC^{+}AML patients compared to nl-HSC$^{-/low}$ AML patients (mean OS: 1177 days *versus* 414 days, mean DFS: 1214 days *versus* 284 days) (Figure 3A). The mutation status of patients with nl-HSC$^{-/low}$ and nl-HSC^{+} AML did not explain the observed differences in survival (*Online Supplementary Table S1*). However, most patients could be stratified into the intermediate-risk category in which there are huge variations in patients' outcomes. We therefore analyzed the outcome of 39 patients for whom we had survival data categorized in the cytogenetic intermediate-risk group, comparing nl-HSC$^{-/low}$ and nl-HSC^{+} AML. This analysis also identified the nl-HSC$^{-/low}$ AML cohort as a subpopulation with higher risk of relapse, with a median OS of 361 days (mean OS: 376 days) and median DFS of 150 days (mean DFS: 272 days), compared to nl-HSC^{+} AML for which the respective median OS has not been reached and the median DFS is 1248 days. Again, mean values for OS and DFS were substantially longer for nl-HSC^{+} AML than for nl-HSC$^{-/low}$ AML (mean OS: 1218 days *versus* 376 days, mean DFS: 1115 days *versus* 272 days) (Figure 3B). Comparative survival analysis of patients with nl-HSC^{+} and nl-HSC$^{-/low}$ AML who underwent allogeneic HSCT revealed that

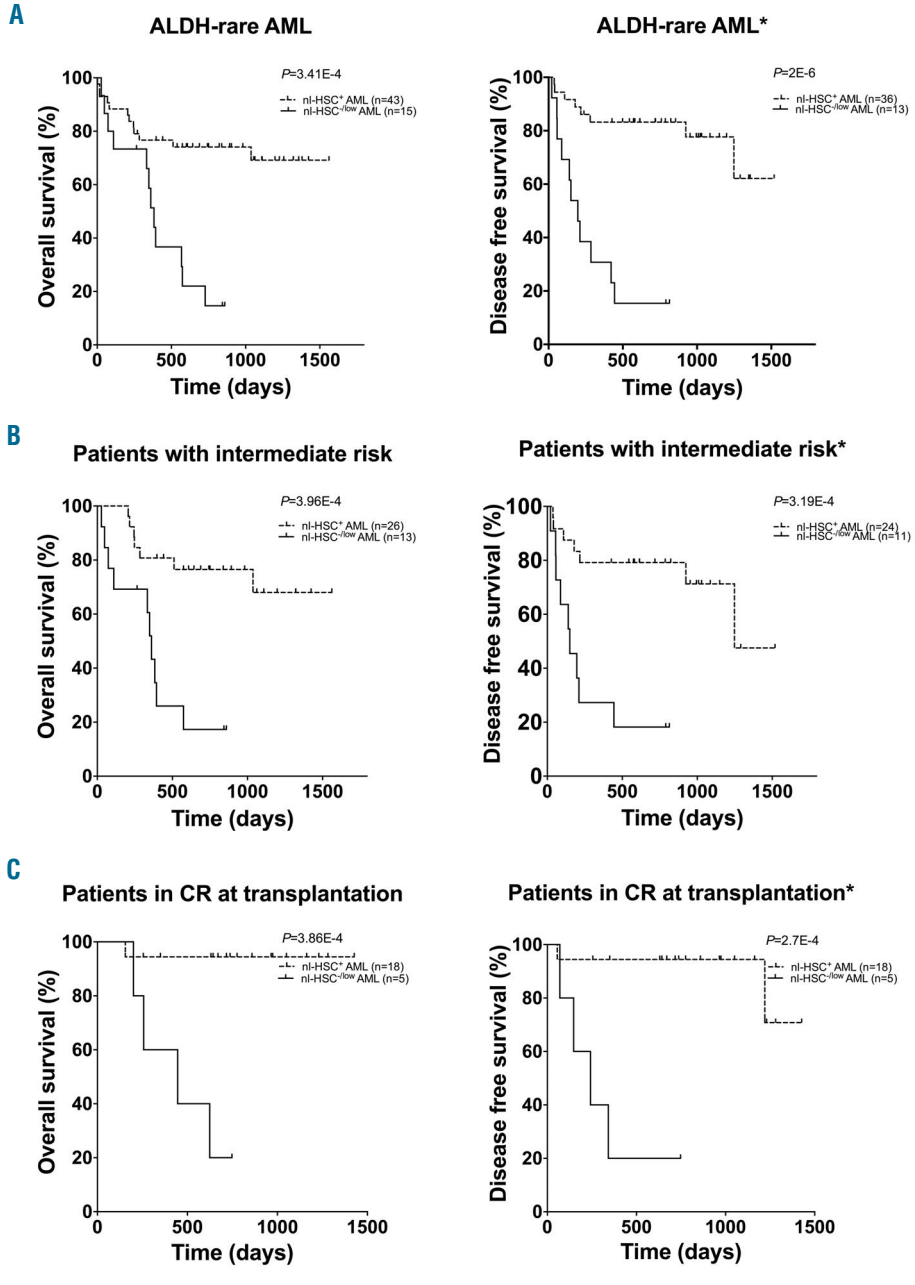

Figure 3. Patients with nl-HSC$^{-/low}$ acute myeloid leukemia have extremely poor survival. (A) Survival analysis revealed significantly shorter overall and disease-free survival for patients with nl-HSC$^{-/low}$ AML compared to those with nl-HSC^{+} AML [OS: nl-HSC^{+} AML (n=43) nl-HSC$^{-/low}$ AML (n=15); DFS nl-HSC^{+} AML (n=36) nl-HSC$^{-/low}$ AML (n=13)]. (B) Within the cytogenetic intermediate-risk group, nl-HSC$^{-/low}$ AML also represent a poor prognosis group [OS: nl-HSC^{+} AML (n=26) nl-HSC$^{-/low}$ AML (n=13); DFS nl-HSC^{+} AML (n=24) nl-HSC$^{-/low}$ AML (n=11)]. (C) For patients undergoing allogeneic HSCT stratification into nl-HSC$^{-/low}$ and nl-HSC^{+} AML identifies nl-HSC$^{-/low}$ AML as a cohort with poor therapy response (P<0.05) [OS: nl-HSC^{+} AML (n=18) nl-HSC$^{-/low}$ AML (n=5); DFS nl-HSC^{+} AML (n=18) nl-HSC$^{-/low}$ AML (n=5)]. *Patients who never achieved a complete remission (CR) were excluded from the DFS analysis.

