

掲載予定雑誌名 : Leukemia

Genetic analysis of myeloid neoplasms with der(1;7)(q10;p10)

Rurika Okuda¹, Yotaro Ochi^{1*}, Ryunosuke Saiki¹, Toshiyuki Yamanaka¹, Chikashi Terao^{2,3,4}, Tetsuichi Yoshizato¹, Masahiro M. Nakagawa¹, Zhao Lanying¹, Kazuma Ohyashiki⁵, Nobuhiro Hiramoto⁶, Masashi Sanada^{1,7}, Hiroshi Handa⁸, Senji Kasahara⁹, Yasushi Miyazaki¹⁰, Nobuo Sezaki¹¹, Lee-Yung Shih¹², Wolfgang Kern¹³, Nobuhiro Kanemura¹⁴, Toshiyuki, Kitano¹⁵, Shinsaku Imashuku¹⁶, Mitsumasa Watanabe¹⁷, Maria Creignou^{18,19}, Kazuhisa Chonabayashi^{20,21}, Kensuke Usuki²², Takayuki Ishikawa⁶, Akihiko Gotoh⁵, Yoshiko Atsuta^{23,24}, Yuichi Shiraishi²⁵, Kinuko Mitani²⁶, Shigeru Chiba²⁷, Akifumi Takaori-Kondo²¹, Satoru Miyano^{25,28}, Yoichiro Kamatani²⁹, Torsten Haferlach¹³, Eva Hellstrom-Lindberg^{18,30}, Koichi Matsuda²⁹, Yoshinori Yoshida²⁰, Hideki Makishima¹, Yasuhito Nannya¹, Seishi Ogawa^{1,31*}

¹Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; ²Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; ³Clinical Research Center, Shizuoka General Hospital, Shizuoka, Japan; ⁴The Department of Applied Genetics, The School of Pharmaceutical Sciences, University of Shizuoka, Shizuoka, Japan; ⁵Department of Hematology, Tokyo Medical University, Tokyo, Japan; ⁶Department of Hematology, Kobe City Medical Center General Hospital, Hyogo, Japan; ⁷Department of Advanced Diagnosis, Clinical Research Center, NHO Nagoya Medical Center, Nagoya, Japan; ⁸Department of Hematology, Graduate School of Medicine, Gunma University, Gunma, Japan; ⁹Department of Hematology, Gifu Municipal Hospital, Gifu, Japan; ¹⁰Japan Adult Leukemia Study Group, and Department of Hematology, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki, Japan; ¹¹Department of Hematology, Chugoku Central Hospital, Hiroshima, Japan; ¹²Division of Hematology-Oncology, Chang Gung Memorial Hospital-Linkou, Chang Gung University, Taoyuan, Taiwan; ¹³MLL Munich Leukemia Laboratory, Munich, Germany; ¹⁴Department of Hematology and Infectious Disease, Gifu University Graduate School of Medicine, Gifu, Japan; ¹⁵Department of Hematology, Tazuke Kofukai, Medical Research Institute, Kitano Hospital, Osaka, Japan; ¹⁶Department of Laboratory Medicine, Uji-Tokushukai Medical Center, Uji, Japan; ¹⁷Department of Hematology, Hyogo Prefectural Amagasaki General Medical Center, Hyogo, Japan; ¹⁸Department of Medicine Huddinge, Center for Hematology and Regenerative Medicine, Karolinska Institutet, Stockholm, Sweden; ¹⁹Center for Clinical Cancer Studies, Phase 1 unit, Karolinska University Hospital, Stockholm, Sweden; ²⁰Department of Cell Growth and Differentiation, Center for iPS Research and Application, Kyoto University, Kyoto, Japan; ²¹Department of Hematology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; ²²Department of Hematology, NTT Medical Center Tokyo, Tokyo, Japan; ²³The Japanese Data Center for Hematopoietic Cell Transplantation, Aichi, Japan; ²⁴Department of Registry Science for Transplant and Cellular Therapy, Aichi Medical University School of Medicine, Aichi, Japan; ²⁵Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; ²⁶Department of Hematology and Oncology, Dokkyo Medical University, Tochigi, Japan; ²⁷Department of Hematology, Institute of Medicine, University of Tsukuba, Ibaraki, Japan; ²⁸M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan; ²⁹Department of Computational Biology and Medical Sciences, Graduate Schools of Frontier Sciences, The University of Tokyo, Tokyo, Japan; ³⁰Department of Hematology, Karolinska University Hospital, Stockholm, Sweden; ³¹Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto, Japan

*Contact information for corresponding authors:

Seishi Ogawa (sogawa-tky@umin.ac.jp) and Yotaro Ochi (ochi.yotaro.24v@kyoto-u.jp).

S.O. and Y.O.: Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto

University, F-building, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan

TEL: +81-75-753-9285, FAX: +81-75-753-9282

Word Count: 1422

To the editor:

der(1;7)(q10;p10) is an unbalanced translocation recurrently found in a variety of myeloid neoplasms, where +1q and -7q are the common consequences¹⁻⁶. Although it represents one of the most frequent chromosomal abnormalities in myeloid neoplasms among Asian populations, the molecular characteristics of der(1;7)(q10;p10)(+) myeloid neoplasms have not been fully elucidated. In this retrospective study, we enrolled 3,385 myeloid neoplasm cases to identify a total of 148 cases with der(1;7)(q10;p10) on the basis of conventional cytogenetics and/or sequencing-based copy number analysis⁷ and investigate their clinical features and mutational profiles in comparison with those cases having -7/del(7q) (n=376) and 1q trisomy (+1q) (n=54) alone, using whole-exome sequencing (WES) and/or targeted-capture sequencing. The remaining 2,808 cases were collectively analyzed as 'OTHER' cases (**Supplementary Methods and Supplementary Table 1**). Through these analyses, we demonstrated that der(1;7)(q10;p10)(+) myeloid neoplasms were characterized by unique clinical and mutational features and therefore, represented a distinct subset of myeloid neoplasms.

In accordance with previous reports^{1,4}, there was an extreme male predominance in der(1;7)(q10;p10)(+) cases (87.8%)($P<0.001$) (**Supplementary Table 1**). der(1;7)(q10;p10) was significantly more prevalent in Asian (54/936) than German (4/944) MDS cases (5.8% vs. 0.4%) (Data not shown). Compared to -7/del(7q) and +1q, der(1;7)(q10;p10) was more enriched (72.3%) for MDS than for AML (23.6%) and MDS/myeloproliferative neoplasm (MPN) (3.4%). MDS with excess blasts (MDS-EB) was less common in der(1;7)(q10;p10)(+) MDS cases than in -7/del(7q)(+) MDS patients, but more frequent compared to +1q cases. Secondary AML was more enriched in der(1;7)(q10;p10)(+) AML cases compared to non-der(1;7)(q10;p10) cases (**Supplementary Fig. 1A**). der(1;7)(q10;p10)(+) MDS cases had significantly higher hemoglobin levels, higher platelet counts, and lower blast counts than -7/del(7q) MDS cases (**Supplementary Fig. 1B**).

der(1;7)(q10;p10)(+) MDS, together with -7/del(7q)(+) MDS, had a significantly shorter overall survival and faster leukemic transformation than OTHER MDS cases, although in accordance with our previous report,¹ der(1;7)(q10;p10)(+) MDS cases tended to have a slightly better OS and slower leukemic progression than -7/del(7q) MDS cases (**Supplementary Fig. 1C**). By contrast, +1q MDS cases showed a trend for better OS. Compared to OTHER MDS cases, der(1;7)(q10;p10)(+) MDS cases tended to die without AML progression. Of note, infection-related deaths explained as many as 45% of non-leukemic deaths among der(1;7)(q10;p10)(+) MDS patients, whereas it explained only 13.9% and 10.8% for -7/del(7q) and OTHER MDS cases, respectively (**Supplementary Fig. 1D**). Interestingly, +1q MDS patients also showed a similar trend with 60% of cases dying from infection-related causes.

