Data file S1: Detected somatic mutations in donors

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	CADD	Deleterious	VAF1	VAF2	patientID	Engrafted	Age
chr11	32456672	G	A	WT1	missense	p.R74W	NA	28.7	Yes	0.011304	0.011202	PID_0589	Yes	30
chr2	25469138	C	T	DNMT3A	nonsense	p.W288X	COSM1130818	40	Yes	0.0004	0.001268	PID_0489	Yes	58
chr2	25470498	G	T	DNMT3A	missense	p.R174S	NA	26.5	Yes	0.011218	0.008888	PID_0489	Yes	58
chr6	75893069	T	A	COL12A1	missense	p.I530L	COSM271996	22.1	Yes	0.000843	0.002573	PID_0489	Yes	58
chr2	25467428	C	T	DNMT3A	missense	p.G398R	COSM256035	30	Yes	0.00142	0.00182	PID_0459	Yes	28
chr2	25965982	C	T	ASXL2	missense	p.R1075Q	COSM6494820	10.77	No	0.004005	0.004873	PID_0450	Yes	26
chr2	61492689	T	C	USP34	missense	p.H1874R	NA	22.5	Yes	0.002401	0.002433	PID_0450	Yes	26
chr2	25470544	A	C	DNMT3A	missense	p.I158M	NA	23.6	Yes	0.000541	0.000614	PID_0655	Yes	36
chr4	106180927	G	A	TET2	splicing	NA	COSM87141	34	Yes	0.000185805	0.001074009	PID_0421	Yes	38
chrX	123179113	T	G	STAG2	missense	p.Y188D	NA	26.3	Yes	0.00164	0.004084	PID_00394	Yes	40
chr2	25469921	T	G	DNMT3A	missense	p.Q222P	NA	26.1	Yes	0.012781	0.01809	PID_0373	Yes	51
chr4	106182995	A	G	TET2	missense	p.Y1345C	NA	32	Yes	0.00133	0.001933	PID_0373	Yes	51
chr16	30721283	T	TGCTTCGCC	SRCAP	indel	NA	NA	29	Yes	0.029	0.0258	PID_0372	Yes	23
chr17	11556271	T	C	DNAH9	silent	p.Y849Y	NA	6.417	No	0.003965	0.000986	PID_0314	Yes	20
chr6	75818737	G	A	COL12A1	silent	p.A1535A	NA	1.438	No	0.001792	0.00254	PID_0268	Yes	53
chr16	3801781	G	A	CREBBP	missense	p.T1204I	NA	25.5	Yes	0.002441	0.001158	PID_0268	Yes	53
chr17	7577094	G	A	TP53	missense	p.R150W	COSM99925	25.7	Yes	0.013423	0.017781	PID_0268	Yes	53
chr2	25459821	T	G	DNMT3A	missense	p.H669P	NA	23.3	Yes	0.005147	0.004734	PID_0268	Yes	53
chr4	187549458	C	T	FAT1	missense	p.D1554N	COSM1429043	25.8	Yes	0.007042	0.004139	PID_0268	Yes	53

Data file S2: Detected somatic mutations in recipients post-HSCT after removing recipient's hematopoietic clones

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	CADD	VAF1	VAF2	Timepoint	patientID	Engrafted or New
chr16	3801781	G	A	CREBBP	missense	p.T1204I	NA	25.5	0.00088	0.001588	D30	PID_0268	Engrafted
chr16	3801781	G	A	CREBBP	missense	p.T1204I	NA	25.5	0.001788	0.001573	D100	PID_0268	Engrafted
chr16	3801781	G	A	CREBBP	missense	p.T1204I	NA	25.5	0.001285	0.001304	D365	PID_0268	Engrafted
chr17	7577094	G	A	TP53	missense	p.R150W	COSM99925	25.7	0.018132	0.018873	D30	PID_0268	Engrafted
chr17	7577094	G	A	TP53	missense	p.R150W	COSM99925	25.7	0.023088	0.018078	D100	PID_0268	Engrafted
chr17	7577094	G	A	TP53	missense	p.R150W	COSM99925	25.7	0.023448	0.022782	D365	PID_0268	Engrafted
chr2	25459821	T	G	DNMT3A	missense	p.H669P	NA	23.3	0.004792	0.00754	D30	PID_0268	Engrafted
chr2	25459821	T	G	DNMT3A	missense	p.H669P	NA	23.3	0.00522	0.001862	D100	PID_0268	Engrafted
