A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia

W Zeijlemaker1, A Kelder1, YJM Oussoren-Brockhoff1, WJ Scholten1, AN Snel1, D Veldhuizen1, J Cloos1,2, GJ Ossenkoppele1 and GJ Schuurhuis1

INTRODUCTION

After achieving complete remission, many patients with acute myeloid leukemia (AML) experience a relapse, resulting in dismal outcome.1 It is generally accepted that chemotherapy-resistant leukemic cells, often referred to as leukemic stem cells (LSCs) or leukemia-initiating cells, are responsible for relapse. The frequency of such minimal or measurable residual disease (MRD) cells offers an important post-remission risk factor in AML.2–7 MRD, determined using immunophenotypical and/or molecular procedures, is implemented in risk classifications of many major AML trial groups. However, relapses do occur in a considerable number (20–70%) of patients with low/negative MRD levels, even in the intermediate-risk group.2–7 To account for these ‘false-negative’ cases, one may focus on these surviving LSCs that are at the basis of outgrowth of MRD cells to overt relapse. Although different cellular compartments (CD34+CD38+, CD34+CD38−) may contain LSCs,8–12 CD34+CD38− stem cells seem to be most therapy resistant and least immunogenic.13–16 This in accordance with the finding that CD34+CD38− or CD34+CD38dim LSC frequency is highly predictive for relapse in AML.9,17 In addition, CD34+CD38− LSC frequency adds important prognostic information to MRD assessment, reducing the number of false-negative MRD cases.9 Overall, the LSC assessment is instrumental in defining remission quality and predicting relapse risk. However, identification of LSCs can be challenging owing to the very low frequency of this stem cell population (frequency 0.2–625 cells per 10⁶ mononuclear cells).14,15 It is known that compared with CD34+CD38+ hematopoietic stem cells (HSCs), CD34+CD38− LSCs often aberrantly express cell surface markers,19 whereby patterns may be very heterogeneous. Markers expressed on LSCs include myeloid antigens CD13, CD33 and CD123,20,21 CLL-1 (also known as CLEC12A)22 and the lineage markers CD2, CD7, CD11b, CD14, CD15, CD19, CD22 and CD56.12,23 The expression of CD96, also known as Tactile, and T-cell Ig mucin 3 (TIM-3), is also higher on LSCs as compared with HSCs.24,25 Although CD44 is already highly expressed on HSCs, LSCs often show overexpression for CD44.26 Using aberrant markers and well-established functional assays like aldehyde dehydrogenase27–29 and the side population,30 the normal and neoplastic nature of such immunophenotypically defined CD34+CD38− LSCs and HSCs could be confirmed using molecular and cytogenetic assays in multiple samples.9,27,30,31 Moreover, current LSC studies have revealed that marker-positive CD34+CD38− cells show leukemia engraftment in different mouse models.12,20,22,24,26,32 Marker expression differs between and within patients9,20,22–24 and therefore different immunophenotypically defined LSC compartments may be associated with specific subpopulations showing different sensitivity for therapy.33,34 Owing to tumor heterogeneity, accurate flow cytometric LSC detection at time of diagnosis requires extensive antibody panels. In addition, besides molecular constitution,34,35 also immunophenotypes can change between diagnosis and relapse.36 As a consequence, a broader panel of markers is also required during follow-up, enabling the detection of populations emerging during disease. Recent studies have shown that HSCs can also harbor leukemia-specific mutations.37 Although these so-called pre-leukemic HSCs are non-leukemic, they are significantly different from real HSCs as they are supposed to be of importance in the process of leukemogenesis.38 In case the relapse aberrancies are very distinct from diagnosis, it can be suggested that the relapse evolved from such a ‘pre-leukemic’ clone.39,40 To facilitate complete and
accurate CD34+CD38− LSC detection in AML patients, often with small amounts of bone marrow (BM) cells available, we composed a single flow cytometric tube. Expression and usefulness of 15 different cell surface markers (apart from the backbone markers, CD45, CD34 and CD38) were compared in a large cohort of AML patients. This enabled the design of a single 8-color LSC detection tube consisting of a cocktail of six markers in one fluorescence channel combined with single channels for the backbone markers (CD45, CD34 and CD38) and stem cell markers (CD44, CD33 and CD123). Overall, we defined a single tube containing in total 13 antibodies, which enables CD34+CD38− LSC detection in a broadly applicable, less expensive and more efficient manner than the current detection strategies.

