TET2 mutational status affects myelodysplastic syndrome evolution to chronic myelomonocytic leukemia

Violaine Tran Quang,[1,2] Benjamin Podvin,[3+] Christophe Desterke,[4+] Sihem Tarfi,[1,2] Quentin Barathon,[2] Bouchra Badaoui,[2] Nicolas Freynet,[2] Vincent Parinet,[1,5] Mathieu Leclerc,[1,5] Sébastien Maury,[1,5] Eric Solary,[6] Dorothée Selimoglu-Buet,[6] Nicolas Duployez,[4] Orianne Wagner-Ballon[1,2#] and Ivan Sloma[1,2#]

[1]Université Paris Est Créteil, INSERM, IMRB, Créteil; [2]AP-HP, Hôpital Henri Mondor, Département d'Hématologie et Immunologie, Créteil; [3]Centre Hospitalier Régional Universitaire de Lille, Laboratoire d'Hématologie, Lille; [4]Université Paris-Sud, Faculté de Médecine Kremlin Bicêtre, INSERM UMS 33, Villejuif; [5]AP-HP, Hôpital Henri Mondor, Département d'Hématologie, Créteil and [6]INSERM Unité Mixte de Recherche (UMR) 1287, Faculté de Médecine, Université Paris-Sud, Gustave Roussy, Villejuif, France

[+]BP and CD contributed equally.
[#]OW-B and IS contributed equally as senior authors.

Correspondence:
I. SLOMA - ivan.sloma@aphp.fr

https://doi.org/10.3324/haematol.2022.282528

TET2 mutational status affects myelodysplastic syndrome evolution to chronic myelomonocytic leukemia

Violaine Tran Quang[1,2], Benjamin Podvin[3]* , Christophe Desterke[4]* , Sihem Tarfi[1,2], Quentin Barathon[2], Bouchra Badaoui[2], Nicolas Freynet[2], Vincent Parinet[1,5], Mathieu Leclerc[1,5], Sébastien Maury[1,5], Eric Solary[6], Dorothée Selimoglu-Buet[6], Nicolas Duployez[4], Orianne Wagner-Ballon[1,2]* and Ivan Sloma[1,2]*

[1]Univ Paris Est Créteil, INSERM, IMRB, F-94010 Créteil, France
[2]AP-HP, Hôpital Henri Mondor, Département d'Hématologie et Immunologie, F-94010 Créteil, France
[3] Centre hospitalier régional Universitaire de Lille, Laboratoire d'hématologie, F-59000 Lille, France
[4] Université Paris-Sud, Faculté de Médecine Kremlin Bicêtre, INSERM UMS 33 Villejuif, France
[5]AP-HP, Hôpital Henri Mondor, Département d'Hématologie, F-94010 Créteil, France
[6]INSERM Unité Mixte de Recherche (UMR) 1287, Faculté de Médecine, Université Paris-Sud, Gustave Roussy, Villejuif, France

Supplementary Data

Supplementary Table legends

Supplementary Table 1. Clinical and biological characteristics of MDS patients of training and validation cohorts.

MDS patients from training and validation cohorts have been divided into two groups according to *TET2* mutational profile. Statistical analysis was performed with a Student t-test with Welch's correction unless otherwise specified: [a] Fisher's exact test – [b] Chi-square test. *WBC: White Blood Count – ANC: Absolute Neutrophil Count – AMC: Absolute Monocyte Count – cMO: classical Monocytes – IPSS-R: Revised International Prognostic Scoring System.*

Supplementary Table 2. Mutations identified in CMML and MDS patients.

List of mutations identified by NGS in CMML and MDS patients. Samples from the learning cohort were analyzed with an NGS panel covering 38 genes. The validation cohort was analyzed by the same panel (n=23) or a panel covering 36 genes (n=16) with an overlap including 30 genes.

Supplementary Figure legends

Supplementary Figure 1. Mutational landscape of MDS patients in the validation cohort.

A. Thirty-nine MDS patients were genotyped. Each column represents a patient. The first line shows cytogenetic data: normal karyotype (light blue square), complex karyotype (dark blue square), and other cytogenetic profiles (blue square). The second line represents prognosis according to the IPSS-R score for MDS patients and CPSS score for CMML patients: low-risk profile (IPSS-R ≤ 3 and CPSS ≤ 1, light purple square) or high-risk profile (IPSS-R >3 and CPSS ≥ 2, dark purple square). In this heatmap, each line represents one gene: absence of mutation is shown by a light grey square, one mutation by a yellow square, and multiple mutations by a red square. Multiple mutations are defined by either several gene mutations or a single mutation with a variant allelic frequency (VAF) above 65%. Analysis not performed or uninterpretable results are shown as a dark grey square. B. Lollipop diagram showing mutation distribution across the TET2 protein identified in MDS patients included in the validation cohort. Thirty-seven mutations from 21 patients are plotted (nonsense mutations are shown in black, frameshift mutations in pink, and missense mutations in green).

