

ctDNA Evaluation Set: Insights

Demonstrating Sequins Compatibility with IDT-Based ctDNA Hybrid-Capture Workflow

Overview

The Sequins™ ctDNA Evaluation Set was assessed in an IDT-based ctDNA hybrid-capture workflow to evaluate compatibility across library preparation, hybrid-capture, sequencing, and variant analysis. The study used the IDT xGen™ cfDNA & FFPE DNA Library Prep v2 Kit, followed by hybrid capture with the xGen™ Hybridization and Wash Reagents v3 Kit, together with an IDT Custom Panel (design ID: 08140653_invertedP250912) targeting the 23-gene Sequins ctDNA Evaluation Set. The Evaluation Set comprises synthetic mirrored reference and variant alleles formulated as a molecular ladder across five variant allele frequency (VAF) levels: 0.1%, 0.5%, 1%, 5%, and 10%. When spiked into normal male cfDNA samples, the Evaluation Set ladder was successfully captured and sequenced, with observed VAFs showing strong concordance with expected VAFs across all tested conditions. These findings support the compatibility of Sequins spike-in controls with an IDT-based ctDNA hybrid-capture workflow and demonstrate their potential as an in-sample companion control for monitoring sensitivity, precision, and quantitative performance.

Background

Circulating tumor DNA (ctDNA) analysis by next-generation sequencing (NGS) is increasingly used in oncology research and precision medicine to detect, monitor, and characterize tumor-derived variants from minimally invasive liquid biopsy samples. However, ctDNA assays are analytically challenging because tumor-derived fragments are often present at low concentrations, resulting in variants observed at very low VAFs within a high background of non-tumor cell-free DNA (cfDNA). Accurate detection of these low-frequency variants can be affected by multiple factors across the workflow, including input quantity and quality of cfDNA, library preparation efficiency, hybrid-capture performance, sequencing depth, error suppression, and downstream bioinformatics analysis.

These challenges create a need for internal reference materials that can be introduced directly into the sample and carried through the same workflow as endogenous ctDNA. Such controls can help laboratories assess whether low-VAF variants are being captured, sequenced, and analyzed as expected within a given assay setting, while also supporting more consistent interpretation of assay sensitivity, precision, and quantitative performance over time. In the context of capture-based ctDNA workflows, this is particularly important because variation in probe design, enrichment efficiency, and sequencing allocation can materially influence analytical performance.

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 VAFs 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.

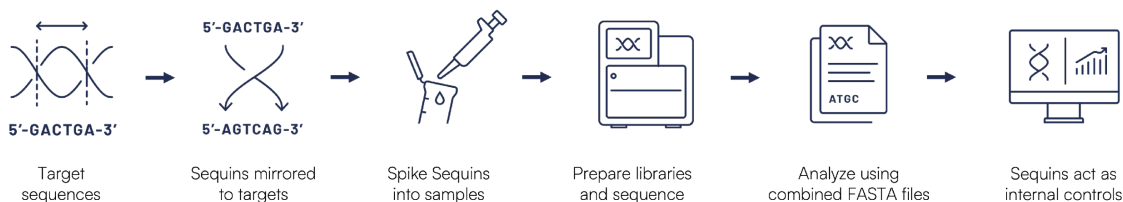


Figure 1: Schematic showing the design and use of Sequins in an NGS workflow.

ctDNA Evaluation Set Overview

Intended Use

The Sequins ctDNA Evaluation Set is a synthetic spike-in control designed for evaluation use in ctDNA assays and is provided for research use only (RUO), enabling laboratories to assess the performance of Sequins technology in their workflows. This includes the mirrored control design, molecular ladder, and associated analyses. The Evaluation Set is intended for initial workflow assessment, to inform the subsequent development of bespoke Sequins control sets tailored to user-specific targets, variants, and VAF requirements, which can then support sample- and run-specific performance assessment, assay development and validation, and routine quality control.

Molecular Ladders

When spiked into every sample, ctDNA molecular ladders provide a framework for generating both sample- and run-specific performance insights within every sequencing run. By including variants at multiple defined allele frequencies, molecular ladders enable assessment of analytical sensitivity, precision, and quantitative performance across a clinically relevant VAF range. Because the ladder is present in every sample, associated performance metrics such as analytical sensitivity, slope and linearity can be trended over time to support monitoring of analytical consistency and detection of shifts related to routine workflow variation.

