# 5-formylcytosine and 5-hydroxymethyluracil as surrogate markers of TET2 and SF3B1 mutations in myelodysplastic syndrome, respectively

*Daniel Gackowski,[1] Maciej Gawronski,[1] Cassandra Kerr,[2] Tomas Radivoyevitch,[3] Ewelina Zarakowska,[1] Marta Starczak,[1] Abdulkadir Abakir,[4] Alexey Ruzov,[4] Jaroslaw P. Maciejewski[2] and Ryszard Olinski[1]*

*[1]Department of Clinical Biochemistry, Faculty of Pharmacy, Nicolaus Copernicus University in Torun, Ludwik Rydygier Collegium Medicum in Bydgoszcz, Poland; [2]Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, USA; [3]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA and [4]Wolfson Centre for Stem Cells, Tissue Engineering and Modelling (STEM), Division of Cancer and Stem Cells, School of Medicine, Centre for Biomolecular Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*

*Correspondence: RYSZARD OLINSKI - ryszardo@cm.umk.pl*

*JAROSLAW P. MACIEJEWSKI - maciejj@ccf.org*

*doi:10.3324/haematol.2019.224030*

**Myelodysplastic syndrome mutation links to epigenetics: 5-formylcytosine is a better marker of *TET2* mutations than 5-hydroxymethylcytosine while 5-hydroxymethyluracilindicates*SF3B1* mutations**

Daniel Gackowski[1], Maciej Gawronski[1], Cassandra Kerr[2], Tomas Radivoyevitch[3], Ewelina Zarakowska[1], Marta Starczak[1], Abdulkadir Abakir[4], Alexey Ruzov[4], Jaroslaw P. Maciejewski[2*], Ryszard Olinski[1*]

[1] Department of Clinical Biochemistry, Faculty of Pharmacy, Nicolaus Copernicus University in Toruń, Ludwik Rydygier Collegium Medicum in Bydgoszcz,

[2] Department of Translational Hematology and Oncology Research, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA,

[3]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH.

[4]Wolfson Centre for Stem Cells, Tissue Engineering and Modelling (STEM), Division of Cancer and Stem Cells, School of Medicine, Centre for Biomolecular Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

