Utilization of FFPE Derived Total RNA for Targeted or Whole Transcriptome RNA-seq using SureSelect Strand Specific RNA Library Preparation Kit

Application Note

Author

Alex Siebold, Ph.D. and
Bahram Arezi, Ph.D.

Abstract

In this application note, we present modifications that should be considered when processing FFPE RNA using Agilent's SureSelect strand specific RNA library preparation kit for targeted RNA-seq. These guidelines were used to process two very low-quality FFPE RNA samples in parallel with matched snap-frozen controls using a custom SureSelect capture < 3 Mb in size. Our data shows that with a few minor adjustments to the protocol, the SureSelect strand specific library preparation kit can be successfully used to generate libraries from highly fragmented FFPE samples.

Introduction

The ability to extract RNA and prepare RNA sequencing (RNA-seq) libraries from Formalin Fixed Paraffin Embedded (FFPE) tissues opens the door to a host of samples that can be used to discover and/or better characterize coding and non-coding transcripts. Similar to DNA isolated from FFPE tissues, RNA derived from FFPE can vary significantly in quality depending on the type of tissue that was preserved, how it was preserved, and how long and/or well it was stored[1]. In short, quality control (QC) of the starting material is particularly critical when working with RNA from FFPE samples for next generation sequencing applications. Based on the integrity of FFPE-derived RNA, the appropriate modifications can be made to Agilent's existing protocols to create the highest quality libraries for whole transcriptome and targeted RNA sequencing. Critical adjustments for FFPE-derived RNA are discussed in detail below, and include how much starting material to use, how long to shear/fragment the samples (if at all), and how to adjust the PCR conditions and SPRI bead clean up steps to maximize the amount of material that is being carried through the workflow.

Agilent Technologies

Recommended methods for quality control and processing of FFPE RNA with Agilent's strand specific library preparation kit

1- Quality control of FFPE RNA starting material. RNA integrity numbers (called a RIN or RINe), generated by either the Agilent 2100 Bioanalyzer or the Agilent 2200 TapeStation systems, respectively, are well-accepted standards for total RNA sample quality. When working with high-quality RNA, Agilent SureSelect protocols recommend that the RNA samples have a minimum RIN or RINe value of 8.0. In contrast, RIN/RINe values obtained from FFPE-derived RNA samples typically fall well below these minimum guidelines. In this case, RNA integrity can be assessed by determining fragment size distribution using the 2100 Bioanalyzer or 2200 TapeStation system. As guidance, when using FFPE RNA in Agilent's RNA-seq workflow, the average fragment size should be at least 100–125 nt. Samples with degradation products below this threshold will not work well since SPRI beads are used for all the sample clean up steps and the Agencourt AMPure XP beads that Agilent recommends will only retain fragments > 100 nt in length[2]. In addition to performing the sizing and RIN/RINe assessments using the 2100 Bioanalyzer or the 2200 TapeStation system, both FFPE and non-FFPE RNA samples should also be measured for sample purity using spectrophotometric analysis (NanoDrop®, for example) to obtain 260/280 and 260/230 ratios. For RNA, these ratios should both be as close to 2.0 as possible[3].

2- Guidelines for starting input amounts. The QC measurements outlined above are important for estimating starting input amounts for RNA-seq. The SureSelect RNA-Seq library prep kit protocol recommends starting with 1 µg of high quality RNA (RIN/RIN^e > 8) for the ribosomal RNA depletion step (note PolyA selection should not be used when working with FFPE RNA samples, refer to section 3 below for details). Starting input amounts <1 µg have not been tested by Agilent since these are below the suggested input amounts of our recommended ribosome depletion workflow, referenced below. In general, the more highly degraded the sample appears, the more starting material should be used to generate libraries of sufficient complexity. For FFPE-derived RNA samples, input amount should be adjusted upward (e.g., 2 µg, 4 µg, or higher) and ideally, determined empirically for each RNA sample as quantity, cost, and time allows.

3- Removal of ribosomal content. Performing PolyA selection using oligo(dT) magnetic beads, or through the use of oligo(dT) primers during cDNA synthesis, are the most popular methods for selecting mRNA that will comprise the library preparation. Agilent's strand specific RNA library prep kit includes oligo(dT) beads and the necessary reagents to perform this purification step with high-quality (non-FFPE) RNA prior to library preparation. PolyA selection is inefficient with degraded FFPE-derived RNA and typically introduces a strong 3' bias in the final sequencing data. Instead, Agilent recommends purifying FFPE mRNA using a ribosomal depletion kit, such as EpiCentre's Ribo-Zero™ Gold kit. According to the manufacturer, a single Ribo-Zero Gold reaction can remove > 99% of 28S, 18S, and 5.8S and > 70% of 5S rRNAs, and > 99% of 12S and 16S mitochondrial rRNAs, from up to 5 µg of total human RNA as assessed by RT-qPCR[4]. When using a Ribo-Zero™ Gold kit to remove ribosomal content from your FFPE material, be sure to select the right kit based on the species of RNA you are working with as these kits are species specific. For additional guidance, refer to the application note titled: Whole transcriptome analysis using Agilent Strand Specific RNA Library Preparation Kits" which can be downloaded from the following link:
http://www.chem.agilent.com/library/quickreference/Public/Whole-transcriptome-analysis--Strand-Specific-RNA-Library-Prep-Kits.pdf