Sequins ctDNA Evaluation Set

Sequins controls are designed as matched pairs, with one molecule representing the wild-type allele and the other the variant allele. These molecules are blended at defined ratios to generate the desired VAFs. The ctDNA Evaluation Set is formulated as a molecular ladder spanning 23 cancer genes, with variants represented at allele frequencies of 0.1%, 0.5%, 1%, 5%, and 10% in a single mix (Table 1). The material contained within the Set is fragmented to represent the size profile of cfDNA. Importantly, the patented mirrored design ensures that ladder-specific variants are distinguishable from sample-derived variants during bioinformatics analysis.

Table 1: Sequins Evaluation Set Content

Gene	ClinVar UID	Variant	VAF %
<i>BRCA1</i>	ClinVar_1076672	c.2219_2220insTAAT (p.Ser741fs)	0.1
<i>WT1</i>	ClinVar_449416	c.1120C>T (p.Arg374Ter)	0.1
<i>SMAD4</i>	ClinVar_24867	c.1587dup (p.His530fs)	0.1
<i>APC</i>	ClinVar_1072211	c.476_488del (p.Tyr158_Tyr159insTer)	0.1
<i>RBI</i>	ClinVar_1073997	c.46_61GCC[2]GCGGAACCCCAGGCACCGCCGCCGCCGCCGCGGAACCCC[1] (p.Pro21fs)	0.1
<i>BRCA2</i>	ClinVar_9342	c.658_659del (p.Val220fs)	0.5
<i>PMS2</i>	ClinVar_1328224	c.741_742insGTGTGTGAAG (p.Ser248fs)	0.5
<i>TSC1</i>	ClinVar_1013340	c.903dup (p.Asn302fs)	0.5
<i>VHL</i>	ClinVar_2218	c.499C>T (p.Arg167Trp)	0.5
<i>MEN1</i>	ClinVar_1070954	c.1375_1391dup (p.Ala467fs)	0.5
<i>MAX</i>	ClinVar_404110	c.211_221del (p.Ile71fs)	0.5
<i>MLH1</i>	ClinVar_89935	c.18_34del (p.Val7fs)	1.0
<i>PTEN</i>	ClinVar_545882	c.865_866insTTCT (p.Lys289fs)	1.0
<i>SDHAF2</i>	ClinVar_532513	c.177dup (p.Asp60Ter)	1.0
<i>TSC2</i>	ClinVar_237966	c.148A>G (p.Met50Val)	1.0
<i>MSH6</i>	ClinVar_1070916	c.3964_3980dup (p.Asn1327delinsLysAsnLeuArgArgTer)	5.0
<i>TP53</i>	ClinVar_1066203	c.934_992dup (p.Ile332fs)	5.0
<i>BMPRIA</i>	ClinVar_529927	c.366_384del (p.Glu123fs)	5.0
<i>TMEM127</i>	ClinVar_126966	c.265_268del (p.Thr89fs)	5.0
<i>MSH2</i>	ClinVar_90711	c.1638_1639dup (p.Asn547fs)	10.0
<i>RET</i>	ClinVar_230926	c.1998G>C (p.Lys666Asn)	10.0
<i>STK11</i>	ClinGen_CA402950689	g.1221314G>A (g.1221314G>A)	10.0
<i>SDHB</i>	ClinVar_428926	c.17_42dup (p.Ala15delinsProSerProTer)	10.0

Capture-Based Approach

Molecular ladder approaches can be used in capture-based workflows. For Sequins, application in capture-based methods requires corresponding capture probes to account for the mirrored molecular design. In this technical note, data is presented from a pilot verification study demonstrating the use of the ctDNA Evaluation Set with IDT probes. This showed that Sequins molecules can be co-captured across all target loci consistently in the IDT probe-based workflow.

Further details on target content, variant composition, and sequencing characteristics are provided in the ctDNA Evaluation Set Information Sheet.

