Supplemental information

Frequent reconstitution of IDH2^{R140Q} mutant clonal multilineage hematopoiesis following chemotherapy for acute myeloid leukemia

Daniel H. Wiseman, Emma L. Williams, Deepti P. Wilks, Hui Sun Leong, Tim D. D. Somerville, Michael W. Dennis, Eduard A. Struys, Abdellatif Bakkali, Gajja S. Salomons and Tim C. P. Somervaille

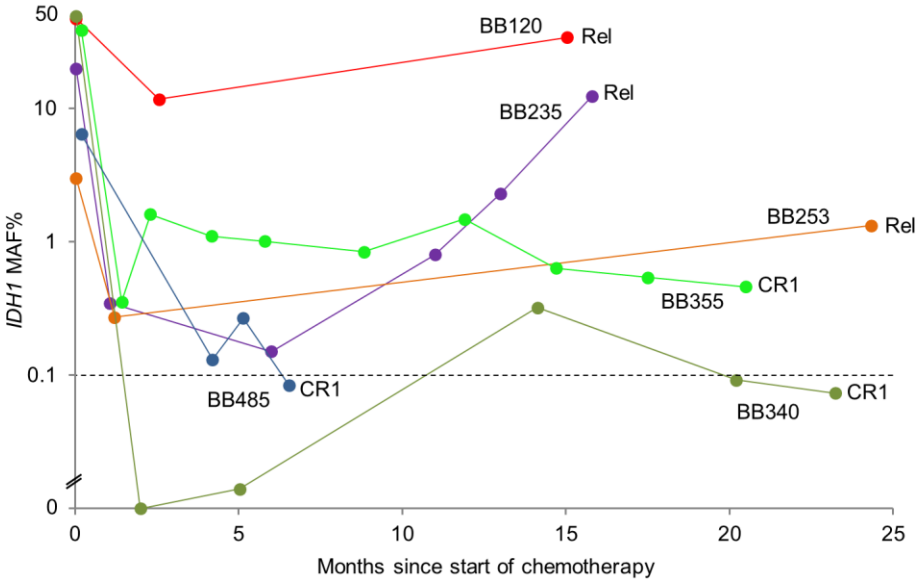

Figure S1. Sequential *IDH1* mutant allele frequencies in AML patients achieving morphologic CR. *IDH1* MAFs at presentation, at complete morphologic remission and following completion of intensive chemotherapy in six *IDH1*^{mut} AML patients treated with chemotherapy alone. Remission status is indicated (CR1, first complete remission; Rel, relapsed). MAFs were determined by digital PCR. Biobank identifiers are shown.

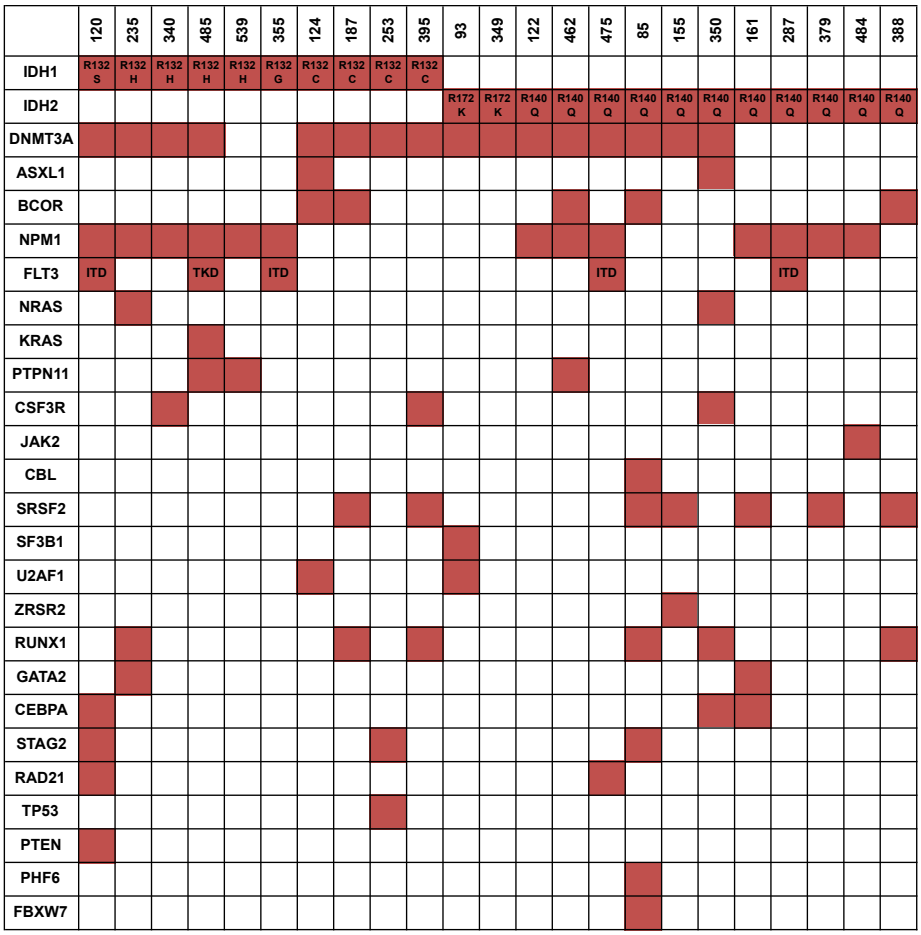

Figure S2. Mutation heat map. Concomitantly mutated genes in presentation samples of AML (n=21) or RAEB (n=2) patients exhibiting an *IDH1* or *IDH2* mutation who achieved complete morphologic remission following chemotherapy.

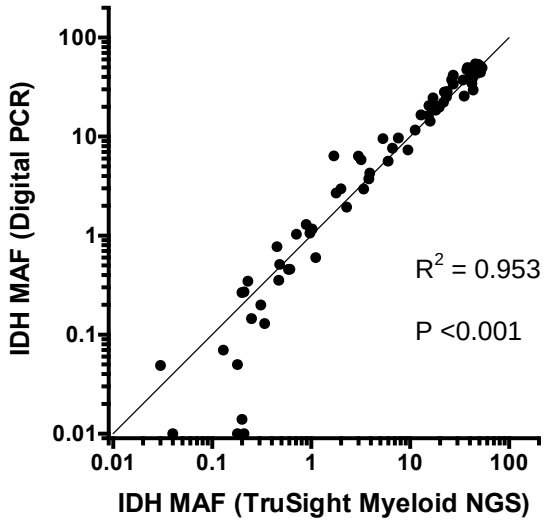

Figure S3. Concordance of digital PCR and targeted next generation sequencing (TruSight Myeloid Panel) for measurement of *IDH* mutant allele frequency. Eighty BM or blood samples were tested by both methods. R^2 = Pearson correlation coefficient. Digital PCR assays were previously demonstrated as quantitative to ≤0.1%.[8] Loss of concordance at lower MAFs reflects the impact of systematic high throughput sequencing noise.[S1]

