ORIGINAL ARTICLE

CLINICAL CYTOMETRY WILEY

Unsupervised cluster analysis and subset characterization of abnormal erythropoiesis using the bioinformatic Flow-Self Organizing Maps algorithm

Anna Porwit[1,2] | Despoina Violidaki[1,2] | Olof Axler[1,2] | Francis Lacombe[3] | Mats Ehinger[1,2] | Marie C. Béné[4]

[1]Department of Clinical Sciences, Oncology and Pathology, Lund University, Faculty of Medicine, Lund, Sweden

[2]Department of Clinical Genetics and Pathology, Skåne University Hospital, Lund, Sweden

[3]Hematology Biology, Bordeaux University Hospital Haut Leveque, Bordeaux, France

[4]Hematology Biology, Nantes University Hospital & CRCINA, Nantes, France

Correspondence
Anna Porwit, Lund University, Department of Clinical Sciences, Division of Oncology and Pathology, Sölvegatan 25b, 22185 Lund, Sweden.
Email: anna.porwit@med.lu.se

Abstract

Background: The Flow-Self Organizing Maps (FlowSOM) artificial intelligence (AI) program, available within the Bioconductor open-source R-project, allows for an unsupervised visualization and interpretation of multiparameter flow cytometry (MFC) data.

Methods: Applied to a reference merged file from 11 normal bone marrows (BM) analyzed with an MFC panel targeting erythropoiesis, FlowSOM allowed to identify six subpopulations of erythropoietic precursors (EPs). In order to find out how this program would help in the characterization of abnormalities in erythropoiesis, MFC data from list-mode files of 16 patients (5 with non-clonal anemia and 11 with myelodysplastic syndrome [MDS] at diagnosis) were analyzed.

Results: Unsupervised FlowSOM analysis identified 18 additional subsets of EPs not present in the merged normal BM samples. Most of them involved subtle unexpected and previously unreported modifications in CD36 and/or CD71 antigen expression and in side scatter characteristics. Three patterns were observed in MDS patient samples: i) EPs with decreased proliferation and abnormal proliferating precursors, ii) EPs with a normal proliferating fraction and maturation defects in late precursors, and iii) EPs with a reduced erythropoietic fraction but mostly normal patterns suggesting that erythropoiesis was less affected. Additionally, analysis of sequential samples from an MDS patient under treatment showed a decrease of abnormal subsets after azacytidine treatment and near normalization after allogeneic hematopoietic stem-cell transplantation.

Conclusion: Unsupervised clustering analysis of MFC data discloses subtle alterations in erythropoiesis not detectable by cytology nor FCM supervised analysis. This novel AI analytical approach sheds some new light on the pathophysiology of these conditions.

KEYWORDS
anemia, erythropoiesis, flow cytometry, myelodysplastic syndromes, unsupervised clustering

1 | INTRODUCTION

The Flow-Self Organizing Maps (FlowSOM) program, created within the open-access Bioconductor open-source R project, applies artificial intelligence (AI) strategies to generate minimal spanning trees (MST) for an unsupervised visualization and interpretation of cytometry data (Van Gassen et al., 2015). This program was developed both for mass and classical cytometry studies. FlowSOM can be programmed to organize multiparameter flow cytometry (MFC) data [both fluorescence (FL) and scatter (SSC)] from list mode (LMD) files in up to 100 nodes displayed as MST. Previous publications have shown how FlowSOM can be used in a novel approach to dissect bone marrow (BM) cell subsets with myeloid- or lymphoid-oriented panels (Lacombe, et al., 2019a; Lacombe, et al., 2019b).

Cytological changes are seen in BM erythropoietic precursors in both anemia of non-clonal origin and in myelodysplastic syndromes (MDS). These changes are still described subjectively through microscopical evaluation of BM smears, although attempts to quantify dysplasia have been made (Della Porta et al., 2015).However, many of the changes described as signs of dysplasia lack specificity (Goasguen et al., 2018). MFC has been used to evaluate abnormal changes in erythropoietic maturation (Arnoulet et al., 2010; Chesney et al., 2011; Eidenschink Brodersen et al., 2015; Violidaki et al., 2020; Westers et al., 2017) but to the best of our knowledge no approach using unsupervised clustering to study the diversity of erythropoiesis in non-clonal anemia and in MDS has been published. We aimed here to explore abnormal erythropoiesis trying to correlate the changes in erythropoietic differentiation with pathophysiology of the disease.

Recently, we have described how FlowSOM can be applied to an erythroid-specific panel to dissect normal stages of maturation in the erythroid compartment (Béné et al., 2020). The panel used was primarily established for an optimal evaluation of erythroid differentiation without any substantial loss of erythroid precursors (EPs) by avoiding lysis (Violidaki et al., 2020). In a reference normal BM (NBM) file built by merging 11 different BM samples, 24 nodes of a FlowSOM MST were identified as belonging to erythropoiesis and could be grouped in six subsets (Béné et al., 2020) consistent with known stages of erythroid maturation. This generation of a normal reference pattern was a prerequisite to the in- depth analysis of BM samples from patients with erythropoiesis disorders.

In the present study, we report on the FlowSOM analysis of MFC files obtained with the same erythroid panel in various pathological BM samples, including BM from patients with MDS. This led to the detection of several previously unknown subsets of erythropoietic precursors and data analysis revealed disturbances in the observed patterns suggesting that FlowSOM may be of help in the diagnosis and follow-up of erythropoietic maturation abnormalities.

2 | MATERIALS AND METHODS

Samples: Eleven BM samples included in the merged NBM file reported in our previous publication (Béné et al., 2020) comprised 3 BM from healthy donors and 8 BM from patients under lymphoma staging but confirmed for the absence of medullary lymphoma involvement. All patients had normal hemoglobin levels, white blood cells count and platelet values, as well as normal BM morphology. Here we analyzed one more NBM sample from a patient with follicular lymphoma with no BM involvement to establish whether the normal pattern was maintained in a sample analyzed separately.

