DNA Methylation Predicts Survival and Response to Therapy in Patients with Myelodysplastic Syndromes

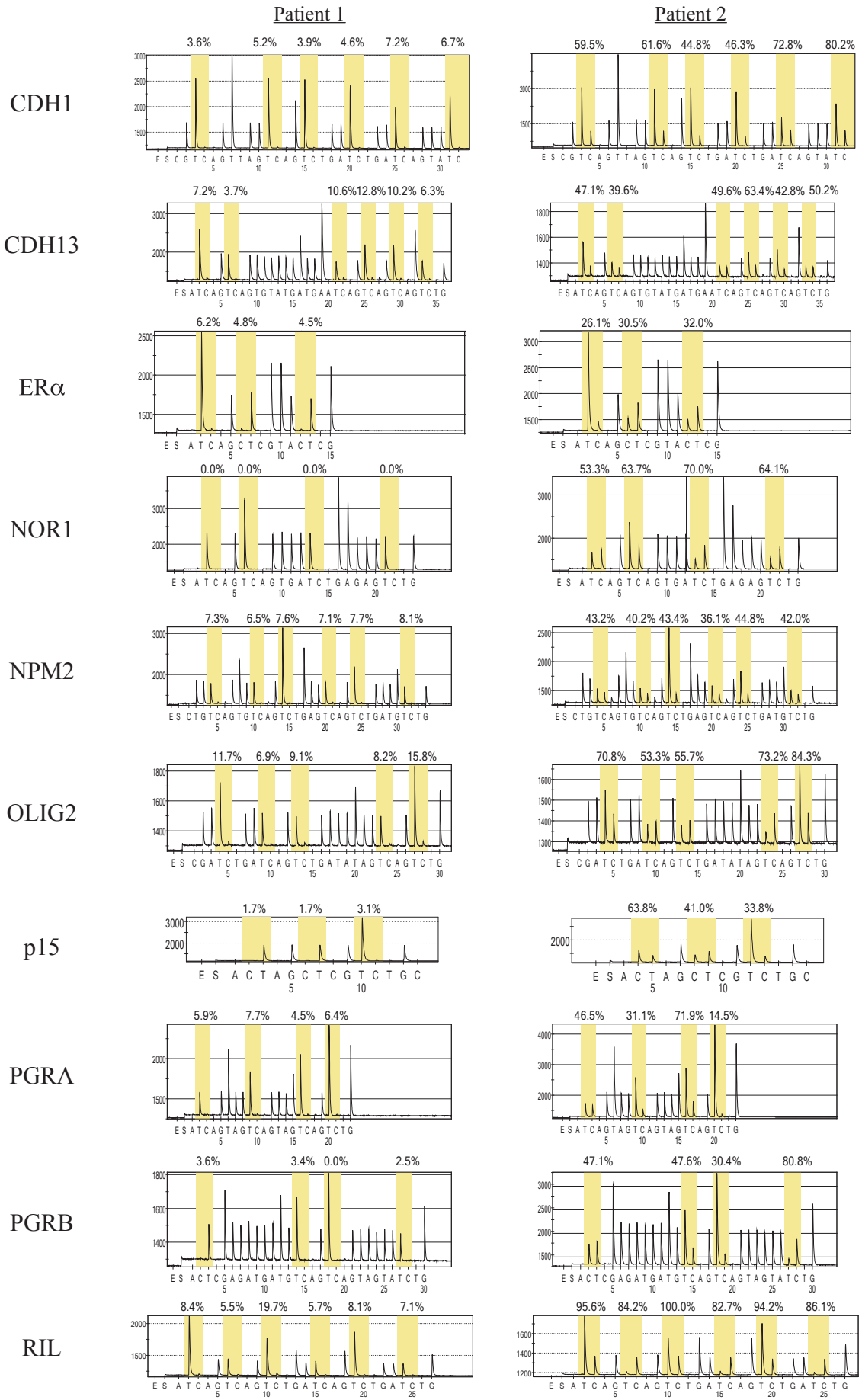
Jean-Pierre J. Issa, et al.