**PATIENTS AND METHODS**

**Patients**

In total 236 patients were screened for initial analyses at time of AML diagnosis. For data concerning marker stability 132 follow-up samples were acquired. All samples were gathered between 2010 and 2015. All patients had a cytopathologically confirmed diagnosis of AML according to the WHO (World Health Organization) classification (excluding acute promyelocytic leukemia) or a diagnosis of refractory anemia with excess of blasts and International Prognostic Scoring System, score > 1.5. Patients were treated according to HOVON/SAKK clinical trials (www.hovon.nl) and provided their written informed consent before entrance into the study (for approval numbers see Supplementary File). BM samples from patients. This enabled the design of a single 8-color LSC detection tool, a scoring system was used to, retrospectively, de

**Immunophenotyping**

Erythrocyte-lysed (Pharm lyse, Becton and Dickinson (BD; San Jose, CA, USA)) fresh BM samples were used to perform 8-color multiparameter flow cytometry using a FACS Canto II from BD. After lysis, cells were washed with phosphate-buffered saline/0.1% human serum albumin and labeled with the appropriate antibodies. Cells and antibodies were incubated for 15 min in the dark at room temperature and subsequently, cells were washed to remove unlabeled antibodies. In general a minimum of 500 000 white blood cells were acquired. Detailed information concerning the used antibodies is provided in the Supplementary Data. Table 1 shows the conventional 7-tube 8-color antibody panel at diagnosis.

**Gating strategy and marker selection**

Samples were analyzed using Infinicyt software, v. 1.7 (Cytognos, Salamanca, Spain). The gating strategy is shown in Supplementary Figures S1A–D. Within the blast population presumed HSCs were defined as CD34+/CD38−/SSClow/SSCint/CD45+/Lin−/CD123−/CD19−/CD122−/CD65−. Presumed LSCs were defined as CD34+−/CD38−+ and CD44+− and/or CD123−− and/or CD33−− and/or CD11−− and/or CD7−− and/or CD16−− and/or CD3−− and/or CD14−− and/or CD1−− and/or CD7−− and/or CD16−− and/or CD3−− as available available, CD45RA+ and CD44++ and/or CD123++ and/or CD33++ and/or CLL-1+, and/or CD19+ and/or CD14+. Following cell surface markers showed aberrant expression on CD34+CD38− LSCs, leukemic stem cells, after injection of this cell fraction in mice, indeed gave multilineage engraftment, while in mice injected with CD34+CD38+ LSCs, leukemic engraftment was seen. To determine which markers were needed for the LSC detection tool, a scoring system was used to, retrospectively, define and compare the usefulness of above mentioned cell surface markers. This scoring system takes into account background autofluorescence of individual markers that were sometimes measured in different fluorochrome channels and consequently had differences in brightness. Therefore, this scoring system was created to ensure a less subjective comparison as compared with solely investigating median fluorescence intensities (MFIs) of the different markers. This scoring system ranked from 0 to 3 points, whereby, in short, one point was attributed to a marker for (a) clear distinction of two populations within the CD34+CD38− fraction, (b) high-negative predictive value of the particular marker (no LSC pollution in marker-negative fraction) and (c) high sensitivity (maximal LSC coverage; Figure 1). Based on the number of points given, one (or sometimes multiple) best marker(s) were defined in each AML sample. The details are outlined in the legend of Supplementary Figure S1. If a marker scored three points, it was by definition a ‘best marker’. In the minority of patients (8/131), however, the ‘best marker’ scored < 3 points. The scores were used to compare overall marker prevalence and usefulness regarding LSC detection in AML in 131 cases in which enough LSCs were present (cluster of ≥ 5 cells) to evaluate marker performance of the in total 236 screened cases. For examples and more detailed information concerning this scoring system, see Supplementary Figure S1.

