High Performance Multiplexed Target Enrichment Sequencing from FFPE Tissues

Richard Gantt[1], Leonardo Arbiza[1], Kristin Butcher[1], Siyuan Chen[1], Hutson Chilton[1], Mark Consugar[1], Sabina Gude[1], Brenton Graham[1], Jim Laugharn[2], Jayne Simon[2], Ulrich Thomann[2], Christina Thompson[1], Martina Werner[2]

[1]Twist Bioscience; [2]Covaris, Inc.

1. Abstract

Library construction for Next Generation Sequencing (NGS) using formalin-fixed paraffin-embedded (FFPE) samples offers unique challenges in acquiring high-quality sequencing data due to wide distribution of sample quality. Differences in formalin fixation methods, storage conditions, and age lead to crosslinked and/or degraded nucleic acid and inconsistent extraction yields. Therefore, FFPE extraction and library construction methods must be carefully considered for target enrichment applications. In collaboration, Covaris and Twist Bioscience demonstrate a complete library preparation and target enrichment solution that generates ready-to-sequence multiplexed libraries directly from FFPE tissue of various qualities.

This workflow leverages the Covaris truXTRAC® FFPE total Nucleic Acid Plus Kit and AFA-TUBE™ TPX shearing with the world-class performance of Twist Bioscience's Target Enrichment Solutions. Covaris, the Gold Standard for mechanical DNA shearing in NGS applications, offers pre-analytical products that leverage Adaptive Focused Acoustics® (AFA®) technology. In this FFPE-specific application, the Covaris truXTRAC FFPE total Nucleic Acid Plus Kit and AFA-TUBE™ TPX shearing on the LE220-plus Focused-ultrasonicator enables full emulsification of paraffin and disaggregation of tissue for highly efficient nucleic acid extraction and generation of size-specific DNA libraries. With the Twist Bioscience Human Core Exome kit, the resulting libraries are indexed, pooled, and target enriched with uniquely optimized DNA probes to generate ready-to-sequence high quality multiplexed libraries.

Using the aforementioned workflow, results from processing numerous FFPE tissue types and qualities with KAPA Q305/Q41 qPCR ratios ranging from 0.34 to 0.02 are presented. With samples presenting Q305/Q41 ratios ≥0.05, sequencing results of 8-plexed libraries demonstrate large improvements in general Picard metrics that include uniformity (Fold_80 ≤ 1.8), sequencing depth (30X coverage ≥88% with 150X down sampling), and duplication rates (≤11%) when compared to similar published studies. These results demonstrate a validated solution for library preparation and targeted exome sequencing of FFPE samples that can be integrated into automated workflows. The truXTRAC kit and AFA® technology from Covaris generate size specific DNA libraries from FFPE samples that, when paired with Twist Bioscience's superior target enrichment workflow, deliver multiplexed libraries for high performance targeted sequencing.

2. Covaris AFA Technology

Mutation detection-based sequencing is becoming increasingly important in both research and the clinic. Sample preparation is recognized as the limiting factor for sensitivity and specificity of biomarker detection. Adaptive Focused Acoustics® (AFA®) is an advanced acoustic technology enabling the mechanical processing of samples by Focused-ultrasonicators. AFA employs highly controlled bursts of focused high-frequency acoustic energy to efficiently and reproducibly process samples in a temperature-controlled and non-contact environment. This focused and efficient delivery requires a minimal amount of energy input avoiding the adverse effects of excess energy such as damaging heat, experimental variability, and sample over-processing typical of ordinary sonicators.

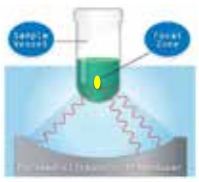
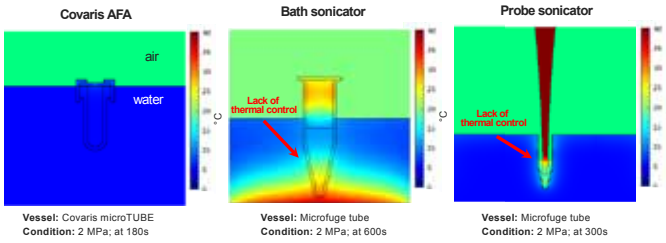


Figure 2.1. Illustrative representation of AFA technology.



Vessel: Covaris microTUBE
Condition: 2 MPa; at 180s

Vessel: Microfuge tube
Condition: 2 MPa; at 600s

Vessel: Microfuge tube
Condition: 2 MPa; at 300s

Figure 2.2 Thermal profile comparison of AFA with probe and bath sonicators. Note the superior thermal profile around the sample with application of AFA.

