SUPPLEMENTARY INFORMATION

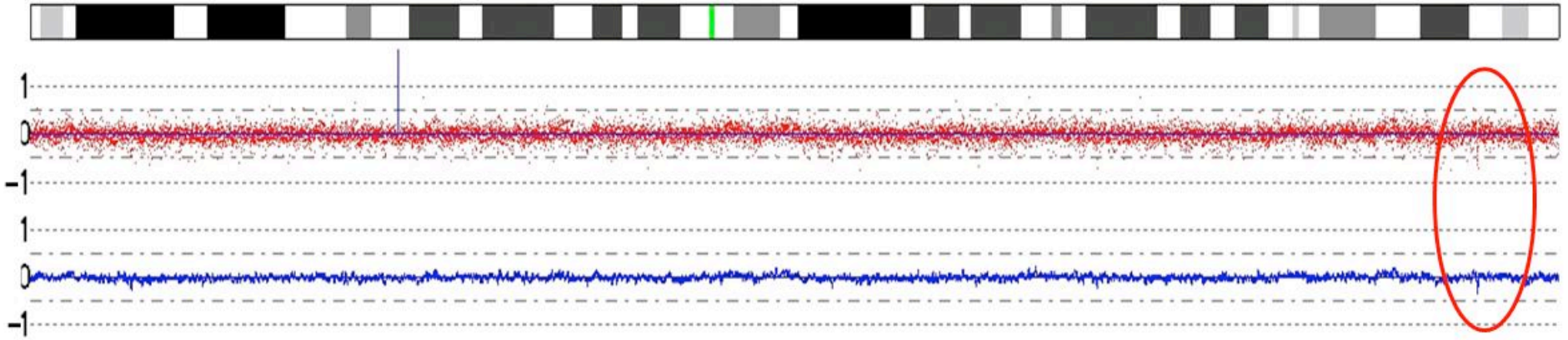
Somatic mutations of the histone methyltransferase gene *EZH2* in myelodysplastic syndromes

Gorica Nikoloski, Saskia M.C. Langemeijer, Roland P. Kuiper, Ruth Knops, Marion Massop, Evelyn R.L.T.M. Tönnissen, Adrian van der Heijden, Theresia N. Scheele, Peter Vandenberghe, Theo de Witte, Bert A. van der Reijden & Joop H. Jansen

Nature Genetics: doi:10.1038/ng.620

Supplementary Figure 1: 250k SNP array profile of chromsome 7 showing a microdeletion at 7q36.1.

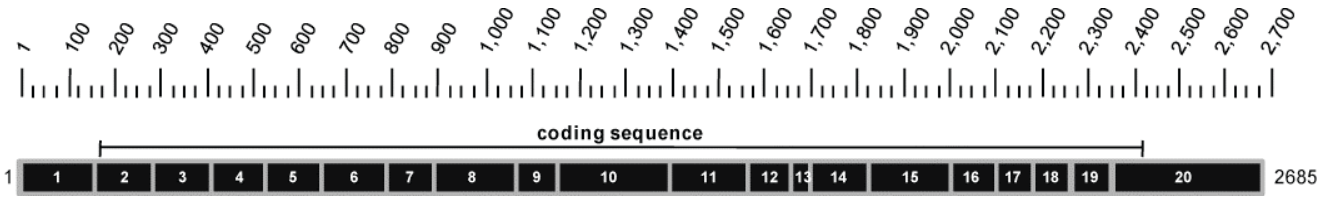
Genomic profile of chromsome 7 in subject 87 showed a microdeletion located on 7q36.1 (red circle). This microdeletion encompassed the *CUL1* and *EZH2* genes (see also Figure 1a). Relative intensities of individual (red dots) and 10 consecutive SNPs (blue line) on chromsome 7 are indicated. For this patient, no T cells could be obtained to show that the deletion was acquired. However, this region was never found to be affected in a large control cohort of 1015 healthy controls that were individually hybridized,.



Nature Genetics: doi:10.1038/ng.620

Supplementary Figure 2: *EZH2* exon organization and sequence primers.

Exon organization of *EZH2* according to GenBank: *EZH2* isoform 1, NM_004456.3, representing the longest isoform. Localization of the primers is indicated in the table. M13 sequence primers are indicated in lower case.



Exon	Amplification primer sequence (5' to 3')	Forward / Reverse	Location	Sequence primer sequence (5' to 3')
2	tgtaaaacgacggccagtAAGTGTTTTAAGGATTTAACCATGCA	Forward	intron 1	tgtaaaacgacggccagt
	caggaaacagctatgaccCCTTTATATTTAGGGAGGCATTTCTG	Reverse	intron 2	caggaaacagctatgacc
3	tgtaaaacgacggccagtGTCAACAAAGAAAGGATGGCAAT	Forward	intron 2	AAGAGTATGTTTAGTTCCA
	caggaaacagctatgaccAGCTAATAGTGTGATCTACAGCAGTCATT	Reverse	intron 3	caggaaacagctatgacc
4	tgtaaaacgacggccagtGGGTAGGCAGCATCTCTTTACAATAT	Forward	intron 3	tgtaaaacgacggccagt
	caggaaacagctatgaccAATTTATACTGTCTTGATTCACCTTGACA	Reverse	intron 4	caggaaacagctatgacc
5	tgtaaaacgacggccagtTTTTCTCAGATATGCTTATTGGTGAGA	Forward	intron 4	tgtaaaacgacggccagt
	caggaaacagctatgaccGTTCAGTCCCTTATAGTTCAGTGCAA	Reverse	intron 5	caggaaacagctatgacc
6	tgtaaaacgacggccagtTTGCCTAACACCAGTCCTGAAA	Forward	intron 5	tgtaaaacgacggccagt
	caggaaacagctatgaccTCCCAAGTGCTAGGATTACAGAGTTAG	Reverse	intron 6	caggaaacagctatgacc
7	tgtaaaacgacggccagtGTAGCAGAGCTGGGAGTAGAACCTA	Forward	intron 6	tgtaaaacgacggccagt
	caggaaacagctatgaccGTAATGCAGAGTACCACAAGTACACATG	Reverse	intron 7	caggaaacagctatgacc
8	tgtaaaacgacggccagtAATTTGATTCTTGATAACACCATGCA	Forward	intron 7	tgtaaaacgacggccagt
	caggaaacagctatgaccCAGAGCAATCCTCAAGCAACAAA	Reverse	intron 8	caggaaacagctatgacc
9	tgtaaaacgacggccagtGAGGAGGAATGGAGAATACGTTGT	Forward	intron 8	tgtaaaacgacggccagt
	caggaaacagctatgaccGAAACAGCATGGGTGAGAAAGC	Reverse	intron 9	caggaaacagctatgacc
10	tgtaaaacgacggccagtATGAGTTTTAGAACTTTGCCCTGATG	Forward	intron 9	tgtaaaacgacggccagt
	caggaaacagctatgaccCACAACACGAACTTTCACAGAACAG	Reverse	intron 10	caggaaacagctatgacc
11	tgtaaaacgacggccagtTGTGAACTACGATGGGTTAGTGTTTT	Forward	intron 10	tgtaaaacgacggccagt
	caggaaacagctatgaccTTAAAAGTCTACATTGGGGAAATTCTG	Reverse	intron 11	caggaaacagctatgacc
12	tgtaaaacgacggccagtGCATCACAGTTCTTGACCAGAATATAA	Forward	intron 11	tgtaaaacgacggccagt
	caggaaacagctatgaccAGTTCTGTTTTTGATGGCAGTTTAAG	Reverse	intron 12	caggaaacagctatgacc
13	tgtaaaacgacggccagtAGCCTGAGCCATCAAGCTGTT	Forward	intron 12	tgtaaaacgacggccagt
	caggaaacagctatgaccCCAGTCAGCCTCCACTTTACAGA	Reverse	intron 13	caggaaacagctatgacc
14	tgtaaaacgacggccagtCCGTGCTTTAGATGGAACTCATAA	Forward	intron 13	tgtaaaacgacggccagt
	caggaaacagctatgaccGCCTCACACACACAGACACACA	Reverse	intron 14	caggaaacagctatgacc
15	tgtaaaacgacggccagtCGCTGTTAATACTTCCATTTCCTTGT	Forward	intron 14	tgtaaaacgacggccagt
	caggaaacagctatgaccACACTTTCTCATCAGTTGCACCTTT	Reverse	intron 15	caggaaacagctatgacc
16	tgtaaaacgacggccagtACAGGTCTGAGGATTTACAGTGATAGC	Forward	intron 15	tgtaaaacgacggccagt
	caggaaacagctatgaccCAGTTTATGGCAATTCATTTCCAAT	Reverse	intron 16	caggaaacagctatgacc
17	tgtaaaacgacggccagtTGGGAAAGAGAACTTGGCTGTAGT	Forward	intron 16	tgtaaaacgacggccagt
	caggaaacagctatgaccCCAGTTCCTTTCAAGCAAGCA	Reverse	intron 17	caggaaacagctatgacc
18	tgtaaaacgacggccagtAACAATAGTGTGTTCTTCCAAATGTCA	Forward	intron 17	tgtaaaacgacggccagt
	caggaaacagctatgaccCTGGTGTCAGTGAGCATGAAGAC	Reverse	intron 18	caggaaacagctatgacc
19	tgtaaaacgacggccagtCCTTTGACGTGAGAATTGGAACT	Forward	intron 18	tgtaaaacgacggccagt
	caggaaacagctatgaccCTGTCAACAGCAGGGTGAGAAAT	Reverse	intron 19	caggaaacagctatgacc
20	tgtaaaacgacggccagtACCCACTATCTTCAGCAGGCTTT	Forward	intron 19	TCAGCACATGTTGGATGG
	caggaaacagctatgaccCTTCCACATATTCACAGGCAGTATTAGT	Reverse	downsteam of exon 20	caggaaacagctatgacc