**RESULTS**

Prevalence of aberrant cell surface markers on CD34+CD38− cells at diagnosis

The aim of this study was to design a single 8-color tube that would allow specific identification of CD34+CD38− LSCs. For 15 different cell surface markers, we compared general prevalence, regardless of the percentage and intensity of expression on CD34+CD38− cells. A marker was defined as present on CD34+CD38− when it scored at least one point according to the scoring system described in the methods section. Figure 2 shows that CD123 and CD33 were most often aberrantly expressed with a prevalence of 82.3% (107/130) for both these markers. Furthermore, the following cell surface markers showed aberrant expression on CD34+CD38− cells in more than half of the cases: CD11− (70%), TIM-3 (62%), CD11b (55%) and CD22 (51%). In addition, CD7 (43%), CD96 (33%), CD56 (32%) and CD15 (30%) were all markers that were present between 30 and 50% of the evaluated cases. CD44 is an exceptional marker as it is highly expressed on HSCs and indeed CD44 was found to be expressed on the total CD34+CD38− compartment in 100% of the evaluated AML cases (n = 131, example in Supplementary Figure S1). Ablent CD44 expression.

<table>
<thead>
<tr>
<th>Tube</th>
<th>FITC</th>
<th>PE</th>
<th>PerCP-CY5.5</th>
<th>PC7</th>
<th>APC</th>
<th>APC-H7</th>
<th>HV450</th>
<th>HV500c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PBS</td>
<td>PBS</td>
<td>CD13</td>
<td>PBS</td>
<td>PBS</td>
<td>CD34</td>
<td>CD45</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>CD44</td>
<td>CD34</td>
<td>CD33</td>
<td>CD17</td>
<td>CD38</td>
<td>CD19</td>
<td>CD34</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>CD7</td>
<td>TIM-3</td>
<td>CD17</td>
<td>CD19</td>
<td>CD38</td>
<td>CD14</td>
<td>CD34</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>CD2</td>
<td>CD133</td>
<td>CD33</td>
<td>CD19</td>
<td>CD38</td>
<td>CD14</td>
<td>CD34</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>CD36</td>
<td>CD123</td>
<td>CD33</td>
<td>CD19</td>
<td>CD38</td>
<td>CD14</td>
<td>CD34</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CD11b</td>
<td>CD96</td>
<td>CD17</td>
<td>CD38</td>
<td>CD14</td>
<td>CD34</td>
<td>CD34</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>CD15</td>
<td>CD22</td>
<td>CD17</td>
<td>CD38</td>
<td>CD14</td>
<td>CD34</td>
<td>CD34</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. 8-Color antibody panel

Abbreviations: APC, allophycocyanin; FITC, fluorescein isothiocyanate; PBS, phosphate-buffered saline; PC7, phycoerythrin cyanin 7; PE, phycoerythrin; PerCP, peridinin chlorophyll protein complex.
CD34+CD38− markers (CLL-1 and CD7) show two different populations within CD34+CD38− fraction. a and b show two different patients whereby both markers (CLL-1 and CD7) show two different populations within CD34+CD38−. Section II shows the second criterion: (b) high-negative predictive value of the marker. CLL-1 expression in patient #1711 is shown (a) in which a clear population of HSCs is present (based on both marker negativity and FSClow/SSClow scatter properties) as shown in b). CLL-1 has only little expression on CD34+CD38− cells and all marker-negative cells are presumed HSCs, implying that there are no/few LSCs present within the CLL-1-negative CD34+CD38− fraction. Finally, section III shows the last criterion: (c) high sensitivity of the marker. Both CD33 (a) and TIM-3 (b) are highly positive on CD34+CD38− cells and almost all LSCs were marker positive. In case the marker of interest fulfilled one of the criteria, one point was given; thereby the total number of points in the scoring system ranged from 0 to 3 points. Gating strategy of CD34+CD38− cells and examples of how scores were defined are shown in Supplementary Figure S1.

