SUPPLEMENTAL MATERIAL

Clonal Trajectories and Cellular Dynamics of Myeloid Neoplasms with *SF3B1* Mutations

Hassan Awada[1], Cassandra M. Kerr[1], Arda Durmaz[1], Vera Adema[1], Carmelo Gurnari[1], Simona Pagliuca[1], Misam Zawit[1], Sunisa Kongkiatkamon[1], Heesun J. Rogers[2], Yogen Saunthararajah[1,3], Mikkael A. Sekeres[3], Hetty Carraway[3], Jaroslaw P. Maciejewski[1,3], Valeria Visconte[1]

[1]Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, USA; [2]Department of Laboratory Medicine, Cleveland Clinic, Cleveland, OH, USA; [3]Leukemia Program, Department of Hematology and Medical Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, USA

Running title: Clonal architecture of *SF3B1* mutations

Keywords: *SF3B1* mutations, clonal hierarchy, myeloid neoplasms

Correspondence to: Valeria Visconte, Ph.D.; Department of Translational Hematology and Oncology Research**,** Lerner Research Institute**,** Cleveland Clinic**,** 9620 Carnegie Ave n building, NE6-250**,** Cleveland, OH 44106; Phone: 216-445-6895, Fax: 216-636-2498, E-mail: visconv@ccf.org

Text: 1,488 words
Tables/ Figures: 0/2
References (main text): 10
Supplemental Tables/ Figures: 3/9
References (supplemental): 7

METHODS

Patients

Data were collected from a total of 3673 patients from the Cleveland Clinic Foundation (CCF) and publicly available databases [The Cancer Genome Atlas (TCGA) [1], The German-Austrian Study Group [2] and Beat AML Master Trial [3]]. In total 209 patients carried *SF3B1* mutations. Written informed consents according to the protocols approved by the Institutional Review Board of the CCF in accordance with the Declaration of Helsinki were obtained for the collection of peripheral blood and/ or bone marrow specimens. Clinical and molecular data were collected after obtaining appropriate material transfer agreements and from resources available online.

Genomic Studies

Targeted sequencing was performed using a TruSeq Custom Amplicon (Illumina, San Diego, CA), a custom cRNA bait library (SureSelect; Agilent Technologies, Santa Clara, CA) or a Nextera technology as previously described [4,5,6]. Sequencing libraries were generated according to Illumina paired-end library protocols. Our collection of targeted sequencing covered a panel of 63 (TruSeq) and 182 (Nextera) genes frequently mutated in myeloid malignancies. The enriched targets were subjected to sequencing using HiSeq 2000 or MiSeq sequencer (Illumina). Variants were called using an in-house bioanalytic pipeline as previously described [4,5]. Mutations were scored as somatic by comparison with healthy controls and mutational databases, such as dbSNP138, 1000 Genomes or ESP 6500 database, and Exome Aggregation Consortium (ExAC). Mapping errors were removed by visual inspection by using the Integral Genomics Viewer.

Clonal Hierarchy

The clonal hierarchy was resolved using our in-house designed VAF-based bioanalytic method. To avoid false-positives, only variants with variant allele frequencies (VAFs) ≥5% were considered for further analysis. For the purpose of this study and given resolution limitations due to sequencing depth, we used a cut-off of at least 5% difference between VAFs to identify founder mutations (ΔVAF >5%). This cut-off difference was previously confirmed by the PyClone pipeline, which showed a high level of

concordance [4]. First, we assigned clonal hierarchy to *SF3B1*^{MT} by using VAFs (adjusted for copy number and zygosity) and then classified the mutations into dominant/ancestral (if a cutoff of at least 5% difference between the highest 2 VAFs existed), secondary/sub-clonal (any subsequent hit that is >5% lower than the highest VAF) and co-dominant (if the difference between the highest 2 VAFs is <5%).

Single cell-DNA sequencing

Bone marrow mononuclear cells from patients with *SF3B1* mutations (n=2) were thawed overnight in Iscove Modified Dulbecco Media (Thermo Fisher Scientific, Waltham, MA) supplemented with 10% fetal bovine serum. The day after cells were washed with 1X phosphate-buffer saline and 3-4,000 cells/µL were resuspended in Tapestri's cell buffer and encapsulated using Tapestri microfluidics cartridges, lysed, and barcoded using a Mission Bio (South San Francisco, CA) [7] sequencer for single cell analysis. Samples were amplified using the 45-myeloid gene panel (Tapestri). PCR products were purified using Ampure XP beads (Beckman Coulter, Pasadena, CA), quantified for size and concentration using an Agilent bioanalyzer and pooled libraries were run on an Illumina NovaSeq at the genomic core of the Lerner Research Institute (Cleveland Clinic). Variants analysis was conducted using Tapestri Insights platform version 2.2. To differentiate high-quality variants from low-quality variants, a series of metrics were applied. Potential false positive were excluded if: a) relatively a low percentage of genotyped cells per sample was observed; b) comparable number of mutated cells were found across unrelated and independent samples; c) VAF by read count versus VAF by cell count was more than 1.5 fold discrepant; d) average clone and variant specific VAFs deviate from associated zygosity values.

Statistical analyses

Fisher's exact test, Chi-square and analysis of variance (ANOVA) tests were used to compare categorical variables. Mann–Whitney U test/ Wilcoxon rank-sum test was used for continuous variables. Univariate logistic regression analysis was performed. Survival analyses were done using Kaplan-Meier

estimator method. All *P*-values were two-sided; those less than 0.05 were considered statistically significant. All statistical computations were performed using R 3.5.1 (www.r-project.org) and GraphPad Prism 8.

References

1. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013; 368: 2059-2074.

2. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. N Engl J Med. 2016; 374: 2209–2221.

3. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. Nature. 2018;562: 526-531.

4. Nagata Y, Makishima H, Kerr CM, et al. Invariant patterns of clonal succession determine specific clinical features of myelodysplastic syndromes. Nat Commun. 2019; 10: 5386.

5. Hirsch CM, Nazha A, Kneen K, et al. Consequences of mutant TET2 on clonality and subclonal hierarchy. Leukemia. 2018; 32:1751-1761.

6. Awada H, Nagata Y, Goyal A, et al. Invariant phenotype and molecular association of biallelic TET2 mutant myeloid neoplasia. Blood Adv. 2019; 3: 339–349.