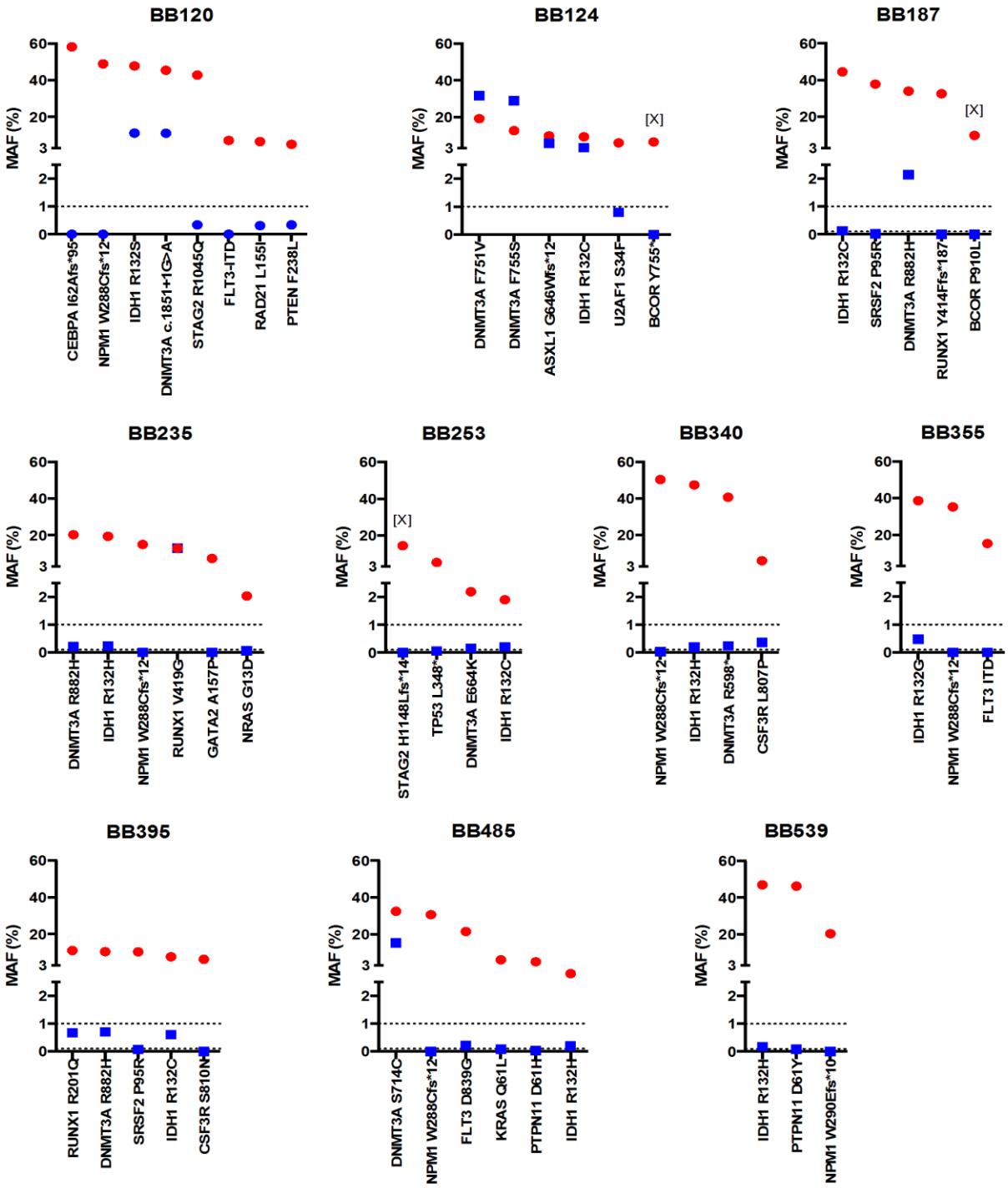

Figure S4. Somatic mutations identified by targeted next generation sequencing (TruSight Myeloid Panel) at presentation and onset of complete morphological remission in *IDH1*-mutated patients. Red circles show MAFs at presentation. Blue squares show MAFs at CR1 following one or two cycles of chemotherapy. [X] indicates location of gene on the X chromosome.

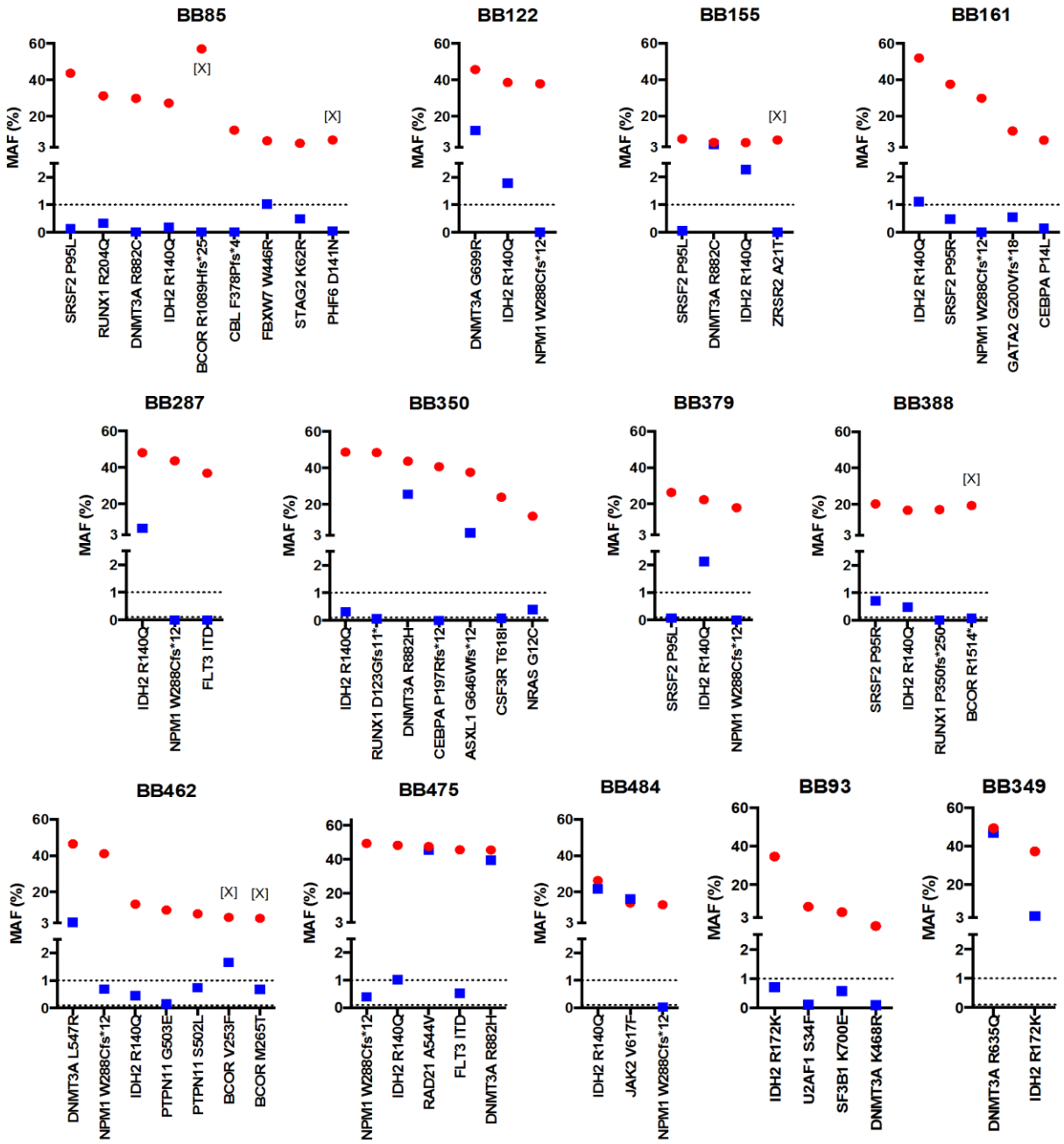

Figure S5. Somatic mutations identified by targeted next generation sequencing (TruSight Myeloid Panel) at presentation and onset of complete morphological remission in *IDH2*-mutated patients. Red circles show MAFs at presentation. Blue squares show MAFs at CR1 following one or two cycles of chemotherapy. [X] indicates location of gene on the X chromosome.

