

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- |                                     |                                     |                                                                                                                                                                                                                                                            |
|-------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of all covariates tested                                                                                                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted<br><i>Give P values as exact values whenever suitable.</i>                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i> ), indicating how they were calculated                                                                                                                                               |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection BD FACSDiva, version 8.01 (<https://www.bdbiosciences.com>)

Data analysis FlowJo 10.7.1 ([www.flowjo.com](http://www.flowjo.com)); GraphPad Prism 10.0 ([www.graphpad.com](http://www.graphpad.com)); R ([www.r-project.org](http://www.r-project.org)); Metascape (<https://metascape.org>); Seurat R package 4.3.0 (<https://satijalab.org/seurat>); MACS2 v2.2.9.1 (<https://pypi.org/project/MACS2/>); Spotfire (<https://www.spotfire.com/>); Image J Software (<https://imagej.net/ij/>); All of the software used has been included in the "Methods" section of the paper.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The datasets generated in this study using scDNA-seq, RNA-seq, long-read scRNA-seq, and ATAC-seq have been deposited under GSE276492 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE276492>), GSE276360 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE276360>), GSE276974 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE276974>).

[www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE276974](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE276974)), and GSE289476 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE289476>), respectively. Source data are provided in this paper. The remaining data are available within the article, supplementary information, or source data file.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

|                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Reporting on sex and gender                                        | We did not perform sex-based or gender based-analyses. Patient samples were selected based on diagnosis regardless of sex and gender. Males and females had comparable distributions in CCUS and healthy donor cohorts. The sex of all samples in each experiment is indicated in Supplementary Data 1. Gender is not relevant to this study.                                                                                                                                                                  |
| Reporting on race, ethnicity, or other socially relevant groupings | Patient samples were selected regardless of their race. Race is not relevant to this study.                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Population characteristics                                         | Covariate-relevant population characteristics were age, sex, cytogenetic and molecular information, and diagnosis.                                                                                                                                                                                                                                                                                                                                                                                             |
| Recruitment                                                        | Bone marrow (BM) aspirates from healthy donors were obtained from Lonza (Morristown, NJ) and AllCells (Alameda, CA) or the Department of Stem Cell Transplantation and Cellular Therapy at MD Anderson Cancer Center. BM aspirates and PB samples from CCUS patients were obtained from the Department of Leukemia at MD Anderson and the Department of Medicine at the University of California San Diego. BM biopsies from CCUS patients were obtained from the Department of Hematopathology at MD Anderson |
| Ethics oversight                                                   | All samples were obtained following approval of the corresponding Institutional Review Board at the University of Texas MD Anderson Cancer Center (Houston, TX, USA) or at the Department of Medicine at the University of California San Diego in accordance with the Declaration of Helsinki. Written informed consent was obtained from all donors. No compensation was provided to the study participants.                                                                                                 |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

|                 |                                                                                                                                                                                                                                                                                                                |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sample size     | Patient samples were used based on availability of fresh or banked samples and selected by diagnosis and therapy. A description of our patient selection criteria is included in the "Methods" section of the manuscript. The summary of patient clinical characteristics is included in Supplementary Data 1. |
| Data exclusions | No data were excluded from the analysis.                                                                                                                                                                                                                                                                       |
| Replication     | Experiment replication other than the use of biological replicates (different patient samples) was not possible. The reproducibility of the experimental findings was validated by using multiple types of analysis.                                                                                           |
| Randomization   | Allocation of human samples into experimental groups was based on diagnosis and treatment as indicated above. Randomization was only possible in studies conducted in research animals.                                                                                                                        |
| Blinding        | Experimental data analysis was performed blindly.                                                                                                                                                                                                                                                              |

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials &amp; experimental systems

|                                     |                                                                 |
|-------------------------------------|-----------------------------------------------------------------|
| n/a                                 | Involved in the study                                           |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines       |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                                 |

## Methods

|                                     |                                                    |
|-------------------------------------|----------------------------------------------------|
| n/a                                 | Involved in the study                              |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Antibodies

