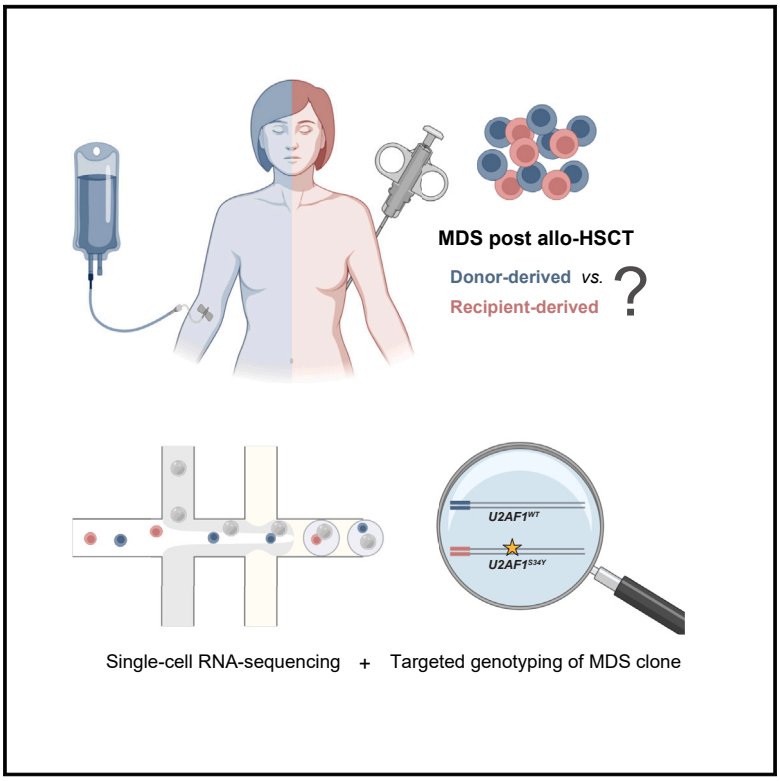
Med

Diagnosing recipient- vs. donor-derived posttransplant myelodysplastic neoplasm via targeted single-cell mutational profiling

Graphical abstract



Highlights

- Distinguishing recipient- from donor-derived MDS post allo-HSCT is critical
- The latter can currently not be addressed with conventional methods
- Single-cell profiling of the *U2AF1*^{S34Y} mutation unravels the origin of MDS
- Targeted single-cell sequencing may serve as an additional pillar for MDS diagnostics

Authors

Jana Ihlow, Livius Penter, Lam Giang Vuong, ..., Nils Blüthgen, David Horst, Samantha D. Praktiknjo

Correspondence

samantha.praktiknjo@bih-charite.de

In brief

Ihlow et al. develop a customized single-cell approach to solve the origin of MDS in a patient after allo-HSCT. Using targeted sequencing of the MDS-specific *U2AF1*^{S34Y} locus, they show that *U2AF1*^{S34Y}-mutated cells exclusively stem from the patient and not from the donor, thus providing an unambiguous recipient-derived posttransplant MDS diagnosis.

Translation to Patients

Ihlow et al., 2025, Med 6, 100548
April 11, 2025 © 2024 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.medj.2024.11.001

CellPress

Med

CellPress
OPEN ACCESS

Article

Diagnosing recipient- vs. donor-derived posttransplant myelodysplastic neoplasm via targeted single-cell mutational profiling

Jana Ihlow,[1,2,3] Livius Penter,[2,3,4,5,6] Lam Giang Vuong,[2] Philip Bischoff,[1,3,7] Benedikt Obermayer,[8] Alexandra Trinks,[9] Olga Blau,[2,10] Anke Behnke,[1] Thomas Conrad,[11] Markus Morkel,[1,9] Catherine J. Wu,[4,5,6] Jörg Westermann,[2,10] Lars Bullinger,[2,3,7,10] Ann-Christin von Brünneck,[1] Nils Blüthgen,[1,7,12] David Horst,[1,7] and Samantha D. Praktiknjo[3,13,*]

[1]Institute of Pathology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
[2]Department of Hematology, Oncology and Cancer Immunology, Campus Virchow Clinic, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
[3]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany
[4]Department of Medical Oncology, Dana-Faber Cancer Institute, Boston, MA, USA
[5]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA
[6]Harvard Medical School, Boston, MA, USA
[7]German Cancer Consortium (DKTK) Partner Site Berlin, German Cancer Research Center (DKFZ), Heidelberg, Germany
[8]Core Unit Bioinformatics, Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany
[9]BIH Bioportal Single Cells, Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany
[10]Labor Berlin Charité Vivantes GmbH, Berlin, Germany
[11]Genomics Technology Platform, Berlin Institute of Health at Charité – Universitätsmedizin Berlin and Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany
[12]Institute for Biology, Humboldt-Universität zu Berlin, Berlin, Germany
[13]Lead contact
*Correspondence: samantha.praktiknjo@bih-charite.de
https://doi.org/10.1016/j.medj.2024.11.001

CONTEXT AND SIGNIFICANCE Myelodysplastic neoplasms (MDS) are characterized by malignant transformation of blood-forming cells and can occur after allogeneic hematopoietic stem cell transplantation. Determining whether this is caused by the patient's own cells (recipient derived) or cells received from the transplant (donor derived) is therefore critical. However, conventional methods fail to adequately address this question. Here, the authors solve the origin of MDS in a 38-year-old patient affected by this scenario. Using a customized single-cell-based strategy that identifies the MDS-associated *U2AF1*^{S34Y} mutation, they reveal a recipient-derived MDS in an unambiguous manner, thus providing the opportunity for an effective treatment. The study emphasizes the need to further develop and implement high-resolution approaches at single-cell resolution to advance personalized medicine.

SUMMARY

Background: Distinguishing donor- vs. recipient-derived myelodysplastic neoplasm (MDS) after allogeneic hematopoietic stem cell transplantation (allo-HSCT) is challenging and has direct therapeutical implications.
Methods: Here, we took a translational approach that we used in addition to conventional diagnostic techniques to resolve the origin of MDS in a 38-year-old patient with acquired aplastic anemia and evolving MDS after first allo-HSCT. Specifically, we used single-cell transcriptional profiling to differentiate between donor- and recipient-derived bone marrow cells and established a strategy that additionally allows identification of cells carrying the MDS-associated *U2AF1*^{S34Y} variant.
Results: The patient exhibited mixed donor chimerism combined with severely reduced erythropoiesis and dysplastic morphology within the granulocytic and megakaryocytic lineage along with the MDS-associated *U2AF1*^{S34Y} mutation in the bone marrow. Single-cell transcriptional profiling together with targeted enrichment of the *U2AF1*^{S34Y}-specific locus further revealed that, while the immune compartment was mainly populated by donor-derived cells, myelopoiesis was predominantly driven by the recipient. Additionally, concordant with

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

recipient-derived MDS, we found that *U2AF1*^{S34Y}-mutated cells were exclusively recipient derived with X but not Y chromosome-specific gene expression.

Conclusion: Our study highlights the clinical potential of integrating high-resolution single-cell techniques to resolve complex cases for personalized treatment decisions.

Funding: The study was funded by intramural resources of the Charité – Universitätsmedizin Berlin and the Berlin Institute of Health.

INTRODUCTION

With an incidence of ~2–6/million,[1] aplastic anemia (AA) is a rare, potentially life-threatening hematopoietic disorder that results in bone marrow failure (BMF) with trilineage cytopenia due to reduction of hematopoietic stem cells.[2–4] In younger patients, allogeneic hematopoietic stem cell transplantation (allo-HSCT) from a human leukocyte antigen (HLA)-identical sibling donor is the therapy of choice, since HLA-mismatched donors pose the risk of severe complications from graft versus host disease (GvHD).[5–9] After immunosuppressive or cytotoxic therapy, secondary myelodysplastic neoplasm (MDS) can evolve from AA[10–18] and is then categorized as myeloid neoplasm post cytotoxic therapy according to the current classification by the World Health Organization (WHO) and International Consensus Classification (ICC).[11,12] In the setting of posttransplant MDS, it is critical to additionally distinguish between secondary MDS and donor MDS, since the latter would preclude the initial donor and require an alternative donor for future allo-HSCTs. However, current methods in routine diagnostics provide only limited accuracy to differentiate between donor- and recipient-derived mutations in bone marrow (BM) samples with mixed donor chimerism. To overcome this limitation, we established a customized single-cell strategy to unambiguously address the origin of MDS after allo-HSCT in a 38-year-old female patient with acquired AA. We show how our implementation of this translational approach provides an opportunity for personalized diagnostics to solve similar clinical questions.

