Clonal hematopoiesis with *DNMT3A* and *PPM1D* mutations impairs regeneration in autologous stem cell transplant recipients

Patrick Stelmach,[1,2,3*] Sarah Richter,[1*] Sandra Sauer,[1*] Margarete A. Fabre,[4,5,6,7*] Muxin Gu,[4,5,6] Christian Rohde,[1] Maike Janssen,[1] Nora Liebers,[1,8] Rumyana Proynova,[1] Niels Weinhold,[1] Marc S. Raab,[1] Hartmut Goldschmidt,[1] Birgit Besenbeck,[1] Petra Pavel,[9] Sascha Laier,[9] Andreas Trumpp,[2,3,10,11] Sascha Dietrich,[1,12#] George S. Vassiliou[4,5,6#] and Carsten Müller-Tidow[1,8,12#]

[1]Department of Medicine V, Heidelberg University Hospital, Heidelberg, Germany; [2]Division of Stem Cells and Cancer, German Cancer Research Center (DKFZ) and DKFZ-ZMBH Alliance, Heidelberg, Germany; [3]Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM GmbH), Heidelberg, Germany; [4]Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK; [5]Department of Hematology, University of Cambridge, Cambridge, UK; [6]Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK; [7]Center for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK; [8]National Center for Tumor Diseases (NCT), Heidelberg, Germany; [9]Stem Cell Laboratory, Institute of Clinical Transfusion Medicine and Cell Therapy Heidelberg GmbH, Heidelberg, Germany; [10]Faculty of Biosciences, Heidelberg University, Heidelberg, Germany; [11]German Cancer Consortium (DKTK), Heidelberg, Germany and [12]Molecular Medicine Partnership Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

**PS, SR, SS and MAF contributed equally as first authors.*
#SD, GSV and CM-T contributed equally as senior authors.

Correspondence: C. Müller-Tidow
carsten.mueller-tidow@med.uni-heidelberg.de
G. S. Vassiliou
gsv20@cam.ac.uk

Received: February 21, 2023
Accepted: June 19, 2023.
Early view: June 29, 2023.

https://doi.org/10.3324/haematol.2023.282992

SUPPLEMENTARY INFORMATION

Gene	Criteria for definition as a driver mutation	Transcript
ASXL1	Frameshift/nonsense/splice-site in exon 11-12	NM_015338
BCOR	Frameshift/nonsense/splice-site	NM_001123385
BCORL1	Frameshift/nonsense/splice-site	NM_021946
BRAF	Missense in aa range p.590-615; Missense at G469	NM_004333
BRCC3	Frameshift/nonsense/splice-site	NM_024332
CALR	Frameshift in exon 9	NM_004343
CBL	Missense in Linker/RING finger domains (p.345-434)	NM_005188
CEBPA	Frameshift/nonsense/splice-site	NM_004364
CREBBP	Frameshift/nonsense/splice-site	NM_004380
CSF1R	Missense at L301 / Y969	NM_005211
CSF3R	T615A, T618I, truncating c.741-791	NM_000760
CTCF	Frameshift/nonsense/splice-site, R377C, R377H, P378A, P378L	NM_006565
CUX1	Frameshift/nonsense/splice-site	NM_181552
DNMT3A	Frameshift/nonsense/splice-site; Missense in PWWP (p.292-350) / ADD (p.482-614) / MTase (p.634-912) domains	NM_022552
ETV6	Frameshift/nonsense/splice-site	NM_001987
EZH2	Frameshift/nonsense/splice-site; Missense in SET domain (p.617-732)	NM_001203247
FLT3	V579A, V592A, V592I, F594L, FY590-591GD, D835Y, D835H, D835E, del835	NM_004119
GATA2	Frameshift/nonsense/splice-site, R293Q, N317H, A318T, A318V, A318G, G320D, L321P, L321F, L321V, Q328P, R330Q, R361L, L359V, A372T, R384G, R384K	NM_001145661
GNAS	Missense at R201 (844)	NM_016592
GNB1	Missense at K57 / I80	NM_002074
IDH1	Missense at R132	NM_005896
IDH2	Missense at R140 / R172	NM_002168
JAK2	V617F; Missense/indel in aa range p.536-547	NM_004972
JAK3	M511T, M511I, A572V, A572T, A573V, R657Q, V715I, V715A	NM_000215
KDM6A	Frameshift/nonsense/splice-site	NM_021140
KIT	ins503, V559A, V559D, V559G, V559I, V560D, V560A, V560G, V560E, del560, E561K, del579, P627L, P627T, R634W, K642E, K642Q, V654A, V654E, H697Y, H697D, E761D, K807R, D816H, D816Y, D816F, D816I, D816V, D816H, del551-559	NM_000222
KMT2A	Frameshift/nonsense/splice-site	NM_005933
KRAS	Missense at G12 / G13 / Q61 / A146	NM_033360
MPL	S505G, S505N, S505C, L510P, del513, W515A, W515R, W515K, W515S, W515L, A519T, A519V, Y591D, W515-518KT	NM_005373
MYD88	L265P	NM_002468
NOTCH1	Frameshift/nonsense/splice-site/missense in exon 26-34	NM_017617
NPM1	Frameshift in exon 12	NM_002520
NRAS	Missense at G12 / G13 / Q61	NM_002524
PHF6	Frameshift/nonsense/splice-site	NM_001015877
PIGA	Frameshift/nonsense/splice-site	NM_002641
PPM1D	Frameshift/nonsense/splice-site in exon 5/6	NM_003620
PRPF40B	Frameshift/nonsense/splice-site	NM_001031698
PTEN	Frameshift/nonsense/splice-site	NM_000314

PTPN11	Missense in aa range p.58-76 and p.491-510	NM_002834
RAD21	Frameshift/nonsense/splice-site	NM_006265
RUNX1	Frameshift/nonsense/splice-site, S73F, H78Q, H78L, R80C, R80P, R80H, L85Q, P86L, P86H, S114L, D133Y, L134P, R135G, R135K, R135S, R139Q, R142S, A165V, R174Q, R177L, R177Q, A224T, D171G, D171V, D171N, R205W, R223C	NM_001001890
SETBP1	D868N, D868T, S869N, G870S, I871T, D880N, D880Q	NM_015559
SF1	Frameshift/nonsense/splice-site	NM_004630
SF3A1	Frameshift/nonsense/splice-site	NM_005877
SF3B1	Missense in terminal HEAT domains (p.529-1201)	NM_012433
SMC1A	Missense at R96 / R586	NM_006306
SMC3	Frameshift/nonsense/splice-site	NM_005445
SRSF2	Missense/deletion involving p.P95	NM_003016
STAG2	Frameshift/nonsense/splice-site	NM_006603
STAT3	Missense in SH2 domain (p.580-670)	NM_139276
TET2	Frameshift/nonsense/splice-site; Missense in conserved domains (p.1104-1481 and p.1843-2002)	NM_001127208
TP53	Frameshift/nonsense/splice-site; Missense in DNA-binding domain (p.95-288); Missense at P72 / R337	NM_001126112
U2AF1	Missense at S34 / R156 / Q157	NM_006758
U2AF2	Missense in RNA recognition motifs domains (p.149-231, p.259-337, p.381-462)	NM_007279
WT1	Frameshift/nonsense/splice-site	NM_024426
ZRSR2	Frameshift/nonsense/splice-site	NM_005089

Supplemental Table 1: Criteria for definition as a driver mutation

ASXL1	*ETV6*	*MPL*	*SF1*
BCOR	*EZH2*	*MYD88*	*SF3A1*
BCORL1	*FLT3*	*NOTCH1*	*SF3B1*
BRAF	*GATA2*	*NPM1*	*SMC1A*
BRCC3	*GNAS*	*NRAS*	*SMC3*
CALR	*GNB1*	*PHF6*	*SRSF2*
CBL	*IDH1*	*PIGA*	*STAG2*
CEBPA	*IDH2*	*PPM1D*	*STAT3*
CREBBP	*JAK2*	*PRPF40B*	*TET2*
CSF1R	*JAK3*	*PTEN*	*TP53*
CSF3R	*KDM6A*	*PTPN11*	*U2AF1*
CTCF	*KIT*	*RAD21*	*U2AF2*
CUX1	*KMT2A*	*RUNX1*	*WT1*
DNMT3A	*KRAS*	*SETBP1*	*ZRSR2*

Supplemental Table 2: Targeted enriched genes (56-gene panel)

PatientID	Chr	Pos	Ref	Alt	Gene	AA	VAF	VariantClass	Type	Transcript
1	2	25234307	G	A	DNMT3A	p.P904L	0,0234	missense	snv	ENST00000264709.7
2	2	25234373	C	T	DNMT3A	p.R882H	0,1198	missense	snv	ENST00000264709.7
3	2	25234373	C	T	DNMT3A	p.R882H	0,0382	missense	snv	ENST00000264709.7
4	2	25234373	C	T	DNMT3A	p.R882H	0,0118	missense	snv	ENST00000264709.7
5	2	25234374	G	A	DNMT3A	p.R882C	0,0118	missense	snv	ENST00000264709.7

6	2	25234374	G	A	DNMT3A	p.R882C	0,0847	missense	snv	ENST00000264709.7
7	2	25234374	G	A	DNMT3A	p.R882C	0,0101	missense	snv	ENST00000264709.7
8	2	25234415	A	G	DNMT3A	p.F868S	0,1689	missense	snv	ENST00000264709.7
9	2	25235726	A	G	DNMT3A	p.W860R	0,0111	missense	snv	ENST00000264709.7
10	2	25235774	T	TG	DNMT3A	K844fs*10	0,3008	frameshift	insertion	ENST00000264709.7
11	2	25235787	GGAGTT	G	DNMT3A	N838fs*14	0,0119	frameshift	deletion	ENST00000264709.7
12	2	25239130	C	T	DNMT3A	p.R803K	0,0359	missense	snv	ENST00000264709.7
13	2	25239137	T	C	DNMT3A	p.M801V	0,011	missense	snv	ENST00000264709.7
14	2	25239137	T	C	DNMT3A	p.M801V	0,011	missense	snv	ENST00000264709.7
15	2	25239199	A	G	DNMT3A	p.I780T	0,0157	missense	snv	ENST00000264709.7
16	2	25240306	A	T	DNMT3A	p.L773H	0,1153	missense	snv	ENST00000264709.7
12	2	25240307	G	C	DNMT3A	p.L773V	0,0642	missense	snv	ENST00000264709.7
17	2	25240312	C	T	DNMT3A	p.R771Q	0,0132	missense	snv	ENST00000264709.7
18	2	25240312	C	T	DNMT3A	p.R771Q	0,0977	missense	snv	ENST00000264709.7
19	2	25240323	C	A	DNMT3A	p.R767S	0,0227	missense	snv	ENST00000264709.7
20	2	25240324	CT	C	DNMT3A	R767fs*11	0,0267	frameshift	deletion	ENST00000264709.7
21	2	25240340	C	A	DNMT3A	p.G762C	0,0132	missense	snv	ENST00000264709.7
22	2	25240353	A	AT	DNMT3A	V758fs*6	0,0195	frameshift	insertion	ENST00000264709.7
23	2	25240391	C	A	DNMT3A	p.E745*	0,0159	nonsense	snv	ENST00000264709.7
24	2	25240417	C	G	DNMT3A	p.R736P	0,0695	missense	snv	ENST00000264709.7
25	2	25240417	C	T	DNMT3A	p.R736H	0,2684	missense	snv	ENST00000264709.7
26	2	25240420	T	G	DNMT3A	p.Y735S	0,0151	missense	snv	ENST00000264709.7
27	2	25240428	AAAG	A	DNMT3A	E732_Del:F	0,012	missense	deletion	ENST00000264709.7
28	2	25240439	G	A	DNMT3A	p.R729W	0,0226	missense	snv	ENST00000264709.7
29	2	25240439	G	C	DNMT3A	p.R729G	0,0138	missense	snv	ENST00000264709.7
30	2	25240650	C	A	DNMT3A	p.K721N	0,0212	missense	snv	ENST00000264709.7
31	2	25240708	T	A	DNMT3A	p.D702V	0,0109	missense	snv	ENST00000264709.7
32	2	25240710	G	T	DNMT3A	p.F701L	0,0317	missense	snv	ENST00000264709.7
33	2	25240714	G	A	DNMT3A	p.P700L	0,0523	missense	snv	ENST00000264709.7
34	2	25240729	A	G	DNMT3A	p.I695T	0,0124	missense	snv	ENST00000264709.7
35	2	25241602	A	C	DNMT3A	p.I681S	0,0478	missense	snv	ENST00000264709.7
36	2	25241608	C	A	DNMT3A	p.G679V	0,0176	missense	snv	ENST00000264709.7
12	2	25241621	C	T	DNMT3A	p.V675M	0,0205	missense	snv	ENST00000264709.7
37	2	25241621	C	T	DNMT3A	p.V675M	0,0782	missense	snv	ENST00000264709.7
38	2	25241645	C	A	DNMT3A	p.E667*	0,0139	nonsense	snv	ENST00000264709.7
39	2	25241674	A	G	DNMT3A	p.V657A	0,0647	missense	snv	ENST00000264709.7
5	2	25241676	C	CA	DNMT3A	V657fs*10	0,028	frameshift	insertion	ENST00000264709.7
40	2	25241685	C	G	DNMT3A	p.L653F	0,0463	missense	snv	ENST00000264709.7
41	2	25241704	A	T	DNMT3A	p.L647H	0,079	missense	snv	ENST00000264709.7
42	2	25243898	CT	C	DNMT3A	L647fs*3	0,0131	frameshift	deletion	ENST00000264709.7
43	2	25243898	C	T	DNMT3A	p.G646R	0,0304	missense	snv	ENST00000264709.7
29	2	25243928	C	T	DNMT3A	p.V636M	0,0124	missense	snv	ENST00000264709.7
44	2	25243928	C	T	DNMT3A	p.V636M	0,0194	missense	snv	ENST00000264709.7