chr2	25459821	T	G	DNMT3A	missense	p.H669P	NA	23.3	0.004636	0.005898	D365	PID_0268	Engrafted
chr2	25462044	G	A	DNMT3A	missense	p.A636V	NA	26.9	0.001243	0.001637	D30	PID_0268	New
chr2	25462044	G	A	DNMT3A	missense	p.A636V	NA	26.9	0.001319	0.000832	D100	PID_0268	New
chr6	75818737	G	A	COL12A1	silent	p.A1535A	NA	1.438	0.001169	0.001683	D30	PID_0268	Engrafted
chr6	75818737	G	A	COL12A1	silent	p.A1535A	NA	1.438	0.00164	0.002633	D100	PID_0268	Engrafted
chr6	75818737	G	A	COL12A1	silent	p.A1535A	NA	1.438	0.002188	0.000939	D365	PID_0268	Engrafted
chr4	106157471	ATT	A	TET2	indel	NA	NA	27.2	0.004	0.0027	D365	PID_0268	New
chr4	187549458	C	T	FAT1	missense	p.D1554N	COSM1429043	25.8	0.003422	0.006051	D30	PID_0268	Engrafted
chr4	187549458	C	T	FAT1	missense	p.D1554N	COSM1429043	25.8	0.004722	0.00372	D100	PID_0268	Engrafted
chr4	187549458	C	T	FAT1	missense	p.D1554N	COSM1429043	25.8	0.005821	0.005068	D365	PID_0268	Engrafted
chr17	11556271	T	C	DNAH9	silent	p.Y849Y	NA	6.417	0.002639	0.001049	D30	PID_0314	Engrafted
chr17	11556271	T	C	DNAH9	silent	p.Y849Y	NA	6.417	0.010758	0.001891	D100	PID_0314	Engrafted
chr17	10427979	T	C	MYH2	missense	p.Q1660R	NA	23.8	0.001193	0.001827	D365	PID_0318	New
chr4	187630237	G	T	FAT1	missense	p.Q249K	NA	22.7	0.004705	0.006863	D100	PID_0318	New
chr4	187630237	G	T	FAT1	missense	p.Q249K	NA	22.7	0.002002	0.003591	D365	PID_0318	New
chr7	50450427	C	T	IKZF1	intronic	NA	NA	0.256	0.003137	0.00269	D365	PID_0335	New
chr7	135322646	A	G	NUP205	intronic	NA	NA	11.51	0.001635	0.001371	D100	PID_0335	New
chr9	139391938	C	A	NOTCH1	missense	p.A2085S	NA	25.1	0.010538	0.011864	D100	PID_0335	New
chr9	139391938	C	A	NOTCH1	missense	p.A2085S	NA	25.1	0.004486	0.002006	D365	PID_0335	New
chr17	7578190	T	C	TP53	missense	p.Y88C	COSM99718	29.5	0.006901	0.00515	D100	PID_0347	New
chr20	31022402	TCACCACTGCCATAGAGAGGCGGC	T	ASXL1	indel	NA	NA	35	0.0142	0.0129	D100	PID_0360	New
chr4	106190797	C	T	TET2	missense	p.R1359C	COSM41649	32	0.004008	0.002231	D100	PID_0361	New
chr2	25458642	T	G	DNMT3A	missense	p.K692T	NA	28.6	0.003823	0.004353	D30	PID_0366	New
chr2	25458642	T	G	DNMT3A	missense	p.K692T	NA	28.6	0.001147	0.00074	D100	PID_0366	New
chrX	44945182	T	C	KDM6A	missense	p.V1090A	NA	27	0.001714	0.002526	D365	PID_0366	New
chr16	30721283	T	TGCTTCGCC	SRCAP	indel	NA	NA	29	0.0266	0.0263	D30	PID_0372	Engrafted
chr16	30721283	T	TGCTTCGCC	SRCAP	indel	NA	NA	29	0.033	0.0255	D100	PID_0372	Engrafted
chr16	30721283	T	TGCTTCGCC	SRCAP	indel	NA	NA	29	0.006	0.0037	D365	PID_0372	Engrafted
chrX	44896997	C	G	KDM6A	intronic	NA	NA	5.46	0.003722	0.003451	D365	PID_0372	New
chr2	25469921	T	G	DNMT3A	missense	p.Q222P	NA	26.1	0.024433	0.025272	D30	PID_0373	Engrafted
chr2	25469921	T	G	DNMT3A	missense	p.Q222P	NA	26.1	0.02903	0.025094	D100	PID_0373	Engrafted

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	CADD	VAF1	VAF2	Timepoint	patientID	Engrafted or New
chr2	25469921	T	G	DNMT3A	missense	p.Q222P	NA	26.1	0.01125	0.008141	D365	PID_0373	Engrafted