Figure 1. Scoring system criteria to evaluate marker performance. Marker-positive and marker-negative cells are shown in red and green, respectively. Section I shows the first scoring system criterion: (a) clear distinction between two populations within the CD34+CD38− fraction. a and b show two different patients whereby both markers (CLL-1 and CD7) show two different populations within CD34+CD38−. Section II shows the second criterion: (b) high-negative predictive value of the marker. CLL-1 expression in patient #1711 is shown (a) in which a clear distinction between presumed CD34−CD38− HSCs and LSCs. TIM-3 (24%) and CD11b (16%) had three points in > 10% of the samples, followed by CD22 (10%), CLL-1 (9%), CD44 (8%), CD96 (8%), CD56 (6%), CD13 (5%) and CD7 (4%). In addition, CD15 (2%), CD2 (2%) and CD19 (2%) were very useful in only few cases. Finally, CD14 had no points scored at all. Note that > 1 marker may score three points in the same AML case.

Essential markers for identifying CD34+CD38− LSCs

Of the 131 evaluable cases there were 48 cases (37%) in which only one best marker for LSC detection was present (not shown). In 47/48 cases the best marker scored three points, however in one case the best marker scored two points. These 48 cases thus provided the most essential information as to defining which particular markers are definitely needed in the single LSC detection tube and which markers were redundant. The definition of such antibodies is that, for each AML case studied, at least one of the other antibodies was better. Of these 48 cases, CD123 was best in 44% of the samples (21/48), CD33 in 33% (16/48), CLL-1 in 25% (12/48), TIM-3 in 24% (11/48), CD22 in 16% (8/48), CD56 in 14% (7/48), CD13 (5%) and CD7 (4%). Consequently, these six markers were considered essential in the formation of a diagnostic tool with the aim to specifically detect LSCs in a maximal number of patients. Therefore, CD2, CD7, CD11b, CD13, CD14, CD15, CD19, CD44 and CD96 were redundant. However, seen the high incidence of CD7 and CD11b (Figure 2), we left these markers in the study for further experiments. CD13 was not only highly expressed on HSCs, but also showed low incidence (Figure 2). This marker was therefore excluded from further analysis.

CD33 and CD123 are expressed on CD34+CD38− HSCs in follow-up BM

To enable specific LSC detection in a follow-up BM it is of importance that marker expression on stem cells remains stable during disease and/or treatment. At follow-up, in most cases the majority of CD34+CD38− cells are presumed HSCs. Upregulation

characterized by an overexpression of CD44 (usually MFI of CD44++ on CD34+CD38− > 25 000) was found in 31% (40/131) of the cases (Figure 2). Finally, CD13 (29%), CD2 (18%), CD19 (8%) and CD14 (2%) were aberrantly expressed in < 30% of the cases.