## Antibodies used

Use, Species, Antigen, Conjugate, Clone, Supplier, Catalog #

Flow cytometry, Human, CD366 (TIM-3), F38-2E2, BioLegend, 345028; 1:80

Flow cytometry, Human, CD223 (LAG-3), 11C3C65, BioLegend, 369314; 1:80

Flow cytometry, Human, TIGIT (VSTM3), A15153G, BioLegend, 372744; 1:80

Flow cytometry, Human, CD279 (PD-1), A17188A, BioLegend, 379210; 1:80

Flow cytometry, Human, CD3, SK7, BioLegend, 317339; 1:80

Flow cytometry, Human, CD4, SK1, BioLegend, 300502; 1:80

Flow cytometry, Human, CD8, SK1, BioLegend, 344760; 1:80

Flow cytometry, Human, CD56, B159, BioLegend, 362544; 1:80

Flow cytometry, Human, CD19, H1B19, BioLegend, 302226 and 302230; 1:200

Flow cytometry, Human, CD14, 63D3, BioLegend, 367114; 1:80

Flow cytometry, Human, CD14, HCD14, BioLegend, 325620; 1:100

Flow cytometry, Human, CD34, 581, BioLegend, 343526 and Horizon 562383; 1:10

Flow cytometry, Human, CD33, P67.6, BioLegend, 303436; 1:80

Flow cytometry, Human, CD28, CD28.2, BioLegend, 302928; 1:80

Flow cytometry, Human, CD244 (2B4), C1.7, BioLegend, 329516; 1:80

Flow cytometry, Human, CD27, O323, BioLegend, 302846; 1:80

Flow cytometry, Human, KLRG1, SA231A2, BioLegend, 367733; 1:80

Flow cytometry, Human, CD57, HNK-1, BioLegend, 359624; 1:80

Flow cytometry, Human, CD4, RPA-T4, BioLegend, 300546; 1:20

Flow cytometry, Human, CD56, 5.1H11, BioLegend, 362544; 1:20

Flow cytometry, Human, CD16, 3G8, BioLegend, 302028; 1:10

Flow cytometry, Human, TNF- $\alpha$ , Mab11, BD Biosciences, 554512; 1:20

Flow cytometry, Human, IFN- $\gamma$ , B27, BD Biosciences, 554702; 1:20

Flow cytometry, Human, Granzyme B, GB11, BioLegend, 515408; 1:20

Flow cytometry, Human, Perforin 1, Dg9, BioLegend, 308130; 1:20

Flow cytometry, Human, CD107a, H4A3, BioLegend, 328644; 1:20

Flow cytometry, Human, NKG2D, 1D11, BioLegend, 320834; 1:10

Flow cytometry, Human, NCR3, P30-15, BioLegend, 325202; 1:20

Flow cytometry, Human, NKG7, Polyclonal, Epigentek, A60028-100; 1:10

Flow cytometry, Mouse, CD45, 30-F11, BioLegend, 103108 and 304096; 1:10

Flow cytometry, Human, CD45, HI30, BD Biosciences, 555483; 1:20

Flow cytometry, mouse Tert119, Ter119, BioLegend, 116205; 1:25

Flow cytometry, Human, CD15, H198, BioLegend, 301907; 1:20

Flow cytometry, Human, CD33, WM33, BioLegend, 303416; 1:20

Flow cytometry, Human, CD71, CY1G4, BioLegend, 334116; 1:20

IF Multiplex, Human, CD3e, D7A6E, Cell Signaling Technology, 85061BF; 1:100

IF Multiplex, Human, CD8, C8/144B, Invitrogen, MA5-13473; 1:25

IF Multiplex Human, CD56, 123C3, Dako (Agilent), M7304; 1:25

IF Multiplex, Human, CD33, BCL-L-CD33, Leica, ZCL223; 1:100

Western Blot, Human, DNMT3A, N/A, Thermo Fisher, MA5-35334; 1:1000

Western Blot, Human, TET2, N/A, Cell Signaling Technology, 1895; 1:1000

Western Blot, Human,  $\beta$ -actin, N/A, Cell Signaling Technology, 5125; 1:2000

## Validation

Every antibody used in this study had been previously validated by the manufacturer. In addition, all of our flow cytometry experiments were performed using single-color and "fluorescence-minus-one" controls.

