Supplementary Table 1. Characteristics of patients found to harbor U2AF1 mutations.

ID	Dx	Cytogenetics	Sex	U2AF1 mutation	Exome seq	RNA seq
1	sAML	46,XX[20]	F	c.101C>A,p.S34Y,c.104G>T,p.R35L	0	0
2	pAML(M2)	45,XY,-7[17]/45,idem,del(12)(p12)[3]	M	c.101C>A, p.S34Y	0	0
3	pAML(M4)	46,XY[20]	M	c.101C>T, p.S34F	1	0
4	pAML(M4)	46,XY[20]	M	c.101C>T, p.S34F	1	0
5	pAML(M4)	46,XY, add(8)	M	c.101C>T, p.S34F	1	1
6	pAML(M5)	46,XY, add(8)	M	c.101C>A, p.S34Y	1	1
7	pAML(M5)	47,XY,add(3)(p21),del(6)(p21.3),del(11)(q23),del(13)(q14q22),add(14)(p10),del(15)(q22),+19,add(19)(p13.3)[18]/46,XY[2].	M	c.467A>G, p.R156Q	0	0
8	pAML(M5)	45,X,-Y[22]/46,XY[2]	M	c.101C>T, p.S34F	0	0
9	pAML(M6)	N/A	M	c.101C>T, p.S34F	1	1
10	sAML	46,XY[20]	M	c.101C>T, p.S34F	1	0
11	sAML	47,XY,+19[20]	M	mutated	0	0
12	pAML	46,XY,-7,+14[17]/46,XY[3]	M	c.470A>C, p.Q157P	0	0
13	sAML	46,XY,del(20)(q11.2)[12].	M	c.101C>T, p.S34F	1	0
14	sAML	46,X,-X[17]/46,XX[3].	F	c.470A>C, p.Q157P	0	0
15	sAML	46,XY,del(7)(q11.2)[10]/46,XY,add(12)(q24.3)[7]/46,XY[3]	M	c.470A>C, p.Q157P	0	0
16	sAML	46,XY,del(20)(q11.2)[6]/46,idem,del(7)(p12)[2]/ 46,XY,ider(20)(q10)del(20)(q11.2)[8]/46,XY[4]	M	c.101C>T, p.S34F	0	0
17	sAML	46,XY[20]	M	c.101C>T, p.S34F	0	0
18	sAML	46,XX,del(17)(q24)[6]/46,XX[24]	F	c.101C>T, p.S34F	0	0
19	pAML(M1)	46,XY,del (7q) / 7q-	M	c.101C>T, p.S34F	1	1
20	pAML(M1)	47,XY, add(8)	M	c.470A>C, p.Q157P	1	0
21	pAML(M1)	46,XY,t(8;21)	M	C>G p.G213A	1	0
22	pAML(M1)	47,XY,+11[20]	M	c.101C>T, p.S34F	0	0
23	pAML(M1)	46,XY,-3,-4,der(4)add(4)(p16)add(4)(q35),del(5)(q13q33),add(5)(p13),-7,del(9)(q31),-13,add(15)(p13),der(17)t(13;17)(q12;p11.2),-21,-22,-22,+mar1,+mar2[cp17]/41,idem,i(13)(q10)[3].	M	whole gene deletion	0	0
24	pAML(M2)	46,XY[20]	M	c.101C>T, p.S34F	1	1
25	pAML(M2)	46,XX[20],del (7q) / 7q-	F	c.101C>A, p.S34Y	1	1
26	sAML	46,XY[18]	M	c.251A>C p.Q84P	1	0
27	CMML-1	46,XY,?inv(20)(q11.2q13)[20]	M	mutated	0	0
28	CMML-1	46,XY[20]	M	c.467A>G, p.R156Q	0	0
29	CMML-1	46,XY[17]	M	mutated	0	0
30	CMML-1	45,XY,-7[18]/46,XY[2]	M	c.470A>C, p.Q157P	0	0
31	CMML-2	46,XY[20]	M	c.470A>C, p.Q157P	0	0
32	CMML-2	46,XY,del(20)(q11.2q13.3)[19]/46,XY[1].	M	c.101C>T, p.S34F	0	0
33	CMML-2	46,XY[20]	M	c.101C>T, p.S34F	0	0
34	RAEB	44,XY,-3,-5,-7,+9,del(11)(p12)[2]/43,idem,-9,-18,+mar[4]/46,XY[8]	M	c.470A>C, p.Q157P	0	0
35	RAEB	47, XX, del(3)(q21),del(5)(q13),del(12)(q21q24.3),+mar1[7]/45,sl,+3,-del(3),-7,-18,-mar1,+mar2[cp12 one is 4n]/46XX[1]	F	c.470A>C, p.Q157P	0	0
36	RAEB	46,XY,del(20)(q11.2)[2]/46,idem,-4,+8[1]/46,XY,der(20)del(20)(p11.2)del(20)(q11.2)[2]/46,XY[6].	M	c.101C>T, p.S34F	0	0
37	RAEB	46,XX,?add(7)(q22) or ?del(7)(q22q22)[5]/46,XX[15]	F	c.101C>T, p.S34F	0	0
38	RARS	Del(5), -7, -20, +19, +22	M	c.101C>T, p.S34F	0	0
39	RCMD	46,XY,del(20)(q11.2)[3]/46,idem,del(5)(q12q33)[15]/46,XY[2]	M	c.101C>T, p.S34F	0	0
40	RCUD MDS	46,XX,del(20)(q11.2)[20]	M	c.470A>G p.Q157R	1	0
41	tMDS	46,XX,add(12)(p13)[6]/46,XX[14]	F	c.470A>C, p.Q157P	0	0