We next conducted targeted-capture sequencing for known driver genes and gene mutations identified through WES of 26 der(1;7)(q10;p10) cases and revealed that 132 of 148 der(1;7)(q10;p10)(+) cases (89.2%) had one or more gene mutations (**Fig. 1A**). The mean number of mutations was significantly higher in der(1;7)(q10;p10)(+) cases (2.8/sample) compared to other sub-groups (range, 1.7-1.9)($P<0.001$) (**Supplementary Fig. 1E**). Most frequently mutated genes in der(1;7)(q10;p10)(+) cases included *RUNX1*, *EZH2*, *ETNK1*, *U2AF1*, *DNMT3A*, *BCOR*, *ETV6*, *TET2*, *GATA2*, *MYB*, *IDH1*, *PHF6*, and *ASXL1*, the majority of which were more frequently mutated in der(1;7)(q10;p10)(+) cases than in other sub-groups (**Fig. 1B and Supplementary Fig. 1F**). Accounting for half of these genes, transcription factor genes, including *RUNX1*,

BCOR, *ETV6*, *GATA2*, *MYB*, and *CEBPA*, represented major mutational targets in der(1;7)(q10;p10)(+) cases (**Fig. 1C**). Interestingly, we observed multiple mutations affecting single genes for some of these transcription factor genes, including *RUNX1*, *ETV6*, *MYB*, and *GATA2*, suggesting a strong selective pressure that favors these mutations within the der(1;7)(q10;p10)(+) populations (**Fig. 1A**). Overall, 62.2% of der(1;7)(q10;p10)(+) cases harbored at least one transcription factor mutation, a much higher frequency than der(1;7)(q10;p10)(-) cases (**Fig. 1C**). +8 and del(20q) were common in der(1;7)(q10;p10)(+) cases. By contrast, *TP53* mutations and del(5q) were rarely seen in der(1;7)(q10;p10)(+) cases, showing a sharp contrast to the very high frequencies of these alterations in -7/del(7q), +1q and OTHER cases (**Supplementary Fig. 1F**).

Among frequent mutational targets of der(1;7)(q10;p10), two genes, *MYB* and *ETNK1*, were rarely mutated in OTHER and -7/del(7q) sub-groups, and therefore, are highly characteristic of der(1;7)(q10;p10)(+) cases. Interestingly, *MYB* mutations were also common in +1q cases (**Fig. 1B**). *MYB* is a proto-oncogene originally identified as the v-Myb oncogene within the genome of avian myeloblastosis virus and E26 virus⁸. Most variants were protein-truncating, and widely distributed along the entire coding sequence and therefore, lead to a loss of function (**Fig. 2A**). *MYB* mutations were found in 13.1% of MDS and 11.4% of AML cases with der(1;7)(q10;p10)(+), but rarely detected in der(1;7)(q10;p10)(-) cases (**Fig. 2B**). There were no significant differences in complete blood counts between patients with and without *MYB* mutations (**Supplementary Fig. 2A**). Mutation profiles did not substantially differ between *MYB*-mutated (mut) and *MYB*-wild type (wt) cases. *MYB*-mut cases showed a shorter OS than *MYB*-wt cases, although not statistically significant. *ETNK1* was another gene uniquely mutated in der(1;7)(q10;p10)(+) cases. We found frequent *ETNK1* mutations in 18% of der(1;7)(q10;p10)(+) cases, which showed a prominent mutational hotspot affecting N244. Mutations were highly specific to MDS cases (24.3%) with only 2.9% in AML cases, which was comparable to der(1;7)(q10;p10)(-) MDS (2.6%) and MDS/MPN (3.3%) cases (**Fig. 2A-B**). As was the case with atypical chronic myeloid leukemia⁹, *ETNK1* mutations significantly co-occurred with *SETBP1* mutations in der(1;7)(q10;p10)(+) myeloid neoplasms ($P=0.042$) (**Fig. 2C**). By contrast, +8 tended to be mutually exclusive with *ETNK1* mutations ($P=0.0078$) (**Fig. 2C**). Notably, *ETNK1*-mut der(1;7)(q10;p10)(+) cases presented with eosinophilia (mean: 6.2% vs. 14.8%) ($P=0.025$) (**Supplementary Fig. 2A**). *ETNK1*-mutated MDS cases showed a poorer prognosis than *ETNK1*-wt cases.

Next, we examined gene expression profiles using RNA-sequencing of ten der(1;7)(q10;p10)(+), 20 -7/del(7q), and 62 OTHER MDS patients. As expected, we confirmed the effect of allelic dosage: elevated and decreased expression across the genes on 1q and 7q, respectively, while the effect was highly variable for individual genes. For example, we noted a consistent reduction of gene expression of *EZH2* and *CUX1*, which are two major tumor suppressors and frequent targets of mutations and focal deletions on 7q in a variety of myeloid neoplasms, while elevated expression was observed for several oncogenes on 1q, such as *AKT3*, *BCL9*, *NCSTN*, *LAMC1*, *MDM4*, and *RIT1* (**Supplementary Fig. 2B**). Thus, deregulation of these tumor suppressors and oncogenes could explain the pathogenesis of der(1;7)(q10;p10)(+) MDS. Of particular interest among these is *MDM4* because *MDM4* is a negative regulator of p53 functions and the lack of *TP53* mutation was another unique feature in der(1;7)(q10;p10)(+) cases. Elevated *MDM4* expression caused by 1q gain has been implicated in the paucity of *TP53* mutations in +1q cases in other tumors^{10–12}. In line with these reports, down-regulated expression of p53 signaling pathway genes observed in *TP53*-mutated MDS with and without -7/del(7q), was also demonstrated in der(1;7)(q10;p10)(+) cases (**Fig. 2D**). *TP53* mutations were also frequent in cases

with +1q alone, however, *TP53*-mutations tended to show a larger mutant cell fraction than that of +1q, suggesting that the +1q clone evolved from within the pre-existing *TP53*-mutant clone (Data not shown), implying a distinct pathogenesis between der(1;7)(q10;p10) and +1q alone.

