A 4-lncRNA Scoring System for Prognostication of Adult Myelodysplastic Syndromes

Supplementary Information

Methodology used to externally validate the 4-lncRNA score

First, as in the training set, we used the Robust Multichip Average (RMA) algorithm for microarray pre-processing and normalization. Then we used the mean and standard deviation (of each of the four lncRNAs) we derived from the training set to "Z-transform" (i.e., each expression value underwent subtraction of sample mean derived from the 176 training-set patients, followed by division by sample standard deviation derived from the 176 training-set patients) of each the four lncRNA probes, namely TC07000551.hg.1, TC08000489.hg.1, TC02004770.hg.1, and TC03000701.hg.1.

Then we used the Z-transformed expression values to calculate the lncRNA score in the validation set with the formula we derived from the training set: lncRNA risk score = [TC07000551.hg.1] x 0.578 + [TC08000489.hg.1] x 0.526 + [TC02004770.hg.1] x 0.289 + [TC03000701.hg.1] x 0.183. We then used the median of the composite lncRNA scores of the training set as the threshold for assigning patients into either high-score or low-score subgroups. And therefore, in the end, of the 30 MDS patients in the validation cohort, 19 belonged to high-score subgroup, while 11 belonged to the low-score subgroup. Subsequently K-M analyses of survival and AML transformation were performed on these two subgroups.

A list of the relevant parameters used for the detailed calculation are tabulated in Supplementary Table S1.

Supplementary Table S1. **Pertinent lncRNA statistics in the training set (N=176)**

Parameter	RNA accession	Range	Mean	S.D.
TC07000551.hg.1	ENST00000419668	4.757 ~ 9.657	6.286	0.901
TC08000489.hg.1	OTTHUMT00000379034	3.493 ~ 6.298	4.325	0.380
TC02004770.hg.1	TCONS_00002678-XLOC_002452	4.201 ~ 5.445	4.699	0.226
TC03000701.hg.1	OTTHUMT00000358308	3.990 ~ 6.531	4.725	0.398
4-lncRNA risk score		-1.857~5.149	-0.124	1.035

Supplementary Table S2. List of the 30 lncRNAs with significant association ($P<1\times10^{-6}$) with overall survival in the univariate Cox proportional regression model

Probe Set ID	RNA Accession	RNA Source	Hazard Ratio
TC07000168.hg.1	ENST00000519694	ENSEMBL	1.938
TC07000551.hg.1	ENST00000419668	ENSEMBL	2.049
TC10000155.hg.1	NR_038921	RefSeq	1.777
TC03000698.hg.1	NR_037893	RefSeq	1.725
TC06003705.hg.1	uc003pes.1	NONCODE	1.583
TC12000583.hg.1	ENST00000545709	ENSEMBL	1.967
TC01004990.hg.1	TCONS_00001254-XLOC_000542	Rinn lincRNA	1.812
TC13001319.hg.1	NR_044995	NONCODE	1.752
TC14000683.hg.1	ENST00000554693	ENSEMBL	1.711
TC03002398.hg.1	TCONS_00006099-XLOC_002710	Rinn lincRNA	1.666
TC10002501.hg.1	TCONS_00018450-XLOC_008747	Rinn lincRNA	1.675
TC03000701.hg.1	OTTHUMT00000358308	Havana transcript	1.699
TC02004770.hg.1	TCONS_00002678-XLOC_002452	Rinn lincRNA	1.73
TC02003715.hg.1	ENST00000415236	NONCODE	1.781
TC16001489.hg.1	ENST00000457283	NONCODE	1.765
TC15002805.hg.1	NR_026858	RefSeq	1.822
TC08000699.hg.1	OTTHUMT00000381603	Havana transcript	1.701
TC09001544.hg.1	OTTHUMT00000053876	Havana transcript	1.744
TC13000470.hg.1	ENST00000411525	ENSEMBL	1.696
TC08000489.hg.1	OTTHUMT00000379034	Havana transcript	1.702
TC02001072.hg.1	ENST00000436616	ENSEMBL	1.739
TC02002657.hg.1	NR_034096	RefSeq	1.669
TC14000928.hg.1	ENST00000535351	ENSEMBL	1.595
TC15002771.hg.1	NR_040057	RefSeq	1.581
TC14000659.hg.1	NR_024149	RefSeq	1.576
TC09001836.hg.1	uc003zmg.2	NONCODE	1.664
TC21000243.hg.1	NR_026943	RefSeq	1.723
TC06002878.hg.1	ENST00000438676	NONCODE	1.738
TC02003490.hg.1	ENST00000336905	NONCODE	1.556
TC19002464.hg.1	NR_038279	NONCODE	1.538

Supplementary Table S3. Detailed information of the four lncRNAs used to construct the lncRNA risk score

Probe Set ID	**RNA Accession**	**RNA Source**	**Chromosome Locus**	**Transcript Length**
TC07000551.hg.1	ENST00000419668	ENSEMBL	chr7:92,484,223-92,546,488	527 bp
TC08000489.hg.1	OTTHUMT00000379034	Havana transcript	chr8:74,964,575-75,012,088	738 bp
TC02004770.hg.1	TCONS_00002678	Rinn lncRNA	chr2:200,625,267-200,715,896	944 bp
TC03000701.hg.1	OTTHUMT00000358308	Havana transcript	chr3:129,831,393-129,838,017	569 bp

Reference Genome: Human GRCh37/hg19

Supplementary Table S4. Comparison of cytogenetic changes between patients with low and high lncRNA scores

