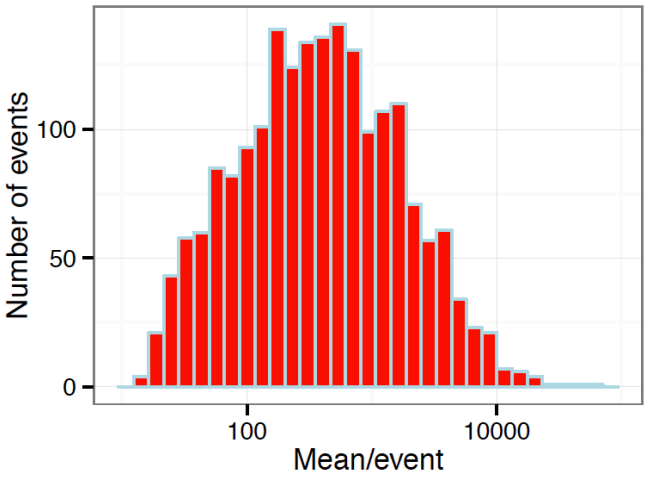
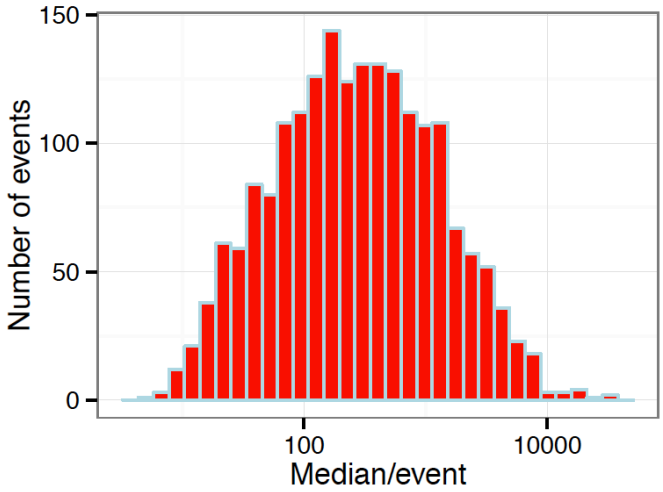
Figure S1

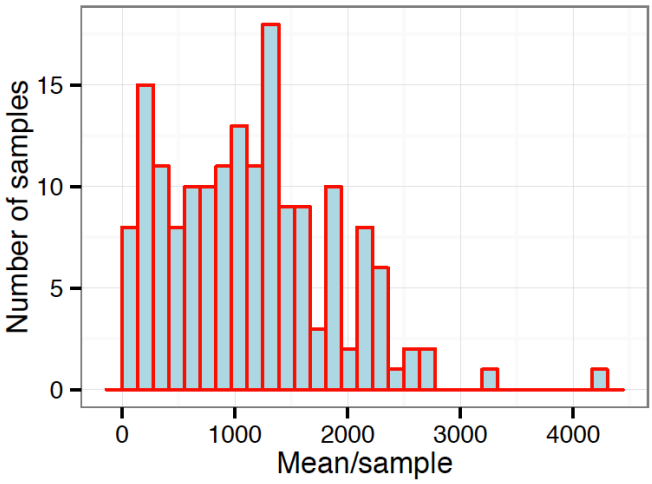
A



B



C



D

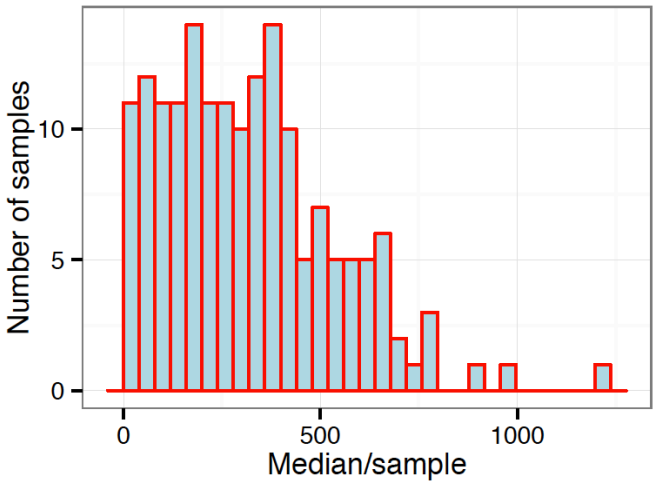


Figure S2

A

Out-of-frame

MDS express 2.3 fold more of 98bp exon 4b containing isoform

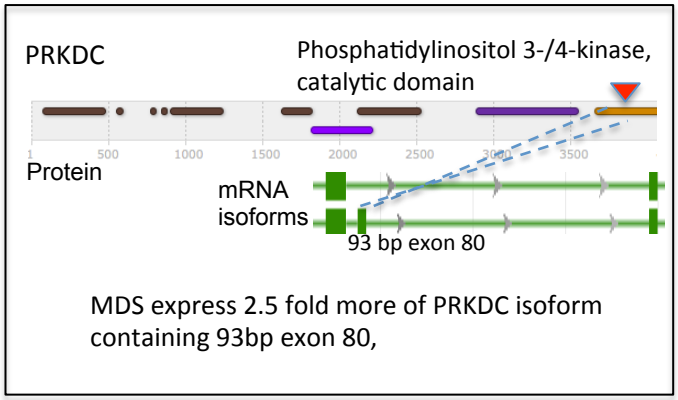
MDS express 3.3 fold more of 70bp exon 31 skipping isoform

MDS express 3.5 fold more of 92bp exon 6 skipping isoform

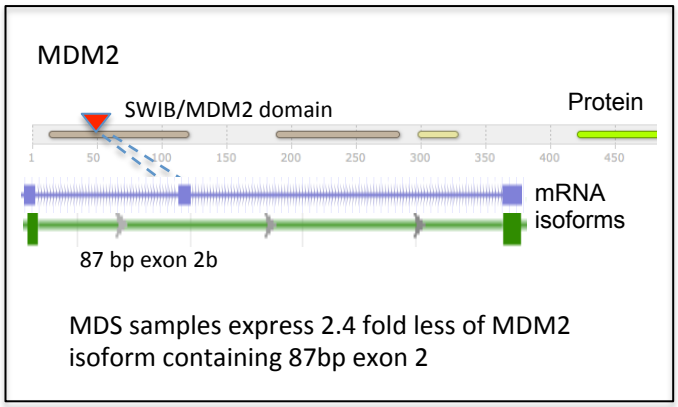
MDS express 2.1 fold more of 133bp exon 3 skipping isoform

