Duplex Sequencing reveals ubiquitous clonal hematopoiesis and complex donor-recipient clonal dynamics following hematopoietic stem cell transplant

Masumi Ueda[1], Jake Higgins[2], Thomas Smith[2], Elizabeth Schmidt[2], Fang Yin Lo[2], Charles C. Valentine III[2], Jesse Salk[2], Rainer Storb[1], and Jerald P. Radich[1]

[1]Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA - [2]TwinStrand Biosciences Inc., Seattle, WA

TWINSTRAND™ BIOSCIENCES

FRED HUTCH CURES START HERE™

Introduction

Clonal hematopoiesis (CH aka CHIP) is the process whereby otherwise healthy individuals accumulate clones bearing low-frequency somatic mutations in hematologic malignancy (HM) driver genes. Historically thought to be a phenomenon of the elderly that is associated with increased risk of HM and cardiovascular disease, CH has been observed in increasingly younger cohorts as sequencing technology has advanced. CH in hematopoietic stem cell transplant (HSCT) donors has also been associated with adverse recipient outcomes.

Duplex Sequencing (DS) is an error-corrected sequencing method that generates double-stranded consensus sequences to virtually eliminate PCR and sequencing errors to enable single-molecule resolution. Here we use DS to refine the definition of CH as requiring the same alternate allele call from 2 or more unique DNA molecules, rather than a variant allele frequency (VAF) cutoff.

Our retrospective study examines CH in donor-recipient pairs surviving up to 45 years post-HSCT and represents one of the longest follow-up of studies of this kind. Using these unique samples, we sought to characterize the acquisition of somatic mutations over time, assess the impact of HSCT on these clonal dynamics, and investigate factors that may modulate clonal dynamics after HSCT.

Sequencing Errors Obscure Truth

Next-Generation Sequencing (NGS)

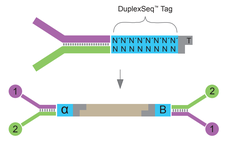
Single Strand Error-Corrected NGS

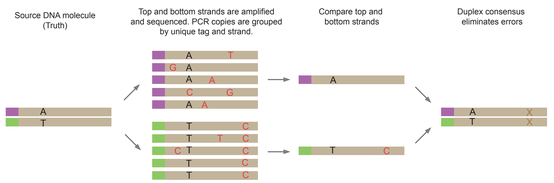
Duplex Sequencing

TwinStrand Duplex Sequencing™ Technology

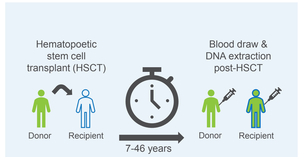
A DuplexSeq™ Adapter has:

1. Identical (or relatable) degenerate tags in each strand.
2. An asymmetry allowing independent strand identification.





Study Cohort



Pair ID	Disease	Stem Cell Source	Donor age at HSCT	Recipient age at HSCT	Years HSCT to sampling	"Hematopoietic age"
1*	MDS	PB	84	48	7	91
1574*	CML	BM	62	35	26	88
32513*	AML	PB	71	42	8	79
1408*	CML	BM	43	30	32	75
13476*	CML	BM	51	45	20	71
3459*	AA	BM	35	3	32	67
136	PNH	BM	18	23	46	64
2707*	ALL	BM	30	24	34	64
1409	CML	BM	27	28	36	63
1956*	CML	BM	25	23	35	60
1115	AA	BM	19	22	38	57
978	AML	BM	14	16	39	53
896	AA	BM	13	12	39	52
1329*	AA	BM	12	14	37	49
3280*	ALL	BM	16	19	32	48
2514*	ALL	BM	12	16	34	46

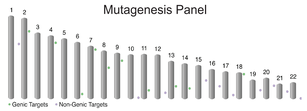
Blood samples were collected from 16 donor-recipient pairs 7-46 years post-HSCT. An additional donor sample was collected at the time of HSCT for 11/16 pairs (*). Patients were being treated for the following hematologic diseases: myelodysplastic syndrome (MDS), chronic myeloid leukemia (CML), acute myeloid leukemia (AML), aplastic anemia (AA), paroxysmal nocturnal hemoglobinuria (PNH) or acute lymphocytic leukemia (ALL). All individuals were healthy at time of sampling.