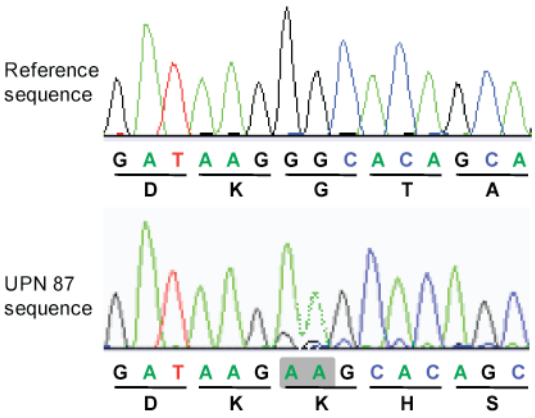
Nature Genetics: doi:10.1038/ng.620

		CAGCTGTTTCAGAGGAGGGGG

Nature Genetics: doi:10.1038/ng.620

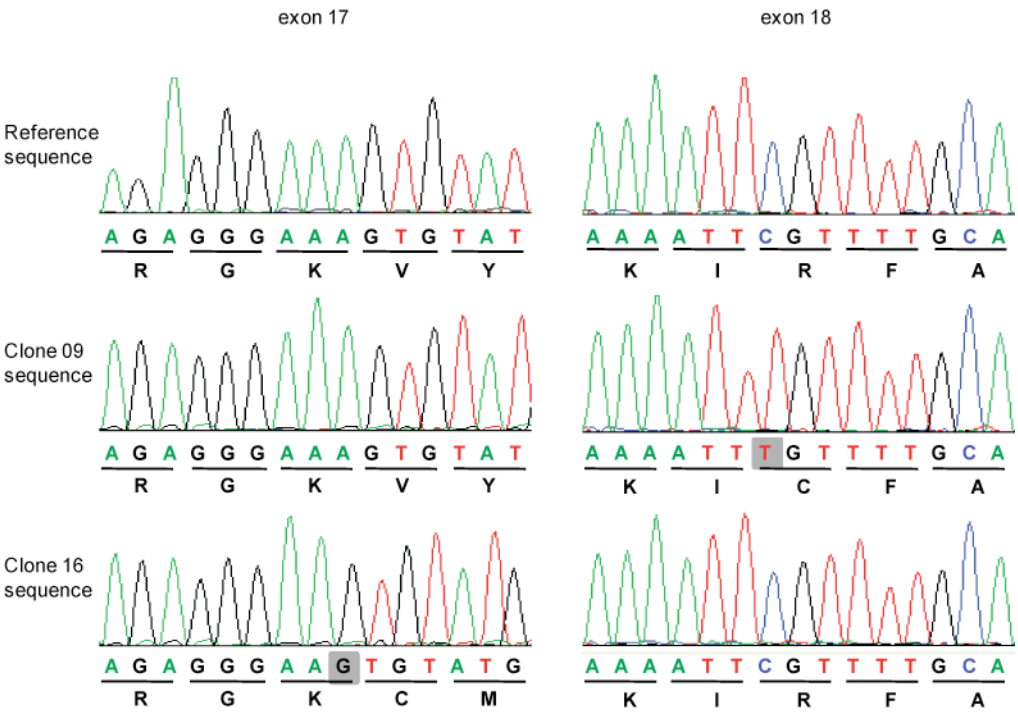
Supplementary Figure 3: Sequence analysis of genomic DNA from subject 87.

SNP array analysis revealed a 130-kb microdeletion at 7q36.1, containing the *CUL1* and *EZH2* genes, in subject 87 (**Fig. 1a**). Subsequent sequence analysis of genomic DNA from this subject showed a c.703delGinsAA mutation in the remaining copy of *EZH2*, predicting the formation of a truncated protein.



Supplementary Figure 4: Bi-allelic point mutations in subject 96.

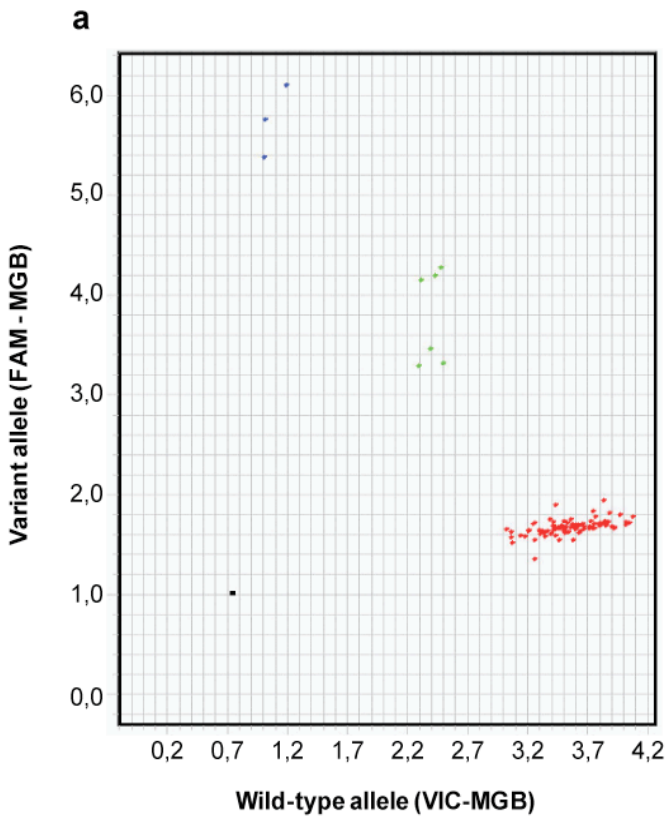
A PCR product covering *EZH2* exons 17 and 18 from subject 96 was cloned and individual clones were sequenced (**Supplementary Methods** online). Clones contained either the c.1983delA mutation in exon 17 (e.g. clone 16) or the c.2068C>T mutation in exon 18 (e.g. clone 9), indicating the bi-allelic nature of these mutations. Mutant nucleotides are marked.