Variables	Total[a]	Low lncRNA score (n=78)[a]	High lncRNA score (n=86)[a]	*P* value
Good karyotype[b]	110	60 (76.9%)	50 (58.1%)	0.013*
Intermediate karyotype[b]	29	11 (14.1%)	18 (20.9%)	0.307
Poor karyotype[b]	25	7 (9.0%)	18 (20.9%)	0.049*
Normal karyotype	100	54 (69.2%)	46 (53.5%)	0.054
Any karyotype abnormality	64	24 (30.8%)	40 (46.5%)	0.054
Loss Y[c]	3	2 (2.6%)	1 (1.2%)	0.605
Del 20q[c]	5	3 (3.8%)	2 (2.3%)	0.669
Del 5q[c]	2	1 (1.3%)	1 (1.2%)	>0.999
Trisomy 8[c]	7	2 (2.6%)	5 (5.8%)	0.447
Monosomy 7[c]	4	1 (1.3%)	3 (3.5%)	0.662
Complex karyotype	18	4 (5.1%)	14 (16.3%)	0.026*
Other abnormalities	25	11 (14.1%)	14 (16.3%)	0.828

[a] Cytogenetic data at diagnosis were available in 164 patients, including78 with low lncRNA scores and 86 with high lncRNA scores

[b] Good, normal karyotype, isolated -Y, del(5q) or del(20q); Poor, complex (≧ 3 abnormalities) or chromosome 7 anomalies; Intermediate, other abnormalities.

[c] As the sole abnormality.

Supplementary Table S5. Comparison of clinical and laboratory features between the training cohort and the validation cohort

Clinical characters	Training cohort (n=176)	Validation cohort (n=30)	*P* value
Sex			0.400
Male	121 (68.8%)	18 (60.0%)	
Female	55 (31.3%)	12 (40.0%)	
Age (year)[a]	67.8 (18.5-94.5)	68.3 (34.5-91.8)	0.811
Laboratory data[a]			
WBC (/μL)	3825 (490-20440)	3630 (1030-19740)	0.526
ANC (/μL)	1768 (103-12728)	1631 (215-15397)	0.299
Hb (g/dL)	8.1 (3.5-14.6)	8.1 (5.8-11.5)	0.633
Platelet (×1,000 /μL)	86 (3-721)	54 (1-376)	0.174
BM blast (%)	3.0 (0-18.8)	6.8 (0.5-17.0)	0.185
2008 WHO classification			
RCUD	42 (23.9%)	6 (20.0%)	0.816
RARS	13 (7.4%)	4 (13.3%)	0.282
RCMD	36 (20.5%)	4 (13.3%)	0.460
RCMD-RS	9 (5.1%)	0 (0%)	0.362
RAEB1	32 (18.2%)	7 (23.3%)	0.462
RAEB2	44 (25.0%)	9 (30.0%)	0.652
Cytogenetic risk[b]			
NK	100 (61.0%)	18 (62.1%)	>0.999
Complex karyotype	18 (11.0%)	7 (24.1%)	0.069
IPSS-R[b,c]			
Very low	6 (3.7%)	0 (0%)	0.594
Low	54 (32.9%)	4 (13.8%)	0.047*
INT	41 (25.0%)	6 (20.7%)	0.815
High	35 (21.3%)	10 (34.5%)	0.152
Very high	28 (17.1%)	9 (31.0%)	0.121

[a] Median (range).

[b] 164 patients in the training cohort, and 29 patients in the validation cohort, had chromosome data at diagnosis.

[c] IPSS-R: Very low, ≦1.5; Low, >1.5-3; intermediate (INT),>3-4.5; High, >4.5-6; and Very high, >6.

**P* value < 0.05

Abbreviations: FAB, French-American-British classification; RARS, refractory anemia with ring sideroblasts; RAEB, refractory anemia with excess blasts; RCUD, refractory cytopenia with unilineage dysplasia; RCMD, refractory cytopenia with multilineage dysplasia; IPSS-R, revised international prognosis scoring system.

Supplementary Table S6. Univariate analysis (Cox regression) of the impact of different variables on the overall survival in the 176 MDS patients

Variable	**Hazard ratio**	**Lower 95% CI**	**Upper 95% CI**	***P* value**
Age [a]	1.019	1.005	1.033	0.007
WBC [b]	1.000	1.000	1.000	0.117
Karyotype [c]	2.171	1.253	3.760	0.006
IPSS-R	1.757	1.454	2.122	<0.001
***FLT3/*ITD**	1.930	0.267	13.944	0.515
NRAS/ KRAS	1.553	0.565	4.270	0.394
RUNX1	1.553	0.836	2.883	0.163
***MLL/*PTD**	1.341	0.186	9.677	0.771
IDH1/2	1.579	0.385	6.471	0.526
ASXL1	1.961	1.205	3.190	0.007
TET2	1.336	0.722	2.472	0.356
DNMT3A	1.307	0.721	2.367	0.378
TP53	10.073	4.794	21.165	<0.001
SETBP1	1.853	0.583	5.888	0.295
EZH2	1.931	0.888	4.200	0.097
SF3B1	0.720	0.397	1.307	0.281
U2AF1	1.044	0.503	2.169	0.908
SRSF2	1.886	1.055	3.371	0.032
ZRSR2	2.646	1.409	4.970	0.002

[a] Age as a continuous variable

[b] WBC as a continuous variable

[c] Unfavorable cytogenetics versus others. Patients without chromosome data were not included in the analysis.

Supplementary Table S7. The 255 genes differentially expressed between patients with the highest (>average +1 SD, n=22) and the lowest (< average − 1 SD, n=22) lncRNA risk scores.

