P21.044: Optimized Library Preparation Kit and Workflow for Improving cfDNA Sequencing Sensitivity

Sean Tighe, Owen Smith, Tong Liu, Tiffany Truong, Tina Han*, Elian Lee, Ramsey Zeitoun, Siyuan Chen

Twist Bioscience, South San Francisco, California, USA *presenting author

Abstract

Next-Generation Sequencing (NGS) of cell-free DNA (cfDNA) has emerged as a promising strategy for early detection, diagnosis, and monitoring of cancer progression. In this process, cfDNA is extracted from a patient's blood and undergoes library preparation prior to NGS. The challenge with cfDNA library preparation is reliably capturing and converting all the fragmented DNA, especially when present in low concentrations within biological samples. Attaining low variant detection thresholds and variant calling confidence demands high-performance NGS libraries and targeted sequencing protocols.

Presented here is a workflow leveraging the Twist cfDNA Library Preparation Kit and an optimized target enrichment protocol to maximize the conversion of duplex, on-target, sequenceable sample molecules. It is demonstrated that by increasing target coverage we can increase detection sensitivity at low variant allele frequencies (VAF). In addition, this workflow improves detection of duplex-molecule families relative to comparable workflows due to more efficient four-point ligation and newly optimized target enrichment. It is also demonstrated that achieving improvements in complexity does not necessitate compromising data fidelity by introducing artifacts or losing uniformity. This improved conversion and sensitivity is applicable to as low as 1ng input samples with both native cfDNA and synthetic cfDNA control.

Straightforward cfDNA Workflow

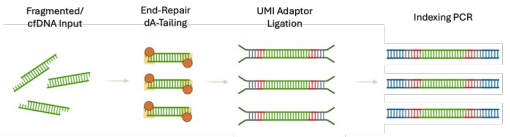


Figure 1. Workflow for Twist cfDNA Library Preparation Kit. With fragmented or cfDNA input, a single end-repair and dA-tailing step converts input DNA into blunt-ended substrates with dA-tails ready for ligation of adaptors with T-overhang. A ligation mastermix is then added on top of the original reaction without the need for purification to attach Twist Unique Molecular Identifier (UMI) adaptors onto the DNA fragments via T/A junction ligation. In the final step, PCR with Twist Unique Dual Index (UDI) primers incorporates sample index sequences for sequencing on Illumina platforms.

High Conversion for UMI and Duplex-UMI Sequencing

- Superior conversion over existing commercial library preparation kits enables more molecules to be sequenced for deeper analysis.
- Higher conversion translates to greater sensitivity and lesser low frequency variants missed in variant calling with both UMI collapse (1-0-0) and Duplex UMI collapse (2-1-1).

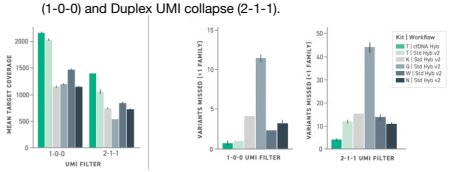


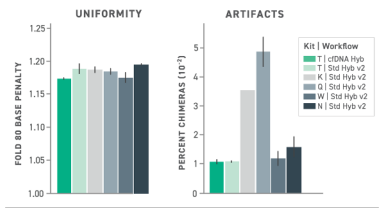
Figure 2. Twist cfDNA Library Prep and Hyb Mix Kit delivers higher library complexity for UMI and Duplex-UMI workflows. NGS libraries were made from 15ng of Twist Pan-Cancer Reference Standards v2 (0.25% VAF) input with various cfDNA kits from vendors in replicates following manufacturer's instructions with 8 cycles of PCR. Libraries were single-plex captured with a 48kb oncology panel targeting variant sites in the cfDNA standards with (Std Hyb V2) or (cfDNA Hyb) TE workflows. Captured libraries were pooled and sequenced 2x74 cycles paired end with at least 80,000x coverage on a Nextseq 500 (Illumina). Picard alignment metrics are reported following UMI collapse with fgbio tool (Fulcrum Genomics). UMI deduplication was performed with fgbio CallMolecularConsensusRead using various --min-reads arguments (i.e. 1-0-0 and 2-1-1).