Supplementary Table 1

	MDS training cohort				MDS validation cohort			
	Total	Multiple *TET2* mutations	≤ 1 *TET2* mutation	p	Total	Multiple *TET2* mutations	≤ 1 *TET2* mutation	p
Patients, n (%)	44	12 (27.3)	32 (72.7)		39	14 (35.9)	25 (64.1)	
Age, y	72.0 ± 1.8	79 ± 2	72 ± 2	≤ 0.05	75 ± 2	74 ± 3	75 ± 2	0.8584
Male, n (%)	26 (55.3)	8 (66.7)	18 (56.3)	0.7328[a]	26 (66.7)	12 (85.7)	14 (56.0)	≤ 0.05[a]
Haemoglobin, g/dL	10.9 ± 0.3	11.1 ± 0.6	10.9 ± 0.4	0.7206	11.1 ± 0.3	12.0 ± 0.6	10.6 ± 0.4	0.0942
Platelets, $x10^9$/L	178 ± 19	136 ± 32	193 ± 24	0.1635	192 ± 17	149 ± 26	215 ± 22	0.0629
WBC, $x10^9$/L	4.7 ± 0.4	7.4 ± 1.1	5.9 ± 0.6	0.4380	4.6 ± 0.4	3.8 ± 0.6	5.0 ± 0.5	0.1114
ANC, $x10^9$/L	2.6 ± 0.3	2.2 ± 0.5	2.8 ± 0.3	0.2768	2.5 ± 0.3	1.5 ± 0.4	3.0 ± 0.4	≤ 0.05
AMC, $x10^9$/L	0.5 ± 0.0	0.6 ± 0.1	0.4 ± 0.0	≤ 0.01	0.6 ± 0.0	0.7 ± 0.1	0.6 ± 0.1	0.3992
Blood monocytes, %	10.6 ± 0.9	15.8 ± 1.8	8.7 ± 0.9	≤ 0.01	14.4 ± 1.3	19.6 ± 2.3	11.7 ± 1.2	≤ 0.01
Marrow monocytes, %	4 ± 1	8 ± 2	3 ± 0	≤ 0.01	4 ± 1	6 ± 1	3 ± 0	≤ 0.01
Marrow blasts, %	7 ± 1	7 ± 1	6 ± 1	0.3870	4 ± 0	5 ± 1	4 ± 1	0.1665
Mean MO1 fraction, %	88.0 ± 1.5	95.5 ± 0.8	86.7 ± 1.9	≤ 0.001				
Patients with cMo ≥ 94%, n (%)	18 (40.9)	10 (83.3)	8 (25)	≤ 0.01[a]				
IPSS-R (%)	38 (86.4)	11 (91.7)	28 (87.5)		32 (82.1)	11 (78.6)	21 (84.0)	
Very low IPSS-R, n (%)	5 (12.5)	0 (0)	5 (17.9)		7 (21.9)	1 (9.1)	6 (28.6)	
Low IPSS-R, n (%)	19 (47.5)	3 (27.3)	14 (50)		17 (53.1)	7 (63.6)	10 (47.6)	
Intermediate IPSS-R, n (%)	9 (22.5)	6 (57.6)	5 (17.9)	0.1244[b]	6 (18.8)	3 (27.3)	3 (14.3)	0.6541[b]
High IPSS-R, n (%)	5 (12.5)	2 (18.1)	2 (7.1)		2 (6.2)	0 (0)	2 (9.5)	
Very high IPSS-R, n (%)	2 (5.0)	0 (0)	2 (7.1)		0 (0)	0 (0)	0 (0)	

