Supplementary Table 1. MDS- and isogenic normal iPSC lines used in this study.

Patient	Diagnosis	BM FISH	Starting cells	Del(7q)-MDS-iPSC lines	Normal iPSC lines
#2	MDS (RAEB)	del(7q) in 72.5%	BMMCs	MDS-2.13 MDS-2.A3*	N-2.8 N-2.12 N-2.A2* N-2.A11*
#3	MDS (RAEB)	del(7q) in 99%	BMMCs	MDS-3.1 MDS-3.4 MDS-3.5	N-3.10

All iPSC lines used in this study. The lines MDS-2.A3, N-2.A2 and N-2.A11, marked with an asterisk, were derived in a separate reprogramming experiment than lines MDS-2.13, N-2.8 and N-2.12 (see Supplementary Table 2). (RAEB: Refractory anemia with excess blasts, according to the FAB classification)

Supplementary Table 2. All reprogramming experiments.

Reprogramming experiment	Patient	Total colonies picked	Colonies tested for chr7q copy number by qPCR	Colonies expanded	iPSC lines studied
1st	#2	11	0	1 del(7q) and 10 normal[a]	1 del(7q) and 2 normal[c]
2nd	#3	16	15 del(7q) and 1 normal	9 del(7q)[b] and 1 normal	3 del(7q) and 1 normal[c]
3rd	#2	>100	17 del(7q) and 10 normal[d]	1 del(7q), 6 normal and 4 del(7q) partially reprogrammed	1 del(7q) and 2 normal[c]

Details of the 3 reprogramming experiments performed to generate the iPSC lines shown in Supplementary Table 1.

[a]In this first experiment, chr7q dosage was determined after cell expansion by karyotyping. (These lines subsequently served as one-copy and two-copy controls to establish the qPCR assay.) One pair of the 10 normal lines was found to harbor the same vector integrations (N-2.1 and N-2.8 in Supplementary Fig. 1b), therefore the number of normal lines derived from different starting cells is 9.

[b]Two pairs of lines were found to harbor the same vector integrations (MDS-3.1 and MDS-3.2 in Supplementary Fig. 1b and one more pair not shown), therefore the number of del(7q) lines derived from different starting cells is 7.

[c]These lines are listed in Supplementary Table 1.

[d]All but one del(7q) clones formed colonies with partially reprogrammed morphology.

Supplementary Table 3. Isogenicity of normal and MDS-iPSCs.

Marker	N-2.12		MDS-2.13		N-3.10		MDS-3.1	
Amelogenin	X	Y	X	Y	X	Y	X	Y
CSF1PO	11		11		ND	ND	ND	ND
D13S317	11		11		8	12	8	12
D16S539	11	13	11	13	ND	ND	ND	ND
D18S51	14	16	ND	ND	13	18	13	18
D19S433	13	15	13	15	ND	ND	ND	ND
D21S11	29	32.2	29	32.2	28	30	28	30
D2S1338	16	23	16	23	ND	ND	ND	ND
D3S1358	14	15	14	15	14	15	14	15
D5S818	12	13	12	13	11		11	
D7S820	10	12	10	12	9	8	9	
D8S1179	12	15	12	15	12	15	12	15
FGA	23		23		21	23	21	23
TH01	6	9.3	6	9.3	ND	ND	ND	ND
TPOX	8	10	8	10	ND	ND	ND	ND
vWA	17		17		16	18	16	18

DNA fingerprinting with CODIS markers in one normal (N-) and one MDS- iPSC line from patients #2 and #3, confirming the common genetic background. (Marker D7S820 is located in band q21.11 of chromosome 7, which is encompassed by the deletion in patient #3, but not in patient #2).

Supplementary Table 5. Coordinates of the chr7q deletion in the MDS–2.A3 iPSC line, before and after spontaneous correction.

MDS-2.A3	
Deletion 7q21.2 - 7q36.2	92,782,520 - 154,384,589
Deletion 7q36.2 - 7q36.3	154,401,554 - 159,128,530
MDS-2.A3C	
Deletion 7q21.3 - 7q21.3	96,205,654 - 97,516,805
Deletion 7q21.3 - 7q22.1	97,598,077 - 99,844,739
Deletion 7q22.1 - 7q22.1	99,938,082 - 100,322,721
Deletion 7q22.1 - 7q32.3	100,341,390 - 131,706,336

Genomic coordinates (human genome assembly hg19) of the chromosome 7 deletions in MDS-iPSC line MDS-2.A3 before and after (MDS-2.A3C) spontaneous compensation for chromosome 7q dosage (shown in Fig. 3h). The median probe spacing of the tiling array used was 795 bp.

Supplementary Table 6. Coordinates of the engineered chromosomal deletions.