B

In-frame, in-domain



MDS express 2.5 fold more of PRKDC isoform containing 93bp exon 80,

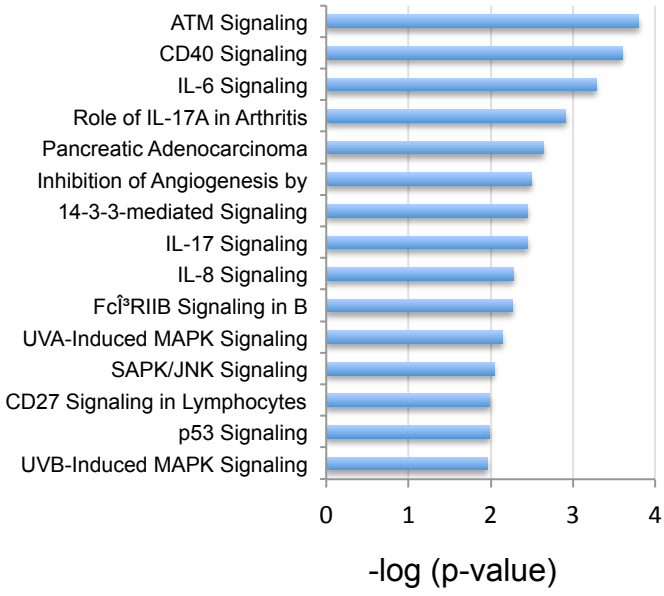


MDS samples express 2.4 fold less of MDM2 isoform containing 87bp exon 2

Figure S3

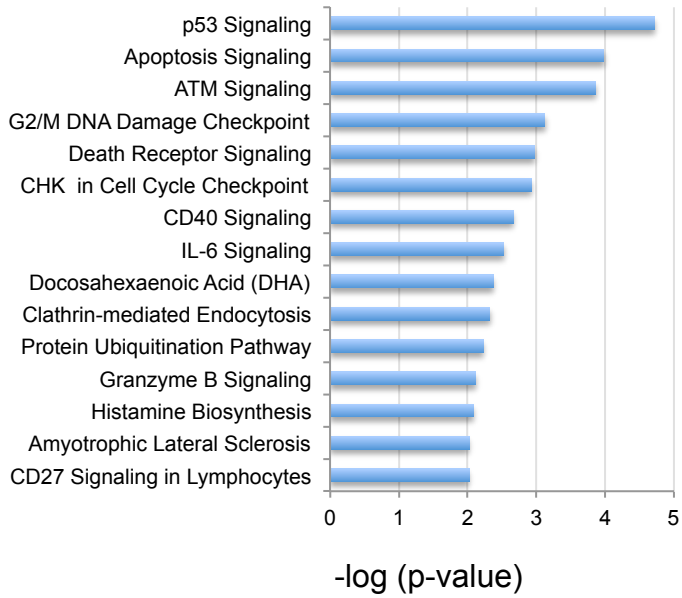
A

SRSF2-mut



C

SF3B1-mut



B

U2AF1-mut

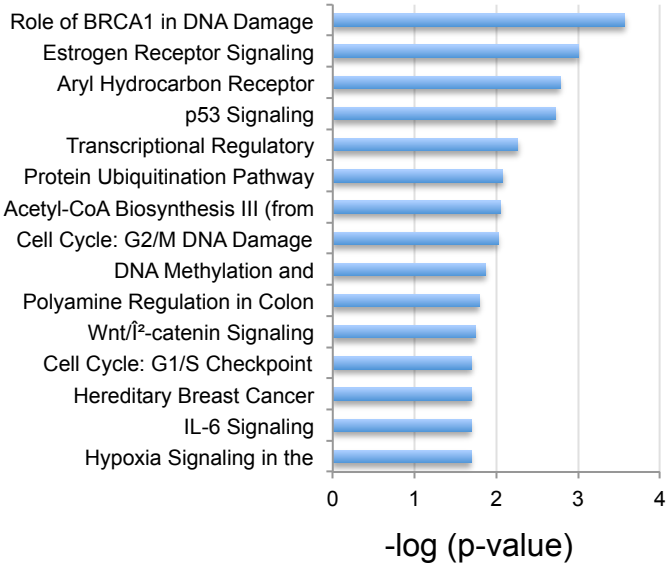
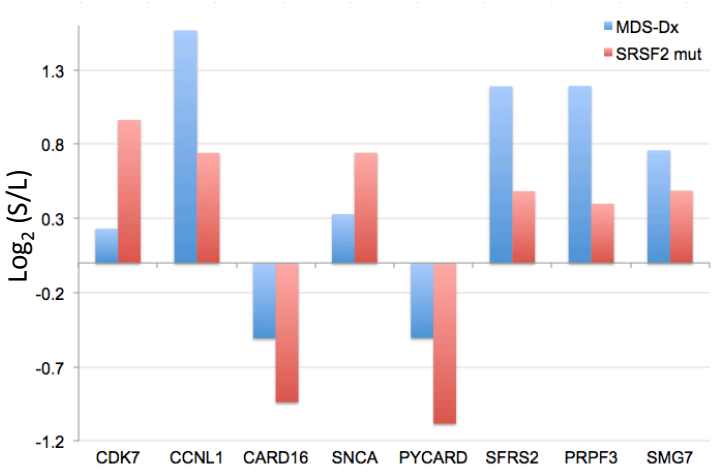