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

|                                                                      |                                                                                                                                                                                                                                                                                                                |
|----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cell line source(s)                                                  | The MDS-L cell line established from a male MDS patient was donated by Dr. Kaoru Tohyama (Department of Laboratory Medicine, Kawasaki Medical School Okayama, Japan). K562 and THP-1 cells were purchased from ATCC (CCL-243 and TIB-202, respectively), and MOLM-14 cells were purchased from DSMZ (ACC 777). |
| Authentication                                                       | The identity of all cell lines was confirmed by short tandem repeat DNA fingerprinting at MD Anderson's Characterized Cell Line Core Facility                                                                                                                                                                  |
| Mycoplasma contamination                                             | Cells were tested negative for mycoplasma contamination at MD Anderson's Characterized Cell Line Core Facility.                                                                                                                                                                                                |
| Commonly misidentified lines<br>(See <a href="#">ICLAC</a> register) | No commonly misidentified lines were used in the study.                                                                                                                                                                                                                                                        |

## Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

|                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Laboratory animals      | Mice were maintained under specific-pathogen-free conditions at MD Anderson and housed in a barrier facility at 25°C under ambient oxygen conditions in a 12-hour light/12-hour dark cycle under 50% humidity, as previously described. All animal experiments were performed with the approval of MD Anderson's Institutional Animal Care and Use Committee. NSGS mice were obtained from The Jackson Laboratory, and MISTRG mice were obtained from Regeneron.                                                                                                          |
| Wild animals            | No wild animals were used in the study.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Reporting on sex        | All NSGS mice used in this study were female, as female recipients have been shown to yield better engraftment rates than male recipients (Schmitt et al. 2022), regardless of donor sex (Mian et al., 2024). NSGS mice were 2-month-old. MISTRG pups were selected exclusively based on availability upon birth as xenotransplant was performed before their sex was distinguishable. Additionally, previous studies have not shown significant sex-related differences in immune reconstitution or hematopoietic function in this model (Song et al., Nat Commun 2019). |
| Field-collected samples | No field collected samples were used in the study.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| Ethics oversight        | All animal experiments were performed with the approval of MD Anderson's Institutional Animal Care and Use Committee under the protocol IACUC-0000931-RN03.                                                                                                                                                                                                                                                                                                                                                                                                               |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Flow Cytometry

### Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

|                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sample preparation        | Sample preparation details are extensively described in the "Methods" section of the paper. Briefly, BM aspirates were collected from research subjects following standard clinical procedures. BM mononuclear cells were isolated from each sample using the standard gradient separation approach with Ficoll-Paque PLUS (GE Healthcare Lifesciences, Pittsburgh, PA). For cell sorting, MNCs were pre-enriched using magnetic sorting with microbead kits for the specific antigens (Miltenyi Biotec, San Diego, CA) and further purified by fluorescence-activated cell sorting (FACS). Mononuclear cells or pre-enriched populations were washed with PBS+10% FBS and stained with the corresponding antibody cocktail. |
| Instrument                | BD Influx Cell Sorter (BD Biosciences); Cytex Aurora (Cytex Biosciences); BD Fortessa (BD Biosciences)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Software                  | BD FACSDiva, version 8.01 ( <a href="https://www.bdbiosciences.com">https://www.bdbiosciences.com</a> )                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Cell population abundance | Every cell sorting experiment was performed using a 2-step purification protocol (magnetic bead enrichment followed by FACS). The purity of the sorted cell fraction could not be determined in order to avoid increasing cell loss in very limited samples. However, our subsequent computational analysis did not detect cells with unexpected identities.                                                                                                                                                                                                                                                                                                                                                                 |

#### Gating strategy

Cytokine secretion and degranulation assays gating strategies were previously described in our prior publication (Montalban-Bravo et al., Cell Reports Medicine 2024). Our flow cytometry analysis of exhaustion markers uses previously validated gating strategy (BioLegend, <https://www.biolegend.com/fr-fr/flow-cytometry/optimized-panels/human-panels/hu-t-cell-differentiation-exhaustion-27c>). As for the new analysis of human BM cell engraftment after transplantation, the gate strategy is shown in Figure S4G.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.