Supplementary Table 2

Patient ID	Cohort	Diagnosis (WHO classification 2017)	Mutations
UPN_1	Learning	CMML-1	TET2 NM_001127208 :c.4553C>G - p.Ser1518* (86.8%) SRSF2 NM_003016.4 :c.284_307del - p.Pro95_Arg102del (35%) ASXL1 NM_015338.5 :c.1934dupG - p.Gly646fs (24.6%)
UPN_2	Learning	CMML-2	TET2 NM_001127208:c.2689C>T - p.Gln897* (47.7%) TET2 NM_001127208:c.3594+2T>C - splice (47.2%) SRSF2 NM_003016.4 :c.284C>A - p.Pro95His ASXL1 NM_015338.5 :c.2274_2301dup - p.Gln768fs (14.6%)
UPN_3	Learning	CMML-0	TET2 NM_001127208:c.4393C>T - p.Arg1465* (44.3%) TET2 NM_001127208:c.5172del - p.Thr1726fs (46%) ASXL1 NM_015338.5:c.1762C>T - p.Gln588* (46.6%)
UPN_4	Learning	CMML-1	SRSF2 NM_003016.4:c.284C>G - p.Pro95Arg (42.8%) TET2 NM_001127208.2:c.2784del - p.Pro929fs (44.7%) TET2 NM_001127208.2:c.4944del - p.Tyr1649fs (47.3%)
UPN_5	Learning	CMML-2	SRSF2 NM_003016.4:c.284C>A - p.Pro95His (45.5%) TET2 NM_001127208.2:c.945del - p.Gln317fs (51.4%) TET2 NM_001127208.2:c.5219_5246dup - p.Met1749fs (15.5%)
UPN_6	Learning	CMML-1	SRSF2 NM_003016.4:c.284C>T - p.Pro95Leu (46.9%) TET2 NM_001127208.2: c.5393C>G - p.Ser1798* (45.7%) TET2 NM_001127208.2:c.5508_5521delinsA - p.Ala1837fs (45.1%) RUNX1 NM_001754.4:c.320G>A - p.Arg107His (48%)
UPN_7	Learning	CMML-1	TET2 NM_001127208.2:c.600_601insG - p.Lys201fs (30.3%) TET2 NM_001127208.2:c.5273C>G - Ser1758* (47.4%) SRSF2 NM_003016.4:c.130T>C - p.Tyr44His (2.8%)
UPN_8	Learning	CMML-0	SF3B1 NM_012433.3:c.1986C>G - p.His662Gln (44.6%) TET2 NM_001127208.2:c.2290dupC - p.Gln764fs (21.3%) TET2 NM_001127208.2:c.5038C>T - p.Gln1680* (5.5%) TET2 NM_001127208.2:c.4827delG - p.Asn1610fs (2.3%) ETNK1 NM_018638.4:c.734G>A - p.Gly245Asp (4%)
UPN_9	Learning	CMML-2	TET2 NM_001127208:c.1970C>A - p.Ser657* (68.5%) TET2 NM_001127208:c.5618T>C - p.Ile1873Thr (3.4%) SF3B1 NM_012433.3:c.1998G>T - p.Lys666Asn (3.2%) KRAS NM_004985(ex2-3):c.149C>T - p.Thr50Ile (3.9%)
UPN_10	Learning	CMML-1	TET2 NM_001127208.2:c.380_381dupT - p.Ser128fs (3.4%) TET2 NM_001127208.2:c.2290dupC - p.Gln764fs (5.4%) TET2 NM_001127208.2:c.2539C>T - p.Gln847* (46.4%) TET2 NM_001127208.2 :c.3356T>G - p.Leu1119* (1.2%) TET2 NM_001127208.2:c.3593G>A - p.Trp1198* (30%) TET2 NM_001127208.2:c.4546C>T - p.Arg1516* (2.4%) ZRSR2 NM_005089.3:c.1207delA - p.Arg403fs (85.7%)
UPN_11	Learning	CMML-2	ASXL2 NM_018263.4:c.1939_1946delACTAGTCC - p.Thr647fs (1%) TET2 NM_001127208.2:c.1249C>T - p.Gln417* (33.3%) TET2 NM_001127208.2:c.2484C>A - p.Cys828* (1.4%) TET2 NM_001127208.2:c.3796A>C - p.Asn1266His (29.8%)
UPN_12	Learning	CMML-1	TET2 NM_001127208.2:c.822del - p.Asn275fs (33.5%) TET2 NM_001127208.2:c.1639G>T - p.Glu547* (44.4%) TET2 NM_001127208.2:c.4139A>G - p.His1380Arg (10.3%) TET2 NM_001127208.2:c.5152G>T - p.Val1718Leu (49.3%)
UPN_13	Learning	CMML-2	TET2 NM_001127208.2:c.3696delG - p.Trp1233fs (51.7%) TET2 NM_001127208.2:c.5712T>A - p.His1904Gln (23.7%)
UPN_14	Learning	CMML-1	TET2 NM_001127208: c.685dupA - p.Thr229fs (20%) TET2 NM_001127208: c.3260_3263del - p.Ser1087fs (2.9%)
UPN_15	Learning	CMML-1	ASXL1 NM_015338.5:c.2242C>T - p.Gln748* (44.6%) CBL NM_005188.3:c.1259G>A - p.Arg420Gln (7.8%) TET2 NM_001127208.2:c.4263C>G - p.Tyr1421* (44.6%)