45	2	25243931	G	A	DNMT3A	p.R635W	0,0134	missense	snv	ENST00000264709.7
46	2	25243931	G	A	DNMT3A	p.R635W	0,0132	missense	snv	ENST00000264709.7
47	2	25243931	G	A	DNMT3A	p.R635W	0,0192	missense	snv	ENST00000264709.7
48	2	25244248	G	C	DNMT3A	p.C586W	0,0108	missense	snv	ENST00000264709.7
49	2	25244264	CA	C	DNMT3A	W581fs*69	0,0422	frameshift	deletion	ENST00000264709.7
50	2	25244274	CCTTAATGGCTGCCTGGGCAG	C	DNMT3A	A571fs*33	0,076	frameshift	deletion	ENST00000264709.7
51	2	25244277	TA	T	DNMT3A	K577fs*73	0,0188	frameshift	deletion	ENST00000264709.7
52	2	25244321	C	T	DNMT3A	p.C562Y	0,0197	missense	snv	ENST00000264709.7
40	2	25244331	AAAAGCACCTGG	A	DNMT3A	R556fs*92	0,0974	frameshift	deletion	ENST00000264709.7
30	2	25244564	A	G	DNMT3A	p.M548T	0,0146	missense	snv	ENST00000264709.7
53	2	25244623	G	T	DNMT3A	p.Y528*	0,0201	nonsense	snv	ENST00000264709.7
35	2	25244643	GA	G	DNMT3A	L522fs*128	0,0133	frameshift	deletion	ENST00000264709.7
22	2	25244647	G	T	DNMT3A	p.C520*	0,0162	nonsense	snv	ENST00000264709.7
45	2	25245265	G	C	DNMT3A	p.C514W	0,016	missense	snv	ENST00000264709.7
54	2	25245288	G	T	DNMT3A	p.P507T	0,0183	missense	snv	ENST00000264709.7
55	2	25245299	G	GT	DNMT3A	T503fs*42	0,0101	frameshift	insertion	ENST00000264709.7
56	2	25246620	C	A	DNMT3A	p.E427*	0,0206	nonsense	snv	ENST00000264709.7
57	2	25247062	C	A	DNMT3A	p.E371*	0,018	nonsense	snv	ENST00000264709.7
58	2	25247611	C	T	DNMT3A	p.G332R	0,0109	missense	snv	ENST00000264709.7
20	2	25247616	C	T	DNMT3A	p.W330*	0,0322	nonsense	snv	ENST00000264709.7
45	2	25247628	C	T	DNMT3A	p.R326H	0,0554	missense	snv	ENST00000264709.7
59	2	25247628	C	T	DNMT3A	p.R326H	0,0108	missense	snv	ENST00000264709.7
60	2	25247634	C	A	DNMT3A	p.G324V	0,0303	missense	snv	ENST00000264709.7
61	2	25247664	C	A	DNMT3A	p.W314L	0,0171	missense	snv	ENST00000264709.7
62	2	25247674	C	A	DNMT3A	p.V311L	0,0311	missense	snv	ENST00000264709.7
62	2	25247674	CA	C	DNMT3A	I310fs*5	0,131	frameshift	deletion	ENST00000264709.7
20	2	25247680	G	A	DNMT3A	p.R309C	0,1523	missense	snv	ENST00000264709.7
63	2	25247715	C	G	DNMT3A	p.W297S	0,0145	missense	snv	ENST00000264709.7
64	2	25248099	CG	C	DNMT3A	V265fs*50	0,0373	frameshift	deletion	ENST00000264709.7
65	2	25248157	AG	A	DNMT3A	P245fs*70	0,0132	frameshift	deletion	ENST00000264709.7
56	2	197398046	G	T	SF3B1	p.H1069N	0,0241	missense	snv	ENST00000335508.10
66	2	197401774	G	T	SF3B1	p.P780T	0,0198	missense	snv	ENST00000335508.10
67	2	197401871	C	A	SF3B1	p.L747F	0,016	missense	snv	ENST00000335508.10
27	2	197401989	C	T	SF3B1	p.G740E	0,0296	missense	snv	ENST00000335508.10
68	2	197403659	G	A	SF3B1	p.R549C	0,014	missense	snv	ENST00000335508.10
69	2	197403702	C	A	SF3B1	p.Q534H	0,0161	missense	snv	ENST00000335508.10
70	3	128486959	G	GT	GATA2	H25fs*159	0,0389	frameshift	insertion	ENST00000341105.6
71	4	105234763	TC	T	TET2	N275fs*17	0,0141	frameshift	deletion	ENST00000540549.5
56	4	105234990	G	T	TET2	p.E350*	0,0151	nonsense	snv	ENST00000540549.5
56	4	105234993	G	T	TET2	p.E351*	0,0184	nonsense	snv	ENST00000540549.5
72	4	105235572	C	T	TET2	p.R544*	0,0925	nonsense	snv	ENST00000540549.5
73	4	105235900	TC	T	TET2	Q654fs*45	0,1727	frameshift	deletion	ENST00000540549.5
8	4	105236082	TC	T	TET2	S714fs	0,0298	frameshift	deletion	ENST00000540549.5

74	4	105236246	TC	T	TET2	Q769fs*43	0,0202	frameshift	deletion	ENST00000540549.5
75	4	105236492	TC	T	TET2	P851fs*21	0,0151	frameshift	deletion	ENST00000540549.5
76	4	105236634	G	T	TET2	p.G898*	0,024	nonsense	snv	ENST00000540549.5
70	4	105236739	C	T	TET2	p.Q933*	0,0469	nonsense	snv	ENST00000540549.5
8	4	105236859	T	TG	TET2	C973fs*2	0,1452	frameshift	insertion	ENST00000540549.5
16	4	105237042	C	T	TET2	p.Q1034*	0,0215	nonsense	snv	ENST00000540549.5
77	4	105237143	ACAAACCACTGCTG	A	TET2	T1069fs*8	0,02	frameshift	deletion	ENST00000540549.5
78	4	105237346	G	A	TET2	p.C1135Y	0,0111	missense	snv	ENST00000540549.5
66	4	105242907	G	T	TET2	p.G1192*	0,0244	nonsense	snv	ENST00000540549.5
79	4	105242911	G	A	TET2	p.C1193Y	0,0147	missense	snv	ENST00000540549.5
80	4	105243621	C	T	TET2	p.R1216*	0,0105	nonsense	snv	ENST00000540549.5
74	4	105243637	G	A	TET2	p.C1221Y	0,0168	missense	snv	ENST00000540549.5
70	4	105243651	AT	A	TET2	I1226fs*1	0,0113	frameshift	deletion	ENST00000540549.5
81	4	105243757	G	A	TET2	p.R1261H	0,0601	missense	snv	ENST00000540549.5
46	4	105259626	T	TG	TET2	C1271fs*28	0,0455	frameshift	insertion	ENST00000540549.5
36	4	105259628	C	A	TET2	p.C1271*	0,2947	nonsense	snv	ENST00000540549.5
82	4	105259672	C	A	TET2	p.S1286Y	0,0225	missense	snv	ENST00000540549.5
83	4	105259678	G	A	TET2	p.G1288D	0,1028	missense	snv	ENST00000540549.5
84	4	105259769	G	T	TET2	p.E1318D	0,0239	missense	snv	ENST00000540549.5
85	4	105261848	G	T	TET2	p.Q1348H	0,023	missense	snv	ENST00000540549.5
86	4	105269634	G	T	TET2	p.E1357*	0,0202	nonsense	snv	ENST00000540549.5
87	4	105269655	G	T	TET2	p.E1364*	0,0173	nonsense	snv	ENST00000540549.5
66	4	105272603	G	T	TET2	p.G1408*	0,0169	nonsense	snv	ENST00000540549.5
76	4	105272716	G	T	TET2	p.Q1445H	0,0191	missense	snv	ENST00000540549.5
88	4	105272735	C	T	TET2	p.R1452*	0,0371	nonsense	snv	ENST00000540549.5
89	4	105272735	C	T	TET2	p.R1452*	0,0739	nonsense	snv	ENST00000540549.5
64	4	105272775	G	A	TET2	p.R1465Q	0,0145	missense	snv	ENST00000540549.5
90	4	105272862	C	G	TET2	p.S1494*	0,0107	nonsense	snv	ENST00000540549.5
21	4	105272882	C	T	TET2	p.Q1501*	0,0472	nonsense	snv	ENST00000540549.5
57	4	105275099	C	CGTAA	TET2	L1531fs	0,0181	frameshift	insertion	ENST00000540549.5
91	4	105275104	C	T	TET2	p.Q1532*	0,044	nonsense	snv	ENST00000540549.5
92	4	105275267	C	A	TET2	p.S1586*	0,0151	nonsense	snv	ENST00000540549.5
93	4	105275599	G	T	TET2	p.G1697*	0,0177	nonsense	snv	ENST00000540549.5
42	4	105276092	G	T	TET2	p.G1861V	0,0554	missense	snv	ENST00000540549.5
94	4	105276128	T	C	TET2	p.I1873T	0,0297	missense	snv	ENST00000540549.5
95	4	105276139	A	G	TET2	p.K1877E	0,0141	missense	snv	ENST00000540549.5
27	4	105276196	AG	A	TET2	I1897fs*10	0,0167	frameshift	deletion	ENST00000540549.5
96	4	105276197	G	C	TET2	p.R1896T	0,0271	missense	snv	ENST00000540549.5
67	4	105276217	C	A	TET2	p.Q1903K	0,0176	missense	snv	ENST00000540549.5
21	4	105276217	C	T	TET2	p.Q1903*	0,0219	nonsense	snv	ENST00000540549.5
97	4	105276221	A	G	TET2	p.H1904R	0,0141	missense	snv	ENST00000540549.5
98	4	105276256	C	A	TET2	p.L1916I	0,0163	missense	snv	ENST00000540549.5
99	4	105276411	C	A	TET2	p.F1967L	0,0184	missense	snv	ENST00000540549.5

69	7	102097435	G	T	CUX1	p.E125*	0,0209	nonsense	snv	ENST00000360264.7
54	7	102202074	C	A	CUX1	p.S937*	0,0113	nonsense	snv	ENST00000360264.7
100	7	102227429	G	T	CUX1	p.E1076*	0,0198	nonsense	snv	ENST00000360264.7
54	7	102227645	G	T	CUX1	p.E1148*	0,0194	nonsense	snv	ENST00000360264.7
66	7	102248443	G	T	CUX1	p.E1318*	0,0176	nonsense	snv	ENST00000360264.7
101	7	102248644	G	T	CUX1	p.E1385*	0,0355	nonsense	snv	ENST00000360264.7
102	7	140753358	C	A	BRAF	p.G593C	0,0257	missense	snv	ENST00000646891.1
103	7	148809096	C	A	EZH2	p.G724C	0,0174	missense	snv	ENST00000320356.6
104	7	148809110	C	A	EZH2	p.R719I	0,0166	missense	snv	ENST00000320356.6
105	7	148814941	C	A	EZH2	p.E549*	0,0189	nonsense	snv	ENST00000320356.6
106	7	148816757	C	A	EZH2	p.E478*	0,0174	nonsense	snv	ENST00000320356.6
67	8	116852600	C	A	RAD21	p.E424*	0,015	nonsense	snv	ENST00000297338.6
54	8	116854307	C	A	RAD21	p.G367*	0,0184	nonsense	snv	ENST00000297338.6
3	8	116861920	C	A	RAD21	p.E99*	0,0203	nonsense	snv	ENST00000297338.6
107	9	5070033	G	T	JAK2	p.R541I	0,0176	missense	snv	ENST00000381652.3
36	9	5073770	G	T	JAK2	p.V617F	0,1757	missense	snv	ENST00000381652.3
108	9	5073770	G	T	JAK2	p.V617F	0,0145	missense	snv	ENST00000381652.3
109	9	5073770	G	T	JAK2	p.V617F	0,0144	missense	snv	ENST00000381652.3
105	9	136502018	C	A	NOTCH1	p.E1819*	0,013	nonsense	snv	ENST00000277541.7
109	10	87957958	TAC	T	PTEN	L247fs*4	0,0108	frameshift	deletion	ENST00000371953.7
110	10	87960963	G	T	PTEN	p.E291*	0,0201	nonsense	snv	ENST00000371953.7
111	10	110577457	G	T	SMC3	p.E79*	0,0193	nonsense	snv	ENST00000361804.4
112	10	110577854	C	A	SMC3	p.S97*	0,0207	nonsense	snv	ENST00000361804.4
113	10	110582640	G	T	SMC3	p.E268*	0,0199	nonsense	snv	ENST00000361804.4
114	10	110583898	G	T	SMC3	p.E343*	0,0233	nonsense	snv	ENST00000361804.4
47	10	110601077	C	A	SMC3	p.S864*	0,0181	nonsense	snv	ENST00000361804.4
115	10	110601801	G	T	SMC3	p.E937*	0,0122	nonsense	snv	ENST00000361804.4
47	10	110602546	G	T	SMC3	p.G1060*	0,0214	nonsense	snv	ENST00000361804.4
66	10	110602849	G	T	SMC3	p.E1108*	0,018	nonsense	snv	ENST00000361804.4
111	11	118471766	G	T	KMT2A	p.E203*	0,0196	nonsense	snv	ENST00000534358.5
116	11	119278265	C	A	CBL	p.L399I	0,0174	missense	snv	ENST00000264033.5
117	11	119278541	G	A	CBL	p.R420Q	0,0155	missense	snv	ENST00000264033.5
118	11	119278555	G	T	CBL	p.G425C	0,0207	missense	snv	ENST00000264033.5
39	12	11839289	C	T	ETV6	p.R105*	0,068	nonsense	snv	ENST00000396373.8
119	12	11869840	G	T	ETV6	p.E294*	0,0163	nonsense	snv	ENST00000396373.8
75	12	25225628	C	T	KRAS	p.A146T	0,0154	missense	snv	ENST00000256078.8
120	12	25227342	T	C	KRAS	p.Q61R	0,0117	missense	snv	ENST00000256078.8
47	12	112489092	C	A	PTPN11	p.Q510K	0,0154	missense	snv	ENST00000635625.1
105	16	3731778	C	A	CREBBP	p.E1630*	0,0157	nonsense	snv	ENST00000262367.9
114	16	3739612	C	A	CREBBP	p.E1416*	0,019	nonsense	snv	ENST00000262367.9
71	16	3757973	C	A	CREBBP	p.E1149*	0,0172	nonsense	snv	ENST00000262367.9
121	16	3780841	C	A	CREBBP	p.G572*	0,0185	nonsense	snv	ENST00000262367.9
71	16	3793610	G	T	CREBBP	p.S331*	0,0195	nonsense	snv	ENST00000262367.9

