ctDNA Evaluation Set: Insights

A new era of ctDNA accuracy: powerful calibration built into every sample

Background

Next-generation sequencing (NGS) can be used to identify genetic variation and disease-associated mutations and has become a principal tool in biomedical research and clinical testing. However, numerous factors influence the accuracy of variant detection using NGS including PCR amplification biases, capture efficiency, sequencing depth, read length and sequencing errors. It is therefore imperative that a system of control standards is incorporated to account for accumulated errors, and improve data quality and interpretation, maximizing genomic insights.

Introduction to Sequins

Sequins™ (sequencing spike-ins) are synthetic nucleic acid controls that directly mirror naturally occurring sequences. Because Sequins retain the same nucleotide composition as the natural sequence, they enable accurate representation of genomic complexity without compromising the integrity of the sample and results. Sequins perform equivalently throughout sequencing workflows, providing a true measure of control (Deveson et al. 2019).

Sequins' innovative design enables the production of synthetic mirrored sequences that directly represent almost any genomic feature, in any organism with a reference genome. This includes common and clinically relevant variants and analytically challenging regions of the genome. By combining Sequins in precise ratios, quantitative features of genome biology, such as variant allele frequencies or copy-number variation, can also be emulated (Deveson et al 2021).

Sequins are simply 'spiked-in' to a sample prior to library preparation and progressed together through a workflow (Figure 1). Sequins controls can then be distinguished from the native sample in the output library by virtue of their mirrored sequence enabling standardization and comparison between samples, runs, laboratories, chemistries, and sequencers.

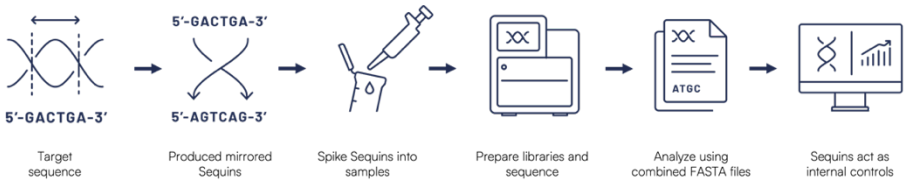

Figure 1: Schematic showing the design and use of Sequins in an NGS workflow.

Insights from testing of the Sequins ctDNA Evaluation Set

Overview

The Sequins ctDNA Evaluation Set is a synthetic reference material designed to be spiked into ctDNA samples at low amount to monitor in-sample sensitivity and precision.

The material is formulated as a molecular ladder comprising twenty-two cancer genes with variant alleles at frequencies of 0.1%, 0.5%, 1%, 5%, and 10%, combined in a single mix. Detailed information on the target content, variant composition, and sequencing performance can be found in the ctDNA Evaluation Set Information Sheet.

Pilot Verification Study

Methodology

Normal male human cell-free DNA (50ng) was spiked with 0.0016% (0.8pg), 0.008% (4pg), 0.04% (20pg), 0.2% (100pg), or 1% (500pg) of the Sequins ctDNA Evaluation Set. Libraries were prepared using the xGen™ cfDNA & FFPE DNA Library Prep v2 Kit (IDT), followed by hybrid capture with the xGen™ Hybridization and Wash Reagents v3 Kit (IDT).

Two capture probe sets were used:

1. IDT xGen™ Human ID Hyb Panel: targeting 76 SNPs with 229 probes
2. Sequins control probe set (herein referred to simply as Sequins): targeting the Sequins ctDNA Evaluation Set with 212 probes

Sequencing was performed on the Illumina NextSeq 2000 P1 flow cell (2 × 150bp) with a minimum of 5Gb per sample.

Key Insights

Compatibility with Existing Panels

Sequins showed no interference with the IDT Human ID Hyb Panel. Coverage uniformity across targets remained consistent whether captured alone or co-captured with Sequins across a range of spike-in levels (Figure 2). The insert size distribution also remained consistent when co-captured with Sequins (Figure 3).

Figure 2: Combined mean coverage for each target in the IDT Human ID Hyb Panel when captured with and without Sequins. Coverage of each target remained consistent when the Sequins were also captured.