7. Mission Bio: https://missionbio.com/

Supplemental Table 1. Baseline and cytogenetic characteristics of *SF3B1* mutant myeloid neoplasms

Variables	***SF3B1*^{MT}** n (%)
Age (y), (median/range)	70 (10-89)
≥ 60 y	135 (78)
Female/Male	94 (45)/115 (55)
2016 WHO Classification	
MDS	89 (43)
MDS-SLD/ MDS-MLD/ isolated del(5q)/ MDS-U	28
MDS-SLD-RS/ MDS-MLD-RS	44
MDS EB-1/2	17
MDS/MPN Overlap	29 (14)
CMML-1/2	5
MDS/MPN-U/ MDS/MPN-RS-T	24
MPN	4 (2)
pAML	45 (22)
sAML	30 (14)
tAML	5 (2)
Karyotype	
normal	119 (59)
-5/del(5q)	16 (8)
-7/del(7q)	10 (5)
-17/del(17q)	2 (1)
del(20q)	8 (4)
+8	14 (7)
-Y	9 (5)
complex	14 (7)

Abbreviations: *SF3B1*^{MT} , *SF3B1* mutant; y, years; WHO, World Health Organization; MDS, myelodysplastic syndrome; MDS-SLD, MDS with single lineage dysplasia; MDS-MLD, MDS with multilineage dysplasia; MDS-U, MDS unclassifiable; MDS-SLD-RS, MDS-SLD with ring sideroblasts; MDS-MLD-RS, MDS-MLD with ring sideroblasts; MDS-EB, MDS with excess blasts; MPN, myeloproliferative neoplasms; MDS/MPN-U, MDS/MPN unclassifiable; MDS/MPN-RS-T, MDS/MPN-RS with thrombocytosis; CMML, chronic myelomonocytic leukemia; pAML, primary acute myeloid leukemia; sAML, secondary acute myeloid leukemia; t-AML, therapy-related acute myeloid leukemia

Supplemental Table 2. Baseline and cytogenetic characteristics of co-dominant versus dominant and secondary *SF3B1* mutant myeloid neoplasms

Variables	*SF3B1*^{COD} n (%)	*vs. SF3B1*^{DOM} P-value	*vs. SF3B1*^{SEC} P-value
Age (y), (median/range)	70 (10-84)	0.8	0.3
≥ 60 y	26 (68)	0.5	**0.03**
Female/Male	19 (42)/26 (58)	0.5	1
2016 WHO Classification			
MDS	21 (47)	1	0.2
MDS-SLD/ MDS-MLD/ isolated del(5q)/ MDS-U	8	0.4	0.4
MDS-SLD-RS/ MDS-MLD-RS	8	0.2	0.6
MDS EB-1/2	5	0.5	0.5
MDS/MPN Overlap	6 (13)	1	0.8
CMML-1/2	1	1	0.6
MDS/MPN-U/ MDS/MPN-RS-T	5	1	1
MPN	1 (2)	1	1
pAML	10 (22)	1	1
sAML	6 (13)	0.6	0.3
tAML	1 (2)	1	0.6
Karyotype			
normal	26 (59)	1	0.8
-5/del(5q)	4 (9)	0.5	1
-7/del(7q)	1 (2)	0.4	0.6
-17/del(17q)	0 (0)	1	1
del(20q)	2 (5)	1	0.6
+8	2 (5)	0.7	0.7
-Y	2 (5)	1	0.6
complex	3 (7)	1	1

Abbreviations: *SF3B1*^{COD}, co-dominant *SF3B1* mutation*; SF3B1*^{COM}, dominant *SF3B1* mutation; *SF3B1*^{SEC}, secondary *SF3B1* mutation; y, years; WHO, World Health Organization; MDS, myelodysplastic syndrome; MDS-SLD, MDS with single lineage dysplasia; MDS-MLD, MDS with multilineage dysplasia; MDS-U, MDS unclassifiable; MDS-SLD-RS, MDS-SLD with ring sideroblasts; MDS-MLD-RS, MDS-MLD with ring sideroblasts; MDS-EB, MDS with excess blasts; MPN, myeloproliferative neoplasms; MDS/MPN-U, MDS/MPN unclassifiable; MDS/MPN-RS-T, MDS/MPN-RS with thrombocytosis; CMML, chronic myelomonocytic leukemia; pAML, primary acute myeloid leukemia; sAML, secondary acute myeloid leukemia; t-AML, therapy-related acute myeloid leukemia

Supplemental Table 3. Clinical characteristics of dominant *vs.* secondary *SF3B1* mutant myeloid neoplasms

Variables	*SF3B1*^{DOM} n (%)	*SF3B1*^{SEC} n (%)	*P*-value
Age (y), (median/range)	69 (29-89)	70 (21-85)	0.1[B]
≥ 60 y	95 (80)	94 (85)	0.3[A]
Female/Male	69 (49)/71 (51)	48 (40)/73 (60)	0.1[A]
2016 WHO Classification			
MDS	46 (33)	22 (18)	**0.007**[A]
MDS-SLD/ MDS-MLD/ isolated del(5q)/ MDS-U	12	8	0.6[A]
MDS-SLD-RS/ MDS-MLD-RS	27	9	**0.006**[A]
MDS EB-1/2	7	5	0.7[A]
MDS/MPN Overlap	12 (9)	11 (9)	1[A]
CMML-1/2	1	3	0.3[A]
MDS/MPN-U/ MDS/MPN-RS-T	11	8	0.8[A]
MPN	1 (1)	2 (2)	0.6[A]
pAML	54 (39)	53 (43)	0.6[A]
sAML	18 (12)	28 (23)	**0.03**[A]
tAML	4 (3)	3 (3)	1[A]
Karyotype			
normal	78 (58)	69 (60)	0.7[A]
-5/del(5q)	9 (7)	8 (7)	1[A]
-7/del(7q)	8 (7)	9 (8)	1[A]
-17/del(17q)	1 (1)	3 (3)	0.3[A]
del(20q)	5 (4)	1 (1)	0.2[A]
+8	8 (6)	8 (7)	0.7[A]
-Y	7 (4)	1 (1)	0.07[A]
complex	13 (10)	11 (9)	1[A]
Hematological parameters (median/range)			
WBCs (x 10^9/L), (median/range)	5.34 (0.3-226)	6 (0.9-207)	0.1[B]
< 3.5 x 10^9/L	36 (28)	33 (31)	0.6[A]
Hgb (g/dL), (median/range)	9.1 (4.4-14.1)	9.1 (5.2-16.1)	0.5[B]
< 10 g/dL	85 (67)	79 (75)	0.2[A]
Plt (x 10^9/L), (median/range)	130 (9-1116)	97 (6-950)	**0.05**[B]
< 150 x 10^9/L	67 (54)	70 (67)	**0.05**[A]
ANC (x 10^9/L), (median/range)	2.1 (0.04-37.3)	1.8 (0.01-38.2)	0.5[B]
< 1.5 x 10^9/L	41 (42)	37 (46)	0.6[A]
RS, (median/range)	37 (0-81)	3 (0-80)	**0.002**[B]
≥ 15%	36 (77)	16 (44)	**0.007**[A]
Cytopenia			
None	18 (14)	8 (8)	0.1[A]
Mono-cytopenia	44 (34)	25 (24)	0.08[A]
Bi-cytopenia	46 (36)	55 (52)	**0.01**[A]
Pan-cytopenia	20 (16)	18 (17)	0.8[A]
¥**BM cellularity and morphology**			
Blasts, (median/range)	26 (0-94)	26 (0-96)	0.1[B]
≥ 5%	78 (61)	77 (71)	0.1[A]
Hypercellular	47 (52)	58 (64)	0.09[A]
Normocellular	42 (46)	26 (29)	**0.02**[A]