No.	Gene symbol	Fold change	Parametric P value	FDR
1	KIT	9.058	< 1.00E-07	< 1e-07
2	CD34	8.731	< 1.00E-07	< 1.00E-07
3	FLT3	6.951	2.00E-07	0.000162
4	PROM1	6.765	< 1.00E-07	< 1.00E-07
5	PRSS57	6.047	3.00E-07	0.000193
6	SOX4	5.142	< 1.00E-07	< 1.00E-07
7	CPA3	4.937	3.82E-05	0.00339
8	OCLN	4.750	8.53E-05	0.00557
9	CFD	4.313	5.93E-05	0.00449
10	ANGPT1	4.090	5.09E-05	0.00404
11	HPGDS	3.468	0.000104	0.00613
12	CYTL1	3.423	1.70E-05	0.00214
13	CPXM1	3.387	< 1.00E-07	0.00011
14	MYB	3.231	2.34E-05	0.00252
15	IGSF10	3.154	1.50E-06	0.000518
16	MGST1	3.144	2.00E-07	0.000162
17	WDR49	3.087	3.40E-06	0.000792
18	NPR3	3.068	0.000266	0.00991
19	ERLIN1	3.060	8.22E-05	0.00545
20	EMP1	3.022	0.000485	0.014
21	SMIM24	3.016	2.60E-06	0.000683
22	ZNF521	2.914	6.87E-05	0.005
23	MAMDC2	2.888	0.000162	0.00764
24	HGF	2.886	0.000168	0.0077
25	HIST2H4B	2.857	1.39E-05	0.00195
26	HOXA5	2.768	0.000209	0.00867
27	MYO5C	2.577	< 1.00E-07	< 1.00E-07
28	MSRB3	2.572	3.20E-06	0.000775
29	TM4SF1	2.571	3.92E-05	0.00345
30	CDK6	2.570	5.00E-07	0.000291
31	ITGA5	2.554	4.05E-05	0.00349
32	JAG1	2.548	1.42E-05	0.00196
33	CD38	2.508	0.000296	0.0106

34	IGFBP7	2.481	2.46E-05	0.00258
35	SPNS3	2.478	0.000357	0.0117
36	RFX8	2.465	2.30E-05	0.00251
37	JUP	2.439	2.00E-06	0.000601
38	HAL	2.438	0.000145	0.00725
39	C1orf186	2.433	2.60E-06	0.000683
40	ETV6	2.362	3.06E-05	0.00296
41	EAF2	2.307	0.000165	0.0077
42	RAB32	2.278	0.000605	0.016
43	TM7SF3	2.260	5.92E-05	0.00449
44	MPL	2.237	0.000268	0.00992
45	CXorf21	2.228	3.40E-06	0.000792
46	CYP7B1	2.211	0.000501	0.0141
47	DOCK5	2.204	0.000222	0.00897
48	SLC15A2	2.171	2.43E-05	0.00256
49	GUCY1A3	2.163	0.000864	0.02
50	SPTLC3	2.160	0.000144	0.00725
51	DOCK1	2.148	1.45E-05	0.00196
52	TNFSF13B	2.148	0.000152	0.00748
53	GNA15	2.131	2.37E-05	0.00254
54	DPY19L2	2.117	1.56E-05	0.00205
55	TARBP1	2.116	< 1.00E-07	0.00011
56	ANKRD50	2.113	0.000273	0.0101
57	B4GALT6	2.105	2.20E-06	0.000621
58	ERG	2.097	0.000836	0.0196
59	EBPL	2.075	9.30E-06	0.00147
60	TNFRSF1A	2.070	1.40E-06	0.000518
61	MFAP4	2.063	1.00E-06	0.000414
62	ME3	2.062	5.10E-06	0.00103
63	MEIS1	2.059	0.000748	0.0183
64	HOXA7	2.035	6.81E-05	0.00498
65	SERPING1	2.034	0.000198	0.00844
66	ZNF660	2.025	< 1.00E-07	< 1.00E-07
67	LAIR1	2.024	0.000122	0.00671
68	SCCPDH	2.022	0.000188	0.00819
69	HEXA	2.020	4.23E-05	0.00358
70	C3orf80	2.019	0.00098	0.0214

71	S100Z	2.018	0.000123	0.00674
72	HOXA6	2.012	0.000207	0.00864
73	RAVER2	2.003	0.000247	0.00956
74	SLC9A7	2.003	9.61E-05	0.00587
75	MKRN1	0.500	0.000235	0.0093
76	CD160	0.497	2.17E-05	0.00241
77	ASCC2	0.493	0.000243	0.00948
78	BTLA	0.491	2.00E-05	0.00232
79	POC1B	0.491	3.98E-05	0.00347
80	MAP2K3	0.489	0.000478	0.0139
81	ENDOD1	0.488	0.000159	0.0076
82	MICALCL	0.488	0.000227	0.0091
83	KLF3	0.487	0.000887	0.0203
84	NCOA4	0.487	0.000319	0.0109
85	PPM1A	0.487	4.36E-05	0.00366
86	ACSL6	0.482	0.000118	0.00657
87	CRTAM	0.479	8.20E-06	0.00136
88	ST6GALNAC4	0.478	0.000883	0.0203
89	CDR2	0.477	0.000331	0.0111
90	RAD23A	0.475	0.000418	0.0127
91	LGALS3	0.475	1.44E-05	0.00196
92	CARM1	0.475	7.52E-05	0.00514
93	DENND4A	0.474	0.000182	0.00801
94	SLAMF1	0.473	0.000106	0.00619
95	CD48	0.469	0.000853	0.0198
96	MARCH3	0.468	0.000316	0.0109
97	USP15	0.466	0.000298	0.0106
98	TMEM57	0.465	0.000594	0.0159
99	PTPN4	0.464	7.20E-06	0.00127
100	RNF123	0.464	0.000462	0.0137
101	LY9	0.462	1.96E-05	0.00228
102	ANKRD9	0.461	0.000361	0.0117
103	SAMD3	0.459	0.000143	0.00725
104	HK1	0.458	9.35E-05	0.00576
105	CTLA4	0.454	0.000677	0.0171
106	AIDA	0.451	0.000183	0.00806
107	WDR26	0.450	1.82E-05	0.0022

