cfDNA Library Prep Kit Datasheet

INTRODUCTION

Next-Generation Sequencing (NGS) of cell-free DNA (cfDNA) has emerged as a promising strategy for various applications, including liquid biopsy-based diagnostics, cancer research, and prenatal screening. However, capturing and converting cfDNA into DNA libraries is challenging, especially when it has low presence within a biological sample. Attaining accurate low-variant detection demands high-performing NGS library preparation and target enrichment protocols.

The Twist cfDNA Library Preparation Kit and the Twist cfDNA Target Enrichment Standard Hybridization workflow address these challenges by offering a solution for the efficient conversion of cfDNA molecules into sequenceable UMI families.

This datasheet details the general performance expectations for cfDNA library preparation. Specifically, we have conducted a thorough analysis that maps the coverage response across different input masses and sequencing conditions to provide guidance on how to achieve a desirable outcome from this workflow.

METHODS

Twist cfDNA Library Preparation Kit was used following manufacturer's guideline. Different input masses of 1 ng, 5 ng, 15 ng, and 30 ng (note: 30 ng exceeds the input guidance described in the manufacturer's instructions) for Twist cfDNA Pan-Cancer Reference Standard (v1, 0.5% VAF) were quantified using the Qubit™ dsDNA High Sensitivity Quantification Assay (Thermo Fisher) and subsequently used as substrates for library preparation. In the ligation and PCR steps, 3 µL of Twist Unique Molecular Identifier (UMI) Adapters and 10 µL of Twist Unique Dual Indexed (UDI) primers were used, respectively. Kapa HiFi HotStart ReadyMix (2X) was used to amplify after ligation. The cfDNA libraries post-elution were quality checked with a DNA 7500 Kit on the Bioanalyzer (Agilent).

After library preparation, cfDNA libraries were hybridized in accordance with the Twist cfDNA Hybridization protocol with variable mass into target capture against a 48 kb oncology panel targeting variant sites in the cfDNA standard material. Kapa HiFi HotStart ReadyMix (2X) was used to amplify after capture. Post-elution capture libraries were quality checked with a DNA High Sensitivity Kit on the Bioanalyzer (Agilent). Capture libraries were pooled and sequenced with 74x74 paired-end reads with at least 80,000x coverage on a NextSeq 550 (Illumina). Capture libraries in Table 2 and Figures 2-4 are sequenced on a NextSeq 2000 (Illumina) to achieve a higher read depth. Error bars or error distributions are derived from the standard deviation of N=2 samples.

UMI deduplication was performed with fgbio CallMolecularConsensusReads[1] using various --min-reads arguments (i.e. 1-0-0, 2-1-1, 4-2-2). Mean target coverage and other sequencing metrics were computed from Picard. Arguments to --min-reads determine how consensus reads are called. A 1-0-0 argument retains singletons, which is ideal for libraries with low duplication rates. On the other hand, a 2-1-1 argument discards singletons and consensus is error corrected with a minimum of 1 Watson and 1 Crick strand, which is ideal for libraries with high duplication rates.

[1] *For additional details please see https://fulcrumgenomics.github.io/fgbio/tools/latest/CallMolecularConsensusReads.html*

RESULTS

LIBRARY PREPARATION YIELD

Obtaining sufficient PCR yield after library preparation is crucial for downstream applications such as target enrichment or whole genome sequencing. Higher yield following PCR at a fixed cycle number is generally correlated with increased initial library complexity. While PCR amplification is necessary for generating sufficient cfDNA for sequencing, excessive PCR cycles can result in amplification bias that distorts the original population and may also introduce sequencing artifacts. Table 1 and Figure 1 provide observed library yields as a reference in support of optimal PCR cycle guidance listed in the Twist cfDNA Library Preparation Kit protocol.

This data also shows the potential pitfalls of overamplification where heteroduplexes can form from inefficient PCR cycles, resulting in potential quantification underestimation[2]. The effect of underestimating mass can be inferred from later capture data presented (see Figure 2, Figure 3, and Table 2).