chr4	106182995	A	G	TET2	missense	p.Y1345C	NA	32	0.003003	0.002545	D30	PID_0373	Engrafted
chr4	106182995	A	G	TET2	missense	p.Y1345C	NA	32	0.001785	0.002803	D100	PID_0373	Engrafted
chr17	10432262	A	G	MYH2	silent	p.T1163T	NA	15.78	0.002419	0.002522	D365	PID_0394	New
chr17	11701030	T	A	DNAH9	silent	p.T2780T	NA	1.371	0.005347	0.005514	D365	PID_0394	New
chr16	30736081	T	C	SRCAP	missense	p.L1779P	NA	23.7	0.002611	0.001984	D100	PID_0394	New
chr16	30736081	T	C	SRCAP	missense	p.L1779P	NA	23.7	0.001275	0.002601	D365	PID_0394	New
chrX	123179113	T	G	STAG2	missense	p.Y188D	NA	26.3	0.003225	0.001184	D30	PID_0394	Engrafted
chr4	24572496	T	G	DHX15	intronic	NA	NA	16.29	0.009922	0.005584	D30	PID_0421	New
chr4	106156485	TC	T	TET2	indel	NA	NA	26.7	0.0289	0.02	D365	PID_0421	New
chr4	106180927	G	A	TET2	splicing	NA	COSM87141	34	0.000737075	0.001845444	D30	PID_0421	Engrafted
chr4	106180927	G	A	TET2	splicing	NA	COSM87141	34	0.000895255	0.001129093	D100	PID_0421	Engrafted
chr4	106180927	G	A	TET2	splicing	NA	COSM87141	34	0.004502	0.004145	D365	PID_0421	Engrafted
chr4	106190860	C	T	TET2	missense	p.H1380Y	COSM87161	27.4	0.00169	0.001286	D365	PID_0421	New
chr16	3832811	G	A	CREBBP	stopgain	p.R445X	COSM255965	38	0.002238	0.003813	D100	PID_0450	New
chr16	3832811	G	A	CREBBP	stopgain	p.R445X	COSM255965	38	0.031195	0.023312	D365	PID_0450	New
chr17	11502173	G	T	DNAH9	missense	p.V120L	NA	11.09	0.002636	0.003343	D365	PID_0450	New
chr2	25965982	C	T	ASXL2	missense	p.R1075Q	COSM6494820	10.77	0.005974	0.007519	D30	PID_0450	Engrafted
chr2	25965982	C	T	ASXL2	missense	p.R1075Q	COSM6494820	10.77	0.007112	0.016319	D100	PID_0450	Engrafted
chr2	25965982	C	T	ASXL2	missense	p.R1075Q	COSM6494820	10.77	0.003058	0.004677	D365	PID_0450	Engrafted
chr2	61492689	T	C	USP34	missense	p.H1874R	NA	22.5	0.003149	0.003222	D30	PID_0450	Engrafted
chr2	61492689	T	C	USP34	missense	p.H1874R	NA	22.5	0.002382	0.001837	D100	PID_0450	Engrafted
chr2	61492689	T	C	USP34	missense	p.H1874R	NA	22.5	0.001734	0.001308	D365	PID_0450	Engrafted
chrX	123181311	C	T	STAG2	stopgain	p.R259X	COSM1598816	36	0.004316	0.004992	D100	PID_0450	New
chr9	139399647	G	A	NOTCH1	intronic	NA	NA	1.322	0.001	0.001415	D100	PID_0450	New
chr9	139399647	G	A	NOTCH1	intronic	NA	NA	1.322	0.001125	0.001466	D365	PID_0450	New
chr2	25467428	C	T	DNMT3A	missense	p.G398R	COSM256035	30	0.001465	0.001914	D30	PID_0459	Engrafted
chr2	25467428	C	T	DNMT3A	missense	p.G398R	COSM256035	30	0.001199	0.002875	D100	PID_0459	Engrafted
chr2	25467428	C	T	DNMT3A	missense	p.G398R	COSM256035	30	0.000982	0.002632	D365	PID_0459	Engrafted
chr4	106190798	G	A	TET2	missense	p.R1359H	COSM42055	33	0.001979	0.004812	D100	PID_0467	New
chr4	187517780	G	A	FAT1	missense	p.A4305V	COSM6056356	33	0.006958	0.008668	D365	PID_0467	New
chr17	10355270	G	T	MYH4	silent	p.V1242V	NA	15.61	0.018793	0.019914	D365	PID_0475	New