108	ARL4A	0.445	0.000241	0.00947
109	S1PR5	0.445	0.000675	0.0171
110	SUCLG2	0.443	2.42E-05	0.00256
111	CCDC109B	0.441	0.00061	0.0161
112	SPOCK2	0.441	0.00058	0.0156
113	TUBB2A	0.437	0.000507	0.0143
114	LCK	0.437	0.0005	0.0141
115	ELL2	0.437	0.000195	0.00839
116	IL32	0.434	0.000242	0.00947
117	SOX6	0.432	0.000994	0.0216
118	EOMES	0.432	5.10E-06	0.00103
119	PPME1	0.432	0.000647	0.0166
120	TRIM23	0.429	2.42E-05	0.00256
121	DNAJC6	0.424	0.000317	0.0109
122	DCUN1D1	0.422	0.000918	0.0208
123	FAM210B	0.420	0.000612	0.0161
124	DPP4	0.420	0.000614	0.0161
125	SLC1A5	0.418	0.000724	0.0179
126	ADIPOR1	0.417	0.000196	0.00839
127	CCR4	0.416	0.000498	0.0141
128	STAT4	0.413	0.000177	0.00793
129	RNF14	0.412	3.30E-06	0.000789
130	GPRIN3	0.411	0.000833	0.0196
131	PCGF5	0.410	2.20E-06	0.000621
132	GPR15	0.408	0.00061	0.0161
133	SEC62	0.407	1.44E-05	0.00196
134	CAMK4	0.407	0.000316	0.0109
135	CAPRIN2	0.406	0.000965	0.0213
136	TMEM63B	0.406	9.40E-05	0.00576
137	NELL2	0.404	0.000538	0.0149
138	FOXO3	0.401	4.10E-05	0.00352
139	CDC34	0.401	3.72E-05	0.00333
140	CLCN3	0.401	0.000388	0.0121
141	GPR18	0.400	0.000554	0.0152
142	CISD2	0.399	0.000665	0.0169
143	OPTN	0.398	0.000479	0.0139
144	WNK1	0.396	0.000839	0.0196

145	CD247	0.395	0.000323	0.011
146	C9orf78	0.393	8.58E-05	0.00557
147	KLF12	0.392	0.000846	0.0197
148	IL6ST	0.392	2.16E-05	0.00241
149	PSME4	0.390	0.000107	0.00619
150	UBE2H	0.383	1.17E-05	0.00174
151	PRDM1	0.383	0.000814	0.0193
152	MYO1D	0.381	3.00E-07	0.000193
153	SAMD9	0.381	0.000116	0.0065
154	SKAP1	0.379	0.000171	0.00777
155	PIP4K2A	0.378	3.68E-05	0.00331
156	KLRC2	0.378	3.43E-05	0.0032
157	RNF11	0.378	9.00E-07	0.000414
158	ABCD2	0.375	1.59E-05	0.00207
159	GIMAP7	0.375	0.000957	0.0212
160	SH2D1B	0.374	0.000924	0.0209
161	KLRG1	0.374	0.000475	0.0139
162	NEDD4L	0.373	2.14E-05	0.0024
163	BCL11B	0.370	0.000632	0.0164
164	EPB41	0.370	0.000282	0.0103
165	FAXDC2	0.367	0.000939	0.021
166	BIRC3	0.364	0.000529	0.0147
167	SLFN12L	0.363	5.36E-05	0.00415
168	SGPP1	0.362	0.00018	0.00795
169	ATG14	0.359	0.000127	0.00684
170	SLFN5	0.359	0.000638	0.0165
171	NCEH1	0.356	0.000368	0.0118
172	CPEB4	0.355	0.000195	0.00839
173	DTHD1	0.347	3.11E-05	0.00297
174	MXI1	0.344	0.000242	0.00947
175	TERF2IP	0.340	4.18E-05	0.00357
176	KDM7A	0.340	9.70E-06	0.00152
177	KAT2B	0.337	0.000114	0.00646
178	THEMIS	0.335	4.00E-05	0.00347
179	GPR171	0.334	4.81E-05	0.00394
180	CD28	0.332	0.000386	0.0121
181	MYBL1	0.332	9.40E-05	0.00576

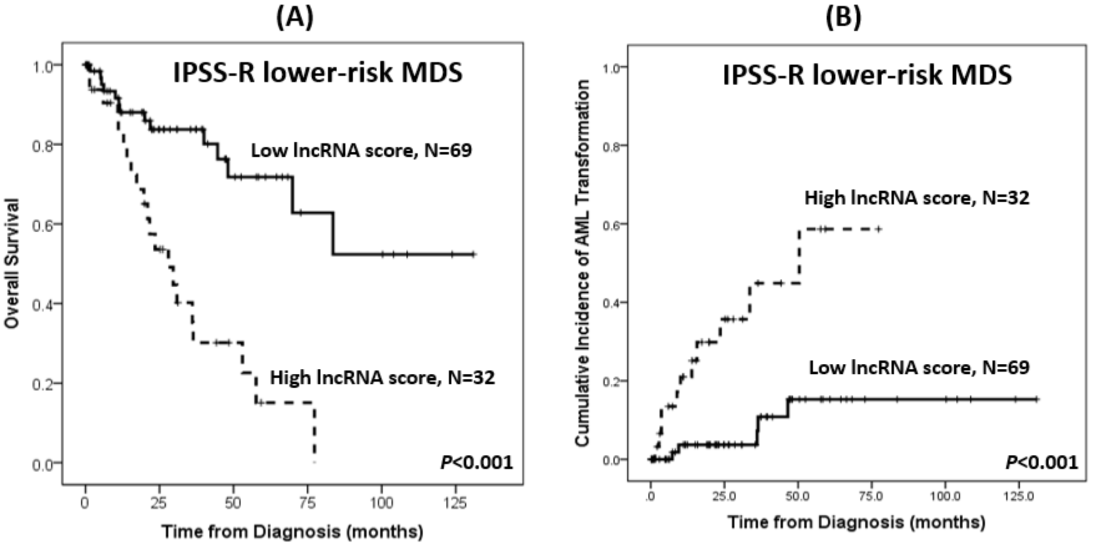
182	PIM1	0.328	0.00018	0.00795
183	SLC6A8	0.326	0.000978	0.0214
184	SLC6A19	0.325	0.000759	0.0185
185	GRAP2	0.323	1.50E-06	0.000518
186	ETS1	0.320	3.49E-05	0.00322
187	RASGRP1	0.319	8.88E-05	0.00562
188	CCNDBP1	0.319	2.40E-06	0.000658
189	ERAP2	0.317	0.000641	0.0166
190	RBM38	0.316	0.000122	0.0067
191	FCRL6	0.314	0.000712	0.0178
192	XPO7	0.313	0.000729	0.018
193	UTS2	0.313	0.000295	0.0106
194	ABCG2	0.312	0.000109	0.00622
195	USP12	0.312	4.27E-05	0.0036
196	PITHD1	0.312	0.000249	0.0096
197	FCRL3	0.310	3.37E-05	0.00317
198	CD3E	0.310	0.000212	0.0087
199	ICOS	0.309	0.000769	0.0185
200	P2RY10	0.308	0.000191	0.00827
201	PYHIN1	0.308	0.000283	0.0103
202	TMCC2	0.307	6.44E-05	0.00478
203	KIR2DL3	0.301	0.000734	0.018
204	E2F2	0.299	2.77E-05	0.00281
205	RORA	0.299	5.80E-06	0.0011
206	MS4A1	0.298	0.000987	0.0215
207	SESN3	0.298	2.11E-05	0.00238
208	TRAT1	0.296	3.00E-06	0.000746
209	SYNE2	0.295	0.000124	0.00676
210	YOD1	0.293	1.65E-05	0.00211
211	TC2N	0.292	4.21E-05	0.00358
212	GPR174	0.292	0.000159	0.0076
213	UBE2O	0.292	7.80E-05	0.00527
214	CD3D	0.291	3.40E-05	0.00318
215	TCP11L2	0.291	7.31E-05	0.0051
216	KLRC3	0.278	6.56E-05	0.00485
217	CD2	0.272	0.000228	0.0091
218	HBM	0.271	0.000645	0.0166

