De novo primary central nervous system pure erythroid leukemia/sarcoma with t(1;16)(p31;q24) *NFIA/CBFA2T3* translocation

Huifei Liu,[1] Terri L. Guinipero,[2] Kathleen M. Schieffer,[3] Chris Carter,[4] Susan Colace,[2] Jeffrey R. Leonard,[5] Brent A. Orr,[6] Samir B. Kahwash,[1] Patrick J. Brennan,[3] James R. Fitch,[3] Benjamin Kelly,[3] Vincent J. Magrini,[3] Peter White,[3] Richard K. Wilson,[3] Elaine R. Mardis,[3] Catherine E. Cottrell[3] and Daniel R. Boué[1]

[1]Department of Pathology and Laboratory Medicine, Nationwide Children's Hospital, The Ohio State University College of Medicine, Columbus, OH; [2]Department of Hematology/Oncology/BMT, Division of Pediatrics, Nationwide Children's Hospital, The Ohio State University College of Medicine, Columbus, OH; [3]The Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH; [4]Indiana University Health Bloomington Hospital, Bloomington, IN; [5]Department of Pediatric Neurosurgery, Nationwide Children's Hospital, Columbus, OH and [6]Pathology Department, St. Jude Children's Research Hospital, Memphis, TN, USA

Correspondence: HUIFEI LIU - huifei.liu@nationwidechildrens.org or liuhuifei@gmail.com

doi:10.3324/haematol.2019.231928

Supplemental Methods

Study Participant

The patient was enrolled as part of Institutional Review Board (IRB) approved studies (IRB16-00777; IRB17-00206) within the Institute for Genomic Medicine (IGM) at Nationwide Children's Hospital. Informed consent was provided by the parents for molecular genetic analysis, including whole genome sequencing and total RNA-sequencing. Peripheral blood was collected by routine venipuncture from the patient and used for genomic DNA extraction. Cerebrospinal fluid was obtained by lumbar puncture for DNA and RNA co-extraction.

RNA-sequencing

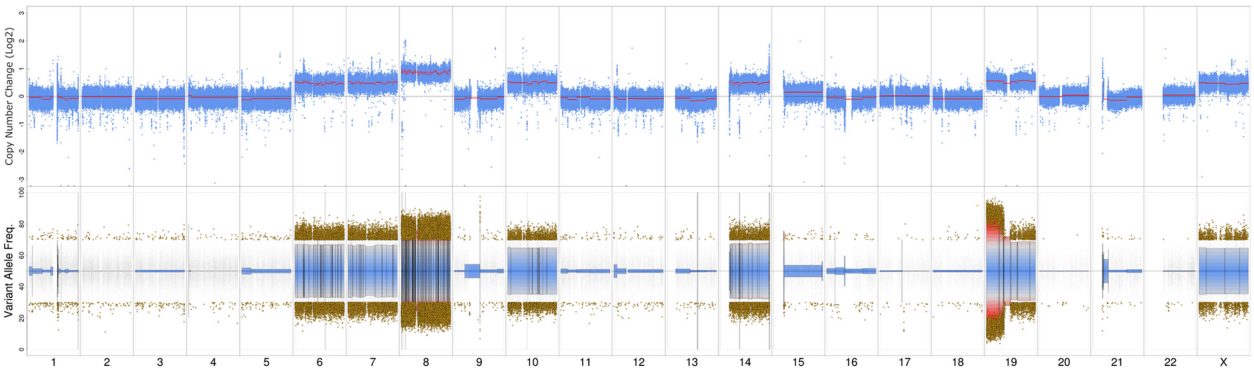
CSF tumor-derived total RNA, 1µg, was DNase-treated and concentrated per the manufacturers protocol (catalog # E1010 and R1015; Zymo Research, Irvine CA). 500ng DNase-treated RNA was incubated with Ribo-Zero and Globin-Zero depletion probes as the precursor to chemical fragmentation. RNA was processed with the TruSeq-stranded total RNA protocol performed with an 8 second chemical fragmentation (Catalog # RS-122-2501; Illumina, Inc., San Diego CA). Whole transcriptome data (RNA-sequencing) was produced as paired-end 151-bp reads generated on the Illumina HiSeq 4000. Reads were aligned to the human genome reference sequence (GRCh37/hg19) with the resultant output representing 114,733,896 mapped reads. RNA-sequencing data were processed using StarFusion which identified the *NFIA/CBFA2T3* fusion and the reciprocal *CBFA2T3/NFIA* fusion.[1] The fusion event was verified using reverse-transcriptase PCR with subsequent Sanger sequencing of the amplified product. Breakpoints were visualized using BLAT for alignment in the UCSC genome browser.