RESULTS

Clinical background of the patient with AA requiring allo-HSCT

The 38-year-old patient with AA in whom we demonstrate our strategy initially presented with severe symptomatic pancytopenia (Figure 1), including life-threatening macrocytic anemia (hemoglobin [Hb] 4.6 g/dL, mean corpuscular volume [MCV] 120 fL, mean corpuscular hemoglobin [MCH] 41.8 pg, mean corpuscular hemoglobin concentration [MCHC] 34.8 g/dL, erythropoietin 510 U/L, reticulocyte production index [RPI] 0.1, lactate dehydrogenase [LDH] 251 U/L) as sign of BMF in 2019. Since disorders associated with anemia and BMF are manifold, comprehensive diagnostic screening was performed.[9,19] Morphological, serological, flow cytometric, and molecular assessment (Figure 2) excluded a range of potential diagnoses. Vitamin B12, folic acid, ferritin, and iron levels were within normal range. Therefore, metabolite deficiency could be ruled out as cause for the anemia. BM biopsy revealed regular cell morphology without dysmorphias, fibrosis, and blasts but a substantial decrease in BM cellularity to 20% with hematopoietic cells <10%. Flow cytometry revealed a Della Porta score[20] of 1, absence of myeloid blasts, and full presence of all glycosylphosphatidylinositol-anchored proteins on hematopoietic cells at the time point of initial diagnosis. In the molecular assessment, there were no mutations present, especially no MDS-typical mutations. Therefore, hypoplastic MDS and paroxysmal nocturnal hemoglobinuria could be ruled out as main differential diagnoses. Cytogenetic assessment showed a normal karyotype (46, XX), making Fanconi anemia unlikely. Both liver parameters (alanine transaminase [ALT] 16 U/L, aspartate transaminase [AST] 34 U/L, gamma-glutamyltransferase [gamma-GT] 20 U/L, bilirubin 0.64 mg/dL, urea 21 mg/dL, albumin 41.9 g/L) and kidney function (creatinine 0.73 mg/dL, natrium 139 mmol/L, potassium 3.9 mmol/L, glomerular filtration rate [GFR] >90) were within normal range, excluding hepatic or renal insufficiency as cause for the BM suppression. Serology was negative for viral infection and anti-nuclear and anti-mitochondrial antibodies, ruling out viral or autoimmune causes. Finally, the beta-glucocerebrosidase enzyme activity assay and serum protein electrophoresis were negative for Gaucher's disease. Gastrointestinal diseases were excluded via esophagogastroscopy and colonoscopy. As a consequence, and based on both her BM biopsy (cellularity <25%, hematopoietic cells <30%) and pancytopenic peripheral blood results (Figures 1C and 2), the patient was diagnosed with non-severe AA (nSAA) in transition to severe AA (SAA) by a team of specialized hematologists and hematopathologists. After this initial diagnosis, the patient showed rapid progression of BMF and pancytopenia, as evidenced by further decrease in BM cellularity (~10%), decreasing neutrophil count (0.52/nL), and decreasing reticulocyte count (7.2/nL) in the peripheral blood. Moreover, her platelet count remained far below the lower normal limit even after several platelet transfusions (41/nL) over the following months, which led to a dynamic transition toward SAA. The patient then ultimately required an allo-HSCT, which she received from her HLA-identical brother 5 months after onset of AA following non-myeloablative conditioning therapy with 2-Gy radiation, fludarabine (4 × 30 mg/m^2), cyclophosphamide (4 × 300 mg/m^2), and anti-thymocyte globulin (Thymoglobulin: 3 × 2.5 mg/kg) (Figures 1A and 1B).

Transition to MDS of unclear origin

Due to beginning graft failure and deterioration of the patient's general condition, a single stem cell boost at month 21 and immunosuppressive therapy with anti-thymocyte globulin (ATGAM) were given to the patient. However, CD34$^+$ donor chimerism declined progressively, ultimately reaching 69% 33 months after initial diagnosis of AA (Figure 1C). Additionally, morphological signs of MDS became detectable in the patient's BM, particularly affecting megakaryopoiesis, as shown by histopathology (Figure 3A). Due to repeated dry puncture, cytology could not be assessed properly. Evaluation by flow cytometry was compatible with MDS according to the scoring system of Della Porta[20]

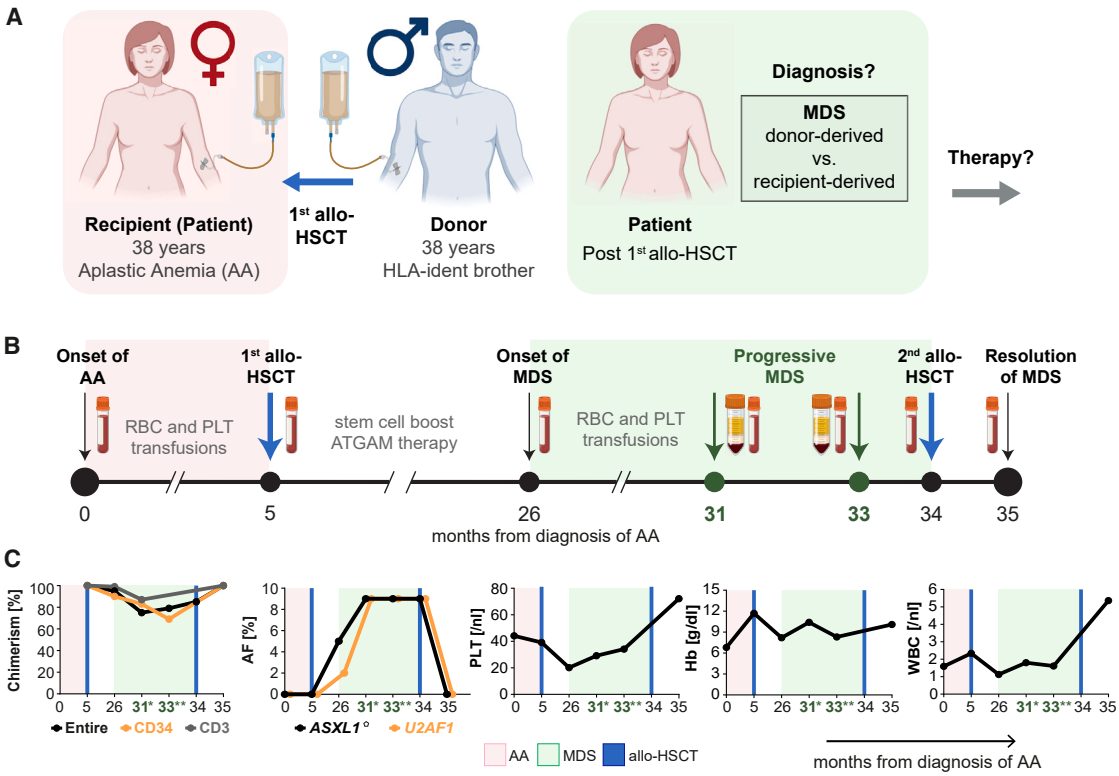


Figure 1. Clinical course of the patient with AA, onset and resolution of MDS

(A) Clinical scenario showing patient characteristics of donor and recipient with unclear origin of MDS.

(B) Timeline of events following onset/diagnosis of AA (month 0) with additional examination of BM samples collected at month 31 (histopathology, XY-fluorescence *in situ* hybridization [FISH], and scRNA-seq with targeted *U2AF1*^{S34Y} enrichment) and month 33 (scRNA-seq).

(C) Development of donor chimerism, mutational status, and blood counts during the course of disease. °Two different *ASXL1* mutations (*ASXL1*^{G651} at month 26 and *ASXL*^{R965*} from month 31 onwards) and a single *U2AF1* mutation (*U2AF1*^{S34Y}) with increasing AF were detected in the patient's BM. The x axes are not to scale.

Allo-HSCT, allogeneic hematopoietic stem cell transplantation; AA, aplastic anemia; MDS, myelodysplastic neoplasm; RBC, red blood cell; PLT, platelet; ATGAM, anti-thymocyte globulin treatment; AF, allele frequency; HLA, human leukocyte antigen; Hb, hemoglobin; WBC, white blood cell count.

(Figures 3C and 3D). In addition, MDS-typical mutations[21–27] were found by bulk DNA sequencing of known mutational hotspots, notably comprising the *U2AF1*^{S34Y} variant, which was longitudinally detected between months 26 and 34 after diagnosis of AA (Figures 1C, 3C, and 3D). Together, this indicated that, following allo-HSCT, the patient had developed MDS that either originated from the donor or the recipient, with the latter likely being a result of previous cytotoxic therapy during allo-HSCT conditioning therapy (myelodysplastic neoplasm post cytotoxic therapy [MDS-pCT] according to WHO/ICC[11,12]). Due to rapid dynamic progression of MDS (Figure 3), it was expected that the patient would require a second allo-HSCT. However, donor-associated MDS would have required the search for a substitute donor, thereby increasing the risk for GvHD. Additional investigation via XY-fluorescence *in situ* hybridization (FISH) on a BM trephine biopsy showed an accumulation of X chromosome signals in dysmorphic cells, particularly in megakaryocytes (Figure 3B). However, further assessment was not possible due to autofluorescence and differences in fluorescence intensity between Y and X chromosome-specific signals. Moreover, diagnosis of lineage dysplasia based on morphological assessment of BM biopsies alone had to be interpreted with caution, since potential cytotoxic cell damage due to antecedent therapy is indistinguishable from true MDS.

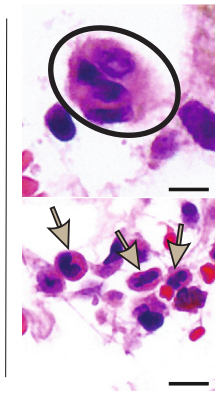
Thus, further treatment decisions were greatly limited by the fact that conventional methods in routine diagnostics could not distinguish between the origin of recipient-derived MDS or donor-cell-derived MDS.