3. Covaris FFPE Pre-Analytical Products

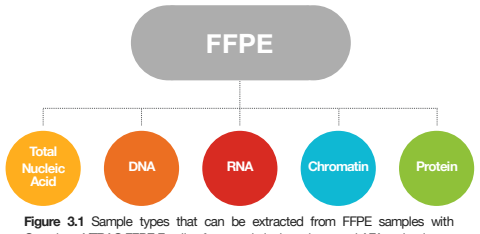


Figure 3.1 Sample types that can be extracted from FFPE samples with Covaris truXTRAC FFPE Family of pre-analytical products and AFA technology.

The truXTRAC FFPE total Nucleic Acid family of kits incorporates the patented AFA technology into the deparaffinization and nucleic acid extraction workflow (Figure 3.1). Fine-tuned AFA energy settings allow RNA and DNA isolation in parallel from the same sample (no splitting), thereby increasing yields and reducing heterogeneity due to separate sample input.

Due to the solvent-free deparaffinization and active extraction process, high quality nucleic acids in sufficient quantity for downstream NGS analysis are obtained. The truXTRAC FFPE total NA Plus Kit is designed for efficient and sequential extraction of total nucleic acids (RNA and DNA) from Formalin-Fixed, Paraffin-Embedded (FFPE) tissue samples using AFA® (Figure 3.2).

Benefits for DNA Extraction:
- Highly amplifiable DNA due to increased fragment length and quality
- Drastically reduced **Q**uantity **N**ot **S**ufficient rates
- More DNA from less tissue - preservation of valuable sample
- Homogeneous extraction guarantees coverage of entire tissue examined

Benefits for RNA Extraction:
- High DV_{200} scores of the extracted & purified RNA
- Drastically reduced **Q**uantity **N**ot **S**ufficient rates
- Exon-shuffling and gene-fusion detection is optimized due to increased length of extracted transcript

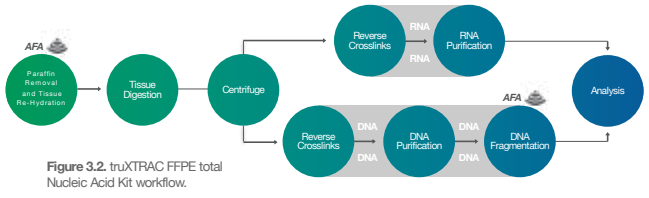


Figure 3.2. truXTRAC FFPE total Nucleic Acid Kit workflow.

4. Twist Target Enrichment Panels

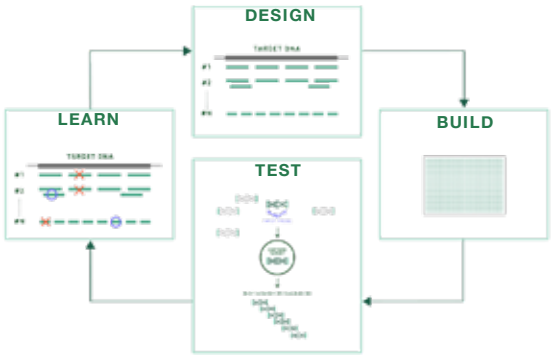


Figure 4.1 Design-Build-Test-Learn: Workflow of design-build-test-learn strategy that was used to generate process for designing a target enrichment system.

A Design-Build-Test-Learn (**DBTL**) strategy was implemented towards developing a framework for generating reproducibly high-performing panels for target enrichment and sequencing (Figure 4.1).

This iterative learning approach requires each step to be performed with reproducible results towards building on results of previous iterations. The reproducibility and expected performance of both the **build** and **test** steps of the DBTL system is presented. The reproducibility data is shown for a representative 800 kb panel consisting of roughly 7,400 probes. Replicates were synthesized 1 month apart.

Build: An NGS quality control step is performed on every custom panel generated where probe representation is measured post-production. This ensures the process completed as expected and the probe content and representation reflects the intended design. Reproducibility between two panels based on NGS probe counting is high and supports DBTL (Figure 4.2).

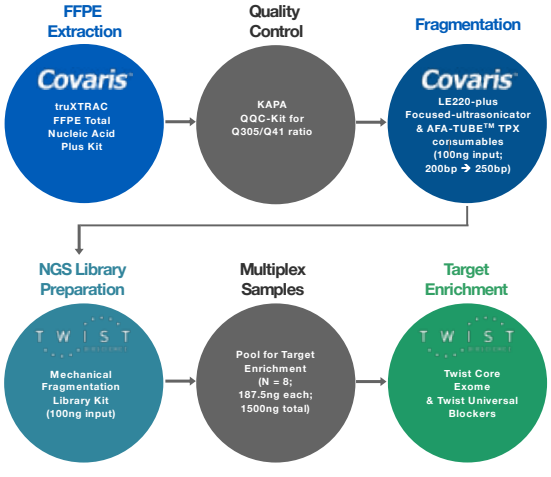
5. Targeted Enrichment of FFPE Samples

Working with large cohorts of FFPE samples presents many challenges including poor yields during extraction, wide distribution of sample quality due to formalin fixation methods, archival storage conditions, and chemical modifications that limit downstream conversion into libraries suitable for NGS. These factors directly impact sequencing quality with lower library diversity, poor uniformity, higher duplication rates, and lower sequence coverage which constrains the number of samples per sequencing run. Additionally, FFPE derived NGS libraries are typically target enriched as single-plex samples due to these factors, limiting high throughput applications.