219	IKZF3	0.270	0.000108	0.00622
220	GSPT1	0.269	1.77E-05	0.00218
221	MMD	0.268	4.77E-05	0.00394
222	CCL5	0.267	0.000132	0.00698
223	LCN2	0.267	0.000368	0.0118
224	KLRF1	0.264	0.00094	0.021
225	ITK	0.257	0.000103	0.00608
226	RIOK3	0.256	1.30E-06	0.000505
227	CD226	0.254	9.32E-05	0.00576
228	BPGM	0.252	5.31E-05	0.00415
229	HEMGN	0.249	0.00094	0.021
230	SLC25A39	0.246	0.000204	0.00856
231	MBNL3	0.244	4.43E-05	0.0037
232	PLBD1	0.244	7.76E-05	0.00526
233	CD3G	0.241	0.000514	0.0144
234	KLRD1	0.237	7.13E-05	0.00504
235	GYPB	0.236	0.000575	0.0156
236	FECH	0.235	0.00031	0.0107
237	SNCA	0.228	0.000269	0.00994
238	SLAMF6	0.221	7.85E-05	0.00528
239	GZMK	0.219	5.70E-06	0.0011
240	DCAF12	0.216	0.000139	0.00719
241	GYPE	0.216	0.000805	0.0191
242	SLFN14	0.215	0.000218	0.00888
243	FHDC1	0.212	0.000622	0.0162
244	STRADB	0.204	0.000128	0.00684
245	GZMA	0.202	3.96E-05	0.00347
246	GMPR	0.198	0.000138	0.00717
247	SLC14A1	0.197	0.000655	0.0168
248	KLRC4-KLRK1	0.192	9.30E-06	0.00147
249	TUBB1	0.165	0.000766	0.0185
250	IL7R	0.159	7.90E-06	0.00133
251	PHOSPHO1	0.153	1.00E-06	0.000414
252	FAM46C	0.151	4.40E-06	0.000954
253	KLRB1	0.147	3.70E-06	0.000841
254	RUNDC3A	0.143	1.50E-06	0.000518
255	IFIT1B	0.115	0.000447	0.0134

Supplementary Table S8. The KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways associated with the 255 differentially expressed genes between the MDS patients with high and low lncRNA scores

Implicated KEGG Pathways	**Gene Count**	**Nominal *P* Value**	**FDR**
Hematopoietic cell lineage	**10**	**1.20E-05**	**1.00E-03**
T cell receptor signaling pathway	**11**	**8.30E-06**	**1.40E-03**
Primary immunodeficiency	5	2.40E-03	1.30E-01
Chagas disease (American trypanosomiasis)	7	8.20E-03	3.00E-01
Natural killer cell mediated cytotoxicity	7	1.70E-02	3.90E-01
Cytokine-cytokine receptor interaction	10	1.60E-02	4.30E-01
HTLV-I infection	10	2.90E-02	5.20E-01
Cell adhesion molecules (CAMs)	7	3.30E-02	5.20E-01
Malaria	4	5.00E-02	6.30E-01
Dorso-ventral axis formation	3	7.60E-02	6.30E-01
Acute myeloid leukemia	4	6.90E-02	6.50E-01
Measles	6	7.50E-02	6.50E-01
Transcriptional misregulation in cancer	7	6.50E-02	6.60E-01
Rheumatoid arthritis	5	6.20E-02	6.70E-01
Central carbon metabolism in cancer	4	9.40E-02	6.80E-01