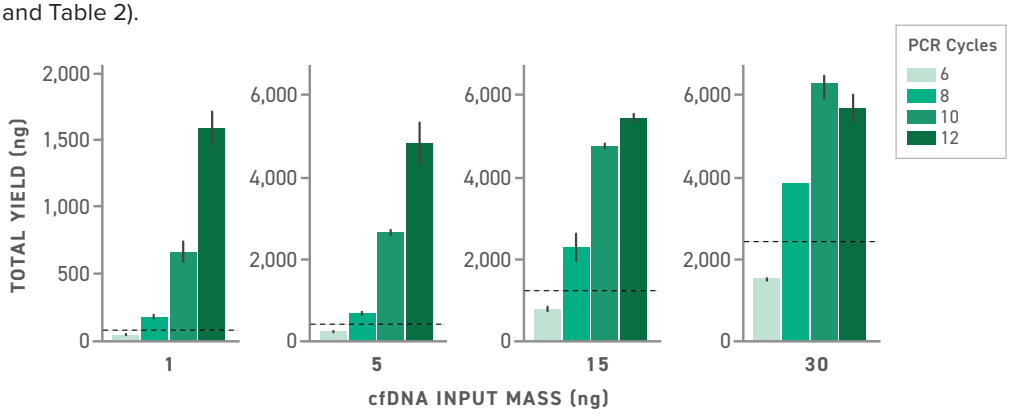

Figure 1. Total NGS Library Yield (ng) after performing Twist cfDNA Library Preparation Kit protocol with various cfDNA input masses. dsDNA concentration after library preparation elution was quantified with Qubit dsDNA Broad Range Quantification Assay (Thermo Fisher). Total yield is calculated by dsDNA concentration (ng/µL) * total elution volume (µL) and reported above. For each specific cfDNA input mass, Twist provides guidance and minimum requirements for input mass into target enrichment and is denoted by the dotted lines in each graph.

cfDNA INPUT MASS (ng)	REQUIRED YIELD (ng)	6 CYCLES (ng)	8 CYCLES (ng)	10 CYCLES (ng)	12 CYCLES (ng)
1	80	**45 ± 3**	180 ± 20	700 ± 100	1600 ± 200
5	400	**216 ± 6**	680 ± 40	2640 ± 60	4800 ± 700
15	1200	**760 ± 50**	2300 ± 500	4720 ± 60	5400 ± 100
30	2400	**1470 ± 20**	3830 ± 40	6200 ± 600	5600 ± 500

Table 1. Total NGS Library Yield (ng) after performing Twist cfDNA Library Preparation Kit protocol with various cfDNA input masses. The dsDNA concentration after library preparation elution was quantified with Qubit dsDNA Broad Range Quantification Assay (Thermo Fisher). Total yield is calculated by dsDNA concentration (ng/µL) * total elution volume (µL) and reported above. The values highlighted in red represent insufficient mass generated for library preparation.

INPUT MASS TO TARGET ENRICHMENT

Library input mass into target enrichment can affect the complexity of deeply sequenced libraries. Twist suggests employing an 80x factor to determine the minimum library mass required for target enrichment, calculated based on the initial cfDNA sample mass. This approach ensures the effectiveness of the enrichment process is maximized.

$$\textit{Minimum Mass Into TE} = 80 \times \textit{Mass Into Library Prep}$$

[2] *https://www.twistbioscience.com/resources/white-paper/heteroduplexes-affect-library-size-determination-without-impacting-targeted*

Figures 2 and 3 show observed mean target coverages and percent duplexing recovered in response to various input masses at near saturation sequencing coverage. These values serve as a reference guide for titrating input mass into target enrichment.