Supplementary Table 2

Patient ID	Cohort	Diagnosis (WHO classification 2017)	Mutations
UPN_16	Learning	CMML-0	ASXL1 NM_015338.5:c.2332C>T - p.Gln778* (47.4%) CBL NM_005188.3:c.1259G>A - p.Arg420Gln (1.6%) IDH2 NM_002168.3:c.419G>A - p.Arg140Gln (45.5%) SRSF2 NM_003016.4:c.284C>A - p.Pro95His (44.6%)
UPN_17	Learning	CMML-2	ASXL1 NM_015338.5:c.2846delG - p.Gly949fs (47.4%) CBL NM_005188.3:c.1259G>A - p.Arg420Gln (3.4%) NRAS NM_002524.4:c.35G>A - p.Gly12Asp (43.7%) SETBP1 NM_015559.2:c.2608G>A - p.Gly870Ser (47%)
UPN_18	Learning	CMML-1	ASXL1 NM_015338.5:c.1934dupG - p.Gly646fs (26.7%) ETNK1 NM_018638.4:c.731A>G - p.Asn244Ser (83.6%) SETBP1 NM_015559.2:c.2602G>A - p.Asp868Asn (48.8%) JAK2 NM_004972.3:c.3323A>G - p.Asn1108Ser (47.3%)
UPN_19	Learning	CMML-2	DNMT3A NM_022552.4:c.2644C>T - p.Arg882Cys (42.5%) TP53 NM_000546.5:c.404G>A - p.Cys135Tyr (72.8%) U2AF1 NM_006758.2:c.471G>T - p.Gln157His (42%)
UPN_20	Learning	CMML-0	MPL NM_005373.2:c.1543T>C - p.Trp515Arg (1.1%) SETBP1 NM_015559.2:c.2608G>A - p.Gly870Ser (18.9%) U2AF1 NM_006758.2:c.470A>C - p.Gln157Pro (45.2%)
UPN_21	Learning	CMML-1	KRAS NM_004985.4:c.34G>T - p.Gly12Cys (38.3%)
UPN_22	Learning	CMML-0	No mutation
UPN_23	Learning	MDS-SLD	SRSF2 NM_003016.4:c.284C>G - p.Pro95Arg (47.5%) ASXL2 NM_018263.4:c.3042_3043insTGGA - p.Ala1015fs (24.9%) CBL NM_005188.3:c.1139T>C - p.Leu380Pro (46.1%) TET2 NM_001127208.2:c.4138C>T - p.His1380Tyr (96.5%) ASXL2 NM_018263.4:c.1489G>A - p.Ala497Thr (46.6%) ETNK1 NM_018638.4:c.742C>A - p.Pro248Thr (46.8%)
UPN_24	Learning	MDS-EB1	TET2 NM_001127208:c.2884C>T - p.Gln962* (44.8%) TET2 NM_001127208:c.2290dupC - p.Gln764fs (30.2%) SRSF2 NM_003016.4:c.284C>T - p.Pro95Leu (41.2%) TP53 NM_000546.5:c.826G>C - p.Ala276Pro (40.9%)
UPN_25	Learning	MDS-EB1	TET2 NM_001127208:c.4532T>A - p.Leu1511* (48%) TET2 NM_001127208:c.521dupC - p.Glu175fs (23.4%) TET2 NM_001127208:c.2068C>T - p.Gln690* (11.7%) SRSF2 NM_003016.4 :c.284C>T - p.Pro95Leu (44.6%)
UPN_26	Learning	MDS-EB1	TET2 NM_001127208.2:c.3530T>G - p.Ile1177Ser (86.7%) SF3B1 NM_012433.3:c.1986C>G - p.His662Gln (35.2%)
UPN_27	Learning	MDS-RS-MLD	TET2 NM_001127208 :c.3764dupA - p.Tyr1255fs (48.2) TET2 NM_001127208 :c.3823G>A - p.Gly1275Arg (48.2%) SF3B1 NM_012433.3 :c.2098A>G - p.Lys700Glu (39.8%)
UPN_28	Learning	MDS-EB1	TET 2 NM_001127208.2: c.2628delA - p.Asp877fs (18.7%) TET2 NM_001127208.2: c.5361delT - p.Asn1787fs (27.9%) ZRSR2 NM_005089.3:c.820T>G - p.Tyr274Asp (12.5%) ZRSR2 NM_005089.3: c.935G>A - p.Cys312Tyr (39.6%)
UPN_29	Learning	MDS-MLD	TET2 NM_001127208.2:c.2239A>T - p.Lys747* (30%) TET2 NM_001127208.2:c.2148dupA - p.His717f (22.5%) ZRSR2 NM_005089.3:c.709delC - p.Leu237fs (59%)
UPN_30	Learning	MDS-EB1	TET2 NM_001127208.2:c.2944A>T - :p.Lys982* (46.7%) TET2 NM_001127208.2:c.3739G>T - p.Glu1247* (45.2%) U2AF1 NM_006758.2:c.101C>T - Ser34Phe (43.5%)
UPN_31	Learning	MDS-EB2	TET2 NM_001127208.2:c.800dupC - p.Ser268fs (25.3%) TET2 NM_001127208.2:c.5666C>G - p.Pro1889Arg (28.8%)
UPN_32	Learning	MDS-EB1	TET2 NM_001127208:c.3263C>G - p.Ser1088* (49.3%) TET2 NM_001127208:c.3333dupA - p.Leu1112fs (22.1%)