Supplementary Table 2. Mutational and karyotypic information on abnormally spliced genes.

Gene	Chromosome	exon start	exon end	Change type	Average Difference (mut - WT)	COSMIC FS	COSMIC miss	COSMIC stop	COSMIC total	COSMIC %	Inhouse FS	Inhouse miss	Inhouse non	Inhouse total	Inhouse %	Deletion (N)	Deletion (all tested)	Deletion %	Mutations in the same exon (N)
RIPK2	chr8	90844193	90844347	Exon skipping	42.20%	0	5	1	528	1.14%	0	0	0	276	0.00%	4	756	0.53%	1
CEP164	chr11	116739354	116739432	Exon skipping	34.40%	2	16	1	86	22.09%	0	2	0	276	0.72%	7	756	0.93%	0
SUOX	chr12	54679383	54679509	Exon skipping	29.70%	1	4	0	95	5.26%	0	0	0	276	0.00%	3	756	0.40%	0
SCARB1	chr12	123833181	123833310	Exon skipping	29.11%	0	7	0	320	2.19%	0	2	0	276	0.72%	6	756	0.79%	0
SNX14	chr6	86305274	86305301	Exon skipping	27.64%	1	7	0	236	3.39%	0	0	0	276	0.00%	3	756	0.40%	0
CCDC18	chr1	93422122	93422291	Exon skipping	24.89%	1	13	2	149	10.74%	0	0	0	276	0.00%	4	756	0.53%	1
TPD52L2	chr20	61977612	61977672	Exon skipping	24.42%	0	4	0	95	4.21%	0	1	0	276	0.36%	1	756	0.13%	1
RTCD1	chr1	100512967	100513109	Exon skipping	24.37%	0	3	1	94	4.26%	0	0	0	276	0.00%	2	756	0.26%	0
C12orf11	chr12	26958607	26958778	Exon skipping	24.23%	1	13	1	168	8.93%	0	0	1	276	0.36%	17	756	2.25%	0
PTBP1	chr19	756491	756569	Exon skipping	24.17%	0	11	1	325	3.69%	0	0	0	276	0.00%	2	756	0.26%	2
SLC12A6	chr15	32337134	32337281	Exon skipping	23.51%	1	14	4	530	3.58%	0	0	0	276	0.00%	4	756	0.53%	0
STRAP	chr12	15927741	15927877	Exon skipping	22.51%	0	3	0	230	1.30%	0	0	0	276	0.00%	40	756	5.29%	1
MNX1	chr7	150564426	150564609	Exon skipping	22.40%	0	2	0	749	0.27%	0	0	0	276	0.00%	146	756	19.31%	0
FGD3	chr9	94821933	94821964	Exon skipping	22.20%	0	6	0	74	8.11%	0	0	0	276	0.00%	13	756	1.72%	0
CEP110	chr9	122896986	122897117	Exon skipping	21.98%	0	16	2	763	2.36%	0	1	0	276	0.36%	7	756	0.93%	0
NARG1L	chr13	40808785	40808892	Exon skipping	21.88%	0	13	1	104	13.46%	0	1	0	276	0.36%	29	756	3.84%	0