CD33 and CD123 perform best in detection of leukemic and normal CD34+CD38− cells

Performance of the individual markers for LSC detection was investigated using the scoring system described in the methods section. Figure 2 gives an overview of the scores of the different markers. CD123 and CD33 not only had the highest prevalence, but scored the maximum number of three points in 60 and 59% of all evaluable cases, respectively. This implies that these markers not only had high expression levels, but also most often showed a clear distinction in antigen expression between presumed CD34+CD38− HSCs and LSCs. TIM-3 (24%) and CD11b (16%) had three points in > 10% of the samples, followed by CD22 (10%), CLL-1 (9%), CD44 (8%), CD96 (8%), CD56 (6%), CD13 (5%) and CD7 (4%). In addition, CD15 (2%), CD2 (2%) and CD19 (2%) were very useful in only few cases. Finally, CD14 had no points scored at all. Note that > 1 marker may score three points in the same AML case.
on HSC of markers used at diagnosis for LSC identification might lead to largely over-estimated LSC frequencies. Table 2 and Supplementary Table S1 show that marker expression on HSCs of all markers is low at diagnosis (CD44 was excluded from these experiments as CD44 is highly expressed on HSCs both at diagnosis and follow-up). High MFI ratios are seen for CD123 in follow-up BM (Table 2), suggesting an upregulation of this marker on HSCs after treatment, especially since CD123 is expressed, albeit at low levels, in normal bone marrow (NBM). Consistent with previously published results, we found CD33 to be clearly expressed on HSCs in NBM (Supplementary Table S1). Table 2 shows that CD33 expression is strongly suppressed on HSCs present in AML diagnosis BM as compared with HSCs in NBM (P < 0.001) and follow-up BM (P < 0.001). Supplementary Figure S2 shows the MFI ratios of Table 2 including MFI values of both the CD34+CD38+ HSCs and lymphocyte sub-fractions. This figure shows that the high MFI ratios as found for CD33 and CD123 in follow-up BM and in the normal control cases cannot be explained by differences in background expression of the control population (lymphocyte subpopulations). Although for many markers MFI ratios are significantly different between the different type of samples, the wide range of ratios as seen for CD33 and CD123 in both AML follow-up and NBM (Table 2) shows that although CD33 and especially CD123 are markers that are important to discriminate LSC and HSC at diagnosis, they should be used with care at follow-up.

Composition and validation of CD34+CD38+ LSC detection tool

Based on above described results concerning performance, prevalence, redundancy and stability on HSCs of the different markers, one tube could be designed that enables specific LSC detection in almost all AML cases, both at time of diagnosis and follow-up. CLL-1, TIM-3, CD7, CD11b, CD22, CD33, CD56 and CD123, chosen for specific LSC detection, had to be included in the LSC tube, together with the backbone markers CD45, CD34 and CD38, making up 11 antibodies. To enable such for a single 8-color tube, we tested whether we could combine multiple markers in one fluorescence channel. To enable this for both diagnosis and follow-up, a prerequisite for the combination of these markers was to be and remain marker negative on HSCs.

Therefore, a cocktail of such stable cell surface markers, including CLL-1, TIM-3, CD7, CD22, CD56 and CD11b, was used in phycoerythrin (PE)-fluorescence channel, which usually provides strong fluorescence for all single markers. Cell surface markers CD33 and CD123 are useful for follow-up despite upregulation of expression on HSCs, since expression differences between LSC and HSC remain due to usual overexpression on LSCs. Nevertheless, due to their instability, these markers should be evaluated in a separate fluorescence channel. Table 3 shows the newly developed single 8-color flow cytometry tube. Accuracy of the marker cocktail in the PE-channel of the one-tube panel was checked by comparing CD34+CD38+ LSC levels (as percentage of the total amount of white blood cells) with the conventional antibody panel from Table 1. In nine NBM samples it was validated that the marker cocktail of the single LSC tube indeed had no expression on HSCs (example in Figure 3a; in all nine samples there was 0% positivity on CD34+CD38+ cells; median MFI ratio of marker cocktail on CD34+CD38+ HSCs was 1.06, range 0.59–1.54). In addition, in six pathological control samples also no expression on HSCs was found (median 0% positivity, median MFI ratio 1.52 and range 1.09–1.92). Furthermore, as negative control, we verified in nine CD34-negative AML samples with only HSCs present,31 that no expression on CD34+CD38+ cells was found with the marker cocktail (example in Figure 3b; 0% positivity, n = 9; median MFI ratio of marker cocktail was 1.24, range 0.46–7.43). These control samples show that the LSC detection tube is highly specific. When comparing 22 AML samples, with CD34+CD38+ LSCs present (example in Figure 3c), a correlation coefficient of 0.968 between LSC measured with the best single marker in the cocktail and the six markers combined in PE-channel, was found (Figure 4a, P < 0.001). This clearly shows that for quantification, the combination of six markers is as good as the best individual marker and in 18 cases slightly better (Figure 4a; median difference in LSC% obtained with the marker cocktail and the best single marker is 0.001%, range 0.000–0.836%). Only in four cases we found that the marker cocktail had slightly lower

