Development of RNA-Seq library preparation method for highly degraded FFPE RNA and ultra-low RNA input

Ekaterina Star, Noreen Karim, Berlin DeGroen, Michael Salgado, David Wang, Jing Li, Katica Ilic, Sarah Beaudoin, Christopher Vakulskas, Bosun Min, Peilin Chen | Integrated DNA Technologies

Introduction

RNA-Seq library preparation from archived Formalin Fixed Paraffin Embedded (FFPE) samples tends to fail due to both, limited quantity and high degradation. We aimed to overcome the sample quality and quantity limitations by leveraging the unique, single-stranded ligation strategy of the IDT xGen™ cfDNA & FFPE DNA Library Prep chemistry with in-house cDNA synthesis modules. This approach delivers high library conversion of input molecules and maximizes high quality data from sequencing.

Methods

To demonstrate the effectiveness of the RNA-Seq lib prep method using in-house cDNA synthesis modules and xGen cfDNA & FFPE DNA lib prep kit, we collected thirty-seven low quality FFPE RNA from lung, breast, colorectal, thyroid tumors and normal samples. In addition, twenty-one frozen tissue RNA samples and nine RNA reference samples were used as controls. Regardless of sample type, 25 ng of total RNA were taken for RNA-Seq library preparation. In-house cDNA synthesis modules were used for the 1st and 2nd strand cDNA synthesis followed by NGS library generation using IDT's xGen cfDNA & FFPE DNA Library Prep Kit. All sample types generated sufficient high-quality libraries for hybridization capture enrichment. The RNA-seq libraries were captured using IDT's custom designed panel for 56 known fusion targets (Table 1) with xGen Hybridization and Wash V2 kit. Final capture libraries were sequenced on a NextSeq2000.

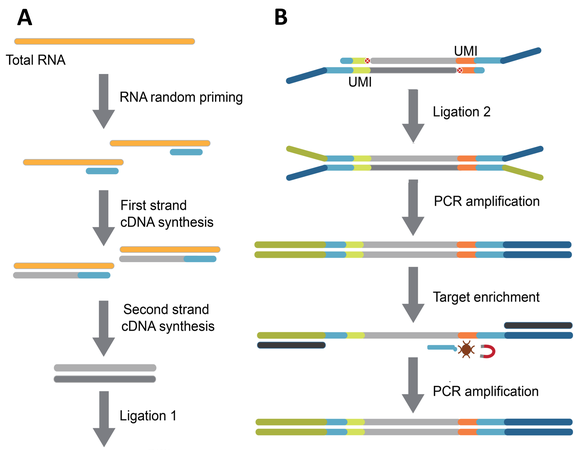

Figure 1. Targeted RNA-Seq library preparation workflow. A. cDNA is generated from total RNA using novel 1st and 2nd strand cDNA synthesis module with a proprietary in-house Reverse Transcriptase. **B.** Double-stranded cDNA is taken through a modified xGen cfDNA & FFPE DNA Library Preparation v2 MC kit workflow without the end-repair step. Ligation 1 Enzyme attaches ligation 1 adapter to the 3' end of the dscDNA. The Ligation 2 Adapter acts as a primer to gap-fill the base complementary to Unique Molecular Identifiers (UMIs), followed by ligation of the 5' end of the cDNA, to create a fully double-stranded product. Finally, PCR amplification with index primers is performed to generate whole transcriptome libraries. For targeted RNA-seq, target regions are enriched using biotin-labeled probes with xGen hybridization capture reagents, followed by PCR amplification.