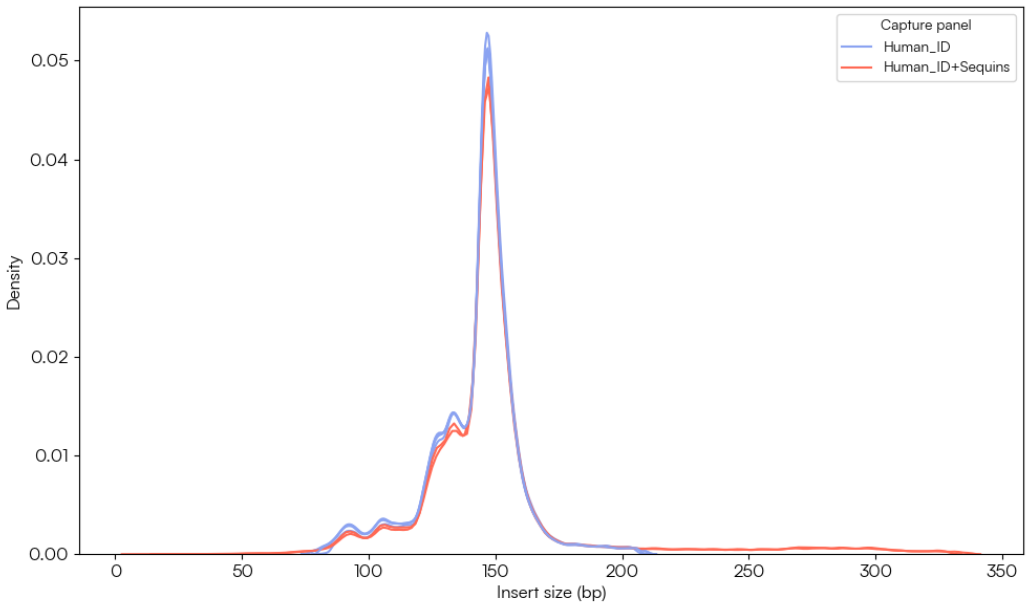

Figure 3: Insert size distribution of the IDT Human ID Hyb Panel when captured with and without Sequins. The distribution of insert sizes was unaffected by the addition of Sequins with the peaks at ~150bp and ~130bp maintained.

Optimal Spike-in Level

Reads mapping to Sequins and the captured sample naturally compete for sequencing output. Performance data showed the Sequins control footprint can be predictably reduced by lowering the spike-in amount while maintaining optimal performance (Figure 4). Decreased Sequins spike-in concentrations resulted in a corresponding increase in the proportion of reads aligning to the sample with a concomitant decrease in Sequins coverage. Importantly, this reduction in spike-in concentration did not affect the performance of the molecular ladder, with the linearity and slope maintained across the dilution series. Beyond a certain spike-in threshold, Sequins coverage dropped below that of the sample. For accurate sensitivity calibration, the Sequins spike-in coverage should remain higher than the captured sample.

The number of reads aligning to Sequins will vary according to panel size and complexity, therefore it's important for users to optimize and titrate Sequins spike-in amounts in combination with their incumbent targeted panel to ensure that appropriate read depth is achieved for both controls and targets.