N-2.12-D-Cre10	
Amplification 7p22.3 - 7p14.3	41,605 - 30,750,671
Deletion 7p13 - 7p11.2	44,704,157 - 56,473,923
Deletion 7p11.2 - 7q36.3	57,654,583 - 159,128,530
N-2.12-D-Cre32	
Amplification 7p22.3 - 7p14.3	41,605 - 30,750,671
Deletion 7p13 - 7p11.2	44,704,157 - 56,473,923
Deletion 7p11.2 - 7q36.3	57,685,793 - 159,128,530
N-2.12-D-2Cre4	
Amplification 7p22.3 - 7p11.2	41,605 - 57,685,793
Deletion 7q11.21 - 7q36.3	61,943,112 - 159,128,530
N-2.12-D-8Cre21	
Deletion 7q11.21 - 7q36.3	62,001,073 - 159,128,530
N-2.12-D-8Cre23	
Amplification 7p22.3 - 7p15.3	41,605 - 23,018,094
Amplification 7q11.21 - 7q11.22	61,979,253 - 68,955,029
Deletion 7q11.22 - 7q36.3	68,955,029 - 159,128,530
N-2.12-D-Cre44	
Deletion 7p12.1	51,128,908 - 53,480,653
Deletion 7q11.21 - 7q31.1	64,839,001 - 112,901,029
N-2.12-D-6Cre6	
Deletion 7q11.1 - 7q31.1	61,075,538 - 112,920,418
H1-D-Cre1	
Deletion 7q11.21 - 7q36.3	62,436,296 - 159,128,530
H1-D-2Cre6	
Deletion 7q21.11 - 7q36.1	83,091,961 - 152,127,281
H1-D-Cre7	
Deletion 7q21.3 - 7q36.3	92,859,256 - 159,128,530

Genomic coordinates (human genome assembly hg19) of hemizygous chr7q deletions engineered in hPSC clones (shown in Fig. 4b). The median probe spacing of the tiling array used was 795 bp.

Supplementary Table 9. Screening experiments.

TRANSDUCTION	CELL LINE	LIBRARY BATCH	%GFP$^+$	SCREENING EXPERIMENT
#1	MDS-2.13 p23+13	#1, low MOI	30%	None (due to library gene silencing)
#2	MDS-2.13 p23+13	#1, high MOI	65%	#2A, #2B
#3	MDS-2.13 p23+20	#1	100%	None
#4	MDS-2.13 p23+21	#1	70%	None (due to spontaneous conversion to diploid 7q state)
#5	MDS-2.13 p23+21	#1	0%	None (due to library gene silencing)
#6	MDS-2.13 p23+41	#2	73.60%	#6A, #6B
#7	MDS-2.13 p23+41	#2	69.30%	#7
#8	MDS-2.13 p23+39	#3	86.70%	#8
#9	MDS-2.13 p23+39	#3	83.10%	#9
#10	MDS-2.A3 p27	#1, low MOI	40%	None
#11	MDS-2.A3 p27	#1, high MOI	62.50%	None
#12	MDS-2.A3 p27	#1, high MOI	72%	#12

List of all independent transductions and screening experiments.

Supplementary Table 10. Library representation and reproducibility of screen readout.

#4A			#4B		
Reads	**Barcode**	**% Representation**	**Reads**	**Barcode**	**% Representation**
0	AGAA	0	1	AGAA	0.00034817
3	TTAC	0.001490417	6	TTAC	0.00208902
26	TGTG	0.012916944	28	TGTG	0.009748761
202	TAGC	0.100354719	253	TAGC	0.088087015
430	ATCT	0.213626382	640	ATCC	0.222828812
456	ATCC	0.226543326	648	ATCT	0.225614172
670	CGCG	0.332859712	932	GCTC	0.324494457
676	GTCC	0.335840545	1011	GTCC	0.351999889
704	GCTC	0.3497511	1037	CGCG	0.361052309
968	GGTC	0.480907763	1557	GGTC	0.542100719
1562	CATG	0.776010254	1900	CATG	0.661523035
2043	TTAT	1.014973719	2729	ACGC	0.95015598
2065	ACGC	1.025903441	2859	TTAT	0.995418083
2775	GCAT	1.378635375	3304	GCAT	1.150353741
3078	CATA	1.529167453	4414	CATA	1.536822461
3359	ACTC	1.66876981	4642	TGAG	1.616205225
3804	TGAG	1.889848276	4893	CGAC	1.7035959
3928	CGAC	1.951452163	5173	ACTC	1.801083505
3992	CCTA	1.983247717	6240	CCTA	2.172580915
4182	TGAT	2.07764077	6295	TGAT	2.191730266
4392	GGAT	2.181969933	6382	AATT	2.222021057
4693	AATT	2.331508401	6398	GGAT	2.227591778
4770	TACT	2.369762428	6790	TACT	2.364074425
5018	TCGT	2.492970202	6832	TCGT	2.378697566
5071	TTAG	2.519300895	7257	CAGT	2.526669823
5224	CAGT	2.595312143	7372	TTAG	2.566709376
5432	CTAC	2.698647695	7533	CAGC	2.622764748
5528	ATTG	2.746341027	7810	TCAG	2.719207844
5612	CAGC	2.788072693	8164	CTAC	2.84246003
5632	TCAG	2.798008803	8347	ATTG	2.906175143
6289	ACTT	3.124410043	9124	ACTT	3.176703248
7124	GACA	3.53924267	9784	CGTG	3.40649546
7224	CGTG	3.588923224	10482	GACA	3.649518133
7387	TATC	3.669902527	11207	TATC	3.901941396
8830	CGTA	4.386792922	13844	CGCA	4.820065734
9937	CGCA	4.936756655	14073	CGTA	4.899796669