Appendix Methods-Statistical Analysis

To evaluate the effect of the overall genes, methylation for each gene among the patients was standardized by the Z score method, and the Z score was calculated by $(X-\mu)/\sigma$: X stands for methylation data of each gene in each sample; μ stands for mean of methylation of each gene among all samples; and σ stands for standard deviation. Each patient was then assigned a methylation "score" based on the average of Z scores for all genes. The mean and standard deviation of methylation for each gene were calculated from 89 patients in the training cohort. We also used these patients as a reference population to calculate the median methylation Z-score for all genes. Patients were classified into two groups: methylation high, if the methylation scores were greater than the median of the reference population; or methylation low, if the methylation scores were lower than the median of the reference population.

Correlations between methylation and clinical parameters were analyzed with Chi-square tests for categorical variables and with Fisher's exact test when testing small samples. The Wilcoxon rank-sum test or Kruskal Wallis non-parametric test were used to compare two groups of independent but continuous variables. Overall survival was calculated from the date of baseline sampling to that of last follow-up or death, and progression-free survival was calculated from the date of baseline sampling to that of AML diagnosis, death, or last follow-up.

Data Supplement Figure 1. Representative bisulfite-pyrosequencing results of ten genes from two MDS patients. Gene name is indicated on the left. The percentage methylation is quantified by pyrosequencer and indicted at the top of each CpG sites analyzed.

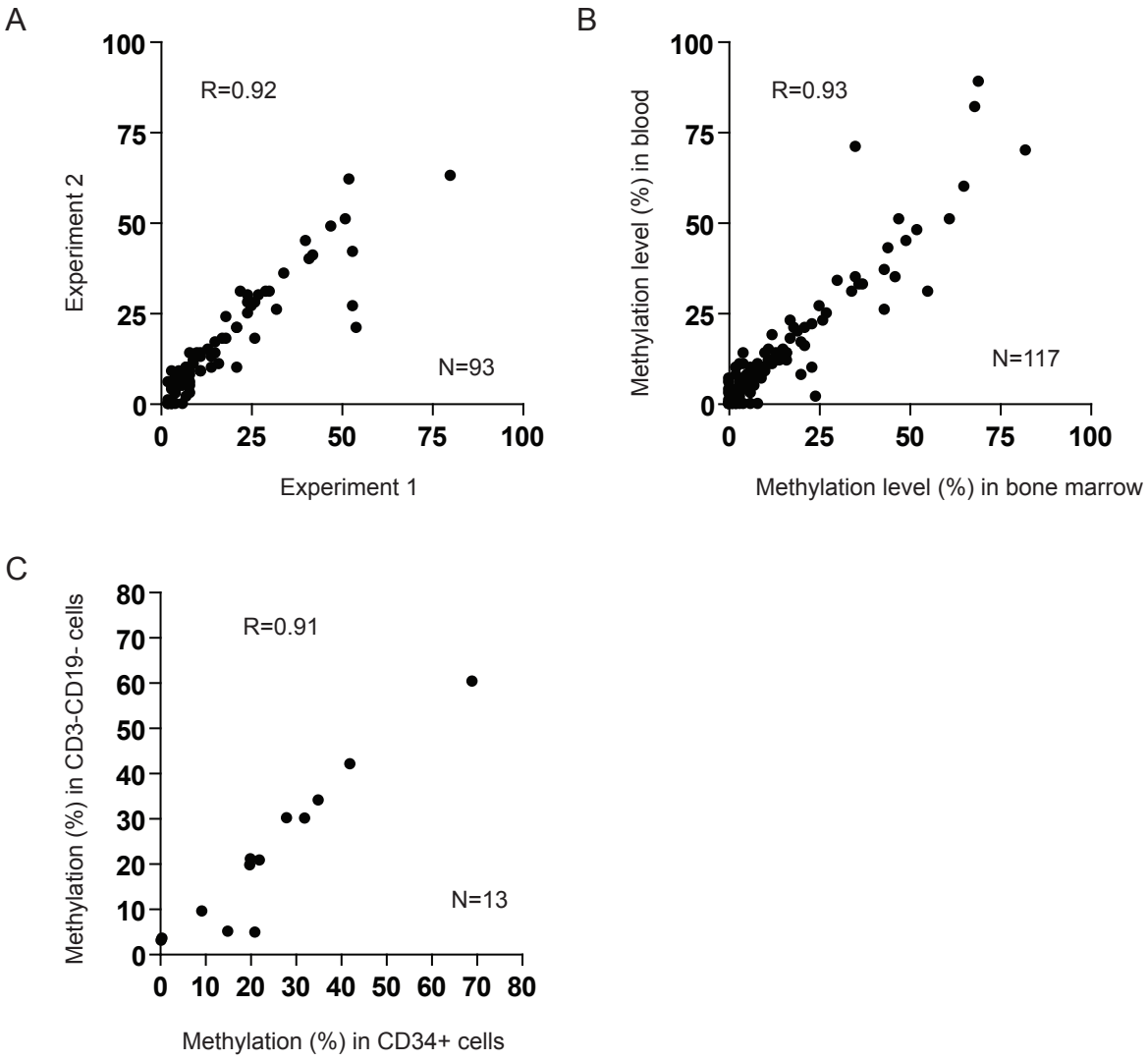


Data Supplement Figure 2.

