Demonstrating the use of Archer VARIANT*Plex*™ and FUSION*Plex*™ assays in observing variants at sub-0.1% allele frequencies and low transcript number fusions in acute myeloid leukemia relevant genes

319

Laura Johnson, Michael Washburn, Paula Kalavakur, Dhruti Legare | Archer, Boulder, CO. USA.

Introduction

The presence of acute myeloid leukemia (AML) residual disease is valuable for understanding cancer progression. Studies have shown that variants detected at allele frequencies as low as 0.01% are useful for stratifying acute myeloid leukemias [1]. Both low limits of detection and high specificity are required for an assay to be useful in this context. Therefore, labs must choose an assay that has inherently low levels of background noise and that employs robust error suppression techniques.

Here we demonstrate the use of IDT's Archer Next Generation Sequencing-based VARIANT*Plex* and FUSION*Plex* assays (RUO) and accompanying software, Archer Analysis, for the application of detecting low allele frequency and low transcript number fusions with high levels of sensitivity and specificity. Currently available Archer assays and software solutions target AML-relevant variants and are compatible with multiple sequencing platforms. Using Archer technology, key AML-related mutations were detected at allele frequencies less than 0.1%, and low transcript numbers of myeloid relevant fusions were identified. In addition, > 95% of the bases in the targeted region of interest were powered to detect variants at allele frequencies of less than 0.1% in the VARIANT*Plex* libraries. Finally, error correction and noise filtering techniques remove many false positive variants in order to reduce the false positive variant calling rate.

Methods

To generate the input material, 1,000 ng total gDNA or 200 ng total RNA from commercially available cell lines containing myeloid relevant variants was diluted into a background of wild type gDNA at a mass ratio of 1:100 or RNA at a mass ratio of between 1:2 and 1:20 to assess analytical sensitivity. Libraries were prepared using the catalog panels VARIANT*Plex* AML Focus or the FUSION*Plex* Pan Heme assay. A prototype VARIANT*Plex* AML MRD panel was also tested using 50-1,000 ng NA12877 input. VARIANT*Plex* libraries were sequenced on Illumina and Element Biosciences sequencing platforms to depths of 10-50 M reads. FUSION*Plex* libraries were sequenced on Illumina NextSeq 500s to depths of 4.5-10M reads. Data were analyzed with Archer Analysis using a pipeline which includes read cleaning, deduplication, error correction, variant calling, variant filtering, and reporting steps.

Variant Filtering Strategies

Variant Caller	Pros	Cons
Tumor Informed (Vision)	Best specificity	Slower turn-around-time and less cost effective
De Novo No normal cohort	Rapid turn-around-time and most cost effective	Lowest specificity
De Novo With normal cohort & outlier detection	Rapid turn-around-time. Improved Specificity	Middle cost-effectiveness

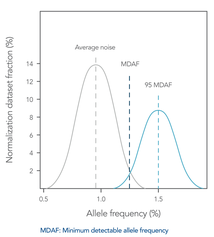

Figure 1. Archer Analysis Variant Filtering Methods. Archer Analysis has many customizable variant filtering options including targeted variant calling (Vision) and positional background error rate modeling using a normal cohort. The advantages of several of the different strategies are described in the table above. The concept of the minimal detectable allele frequency (MDAF), determined by weighing both the positional coverage and the positional background noise found in a normal cohort, is described to the left. Users can filter out variants that do not have allele frequencies significantly above the positional background noise.