38	16	67610863	G	T	CTCF	p.E11*	0,0196	nonsense	snv	ENST00000646076.1
47	16	67637747	G	T	CTCF	p.E687*	0,016	nonsense	snv	ENST00000646076.1
122	17	7673767	C	T	TP53	p.E285K	0,0111	missense	snv	ENST00000269305.8
123	17	7673773	G	A	TP53	p.R283C	0,4934	missense	snv	ENST00000269305.8
120	17	7673802	C	T	TP53	p.R273H	0,015	missense	snv	ENST00000269305.8
96	17	7674221	G	A	TP53	p.R248W	0,0631	missense	snv	ENST00000269305.8
124	17	7674226	A	G	TP53	p.M246T	0,0207	missense	snv	ENST00000269305.8
80	17	7674924	C	T	TP53	p.V203M	0,0136	missense	snv	ENST00000269305.8
125	17	7675148	G	A	TP53	p.T155I	0,0114	missense	snv	ENST00000269305.8
126	17	42322387	G	T	STAT3	p.L666M	0,0176	missense	snv	ENST00000264657.9
54	17	42322409	C	A	STAT3	p.K658N	0,0168	missense	snv	ENST00000264657.9
44	17	42323014	C	A	STAT3	p.K626N	0,021	missense	snv	ENST00000264657.9
127	17	60656815	T	TC	PPM1D	C414fs*19	0,0102	frameshift	insertion	ENST00000305921.7
46	17	60663004	G	T	PPM1D	p.E424*	0,0214	nonsense	snv	ENST00000305921.7
128	17	60663073	G	T	PPM1D	p.E447*	0,02	nonsense	snv	ENST00000305921.7
129	17	60663077	AT	A	PPM1D	L450fs	0,0148	frameshift	deletion	ENST00000305921.7
130	17	60663137	C	A	PPM1D	p.S468*	0,0196	nonsense	snv	ENST00000305921.7
131	17	60663160	GA	G	PPM1D	N477fs*5	0,2313	frameshift	deletion	ENST00000305921.7
53	17	60663166	TG	T	PPM1D	C478fs*4	0,0198	frameshift	deletion	ENST00000305921.7
132	17	60663175	GC	G	PPM1D	L482fs	0,0287	frameshift	deletion	ENST00000305921.7
129	17	60663182	CT	C	PPM1D	L484fs	0,0107	frameshift	deletion	ENST00000305921.7
133	17	60663185	T	A	PPM1D	p.L484*	0,1667	nonsense	snv	ENST00000305921.7
134	17	60663203	T	A	PPM1D	p.L490*	0,0893	nonsense	snv	ENST00000305921.7
132	17	60663262	C	CA	PPM1D	N512fs*15	0,2261	frameshift	insertion	ENST00000305921.7
135	17	60663307	GA	G	PPM1D	I526fs*12	0,0176	frameshift	deletion	ENST00000305921.7
136	17	60663370	CT	C	PPM1D	L546fs*1	0,0114	frameshift	deletion	ENST00000305921.7
20	17	60663382	CA	C	PPM1D	H550fs*5	0,1401	frameshift	deletion	ENST00000305921.7
20	17	60663475	C	T	PPM1D	p.R581*	0,0125	nonsense	snv	ENST00000305921.7
137	19	33302048	C	A	CEBPA	p.G123*	0,0202	nonsense	snv	ENST00000498907.2
138	19	33302048	C	A	CEBPA	p.G123*	0,0206	nonsense	snv	ENST00000498907.2
130	20	32433356	G	GT	ASXL1	C387fs*22	0,0127	frameshift	insertion	ENST00000375687.9
100	20	32433732	C	T	ASXL1	p.Q512*	0,1033	nonsense	snv	ENST00000375687.9
117	20	32433787	C	CTCAA	ASXL1	A531fs*13	0,0173	frameshift	insertion	ENST00000375687.9
139	20	32434461	G	A	ASXL1	p.W583*	0,0744	nonsense	snv	ENST00000375687.9
3	20	32434599	CCACCACTGCCATAGAGAGGCGGC	C	ASXL1	E635fs*14	0,0284	frameshift	deletion	ENST00000375687.9
140	20	32434638	A	AG	ASXL1	G646fs*11	0,0116	frameshift	insertion	ENST00000375687.9
141	20	32435717	CCT	C	ASXL1	S1003fs	0,0117	frameshift	deletion	ENST00000375687.9
142	20	32436178	G	T	ASXL1	p.G1156*	0,0159	nonsense	snv	ENST00000375687.9
3	20	32436950	T	TA	ASXL1	P1414fs*9	0,0151	frameshift	insertion	ENST00000375687.9
143	20	58909366	G	A	GNAS	p.R844H	0,1264	missense	snv	ENST00000371100.8
16	21	34799403	C	A	RUNX1	p.G289*	0,0206	nonsense	snv	ENST00000300305.7
144	21	34834536	C	A	RUNX1	p.E227*	0,0374	nonsense	snv	ENST00000300305.7
145	X	15822973	G	T	ZRSR2	p.E394*	0,0227	nonsense	snv	ENST00000307771.7

114	X	40052153	C	A	BCOR	p.E1742*	0,0273	nonsense	snv	ENST00000378444.8
57	X	40071045	C	A	BCOR	p.G1056*	0,0205	nonsense	snv	ENST00000378444.8
146	X	40071684	C	A	BCOR	p.G1002*	0,0257	nonsense	snv	ENST00000378444.8
147	X	40074543	G	T	BCOR	p.S268*	0,0192	nonsense	snv	ENST00000378444.8
115	X	40077922	G	T	BCOR	p.S3*	0,0275	nonsense	snv	ENST00000378444.8
113	X	45020716	G	T	KDM6A	p.E184*	0,0285	nonsense	snv	ENST00000377967.8
21	X	45069819	G	T	KDM6A	p.G722*	0,0333	nonsense	snv	ENST00000377967.8
148	X	45069861	G	T	KDM6A	p.E736*	0,0242	nonsense	snv	ENST00000377967.8
149	X	45069891	G	T	KDM6A	p.G746*	0,0328	nonsense	snv	ENST00000377967.8
106	X	45079332	C	A	KDM6A	p.S1042*	0,0303	nonsense	snv	ENST00000377967.8
150	X	124031069	G	T	STAG2	p.G78*	0,0189	nonsense	snv	ENST00000218089.13
151	X	124042634	G	T	STAG2	p.E151*	0,025	nonsense	snv	ENST00000218089.13
152	X	124047419	C	T	STAG2	p.Q245*	0,0101	nonsense	snv	ENST00000218089.13
102	X	124061317	G	T	STAG2	p.E504*	0,0231	nonsense	snv	ENST00000218089.13
37	X	124077972	G	T	STAG2	p.G897*	0,0265	nonsense	snv	ENST00000218089.13
80	X	124095390	C	T	STAG2	p.R1242*	0,0174	nonsense	snv	ENST00000218089.13
122	X	124095390	C	T	STAG2	p.R1242*	0,0396	nonsense	snv	ENST00000218089.13
116	X	130013097	G	T	BCORL1	p.E109*	0,0206	nonsense	snv	ENST00000540052.5
54	X	134415067	G	T	PHF6	p.G261*	0,0392	nonsense	snv	ENST00000332070.7

Supplemental Table 3: Called variants (n=152 unsorted myeloma stem cell grafts)

	C0 (n = 305)	C1 (n = 20)	C2 (n = 62)	C3 (n = 47)	All (n = 434)
Myeloma subtype					
IgA only	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
IgA kappa	29 (9.5%)	1 (5.0%)	5 (8.1%)	2 (4.3%)	37 (8.5%)
IgA lambda	27 (8.9%)	2 (10.0%)	3 (4.8%)	4 (8.5%)	36 (8.3%)
IgG only	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)
IgG kappa	132 (43.3%)	13 (65.0%)	29 (46.8%)	23 (48.9%)	197 (45.4%)
IgG lambda	52 (17.0%)	2 (10.0%)	14 (22.6%)	7 (14.9%)	75 (17.3%)
IgM kappa	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
IgM lambda	0 (0%)	0 (0%)	1 (1.6%)	0 (0%)	1 (0.2%)
IgD kappa	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)
IgD lambda	6 (2.0%)	0 (0%)	2 (3.2%)	0 (0%)	8 (1.8%)

	C0 (n = 305)	C1 (n = 20)	C2 (n = 62)	C3 (n = 47)	All (n = 434)
Light chain myeloma kappa	32 (10.5%)	1 (5.0%)	4 (6.5%)	3 (6.4%)	40 (9.2%)
Light chain myeloma lambda	18 (5.9%)	1 (5.0%)	2 (3.2%)	7 (14.9%)	28 (6.5%)
Non-secretory	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NA	7 (2.3%)	0 (0%)	2 (3.2%)	1 (2.1%)	10 (2.3%)
ISS at diagnosis					
I	70 (23.0%)	2 (10.0%)	8 (12.9%)	9 (19.1%)	89 (20.5%)
II	27 (8.9%)	5 (25.0%)	5 (8.1%)	7 (14.9%)	44 (10.1%)
III	22 (7.2%)	0 (0%)	5 (8.1%)	1 (2.1%)	28 (6.5%)
NA	186 (61.0%)	13 (65.0%)	44 (71.0%)	30 (63.8%)	273 (62.9%)
Number of osteolysis at diagnosis					
None	54 (17.7%)	1 (5.0%)	15 (24.2%)	11 (23.4%)	81 (18.7%)
≥ 1	148 (48.5%)	15 (75.0%)	30 (48.4%)	23 (48.9%)	216 (49.8%)
≥ 3	68 (22.3%)	2 (10.0%)	8 (12.9%)	9 (19.1%)	87 (20.0%)
NA	35 (11.5%)	2 (10.0%)	9 (14.5%)	4 (8.5%)	50 (11.5%)
Induction chemotherapy					
VD	4 (1.3%)	0 (0%)	1 (1.6%)	0 (0%)	5 (1.2%)
VCD	80 (26.2%)	7 (35.0%)	22 (35.5%)	16 (34.0%)	125 (28.8%)
VTD	0 (0%)	0 (0%)	1 (1.6%)	0 (0%)	1 (0.2%)
VAD	63 (20.7%)	4 (20.0%)	10 (16.1%)	7 (14.9%)	84 (19.4%)
PAD	102 (33.4%)	3 (15.0%)	15 (24.2%)	14 (29.8%)	134 (30.9%)
TAD	11 (3.6%)	0 (0%)	2 (3.2%)	4 (8.5%)	17 (3.9%)
RD	6 (2.0%)	1 (5.0%)	1 (1.6%)	0 (0%)	8 (1.8%)

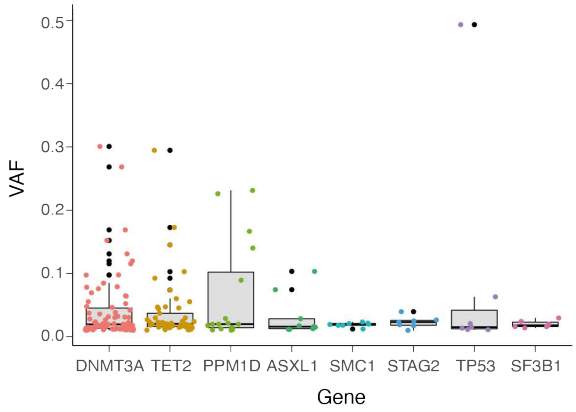
	C0 (n = 305)	C1 (n = 20)	C2 (n = 62)	C3 (n = 47)	All (n = 434)
VRD	20 (6.6%)	4 (20.0%)	6 (9.7%)	6 (12.8%)	36 (8.3%)
Other	15 (4.9%)	1 (5.0%)	3 (4.8%)	0 (0%)	19 (4.4%)
NA	4 (1.3%)	0 (0%)	1 (1.6%)	0 (0%)	5 (1.2%)
Remission before ASCT					
CR	100 (32.8%)	2 (10.0%)	13 (21.0%)	21 (44.7%)	136 (31.3%)
SD	58 (19.0%)	4 (20.0%)	15 (24.2%)	8 (17.0%)	85 (19.6%)
PD	21 (6.9%)	0 (0%)	5 (8.1%)	1 (2.1%)	27 (6.2%)
PR	126 (41.3%)	14 (70.0%)	29 (46.8%)	17 (36.2%)	186 (42.9%)
Remission after ASCT					
CR	71 (23.3%)	4 (20.0%)	15 (24.2%)	16 (34.0%)	106 (24.4%)
MR	20 (6.6%)	2 (10.0%)	3 (4.8%)	1 (2.1%)	26 (6.0%)
PD	3 (1.0%)	0 (0%)	1 (1.6%)	0 (0%)	4 (0.9%)
PR	128 (42.0%)	7 (35.0%)	25 (40.3%)	21 (44.7%)	181 (41.7%)
SD	10 (3.3%)	0 (0%)	4 (6.5%)	0 (0%)	14 (3.2%)
VGPR	70 (23.0%)	7 (35.0%)	13 (21.0%)	9 (19.1%)	99 (22.8%)
NA	3 (1.0%)	0 (0%)	1 (1.6%)	0 (0%)	4 (0.9%)
Tandem ASCT					
No	120 (39.3%)	9 (45.0%)	32 (51.6%)	21 (44.7%)	182 (41.9%)
Yes	185 (60.7%)	11 (55.0%)	30 (48.4%)	26 (55.3%)	252 (58.1%)
Maintenance therapy type					
Lenalidomide	61 (20.0%)	3 (15.0%)	14 (22.6%)	6 (12.8%)	84 (19.4%)
Thalidomide	70 (23.0%)	3 (15.0%)	15 (24.2%)	14 (29.8%)	102 (23.5%)