A. Comparison of methylation analysis between the two independent bisulfite-pyrosequencing results using same sample processed at different times. We found the correlation between the two duplicates was 0.92 (N=93), indicating that the method is highly reproducible.

B. Comparison of methylation between bone marrow and blood. We compared DNA methylation in bone marrow versus blood from 25 patients obtained at same time, a significant correlation in DNA methylation was found between these two types of samples (N=117, R=0.93).

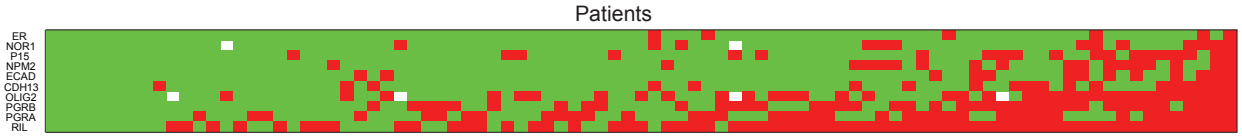
C. Comparison of methylation between CD34+ and negative cells. A perfect correlation was observed in methylation between sorted CD34+ and CD3/19- cells from the same patients (N=13, R=0.91).



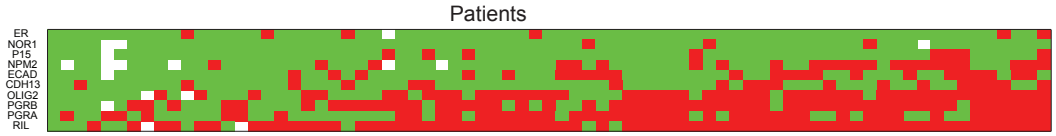
Data Supplement Figure 3. DNA methylation profiling in MDS patients.

The methylation status of ten methylation markers was determined in pretreatment samples from the training cohort (A), first validation cohort (B) and second validation cohort (C). Each column represents a unique patient sample. Each row represents one of the ten genes analyzed (gene names are listed at the left of the figure). Methylation level is shown in two scales: red: methylation level greater than 15%; green: methylation level less than 15%. White indicates no data.