Hypocellular	2 (2)	6 (7)	0.1[A]
Dysplastic myeloid lineages	18 (29)	25 (53)	**0.01**[A]
No dysplasia	9 (15)	4 (8)	0.3[A]
Uni-lineage	24 (39)	12 (25)	0.1[A]
Bi-lineage	16 (26)	22 (47)	**0.02**[A]
Tri-lineage	13 (21)	9 (20)	1[A]

Abbreviations: *SF3B1*^{DOM} , dominant *SF3B1* mutation; *SF3B1*^{SEC}, secondary *SF3B1* mutation; y, years; WHO, World Health Organization; MDS, myelodysplastic syndrome; MDS-SLD, MDS with single lineage dysplasia; MDS-MLD, MDS with multilineage dysplasia; MDS-U, MDS unclassifiable; MDS-SLD-RS, MDS-SLD with ring sideroblasts; MDS-MLD-RS, MDS-MLD with ring sideroblasts; MDS-EB, MDS with excess blasts; MPN, myeloproliferative neoplasms; MDS/MPN-U, MDS/MPN unclassifiable; MDS/MPN-RS-T, MDS/MPN-RS with thrombocytosis; CMML, chronic myelomonocytic leukemia; pAML, primary acute myeloid leukemia; sAML, secondary acute myeloid leukemia; t-AML, therapy-related acute myeloid leukemia

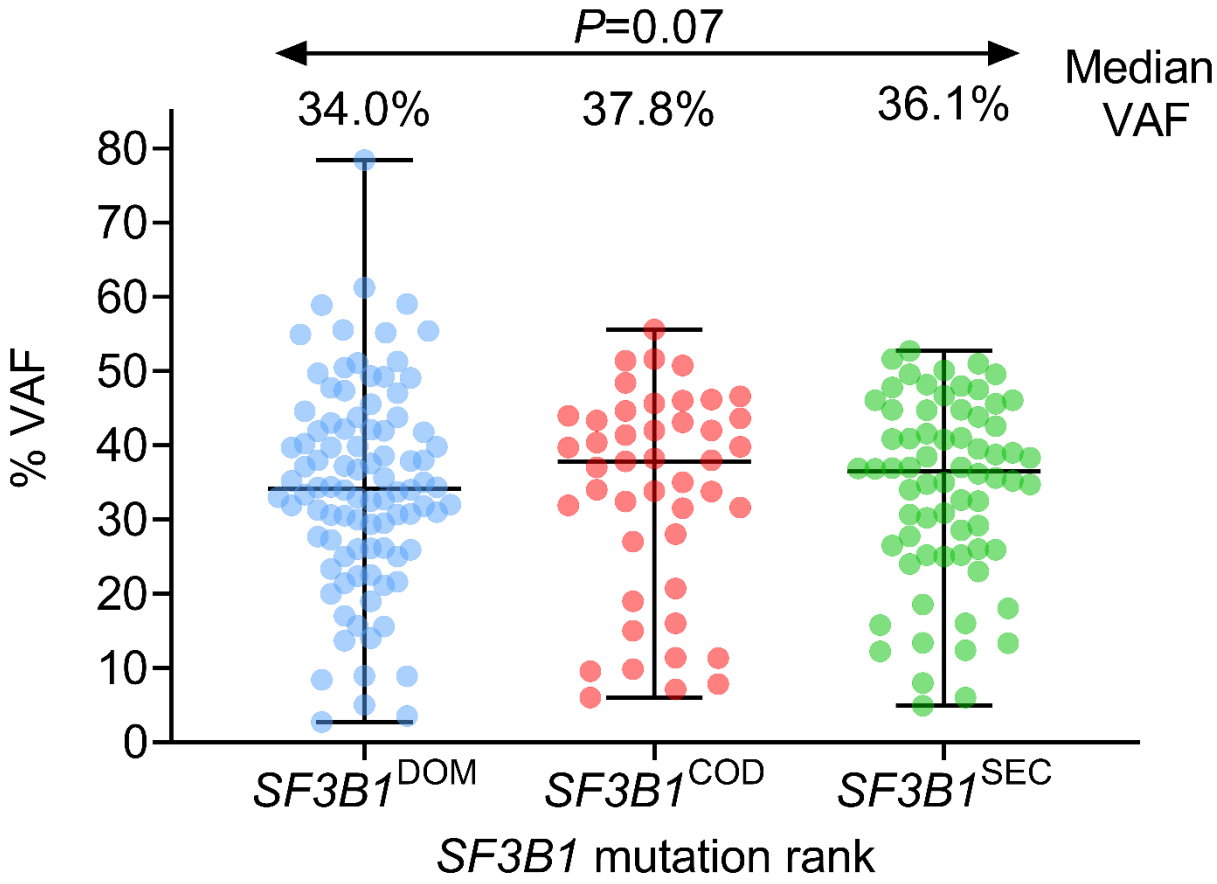
WBCs, white blood cells; Hgb, hemoglobin; RS, ring sideroblasts; BM, bone marrow; Plt, platelet count; ANC, absolute neutrophil count