	C0 (n = 305)	C1 (n = 20)	C2 (n = 62)	C3 (n = 47)	All (n = 434)
Bortezomib	34 (11.1%)	0 (0%)	1 (1.6%)	4 (8.5%)	39 (9.0%)
Interferon	30 (9.8%)	2 (10.0%)	4 (6.5%)	3 (6.4%)	39 (9.0%)
Elotuzumab/lenalidomide	10 (3.3%)	3 (15.0%)	3 (4.8%)	3 (6.4%)	19 (4.4%)
Missing	100 (32.8%)	9 (45.0%)	25 (40.3%)	17 (36.2%)	151 (34.8%)

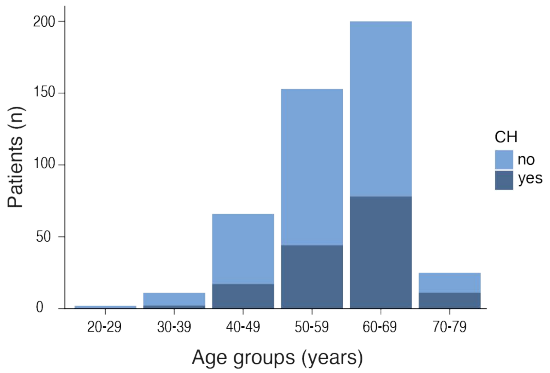
Supplemental Table 4: Additional patient characteristics per cluster

Abbreviations: MM = multiple myeloma. Induction chemotherapy: V/P = bortezomib, C = cyclophosphamide, D = dexamethasone, T = thalidomide, A = doxorubicin, R = lenalidomide. Response: CR = complete remission, VGPR = very good partial remission, PR = partial remission, MR = minimal response, SD = stable disease, PD = progressive disease.

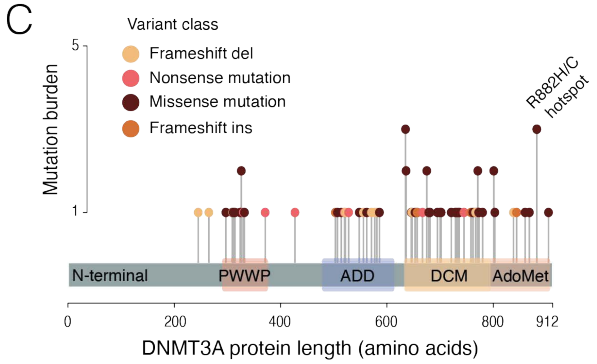
S1A


VAF
0.5
0.4
0.3
0.2
0.1
0.0
DNMT3A
TET2
PPM1D
ASXL1
SMC1
STAG2
TP53
SF3B1
Gene

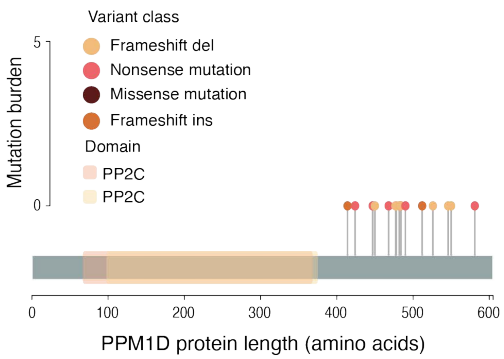
B


Patients (n)
200
150
100
50
0
20-29
30-39
40-49
50-59
60-69
70-79
Age groups (years)
CH
no
yes

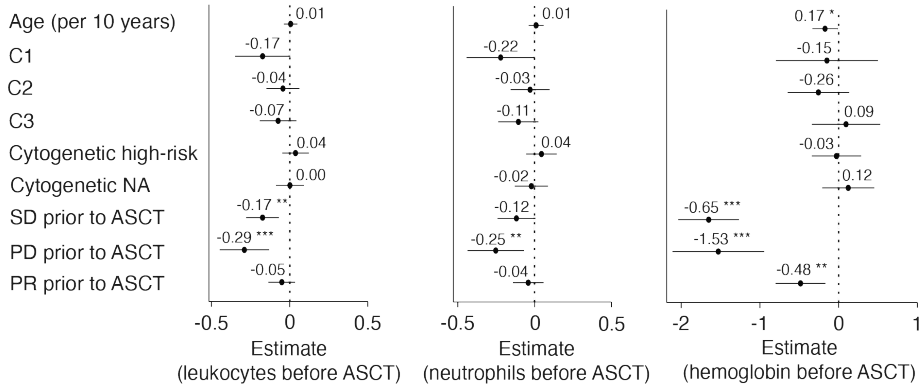
C


Variant class
Frameshift del
Nonsense mutation
Missense mutation
Frameshift ins
Mutation burden
5
1
R882H/C
hotspot
N-terminal
PWWP
ADD
DCM
AdoMet
0
200
400
600
800
912
DNMT3A protein length (amino acids)

D


Variant class
Frameshift del
Nonsense mutation
Missense mutation
Frameshift ins
Domain
PP2C
PP2C
Mutation burden
5
0
0
100
200
300
400
500
600
PPM1D protein length (amino acids)

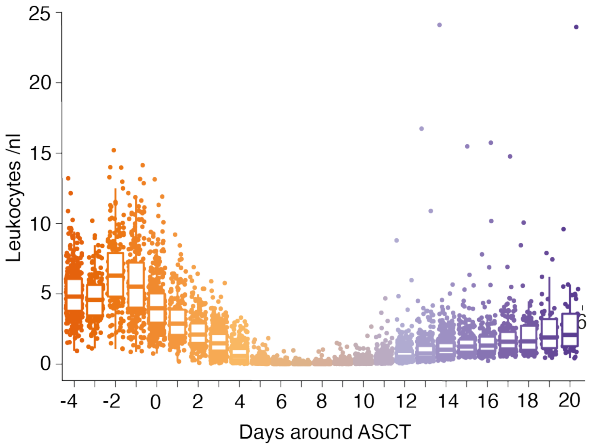
E


Age (per 10 years)
C1
C2
C3
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
0.01
-0.17
-0.04
-0.07
0.04
0.00
-0.17 **
-0.29 ***
-0.05
-0.5
0
0.5
Estimate
(leukocytes before ASCT)
0.01
-0.22
-0.03
-0.11
0.04
-0.02
-0.12
-0.25 **
-0.04
-0.5
0
0.5
Estimate
(neutrophils before ASCT)
0.17 *
-0.15
-0.26
0.09
-0.03
0.12
-0.65 ***
-1.53 ***
-0.48 **
-2
-1
0
1
Estimate
(hemoglobin before ASCT)

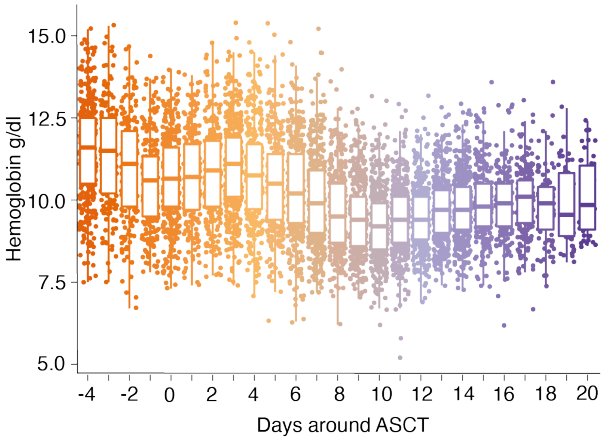
Supplemental Figure 1: Mutational landscape and blood cell counts before transplant

A) Variant allele frequency (VAF) distribution boxplot including the eight most frequently mutated genes.
B) Bar chart illustrating the age distribution and the prevalence of CH (stacked).
C) Lollipop plot illustrating the distribution of mutated amino acid positions along the DNMT3A protein. The distance to the protein sequence indicates the frequency (mutation burden) and the variant class is specified by color. The R882H/C hotspot is highlighted. The plot refers to transcript NM_022552 and shows the DNMT3A domain architecture.
D) Lollipop plot illustrating the distribution of mutated amino acid positions along the PPM1D protein. The plot refers to transcript NM_003620.
E) Forest plot visualizing the output of a linear regression model for leukocyte counts, neutrophil counts and hemoglobin levels before ASCT including the specified independent variables. The leukocyte counts were transformed by square root transformation. The plot illustrates the estimates/coefficients and their respective confidence intervals and statistical significance is indicated if the value is flagged with one or more stars. The reference for C1-3 is the patient group without CH mutations (C0).

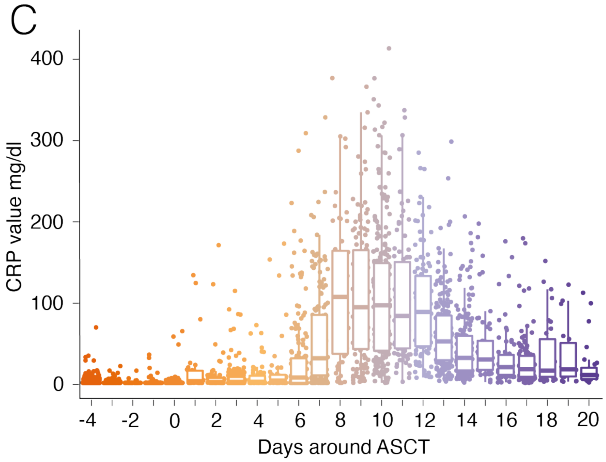
S2A


Leukocytes /nl
25
20
15
10
5
0
-4 -2 0 2 4 6 8 10 12 14 16 18 20
Days around ASCT

B

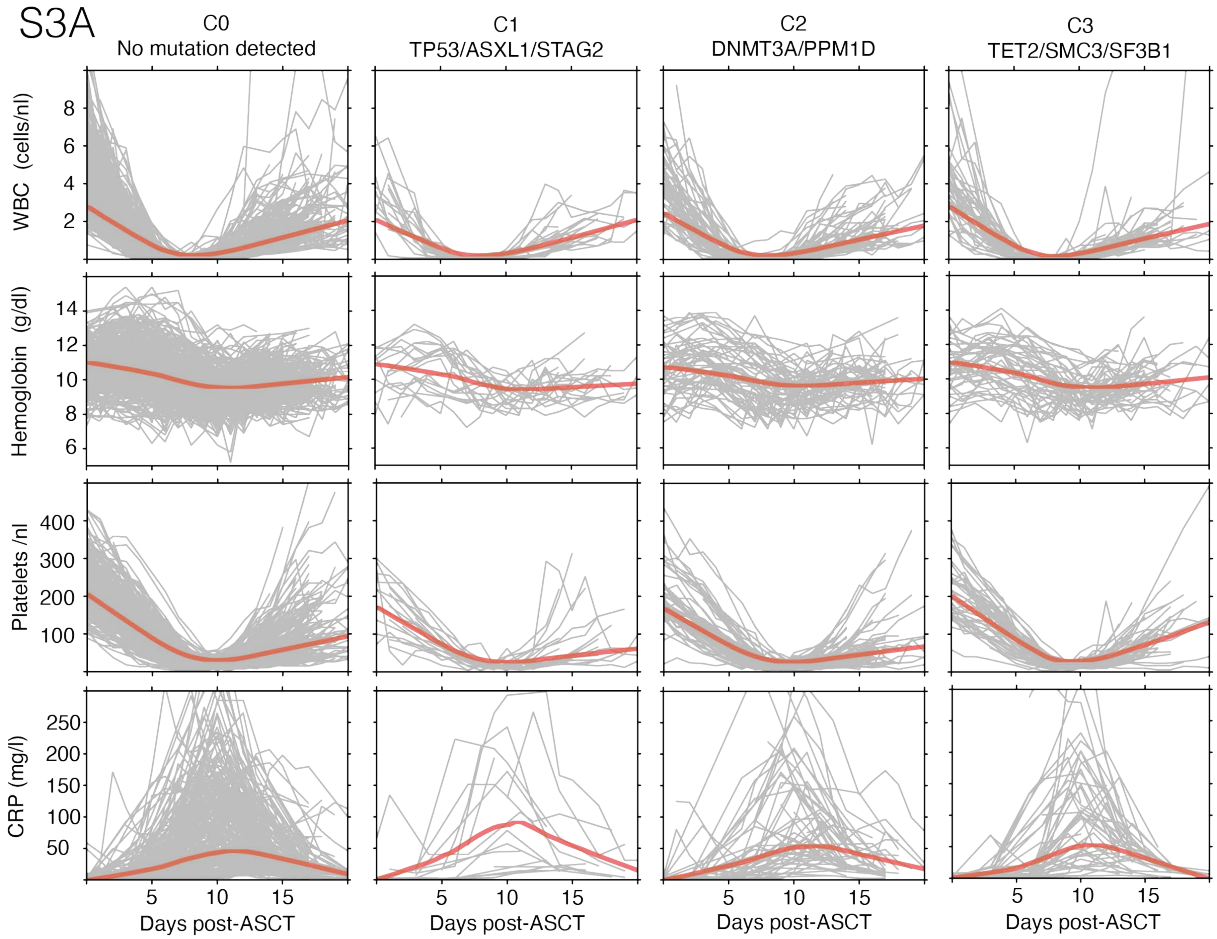

Hemoglobin g/dl
15.0
12.5
10.0
7.5
5.0
-4 -2 0 2 4 6 8 10 12 14 16 18 20
Days around ASCT