A customized single-cell approach enables MDS precision diagnostics

We thus used a single-cell strategy to differentiate between donor- and recipient-derived BM cells and established an approach that additionally allows identification of cells carrying the *U2AF1*^{S34Y} variant (Figure 4A). Specifically, we performed standard 3′ single-cell gene expression profiling of BM-derived cells collected 31 and 33 months after diagnosis of AA and onset of MDS in the patient. To enable mutational profiling, we modified the standard workflow and used the barcoded single-cell cDNA pool from the sample collected at month 31 to enrich *U2AF1*-specific amplicons that comprise the locus of the mutant variant while inserting a sequencing handle via targeted amplification.

A

MORPHOLOGY



FLOW CYTOMETRY

Parameter	Cutoff	Value	Points
CD34+ myeloblast-related cluster size	≥ 2%	0.26%	0
B-progenitor-related cluster size	≤ 5%	10.2%	0
Myeloblast CD45 expression (ratio)	≤ 4	3.5	1
Granulocyte side scatter value (ratio)	≤ 6	6.97	0
Total Della Porta Score			**1**

MOLECULAR ASSESSMENT

- Cytogenetics: 46, XX
- Mutational status: WT

B

Parameter	Patient	LLN	ULN	nSAA lim.	SAA lim.	vSAA lim.	Classification
Neutrophils, /nl	0.77	1.5	7.7	< 1.2	< 0.5	< 0.2	nSAA
Platelets, /nl	34	150	370	< 70	< 20	< 20	
Reticulocytes, /nl	12	25	150	< 60	< 20	< 20	SAA
Age adj. BM cell., %	20	52	72	≥ 25	≤ 25	≤ 25	
Hematopoiesis, %	< 10			> 30	< 30	< 30	

Figure 2. Clinical characteristics of the 38-year-old patient with AA at initial diagnosis requiring allo-HSCT

(A) Histopathology of the trephine biopsy reveals severely hypocellular BM with a cellularity of 20%, less than 10% hematopoietic cells, a decrease of all three cell lines, and nearly complete absence of erythropoiesis. However, both megakaryopoiesis and granulopoiesis show lineage maturation. Megakaryocytes (encircled) are quantitatively reduced but show regular nuclei with normal lobulation. Granulocytes show band formed and segmented nuclei (arrows) as a sign of maturation. Flow cytometry-based results were not compatible with MDS according to the Della Porta scoring system.[20] Karyotype and mutational status do not show any deviations or signs of clonal hematopoiesis. Scale bar: 30 μm.

(B) Peripheral blood parameters and BM biopsy result of the patient in relation to diagnostic requirements for nSAA, SAA, and vSAA.[7,8] The patient is at the border of nSAA toward SAA.

WT, wild type; adj., adjusted; BM, bone marrow; LLN, lower limit of normal; ULN, upper limit of normal; nSAA, non-severe AA; SAA, severe AA; vSAA, very severe AA.

To evaluate our approach, we benchmarked the targeted strategy against the standard 3′ single-cell RNA sequencing (scRNA-seq) run and an additional PCR approach that was performed on the cDNA from the initial targeted enrichment. We found that sequences used for the targeted or additional amplification strategies were specifically abundant in the respective enrichment library but practically absent in the native 3′ scRNA-seq library (Figure S1A). We further evaluated the overall performance of our enrichment strategy, which showed that the distribution of counts for all (non-*U2AF1*) genes was strongly reduced (Figures S1B and S1C), whereas *U2AF1*^{S34Y} locus-specific coverage was greatly increased in the targeted and additional PCR libraries (Figures S1D and S2). While the additional PCR did not outperform the targeted approach, we conclude that both enrichment strategies are highly comparable (Figures S1E–S1G), overall highly efficient, and consistent in providing *U2AF1*^{S34Y} locus-specific information.

Unambiguous recipient-derived posttransplant MDS diagnosis enables personalized treatment choice and disease resolution

Having established the methodological framework, we first used the 3′ scRNA-seq libraries to identify major cell types (Figure 4B) and assign individual cells to their origin (recipient vs. donor) via single-nucleotide polymorphism (SNP)-based deconvolution (Figure 4C) and expression of sex-specific genes (Figure 4F). This showed that the immune compartment predominantly consisted of male cells, whereas megakaryocytes and monocytes were almost all female (Figure 4D), indicating that, in the chimeric BM, immediate myelopoiesis mainly originated from the recipient, while lymphocytes stemmed from the donor. This was consistent with secondary rather than donor-derived MDS, especially since MDS can be accompanied by monocytosis.[29] Moreover, CD34^{+} donor chimerism further decreased at month 33 compared to month 31 (Figure 1C) together with a recipient-derived increase and decrease in megakaryocytes and monocytes, respectively (Figures S1H and S1I). This indicated progressive expansion of the MDS clone via megakaryocytic proliferation, which is in line with the morphologic dysplasia we observed in this lineage.

We then leveraged the targeted enrichment libraries to further genotype cells based on their *U2AF1*-specific mutational status, which enabled us to identify a total of 25 and 3,019 cells carrying either the mutated (*U2AF1*^{S34Y}) or wild-type (*U2AF1*^{WT}) variant, respectively (Figure 4E). Finally, integration of these findings with our previous analyses revealed that mutated *U2AF1*^{S34Y} cells were recipient-derived female, while *U2AF1*^{WT}-cells were donor-derived male, as illustrated for monocytes and megakaryocytes (Figure 4G).

Following our results, the patient was scheduled for a second allo-HSCT from her HLA-identical brother at month 34 instead of receiving an allo-HSCT from a partially mismatched unrelated donor, thereby reducing the risk of severe GvHD. So far, longitudinal examination of the recipient's BM showed complete remission of MDS with normal flow cytometry results (Della Porta score 0), absence of MDS-related mutations, and a donor chimerism of 100% (Figures 1C and S3).

DISCUSSION

The outcome of SAA has improved substantially since allo-HSCT emerged as curative upfront treatment and outperforms immunosuppressive therapy in younger patients.[30–32] However,

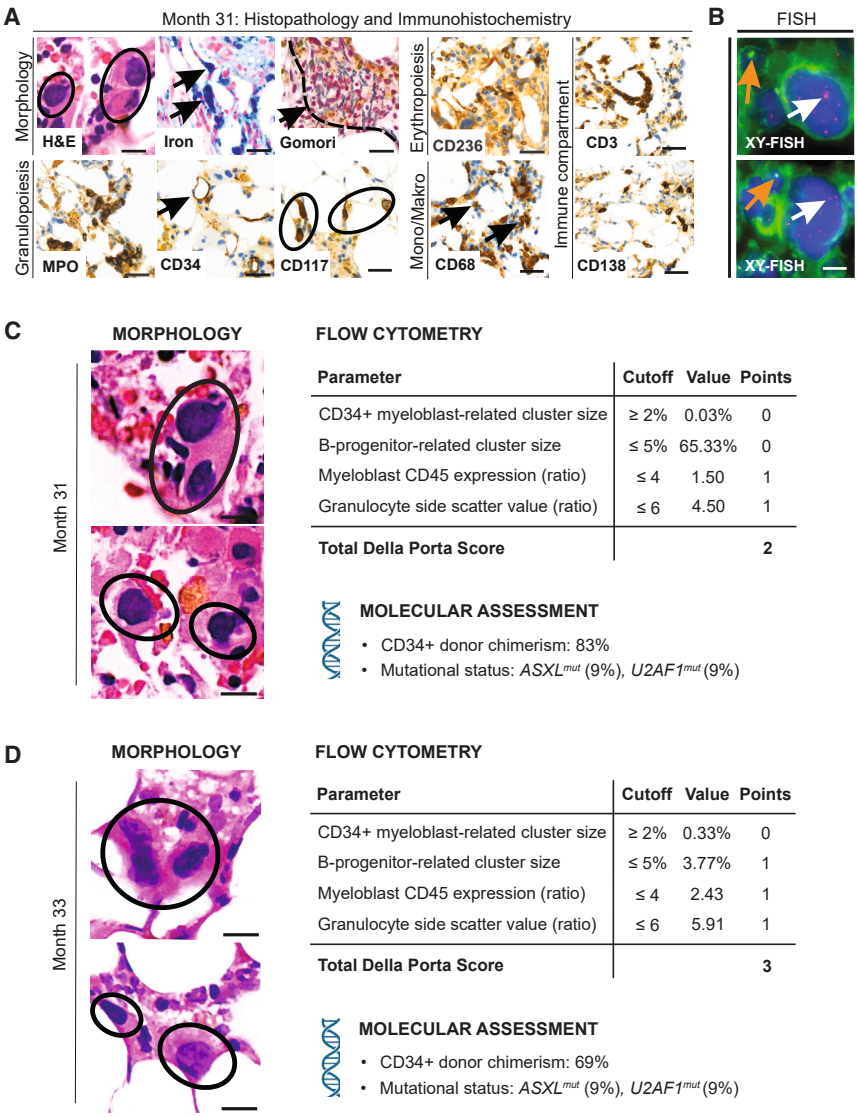


Figure 3. Clinical manifestation of posttransplant onset and progression of MDS

(A) Histopathology of the recipient's BM 31 months after initial diagnosis of AA. H&E stain shows increased cellularity as compared to initial diagnosis with a left shift of granulopoiesis, dysmorphic small megakaryocytes with hypolobulated nuclei, abnormal nucleus:cytoplasm ratio, and condensed chromatin (encircled). However, with a cellularity of 40%, the BM is still hypocellular. Prussian blue stain reveals severe iron overload in stromal macrophages (arrows). Gomori silver stain shows mild to moderate expansion of reticulin fibers (right area next to dashed line). CD236 stain is mostly negative, since erythropoiesis is nearly absent. Myeloperoxidase stain marks granulocytes. CD34 is expressed by capillary endothelia. Myeloblasts are not visible. CD117 stain shows a few randomly distributed mast cells (circles). CD68 stain reveals mostly mature stromal macrophages but also some monocytic cells (arrows). CD3 highlights small aggregates of T lymphocytes that were enriched in the recipient as a morphological correlate for T cell-dominant lymphoplasmacytosis. CD138 stain highlights slightly enriched plasma cells. Scale bars: 70 μm (H&E) and 100 μm (immunohistochemistry).