PM20D2	chr6	83920331	83920481	Exon skipping	21.42%	0	4	0	73	5.48%	0	0	0	276	0.00%	3	756	0.40%	0
AP2M1	chr3	185381126	185381132	Exon skipping	19.81%	0	6	0	97	6.19%	0	0	0	276	0.00%	5	756	0.66%	0
ARID2	chr12	44516790	44517041	Exon skipping	19.66%	13	34	15	278	22.30%	1	0	3	276	1.45%	10	756	1.32%	6
TLE4	chr9	81511481	81511634	Exon skipping	19.44%	0	19	2	544	3.86%	0	0	0	276	0.00%	13	756	1.72%	2
PPWD1	chr5	64903421	64903869	Exon skipping	19.08%	0	6	1	98	7.14%	0	0	0	276	0.00%	32	756	4.23%	0
TMEM131	chr2	97777824	97778010	Exon Skipping	18.41%	1	20	1	91	24.18%	0	2	0	276	0.72%	0	756	0.00%	0
KCTD20	chr6	36550543	36550817	Exon Skipping	16.43%	2	6	0	98	8.16%	0	0	0	276	0.00%	1	756	0.13%	1
ZSCAN29	chr15	41448412	41448617	Exon Skipping	16.39%	1	10	0	101	10.89%	0	0	0	276	0.00%	3	756	0.40%	3
PABPC4	chr1	39802094	39802181	Exon Skipping	15.80%	0	11	0	79	13.92%	0	0	0	276	0.00%	2	756	0.26%	0
EHMT1	chr9	139766603	139766681	Exon Skipping	15.69%	0	15	0	168	8.93%	0	2	0	276	0.72%	3	756	0.40%	0
WAC	chr10	28924667	28924976	Exon Skipping	15.11%	3	9	0	168	7.14%	1	0	0	276	0.36%	3	756	0.40%	1
UPF3B	chrX	118858636	118858675	Exon Retention	-15.37%	1	5	3	141	6.38%	0	0	0	276	0.00%	8	756	1.06%	0
SLC25A32	chr8	104494826	104494935	Exon Retention	-18.62%	0	5	0	95	5.26%	0	0	0	276	0.00%	3	756	0.40%	0
OPA1	chr3	194818266	194818320	Exon Retention	-19.57%	0	6	0	74	8.11%	0	0	0	276	0.00%	4	756	0.53%	0
GNAS	chr20	56907390	56907435	Exon Retention	-23.36%	0	420	2	5439	7.76%	1	1	0	276	0.72%	6	756	0.79%	0
PIKFYVE	chr2	208877814	208877982	Exon Retention	-25.64%	0	39	5	708	6.21%	0	2	0	276	0.72%	5	756	0.66%	0
NUCB2	chr11	17269362	17269488	Exon Retention	-26.46%	0	3	1	94	4.26%	0	2	0	276	0.72%	6	756	0.79%	0
ITGB3BP	chr1	63685823	63685873	Exon Retention	-33.57%	0	3	1	95	4.21%	0	0	0	276	0.00%	3	756	0.40%	0
ATR	chr3	143651997	143652134	Exon Retention	-45.38%	3	47	5	1160	4.74%	0	0	0	276	0.00%	14	756	1.85%	0