The cohort also included 8 patients with other disorders including acute promyelocytic leukemia, acute lymphoblastic leukemia and aplastic anemia

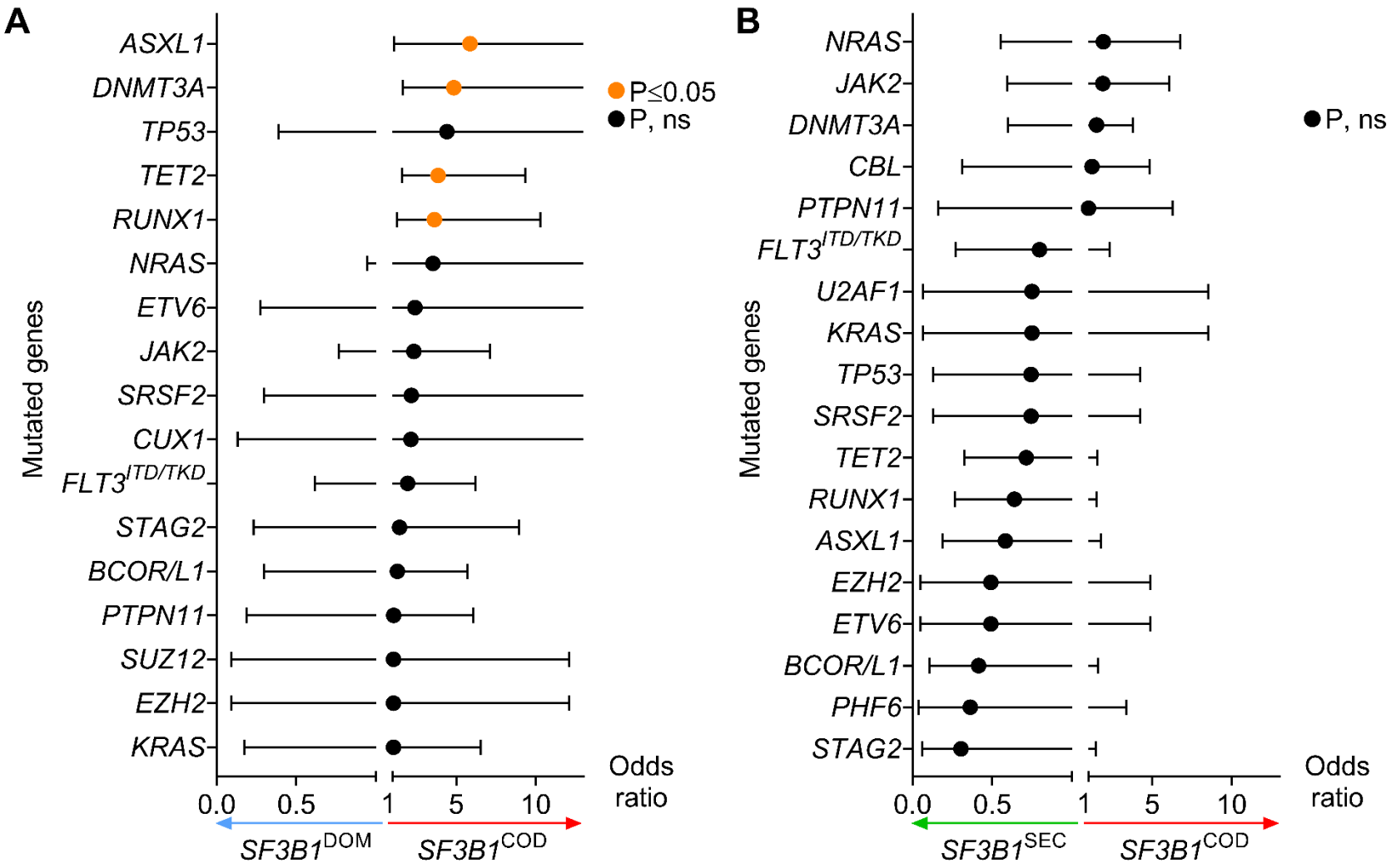
[¥]Bone marrow cellularity was adjusted to age