Results: VARIANT*Plex* Analytical Sensitivity, Coverage and Error Rates

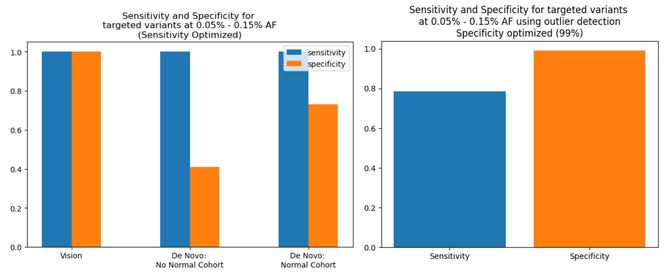

Figure 2. Sensitivity and Specificity. Sensitivity for detecting the 5 expected indels and 2 expected FLT3 internal tandem duplications in the diluted SeraCare Myeloid samples prepped with a VariantPlex AML MRD (in-development) panel using various variant filtering methods. When optimizing for sensitivity (100%), the specificity is 100% when using a tumor informed approach (Vision). Specificity is second highest with the use of a normal cohort and outlier detection which compares position specific background noise to the alternate observations at the matched position in the interrogated sample. B) When updating the Archer Analysis filter settings to optimize specificity (>99%), 4/5 expected SNVs and 2/2 FLT3 ITDS were detected. The library was prepared with 1,000 ng input, sequenced to 50 M reads, and Archer Analysis variant grid filters including number of alternate observations, minimum allele frequency, gnomad population AF were applied.

variant	expected AF	NextSeq observed AF	Aviti observed AF
IDH1 p.R132C	0.0005	0.001432	0.001278
NPM1 p.W288fs*12	0.0005	0.000420	0.000175
FLT3 p.D835Y	0.0010	0.001506	0.001558
CEBPA p.K313_V314insK	0.0015	0.001265	0.001335
CEBPA p.H24fs*84	0.0015	0.001491	0.001907

Figure 3. Expected and Observed Variant Allele Frequencies in NextSeq 500 and Aviti fastqs. Libraries prepared with 1,000 ng input of the 1:100 diluted SeraCare Myeloid cell line and with the VariantPlex AML Focus panel were sequenced to 50 M reads on both the NextSeq 500 and the Aviti sequencers. D*e novo* variant calling was performed by Archer Analysis. 9/9 expected indels, and 2/2 FLT3 ITDs were detected with both platforms.

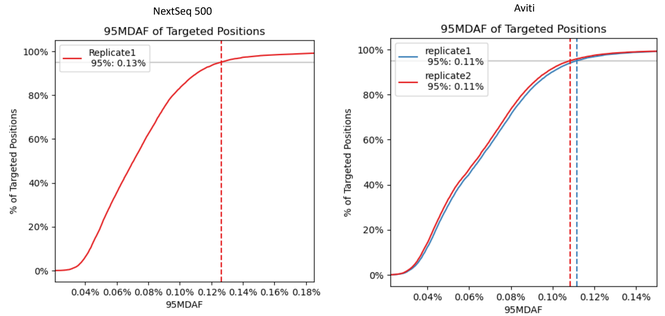

Figure 4. Variant Calling Power Analysis. The 95% minimal detectable allele frequency (95MDAF) is a metric that is calculated by calculating the per position background error rates and per position coverage to determine the minimal variant allele frequency the data is powered to detect >=95% of the time. The above graphs demonstrate the percent of the targeted bases that are powered to detect a given 95MDAF. The fastqs generated by both the NextSeq500 (left) and the Aviti (right) sequencers are powered to detect variant allele frequencies of near 0.1% for 95% of the bases in the region of interest.

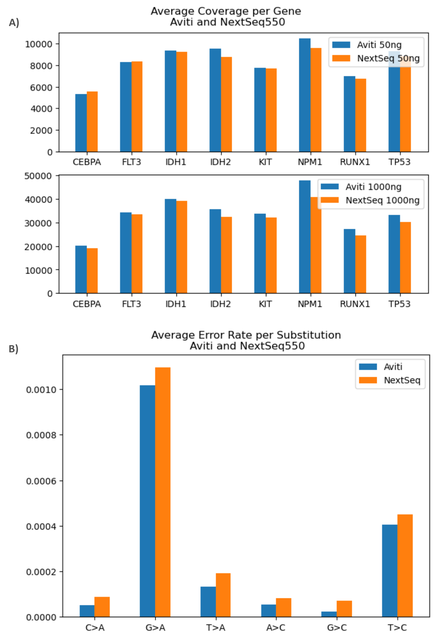

Figure 5. Coverage and Background Substitution Error Rates. A) Unique coverage over target genes with 50 or 1,000 ng Genome in a Bottle gDNA input. Libraries prepared with a development VariantPlex AML MRD panel and sequenced to 10M reads. B) Substitution background error rates in error corrected unique reads.