Abnormal BM samples were obtained from 5 patients with anemia of various origin and 11 patients at MDS diagnosis (Table 1). Four additional samples were collected during follow-up for one of the MDS patients (#16, Table 1). All diagnostic BM samples analyzed here had been previously included in the ERY-tube validation study (Violidaki et al., 2020). LMDs were reanalyzed in this in silico study.

FlowSOM analysis: Analysis was performed with the Kaluza® Analysis Software (BC) equipped with FlowSOM plug-ins as previously described (Lacombe, et al., 2019b). Briefly, all LMD.fcs files were checked for proper compensation and fluorescence was normalized by comparison to lymphocytes with a dedicated R script as reported (Lacombe, et al., 2019b). The additional NBM file and patient samples were analyzed individually. The FlowSOM set-seed script was run for each sample requesting 100 nodes and 6 MST proposals. Each of the 6 MST for each sample contained the same number of nodes that belonged to erythropoiesis after it was visually checked using backgating and color-coding of the cell subsets on a standard CD45/SSC histogram (Arnoulet et al., 2010; Béné et al., 2020). The best display segregating the erythroid lineage as a separate ("erythroid") branch of the MST was chosen for further analysis. Each erythroid node was subsequently analyzed in individual samples on the selected MST using a specific Kaluza protocol displaying the normalized mean fluorescence intensity (MFI) of each marker, normalized scatter and cell numbers. Non-relevant nodes, based on SSC/CD45 plots and other marker expression, that were observed in some samples, were not taken into consideration. Erythropoietic nodes with marker expression levels and scatter similar to those identified in the normal reference BM MST were recognized and color coded as such. The previously published MFI thresholds of CD36 and CD71 expression were used (Béné et al., 2020). Thus, CD36 and CD71 normalized MFIs above 4.1 and 4.59, respectively, identified nodes with bright (HIGH) expression. Dim (LOW) expression of these markers was assigned to nodes below the lowest normalized MFI in normal BM, respectively 3.83 for CD36 and 2.52 for CD71. Values between these thresholds identified intermediate (INT) expression. The cut-off for "high Draq5" normalized MFI was 0.1 while very low Draq5 was between 0.01 and 0.06. Thus, MFI of the DNA dye Draq5 that parallels DNA content could be used to identify proliferative subsets and correlate it with their immunophenotype. Three different subsets of dividing cells were identified by their high percentage of Draq5^{hi} cells (>90%).

The analysis of pathological samples disclosed nodes with expression patterns not seen in the reference merged NBM MST. Whenever possible, the nodes with new characteristics were grouped into new subpopulations of erythropoietic cells. The various identified normal and abnormal erythroid subsets and their color-coding strategy are given in Table 2.

TABLE 1 Characteristics of studied patients

Case	Final diagnosis	Age	Sex	% of erythropoiesis[a] by FCM	Number of Ery nodes[b]	% CD34+ cells[c]	Hb	MCV	Cytogenetics
1/18HL00940	Reactive bone marrow in patient with Hodgkin lymphoma	79	M	15	18		117	89	ND
2/18HL01935	B12 deficiency	68	M	44	43	1.75	79	126	ND[d]
3/17HL00463	Alcohol abuse and thrombocytopenia	65	F	46	42	2.6	151	93	46 XX
4/17HL00947	Treated Iron deficiency	58	F	9	12	1.6	110	92	ND
5/17HL00987	Sarcoidosis, kidney insufficiency	70	M	18	20	0.4	107	92	ND
6/17HL02192	MDS-MLD-RS[e]	85	F	23	24	0.7	91	105	46,XX,del(11)(q14) [7]/46,XX [14]
7/17HL01484	MDS-MLD	58	M	59	63	5	121	109	47,XY,+8
8/17HL00874	MDS-MLD	81	M	52	53	1	69	109	46,XY,del(20)(q11) [20/46,XY [5]
9/17HL00785	MDS-MLD	69	M	14	15	2	96	90	47, XY, +11
10/17HL001550	MDS-MLD	71	F	10	14	1.5	120	104	45,X,-X[4]/46,X,?del(x) (q28)[3]/46,X,idic[X] (q13)[14]/47,X,idic (X)(q13)x2[2]
11/17HL00450	MDS-EB1	78	F	9	10	4.3	90	94	46 XX
12/17HL00471	MDS-EB1 and B-NHL	84	M	51	48	2.7	87	99	46 XY
13/17HL01164	MDS-EB1	69	M	52	50	17	72	94	47,XY,+19,del(20)(q12) [18]/48,idem,+10[7]
14/17HL00743	MDS-EB1	66	M	49	49	9	65	97	46,XY
15/17HL00979	MDS-EB2	75	F	36	39	9	109	104	ND
16a/17HL00283[f]	MDS-EB2 diagnosis	27	F	39	45	7	104	102	t(3,5,7)(q25;q34;q21)
16b/17HL01002	MDS-EB2 after 2 Vidaza	27	F	48	54	4.6	104	ND	ND
16c/17HL02490	MDS-EB2 3 m after transplant	27	F	1.5	2	1.2	80	ND	ND
16d/18HL00314	MDS-EB2 8 m after transplant	27	F	26	32	0.8	90	ND	ND
16e/18HL01862	MDS-EB2 12 months after transplant	27	F	21	25	0.5	123	ND	ND

[a]In non-lysed BM sample, gated on CD36/CD71.

[b]In FlowSOM analysis.

[c]In lysed BM sample stained with a different panel (Rajab & Porwit, 2015).

[d]ND: not done.

[e]MDS: myelodysplastic syndrome, MLD: multilineage dysplasia, RS: ring sideroblasts; EB: excess of blasts; B-NHL: B-cell non-Hodgkin lymphoma.