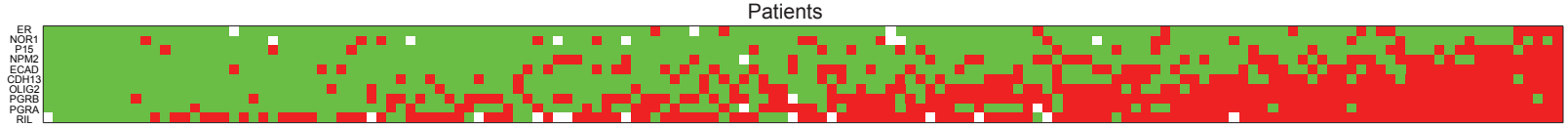
A. Training cohort



B. Validation cohort1



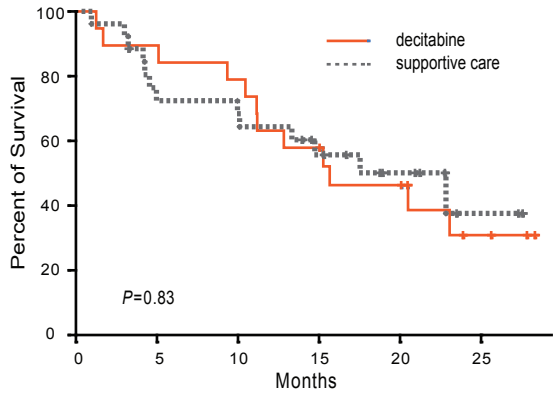
C. Validation cohort2



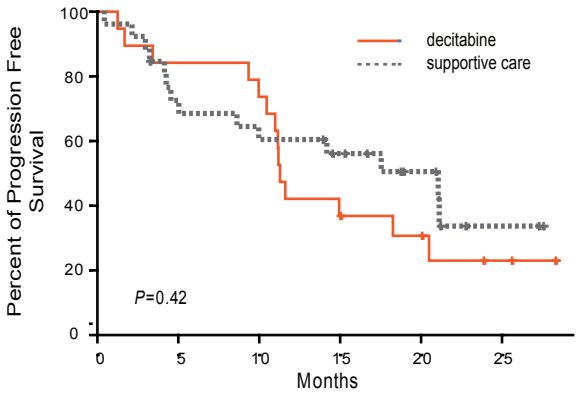
Data Supplement Figure 4.

Kaplan-Meier survival analysis based on baseline methylation score and treatment in the training cohort. Overall survival and progression-free survival are shown for the 45 methylation low patients (Panel A and C, respectively) and the 44 methylation high patients (Panel B and D, respectively). In each panel, patients are grouped based on treatment with decitabine (orange) or supportive care alone (grey). P values are based on the log-rank test.