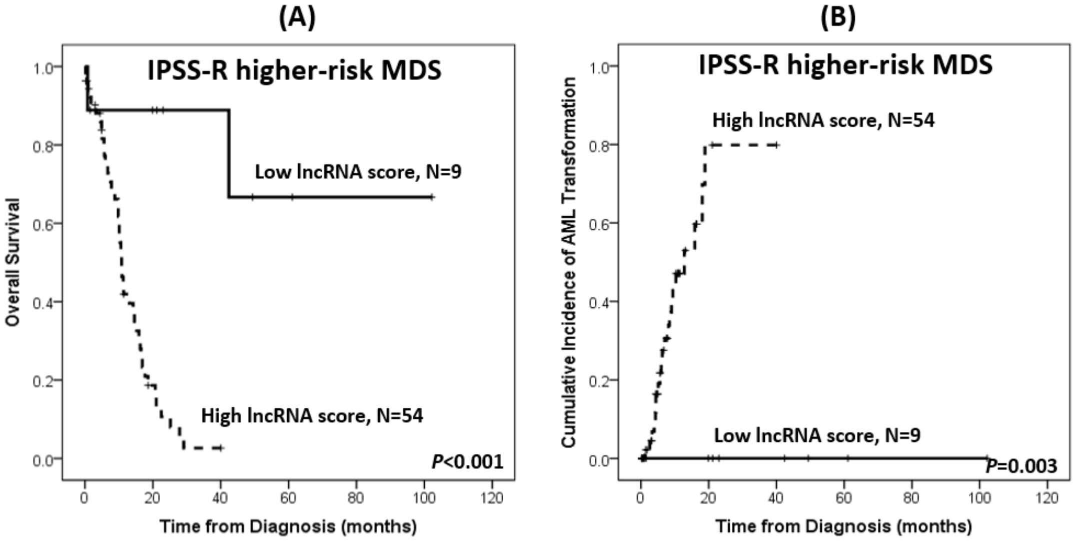
Supplementary Figure S1:



Supplementary Figure S1 Legend:

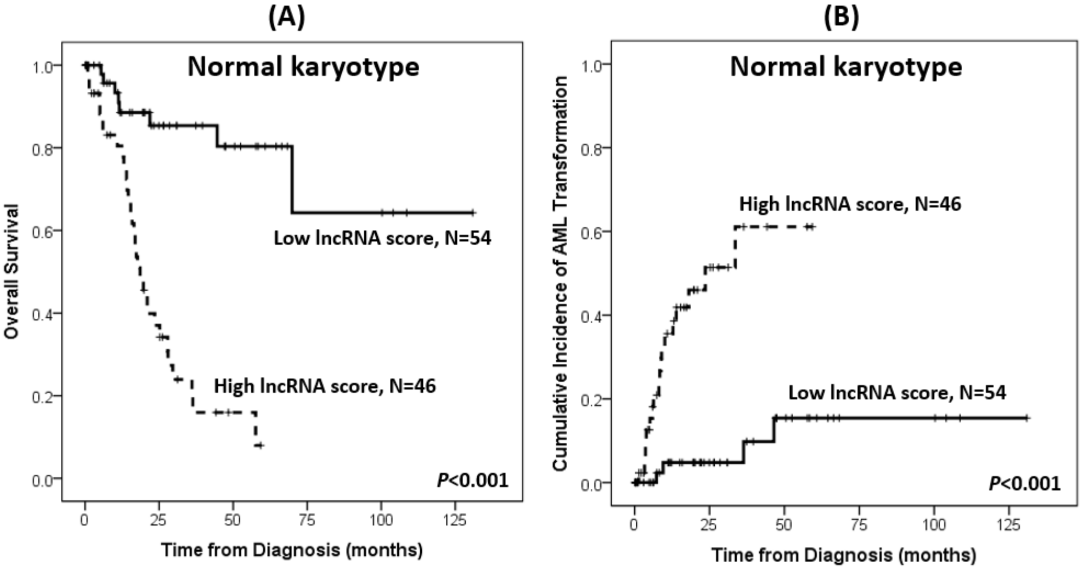
Kaplan-Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 101 MDS patients who had lower-risk IPSS-R (very low, low and intermediate risks) in the training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores.

Supplementary Figure S2:



Supplementary Figure S2 Legend:
Kaplan-Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 63 MDS patients who had higher-risk IPSS-R (high and very high risks) in the training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores.

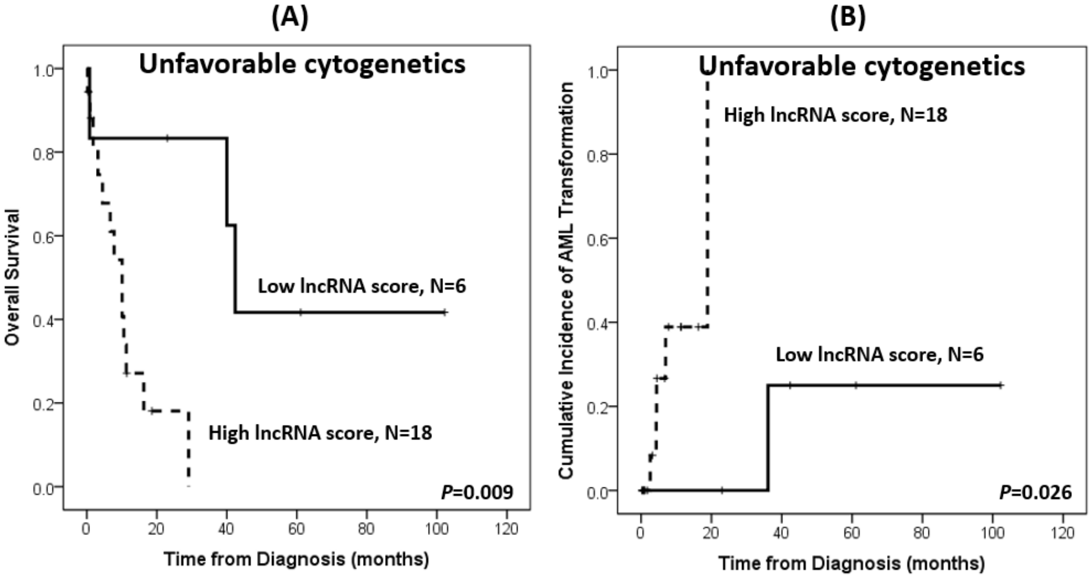
Supplementary Figure S3:



Supplementary Figure S3 Legend:

Kaplan-Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 100 MDS patients who had normal karyotype in the training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores.