[f]Diagnostic bone marrow sample and 4 BM samples under/after treatment were studied (see Table S5; Figure 4).

3 | RESULTS

Normal erythropoiesis: The separate individual analysis of an additional NBM sample retrieved the same six subsets of erythropoietic cells as in the reference merged file (11) (Supporting Methods, Figure S1a,b, Tables S1 and S2) except for the DRAQ LOW subpopulation that was absent since it could be too small to identify in a single sample. Consequently, the total number of erythropoietic nodes was slightly lower in this BM by comparison to the merged file (19 vs. 24 nodes, respectively). Also, the number of nodes within each normal subset was somewhat different when compared to the reference merged file (11) (Figure S1a,b, Tables S1 and S2). The frequency of cells belonging to various subsets shifted to the later stages of erythropoiesis in this sample. However, no abnormal erythropoietic nodes were identified.

TABLE 2 Color coding of nodes identified as various subsets of erythropoiesis in normal and pathological bone marrow samples [Color table can be viewed at wileyonlinelibrary.com]

| Subpopulation	Characteristics	Color	Normal (N)[a] versus abnormal (A)
EARLY DIV	CD117$^+$ CD105$^{+/-}$ CD36^{High} CD71^{High} DRAQ5^{high} > 50%	Sky blue	N
INT DIV	CD117$^-$ CD105$^+$ CD36^{High} CD71^{High} DRAQ5^{high} > 50%	Navy Blue	N
LATE DIV	CD117$^-$ CD105$^-$ CD36^{High} CD71^{High} DRAQ5^{high} > 50%	Amber	N
HIGH 36/HIGH71	CD117$^-$CD105$^-$ CD36^{High} CD71^{High} DRAQ5^{high} < 50%	Green	N
INT 36/INT 71	CD117$^-$CD105$^-$ CD36^{Int} CD71^{Int} DRAQ5^{high} < 50%	Violet	N
DRAQ low	CD117$^-$CD105$^-$ CD36$^{Int/low}$ CD71$^{Int/low}$ DRAQ5 < 0.07	Crimson	N
EARLY 0% ≤ 50% DIV	CD117$^+$ CD105$^{+/-}$ CD36^{High} CD71^{High} DRAQ5^{high} > 20% < 50%	Azure	A
HIGH36/HIGH71 DIV >20% <50%	CD117$^-$CD105$^-$ CD36^{High} CD71^{High} DRAQ5^{high} > 20%,<50%	Sea blue	A
HIGH 36/INT 71 DIV	CD117$^-$CD105$^-$ CD36^{High} CD71^{Int} DRAQ5^{high} > 50%	Gray-green	A
LOW 36/HIGH-INT 71 DIV	CD117$^-$CD105$^-$ CD36^{Low} CD71^{High} DRAQ5^{high} > 50%	Aquamarine	A
LOW 36/LOW 71 DIV	CD117$^-$CD105$^-$ CD36^{Low} CD71^{Low} DRAQ5^{high} > 20%	Apricot	A
HIGH 36/INT 71	CD117$^-$CD105$^-$ CD36^{High} CD71^{Int} DRAQ5^{high} < 50%	Yellow	A
HIGH 36/LOW 71	CD117$^-$ CD105$^-$ CD36^{High} CD71^{Int} DRAQ5^{high} < 50%	Pink	A
INT 36/HIGH 71	CD117$^-$CD105$^-$ CD36^{Int} CD71^{High} DRAQ5^{high} < 50%	Olive	A
INT36/LOW 71	CD117$^-$ CD105$^-$ CD36^{Int} CD71^{Low} DRAQ5^{high} < 50%	Pear green	A
LOW 36/HIGH 71	CD117$^-$ CD105$^-$ CD36^{Low} CD71^{High} DRAQ5^{high} < 50%	Emerald	A
LOW 36/INT 71	CD117$^-$ CD105$^-$ CD36^{Low} CD71^{Int} DRAQ5^{high} < 50%	Peanut	A
LOW 36/LOW 71	CD117$^-$ CD105$^-$ CD36^{Low} CD71^{Low} DRAQ5^{high} < 50%	Red	A
HIGH 36/HIGH 71 HIGH SC	CD117$^-$ CD105$^-$ CD36^{High} CD71^{High} DRAQ5^{high} < 50% SSC > 70	Army green	A
HIGH 36/INT71 HIGH SC	CD117$^-$ CD105$^-$ CD36^{high} CD71^{Int} DRAQ5^{high} < 50% SSC > 70	Royal	A
INT 36/HIGH/INT 71 HIGH SC	CD117- CD105$^-$ CD36^{Low} CD71^{High} DRAQ5^{high} < 50% SSC > 70	Brown	A
INT36/LOW 71 HIGH SC	CD117$^-$ CD105$^-$ CD36^{Low} CD71^{low} DRAQ5^{high} < 50% SSC > 70	Dark brown	A
LOW 36/INT 71 HIGH SC	CD117$^-$ CD105$^-$ CD36^{Low} CD71^{Int} DRAQ5^{high} < 50% SSC > 70	Helio	A
LOW 36/LOW 71 HIGH SC	CD117$^-$ CD105$^-$ CD36^{Low} CD71^{Low} DRAQ5^{high} < 50% SSC > 70	Rose	A

[a]N present in normal BM, A, abnormal.

This observation supports the use of a merged reference file as the basis for analysis to minimize slight individual variations.

Erythropoiesis in non-clonal cytopenia: BM samples from five patients with non-clonal cytopenia were analyzed with FlowSOM and the MSTs are illustrated in Figure 1a–e. Detailed analysis of the erythropoietic compartments in these cases is given in Table S3. Seven new EP subsets could be defined in these samples, while some normal compartments had disappeared in several pathological BMs.