Results

cDNA libraries constructed from severely degraded samples

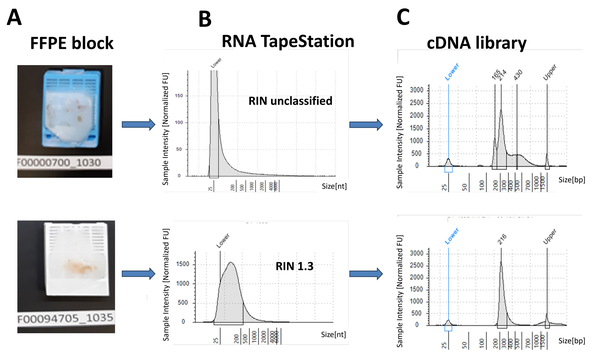

Figure 2. cDNA libraries constructed from severely degraded samples. A. FFPE blocks were sliced into 10 µm curls and RNA was extracted using QIAGEN AllPrep DNA/RNA extraction kit. **B.** Agilent TapeStation High Sensitivity RNA Screen Tape analysis of extracted RNA from FFPE blocks. **C.** cDNA library electropherograms from highly degraded FFPE RNA. 25 ng of total RNA of each sample was used to construct cDNA libraries with our newly developed in-house cDNA synthesis module followed by cfDNA & FFPE DNA Library Prep kit (Figure 1). Libraries were successfully constructed even from the most challenging FFPE blocks with low RNA yield and highly cross-linked samples.

High library yields for low quality FFPE samples

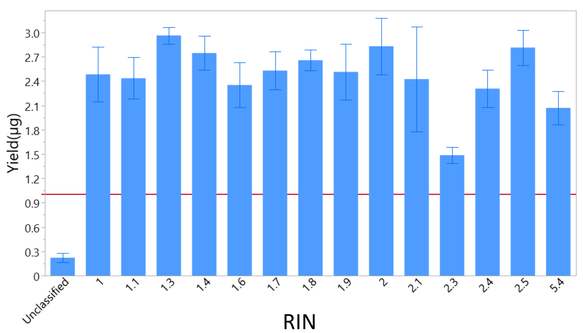

Figure 3. High cDNA library yields for low quality FFPE samples. cDNA libraries were constructed from total RNA of 37 FFPE samples with low RIN scores. All samples with detectable RIN values achieved over 1 µg of library yield with 14 PCR amplification cycles. The sample with very low RNA conc. post extraction and unclassifiable RIN (Figure 2) generated enough cDNA library for downstream applications.

Abundant gene detection and mapping for RNA samples with diverse quality

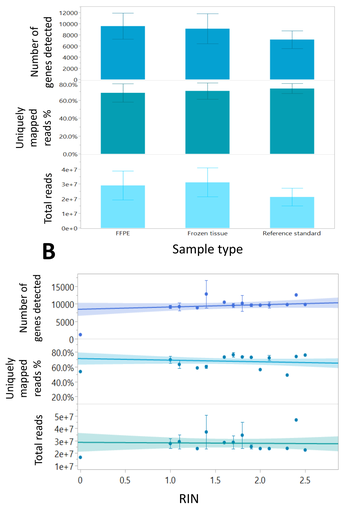

Figure 4. FFPE samples cDNA libraries exhibit similar gene detection and mapping reads as the higher quality ones. All cDNA libraries were captured using IDT's xGen Hybridization and Wash kit V2, and IDT's custom designed panel with 56 known fusion targets. Total cDNA library yield was used for overnight hybridization, followed by capture and amplification steps. Libraries were sequenced on a NextSeq2000. **A.** Sequencing metrics: All libraries had a similar high % of uniquely mapped reads and we detected similar diversity in transcripts between sample types **B.** RIN and number of genes detected: The low-quality samples (RIN 0-3) libraries demonstrated relatively good complexity that varied with individual samples but didn't seem to correlate with the RIN quality score.