even transplantation in complete remission was not sufficient to prevent relapse in most nl-HSC$^{-/low}$ cases. For patients with nl-HSC$^{-/low}$ AML, the median OS was 446 days (mean OS: 455 days) and the median DFS 242 days (mean DFS: 310 days), while median OS and DFS have not been reached for patients with nl-HSC^{+} AML (mean OS: 1356 days, mean DFS: 1302 days) (Figure 3C).

Non-leukemic hematopoietic stem cell frequency predicts leukemic *versus* multi-lineage human engraftment in NSG mice and correlates with non-leukemic hematopoietic stem cell *in vitro* function

To test whether nl-HSC frequencies are correlated with leukemia-initiating potential we transplanted 39 diagnostic AML cases from our cohort of ALDH-rare AML (28 nl-HSC^{+} AML and 11 nl-HSC$^{-/low}$ AML) in immune-deficient NSG mice (Figure 4A).

Most nl-HSC^{+} AML (24/28 cases) gave rise to multi-lineage engraftment and only rarely resulted in AML (3/28 cases) or non-engraftment (1/28 case). In contrast, nl-HSC$^{-/low}$ AML always gave rise to abnormal engraftment with 5/11 cases leading to AML and 6/11 cases not engrafting at all (Figure 4A). To test whether nl-HSC derived from nl-HSC$^{-/low}$ AML are functionally impaired we analyzed and compared the *in vitro* potential of CD34^{+}CD38^{-}ALDH^{+} cells derived from nl-HSC$^{-/low}$ and nl-HSC^{+} AML. Results from LTC-IC assays showed that LTC-IC frequencies of nl-HSC^{+} AML (1 in 2 cells) were comparable to those of normal bone marrow and higher than those of nl-HSC$^{-/low}$ AML (1 in 7 cells) (Figure 4B). In line with these results, the colony-forming potential of CD34^{+}CD38^{-}ALDH^{+} cells was also higher compared to that of cells derived from nl-HSC$^{-/low}$ AML suggesting a functional impairment of nl-HSC in nl-HSC$^{-/low}$ AML (Figure 4C).

Non-leukemic hematopoietic stem cell frequency in patients' bone marrow correlates with acute myeloid leukemia disease status

To understand the behavior of nl-HSC during therapy, matched sample series from diagnosis and various remission time points were analyzed and the results correlated with those of routine bone marrow and MRD analysis, if available (for detailed characteristics of the patients, see *Online Supplementary Table S2*). Initially, we focused on four patients who achieved durable remissions after chemotherapy, which were followed by allogeneic HSCT