Figure 3. Twist cfDNA Library Prep and Hyb Mix Kit improves sensitivity and detection of low variant allele frequencies (VAF). Data from capture with the 48kb oncology panel targeting the verified mutations in the Twist Pan-Cancer Reference Standards v2 (0.25% VAF) were used. Variants were identified from post umi-collapse (1-0-0/2-1-1) bams via samtools mpileup targeting variant sites for verified SNPs. Variants missed are reported if less than 1 family of alternative allele was identified.

Reduced Bias and Artifacts

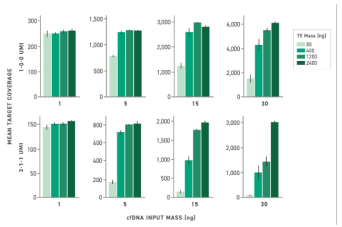
- Improved performance over existing commercial kits comes with no trade-off to data fidelity.
- High conversion maintains uniformity and low chimeric artifacts.

Figure 4. Twist cfDNA Library Prep and Hyb Mix Kit yields high uniformity with minimal artifacts. Data from capture with the 48kb oncology panel targeting the verified mutations in the Twist Pan-Cancer Reference Standards v2 (0.25% VAF) were used. Picard hybrid selection and alignment metrics are reported following UMI collapse with fgbio using 1-0-0 filter.



Target Enrichment Mass Input and Sequencing Characterization

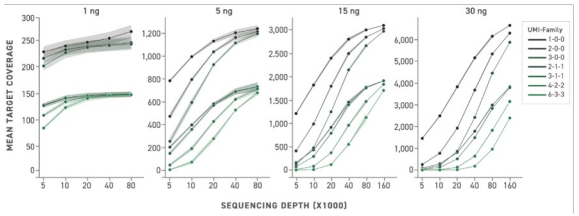
- Mean Target Coverage is influenced by mass input of both library prep and target enrichment



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	140 × 6	48.8 × 0.8
1	400	256 × 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	1290 × 10	806 × 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 × 600	90 × 90	5.7 × 2
30	400	4300 × 600	1000 × 400	25 × 5
30	1200	5500 × 300	1400 × 300	40 × 5
30	2400	6100 × 90	3020 × 60	50.0 × 0.2

Figure 5. Twist cfDNA Library Prep and Hyb Mix Kit UMI and duplex-UMI performance with varying Target Enrichment input mass. NGS Libraries were made from 1ng, 5ng, 15ng and 30ng of Twist Pan-Cancer Reference Standards v1 (0.5% VAF) input with Twist cfDNA Library Prep Kit in replicates with respective recommended PCR cycles and extra PCR cycles to hit high target enrichment mass input yields (i.e. 2400 ng). Libraries were single-plex captured with a 48kb oncology panel targeting variant sites in the cfDNA standards with updated TE. Captured libraries were pooled and sequenced 2x74 cycles paired end with at least 80,000x coverage on a Nextseq 2000 (Illumina). Mean Target Coverage is reported from Picard hybrid selection with fgbio CallMolecularConsensusRead using various --min-reads arguments (i.e. 1-0-0 and 2-1-1).

Table 1. Twist cfDNA Library Prep and Hyb Mix Kit UMI and duplex-UMI target coverage with varying Target Enrichment input mass. Fastqs from post NGS capture with the 48kb oncology panel were downsampled to 80,000x and analyzed. Relationship of Mean Target Coverage corresponding to each cfDNA mass, UMI-filter and capture mass is reported.



- A titration of stringent UMI-duplex filtering highlights a trade-off accuracy between and sensitivity within available sequencing means