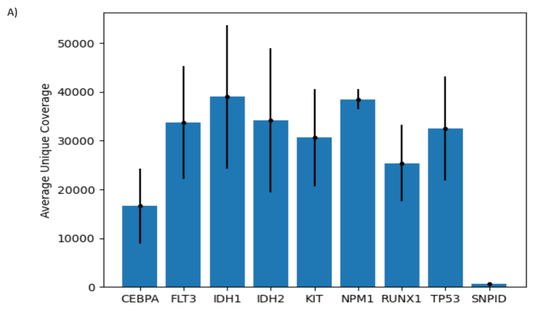

Figure 6. Gene coverage for development AML MRD panel sequenced on NextSeq 550. Unique coverage over target genes with 50 ng Genome in a Bottle gDNA input sequenced to 10 M reads. SNP_ID primers are designed to target positions of heterogeneity within the population, are targeting germline variants which need much less coverage in order to make a homozygous reference, homozygous alternate, or heterozygous variant call. Therefore, to save sequencing costs, these primers are down-balanced in the panel. SNP_ID primers are designed to target positions of heterogeneity within the population, thus allowing each library to be fingerprinted for longitudinal tracking.

Results: FUSION*Plex* Analytical Sensitivity

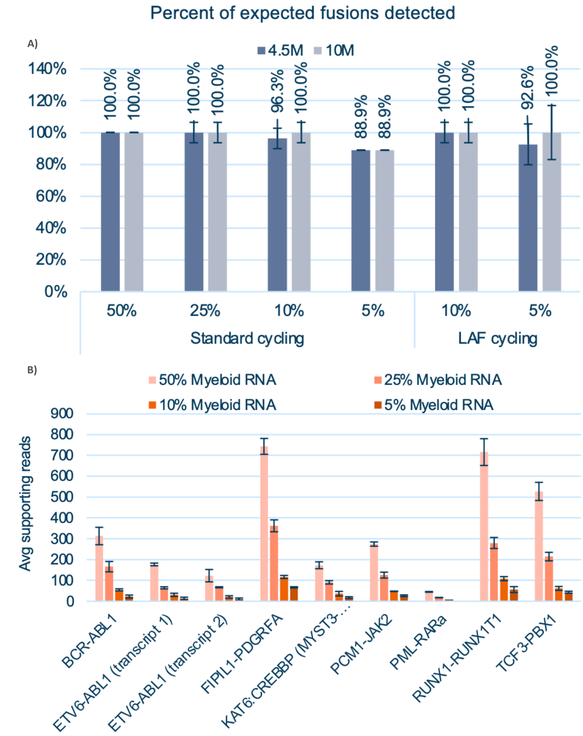

Figure 7. Fusion Positive Cell Line Dilutions. A) Percent of expected fusions detected from input mass dilutions of extracted RNA from cell line containing expected myeloid malignancy relevant fusions. Input mass ratios as low as 5% were detected. An input mass ratio of 5% is not equivalent to 5% cellular dilution level, however, the lower the input mass ratio, the lower the expected number of transcripts containing the fusion breakpoint. Fusion positive Input: Seraseq® Myeloid Fusion RNA Mix (0710-0407), Background: Seraseq® WT RNA (0710-1580), 200ng input mass, *n* = 3 per input. FusionPlex Pan Heme libraries prepared with new liquid chemistry with standard or Low Allele Frequency (LAF) optimized PCR cycling. B) Number of supporting unique reads for each expected fusion across multiple dilution levels.

Conclusions

In conclusion, the VARIANT*Plex* AML Focus assay used with Illumina or Element sequencers demonstrated the ability to detect variants down to allele frequencies of 0.05%, and the FUSION*Plex* Pan Heme assay was used to detect low transcript number fusions which may be useful for AML minimal residual disease research.

References

1. Dillon et. al. JAMA. 2023;329(9):745-755.

Laura Johnson
Scientist, Archer (Integrated DNA Technologies)
ljohnson@idtdna.com
+1 (319) 626-8438