Finally, to investigate the evolution of der(1;7)(q10;p10)(+) clones, we inferred the order of acquisition of genetic alterations by evaluating their variant allele frequency. der(1;7)(q10;p10), together with del(20q) and mutations in *DNMT3A* and *ETNK1*, represented the major clone in most cases, suggesting their early origin during clonal evolution. By contrast, characteristic transcription factor gene mutations in der(1;7)(q10;p10), such as *ETV6*, *GATA2*, and *MYB*, were relatively late events (**Supplementary Fig. 2C-D**). Furthermore, der(1;7)(q10;p10) was identified as clonal mosaicism in healthy individuals before they developed AML/MDS. When we surveyed CNAs among 179,417 healthy individuals from the BioBank Japan based on SNP array data^{13,14}, we found 29 individuals harboring concomitant gain of 1q and loss of 7q suggestive of der(1;7)(q10;p10), of which 4 later died of AML or MDS (**Supplementary Fig. 2E**). Another interesting finding was the identification of *ETNK1* hotspot mutations in only individuals with der(1;7)(q10;p10) when we analyzed mutations by duplex-sequencing in 64 individuals including 6 with der(1;7)(q10;p10) (**Supplementary Fig. 2F**). Droplet Digital PCR (ddPCR) of 146 BBJ samples also identified these hotspot mutations in only those with der(1;7)(q10;p10) (Data not shown). Given the strong and unique association between *ETNK1* mutations and der(1;7)(q10;p10), *ETNK1* mutations may be acquired during early evolution of der(1;7)(q10;p10)(+) myeloid neoplasms as clonal hematopoiesis.

In conclusion, der(1;7)(q10;p10)(+) myeloid neoplasms are characterized by a number of unique genetic features and distinct hematological profiles and are therefore considered to represent a distinct entity of myeloid neoplasms, as supported by the recent findings from the IWG cohort¹⁵. Identification of this subset of myeloid neoplasms may allow for better prognostication and treatment for these patients. An elevated MDM4 expression associated with 1q gain indicates a possible role of MDM4 inhibitor for these neoplasms which needs to be tested in the clinical setting in the future.

References

- 1 Sanada M, Uike N, Ohyashiki K, Ozawa K, Lili W, Hangaishi A *et al.* Unbalanced translocation der(1;7)(q10;p10) defines a unique clinicopathological subgroup of myeloid neoplasms. *Leukemia* 2007; **21**: 992–997.
- 2 Ganster C, Müller-Thomas C, Haferlach C, Strupp C, Ogata K, Germing U *et al.* Comprehensive analysis of isolated der(1;7)(q10;p10) in a large international homogenous cohort of patients with myelodysplastic syndromes. *Genes Chromosomes Cancer* 2019; **58**: 689–697.
- 3 Wang L, Ogawa S, Hangaishi A, Qiao Y, Hosoya N, Nanya Y *et al.* Molecular characterization of the recurrent unbalanced translocation der(1;7)(q10;p10). *Blood* 2003; **102**: 2597–2604.
- 4 Zhang T, Xu Y, Pan J, Li H, Wang Q, Wen L *et al.* High frequency of RUNX1 mutation in myelodysplastic syndrome patients with whole-arm translocation of der(1;7)(q10;p10). *Leukemia* 2017; **31**: 2257–2260.
- 5 Mauritzson N, Albin M, Rylander L, Billström R, Ahlgren T, Mikoczy Z *et al.* Pooled analysis of clinical and cytogenetic features in treatment-related and de novo adult acute myeloid leukemia and myelodysplastic syndromes based on a consecutive series of 761 patients analyzed 1976-1993 and on 5098 unselected cases reported in the literature 1974-2001. *Leukemia* 2002; **16**: 2366–2378.
- 6 Horiike S, Taniwaki M, Misawa S, Nishigaki H, Okuda T, Yokota S *et al.* The Unbalanced 1;7 Translocation in De Novo Myelodysplastic Syndrome and Its Clinical Implication Chromosome Studies. *Cancer* 1989; **65**: 1350–1354.
- 7 Yoshizato T, Nannya Y, Atsuta Y, Shiozawa Y, Iijima-Yamashita Y, Yoshida K *et al.* Genetic abnormalities in myelodysplasia and secondary acute myeloid leukemia: impact on outcome of stem cell transplantation. *Blood* 2017; **129**: 2347–2358.
- 8 Frampton J, Ramqvist T, Graf T. v-Myb of E26 leukemia virus up-regulates bcl-2 and suppresses apoptosis in myeloid cells. *Genes Dev* 1996; **10**: 2720–2731.
- 9 Gambacorti-Passerini CB, Donadoni C, Parmiani A, Pirola A, Redaelli S, Signore G *et al.* Recurrent ETNK1 mutations in atypical chronic myeloid leukemia. *Blood* 2015; **125**: 499–503.
- 10 Sebert M, Gachet S, Leblanc T, Rousseau A, Bluteau O, Kim R *et al.* Clonal hematopoiesis driven by chromosome 1q/MDM4 trisomy defines a canonical route toward leukemia in Fanconi anemia. *Cell Stem Cell* 2023; **30**: 153–170.e9.
- 11 Girish V, Lakhani AA, Thompson SL, Scaduto CM, Brown LM, Hagenson RA *et al.* Oncogene-like addiction to aneuploidy in human cancers. *Science* 2023; **381**: 848.
- 12 Hülle J, Słabicki M, Rosolowski M, Jethwa A, Habringer S, Tomska K *et al.* MDM4 Is Targeted by 1q Gain and Drives Disease in Burkitt Lymphoma. *Cancer Res* 2019; **79**: 3125–3138.
- 13 Terao C, Suzuki A, Momozawa Y, Akiyama M, Ishigaki K, Yamamoto K *et al.* Chromosomal alterations among age-related haematopoietic clones in Japan. *Nature* 2020; **584**: 130–135.
- 14 Saiki R, Momozawa Y, Nannya Y, Nakagawa MM, Ochi Y, Yoshizato T *et al.* Combined landscape of

single-nucleotide variants and copy number alterations in clonal hematopoiesis. *Nat Med* 2021; **27**: 1239–1249.

- 15 Bernard E, Hasserjian RP, Greenberg PL, Arango Ossa JE, Creignou M, Tuechler H *et al.* Molecular taxonomy of myelodysplastic syndromes and its clinical implications. *Blood* 2024; : blood.2023023727.

Figure Captions

Figure 1. Genetic characteristics of der(1;7)(q10;p10)(+) myeloid neoplasms. (A) Landscape of gene mutations and copy number alterations in 148 der(1;7)(q10;p10)(+) myeloid neoplasm cases identified through targeted-capture sequencing. Number of genetic mutations per case shown as bar plots at top. Genetic mutation type and disease depicted as colors. Frequency of each mutation (>1%) and number of patients with the mutation are shown in left and right, respectively. Rows are ordered by functional categories of affected genes. (B) Frequency bar plot of targeted-sequencing gene mutation and copy number alteration for der(1;7)(q10;p10)(+) vs. -7/del(7q) vs. +1q vs others. Type of mutation depicted by differing colors. False discovery rate (q-values) shown by asterisks (*). Transcription factor genes in red. (C) Bar graph of the frequency of the number of transcription factor mutations per case for each group (P-values calculated by Fisher's Exact test).