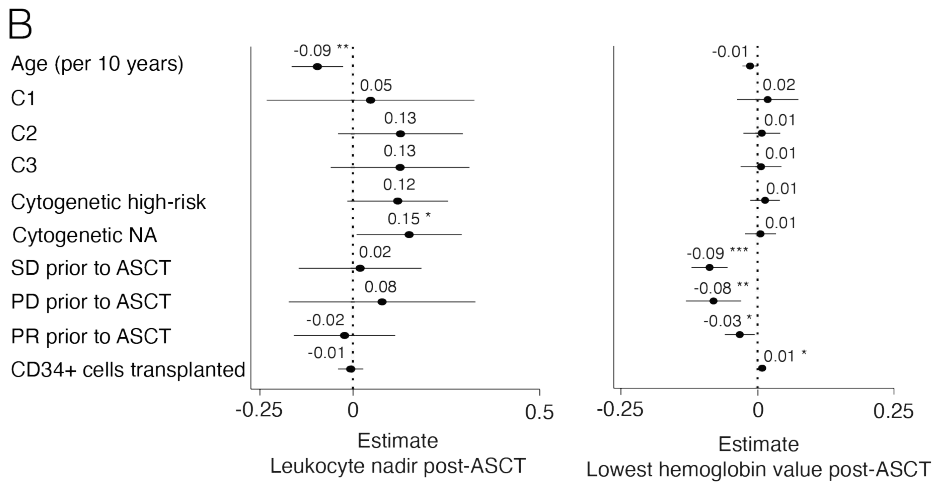
C


CRP value mg/dl
400
300
200
100
0
-4 -2 0 2 4 6 8 10 12 14 16 18 20
Days around ASCT

Supplemental Figure 2: High density data warehouse data

A) Density of data points for leukocyte counts in the myeloma cohort in 25 days around transplant (from day -4 to day +20) sourced from our data warehouse.
B) Density of data points for hemoglobin values in the myeloma cohort in 25 days around transplant (from day -4 to day +20) sourced from our data warehouse.
C) Density of data points for serum CRP values in the myeloma cohort in 25 days around transplant (from day -4 to day +20) sourced from our data warehouse.


S3A
C0
No mutation detected
C1
TP53/ASXL1/STAG2
C2
DNMT3A/PPM1D
C3
TET2/SMC3/SF3B1
WBC (cells/nl)
Hemoglobin (g/dl)
Platelets /nl
CRP (mg/l)
Days post-ASCT

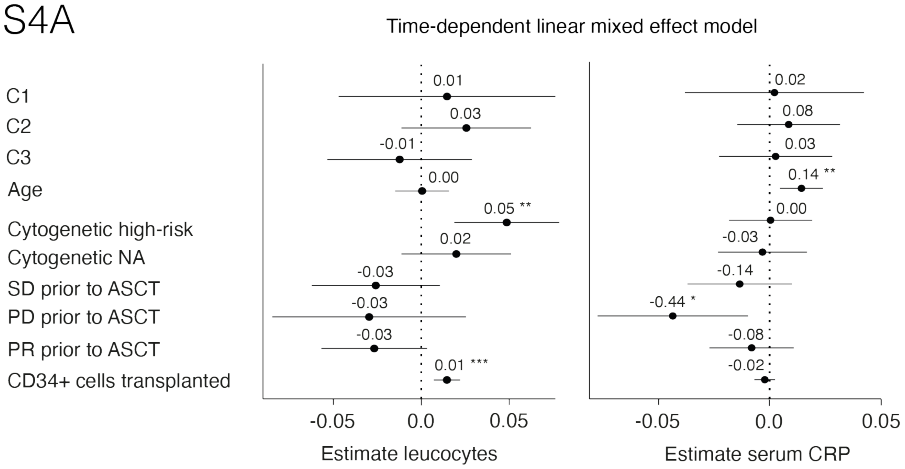

B
Age (per 10 years)
C1
C2
C3
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
-0.09 **
0.05
0.13
0.13
0.12
0.15 *
0.02
0.08
-0.02
-0.01
Estimate
Leukocyte nadir post-ASCT
-0.01
0.02
0.01
0.01
0.01
0.01
-0.09 ***
-0.08 **
-0.03 *
0.01 *
Estimate
Lowest hemoglobin value post-ASCT

Supplemental Figure 3: Trajectories of blood values after transplant by cluster sourced from our data warehouse

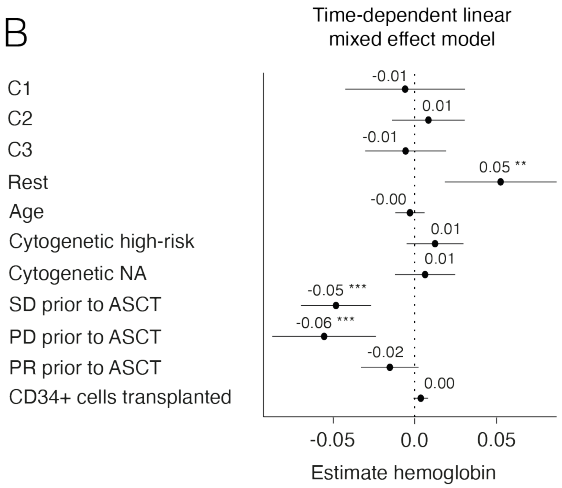
A) The trajectories (grey lines) visualize the effect of time on peripheral leukocyte and platelet counts and hemoglobin and CRP values in 20 days post-ASCT (bold red line = average) per cluster (C1-3) and in patients in whom no mutation was detected (C0).

B) Forest plot visualizing a linear regression model for the lowest leukocyte count (nadir) and the lowest hemoglobin value following high-dose chemotherapy and transplant including the specified independent variables. The reference for C1-3 patients are the patients without CH mutations (C0) and the reference for the shown remission states prior to ASCT is the patient group that achieved a complete remission (CR) after induction chemotherapy. Shown are the estimates/coefficients and their respective confidence intervals and statistical significance is indicated if the value is flagged with one or more stars.

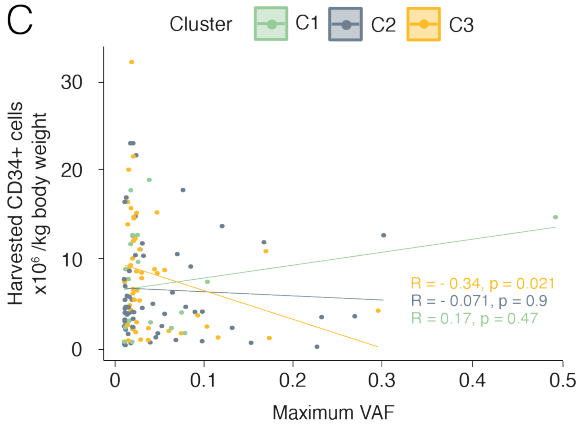
S4A


Time-dependent linear mixed effect model
C1
C2
C3
Age
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
0.01
0.03
-0.01
0.00
0.05 **
0.02
-0.03
-0.03
-0.03
0.01 ***
0.02
0.08
0.03
0.14 **
0.00
-0.03
-0.14
-0.44 *
-0.08
-0.02
-0.05
0.0
0.05
Estimate leucocytes
Estimate serum CRP

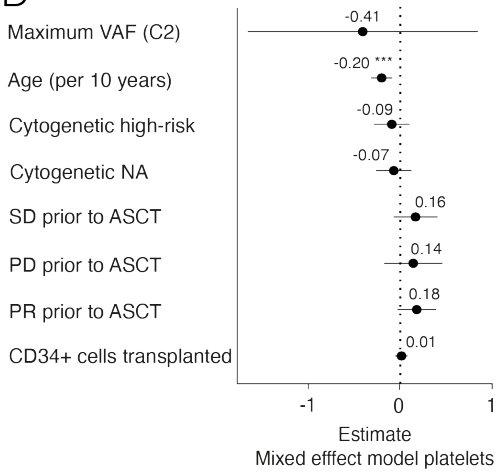
B


Time-dependent linear mixed effect model
C1
C2
C3
Rest
Age
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
-0.01
0.01
-0.01
0.05 **
-0.00
0.01
0.01
-0.05 ***
-0.06 ***
-0.02
0.00
-0.05
0.0
0.05
Estimate hemoglobin

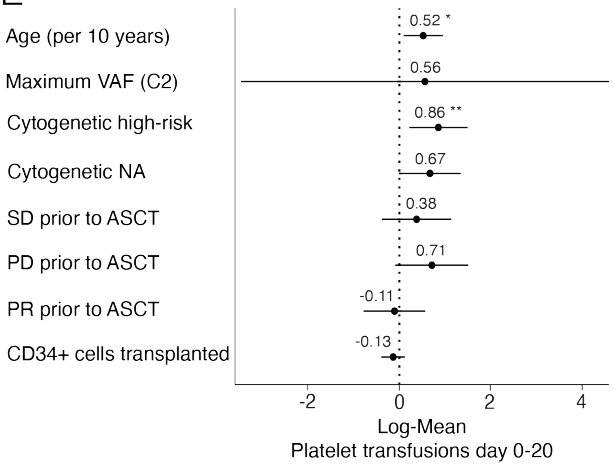
C


Cluster
C1
C2
C3
Harvested CD34+ cells
x10⁶ /kg body weight
30
20
10
0
0
0.1
0.2
0.3
0.4
0.5
Maximum VAF
R = - 0.34, p = 0.021
R = - 0.071, p = 0.9
R = 0.17, p = 0.47

D

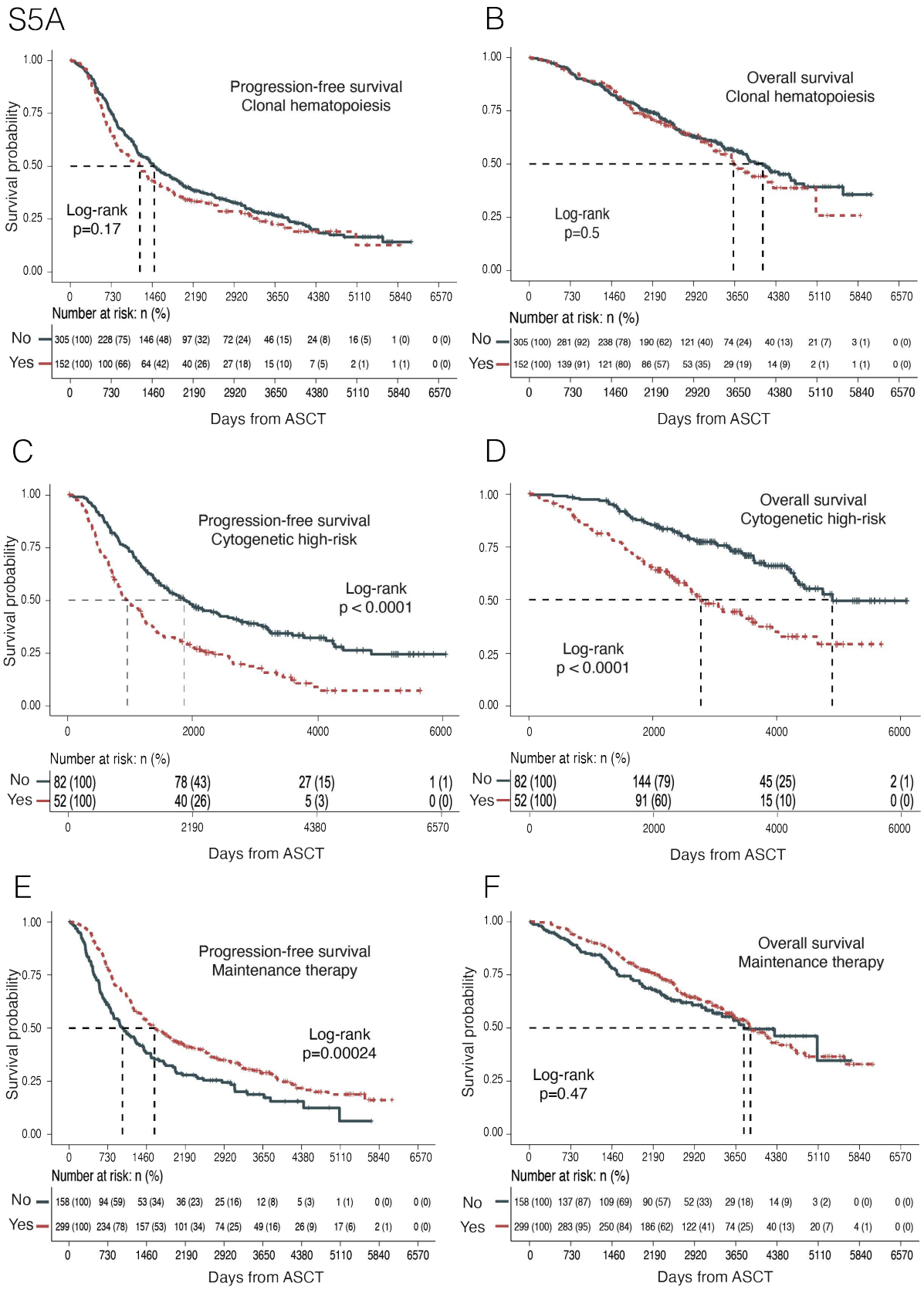

Maximum VAF (C2)
Age (per 10 years)
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
-0.41
-0.20 ***
-0.09
-0.07
0.16
0.14
0.18
0.01
-1
0
1
Estimate
Mixed efffect model platelets