chr16	3778533	T	G	CREBBP	missense	p.N2134T	NA	22.7	0.000779	0.001114	D30	PID_0489	New
chr16	3778533	T	G	CREBBP	missense	p.N2134T	NA	22.7	0.000768	0.000833	D100	PID_0489	New
chr2	25469138	C	T	DNMT3A	stopgain	p.W288X	COSM1130818	40	0.001723	0.001138	D30	PID_0489	Engrafted
chr2	25469138	C	T	DNMT3A	stopgain	p.W288X	COSM1130818	40	0.000715	0.00153	D100	PID_0489	Engrafted
chr2	25469138	C	T	DNMT3A	stopgain	p.W288X	COSM1130818	40	0.001764	0.001629	D365	PID_0489	Engrafted
chr2	25470498	G	T	DNMT3A	missense	p.R174S	NA	26.5	0.015894	0.017054	D30	PID_0489	Engrafted
chr2	25470498	G	T	DNMT3A	missense	p.R174S	NA	26.5	0.011207	0.01087	D100	PID_0489	Engrafted
chr2	25470498	G	T	DNMT3A	missense	p.R174S	NA	26.5	0.011979	0.007639	D365	PID_0489	Engrafted
chr6	75893069	T	A	COL12A1	missense	p.I530L	COSM271996	22.1	0.002051	0.003665	D100	PID_0489	Engrafted

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	CADD	VAF1	VAF2	Timepoint	patientID	Engrafted or New
chr20	31022592	CG	C	ASXL1	indel	NA	NA	34	0.0062	0.0053	D365	PID_0495	New
chr17	7681670	C	A	DNAH2	silent	p.G1808G	NA	5.337	0.004119	0.004819	D100	PID_0589	New
chr17	7681670	C	A	DNAH2	silent	p.G1808G	NA	5.337	0.008886	0.008291	D365	PID_0589	New
chr19	13445165	C	T	CACNA1A	intronic	NA	NA	12.07	0.004625	0.001711	D365	PID_0589	New
chr11	32456672	G	A	WT1	missense	p.R74W	NA	28.7	0.013294	0.015119	D30	PID_0589	Engrafted
chr11	32456672	G	A	WT1	missense	p.R74W	NA	28.7	0.009875	0.009044	D100	PID_0589	Engrafted
chr11	32456672	G	A	WT1	missense	p.R74W	NA	28.7	0.01129	0.009187	D365	PID_0589	Engrafted
chr2	25470544	A	C	DNMT3A	missense	p.I158M	NA	23.6	0.00058	0.000579	D100	PID_0655	Engrafted

Data file S3: Shared variants in pre-HSCT and post-HSCT recipient samples due to incomplete clearance of recipient's hematopoietic clones post-HSCT.

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	VAF1	VAF2	Timepoint	patientID
chr18	7002407	A	C	LAMA1	intronic	NA	NA	0.434713	0.356371	Pre	PID_0576
chr18	7002407	A	C	LAMA1	intronic	NA	NA	0.021635	0.024166	D30	PID_0576
chr18	7002407	A	C	LAMA1	intronic	NA	NA	0.013414	0.018904	D100	PID_0576
chr7	135285707	G	A	NUP205	silent	p.V430V	COSM40403	0.505976	0.512114	Pre	PID_0499
chr7	135285707	G	A	NUP205	silent	p.V430V	COSM40403	0.019954	0.021033	D30	PID_0499
chr7	135285707	G	A	NUP205	silent	p.V430V	COSM40403	0.022706	0.025451	D100	PID_0499
chr11	32456726	G	A	WT1	missense	p.R56W	NA	0.404366	0.403446	Pre	PID_0495
chr11	32456726	G	A	WT1	missense	p.R56W	NA	0.000908	0.000609	D30	PID_0495
chr16	30736222	C	T	SRCAP	missense	p.S1826L	COSM5850742	0.507346	0.496049	Pre	PID_0495
chr16	30736222	C	T	SRCAP	missense	p.S1826L	COSM5850742	0.04389	0.039401	D30	PID_0495
chr16	30736222	C	T	SRCAP	missense	p.S1826L	COSM5850742	0.006037	0.003924	D100	PID_0495
chr17	11550421	G	A	DNAH9	missense	p.R668Q	COSM1236009	0.910698	0.90111	Pre	PID_0495
chr17	11550421	G	A	DNAH9	missense	p.R668Q	COSM1236009	0.042462	0.055456	D30	PID_0495
