Clonal hematopoiesis in the donor does not adversely affect long-term outcomes following allogeneic hematopoietic stem cell transplantation: result from a 13-year follow-up

Kyoung Ha Kim,[1,2*] TaeHyung Kim,[1,3,4*] Igor Novitzky-Basso,[1] Hyewon Lee,[1,5] Youngseok Yoo,[1] Jae-Sook Ahn,[4,6] Ivan Pasic,[1] Arjun Law,[1] Wilson Lam,[1] Fotios V. Michelis,[1] Armin Gerbitz,[1] Auro Viswabandya,[1] Jeffrey Lipton,[1] Rajat Kumar,[1] Jonas Mattsson,[7] Zhaolei Zhang,[3,4,8] Nathali Kaushansky,[9] Yardena Brilon,[9] Noa Chapal-Ilani,[9] Tamir Biezuner,[9] Liran I. Shlush[9#] and Dennis Dong Hwan Kim[1,10#]

[1]Division of Medical Oncology and Hematology, Princess Margaret Cancer Center, Toronto, Ontario, Canada; [2]Department of Internal Medicine, Soonchunhyang University College of Medicine, Soonchunhyang University Hospital, Seoul, Korea; [3]Department of Computer Science, University of Toronto, Toronto, Ontario, Canada; [4]The Donnelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada; [5]Division of Rare and Refractory Cancer, Division of Hemato-Oncology, and Center for Hematologic Malignancy Research Institute and Hospital, National Cancer Center, Goyang, Korea; [6]Department of Internal Medicine, Chonnam National University Hwasun Hospital, Chonnam National University, Gwangju, Korea; [7]Gloria and Seymour Epstein Chair in Cell Therapy and Transplantation, Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [8]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; [9]Department of Immunology, Weizmann Institute of Science, Rehovot, Israel and [10]Institute for Medical Science, Faculty of Medicine, University of Toronto, Toronto, Canada

**KHK and TK contributed equally as co-first authors.*
#LS and DK contributed equally as co-senior authors.

Correspondence: D. D. H. Kim
dr.dennis.kim@uhn.ca

L. Shlush
liran.shlush@weizmann.ac.il

Received: July 24, 2022.
Accepted: January 26, 2023.
Early view: February 2, 2023.

https://doi.org/10.3324/haematol.2022.281806

Supplementary Methods

Improved MIP (iMIP) capture protocol

One µl DNA template was added to a hybridization mix together with a MIP pool (final concentration of 0.04pM per probe) in 0.85x Ampligase buffer. Mix was incubated in a thermal cycler at 98℃ for 3 minutes, followed by 85℃ for 30 minutes, 60℃ for 60 minutes and 56℃ for 60 minutes. Product was mixed with (final concentration in brackets): dNTPs (14pM), Betaine (375 mM), NAD+ (1 mM), additional Ampligase buffer (0.5x), Ampligase (total of 1.25U) and Q5 High-Fidelity DNA Polymerase (0.4 U, New England Biolabs). All the product of the hybridization was incubated at 56℃ for 5 minutes followed by 72℃ for 5 minutes. Enzymatic digestion of linear probes was performed by adding Exonuclease I (8U) and Exonuclease III (50U). Mixture was incubated at 37℃ for 10 minutes hours, followed by 80℃ for 20 minutes. Final product was amplified using NEBNext Ultra II Q5 Master Mix (New England Biolabs). Samples were pooled and concentrated using AMPure XP beads at 0.75x volumetric concentration and sequenced as abovementioned described.

Data preprocessing and variant calling

Paired-end 2X151bp sequencing data were converted to fastq format. Reads were merged using BBmerge v38.62 (13) with default parameters, followed by trimming of the ligation and extension arm using Cutadapt v2.10 (14). Unique Molecular Identifiers (UMI) were trimmed and assigned to each read header. Processed reads were aligned using BWA-MEM(15) to a custom reference genome, comprised of the MIP ARCH panel sequences ± 150 bases extracted from broad HG19 [https://gatk.broadinstitute.org/hc/en-us/articles/360035890711-GRCh37-hg19-b37-humanG1Kv37-Human-Reference-Discrepancies#b37].. Aligned files were sorted, converted to BAM (SAMTools V1.9 (16), followed by Indel realignment using AddOrReplaceReadGroups (Picard tools) and later IndelRealigner (GATK v.3.7 (17)). Variant calling was done using mpileup for the single nucleotide variant (SNVs), and Varscan2 v2.3.9 (18) and Platypus v0.8.1 (19) for indels. Variants were annotated using ANNOVAR(20).