Figure 2. Characteristic mutations in der(1;7)(q10;p10)(+) myeloid neoplasms. (A) Lollipop plot of *MYB* and *ETNK1* gene mutations in der(1;7)(q10;p10) cases. Domains for each gene and type of mutations are depicted. (B) Frequency of *MYB* (left) and *ETNK1* mutations (right) in AML, MDS, MDS/MPN and MPN for der(1;7)(q10;p10) and non-der(1;7)(q10;p10) cases. (C) Presence of indicated mutations according to *ETNK1* mutation status. RAS/RTK genes: *FLT3*, *JAK2*, *MPL*, *CALR*, *CSF3R*, *PTPN11*, *NF1*, *NRAS*, *KRAS*, and *CBL*. (D) Boxplot of *MDM4* gene expression (left) and expression level of p53 pathway genes (right) for each group. The median and 1st and 3rd quartiles are indicated, and whiskers extend to the furthest value within 1.5× the interquartile range. P-values calculated by Wilcoxon test. **P*<0.05, ***P*<0.01

Acknowledgments

We thank Maki Nakamura, Takeshi Shirahari, Kazuhide Onishi, Ai Takatsu, and Natsumi Sakamoto (Department of Pathology and Tumor Biology, Kyoto University) for technical assistance. We thank BioBank Japan for providing patients samples and clinical data. We used the questionnaire information of the Vital Statistics of Japan. This work was supported by grants from the Japan Agency for Medical Research and Development (AMED) (no. 19ck0106470h0001 to H. Makishima; and nos. JP15cm0106056h0005, JP19cm0106501h0004, JP16ck0106073h0003, and JP19ck0106250h0003 to S. Ogawa); the Ministry of Education, Culture, Sports, Science and Technology of Japan; the High Performance Computing Infrastructure System Research Project (nos. hp160219, hp170227, hp180198, hp190158, hp200138, and hp210167 to S. Ogawa) (This research used computational resources of the K computer provided by the RIKEN Advanced Institute for Computational Science through the HPCI System Research project); the Japan Society for the Promotion of Science; Scientific Research on Innovative Areas

(no. JP15H05909 to S. Ogawa); KAKENHI (nos. JP26221308 and JP19H05656 to S. Ogawa, JP22K16320 to Y. Ochi, JP19H01053 to H. Makishima, and JP21J23598 to R. Okuda); the Japan Science and Technology Agency (JST) Fusion Oriented REsearch for disruptive Science and Technology (FOREST) Program (grant no. JPMJFR220L to Y. Ochi); and the Takeda Science Foundation to S. Ogawa. S. Ogawa is a recipient of the JSPS Core-to-Core Program A: Advanced Research Networks. This study was supported by the Moonshot Research and Development Program (grant no. JP22zf0127009) from the Japan Agency for Medical Research and Development.

Authorship Contributions

R.O., Y.O., and S.O. designed the study, reviewed data analysis, and wrote the manuscript; R.O. curated clinical datasets, and performed overall bioinformatic analysis; R.O., Y.O., S.M., K.C., Y.Y., H.M., and Y.N. performed statistical analysis; R.S., C.T., T.Yamanaka, and T.Yoshizato performed copy number analysis; M.N., Z.L., K.O., N.H., M.S., H.H., S.K., Y.M., N.S., L.S., W.K., N.K., T.K., S.I., M.W., M.C., K.U., T.I., A.G., Y.A., Y.S., K.Mitani, S.C., A.T., Y.K., T.H., E.H.L., and K.Matsuda. prepared patient samples. All authors reviewed the manuscript during its preparation.

Disclosure of Conflicts of Interest

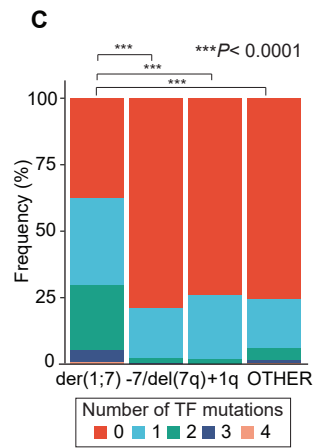
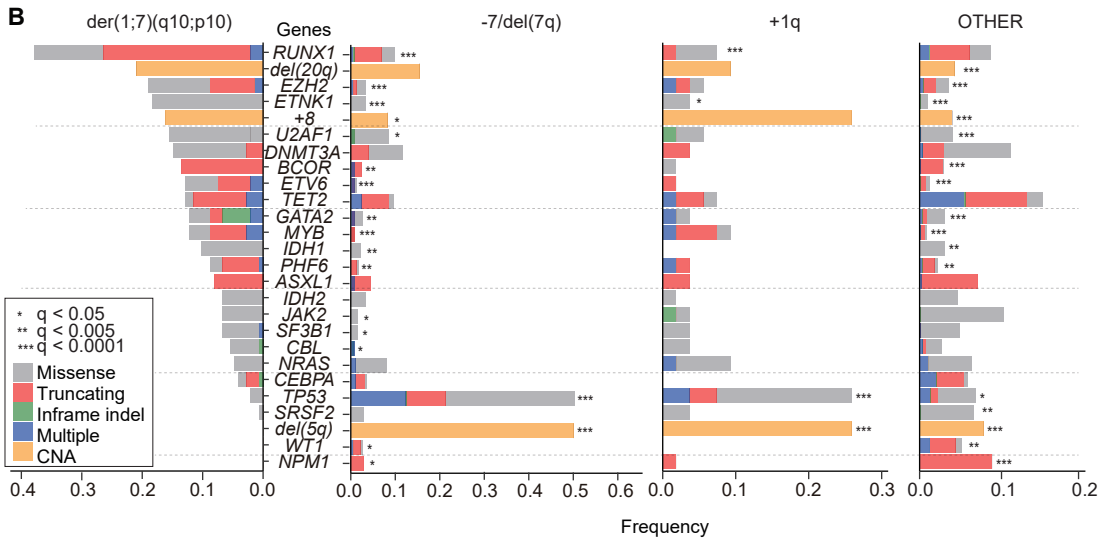
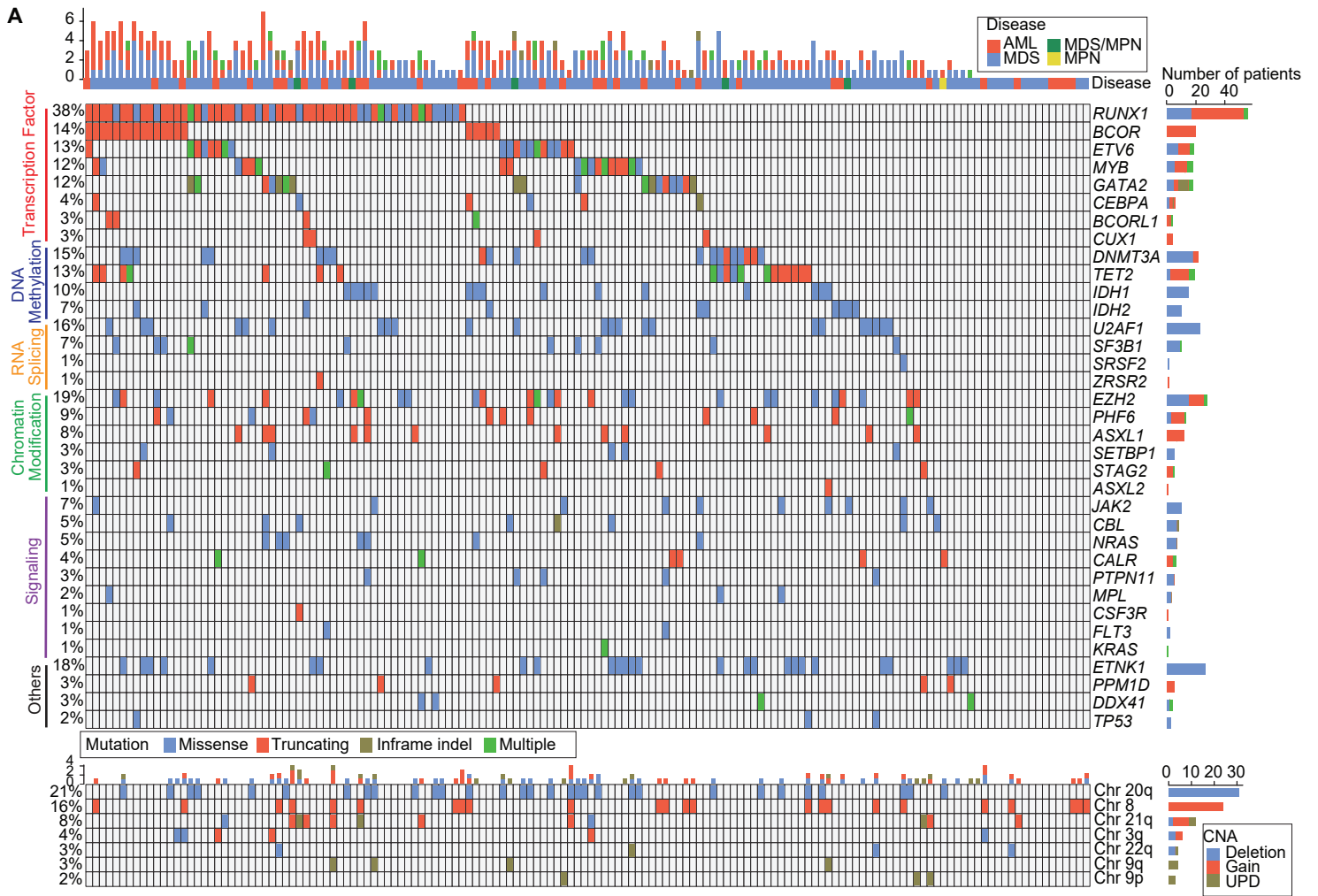
Authors declare nothing to disclose.

