Clonal hematopoiesis of indeterminate potential among cancer survivors exposed to myelotoxic chemotherapy

Authors: Adam J Olszewski, MD [1,2]
Anna D Chorzalska, PhD [1,2,3]
Annette S Kim, MD, PhD [4]
Peter J Quesenberry, MD [1,2]
Mary L Lopresti, DO [1,2]
Mary Anne Fenton, MD [1,2]
John L Reagan, MD [1,2]
James N Butera, MD [1,2]
Ilyas Sahin, MD [1,2]
Celine Hamel, CRA [2]
Jordan Robison, CRA [2]
Max Petersen [1,3]
Patrycja M Dubielecka, PhD [1,2,3]

Affiliations
[1] The Warren Alpert Medical School of Brown University, Providence, RI
[2] Division of Hematology-Oncology, Rhode Island Hospital, Providence, RI
[3] Signal Transduction Lab, Rhode Island Hospital, Providence, RI
[4] Department of Pathology, Brigham and Women's Hospital, Boston, MA

Corresponding author: Adam J Olszewski, MD
Rhode Island Hospital
Address: 593 Eddy St
Providence, RI, 02903
Phone: (401) 444 5435
Email: adam_olszewski@brown.edu

Supplementary Methods.

Patient identification and clinical data. Eligible patients were identified by screening outpatient visits in the oncology clinic at Rhode Island Hospital, between October 18, 2017, and July 5, 2018. Patients were eligible if they had completed curative therapy for breast cancer or and aggressive lymphoma (diffuse large B-cell lymphoma, Burkitt lymphoma, peripheral T-cell lymphoma, or Hodgkin lymphoma), which included adjuvant or primary chemotherapy containing an anthracycline or an alkylating agent, if they were 45 to 75 years of age, had no known hematologic malignancy (other than the primary lymphoma) or unexplained cytopenias, and if they were clinically free of cancer at the time of assessment. All subjects provided a written informed consent. The study was approved by the Institutional Review Board at Rhode Island Hospital. Clinical data (including pathology report, prior blood counts, and record of chemotherapy) were extracted from the medical records by the investigators.

Sample collection and processing. Patients provided 4-10 mL of whole blood collected in tubes containing potassium ethylenediaminetetraacetic acid (K2-EDTA), from which a complete blood count was obtained. We isolated cells of hematopoietic origin expressing leukocyte common antigen (CD45) using Ficoll-Paque™ PLUS gradient centrifugation followed by CD45 MicroBead kit and QuadroMACS™ separation (Miltenyi Biotec, Inc., San Diego, CA). The highly enriched CD45-positive cell population was subjected to genomic DNA extraction using QIAmp DNA Micro kit (Qiagen GmbH, Germany) according to the manufacturer's protocol. The purity of gDNA was assessed by measuring 260/280 (values between 1.8-2.0) and 260/230 ratio (values between 1.7-2.2).

Sequencing analysis. Library preparation utilized the Illumina TruSeq Custom Amplicon chemistry with custom designed primers that result in 1330 amplicons with an average length of approximately 250 bp, covering 757 exons of 95 different genes, as previously described.(Kluk et al, 2016) The libraries from each of 32 samples (30 patient samples, 1 pooled cell line control and 1 negative control) are pooled using i5 and i7 sample indices and run at 12 pM (concentration of the combined library) on MiSeq v2.5.0.5, with a targeted cluster density of 1100 k/mm^2 and a cluster passing filter rate of >80%. gDNA samples from isolated CD45-positive cells were sequenced alone, without any matched samples.

Paired end-reads of 150 bp in length were de-multiplexed and individual BAM files were aligned to hg19 by the Illumina MiSeq Reporter software v2.5.1 with the maximum indel size changed to 300 nucleotides. The genomic vcf files were filtered for variants that did not contain >9 variant reads (regardless of variant allele fraction, VAF) or positions at which the variant read count was 5 to 9 but the VAF was not >33%. Assay-specific noise was filtered through comparison to the run controls and plate-wide statistics. A must call list was utilized to identify any filtered variants at key mutational hotspots for manual review and potential recovery.

Mutation calling. For this study, a cut-off of 2% VAF was employed unless there was a key pathogenic variant at a lower VAF that was deemed by the pathologist to be sufficiently above the background noise on a per nucleotide basis. Internal tandem duplications in FLT3 were assessed in all cases through three mechanisms: 1) MiSeq Reporter software (for short ITDs in the center of an aligned read), 2) unaligned reads where the pair was mapped to FLT3 exons 14 and 15 were assessed for duplications of >10 bp and then evaluated for alignment to FLT3 (for detection of large ITDs preventing alignment of the paired read), and 3) manual review of exons 14 and 15 of FLT3 (for the detection of potential ITDs that clipped the end of the aligned read). Read count analysis was performed through calculation of the fractional read count of each amplicon (amplicon read count/total read count) which was divided by the fractional read count of the corresponding normal control amplicon. The log2 of the sample/normal ratio was then normalized to 0.