chr17	7669761	G	A	DNAH2	missense	p.E1213K	NA	0.912955	0.915971	Pre	PID_0495
chr17	7669761	G	A	DNAH2	missense	p.E1213K	NA	0.052973	0.051185	D30	PID_0495
chr8	128750945	C	T	MYC	missense	p.S161L	COSM1454792	0.40811	0.402829	Pre	PID_0495
chr8	128750945	C	T	MYC	missense	p.S161L	COSM1454792	0.002216	0.001834	D30	PID_0495
chr9	134072817	C	G	NUP214	silent	p.T138T	NA	0.507812	0.506112	Pre	PID_0495
chr9	134072817	C	G	NUP214	silent	p.T138T	NA	0.042121	0.039321	D30	PID_0495
chr9	134072817	C	G	NUP214	silent	p.T138T	NA	0.001785	0.004946	D100	PID_0495
chr17	7697727	G	A	DNAH2	intronic	NA	NA	0.491119	0.492356	Pre	PID_0489
chr17	7697727	G	A	DNAH2	intronic	NA	NA	0.001426	0.00106	D30	PID_0489
chr9	134074118	G	C	NUP214	missense	p.S572T	NA	0.507816	0.517495	Pre	PID_0489
chr9	134074118	G	C	NUP214	missense	p.S572T	NA	0.001387	0.001787	D30	PID_0489
chr17	7577644	C	G	TP53	intronic	NA	COSN26958754	0.507881	0.508418	Pre	PID_0475
chr17	7577644	C	G	TP53	intronic	NA	COSN26958754	0.000881	0.000709	D30	PID_0475
chr5	180057843	A	G	FLT4	intronic	NA	NA	0.519392	0.502796	Pre	PID_0475
chr5	180057843	A	G	FLT4	intronic	NA	NA	0.001027	0.001216	D30	PID_0475
chr13	28608020	G	C	FLT3	intronic	NA	NA	0.505227	0.483484	Pre	PID_0459

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	VAF1	VAF2	Timepoint	patientID
chr13	28608020	G	C	FLT3	intronic	NA	NA	0.032958	0.028085	D30	PID_0459
chr13	28608020	G	C	FLT3	intronic	NA	NA	0.056743	0.058917	D100	PID_0459
chr4	106180845	G	T	TET2	missense	p.W1291C	COSM4383925	0.175306	0.185252	Pre	PID_0459
chr4	106180845	G	T	TET2	missense	p.W1291C	COSM4383925	0.007234	0.00947	D30	PID_0459
chr4	106180845	G	T	TET2	missense	p.W1291C	COSM4383925	0.004535	0.003971	D100	PID_0459
chr4	106157845	C	T	TET2	nonsense	p.Q916X	COSM3733079	0.171096	0.191542	Pre	PID_0459
chr4	106157845	C	T	TET2	nonsense	p.Q916X	COSM3733079	0.004234	0.003035	D30	PID_0459
chr4	106157845	C	T	TET2	nonsense	p.Q916X	COSM3733079	0.007722	0.004722	D100	PID_0459
chr4	55575669	G	A	KIT	missense	p.V399I	COSM51494	0.485625	0.495795	Pre	PID_0459
chr4	55575669	G	A	KIT	missense	p.V399I	COSM51494	0.029492	0.03654	D30	PID_0459
chr4	55575669	G	A	KIT	missense	p.V399I	COSM51494	0.05834	0.047988	D100	PID_0459
chr5	180057876	C	T	FLT4	intronic	NA	NA	0.509449	0.501632	Pre	PID_0459
chr5	180057876	C	T	FLT4	intronic	NA	NA	0.033549	0.029598	D30	PID_0459
chr5	180057876	C	T	FLT4	intronic	NA	NA	0.063978	0.057464	D100	PID_0459
chr5	180046495	A	C	FLT4	intronic	NA	NA	0.530321	0.566294	Pre	PID_0459
chr5	180046495	A	C	FLT4	intronic	NA	NA	0.028811	0.03513	D30	PID_0459
chr5	180046495	A	C	FLT4	intronic	NA	NA	0.067174	0.070913	D100	PID_0459
chr3	128200072	C	T	GATA2	silent	p.A397A	COSM5019736	0.501648	0.504831	Pre	PID_0467
chr3	128200072	C	T	GATA2	silent	p.A397A	COSM5019736	0.150707	0.155821	D30	PID_0467
chr6	138632652	G	A	ARFGEF3	intronic	NA	NA	0.001567	0.001827	Pre	PID_0467