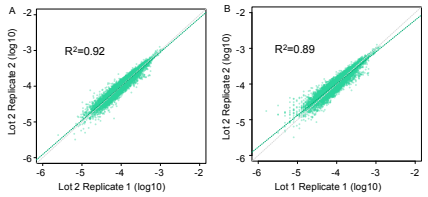
Quantitative scores utilizing qPCR can be utilized to identify low quality FFPE extracted DNA samples and eliminate them from evaluation to optimize sequencing resources. One such kit is the KAPA® hgDNA Quantification and QC Kit (QQC Kit). An independent evaluation of this kit suggests that FFPE extracted DNA with Q305/Q41 ratios <0.2 should not be carried forward for sequencing applications with insert sizes >150bp[1].

Covaris truXTRAC FFPE total Nucleic Acid Plus Kit and AFA-TUBE™ TPX shearing on a LE220-plus Focused-ultrasonicator was combined with Twist Bioscience's Target Enrichment Solutions to address these current limitations around extraction, target capture, and sequencing performance. With only 100 ng of gDNA input required for NGS library creation, this workflow alleviates concerns around FFPE extraction efficiency and variability of library fragment size (**Figure 7.1**). The demonstrated automation friendly workflow enables multiplex target enrichment with high quality sequencing performance on an exome panel (33.1Mb) across a variety of tissue samples with Q305/Q41 ratios well below current recommendations (**Figure 7.2**).

[1]de Abreu, F, *et al.* (AGBT 2015) *The KAPA Human Genomic DNA Quantification and QC Kit Enables Prediction of Sequencing Performance Through User-Defined Metrics*, Marco Island, FL



Experimental Workflow. Multistage workflow combining Covaris AFA Technology for FFPE extraction and fragmentation with Twist Library Creation and Target Enrichment. Various FFPE samples (N = 6) were extracted a single time and their Q305/Q41 ratios determined. Extracted samples were then independently fragmented multiple times to a targeted range of 200bp to 250bp (N ≥ 3) and carried through the remaining steps of the workflow.



Test: An NGS target enrichment experiment to compare probe to probe performance was executed to ensure reproducible capture and testing of the built panel (Figure 4.3). The overall sequencing HS metrics also showed high concordance between lots.

Figure 4.2 Lot to Lot Variability From Build: Each synthesis involves amplification step. A panel containing roughly 7,400 probes (800 kb) was re-synthesized ~1 month apart (Lot1 and Lot2), with two amplification replicates in each Lot (Replicate 1 and 2) **A)** Reproducibility of probe representation within same synthesis, different amplifications. **B)** Reproducibility of probe representation between syntheses.

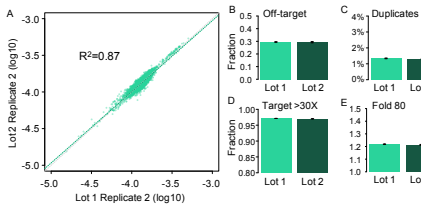


Figure 4.3 Lot to Lot Variability From Test: Data was down-sampled to 150x of target size and analyzed using Picard Metrics with a mapping quality of 20; N = 2. **A)** Lot to lot reproducibility capture per probe. **B-E)** Reproducibility of probe target enrichment performance between syntheses.

Following the optimization of each portion of the cycle the results were used to design high-performance panels in a first attempt. Six panels ranging from 0.02 Mb to 13 Mb were synthesized and shown to have high coverage metrics (30x coverage) which was made possible by a multivariate optimization of key metrics (Figure 4.4).