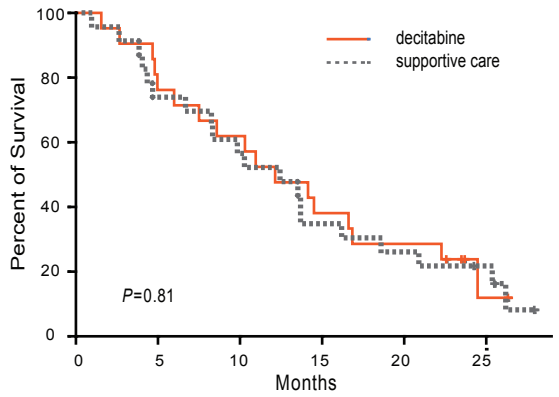
A. Methylation low group



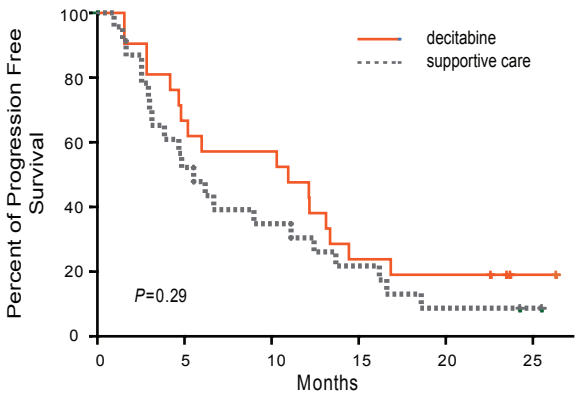
C. Methylation low group



B. Methylation high group

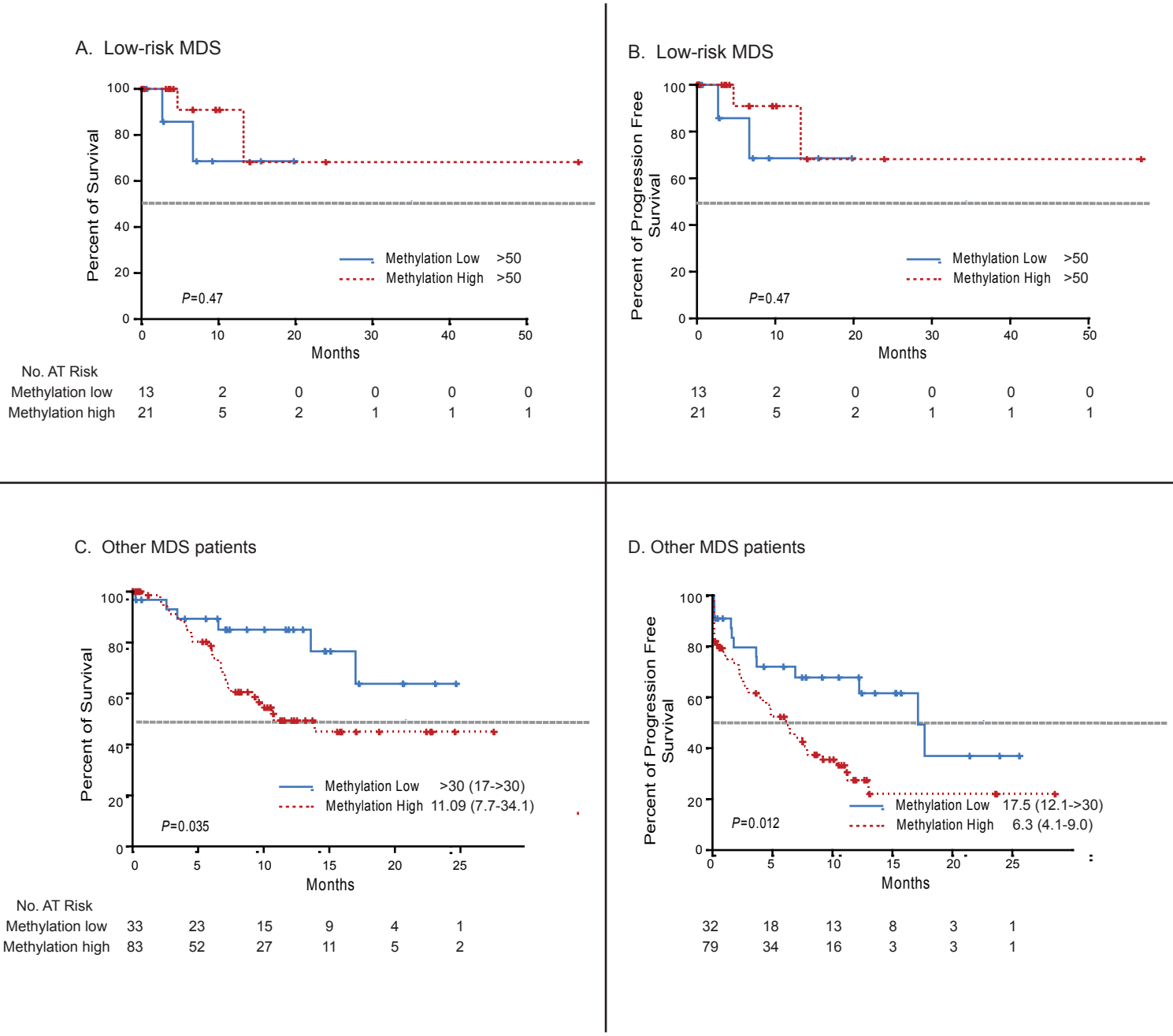


D Methylation high group

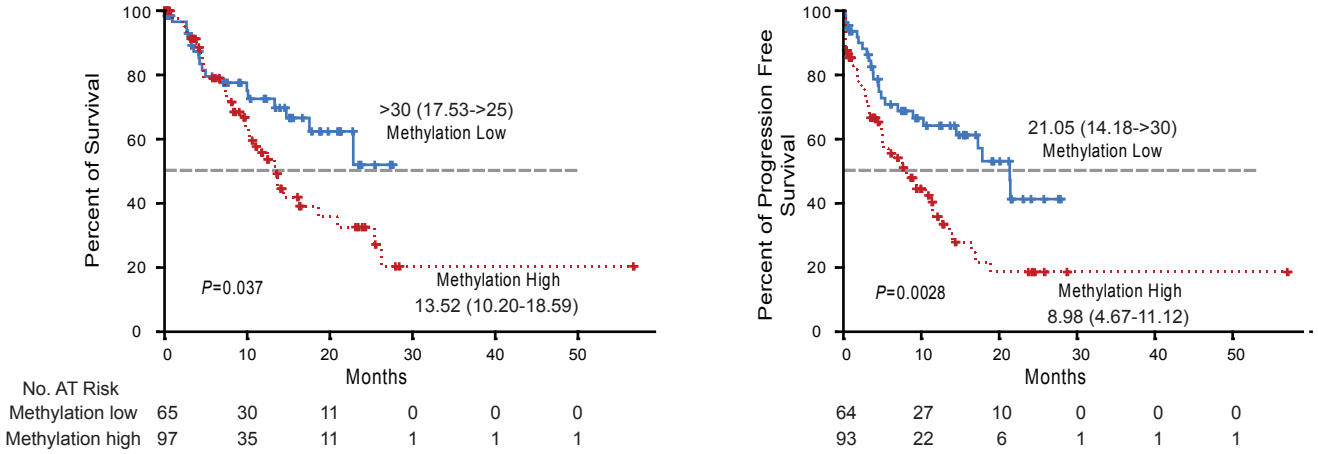


Data Supplement Figure 5.