# Supplementary table 1

# Materials and methods

# Supplementary figures 1-3

**Supplementary table 1. Overview of studied subjects mutational status**

| No. | Sex | Age | Diagnosis | Cytogenetics | TET2 mutation status (VAF%) | SF mutation status (VAF%) | Other mutations |
|---|---|---|---|---|---|---|---|
| 1 | M | 76 | CMML-1 | 46,XY | p.D1384Y (39%) | non-SFmt | ASXL1 (p.545del, 44%) |
| 2 | M | 62 | pAML | 46,XY[20] | p.R1404X (18.1%) | non-SFmt | None |
| 3 | F | 54 | pAML | 46,XX[20] | p.Y1245C; p.W1917X (40%, 44%) | non-SFmt | DNMT3A (p.R582X, 44%) |
| 4 | F | 67 | sAML | 49-51,XX,+6,+8,+11,+19[cp2]/ 46,XX[31]/ 92,XXXX[2] | p.E1477V (55%) | non-SFmt | ASXL1 (p.P941fs, 64%), PTPN11 (p.G503E, 51%) |
| 5 | F | 64 | MDS-RARS | 47,XX,+8[18]/ 46,XX[2] | p.G1275R (40%) | SF3B1 p.K700E (29%) | none |
| 6 | M | 73 | MDS-RAEB1 | 46,XY[20] | p.E1178fs (47%) | non-SFmt | KIT (p.Q427R, 47%), STAG2 (p.573X, 10%; p.S704X,87% ) |
| 7 | M | 46 | pAML | 46,XY[20] | p.R1465X (77%) | non-SFmt | DNMT3A (p.R693C, 37%), EZH2 (p.Q468R, 8%), NPM1 (p.L258fs, 8%), RUNX1 (p.F89L, 6%) |
| 8 | F | 69 | pAML | 46,XX,t(2;11)(q31;p15)[19]/ 44,XX,-14,-21[1] | p.K1438X (46%) | non-SFmt | RUNX1 (p.V97G, 33%) |
| 9 | M | 69 | pAML | 45,X,-Y,del(20)(q13.1 | p.R550X, p.S716fs (42%, 48%) | non-SFmt | U2AF1 (p.S34F, 43%) |
| 10 | M | 79 | MDS/MPN | 46,XY[20] | p.M611fs (29%) | SF3B1 K666T (43%) | CBL(p.Q302X , 36%), PRPF8 (p.V1015D, 23%) |
| 11 | F | 78 | MDS-RARS | 46,XX,add(16)(p13.3)[11]/ 46,XX,der(16)add(16)(p13.3)add(16)(q24)[4]/ 46,XX,add(16)(p13.3)x2[2]/46,XX[3] | p.T1270fs (39%) | SF3B1 H662Q (56%) | CUX1 (p.I155L, 31%), |
| 12 | M | 56 | MDS RCMD | 46,XY[20] | p.W564fs (43%) | SF3B1 p.K700E (55%) | NPM1 (p.L258F, 8%), |
| 13 | F | 43 | MDS RCMD | 46, XX,del(9)(q21q22),del(12)(p12p13)[6]/ 46,XX,del(7)(q22)[4]/ 46,XX,del(7)(q22),del(9)(q21q22),del(12)(p12p13)[2]/ 46,XX[10] | p.H974fs, pR1465X (35%, 51%) | non-SFmt | ASXL1(p.N842fs, 47%), CEBPA(p.A295E, 47%), RUNX1(p.A160fs, 47%), WT1(p.S364fs, 39%) |
| 14 | M | 60 | MDS RCMD-RS | 46,XY[20] | non-TET2mt | SF3B1 K700E (13%) | none |
| 15 | M | 66 | MDS/MPN | 46,XY,del(5)(q12q33),del(13)(q12q14)[18]/ 46,XY[2]. | non-TET2mt | non-SFmt | None |
| 16 | F | 89 | AA | 46,XX[18] | p.P1092T (21%) | non-SFmt | ASXL1(p.Q1433H, 80%), STAG2(p.I1182F, 17%, p.G1180E, 22%) |
| 17 | F | 64 | AML | 47,XX,inv(16)(p13.1q22),+22[19]/ 46,XX[1] | p.R73L (46%) | non-SFmt | ETV6(p.L117F, 31%) |
| 18 | F | NA | Normal BM | NA | NBM | NBM | NA |
| 19 | M | 44 | Normal BM | NA | NBM | NBM | NA |
| 20 | F | 43 | Normal BM | NA | NBM | NBM | NA |
| 21 | F | 21 | Normal BM | NA | NBM | NBM | NA |
| 22 | M | 53 | Normal BM | NA | NBM | NBM | NA |

# Materials and methods

**Leukemic cell isolation and storage.**

Bone marrow aspirates were collected from patients with myeloid neoplasms according to protocols approved by the Cleveland Clinic IRB and the Declaration of Helsinki. Mononuclear cells were isolated from bone marrow specimens with a density gradient media, frozen in FBS 10% DMSO and stored in liquid nitrogen.

**Cell culture and flow cytometry sorting.**

HUES7 hESCs were maintained in Essential 8™ (E8) medium with supplement (#A1517001) on Matrigel™-coated tissue culture flasks at 37 °C with 5 % CO2. Cells were passaged every 3–4 d using TrypLE™ Select Enzyme (#12563029). hiPSCs were treated with 1:100 dilution of KaryoMAX® Colcemid™ Solution (Thermo Fisher Scientific, catalogue number 15212012) for 3 h.G0/G1 and G2/M phases flow cytometry sorting was performed according to the previously described method[16]. Briefly, enzymatically dissociated hESCs were washed in PBS and fixed in 70 % ethanol for 2 h, washed with PBS again and stained with 10 μg/ml propidium iodide (PI) (Sigma-Aldrich, catalogue number P3566) in PBS supplemented with 0.1 % Triton X-100 and 100 μg/ml RNase A (Qiagen, catalogue number 19101). PI treated hESCs were sorted based on the DNA content into G0/G1 and G2/M cells using Beckman Coulter Astrios EQ and the WEASEL software.