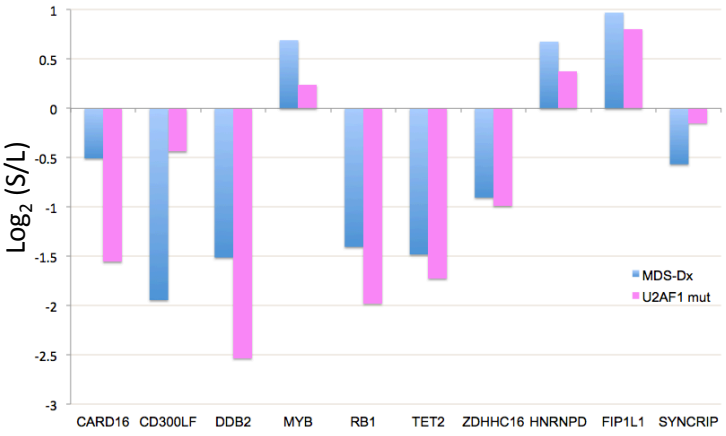
Figure S4

A



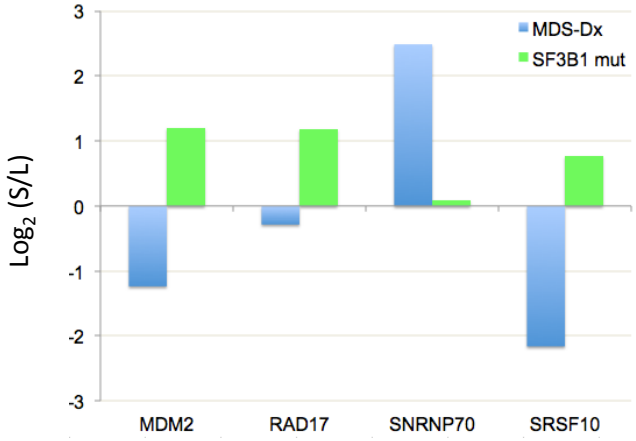
Overlapped events

B



Overlapped events

C



Overlapped events

Figure S5

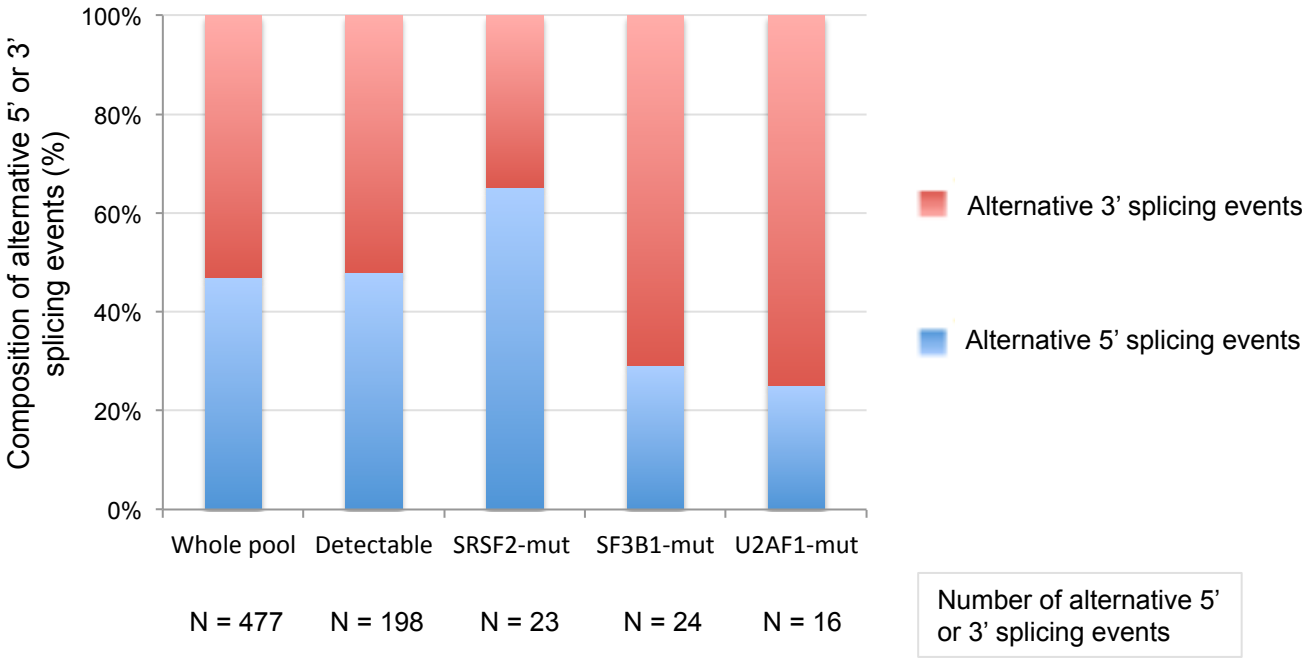

Composition of alternative 5' or 3' splicing events (%)
100%
80%
60%
40%
20%
0%
Whole pool
Detectable
SRSF2-mut
SF3B1-mut
U2AF1-mut
N = 477
N = 198
N = 23
N = 24
N = 16
Alternative 3' splicing events
Alternative 5' splicing events
Number of alternative 5' or 3' splicing events

Figure S6

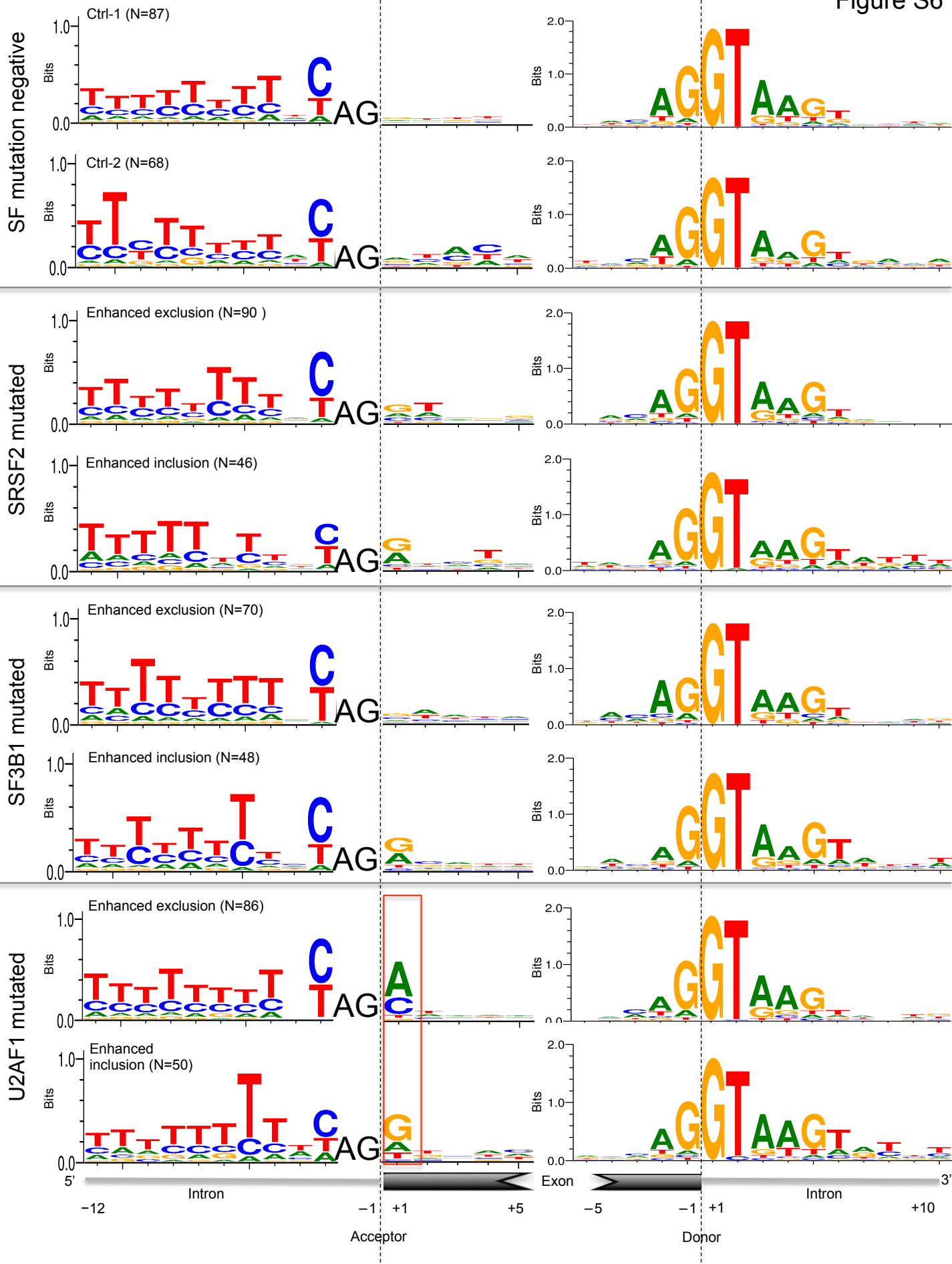
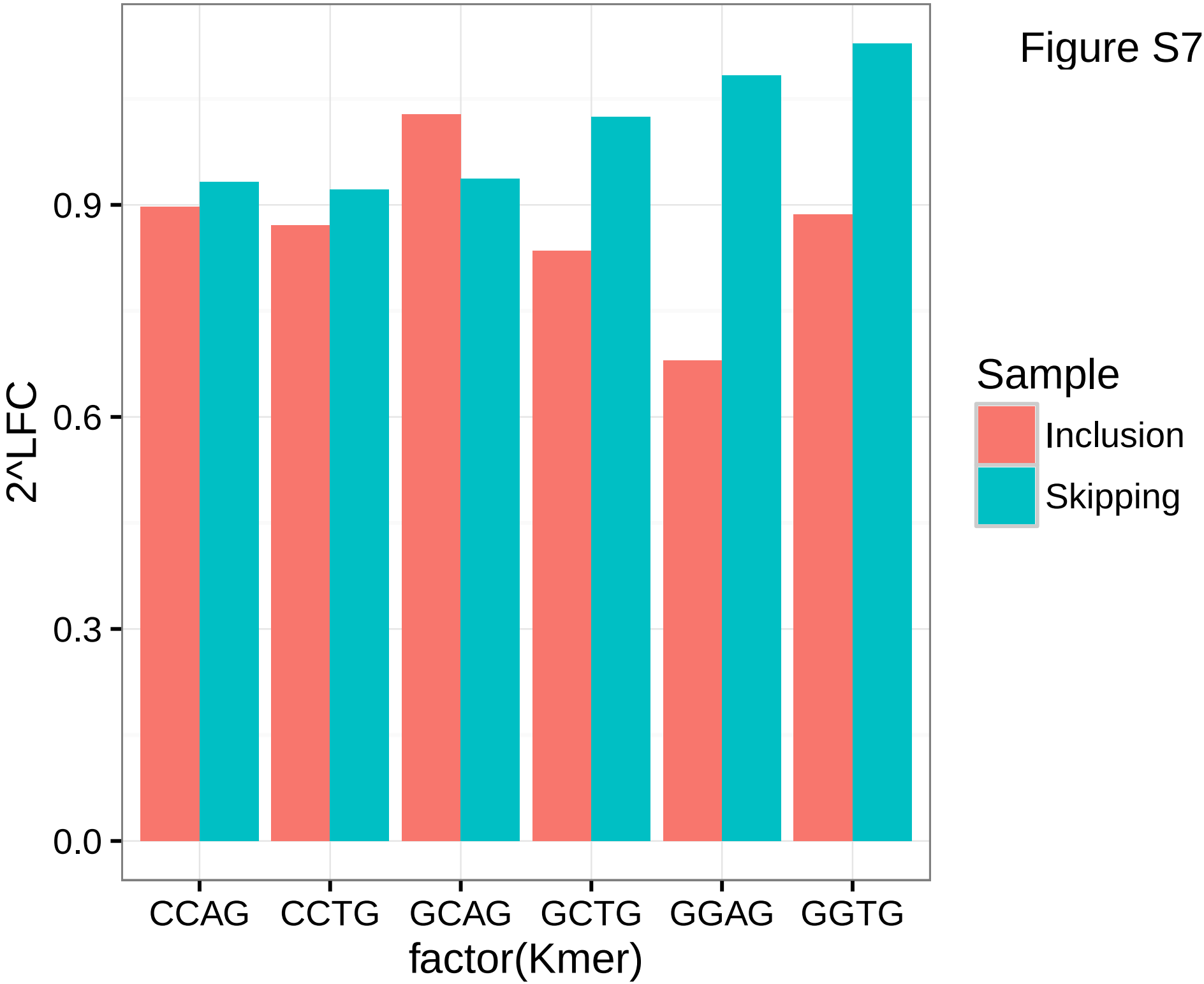
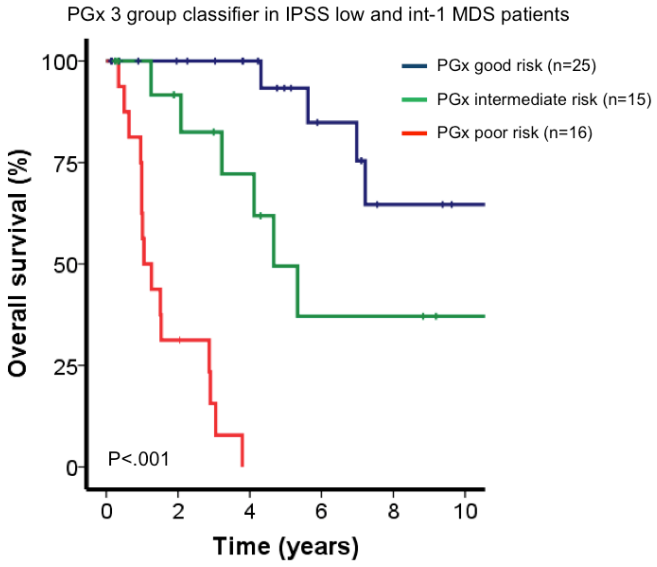

