Myelodysplastic syndromes are multiclonal diseases derived from hematopoietic stem and progenitor cells

Bingqing Luo1, Fang Dong1, Tiejun Qin2, Qingyun Zhang1, Haitao Bai1, Jinhong Wang1, Yujiao Jia1, Shihui Ma1, Erlie Jiang1, Tao Cheng1, Zhijian Xiao2 and Hideo Ema1*

Abstract
Myelodysplastic syndromes (MDS) are generally considered as a group of clonal diseases derived from hematopoietic stem cells, but a number of studies have suggested that they are derived from myeloid progenitor cells. We aimed to identify the cell of origin in MDS by single-cell analyses. Targeted single-cell RNA sequencing, covering six frequently mutated genes (U2AF1, SF3B1, TET2, ASXL1, TP53, and DNMT3A) in MDS, was developed and performed on individual cells isolated from the CD34+ and six lineage populations in the bone marrow of healthy donors (HDs) and patients with MDS. The detected mutations were used as clonal markers to define clones. By dissecting the distribution of clones in six lineages, the clonal origin was determined. We identified three mutations both in HDs and patients with MDS, termed clonal hematopoiesis (CH) mutations. We also identified fifteen mutations only detected in patients with MDS, termed MDS mutations. Clonal analysis showed that CH clones marked by CH mutations and MDS clones marked by MDS mutations were derived from hematopoietic stem cells as well as various hematopoietic progenitor cells. Most patients with MDS showed the chimeric state with CH clones and MDS clones. Clone size analysis suggested that CH mutations may not contribute to clonal expansion of MDS. In conclusion, MDS comprise multiple clones derived from hematopoietic stem and progenitor cells.

Keywords: Myelodysplastic syndromes (MDS), Clonal hematopoiesis (CH), Cell of origin, Hematopoietic stem cells (HSCs), Hematopoietic progenitor cells (HPCs), Single-cell analysis, Targeted single-cell RNA sequencing
Fig. 1 Comparison of mutations and clones between HD and MDS groups. A Data include all single cells analyzed by single-cell RNA-seq. Data of single cells in the HPCs (H), neutrophils (n), monocytes (m), erythroblasts (E), megakaryocytes (M), B cells (B), and T cells (T) populations from individual donors are shown in columns. Detected mutations are shown in rows. Mutations are shown in red. WT is shown as blank. When mutation sites had <10 reads, we could not determine whether mutations exist and data are shown in gray as NA (not available). B The percentage of mutant cells among the total single cells examined per donor. Data are presented as the medians with interquartile ranges. C Number of identified mutations per donor. Data are presented as the mean ± S.E. D The clonal analysis of PT2.7 is shown as an example. Eleven distinct clones were detected based on combinations of four mutations and some clones were detected in multiple populations. The upper panel shows the number of cells detected in a clone. Clones are identified with the combination of mutations. The lower panel shows the combinations of mutations detected in each clone. E Number of clones per donor in the HD group and MDS group. F Clonal diversity examined using the Shannon diversity index. Mann–Whitney U test was used in Fig. 1B, E, F, and two-sided Student’s t test was used in Fig. 1C. ns, p > 0.05; *, p < 0.05; **, p < 0.01; and ***, p < 0.001. D–F Wild type (WT) clones without mutations were excluded.
Single cells were sorted from CD34+ HPCs and six lineage populations: neutrophils, monocytes, erythroblasts, megakaryocytes, B cells, and T cells (Additional file 2: Fig. S1). These populations were then subjected to a targeted single-cell RNA-sequencing mutation analysis, covering U2AF1, SF3B1, TET2, ASXL1, TP53, and DNMT3A, the most frequently mutated genes in MDS [1, 2] (Additional file 2: Fig. S2A–C). After filtration, we obtained 3,549 single cells. A total of eighteen mutations were detected from all donors. All mutations were reported in COSMIC database. Three mutations detected in both the HD and MDS groups were termed CH mutations while the remaining fifteen mutations only detected in MDS group were termed MDS mutations (Fig. 1A, Additional file 1: Table S2), although the border between CH related mutations and MDS/AML related mutations is still blurry [8]. The fraction of cells carrying CH mutations was similar to that carrying MDS mutations (Fig. 1B). Greater mutation number in the MDS group than that in the HD group was observed (Fig. 1C).

Clones were defined as cells with identical combinations of mutations [9] (Fig. 1D). A total of eighty-nine clones were detected in all donors. CH clones were marked with CH mutations while MDS clones were marked with MDS mutations alone or both CH mutations and MDS mutations (Additional file 2: Fig. S4A). Greater clone number and clonal diversity in the MDS group than that in the HD group were observed (Fig. 1E, F), suggesting the accumulation of genetic lesions with more complexity during the progression to malignant diseases [10].