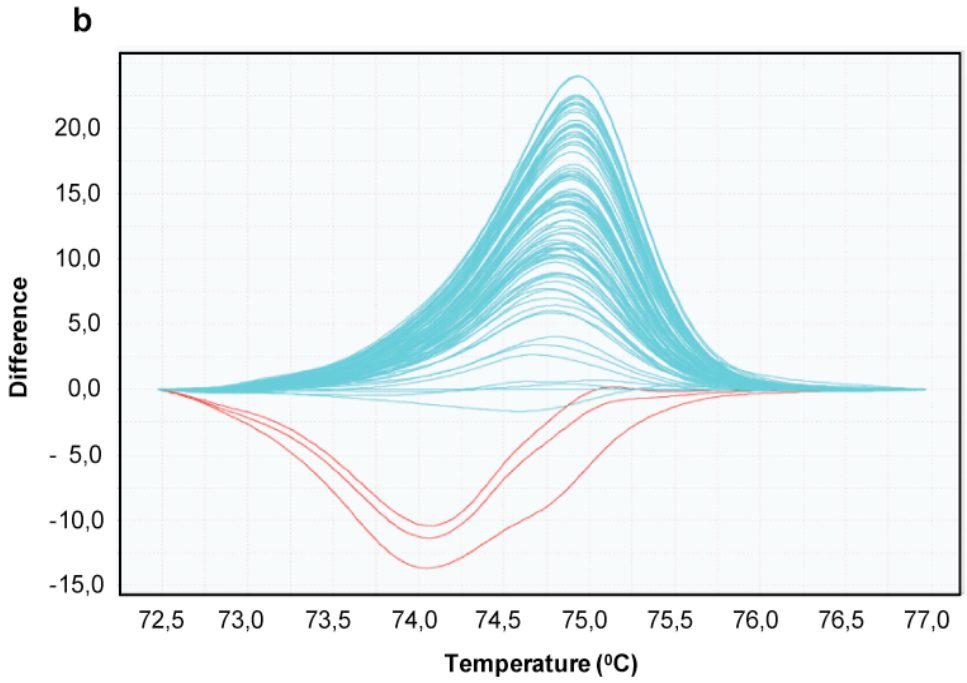


Nature Genetics: doi:10.1038/ng.620

Supplementary Figure 5: Missense mutations are absent in healthy donors.

Allelic discrimination assays (**Supplementary Table 2** and **Supplementary Methods** online) were performed to screen genomic DNA from white blood cells of healthy donors (n = 250). **(a)** Representative result of allelic discrimination assays (PCR with allele-specific probes). The missense variant c.2025C>A (exon 17, subject 97) is absent in 250 healthy donors as determined by this assay. Blue dots represent subject 97 carrying this mutation (triplicate). Red dots represent 80 healthy donor samples and two MDS patients without this mutation (subjects 63 and 73, both in triplicate). Green dots represent 1:1 DNA mixtures of subject 97 and subject 63 (triplicate) and of subject 97 and subject 73 (triplicate) creating artificial heterozygotes. The black dot represents the No Template Control. Variants were auto-called by the SDS 2.3.2. allelic discrimination program (Applied Biosystems). **(b)** Representative result of allelic discrimination assays (PCR followed by HRM analysis). The missense variant c.2068C>T (exon 18, subject 96) is absent in 250 healthy donors. Blue lines represent 83 healthy donor samples and two MDS patients without this mutation (subject 63 and subject 73, both in triplicate). Red lines represent subject 96 carrying this mutation. Variants were auto-called by the HRMv1.0 software (Applied Biosystems).

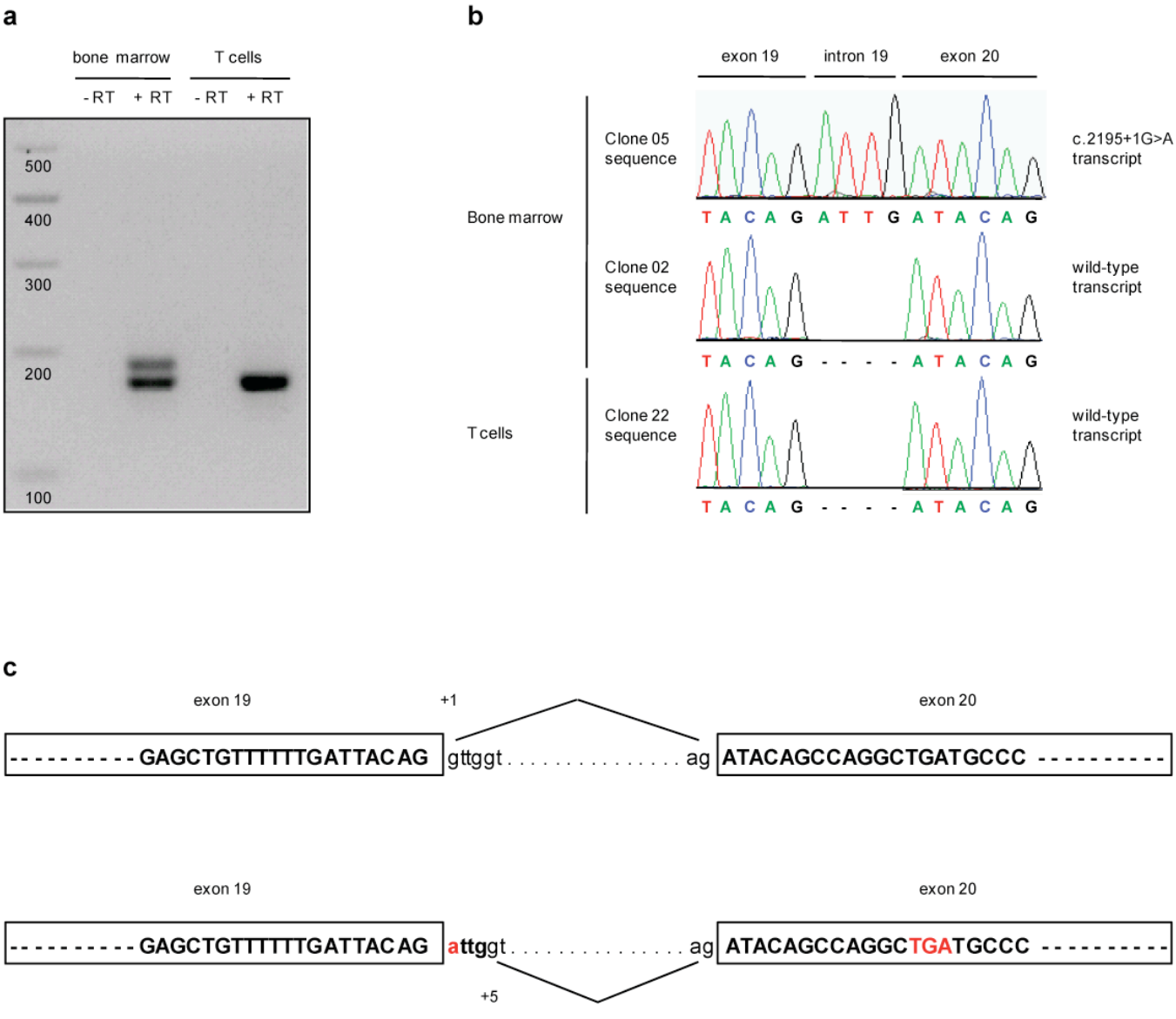




Nature Genetics: doi:10.1038/ng.620

Supplementary Figure 6: Expression of wild-type and mutant alleles in subjects 10 and 127.