(B) XY-FISH of the recipient's BM 31 months after initial diagnosis of AA reveals an accumulation of X chromosome signals (red signal, highlighted by a white arrow) in dysmorphic cells, particularly in polyploid megakaryocytes. Y chromosome signals (green signal, highlighted by an orange arrow) were mostly visible in small regular-shaped immune cells. Scale bar: 15 μm.

(C) Further routine BM assessment at month 31 reveals clear signs of MDS, including severe dysmorphias in the megakaryocytic lineage as described in (A), MDS-compatible flow cytometry results (Della Porta score[20] of 2), decrease in donor chimerism, and mutations in the *ASXL1* and *U2AF1* genes. Scale bar: 30 μm.

(D) Routine BM assessment at month 33 shows progression of MDS with severe dysmorphias in the megakaryocytic lineage as described in (A), an increased Della Porta score[20] of 3, a further decrease in donor chimerism, and maintenance of allele frequencies for mutations in the *ASXL1* and *U2AF1* genes. Scale bar: 30 μm. mut, mutated.

a considerable residual risk of disease progression, relapse or clonal evolution to MDS, or acute myeloid leukemia remains.[5,8,16,33,34] The choice of an adequate donor is particularly important, since GvHD is considered a main complication and cause of mortality after allo-HSCT. If available, allo-HSCT from an HLA-identical sibling donor is therefore the therapy of choice for SAA.[5,6,9,30] Nevertheless, our study shows that additional complications can arise in these cases with the occurrence of clonal disorders secondary to AA, such as posttransplant MDS, requiring a second or third allo-HSCT. Clinical routine diagnostic tools comprise morphological (cytology and/or histology) or flow cytometry-based assessment for this scenario.[11,12,20] Additionally, chimerism analysis, panel sequencing, and minimal residual disease analysis are used as an approximation of the amount of the patient's remaining autologous cells and presence of MDS-associated mutations.[11,12,35,36] However, these measures do not allow us to distinguish therapy-related dysmorphia from disease-associated dysplasia and, in the case of mixed chimerism, cannot determine whether the MDS clone originates from the patient (recipient derived) or cells received from the transplant (donor derived). Although donor-derived myeloid neoplasia is usually rare,[10] the distinction between donor- and recipient-derived MDS is of utmost importance for further therapeutic strategies, especially in patients who are in need of repeated allo-HSCT after initial allo-HSCT from an HLA-identical sibling. In the case of donor-derived MDS, an alternative donor needs to be found that might potentially hold less HLA compatibility than the sibling, therefore likely increasing the risk of GvHD.

More recent molecular tools hold the potential to remedy the difficulty of solving recipient- vs. donor-derived MDS. While bulk approaches generate only average readouts that do not allow us to disentangle information from individual cells,[37,38] single-cell

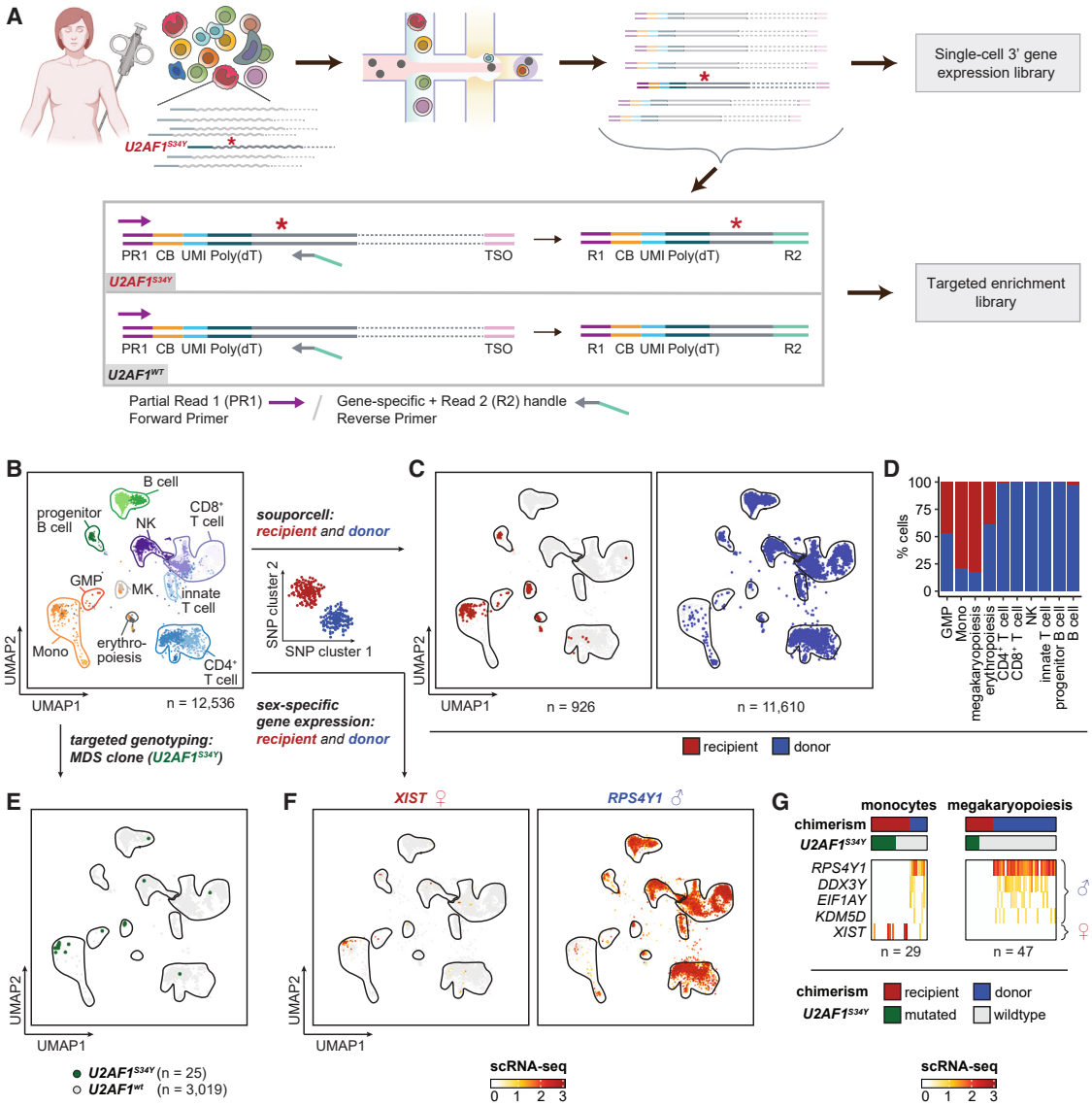


Figure 4. Identification of MDS clones using individualized single-cell genomics

(A) Single-cell strategy used to distinguish recipient- from donor-derived cells. BM-derived cells from the patient were used to generate barcoded cDNA from which single-cell whole-transcriptome and *U2AF1*-targeted enrichment libraries were constructed. PR1, partial read 1; CB, cell barcode; UMI, unique molecular identifier; TSO, template switch oligo; R1, read 1; R2, read 2.

(B) UMAP projection of BM-derived scRNA-seq profiles at 31 (n = 6,494 cells) and 33 (n = 6,042 cells) months following diagnosis of AA and onset of MDS. Major cell types are indicated by colors and labels.

UMAP, uniform manifold approximation and projection.

(C and D) Annotation of recipient-derived (red, n = 926 cells) and donor-derived (blue, n = 11,610 cells) single-cell profiles by genetic deconvolution using expressed SNPs (souporcell[28]). Shown are projection of cells on UMAP (C) and the amount of recipient or donor chimerism across cell types at 31 months post AA diagnosis (D).

(E) Identification of *U2AF1*^{S34Y}-mutated (dark green, n = 25) and *U2AF1*^{WT} (gray, n = 3,019) cells using targeted sequencing at 31 months post AA diagnosis and projection onto UMAP.

(F) Expression of the sex-specific genes *XIST* (female specific) and *RPS4Y1* (male specific).