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	680 × 40	260 × 40	210 × 20	150 × 20	50 × 10	9 × 6
15	5	1210 × 20	609 × 1	119 × 2	159 × 2	71.5 × 0.6	10.8 × 0.2	0
30	5	1480 × 20	253 × 6	38 × 2	87 × 1	21.7 × 0.6	1.4 × 0.2	0
1	10	240 × 10	230 × 10	220 × 10	138 × 6	138 × 6	131 × 5	120 × 5
5	10	976 × 8	800 × 20	600 × 40	400 × 10	360 × 20	195 × 30	70 × 20
15	10	1810 × 30	986 × 8	474.3 × 0.8	431 × 9	290 × 3	81.8 × 0.7	10.5 × 0.4
30	10	2480 × 30	770 × 10	298 × 6	287 × 6	120 × 2	16.4 × 0.3	0
1	20	260 × 20	260 × 10	230 × 10	143 × 6	143 × 6	141 × 6	137 × 6
5	20	1130 × 20	1039 × 9	926 × 8	590 × 10	571 × 9	438 × 10	280 × 30
15	20	2390 × 40	1790 × 20	1260 × 10	910 × 30	790 × 20	369 × 6	116 × 2
30	20	3810 × 50	1920 × 20	898 × 9	805 × 9	512 × 5	128 × 1	14.43 × 0.03
1	40	250 × 20	260 × 10	260 × 10	143 × 6	144 × 6	144 × 6	143 × 6
5	40	1210 × 40	1160 × 30	1120 × 20	699 × 40	699 × 40	629 × 20	530.1 × 0.8
15	40	2790 × 50	2490 × 40	2140 × 30	1430 × 60	1390 × 50	960 × 20	868 × 7
30	40	5140 × 70	3670 × 40	2670 × 20	1780 × 30	1680 × 20	639 × 6	176.6 × 0.7
1	80	260 × 20	260 × 20	260 × 20	146 × 6	145 × 6	145 × 6	144 × 6
5	80	1260 × 40	1210 × 40	1200 × 30	730 × 50	730 × 50	710 × 40	680 × 30
15	80	2990 × 20	2860 × 20	2650 × 50	1770 × 20	1750 × 10	1440 × 40	1120 × 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	1780
30	160	6628	6273	5852	3815	3777	3140	2385

- Sequencing depth to achieve desired Mean Target Coverage is characterized

Figure 6. Twist cfDNA Library Prep and Hyb Mix Kit UMI and duplex-UMI performance with varying sequencing depth. Libraries were made from 1ng, 5ng, 15ng and 30ng of Twist Pan-Cancer Reference Standards v1 (0.5% VAF) input with Twist cfDNA Library Prep Kit in replicates with 8 cycles of PCR. Libraries were single-plex captured with a 48kb oncology panel targeting variant sites in the cfDNA standards with updated TE workflow. Captured libraries were pooled and sequenced 2x74 cycles paired end with at least 160,000x coverage on a Nextseq 2000 (Illumina). Mean Target Coverage is reported from Picard hybrid selection with fgbio CallMolecularConsensusRead using various --min-reads arguments (i.e. 1-0-0, 2-0-0, 2-1-1, 3-1-1, etc.).

Table 2. Twist cfDNA Library Prep and Hyb Mix Kit UMI and duplex-UMI target coverage with varying sequencing depth. Fastqs from capture with the 48kb oncology panel were downsampled to various sequencing depth (X1000) in reference to panel size and analyzed. Relationship of Mean Target Coverage corresponding to each cfDNA mass, UMI-filter and sequencing depth is reported.

Native cfDNA Concordance

- Twist cfDNA control performance is replicated with native cfDNA
- Improved conversion in native cfDNA samples is demonstrated down to 1ng input

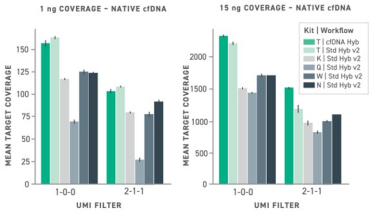


Figure 7. Twist cfDNA Library Prep and Hyb Mix Kit delivers higher Library Complexity for UMI and Duplex-UMI workflows replicated with real native cfDNA samples. Libraries were made from 1ng or 15ng of donor cfDNA samples extracted from plasma with various cfDNA kits from vendors in replicates following manufacturer's instructions with 8 cycles of PCR. Libraries were single-plex captured with a 48kb oncology panel targeting variant sites in the cfDNA standards with present or updated TE recommendations. Libraries captured were pooled and sequenced 2x74 cycles paired end with at least 80,000x coverage on a Nextseq 500 (Illumina). Picard alignment metrics are reported following UMI collapse with fgbio. UMI deduplication was performed with fgbio CallMolecularConsensusRead using various --min-reads arguments (i.e. 1-0-0 and 2-1-1).

Conclusions

The Twist cfDNA Library Prep and Target Enrichment Workflow of 1ng to 30ng showcases strong performance particularly with accurate variant calling and high mean target coverage with and without UMI deduplication. In order to get the most out of this kit, careful consideration is necessary when weighing variables such as library preparation input mass, target enrichment input mass, desired sensitivity at target VAF, and available sequencing resources. The characterizations above should serve as a guide in achieving reliable and expected performance.

Conflict of Interest Statement

All authors are employees and shareholders of Twist Bioscience.