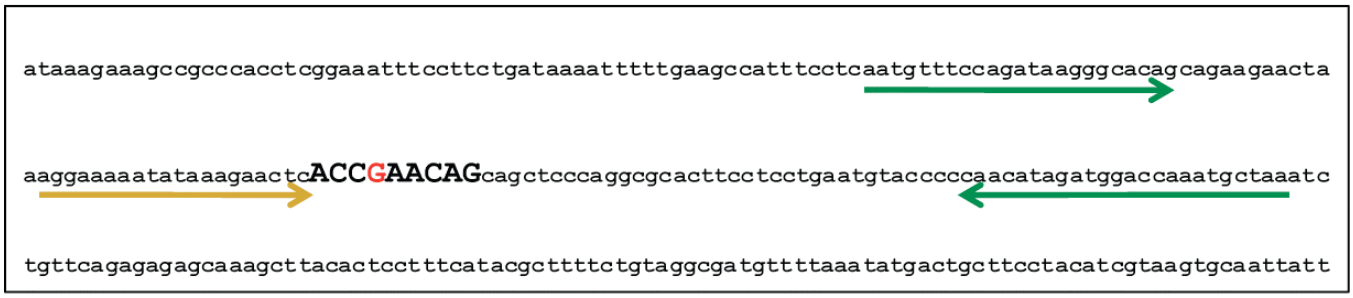
(a-h) Expression of wild-type and mutant alleles in subjects 10 and 127. **(a-c)** Subject 10 harbored a heterozygous mutation c.2195+1G>A in the donor splice site of intron 19. This resulted in aberrant mRNA splicing as shown by agarose gel electrophoresis and sequence analysis of RT-PCR products (**Supplementary Methods** online). **(a)** Bone marrow cDNA from subject 10 was amplified using primers directed against the exon 18/19 boundary and exon 20. The PCR product demonstrated the presence of two different transcripts. Amplification of cDNA from the non-neoplastic T cells showed the presence of only one transcript. **(b)** Sequence analysis of the cloned PCR products revealed an aberrant transcript in bone marrow cells, containing an insertion of four intronic base pairs G>A, T, T and G between exon 19 and 20. Wild-type transcripts were present in both bone marrow cells and non-neoplastic T cells. **(c)** Wild-type RNA was spliced at the r.2195+1 donor splice site. Mutant RNA was aberrantly spliced at r.2195+5.



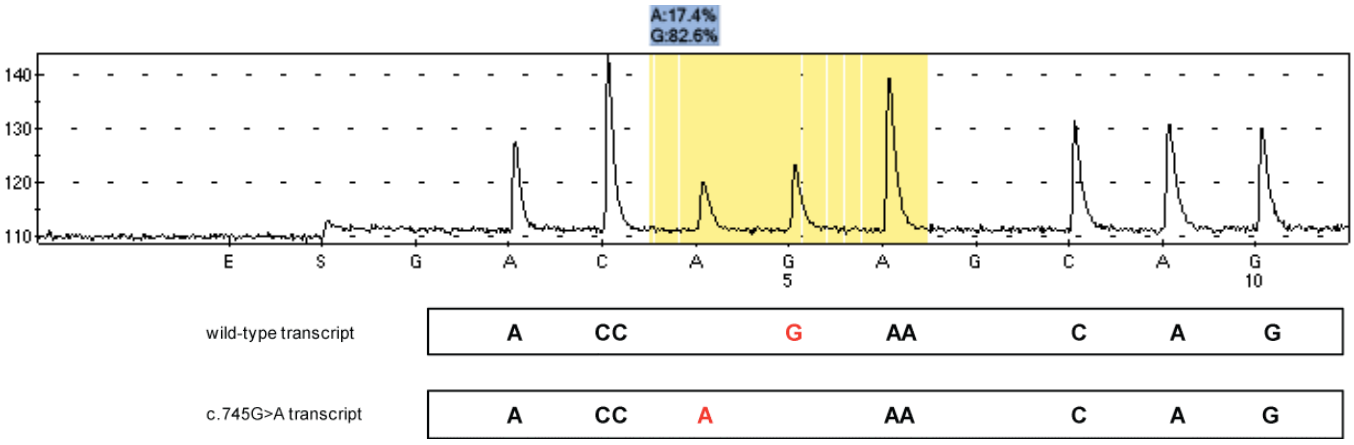
Nature Genetics: doi:10.1038/ng.620

(d-e) Subject 10 harbored the heterozygous missense mutation c.745G>A in exon 8. Transcripts carrying this mutant were expressed as determined by pyrosequencing (**Supplementary Methods** online). **(d)** Bone marrow cDNA from subject 10 was amplified using primers directed against exon 7 and exon 8 (green arrows), spanning the c.745G>A mutation (indicated in red). Pyrosequencing of the PCR products was performed using a nested primer directed against the exon 7/8 boundary (yellow arrow). **(e)** Pyrosequence analysis revealed the presence of both wild-type transcripts (83 %) and mutant transcripts (17 %), as determined by the GeneScan software.