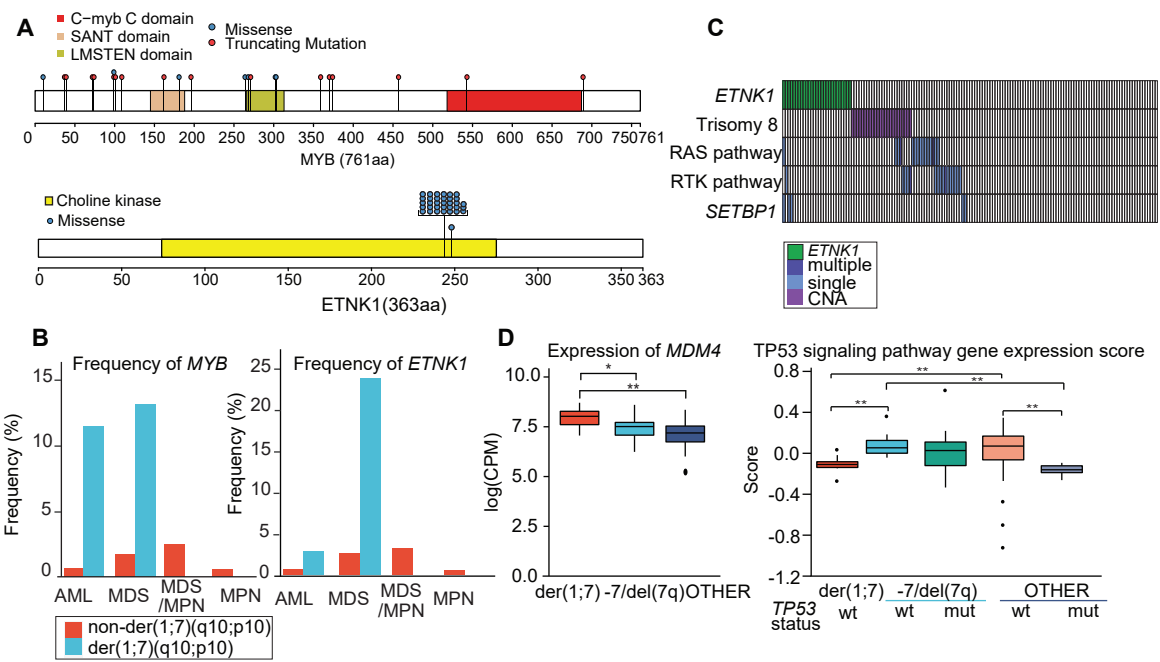
Ethics approval and consent to participate

The authors confirm that this study was approved by the institutional ethics committees of Kyoto University and participating institutions (G608) and was performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from each participant.

Data Sharing Statement

Datasets of WES and RNA-seq data are available in the European Genome-phenome Archive database (Accession ID: EGAS50000000704 and EGAS50000000705).





Supplementary Methods

Patient samples

We enrolled 148 der(1;7)(q10;p10) harboring myeloid neoplasm cases and 3,238 non-der(1;7)(q10;p10) myeloid neoplasm cases from collaborating institutes in this study (Supplementary Table 1). An additional 944 MDS cases from the German Munich Leukemia Laboratory was also included for the inter-racial comparison¹. Whole bone marrow (WBM) samples, peripheral blood (PB) samples, and oral mucosa swabs were collected along with written informed consent in accordance with the Declaration of Helsinki and appropriate Ethics Committee approvals from patients from the following institutions: Kyoto University, Japan Marrow Donor Program, Tokyo Medical University, Kobe City Medical Center General Hospital, Gunma University, Gifu Municipal Hospital, Japan Adult Leukemia Study Group, Chang Gung University, Chugoku Central Hospital, University of Tsukuba, Gifu University Graduate School of Medicine, Kitano Hospital, Uji-Tokushukai Medical Center, Hyogo Prefectural Amagasaki General Medical Center, Dokkyo University, Karolinska Institute, NTT Medical Center Tokyo, Tokyo Metropolitan Cancer and Infectious Diseases Center Komagome Hospital, National Hospital Organization Nagoya Medical Center, Kitano Hospital.

Determination of der(1;7)(q10;p10) cases were conducted by utilizing sequenced-based copy number analysis of targeted-capture sequencing².