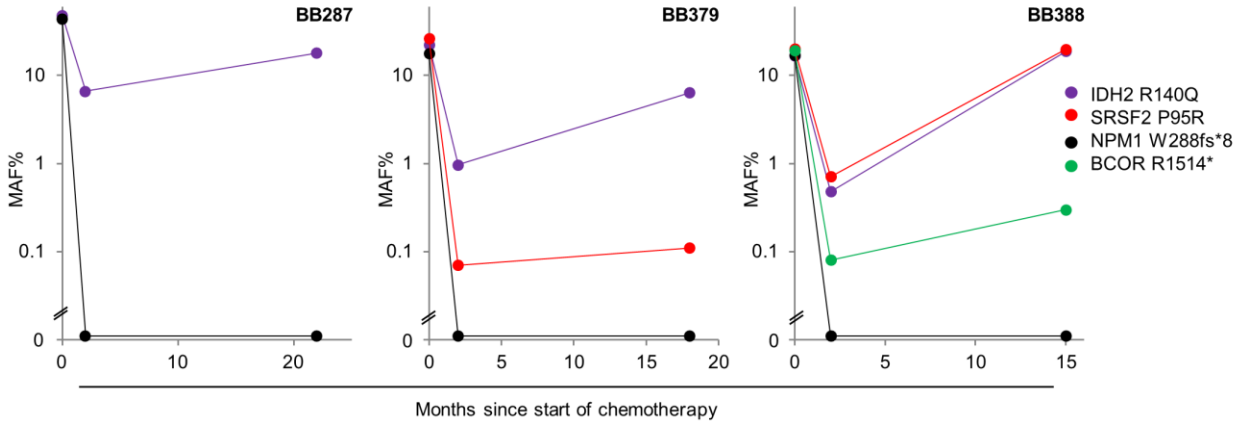

Figure S6. Sequential mutant allele frequencies in selected IDH2^{R140Q} mutated patients in CR1. Graphs show MAFs for somatic mutations identified by targeted next generation sequencing (TruSight Myeloid Panel) at presentation, onset of complete morphological remission and at the indicated follow up time points during sustained remissions in *IDH2*-mutated patients.

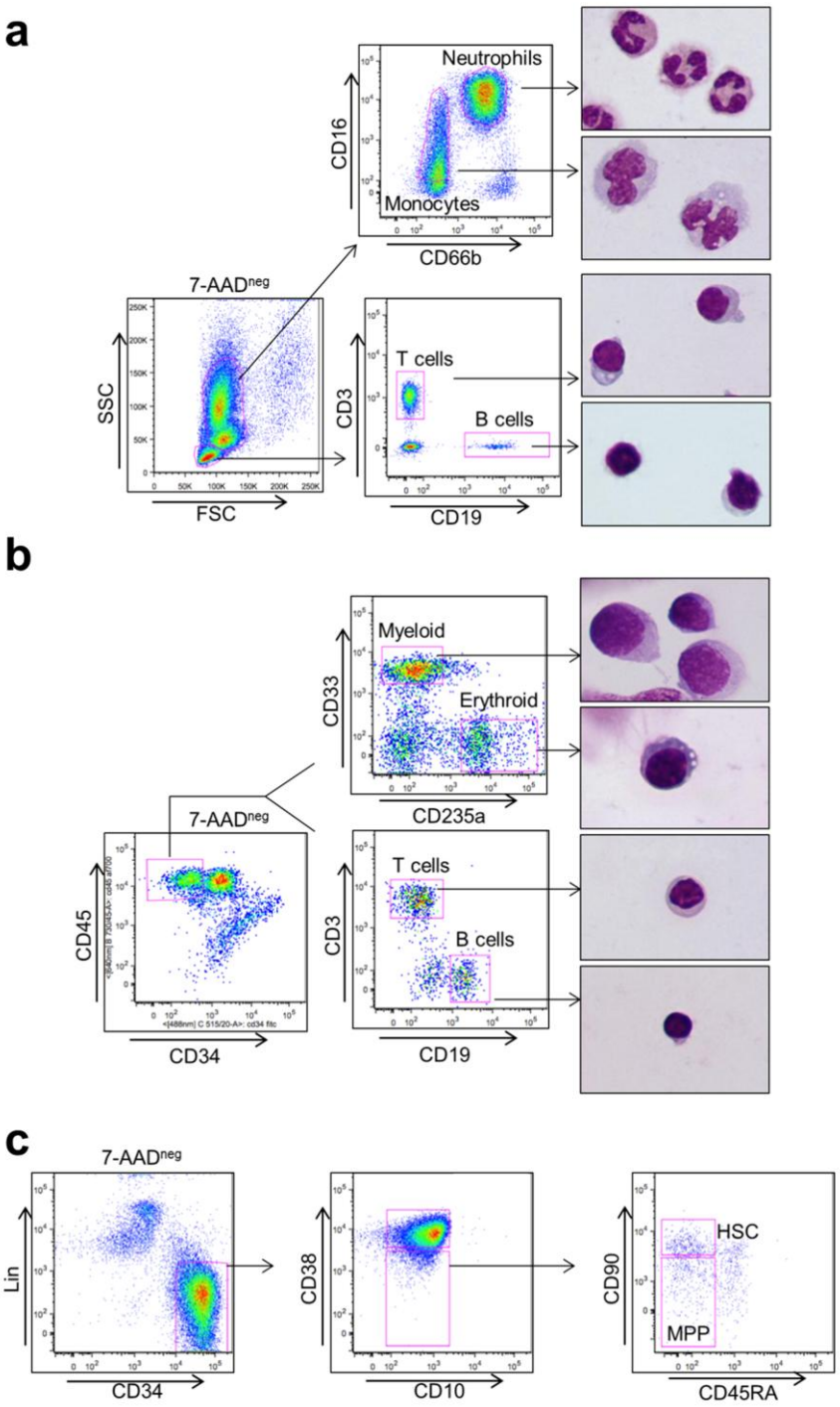

Figure S7. Gating strategies and outcomes for flow sorting experiments. Flow cytometry plots show gating strategies for sorting experiments with (a) blood mononuclear cells, (b) bone marrow $CD34^{-}$ cells and (c) bone marrow immunophenotypic hematopoietic stem cells (HSC) and multipotent progenitor cells (MPP). Purities of sorted populations from (a) and (b) were 97-100%, as determined by morphological analysis of cytospin preparations and post-sort immunophenotypic analyses. In view of the small number of sorted cells in (c) post-sort analyses were not possible.

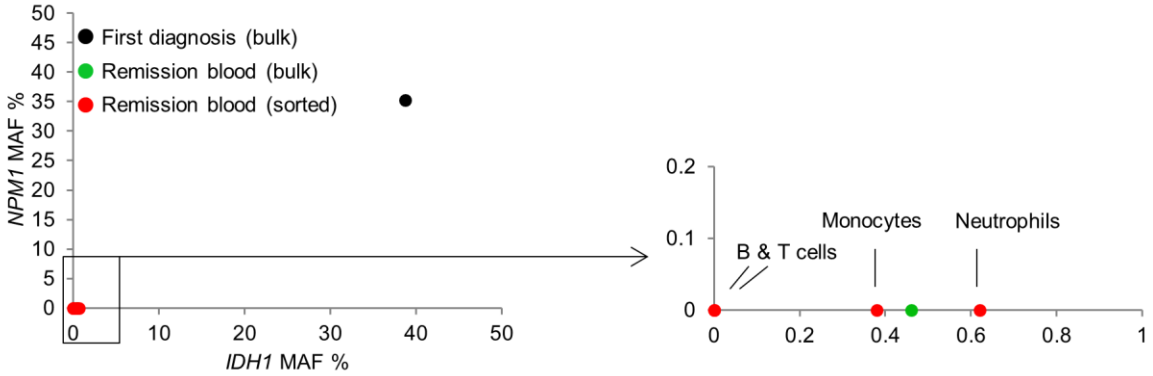

Figure S8. Multilineage contribution of IDH1^{R132G} mutant clonal hematopoiesis in remission. Graphs show IDH1^{R132G} versus NPM1 MAFs at the indicated time points and in the indicated cell populations in patient BB355. The remission sample was collected 22 months following presentation.

Table S1. Characteristics and outcome data for an *IDH*-mutant AML/RAEB cohort achieving complete morphologic remission (n=23).