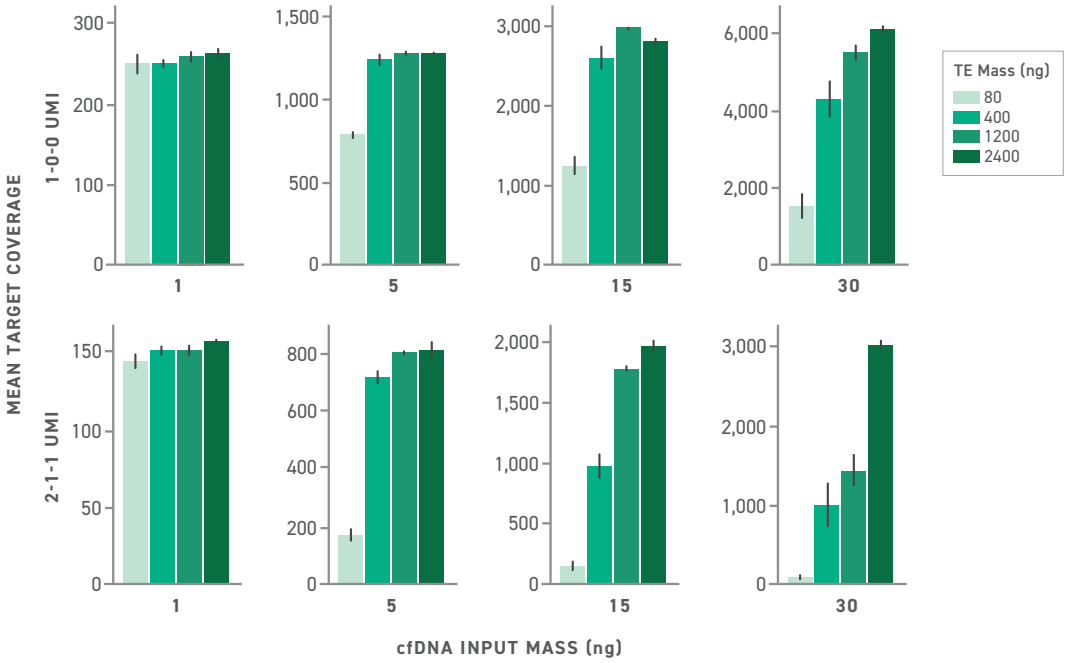

Figure 2. Mean Target Coverages after applying 1-0-0 and 2-1-1 UMI family filters with Twist cfDNA Hybridization workflow over various hybridization input masses. Various input masses were used for target enrichment across a range of 1-30 ng cfDNA sample mass using the Twist cfDNA Target Enrichment Standard Hybridization workflow. Samples were sequenced on a NextSeq 2000 and aligned for Picard metrics after fgbio UMI filtering. For each specific cfDNA sample mass, recovered Mean Target Coverage after 1-0-0 or 2-1-1 duplex deduplication with 80,000x sequencing depth is displayed in response to increasing target enrichment input mass for referencing.

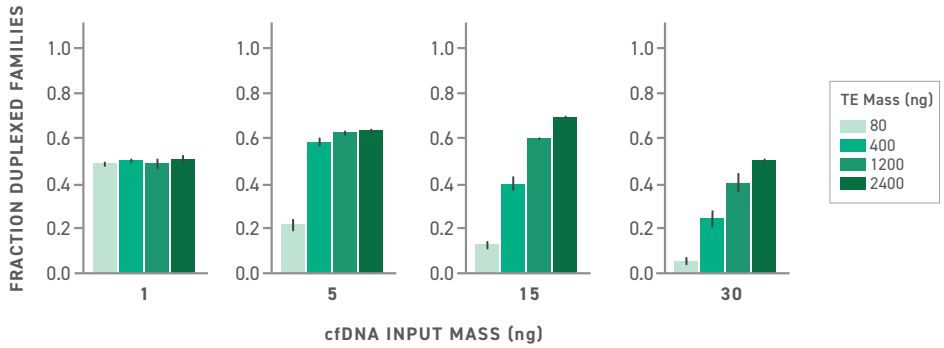

Figure 3. Proportion of all identified families duplexed after applying a 2-1-1 UMI filter with Twist cfDNA Hybridization workflow over various hybridization input masses. Similar to Figure 2, samples were sequenced on a NextSeq 2000 and aligned after fgbio UMI filtering. Duplex sequence metrics were collected after 2-1-1 UMI filtering and the percentages of total families identified that are duplexed are reported. To classify as a duplex, each tag family requires at least one observation of each strand (top and bottom).