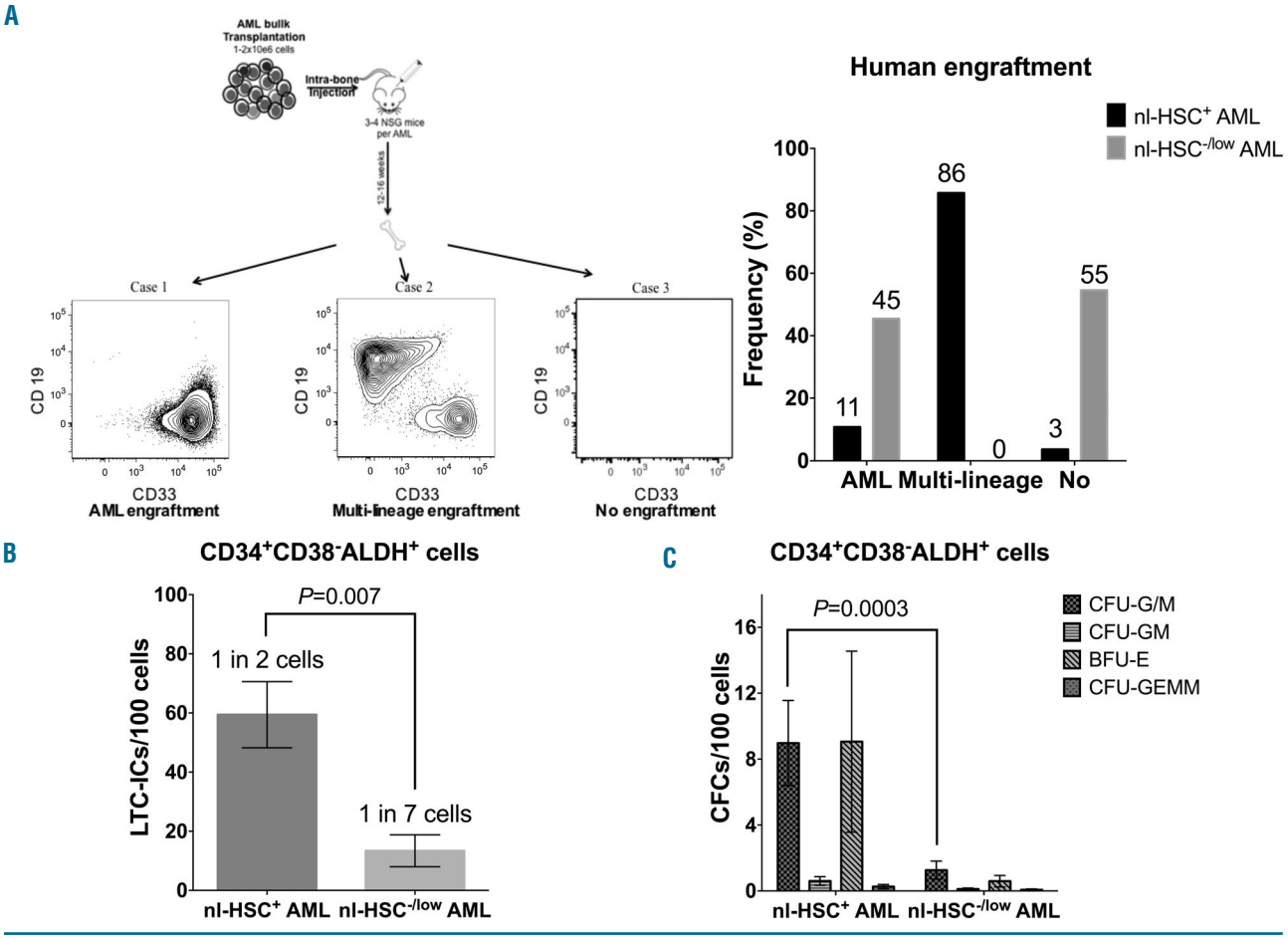

Figure 4. Frequencies of non-leukemic hematopoietic stem cells predict leukemic *versus* non-leukemic engraftment and negatively correlate with *in vitro* hematopoietic stem cell function. (A) Mouse transplantation strategy with examples of AML engraftment (case 1), multi-lineage engraftment (case 2) or no engraftment (case 3). nl-HSC^{+} AML mostly gave rise to multi-lineage engraftment (24/28) and rarely resulted in AML (3/28) or non-engraftment (1/28), whereas nl-HSC$^{-/low}$ AML only gave rise to abnormal engraftment with 5/11 cases leading to AML and 6/11 cases not engrafting at all. (B) LTC-IC frequencies of sorted CD34^{+}CD38^{-}ALDH^{+} cells derived from nl-HSC^{+} AML (n=13) and nl-HSC$^{-/low}$ AML (n=9) revealed impaired *in vitro* function of CD34^{+}CD38^{-}ALDH^{+} cells from nl-HSC$^{-/low}$ AML. Results are shown in LTC-IC per 100 plated cells. (C) Comparison of CFC frequencies of CD34^{+}CD38^{-}ALDH^{+} cells derived from nl-HSC^{+} AML (n=12) and nl-HSC$^{-/low}$ AML (n=10). Data are shown as mean ± SEM.

in three cases. We found a consistent pattern in all cases with nl-HSC proportions increasing over time (Figure 5).

As nl-HSC frequencies were measured by proportions of total MNC, which can be misleading for time points after chemotherapy with aplastic marrow conditions or diagnostic time-points with hyperplastic bone marrow, we also calculated nl-HSC numbers on the basis of available volume data (cells/mL). This analysis showed an increase of nl-HSC numbers in patients who achieved complete remissions (Figure 5A-D). MRD data were available for one patient and confirmed the observations that nl-HSC recover when disease burden goes down (Figure 5A). Some patients had persistent disease after the first course of induction chemotherapy which correlates with poor prognosis.[5] In these cases we found that nl-HSC did not recover (Figure 5B and *Online Supplementary Figure S4*). In one patient for whom we had many follow-up samples available we found that nl-HSC remained rare at time points of persistent disease and only recovered when a durable remission was reached after allogeneic HSCT (Figure 5B).

Decline of non-leukemic hematopoietic stem cell frequency predicts acute myeloid leukemia relapse

We next sought to examine the behavior of nl-HSC in patients with recurrent disease by analyzing five matched sample series from diagnosis, various remission time points and the time of relapse. In all analyzed samples nl-HSC frequencies and numbers decreased at the time of relapse (Figure 6). In this context, nl-HSC numbers (cells/mL) represent the more accurate measurement of nl-HSC compared to nl-HSC percentages as proportions can be skewed by large total cell numbers. Our analysis showed that nl-HSC numbers decreased during therapy and never recovered in patients who finally relapsed. For two sample series MRD data were available with molecular relapse preceding frank leukemic relapse. In both cases nl-HSC numbers and frequencies already decreased to low levels at the time of molecular relapse suggesting that reduced nl-HSC can predict leukemic relapse (Figure 6A,B). For the remaining two cases no MRD data were available. However, both cases finally relapsed with nl-HSC numbers and frequencies decreasing before frank relapses were detected (Figure 6C,D). For detailed nl-HSC numbers and frequencies during therapy see *Online Supplementary Table S3*.

Mathematical modeling suggests that competition of non-leukemic and leukemic stem cells for niche spaces is responsible for the decline of non-leukemic hematopoietic stem cells before relapse

According to our observations a decrease in nl-HSC numbers is associated with the presence of MRD and always precedes a full-blown AML relapse. Based on this