chr6	138632652	G	A	ARFGEF3	intronic	NA	NA	0.001096	0.000721	D30	PID_0467
chr16	15814717	C	T	MYH11	silent	p.K1590K	COSM5020106	0.499401	0.496203	Pre	PID_0437
chr16	15814717	C	T	MYH11	silent	p.K1590K	COSM5020106	0.002947	0.003508	D100	PID_0437
chr16	15814717	C	T	MYH11	silent	p.K1590K	COSM5020106	0.00276	0.003484	D360	PID_0437
chrX	44922982	C	G	KDM6A	missense	p.L536V	NA	0.509905	0.497262	Pre	PID_0437
chrX	44922982	C	G	KDM6A	missense	p.L536V	NA	0.001176	0.002106	D30	PID_0437
chrX	44922982	C	G	KDM6A	missense	p.L536V	NA	0.002725	0.001689	D100	PID_0437
chr4	106156384	G	A	TET2	missense	p.G429R	COSM219042	0.493444	0.49112	Pre	PID_0409
chr4	106156384	G	A	TET2	missense	p.G429R	COSM219042	0.00807	0.008017	D30	PID_0409
chr4	106156384	G	A	TET2	missense	p.G429R	COSM219042	0.013121	0.011373	D100	PID_0409
chr16	3778424	T	G	CREBBP	missense	p.Q2170H	COSM96470	0.476022	0.503146	Pre	PID_0409
chr16	3778424	T	G	CREBBP	missense	p.Q2170H	COSM96470	0.008232	0.008722	D30	PID_0409

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	VAF1	VAF2	Timepoint	patientID
chr16	3778424	T	G	CREBBP	missense	p.Q2170H	COSM96470	0.01268	0.015275	D100	PID_0409
chr16	3779594	C	T	CREBBP	silent	p.V1780V	COSM5019155	0.489053	0.493702	Pre	PID_0409
chr16	3779594	C	T	CREBBP	silent	p.V1780V	COSM5019155	0.008906	0.010031	D30	PID_0409
chr16	3779594	C	T	CREBBP	silent	p.V1780V	COSM5019155	0.01582	0.015028	D100	PID_0409
chr16	30727853	C	G	SRCAP	intronic	NA	NA	0.48963	0.496588	Pre	PID_0394
chr16	30727853	C	G	SRCAP	intronic	NA	NA	0.03362	0.032133	D30	PID_0394
chr17	7684131	C	T	DNAH2	intronic	NA	NA	0.014319	0.012846	D30	PID_0361
chr17	7684131	C	T	DNAH2	intronic	NA	NA	0.503446	0.507958	Pre	PID_0361
chr20	31024704	G	A	ASXL1	missense	p.G1397S	COSM133033	0.016931	0.018797	D30	PID_0361
chr20	31024704	G	A	ASXL1	missense	p.G1397S	COSM133033	0.498835	0.498557	Pre	PID_0361
chrX	39937244	A	C	BCOR	intronic	NA	NA	0.011781	0.015437	D30	PID_0361
chrX	39937244	A	C	BCOR	intronic	NA	NA	1	0.998739	Pre	PID_0361
chr2	209113113	G	A	IDH1	missense	p.R132C	COSM28747	0.009563	0.005514	Pre	PID_0360
chr2	209113113	G	A	IDH1	missense	p.R132C	COSM28747	0.032294	0.034172	D100	PID_0360
chr20	31024260	A	G	ASXL1	missense	p.M1249V	COSM6498418	0.513097	0.490742	Pre	PID_0360
chr20	31024260	A	G	ASXL1	missense	p.M1249V	COSM6498418	0.039568	0.033524	D30	PID_0360
chr20	31024260	A	G	ASXL1	missense	p.M1249V	COSM6498418	0.057228	0.074445	D100	PID_0360
chr20	31022444	G	C	ASXL1	silent	p.G643G	NA	0.609968	0.650678	Pre	PID_0360
chr20	31022444	G	C	ASXL1	silent	p.G643G	NA	0.064838	0.070318	D30	PID_0360
chr20	31022444	G	C	ASXL1	silent	p.G643G	NA	0.094106	0.115806	D100	PID_0360
chr21	36252865	C	T	RUNX1	missense	p.R139Q	COSM36055	0.005559	0.003582	Pre	PID_0360
chr21	36252865	C	T	RUNX1	missense	p.R139Q	COSM36055	0.028548	0.027494	D100	PID_0360
chr3	47125604	A	G	SETD2	missense	p.M1845T	NA	0.501668	0.513251	Pre	PID_0360