The data also underscores the challenges encountered with duplex sequencing when inadequate NGS library mass is introduced into target enrichment. In cases with low capture input masses, the recovery of opposite strands post-capture is inefficient, leading to a notable reduction in target coverage following UMI duplexing. This issue poses a potential obstacle to the broader utility of UMIs in sequencing applications but can be alleviated by following the recommendations in the Twist cfDNA Target Enrichment Standard Hybridization protocol.

cfDNA INPUT MASS (ng)	CAPTURE MASS (ng)	1-0-0 MEAN TARGET COVERAGE	2-1-1 MEAN TARGET COVERAGE	DUPLEX RATE (%)
1	80	250 ± 20	145 ± 6	48.8 ± 0.8
1	400	250 ± 5	151 ± 3	50.3 ± 0.3
1	1200	258 ± 9	151 ± 4	49 ± 2
1	2400	261 ± 8	157.2 ± 0.6	51 ± 1
5	80	780 ± 30	170 ± 30	22 ± 3
5	400	1230 ± 40	720 ± 30	58 ± 2
5	1200	1280 ± 10	804 ± 4	63 ± 1
5	2400	1270 ± 9	810 ± 40	63.3 ± 0.4
15	80	1200 ± 200	160 ± 50	13 ± 2
15	400	2600 ± 200	1000 ± 100	40 ± 4
15	1200	2970 ± 10	1780 ± 20	60 ± 0.2
15	2400	2800 ± 60	1970 ± 60	69.4 ± 0.5
30	80	1500 ± 400	90 ± 50	5.7 ± 2
30	400	4300 ± 600	1000 ± 400	25 ± 5
30	1200	5500 ± 300	1400 ± 300	40 ± 5
30	2400	6130 ± 90	3020 ± 60	50.8 ± 0.2

Table 2. Coverage metrics with Twist cfDNA Target Enrichment Standard Hybridization workflow over various hybridization input masses. Various input masses were used for target enrichment across a range of 1-30 ng cfDNA sample mass using the Twist cfDNA Target Enrichment Standard Hybridization workflow. Samples were sequenced on a NextSeq 2000 and aligned for Picard metrics after fgbio UMI filtering.

SEQUENCING REQUIREMENTS

A large cost driver for any sequencing assay is the number of sequencing reads or coverage required. Having a good understanding of the necessary read amount helps enable assessment and optimization of assay costs.