Most patients with MDS presented the mixture of multiple CH clones and MDS clones (Fig. 2A). The clone size of CH clones was significantly greater than that of MDS clones (Fig. 2B), suggesting the acquisition of CH mutations occurs earlier than that of MDS mutations [11]. No statistical difference in the clone size of CH clones between HD and MDS groups was observed (Fig. 2C). Both CH mutant clones and MDS mutant clones were larger in clone size than double mutant clones from MDS group, suggesting that CH mutations do not provide a significant growth advantage to MDS mutations (Fig. 2D).

CH clones and MDS clones were detected in a variety of lineage combinations (Fig. 2E). The number of cell populations involved in CH clones was significantly greater than that in MDS clones (Fig. 2F). When the same clones were detected in all six lineage populations regardless of H population, they were considered to originate from HSCs, otherwise from HPCs. 27% CH clones were derived from HSCs while 10% MDS clones were derived from HSCs (Fig. 2G). Of note, exact HPC types could not be defined by our strategy. 10 out of 37 HPC-derived MDS clones were found in MDS-EB1 patients and 19 out of 37 HPC-derived MDS clones was found in MDS-EB2 and sAML patients. MDS with higher blasts seems to have more chance to have HPC origin (Additional file 1: Table S3). But the percentage of blasts was not related to clone distribution (Additional file 2: Fig. S4C). A small number of single cells such as 18–26 cells per population were examined in previous single-cell analysis of MDS/AML [12]. In two MDS patients (PT2.13 and PT2.14), more single cells were analyzed, confirming the HSC and HPC origins for most MDS clones.

The number of HSC-derived clones might be underestimated, because a limited number of cells per population was analyzed. Nevertheless, if not all, MDS appeared to be the mixture of CH clones and MDS clones originating from HSCs and HPCs.
Fig. 2 (See legend on previous page.)
Abbreviations
MDS: Myelodysplastic syndromes; BM: Bone marrow; CH: Clonal hematopoiesis; sAML: Secondary acute myeloid leukemia; HD: Healthy donor; HSC: Hematopoietic stem cell; HPC: Hematopoietic progenitor cell; RT-PCR: Reverse transcription-polymerase chain reaction.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40164-022-02280-3.

Additional file 1: Table S1. BM donors. Table S2. Comparison of mutations detected by targeted bulk DNA sequencing and targeted single cell RNA sequencing. Table S3. The number of HSC- and HPC-derived clones in individual donors. Table S4. Primers. Table S5. Barcode, index and adaptor sequences. Table S6. Targeted genes in DNA sequencing.

Additional file 2: Figure S1. Gating strategy for seven cell populations. FACS strategy used to isolate single cells from the seven populations: H, CD34+CD45+ cells; E, CD34+CD5-CD235a+CD71+ cells; M, CD45+CD235a+CD41+ cells; m, CD45+CD33+CD14+CD16- cells; n, CD45-CD33+CD14-CD16- cells; B, CD45-CD19-CD20+ cells; and T, CD45-CD3+ cells. The upper panels represent HD data, and the lower panels represent data from a patient with MDS. Figure S2. Targeted single-cell RNA-seq method. A Schematic diagram of the targeted single-cell RNA-seq method. Single H, n, m, E, M, B, and T were sorted by flow cytometry into a 96-well plate. One-step RT-PCR was performed with the first-round primer mix. The second round of PCR was performed with PCR primer mix 1 or mix 2. Finally, a sequencing library was constructed with p5 and p7 adaptors. B Frequency at which the DNMT3A p.R301W mutation was detected in 32 single JURKAT cells. The positive rate was 78.1%. C Electrophoretogram of the representative libraries from single cells and 20 cells from a patient with MDS. One pair of each primer (U2AF1, SF3B1, ASXL1, TET2, DNMT3A) from the PCR primer mix (PCR primer mix 1 or PCR primer mix 2) was used for the second round of PCR, followed by library construction. PCR products were separated on a gel. Library construction performed without cells served as a negative control (neg).

Figure S3. Qualification of single cell RNA-seq data. A The results of single-cell filtration. The number of qualified and unqualified single cells is shown for individual donors. Only qualified cell data were used in the analysis. B Coverage of each mutation in single cells after filtration. The total number of reads covering each mutation site in qualified single cells is shown as the coverage. The threshold of coverage was 10x and is marked by a horizontal dotted line. Data are presented as the medians with interquartile ranges. Data points that fell outside of the upper and lower whiskers were considered outliers and are not shown. C Variant allele frequency (VAF) of mutations after filtration. Each dot represents the VAF of each mutation from single cells of donors carrying this mutation after filtration. Figure S4. The clone landscape in all donors. A Data include all clones identified. Eighteen identified mutations were used to detect eighty-nine clones from all donors. Each column represents a clone. CH clones were identified with CH mutations. MDS clones were identified with MDS mutations regardless of the coexistence of CH mutations. The lineage distributions of clones are shown in orange in the upper panel. The composition of mutations in each clone is shown in the lower panel. The presence of mutations is shown in red. WT is shown as blank. Mutation sites without enough reads to determine genotypes are shown in gray as not available (NA). The bottom panel shows the clone size. Wild type (WT) clones without mutations were excluded. Data of HD3 and HD5 are not shown since neither mutations nor clones were detected. B Clone size of CH clones and MDS clones from all donors. **, p < 0.001 (Mann–Whitney U test). C Clone size and number of populations involved in clones were compared between low blasts and high blasts groups. Fourteen patients with MDS were classified into low blasts (1%–9%) group (n = 8) mainly consisting of EB1 and high blasts (>9%) group (n = 6) mainly consisting of EB2. ns, p > 0.05 (Mann–Whitney U test).