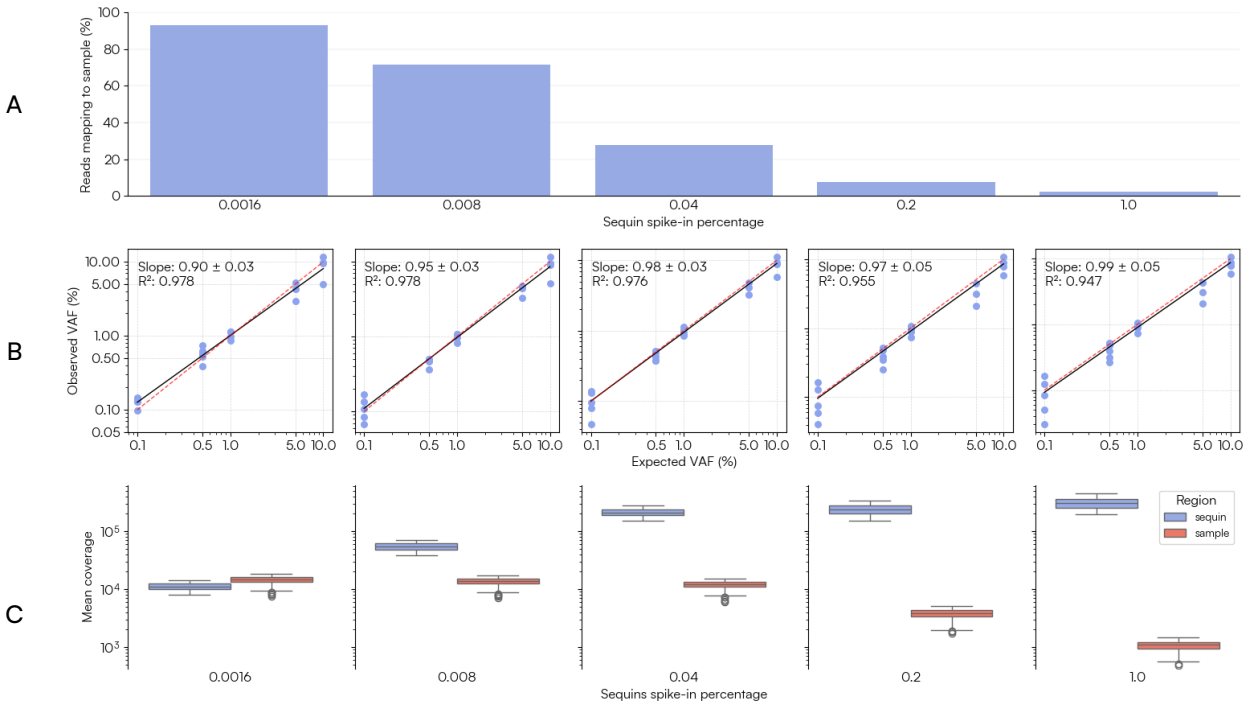

Figure 4: Titration of the Sequins ctDNA Evaluation Set spiked into cfDNA co-captured with the IDT Human ID Hyb Panel. (A) The percentage of reads mapping to the human sample (blue bars) increased with decreasing Sequins spike-in concentration (from 1% to 0.0016%). **(B)** The molecular ladder, as demonstrated by the correlation between the observed and expected variant allele frequencies (VAFs) across the dilution series, maintained linearity across all spike-in concentrations. **(C)** The coverage distribution for the Sequins ctDNA Evaluation Set (blue) and human sample (red) demonstrated that Sequins coverage decreased with decreasing spike-in amounts.

Comparison of called IDT Human ID panel variants

For IDT Human ID Panel Only samples, the called variants were extracted and the union combined to establish a baseline of 53 variants. All 53 variants were called correctly in all Human ID Panel Only samples and all were called correctly in the Human ID panel + Sequins samples when sufficient coverage was achieved. No false positives were detected in the samples including Sequins. Drop out of variant calls only occurred when sequence coverage fell to inadequately low levels. These results further confirm that Sequins spike-in controls do not interfere with the sample or analyses and optimization of coverage is an important parameter to consider for experimental design.

Table 1: Comparison of variants called in the IDT Human ID Hyb Panel with and without Sequins. At mean coverage of >160x all 53 variants were called. At very low mean coverage, drop out of called variants was observed.

Sample	Variants called	Sequin conc	Capture plex	Capture panel	Mean coverage
SMP01	53	0.2	1-plex	Human_ID	9,459
SMP06	53	0.04	3-plex	Human_ID	2,796
SMP07	53	0.2	3-plex	Human_ID	2,895
SMP08	53	1	3-plex	Human_ID	2,784
SMP16	53	0.0016	1-plex	Human_ID+Sequins	9,272
SMP15	53	0.008	1-plex	Human_ID+Sequins	8,423
SMP03	53	0.04	1-plex	Human_ID+Sequins	3,302
SMP17	53	0.04	1-plex	Human_ID+Sequins	5,650
SMP04	53	0.2	1-plex	Human_ID+Sequins	161
SMP05	50	1	1-plex	Human_ID+Sequins	10
SMP12	47	0.04	3-plex	Human_ID+Sequins	8
SMP13	43	0.2	3-plex	Human_ID+Sequins	7
SMP14	38	1	3-plex	Human_ID+Sequins	7

Conclusion

This study demonstrates that the Sequins ctDNA Evaluation set, which comprises a synthetic molecular ladder, has the potential to provide a robust and valuable internal reference control for assessing sensitivity and precision in ctDNA workflows.

Our results showed that:

- Sequins did not impact sample coverage uniformity or variant calls when co-captured with a widely used commercial panel
- Sequins provided stable molecular ladder performance through library prep, hybrid capture and sequencing
- Sequins spike-in levels should be optimized in line with user panel size, complexity and desired read depth to achieve slightly higher coverage than the sample

Contact Information

Enquiries and Support

enquiries@sequins.bio
support@sequins.bio

References

Deveson, I.W., Madala, B.S., Blackburn, J. et al. (2019) Chiral DNA sequences as commutable controls for clinical genomics. *Nat Commun* . 10, 1342.

Deveson, I.W., Gong, B., Lai, K. et al. (2021) Evaluating the analytical validity of circulating tumor DNA sequencing assays for precision oncology. *Nat Biotechnol*. 39, 1115—1128