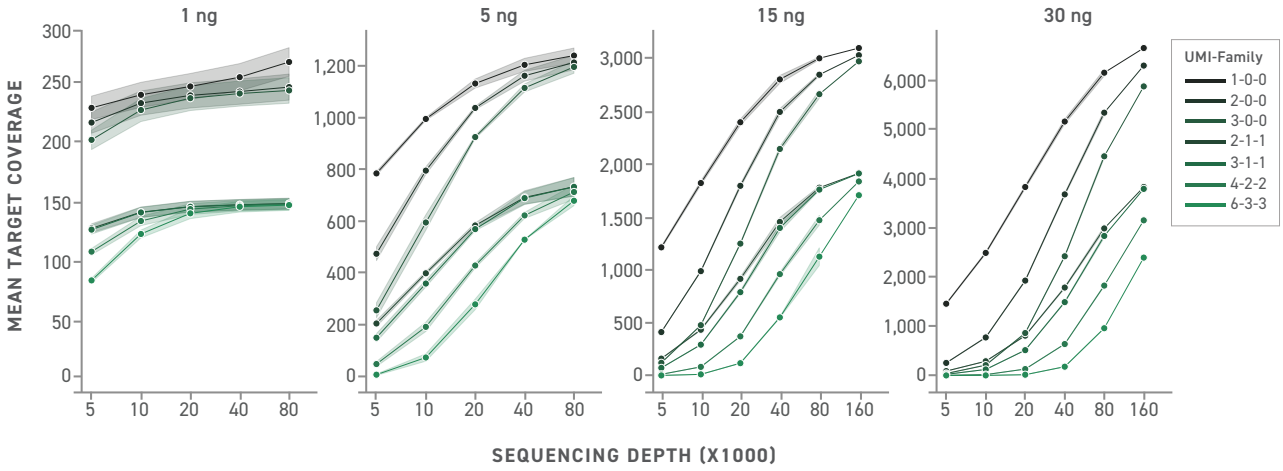

Figure 4. Mean Target Coverage in response to sequencing depth and UMI-family filter. Replicate captures were performed following recommended input mass (80 ng, 400 ng, 1200 ng, 2400 ng) into target enrichment across a set of 1, 5, 15, and 30 ng cfDNA sample mass, respectively. Samples were sequenced on a NextSeq 2000 and aligned for Picard metrics after fgbio UMI filtering.

cfDNA INPUT MASS (ng)	SEQUENCING DEPTH (1000X)	MEAN TARGET COVERAGE						
		1-0-0	2-0-0	3-0-0	2-1-1	3-1-1	4-2-2	6-3-3
1	5	230 ± 10	210 ± 10	200 ± 10	125 ± 5	124 ± 5	105 ± 4	82 ± 3
5	5	790 ± 10	480 ± 40	260 ± 40	210 ± 20	150 ± 20	50 ± 10	9 ± 4
15	5	1210 ± 20	409 ± 1	119 ± 2	159 ± 2	71.5 ± 0.6	10.8 ± 0.2	0
30	5	1450 ± 20	253 ± 4	38 ± 2	87 ± 1	21.7 ± 0.4	1.6 ± 0.2	0
1	10	240 ± 10	230 ± 10	220 ± 10	138 ± 6	138 ± 6	131 ± 5	120 ± 5
5	10	996 ± 8	800 ± 20	600 ± 40	400 ± 10	360 ± 20	190 ± 30	70 ± 20
15	10	1810 ± 30	984 ± 8	474.3 ± 0.8	431 ± 9	290 ± 3	81.0 ± 0.7	10.5 ± 0.6
30	10	2480 ± 30	770 ± 10	208 ± 6	287 ± 4	120 ± 2	16.4 ± 0.3	0
1	20	240 ± 20	240 ± 10	230 ± 10	143 ± 6	143 ± 6	141 ± 6	137 ± 6
5	20	1130 ± 20	1039 ± 9	926 ± 8	590 ± 10	571 ± 9	430 ± 10	280 ± 30
15	20	2390 ± 40	1790 ± 20	1240 ± 10	910 ± 30	790 ± 20	369 ± 4	116 ± 2
30	20	3810 ± 50	1920 ± 20	858 ± 9	805 ± 9	512 ± 5	128 ± 1	14.43 ± 0.03
1	40	250 ± 20	240 ± 10	240 ± 10	145 ± 6	144 ± 6	144 ± 6	143 ± 6
5	40	1210 ± 40	1160 ± 30	1120 ± 20	690 ± 40	690 ± 40	620 ± 20	530.1 ± 0.8
15	40	2790 ± 50	2490 ± 40	2140 ± 30	1450 ± 60	1390 ± 50	960 ± 20	548 ± 7
30	40	5140 ± 70	3670 ± 40	2410 ± 20	1780 ± 30	1480 ± 20	635 ± 6	176.6 ± 0.7
1	80	260 ± 20	240 ± 20	240 ± 20	146 ± 6	145 ± 6	145 ± 6	144 ± 6
5	80	1240 ± 40	1210 ± 40	1200 ± 30	730 ± 50	730 ± 50	710 ± 40	680 ± 30
15	80	2990 ± 20	2840 ± 20	2650 ± 50	1770 ± 20	1750 ± 10	1460 ± 60	1100 ± 100
30	80	6130 ± 90	5320 ± 70	4440 ± 50	2980 ± 60	2820 ± 50	1820 ± 20	956 ± 7
15	160	3087	3020	2962	1906	1905	1829	1700
30	160	6628	6273	5852	3815	3777	3140	2385

Table 3. Mean Target Coverage in response to sequencing depth and UMI-family filter. Replicate captures were performed following recommended input mass (80 ng, 400 ng, 1200 ng, 2400 ng) into target enrichment across a set of 1, 5, 15, and 30 ng cfDNA sample mass, respectively.