ID #	Age/ Gender	Disease	Blast (%)	IDH MUTATION			MAF (%)	Karyotype	Treatment	Follow Up (Months)	Outcome
				Gene	CDS	AA					
85	66/M	RAEB	13%^{M}	*IDH2*	c.419G>A	p.R140Q	41.8^{M}	Normal	ADE x2 → CR1 Rel1: Azacitidine	25.2	Died (Relapse)
93	58/M	AML	35%^{M}	*IDH2*	c.515G>A	p.R172K	25.7^{M}	Normal	ADE x1 (Ref) → FLAGIda x1 → CR1; HSCT in CR1	17.2	Died (Relapse)
120	68/F	AML	97%^{M}	*IDH1*	c.394C>A	p.R132S	47.2^{M}	Normal	DA x2 → CR1 Rel1: FLAGIda x1 + HDAC x1 → CR2; (*unfit for HSCT*) Rel2: Azacitidine	33.7	Died (Relapse)
122	66/M	AML	28%^{M}	*IDH2*	c.419G>A	p.R140Q	48.8^{B}	Normal	ADE+GO (Ref) → Clofarabine x1 → CR1; HSCT in CR1	6.1	Died (in CR1): *Early HSCT complications*
124	54/M	RAEB	14%^{M}	*IDH1*	c.394C>T	p.R132C	7.3^{B}	del(20q)	DA+GO x1, FLAGIda x1 → CR1; HSCT in CR1 Rel1: Azacitidine	37.5	Died (Relapse)
155	63/M	AML	80%^{M}	*IDH2*	c.419G>A	p.R140Q	10.8^{B}	Normal	ADE+GO x1, ADE x1 → CR1; HSCT in CR1	53.4	Alive (CR1)
161	66/F	AML	80%^{M}	*IDH2*	c.419G>A	p.R140Q	51.5^{M}	Normal	ADE+GO x1, ADE x1, HDAC x2 → CR1	52.0	Alive (CR1)
187	61/M	AML	62%^{M}	*IDH1*	c.394C>T	p.R132C	26.0^{B}	+8	DA x2 → CR1; HSCT in CR1	17.9	Died (in CR1): *Late HSCT complications*
235	31/M	AML	20%^{M}	*IDH1*	c.395G>A	p.R132H	19.9^{B}	Normal	DA x2, HDAC x2 → CR1 Rel1: FLAGIda x1 → CR2; HSCT in CR2	39.5	Alive (CR2)
253	73/M	AML	44%^{M}	*IDH1*	c.394C>T	p.R132C	3.0^{B}	Normal	DA x1 → CR1 (*unfit for further intensive chemotherapy*) Rel1: Azacitidine	27.7	Died (Relapse)
287	47/F	AML	90%^{M}	*IDH2*	c.419G>A	p.R140Q	53.7^{M}	Normal	DA x2, HDAC x2 → CR1	31.1	Alive (CR1)
340	57/F	AML	93%^{M}	*IDH1*	c.395G>A	p.R132H	49.1^{M}	Normal	DA x2, HDAC x2 → CR1	24.4	Alive (CR1)
349	43/M	AML	56%^{M}	*IDH2*	c.515G>A	p.R172K	37.3^{M}	Normal	DA (*refractory*) → FLAGIda x2 → CR1; (*no HSCT donor option*) Rel1: Supportive care	19.8	Alive (Relapse)
350	69/F	AML	65%^{M}	*IDH2*	c.419G>A	p.R140Q	52.0^{M}	Normal	DA x1, HDAC x2 → CR1 Rel1: Azacitidine + Vorinostat (*ongoing*)	22.6	Alive (Relapse)
355	65/M	AML	82%^{M}	*IDH1*	c.394C>G	p.R132G	38.6^{B}	Normal	DA x1 → CR1	21.8	Alive

									(unfit for further intensive chemotherapy)		(CR1)
379	65/M	AML	50%^{M}	*IDH2*	c.419G>A	p.R140Q	28.3^{B}	Normal	DA x2, HDAC x1 → CR1	20.5	Alive (CR1)
388	58/M	AML	36%^{M}	*IDH2*	c.419G>A	p.R140Q	24.7^{B}	+8	DA (*refractory*) → FLAGIda x2 → CR1; (*no HSCT donor option*)	19.3	Alive (CR1)
395	60/M	AML	23%^{M}	*IDH1*	c.394C>T	p.R132C	10.6^{B}	Normal	DA x2, HDAC x1 → CR1; HSCT in CR1	18.8	Alive (CR1)
462	48/M	AML	35%^{M}	*IDH2*	c.419G>A	p.R140Q	16.6^{M}	Normal	DA x2, HDAC x2 → CR1	11.0	Alive (CR1)
475	33/M	AML	95%^{B}	*IDH2*	c.419G>A	p.R140Q	50.3^{B}	Normal	DA x2, HDAC x2 → CR1; Rel1: FLAGIda x1 → CR2; (*HSCT in CR2 planned but early relapse*) Rel2: Sorafenib	9.4	Alive (Relapse)
484	72/M	AML	39%^{M}	*IDH2*	c.419G>A	p.R140Q	37.5^{M}	+Y	LDAC x7 → CR1 after 2nd cycle (*ongoing*)	9.0	Alive (CR1)
485	74/M	AML	35%^{M}	*IDH1*	c.395G>A	p.R132H	6.4^{B}	Normal	DA+Ganetespib x3 → CR1	8.2	Alive (CR1)
539	18/M	AML	95%^{M}	*IDH1*	c.395G>A	p.R132H	43.6^{M}	Normal	DA x2 → CR1 (*ongoing*); (*Awaiting HSCT in CR1*)	2.6	Alive (CR1)

For blast percentages and IDH mutant allele frequencies (MAF) the results from the presentation bone marrow (^{M}) or blood sample (^{B}) are shown. Blood samples were used for analyses only where a bone marrow test was not performed at presentation, or where there was insufficient surplus bone marrow sample for biobanking. Blast percentages are from morphologic analysis of bone marrow or blood smears. IDH MAF was determined by digital PCR. Treatment regimens: ADE: cytarabine, daunorubicin, etoposide; DA: daunorubicin, cytarabine; HDAC: high dose cytarabine; LDAC: low dose cytarabine; FLAGIda: fludarabine, cytarabine, idarubicin, granulocyte-colony stimulating factor; GO: gemtuzumab ozogamicin; HSCT: hematopoietic stem cell transplant. Disease status is recorded on 1st February 2016. All deaths were due to relapsed/refractory disease unless indicated. Other abbreviations: AML = acute myeloid leukemia; RAEB = refractory anemia with excess blasts; CR = complete remission; Rel = relapsed disease.

Table S2. Blood counts of individuals with *IDH*-mutant clonal hematopoiesis during sustained remission, at the latest time point indicated in Figure 1f. All values are within age and sex specific reference ranges except those in pink colored boxes, where normal ranges are shown.

Biobank ID	Hemoglobin (g/l)	Total white cell count ($x10^9$/l)	Neutrophils ($x10^9$/l)	Monocytes ($x10^9$/l)	Platelets ($x10^9$/l)
IDH2^{R140Q}					
BB161	113 (115-165)	6.1	3.3	1.0 (0.2-0.8)	331
BB287	147	6.8	3.2	0.3	169
BB379	155	4.5	2.7	0.2	202
BB388	133	4.1	2.3	0.2	77 (150-400)
BB462	145	4.9	3.0	0.4	172
BB484	144	8.1	4.1	0.7	263
IDH1^{R132H}					
BB340	141	8.5	4.8	0.4	206
IDH1^{R132G}					
BB355	159	4.5	3.0	0.3	182

Table S3: Putative somatic mutations identified in 23 RAEB/AML genomes by targeted resequencing using Illumina TruSight Myeloid Panel.

Shown for each genomic variant are genomic location, the consequence of the base change on the protein coding sequence, the COSMIC ID (if found in the Catalogue of Somatic Mutations in Cancer Database) [S2] and the calculated mutant allele frequency (MAF%) at presentation and remission (after one or two cycles of induction chemotherapy).