d

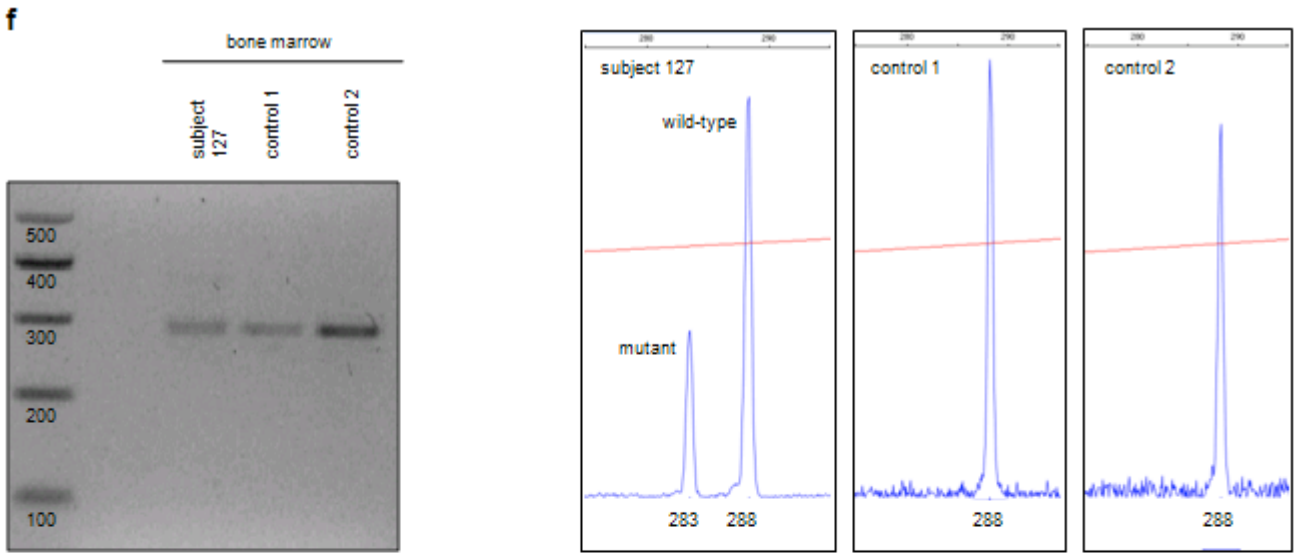


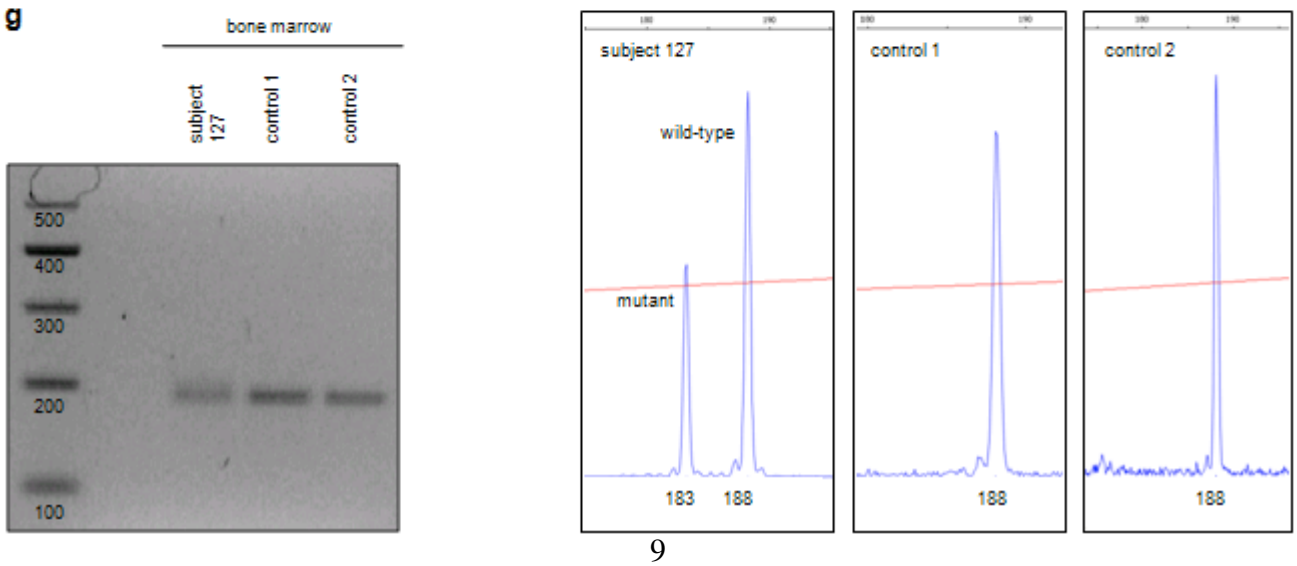
e



Nature Genetics: doi:10.1038/ng.620

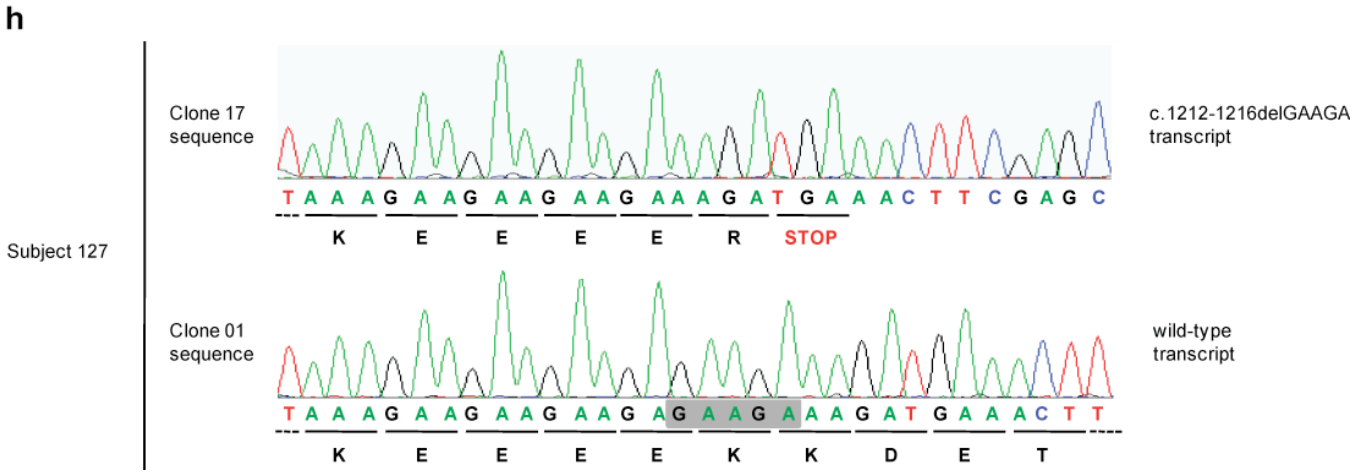
(f-h) Subject 127 carried the heterozygous deletion c.1212-1216delGAAGA in exon 10 resulting in a frameshift. Transcripts carrying this mutant were expressed as determined by GeneScan and sequence analysis of RT-PCR products (**Supplementary Methods** online). **(f)** Bone marrow cDNA from subject 127 was amplified using primers directed against exon 9 and exon 11. GeneScan analysis unveiled the presence of two different transcripts: a wild-type transcript of the expected 288 bp and a mutant transcript of 283 bps (ratio = 2,5 : 1). GeneScan analysis of amplified cDNA from bone marrow cells of two healthy donors (controls 1 and 2) showed the presence of the wild-type transcript only. **(g)** Bone marrow cDNA from subject 127 was amplified using primers directed against exon 10 and exon 11. GeneScan analysis detected two different transcripts: a wild-type transcript of the expected 188 bp and a mutant transcript of 183 bp (ratio = 1,8 : 1). GeneScan analysis of cDNA from bone marrow cells of two healthy controls showed the presence of the wild-type transcript only.





Nature Genetics: doi:10.1038/ng.620

(h) Sequence analysis of the cloned PCR products confirmed the presence of two different transcripts in the bone marrow cells of subject 127. The deleted nucleotides are marked.

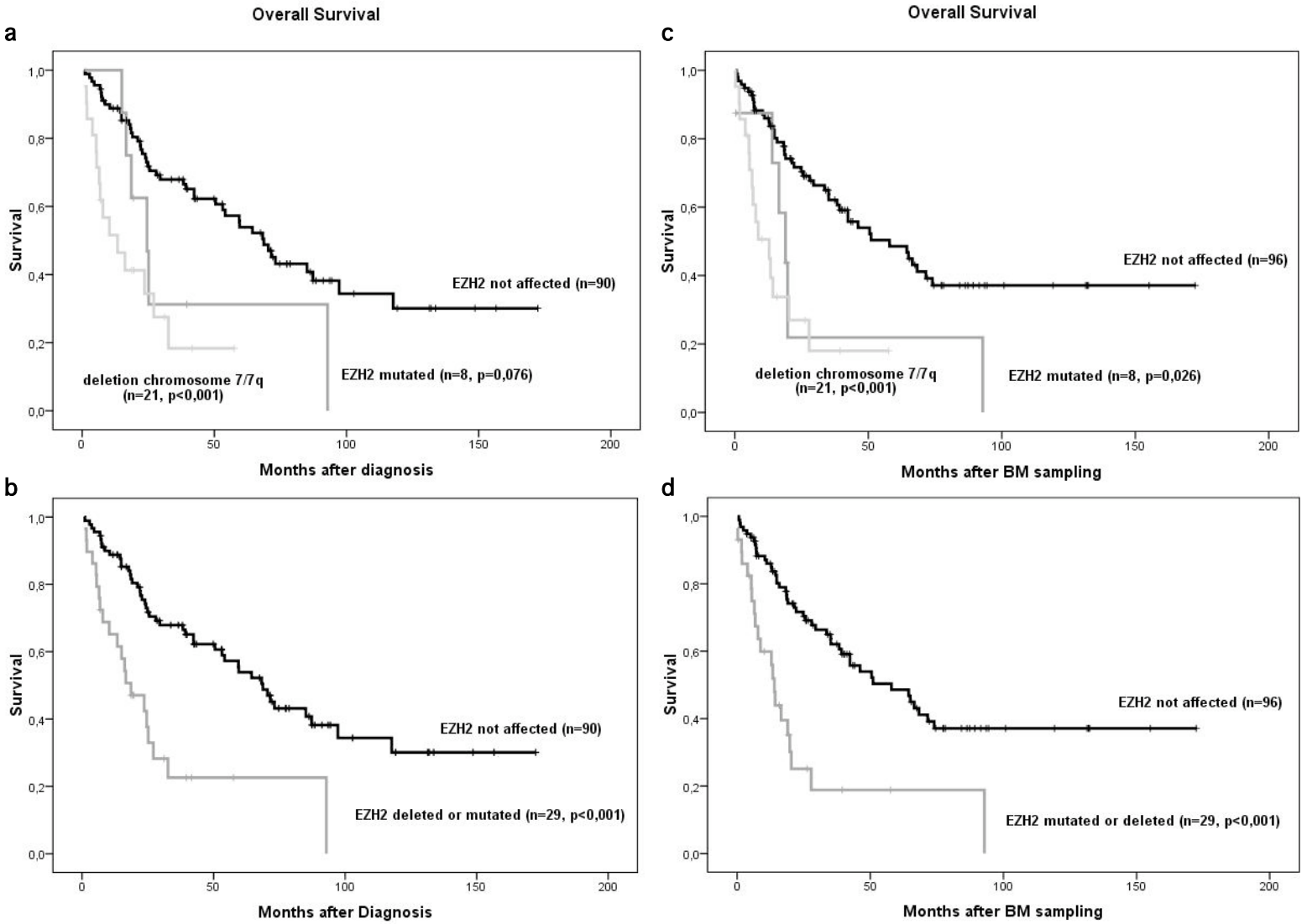


Nature Genetics: doi:10.1038/ng.620

Supplementary Figure 7: Aberrations of 7q36.1 correlate with poor survival in MDS.