**DNA extraction hydrolysis and mass spectrometry.**

DNA was isolated according to standard phenolic method as described previously[17]. DNA samples were incubated with 1 U of nuclease P1 (Sigma-Aldrich) and tetrahydrouridine (Calbiochem) (cytidine deaminase inhibitor, 10 μg per sample) for 1 h at 37°C followed by addition of 12 μl of 5 % (v/v) $NH_4OH$ (JT Baker) and 1.3 U of alkaline phosphatase (Sigma-Aldrich) and additional 1 h incubation at 37 °C. The DNA hydrolysates were acidified with $CH_3COOH$ (Sigma-Aldrich) to final v/v concentration of 2 % and ultrafiltered prior to injection. The 2D-UPLC–MS/MS analyses were performed according to the method described previously[6]. Briefly, DNA hydrolysates were spiked with a mixture of internal standards in volumetric ratio 4:1, to concentration of 50 fmols/μL of [$D_3$]-5-(hydroxymethyl)-2'-deoxycytidine, [$^{13}C_{10}$, $^{15}N_2$]-5-formyl-2'-deoxycytidine, [$^{13}C_{10}$, $^{15}N_2$]-5-carboxyl-2'-deoxycytidine (5-cadC), [$^{13}C_{10}$, $^{15}N_2$]-5-(hydroxymethyl)-2'-deoxyuridine, and [$^{15}N_5$]-8-oxo-2'-deoxyguanosine. Chromatographic separation was performed with a Waters Acquity 2D-UPLC system with photo-diode array detector, for the first-dimension chromatography (used for quantification of unmodified deoxynucleosides (dN) and 5-methyl-2'-deoxycytidine), and Xevo TQ-S tandem quadrupole mass spectrometer for second dimension chromatography. At-column-dilution technique was used between first and second dimension for improving retention at trap/transfer column. The columns used were: a Phenomenex Kinetex C18 column (150 mm×2.1 mm, 1.7 μm) at the first dimension, a Waters X-select C18 CSH (30 mm×2.1 mm, 1.7 μm) at the second dimension and Waters X-select C18 CSH (30 mm×2.1 mm, 1,7 μm) as trap/transfer column. Chromatographic system operated in heart-cutting mode, indicating that selected parts of effluent from the first dimension were directed to

trap/transfer column via 6-port valve switching, which served as “injector” for the second-dimension chromatography system. The flow rate at the first dimension was 0.25 mL/min and the injection volume was 0.5-2 µL. The separation was performed with a gradient elution for 10 minutes using a mobile phase 0.1% acetate (A) and acetonitrile (B) (1-5% B for 5 minutes, column washing with 30% acetonitrile and re-equilibration with 99% A for 3.6 minutes). Flow rate at the second dimension was 0.35 mL/min. The separation was performed with a gradient elution for 10 minutes using a mobile phase 0.01% acetate (A) and methanol (B) (4-50% B for 4 minutes, isocratic flow of 50% B for 1.5 minutes, and re-equilibration with 96% A up to next injection). All samples were analyzed in three to five technical replicates of which technical mean was used for further calculation. Mass spectrometric detection was performed using the Waters Xevo TQ-S tandem quadrupole mass spectrometer, equipped with an electrospray ionization source. Collision-induced dissociation was obtained using argon 6.0 at 3 x $10^{-6}$ bar pressure as the collision gas. Transition patterns for all the analyzed compounds, as well as specific detector settings were determined using the MassLynx 4.1 Intelli-Start feature.

**Next Generation Sequencing Methods**

DNA extracted from bone marrow specimens were subjected to multi-amplicon targeted deep sequencing for a panel of 60 genes most commonly mutated in hematologic malignancies[18-20.] Paired end libraries were sequenced on the Miseq (Illumina, San Diego CA) according to Illumina protocols. Somatic variants were extracted using the GATK3.3 best practices.