Table 2. Median MFI ratios of CD34+CD38+ HSC population as compared with lymphocytes

<table>
<thead>
<tr>
<th></th>
<th>AML diagnosis BM</th>
<th>AML follow-up BM</th>
<th>NBM</th>
<th>Diagnosis vs f-up</th>
<th>Diagnosis vs NBM</th>
<th>Follow-up vs NBM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median (range)</td>
<td>n</td>
<td>Median (range)</td>
<td>n</td>
<td>Median (range)</td>
<td>n</td>
</tr>
<tr>
<td>CLL-1</td>
<td>0.7 (0.1–2.1)</td>
<td>27</td>
<td>4.1 (0.3–14.0)</td>
<td>30</td>
<td>1.6 (0.3–3.0)</td>
<td>10</td>
</tr>
<tr>
<td>TIM-3</td>
<td>0.6 (0.2–4.9)</td>
<td>32</td>
<td>2.2 (0.3–10.8)</td>
<td>28</td>
<td>0.8 (0.6–1.6)</td>
<td>9</td>
</tr>
<tr>
<td>CD7</td>
<td>1.0 (0.4–3.1)</td>
<td>32</td>
<td>2.4 (0.4–5.9)</td>
<td>37</td>
<td>1.7 (0.2–3.6)</td>
<td>10</td>
</tr>
<tr>
<td>CD11b</td>
<td>1.0 (0.2–2.7)</td>
<td>24</td>
<td>3.1 (0.5–5.9)</td>
<td>22</td>
<td>2.5 (0.7–5.4)</td>
<td>9</td>
</tr>
<tr>
<td>CD22</td>
<td>2.0 (0.9–6.6)</td>
<td>12</td>
<td>3.5 (1.4–10.6)</td>
<td>9</td>
<td>5.4 (2.0–11.0)</td>
<td>9</td>
</tr>
<tr>
<td>CD33</td>
<td>2.1 (0.6–23.8)</td>
<td>30</td>
<td>5.5 (0.1–910.6)</td>
<td>35</td>
<td>19.4 (1.0–130.5)</td>
<td>21</td>
</tr>
<tr>
<td>CD56</td>
<td>1.1 (0.3–4.0)</td>
<td>25</td>
<td>4.3 (0.7–26.4)</td>
<td>24</td>
<td>2.5 (1.2–6.8)</td>
<td>9</td>
</tr>
<tr>
<td>CD123</td>
<td>3.9 (0.4–17.7)</td>
<td>30</td>
<td>14.5 (1.3–40.1)</td>
<td>25</td>
<td>5.8 (0.8–56.3)</td>
<td>20</td>
</tr>
</tbody>
</table>

Abbreviations: AML, acute myeloid leukemia; BM, bone marrow; HSC, hematopoietic stem cells; MFI, median fluorescence intensity; n, number of samples; NBM, normal bone marrow. P-values were calculated using the Mann–Whitney U-test. MFI ratios were calculated as follows: MFI of marker on CD34+CD38+ divided by MFI of marker on negative control population (lymphocytes).

Table 3. 8-Color LSC detection tube

<table>
<thead>
<tr>
<th>Tube</th>
<th>FITC</th>
<th>PE</th>
<th>PerCP-CY5.5</th>
<th>Pe-Cy7</th>
<th>APC</th>
<th>APC-H7</th>
<th>BV421</th>
<th>HV500c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CD45RA</td>
<td>CD1-1</td>
<td>TIM-3</td>
<td>CD7</td>
<td>CD11b</td>
<td>CD22</td>
<td>CD33</td>
<td>CD38</td>
</tr>
</tbody>
</table>

