Whole Exome Sequencing
Acute Leukemia of Ambiguous Lineage (ALAL)

*Protocols were performed at St. Jude Children's Research Hospital.

Library construction utilized DNA tagmentation (fragmentation and adapter attachment) performed using the reagent provided in the Illumina Nextera rapid exome kit (version 1.2) and was performed using the Caliper Biosciences (Perking Elmer) Sciclone G3. First-round PCR (10 cycles) was performed using Illumina Nextera kit v1.2 reagents, and clean-up steps employ BC/Agencourt AMPure XP beads. Target capture utilized Illumina Nextera rapid capture exome kit v1.2 and supplied hybridization and associated reagents. The pre-hybridization pool size was 12 samples, and second round PCR (10 cycles) performed with Nextera kit v1.2 reagents. Library quality control was performed using a Victor fluorescence plate reader with Quant-it dsDNA reagents for pre-pool quantitation, and Agilent Bio-analyzer 2200 for final library quantitation. Paired-end sequencing was performed using Illumina HiSeq 2500 with read length 100 bp.

Whole Exome Sequencing Analysis

*Protocols were performed at St. Jude Children's Research Hospital.

Alignment. Paired-end WXS data were aligned to the human reference genome GRCh37 by BWA[1] (version 0.7.12). Samtools[2] (version 1.3.1) were used to generate chromosomal coordinate-sorted and indexed bam files, and then Picard (version 1.129) MarkDuplicates module was used for marking PCR duplication.

SNV/indel calling and filter workflow. The GATK UnifiedGenotyper module was used to identify SNVs and indels from leukemia and germline samples, which were filtered by a homemade pipeline, excluding: 1) reported common SNPs/indels from UCSC dbSNP v142; 2) germline mutations detected from matched germline control samples. All the non-silent SNVs/indels yield from the filtering pipeline were manually reviewed and only the highly reliable somatic ones were reported. Meanwhile, adjacent nucleotide changes on the same allele were merged into a single mutation.

For patients with flow sorted subpopulations of leukemia cells sequenced, the mutation calling for each population was performed de novo. Mutations detected from some/one of the samples were checked across the other samples from the same patient.

References

1. Li H, *et al*. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* **25**(14):1754-60. (PMID: 19451168)

2. Li H, *et al.* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25**(16):2078-9. (PMID: 19505943)