Kaplan-Meier analysis showing overall survival of MDS patients with and without *EZH2*/7q36.1 aberrations. Since diagnosis material was not available in a subset of patients, overall survival was calculated from the date of diagnosis **(a** and **b)** as well as from the date of bone marrow (BM) sampling **(c** and **d).**

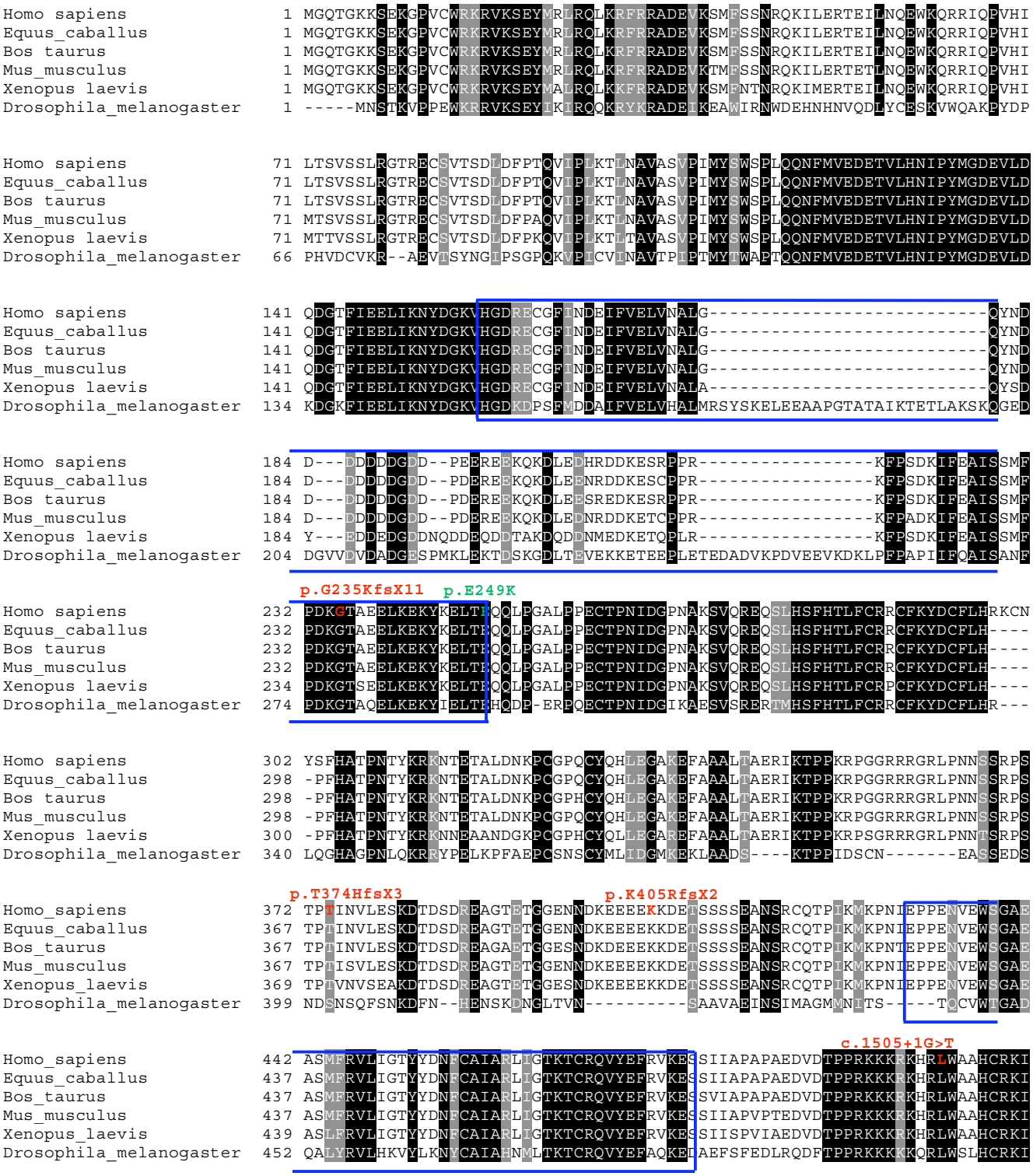
p-values were calculated comparing patients with a deletion or mutation of *EZH2* to patients in whom *EZH2* was not affected.



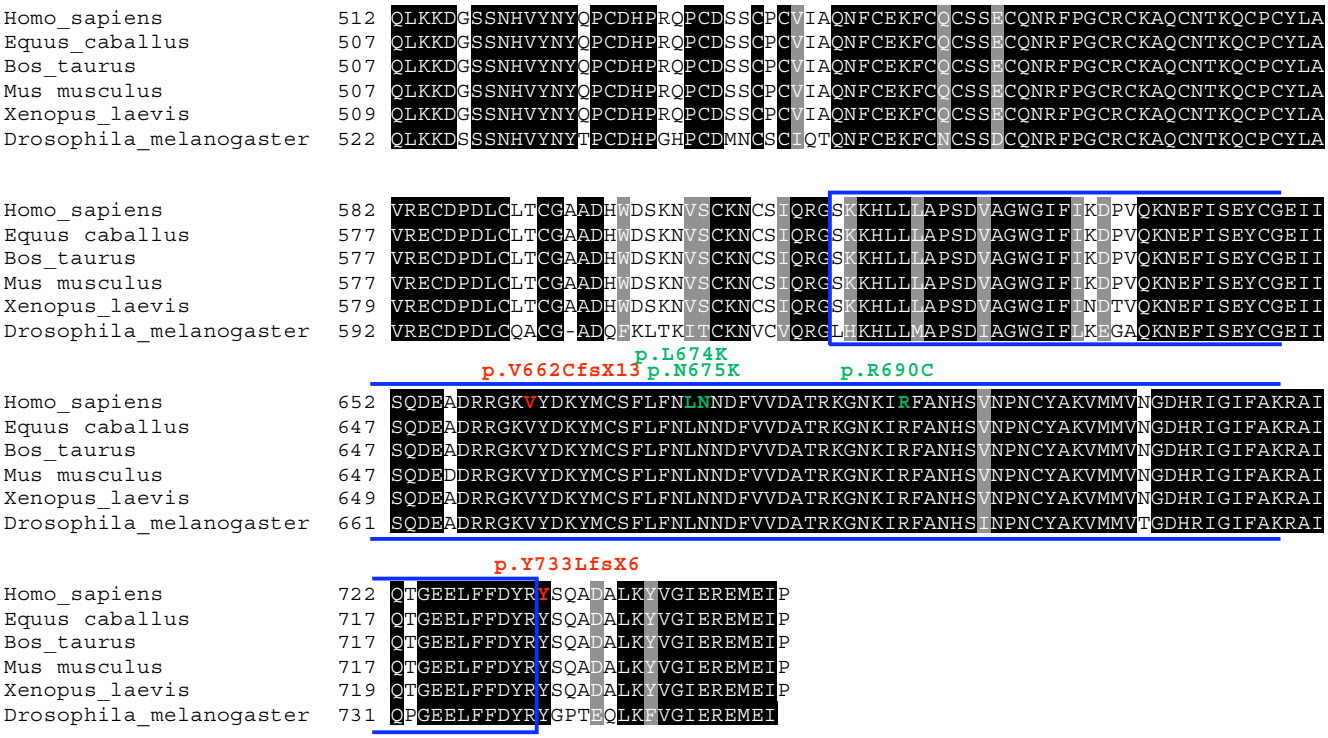
Nature Genetics: doi:10.1038/ng.620

Supplementary Figure 8: Alignment of orthologous EZH2 proteins and summary of point mutations.