Kaplan-Meier survival estimates in MDS patients from the second validation cohort. Overall survival and progression-free survival are shown for the 34 low risk MDS patients (Panel A and Panel B, respectively) and the 116 high risk MDS patients (Panel C and Panel D, respectively). In each panel, patients are grouped into methylation low (blue) and methylation high (red) according to their combined methylation Z-scores. Median survivals (95% CI) of each group in each panel are shown in the figure.



Data Supplement Figure 6. Kaplan-Meier survival estimates in MDS patients who did not receive treatment with hypomethylating agents. Left panel shows overall survival and right panel shows progression-free survival.



Data Supplement Table 1. Primer Design and Conditions for Quantitative Bisulfite-Pyrosequencing Analyses

Region	Primer sequences and annealing temperature	Product length (bp)	Sequencing primers	Sequence length (bp)
P15	Forward GTTTTTTTTTAGAAGTAATTTAGG	113	TTTTTAGAAGTAATTTAGG	8
	Reverse (5'-Biotin) TCCTTCTACRACTTAAAACC 50°C			
CDH1	Forward GGAATTGTAAAGTATTTGTGAGTTT	128	GGAAGTTAGTTTAGATTTTA	34
	Reverse-Universal GGGACACCGCTGATCGTTTATCCAAAAACCCATAACTAACC Universal (5'-Biotin) GGGACACCGCTGATCGTTTA 55°C			
Erα	Forward TGTGTTTTTTTTTTAGGTGG	125	GGATACGGTTTGTATTTTG	13
	Reverse (5'-Biotin) AACCATCCCAAATACTTTAATA 58/56/54/52°C			
RIL PDLIM4	Forward TTTGTGAGTTTGGATTGGT	270	Seq1: TTTATTTAGTTTTTAGAGAT	25
	Reverse (5'-Biotin) CCCCAAATAAACCCTCCAT 58/56/54/52°C (antisense)		Seq2: TGTGAGTTTGGATTG	14
PGRA	Forward TGATTGAGTTGAAGGTAAAG	225	Seq1: GGATTTTTATTGTTGTGT	27
	Reverse-Universal GGGACACCGCTGATCGTTTAAAATCCTATCCCTAACAAAA Universal (5'-Biotin) GGGACACCGCTGATCGTTTA 50°C		Seq2: GGGGAGTTAGATTT	24
PGRB	Forward TGTGGGTGGTATTTTTAATGAGA	227	GGGATTTGAGATTTT	27
	Reverse-Universal GGGACACCGCTGATCGTTTACCCCCTCACTAAAACCCTAAA Universal (5'-Biotin) GGGACACCGCTGATCGTTTA 58/56/54/52°C			
CDH13	Forward			

	TTTGGGAAGTTGGTTGGTTG	186	Seq1:	
	Reverse (5’-Biotin)		GGAAAATATGTTTAGTGTAG	31
	ACAACCCCTCTTCCCTACct		Seq2:	
	55°C		TTTTTTTTAAAGTTTGGTTT	11
NOR1	Forward			
	GAGTTGGATTGGTGAAGAGTT	151	TTAGGTTGTTGGGGTAA	19
	Reverse-Universal			
	GGGACACCGCTGATCGTTTAACCAAACCCCCTTCTAATT			
	Universal (5’-Biotin)			
	GGGACACCGCTGATCGTTTA			
	50°C			
NPM2	Forward			
	TTTTTAGGTTAGAGGGGATGAG	195	Seq1:	
	Reverse-Universal		GTTAGAGGGGATGAGGT	32
	GGGACACCGCTGATCGTTTATAAACCTAACTCAAAAACCTCTATA		Seq2:	
	Universal (5’-Biotin)		GGGAGTTGGGATTTAAG	9
	GGGACACCGCTGATCGTTTA			
	58/56/54/52°C			
OLIG2	Forward			
	TTTTAAAGGTGAGGATGTTTATTAT	229	AGGTGAGGATGTTTATTATA	24
	Reverse-Universal			
	GGGACACCGCTGATCGTTTAAAAAATCCAAACCCCCTATAT			
	Universal (5’-Biotin)			
	GGGACACCGCTGATCGTTTA			
	50°C			

Data Supplement Table 2. Initial Screening of Methylation Markers in 24 MDS Patients