For an inter-racial incidence comparison, only Asian cases sequenced in consecutive cases were included to remove potential selection bias.

Whole-exome sequencing

For whole-exome sequencing, SureSelect Human All Exon v6 kits (Agilent Technologies) or xGen™ Exome Hyb Panel v2 were used to prepare 50ng of above patient DNA. Using Brrows-Wheeler Aligner, the resulting sequencing reads were aligned to the human genome reference (GRCh37/hg19)³. Mutations were called using GenomonPipeline⁴, adopting parameters used in previous studies².

Targeted-capture sequencing

For targeted-capture sequencing, 200ng of above patient DNA were used for hybridization-based capture via SureSelect custom kit (Agilent Technologies) and sequenced on HiSeq 2500, NovaSeq6000 (Illumina) or MGI sequencer (MGI). Using Brrows-Wheeler Aligner, the resulting sequencing reads were aligned to

the human genome reference (GRCh37/hg19)³. Mutations were called using GenomonPipeline⁴, adopting parameters used in previous studies².

RNA-sequencing

For RNA-sequencing, RNA was extracted from ten der(1;7)(q10;p10)(+), 20 -7/del(7q), and 62 OTHER MDS patient BM samples, and CD34+ cell selection was conducted using CD34 MicroBead Kit (Miltenyi Biotec) according to the manufacturer's instructions. Libraries were prepared using NEBNext Ultra II RNA Library Prep Kit for Illumina (New England BioLabs) and sequenced with NovaSeq 6000 instrument (Illumina). The reads were mapped to a reference genome (hg19) using STAR(v2.5.3)⁵, counted using featureCounts⁶, and differentially expressed genes were identified by edgeR⁷ at FDR<0.05. 'Score' for TP53 pathways were calculated using the log₂(CPM) values for genes listed in the 'HALLMARK_P53_PATHWAY' from the Molecular Signatures Database (MSigDB)⁸: and using the R function 'scale' to scale the log₂(CPM) values.

Duplex-sequencing and Digital droplet PCR (ddPCR)

Sequencing was conducted upon 54 samples from healthy individuals in the BBJ cohort using xGen Prism prep kit (IDT), following the manufacturer's instructions. For target enrichment, we designed a custom IDT xGen bait to cover 29 genes commonly mutated in der(1;7)(q10;p10)(+) myeloid neoplasms and genes relating to clonal hematopoiesis (CH). Sequencing was performed on DNBSEQ-G400RS (MGI Tech) in 100-150 bp paired-end mode, according to the manufacturer's instructions. Original sequence reads were mapped by BWA (<https://github.com/lh3/bwa>), followed by conversion into Bam via Picard (<https://broadinstitute.github.io/picard/>). Duplex consensus sequences were extracted and processed according to the manufacturer's instructions. Sequences were collapsed into families by fgbio CallDuplexConsensusReads and mutations were called using Genomon Pipeline v2.6.2⁴. Copy number alterations (CNAs) on chromosomes 1q, 2, 7q, 8, and 20q were evaluated by B-allele frequencies (BAF) of heterozygous SNPs calculated by GenomonMutationFilter (<https://github.com/Genomon-Project/GenomonMutationFilter>) and normalization with the average BAF of SNPs on chr2, which were not affected by CNAs in analyzed cases. For ddPCR, predesigned *ETNK1* p.N244S c.731A>G probes (Catalog # 10049047) were purchased from Bio-Rad. Droplet generation was conducted on the QX200 Droplet Generator for each DNA/enzyme mix (50ng of gDNA + ddPCR Supermix for Probes (no dUTP), Bio-Rad + HaeIII). PCR annealing temperatures were set at 55°C. Droplets were analyzed using QX200 system and QuantaSoft1.7 (Bio-Rad) according to the manufacturer's instructions.

Definitions for gene groups

Genes were categorized as follows according to the gene function⁹.

Transcription factor: *RUNX1*, *BCOR*, *ETV6*, *GATA2*, *MYB*, *CEBPA*, *BCORL1*, and *CUX1*.

RAS/RTK genes: *FLT3*, *JAK2*, *MPL*, *CALR*, *CSF3R*, *PTPN11*, *NF1*, *NRAS*, *KRAS*, and *CBL*.

Clonal hematopoiesis (CH)-related genes¹⁰: *DNMT3A*, *TET2*, *ASXL1*, *SF3B1*, *PPM1D*, *JAK2*, and *CBL*.

Data Sharing Statement

Datasets of WES and RNA-seq data are available in the European Genome-phenome Archive database (Accession ID: EGAS50000000704 and EGAS50000000705).

Supplementary References

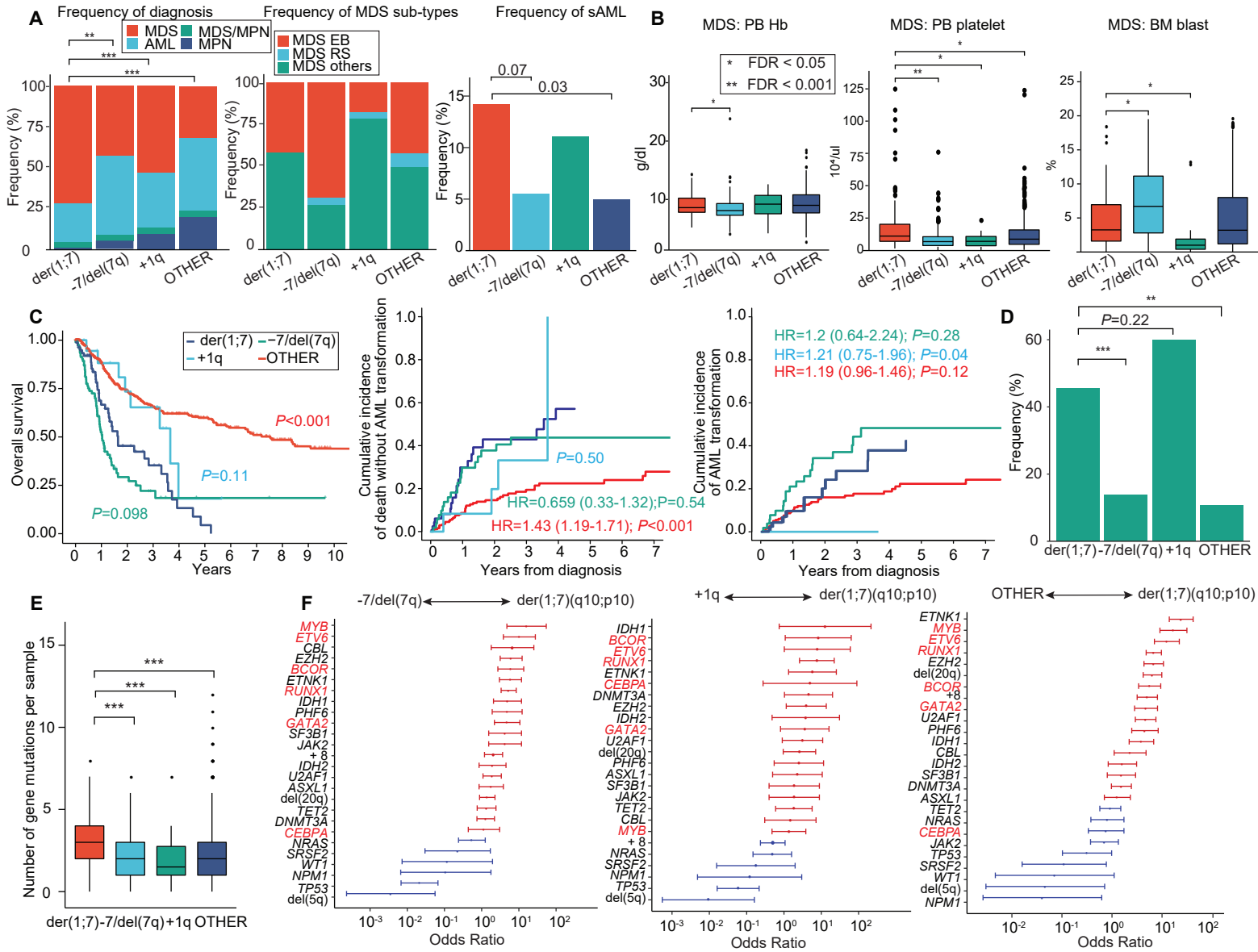
1. Haferlach, T. *et al.* Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* **28**, 241–247 (2014).
2. Yoshizato, T. *et al.* Genetic abnormalities in myelodysplasia and secondary acute myeloid leukemia: impact on outcome of stem cell transplantation. *Blood* **129**, 2347–2358 (2017).
3. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
4. Shiraishi, Y. *et al.* An empirical Bayesian framework for somatic mutation detection from cancer genome sequencing data. *Nucleic Acids Res.* **41**, (2013).
5. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
6. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
7. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
8. Subramanian, A. *et al.* *Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles.* (2005).
9. Ogawa, S. Genetics of MDS. *Blood* **133**, 1049–1059 (2019).
10. Saiki, R. *et al.* Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis. *Nat. Med.* **27**, 1239–1249 (2021).