Leukemia stem cell tool in acute myeloid leukemia

W Zeijlemaker et al
expression than the highest individual marker (median difference in LSC% obtained with the best single marker and the marker cocktail is 0.459%, range 0.000–3.538%). An example of the efficacy of LSC detection using our single tube is shown in Figure 5. Supplementary Figures S3 and S4 illustrates the efficacy of the marker cocktail in case of a stable or, on the other hand, instable immunophenotype during disease progression. CD45RA was included in the LSC tube, not only to objectify possible clonal differences within the LSC fraction but also as it is an important marker for LSC detection in the majority of AML patients. Expression of CD45RA was studied in the same 22 samples used for the comparison of single markers with the cocktail (Figure 4a). LSC results obtained with the marker cocktail and CD45RA were in general comparable, however, the marker cocktail being slightly better than CD45RA (Figure 4b, Median difference between marker cocktail and CD45RA: 0.002%, range 0.000–1.996%). In the seven samples where CD45RA had higher LSC% the median difference between CD45RA and the marker cocktail was 0.010%, range 0.000–0.797%.

DISCUSSION

In this study we developed a broadly applicable one-tube approach to identify specific immunophenotypic CD34+CD38− LSC. Previous reports have demonstrated the prognostic importance of CD34+CD38− LSC frequency both at time of diagnosis and after treatment. Moreover, they further refine current MRD-based risk classification, hence improving the prediction of an emerging relapse. We speculate that the CD34+CD38− LSC frequency will be incorporated in future risk classification in multicenter trials and in monitoring allogeneic transplantation. Consequently, a CD34+CD38− LSC detection approach that is broadly applicable is very desirable. For reliable LSC detection marker expression should not only be stable on HSCs (Table 2, Supplementary Table S1, Supplementary Figure S2), but also on LSCs. However, losses and gains of individual antigens during disease have been frequently observed. For CLL-1 and the lineage markers it has been shown, in small patient groups (n = 2–9), that expression on CD34+CD38− was relatively stable between diagnosis and relapse. Using a marker cocktail in one fluorescence channel, as included in the LSC detection tube, any immunophenotypic shift between the markers in that channel does not affect accurate LSC measurement (Supplementary Figures S3 and S4). This LSC tube thus also anticipates on new LSC populations that can emerge during treatment and/or disease. Importantly, by using this marker cocktail we were able to achieve similar or even better results as compared with the conventional seven tube antibody panel (Figure 4a). Moreover, our LSC detection tube enables to include additional newly discovered LSC markers of interest, provided that these should not be upregulated on HSCs. Although not shown in this paper, the specificity of CD34+CD38− LSC detection can be further improved

Figure 3. Specificity of marker cocktail in the LSC detection tube. This figure shows the absence of expression of the marker cocktail on CD34+CD38− cells in NBM (a) and a CD34-negative AML (b) in this type of AML only CD34+CD38− HSCs are present). In c expression of the marker cocktail is shown in a CD34-positive AML in which both HSCs (marker cocktail negative) and LSCs (marker cocktail positive, 0.0586%) are present. Percentages represent the amount of marker-positive CD34+CD38− cells as compared with the total amount of white blood cells (WBCs).

Figure 4. Performance of marker cocktail: comparison with best single marker and with CD45RA expression. Comparison of LSC results (LSC percentage calculated as percentage of the total amount of white blood cell (WBC), log transformed) in 22 AML diagnosis samples with CD34+CD38− LSC present. Results are obtained with the best individual marker of the conventional panel (x-axis, a) and obtained with the marker cocktail (y-axis, a). The dashed gray line represents the y = x line. (b) shows a comparison of LSC results obtained with the marker cocktail and CD45RA, a potential stem cell marker. Samples are ordered according to increasing percentages obtained with the cocktail.
Leukemia stem cell tool in acute myeloid leukemia