Statistical analysis of SNVs for MIPs and amplicon

The depth for reference calls and all possible variants of all positions was retrieved from the mpileup files. Only positions with depth>100 were included. To estimate background error rate at each position first we calculated the total read depth across all samples (DEPTH_SUM) and the alternate supporting reads (ALT_READS_SUM) (Supplementary Tables S3A, S3B). Next, the number of alternate reads in a sample (n) and the total depth for the sample in that position (N) were analyzed followed by the calculate

of m = ALT_READS_SUM – n and M = DEPTH_SUM –N. For MIPs this was done separately on each technical duplicate. To test whether a specific VAF is significantly different from the background error rate we approximated the distribution of the variant using Poisson distribution and used Poisson exact test on each variant estimation (stats R package), and corrected for multiple hypothesis testing with Benjamini Hochberg (BH)(21) test per p-value to get a BH score.

Supplementary Table 1. The list of genes

Gene ID	Chromosome	region
GNB1	**1**	Missense: positions 57 ,76 and 80 (a.a)
MPL	**1**	Indel: exon10
NRAS	**1**	Missense: positions 12 ,13,61 (a.a)
DNMT3A	**2**	Whole gene
IDH1	**2**	Missense: positions 132 (a.a)
SF3B1	**2**	Missense: positions 625 ,666,700 (a.a)
MYD88	**3**	Missense: position 265 (a.a)
KIT	**4**	Missense: position 816 (a.a)
TET2	**4**	Whole gene
BRAF	**5**	Missense: position 600 (a.a)
NPM1	**5**	indel: position 28 8 (a.a)
EZH2	**7**	Whole gene
JAK2	**9**	Missense: position 617 (a.a)
NOTCH1	**9**	Indel: exon 0
SMC3	**10**	Whole gene
CBL	**11**	Whole gene
WT1	**11**	Whole gene
CALR	**12**	indel: positions 360 to 380 (a.a)
FLT3	**12**	Missense: positions 835 (a.a) and exons: 9-11
KRAS	**12**	Missense: positions 12 ,13,61 (a.a)
PTPN11	**12**	Whole gene
IDH2	**15**	Missense: positions 140 ,172 (a.a)
PPM1D	**17**	Whole gene
SRSF2	**17**	Missense: position 95 (a.a)

TP53	**17**	Whole gene
SETBP1	**18**	Missense: position 870 (a.a)
CEBPA	**19**	Whole gene
ASXL1	**20**	Whole gene
GNAS	**20**	Whole gene
RUNX1	**21**	Whole gene
U2AF1	**21**	Missense: position 34 (a.a)
SMC1A	**X**	Whole gene
STAG2	**X**	Whole gene

Supplementary Table 2. Risk factor analysis of overall survival (A), relapse incidence (B) and non-relapse mortality (C)

A. Overall survival

				Univariate analysis		Multivariate analysis		Multivariate analysis with time dependent covariate	
Variable	**Group**	**No of pts**	**OS rate at 10 yrs (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	48.0 (27.8-65.6)	0.969	1.010 [0.599- 1.706]	0.201	0.709 [0.418-1.201]	0.974	1.008 [0.612-1.660]
	No CHIP	347	41.0 (35.7-46.1)		1.000		1.000		1.000
Chronic GVHD	cGVHD	235	56.2 (49.6-62.3)	<0.001	0.238 [0.178-0.319]	< 0.001	0.240 [0.184-0.314]	0.021	0.662 [0.467-0.939]
	No cGVHD	137	16.0 (10.4-22.7)		1.000		1.000		1.000
Age (per decade)		372		0.004	1.173 [1.051-1.310]	0.010	1.155[1.035-1.288]	0.004	1.175 [1.052-1.312]