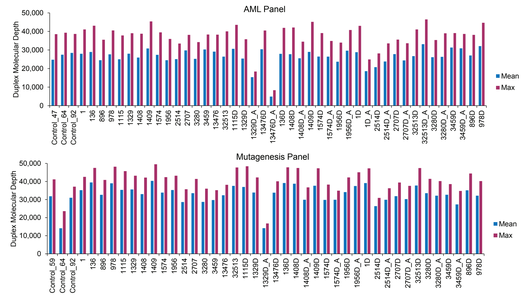
Probe Panels & Duplex Sequencing

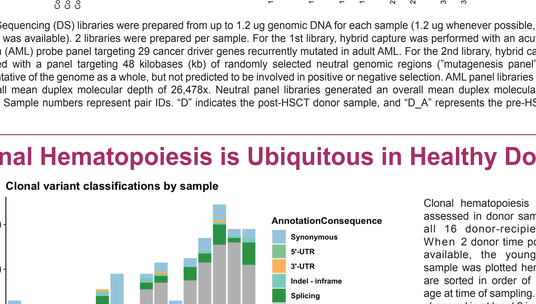
AML Panel

ASXL1	FLT3	KRAS	PTEN	STAG2
CBL	GATA2	PHF6	PTPN11	TET2
CEBPA	HNRNPK	SMC3	RAD21	TP53
DNMT3A	IDH1	NPM1	RUNX1	U2AF1
EZH2	IDH2	NRAS	SMC1A	WT1
FAM5C (BRINP3)	MLL-X (KMT2A-X)	MYH11-CBFB	KIT	

Mutagenesis Panel

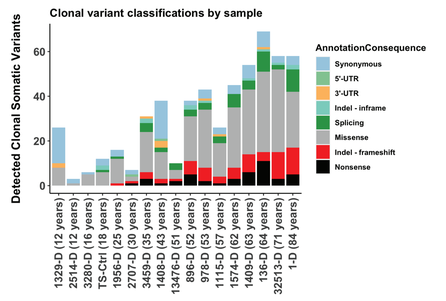






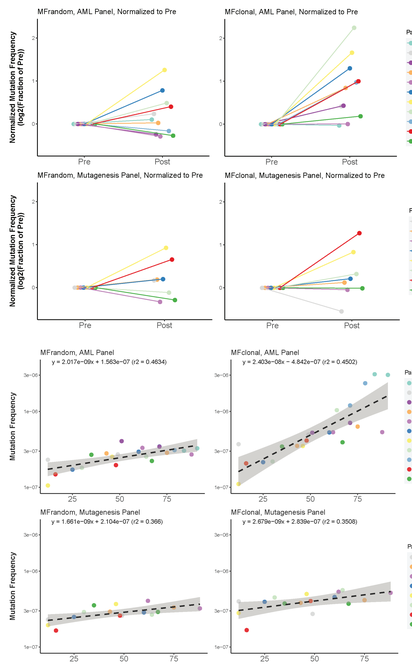
Duplex Sequencing (DS) libraries were prepared from up to 1.2 ug genomic DNA for each sample (1.2 ug whenever possible, or else as much as was available). 2 libraries were prepared per sample. For the 1st library, hybrid capture was performed with an acute myeloid leukemia (AML) probe panel targeting 29 cancer driver genes recurrently mutated in adult AML. For the 2nd library, hybrid capture was performed with a panel targeting 48 kilobases (kb) of randomly selected neutral genomic regions ("mutagenesis panel") that are representative of the genome as a whole, but not predicted to be involved in positive or negative selection. AML panel libraries generated an overall mean duplex molecular depth of 26,478x. Neutral panel libraries generated an overall mean duplex molecular depth of 32,812x. Sample numbers represent pair IDs. "D" indicates the post-HSCT donor sample, and "D_A" represents the pre-HSCT donor sample.

Clonal Hematopoiesis is Ubiquitous in Healthy Donors



Clonal hematopoiesis (CH) was assessed in donor samples from all 16 donor-recipient pairs. When 2 donor time points were available, the younger donor sample was plotted here. Donors are sorted in order of increasing age at time of sampling. Mutations observed in at least 2 independent DNA molecules are plotted and annotated with the predicted functional consequence in the canonical transcript. Every sample harbors between several and several dozen CH mutations. The majority are missense. Additional intronic mutations are not shown.

CH Clones Expand in Healthy Donors over Time

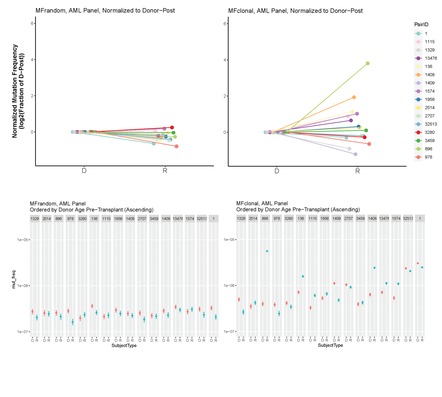


Mutation frequency (MF) was calculated 2 different ways for all samples. For MF_{random}, the total number of unique somatic mutations was divided by the total number of duplex error-corrected bases. For MF_{clonal}, the total number of mutant allele counts was divided by the total number of duplex error-corrected bases. MF_{random} reflects the overall number of different mutations present in a sample, while MF_{clonal} can reflect large or expanding clones. AML and mutagenesis panels are shown for each MF method.