SF mutation negative
Ctrl-1 (N=87)
Ctrl-2 (N=68)
SRSF2 mutated
Enhanced exclusion (N=90)
Enhanced inclusion (N=46)
SF3B1 mutated
Enhanced exclusion (N=70)
Enhanced inclusion (N=48)
U2AF1 mutated
Enhanced exclusion (N=86)
Enhanced inclusion (N=50)
Bits
5'
Intron
Exon
3'
−12
−1
+1
+5
−5
−1
+1
+10
Acceptor
Donor

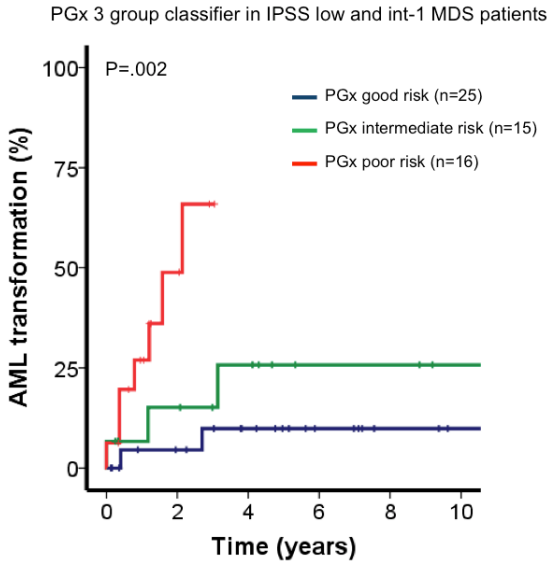
Figure S7


2^LFC
0.0
0.3
0.6
0.9
CCAG
CCTG
GCAG
GCTG
GGAG
GGTG
factor(Kmer)
Sample
Inclusion
Skipping

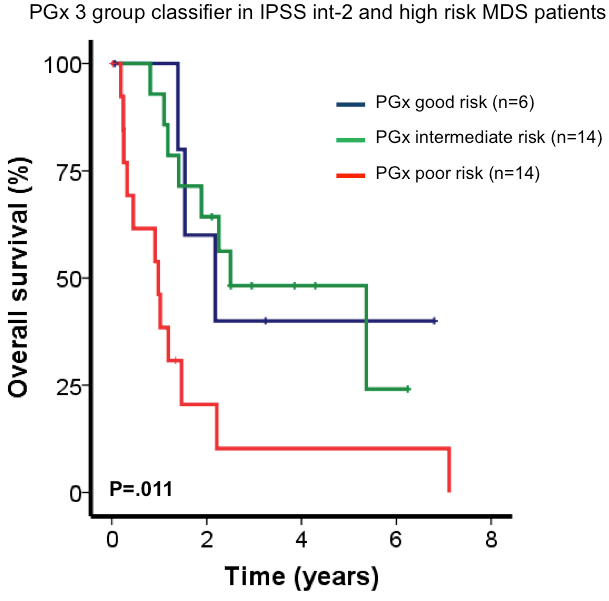
A


PGx 3 group classifier in IPSS low and int-1 MDS patients
PGx good risk (n=25)
PGx intermediate risk (n=15)
PGx poor risk (n=16)
Overall survival (%)
100
75
50
25
0
P<.001
0
2
4
6
8
10
Time (years)

C


PGx 3 group classifier in IPSS low and int-1 MDS patients
P=.002
PGx good risk (n=25)
PGx intermediate risk (n=15)
PGx poor risk (n=16)
AML transformation (%)
100
75
50
25
0
0
2
4
6
8
10
Time (years)

B


PGx 3 group classifier in IPSS int-2 and high risk MDS patients
PGx good risk (n=6)
PGx intermediate risk (n=14)
PGx poor risk (n=14)
Overall survival (%)
100
75
50
25
0
P=.011
0
2
4
6
8
Time (years)