Pilot Verification Study

Methodology

Normal male human cell-free DNA (50ng) (Anchor Molecular) 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. IDT custom control probe set (herein referred to simply as Sequins): targeting the Sequins ctDNA Evaluation Set with 218 probes (manufacturer ID: 08140653_invertedP250912).

Sequencing was performed on the Illumina NextSeq 2000 P1 flow cell (2 × 150bp) with a minimum of 5Gb per sample. Sequencing data were processed using a UMI-aware bioinformatics workflow with alignment to the GRCh38 reference genome and the Sequins ctDNA Evaluation Set decoy chromosome.

VAF ladder (Figure 2) and IDT Human ID Hyb Panel variant calling (Table 2) were assessed using duplex consensus reads. Coverage profiles, insert size distributions, and spike-in titration behavior (Figures 3-5) were evaluated using UMI-deduplicated reads. Mutect2 was used for variant calling, with targeted sites specified for VAF ladder evaluations and unrestricted calling across the targeted panel regions for the IDT Human ID Hyb Panel.

Key Insights

Compatibility with Existing Panels

The captured Evaluation Set ladder demonstrated excellent quantitative performance and showed no interference with the IDT Human ID Hyb Panel.

Using the IDT hybrid capture workflow and custom IDT-designed probe set targeting Evaluation set variants, all variants were detected and demonstrated excellent linearity with an observed R^2 of 0.99 and a slope of 1.03 (Figure 2)

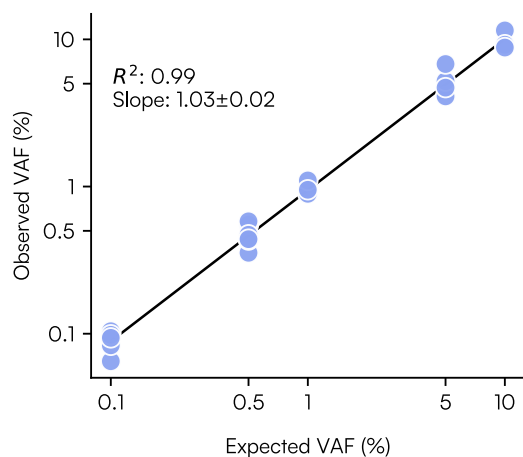


Figure 2: Performance of the Sequins ctDNA Evaluation Set in IDT capture-based workflows. Evaluation Set ladder showing observed versus expected variant allele frequency (VAF) in IDT workflow. Slope was close to 1 and R^2 value was 0.99, demonstrating a strong linear relationship between observed and expected frequencies, indicating accurate and consistent variant quantification.

To assess whether inclusion of the Evaluation Set co-capture probe set affected the capture efficiency of the sample panel probe set, the sample targets were captured using the IDT workflow either with Human ID Hyb panel alone (Human ID) or in combination with Evaluation Set probe set (Human ID + Sequins). No interference with the IDT Human ID Hyb Panel was observed (Figure 3). Coverage uniformity across sample targets remained consistent regardless of whether targets were captured alone or co-captured with the Evaluation Set over a range of spike-in levels (Figure 3). Similarly, the insert size distribution remained unchanged following co-capture with the Evaluation Set (Figure 4).