DNA sequencing

Whole genome sequencing (WGS) was performed on DNA extracted from peripheral blood and the disease-involved specimen (CSF). Libraries were prepared by fragmenting 500ng of genomic DNA (Covaris conditions: 75 W Peak Power, 20% Duty Factor, 1000 Cycles per burst, and a 20s duration). Fragmentation was followed by end repair, 5' phosphorylation, A-tailing, and sequencing adapter ligation using NEB Ultra II reagents (New England Biolabs, Ipswich, MA). PCR-Free libraries were sequenced on

the Illumina HiSeq 4000 in the form of paired-end 151-bp reads. Secondary analysis was performed using Churchill,[2] a comprehensive workflow encompassing raw sequencing reads from alignment through to constitutional and somatic variants calls. Reads were aligned to the human genome reference sequence (build GRCh37) using BWA (v0.7.15). Sequence alignments were refined according to community-accepted guidelines for best practices (https://www.broadinstitute.org/gatk/guide/best-practices). Duplicate sequence reads were removed using samblaster-v.0.1.22, local realignment was performed on the aligned sequence data using the Genome Analysis Toolkit (v3.7–0), and Churchill's own deterministic implementation of base quality score recalibration was applied. Constitutional variants were called using GATK's HaplotypeCaller.

Unique, on-target sequencing coverage depth derived from WGS was 81X for the tumor and 31X for the comparator normal. Somatic single nucleotide variation (SNV) and indel detection was performed using GATK's MuTect2.[3] Somatic SNVs were filtered for quality (minimum site quality ≥100), population frequency (gnomAD maximum population frequency <0.001), representation in comparator normal tissue (inclusion if total depth ≥8 reads and alternate reads=0 reads), somatic alternate read depth (inclusion if ≥4 reads), minimum tumor variant allele frequency (VAF) ≥5%, and genomic location (within a coding or splice site (≤3-bp) region). Additionally, constitutional and somatic variants were analyzed relative to known cancer-associated genes, somatic hotspots described in cancer, as well as publicly available and internally developed databases.[4] Copy number variation (CNV) was assessed using GATK and VarScan2.[5] Lollipop plots were generated using MutationMapper from WGS data and publicly available data within cBioPortal associated with myeloid neoplasms, including acute myeloid leukemia, histiocytic and dendritic cell neoplasms, myelodysplastic syndromes, and myeloproliferative neoplasms.[6,7]



Supplemental Figure 1. Copy number alterations derived from whole genome sequencing. Top plot: Tumor copy number relative to a sex-matched panel of normals, in addition to the patient-matched normal in log2 scale. Blue points represent log2 values based on sequence depth in 100-bp windows. Red lines indicate segmented CNV calls. Bottom plot: Tumor variant allele frequency for variants that are non-reference in the tumor. Points in red indicate significant loss of heterozygosity (LOH). The x-axis denotes the chromosome number.

Supplemental Table 1: High-quality somatic alterations identified by whole genome sequencing