Ultralow RNA input

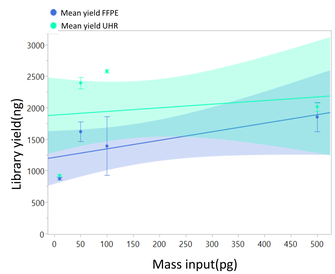

Figure 5. Library yields constructed from as low as 10 pg of total RNA. High quality Human Universal Reference (UHR) RNA (Agilent CN 750500), and a low quality FFPE RNA sample (RIN 2.2) were used to construct RNA-Seq libraries with inputs between 10-500 pg. Libraries were quantified using Qubit High Sensitivity dsDNA assay.

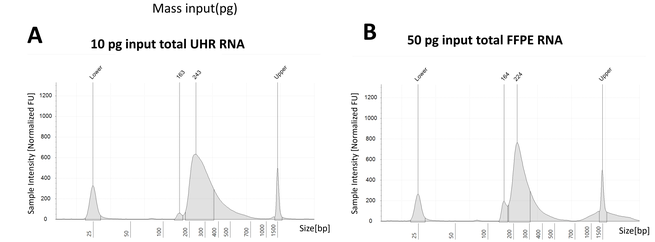

Figure 6. Electropherogram of cDNA libraries generated from ultra low RNA input. A. Library generated from 10 pg high quality UHR RNA. **B**. Library generated from 50 pg of FFPE RNA with RIN 2.2.

Detection of fusion transcripts in targeted RNA-seq

ABL1	*BRCA1*	*ERG*	*FGFR1*	*JAK2*	*MSH2*	*NTRK2*	*PPARG*
AKT3	*BRCA2*	*ESR1*	*FGFR2*	*KDR*	*MYC*	*NTRK3*	*RAF1*
ALK	*CDK4*	*ETS1*	*FGFR3*	*KIF5B*	*NOTCH1*	*PAX3*	*RET*
AR	*CSF1R*	*ETV1*	*FGFR4*	*KIT*	*NOTCH2*	*PAX7*	*ROS1*
AXL	*EGFR*	*ETV4*	*FLI1*	*KMT2A*	*NOTCH3*	*PDGFRA*	*RPS6KB1*
BCL2	*EML4*	*ETV5*	*FLT1*	*MET*	*NRG1*	*PDGFRB*	*TMPRSS2*
BRAF	*ERBB2*	*EWSR1*	*FLT3*	*MLLT3*	*NTRK1*	*PIK3CA*	*TP53*

Table 1. Fusion genes included in IDT's custom panel target space. Total of 56 known fusion genes included in custom designed xGen Hybridization Capture panel.

Sample Type	Total Fusion events	
	Validated fusion events	non-validated fusion events
FFPE	4	33
Reference Standard	7	2
Frozen Tissue	14	7

Table 2. Fusion events detected in Hybridization Captured RNA-Seq libraries. All constructed libraries were captured using IDT's custom designed panel with 56 known fusion targets. Non-subsampled reads were analyzed for fusion detection using STAR-fusion analysis and further filtered using FusionInspector and in-house pre-trained model. All fusions in the reference standard samples that were in the panel target space were detected. Additional previously validated fusion events were detected in the FFPE and frozen tissue samples. Many of the fusion events detected in FFPE and tissue samples need to be validated with a different method.

Conclusions

1) We demonstrate a highly effective workflow for constructing libraries from severely degraded FFPE RNA samples by combining in-house developed 1st strand cDNA synthesis module using in-house proprietary Reverse Transcriptase with IDT's high conversion rate xGen cfDNA & FFPE DNA Library Prep Kit workflow. The method would rescue degraded FFPE RNA samples for RNA-Seq.

2) The RNA-Seq library prep workflow can handle ultralow input and generate libraries from as low as 10 pg of high-quality RNA and 50 pg of degraded FFPE RNA, which could be leveraged for single cell and cfRNA applications.

3) Combining the RNA-Seq library prep with downstream target enrichment using IDT's xGen Hybridization and wash kit V2 provides a complete and flexible solution for targeted RNA-Seq.

Ekaterina Star
Research Scientist,
Integrated DNA Technologies
estar@idtdna.com