Supplementary Table 1

Supplementary Table 1. Summary of clinical information of cohort analyzed.

	der(1;7)(q10;p10)				-7/del(7q)				P-value	+1q				P-value	OTHER				P-value
No. of patients	148.0				376.0					54.0					2808.0				
Age at diagnosis (median (range)) - yrs.	63.8 (16-90)				64.9 (17-90)				0.42	61.4 (17-87)				0.264	59.7 (10-94)				0.006
Gender - no. (%)																			
Male	130.0 (87.84)				269.0 (71.54)				< 0.001	36.0 (66.67)				0.001	1695.0 (60.36)				< 0.001
Female	18.0 (12.16)				107.0 (28.46)					18.0 (33.33)					1113.0 (39.64)				
WHO classification at sampling - no. (%)																			
MDS	(n=107)	(% Total / % Disease)			(n=163)	(% Total / % Disease)				(n=29)	(% Total / % Disease)				(n=807)	(% Total / % Disease)			
isolated del(5q)	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)			na	14.0	(0.50 / 1.73)	0.392		
MDS_SLD	17.0	(11.49 / 15.89)			4.0	(1.06 / 2.45)			< 0.001	4.0	(7.41 / 13.79)			1	112.0	(3.99 / 13.88)	0.556		
MDS_MLD	34.0	(22.97 / 31.78)			31.0	(8.24 / 19.02)			0.02	18.0	(33.33 / 62.07)			0.005	217.0	(7.73 / 26.89)	0.300		
MDS_RS	0.0	(0.00 / 0.00)			1.0	(0.27 / 0.61)			1	0.0	(0.00 / 0.00)			na	4.0	(0.14 / 0.50)	1		
MDS_RS_SLD	0.0	(0.00 / 0.00)			2.0	(0.53 / 1.23)			0.52	0.0	(0.00 / 0.00)			na	32.0	(1.14 / 3.97)	0.025		
MDS_RS_MLD	0.0	(0.00 / 0.00)			4.0	(1.06 / 2.45)			0.155	1.0	(1.85 / 3.45)			0.213	24.0	(0.85 / 2.97)	0.100		
MDS_EB	43.0	(29.05 / 40.19)			108.0	(28.72 / 66.26)			< 0.001	5.0	(9.26 / 17.24)			0.028	323.0	(11.50 / 40.02)	1		
MDS_U	8.0	(5.41 / 7.48)			6.0	(1.60 / 3.68)			0.261	0.0	(0.00 / 0.00)			0.202	32.0	(1.14 / 3.97)	0.125		
unknown	5.0	(3.38 / 4.67)			7.0	(1.86 / 4.29)				1.0	(1.85 / 3.45)				49.0	(1.75 / 6.07)			
AML	(n=35)	(% Total / % Disease)			(n=181)	(% Total / % Disease)				(n=18)	(% Total / % Disease)				(n=1232)	(% Total / % Disease)			
AML_recurrent_genetic_abnormality	0.0	(0.00 / 0.00)			16.0	(4.26 / 8.84)			0.081	2.0	(3.70 / 11.11)			0.111	281.0	(10.01 / 22.81)	< 0.001		
AML_MRC	28.0	(18.92 / 80.00)			162.0	(43.09 / 89.50)			0.151	10.0	(18.52 / 55.56)			0.106	430.0	(15.31 / 34.90)	< 0.001		
tAML	3.0	(2.03 / 8.57)			3.0	(0.80 / 1.66)			0.055	0.0	(0.00 / 0.00)			0.543	12.0	(0.43 / 0.97)	0.007		
AML_NOS	2.0	(1.35 / 5.71)			0.0	(0.00 / 0.00)			0.026	4.0	(7.41 / 22.22)			0.164	355.0	(12.64 / 28.81)	0.002		
unknown	2.0	(1.35 / 5.71)			0.0	(0.00 / 0.00)				2.0	(3.70 / 11.11)				154.0	(5.48 / 12.50)			
MDS/MPN	(n=5)	(% Total / % Disease)			(n=13)	(% Total / % Disease)				(n=2)	(% Total / % Disease)				(n=107)	(% Total / % Disease)			
MDS/MPN_atypical CML	0.0	(0.00 / 0.00)			2.0	(0.53 / 15.38)			1	1.0	(1.85 / 50.00)			0.268	8.0	(0.28 / 7.48)	1		
MDS/MPN_CMML	3.0	(2.03 / 60.00)			6.0	(1.60 / 46.15)			0.624	1.0	(1.85 / 50.00)			0.846	65.0	(2.31 / 60.75)	0.974		
MDS/MPN_RS_T	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)			na	4.0	(0.14 / 3.74)	1		
MDS/MPN_U	2.0	(1.35 / 40.00)			5.0	(1.33 / 38.46)			1	0.0	(0.00 / 0.00)			na	29.0	(1.03 / 27.10)	0.616		
unknown	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)			na	1.0	(0.04 / 0.93)			
MPN	(n=1)	(% Total / % Disease)			(n=19)	(% Total / % Disease)				(n=5)	(% Total / % Disease)				(n=662)	(% Total / % Disease)			
MPN_ET	0.0	(0.00 / 0.00)			7.0	(1.86 / 36.84)			1	3.0	(5.56 / 60.00)			1	216.0	(7.69 / 32.63)	1		
MPN_PV	0.0	(0.00 / 0.00)			1.0	(0.27 / 5.26)			1	0.0	(0.00 / 0.00)			na	114.0	(4.06 / 17.22)	1		
MPN_PMF	1.0	(0.68 / 100.00)			0.0	(0.00 / 0.00)			0.05	1.0	(1.85 / 20.00)			0.333	51.0	(1.82 / 7.70)	0.078		
MPN_CNL	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)			na	7.0	(0.25 / 1.06)	1		
MPN_CML_BCR/ABL	0.0	(0.00 / 0.00)			10.0	(2.66 / 52.63)			1	1.0	(1.85 / 20.00)			1	240.0	(8.55 / 36.25)	1		
MPN_CEL_NOS	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)			na	1.0	(0.04 / 0.15)	1		
MPN_NOS	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)			na	4.0	(0.14 / 0.60)	1		
MPN_U	0.0	(0.00 / 0.00)			1.0	(0.27 / 5.26)			1	0.0	(0.00 / 0.00)			na	21.0	(0.75 / 3.17)	1		
unknown	0.0	(0.00 / 0.00)			0.0	(0.00 / 0.00)			na	0.0	(0.00 / 0.00)				8.0	(0.28 / 1.21)			
Peripheral Blood [MDS]	(n=106)			Supplementary Table 1: Summary of clinical information of cohort analyzed.						(n=29)				(n=807)					
WBC (mean (range)) /uL	3238.6	(600-17800)			4362.7	(700-44900)			0.101	3148.8	(200-9290)			0.871	4470.4	(230-84600)	0.060		
HB (mean (range)) g/dL	9.0	(4.7-14.3)			8.2	(4.1-23.8)			0.027	8.6	(4.2-12.6)			0.359	9.2	(2.7-18.4)	0.365		
Plt (mean (range)) x10 ⁹ /L	19.8	(0.5-124.7)			9.6	(0.1-92.8)			< 0.001	7.2	(0.09-22.9)			0.015	12.8	(0.1-177.8)	< 0.001		
Bone Marrow [MDS]	(n=106)									(n=29)				(n=807)					
Blast (mean (range)) %	4.8	(0-19.4)			7.4	(0-19.5)			< 0.001	2.6	(0-13.2)			0.017	5.0	(0-19.6)	0.693		
Peripheral Blood [AML]	(n=35)									(n=18)				(n=1232)					
WBC (mean (range)) /uL	3787.2	(600-33100)			16069.8	(50-222880)			0.045	21582.7	(670-120500)			0.024	36615.6	(230-676000)	0.013		
HB (mean (range)) g/dL	7.7	(4-11.5)			8.0	(4-26.3)			0.529	10.1	(4.8-15.3)			0.005	9.1	(3.2-79)	0.076		
Plt (mean (range)) x10 ⁹ /L	4.1	(1-14.8)			7.9	(0.3-91.3)			0.065	5.4	(0.2-17.3)			0.371	9.9	(0.2-505)	0.192		
Bone Marrow [AML]	(n=35)									(n=18)				(n=1232)					
Blast (mean (range)) %	34.2	(20.7-73)			43.6	(20-98.4)			0.055	49.2	(22.3-90)			0.19	39.4	(20-100)	0.034		