A patient with anemia and reactive BM (staging of Hodgkin lymphoma, #1, Figure 1a) showed a globally normal pattern yet without a detectable DRAQ low compartment and about 10% of cells with a previously not described CD36^{HI}/CD71^{INT} immunophenotype. Two patients (# 2 and 3, Table 1) presented with an expanded erythropoietic fraction of BM cells (>40%). In the patient with B12 deficiency (# 2), normal late stages of maturation were absent and a large population (46.2%) of EPs with low CD36 and low CD71, absent from the
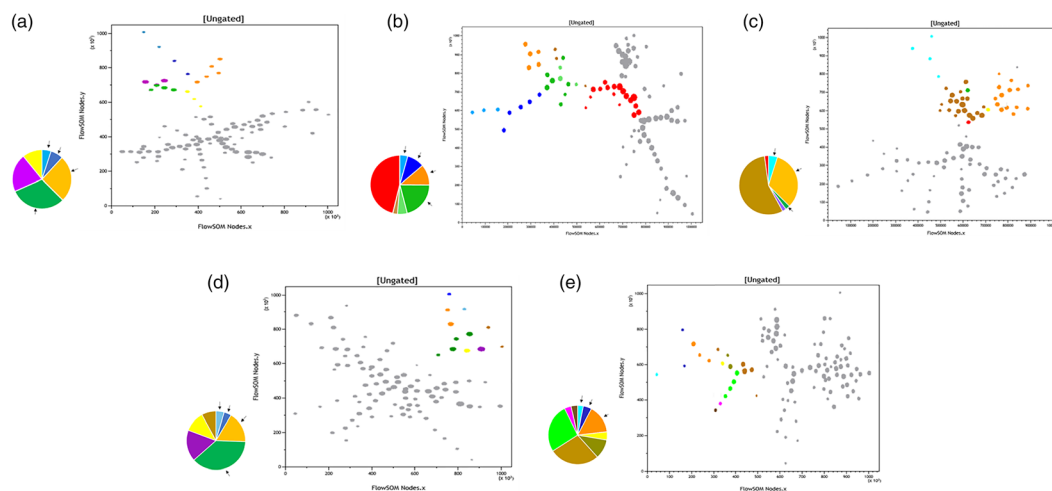
FIGURE 1 Minimal spanning trees and proportions of identified erythroid subsets in patients with non-clonal abnormal hematopoiesis (#1–5, Table 1). Colors are as indicated in Table 2 and full data are given in Table S3. Uncolored (gray) nodes represent non-erythroid hematopoietic subsets. Normal BM compartments (see Figure S1) are indicated by arrows on the pie graphs. (a) Case #1: Anemia and reactive BM; (b) Case #2 patient with vitamin B12 deficiency. The largest subset consists of non-proliferating CD36^{LO} CD71^{LO} cells. (c) Case #3: Alcohol abuse and thrombocytopenia; (d) Case #4: Iron deficiency under treatment. (e) Case #5, patient with renal failure where normal stages of non-dividing cells have been lost, the majority of cells displaying an abnormal CD36^{LO} pattern and variable CD71 [Color figure can be viewed at wileyonlinelibrary.com]

reference NBM, was noted, indicating maturation asynchrony (Figure 1b). The fraction of dividing cells was lower in B12 deficiency BM than in reference NBM and other non-clonal anemias. In the patient with alcohol abuse and thrombocytopenia but no anemia (# 3), there was no INT DIV population. Instead, a large abnormal non-proliferating population of CD36^{LOW}/CD71^{INT} EP was found (55.7% of erythropoiesis) (Figure 1c). In the BM of the patient with iron deficiency under treatment (#4, Figure 1d), the main population (~53%) was the immunophenotypically normal CD36^{HI}/CD71^{HI} non-dividing compartment. Two new non-proliferating subsets with intermediate CD71 and asynchronously high or low CD36, respectively, comprised about 27% of the hematopoiesis. In the patient with anemia due to renal failure (#5, Figure 1e), the normal dividing compartments represented 23% of erythropoiesis, and the normal stages of non-dividing cells were absent. This case presented with a higher number of anomalies and six new populations. The majority or erythropoietic cells (56%) were CD36^{LOW} with variable expression of CD71. Some cells also displayed an unexpectedly high SSC.

Erythropoiesis in myelodysplastic syndromes: Bone marrow samples from 11 MDS patients at diagnosis were investigated. The best selected individual MSTs for each case are shown in Figure 2a–e and Figure S2a–e. Detailed characteristics of identified nodes are given in Table S4. Comparison of MSTs showed that erythropoiesis comprised from 10 to 63 nodes among the 100 displayed. Analysis of individual MST showed substantial variability. A total of 24 different subsets (groups of nodes where detailed manual analysis revealed common characteristics) could be singled out, including 18 abnormal subsets absent in normal BM (Table 2 and Table S4).

In five samples (# 7, 8, 12, 13, 14), erythropoiesis represented more than 40% of all BM cells (Table 1). In three MDS cases (# 6, 9, and 11),

most EP belonged to normal subsets (Figure 2a; Figure S3a,b, Table S4). In other cases, most EP were abnormal, and only a small fraction of BM cells showed a normal erythropoietic immunophenotype (for example, 3%–4% of BM cells in # 10 and 12) (Figure 2b, Figure S3c, Table S4).

Eleven abnormal cell subsets were identified, not seen in normal BM nor non-clonal anemias, most in the proliferating cell compartment. There was substantial variability in the patterns of abnormal erythropoiesis. The most common finding was of CD36^{LOW} precursors with different levels of CD71 expression (Table S4). However, in two MDS cases with del.20q (#8 and #13, Table 1 & Table S4), there was a significant population of CD36^{HI} CD71^{INT} EP (Figure 2c; Figure S3d). In four MDS cases (# 6,7,14,15), a detailed manual analysis revealed nodes corresponding to subsets of EPs with abnormally high SSC properties (Figure 2a,d; Figure S3e,f).