10129	TGTC	5.032143318	14719	TGTC	5.124714501
10323	CCAG	5.128523593	15952	CCAG	5.554008133
14255	CTAT	7.081962978	19860	CTAT	6.914656565
16709	GTAC	8.301123774	21539	GTAC	7.499234026
16784	TAGT	8.33838419	25185	TAGT	8.768661913
Total:			Total:		
201286			287216		

Number of sequencing reads and representation (calculated as percentage of total reads) of all barcodes included in batch #1 in two independently processed samples of the same transduced cell pool (samples #4A and 4B). The samples were processed independently in all steps from gDNA extraction, barcode PCR and high-throughput sequencing.

Supplementary Table 12. Primers and probes for chr7q copy number quantitation by qPCR

Name	Sequence
hALB-F	TGA AAC ATA CGT TCC CAA AGA GTT T
hALB-R	CTC TCC TTC TCA GAA AGT GTG CAT AT
hALB-P	/56-JOEN/TGC TGA AAC ATT CAC CTT CCA TGC AGA /36-TAMSp/
7p14.3-F	GCC TAC CAA CTA CCC TGT TT
7p14.3-R	TTC ATA GTG GGC CTT TGT ATC A
7p14.3-P	/56-FAM/TGA GAA ATT GGA GAG GGT GGG TGC /36-TAMSp/
7q11.22-F	TCA CCG TAA GTG CTG TCA AC
7q11.22-R	CCT TAC ACA CCC AGA GAC AAT C
7q11.22-P	/56-FAM/TCG GAT GCC CAG CAA ACA TCT GTA /36-TAMSp/
7q21.12-F	GAC AGC GTT TCC CTG TAA GT
7q21.12-R	GTG TCT GAC CCT GTG AGT ATT G
7q21.12-P	/56-FAM/AGA GAA GGT GGA ATG AAG CTC GCA /36-TAMSp/
7q31.2-F	AGG AAA GTG GGT GAC TGG GTT TCA
7q31.2-R	TAG TGT TTG CTG GCA ATC CTT GGC
7q31.2-P	/56-FAM/TCC AAT CAG GAA AGG CAA GCC AAG GA/36-TAMSp/
7q35-F	GAG ACC AAA GGA GAA GTG AAG G
7q35-R	CCT CCC AGC CCA ATG TAA TAG
7q35-P	/56-FAM/TAG AAT TGG CAG CAC AGG TGG AGG /36-TAMSp/

Supplementary Table 13. Primers used for qRT-PCR

Name	Sequence
ADCK2-F	ACC TCA TCT CCG TGG CAG TGA A
ADCK2-R	CCT CCA CAA TCT CAG GCA AGC T
AGK-F	GAG GTC TGG AAA GAT GTG CAG C
AGK-R	CTC CTT TGC TGA TGG TGT CAG G
ATP6V0E2-F	ACC GCC GTC TGC TGT TAC CTC
ATP6V0E2-R	AAG CGC ACG TAC CAG ATG GTC T
EZH2-F	GAC CTC TGT CTT ACT TGT GGA GC
EZH2-R	CGT CAG ATG GTG CCA GCA ATA G
GALNT11-F	TTG GGC TAC CAC AGA GAT GTG C
GALNT11-R	CAC TGT CCG AAG CAA GGC AGA A
HIPK2-F	AGC GTC ATC ACC ATC AGC AGT G
HIPK2-R	AGT CGT GGA CTG TGA CAC AGC T
LUC7L2-F	GGA GCA GAG AAC GAT CCA AGA G
LUC7L2-R	TCC GAT CTC TGT CAC GAG GTG A
SSBP1-F	CTG TCT TGA GAC AGG TGG AAG G
SSBP1-R	CTG TGC CAT GTT GTC TTT TGA CTG
ACTIN-F	TGA AGT GTG ACG TGG ACA TC
ACTIN-R	GGA GGA GCA ATG ATC TTG AT

Supplementary Table 14: **shRNA and CRISPR gRNA sequences.**

Name	Sequence
shRNA-scramble	CGA GGG CGA CTT AAC CTT AGG
shRNA-HIPK2	ATT GGG CTG GAT ACT GAC TCG
shRNA1-ATP6V0E2	TGA CCA GAC ATA GTT GTG TGG
shRNA2- ATP6V0E2	ATG AGC CAG AAG AGG TAA CAG
shRNA-LUC7L2	ATT CCG ATA AAC TTC CTC TGC
shRNA1-EZH2	TTT GGT CCC AAT TAA CCT AGC
shRNA2-EZH2	TAA TGG GAT GAC TTG TGT TGG
gRNA-EZH2	CTG AGA AGG GAC CAG TTT GTT GG

The mature antisense shRNA sequences are given 5’ to 3’.