E


Age (per 10 years)
Maximum VAF (C2)
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
0.52 *
0.56
0.86 **
0.67
0.38
0.71
-0.11
-0.13
-2
0
2
4
Log-Mean
Platelet transfusions day 0-20

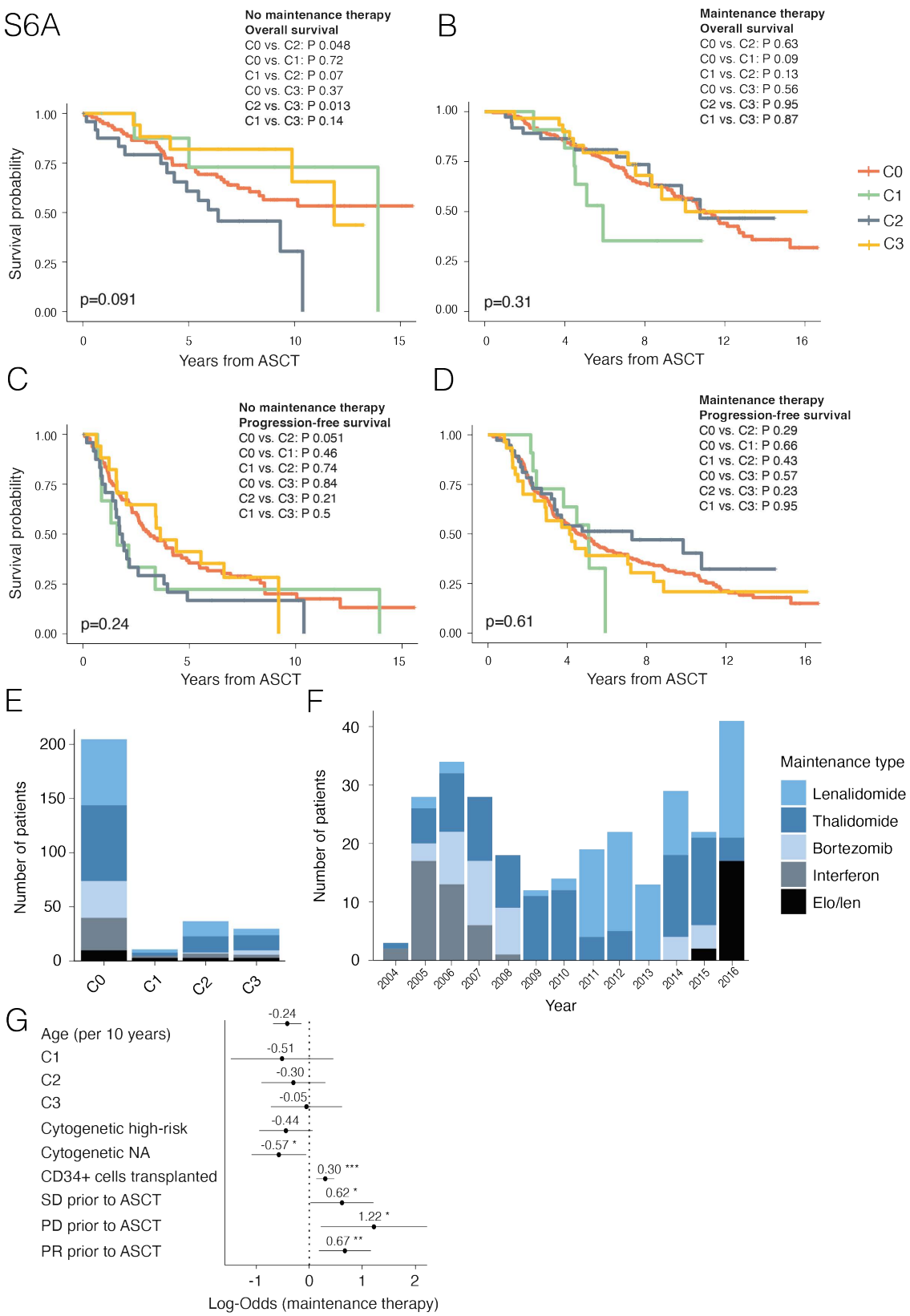
Supplemental Figure 4: Time-dependent linear mixed effect model for leukocyte counts, hemoglobin values and serum CRP values after transplant and VAF dependency of stem cell yields and lower platelet counts

A) Time-dependent linear mixed effect model for leukocyte counts and serum CRP values post-ASCT including the specified variables. The effect estimates are shown in the forest plot. Statistical significance is indicated if the value is flagged with one or more stars.
B) Time-dependent linear mixed effect model for hemoglobin values post-ASCT.
C) Scatter plot of harvested stem cells and maximum VAFs with regression lines per cluster. The Spearman correlation coefficient (R) and the respective p-value for the correlation per cluster is indicated.
D) Time-dependent linear mixed effect model for platelet counts post-ASCT including the specified independent variables. Maximum VAF mutation (C2) = mutation with the highest VAF within C2.
E) Forest plot visualizing a Poisson regression for the number of platelet transfusions within 20 days following ASCT including the specified independent variables. The platelet values have been log-transformed to obtain a normal distribution. The log-mean values and the respective confidence intervals are shown. Maximum VAF mutation (C2) = mutation with the highest VAF within C2.


S5A
Progression-free survival
Clonal hematopoiesis
Log-rank
p=0.17
Survival probability
Number at risk: n (%)
No 305 (100) 228 (75) 146 (48) 97 (32) 72 (24) 46 (15) 24 (8) 16 (5) 1 (0) 0 (0)
Yes 152 (100) 100 (66) 64 (42) 40 (26) 27 (18) 15 (10) 7 (5) 2 (1) 1 (1) 0 (0)
0 730 1460 2190 2920 3650 4380 5110 5840 6570
Days from ASCT
B
Overall survival
Clonal hematopoiesis
Log-rank
p=0.5
Number at risk: n (%)
No 305 (100) 281 (92) 238 (78) 190 (62) 121 (40) 74 (24) 40 (13) 21 (7) 3 (1) 0 (0)
Yes 152 (100) 139 (91) 121 (80) 86 (57) 53 (35) 29 (19) 14 (9) 2 (1) 1 (1) 0 (0)
Days from ASCT
C
Progression-free survival
Cytogenetic high-risk
Log-rank
p < 0.0001
Number at risk: n (%)
No 82 (100) 78 (43) 27 (15) 1 (1)
Yes 52 (100) 40 (26) 5 (3) 0 (0)
Days from ASCT
D
Overall survival
Cytogenetic high-risk
Log-rank
p < 0.0001
Number at risk: n (%)
No 82 (100) 144 (79) 45 (25) 2 (1)
Yes 52 (100) 91 (60) 15 (10) 0 (0)
Days from ASCT
E
Progression-free survival
Maintenance therapy
Log-rank
p=0.00024
Number at risk: n (%)
No 158 (100) 94 (59) 53 (34) 36 (23) 25 (16) 12 (8) 5 (3) 1 (1) 0 (0) 0 (0)
Yes 299 (100) 234 (78) 157 (53) 101 (34) 74 (25) 49 (16) 26 (9) 17 (6) 2 (1) 0 (0)
Days from ASCT
F
Overall survival
Maintenance therapy
Log-rank
p=0.47
Number at risk: n (%)
No 158 (100) 137 (87) 109 (69) 90 (57) 52 (33) 29 (18) 14 (9) 3 (2) 0 (0) 0 (0)
Yes 299 (100) 283 (95) 250 (84) 186 (62) 122 (41) 74 (25) 40 (13) 20 (7) 4 (1) 0 (0)
Days from ASCT

Supplemental Figure 5: Clinical outcomes (entire cohort) stratified by overall CH (all mutations), presence of MM cytogenetic high-risk lesions and maintenance therapy

A) PFS stratified by CH.
B) OS stratified by CH.
C) PFS stratified by MM cytogenetic risk.
D) OS stratified by MM cytogenetic risk.
E) PFS stratified by maintenance therapy.
F) OS stratified by maintenance therapy.

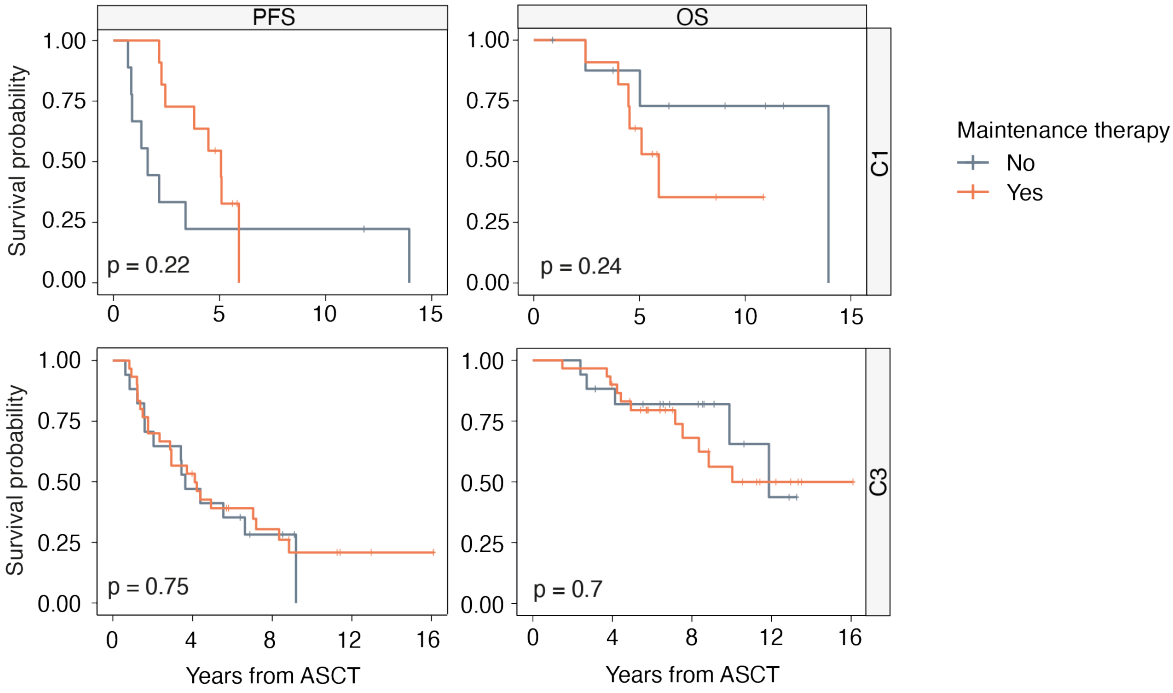

S6A
No maintenance therapy
Overall survival
C0 vs. C2: P 0.048
C0 vs. C1: P 0.72
C1 vs. C2: P 0.07
C0 vs. C3: P 0.37
C2 vs. C3: P 0.013
C1 vs. C3: P 0.14
Survival probability
p=0.091
Years from ASCT
B
Maintenance therapy
Overall survival
C0 vs. C2: P 0.63
C0 vs. C1: P 0.09
C1 vs. C2: P 0.13
C0 vs. C3: P 0.56
C2 vs. C3: P 0.95
C1 vs. C3: P 0.87
p=0.31
Years from ASCT
C0
C1
C2
C3
C
No maintenance therapy
Progression-free survival
C0 vs. C2: P 0.051
C0 vs. C1: P 0.46
C1 vs. C2: P 0.74
C0 vs. C3: P 0.84
C2 vs. C3: P 0.21
C1 vs. C3: P 0.5
Survival probability
p=0.24
Years from ASCT
D
Maintenance therapy
Progression-free survival
C0 vs. C2: P 0.29
C0 vs. C1: P 0.66
C1 vs. C2: P 0.43
C0 vs. C3: P 0.57
C2 vs. C3: P 0.23
C1 vs. C3: P 0.95
p=0.61
Years from ASCT
E
Number of patients
C0
C1
C2
C3
F
Number of patients
Year
Maintenance type
Lenalidomide
Thalidomide
Bortezomib
Interferon
Elo/len
G
Age (per 10 years)
C1
C2
C3
Cytogenetic high-risk
Cytogenetic NA
CD34+ cells transplanted
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
-0.24
-0.51
-0.30
-0.05
-0.44
-0.57 *
0.30 ***
0.62 *
1.22 *
0.67 **
Log-Odds (maintenance therapy)

Supplemental Figure 6: Clinical outcomes per cluster and maintenance therapy

A) Overall survival (OS) of patients not treated with maintenance therapy per cluster (C1-3) and for patients without CH (C0).
B) Overall survival (OS) of patients treated with maintenance therapy per cluster (C1-3) and for patients without CH (C0).
C) Progression-free survival (PFS) of patients not treated with maintenance therapy per cluster (C1-3) and for patients without CH (C0).
D) Progression-free survival (PFS) of patients treated with maintenance therapy per cluster (C1-3) and for patients without CH (C0).
E) Bar chart illustrating the maintenance therapy type per cluster.
F) Bar chart illustrating the distribution of years and chosen maintenance therapy types (stacked).
G) Forest plot visualizing a logistic regression for the treatment with maintenance therapy following ASCT including the specified independent variables. The reference for C1-3 is the patient group without CH mutations (C0) and the reference for the shown remission states prior to ASCT is the patient group that achieved a complete remission (CR) prior to ASCT. Shown are estimates/coefficients and their respective confidence intervals and statistical significance is indicated if the value is flagged with one or more stars.

Abbreviations: Elo/len = elotuzumab/lenalidomide.