The upper 4 panels show donor MF pre- and post-HSCT, regardless of time since transplant, and normalized to pre-HSCT MF. The y-axis is log2-transformed. The lower 4 panels represent the actual MF and donor age at pre- and post-HSCT time points for each donor.

MF_{clonal} for the AML panel showed the greatest overall increase in individual donors over time, as well as with overall donor age, suggesting widespread clonal expansion of driver gene mutations. MF_{random} for the AML panel and MF_{random} and MF_{clonal} for the mutagenesis panel showed a slight overall increase with donor age, but showed inconsistent patterns over time on a per-donor basis.

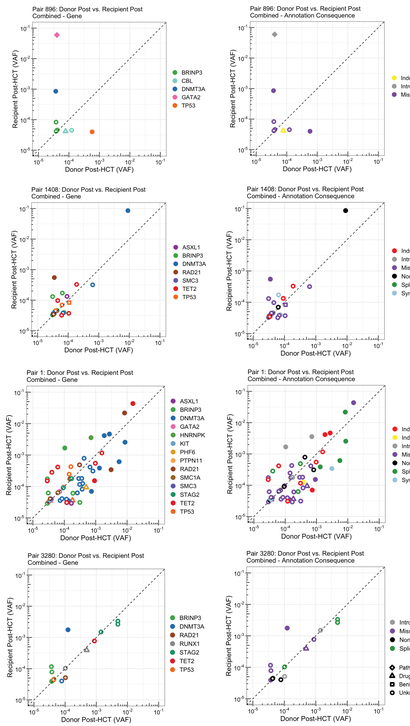
Most HSCT Recipients do not Exhibit Broad Clonal Expansions Relative to Donors



MF_{random} and MF_{clonal} were calculated for all post-HSCT donors (D) and paired recipients (R). In the upper 2 panels, MF was normalized to post-HSCT donor and log2-transformed. The lower 2 panels represent the actual MF for each sample, with 95% Wilson binomial confidence intervals. Only the AML panel is shown for the post-HSCT samples.

There is little change in MF_{random} in recipients relative to donors, suggesting a similar number of total unique mutations. There is not a consistent increase in MF_{clonal} in recipients, suggesting that most CH clones do not preferentially expand in recipients post-HSCT. Recipients in pairs 1408 and 896 have the largest increase in MF_{clonal} relative to post-HSCT donor.

Specific Clones Expand in HSCT Recipients



All individual mutations shared between representative post- HSCT donor-recipient pairs are shown in scatterplots. Genes are color-coded in the left column of plots, and annotation consequences in the canonical transcript are color-coded in the right column of plots (each donor-recipient pair is represented by 2 plots). Shapes indicate ClinVar predicted pathogenicity. Filled in shapes represent mutations with significantly increased or decreased VAF in the recipient relative to the donor post-HSCT (Fisher's exact test, Bonferroni-adjusted $p \leq 0.05$). Hollow shapes are not significantly different in the recipient.

Pair 896, which has substantially higher overall MF_{clonal} in the recipient, has 2 individual mutations that increase significantly in the recipient, but also 1 that significantly decreases.

The recipient in pair 1408, which also has elevated MF_{clonal} relative to the donor, has 2 mutations that significantly increase. But there are many mutations at low VAF that do not significantly change.

In pair 1, which has similar MF_{clonal} in both post-HSCT samples, an equal number of mutations significantly increase or decrease in the recipient.

Pair 3280 has similar MFclonal in the post-HSCT donor and recipient. A single mutation is significantly increased in the recipient, and the rest of the shared variants do not change significantly.

Conclusions

- Clonal hematopoiesis (CH) is observed in 100% of individuals, with multiple clonal somatic mutations observed even in teenagers.
- We propose redefining CH as mutations observed in ≥2 DNA molecules, rather than using a variant allele frequency (VAF) cutoff.
- The number of CH mutations increases with age.
- CH clones expand over time in healthy donors.
- In general, CH clones do not preferentially expand in HSCT recipients relative to donors, though there are outliers.
- Duplex Sequencing will be a valuable research tool for reassessing the risk of hematologic malignancies and cardiovascular disease associated with CH. Rather than a binary assignment of CH-positive or -negative, it is likely that specific genes, mutations and VAFs will need to be considered.

Contact: https://twinstrandbio.com/contact/
jhiggins@twinstrandbio.com