Chr	Start	Gene	Effect	HGVS	Tumor VAF	Comparator Normal VAF
1	**27068455**	***ARID1A***	**CN deletion**	**ARID1A:NG_029965.1(NM_006015.4):c.1920+9172_2878+15del**	**allelic loss**	**0**
1	**27087944**	***ARID1A***	**nonsense**	**ARID1A:NM_006015.4:c.2231C>G:p.Ser744***	**0.7446**	**0**
19	**11488871**	***EPOR***	**nonsense**	**EPOR:NM_000121.3:c.1316G>A:p.Trp439***	**0.5158**	**0**
16	3025782	*PKMYT1*	missense	PKMYT1:NM_004203.4:c.410C>T:p.Ala137Val	0.3932	0
6	55113524	*HCRTR2*	missense	HCRTR2:NM_001526.3:c.311C>G:p.Thr104Ser	0.2718	0
10	29775360	*SVIL*	missense	SVIL:NM_021738.2:c.4612G>T:p.Gly1538Cys	0.2272	0
11	47469531	*RAPSN*	missense	RAPSN:NM_005055.4:c.364G>A:p.Gly122Arg	0.1666	0
14	105995182	*TMEM121*	missense	TMEM121:NM_025268.2:c.11C>T:p.Pro4Leu	0.1654	0
18	25756980	*CDH2*	missense	CDH2:NM_001792.4:c.7C>T:p.Arg3Trp	0.1639	0
15	85788182	*GOLGA6L3*	missense	GOLGA6L3:NM_001310153.1:c.961C>T:p.Arg321Cys	0.1632	0
9	133556992	*PRDM12*	nonframeshift deletion	PRDM12:NM_021619.2:c.1068_1076delCGCCGCCGC:p.Ala357_Ala359del	0.1451	0
1	48701469	*SLC5A9*	missense	SLC5A9:NM_001135181.1:c.1285A>G:p.Ser429Gly	0.1428	0
2	160027149	*TANC1*	missense	TANC1:NM_033394.2:c.1184T>A:p.Ile395Lys	0.1408	0
1	23751111	*TCEA3*	missense	TCEA3:NM_003196.2:c.16G>A:p.Glu6Lys	0.1392	0
11	103326007	*DYNC2H1*	missense	DYNC2H1:NM_001377.2:c.12550C>T:p.Arg4184Cys	0.1320	0
21	45534069	*PWP2*	missense	PWP2:NM_005049.2:c.236C>T:p.Ala79Val	0.1216	0
18	48723279	*MEX3C*	missense	MEX3C:NM_016626.4:c.412C>G:p.Arg138Gly	0.1111	0
17	33771703	*SLFN13*	missense	SLFN13:NM_144682.5:c.997G>A:p.Val333Met	0.1058	0
X	99854064	*TNMD*	missense	TNMD:NM_022144.2:c.629C>A:p.Ala210Asp	0.1046	0
11	124412680	*OR8B12*	missense	OR8B12:NM_001005195.1:c.871T>A:p.Leu291Met	0.1000	0
12	104147083	*STAB2*	missense	STAB2:NM_017564.9:c.6666C>A:p.Asn2222Lys	0.0989	0
17	34581565	*TBC1D3H* *TBC1D3F* *TBC1D3G*	frameshift	TBC1D3H:NM_001123392.3:c.1484dupC:p.Glu496fs TBC1D3G:NM_001291462.1.3:c.1484dupC:p.Glu496fs TBC1D3F:NM_032258.4.2:c.1484dupC:p.Glu496fs	0.0980	0
11	16838679	*PLEKHA7*	missense	PLEKHA7:NM_175058.4:c.1534C>T:p.Arg512Cys	0.0941	0
5	16701257	*MYO10*	missense	MYO10:NM_012334.2:c.3247C>A:p.Pro1083Thr	0.0909	0
15	45467482	*SHF*	missense	SHF:NM_138356.2:c.587G>T:p.Arg196Met	0.0898	0
22	32587002	*RFPL2*	missense	RFPL2:NM_001098527.2:c.894C>A:p.Phe298Leu	0.0897	0
19	16504806	*EPS15L1*	missense	EPS15L1:NM_001258374.1:c.1922C>T:p.Pro641Leu	0.0887	0
12	31106930	*TSPAN11*	missense	TSPAN11:NM_001080509.2:c.5C>T:p.Ala2Val	0.0869	0
12	14832650	*GUCY2C*	missense	GUCY2C:NM_004963.3:c.771G>C:p.Lys257Asn	0.0860	0
9	116136441	*HDHD3*	missense	HDHD3:NM_031219.3:c.194T>G:p.Leu65Arg	0.0843	0
19	23171249	*ZNF728*	missense	ZNF728:NM_001267716.1:c.8C>T:p.Ser3Leu	0.0842	0
4	156787369	*ASIC5*	missense	ASIC5:NM_017419.2:c.10A>G:p.Thr4Ala	0.0833	0
12	86276089	*NTS*	missense	NTS:NM_006183.4:c.449G>A:p.Arg150Gln	0.0833	0
7	1784972	*ELFN1*	missense	ELFN1:NM_001128636.2:c.740A>G:p.Gln247Arg	0.0826	0
4	1345606	*UVSSA*	missense	UVSSA:NM_020894.2:c.533C>T:p.Ala178Val	0.0813	0
19	17273202	*MYO9B*	missense	MYO9B:NM_004145.3:c.1462G>A:p.Ala488Thr	0.0810	0