Figure 3 summarizes the partition of the various subsets in proliferating compartments of erythropoiesis in the pathological BM samples of this study compared to the reference merged NBM. All four MDS-MLD had an approximately normal fraction of proliferating cells, most of the normal LATE DIV type. The MDS MLD-RS had a lower proliferative fraction than normal BM yet with a normal distribution of erythroid precursors. All four MDS EB1 and one of MDS EB2 also had a lower than normal fraction of proliferating erythropoietic cells. In one MDS EB1 (#11) and one MDS EB2 (#15) case, the proliferating cells were mostly of LATE DIV immunophenotype. In three MDS-EB1 cases, most proliferating cells displayed an abnormal immunophenotype. In one of these MDS EB1 cases (# 13, Figure 2c; Table S4), a subset of four nodes contained 20 to 50% of DRAQ5^{HI} cells, a feature not seen in normal BM. Finally, one MDS EB2 had a significantly larger than normal compartment of proliferating cells, composed of CD117+ EARLY DIV nodes (#16, Figure 4a).
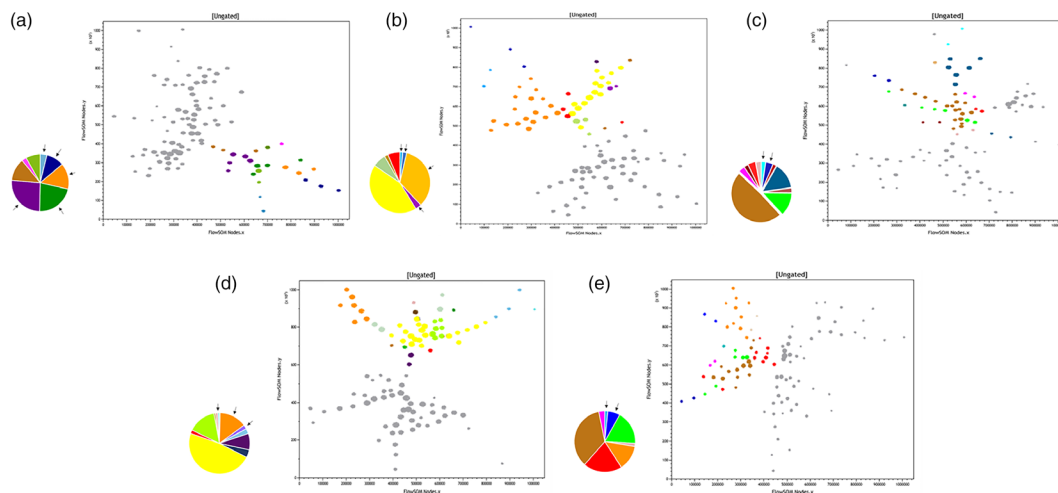
FIGURE 2 Minimal spanning trees and proportions of identified erythroid subsets in five patients with MDS. Colors are as indicated in Table 2 and full data are given in Table S4. Uncolored (gray) nodes represent non-erythroid hematopoietic subsets. Normal BM compartments (see Figure S1) are indicated by arrows on the pie graphs. Patients in panels (a) to (e) are respectively # 6, 8, 12, 13, and 14 in Table 1. MST of remaining MDS patients are given in Figure 4 (#16) and Figure S2 [Color figure can be viewed at wileyonlinelibrary.com]
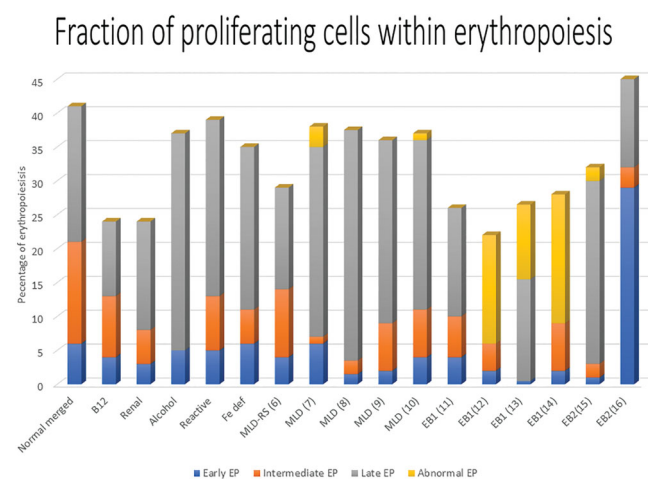

FIGURE 3 Comparative fractions of proliferating erythropoietic subsets in non-clonal anemia and MDS. Proportions of proliferating cells including subsets of normal and abnormal erythroid precursors are shown in comparison to merged normal BM for the 16 patients of the study [Color figure can be viewed at wileyonlinelibrary.com]

Follow-up after treatment: The diagnostic sample of patient #16 was compared with four follow-up samples (Figure 4a–e, Table 1, and Table S5). This young patient (age 27) with MDS-EB2 had at diagnosis extremely left shifted erythropoiesis with a predominance of abnormal CD117+ EP. FlowSOM analysis identified two subsets of CD117+ erythroid cells. One had an immunophenotype corresponding to the EARLY DIV population in normal BM and one had a similar immunophenotype but low or no proliferation (EARLY NON/LOW DIV). Erythropoiesis showed very little maturation (Figure 4a). The follow-up samples were taken after two cycles of azacytidine, and then 3, 8, and 12 months after allogeneic hematopoietic stem cell transplantation, respectively. After azacytidine

(Figure 4b), there was maturation in erythropoiesis, most EP showing low CD36 expression (Table S5) but a fraction corresponding to normal erythropoiesis was also noted. Erythropoiesis was very scarce at the first control time-point 3 months after allogeneic HSCT (Figure 4c), but then a regeneration of erythropoiesis occurred and at 8 (Figure 4d) and 12 months (Figure 4e),where more than half of the EP belonged to normal subsets. However, a minor fraction of CD117+ EARLY NON/LOW DIV EP (0.98% of BM cells) was seen at 12 months post allogeneic HSCT (Table S5). In addition, some abnormal subsets not seen in normal BM were present both at 8 (Figure 4d) and 12 months (Figure 4e). The patient remains in complete remission 2 years after allogeneic HSCT. Nevertheless, the finding of some abnormal "patient-specific" subsets after allogeneic HSCT suggests that residual disease may be present in the hematopoietic niche. Unfortunately we do not have any information on the mutational status of this patient and do not know whether there are some variants still present.