Statistical analysis. Distribution of variables was described as percentages, means and standard deviations, or medians with interquartile range (IQR), as appropriate. The association between presence of CHIP, or the VAF of CHIP-associated mutations, and a pre-defined set of clinical variables, was examined in univariate generalized linear models (GLM) using a canonical link and binomial distribution (for binary endpoint) or Poisson distribution (for proportion) with robust standard error. When studying VAF of mutations, we used a random effect to account for clustering of standard errors within patients. All analyses were conducted using Stata/MP 15.1 (College Station, TX).

References

Kluk, M.J., Lindsley, R.C., Aster, J.C., Lindeman, N.I., Szeto, D., Hall, D. & Kuo, F.C. (2016) Validation and Implementation of a Custom Next-Generation Sequencing Clinical Assay for Hematologic Malignancies. *J Mol Diagn,* **18,** 507-515.

Supplemental Table S1. List of genes and exons included in Brigham and Women's Rapid Heme Panel next-generation sequencing assay used in the study.

Gene	Exon	Gene	Exon
ABL1	e2-e10	*LUC7L2*	e3-e11
ASXL1	e1-e13	*MAP2K1*	e2-e3
ATM	e2-e63	*MEF2B*	e3
BCL11B	e4	*MPL*	e10
BCOR	e2-e15	*MYD88*	e5
BCORL1	e1-e12	*NOTCH1*	e24-e28, e34
BRAF	e15	*NOTCH2*	e24-e28, e34
BRCC3	e3-e11	*NOTCH3*	e25-e26, e33
CALR	e9	*NPM1*	e10-e11
CBL	e7-e8	*NRAS*	e2-e5
CBLB	e9-e11	*NT5C2*	e9, e11, e13, e15, e17
CD79B	e5-e6	*PAX5*	e3, e6-e7
CEBPA	e1	*PDGFRA*	e10-e21, e23
CNOT3	e1-e2	*PDS5B*	e3-e35
CREBBP	e2-e21, e23-31	*PHF6*	e2-e10
CRLF2	e6	*PIGA*	e2-e6
CSF1R	e22	*PIK3CA*	e2, e10, e21
CSF3R	e14-e18	*PIM1*	e1-e6
CTCF	e3-e12	*PRPF40B*	e2-e26
CTNNB1	e2-e4	*PRPF8*	e2-e43
CUX1	e1-e21	*PTEN*	e1-e9
CXCR4	e2	*PTPN11*	e1-e15
DNMT3A	e2-e23	*RAD21*	e2-e14
DNMT3B	e2-e23	*RET*	e7
EED	e1-e12	*RIT1*	e1-e6
EGFR	e18-e21	*RPL10*	e5
EP300	e18-e27	*RUNX1*	e2-e9
ETV6	e1-e8	*SETBP1*	e4
EZH2	e2-e8, e11-e20	*SETD2*	e1-e4, e6-e21
FANCL	e1-e14	*SF1*	e1-e10, e13
FBXW7	e8-e12	*SF3A1*	e1-e2, e5-e16
FLT3	e14-e16, e20	*SF3B1*	e12-e16
GATA1	e2-e6	*SH2B3*	e2-e8
GATA2	e2-e6	*SMC1A*	e1-e25
GATA3	e4-e6	*SMC3*	e2-e29
GNAS	e8-e9	*SRSF2*	e1
GNB1	e5-e6	*STAG2*	e3-e35
IDH1	e4	*STAT3*	e2-e17, e21-23
IDH2	e4	*TET2*	e3-e11
IKZF1	e2-e8	*TLR2*	e1
IKZF2	e1-e8	*TP53*	e2-e11
IKZF3	e1-e8	*U2AF1*	e2, e6
IL7R	e6	*U2AF2*	e1-e12
JAK1	e10-e25	*WHSC1*	e17-e18
JAK2	e12, e14	*WT1*	e1-e10
JAK3	e11-e24	*XPO1*	e15-e16
KIT	e8-9, e11, e17	*ZRSR2*	e1-e11
KRAS	e2-e5		

Supplemental Table S2. Clinical characteristics of the study cohort.