13	103387052	*CCDC168*	missense	CCDC168:NM_001146197.1:c.15995C>T:p.Pro5332Leu	0.0789	0
12	80730300	*OTOGL*	missense	OTOGL:NM_173591.3:c.4681A>G:p.Ile1561Val	0.0781	0
11	118949925	*VPS11*	missense	VPS11:NM_021729.5:c.2350C>G:p.Gln784Glu	0.0777	0
6	32084485	*ATF6B*	missense	ATF6B:NM_004381.4:c.1793G>A:p.Arg598Gln	0.0769	0
7	148802523	*ZNF425*	missense	ZNF425:NM_001001661.2:c.440G>C:p.Ser147Thr	0.0754	0
13	78235644	*LOC100129307*	missense	LOC100129307:NM_001310140.1:c.1118G>A:p.Gly373Glu	0.0746	0
16	74808590	*FA2H*	missense	FA2H:NM_024306.4:c.64G>A:p.Ala22Thr	0.0744	0
19	5914970	*CAPS*	missense	CAPS:NM_004058.4:c.539G>A:p.Arg180Gln	0.0740	0
9	**5089753**	***JAK2***	**missense**	**JAK2:NM_004972.3:c.2651T>C:p.Leu884Pro**	**0.0724**	**0**
17	77097686	*RBFOX3*	missense	RBFOX3:NM_001082575.2:c.548C>T:p.Thr183Met	0.0721	0
6	99282830	*POU3F2*	missense	POU3F2:NM_005604.3:c.81G>A:p.Met27Ile	0.0693	0
11	1857428	*SYT8*	missense	SYT8:NM_138567.4:c.472G>A:p.Val158Ile	0.0689	0
16	81902872	*PLCG2*	missense	PLCG2:NM_002661.4:c.533G>A:p.Ser178Asn	0.0674	0
9	34978078	*PHF24*	missense	PHF24:NM_015297.2:c.1173C>A:p.Ser391Arg	0.0657	0
5	137762673	*KDM3B*	missense	KDM3B:NM_016604.3:c.4421G>T:p.Arg1474Leu	0.0655	0
2	71607662	*ZNF638*	missense	ZNF638:NM_014497.4:c.2344A>G:p.Lys782Glu	0.0649	0
2	47233088	*TTC7A*	missense	TTC7A:NM_020458.3:c.1093G>A:p.Val365Met	0.0641	0
1	62299365	*INADL*	missense	INADL:NM_176877.2:c.2020G>T:p.Asp674Tyr	0.0632	0
2	24406441	*FAM228A*	missense	FAM228A:NM_001040710.2:c.328A>G:p.Thr110Ala	0.0632	0
20	43992179	*SYS1*	missense	SYS1:NM_033542.3:c.8G>C:p.Gly3Ala	0.0632	0
19	40724237	*TTC9B*	missense	TTC9B:NM_152479.5:c.52C>A:p.Pro18Thr	0.0614	0
2	131521985	*AMER3*	missense	AMER3:NM_152698.2:c.2340G>C:p.Glu780Asp	0.0595	0
8	67366311	*ADHFE1*	missense	ADHFE1:NM_144650.2:c.905A>G:p.Asp302Gly	0.0588	0
17	35345870	*AATF*	nonsense	AATF:NM_012138.3:c.1000C>T:p.Arg334*	0.0579	0
1	151139880	*SCNM1* *TNFAIP8L2-SCNM1*	missense	SCNM1:NM_001204856.1:c.283C>T:p.Arg95Trp SCNM1:NM_024041.3:c.388C>T:p.Arg130Trp TNFAIP8L2-SCNM1:NM_001204848.1:c.283C>T:p.Arg95Trp	0.0574	0
16	50756540	*NOD2*	missense	NOD2:NM_022162.2:c.2722G>T:p.Gly908Cys	0.0563	0
1	21031159	*KIF17*	missense	KIF17:NM_020816.3:c.904G>A:p.Gly302Ser	0.0555	0
8	144654674	*MROH6*	missense	MROH6:NM_001100878.1:c.211C>T:p.Arg71Cys	0.0552	0
6	7575014	*DSP*	missense	DSP:NM_004415.2:c.2422C>T:p.Arg808Cys	0.0550	0
14	53187660	*PSMC6*	missense	PSMC6:NM_002806.3:c.901C>T:p.Pro301Ser	0.0549	0
1	26507327	*CNKSR1*	missense	CNKSR1:NM_006314.2:c.332T>A:p.Val111Asp	0.0533	0
4	37446945	*NWD2*	missense	NWD2:NM_001144990.1:c.3335A>G:p.Tyr1112Cys	0.0519	0
11	125864262	*CDON*	missense	CDON:NM_016952.4:c.2567A>G:p.Asn856Ser	0.0519	0
6	90491288	*MDN1*	missense	MDN1:NM_014611.2:c.1473C>A:p.Ser491Arg	0.0515	0
2	130899828	*CCDC74B*	missense	CCDC74B:NM_207310.2:c.422G>A:p.Arg141Gln	0.0512	0
2	110015350	*SH3RF3*	missense	SH3RF3:NM_001099289.2:c.1250C>T:p.Ala417Val	0.0506	0

CN: copy number

References

1. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* 2019;20(1):213.

2. Kelly BJ, Fitch JR, Hu Y, et al. Churchill: an ultra-fast, deterministic, highly scalable and balanced parallelization strategy for the discovery of human genetic variation in clinical and population-scale genomics. *Genome Biol.* 2015;16:6.

3. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.* 2013;31(3):213-219.

4. Zhang J, Walsh MF, Wu G, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med.* 2015;373(24):2336-2346.

5. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012;22(3):568-576.

6. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2(5):401-404.

7. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013;6(269):pl1.