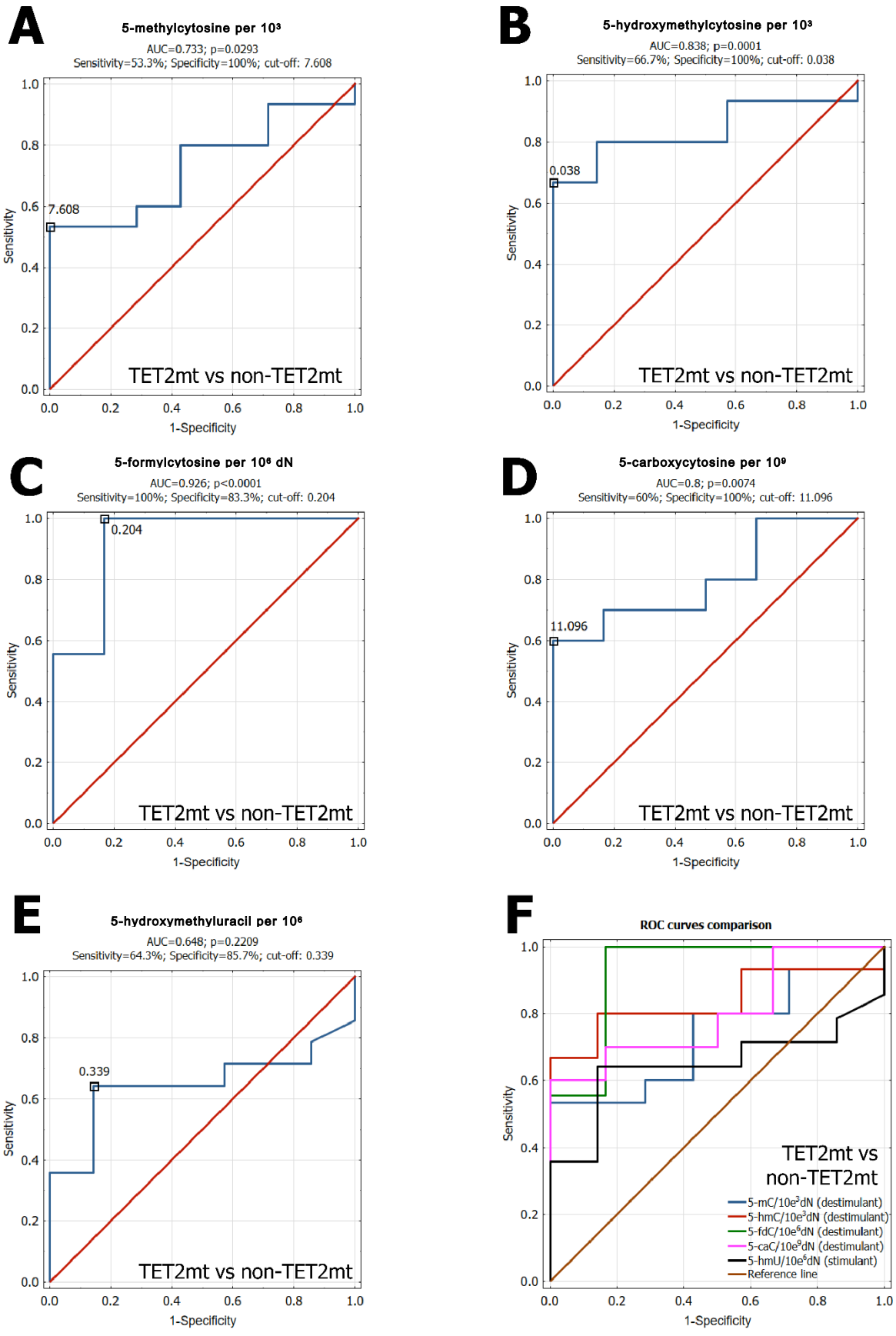
**Statistical methods**

The results are presented as medians, interquartile range and individual data points. Normal distribution of the study variables was verified with Kolmogorov-Smirnov test with Lilliefors correction.One-way analysis of variance (ANOVA) followed by Fisher's least significant difference (LSD) method was used toidentifydifferences between studied groups for variables with normal distribution, while Mann-Whitney U test was used for other variables. Diagnostic power was assessed using ROC curve analysis, allowing determination of cut-off values assuring optimal sensitivity and specificity (using Youden’s index). For *TET2* mutation modeling all variables except 5-hmU were treated as destimulants; in *SF3B1*mutation modelling all variables were treated as stimulants. All statistical analyses were carried out with STATISTICA 13.1 PL [Dell Inc. (2016). version 13. software.dell.com.]. Results were considered statistically significant if*P* values were lessthan 0.05