ID #	Chr	Position	Gene	Variant	Type	CDS	AA Change	COSMIC	PRESENTATION			REMISSION		
									Reads	Variant	MAF%	Reads	Variant	MAF%
85	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	4405	1201	**27.26**	7795	14	**0.18**
85	2	25457243	*DNMT3A*	G>A	Missense	c.2644C>T	p.R882C	COSM53042	4469	1336	**29.89**	4850	1	**0.02**
85	X	39923823	*BCOR*	-AC/ +TGAT	Complex frameshift (elongation)	c.3266_3267GT >ATCA	p.R1089Hfs*25	COSM520115	594	339	**57.07**	1397	3	**0.21**
85	11	119148910	*CBL*	-AT	Frameshift del (truncation)	c.1131_1132del AT	p.F378Pfs*4	-	542	67	**12.36**	7732	1	**0.01**
85	4	153249442	*FBXW7*	A>T	Missense	c.1336T>A	p.W446R	-	382	25	**6.54**	5159	53	**1.03**
85	X	133547523	*PHF6*	G>A	Missense	c.421G>A	p.D141N	COSM1490507	545	38	**6.97**	4019	2	**0.05**
85	21	36231773	*RUNX1*	C>T	Missense	c.611G>A	p.R204Q	COSM24731	3460	1081	**31.24**	4282	14	**0.33**
85	17	74732959	*SRSF2*	G>A	Missense	c.284C>T	p.P95L	COSM146288	2493	1091	**43.76**	2992	4	**0.13**
85	X	123164872	*STAG2*	A>G	Missense	c.185A>G	p.K62R	-	664	34	**5.12**	7625	37	**0.49**
93	15	90631838	*IDH2*	C>T	Missense	c.515G>A	p.R172K	COSM33733	4068	1410	**34.66**	2100	15	**0.71**
93	2	198266834	*SF3B1*	T>C	Missense	c.2098A>G	p.K700E	COSM84677	6211	357	**5.75**	3497	20	**0.57**
93	21	44524456	*U2AF1*	G>A	Missense	c.101C>T	p.S34F	COSM166866	9341	803	**8.60**	5694	6	**0.11**
93	2	25469055	*DNMT3A*	T>C	Missense	c.1403A>G	p.K468R	COSM1318923	2079	55	**2.65**	1116	1	**0.09**
120	2	209113113	*IDH1*	G>T	Missense	c.394C>A	p.R132S	COSM28748	7588	3625	**47.77**	10328	1160	**11.23**
120	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTC TG	p.W288Cfs*12	COSM158604	809	396	**48.95**	3130	0	**0.00**
120	2	25467023	*DNMT3A*	C>T	Splice donor variant	c.1851+1G>A	n/a	-	5065	2304	**45.49**	6403	713	**11.14**
120	13	28608253 28608282 28608314	*FLT3*	+36bp +56bp +64bp	ITD (x3)	n/a	n/a (ITD)	-	n/a	n/a	**7.21**	n/a	n/a	**0.00**
120	19	33793131	*CEBPA*	-GTCGATGGAC	Frameshift del (truncation)	c.180_189delGT CCATCGAC	p.I62Afs*95	COSM29209	644	375	**58.23**	1702	0	**0.00**
120	10	89717687	*PTEN*	T>C	Missense	c.712T>C	p.F238L	COSM36318	837	43	**5.14**	8040	27	**0.34**
120	8	117870609	*RAD21*	A>T	Missense	c.463T>A	p.L155I	-	557	37	**6.64**	9598	30	**0.31**
120	X	123220477	*STAG2*	G>A	Missense	c.3134G>A	p.R1045Q	COSM1227727	11236	4816	**42.86**	15063	51	**0.34**

122	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	16641	6420	**38.58**	12146	217	**1.79**
122	2	25463587	*DNMT3A*	C>G	Missense	c.2095G>C	p.G699R	-	5098	2332	**45.74**	3554	434	**12.21**
122	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	4762	1804	**37.88**	2901	0	**0.00**
124	2	209113113	*IDH1*	G>A	Missense	c.394C>T	p.R132C	COSM28747	12222	1134	**9.28**	11405	386	**3.38**
124	2	25463242	*DNMT3A*	A>C	Missense	c.2251T>G	p.F751V	-	28707	5511	**19.20**	24163	7664	**31.72**
124	2	25463229	*DNMT3A*	A>G	Missense	c.2264T>C	p.F755S	-	28642	3630	**12.67**	24104	6994	**29.02**
124	21	44524456	*U2AF1*	G>A	Missense	c.101C>T	p.S34F	-	25577	1539	**6.02**	19271	154	**0.80**
124	X	39932334	*BCOR*	+T	Frameshift ins (elongation)	c.2264dupA	p.Y755*	-	13917	910	**6.54**	10853	0	**0.00**
124	20	31022441	*ASXL1*	+G	Frameshift ins (elongation)	c.1926_1927insG	p.G646Wfs*12	COSM1411076	10720	1047	**9.77**	9255	534	**5.77**
155	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	18120	980	**5.41**	12497	285	**2.28**
155	2	25457243	*DNMT3A*	G>A	Missense	c.2644C>T	p.R882C	COSM53042	13772	751	**5.45**	7410	350	**4.72**
155	17	74732959	*SRSF2*	G>A	Missense	c.284C>T	p.P95L	COSM146288	7397	551	**7.45**	3293	2	**0.06**
155	X	15809076	*ZRSR2*	G>A	Missense	c.61G>A	p.A21T	-	638	44	**6.90**	859	0	**0.00**
161	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	8039	4180	**52.00**	9377	104	**1.11**
161	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	2447	730	**29.83**	2201	0	**0.00**
161	17	74732959	*SRSF2*	G>T	Missense	c.284C>G	p.P95R	COSM211661	1123	422	**37.58**	2284	11	**0.48**
161	19	33793280	*CEBPA*	G>A	Missense	c.41C>T	p.P14L	-	622	42	**6.75**	727	1	**0.14**
161	3	128204841	*GATA2*	-C	Frameshift del (truncation)	c.599delG	p.G200Vfs*18	COSM1418772	330	39	**11.82**	364	2	**0.55**
187	2	209113113	*IDH1*	G>A	Missense	c.394C>T	p.R132C	COSM28747	10770	4797	**44.54**	11990	15	**0.13**
187	2	25457242	*DNMT3A*	C>T	Missense	c.2645G>A	p.R882H	COSM52944	7156	2434	**34.01**	9132	195	**2.14**
187	X	39931870	*BCOR*	G>A	Missense	c.2729C>T	p.P910L	-	558	55	**9.86**	395	0	**0.00**
187	21	36164634	*RUNX1*	+AGGA	Frameshift ins (elongation)	c.1237_1240dupTCCT	p.Y414Ffs*187	-	2468	805	**32.62**	3946	0	**0.00**
187	17	74732959	*SRSF2*	G>T	Missense	c.284C>G	p.P95R	COSM211661	3059	1158	**37.86**	4680	1	**0.02**
235	2	209113112	*IDH1*	C>T	Missense	c.395G>A	p.R132H	COSM28746	9064	1766	**19.48**	9053	21	**0.23**
235	2	25457242	*DNMT3A*	C>T	Missense	c.2645G>A	p.R882H	COSM52944	7550	1532	**20.29**	8266	17	**0.21**
235	3	128204972	*GATA2*	C>G	Missense	c.469G>C	p.A157P	-	353	26	**7.37**	471	0	**0.00**
235	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	2854	428	**15.00**	3463	0	**0.00**
235	21	36164619	*RUNX1*	A>C	Missense	c.469G>C	p.V419G	-	1086	140	**12.89**	1200	155	**12.92**
235	1	115258744	*NRAS*	C>T	Missense	c.38G>A	p.G13D	COSM573	13967	284	**2.03**	13587	8	**0.06**
253	2	209113113	*IDH1*	G>A	Missense	c.394C>T	p.R132C	COSM28747	2416	46	**1.90**	10142	21	**0.21**