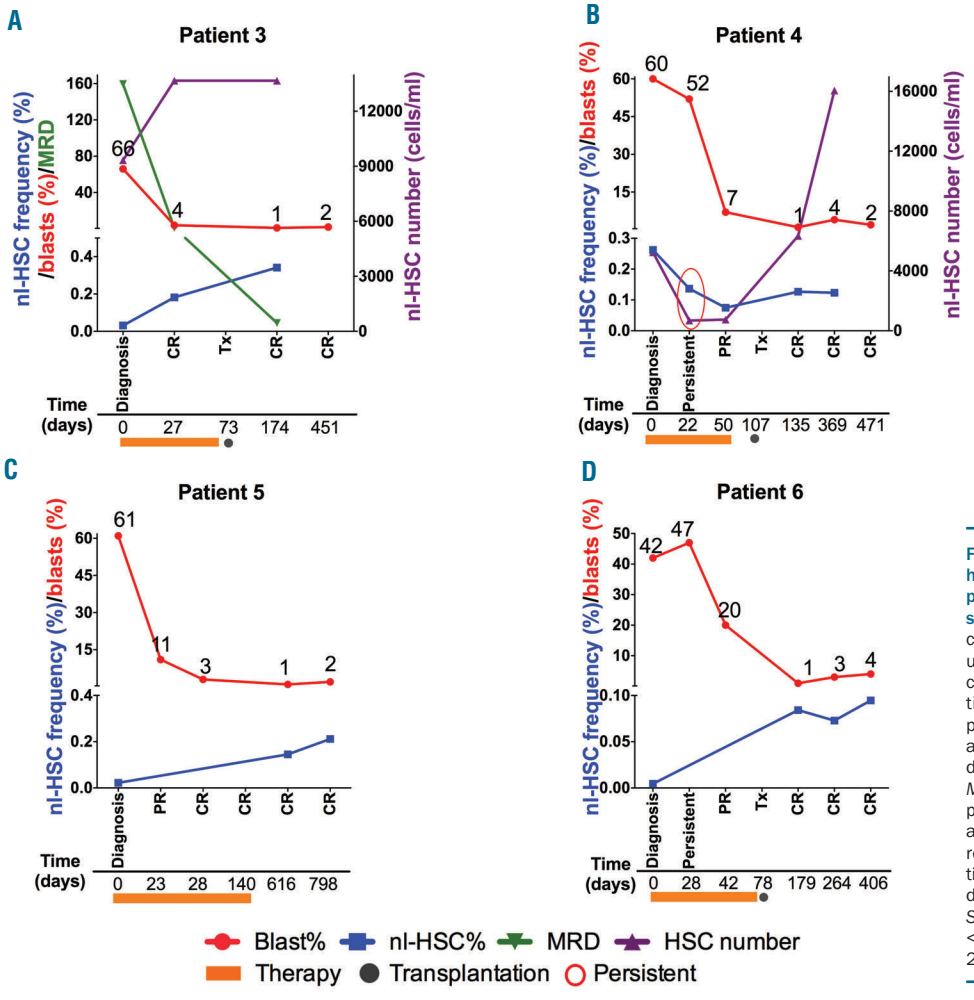

Figure 5. Frequencies of non-leukemic hematopoietic stem cells recover in patients who achieved complete remissions. Blast percentages and nl-HSC percentages at diagnosis and various follow-up time-points are shown with the percentage contribution of these populations in total bone marrow MNC of patients 3-6. nl-HSC numbers (if available) are shown as cells/mL and MRD data of patient 3 are shown as *CBFb-MYH11/ABL* ratio. Treatment times, time points of persistence and the event of allogeneic HSCT are indicated on the respective time line. Detailed characteristics on this and other AML patients are described in *Online Supplementary Table S2*. CR (complete remission): blast% <5%; PR (partial remission): blast% 5-25%; relapse: loss of CR with blasts ≥5%

observation, we hypothesized that this phenomenon reflects competition of nl-HSC and LSC for a limited number of bone marrow niche spaces. Part of this hypothesis is that LSC have a higher affinity than nl-HSC for such niches. This increased affinity enables LSC to dislodge nl-HSC from the protective niche and to occupy their spaces. Due to a lack of direct experimental confirmation, we developed a novel mathematical model to validate this hypothesis. The presented model can simulate different possible scenarios of interaction between leukemic and hematopoietic cells. These scenarios include: (i) competition of nl-HSC and LSC for a joint niche and dislodgement of nl-HSC by LSC; (ii) separate niches for LSC and nl-HSC, impairment of nl-HSC self-renewal and/or progenitor cell expansion by the leukemic cell bulk; (iii) competition of healthy and leukemic cells for systemic factors such as cytokines, including cytokine dependence of leukemic cells; and (iv) increased death rates of hematopoietic cells in the presence of leukemic cells. Notably, the early decline of nl-HSC before overt relapse was only reproduced by model simulations in the presence of a niche with dislodgement of nl-HSC by LSC (Figure 7A,B). In all other considered versions of the model, the leukemic blast count increases before the decline of nl-HSC counts, which is in contradiction to the clinical observations (Figure 7C,D, *Online Supplementary Figures S7* and *S8*). This modeling result, together with the clinical observations, supports the hypothesis that expansion of LSC in the bone marrow niche and dislodgement of nl-HSC are responsible for the decline of nl-HSC before frank relapse. This mechanism also explains the observed correlation of high MRD levels and reduced nl-HSC frequencies. Furthermore the model helps to understand possible functional differences of LSC in nl-HSC^{+} and nl-HSC$^{-/low}$ AML. Simulations indicate that the niche affinity of LSC in nl-HSC$^{-/low}$ AML is higher than in nl-HSC^{+}AML. This may explain the poor prognosis of nl-HSC$^{-/low}$ AML and its resistance to treatment (for details see the *Online Supplement*).

Discussion

Previous studies have shown that ALDH activity alone or in combination with other markers can be used to identify nl-HSC in a subgroup of AML patients.[6,18,26] In this context, we previously established a prospective stratification method based on the level of ALDH^{+} cells separating ALDH-numerous (≥1.9%) and ALDH-rare (<1.9%) AML. We showed that in ALDH-rare AML, nl-HSC can be separated from leukemic cells by using the CD34^{+}ALDH^{+}