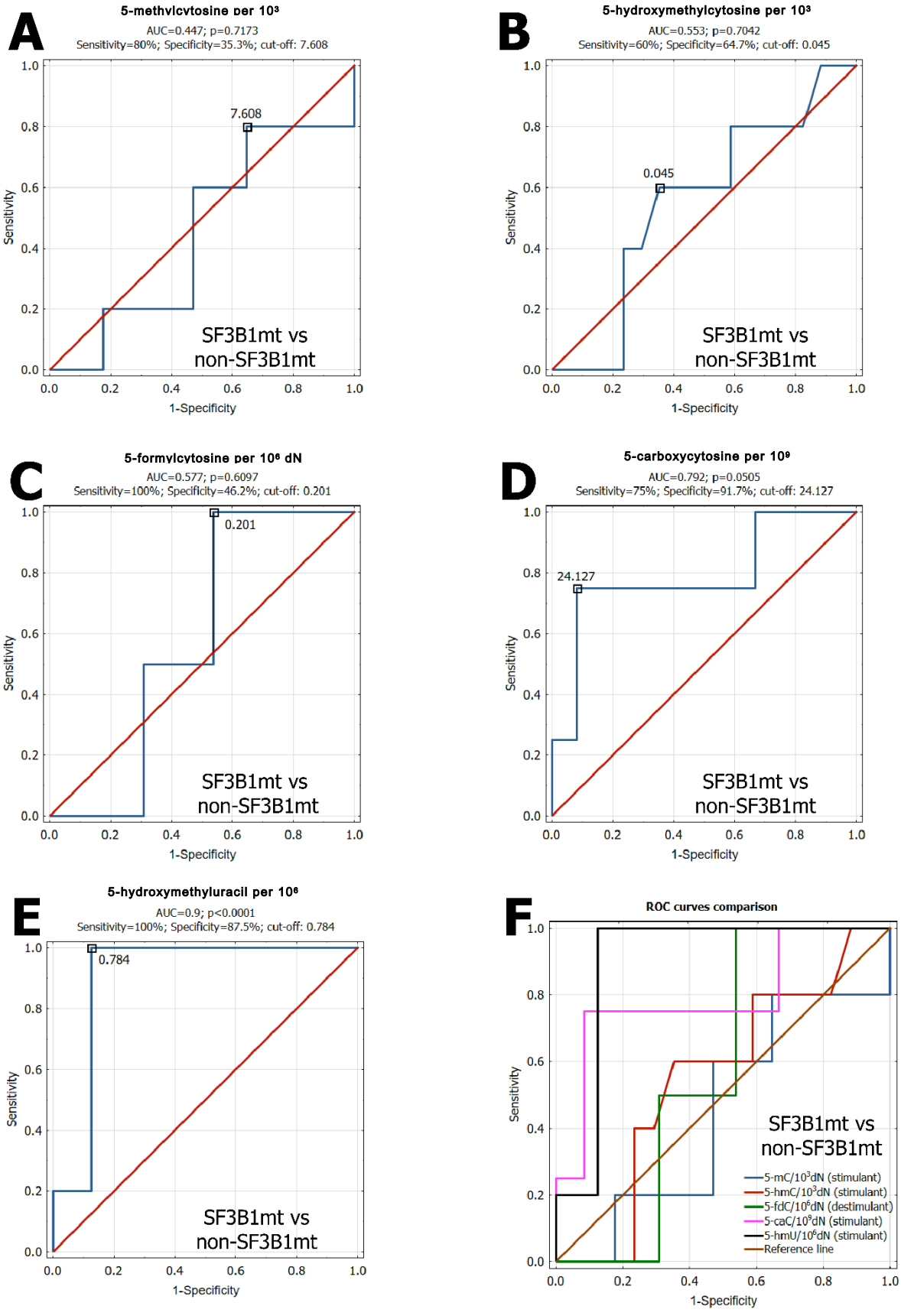
**Supplementary references:**


16. Pozarowski P, Darzynkiewicz Z. Analysis of cell cycle by flow cytometry. Methods Mol Biol 2004; 281:301-311
17. Starczak M, Zarakowska E, Modrzejewska M, Dziaman T, Szpila A, Linowiecka K, et al. In vivo evidence of ascorbate involvement in the generation of epigenetic DNA modifications in leukocytes from patients with colorectal carcinoma, benign adenoma and inflammatory bowel disease. J Transl Med 2018; 16: 204.
18. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia 2014 F; 28: 241-247.
19. Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, et al. Clinical Effect of Point Mutations in Myelodysplastic Syndromes. New Engl J Med 2011; 364: 2496-2506.
20. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 2013; 122: 3616-3627.

**Supplementary figure 1.** ROC analysis (TET2 mutated vs TET2 non-mutated) for each individual epigenetic DNA modification (5-methylcytosine, 5-hydroxymethylcytosine, 5-formylcytosine, 5-carboxycytosine, and 5-hydroxymethyluracil – A, B, C, D, and E respectively). F – Comparison of all ROC curves. Cut-off values were calculated using Youden's index.



**Supplementary figure 2.** ROC analysis (SF3B1 mutated vs non-SF3B1 mutated) for each individual epigenetic DNA modification (5-methylcytosine, 5-hydroxymethylcytosine, 5-formylcytosine, 5-carboxycytosine, and 5-hydroxymethyluracil – A, B, C, D, and E respectively). F – Comparison of all ROC curves. Cut-off values were calculated using Youden's index.



**Supplementary figure 3.** 5-Hydroxymethyluracil is enriched in G2/M cell cycle phase,compared to G0/G1 (FACS sorted hUES7 hESCscells).