D

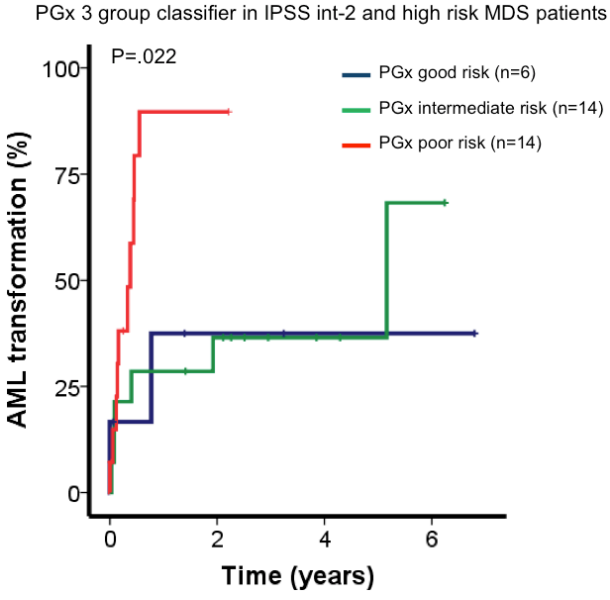

PGx 3 group classifier in IPSS int-2 and high risk MDS patients
P=.022
PGx good risk (n=6)
PGx intermediate risk (n=14)
PGx poor risk (n=14)
AML transformation (%)
100
75
50
25
0
0
2
4
6
Time (years)

Figure S9

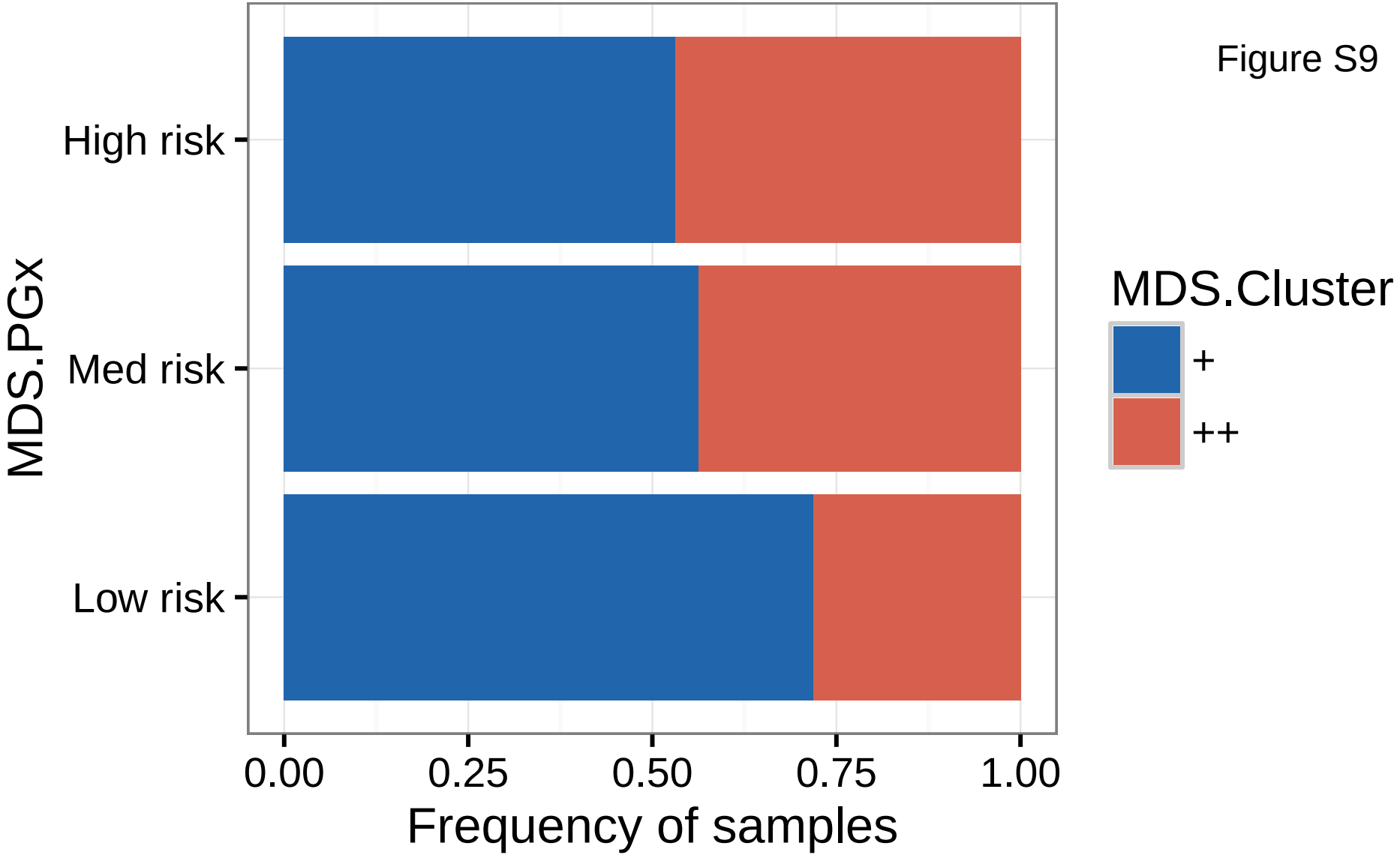

MDS.PGx
High risk
Med risk
Low risk
0.00
0.25
0.50
0.75
1.00
Frequency of samples
MDS.Cluster
+
++

Table S1. Overview of sample characteristics

Sample origin	No.	Selection/phenotype	Antibodies
MDS PB	22	Ficoll separated mononuclear cells from PB	
MDS BM	93	Ficoll separated mononuclear cells from BM	
Healthy			
BM	8	Ficoll separated mononuclear cells from BM	
CD34+_BM	5	CD34+	CD34 microbead kit
Stem cell	1	Lin-CD34+CD38-	CD34-PEcy7, CD38, FITC
CMP	2	CD34+, CD38+, CD123lo, CD45Ra	CD34-PEcy7, CD38 FITC, CD3 PECy7, CD123-APC
T cell	5	CD8-, CD56-, CD3+, CD4+, Lin-	CD8 FITC, CD56 PE, CD3 PECy7, CD123-APC
B cell	5	CD19+	CD19 APC
Granulocyte	5	CD66b+, CD15+	CD66b FITC, CD15 PE
Monocyte	5	CD14+, CD66b-, CD3-, CD56-	CD14PE, CD3 FITC
PB	18	Ficoll separated mononuclear cells from PB	

Note: 3 patients were analyzed in duplicate. PB: peripheral blood; BM: bone marrow

Table S2. Clinical characteristics of MDS patients

Characteristic	All (n=112)
Age, years	
Median	67
range	26-92
Sex	
male - no. (%)	71 (63)
female - no. (%)	41 (37)
WHO-Subtype	
del (5q) - no. (%)	11 (10)
RA - no. (%)	16 (14)
RARS - no. (%)	16 (14)
RCMD - no. (%)	18 (16)
RAEB-1 - no. (%)	11 (10)
RAEB-2 - no. (%)	25 (22)
MDS-U - no. (%)	2 (2)
Secondary AML, CMML - no. (%)	2 (2)
missing – no. (%)	10 (9)
Karyotype risk according to IPSS	
Good - no. (%)	65 (58)
Intermediate - no. (%)	10 (9)
Poor - no. (%)	19 (17)
Missing data - no. (%)	18 (16)
IPSS risk score	
Low - no. (%)	29 (26)
Int-1 - no. (%)	30 (27)
Int-2 - no. (%)	25 (22)
High - no. (%)	10 (9)
Missing data - no. (%)	18 (16)
Transfusion dependence	
Yes - no. (%)	74 (66)
No - no. (%)	22 (20)
Missing data - no. (%)	16 (14)
Transformation into AML	
Yes - no. (%)	31 (28)
No - no. (%)	64 (57)
Missing data - no. (%)	17 (15)
Allogeneic Transplantation	
yes - no. (%)	16 (14)
no - no. (%)	87 (78)
Missing data - no. (%)	9 (8)