Acknowledgments
The authors thank Toshio Suda for critical reading of the paper, and Peng Wu, Wanjun Lei, Haizhao Fan, and Jingyuan Tong for advice on bioinformatics analysis.

Author contributions
BL and HE designed the study. BL, FD, QZ, HB, and JW performed experiments and analyzed data. YJ analyzed targeted bulk DNA sequencing data. TQ and ZX provided patient materials and clinical data. SM and EJ helped to obtain healthy donor materials. TC, ZX, and HE supervised the study. BL and HE wrote the paper. All the authors have read and approved the final manuscript.

Funding
This work was supported by grants from the National Key R&D Program of China (2019YFA0110203 and 2021YFA1100008), the National Natural Science Foundation of China (81970119, 82070112, 81800990, and 82170118), the Haihe Laboratory of Cell Ecosystem Innovation Fund (HH22KYZIX002), the PUMC Project for Building Subjects (201920101401), and the CAMS Initiative Fund for Innovative Medicine (2021-2QM-1-073 and 2021-2QM-1-040).

Availability of data and materials
Supplementary information including Additional file 2: Figs. S1–S4 and Additional file 1: Tables S1–S6 are provided with the online version of this paper. The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Declarations
Ethics approval and consent to participate
Human samples were approved by the Ethics Committee, Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College and met all requirements of the Declaration of Helsinki.

Consent for publication
All authors provide their consent for publication.

Competing interests
The authors have no conflicts of interest to report.

Author details
1 State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haie Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College and met all requirements of the Declaration of Helsinki.

Received: 18 February 2022 Accepted: 16 April 2022
Published online: 16 May 2022

References
1. Pappaemmanuil E, Gerstung M, Malcovati L, Tauris S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013;122(22):3616–27; quiz 3699.
2. Li B, Liu J, Ya J, Wang X, Xu Z, Qin T, et al. Clinical features and biological implications of different U2AF1 mutation types in myelodysplastic syndromes. Genes Chromosomes Cancer. 2018;57(2):80–88.
3. Woll PS, Kallquist U, Chowdhury O, Doolittle H, Wedge DC, Thonguea S, et al. Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. Cancer Cell. 2014;25(6):794–808.
4. Rouault-Pierre K, Smith AE, Mian SA, Pizzitola I, Kulasekararaj AG, Mufti GJ, et al. Myelodysplastic syndrome can propagate from the multipotent progenitor compartment. Haematologica. 2017;102(1):e7-e10.
5. van Kamp H, Fibe WE, Jansen RP, van der Keur M, de Graaff E, Willemsen R, et al. Clonal involvement of granulocytes and monocytes, but not of T and B lymphocytes and natural killer cells in patients with myelodysplasia: analysis by X-linked restriction fragment length polymorphisms and
polymerase chain reaction of the phosphoglycerate kinase gene. Blood. 1992;80(7):1774–80.
6. Lee-Six H, Obro NF, Shepherd MS, Grossmann S, Dawson K, Belmonte M, et al. Population dynamics of normal human blood inferred from somatic mutations. Nature. 2018;561(7724):473–478.
7. Scala S, Aiuti A. In vivo dynamics of human hematopoietic stem cells: novel concepts and future directions. Blood Adv. 2019;3(12):1916–1924.
8. Nagata Y, Makishima H, Kerr CM, Przychodzen BP, Aly M, Goyal A, et al. Invariant patterns of clonal succession determine specific clinical features of myelodysplastic syndromes. Nat Commun. 2019;10(1):3386.
9. Scala S, Aiuti A. In vivo dynamics of human hematopoietic stem cells: novel concepts and future directions. Blood Adv. 2019;3(12):1916–1924.
10. Nagata Y, Makishima H, Kerr CM, Przychodzen BP, Aly M, Goyal A, et al. Invariant patterns of clonal succession determine specific clinical features of myelodysplastic syndromes. Nat Commun. 2019;10(1):3386.
11. Watson CI, Papula AL, Poon GY, Wong WH, Young AL, Druley TE, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. Science. 2020;367(6485):1449–1454.
12. Chen J, Kao YR, Sun D, Todorova TI, Reynolds D, Narayangari SR, et al. Myelodysplastic syndrome progression to acute myeloid leukemia at the stem cell level. Nat Med. 2019;25(1):103–110.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:
• fast, convenient online submission
• thorough peer review by experienced researchers in your field
• rapid publication on acceptance
• support for research data, including large and complex data types
• gold Open Access which fosters wider collaboration and increased citations
• maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.
Learn more biomedcentral.com/submissions