B. Relapse

				Univariate analysis		Multivariate analysis	
Variable	**Group**	**No of pts**	**Relapse at 10 yrs (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	16.0 (4.8-33.1)	0.600	0.788 [0.323-1.918]	0.420	0.701 [0.299-1.648]
	No CHIP	347	24.4 (20.0-29.0)		1.000		1.000
Chronic GVHD	cGVHD	235	17.1 (12.6-22.3)	<0.001	0.410 [0.272 -0.619]	<0.001	0.413 [0.274-0.622]
	No cGVHD	137	35.5 (27.5-43.7)		1.000		1.000
Conditioning intensity	MAC	267	18.5 (14.1-23.4)	<0.001	0.412 [0.272-0.625]	<0.001	0.453 [0.297-0.692]
	RIC	105	37.5 (28.2-46.8)		1.000		1.000

C. Non-relapse mortality

				Univariate analysis		Multivariate analysis	
Variable	**Group**	**No of pts**	**NRM at 10 yrs (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	36.0 (17.8-54.7)	0.570	1.197 [0.639-2.242]	0.970	0.986 [0.499-1.948]
	No CHIP	347	37.7 (32.6-42.9)		1.000		1.000
Chronic GVHD	cGVHD	235	30.0 (24.2-36.0)	<0.001	0.447 [0.320-0.624]	<0.001	0.392 [0.274-0.559]
	No cGVHD	137	50.6 (41.9-58.7)		1.000		1.000
Age (per decade)		372		0.062	1.132 [0.994-1.289]	0.001	1.029 [1.012-1.047]
Conditioning intensity	MAC	267	40.3 (34.3-46.1)	0.034	1.486 [1.030-2.143]	<0.001	2.605 [1.647-4.121]
	RIC	105	31.0 (22.3-40.2)		1.000		1.000

*Abbreviation: CHIP, clonal hematopoiesis of indeterminate potential; OS, overall survival; HR, hazard ratio; 95% CI, 95% confidence interval; GVHD, graft-versus-host disease; cGVHD, chronic graft versus host disease; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; NRM, non-relapse mortality

Supplementary Table 3. Risk factor analysis of neutrophil (A) and platelet engraftment (B)

A. Neutrophil engraftment

				Univariate analysis		Multivariate analysis	
Variable	**Group**	**No of pts**	**Engraftment at 30 days (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	88.0 (64.0-96.4)	0.460	0.843 [0.534-1.331]	0.550	0.847 [0.493-1.453]
	No CHIP	347	91.6 (88.2-94.1)		1.000		1.000
Conditioning intensity	MAC	267	90.6 (86.4-93.6)		0.661 [0.520-0.840]	0.050	0.778 [0.606-1.000]
	RIC	105	93.3 (86.2-96.9)	<0.001	1.000		1.000
Source of stem cells	PBSC	259	94.6 (91.0-96.8)	<0.001	2.047 [1.667-2.513]	<0.001	2.047 [1.667-2.513]
	Bone marrow	113	84.1 (75.8-0.89.7)		1.000		1.000

B. Platelet engraftment

				Univariate analysis		Multivariate analysis	
Variable	**Group**	**No of pts**	**Engraftment at 30 days (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	72.0 (48.8-86.0)	0.250	0.751 [0.461-1.224]	0.280	0.711 [0.380-1.329]
	No CHIP	347	82.1 (77.7-86.0)		1.000		1.000
Conditioning intensity	MAC	267	76.8 (71.2-81.4)		0.368 [0.273-0.495]	<0.001	0.417 [0.305-0.571]
	RIC	105	93.3 (86.2-96.9)	<0.001	1.000		1.000
Source of stem cells	PBSC	259	87.6 (82.9-91.1)	<0.001	2.220 [1.805-2.730]	<0.001	1.983 [1.604-2.451]
	Bone marrow	113	67.3 (57.7-75.1)		1.000		1.000

*Abbreviation: CHIP, clonal hematopoiesis of indeterminate potential; HR, hazard ratio; 95% CI, 95% confidence interval; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; NRM, non-relapse mortality; PBSC, peripheral blood stem cell

Supplementary Table 4. Risk factor analysis of acute GVHD (A) and chronic GVHD (B)