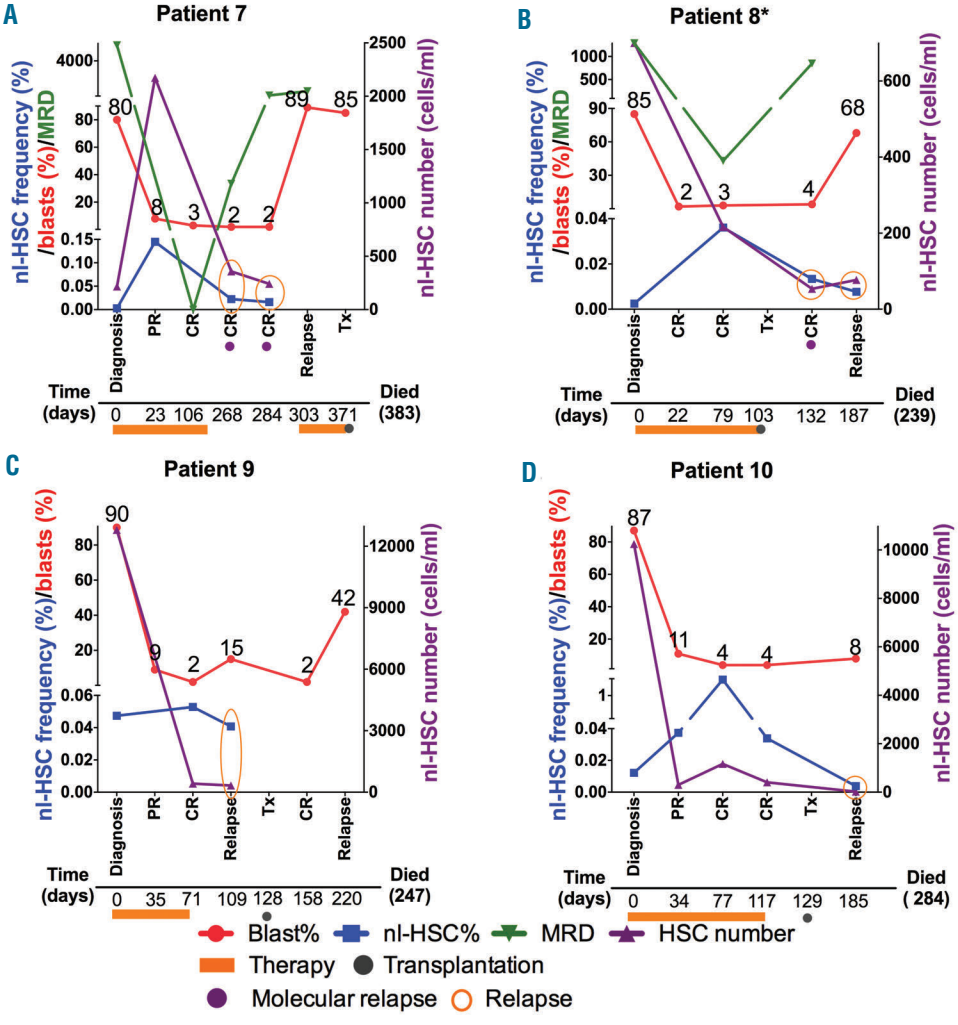

Figure 6. Frequencies of non-leukemic hematopoietic stem cells correlate with disease status and predict relapse. Blast percentages and nl-HSC percentages at diagnosis and various follow-up time-points are shown with the percentage contribution of these populations in total bone marrow MNC of patients 7, 8, 9, 10. nl-HSC numbers (if available) are shown as cells/mL and MRD data for patients 7 and 8 are shown as *NPM1/ABL* ratio. Treatment times, time-points of persistence/relapse and the event of allogeneic HSCT are indicated on the respective time line. Detailed characteristics of this and other AML patients are described in *Online Supplementary Table S2*. CR (complete remission): blast% <5%; PR (partial remission): blast% 5-25%; Persistent: blast% >25%; relapse: loss of CR with blasts ≥5%. * Patient for whom only peripheral blood was available at diagnosis.

phenotype.[14] The present study confirms these data and demonstrates that the $CD34^+CD38^-ALDH^+$ phenotype further enriches for HSC potential in ALDH-rare AML. However, despite these functional results it cannot be excluded that these nl-HSC already carry some leukemia-specific mutations. These so-called pre-leukemic HSC have been shown to be functionally normal with the ability to give rise to multi-lineage engraftment.[27,28]

In our detailed analyses of 61 patients with ALDH-rare AML we found highly variable nl-HSC frequencies including two cases with no detectable nl-HSC. These patients had extremely poor survival leading to the hypothesis that nl-HSC correlate with outcome in AML. We, therefore, stratified all the patients with ALDH-rare AML, based on their nl-HSC frequencies, into those with nl-$HSC^{-/low}$ (<0.01%) or nl-HSC^+ (≥0.01%) AML and found that this stratification was correlated with disease outcome.

In this analysis, patients with nl-$HSC^{-/low}$ AML had poor OS and DFS which were significantly worse than those of the group of nl-HSC^+ AML patients. Risk stratification is particularly important for the identification of patients who require more intensive treatment strategies, such as allogeneic HSCT. Classical risk stratification relies on cytogenetic markers that reliably distinguish patients with a good from those with a poor prognosis. However, some patients do not display either good- or poor-risk cytogenetics resulting in a large cohort of cytogenetic intermediate-risk patients who have variable outcomes. The uncertainty in risk prediction for so many patients highlights the need to improve existing risk stratification approaches.[1] We, therefore, also analyzed the cytogenetic intermediate-risk group according to our nl-$HSC^{-/low}$ and nl-HSC^+ AML criteria and detected a clear separation in outcomes with extremely poor survival for nl-$HSC^{-/low}$ and very good outcome for nl-HSC^+ AML. We also found that nl-$HSC^{-/low}$ AML do not just have poor survival but are also at such a high risk of relapse that allogeneic HSCT in complete remission appears to have virtually no effect on DFS. These results suggest distinct biological properties of these subgroups of patients from the cohort of ALDH-rare AML and may explain the differences observed in our xenotransplantation assays. More importantly and also clinically relevant is the conclusion that these patients seem to require new treatment strategies in order to avoid leukemic relapse.