Source	Genes	Accession No.	Chr. Band	Function	Methylation % positive cases	Methylation Average level (SD)
Genes identified by genome-wide MCA/RDA method						
	PGRB	NM_000926	11q22	a member of the steroid receptor superfamily	63	18 (11.6)
	RIL	NM_003687	5q31.1	PDZ and LIM domain 4	58	41 (33)
	CDH13	NM_001257	16q24.2	a calcium dependent cell-cell adhesion glycoprotein	38	16 (12.7)
	OLIG2	NM_005806	11q11	oligodendrocyte lineage transcription factor 2	33	12 (10.9)
	PGRA	NM_000926	11q22	a member of the steroid receptor superfamily	26	12 (9.4)
	NOR1	NM_145047	1p34.3	oxidored-nitro domain-containing protein, C1orf102	25	15 (23.3)
	NPM2	NM_181345	8p21.3	nucleolar organization and embryonic development	25	12 (13.1)
	SCGB3A1	NM_052863	5q35	secretoglobin, negative regulation of cell growth and proliferation	10	9 (5.1)
	TCEA3	NM_003196	1q36.12	transcription elongation factor A	7	3 (5.0)
	EDG4	NM_004720	19p12	lysophosphatidic acid receptor 2, G protein-coupled receptor	5	7 (11.7)
	TERT	NM_198253	5p15.33	telomerase reverse transcriptase	5	4 (9.0)
	ECGF1	NM_001113755	22q13.33	thymidine phosphorylase, encodes an angiogenic factor	0	4 (1.0)
	FADS2	NM_004265	11q12	a member of the fatty acid desaturase gene family	0	1 (0.4)
	FLJ36116	AK093435	1p13.2	hypothetical protein LOC388666	0	3 (2.1)
	NKD2	NM_033120	5p15.3	naked cuticle homolog 2, negative regulator of the Wnt-catenin-Tcf pathway	0	4 (2.1)
	SLC26A4	NM_000441	7q31	solute carrier family 26, homology to sulfate transporters	0	3 (2.2)
	TYSND1	NM_001040273	10q22.1	trypsin domain containing 1, involved in beta-oxidation of fatty acids	0	1 (0)
Reports identifying aberrant methylated genes in MDS						
	$p15^{INK4b}$	NM_004936	9p21	cyclin-dependent kinase inhibitor 2B	54	20 (15.9)
	CDH1	NM_004360	16q22.1	E-cadherin, a calcium dependent cell-cell adhesion glycoprotein	38	14 (16.2)
	ERa	NM_000125	6q25.1	a ligand-activated transcription factor	12	6 (7.44)
	$p16^{INK4a}$	NM_000077	9p21	cyclin-dependent kinase inhibitor 2A	0	2 (3.9)
	p73	NM_001126240	1p36.3	tumor protein p73, a member of the p53 family of transcription factors	0	2 (1.0)
	DAPK	NM_004938	9q34.1	death-associated protein kinase 1 involved in interferon induced cell death	0	1 (0.6)
	RASSF1A	NM_007182	3p21.3	a protein similar to the RAS effector proteins	0	0 (0.4)

Data Supplement Table 3. A Pairwise Correlation Analysis between Methylaiton of All Genes in the Training Cohort

	CDH1	ER	RIL	PGRA	PGRB	CDH13	NOR1	NPM2	OLIG2
P15	R=0.56	R=0.22	R=0.43	R=0.42	R=0.3	R=0.44	R=0.33	R=0.43	R=0.44
	P <0.0001	*P* =0.03	*P* <0.0001	*P* <0.0001	*P* =0.007	*P* <0.0001	*P* =0.002	*P* =0.002	*P* <0.0001
CDH1		R=0.45	R=0.46	R=0.66	R=0.41	R=0.59	R=0.47	R=0.46	R=0.69
		P <0.0001	*P* <0.0001	*P* <0.0001	*P* <0.0001	*P* <0.0001	*P* <0.0001	*P* <0.0001	*P* <0.0001
ER			R=0.20	R=0.25	R=0.14	R=0.44	R=0.49	R=0.33	R=0.38
			P =0.05	*P* =0.02	*P* =0.18	*P* <0.0001	*P* <0.0001	P=0.002	*P* =0.0003
RIL				R=0.17	R=0.17	R=0.19	R=0.52	R=0.52	R=0.38
				P =0.11	*P* =0.10	*P* =0.07	*P* <0.0001	*P* <0.0001	*P* =0.0004
PGRA					R=0.32	R=0.40	R=0.25	R=0.42	R=0.47
					P =0.002	*P* <0.0001	*P* =0.02	*P* <0.0001	*P* <0.0001
PGRB						R=0.39	R=0.18	R=0.31	R=0.46
						P <0.0001	*P* =0.09	*P* =0.003	*P* <0.0001
CDH13							R=0.35	R=0.33	R=0.57
							P =0.001	*P* =0.002	*P* <0.0001
NOR1								R=0.48	R=0.54
								P <0.0001	*P* <0.0001
NPM2									R=0.42
									P <0.0001

* all of the correlation coefficients were positive and statistically significant.