A. Acute GVHD

				Univariate analysis		Multivariate analysis	
Variable	**Group**	**No of pts**	**Incidence at day 100 (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	80.0 (56.3-91.7)	0.490	1.179 [0.735-1.893]	0.360	1.237 [0.783-1.954]
	No CHIP	347	77.1 (72.2-81.2)		1.000		1.000
HLA disparity	HLA-Mismatch	26	88.5 (65.0-96.6)	0.011	1.730 [1.132-2.644]	0.064	1.517 [0.976-2.356]
	HLA-Match	346	76.4 (71.5-80.6)		1.000		1.000
Conditioning intensity	MAC	267	82.5 (77.2-86.6))	<0.001	1.952 [1.522-2.504]	<0.001	1.740 [1.356-2.233]
	RIC	105	64.1 (54.0-72.6)		1.000		1.000
GVHD prophylaxis	TCD	47	51.4 (36.0-64.8)	<0.001	0.432 [0.285-0.656]	0.004	0.524 [0.336-0.817]
	No TCD	325	81.0 (76.2-84.9)		1.000		1.000

B. Chronic GVHD

				Univariate analysis		Multivariate analysis	
Variable	**Group**	**No of pts**	**Incidence at 3 years (%)**	**p-value**	**HR [95% CI]**	**p-value**	**HR [95% CI]**
Donor CHIP	CHIP	25	48.0 (26.9-66.3)	0.22	0.685 [0.371-1.261]	0.310	0.731 [0.400-1.334]
	No CHIP	347	64.0 (58.7-68.8)		1.000		1.000
Source of stem cells	PBSC	259	68.7 (62.7-74.0)	<0.001	1.720 [1.277-2.317]	<0.001	1.678 [1.244-2.262]
	Bone marrow	113	49.6 (39.9-58.4)		1.000		1.000
GVHD prophylaxis	TCD	47	40.4 (26.2-54.2)	0.005	0.502 [0.309-0.816]	0.009	0.523 [0.320-0.856]
	No TCD	325	66.2 (60.7-71.0)		1.000		1.000

*Abbreviation: GVHD, graft-versus-host disease; CHIP, clonal hematopoiesis of indeterminate potential; HR, hazard ratio; 95% CI, 95% confidence interval; HLA**,** human leukocyte antigen; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; TCD, T-cell depletion; PBSC, peripheral blood stem cell

Supplementary Table 5. Comparison of organ involvement by GVHD according to the presence of donor CHIP

No of pts (%)	CHIP in donor (n=25)	No CHIP in donor (n=347)	p-value
Organ involvement acute GVHD			
Skin involvement	15 (60.0)	220 (63.4)	0.831
Gut involvement	7 (28.0)	127 (36.6)	0.518
Liver involvement	10 (40.0)	159 (45.8)	0.679
Organ involvement of chronic GVHD			
Skin involvement	11(44.0)	171(49.3)	0.681
Mouth involvement	8 (32.0)	130 (37.5)	0.672
Eye involvement	2 (8.0)	74 (21.3)	0.129
Lung involvement	3 (12.0)	49 (14.1)	1.000
GI tract involvement	6 (24.0)	78(22.5)	0.808
Liver involvement	11 (44.0)	166 (47.8)	0.836
Musculoskeletal system involvement	3 (12.0)	46(13.3)	1.000

*Abbreviation: CHIP, clonal hematopoiesis of indeterminate potential; GVHD, graft-versus-host disease

Supplementary Table 6. Clinical and transplantation characteristics of donors and recipients according to development secondary malignancies