MULTIPLEXING

For users who are interested in sample multiplexing to increase throughput, Twist has also validated the input mass recommendations to support its usage. Traditionally, protocols would recommend lowering input mass into a multiplexed capture based on the plexing sizes. Despite initial concerns, careful examination has revealed that decreasing capture input will hurt the efficiency and performance of multiplexed samples as compared to singleplex data. Through experimentation and data provided below, we tested up to 8-plex capture with 15 ng of cfDNA input using a total of 9.6 µg of input NGS library mass. Next, we compared this data with a regular singleplex capture of 15 ng cfDNA input with a capture input mass of 1.2 µg.

A valid concern is that increasing library input may lead to overload and compromise data quality. The verification shows that the optimized approach not only maximizes sequencing output, but it also upholds data fidelity regarding uniformity and off-target performance. In essence, the data reassures that no sacrifice or tradeoffs need to be made when it comes to multiplexing for target enrichment.

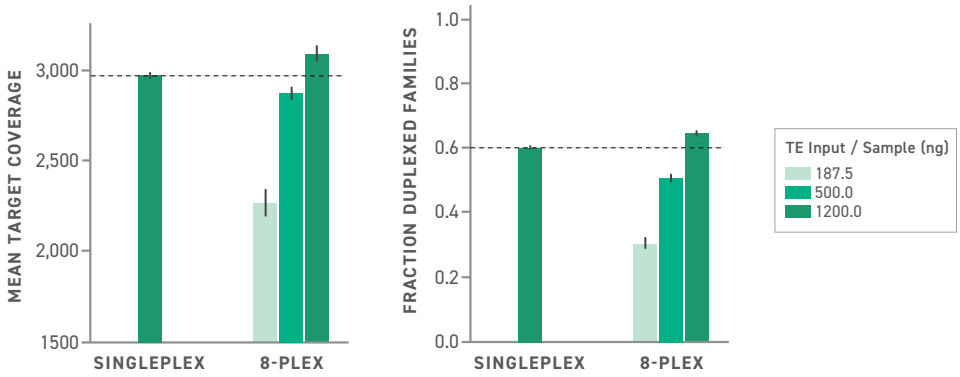

Figure 5. Mean Target Coverages after applying a 1-0-0 UMI family filter with Twist cfDNA Target Enrichment Standard Hybridization workflow over various hybridization input masses with single or 8-plex captures of 15 ng input cfDNA samples. Replicate multiplex captures were performed following input mass of 1200 ng, 500 ng, or 187.5 ng into target enrichment with 15 ng cfDNA sample mass. Individual target enrichment input masses also equate to total 9.6 µg, 4 µg, or 1.5 µg of total input mass. For each metric, a comparison is done against a singleplex capture as denoted by the dotted lines in each graph.