There are a number of possible explanations for the poor survival of patients with nl-$HSC^{-/low}$ AML but the most plausible is competition for identical bone marrow niches between nl-HSC and LSC which has already been suggested for acute lymphobastic leukemia and AML.[29-31]

In our ALDH-rare AML representative xenotransplantation assays we found that mice transplanted with cells from nl-$HSC^{-/low}$ AML either failed to engraft or gave rise to AML which corresponded to their low nl-HSC frequencies. In contrast, nl-HSC^+ AML almost exclusively gave rise to normal multi-lineage engraftment as might be expected from samples with high nl-HSC frequencies. However, this difference was most likely not just a result of increased nl-HSC frequencies but also due to impaired HSC functionality of nl-$HSC^{-/low}$ AML which we demonstrated *in vitro*. The observed *in vitro* functional impairment might be the result of pre-leukemic mutations which have been shown to have variable penetrance within the nl-HSC compartment affecting up to all tested nl-HSC.[27] For *DNMT3A*-mutated nl-HSC a competitive growth advantage has been shown in xentotransplantation assays but for most pre-leukemic mutations the exact functional impact remains unclear.[28] It is, therefore, possible that in nl-$HSC^{-/low}$ AML the burden of pre-leukemic mutations is higher with resulting functional impairment. In order to clarify this observation more detailed functional studies at the single cell level combined with genetic analysis are needed.

As transplantation of leukemic bulk mimics competition of nl-HSC and LSC for niche spaces *in vivo*, we can infer from our data that in nl-$HSC^{-/low}$ AML LSC do indeed outcompete nl-HSC. However, this conclusion is limited by the fact that nl-HSC numbers are reduced and we do not

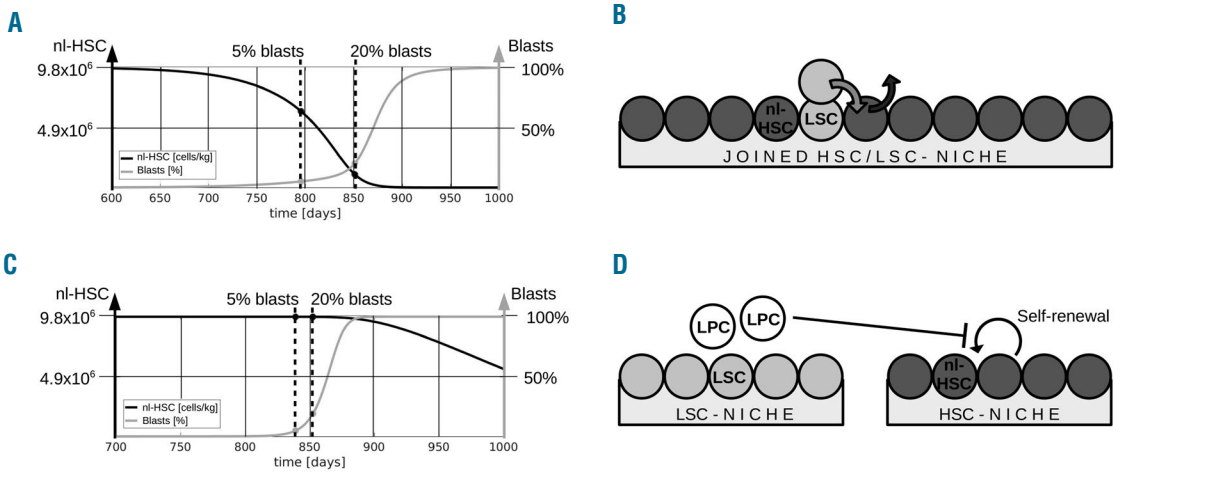

Figure 7. Mathematical modeling suggests that leukemic stem cell niche affinity is responsible for relapse and early decrease of non-leukemic hematopoietic stem cell numbers. (A) Simulation of nl-HSC counts and blast fractions reproduce the early decrease of HSC before overt relapse. (B) The simulations are based on the assumption that LSC and nl-HSC share identical stem cell niche spaces. Daughter LSC emerging from divisions can dislodge nl-HSC from the niche and occupy their spaces. The dislodged nl-HSC differentiate and lose HSC potential. (C) Scenarios without direct niche competition and without nl-HSC dislodgement cannot reproduce the early decrease of nl-HSC counts. (D) The simulation depicted in (C) is based on the assumption that nl-HSC and LSC reside in different niches and that leukemic cells inhibit nl-HSC self-renewal.

know the exact LSC number for the cases we transplanted.

A better and less artificial way to approach competition between nl-HSC and LSC is the analysis of serially collected samples from patients undergoing chemotherapy. In this analysis we found clear patterns for patients who achieved long-term remissions and those suffering from relapse. In cases of long-term remission nl-HSC frequencies and numbers recovered, even when strongly suppressed initially. In contrast, in patients who relapsed after a short period of remission nl-HSC frequencies and numbers decreased when leukemic activity was detectable, even at time points when only MRD could be measured. Suppressed nl-HSC numbers were always followed by leukemic relapse with some cases in which nl-HSC never recovered and other cases in which nl-HSC decreased after a short time of recovery. Comparison of matched diagnostic and relapse samples showed that nl-HSC numbers (cells/mL) were always massively reduced at frank relapse compared to diagnostic time-points, which may be the result of clonal selection of more aggressive subclones (*Online Supplementary Table S3*).

To support our hypothesis of nl-HSC and LSC competition for identical bone marrow niches we developed a mathematical model that enables clinical data to be linked to processes in the stem cell niche that cannot be observed directly. We used computer simulations of models describing different mechanisms to conclude that niche competition with active dislodgement of nl-HSC (independent of HSC cell divisions) is required for the early decline of nl-HSC as observed in this study. Scenarios lacking niche competition (e.g. separate niches to support nl-HSC and LSC) do not reproduce the experimental observation. Importantly, this result is independent of cytokine-dependence of leukemic cells. The dislodgement of nl-HSC from the niche explains the correlation of reduced nl-HSC counts and MRD. Model simulations also suggest that the niche affinity of leukemic cells is higher in nl-HSC$^{-/low}$ AML than in nl-HSC^{+} AML.