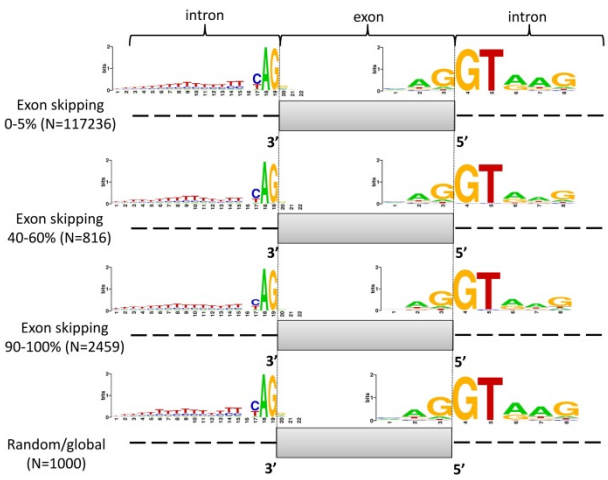
Supplementary Table 3. U2AF1 mutational status of patients with myeloid malignancies

WHO classification	U2AF1 mut	tested
MDS	12	116
Low grade	7	79
RA/RCMD	5	48
RARS/RCMD-RS	2	31
Advanced grade	5	37
RAEB1/2	5	37
MDS/MPN	12	76
CMML1/2	12	76
AML	22	332
Primary AML	15	259
Secondary AML	7	73

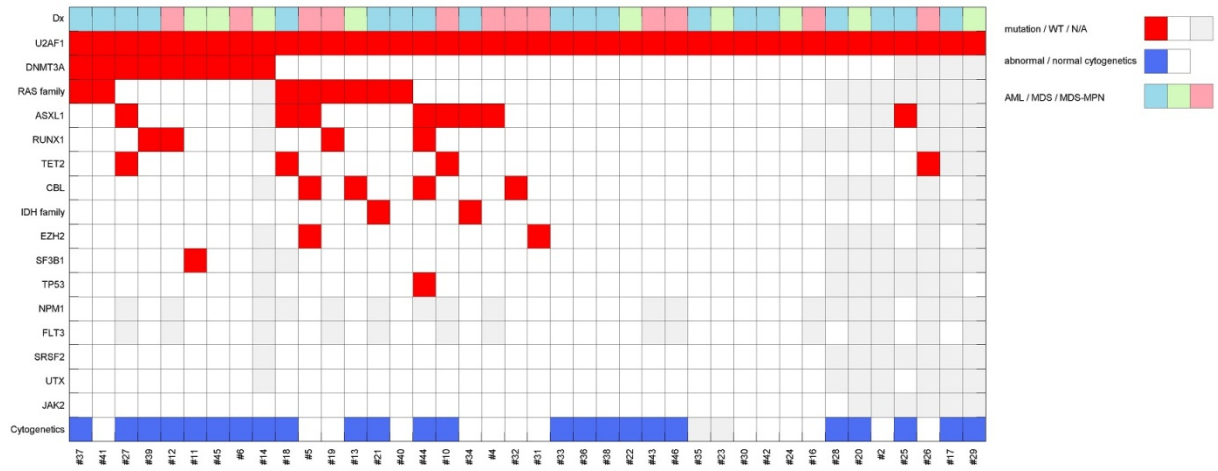
Supplementary Table 1. Characteristics and data types available for patients that were found to carry U2AF1 mutations.