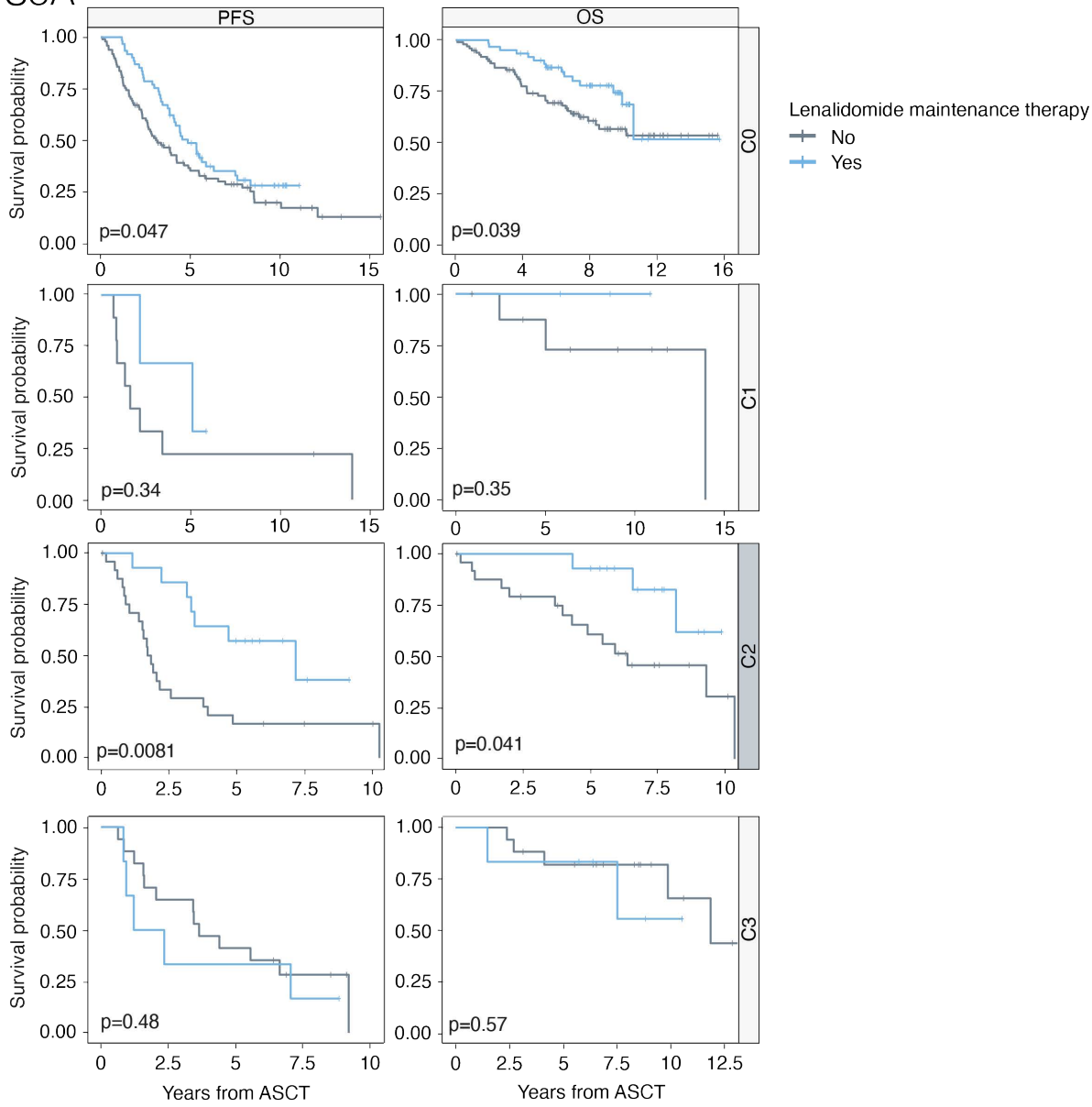
S7A


PFS
OS
C1
C3
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.22
p = 0.24
p = 0.75
p = 0.7
0
5
10
15
4
8
12
16
Years from ASCT
Maintenance therapy
No
Yes

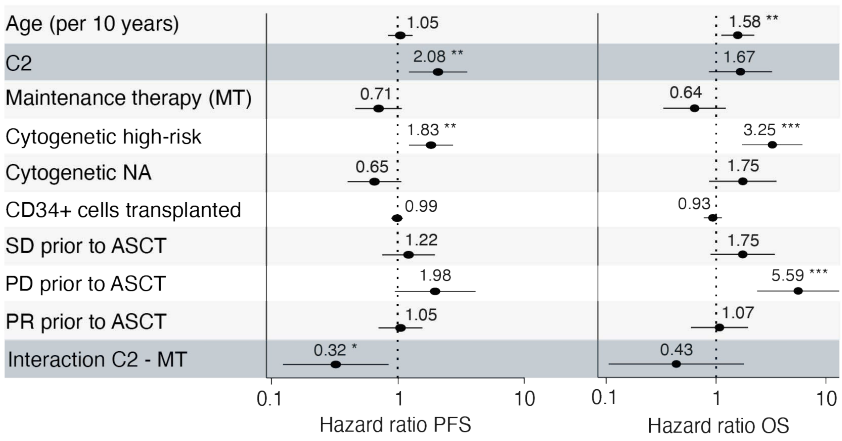
Supplemental Figure 7: PFS and OS stratified by maintenance therapy and cluster

A) PFS and OS from the day of transplant for C1 and C3 stratified by maintenance therapy (grey = no maintenance therapy, blue = lenalidomide maintenance therapy). The respective p-value calculated by log-rank test is indicated.

S8A



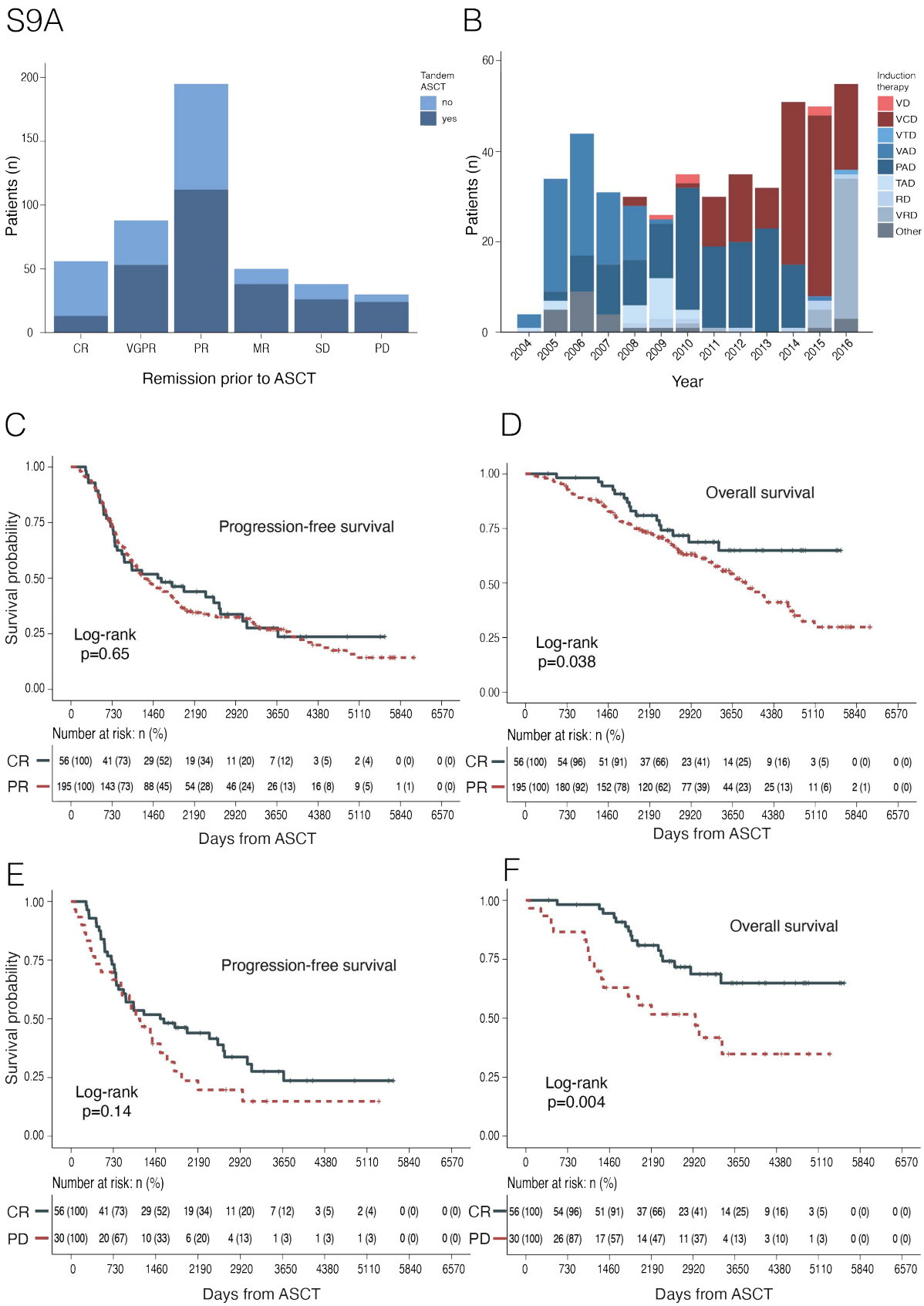
B Multivariate cox regression model
including C0- and C2 patients treated with lenalidomide maintenance therapy (MT)



Supplemental Figure 8: Subgroup of patients treated with lenalidomide as maintenance therapy type

A) Progression-free survival (PFS) and overall survival (OS) from the day of transplant per cluster (C1-3) stratified by maintenance therapy (grey = no maintenance therapy, blue = maintenance therapy) for C0- and C2 patients. C0 = patients in whom no CH mutation was detected. The respective p-value calculated by log-rank test is indicated.

B) Multivariate cox regression model (calculated from day +90 to overcome immortal time bias) including C2 patients and patients without CH mutations (C0) treated with lenalidomide as maintenance therapy type (n = 75). C0 patients are the reference and therefore not shown. The model shows a significant interaction between C2 patients and lenalidomide maintenance therapy demonstrating that C2 patients benefit particularly strong regarding PFS. The presence of cytogenetic high-risk lesions was included as a covariate, thereby correcting for cytogenetic high-risk status.

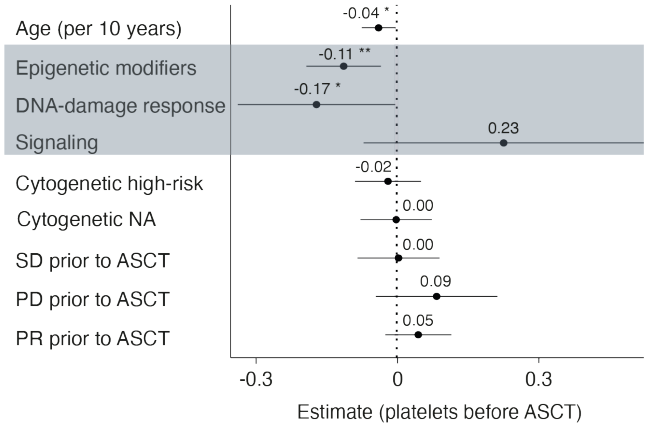

S9A
Patients (n)
Tandem ASCT
no
yes
CR
VGPR
PR
MR
SD
PD
Remission prior to ASCT
B
Patients (n)
Induction therapy
VD
VCD
VTD
VAD
PAD
TAD
RD
VRD
Other
Year
C
Survival probability
Progression-free survival
Log-rank p=0.65
Number at risk: n (%)
CR 56 (100) 41 (73) 29 (52) 19 (34) 11 (20) 7 (12) 3 (5) 2 (4) 0 (0) 0 (0)
PR 195 (100) 143 (73) 88 (45) 54 (28) 46 (24) 26 (13) 16 (8) 9 (5) 1 (1) 0 (0)
Days from ASCT
D
Overall survival
Log-rank p=0.038
Number at risk: n (%)
CR 56 (100) 54 (96) 51 (91) 37 (66) 23 (41) 14 (25) 9 (16) 3 (5) 0 (0) 0 (0)
PR 195 (100) 180 (92) 152 (78) 120 (62) 77 (39) 44 (23) 25 (13) 11 (6) 2 (1) 0 (0)
Days from ASCT
E
Survival probability
Progression-free survival
Log-rank p=0.14
Number at risk: n (%)
CR 56 (100) 41 (73) 29 (52) 19 (34) 11 (20) 7 (12) 3 (5) 2 (4) 0 (0) 0 (0)
PD 30 (100) 20 (67) 10 (33) 6 (20) 4 (13) 1 (3) 1 (3) 1 (3) 0 (0) 0 (0)
Days from ASCT
F
Overall survival
Log-rank p=0.004
Number at risk: n (%)
CR 56 (100) 54 (96) 51 (91) 37 (66) 23 (41) 14 (25) 9 (16) 3 (5) 0 (0) 0 (0)
PD 30 (100) 26 (87) 17 (57) 14 (47) 11 (37) 4 (13) 3 (10) 1 (3) 0 (0) 0 (0)
Days from ASCT

Supplemental Figure 9: Remission after induction chemotherapy (before ASCT) and clinical outcome

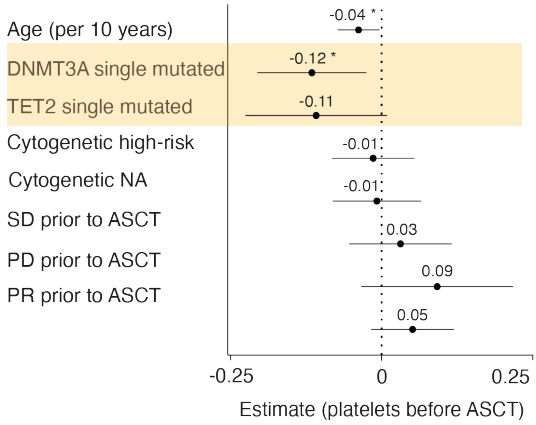
A) Bar chart including the remission status achieved before ASCT and performance of tandem transplant (stacked).
B) Bar chart illustrating the distribution of years and chosen induction therapies (stacked).
C) PFS stratified by response (CR vs. PR).
D) OS stratified by response (CR vs. PR).
E) PFS stratified by response (CR vs. PD).
F) OS stratified by response (CR vs. PD).