[A]Fisher's exact test; [B]Mann–Whitney U test

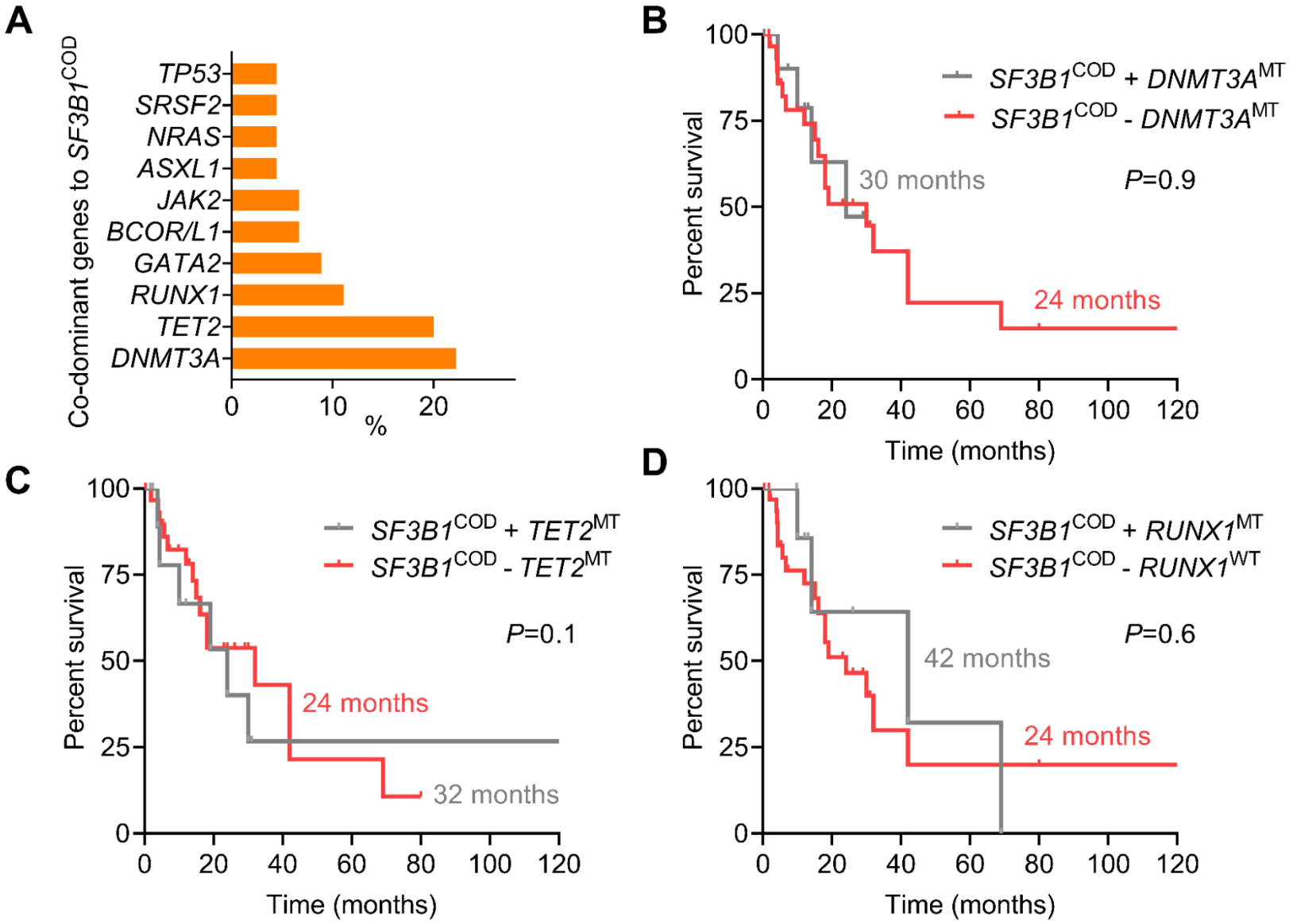
Supplementary Figure 1



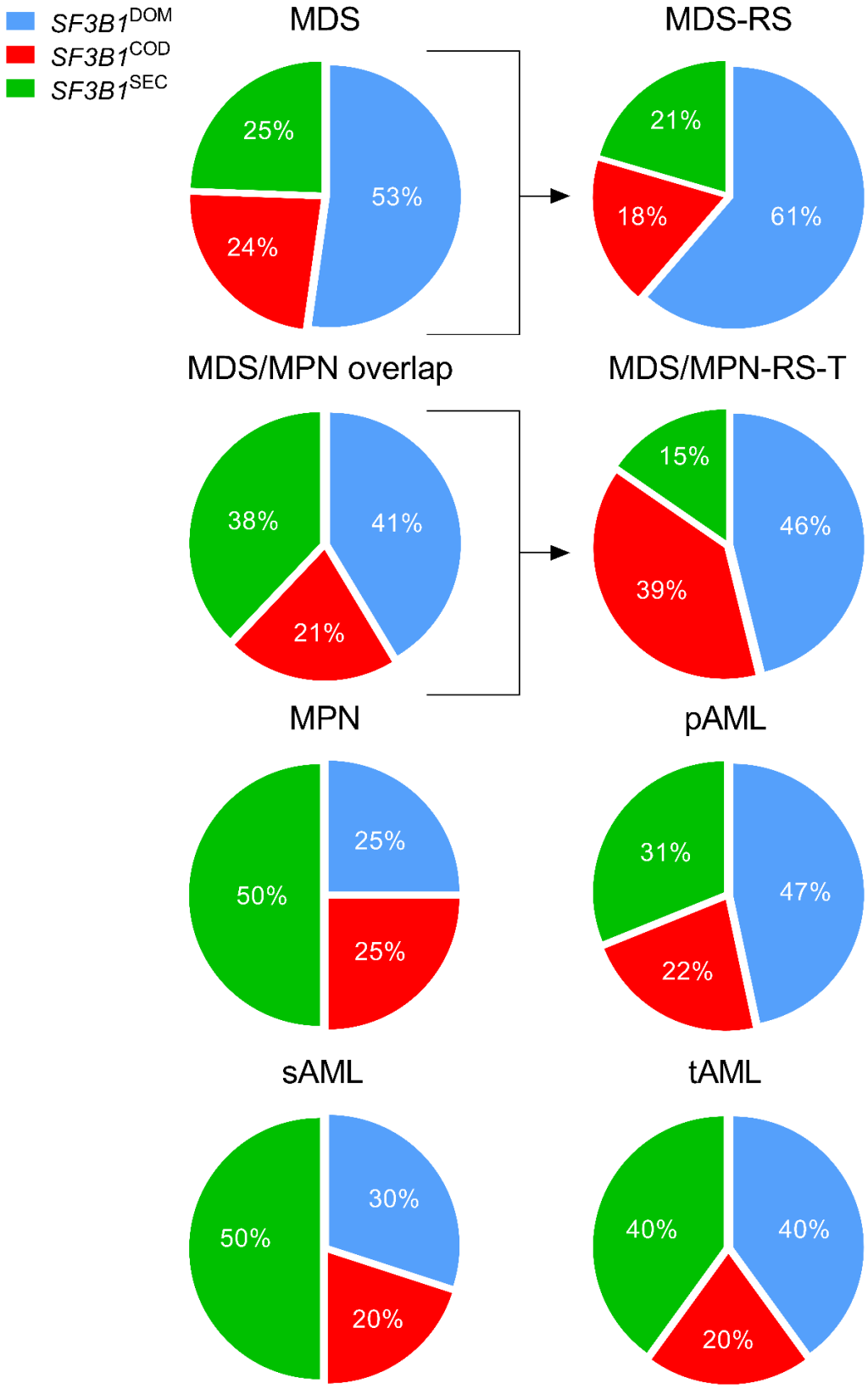
Supplementary Figure 2



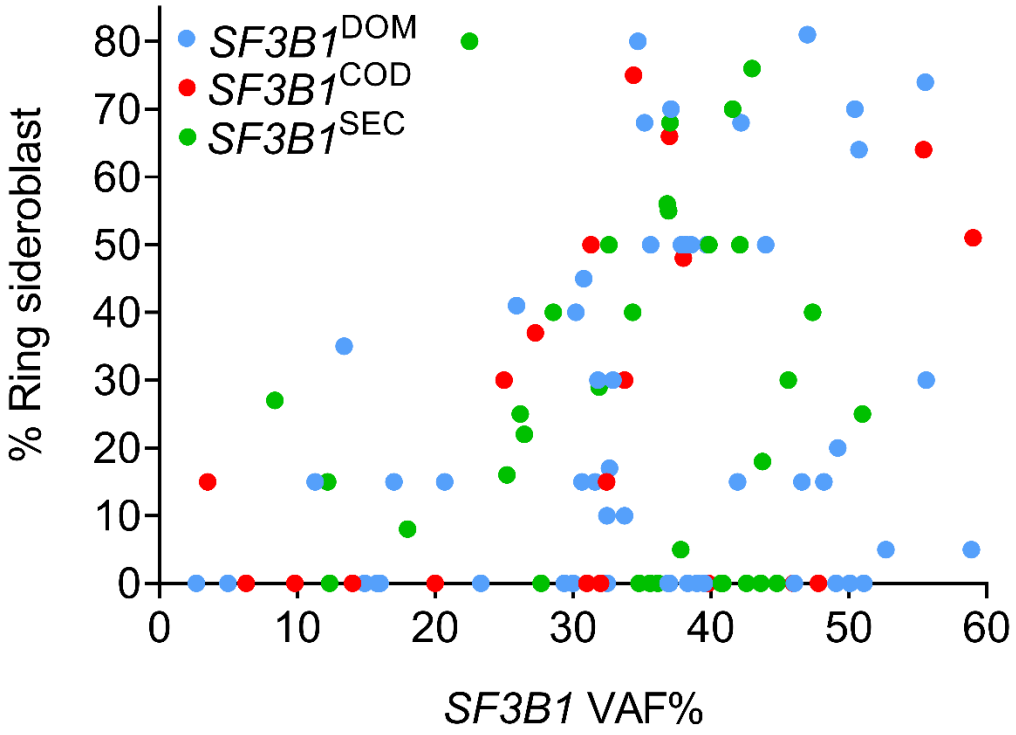
Supplementary Figure 3



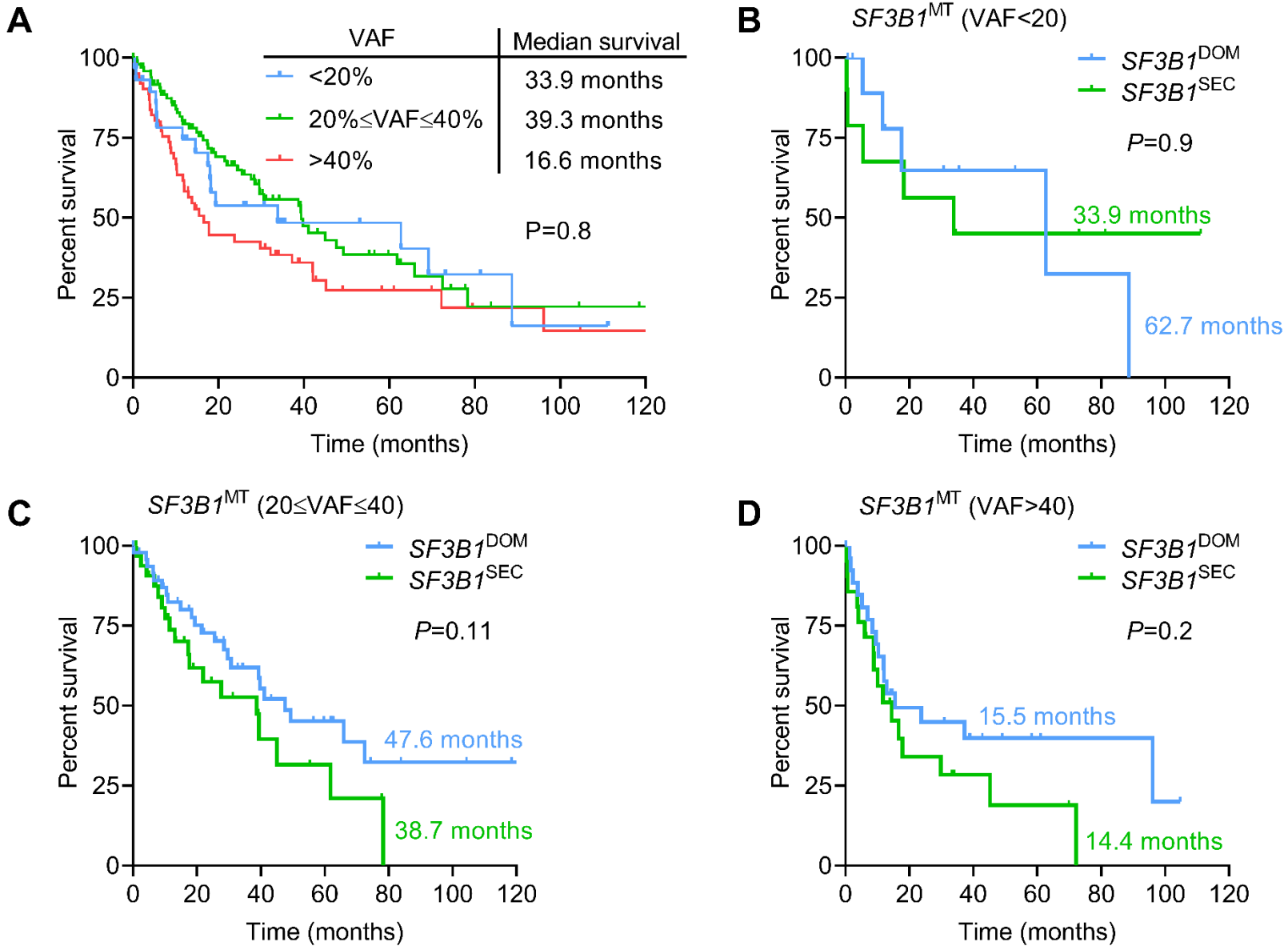
Supplementary Figure 4



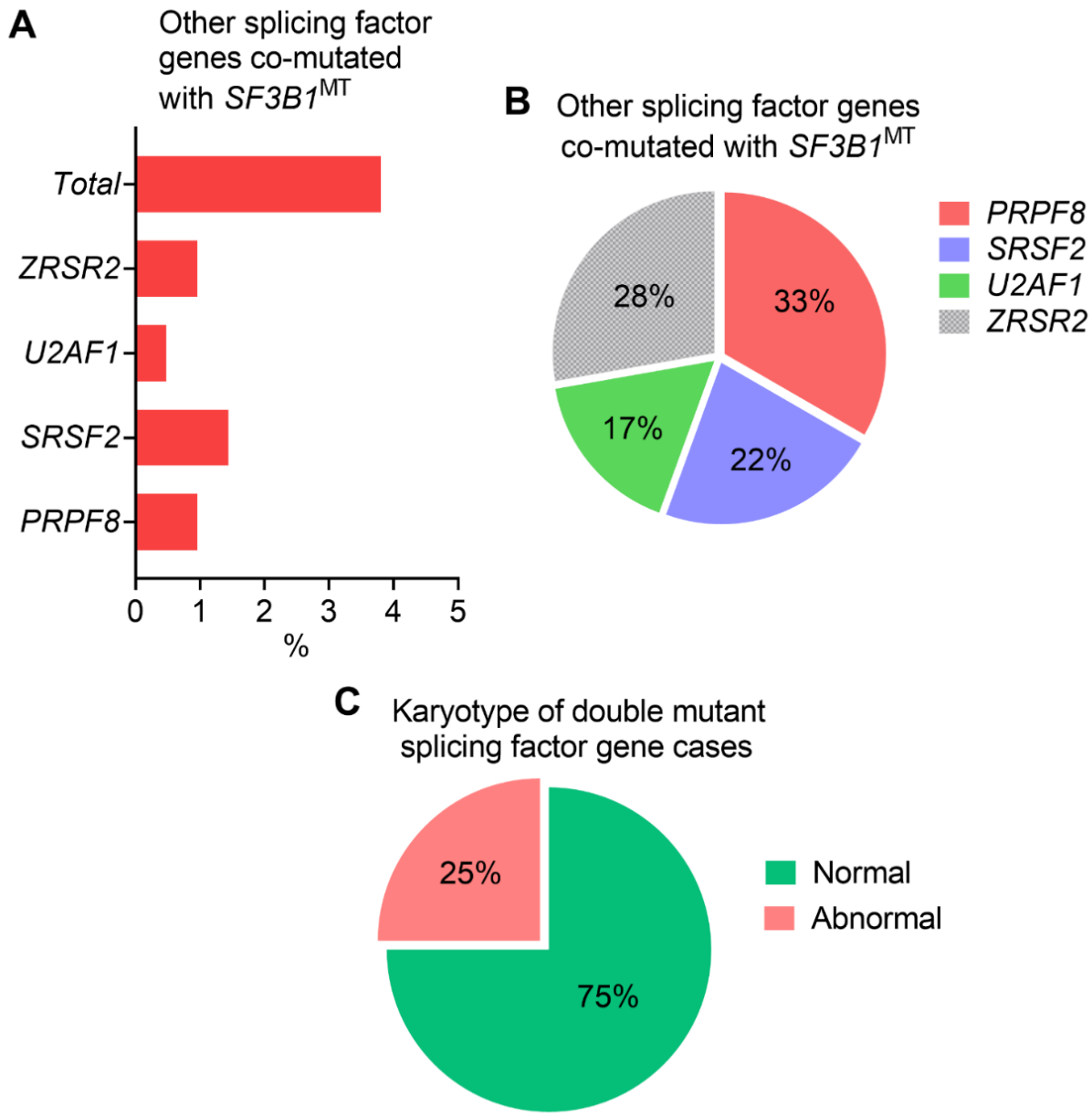
Supplementary Figure 5



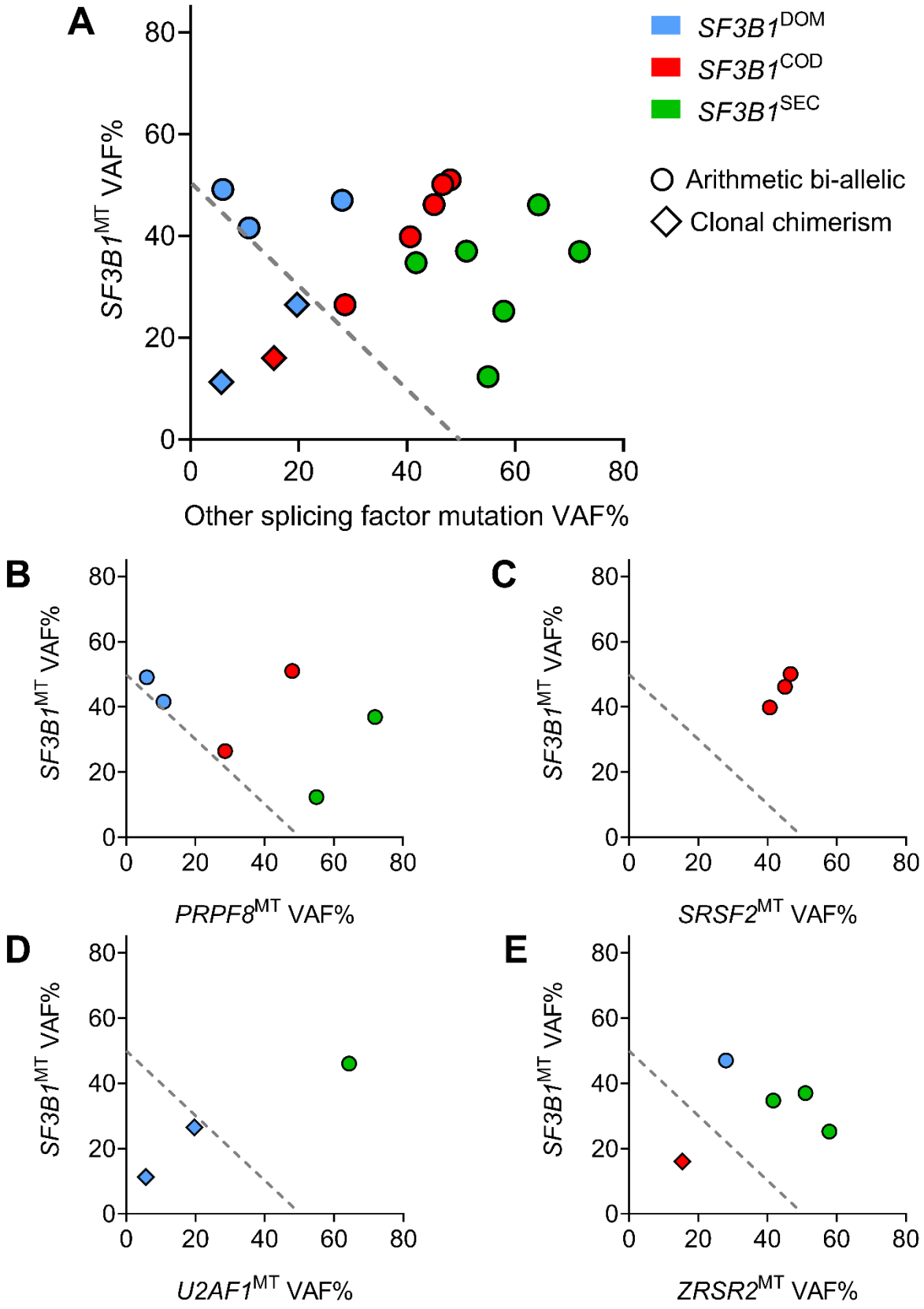
Supplementary Figure 6



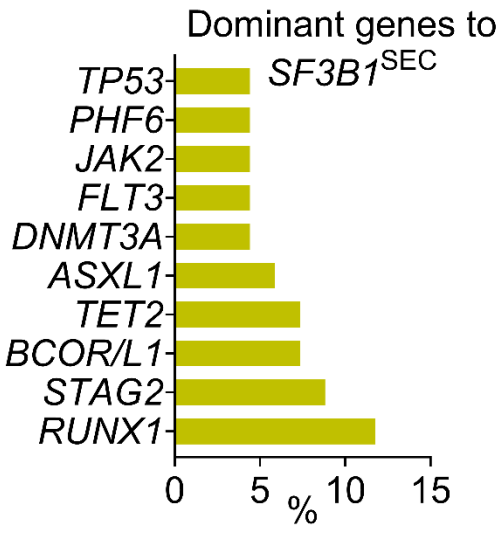
Supplementary Figure 7



Supplementary Figure 8



Supplementary Figure 9



Supplementary Table Legends

Supplemental Table S1. Baseline and cytogenetic characteristics of *SF3B1* mutant myeloid neoplasms. Medical demographics including cytogenetics of 209 patients with myeloid neoplasms carrying *SF3B1* mutations are summarized.

Supplemental Table S2. Baseline and cytogenetic characteristics of co-dominant versus dominant and secondary *SF3B1* mutant myeloid neoplasms. Comparisons of baseline characteristics and cytogenetics between *SF3B1* mutations in co-dominat versus dominant and secondary status.

Supplemental Table S3. Clinical characteristics of dominant *vs.* secondary *SF3B1* mutant myeloid neoplasms. Comparisons of demographics, baseline and morphologic parameters and cytogenetics of dominant versus secondary *SF3B1* mutations.