Supplementary Table 2. Exon usage levels in genes that were found to be differentially spliced in U2AF1 mutants. Provided is additional information about the number of mutations and copy number abnormalities found in these genes in public databases (COSMIC) as well as from an in-house database of patients.

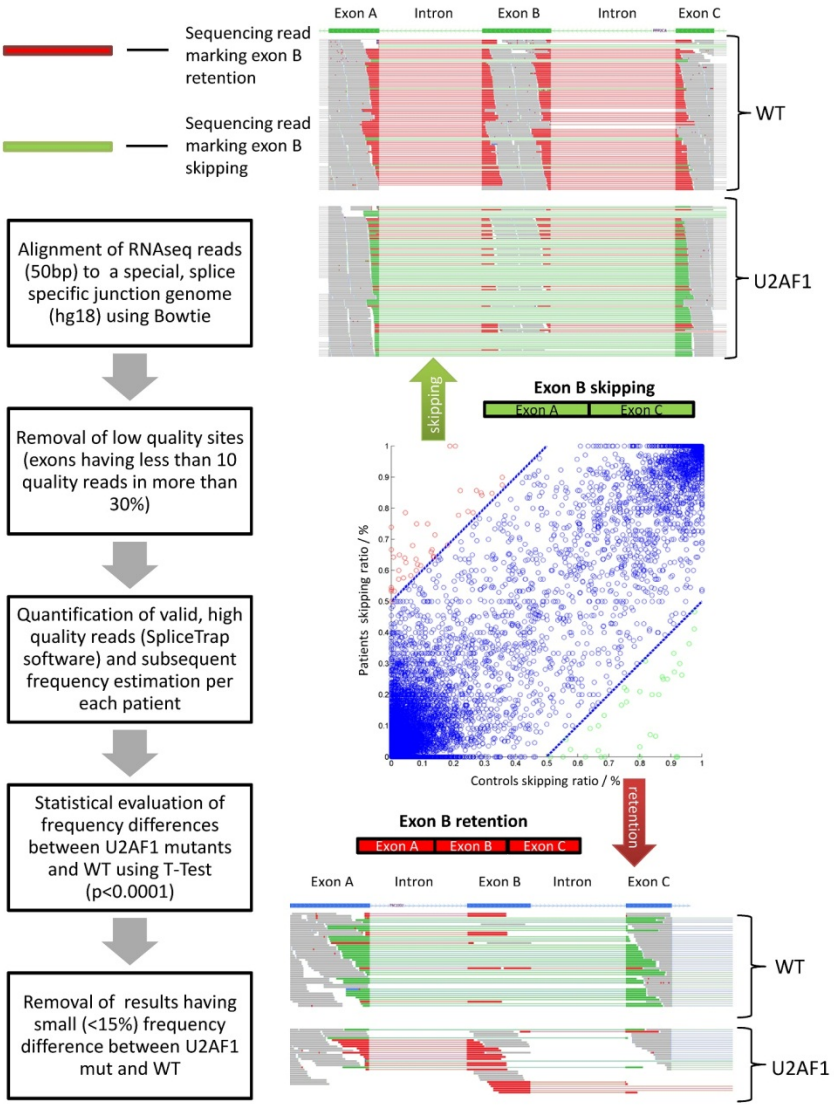
Supplementary Table 3. Characteristics of patients used in this study. The total number of patients studied is 524, with 116 patients diagnosed with MDS, 76 with MDS/MPN and 332 with AML. The total number of patients harboring U2AF1 mutations is 46 (12 MDS, 12 MDS/MPN and 22 AML).



Supplementary Figure 1. Frequencies of nucleotides at 3' and 5' splice sites flanking exons in control RNA samples. The top three lines show all exons grouped by their frequency of skipping. The bottom line shows a randomly selected set of 1,000 exons. Nucleotide frequencies are represented using WebLogo software. The height of each stack represents the information content of that position in bits. The height of each letter represents the frequency of occurrence of each nucleotide at each position.