Abbreviations: CR = complete remission, VGPR = very good partial remission, PR = partial remission, MR = minimal response, SD = stable disease, PD = progressive disease, V/P = bortezomib, C = cyclophosphamide, D = dexamethasone, T = thalidomide, A = doxorubicin, R = lenalidomide.

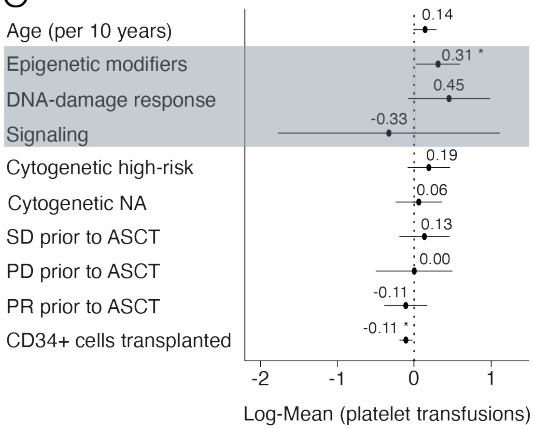
S10A


Age (per 10 years)
Epigenetic modifiers
DNA-damage response
Signaling
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
-0.04 *
-0.11 **
-0.17 *
0.23
-0.02
0.00
0.00
0.09
0.05
-0.3
0
0.3
Estimate (platelets before ASCT)

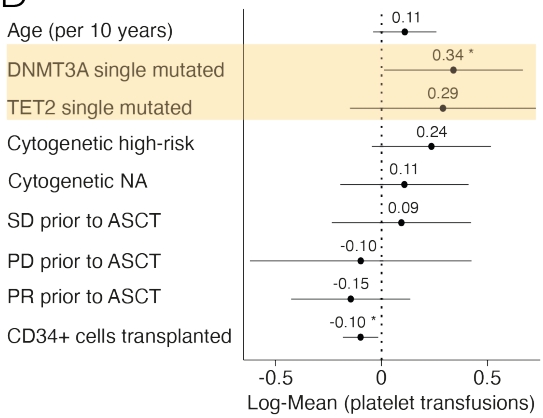
B


Age (per 10 years)
DNMT3A single mutated
TET2 single mutated
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
-0.04 *
-0.12 *
-0.11
-0.01
-0.01
0.03
0.09
0.05
-0.25
0
0.25
Estimate (platelets before ASCT)

C


Age (per 10 years)
Epigenetic modifiers
DNA-damage response
Signaling
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
0.14
0.31 *
0.45
-0.33
0.19
0.06
0.13
0.00
-0.11
-0.11 *
-2
-1
0
1
Log-Mean (platelet transfusions)

D


Age (per 10 years)
DNMT3A single mutated
TET2 single mutated
Cytogenetic high-risk
Cytogenetic NA
SD prior to ASCT
PD prior to ASCT
PR prior to ASCT
CD34+ cells transplanted
0.11
0.34 *
0.29
0.24
0.11
0.09
-0.10
-0.15
-0.10 *
-0.5
0
0.5
Log-Mean (platelet transfusions)

Supplemental Figure 10: Blood platelet counts and transfusion dependency (patients grouped by most prevalent driver genes and class of mutated genes)

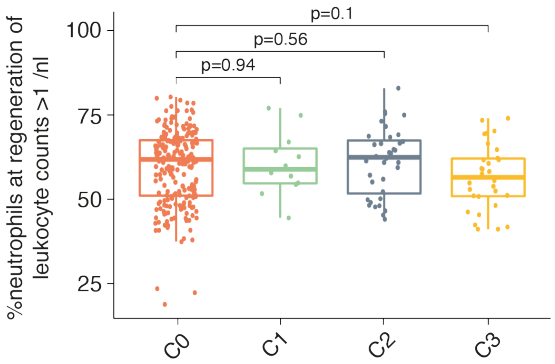
A) Forest plot visualizing the output of a linear regression model for platelet counts before transplant including the specified independent variables. The patients are grouped by class of mutated gene (epigenetic modifiers: DNMT3A/TET2 [n=71], DNA-damage response: TP53/PPM1D/PTEN [n=13] and signaling: JAK2/GNAS/CBL [n=4]). The platelet counts have been log-transformed to obtain a normal distribution. The plot illustrates the estimates/coefficients and their respective confidence intervals and statistical significance is indicated if the value is flagged with one or more stars.

B) Forest plot visualizing the output of a linear regression model for platelet counts before transplant including the specified independent variables. The patients are grouped by the most prevalent driver genes (*DNMT3A* single-mutated [n=46]), *TET2* single-mutated [n=25])

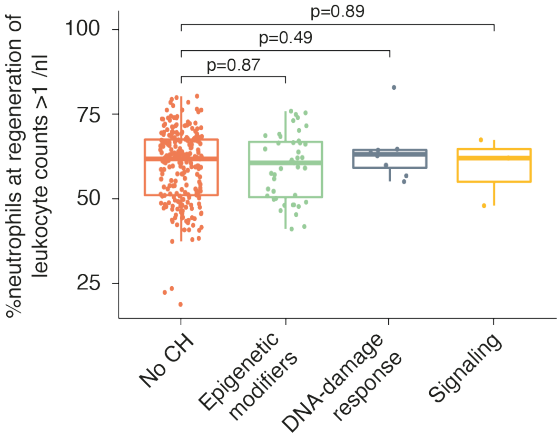
C) Forest plot visualizing a Poisson regression for the number of platelet transfusions within 20 days after transplant including the specified independent variables. The platelet values have been log-transformed to obtain a normal distribution. The log-mean values and the respective confidence intervals are shown. The patients are grouped by class of mutated gene.

D) Forest plot visualizing a Poisson regression for the number of platelet transfusions within 20 days after transplant. The patients are grouped by the most prevalent driver genes.

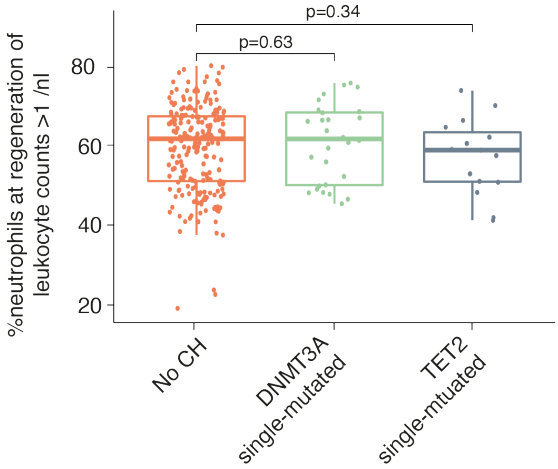
S11A


p=0.1
p=0.56
p=0.94
%neutrophils at regeneration of leukocyte counts >1 /nl
100
75
50
25
C0
C1
C2
C3

B


p=0.89
p=0.49
p=0.87
%neutrophils at regeneration of leukocyte counts >1 /nl
100
75
50
25
No CH
Epigenetic modifiers
DNA-damage response
Signaling

C


p=0.34
p=0.63
%neutrophils at regeneration of leukocyte counts >1 /nl
80
60
40
20
No CH
DNMT3A single-mutated
TET2 single-mtuated

Supplemental Figure 11: Neutrophil counts upon regeneration after post-transplant aplasia

A) Boxplot illustrating the relative proportion of neutrophils (of leukocyte counts) at regeneration of leukocyte counts > 1 /nl after transplant-related aplasia per cluster.

B) Boxplot illustrating the relative proportion of neutrophils (of leukocyte counts) at regeneration of leukocyte counts > 1 /nl after transplant-related aplasia. The patients are grouped by class of mutated gene.

E) Boxplot illustrating the relative proportion of neutrophils (of total leukocyte counts) at regeneration of leukocyte counts > 1 /nl after transplant-related aplasia. The patients are grouped by the most prevalent driver genes.

Supplemental method descriptions

Variant calling and filtering

NGS reads were aligned to the human genome assembly GRCh38 using BWA mem 0.7.17[1] and marked for duplicates using Samtools 1.9[2]. Variant calling was performed by the Mutect2 module of GATK version 4.1.8.1[3] using "-tumor" without matched normal sample and the Genome Aggregation Database as "--germline-resource" and the 1000 Genomes Project as "--panel-of-normals", all provided by the GATK resources (gs://gatk-best-practices/somatic-hg38/). Variants were functionally annotated by the Funcotator module of GATK.

To remove noise from called variants, non-coding and synonymous mutations were first removed and calls with variant allele frequency (VAF) > 0.05 were extracted. For each variant, we calculated the recurrence in other samples with $VAF > 0.005$ ($N_{0.005}$) and $VAF > 0.01$ ($N_{0.01}$), and the recurrence in the COSMIC database (N_{cosmic})[4]. Here, mutations were divided into two groups: substitutions and insertions/deletions (indels), to which different filtering strategies were applied. Substitutions were retained if satisfying any one of the three criteria: A) $N_{0.005} \geq 100$ and $N_{0.01} < 5$, B) $5 \leq N_{0.005} < 100$ and $N_{cosmic} \geq 30$, or C) $N_{0.005} < 5$. Retained substitutions were flagged as oncogenic if satisfying either: A) $N_{cosmic} \geq 10$, or B) resulting in a gain of stop codon and otherwise flagged as non-oncogenic. Flagged oncogenic/non-oncogenic mutations were manually corrected if existing annotations by ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) classed them as pathogenic/non-pathogenic. For indel filtering, any recurrent indel with $N_{0.01} \geq 2$ were removed, with the exception of NPM1 indels resulting in p.W288 frameshifts and ASXL1 indels resulting in p.G646 frameshifts.

NMF clustering and statistical analyses

To robustly identify patients with shared CH mutations, we applied a non-negative matrix factorization (NMF) clustering algorithm on the co-occurrence matrix of identified mutations. We excluded from the clustering mutant genes with a frequency less than 6, because of the low statistical power and induced convergence problems of the NMF clustering. Due to the generally low number of patients with CH mutations, we decided on three clusters (C1-C3) to keep the number of patients per cluster high for further analyses, whilst still ensuring sufficient differentiation. The NMF clustering provided the cluster membership for each patient. Patients without CH mutations were assigned to cluster C0.

Univariate and multivariable analyses of time-to-event endpoints were performed using Cox regression. OS was defined as time from ASCT, maintenance therapy or platelet nadir post ASCT until death from any cause. Subjects not confirmed dead were censored at the time last

known to be alive. PFS was defined as time from ASCT, maintenance therapy or platelet nadir post-ASCT until the earliest time of progression or death from any cause and censored at time last known to be alive and free of progression. Since we wanted to avoid the immortal time bias, but the start date of maintenance therapy was often not accurately reported, we chose day 90 post ASCT to calculate hazard ratios (HRs) for OS and PFS with respect to maintenance therapy. HRs with 95% confidence intervals (CI) and Wald P values were reported for model covariates, and likelihood-ratio tests and P values were reported for multivariable models. Median event times were estimated using the method of Kaplan and Meier (KM) and reported with 95% CIs. Greenwood's formula was used to approximate the variance of KM estimates. Differences in survival curves were assessed using log-rank tests.

Multivariate linear, logistic and Poisson regression were used for continuous, binary and count variables, respectively. Potential predictors were calculated in a complete case analysis by prior exploratory univariate regressions. Further, predictors considered clinically relevant were included. Multivariate regressions to predict response variables before ASCT were adjusted for the identified clusters, age (per 10 years), remission prior to ASCT and MM cytogenetic risk. In addition to these, regressions for response variables after ASCT were adjusted for transplanted CD34+ stem cells. Poisson regression was adjusted for overdispersion. Because C-reactive protein (CRP) values <2 mg/L were not reported, we used Tobit regression left-censored at 2 mg/L to test for associations between CRP and potential confounders[5]. Since stem cells were harvested for 1-5 days to ensure that an adequate number of CD34+ was collected, the total number of harvested stem cells was normalized for the number of harvest days. Because of the bimodal distribution of normalized harvested cell numbers, median regression was used. Fit of the linear models was assessed visually using plotted residuals. Fit of logistic regression was assessed with the Hosmer and Lemeshow goodness-of-fit test.

To assess the trajectories with multiple measurements per patient, mixed effects models were used to investigate the pattern of platelets, leukocytes, hemoglobin and CRP after ASCT. To satisfy the model assumptions, outcome variables (platelets, leukocytes, hemoglobin, CRP) were transformed with the natural logarithm or square root. As an independent variable, time after ASCT was included and modeled with a natural cubic spline, to allow a non-linear effect and to capture the trend in the data with more precision. We modelled outcome variables up to day 20 and up to day 50 after ASCT. For models to day 20 one spline knot was set on day 9. Models to day 50 used two spline knots on day 7 and 12. Residual plots were used to validate the models' assumptions.

Continuous and categorical data are reported as mean (standard deviation) and count (percent), respectively. Fisher's exact test was used to test for associations between categorical variables. The Mann-Whitney-U or Kruskal-Wallis rank-sum test were used to assess a location shift in the distribution of continuous variables between two or more than two groups, respectively.

References

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.

2. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079.

3. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-1303.

4. Tate JG, Bamford S, Jubb HC, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019;47(D1):D941-D947.

5. Ahmadi H, Granger DA, Hamilton KR, Blair C, Riis JL. Censored data considerations and analytical approaches for salivary bioscience data. Psychoneuroendocrinology. 2021;129(105274.