Supplementary Figure Legends

Supplemental Figure 1. Comparison of variant allele frequencies between *SF3B1* dominant, co-dominant and secondary. Scatter plot describing the variant allele frequency in percentage corresponding to individual patients with *SF3B1* dominant (light grey), co-dominant (orange) and secondary (green). Median variant allele frequency in percentage is indicated for every *SF3B1* clonal status. Levels of statistical significance was calculated by *P* values.

Supplemental Figure 2. Molecular associations of *SF3B1* co-dominant versus dominant and secondary. Odds ratio measuring the associations of mutated myeloid genes and *SF3B1* clonal status by comparison of **(A)** *SF3B1* dominant versus co-dominant and **(B)** *SF3B1* secondary versus co-dominant. Levels of statistical significance was calculated by *P* values.

Supplemental Figure 3. Molecular and clinical features of patients with *SF3B1* co-dominant. (A) Bar graph depicting the frequency of the most frequent gene mutations in patients with *SF3B1* co-dominant. Survival analysis of patients with myeloid neoplasia and *SF3B1* mutations. Kaplan-Meier curves representing the percent survival of patients with myeloid neoplasia with *SF3B1* co-dominant

carrying **(B)** *DNMT3A*, *TET2* **(C)** and *RUNX1* **(D)** mutations compared to patients with myeloid neoplasia with *SF3B1* co-dominant mutations only.

Supplemental Figure 4. Frequency of *SF3B1* mutations in myeloid neoplasia. Pie charts representing *SF3B1* mutations according to clonal status (dominant, blue; co-dominant, red; secondary, green) in different disease subtypes. MDS, myelodysplastic syndromes; MDS-RS, myelodysplastic syndromes ringed sideroblasts; MPN, myeloproliferative neoplasms; MDS-MPN-RS-T, Myelodysplastic syndrome/myeloproliferative neoplasm with ring sideroblasts and thrombocytosis; pAML, primary acute myeloid leukemia; sAML, secondary acute myeloid leukemia; tAML, therapy-related acute myeloid leukemia.

Supplemental Figure 5. Relation between ringed sideroblasts and *SF3B1* mutations. (A) Scatter plot representing the correlation between the percentage of ringed sideroblasts versus variant allele frequencies of patients with *SF3B1* dominant (blue), secondary (green) and co-dominant (red) mutations.