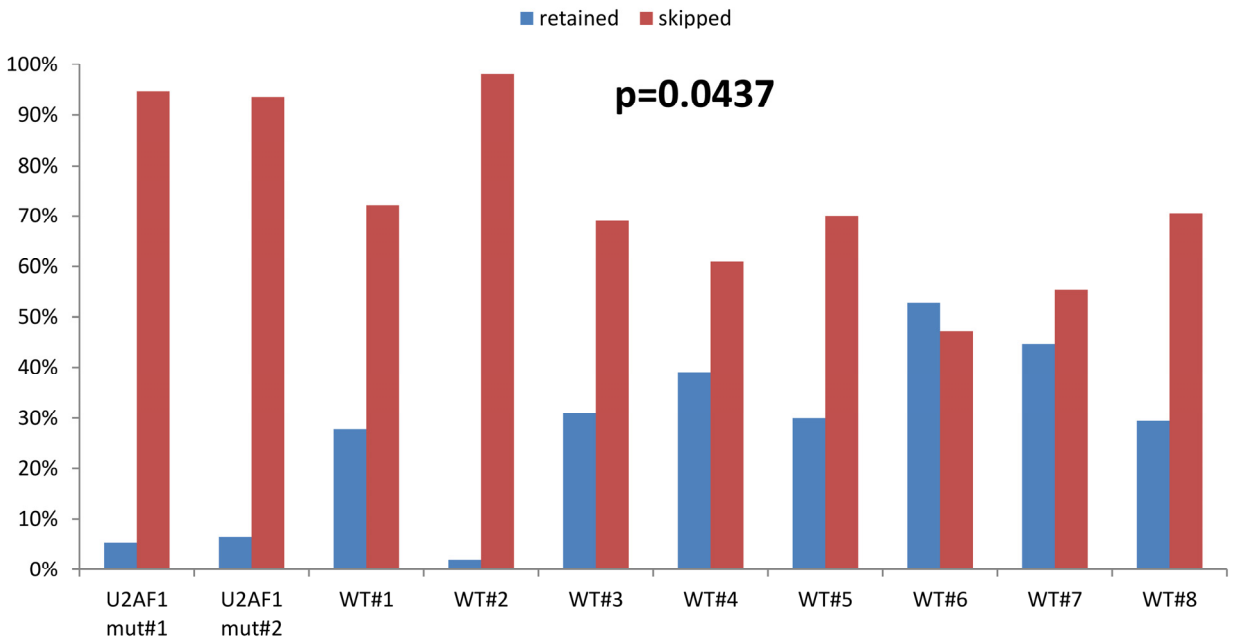


Supplementary Figure 2. Additional mutations found in patients with U2AF1 mutations. Each column represents one patient. The top row represents the diagnosis of the patient (AML – blue, MDS/MPN – pink, MDS – green). The bottom row represents the individual karyotype of each patient (abnormal – blue, normal – white). The remaining rows represent genes that were screened for mutations in the U2AF1 mutated patients. Each row is a separate gene as listed. Red color indicates the presence of a mutation, white color indicates a WT sequence and light grey indicates an unknown status (sequencing not done).



Supplementary Figure 3. Work flow for the identification of candidate exons harboring differential exon usage between U2AF1 mutants and WT. Two major patterns of changes in alternative splicing were identified: 1) increased exon skipping in patients with U2AF1 mutations as shown by an excess of green reads supporting exclusion of the middle exon as depicted in the upper panel and 2) increased exon retention in patients with U2AF1 mutation as shown by an excess of red reads supporting inclusion of the middle exon as depicted in the lower panel. The center panel shows a scatter plot of exon skipping in

control cell RNA versus U2AF1 mutant cell RNA. The lines show the 15% difference cutoff limit used in selecting the most affected exons.



Supplementary Figure 4. Patients exon usage ratio of CEP164 gene (exon7) using intensity of bands on a gel.