Variable	Distribution	
N	80	
Age, mean (SD)	62	(7)
Age, median (range)	62	(47-75)
Female sex, *N* (%)	62	(78)
Race / ethnicity, *N* (%)		
White non-Hispanic	70	(88)
Hispanic	3	(4)
Black	3	(4)
Asian	3	(4)
Unspecified	1	(1)
Cancer type, *N* (%)		
Breast	46	(57)
Lymphoma [a]	34	(43)
Chemotherapy regimen, *N* (%)		
RCHOP	29	(36)
AC	21	(26)
AC + paclitaxel	12	(15)
Docetaxel + cyclophosphamide	9	(11)
ABVD	3	(4)
TCH or PTCH	3	(4)
AC + bevacizumab	1	(1)
CHOP	1	(1)
DA-EPOCH-R	1	(1)
Any anthracycline	68	(85)
Blood counts, median (IQR)		
White blood cells ($x10^3/mm^3$)	6.1	(4.9-7.8)
Hemoglobin (g/dL)	13.4	(12.6-14.4)
Platelet count ($x10^3/mm^3$)	229	(185-262)

[a] Eligible histologies included diffuse large B-cell lymphoma, high grade B-cell lymphoma, classical Hodgkin lymphoma, and peripheral T-cell lymphoma.

ABVD: doxorubicin, bleomycin, vinblastine, and dacarbazine; AC: doxorubicin and cyclophosphamide, DA-EPOCH-R: dose-adjusted etoposide, prednisone, vincristine, cyclophosphamide, doxorubicine, and rituximab; IQR: interquartile range; RCHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; SD: standard deviation; TCH: trastuzumab, carboplatin, docetaxel; PTCH: pertuzumab, trastuzumab, carboplatin, and docetaxel.

Supplemental Table S3. List of CHIP-defining mutations identified in the study.

Gene	Accession	cDNA change	Protein change	VAF (%)	Number of reads	COSMIC ID
ASXL1	NM_015338	c.2119_2120delAC	p.H706fs	7.1	225	NA
ASXL1	NM_015338	c.2323delT	p.R774fs	20.7	2286	6908493
ASXL1	NM_015338	c.2467delT	p.T822fs	2.2	1411	4385107
BCORL1	NM_021946	c.3158_3159insG	p.K1053fs	3.6	450	NA
DNMT3A	NM_175629	c.1489T>A	p.C497S	5.7	299	NA
DNMT3A	NM_175629	c.1685G>T	p.C562F	6.4	250	NA
DNMT3A	NM_175629	c.2092T>A	p.W698R	18.4	593	NA
DNMT3A	NM_175629	c.2195T>A	p.F732Y	3.1	781	NA
DNMT3A	NM_175629	c.2202delT	p.F735fs	5.0	1061	NA
DNMT3A	NM_175629	c.2339T>C	p.I780T	9.0	667	1583121
DNMT3A	NM_175629	c.2342A>G	p.D781G	3.8	1164	1318941
DNMT3A	NM_175629	c.2357C>A	p.S786*	5.5	943	NA
DNMT3A	NM_175629	c.2521A>T	p.K841*	21.4	1821	NA
DNMT3A	NM_175629	c.2644C>A	p.R882S	3.1	1505	87001
DNMT3A	NM_175629	c.2645G>A	p.R882H	3.6	419	52944
DNMT3A	NM_175629	c.2665_2666delTG	p.L888fs	3.7	625	NA
GNAS	NM_080425	c.2531G>A	p.R201H	11.6	370	94388
KIT †	NM_000222	c.2452A>G	p.K818E	1.4	1189	NA
SRSF2	NM_001195427	c.284C>A	p.P95H	2.6	271	211504
SRSF2	NM_001195427	c.284C>A	p.P95H	9.5	148	211504
TET2	NM_001127208	c.1552_1553insTT	p.I518fs	29.9	1439	NA
TET2	NM_001127208	c.944delC	p.S315fs	2.0	3010	4383771
TET2	NM_001127208	c.4094G>T	p.G1365V	8.0	610	5985194
TET2	NM_001127208	c.4367delT	p.M1456fs	22.2	522	NA
TP53	NM_000546	c.727A>G	p.M243V	3.3	364	44844
ZRSR2	NM_005089	c.313A>T	p.R105*	25.3	170	NA

† Mutation was called with a VAF of <2%, as it was present in >30 reads and the patient had an additional pathogenic *SRSF2* mutation.

Supplemental Figure S1. Distribution of CHIP-associated pathogenic mutations among patients with breast cancer or lymphoma; *P*-values for the 4 top mutated genes are derived from exact Fisher's test.