Supplemental Figure 6. Survival outcomes of patients with myeloid neoplasia harboring *SF3B1* mutations with different variant allele frequencies. (A-D) Kaplan-Meier curves representing the percent survival of patients *SF3B1* mutations and variant allele frequencies <20%, between 20-40% and >40%.

Supplemental Figure 7. Concomitant presence of other splicing factor mutations. (A-B) Bar graph and pie chart describing the frequency of other splicing factor in our cohort of *SF3B1* mutant myeloid neoplasia patients. **(C)** Pie chart describing the percentage of patients with more than one splicing factor mutations with normal or abnormal karyotype.

Supplemental Figure 8. Variant allele frequency of mutations in splicing factors. Scatter plots indicating variant allele frequencies of *SF3B1* (Y-axis) and variant allele frequencies of all the splicing factors (X-axis) **(A)** and **(B-E)** variant allele frequencies of each splicing factor (X-axis). We used a cut-off of combined variant allele frequency (VAF) sum of 50% to determine i) arithmetic bi-allelic double splicing factor mutations (VAF sum >50%) *vs*. ii) clonal chimerism (VAF sum <50%).

Supplemental Figure 9. Dominant gene mutations to secondary *SF3B1*. A bar graph representing the frequency (in percent) of the most common dominant gene mutations to *SF3B1* secondary (*SF3B1*^{SEC}) mutation.