253	2	25464523	*DNMT3A*	C>T	Missense	c.1990G>A	p.E664K	COSM256045	2596	57	**2.20**	11675	19	**0.16**
253	X	123224582	*STAG2*	+ACT	Frameshift ins (elongation)	c.3435_3436ins ACTGA	p.H1148Lfs*14	-	1795	258	**14.37**	5569	0	**0.00**
253	17	7573984	*TP53*	A>T	Stop-gain	c.1043T>A	TP53 L348*	COSM46015	548	29	**5.29**	1804	1	**0.06**
287	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	10543	5081	**48.19**	9861	652	**6.61**
287	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTC TG	p.W288Cfs*12	COSM158604	2382	1040	**43.66**	2438	0	**0.00**
287	13	28608266	*FLT3*	+31bp	ITD	n/a	n/a (ITD)	-	n/a	n/a	**36.90**	n/a	n/a	**0.00**
340	2	209113112	*IDH1*	C>T	Missense	c.395G>A	p.R132H	-	7154	3389	**47.37**	10525	21	**0.20**
340	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTC TG	p.W288Cfs*12	COSM158604	2726	1372	**50.33**	3608	1	**0.03**
340	2	25467083	*DNMT3A*	G>A	Stop-gain	c.1792C>T	p.R598*	COSM133736	5074	2066	**40.72**	7592	18	**0.24**
340	1	36932130	*CSF3R*	A>G	Missense	c.2420T>C	p.L807P	-	459	28	**6.10**	1361	5	**0.37**
349	15	90631838	*IDH2*	C>T	Missense	c.515G>A	p.R172K	COSM33733	8237	2828	**34.33**	8981	344	**3.83**
349	2	25466799	*DNMT3A*	C>T	Missense	c.1904G>A	p.R635Q	-	5734	2832	**49.39**	5852	2747	**46.94**
350	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	12789	6239	**48.78**	14720	45	**0.31**
350	2	25457242	*DNMT3A*	C>T	Missense	c.2645G>A	p.R882H	COSM52944	7915	3460	**43.71**	8700	2223	**25.55**
350	21	36252994	*RUNX1*	+CC	Frameshift ins (elongation)	c.366_367dupG G	p.D123Gfs11*	COSM1318809	11595	5623	**48.50**	11579	7	**0.06**
350	1	115258748	*NRAS*	C>A	Missense	c.34G>T	p.G12C	COSM562	14193	1909	**13.45**	14169	56	**0.40**
350	19	33792730	*CEBPA*	-G	Frameshift del (truncation)	c.590delC	p.P197Rfs*12	-	482	196	**40.66**	689	0	**0.00**
350	1	36933434	*CSF3R*	G>A	Missense	c.1853C>T	p.T618I	-	1892	453	**23.94**	2471	2	**0.08**
350	20	31022441	*ASXL1*	+G	Frameshift ins (elongation)	c.1926_1927ins G	p.G646Wfs*12	COSM1411076	9117	3435	**37.68**	10117	435	**4.30**
355	2	209113113	*IDH1*	G>C	Missense	c.394C>G	p.R132G	COSM28749	13374	5178	**38.72**	14393	67	**0.47**
355	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTC TG	p.W288Cfs*12	COSM158604	4334	1530	**35.30**	4322	0	**0.00**
355	13	28608299	*FLT3*	+61bp	ITD	n/a	n/a (ITD)	-	n/a	n/a	**15.36**	n/a	n/a	**0.00**
379	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	11653	2604	**22.35**	16872	164	**0.97**
379	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTC TG	p.W288Cfs*12	COSM158604	3133	562	**17.94**	3646	0	**0.00**
379	17	74732959	*SRSF2*	G>A	Missense	c.284C>T	p.P95L	COSM146288	4196	1106	**26.36**	7036	5	**0.07**
388	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	10443	1749	**16.75**	11151	54	**0.48**
388	21	36164827	*RUNX1*	+A	Frameshift ins (elongation)	c.1047dupT	p.P350fs*250	-	14665	2502	**17.06**	16166	0	**0.00**
388	17	74732959	*SRSF2*	G>T	Missense	c.284C>G	p.P95R	COSM211661	2858	579	**20.26**	2943	21	**0.71**
388	X	39916463	*BCOR*	G>A	Stop-gain	c.4540C>T	p.R1514*	COSM1599459	993	192	**19.34**	1243	1	**0.08**
395	2	209113113	*IDH1*	G>A	Missense	c.394C>T	p.R132C	COSM28747	7730	586	**7.58**	9699	59	**0.61**

395	1	36932121	*CSF3R*	C>T	Missense	c.2429G>A	p.S810N	-	546	34	**6.23**	1387	0	**0.00**
395	2	25457242	*DNMT3A*	C>T	Missense	c.2645G>A	p.R882H	COSM52944	6690	699	**10.45**	8271	58	**0.70**
395	21	36231782	*RUNX1*	C>T	Missense	c.602G>A	p.R201Q	COSM24805	6248	687	**11.00**	7509	50	**0.67**
395	17	74732959	*SRSF2*	G>T	Missense	c.284C>G	p.P95R	COSM211661	3145	324	**10.30**	4975	4	**0.08**
462	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	13166	1753	**13.31**	10481	47	**0.45**
462	2	25467436	*DNMT3A*	A>C	Missense	c.1640T>G	p.L547R	-	5044	2345	**46.49**	3656	122	**3.34**
462	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	3342	1375	**41.14**	3187	22	**0.69**
462	12	112926885	*PTPN11*	C>T	Missense	c.1505C>T	p.S502L	-	33285	2671	**8.02**	24279	179	**0.74**
462	12	112926888	*PTPN11*	G>A	Missense	c.1508G>A	p.G503E	COSM13021	33335	3344	**10.03**	24299	36	**0.15**
462	X	39933842	*BCOR*	C>A	Missense	c.757G>T	p.V253F	-	452	27	**5.97**	302	5	**1.66**
462	X	39933805	*BCOR*	A>G	Missense	c.794T>C	p.M265T	-	456	25	**5.48**	300	2	**0.67**
475	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	10150	4910	**48.37**	13970	142	**1.02**
475	2	25457242	*DNMT3A*	C>T	Missense	c.2645G>A	p.R882H	COSM52944	4753	2167	**45.59**	7180	2836	**39.50**
475	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	2495	1235	**49.50**	3823	18	**0.47**
475	13	28608285	*FLT3*	+41bp	ITD	n/a	n/a (ITD)	-	n/a	n/a	**45.63**	n/a	n/a	**0.53**
475	8	117861258	*RAD21*	G>A	Missense	c.1631C>T	p.A544V	-	3119	1485	**47.61**	4440	2013	**45.34**
484	15	90631934	*IDH2*	C>T	Missense	c.419G>A	p.R140Q	COSM41590	12158	3201	**26.33**	14896	3229	**21.68**
484	9	5073770	*JAK2*	G>T	Missense	c.1849G>T	p.V617F	COSM12600	13212	1838	**13.91**	14400	2323	**16.13**
484	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	3392	440	**12.97**	3401	1	**0.03**
485	2	209113112	*IDH1*	C>T	Missense	c.395G>A	p.R132H	COSM28746	11310	332	**2.94**	10507	21	**0.20**
485	2	25463541	*DNMT3A*	G>C	Missense	c.2141C>G	p.S714C	COSM87011	4533	1471	**32.45**	6225	951	**15.28**
485	13	28592629	*FLT3*	T>C	Missense	c.2516A>G	p.D839G	COSM1166729	18160	3887	**21.40**	17567	39	**0.22**
485	12	25380276	*KRAS*	T>A	Missense	c.182A>T	p.Q61L	COSM553	5256	317	**6.03**	5407	5	**0.09**
485	5	170837543	*NPM1*	+TCTG	Frameshift ins (elongation)	c.859_860insTCTG	p.W288Cfs*12	COSM158604	4058	1242	**30.61**	3555	0	**0.00**
485	12	112888165	*PTPN11*	G>C	Missense	c.181G>C	p.D61H	COSM20900	17158	851	**4.96**	17198	7	**0.04**
539	2	209113112	*IDH1*	C>T	Missense	c.395G>A	p.R132H	COSM28746	9363	4395	**46.94**	11091	20	**0.18**
539	5	170837551	*NPM1*	-TGGAGGAAG/ +GAAGTTTCTCGCC	Complex frameshift (elongation)	c.868_876TGGAGGAAG>GAAGTTTCTCGCC	p.W290Efs*10	-	1687	343	**20.33**	3703	0	**0.00**
539	12	112888165	*PTPN11*	G>T	Missense	c.181G>T	p.D61Y	COSM13011	14545	6722	**46.22**	17077	16	**0.09**