Table S3. Event types detected by RASL-seq oligo pool

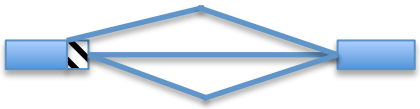
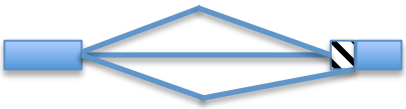
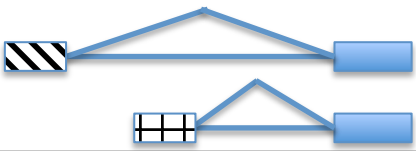
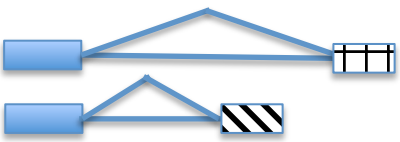
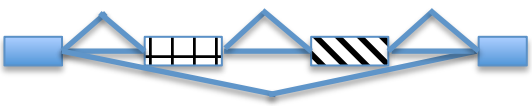
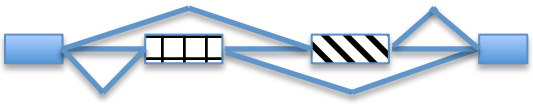
Event type	Event sketch	Event number
Cassette exon		N=4667
Alternative 5' splice site		N=224
Alternative 3' splice site		N=253
Alternative trans-cription initiation		N=136
Alternative trans-cription termination		N=40
Alternative 5' Cassette exon		N=2
Alternative 3' Cassette exon		N=3
Multi-exon skipping		N=113
Mutually exclusive exon		N=64

Table S4. RT-PCR validation primers

Primer Name	Primer sequence
DDX50_F	GACATTATGGAGCTGGAAGC
DDX50_R	CATCCAGGTCATCTGTAACAC
TFAM_F	GGAGGGAACTTCCTGATTC
TFAM_R	TCCTTTTCAGAGTCAGACAG
NUP214_F	CTCTGCTGTCAAAGTCAACC
NUP214_R	TCCTGGGTGATTTCTGTACC
RUNX1_F	CTCAGCCTCAGAGTCAGATG
RUNX1_R1	GAGAGTCGACTGGAAAGTTCTG
RUNX1_R2	CTGAGTCTCTTCTGAGGATGAG

Table S5B. Genes in pathways enriched in Hemo-SP

Ingenuity Canonical Pathways	Genes
PI3K Signaling in B Lymphocytes	FYN,PLCB1,NFATC2,MALT1,PLEKHA1,ITPR1
Phospholipase C Signaling	FYN,GNAS,PLCB1,NFATC2,ARHGEF1,ITPR1,FNBP1
Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	SMAD2,FYN,NFATC2,MALT1
HIPPO signaling	SMAD2,TJP2,YWHAZ,SKP2
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	YWHAZ,TRIP12,SKP2
GÎ±q Signaling	GNAS,PLCB1,NFATC2,ITPR1,FNBP1
fMLP Signaling in Neutrophils	GNAS,PLCB1,NFATC2,ITPR1

Table S6B. Genes in pathways enriched in MDS-Dx

Ingenuity Canonical Pathways	Genes
Role of CHK Proteins in Cell Cycle Checkpoint Control	E2F6, RAD17, MRE11A, TLK1, PPP2R5E, CHEK2, ATM
Role of BRCA1 in DNA Damage Response	RB1, E2F6, MRE11A, CHEK2, MLH1, ATM
p53 Signaling	PRKDC, RB1, ADCK3, MDM2, BAX, CHEK2, ATM
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	PRKDC, CDK7, MDM2, CHEK2, ATM
Hereditary Breast Cancer Signaling	RB1, XPC, MRE11A, DDB2, CHEK2, MLH1, ATM
ATM Signaling	MDM2, MRE11A, TLK1, CHEK2, ATM
DNA Double-Strand Break Repair by Homologous Recombination	ATRX, MRE11A, ATM
DNA Double-Strand Break Repair by Non-Homologous End Joining	PRKDC, MRE11A, ATM
Cyclins and Cell Cycle Regulation	RB1, E2F6, CDK7, PPP2R5E, ATM
Aryl Hydrocarbon Receptor Signaling	RB1, TRIP11, MDM2, BAX, CHEK2, ATM
Molecular Mechanisms of Cancer	PRKDC, RB1, E2F6, CDK7, MDM2, BAX, NFKBIB, CHEK2, CTNND1, ATM
Protein Ubiquitination Pathway	USP3, DNAJC21, DNAJC8, HSPH1, DNAJC3, MDM2, DNAJC7, VHL
Role of p14/p19ARF in Tumor Suppression	RB1, MDM2, ATM
Cell Cycle: G1/S Checkpoint Regulation	RB1, E2F6, MDM2, ATM

Table S10. Events overlapped in SRSF2, U2AF1, and SF3B1 mutation related splicing programs

Gene_symbol	Description	Chromosome Location	Event type	Event_position
MICAL3	microtubule associated monooxygenase, calponin and LIM domain containing 3	22q11.21	Cassette	chr22:16734789-16748643:-
MAP4K4	mitogen-activated protein kinase kinase kinase kinase 4	2q11.2-q12	Cassette	chr2:101853309-101856540:+
GABPB1	GA binding protein transcription factor, beta subunit 1	15q21.2	Cassette	chr15:48380857-48383454:-
PCBP2	poly(rC) binding protein 2	12q13.13	Cassette	chr12:52147344-52148827:+
USO1	USO1 vesicle transport factor	4q21.1	Cassette	chr4:76934078-76939798:+
ITSN	intersectin 1 (SH3 domain protein)	21q22.1-q22.2	Cassette	chr21:34113497-34120991:+
SLC35A1	solute carrier family 35 (CMP-sialic acid transporter), member A1	6q15	Cassette	chr6:88243976-88267594:+
TNFSF12-TNFS	TNFSF12-TNFSF13 readthrough	17p13.1	Cassette	chr17:7395019-7400818:+
PAN3	PAN3 poly(A) specific ribonuclease subunit	13q12.2	Cassette	chr13:27650072-27692367:+
SETD5	SET domain containing 5	3p25.3	Cassette	chr3:9452590-9457139:+

Table S11. Genes in pathways enriched in SRSF2 mutation-related splicing program

Ingenuity Canonical Pathways	Molecules
ATM Signaling	MAPK8, MAPK9, TLK1, TDP1, ATM
CD40 Signaling	TANK, MAPK8, MAPK9, MAPKAPK2, ATM
IL-6 Signaling	VEGFA, MAPK8, MAPK9, MAPKAPK2, MAP4K4, ATM
Role of IL-17A in Arthritis	MAPK8, MAPK9, MAPKAPK2, ATM
Pancreatic Adenocarcinoma Signaling	VEGFA, BCL2L1, MAPK8, MAPK9, ATM
Inhibition of Angiogenesis by TSP1	VEGFA, MAPK8, MAPK9
14-3-3-mediated Signaling	STRADA, MAPK8, MAPK9, SNCA, ATM
IL-17 Signaling	MAPK8, MAPK9, MAPKAPK2, ATM
IL-8 Signaling	VEGFA, BCL2L1, MAPK8, MAPK9, MAP4K4, ATM
FcÎ³RIIB Signaling in B Lymphocytes	MAPK8, MAPK9, ATM
UVA-Induced MAPK Signaling	BCL2L1, MAPK8, MAPK9, ATM
SAPK/JNK Signaling	MAPK8, MAPK9, MAP4K4, ATM
CD27 Signaling in Lymphocytes	BCL2L1, MAPK8, MAPK9
p53 Signaling	BCL2L1, MAPK8, HIPK2, ATM
UVB-Induced MAPK Signaling	MAPK8, MAPK9, ATM