4 | DISCUSSION

This study aimed at deepening knowledge about the alterations of erythropoiesis in various diseases, already hinted at via classical flow cytometry analysis (Arnoulet et al., 2010; Chesney et al., 2011; Eidenschink Brodersen et al., 2015; Violidaki et al., 2020; Westers et al., 2017). The advantage of unsupervised analysis lies in the ability of the software to segregate subsets (Van Gassen et al., 2015) that are merged in the image of a continuum presented in classical bi-parametric plots and even in the more sophisticated radar representation (Violidaki et al., 2020). FlowSOM can be applied retrospectively to already acquired data sets and thus disclose important information without the need to perform new experiments. In the present study, as in previous analyses of myeloid and lymphoid compartments
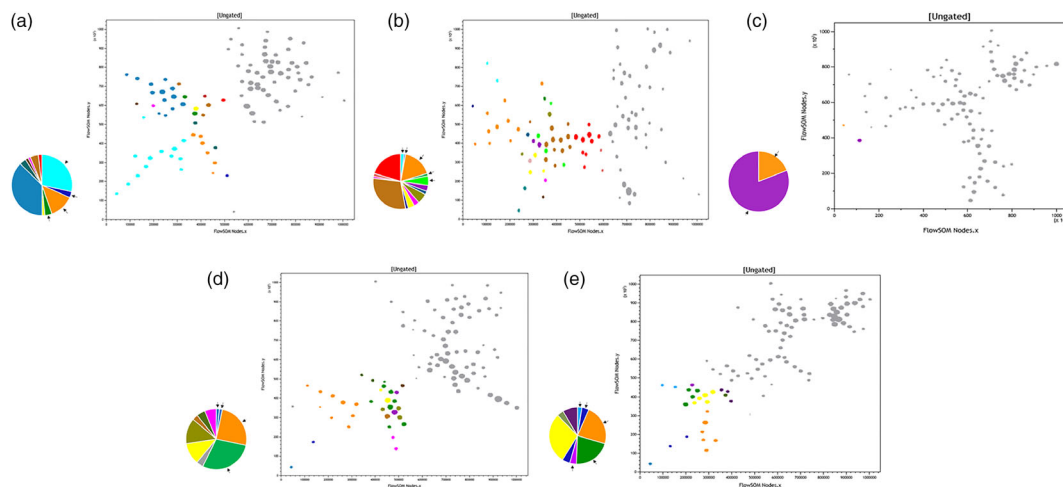
FIGURE 4 Minimal spanning trees and proportions of identified erythroid subsets at different follow-up time-points of patient with MDS-EB2 under treatment (#16). Colors are as indicated in Table 2 and full data are given in Table S5. Uncolored (gray) nodes represent non-erythroid hematopoietic subsets. Normal BM compartments (see Figure S1) are indicated by arrows on the pie graphs. (a) Case #16 at diagnosis; (b) after 2 cycles of Azacytidine; (c) 3 months post Allo-HSCT; (d) 8 months post Allo-HSCT; (e) 12 months post Allo-HSCT [Color figure can be viewed at wileyonlinelibrary.com]

(Lacombe, 2019a; Lacombe, 2019b), we chose to use MST with 100 nodes. If the number of nodes would be lower, visualization could be easier to follow but some populations of interest might be lost. With a high number of nodes, new populations are easier to detect. We have not used the method of merging the abnormal sample with the NBM reference file as described in the study aimed on detecting minimal residual disease (Vial et al., 2021) to avoid a bias towards normal subsets and to enable the discovery of abnormal differentiation in erythropoiesis.

In the recent study by Duetz et al. (2021), FlowSOM was applied together with a machine learning classifier (Random Forest) as a tool for evaluation of flow cytometry results in suspected MDS. Although the aim of this study, using 6 different tubes with 8-color panels, was different than ours, the erythroid tube (no 3) provided the highest sensitivity and specificity. However, the erythroid compartment diversity was not as obvious using a classifier that tried to apply a defined number of erythroid nodes in each sample.

The present analysis of BM samples from patients with anemia of various origin and MDS revealed many new subsets of erythropoietic precursors not found in NBM. Abnormal subsets found in both non-clonal anemia and MDS patients displayed either an asynchronous expression of CD36 and CD71 or abnormally low expression of these markers. In reactive BM, a subset of EP with high CD36 and intermediate expression of CD71 was noted. This subset was also seen in the BM of a patient with treated iron deficiency and it was highly represented in MDS cases with del 20q that often present with only borderline morphological dysplasia (Gupta et al., 2007).

Late erythropoietic precursors in a patient with Vitamin B12 deficiency were mostly abnormal with a significant new population of CD36^{LOW}/CD71^{LOW} EPs. The presence of this population is consistent with the nuclear/cytoplasmic dissociation described in Vitamin B12 deficiency, and ineffective erythropoiesis due to impaired DNA synthesis and maturation arrest (Green, 2017). The EP compartment was increased but showed a lower proliferative fraction than that in NBM. The CD36^{LOW}/CD71^{LOW} population was also found in seven of 11 studied MDS cases but at a lower frequency than in the BM sample from a patient with B12 deficiency. This low expression of CD36 and CD71, two surface antigens that disappear upon final maturation in normal erythropoiesis, could be related to a normal level of cytoplasmic hemoglobin, known to induce the arrest of proliferation. In the case of Vitamin B12 deficiency, this occurs before enough divisions have been completed.