Supplementary Table 2

Patient ID	Cohort	Diagnosis (WHO classification 2017)	Mutations
UPN_33	Learning	MDS-RS-MLD	ASXL1 NM_015338.5:c.4243C>T - p.Arg1415* (36.7%) ASXL1 NM_015338.5:c.3824C>A - p.Ser1275* (5.1%) MPL NM_005373.2:c.1543T>C - p.Trp515Arg (1.8%) SF3B1 NM_012433.3:c.1997A>C - p.Lys666Thr (44%) TET2 NM_001127208.2 - Variant1 :c.5630delA - p.Lys1877fs (35.1%)
UPN_34	Learning	MDS-RS-MLD	TET2 NM_001127208.2:c.444_447delAGAA - p.Lys148fs (46.8%) SF3B1 NM_012433.3:c.1866G>C - p.Glu622Asp (39%) DNMT3A NM_022552.4:c.2553delC - p.Phe851fs (35.2%)
UPN_35	Learning	MDS-MLD	ASXL1 NM_015338.5:c.1934dupG - p.Gly646fs (19.8%) RUNX1 NM_001754.4:c.958C>T - p.Arg320* (5%) SRSF2 NM_003016.4:c.284C>A - p.Pro95His (1.9%) ZRSR2 NM_005089.3:c.294delA - p.Lys98fs (37%) ZRSR2 NM_005089.3 - Variant2 :c.812A>G - p.Tyr271Cys (4.2%)
UPN_36	Learning	MDS-EB2	ASXL2 NM_018263.4:c.2566G>T - p.Glu856* (44.2%) ETV6 NM_001987.4:c.403delC - p.His135fs (45.6%) SETBP1 NM_015559.2:c.2608G>A - p.Gly870Ser (42.3%) SRSF2 NM_003016.4:c.284C>G - p.Pro95Arg (60.6%)
UPN_37	Learning	MDS-EB1	ASXL1 NM_015338.5 :c.1534C>T - p.Gln512* (31.1%) U2AF1 NM_006758.2 :c.470A>C - p.Gln157Pro (28.7%) SETBP1 NM_015559.2:c.2603A>G - p.Asp868Gly (2.1%) SETBP1 NM_015559.2:c.2608G>A - p.Gly870Ser (1.1%)
UPN_38	Learning	MDS-EB2	TP53 NM_000546.5:c.838A>G - p.Arg280Gly (80.8%)
UPN_39	Learning	MDS-EB1	TP53 NM_000546.5:c.916C>T - p.Arg306* (47.4%) TP53 NM_000546.5:c.584T>C - p.Ile195Thr (50.2%)
UPN_40	Learning	MDS-EB1	RUNX1 NM_001754:exon8:c.C958T:p.R320X (11%) STAG2 NM_001042750:exon28:c.A2845T:p.K949 (10%) STAG2 NM_001042750:exon30:c.C3097T:p.R1033X (6.5%) STAG2 NM_001042750:exon8:c.624dupT:p.S208fs (4.3%)
UPN_41	Learning	MDS-EB2	TET2 NM_001127208:c.5618T>C - p.Ile1873Thr (90.6%) EZH2 NM_004456:c.2005A>G - p.Ser669Gly (90.8%) ASXL1 NM_015338.5:c.1772dupA -p.Tyr591fs (29.3%) RUNX1 NM_001754:c.1026_1027insA - p.Ser343fs (26.3%) ZRSR2 NM_005089.3:c.827+2dupT - splice (44.5%) JAK2 NM_004972.3:c.3323A>G - p.Asn1108Ser (1.8%)
UPN_42	Learning	MDS-MLD	TET2 NM_001127208.2:c.5152G>T - p.Val1718Leu (49.4%) TET2 NM_001127208.2:c.4082G>C - p.Gly1361Ala (4.8%) TET2 NM_001127208.2:c.3954+1G>T - splice (1.5%)
UPN_43	Learning	MDS-RS-MLD	DNMT3A NM_022552.4:c.2173+1G>A (30.2%) - splice SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (24.3%) TET2 NM_001127208.2:c.2290dupC - p.Gln764fs (16.1%) JAK2 NM_004972.3:c.3188G>A - p.Arg1063His (48.1%)
UPN_44	Learning	MDS-RS-MLD	SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (44.3%) TET2 NM_001127208.2:c.2872C>T - p.Gln958* (44.3%)
UPN_45	Learning	MDS-MLD	DNMT3A NM_022552.4 :c.2645G>A - p.Arg882His (38%) TET2 NM_001127208 :c.967C>T - p.Gln323* (32.5%)
UPN_46	Learning	MDS-EB1	DNMT3A NM_022552.4:c.976C>T - p.Arg326Cys (12.8%) TET2 NM_001127208.2:c.5618T>C - p.Ile1873Thr (2.9%)
UPN_47	Learning	MDS-EB1	CBL NM_005188.3:c.1311delA - p.Ser439fs (41%) CBL NM_005188.3:c.1211G>A - p.Cys404Tyr (2.5%) TET2 NM_001127208.2:c.2479_2489dupGCATGCAAAAT - p.Ile830fs (42.8%)
UPN_48	Learning	MDS-EB1	TET2 NM_001127208.2:c.4461_4464delTAAA - p.Asn1487fs (34.1%)
UPN_49	Learning	MDS-RS-MLD	TET2 NM_001127208.2:c.3689T>G - p.Ile1230Ser (32.5%)
UPN_50	Learning	MDS-EB2	ASXL1 NM_015338.5:c.1934dupG - p.Gly646fs (25%) RUNX1 NM_001754.4:c.808dupA - p.Thr270fs (30%) SRSF2 NM_003016.4:c.284C>A - p.Pro95His (48.2%) STAG2 NM_001042750.1:c.1216dupA - p.Thr406fs (32.9%)