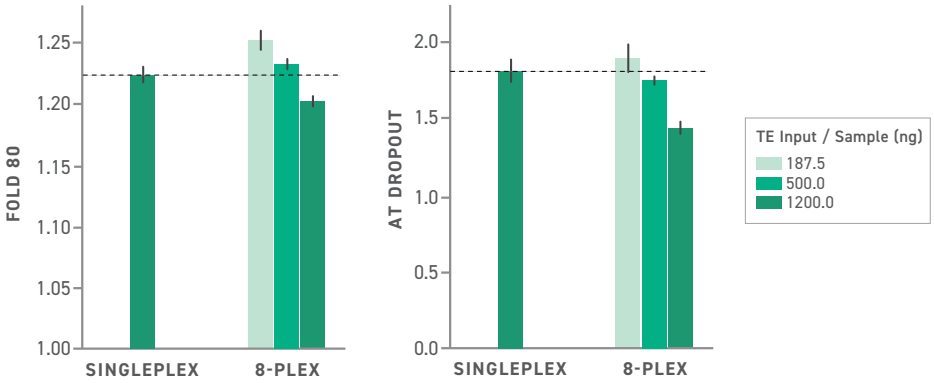

Figure 6. Uniformity metrics after applying a 1-0-0 UMI family filter with Twist cfDNA Target Enrichment Standard Hybridization workflow over various hybridization input masses with single or 8-plex captures of 15 ng input cfDNA samples. Replicate multiplex captures were performed following input mass of 1200 ng, 500 ng, or 187.5 ng into target enrichment with 15 ng cfDNA sample mass. Individual target enrichment input masses also equate to total 9.6 µg, 4 µg, or 1.5 µg of total input mass. For each metric, a comparison is done against a singleplex capture as denoted by the dotted lines in each graph.

cfDNA INPUT MASS (ng)	CAPTURE MASS / SAMPLE (ng)	MULTIPLEX	1-0-0 MEAN TARGET COVERAGE	DUPLEXES (X1,000)	DUPLEX RATE (%)	FOLD 80	AT DROPOUT	GC DROPOUT
15	1200	8-plex	3080 ± 40	1100 ± 20	65.2 ± 0.5	1.21 ± 0.01	1.50 ± 0.05	0
15	500	8-plex	2860 ± 70	760 ± 20	49 ± 2	1.23 ± 0.01	1.75 ± 0.05	0
15	187.5	8-plex	2140 ± 80	330 ± 30	28 ± 1	1.26 ± 0.01	2.07 ± 0.06	0
15	1200	8-plex	3100 ± 100	1070 ± 20	63 ± 1	1.20 ± 0.01	1.37 ± 0.04	0
15	500	8-plex	2880 ± 40	820 ± 20	51.9 ± 0.9	1.23 ± 0.01	1.75 ± 0.05	0
15	187.5	8-plex	2390 ± 60	430 ± 30	32.6 ± 2	1.24 ± 0.01	1.72 ± 0.06	0
15	1200	1-plex	2970 ± 20	960 ± 10	60.0 ± 0.2	1.22 ± 0.01	1.80 ± 0.10	0

Table 4. Post-capture Performance Metrics: 8-plex vs singleplex. Replicate multiplex captures were performed following input mass of 1200 ng, 500 ng, or 187.5 ng into target enrichment with 15 ng cfDNA sample mass. Individual target enrichment input masses also equate to total 9.6 µg, 4 µg, or 1.5 µg of total input mass. Samples were sequenced on a NextSeq 2000 and aligned for Picard metrics after fgbio UMI filtering.

STANDARD HYBRIDIZATION VS FAST HYBRIDIZATION

Users may be interested in leveraging Twist's Fast Hybridization and Wash Kit that enables speed and helps accommodate busy schedules. Twist has conducted experiments comparing the performance of singleplexes in both 1-hour and 4-hour hybridization against our recommended Twist cfDNA Target Enrichment Standard Hybridization workflow. The results revealed remarkably similar performance between the fast hybridization and the recommended workflows, highlighting the versatility of our solutions. Though the differences are subtle, we maintain our recommendation of the Twist cfDNA Target Enrichment Standard Hybridization workflow for those seeking maximum performance from the kit.

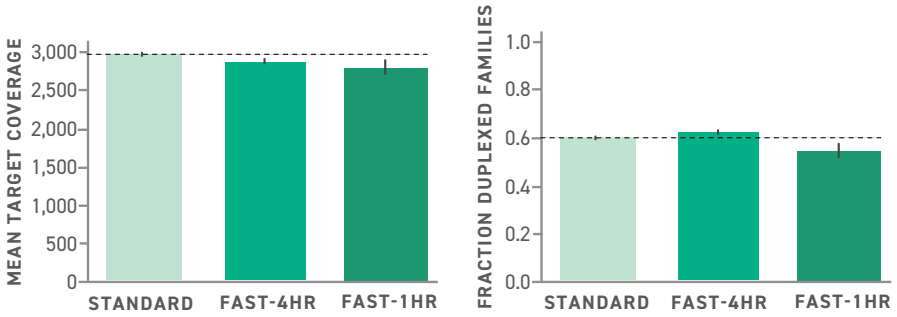

Figure 7. Mean Target Coverages after applying a 1-0-0 UMI family filter with Twist cfDNA Target Enrichment Standard Hybridization vs 1-hr and 4-hr Fast Hybridization workflow. Replicate singleplex captures were performed with target enrichment input mass of 1200 ng with 15 ng reference cfDNA sample mass. For each metric, comparison is focused against singleplex capture with our recommended Twist cfDNA Target Enrichment Standard Hybridization workflow as denoted by the dotted lines in each graph.