chr3	47125604	A	G	SETD2	missense	p.M1845T	NA	0.03326	0.031476	D30	PID_0360
chr3	47125604	A	G	SETD2	missense	p.M1845T	NA	0.066519	0.069635	D100	PID_0360
chr7	148523590	C	T	EZH2	missense	p.R249Q	COSM1449004	0.00326	0.004129	Pre	PID_0360
chr7	148523590	C	T	EZH2	missense	p.R249Q	COSM1449004	0.014485	0.015988	D100	PID_0360
chr16	30735348	C	G	SRCAP	missense	p.P1535A	NA	0.491278	0.505079	Pre	PID_0347
chr16	30735348	C	G	SRCAP	missense	p.P1535A	NA	0.013502	0.012958	D30	PID_0347
chr16	30735348	C	G	SRCAP	missense	p.P1535A	NA	0.02768	0.030115	D100	PID_0347
chr17	7578535	T	C	TP53	missense	p.K93R	COSM3388223	0.002619	0.002549	Pre	PID_0347
chr17	7578535	T	C	TP53	missense	p.K93R	COSM3388223	0.003423	0.003649	D30	PID_0347

Chr	Start	Ref	Alt	Gene	Type	AA	COSMIC ID	VAF1	VAF2	Timepoint	patientID
chr17	7578406	C	T	TP53	missense	p.R43H	COSM10648	0.001863	0.000881	Pre	PID_0347
chr17	7578406	C	T	TP53	missense	p.R43H	COSM10648	0.002412	0.001305	D30	PID_0347
chr17	7578406	C	T	TP53	missense	p.R43H	COSM10648	0.001066	0.002532	D100	PID_0347
chr17	7578263	G	A	TP53	nonsense	p.R64X	COSM3378446	0.040574	0.037321	Pre	PID_0347
chr17	7578263	G	A	TP53	nonsense	p.R64X	COSM3378446	0.016537	0.012518	D30	PID_0347
chr17	7578263	G	A	TP53	nonsense	p.R64X	COSM3378446	0.052481	0.039326	D100	PID_0347
chr2	198266834	T	C	SF3B1	missense	p.K700E	COSM84677	0.029823	0.030639	Pre	PID_0347
chr2	198266834	T	C	SF3B1	missense	p.K700E	COSM84677	0.01055	0.007808	D30	PID_0347
chr2	198266834	T	C	SF3B1	missense	p.K700E	COSM84677	0.043792	0.031544	D100	PID_0347
chr17	29677184	C	T	NF1	intronic	NA	NA	0.014122	0.013542	D30	PID_0335
chr17	29677184	C	T	NF1	intronic	NA	NA	0.501188	0.487533	Pre	PID_0335
chr11	32417962	A	G	WT1	intronic	NA	COSM6494031	0.506841	0.483758	Pre	PID_0268
chr11	32417962	A	G	WT1	intronic	NA	COSM6494031	0.00479	0.005205	D30	PID_0268
chr11	32417962	A	G	WT1	intronic	NA	COSM6494031	0.00144	0.001023	D100	PID_0268
chr17	7727547	C	T	DNAH2	missense	p.R3863C	NA	0.497982	0.518192	Pre	PID_0268
chr17	7727547	C	T	DNAH2	missense	p.R3863C	NA	0.004903	0.00291	D30	PID_0268
chr17	7727547	C	T	DNAH2	missense	p.R3863C	NA	0.002456	0.002404	D100	PID_0268
chr2	25991798	A	G	ASXL2	intronic	NA	NA	0.559502	0.508356	Pre	PID_0268
chr2	25991798	A	G	ASXL2	intronic	NA	NA	0.004492	0.003673	D30	PID_0268
chr2	25991798	A	G	ASXL2	intronic	NA	NA	0.003618	0.000983	D100	PID_0268
chr9	139399343	C	T	NOTCH1	silent	p.L1600L	NA	0.506623	0.51215	Pre	PID_0268
chr9	139399343	C	T	NOTCH1	silent	p.L1600L	NA	0.005394	0.003518	D30	PID_0268
chr9	139399343	C	T	NOTCH1	silent	p.L1600L	NA	0.003461	0.002481	D100	PID_0268

Data file S4: Analysis of recipient clinical outcomes in relation to engraftment of donor-derived mutations			
Mutation engraftment variable	Disease characteristic	Test	p-value
Pathogenic donor mutation	Acute GvHD	Fisher's Exact test	0.54
	Cytopenia	Fisher's Exact test	0.24
	Duration of cytopenia	Wilcoxon rank-sum test	0.28