Erythropoietic precursors in the BM of a patient with anemia due to renal insufficiency showed an abnormal pattern of marker expression and populations with abnormal scatter in late EPs consistent with maturation asynchrony. The primary mechanism of anemia in renal failure is due to a low erythropoietin level (Means, 2019; Shih et al., 2018). Erythropoietin acts primarily by promoting the survival of late EPs allowing for the differentiation and emergence of mature red cells, but it has little or no effect on the proliferation and differentiation of early EPs. The low proliferative fraction found in the erythropoiesis of this patient is consistent with this pathophysiological mechanism.

Abnormal EP precursors were also noted in a hospital control patient with alcohol abuse and normal hemoglobin but macrocytosis and thrombocytopenia. A large population of LOW CD36/INT CD71 with absence of the normal INT DIV population and thrombocytopenia raises the possibility of CD36 deficiency in this patient (Hirano et al., 2003).

Abnormal subsets seen in MDS patients and also found in anemic hospital controls were due to abnormally low CD36 expression with variable CD71 and belonged mostly to the late EP compartment. Additional abnormal subsets seen only in MDS patients were found in the proliferating cell compartment. These subsets were more often

seen in cases with an excess of blasts. Although the studied MDS cohort was probably not large enough to draw definitive conclusions, three patterns could be outlined:

1. EPs with decreased proliferation and phenotypically abnormal proliferating precursors (patients # 12, 13, 14, 16)
2. EPs with normal proliferating fraction and maturation defects in late EPs (patient #7, 8, 15)
3. EPs with low erythropoietic fraction and a mostly normal pattern suggesting that erythropoiesis was less affected (patients #6,9,10,11).

Of note, mutational analysis was not as yet performed routinely for MDS patients at the time BM samples for the current study were collected. A larger cohort of patients with available mutational analysis data is under investigation in search for a correlation between the observed patterns and mutational status.

Nevertheless, comparison of the results of this study and of our previous analysis with the help of radar display in the same cases, clearly shows that the unsupervised analysis performed by FlowSOM is more powerful than any analyses that can be achieved with classical software. The abnormal maturation patterns disclosed in bi-parametric or radar assessments are segmented by FlowSOM in subsets with subtle modifications of marker expression (mostly CD36 and CD71), in a novel and unsuspected way. The previously reported abnormal CVs of these two markers' MFIs (Eidenschink Brodersen et al., 2015; Mathis et al., 2013; Westers et al., 2017) gain here a much higher level of discrimination, highlighting further the heterogeneity of erythropoietic disturbances in MDS. Rather than trying to perform wholly automated analysis as in Duetz et al. (2021), we tried to take advantage of the ability of the clustering algorithm to identify the complexity of the erythroid compartment in diseased samples, which was possible by combining the unsupervised clustering with manual characterization and classification of the nodes.

Another interesting point is that unsupervised analysis allowed to retrieve at least some of normal erythropoietic populations in most MDS and non-clonal hematopoietic diseases, which may help to better understand where homeostasis went wrong. Yet, the most promising results were to be seen in the patient with follow-up samples, where the induction of maturation with azacytidine treatment and the recovery of nearly-normal erythropoiesis after allogeneic HSCT was documented. This observation also deserves further investigation in a larger cohort (ongoing) but appears as a promising way to evaluate hematological reconstitution in MDS patients, whether transplanted or not.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

ORCID

Anna Porwit https://orcid.org/0000-0003-0709-2713
Despoina Violidaki https://orcid.org/0000-0002-2229-4260
Marie C. Béné https://orcid.org/0000-0002-6569-7414

REFERENCES

Arnoulet, C., Béné, M. C., Durrieu, F., Feuillard, J., Fossat, C., Husson, B., Jouault, H., Maynadié, M., & Lacombe, F. (2010). Four- and five-color flow cytometry analysis of leukocyte differentiation pathways in normal bone marrow: A reference document based on a systematic approach by the GTLLF and GEIL. *Cytometry. Part B, Clinical Cytometry*, *78*, 4–10. https://doi.org/10.1002/cyto.b.20484

Béné, M. C., Axler, O., Violidaki, D., Lacombe, F., Ehinger, M., & Porwit, A. (2020). Definition of erythroid differentiation subsets in normal human bone marrow using FlowSOM unsupervised cluster analysis of flow cytometry data. *Hema*, *5*, e512. https://doi.org/10.1097/HS9.0000000000000512

Chesney, A., Good, D., & Reis, M. (2011). Clinical utility of flow cytometry in the study of erythropoiesis and nonclonal red cell disorders. *Methods in Cell Biology*, *103*, 311–332. https://doi.org/10.1016/B978-0-12-385493-3.00013-9

Della Porta, M. G., Travaglino, E., Boveri, E., Ponzoni, M., Malcovati, L., Papaemmanuil, E., Rigolin, G. M., Pascutto, C., Croci, G., Gianelli, U., Milani, R., Ambaglio, I., Elena, C., Ubezio, M., da Via', M. C., Bono, E., Pietra, D., Quaglia, F., Bastia, R., ... Rete Ematologica Lombarda (REL) Clinical Network. (2015). Rete Ematologica Lombarda (REL) clinical network. Minimal morphological criteria for defining bone marrow dysplasia: A basis for clinical implementation of WHO classification of myelodysplastic syndromes. *Leukemia*, *29*, 66–75. https://doi.org/10.1038/leu.2014.161