Table S12. Genes in pathways enriched in U2AF1 mutation-related splicing program

Ingenuity Canonical Pathways	Molecules
Role of BRCA1 in DNA Damage Response	RB1, FANCC, BRCC3, CHEK2, SMARCA4
Estrogen Receptor Signaling	TAF1, CCNC, MED15, HNRNPD, NCOR2, SMARCA4
Aryl Hydrocarbon Receptor Signaling	RB1, MDM2, NCOR2, CHEK2, SMARCA4, ARNT
p53 Signaling	RB1, TP53INP1, MDM2, CTNNB1, CHEK2
Transcriptional Regulatory Network in Embryonic Stem Cells	RIF1, MEIS1, STAT3
Protein Ubiquitination Pathway	USO1, USP15, FBXW7, MDM2, BTRC, DNAJC1, UBE2F
Acetyl-CoA Biosynthesis III (from Citrate)	ACLY
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	MDM2, BTRC, CHEK2
DNA Methylation and Transcriptional Repression Signaling	MTA1, DNMT1
Polyamine Regulation in Colon Cancer	SAT1, CTNNB1
Wnt/Î²-catenin Signaling	MAP3K7, MAP4K1, MDM2, BTRC, CTNNB1
Cell Cycle: G1/S Checkpoint Regulation	RB1, MDM2, BTRC
Hereditary Breast Cancer Signaling	RB1, FANCC, CHEK2, SMARCA4
IL-6 Signaling	IL6R, MAP3K7, STAT3, MAP4K4
Hypoxia Signaling in the Cardiovascular System	MDM2, UBE2F, ARNT

Table S13. Genes in pathways enriched in SF3B1 mutation-related splicing program

Ingenuity Canonical Pathways	Molecules
p53 Signaling	BCL2L1, MDM4, PIK3R1, APAF1, MDM2, CHEK2, FAS
Apoptosis Signaling	ACIN1, BCL2L1, APAF1, MAP4K4, FAS, MCL1
ATM Signaling	MDM4, MDM2, MRE11A, TLK2, CHEK2
G2/M DNA Damage Checkpoint Regulation	MDM4, CDK7, MDM2, CHEK2
Death Receptor Signaling	ACIN1, TANK, APAF1, MAP4K4, FAS
CHK in Cell Cycle Checkpoint Control	RAD17, MRE11A, TLK2, CHEK2
CD40 Signaling	TANK, PIK3R1, MAP2K3, STAT3
IL-6 Signaling	PIK3R1, MAP2K3, STAT3, MAP4K4, MCL1
Docosahexaenoic Acid (DHA) Signaling	BCL2L1, PIK3R1, APAF1
Clathrin-mediated Endocytosis Signaling	PICALM, PIK3R1, SH3GLB1, DAB2, MDM2, SH3KBP1
Protein Ubiquitination Pathway	USO1, DNAJC21, UBE2L3, DNAJC3, MDM2, PSMA3, VHL
Granzyme B Signaling	NUMA1, APAF1
Histamine Biosynthesis	HDC
Amyotrophic Lateral Sclerosis Signaling	BCL2L1, PIK3R1, CASP1, APAF1
CD27 Signaling in Lymphocytes	BCL2L1, APAF1, MAP2K3

Table S14. MDS-PGx Associated With Patient Survival Identified Using Coxnet

Gene_symbol	Description	Event_type	Event_position	In or out of frame*	Within frame (Yes/No)
BCAS3	breast carcinoma amplified sequence 3	Cassette	chr17:56800637-56824119:+	In frame	No
PROM1	prominin 1	Cassette	chr4:15644230-15649666:-	In frame	N/A
RNF14	ring finger protein 14	Cassette	chr5:141333491-141338051:+	Out of frame	
MBTD1	mbt domain containing 1	Cassette	chr17:46657569-46692194:-	Out of frame	
CSNK1E	casein kinase 1, epsilon	Cassette	chr22:37087511-37106320:-	In frame	N/A
GOLM1	golgi membrane protein 1	Cassette	chr9:87851307-87882146:-	Out of frame	
CCM2	cerebral cavernous malformation 2	Cassette	chr7:45006487-45070041:+	In frame	Yes
CDCA2	cell division cycle associated 2	Cassette	chr8:25383413-25393345:+	In frame	No
TAF4B	TAF4b RNA polymerase II, TATA box binding protein (TBP)-associated factor, 105kDa	Cassette	chr18:22127493-22149190:+	In frame	Yes
SLC35A1	solute carrier family 35 (CMP-sialic acid transporter), member A1	Cassette	chr6:88243976-88267594:+	Out of frame	
ABI2	abl-interactor 2	Cassette	chr2:203940012-203953175:+	Out of frame	

* related events are in or out of frame

** For in frame events, if they are located within a reported functional protein domain/ N/A: no available domain information

Table S15. Descriptive statistics for clinical covariates among risk groups classified by MDS-PGx

		PGx good risk (n=29)	PGx intermediate risk (n=29)	PGx poor risk (n=30)	Total (n=88)*	p-value
IPSS						0.009**
	low risk	14(48.3%)	10(34.5%)	2(6.7%)	26(29.5%)	
	intermediate-1 risk	8(27.6%)	6(20.7%)	14(46.7%)	28(31.8%)	
	intermediate-2 risk	6(20.7%)	10(34.5%)	10(33.3%)	26(29.5%)	
	high risk	1(3.4%)	3(10.3%)	4(13.3%)	8(9.1%)	
Age		60.34(11.57)	67.83(9.50)	67.40(10.86)	65.22(11.10)	0.014*
Hgb		9.94(1.55)	9.09(2.02)	8.80(1.42)	9.27(1.73)	0.030*
Blasts.BM						0.304
	<5%	22(75.9%)	15(51.7%)	17(56.7%)	54(61.4%)	
	5%-10%	2(6.9%)	7(24.1%)	5(16.7%)	14(15.9%)	
	10%-20%	5(17.2%)	7(24.1%)	8(26.7%)	20(22.7%)	
	>20%	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	
Cytogenetics						0.718
	low risk	20(69.0%)	21(72.4%)	19(63.3%)	60(68.2%)	
	intermediate risk	5(17.2%)	2(6.9%)	4(13.3%)	11(12.5%)	
	high risk	4(13.8%)	6(20.7%)	7(23.3%)	17(19.3%)	

* only 88 out of 96 patients with all the clnical information were analyzed

Mean (SD) for continuous data. Count (%) for categorical data.

p-values are calculated based on a F-test for continuous variable, and a Fisher's exact test for categorical variable.

Hgb: Hemoglobin; BM: bone marrow

Table S16. Univariate Cox regression model analysis for MDS-PGx, IPSS, Age, Hgb, Blasts-BM, Cytogenetics, and Gender

	Comparator	Odds Ratio(OR)	95% CI for OR	Pr(>|z|)	LRT
MDS-PGx	good risk				<.001
intermediate risk		2.4	[0.91,6.33]	0.077	
poor risk		12.91	[5.16,32.32]	<.001	
IPSS	low risk				0.002
intermediate-1 risk		2.31	[1.01,5.30]	0.049	
intermediate-2 risk		2.63	[1.14,6.05]	0.023	
high risk		7.74	[2.77,21.63]	<.001	
Age	< 67				0.032
≥ 67		1.91	[1.05,3.48]	0.034	
Hgb	< 8				0.092
8 ≤ Hgb <10		0.57	[0.28,1.16]	0.121	
Hgb ≥10		0.4	[0.18,0.90]	0.026	
Blasts-BM	< 5%				0.079
5%-10%		2.19	[0.96,4.96]	0.061	
> 10%-20%		1.88	[0.96,3.71]	0.068	
Cytogenetics	low risk				0.108
intermediate risk		1.43	[0.59,3.47]	0.433	
high risk		2.14	[1.08,4.22]	0.028	
Gender	male				0.218
female		0.68	[0.36,1.27]	0.228	

LTR: likelihood ratio test (LRT)

Variables with more than 2 categories: Odds ratios greater than or less than 1 indicate an increased or decreased risk, respectively, of an event for the category listed compared to the category listed in the first row of each variable, which has a OR of 1.0.