Supplementary Table 2

Patient ID	Cohort	Diagnosis (WHO classification 2017)	Mutations
UPN_51	Learning	MDS-EB2	SF3B1 NM_012433.3:c.1998G>T - p.Lys666Asn (26%) SRSF2 NM_003016.4:c.284C>T - p.Pro95Leu (11.5%)
UPN_52	Learning	MDS-RS-MLD	IDH2 NM_002168.3:c.419G>A - p.Arg140Gln (34.9%) KRAS NM_004985.4:c.64C>A - p.Gln22Lys (1.3%) SRSF2 NM_003016.4:c.284C>T - p.Pro95Leu (35.5%)
UPN_53	Learning	MDS-EB2	IDH2 NM_002168.3 : c.419G>A - p.Arg140Gln (43.8%) SRSF2 NM_003016.4 : c.284C>T - p.Pro95Leu (43.1%)
UPN_54	Learning	MDS-EB1	SF3B1 NM_012433.3:c.1873C>T - p.Arg625Cys (39.3%)
UPN_55	Learning	MDS-RS-MLD	SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (24.9%) PPM1D NM_003620.3:c.1637del - p.Leu546fs (3.7%) GATA2 NM_032638 (ex4-5) :c.1096G>C - p.Gly366Arg (3.3%)
UPN_56	Learning	MDS-EB1	JAK2 NM_004972.3:c.1849G>T - p.Val617Phe (2.3%) DNMT3A NM_022552.4:c.2385G>A - p.Trp795* (1.5%) DNMT3A NM_022552.4:c.1077C>A - p.Tyr359* (1.3%)
UPN_57	Learning	MDS with del(5c	DNMT3A NM_022552.4:c.2390A>G - p.Asn797Ser (2.6%)
UPN_58	Learning	MDS-EB2	TP53 NM_000546.5:c.814G>A - p.Val272Met (15.4%) ASXL1 NM_015338.5: c.1945G>T p.Gly649* (2.5%)
UPN_59	Learning	MDS-MLD	ZRSR2 NM_005089.3:c.328C>T - p.Gln110* (29.4%) ZRSR2 NM_005089.3:c.92G>C - p.Arg31Pro (21.7%)
UPN_60	Learning	MDS-EB1	EZH2 NM_004456.4:c.2197T>G - p.Tyr733Asp (12.5%)
UPN_61	Learning	MDS-EB1	No mutation
UPN_62	Learning	MDS-EB2	No mutation
UPN_63	Learning	MDS-EB1	No mutation
UPN_64	Learning	MDS-EB1	No mutation
UPN_65	Learning	MDS-SLD	No mutation
UPN_66	Learning	MDS-MLD	No mutation
UPN_67	Validation	MDS	TET2 NM_001127208:c.507_508delTA - p.His169fs (88%) ASXL1 NM_015338.5:c.1585C>T - p.Gln529* (44%) SRSF2 NM_003016.4:c.284C>A - p.Pro95His (43%) KIT NM_000222.2:c.2447A>T - p.Asp816Val (38%) JAK2 NM_004972.3:c.1849G>T - p.Val617Phe (2%) TET2 NM_001127208:c.3594+5G>A - splice (2%)
UPN_68	Validation	MDS-EB1	TET2 NM_001127208: c.5582G>A - p.Gly1861Glu (92%) SRSF2 NM_003016.4:c.284C>A - p.Pro95His (47%) CBL NM_005188.3:c.1380_1382dupTGA - p.Asp460dup (32%) ASXL1 NM_015338.5:c.2077C>T - p.Arg693* (3%)
UPN_69	Validation	MDS-MLD	ASXL1 NM_015338.5 :c.2101_2105del - p.Pro701fs (20.7%) TET2 NM_001127208 :c.5618T>C - p.Ile1873Thr (11.8%) TET2 NM_001127208:c.3621_3658del - p.Glu1207fs (9.2%) TET2 NM_001127208:c.4546C>T - p.Arg1516* (2.4%) JAK2 NM_004972.3 :c.1849G>T - p.Val617Phe (5.2%)
UPN_70	Validation	MDS-MLD	TET2 NM_001127208:c.1947_1948delAC - p.His650fs (25%) SF3B1 NM_012433.3:c.1986C>G - p.His662Gln (37%) TET2 NM_001127208:c.2862G>A - p.Trp954* (42%) JAK2 NM_004972.3:c.1849G>T - p.Val617Phe (12%)
UPN_71	Validation	MDS-MLD	SRSF2 NM_003016.4 :c.284C>G - p.Pro95Arg (44.1%) TET2 NM_001127208:c.3151C>T - p.Gln1051* (52.1%) TET2 NM_001127208:c.2862G>A - p.Trp954* (22.9%) TET2 NM_001127208:c.2290dupC - p.Gln764fs (13.1%) JAK2 NM_004972.3 :c.1849G>T - p.Val617Phe (0.9%) CBL NM_005188.3:c.1139T>C - p.Leu380Pro (5.3%) NRAS NM_002524.4:c.35G>A - p.Gly12Asp (0.7%) NRAS NM_002524.4 :c.190T>G - p.Tyr64Asp (1.5%) CBL NM_005188.3 :c.1255T>C - p.Cys419Arg (1%)