Duetz, C., Van Gassen, S., Westers, T. M., van Spronsen, M. F., Bachas, C., Saeys, Y., & van de Loosdrecht, A. A. (2021). Computational flow cytometry as a diagnostic tool in suspected-myelodysplastic syndromes. *Cytometry. Part A*, *99*, 814–824. https://doi.org/10.1002/cyto.a.24360

Eidenschink Brodersen, L., Menssen, A. J., Wangen, J. R., Stephenson, C. F., de Baca, M. E., Zehentner, B. K., Wells, D. A., & Loken, M. R. (2015). Assessment of erythroid dysplasia by "difference from normal" in routine clinical flow cytometry workup. *Cytometry. Part B, Clinical Cytometry*, *88*, 125–135. https://doi.org/10.1002/cyto.b.21199

Goasguen, J. E., Bennett, J. M., Bain, B. J., Brunning, R., Vallespi, M. T., Tomonaga, M., Zini, G., Renault, A., & The International Working Group on Morphology of MDS. (2018). Dyserythropoiesis in the diagnosis of the myelodysplastic syndromes and other myeloid neoplasms: Problem areas. *British Journal of Haematology*, *182*, 526–533. https://doi.org/10.1111/bjh.15435

Green, R. (2017). Vitamin B$_{12}$ deficiency from the perspective of a practicing hematologist. *Blood*, *129*, 2603–2611. https://doi.org/10.1182/blood-2016-10-569186

Gupta, R., Soupir, C. P., Johari, V., & Hasserjian, R. P. (2007). Myelodysplastic syndrome with isolated deletion of chromosome 20q: An indolent disease with minimal morphological dysplasia and frequent thrombocytopenic presentation. *British Journal of Haematology*, *139*, 265–268. https://doi.org/10.1111/j.1365-2141.2007.06776.x

Hirano, K., Kuwasako, T., Nakagawa-Toyama, Y., Janabi, M., Yamashita, S., & Matsuzawa, Y. (2003). Pathophysiology of human genetic CD36 deficiency. *Trends in Cardiovascular Medicine*, *13*, 136–141. https://doi.org/10.1016/s1050-1738(03)00026-4

Lacombe, F., Dupont, B., Lechevalier, N., Vial, J. P., & Béné, M. C. (2019a). Innovation in flow cytometry analysis: A new paradigm delineating normal or diseased bone marrow subsets through machine learning. *Hema*, *3*, e173. https://doi.org/10.1097/HS9.0000000000000173

Lacombe, F., Lechevalier, N., Vial, J. P., & Béné, M. C. (2019b). An R-derived FlowSOM process to analyze unsupervised clustering of normal and malignant human bone marrow classical flow cytometry data. *Cytometry. Part A*, *95*, 1191–1197. https://doi.org/10.1002/cyto.a.23897

Mathis, S., Chapuis, N., Debord, C., Rouquette, A., Radford-Weiss, I., Park, S., Dreyfus, F., Lacombe, C., Béné, M. C., Kosmider, O., Fontenay, M., & Bardet, V. (2013). Flow cytometric detection of dyserythropoiesis: A sensitive and powerful diagnostic tool for myelodysplastic syndromes. *Leukemia*, *27*, 1981–1987. https://doi.org/10.1038/leu.2013.178

Means, R. T. (2019). Anemias secondary to inflammation/chronic disease and systemic disorders. In J. P. Greer, D. A. Arber, B. E. Glader, A. F. List, R. T. Means, Jr., & G. M. Rodgers (Eds.), *Wintrobe's clinical hematology* (14th ed.). Wolters Kluwer.

Rajab, A., & Porwit, A. (2015). Screening bone marrow samples for abnormal lymphoid populations and myelodysplasia-related features with one 10-color 14-antibody screening tube. *Cytometry. Part B, Clinical Cytometry*, *88*, 253–260. https://doi.org/10.1002/cyto.b.21233

Shih, H. M., Wu, C. J., & Lin, S. L. (2018). Physiology and pathophysiology of renal erythropoietin-producing cells. *Journal of the Formosan Medical Association*, *117*, 955–963. https://doi.org/10.1016/j.jfma.2018.03.017

Van Gassen, S., Callebaut, B., van Helden, M., Lambrecht, B. N., Demeester, P., Dhaene, T., & Saeys, Y. (2015). FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry. Part A*, *87*, 636–645. https://doi.org/10.1002/cyto.a.22625

Vial, J. P., Lechevalier, N., Lacombe, F., Dumas, P. Y., Bidet, A., Leguay, T., Vergez, F., Pigneux, A., & Béné, M. C. (2021). Unsupervised flow cytometry analysis allows for an accurate identification of minimal residual disease assessment in acute myeloid leukemia. *Cancers*, *13*, 629. https://doi.org/10.3390/cancers13040629

Violidaki, D., Axler, O., Jafari, K., Bild, F., Nilsson, L., Mazur, J., Ehinger, M., & Porwit, A. (2020). Analysis of erythroid maturation in the nonlysed bone marrow with help of radar plots facilitates detection of flow cytometric aberrations in myelodysplastic syndromes. *Cytometry. Part B, Clinical Cytometry*, *98*, 399–411. https://doi.org/10.1002/cyto.b.21931

Westers, T. M., Cremers, E. M., Oelschlaegel, U., Johansson, U., Bettelheim, P., Matarraz, S., Orfao, A., Moshaver, B., Brodersen, L. E., Loken, M. R., Wells, D. A., Subirá, D., Cullen, M., te Marvelde, J. G., van der Velden, V., Preijers, F. W., Chu, S. C., Feuillard, J., Guérin, E., ... IMDSFlow Working Group. (2017). Immunophenotypic analysis of erythroid dysplasia in myelodysplastic syndromes. A report from the iMDSFlow working group. *Haematologica*, *102*, 308–319. https://doi.org/10.3324/haematol.2016.147835

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.