Variabels with 2 categories: Odds ratios greater than or less than 1 indicate an increased or decreased risk, respectively, of an event for the first category listed.

Hgb: Hemoglobin, BM: bone marrow

Table S17. Baseline characteristics of patients according to MDS-PGx score

	good risk	intermediate risk	poor risk	p value
	(n=32)	(n=32)	(n=32)	
Age				0.17
median (years)	65	68	66	
Sex				0.097
male - no. (%)	19 (59)	17 (53)	25 (78)	
female - no. (%)	13 (41)	15 (47)	7 (22)	
IPSS risk score				0.006
low - no. (%)	16 (52)	9 (31)	2 (7)	
intermediate I - no. (%)	9 (29)	6 (21)	14 (47)	
intermediate II - no. (%)	5 (16)	10 (35)	10 (33)	
high - no. (%)	1 (3)	4 (14)	4 (13)	
WHO type				0.274
RA/RCMD - no. (%)	13 (41)	5 (16)	9 (28)	
RARS/RCMD-RS - no. (%)	7 (22)	4 (13)	5 (16)	
5q- syndrome - no. (%)	3 (9)	5 (16)	3 (9)	
RAEBI or II - no. (%)	6 (19)	16 (50)	12 (38)	
Unknown - no. (%)	3 (9)	2 (6)	3 (9)	
Blasts in BM				0.113
<5% - no. (%)	26 (81)	16 (52)	18 (56)	
5-9% - no. (%)	2 (6)	7 (23)	5 (16)	
10-20% - no. (%)	4 (13)	8 (26)	9 (28)	
Blasts in PB				0.533
<5% - no. (%)	32 (100)	29 (94)	31 (97)	
5-9% - no. (%)	0 (0)	1 (3)	1 (3)	
10-20% - no. (%)	0 (0)	1 (3)	0 (0)	
Cytogenetic risk (IPSS)				0.687
good - no. (%)	21 (70)	20 (69)	20 (65)	
intermediate - no. (%)	5 (17)	2 (7)	4 (13)	
high risk - no. (%)	4 (13)	7 (24)	7 (23)	
Transfusion dependency				0.115
no - no. (%)	12 (38)	7 (22)	5 (16)	
yes - no. (%)	20 (63)	25 (78)	27 (84)	
Allogeneic Transplantation				0.377
no - no. (%)	25 (78)	26 (81)	29 (91)	
yes - no. (%)	7 (22)	6 (19)	3 (9)	
SF3B1				0.43
wildtype - no. (%)	22 (69)	22 (69)	26 (81)	
mutated - no. (%)	10 (31)	10 (31)	6 (6)	
SRSF2				0.006
wildtype - no. (%)	29 (91)	25 (78)	18 (56)	
mutated - no. (%)	3 (9)	7 (22)	14 (44)	
U2AF1				0.751
wildtype - no. (%)	28 (88)	29 (91)	27 (84)	
mutated - no. (%)	4 (13)	3 (9)	5 (16)	

ZRSR2				0.594
wildtype - no. (%)	18 (95)	19 (86)	26 (93)	
mutated - no. (%)	1 (5)	3 (14)	2 (7)	
ASXL1				0.299
wildtype - no. (%)	28 (88)	24 (77)	23 (72)	
mutated - no. (%)	4 (12)	7 (23)	9 (28)	
RUNX1				0.332
wildtype - no. (%)	32 (100)	27 (96)	29 (100)	
mutated - no. (%)	0 (0)	1 (3)	0 (0)	
IDH1				0.587
wildtype - no. (%)	31 (97)	30 (94)	29 (91)	
mutated - no. (%)	1 (3)	2 (6)	3 (9)	
IDH2				n.a.
wildtype - no. (%)	32 (100)	32 (100)	32 (100)	
mutated - no. (%)	0 (0)	0 (0)	0 (0)	
NPM1				0.352
wildtype - no. (%)	32 (100)	30 (97)	32 (100)	
mutated - no. (%)	0 (0)	1 (3)	0 (0)	
DNMT3A				0.595
wildtype - no. (%)	30 (97)	32 (100)	31 (97)	
mutated - no. (%)	1 (3)	0 (0)	1 (3)	

Supplemental information

Figure Legends

Figure S1. Overview of the distribution of RASL-seq sequence counts by samples and events. (*A*) The mean and (*B*) median distribution of RASL-seq sample counts per event. (*C*) The mean and (*D*) median distribution of RASL-seq event counts per sample.

Figure S2. Effect of alternatively spliced events in the MDS-Dx gene panel on gene product function. (*A*) Demonstration of the four examples of out-of-frame, alternatively spliced events found in the MDS-Dx panel. (*B*) Demonstration of two examples of in-frame events from the MDS-Dx panel, whose alternatively spliced events are located inside of functional protein domains. Green: coding region; light green: non-coding region. Violet: non-coding RNA. Grey arrowhead: transcription direction. Red triangle: alternatively spliced exon. The top isoform shows the events enriched in normal samples, while the bottom isoform shows the events enriched in MDS samples.

Figure S3. Pathway analysis of splicing factor mutation-related programs. Ingenuity IPA analysis of *SRSF2* (*A*), *U2AF1* (*B*), and *SF3B1* (*C*) mutation related splicing programs, respectively.

Figure S4. Comparison of splicing factor mutation-related programs with MDS-Dx. Events related to disease progression in *SRSF2*-mutated (*A*), *U2AF1*-mutated (*B*), and *SF3B1*-mutated (*C*) programs were compared to that of MDS-Dx.

Figure S5. Influences of splicing factor mutations on alternative 5' and 3' splicing. The number of alternative 5' or alternative 3' splicing events in control (total pool, and detectable pool) was compared with that in different splicing factor mutation-related programs.

Figure S6. Acceptor and donor site consensus sequence analysis of splicing factor mutation-related splicing events. The consensus acceptor and donor sequences of the isoforms in specific splicing factor-mutation induced programs (with average event-ratio change > 1.5 fold) were collected by extracting 30 bps of exon sequences neighboring the acceptor and donor junctions together with related 200 bps of upstream and downstream intron sequences. The logo of base enrichment was generated using Sequencing Logo. Ctrl: control.

Figure S7. Indemnification of enriched motifs in the *SRSF2* mutation-affected splicing program. We first identified 140 increased exon inclusion and 123 increased skipping events that were linked to *SRSF2* mutations ($p<0.01$, $|\text{Fold Change}| >= 2$). We next calculated relative 4-mer enrichment over random background between the two groups, showing selective enrichment of GGNG motifs in induced exon skipping events.

Figure S8. IPSS lower risk and higher risk MDS patients restratified by their MDS-PGx score. (*A*) Overall survival of IPSS lower risk patients (IPSS-Low and Intermediate-1) reclassified into 3 risk groups by the MDS-PGx score. (*B*) Overall survival of IPSS

higher risk patients (IPSS Intermediate-2, and High) reclassified into 3 risk groups by the MDS-PGx score. (*C*) Time to AML transformation of IPSS lower risk patients (IPSS-Low and Intermediate-1) reclassified into 3 risk groups by the MDS-PGx score. (*D*) Time to AML transformation of IPSS higher risk patients (IPSS Intermediate-2, and High) reclassified into 3 risk groups by the MDS-PGx score.

List of Supplemental Tables