Translated into our setting, HSC-negativity would represent cases in which LSC have occupied *bona fide* HSC bone marrow niches, which have been described to be of limited number, thereby squeezing out nl-HSC.[32-34] Together with clinical data the modeling results lead to the conclusion that dislodgement of nl-HSC from the niche is a relevant mechanism in nl-HSC$^{-/low}$ AML

Supporting our hypothesis, real-time imaging studies performed in xenotransplantation assays demonstrated that acute lymphoblastic leukemia cells can displace nl-HSC from their marrow niche by inducing changes to the microenvironment.[30]

It remains unclear whether the effect of nl-HSC suppression is a distinct biological property or just a sign of disease stage. It might be that uncontrolled AML growth ultimately dislodges all HSC with the differences we observe just reflecting disease-stage at the time of diagnosis. However, judging from our data this seems to be unlikely as (i) some nl-HSC^{+} AML display extremely high white blood counts and (ii) nl-HSC suppression can be observed in follow-up and early relapse samples from time points when disease burden is still moderate. Our data show that at the time of hematologic relapse nl-HSC numbers decrease massively compared to the time of diagnosis with no nl-HSC-recovery between these time points. This suggests that in these cases LSC survived chemotherapy by occupying the protective bone marrow niches. Importantly, even allogeneic HSCT did not improve patients' overall poor outcome which was potentially caused by a niche occupation of therapy-resistant LSC. Supporting this hypothesis is the observation that engraftment of nl-HSC$^{-/low}$ AML patients who relapsed after allogeneic HSCT was delayed in comparison to that of nl-HSC^{+} AML patients who remained in remission after allogeneic HSCT. For this analysis we compared three of four possible nl-HSC$^{-/low}$ AML patients who relapsed after allogeneic HSCT. One patient was excluded because this patient only received a reduced conditioning regime with 2 Gy total body irradiation/fludarabine and never became aplastic. These patients were compared to 11 of 12 nl-HSC^{+} AML patients who remained in remission after allogeneic HSCT for whom engraftment data were available. For detailed information see *Online Supplementary Table S5* and *Online Supplementary Figure S9*.

Recently a study has been published analyzing the effects of human AML xenotransplantation on murine HSC and progenitors.[35] The authors observed that after xenotransplantation of human AML in immune-deficient mice, murine HSC were not depleted but a differentiation block at the progenitor stage occurred. These results were confirmed by analyzing a cohort of primary human AML samples.

There are important differences between our study and the above-described results. We do not consider that a xenotransplantation model examining murine hematopoiesis can be used to draw conclusions on human hematopoiesis in the human bone marrow. Another difference between the studies is that they only used a certain subgroup of AML with very low CD34 expression for their analysis of human bone marrow. In these leukemias, which are enriched for good prognosis nucleophosmin-mutant AML, non-leukemic cells can be isolated by sorting for CD34^{+} cells. In contrast, in our study we analyzed all subsets of AML and did not limit the analysis to such a rare and good-risk enriched subtype. Especially in the poor prognosis nl-HSC$^{-/low}$ AML cohort we found a different scenario with decreased nl-HSC numbers. We therefore think that their conclusions cannot be generalized to all AML cases, and in particular to high-risk cases. Collectively these data suggest that AML show different behaviors in influencing non-leukemic hematopoiesis which seem to vary according to their aggressiveness and niche affinity.

One limitation of our study is that the results presented are only representative of the group of ALDH-rare AML in which HSC isolation is possible and which accounts for about 77% of all AML cases, whereas in ALDH-numerous cases this isolation is not feasible.[14] The reason for this is most likely that in these cases normal HSC and leukemic cells both exhibit high ALDH-activity. In order to use our approach for outcome prediction in all AML cases a different HSC-isolation strategy is necessary. Various approaches for HSC isolation from AML samples have been described.[27,28] However, these studies only analyzed small series and ALDH status was not reported. Future studies should therefore compare these approaches in order to test whether nl-HSC can be isolated from all AML and whether the observed behavior of nl-HSC is universal or specific to the subgroup of ALDH-rare AML.

Taken together our results suggest that nl-HSC can serve as a risk stratification tool and as qualitative marker of MRD in a subgroup of AML patients. These cases may

also provide a starting point to understand the interactions of LSC and the protective bone marrow niche.

Acknowledgments

This work was supported by research funding from the German Research Foundation DFG (SFB 873; subprojects A13 to CL, B04 to AT, B07 to ADH, B08 to AMC and Z02 to VE), the SyTASC consortium funded by the Deutsche Krebshilfe and the Dietmar Hopp Foundation (both to AT).

The authors thank Karina Borowski for her technical assistance (FACS), Michaela Brough for her performance of FISH, Sandra Blaszkiewicz for her help in mouse experiments, Katrin Miesala for the isolation of primary cells and help in tissue culture and Anke Diehlmann for the isolation and culture of MSC. We are grateful to all clinicians in the Hematology Department of Heidelberg University Hospital for their contribution to collection of bone marrow material and patients and healthy donors who participated in this study.