Alignment of EZH2 orthologs in various species. Blue boxes represent two the SANT domains and the SET domain respectively. Positions of donor splice site and frameshift mutations are indicated in red. In case of donor splice site mutation c.1505+1G>T, the first amino acid of the next exon is highlighted in red, since the deduced protein sequence is unknown. Missense mutations are indicated in green.



Nature Genetics: doi:10.1038/ng.620



Nature Genetics: doi:10.1038/ng.620

Supplementary Table 1: Characteristics of MDS patients UPN106 - UPN129.

UPN	FAB	WHO	IPSS	Karyotype	*TET2* mutation
106	RA	RA	int-2	45,XY,-7[6]/46,XY[4]	no
107	RARS	RCMD-RS	int-1	46,XX,t(3;3)(q21;q26)[10]	no
108	RAEB	RAEB-1	int-1	46,XX,del(5)(q12q33), add(7)(q11), del(8)(q22), add(9)(q34), ?10 [6]/ 46,sl,add(14)(q32), add(19)(p13) [3]/46,XX[1]	no
109	RA	RCMD	int-1	45,X,-Y [4]/46,XY[16]	no
110	RARS	RCMD-RS	int-1	normal	yes
111	RA	5q-	unknown	46,XY, 5q-	no
112	RARS	RCMD-RS	int-1	45,XX,del(5)(q13q33),-7 [5]/ 45,XX,idem,del(4)(q25q31),-17,der(20)t(17;20)(q12;q13.1),+mar[5]/46,XX[11]	no
113	RAEB	RAEB-1	int-2	45,XX, -7 [8]/46,XX[2]	no
114	RA	RCMD	int-2	46,XY,der(7)t(1;7)(p10;q10)[9]/46,XY[1]	yes
115	RAEB	RAEB-2	int-2	normal	no
116	RA	RA	int-1	normal	no
117	RAEB-t	RAEB-2	int-2/high	no metaphases, FISH: no deletion 5/7 or trisomy 8/21	no
118	RA	RCMD	low	normal	no
119	RAEB	RAEB-1	int-2	46,XX,del(5)(q15q33),der(7)del(7)(p11)add(7)(q3?3) [8]/46,XX [2]	no
120	RARS	RARS	low	normal	no
121	RAEB	RAEB-1	int-2	44~47,XY,-2,-2,-4, del(4)(q31),-5,-7,+15,-17,del(20)(q11), +mar1, +mar2,+mar3 [cp2]/ 44,sl.add(9)(q34),add(12)(p13),-13,-15,-15,-16,add(17)(p13), +mar4,+mar5,+mar6,+mar7 [cp6]/46XY [2]	yes
122	RARS	RARS	low	normal	no
123	RA	RCMD	int-1	46,X,t(Y;11;20)(q11;p15;p11)[10]	no
124	RA	5q-	int-1	46,XX,del(5)(q12q33) [7]/46,XX [3]	yes
125	RAEB	RAEB-1	int-1	normal	yes
126	RA	RCMD	int-2	47,XX,-7,+21,+22[5]	no
127	RA	RA	low	normal	no
128	RAEB	RAEB-1	int-2	46,XY,-3,add(5)(q?31),der(9)t(3;9)(q13;q22), add(21)(p11), add(22)(q13),+mar1 [16]/ 47,sl,+mar2 [4]	no
129	RAEB-t	RAEB-2/AML	int-2/high	46,XX,del(5)(q22q33)[9]/46,XX[1]	no

Nature Genetics: doi:10.1038/ng.620

Supplementary Table 2: Primers and probes for allelic discrimination.

UPN	Variant	Assay	Amplification primer and probe sequence (5' to 3')	Forward / Reverse	Reporter/ Quencher	Wild-type/ Variant	Location
97	c.2025C>A exon 17	Allele-specific probes	GAAGCTGACAGAAGAGGGAAAGTG	Forward			exon 17
			TGCCACATGCAACTCAGGAA	Reverse			intron 17
			TCAACTTGAA**C**AATGGT	Probe	VIC / MGB	Wild-type	exon 17 - intron 17
			TCAACTTGAA**A**AATGGT	Probe	FAM / MGB	Variant	exon 17 - intron 17
96	c.2068C>T exon 18	HRM	AAATTATTCACTGGGCTGTGCTTACT	Forward			intron 17
			TACCTTTTGCATAGCAGTTTGGATT	Reverse			exon 18 - intron 18

Supplementary Table 3: Identified SNPs in the coding sequence and intron-exon boundaries of *EZH2* in MDS patients.

Substitution	SNP ID	Type of substitution	Amino acid change	Frequency in MDS cohort (%)	Present in T-cell fraction
c.87C>T		synonymous	p.L29L	1/126 (0,8)	N/A
c.396T>C	rs61732845	synonymous	p.P132P	3/126 (2,4)	N/A
c.553G>C	rs2302427	non synonymous	p.D185H	20/126 (16,0)	N/A
c.623A>G	rs61753264	non synonymous	p.D208G	1/126 (0,8)	N/A
c.933T>C		synonymous	p.Y311Y	1/126 (0,8)	N/A
c.1457C>T		synonymous	p.P486P	2/126 (1,6)	N/A
c.1731G>A	rs41277437	synonymous	p.P577P	7/126 (5,6)	N/A
c.2113-6C >T		non-coding region	non-coding region	2/126 (1,6)	yes
c.2113-9A>G		non-coding region	non-coding region	1/126 (0,8)	yes
c.2276+6T >G	rs41277434	non-coding region	non-coding region	12/126 (9,5)	N/A

Nature Genetics: doi:10.1038/ng.620

Supplementary Methods

Study participants. The characteristics of MDS patients (n = 102), the collection of blood and/or bone marrow after obtaining informed consent and the isolation of cells have been described previously[3]. The characteristics of 24 additional MDS patients are described in this paper (**Supplementary Table 1** online). Consecutive subjects from the Radboud University Nijmegen Medical Center (n = 107) belonging to all different WHO and IPSS categories and cytogenetic subgroups were included, if bone marrow was available, as well as a second cohort of karyotypically normal cases from the University Hospital Leuven (n = 19). For control experiments, blood from 250 healthy individuals and bone marrow from two healthy donors was obtained after informed consent. Cells were frozen in liquid nitrogen until further use. For SNP array analysis and genomic sequencing, DNA was extracted from FACS-sorted lymphocyte-depleted bone marrow (n = 41) and Ficoll density gradient– isolated mononuclear cells from bone marrow (n = 85). DNA from healthy donors (controls, n = 250) was isolated from white blood cells after NH_4Cl lysis. RNA from healthy donors (controls, n = 2) was isolated from total bone marrow as described.

SNP array analysis. The SNP array analysis of genomic DNA from 102 MDS patients was performed and described previously[3].