(G) Co-analysis of chimerism analysis, targeted sequencing of *U2AF1*, and sex-specific gene expression for monocytes (n = 29) and megakaryocytes (n = 47). *U2AF1*^{S34Y}-mutated cells are recipient derived and do not have Y chromosome-specific gene expression.

approaches have provided an unprecedented opportunity to investigate heterogeneous cell populations in the human BM.[39,40] However, standard and commercially available protocols are mostly biased toward limited sequence information from the 3′ or 5′ end of the mRNA,[41–43] thus often not capturing or providing sufficient coverage of the mutational locus that would be required

for the identification of malignant cells. In principle, this could be addressed with long-read sequencing strategies, although, in practice, their use remains very limited, either because high error rates hamper robust detection of mutations at single-nucleotide resolution or because elaborate sample preparation, high costs, and the scarcity of instruments make them unfeasible in the clinical context.[44–46]

Our work highlights the real-life setting of a young patient at our hospital affected by this clinical scenario but for whom the origin of MDS could not be diagnosed. This prompted us to develop an approach that could be readily implemented within a short time frame. Notably, our single-cell strategy leverages existing single-cell gene expression protocols that are widely available in research hospitals to enable additional genotyping of each cell based on its MDS-specific *U2AF1*$^{WT/S34Y}$ status with high accuracy. Ultimately, this allowed us to unambiguously diagnose recipient-derived MDS in our patient. Although further optimization and benchmarking would be required, our single-cell approach can, in principle, be readily adapted for the profiling of other disease-specific mutations, thus offering an unprecedented opportunity for precision diagnostics and personalized therapy.

Limitations of the study

Our study highlights the clinical potential of implementing customized high-resolution techniques to solve complex cases that would otherwise not be possible with typically available methods. While our *U2AF1*^{S34Y}-based single-cell strategy allowed us to effectively address the origin of MDS in our patient, an individual approach and further investigation are required for each case that take into account patient and sample diversity. The study of other disease-specific mutations also highly depends on the feasibility of designing specific primers around the locus of interest and the need to optimize PCR and DNA purification steps. Additionally, any methodology for personalized diagnostics that is not part of clinical routine is currently not funded by public health care. To make these approaches more accessible, future research may be needed to develop cost-effective and multiplexed strategies, such as via the availability of single-cell-resolved MDS-specific gene panels.

RESOURCE AVAILABILITY

Lead contact

Requests for further information, resources, and reagents should be directed to the lead contact, Samantha Praktiknjo (samantha.praktiknjo@bih-charite.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- scRNA-seq count data have been deposited as listed in the key resources table and are publicly available as of the date of publication.
- Code for scRNA-seq data analyses is available as listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We are deeply grateful to the patient and the donor who made this work possible. We thank Anja Sieber and Caroline Bräuning for helpful discussions and technical support. J.I. and P.B. are participants in the BIH Charité Clinician Scientist Program funded by the DFG, the Charité – Universitätsmedizin Berlin, and the Berlin Institute of Health at Charité (BIH). J.I. receives funding from the Berliner Krebsgesellschaft (IHFF202110) and the Else Kröner-Fresenius-Stiftung (2023_EKEA.176). L.P. was a Scholar of the American Society of Hematology; is a participant in the BIH Charité Digital Clinician Scientist Program funded by the DFG, the Charité – Universitätsmedizin Berlin, and the Berlin Institute of Health at Charité (BIH); and is supported by the Max-Eder Program of the German Cancer Aid (Deutsche Krebshilfe), the Else Kröner-Fresenius-Stiftung (2023_EKEA.102), and the DKMS John Hansen Research Grant. C.J.W. is supported by National Institutes of Health National Cancer Institute grant P01CA229092. Icons were created with BioRender.

AUTHOR CONTRIBUTIONS

J.I. and S.D.P. designed the study and had unrestricted access to all data. L.G.V., J.W., and L.B. provided clinical samples and performed cytological assessment, FACS analysis, clinical workup, and input for data interpretation. J.I., A.-C.v.B., and A.B. performed histopathological BM assessment and FISH analysis. O.B. performed bulk DNA sequencing and chimerism analyses. J.I., P.B., A.T., and T.C. performed single-cell experiments. L.P., B.O., and P.B. conducted computational analysis of single-cell data. D.H., A.-C.v.B., M.M., L.B., C.J.W., and N.B. contributed resources and conceptual input. J.I., L.P., and S.D.P. analyzed data and prepared figures. S.D.P. established the single-cell strategy and supervised the study. J.I. and S.D.P. wrote the manuscript with input from all authors. All authors read and approved the final article and take responsibility for its content.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Histopathology and XY-FISH
  - Flow cytometry
  - Cytology
  - DNA sequencing
  - STR-based donor chimerism analysis
  - Single-cell sequencing and targeted enrichment
  - Computational analysis of single-cell data

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2024.11.001.

Received: October 10, 2023
Revised: June 13, 2024
Accepted: November 1, 2024
Published: December 6, 2024

REFERENCES

1. Ahmed, P., Chaudhry, Q.U.N., Satti, T.M., Mahmood, S.K., Ghafoor, T., Shahbaz, N., Khan, M.A., Satti, H.S., Akram, Z., and Iftikhar, R. (2020).

Epidemiology of aplastic anemia: a study of 1324 cases. Hematology *25*, 48–54. https://doi.org/10.1080/16078454.2019.1711344.