A

Panel Name	Size (Mb)	Probes	Genes
Mitochondrial DNA	0.02	139	37
Cancer Hotspot	0.04	384	50
Neurodegenerative	0.6	6,024	118
Cancer + Hotspot	0.8	7,446	127
Actionable Cancer	1.7	15,661	522
Pan-Cancer	3.2	31,002	578
Exploratory Cancer	13.3	136,937	5,442

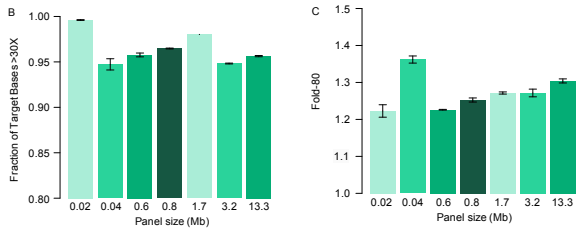
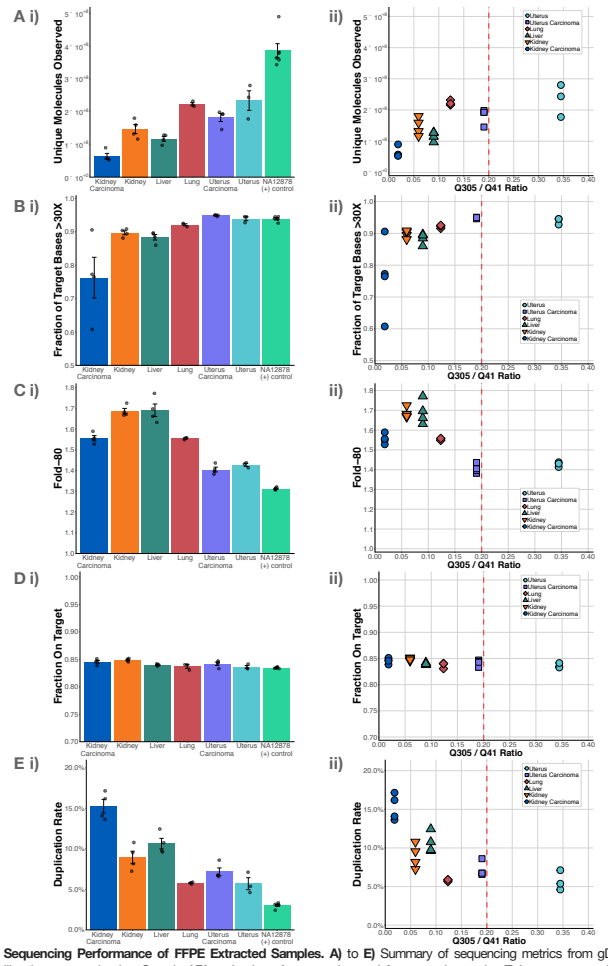


Figure 4.4 First Pass Capture Performance*: Information of capture performance across 6 different panels. **A)** Description of panels and size. **B)** Uniformity (Fold 80) **C)** 30x Coverage performance of each panel as defined by Picard HS metrics. *Hybrid capture was performed using several target enrichment panels (Twist Bioscience) using 500 ng of gDNA (NA12878; Coriell) per single-plex pool following manufacturer's recommendations. Sequencing was performed with a NextSeq® 500/550 High Output v2 kit (Illumina®) to generate 2x76 paired end reads. Data was downsampled to 150x of target size and analyzed using Picard Metrics with a mapping quality of 20; N = 2.



Sequencing Performance of FFPE Extracted Samples. A) to **E)** Summary of sequencing metrics from gDNA libraries prepared using Covaris AFA technology for extraction and fragmentation and a Twist exome panel for capture (see **Figure 7.1**). Experimental workflow was carried out according to manufacturer's recommendations for the respective step (**Figure 7.1**). **i)** Bar graphs of sequencing metrics by tissue type and **ii)** scatter plots of quantitative scores of library integrity (Q305/Q41 ratios; KAPA hgDNA Quantification and QC Kit) versus sequencing metrics. Note that samples with Q305/Q41 ratios of <0.2 (dashed red lines) are typically not recommended for sequencing applications with inserts >150bp[1]. Samples were sequenced a NextSeq 500/550 High Output v2 kit (Illumina) to generate 2 x 76 paired-end reads and downsampled to 150x of targeted bases for evaluation. Picard HS_metrics tools with a mapping quality of 20 were utilized for sequence analysis. Positive control sheared with AFA but not subject to FFPE extraction or Q305/Q41 ratio determination. N ≥ 3 for all observations; error bars denote standard deviation.

8. Summary

- Covaris truXTRAC products and AFA® technology enable high quality extraction and shearing of DNA from FFPE samples.
- Twist target enrichment workflows and panels provide reliable multiplex performance across a wide range of panel sizes.
- Samples with KAPA QQC Q305/Q41 ratios of <0.2 are typically binned as 'low-quality' and not suitable for sequencing.
- Combining Covaris AFA® FFPE extraction products and AFA-TUBE™ TPX shearing with Twist library creation and multiplex target enrichment allows for sequencing of challenging FFPE samples (Q305/Q41 ratios ≥0.05) with uniformity and 30X depth of coverage values that are currently unmatched in the literature for exome-sized target enrichment panels.
- Improved uniformity for sequencing of FFPE samples translates directly to more samples per lane and reduced sequencing costs for a desired depth of coverage.