References

1. Grimwade D, Walker H, Oliver F, et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. Blood. 1998;92(7):2322-2333.
2. Lowenberg B, Downing JR, Burnett A. Acute myeloid leukemia. N Engl J Med. 1999;341(14):1051-1062.
3. Estey E, Dohner H. Acute myeloid leukaemia. Lancet. 2006;368(9550):1894-1907.
4. Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. J Clin Oncol. 2003;21(24):4642-4649.
5. Dohner 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.
6. Gerber JM, Smith BD, Ngwang B, et al. A clinically relevant population of leukemic CD34(+)CD38(-) cells in acute myeloid leukemia. Blood. 2012;119(15):3571-3577.
7. Papadaki C, Dufour A, Seibl M, et al. Monitoring minimal residual disease in acute myeloid leukaemia with NPM1 mutations by quantitative PCR: clonal evolution is a limiting factor. Br J Haematol. 2009;144(4):517-523.
8. Yanada M, Matsuo K, Suzuki T, Kiyoi H, Naoe T. Prognostic significance of FLT3 internal tandem duplication and tyrosine kinase domain mutations for acute myeloid leukemia: a meta-analysis. Leukemia. 2005;19(8):1345-1349.
9. Gruszka AM, Lavorgna S, Consalvo MI, et al. A monoclonal antibody against mutated nucleophosmin 1 for the molecular diagnosis of acute myeloid leukemias. Blood. 2010;116(12):2096-2102.
10. Gabert J, Beillard E, van der Velden VH, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. Leukemia. 2003;17(12): 2318-2357.
11. Buccisano F, Maurillo L, Del Principe MI, et al. Prognostic and therapeutic implications of minimal residual disease detection in acute myeloid leukemia. Blood. 2012;119(2):332-341.
12. Al-Mawali A, Gillis D, Hissaria P, Lewis I. Incidence, sensitivity, and specificity of leukemia-associated phenotypes in acute myeloid leukemia using specific five-color multiparameter flow cytometry. Am J Clin Pathol. 2008;129(6):934-945.
13. Inaba H, Coustan-Smith E, Cao X, et al. Comparative analysis of different approaches to measure treatment response in acute myeloid leukemia. J Clin Oncol. 2012;30(29):3625-3632.
14. Hoang VT, Buss EC, Wang W, et al. The rarity of ALDH(+) cells is the key to separation of normal versus leukemia stem cells by ALDH activity in AML patients. Int J Cancer. 2015;137(3):525-536.
15. Schuurhuis GJ, Meel MH, Wouters F, et al. Normal hematopoietic stem cells within the AML bone marrow have a distinct and higher ALDH activity level than co-existing leukemic stem cells. PLoS One. 2013;8(11):e78897.
16. Ran D, Schubert M, Pietsch L, et al. Aldehyde dehydrogenase activity among primary leukemia cells is associated with stem cell features and correlates with adverse clinical outcomes. Exp Hematol. 2009;37(12):1423-1434.
17. Ran D, Schubert M, Taubert I, et al. Heterogeneity of leukemia stem cell candidates at diagnosis of acute myeloid leukemia and their clinical significance. Exp Hematol. 2012;40(2):155-165.e151.
18. Pearce DJ, Taussig D, Simpson C, et al. Characterization of cells with a high aldehyde dehydrogenase activity from cord blood and acute myeloid leukemia samples. Stem Cells. 2005;23(6):752-760.
19. Whichard ZL, Sarkar CA, Kimmel M, Corey SJ. Hematopoiesis and its disorders: a systems biology approach. Blood. 2010;115(12):2339-2347.
20. Catlin SN, Busque L, Gale RE, Guttorp P, Abkowitz JL. The replication rate of human hematopoietic stem cells in vivo. Blood. 2011;117(17):4460-4466.
21. Wodarz D, Garg N, Komarova NL, et al. Kinetics of CLL cells in tissues and blood during therapy with the BTK inhibitor ibrutinib. Blood. 2014;123(26):4132-4135.
22. Stiehl T, Ho AD, Marciniak-Czochra A. The impact of CD34+ cell dose on engraftment after SCTs: personalized estimates based on mathematical modeling. Bone Marrow Transplant. 2014;49(1):30-37.
23. Stiehl T, Baran N, Ho AD, Marciniak-Czochra A. Cell division patterns in acute myeloid leukemia stem-like cells determine clinical course: a model to predict patient survival. Cancer Res. 2015;75(6):940-949.
24. Wagner W, Horn P, Castoldi M, et al. Replicative senescence of mesenchymal stem cells: a continuous and organized process. PLoS One. 2008;3(5):e2213.
25. Taussig DC, Vargaftig J, Miraki-Moud F, et al. Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. Blood. 2010;115(10):1976-1984.
26. Lioznov MV, Freiberger P, Kroger N, Zander AR, Fehse B. Aldehyde dehydrogenase activity as a marker for the quality of hematopoietic stem cell transplants. Bone Marrow Transplant. 2005;35(9):909-914.
27. Jan M, Snyder TM, Corces-Zimmerman MR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med. 2012;4(149):149ra118.
28. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014;506(7488):328-333.
29. Boyd AL, Campbell CJ, Hopkins CI, et al. Niche displacement of human leukemic stem cells uniquely allows their competitive replacement with healthy HSPCs. J Exp Med. 2014;211(10):1925-1935.
30. Colmone A, Amorim M, Pontier AL, Wang S, Jablonski E, Sipkins DA. Leukemic cells create bone marrow niches that disrupt the behavior of normal hematopoietic progenitor cells. Science. 2008;322(5909):1861-1865.
31. Sipkins DA, Wei X, Wu JW, et al. In vivo imaging of specialized bone marrow endothelial microdomains for tumour engraftment. Nature. 2005;435(7044):969-973.
32. Colvin GA, Lambert JF, Abedi M, et al. Murine marrow cellularity and the concept of stem cell competition: geographic and quantitative determinants in stem cell biology. Leukemia. 2004;18(3):575-583.
33. Czechowicz A, Kraft D, Weissman IL, Bhattacharya D. Efficient transplantation via antibody-based clearance of hematopoietic stem cell niches. Science. 2007;318(5854):1296-1299.
34. Ishikawa F, Yoshida S, Saito Y, et al. Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. Nat Biotechnol. 2007;25(11):1315-1321.
35. Miraki-Moud F, Anjos-Afonso F, Hodby KA, et al. Acute myeloid leukemia does not deplete normal hematopoietic stem cells but induces cytopenias by impeding their differentiation. Proc Natl Acad Sci USA. 2013;110(33):13576-13581.