	Cumulative incidence of of chronic GvHD	Fine-Gray subdistribution hazard model, Gray's test	0.22
	Leukemia Free Survival	Kaplan-Meier model, log-rank test	0.57
	CMV reactivation	Fisher's Exact test	0.092
	Cardiac event	Fisher's Exact test	0.99
	Mixed chimerism	Repeated measure logistic regression	0.23
	Neutrophil engraftment	No test, all engrafted	-
Persistent donor engraftment	Acute GvHD	Fisher's Exact test	0.23
	Cytopenia	Fisher's Exact test	0.21
	Duration of cytopenia	Wilcoxon rank-sum test	0.054
	Cumulative incidence of of chronic GvHD	Fine-Gray subdistribution hazard model, Gray's test	0.23
	Leukemia Free Survival	Kaplan-Meier model, log-rank test	0.72
	CMV reactivation	Fisher's Exact test	0.099
	Cardiac event	Fisher's Exact test	0.57
	Mixed chimerism	Repeated measure logistic regression	0.33
	Neutrophil engraftment	No test, all engrafted	-
Persistent engraftment of COSMIC-related donor mutation	Acute GvHD	Fisher's Exact test	0.5
	Cytopenia	Fisher's Exact test	0.99
	Duration of cytopenia	Wilcoxon rank-sum test	0.2
	Cumulative incidence of of chronic GvHD	Fine-Gray subdistribution hazard model, Gray's test	0.14
	Leukemia Free Survival	Kaplan-Meier model, log-rank test	0.43
	CMV reactivation	Fisher's Exact test	0.33
	Cardiac event	Fisher's Exact test	0.55
	Mixed chimerism	Repeated measure logistic regression	0.65
	Neutrophil engraftment	No test, all engrafted	-

Data file S5: Recurrently mutated genes in adult and pediatric AML

Adult	ABL1	ASXL1	ATRX	BCOR	BRAF	CALR	CBL
	CBLB	CBLC	CDKN2A	CEBPA	CSF3R	DNMT3A	ETV6
	EZH2	FBXW7	FLT3	GATA1	GATA2	GNAS	HRAS
	IDH1	IDH2	IKZF1	JAK2	JAK3	KDM6A	KIT
	KRAS	MPL	MYD88	NOTCH1	NPM1	NRAS	PDGFRA
	PHF6	PTEN	PTPN11	RAD21	RUNX1	SETBP1	SF3B1
	SMC1A	STAG2	TET2	TP53	U2AF1	WT1	ZRSR2
Pediatric	AFF3	ARFGEF3	ASXL2	CACNA1A	CCND3	COL12A1	CREBBP
	DHX15	DNAH2	DNAH9	FAT1	FLT4	KMT2A	LAMA1
	MED23	MFSD11	MYC	MYH11	MYH2	MYH4	NCOR1
	NF1	NUP205	NUP214	RELN	SETD2	SPI1	SRCAP
	TRIM24	TRRAP	USP34				

Data file S6: ddPCR probe sequences	
Mutation	Probe sequence
STAG2 p.R259*	CATTAATATGGATAATACACAAAGACAATATGAAGCAGAACG GAATAAAATGATTGGAAAA[C/T]GAGCCAATGAGAGGCTAGA ACTCCTGCTACAAAAGCGGAAAGAGGTAAACTTTTATATTGA
DNAH2 p.G1808G	CACTGACCACGGCATTGCACCTGCACCGAGGGGGCTCCCCCA AAGGCCCTGCAGGCACAGG[C/A]AAGACCGAGACCGTCAAGG ACCTGGGCAAGGCCCTGGGCATATATGTCATTGTGGTCAACT
CREBBP p.R445X	AGGGCAACAGAATGCCACTTCTTTAAGTAACCCAAATCCCATA GACCCCAGCTCCATGCAG[C/T]GAGCCTATGCTGCTCTCGGACT CCCCTACATGAACCAGCCCCAGACGCAGCTGCAGCCTCA
TET2 p.R1359H	TTCGCATTCACACACACTTTTATTTTTCAGATTGAATATGAACA CAGAGCACCAGAGTGCC[G/A]TCTGGGTCTGAAGGAAGGCCGT CCATTCTCAGGGGTCACTGCATGTTTGGACTTCTGTGCT
SRCAP p.L1779P	CCAGTGGGCCCAGCCCCAGCTCACACGCTGACTTTGGCTCCAG CATCGTCATCTGCTTCAC[T/C]CCTGGCCCCAGCTTCAGTGCAG ACACTGACCTTGAGCCCTGCCCCAGTTCCTACCCTGGGC
FAT1 p.Q249K	CTGTCCAGTTCTGATGGTGACAATGTCACTGCTGTTATCACCG GAGCACATTCATTGGCCT[G/T]TTCGATGTGCACCGTTAGCTTG GCCATGCTGCTGATGCCACTGCTCCCATACAACTTCATG