Shared cell of origin in a patient with Erdheim-Chester disease and acute myeloid leukemia

Armin Ghobadi,[1,] Christopher A. Miller,[1,2*] Tiandao Li,[2] Michelle O'Laughlin,[2] Yi-Shan Lee,[3] Mohga Ali,[3] Peter Westervelt,[1] John F. DiPersio[1] and Lukas Wartman[1]*

[1]Division of Oncology, Department of Medicine, Washington University School of Medicine; [2]McDonnell Genome Institute, Washington University School of Medicine and [3]Department of Pathology, Washington University School of Medicine, St Louis, MO, USA

**AG and CAM co-first authors.*

Correspondence: ARMIN GHOBADI.
arminghobadi@wustl.edu.
doi:10.3324/haematol.2019.217794

Supplementary Appendix

Shared cell of origin in a patient with Erdheim Chester disease and acute myeloid leukemia

Armin Ghobadi, Christopher A. Miller, Tiandao Li, Michelle O'Laughlin, Yi-Shan Lee, Mohga Ali, Peter Westervelt, John F. DiPersio, and Lukas Wartman,

CONTENTS

METHODS:

Exome sequencing data were aligned to reference sequence GRCh37-lite-build37 using BWA-mem [Li, H. arXiv:1303.3997] version 0.7.10 (params: -t 8), then merged and deduplicated using picard version 1.113 (https://broadinstitute.github.io/picard/). Somatic variants were called from the combined data using our Genome Modeling System[1] as follows:

SNVs were detected using the union of four callers: 1) samtools[2] version r982 (params: mpileup -BuDs) intersected with Somatic Sniper[2] version 1.0.4 (params: -F vcf –G -L -q 1 -Q 15) and processed through false-positive filter v1 (params: --bam-readcount-version 0.4 --bam-readcount-min-base-quality 15 --min-mapping-quality 40 --min-somatic-score 40), 2) VarScan[3] version 2.3.6 filtered by varscan-high-confidence filter version v1 and processed through false-positive filter v1 (params: --bam-readcount-version 0.4 --bam-readcount-min-base-quality 15), 3) Strelka[4] version 1.0.11 (params: isSkipDepthFilters = 1), and 4) Mutect[5] v1.1.4.

Indels were detected using the union of 4 callers: 1) GATK[6] somatic-indel (version 5336), 2) pindel[7] version 0.5 filtered with pindel somatic calls and VCF filters (params: --variant-freq-cutoff=0.08), and pindel read support, 3) VarScan version 2.3.6 filtered by varscan-high-confidence- indel version v1 and 4) Strelka version 1.0.11 (params: isSkipDepthFilters = 1).

SNVs and Indels were further filtered by requiring 20x sequence coverage, removing artifacts found in a panel of 905 normal exomes, removing sites that exceeded 0.1% frequency in the 1000 genomes or NHLBI exome sequencing projects, and then using a Bayesian classifier (https://github.com/genome/genome/blob/master/lib/perl/Genome/Model/Tools/Validation/IdentifyOutliers.pm) and retaining variants classified as somatic with a binomial log-likelihood of at least 10 in the AML sample, and 5 in the ECD (due to lower purity). Finally, all calls were manually reviewed to remove false positive variants caused by read mapping issues, slippage at homopolymer runs, and other sequencing artifacts.

Clonal inference was performed using sciClone.[8]

RESULTS

Table S1: Mutations identified via exome sequencing. Mutations in up- or down-stream regions are annotated with the nearest gene.

Class	Chr	Pos	Reference	Variant	Gene name	Transcript name	AA change	Effect	Called in	Nrm ref count	Nrm var count	Nrm VAF	AML ref count	AML var count	AML VAF	Lung ref count	Lung var count	Lung VAF
Shared founding clone	2	84960545	A	G	DNAH6	ENST00000237449	e61-23	-	AML	92	0	0	54	13	19.4	732	32	4.19
Shared founding clone	3	120321032	G	A	NDUFB4	ENST00000485064	NULL	-	AML	75	0	0	36	12	25	830	7	0.84
Shared founding clone	3	164760946	A	G	SI	ENST00000264382	p.C635	silent	AML,Lung	351	0	0	53	13	19.7	1603	44	2.67
Shared founding clone	7	131650156	C	T	ENSG00000252849	ENST00000517040	NULL	-	AML	79	0	0	29	11	27.5	532	8	1.48
Shared founding clone	7	140453136	A	T	BRAF	ENST00000288602	p.V600E	missense	AML,Lung	126	0	0	62	23	27.06	889	58	6.12
AML specific	3	169569336	A	G	LRRC31	ENST00000316428	e7+71	-	AML	61	0	0	35	12	25.53	479	0	0
AML specific	10	60573728	C	T	BICC1	ENST00000373886	p.R839C	missense	AML	93	0	0	80	17	17.53	928	2	0.22
AML specific	11	47376952	CCAGCGGTGCG	-	SPI1	ENST00000227163	p.A211fs	frameshift_del	AML	33	0	0	58	7	10.77	430	0	0
AML specific	X	34149573	C	A	FAM47A	ENST00000346193	p.D275Y	missense	AML	291	0	0	42	7	14.29	688	0	0
AML specific	X	41007718	AA	-	USP9X	ENST00000324545	p.K506fs	frameshift_del	AML	289	0	0	60	18	23.08	1077	0	0
AML specific	X	70339254	G	A	MED12	ENST00000333646	p.G44D	missense	AML	214	1	0.46	58	25	30.12	1066	0	0
Lung specific	7	157897493	A	T	PTPRN2	ENST00000404321	e10+6028	-	Lung	281	1	0.35	159	0	0	685	45	6.11
Distinct clonal hematopoiesis	2	113500302	G	T	CKAP2L	ENST00000302450	p.D601E	missense	Lung	109	0	0	39	3	7.14	449	37	7.6
Distinct clonal hematopoiesis	15	90631934	C	T	IDH2	ENST00000330062	p.R140Q	missense	Lung	460	1	0.22	162	6	3.57	1262	52	3.95

Chr indicates chromosome; Pos, position; AA, amino acid; Nrm, normal; ref, reference; VAF, variant allele frequency.

REFERENCES

1. Griffith M, Griffith OL, Smith SM, et al. Genome Modeling System: A Knowledge Management Platform for Genomics. Plos Comput Biol 2015;11.
2. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.
3. Koboldt DC, Zhang QY, Larson DE, et al. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012;22:568-76.
4. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics 2012;28:1811-7.
5. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213-9.
6. 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;20:1297-303.
7. Ye K, Schulz MH, Long Q, Apweiler R, Ning ZM. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics 2009;25:2865-71.
8. Miller CA, White BS, Dees ND, et al. SciClone: Inferring Clonal Architecture and Tracking the Spatial and Temporal Patterns of Tumor Evolution. Plos Comput Biol 2014;10.