No of pts (%)	SM (n=56)	No SM (n=316)	p-value
Donor age in years, range	51 (11-75)	48(16-71)	0.083
Donor gender,			
Male	25 (51.0)	120 (48.0)	0.756
Female	24 (49.0)	130 (52.0)	
Recipient age in years, range	52 (26-66)	47 (17-71)	0.091
Recipient gender			
Male	35 (62.5)	188 (59.5)	0.768
Female	21 (37.5)	128 (40.5)	
Diagnosis			
Aplastic anemia	2 (3.6)	14 (4.4)	0.297
AML/MDS/MPN	16 (28.6)/3 (5.4)/3 (5.4)	122 (38.6) /25 (7.9)/17 (5.4)	
ALL/CLL/NHL	5 (8.9)/6 (10.7)/12 (21.4)	46 (14.6)/19 (6.0)/35 (11.1)	
CML/MM/Others	8 (14.3)/1 (1.8)/ 0 (0)	31 (9.8)/3 (0.9)/ 4 (1.3)	
Conditioning intensity			
MAC	37 (66.1)	230 (72.8)	0.334
RIC	19 (33.9)	86 (27.2)	
Source of stem cell			
Bone marrow	14 (25.0)	99 (31.3)	0.431
PBSC	42 (75.0)	217 (68.7)	
Donor type			
Matched related donor	46 (82.1)	226 (71.5)	0.275
Matched unrelated donor	7 (12.5)	67 (21.2)	
Alternative donor	3 (5.4)	23 (7.3)	
TBI			
No TBI	17 (30.4)	72 (22.8)	0.236
TBI	39 (69.6)	244 (77.2)	
aGVHD			
No	10 (17.9)	76 (24.1)	0.390
Yes	46 (82.1)	240 (75.9)	
cGVHD			
No	4 (7.1)	133 (42.1)	<0.001
Yes	52 (92.9)	183 (57.9)	
GVHD prophylaxis			
No TCD	49 (87.5)	276 (87.3)	1.000
TCD	7 (12.5)	40 (12.7)	
CHIP			
No	54 (96.4)	293 (92,7)	0.398
Yes	2 (3.6)	23 (7.3)	

*Abbreviation: SM, secondary malignancy; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; MPN, myeloproliferative neoplasia; ALL, acute lymphoblastic leukemia; CLL, chronic lymphocytic leukemia; NHL non-Hodgkin's lymphoma; MM, multiple myeloma; CML, chronic myeloid leukemia; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; PBSC, peripheral-blood stem cell; TBI, total body irradiation; aGVHD, acute graft versus host diseases; cGVHD, chronic graft versus host disease; TCD,T cell depletion; CHIP, clonal hematopoiesis of indeterminate potential.

Supplementary Table 7A. Summary and subtypes of diagnosed secondary malignancies after allogeneic HCT segregated by presence of donor CHIP

No of pts (%)	Total Patients (n=372)	CHIP in donor (n=25)	No CHIP in donor (n=347)
Overall cases of SM	56 (15.1)	2 (8.0)	54 (15.6)
Skin, SCC	16 (4.3)	-	16 (4.6)
Skin, BCC	11 (3.0)	2 (8.0)	9 (2.6)
Hematological cancer	5 (1.3)	-	5 (1.4)[1]
Prostate cancer	5 (1.3)	-	5 (1.4)
Lung cancer	5 (1.3)	-	5 (1.4)
Head and neck cancer	4 (1.1)	-	4 (1.1)
PTLD	2 (0.5)	-	2 (0.6)
Melanoma	2 (0.5)	-	2 (0.6)
Cervical cancer	1 (0.3)	-	1 (0.3)
Kidney cancer	1 (0.3)	-	1 (0.3)
Pancreatic cancer	1 (0.3)	-	1 (0.3)
Colon cancer	1 (0.3)	-	1 (0.3)
Esophageal cancer	1 (0.3)	-	1 (0.3)
Breast cancer	1 (0.3)	-	1 (0.3)

*Abbreviation: HCT, hematopoietic stem cell transplantation; CHIP, clonal hematopoiesis of indeterminate potential; SM, secondary malignancy, SCC, squamous cell carcinoma; BCC, basal cell carcinoma; PTLD, post-transplant lymphoproliferative disorder