Sequence analysis. DNA was extracted using Qiagen spin-columns and DNA quality was checked by gel analysis and nanodrop ND-1000 spectrophotometer analysis. We conducted bidirectional sequence analysis on PCR-amplified genomic DNA fragments spanning the entire coding region and splice sites (**Supplementary Fig. 3** online). Sequence variations were confirmed by independent PCR-amplification and sequencing of the original DNA. To analyze whether the observed variants were acquired, DNA from cultured non-neoplastic T cells was amplified and sequenced, whenever available (n = 3). To minimize the chance that the remaining putative variants represent SNPs, we compared the observed variants with public databases (NCBI SNP and EST database) and investigated the presence of missense variants in healthy donors (n = 250) by allelic discrimination assays. Detected mutations were described according to the nomenclature explained at http://www.hgvs.org/mutnomen[5].

Cloning of variant sequences. Subject 96 harbored two heterozygous variants (c.1983delA in exon 17 and c.2068C>T in exon 18) in close proximity, enabling DNA amplification using primers spanning both variants: forward primer 5' - tgtaaaacgacggccagtTGGGAAAGAGAACTTGGCTGTAGT - 3' (intron 16) and reverse primer 5'- caggaaacagctatgaccCTGGTGTCAGTGAGCATGAAGAC - 3' (intron 18). The generated amplicon was cloned into the pDrive cloning vector (Qiagen). To determine whether the variants were mono- or bi-allelic, individual clones were sequenced using primers
5' -TGGGAAAGAGAACTTGGCTGTAGT - 3' (forward primer, intron 16),
5' -CCAGTTCCTTTCAAGCAAGCA- 3' (reverse primer, intron 17),
5' -AACAATAGTGTGTTCTTCCAAATGTCA- 3' (forward primer, intron 17) and

Nature Genetics: doi:10.1038/ng.620

5'- CTGGTGTCAGTGAGCATGAAGAC - 3' (reverse primer, intron 18).

T cell cultures. To obtain DNA from non-malignant cells, we conducted in vitro expansion of polyclonal T cells from frozen blood or bone marrow samples. These samples were depleted of myeloid cells by adherence to tissue culture flasks. The remaining cells were cultured for 7-14 days in IMDM (Gibco Invitrogen Corporation) supplemented with human serum (10%, PAA Laboratories GMBH), IL-2 (100 IU/ml) and CD3/CD28 coated beads (Dynabeads, Invitrogen). The purity of T-cells was determined by FACS analysis of CD3+ cells. DNA was extracted from T cells when the purity was > 90%.

Allelic discrimination assay: PCR with allele-specific probes. We examined DNA obtained from 250 healthy donors for the presence of the identified homozygous missense variant c.2025C>A (subject 97). Probes (**Supplementary Table 2** online) were designed for recognition of the variant (FAM-MGB probes, Applied Biosystems) and wild-type (VIC-MGB probes, Applied Biosystems) sequences. Primers (**Supplementary Table 2** online) spanning the site of variation were used in combination with these probes to generate a PCR product of 115 bp using the 7900HT Fast Real-Time PCR System (Applied Biosystems). MDS samples without this mutation (subjects 63 and 73) and the MDS sample with this homozygous missense mutation (subject 97) were included in each run in triplicate. Artificial heterozygotes were created by 1:1 mixing of DNA from subject 97 and subject 63 (triplicate) and subject 97 and subject 73 (triplicate). End point VIC/FAM measurements were performed and values were corrected for passive reference ROX values. Corrected values were plotted automatically by the SDS 2.3.2. allelic discrimination program (Applied Biosystems).

Allelic discrimination assay: PCR followed by high-resolution melting (HRM) analysis. We examined the DNA of 250 healthy donors for the presence of the newly identified missense variant c.2068C>T (subject 96). Primers (**Supplementary Table 2** online) spanning the site of variation and the MeltDoctor™ HRM Master Mix (Applied Biosystems) were used to generate a PCR product of 123 bp using the 7500 Fast Real-Time PCR System (Applied Biosystems). MDS samples without this mutation (subjects 63 and 73) and the MDS sample containing this heterozygous missense mutation (subject 96) were included in each run in triplicate. Melting temperatures of the generated amplicons were recorded during the subsequent dissociation stage by the 7500 System SDS Software. Variants were auto-called and grouped by the HRMv1.0 software (Applied Biosystems).

Nature Genetics: doi:10.1038/ng.620

RNA isolation and reverse transcriptase reactions. RNA was extracted from total bone marrow or expanded T cells using TRIzol Reagent (Invitrogen) or RNA-Bee (Bio-Connect). Traces of genomic DNA were depleted from the isolated RNA by DNAse I treatment (Invitrogen) according to the manufacturer's instructions. Reverse transcriptase (+RT) reactions on 1 g of RNA were performed using M-MLV reverse transcriptase (Invitrogen). Negative controls containing no reverse transcriptase (-RT) were included to detect remaining traces of genomic DNA.

Confirmation of the c.2195+1G>A mutation (subject 10) at the transcript level. Subject 10 carried a heterozygous donor splice site mutation in intron 19 (c.2195+1G>A). To determine the effect on splicing, we designed PCR primers spanning the affected intron: forward primer 5' - CCAAACTGCTATGCAAAAGTTATGAT - 3' (exon 18/19 boundary) and reverse primer 5' - CAGATGTCAAGGGATTTCCATTTC - 3' (exon 20). PCR reactions were performed on cDNA (-RT and +RT). PCR products were analyzed by agarose gel (2%) electrophoresis and by sequence analysis of products that were cloned into the pDrive cloning vector (Qiagen).

Confirmation of the c.745G>A mutation (subject 10) at the transcript level. Subject 10 carried a heterozygous missense mutation in exon 8 (c.745G>A). To determine the presence of this mutation in RNA transcripts, we designed primers for amplification and pyrosequencing of cDNA, spanning the c.745G>A mutation:

forward amplification primer (exon 7)	5' - AATGTTTCCAGATAAGGGCACAG - 3'
reverse amplification primer (exon 8)	5' - TTAGCATTTGGTCCATCTATGTTG - 3'
forward pyrosequencing primer (exon 7/8 boundary)	5' - AGGAAAAATATAAAGAACTC - 3'

Confirmation of the c.1505+1G>T mutation (subject 73) at the transcript level. Subject 73 carried a donor splice site mutation in intron 12 (c.1505+1G>T). To determine the effect on splicing, we designed two PCRs. For the 1st PCR, we used primers spanning the affected intron: forward primer 5' - GGACCAAAACATGTAGACAGGTGTA - 3' (exon 11/12 boundary) and reverse primer 5' - GCAGTGTGCAGCCCACAA - 3' (exon 13). For the 2nd PCR, we used primers spanning the entire 3' coding region (exon 11/12 to exon 20): forward primer 5' - GGACCAAAACATGTAGACAGGTGTA - 3' (exon 11/12 boundary) and reverse primer 5' - CAGATGTCAAGGGATTTCCATTTC - 3' (exon 20). PCR reactions were performed on cDNA (-RT and +RT). PCR products were analyzed by agarose gel (2%) electrophoresis and by sequence analysis of PCR products.

Confirmation of the c.1212-1216delGAAGA mutation (subject 127) at the transcript level. Subject 127 harbored a heterozygous c.1212-1216delGAAGA mutation in exon 10. To assess the presence of this mutation, we designed two PCR assays spanning the affected exon:

Nature Genetics: doi:10.1038/ng.620

forward primer A 5' - GACCACAGTGTTACCAGCATTTG - 3' (exon 9),
forward primer B 5' - AAGAGGACGGCTTCCCAATAA - 3' (exon 10) and
reverse primer 5' - TGTTTGACACCGAGAATTTGCT - 3' (exon 11, coupled to a 5' 6FAM label).
PCR reactions were performed on cDNA (-RT and +RT). PCR products were analyzed by agarose gel (2%) electrophoresis and by sequence analysis of products that were cloned into the pDrive cloning vector (Qiagen). GeneScan analysis was performed using the 6FAM-coupled PCR products in combination with the GeneScan™ 500 LIZ® Size Standard (Applied Biosystems) on a 3730 DNA Analyzer (Applied Biosystems).

Nature Genetics: doi:10.1038/ng.620