Supplementary Table 2

Patient ID	Cohort	Diagnosis (WHO classification 2017)	Mutations
UPN_72	Validation	MDS-MLD	SRSF2 NM_003016.4 : c.284C>A - p.Pro95His (23.2%) TET2 NM_001127208:c.4546C>T - p.Arg1516* (14.1%) TET2 NM_001127208:c.4393C>T - p.Arg1465* (4.6%)
UPN_73	Validation	MDS-EB1	TET2 NM_001127208:c.2474del - p.Ser825fs (52.3%) TET2 NM_001127208:c.2245C>T - p.Gln749* (38.1%) SRSF2 NM_003016.4 :c.43_44inv - p.Ser15Asp (35.6%)
UPN_74	Validation	MDS-EB1	TET2 NM_001127208:c.2253_2256dupGAAT - p.Lys753fs (44.6%) TET2 NM_001127208: c.2746C>T - p.Gln916* (35.8%) SRSF2 NM_003016.4 : c.284C>T - p.Pro95Leu (3.7%)
UPN_75	Validation	MDS-EB2	ZRSR2 NM_005089.3:c.205C>T - p.Gln69* (50%) TET2 NM_001127208:c.3997dupA - p.Met1333fs (46%) TET2 NM_001127208:c.945delC - p.Gln317fs (43%) ZRSR2 NM_005089.3:c.1207delA - p.Arg403fs (28%) CBL NM_005188.3: c.1096-7A>G - splice (2%)
UPN_76	Validation	MDS-SLD	TET2 NM_001127208:c.3869C>G - p.Ser1290* (46%) TET2 NM_001127208:c.5618T>C - p.Ile1873Thr (28%) TET2 NM_001127208:c.5538delC - p.Trp1847fs (14%) ZRSR2 NM_005089.3:c.204-1_246dup - p.Gln83fs (9%) ZRSR2 NM_005089.3:c.515G>T - p.Cys172Phe (2%) NRAS NM_002524.4: c.35G>A - p.Gly12Asp (1%)
UPN_77	Validation	MDS-SLD	TET2 NM_001127208:c.3443A>G - p.Tyr1148Cys (75%) ZRSR2 NM_005089.3:c.710T>C - p.Leu237Pro (24%) ZRSR2 NM_005089.3:c.572A>G - p.His191Arg (13%) ZRSR2 NM_005089.3:c.868C>T - p.Arg290* (17%) ZRSR2 NM_005089.3: c.883C>T - p.Arg295* (5%)
UPN_78	Validation	MDS-MLD	TET2 NM_001127208 :c.2457T>A - p.Tyr819* (41.8%) ZRSR2 NM_005089.3 :c.399+2T>G - splice (50%) TET2 NM_001127208 :c.3481A>G - p.Arg1161Gly (3.5%)
UPN_79	Validation	MDS-EB1	ZRSR2 NM_005089.3 :c.952C>T - p.Gln318* (73.2%) TET2 NM_001127208 :c.3368del - p.Pro1123fs (36.8%) TET2 NM_001127208 :c.2290dupC - p.Gln764fs (35.5%) TET2 NM_001127208 :- Variant3 : c.3824G>C - p.Gly1275Ala (1.1%)
UPN_80	Validation	MDS	TET2 NM_001127208:c.4075C>T - p.Arg1359Cys (34%) TET2 NM_001127208:c.2243T>G - p.Leu748* (32%)
UPN_81	Validation	MDS-EB1	SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (33.1%) ASXL1 NM_015338.5:c.2324del - p.Leu775fs (35.4%) ASXL1 NM_015338.5:c.1934dupG - p.Gly646fs (1.5%) ASXL1 NM_015338.5:c.3155dupT - p.Val1053fs (1.2%) TET2 NM_001127208 : c.4567C>T -p.Gln1523* (1%)
UPN_82	Validation	MDS-MLD	ZRSR2 NM_005089.3: c.859delT - p.Phe287Leu* (50%) TET2 NM_001127208: c.5582G>A - p.Gly1861Glu (37%) EZH2 NM_004456.4:c.1870T>C - p.Ser624Pro (23%) ASXL1 NM_015338.5:c.2385delC - p.Trp796fs (19%)
UPN_83	Validation	MDS-RS-MLD	SF3B1 NM_012433.3:c.1998G>C - p.Lys666Asn (48.6%) DNMT3A NM_022552.4:c.2086C>T - p.Gln696* (47%) TET2 NM_001127208:c.3823G>C - p.Gly1275Arg (1.6%)
UPN_84	Validation	MDS-RS-MLD	SF3B1 NM_012433.3 :c.2098A>G - p.Lys700Glu (45.5%) TET2 NM_001127208 :c.2872C>T - p.Gln958* (47.2%)
UPN_85	Validation	MDS-RS-MLD	SF3B1 NM_012433.3 :c.2098A>G - p.Lys700Glu (32.6%) TET2 NM_001127208 :c.5602C>T - p.His1868Tyr (2.5%)
UPN_86	Validation		TET2 NM_001127208:c.4087_4088delAAinsG - p.Lys1363fs* (46%) SRSF2 NM_003016.4:c.284C>A - p.Pro95His (32%)
UPN_87	Validation	MDS-EB1	DNMT3A NM_022552.4:c.2339T>C - p.Ile780Thr (41.7%) IDH2 NM_002168.3:c.419G>A - p.Arg140Gln (40.2%) TET2 NM_001127208:c.5081dupT - p.Leu1694fs (1.0%)