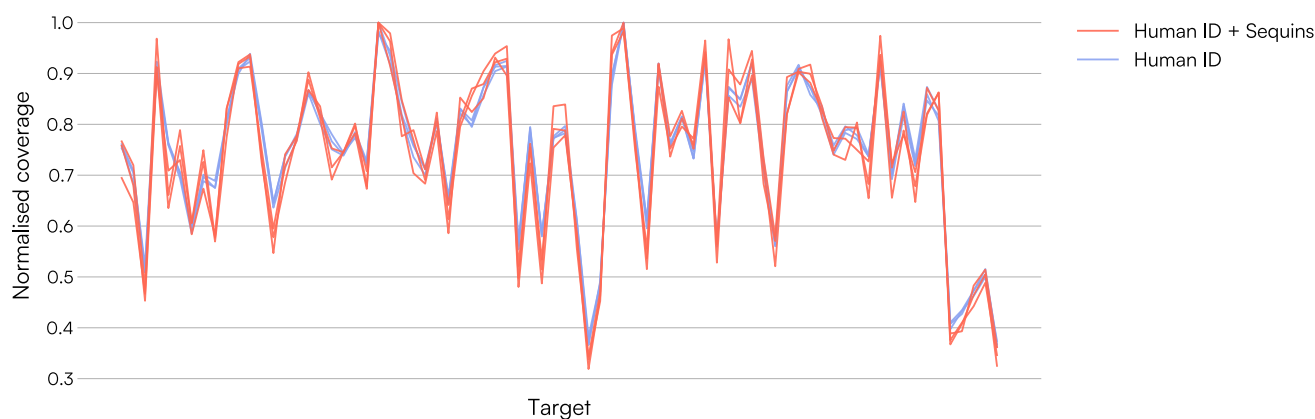


Figure 3: Combined mean coverage for each target in the IDT Human ID Hyb Panel when captured with and without Sequins. Coverage of each target remained consistent when the Sequins were also captured

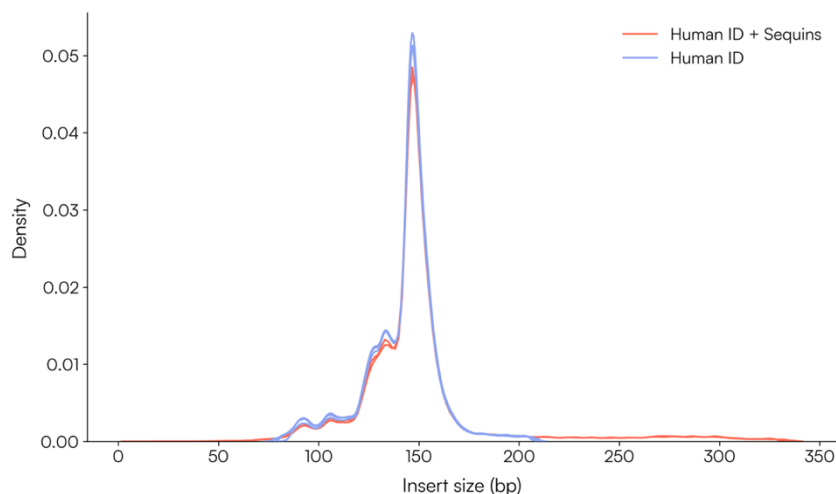


Figure 4: Insert size distribution of the IDT Human ID Hyb Panel when captured with and without Sequins. The distribution of insert sizes was unaffected by the addition of Sequins with the peaks at ~150bp and ~130bp maintained.

Importantly, the results further demonstrated that Sequins spike-in controls did not interfere with the sample or downstream analyses. For the IDT Human ID Hyb Panel Only samples, called variants were extracted and the union combined to establish a baseline of 53 variants. All 53 variants were correctly detected in every Human ID Panel Only sample and all were also correctly detected in the Human ID panel + Sequins samples when sufficient coverage was achieved (Table 2). No false positives were detected in samples containing Sequins. Drop out of variant calls only occurred when sequence coverage fell to inadequately low levels, highlighting coverage optimization as an important consideration for experimental design.

Table 2: 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

Optimal Spike-in Level

Reads mapping to the Evaluation Set and the captured sample naturally compete for sequencing output. It is, therefore, important to balance the Evaluation Set spike-in level so that sufficient control coverage is achieved without disproportionately increasing its share of reads relative to the sample. Performance data showed the Evaluation Set footprint can be predictably reduced by lowering the spike-in amount while maintaining optimal performance (Figure 5). Decreased Evaluation Set spike-in concentrations resulted in a corresponding increase in the proportion of reads aligning to the sample with a concomitant decrease in Evaluation Set 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. However, below genome equivalence, where Evaluation Set copy number fell below sample copy number, control coverage dropped below sample coverage. At 0.0016% it no longer provided the coverage excess needed for calibration (Figure 5B).

For accurate sensitivity calibration, the Evaluation Set should therefore be spiked in at a level that maintains control coverage slightly above sample coverage. As this relationship will vary with incumbent panel size and complexity, spike in levels should be optimized for each target panel.