W Zeijlemaker et al


by using secondary gating strategies based on differences in CD34 and/or CD45 expression and differences in light scatter properties. Importantly, in this paper we define HSCs both on the absence of aberrant marker expression and specific light scatter characteristics (FSC<sub>low</sub>/SSC<sub>low</sub>). We have demonstrated that injection of such defined HSCs indeed gave multilineage engraftment in mice. It has however to be emphasized that recent studies have shown that these HSCs can harbor leukemia-specific mutations such as IDH1 and IDH2, TET2 and DNMT3a. The establishment of the role of these so-called pre-leukemic HSCs in leukemogenesis and relapse development needs further investigation. Furthermore, it must be taken into account that leukemia-initiating capacity is not only present in CD34<sup>+</sup>CD38<sup>-</sup> cells, but also in CD34<sup>+</sup>CD38<sup>-</sup> and CD34<sup>-</sup> immunophenotypic compartments. However, the CD34<sup>+</sup>CD38<sup>-</sup> population in a CD34<sup>-</sup>-positive AML has the highest engraftment potential in mice and is most therapy resistant in vitro and in vivo. In addition, in diagnosis AML there was a very strong prognostic value for CD34<sup>+</sup>CD38<sup>-</sup> LSCs, but no impact whatsoever for CD34<sup>+</sup>CD38<sup>+</sup> and CD34<sup>-</sup> cells. Importantly, final analyses of a large clinical study are currently ongoing, wherein the prognostic value of the CD34<sup>+</sup>CD38<sup>-</sup> LSC frequency, as determined using the here described method, should be confirmed. In the absence of CD34<sup>+</sup>CD38<sup>-</sup> LSCs, it is however plausible that the CD34<sup>+</sup>CD38<sup>+</sup> and/or CD34<sup>-</sup> cells cause the leukemic engraftment. In CD34-negative patients (~20%) the leukemia-initiating cells are present in the CD34-negative compartment. As this compartment makes up the vast majority of the blast cells, the putative stem cells here may be hidden in a sub-compartment that may be identified by functional assays like the side population. Interestingly, the LSC detection tube also enables identification of CD34-negative patients, who experience fewer relapses and have a longer survival as compared with CD34-positive patients. Overall, our results clearly demonstrate that the use of both a cocktail of markers in one fluorescence channel and of single markers with specific properties in the other fluorescence channels is feasible.

Figure 5. Efficacy of LSC detection tube in diagnosis AML (#2128). After labeling of the cells with antibodies, CD34<sup>+</sup> blasts were identified as shown in Supplementary Figure S1I (A–D) and subsequently CD34<sup>+</sup>CD38<sup>-</sup> cells were gated within this fraction (CD34<sup>+</sup>CD38<sup>-</sup> blasts shown in dark gray). Figure I (a–f) shows the expression of the six individual markers (CLL-1/CLEC12A, TIM-3, CD7, CD11b, CD22 and CD56) on CD34<sup>+</sup> CD38<sup>-</sup> cells. The percentage of identified LSCs is shown for each of these individual markers. In this particular patient, TIM-3 (0.0282% LSCs) and especially CD56 (0.0257% LSCs) are most useful markers for LSC detection. Figure II shows the result of the marker cocktail in the LSC tube, whereby CLL-1, TIM-3, CD7, CD11b, CD22 and CD56 are combined in the PE-fluorescence channel. Although individual expression patterns are no longer visible, the result of the marker cocktail in the LSC tube (Figure S1II) is as good as the best individual marker (here CD56 in Figure II). Note that expression percentages of the best individual markers (TIM-3 and CD56) and the cocktail are very similar: 0.0282, 0.0257 and 0.0292%, respectively. Percentage of LSCs identified by the rest of the markers in the LSC detection tube (markers not shown) were less accurate in this sample as compared with the marker cocktail (CD45RA 0.0123%, CD33 0.0043%, CD123 0.0008% and CD44 0.0000%).
for accurate LSC detection, which is broadly applicable in a multi-institutional setting. For specific purposes, the panel can be extended with additional antibodies in the PE-channel. The LSC detection tube is manufactured and validation experiments are currently ongoing for an improved standardized LSC detection procedure in the near future.

CONFLICT OF INTEREST

Financial support for part of this research was received by BD Biosciences.

ACKNOWLEDGEMENTS

We thank all participating study centers for including their patients in the HOVON trials.

REFERENCES


Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)