Table S4. Flow sorting antibody combinations

	Fluorochrome	***Clone***	***Supplier***
Blood			
Anti-CD66B	FITC	G10F5	BD Biosciences
Anti-CD16	PE-Cy7	3G8	BD Biosciences
Anti-CD19	PE	SJ25C1	eBioscience
Anti-CD3	APC-eFluor780	UCHT1	eBioscience
Bone Marrow *(CD34-depleted)*			
Anti-CD34	FITC	AC136	Miltenyi Biotec
Anti-CD33	PE-Cy7	WM-53	eBioscience
Anti-CD3	APC-eFluor780	UCHT1	eBioscience
Anti-CD19	eFluor450	SJ25C1	eBioscience
Anti-CD235a	PE	HIR2	eBioscience
Bone Marrow *(CD34-enriched)*			
(Primary incubation)			
lineage cocktail			
Anti-CD2	PE	RPA-2.10	BD Biosciences
Anti-CD3	PE	UCHT1	eBioscience
Anti-CD4	PE	L3T4	eBioscience
Anti-CD7	PE	eBio124-1D1	eBioscience
Anti-CD8	PE	OKT8	eBioscience
Anti-CD11b	PE	ICRF44	eBioscience
Anti-CD14	PE	61D3	eBioscience
Anti-CD19	PE	SJ25C1	eBioscience
Anti-CD20	PE	2H7	eBioscience
Anti-CD56	PE	B159	BD Biosciences
Anti-CD235a	PE	HIR2	eBioscience
(Secondary incubation)			
Anti-CD34	FITC	AC136	Miltenyi Biotec
Anti-CD38	PE-Cy7	HIT2	eBioscience
Anti-CD10	BV-421	HI10a	BD Biosciences
Anti-CD90	BV-510	5E10	BD Biosciences
Anti-CD45RA	AlexaFluor700	HI100	BD Biosciences
Anti-CD110	APC	1.6.1	BD Pharmingen
Anti-CD123	BV-786	7G3	BD Biosciences

Table S5. TruSight Myeloid Sequencing Panel genes

Gene	Targeted Region / Exons	Gene	Targeted Region / Exons
ABL1	4-6	JAK3	13
ASXL1	12	KDM6A	Full
ATRX	8-10; 17-31	KIT	2; 8-11; 13; 17
BCOR	Full	KRAS	2-3
BCORL1	Full	MLL (KMT2A)	5-8
BRAF	15	MPL	10
CALR	9	MYD88	3-5
CBL	8-9	NOTCH1	26-28; 34
CBLB	9-10	NPM1	12
CBLC	9-10	NRAS	2-3
CDKN2A	Full	PDGFRA	12; 14; 18
CEBPA	Full	PHF6	Full
CSF3R	14-17	PTEN	5; 7
CUX1	Full	PTPN11	3; 13
DNMT3A	Full	RAD21	Full
ETV6	Full	RUNX1	Full
EZH2	Full	SETBP1	4 (partial)
FBXW7	9-11	SF3B1	13-16
FLT3	14-15; 20	SMC1A	2; 11; 16-17
GATA1	2	SMC3	10; 13; 19; 23; 25; 28
GATA2	2-6	SRSF2	1
GNAS	8-9	STAG2	Full
HRAS	2-3	TET2	3-11
IDH1	4	TP53	2-11
IDH2	4	U2AF1	2; 6
IKZF1	Full	WT1	7; 9
JAK2	12; 14	ZRSR2	Full

Materials and methods

Study design, genotyping, digital PCR and plasma 2-HG measurements

The 23 *IDH* mutant cases were from a cohort of 241 adult patients with MDS or AML presenting to The Christie NHS Foundation Trust, Manchester, UK, of whom 46 were genotyped by Sanger sequencing as *IDH1*- or *IDH2*-mutated. *IDH* genotyping was performed as described.[8] Tissue and plasma samples were from the Manchester Cancer Research Centre Biobank, instituted with the approval of the South Manchester Research Ethics Committee. Analyses and experiments were ethically approved by the Biobank's scientific sub-committee. Bone marrow (BM) and blood mononuclear cells (MNCs) and platelet-poor plasma samples were cryopreserved at presentation and multiple time points during follow up, as described.[S3,S4] All *IDH*-mutated patients for whom at least presentation and remission samples were available, and who received chemotherapy to complete remission (CR) (n=23), were included in the study. *IDH* (and *NPM1*, where present) mutant allele frequency (MAF) was determined for each serial MNC sample in the study cohort by quantitative digital PCR (dPCR), as previously described.[8] Plasma 2-hydroxyglutarate was assayed as previously described.[8,S5]

Flow sorting of hematopoietic stem/progenitor and downstream lineage populations

Fresh BM and blood were obtained from two cases with residual clonal $IDH2^{R140Q}$ hematopoiesis (BB161, BB287) at remission time points 40 and 18 months following presentation, respectively. Analyses were also performed on cryopreserved remission material from BB161 20 months following presentation. Fresh blood was also obtained from one case with residual clonal $IDH1^{R132G}$ hematopoiesis (BB355) 22 months after presentation. For BM experiments, MNCs were isolated by density gradient centrifugation (Lymphoprep, Stem Cell Technologies, Cambridge, UK). The $CD34^+$ compartment was isolated by magnetic bead separation using an autoMACS Pro (Miltenyi Biotec, Bisley, UK), according to manufacturer protocols. Blood leukocytes were isolated by buffy coat separation to preserve neutrophil content. The CD34 enriched BM, CD34-depleted BM and peripheral blood cell fractions were each separately sorted by fluorescence activated cell sorting (FACS) on a BD FACSAria™ III cell sorter (BD Biosciences, Oxford, UK). Live-dead discrimination was by inspection of forward/side scatter and 7-AAD incorporation. Antibody panels and gating strategies are indicated in Table S4 and Figure S7 respectively. Population purities were confirmed, where cell numbers permitted, by post-sort flow analyses and inspection of morphology on cytospin preparations.

Colony forming assays, single colony expansion and colony genotyping

Fresh $CD34^+$ BM cells from BB161 and BB287 were harvested and plated in Methocult H4230 (Stem Cell Technologies) (2000 cells/ml, in triplicate) supplemented with the following cytokines: IL-3 20ng/mL, IL-6 20ng/mL, IL-11 10ng/mL, FLT3L 50ng/mL, SCF 50ng/mL, TPO 50ng/mL (all from Peprotech, London, UK), G-CSF 50ng/mL (Chugai, London, UK) and EPO 4U/mL (Janssen Cilag, High Wycombe, UK). For single cell sorting of immunophenotypic HSCs, $CD34^+38^-90^+Lin^-$ cells were flow sorted directly into U-bottom 96 well plates containing 100uL of Methocult per well supplemented with cytokines. Plates were incubated at 37°C

and 5% CO_2 for 14 days. Colonies were typed by inverted microscopy, plucked and genomic DNA extracted using the Taqman® Sample-To-SNP Kit (Life Technologies, Paisley, UK) according to manufacturer's instructions. DNA was pre-amplified using Taqman® PreAmp MasterMix (Life Technologies), according to the manufacturer's protocol. Amplified DNA was genotyped by allelic discrimination dPCR. 5µL reaction mixes containing 1µL DNA with 2.5µL 2X Taqman® GTXpress Master Mix, 0.25µL 20X Taqman® $IDH2^{R140Q}$ SNP genotyping assay and 1.25µL water were subjected to the following thermal cycling conditions on an ABI Prism 7900HT (Applied Biosystems): 95°C for 20s, followed by 40 cycles of 95°C for 3s and 60°C for 20s. ROX was used as passive reference.