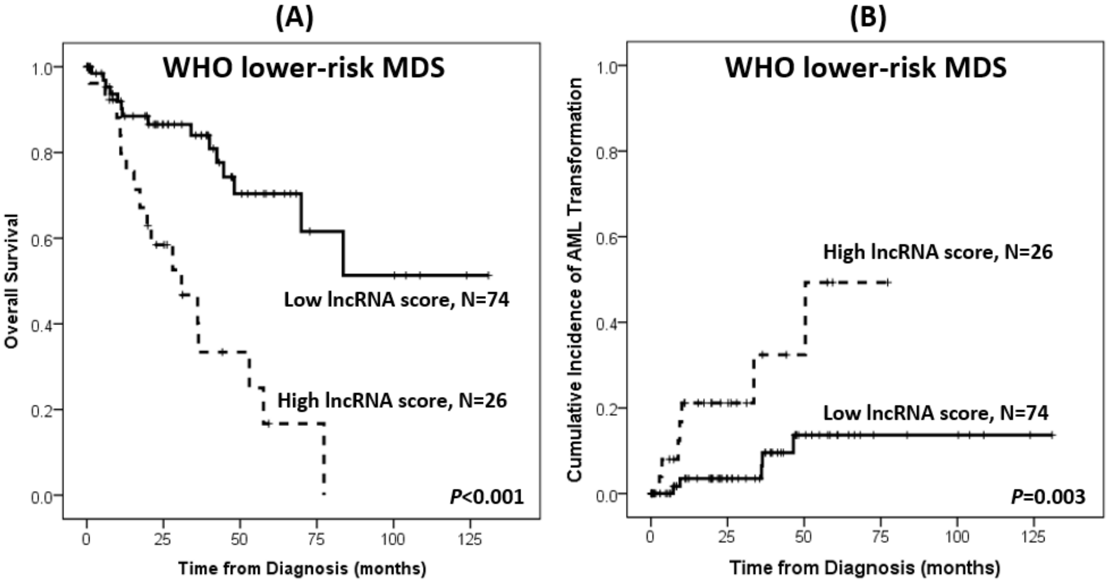
Supplementary Figure S4:



Supplementary Figure S4 Legend:

Kaplan-Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 24 MDS patients who had unfavorable cytogenetics in the training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores.

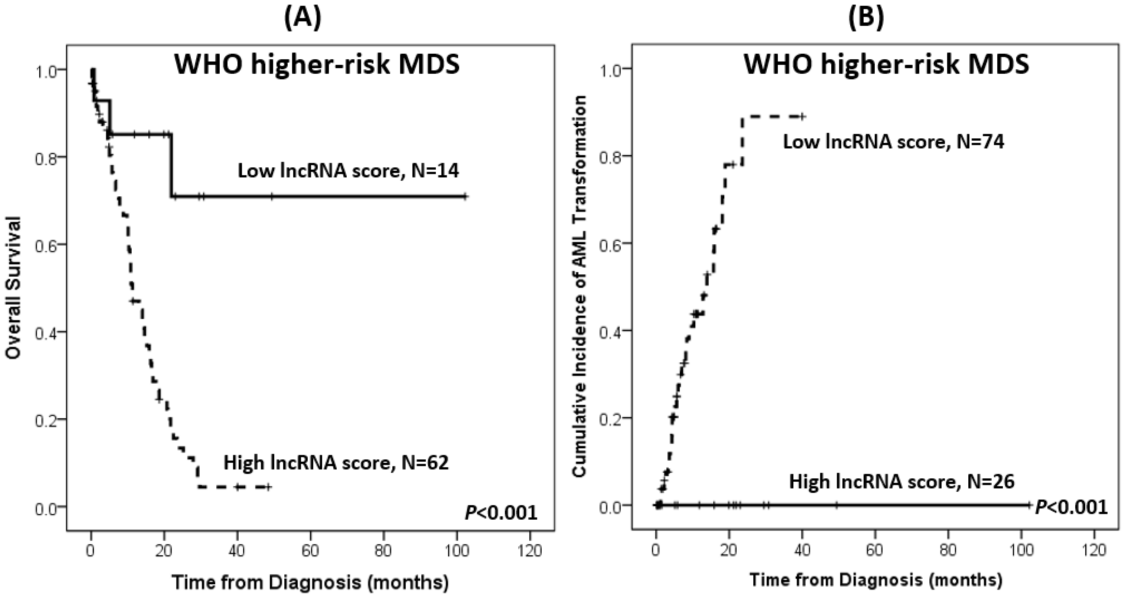
Supplementary Figure S5:



Supplementary Figure S5 Legend:

Kaplan-Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 100 MDS patients who had lower-risk WHO classifications (RCUD, RARS, RCMD, and RCMD-RS) in the training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores.

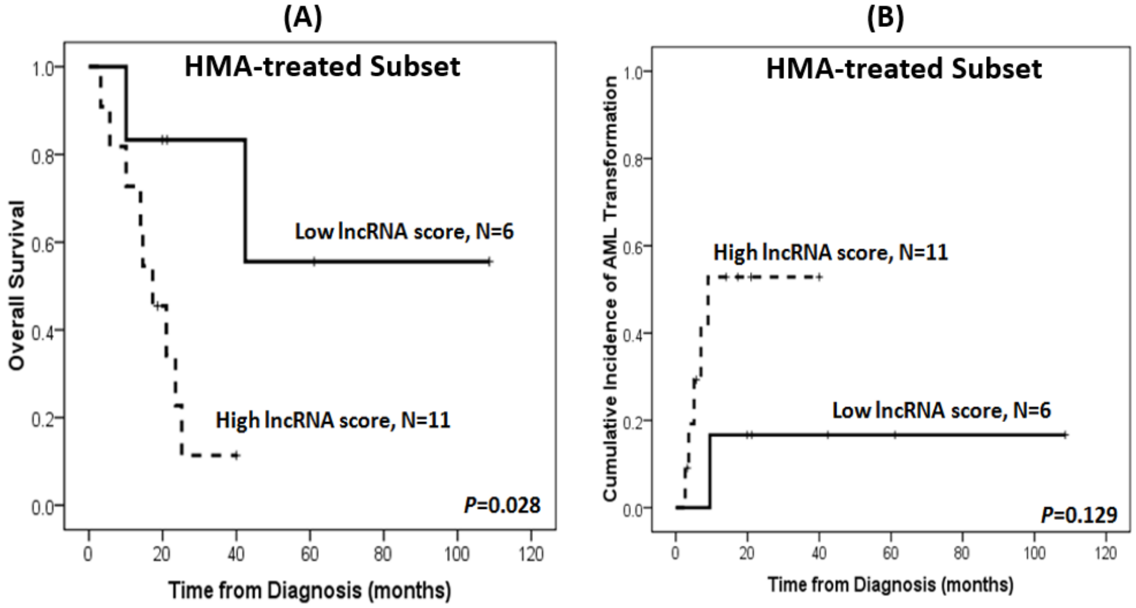
Supplementary Figure S6:



Supplementary Figure S6 Legend:

Kaplan-Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 76 MDS patients who had higher-risk WHO classifications (RAEB-1 and RAEB-2) in the training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores.

Supplementary Figure S7:

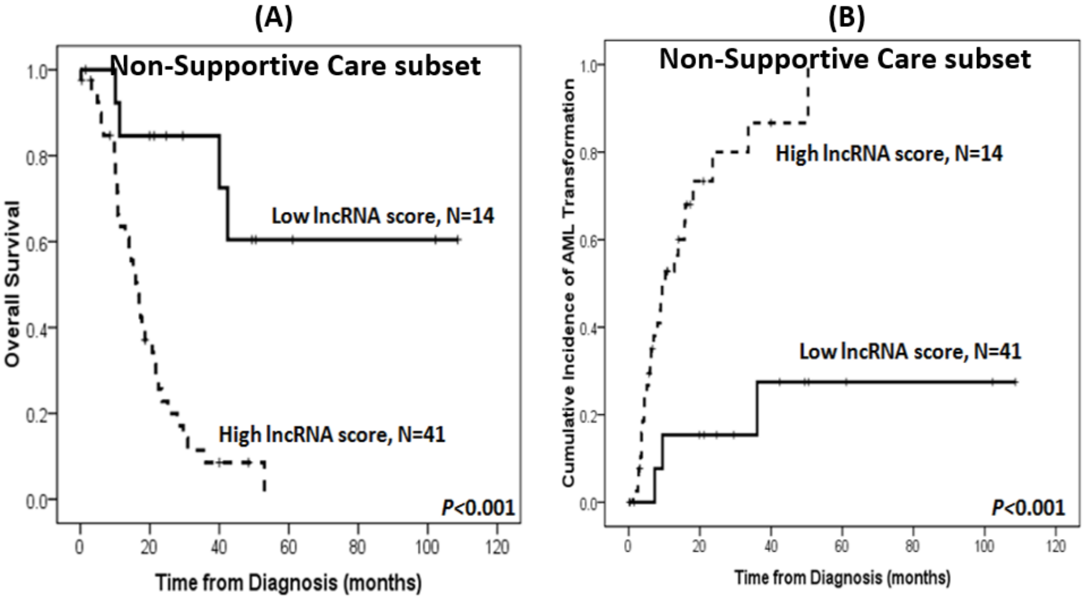


Supplementary Figure S7 Legend:

Kaplan–Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 17 MDS patients in our training cohort who had received HMA treatment, stratified by lncRNA scores. Patients with higher lncRNA scores had statistically significant shorter median, along with a marginally higher incidence of projected AML transformation at five years, than those with lower scores. (Median OS: 17.3 months vs. NR, P=0.028; AML transformation rate at five years: 52.9% vs. 16.7%, P=0.129)

(Please note: Although HMAs have become the mainstay treatment for high-risk MDS patients since the results of the AZA-001 trials had been published in 2009, in Taiwan, the national reimbursement of azacitidine only started after year 2013. Therefore, as most of the patients' samples in this retrospective cohort of 176 adult primary MDS patients were collected at time points (November 1992 ~ December 2010) earlier than the widespread use of azacitidine in Taiwan, most of the patients did not have access to this medication. Only few patients who were economically self-sufficient could afford azacitidine treatment back then. So, in this cohort of 176 patients, only 17 of them had received azacitidine treatment.)

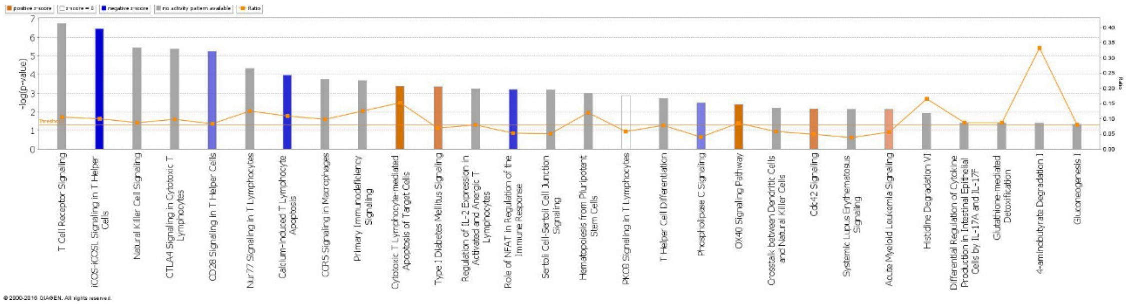
Supplementary Figure S8:



Supplementary Figure S8 Legend:

Kaplan–Meier plots of OS (A) and cumulative incidence of AML transformation (B) of the 55 non-supportive care MDS patients in our training cohort, stratified by lncRNA scores. Patients with higher lncRNA scores had worse clinical outcomes than those with lower scores. (Median OS: 16.3 months vs. NR, P<0.001; AML transformation rate at five years: 100.0% vs. 27.5%, P<0.001)

Supplementary Figure S9:



Supplementary Figure S9 Legend:

The implicated IPA pathways associated the 255 differentially expressed genes between patients with the highest (>average +1 SD, n=22) and the lowest (< average − 1 SD, n=22) lncRNA risk scores.