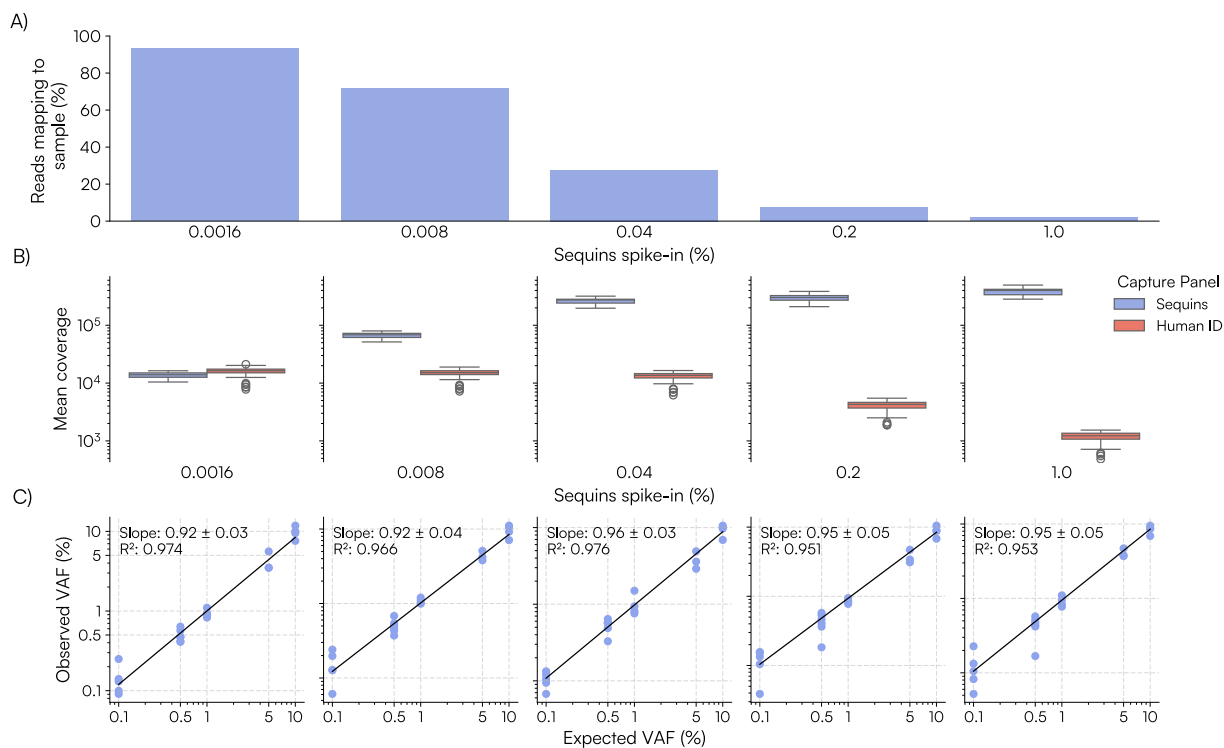


Figure 5: Titration of the Sequins ctDNA Evaluation Set spiked into cfDNA co-captured with the IDT Human ID Hyb Panel. (A) The percentage of reads mapping to the human sample (blue bars) increased with decreasing Sequins spike-in concentration (from 1% to 0.0016%). **(B)** The coverage distribution for the ctDNA Evaluation Set (blue) and human sample (red) demonstrated that Sequins coverage decreased with decreasing spike-in amounts. **(C)** The molecular ladder showed a linear relationship between observed and expected variant allele frequencies (VAFs) across all spike-in concentrations, with slopes close to 1 and high R^2 values.

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

Access to the ctDNA Evaluation Set

The Sequins ctDNA Evaluation Set is available through the Sequins Technology Access Program for laboratories interested in evaluating Sequins in their own ctDNA workflows. The Set is designed for use alongside targeted gene panels with an appropriate Sequins-specific capture probe set and can also serve as a foundation for further customization through the Sequins On-Demand program. To enquire about access, apply via the Sequins website (www.sequins.bio) or contact our team directly at enquiries@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

Disclaimer

IDT and Sequins products are intended for Research Use Only. Not intended for use in diagnostic procedures.