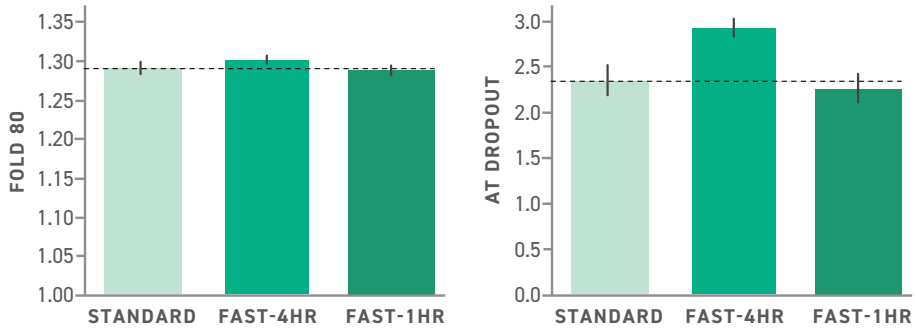

Figure 8. Uniformity metrics after applying a 1-0-0 UMI family filter with Twist cfDNA Target Enrichment Standard Hybridization vs 1-hr and 4-hr Fast Hybridization workflow. Replicate singleplex captures were performed with target enrichment input mass of 1200 ng with 15 ng reference cfDNA sample mass. For each metric, comparison is focused against singleplex capture with our recommended Twist cfDNA Target Enrichment Standard Hybridization workflow as denoted by the dotted lines in each graph.

cfDNA INPUT MASS (ng)	CAPTURE MASS (ng)	HYBRIDIZATION	1-0-0 MEAN TARGET COVERAGE	2-1-1 MEAN TARGET COVERAGE	DUPLEX RATE (%)	FOLD 80	AT DROPOUT	GC DROPOUT
15	1200	Standard Hyb	2970 ± 20	1780 ± 10	60 ± 0.2	1.29 ± 0.01	2.3 ± 0.2	0
15	1200	Fast Hyb - 4hr	2870 ± 20	1810 ± 20	62.2 ± 0.9	1.30 ± 0.01	2.9 ± 0.1	0
15	1200	Fast Hyb - 1hr	2790 ± 100	1500 ± 200	54 ± 3	1.29 ± 0.01	2.2 ± 0.2	0

Table 5. Post-capture Performance Metrics: singleplex with Twist cfDNA Target Enrichment Standard Hybridization vs 1-hr and 4-hr Fast Hybridization workflow. Replicate singleplex captures were performed with target enrichment input mass of 1200 ng with 15 ng reference cfDNA sample mass. Samples were captured and sequenced on a NextSeq 2000 and aligned for Picard metrics after fgbio UMI filtering.

CONCLUSIONS

Implementation of the Twist cfDNA Target Enrichment Standard Hybridization workflow with cfDNA input masses of 1 ng to 30 ng showcases strong performance, particularly with high mean target coverage following UMI deduplication. This advancement not only facilitates throughput with singleplex and 8-plex captures, but also maintains performance with optimized input masses. However, careful consideration is necessary when weighing variables such as input cfDNA sample mass, desired sensitivity at target VAF that is dependent on target coverage, available sequencing resources, and time. This is particularly evident in higher cfDNA input masses, where the required sequencing becomes more prohibitive since the Twist cfDNA Library Preparation Kit's molecular biology is no longer the limiting factor to the increased mean target coverage observed.

LEARN MORE

TWISTBIOSCIENCE.COM/NGS
SALES@TWISTBIOSCIENCE.COM