2. Scopes, J., Bagnara, M., Gordon-Smith, E.C., Ball, S.E., and Gibson, F.M. (1994). Haemopoietic progenitor cells are reduced in aplastic anaemia. Br. J. Haematol. *86*, 427–430. https://doi.org/10.1111/j.1365-2141.1994.tb04761.x.
3. Maciejewski, J.P., Anderson, S., Katevas, P., and Young, N.S. (1994). Phenotypic and functional analysis of bone marrow progenitor cell compartment in bone marrow failure. Br. J. Haematol. *87*, 227–234. https://doi.org/10.1111/j.1365-2141.1994.tb04903.x.
4. Brodsky, R.A., and Jones, R.J. (2005). Aplastic anaemia. Lancet *365*, 1647–1656. https://doi.org/10.1016/S0140-6736(05)66515-4.
5. Bacigalupo, A. (2017). How I treat acquired aplastic anemia. Blood *129*, 1428–1436. https://doi.org/10.1182/blood-2016-08-693481.
6. Schrezenmeier, H., Körper, S., and Höchsmann, B. (2015). [Aplastic anemia: Current state of diagnosis and treatment]. Internist *56*, 989–999. https://doi.org/10.1007/s00108-015-3662-7.
7. Kojima, S., Nakao, S., Tomonaga, M., Hows, J., Marsh, J., Gerard, S., Bacigalupo, A., and Mizoguchi, H. (2000). Consensus Conference on the Treatment of Aplastic Anemia. Int. J. Hematol. *72*, 118–123.
8. Killick, S.B., Bown, N., Cavenagh, J., Dokal, I., Foukaneli, T., Hill, A., Hillmen, P., Ireland, R., Kulasekararaj, A., Mufti, G., et al. (2016). Guidelines for the diagnosis and management of adult aplastic anaemia. Br. J. Haematol. *172*, 187–207. https://doi.org/10.1111/bjh.13853.
9. Peslak, S.A., Olson, T., and Babushok, D.V. (2017). Diagnosis and Treatment of Aplastic Anemia. Curr. Treat. Options Oncol. *18*, 70. https://doi.org/10.1007/s11864-017-0511-z.
10. Shah, N.N., Bacher, U., Fry, T., Calvo, K.R., Stetler-Stevenson, M., Arthur, D.C., Kurlander, R., Baird, K., Wise, B., Giralt, S., et al. (2012). Myelodysplastic syndrome after allogeneic hematopoietic stem cell transplantation: diagnostic and therapeutic challenges. Am. J. Hematol. *87*, 916–922. https://doi.org/10.1002/ajh.23174.
11. Khoury, J.D., Solary, E., Abla, O., Akkari, Y., Alaggio, R., Apperley, J.F., Bejar, R., Berti, E., Busque, L., Chan, J.K.C., et al. (2022). The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia *36*, 1703–1719. https://doi.org/10.1038/s41375-022-01613-1.
12. Arber, D.A., Orazi, A., Hasserjian, R.P., Borowitz, M.J., Calvo, K.R., Kvasnicka, H.M., Wang, S.A., Bagg, A., Barbui, T., Branford, S., et al. (2022). International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. Blood *140*, 1200–1228. https://doi.org/10.1182/blood.2022015850.
13. Kaito, K., Kobayashi, M., Katayama, T., Masuoka, H., Shimada, T., Nishiwaki, K., Sekita, T., Otsubo, H., Ogasawara, Y., and Hosoya, T. (1998). Long-term administration of G-CSF for aplastic anaemia is closely related to the early evolution of monosomy 7 MDS in adults. Br. J. Haematol. *103*, 297–303. https://doi.org/10.1046/j.1365-2141.1998.01014.x.
14. Ogawa, S. (2016). Clonal hematopoiesis in acquired aplastic anemia. Blood *128*, 337–347. https://doi.org/10.1182/blood-2016-01-636381.
15. Young, N.S., and Ogawa, S. (2015). Somatic Mutations and Clonal Hematopoiesis in Aplastic Anemia. N. Engl. J. Med. *373*, 1675–1676. https://doi.org/10.1056/NEJMc1509703.
16. Kulasekararaj, A.G., Jiang, J., Smith, A.E., Mohamedali, A.M., Mian, S., Gandhi, S., Gaken, J., Czepulkowski, B., Marsh, J.C.W., and Mufti, G.J. (2014). Somatic mutations identify a subgroup of aplastic anemia patients who progress to myelodysplastic syndrome. Blood *124*, 2698–2704. https://doi.org/10.1182/blood-2014-05-574889.
17. Socie, G., Rosenfeld, S., Frickhofen, N., Gluckman, E., and Tichelli, A. (2000). Late clonal diseases of treated aplastic anemia. Semin. Hematol. *37*, 91–101.
18. Kulasekararaj, A., Cavenagh, J., Dokal, I., Foukaneli, T., Gandhi, S., Garg, M., Griffin, M., Hillmen, P., Ireland, R., Killick, S., et al. (2024). Guidelines for the diagnosis and management of adult aplastic anaemia: A British Society for Haematology Guideline. Br. J. Haematol. *204*, 784–804. https://doi.org/10.1111/bjh.19236.
19. Gondek, L.P., and DeZern, A.E. (2014). I walk the line: how to tell MDS from other bone marrow failure conditions. Curr. Hematol. Malig. Rep. *9*, 389–399. https://doi.org/10.1007/s11899-014-0224-3.
20. Della Porta, M.G., Picone, C., Pascutto, C., Malcovati, L., Tamura, H., Handa, H., Czader, M., Freeman, S., Vyas, P., Porwit, A., et al. (2012). Multicenter validation of a reproducible flow cytometric score for the diagnosis of low-grade myelodysplastic syndromes: results of a European LeukemiaNET study. Haematologica *97*, 1209–1217. https://doi.org/10.3324/haematol.2011.048421.
21. Graubert, T.A., Shen, D., Ding, L., Okeyo-Owuor, T., Lunn, C.L., Shao, J., Krysiak, K., Harris, C.C., Koboldt, D.C., Larson, D.E., et al. (2011). Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat. Genet. *44*, 53–57. https://doi.org/10.1038/ng.1031.
22. Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. Nature *478*, 64–69. https://doi.org/10.1038/nature10496.
23. Thol, F., Kade, S., Schlarmann, C., Löffeld, P., Morgan, M., Krauter, J., Wlodarski, M.W., Kölking, B., Wichmann, M., Görlich, K., et al. (2012). Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood *119*, 3578–3584. https://doi.org/10.1182/blood-2011-12-399337.
24. Xu, L., Gu, Z.H., Li, Y., Zhang, J.L., Chang, C.K., Pan, C.M., Shi, J.Y., Shen, Y., Chen, B., Wang, Y.Y., et al. (2014). Genomic landscape of CD34+ hematopoietic cells in myelodysplastic syndrome and gene mutation profiles as prognostic markers. Proc. Natl. Acad. Sci. USA *111*, 8589–8594. https://doi.org/10.1073/pnas.1407688111.
25. Gelsi-Boyer, V., Brecqueville, M., Devillier, R., Murati, A., Mozziconacci, M.J., and Birnbaum, D. (2012). Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. J. Hematol. Oncol. *5*, 12. https://doi.org/10.1186/1756-8722-5-12.
26. Tefferi, A., Mudireddy, M., Finke, C.M., Nicolosi, M., Lasho, T.L., Hanson, C.A., Patnaik, M.M., Pardanani, A., and Gangat, N. (2018). U2AF1 mutation variants in myelodysplastic syndromes and their clinical correlates. Am. J. Hematol. *93*, E146–E148. https://doi.org/10.1002/ajh.25084.
27. Negoro, E., Nagata, Y., Clemente, M.J., Hosono, N., Shen, W., Nazha, A., Yoshizato, T., Hirsch, C., Przychodzen, B., Mahfouz, R.Z., et al. (2017). Origins of myelodysplastic syndromes after aplastic anemia. Blood *130*, 1953–1957. https://doi.org/10.1182/blood-2017-02-767731.
28. Heaton, H., Talman, A.M., Knights, A., Imaz, M., Gaffney, D.J., Durbin, R., Hemberg, M., and Lawniczak, M.K.N. (2020). Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. Nat. Methods *17*, 615–620. https://doi.org/10.1038/s41592-020-0820-1.
29. Wu, A., Gao, P., Wu, N., Shi, C., Huang, Z., Rong, C., Sun, Y., Sheng, L., Ouyang, G., and Mu, Q. (2021). Elevated mature monocytes in bone marrow accompanied with a higher IPSS-R score predicts a poor prognosis in myelodysplastic syndromes. BMC Cancer *21*, 546. https://doi.org/10.1186/s12885-021-08303-8.
30. Drexler, B., Zurbriggen, F., Diesch, T., Viollier, R., Halter, J.P., Heim, D., Holbro, A., Infanti, L., Buser, A., Gerull, S., et al. (2020). Very long-term follow-up of aplastic anemia treated with immunosuppressive therapy or allogeneic hematopoietic cell transplantation. Ann. Hematol. *99*, 2529–2538. https://doi.org/10.1007/s00277-020-04271-4.
31. Locasciulli, A., Oneto, R., Bacigalupo, A., Socié, G., Korthof, E., Bekassy, A., Schrezenmeier, H., Passweg, J., and Führer, M.; Severe Aplastic Anemia Working Party of the European Blood and Marrow Transplant Group (2007). Outcome of patients with acquired aplastic anemia given first line bone marrow transplantation or immunosuppressive treatment in the last decade: a report from the European Group for Blood and Marrow Transplantation (EBMT). Haematologica *92*, 11–18. https://doi.org/10.3324/haematol.10075.

32. Bacigalupo, A., Giammarco, S., and Sica, S. (2016). Bone marrow transplantation versus immunosuppressive therapy in patients with acquired severe aplastic anemia. Int. J. Hematol. *104*, 168–174. https://doi.org/10.1007/s12185-016-2037-8.

33. Sun, L., and Babushok, D.V. (2020). Secondary myelodysplastic syndrome and leukemia in acquired aplastic anemia and paroxysmal nocturnal hemoglobinuria. Blood *136*, 36–49. https://doi.org/10.1182/blood.2019000940.

34. Georges, G.E., Doney, K., and Storb, R. (2018). Severe aplastic anemia: allogeneic bone marrow transplantation as first-line treatment. Blood Adv. *2*, 2020–2028. https://doi.org/10.1182/bloodadvances.2018021162.

35. Tozzo, P., Delicati, A., Zambello, R., and Caenazzo, L. (2021). Chimerism Monitoring Techniques after Hematopoietic Stem Cell Transplantation: An Overview of the Last 15 Years of Innovations. Diagnostics *11*, 621. https://doi.org/10.3390/diagnostics11040621.

36. Schwind, S., Jentzsch, M., Kubasch, A.S., Metzeler, K.H., and Platzbecker, U. (2021). Myelodysplastic syndromes: Biological and therapeutic consequences of the evolving molecular aberrations landscape. Neoplasia *23*, 1101–1109. https://doi.org/10.1016/j.neo.2021.09.002.

37. Carter, B., and Zhao, K. (2021). The epigenetic basis of cellular heterogeneity. Nat. Rev. Genet. *22*, 235–250. https://doi.org/10.1038/s41576-020-00300-0.

38. Lim, J., Chin, V., Fairfax, K., Moutinho, C., Suan, D., Ji, H., and Powell, J.E. (2023). Transitioning single-cell genomics into the clinic. Nat. Rev. Genet. *24*, 573–584. https://doi.org/10.1038/s41576-023-00613-w.

39. van Galen, P., Hovestadt, V., Wadsworth Ii, M.H., Hughes, T.K., Griffin, G.K., Battaglia, S., Verga, J.A., Stephansky, J., Pastika, T.J., Lombardi Story, J., et al. (2019). Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. Cell *176*, 1265–1281.e24. https://doi.org/10.1016/j.cell.2019.01.031.

40. Triana, S., Vonficht, D., Jopp-Saile, L., Raffel, S., Lutz, R., Leonce, D., Antes, M., Hernández-Malmierca, P., Ordoñez-Rueda, D., Ramasz, B., et al. (2021). Single-cell proteo-genomic reference maps of the hematopoietic system enable the purification and massive profiling of precisely defined cell states. Nat. Immunol. *22*, 1577–1589. https://doi.org/10.1038/s41590-021-01059-0.

41. Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002.

42. Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044.

43. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. https://doi.org/10.1038/ncomms14049.

44. Probst, V., Simonyan, A., Pacheco, F., Guo, Y., Nielsen, F.C., and Bagger, F.O. (2022). Benchmarking full-length transcript single cell mRNA sequencing protocols. BMC Genom. *23*, 860. https://doi.org/10.1186/s12864-022-09014-5.

45. Lebrigand, K., Magnone, V., Barbry, P., and Waldmann, R. (2020). High throughput error corrected Nanopore single cell transcriptome sequencing. Nat. Commun. *11*, 4025. https://doi.org/10.1038/s41467-020-17800-6.

46. Volden, R., and Vollmers, C. (2022). Single-cell isoform analysis in human immune cells. Genome Biol. *23*, 47. https://doi.org/10.1186/s13059-022-02615-z.

47. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

48. Zorita, E., Cuscó, P., and Filion, G.J. (2015). Starcode: sequence clustering based on all-pairs search. Bioinformatics *31*, 1913–1919. https://doi.org/10.1093/bioinformatics/btv053.

49. Melsted, P., Booeshaghi, A.S., Liu, L., Gao, F., Lu, L., Min, K.H.J., da Veiga Beltrame, E., Hjörleifsson, K.E., Gehring, J., and Pachter, L. (2021). Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat. Biotechnol. *39*, 813–818. https://doi.org/10.1038/s41587-021-00870-2.

50. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

51. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag).

52. Penter, L., Borji, M., Nagler, A., Lyu, H., Lu, W.S., Cieri, N., Maurer, K., Oliveira, G., Al'Khafaji, A.M., Garimella, K.V., et al. (2024). Integrative genotyping of cancer and immune phenotypes by long-read sequencing. Nat. Commun. *15*, 32. https://doi.org/10.1038/s41467-023-44137-7.

53. Tzankov, A., Dirnhofer, S., and Beham-Schmid, C. (2012). Normales Knochenmark und häufige reaktive Veränderungen. Pathologe *33*, 496–507. https://doi.org/10.1007/s00292-012-1649-x.

54. Penter, L., Liu, Y., Wolff, J.O., Yang, L., Taing, L., Jhaveri, A., Southard, J., Patel, M., Cullen, N.M., Pfaff, K.L., et al. (2023). Mechanisms of response and resistance to combined decitabine and ipilimumab for advanced myeloid disease. Blood *141*, 1817–1830. https://doi.org/10.1182/blood.2022018246.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Histology: Glycophorin C/CD236, Ret40f	Dako/Agilent	Cat#M0820; RRID:AB_2248164
Histology: Myeloperoxidase, polyclonal	Dako/Agilent	Cat#A0398; RRID:AB_2335676
Histology: CD34, Qbend/10	Leica	Cat#NCL-L-END; RRID:AB_563552
Histology: CD117, polyclonal	Dako/Agilent	Cat#A4502; RRID:AB_2335702
Histology: CD14, SP192	Abcam	Cat#ab183322; RRID:AB_2909463
Histology: CD68, PG-M1	Dako/Agilent	Cat#M0876; RRID:AB_2074844
Histology: CD61, 2f2	Leica	Cat#PA0312; RRID:AB_10554439
Histology: CD20, L26	Dako/Agilent	Cat#M0755; RRID:AB_2282030
Histology: CD3, polyclonal	Dako/Agilent	Cat#A0452; RRID:AB_2335677
Histology: CD138, MI15	Dako/Agilent	Cat#M7228; RRID:AB_2254116
FACS: CD45, Krome Orange	Beckman Coulter	Cat#B36294; RRID:AB_2833027
FACS: CD34, APC-A700	Beckman Coulter	Cat#B92417
FACS: CD33, PE, D3HL60.251	Beckman Coulter	Cat#A07775
FACS:CD13-PC5.5, Immu103.44	Beckman Coulter	Cat#B49196
FACS: CD117, PE, 95C3	Beckman Coulter	Cat#IM1360U
FACS: CD38-PB, LS198-4-3	Beckman Coulter	Cat#B92396
FACS: HLA-DR-PE, Immu-357	Beckman Coulter	Cat#IM1639; RRID:AB_131284
FACS: CD15-PB	Beckman Coulter	Cat#B49218
FACS: CD11b, APC, Bear1	Beckman Coulter	Cat#A87782; RRID:AB_2938661
FACS: CD25-ECD, B1.49.9	Beckman Coulter	Cat#6607112
FACS: CD14-ECD, RMO52	Beckman Coulter	Cat#B92391
FACS: CD19-ECD, J3-119	Beckman Coulter	Cat#A07770; RRID:AB_2940901
FACS: CD10, FITC, ALB1	Beckman Coulter	Cat#A07759
Isotype control FITC IgG1	Beckman Coulter	Cat#A07795; RRID:AB_2832964
Isotype control PE IgG1	Beckmann Coulter	Cat#A07796; RRID:AB_2832963
Isotype control ECD IgG1	Beckmann Coulter	Cat#A07797
Isotype control PC 5.5 IgG1	Beckmann Coulter	Cat#A62833
Isotype control PC7, IgG1	Beckmann Coulter	Cat#737662
Isotype control APC IgG1	Beckmann Coulter	Cat#IM2475; RRID:AB_130786
Isotype control PacBlue IgG1	Beckmann Coulter	Cat#74764
Biological samples		
Patient-derived bone marrow samples	This paper	N/A
Chemicals, peptides, and recombinant proteins		
EDTA soft decalcifier	Roth	Cat#6484.2
Haemalaun Mayer	Waldeck	Cat#2E−038
Eosin Y Solution	Sigma-Aldrich	Cat#HT 110132-1L
Schiff's reagent	Sigma-Aldrich	Cat#1.09033.0500
Giemsa	Sigma-Aldrich	Cat#1.09204.0500
Periodic Acid Solution	Sigma-Aldrich	Cat#395132-1L
Weigert A	Sigma-Aldrich	Cat#HT107-500ML
Weigert B	Sigma-Aldrich	Cat#HT109-500ML
Goldner 1	Dr. K. Hollborn	Cat#G19-1000
Goldner 2	Dr. K. Hollborn	Cat#G20-1000

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Goldner 3	Dr. K. Hollborn	Cat#G21-1000
TOMO slides	TOMO	Cat#TOM-1190
Iron Stain Kit	Agilent (DAKO)	Cat#AR158
Reticulin-Nuclear Fast Red	Agilent (DAKO)	Cat#AR179
Pharm Lyse lysing buffer	BD Biosciences	Cat#555899
Lymphoprep™	Stemcell Technologies	Cat#07801
ACK Lysing buffer	Thermo Fisher Scientific	Cat#A1049201
Critical commercial assays		
Chromium Next GEM Single Cell 3' Reagent Kit v3.1	10X Genomics	Prod#1000128
Chromium Next GEM Chip G Single Cell Kit	10X Genomics	Prod#1000127
KAPA HiFi HotStart ReadyMix	Roche	Cat#KK2601
SPRIselect Reagent	Beckmann Coulter	Cat#B23317
NextSeq 500/550 v2.5 Kit	Illumina	Cat#20024907
Novaseq 6000 SP Reagent Kit v1.5	Illumina	Cat#20040719
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat#32851
AmpFLSTR™ Identifier™ PCR Amplification Kit	Applied Biosystems	Cat#4322288
QIAamp DNAmini Kit	Qiagen	Cat#56304
High Sensitivity D5000 ScreenTape Assay	Agilent	Cat#5067-5592
NEXTflex Myeloid Amplicon Panel	Perkin Elmer	Cat#NOVA-4260
BOND Polymer Refine Detection	Leica	Cat#DS9800
Deposited data		
scRNA-seq count data	This paper	GEO: GSE228278
Oligonucleotides		
10X Genomics cDNA forward primer: CTACACGACGCTCTTCCGATCT	Integrated DNA Technologies	N/A
Custom *U2AF1*-specific reverse primer for targeted PCR: GTGACTGGAGTTCA GACGTGTGCTCTTCCGATCTAGCAT GTCGTCATGGAGACA	Integrated DNA Technologies	N/A
Custom *U2AF1*-specific reverse primer for additional PCR: GTGACTGGAGTTCA GACGTGTGCTCTTCCGATCTGTTCAG ACGTGTGCTCTTCC	Integrated DNA Technologies	N/A
Software and algorithms		
MiSeq Reporter 2.6	Illumina	https://support.illumina.com/sequencing/sequencing_software/miseq_reporter/downloads.html
VariantStudio 3.0	Illumina	https://support.illumina.com/sequencing/sequencing_software/variantstudio/downloads.html
GeneMapper 3.7	Applied Biosystems	https://genemapper.software.informer.com/3.7/
TapeStation Software 3.1.1	Agilent Technologies	https://www.agilent.com/en/product/automated-electrophoresis/tapestation-systems/tapestation-software/tapestation-software-379381
Kaluza Software 2.1	Beckman Coulter	https://www.beckman.com/flow-cytometry/software/kaluza

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Code for scRNA-seq data analyses	This paper	https://github.com/liviuspenter/MDS https://doi.org/10.5281/zenodo.13988685
Cell Ranger 7.1.0	10X Genomics	https://www.10xgenomics.com/support/software/cell-ranger/latest
R 4.2.2	N/A	https://cran.r-project.org/
R studio 2022.12	N/A	https://github.com/rstudio/rstudio
Seurat 4.3.0	Hao et al.[47]	https://github.com/satijalab/seurat
Souporcell 2.4	Heaton et al.[28]	https://github.com/wheaton5/souporcell
Pysam 0.2	N/A	https://github.com/pysam-developers/pysam
Starcode 1.3	Zorita et al.[48]	https://github.com/gui11aume/starcode
kallisto bustools 0.48	Melsted et al.[49]	https://github.com/pachterlab/kallisto
bedtools 2.30.0	Quinlan & Hall[50]	https://github.com/arq5x/bedtools2
ggplot2 3.4.0	Wickham[51]	https://github.com/tidyverse/ggplot2
nanoranger	Penter et al.[52]	https://github.com/mehdiborji/nanoranger
Other		
NovaSeq 6000	Illumina	Cat#20012850
NextSeq 500	Illumina	Cat#SY-415-1001
FACSAria III Instrument	BD Biosciences	Cat#23-11585-02
MiSeq System	Illumina	Cat#SY-410-1003
Chromium Controller	10x Genomics	Prod#1000204
40 μm mesh	Falcon	Cat #352340
4200 TapeStation system	Agilent	Part#G2991BA
Olympus BX46F	Olympus	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The patient was a 38-year-old Caucasian female. Her donor was her 38-year-old Caucasian male twin brother. This study is in accordance with the declaration of Helsinki and was approved by the local ethics committee (EA2/147/22). Both the patient and her donor gave informed consent for the publication of clinical results.