Targeted next generation sequencing

Illumina TruSight Myeloid Panel

Genomic DNA from BM or blood MNCs collected at presentation and remission time points was subjected to targeted next generation sequencing (NGS) using a NextSeq 500 sequencer (Illumina, San Diego, CA, USA). Illumina's amplicon-based TruSight Myeloid Sequencing Panel interrogates 56 genes recurrently mutated in myeloid neoplasms (Table S5), using a proprietary multiplexed oligonucleotide pool covering each region of interest. Each primer includes universal adaptor sequences, subsequently incorporated in the amplification reaction. Amplicons are of standardized length (~250bp) for an overall library size of 141 kb. Libraries were generated according to the manufacturer's protocol. Briefly, genomic DNA was quantified using a Qubit DNA BR assay kit (Life Technologies, Carlsbad, CA, USA) and diluted to 50ng in 96 well plates. Oligonucleotides were hybridized to regions of interest, followed by an extension-ligation reaction and PCR amplification, incorporating unique combinations of i5/i7 index sequences to permit multiplexing up to 96 samples per sequencing run. Successful amplification was confirmed using a DNA 1000 kit and the 2100 Bioanalyzer system (Agilent Technologies, CA, USA). Libraries were purified using AMPure magnetic beads (Agencourt, Brea, CA, USA) and bead-normalized according to the TruSight protocol. Libraries were pooled (5µL per library, 96 per pool) and quantified by PCR to determine molarity for loading onto the NextSeq flow cell to achieve optimal cluster density (170-220k/m^2). The pooled library was denatured, diluted and loaded onto a reagent cartridge according to Illumina's protocol. Paired end (150bp) sequencing was performed on the NextSeq 500 sequencer (Illumina) with 96 samples multiplexed on a single NextSeq 500 High Output run (300 cycles).

Alignments, variant calling and annotation

Data analysis was performed within Illumina's online BaseSpace genomics analysis platform. FASTQ files were aligned to human genome reference GRCh37/hg19 by the TruSeq Amplicon App (v2.0; Illumina; https://basespace.illumina.com/apps/2005003/TruSeq-Amplicon) using a banded Smith-Waterman algorithm. Variant calling was performed by Somatic Variant Caller (v4.0.13.1; Illumina) using default parameters. The resulting gVCF files were uploaded to Variant Studio (v2.2.3; Illumina; https://basespace.illumina.com/apps/639639/VariantStudio-App) for downstream filtering and annotation (from RefSeq database) of high confidence variants. Integrative Genomics Viewer (v2.1.2; Broad

Institute/Illumina; https://basespace.illumina.com/apps/1886885/Integrative-Genomics-Viewer) [S6] was used to confirm regional coverage, visualize read alignments and confirm variant calls.

Variant allele frequency was calculated as the fraction of mutated reads versus total number of reads covering that base. For mutation discovery on presentation samples, sequencing coverage of 250X (bi-directional) and minimum variant frequency of 5% were used as lower thresholds for variant calling, as employed previously.[S7] Filters were applied to exclude known germline SNPs with population frequency >1%. However, since unequivocally pathogenic/somatic SNVs also have dbSNP database entries (including all common AML-associated IDH1/2 mutations), presence on dbSNP was not itself an exclusion criterion. All variants passing initial filtering were manually interrogated and those suspicious for germline polymorphisms were excluded. To screen for lower frequency variants, all those detected at 1-5% frequency were manually interrogated and included if they (a) corresponded to a COSMIC database entry [S1] with multiple reports in hematopoietic tissue of a confirmed somatic variant; (b) were supported by >40 individual variant reads; and (c) were convincingly absent (or present at <20% of the presentation sample level) in the corresponding remission sample. Among the six additional variants included by this approach were two known sub-clonal IDH1 mutations (BB253: $IDH1^{R132C}$ 1.9%; BB485: $IDH1^{R132H}$ 2.9%), both previously detected and quantitated to a similar level by highly sensitive digital PCR and both of which had been initially excluded by the TruSight analysis pipeline. A single presentation sample (BB85) displayed significantly lower coverage (total 0.89 million reads vs average 3.69 million reads per sample), and yielded an unusually high number of variants passing the standard filters (n=16, versus median of four high confidence mutations for the remainder of the cohort). For this case more stringent, manual filtering was applied to minimize false positive calls leading to exclusion of seven possible variants.

FLT3 internal tandem duplications (ITDs) are not detected by the standard alignment and variant calling pipelines for the TruSight Myeloid Panel. Paired reads were re-aligned using BWA-MEM (v0.7.12; https://basespace.illumina.com/apps/2084082/BWA-Aligner) [S8] and the revised BAM files were uploaded to two alternative variant callers: Pindel (v0.2.4) [S9] and ITDseek,[S10] both of which have been reported to call *FLT3*-ITD successfully from TruSight/TruSeq targeted sequencing data.[S7,S10] ITDseek outperformed Pindel, calling all *FLT3*-ITDs that were concurrently identified through standard PCR and gel electrophoresis on corresponding presentation gDNA samples, as previously described.[S11] *FLT3*-ITD burden was quantified by running PCR products on a Bioanalyzer 2100 High Sensitivity DNA chip (Agilent Technologies, CA, USA), as previously described.[S12]

Supplemental references

S1. Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X *et al.* The challenges of sequencing by synthesis. *Nat Biotechnol* 2009; **27**: 1013–1023.

S2. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res* 2015; **43**: D805–11.

S3. Greystoke BF, Huang X, Wilks DP, Atkinson S, Somervaille TCP. Very high frequencies of leukaemia-initiating cells in precursor T-acute lymphoblastic leukaemia may be obscured by cryopreservation. *Br J Haematol* 2013; **163**: 538–541.

S4. Wiseman DH, Small HF, Wilks DP, Waddell ID, Dennis MW, Ogilvie DJ, Somervaille TCP. Elevated plasma 2-hydroxyglutarate in acute myeloid leukaemia: association with the IDH1 SNP rs11554137 and severe renal impairment. *Br J Haematol* 2014; **166**: 145–148.

S5. Struys EA, Jansen EEW, Verhoeven NM, Jakobs C. Measurement of urinary D- and L-2-hydroxyglutarate enantiomers by stable-isotope-dilution liquid chromatography-tandem mass spectrometry after derivatization with diacetyl-L-tartaric anhydride. *Clin Chem* 2004; **50**: 1391–1395.

S6. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011; **29**: 24–26.

S7. Luthra R, Patel KP, Reddy NG, Haghshenas V, Routbort MJ, Harmon MA *et al.* Next-generation sequencing-based multigene mutational screening for acute myeloid leukemia using MiSeq: applicability for diagnostics and disease monitoring. *Haematologica* 2014; **99**: 465–473.

S8. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* 2010; **26**: 589–595.

S9. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 2009; **25**: 2865–2871.

S10. Au CH, Wa A, Ho DN, Chan TL, Ma ESK. Clinical evaluation of panel testing by next-generation sequencing (NGS) for gene mutations in myeloid neoplasms. *Diagn Pathol* 2016; **11**: 11.

S11. Shih L-Y, Huang C-F, Wu J-H, Lin T-L, Dunn P, Wang P-N *et al.* Internal tandem duplication of FLT3 in relapsed acute myeloid leukemia: a comparative analysis of bone marrow samples from 108 adult patients at diagnosis and relapse. *Blood* 2002; **100**: 2387–2392.

S12. Mills KI, Gilkes AF, Walsh V, Sweeney M, Gale R. Rapid and sensitive detection of internal tandem duplication and activating loop mutations of FLT3. *Br J Haematol* 2005; **130**: 203–208.