Data Supplement Table 4. Correlation between Average Methylation (by z-score) and Clinical Features According to Treatment

Characteristic	without decitabine/azacitidine treatment			with decitabine/azacitidine treatment		
	N	mean (SD)	*P* value	N	mean (SD)	*P* value
Age, years						
< 45	10	0.37 (0.62)	0.25	2	-0.31 (0.10)	0.24
45-65	71	0.18 (0.67)		60	0.24 (0.56)	
> 65	81	0.06 (0.59)		93	0.14 (0.51)	
Gender						
Male	106	0.07 (0.54)	0.10	114	0.20 (0.57)	0.26
Female	56	0.25 (0.77)		41	0.11 (0.39)	
Hemoglobin level, Hg (g/dL)						
< 10	90	0.15 (0.66)	0.75	104	0.19 (0.54)	0.68
≥ 10	70	0.12 (0.59)		49	0.15 (0.51)	
Absolute neutrophil count, ANC ($x10^9$/l)						
< 1.8	84	0.14 (0.59)	0.92	104	0.22 (0.53)	0.41
≥ 1.8	76	0.13 (0.68)		45	0.14 (0.53)	
Platelet count ($x10^9$/l)						
< 100	103	0.17 (0.64)	0.39	106	0.19 (0.55)	0.79
≥ 100	58	0.08 (0.61)		46	0.16 (0.49)	
Marrow blast percentage						
< 5	81	0.01 (0.52)	0.01	42	0.14 (0.55)	0.64
5-10	32	0.08 (0.60)		50	0.20 (0.47)	
11-20	30	0.37 (0.87)		46	0.14 (0.51)	
21-30	18	0.41 (0.51)		15	0.33 (0.78)	
IPSS risk category						
Low	30	-0.005 (0.55)	0.006	6	0.19 (0.71)	0.68
Intermediate 1	60	-0.04 (0.52)		57	0.17 (0.52)	
Intermediate 2	44	0.31 (0.79)		66	0.12 (0.43)	
High	26	0.34 (0.53)		25	0.27 (0.70)	
FAB type						
Refractory anemia (RA)	54	0.04 (0.52)	0.08	17	0.17 (0.50)	0.11
Refractory anemia with ringed sideroblasts (RAS)	17	-0.12 (0.56)		8	-0.15 (0.31)	
Refractory anemia with excess blasts (RAEB, RAEBT)	75	0.25 (0.70)		104	0.23 (0.56)	
Chronic myelomonocytic leukemia (CML, CMML)	16	0.18 (0.58)		23	0.03 (0.43)	
Karyotype						
Good	81	0.07 (0.56)	0.25	78	0.16 (0.56)	0.86
Intermediate	30	0.26 (0.62)		42	0.22 (0.59)	
Poor	39	0.16 (0.52)		28	0.18 (0.39)	
MDS type						
De novo	121	0.15 (0.66)	0.58	115	0.14 (0.56)	0.18
Secondary	41	0.09 (0.54)		40	0.26 (0.44)	

Data Supplement Table 5. Correlation between Methylation of All Genes at Baseline and Clinical Responses to Decitabine Treatment

	Phase 3 trial (N=89)					Phase 2 trail (N=75)				
			Methylation Z score					Methylation Z score		
Responses	N	%	Mean	SD	P value	N	%	Mean	SD	P value
CR/PR	8	9	-0.08	0.39	0.09	35	47	0.16	0.42	0.29
CB	9	10	-0.10	0.53		25	33	0.17	0.28	
SD	29	33	-0.15	0.53		2	3	-0.25	0.03	
PD	30	34	0.19	0.80		7	9	0.33	0.55	
NE	13	14	0.03	0.42		6	8	0.21	0.44	

CR: complete remission; PR: partial remission; CB: clinical benefit; SD: stable disease; PD: progress disease; NE: not evaluated
NE cases was excluded from statisitic analysis