Supplementary Fig. 1



Supplementary Fig. 1. Clinical and genetic characteristics of *der(1;7)(q10;p10)(+)* myeloid neoplasms. (A) Myeloid neoplasm disease distribution (left), MDS diagnosis distribution (middle), and Frequency of secondary AML (right) for *der(1;7)(q10;p10)(+)*, *-7/del(7q)*, *+1q*, and OTHER (Fisher's Exact Test). ** $P < 0.01$, *** $P < 0.001$ (B) Box plot of peripheral blood hemoglobin (Hb) count, platelet count and bone marrow blast count of *der(1;7)(q10;p10)(+)*, *-7/del(7q)*, *+1q* and OTHER. Boxplot depicts median, 1st and 3rd quartile. False discovery rate (FDR) calculated by Benjamin-Hochberg procedure. (C) Kaplan-Meier curves for overall survival (left), cumulative incidence of death without AML transformation (middle), and cumulative incidence of AML transformation (right) of *der(1;7)(q10;p10)(+)*, *-7/del(7q)*, *+1q*, and OTHER. Number at risk shown below the curve for Kaplan-Meier curve. Hazard ratios (HR) and P-values relative to *der(1;7)(q10;p10)(+)* for cumulative incidence curves. (D) Bar graph showing the fraction of infection-related deaths in *der(1;7)(q10;p10)(+)* vs. *-7/del(7q)* vs. *+1q* vs. OTHER MDS cases (Fisher's Exact Test). ** $P < 0.01$, *** $P < 0.001$ (E) Boxplot for number of additional genetic events per case. The median and 1st and 3rd quartiles are indicated, and whiskers extend to the furthest value within $1.5 \times$ the interquartile range. P-values calculated by Wilcoxon test. *** $P < 0.001$ (F) Forest plot of mutation frequency odds ratio. Each sub-group vs *der(1;7)(q10;p10)(+)*. Error bars indicate 95% confidence interval. Transcription factor genes in red.

A



Supplementary Fig. 2. Characteristic gene mutations of der(1;7)(q10;p10). (A) Boxplot of peripheral blood white blood cell (WBC) count, hemoglobin (Hb) level, platelet count, neutrophil count, eosinophil level, and bone marrow blast percentage count (Wilcoxon test P-values) and Kaplan-Meier curve for overall survival of all MDS cases comparing wild type (wt)(red) vs. mutated (mut)(blue) cases for *MYB* (above) and *ETNK1* (below). (B) Dot plot depicting log ratio of expression of genes comparing der(1;7)(q10;p10) with OTHER group. Genes on chromosomes 1q and 7q are depicted in (left) and (right), respectively. Key genes colored as follows: Oncogene (blue) and p38 MAPK-signaling (green) on (left); Tumor suppressor (red) and Other known genes (blue) on (right). (C) Bradley-Terry analysis result for gene mutations and CNAs involved in at least five pairwise precedence. The relative temporal order shown with respect to *ASXL1*. Standard errors shown as horizontal bars. (D) Scatter plot of tumor cell fraction (TCF) of der(1;7)(q10;p10) (x-axis) and adjusted variable allele frequency (aVAF) of other mutations (y-axis). y-axis represents aVAF of CH related genes (*DNMT3A*, *TET2*, *ASXL1*, *SF3B1*, and *PPM1D*) in (right), and transcription factor genes (*RUNX1*, *BCOR*, *ETV6*, *GATA2*, *MYB*, and *PHF6*) (TF), *ETNK1*, and other genes on the left. Each dot represents each mutation. (E) Pie chart showing the development of MDS or AML for der(1;7)(q10;p10), del(7q) and +1q harboring individuals. (F) Table showing number of *ETNK1* mutations found in each sub-group through duplex-sequencing. (Fisher's Exact Test)