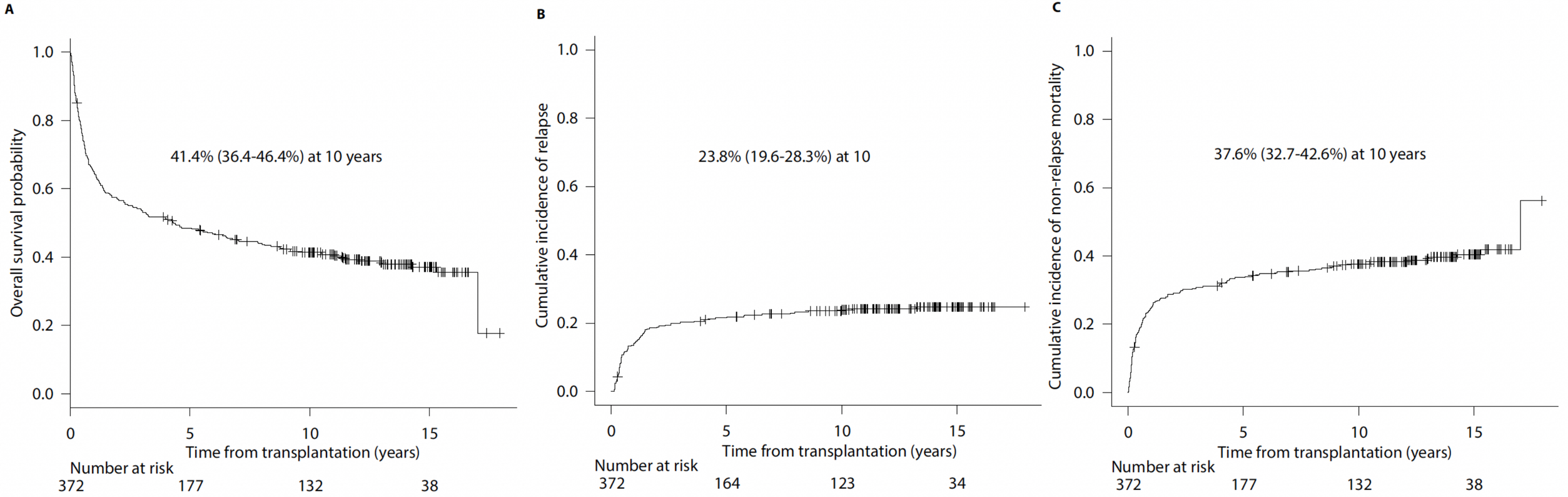
1. These 5 patients include a patient diagnosed MDS at 17.4 years post-HCT with a primary diagnosis of CML, a patient diagnosed as AML at 6.8 years post-HCT with a primary diagnosis of MM, a patient diagnosed as ALL at 7.9 years post-HCT with a primary diagnosis of NHL (follicular lymphoma), a patient diagnosed as AML at 10.6 years post-HCT with a primary diagnosis of Waldenstrom's macroglobulinemia and a patient diagnosed as MDS at 10 years post-HCT with a primary diagnosis of AML.

Supplementary Table 7B. Patients with secondary malignancies in donor CHIP group

SM subtype	Recipient gender	Recipient diagnosis	Recipient age	Recipient mutation	Donor age	Donor gender	Donor type	Donor mutation	Time to SM dx (year)	death
Skin, BCC	Female	AML	54	*DNMT3A, TET2*	52	F	MSD	*TET2*	13.0	alive
Skin, BCC	Male	NHL	51	negative	56	M	MSD	*EZH2*	11.0	alive

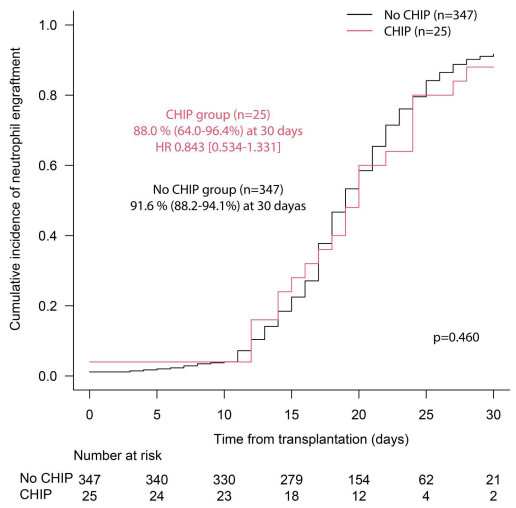
*Abbreviation: CHIP, clonal hematopoiesis of indeterminate potential; SM, secondary malignancy; BCC, basal cell carcinoma; NHL non-Hodgkin's lymphoma; AML, acute myeloid leukemia