METHOD DETAILS

Histopathology and XY-FISH

Histopathology was performed using formalin-fixed and paraffin-embedded EDTA decalcified BM trephine biopsies obtained 26, 31 and 33 months after onset of AA. All biopsies were stained with hematoxylin-eosin (H&E), Periodic-acid Schiff stain (PAS), Giemsa-, Gomori-Silver and Prussian blue stain. Additionally, immunohistochemical staining was performed using the LEICA Bond MAX (Leica Biosystems, Buffalo, NY, USA) immunostainer. Briefly, 2 μm tissue sections were deparaffinized, rehydrated, and subjected to heat-induced epitope retrieval and endogenous peroxidase blocking with H_2O_2. Subsequently, BM slides were incubated for 30 min with the following primary antibodies: CD236 (1:400, Dako, Ret40f), Myeloperoxidase (MPO, 1:3000, Dako, polyclonal), CD34 (1:50, Leica, Qbend/10), CD117 (1:400, Dako, polyclonal), CD14 (1:200, Abcam, SP192), CD68 (1:200, Dako, PG-M1), CD61 (1:100, Leica, 2f2), CD20 (1:750, Dako, L26), CD3 (1:100, Dako, polyclocal), CD138 (1:25, Dako, MI15). All slides were examined with an Olympus BX46F and a 40x ocular lens by two hematopathologists. BM evaluation was conducted according to the standardized procedure for iliac trephine biopsies.[53] Interphase FISH analysis was performed using a CEP X Spectrum orange/CEP Y Spectrum green probe kit (Abott), targeting chromosomal regions Xp11.1-q11.1/Yq12 according to the manufacturer's instruction.

Flow cytometry

100 μL of fresh EDTA-anticoagulated BM aspirate was collected 5, 26, 31 and 33 months after onset of AA. Staining (whole blood staining method) was performed using the following antibodies for 20 min on ice: CD45-Krome Orange, CD34-APC, CD33-PE, CD13-PC5.5, CD117-PE, CD38-Pacific Blue, HLA-DR-Pacific Blue, CD15-Pacific Blue, CD11b-APC, CD25-ECD, CD4-ECD, CD19-PE, CD10-FITC and isotype controls (all Beckman Coulter, Brea, CA, USA). Red blood cells were subsequently lysed using

commercial lysing buffer (Pharm Lyse, BD Biosciences). Flow cytometry data were analyzed using the Kaluza software (Beckman Coulter) and according to the method described by Della Porta et al. for MDS.[20]

Cytology

For cytological assessment, EDTA chelated BM aspirate was collected 5, 26, 31 and 33 months after onset of AA. BM smears were prepared and stained according to Pappenheim and with Prussian blue stain. All slides were microscopically evaluated by two hematologists.

DNA sequencing

100 μL of fresh heparinized BM aspirate was collected at initial diagnosis of AA and 5, 26, 31, 33 and 34 months after onset of AA. Mononuclear cells were isolated via density gradient centrifugation (Lymphoprep, Stemcell Technologies). DNA was isolated using the QIAamp DNAmini Kit (Qiagen) according to the manufacturer's instructions. 10–40 ng DNA was used per sample. Polymerase chain reaction (PCR), purifications, adapter ligation, library quantification and pooling were conducted as outlined in the recommended BIOO Scientific protocol "NEXTflex Myeloid Amplicon Panel" (PerkinElmer). Sequencing data was automatically analyzed on the Illumina MiSeq instrument using the MiSeq Reporter. FASTQ files were generated as output and reads aligned to the reference human genome (hg19) with deviations from the reference sequences stored as variants. The Variant Call Format (VCF) files were opened and annotated in Illumina Variant Studio.

STR-based donor chimerism analysis

DNA was extracted using a standard DNA extraction method (QIAamp; Qiagen) as recommended by the manufacturer. Chimerism analyses were conducted based on the discrimination of donor and recipient alleles on short tandem repeats (STRs) using the AmpFLSTR Identifier PCR Amplification Kit (Applied Biosystems). For quantification of chimerism, the areas under the curves were calculated using Genemapper Version 3.7 software (Applied Biosystems).

Single-cell sequencing and targeted enrichment

BM aspirates were collected from the patient at two time points (31 and 33 months) after initial diagnosis of AA. Samples were filtered through a 40 μM mesh and mononuclear cells isolated via density gradient centrifugation (Lymphoprep, Stemcell Technologies) and an additional erythrocyte removal step using the ACK Lysing Buffer (Thermo Fisher Scientific).

For the sample collected at month 31, cells were processed using the Chromium Next GEM Single Cell 3' Reagent Kit v3.1 and the Chromium Controller (10x Genomics) to generate barcoded, full-length cDNA from which a whole-transcriptome 3′-scRNA-seq library was produced according to the manufacturer's instructions.

For targeted enrichment of the *U2AF1*^{S34Y} locus, 30 ng of total barcoded cDNA were amplified in a 50 μL PCR reaction using 1x KAPA HiFi HotStart ReadyMix (Kapa Biosystems), 0.4 μM of a cDNA forward primer containing a partial read 1 sequence (CTACACG ACGCTCTTCCGATCT) and 1 μM of a *U2AF1*-specific reverse primer containing a read 2 handle (GTGACTGGAGTTCAGACG TGTGCTCTTCCGATCTAGCATGTCGTCATGGAGACA) with the following cycling conditions as outlined in the Drop-seq protocol[41]: 95°C for 3 min; 4 cycles of: 98°C for 20 s, 65°C for 45 s, 72°C for 3 min; 9 cycles of: 98°C for 20 s, 67°C for 20 s, 72°C for 3 min; 1 cycle of 72°C for 5 min followed by cooling to 4°C. The PCR product was completed to 100 μL with H_2O and then purified using a SPRIselect Reagent (Beckman Coulter) to PCR mix ratio of 0.6x as outlined in the cDNA cleanup step of the 10x Genomics Chromium Next GEM Single Cell 3' User Guide (#CG000315 Rev C).

An additional PCR was performed using 100 ng purified DNA from the targeted enrichment with 0.4 μM of the cDNA forward primer containing a partial read 1 sequence and a reverse primer containing a partial read 2 sequence and a read 2 handle (GTGA CTGGAGTTCAGACGTGTGCTCTTCCGATCTGTTCAGACGTGTGCTCTTCC) with the same cycling and purification steps as described previously. cDNA samples from the targeted and additional PCRs were then separately indexed following the steps as outlined in the "Sample Index PCR" section of the 10x Genomics Chromium Next GEM Single Cell 3' User Guide using a total of 13 cycles. The targeted library was purified via double-sided size selection using SPRIselect to PCR mix ratios of 0.48x and 0.68x to remove unspecific large and small fragments, respectively. The additional PCR library was cleaned up using an SPRIselect to PCR mix ratio of 0.48x for the removal of unspecific large fragments. DNA fragment size distributions and DNA concentrations were determined on a TapeStation instrument using the High Sensitivity D5000 ScreenTape Assay (Agilent) and a Qubit Fluorometer using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), respectively.

The 3′-scRNA-seq, targeted and additional PCR libraries from the sample collected at month 31 were sequenced on a NextSeq 500 sequencer (Illumina) using the High Output v2.5 kit (150 cycles) in the following sequencing mode: Read 1 (28 cycles), i7 index (10 cycles), i5 index (10 cycles), Read 2 (102 cycles).

For the sample collected at month 33, mononuclear cells were sorted on a FACS Aria III Instrument with a 100 μm nozzle configuration, using PI staining (Invitrogen) to exclude dead cells. The 3′-scRNA-seq library was then generated using the Chromium Next GEM Single Cell 3' Reagent Kits v3.1 and sequenced on a NovaSeq 6000 instrument (Illumina) with the SP Reagent Kit v1.5 (200 cycles) and the same sequencing mode used previously.

Computational analysis of single-cell data

Following alignment against GRCh38 and generation of count matrices with 10x Genomics Cell Ranger, single cell gene expression profiles were analyzed using the Seurat package 4.3.0[47] as previously described.[54] Briefly, after filtering of high-quality cells (genes >200, UMIs <5000, %mitochondrial reads <20) and normalization of data, single cell profiles were annotated based on a normal human reference BM dataset provided by the Seurat package using weighted-nearest neighbor analysis.[47] Donor- and recipient-derived cells were identified through clustering based on expressed single nucleotide polymorphisms using souporcell[28] with k = 2 and information on T cell chimerism.

Following processing of amplicon libraries using 10x Genomics Cell Ranger, *U2AF1*^{S34Y} and *U2AF1*^{wt} cells were identified by analysis of reads mapping to the mutation locus (chr21:43,104,346) using pysam (https://github.com/pysam-developers/pysam) and collapsing of UMIs using starcode[48] with a Levenshtein distance of 3 as described.[52] PCR handles were quantified from raw sequencing data using kallisto bustools.[49] Pseudobulk coverage was computed using bedtools v.2.30.0.[49] Plots were visualized using ggplot2 v.3.4.0.[51]