Supplementary Table 2

Patient ID	Cohort	Diagnosis (WHO classification 2017)	Mutations
UPN_88	Validation	MDS-EB1	SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (31%) ASXL1 NM_015338.5:c.3796del - p.Thr1267fs (5.9%) DNMT3A NM_022552.4:c.1846G>T - p.Glu616* (1.3%) DNMT3A NM_022552.4:c.1507dupA - p.Thr503fs (1%)
UPN_89	Validation	MDS-RS-MLD	SF3B1 NM_012433.3 - Variant1 :c.1986C>A - p.His662Gln (39.8%) ASXL1 NM_015338.5 - Variant1 :c.1934dupG - p.Gly646fs (11%)
UPN_90	Validation	MDS-RS-MLD	ASXL1 NM_015338.5:c.3202C>T - p.Arg1068* (44.1%) IDH2 NM_002168.3:c.418C>T - p.Arg140Trp (44.3%) SRSF2 NM_003016.4:c.284_307del - p.Pro95_Arg102del (38%) STAG2 NM_001042750.1 - Variant1 :c.288+2dupT - splice (43.8%)
UPN_91	Validation	MDS-SLD	ASXL1 NM_015338.5:c.2293delC - p.Leu765fs (8%) CBL NM_005188.3:c.1139T>C - p.Leu380Pro (3%) CBL NM_005188.3:c.1211G>A - p.Cys404Tyr (3%) CBL NM_005188.3:c.1291G>C - p.Val431Leu (2%)
UPN_92	Validation	MDS-MLD	ASXL1 NM_015338.5 :c.4460C>A - p.Ser1487* (29.2%) PPM1D NM_003620.3 :c.1535dupA - p.Asn512fs (26.5%)
UPN_93	Validation	MDS-MLD	ASXL1 NM_015338.5:c.1934dupG - p.Gly646fs (25.7%)
UPN_94	Validation	MDS-MLD	ASXL1 NM_015338.5:c.1934dupG - p.Gly646fs
UPN_95	Validation	MDS-MLD	ASXL1 NM_015338.5 :c.1934dupG - p.Gly646fs
UPN_96	Validation	MDS-MLD	EZH2 NM_004456.4:c.1625_1626delTG - p.Val542fs (20.1%) PPM1D NM_003620.3:c.1535delA - p.Asn512fs (1.1%) SF3B1 NM_012433.3:c.2098A>G - Lys700Glu (46.3%) TP53 NM_000546.5:c.527G>A - p.Cys176Tyr (1.7%) TP53 NM_000546.5:c.731G>T - p.Gly244Val (18.1%) TP53 NM_000546.5:c.736A>G - p.Met246Val (5.8%) TP53 NM_000546.5:c.743G>A - p.Arg248Gln (2%) TP53 NM_000546.5:c.797G>A - p.Gly266Glu (5.3%) TP53 NM_000546.5:c.832C>T - p.Pro278Ser (2.1%) ETV6 NM_001987.4:c.1022T>C - p.Leu341Pro (6.5%) DNMT3A Variant1 :c.2578T>C - p.Trp860Arg (1.3%)
UPN_97	Validation	MDS-RS-SLD	DNMT3A NM_022552.4:c.2113_2124delinsCTTG - p.Ile705fs (31%) RUNX1 NM_001754:c.404G>T - p.Arg135Met (20%) SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (40%)
UPN_98	Validation	MDS-MLD	DNMT3A NM_022552.4 : c.1103C>T - p.Ala368Val (41%) SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (40%)
UPN_99	Validation	MDS-MLD	IDH2 NM_002168.3:c.419G>A - p.Arg140Gln (45.3%) SF3B1 NM_012433.3:c.2098A>G - p.Lys700Glu (45.2%)
UPN_100	Validation	MDS-MLD	SF3B1 NM_012433.3:c.1874G>T - p.Arg625Leu (35.2%)
UPN_101	Validation	MDS-EB2	JAK2 NM_004972.3 - Variant1 :c.1849G>T - p.Val617Phe (68.2%) IDH1 NM_005896.3 - Variant1 :c.394C>A - p.Arg132Ser (33%) DNMT3A NM_022552.4 - Variant1 :c.1232T>C - p.Leu411Pro (1.1%)
UPN_102	Validation	MDS-MLD	DNMT3A NM_022552.4:c.1903C>T : p.Arg635Trp (1%)
UPN_103	Validation	MDS-MLD	TP53 NM_000546.5:c.713G>A - p.Cys238Tyr (59%) TP53 NM_000546.5:c.1009C>T - p.Arg337Cys (2%)
UPN_104	Validation	MDS-EB1	U2AF1 NM_006758.2 :c.101C>T - p.Ser34Phe (14%)
UPN_105	Validation	MDS-MLD	No mutation

Supplementary Figure 1