Supplementary Figure 1

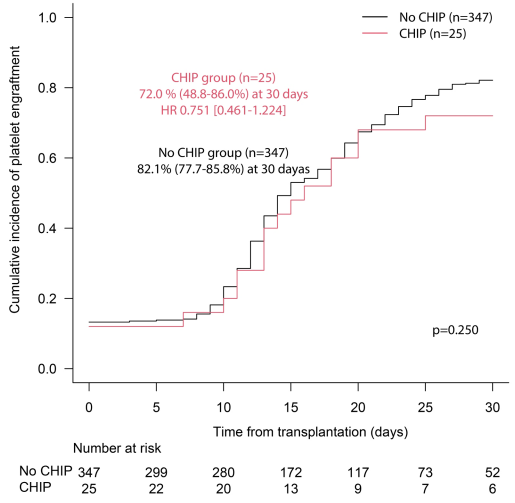


Supplementary Figure 2

A

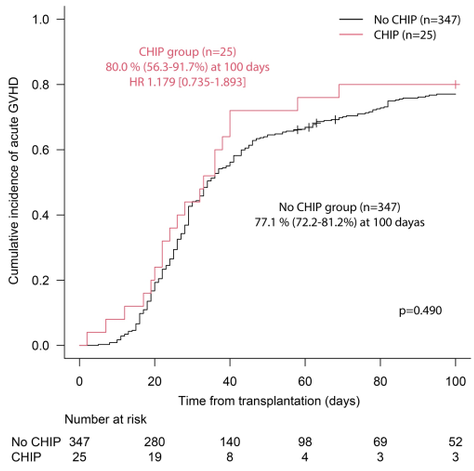


B

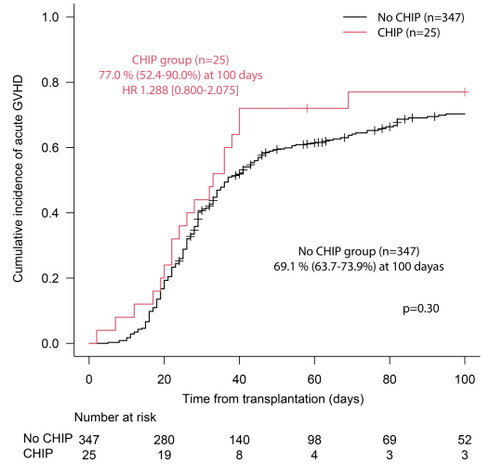


Supplementary Figure 3

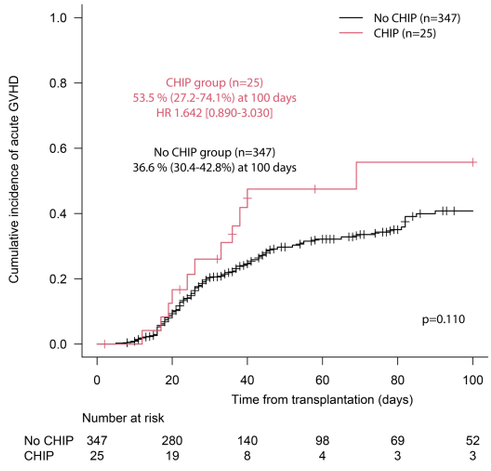
A



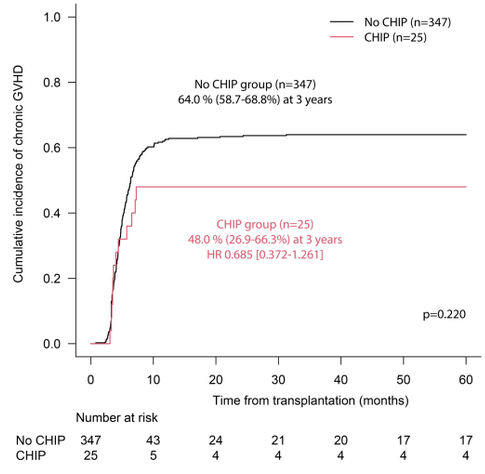
B



C



D



Supplementary Figure 4