4- RNA fragmentation. RNA shearing is the first step in the library preparation workflow. In the SureSelect RNA-Seq kit, fragmentation of high-quality RNA is achieved by treatment with metal ions at elevated temperature (8 minutes at 94°C). Figure 1 (adapted from TruSeq's Library Prep Protocol) shows median fragment length as a function of incubation time at 94°C, where insert length is determined after clustering and paired end sequencing. Careful consideration is required when performing this step with FFPE-derived RNA. In many cases, the FFPE-derived RNA is already degraded to a point where either no fragmentation is needed or an abbreviated shearing protocol should be run. The two bullet points below outline these two strategies that can be used for highly (< 300 bp) or partially (> 300 bp) degraded FFPE RNA samples.

- **Highly degraded samples (fragments ≤ 300 nt)**. Following ribosomal RNA depletion of FFPE-derived RNA, analyze the distribution of fragment sizes using the 2100 Bioanalyzer or 2200 TapeStation system. If the bulk of the fragments are ≤ 300 nt, replace the shearing step with a denaturation step outlined in Table 1 to denature RNA secondary structure and permit annealing of the cDNA synthesis primers.
- **Partially degraded samples (fragments > 300 nt)**. If the RNA fragment size appears to be > 300 nt in size, then an abbreviated fragmentation reaction should be run. As an example, FFPE-derived samples with fragments in the range of 500-800 nt may only need to incubate at 94 ˚C for as few as two minutes followed by a hold at 4 ˚C, whereas samples with larger starting fragment sizes may benefit from a 3 to 6 minute incubation, prior to the hold at 4 ˚C. The exact duration of the incubation step at 94 ˚C should be determined empirically by the user. Ideal fragment sizes to aim for following the fragmentation reaction for RNA should be between 150–200 nt in size.

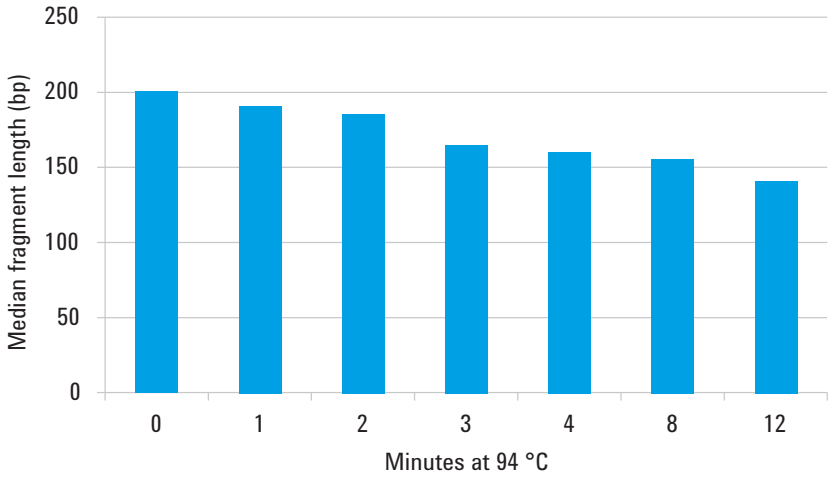

Figure 1. *RNA fragment length decreases as incubation time at 94°C increases.*

Table 1. *Denaturation protocol for highly degraded RNA samples (Replaces the fragmentation step in SureSelect RNA-Seq library prep protocol).*

Step	Temperature	Time
Step 1	65˚C	5 minutes
Step 2	4˚C	Hold

5- Adjustments to SPRI bead clean up steps. The SureSelect RNA-Seq kit employs several SPRI bead clean up steps with either a 1.8:1 or 1.2:1 SPRI bead volume to sample volume ratio. For example, if the volume of cDNA to be cleaned up is 50 µl, 1.8 times that volume is 90 µl, which is the volume of SPRI beads that should be used to ensure maximal recovery. A 1.2:1 ratio is used for SPRI bead clean up steps that occur following the adapter ligation and PCR of adapter-ligated products to ensure removal of smaller contaminants such as adapter dimers, primer dimers, and PCR primers. However, when working with FFPE RNA samples, particularly highly degraded ones such as those shown in Figure 2, all SPRI bead clean up steps should utilize a 1.8:1 ratio to limit loss of small RNA fragments at or below the binding limits of SPRI beads (< 100 nt in size[2]).

6- PCR cycling conditions. The final variable that can be adjusted to improve performance with FFPE-derived RNA is PCR cycle number. The SureSelect RNA-Seq kit protocol recommends 9-11 cycles for pre-cap PCR during the targeted RNA-seq workflow when starting with 2–4 µg of high quality RNA. When FFPE RNA is used, the number of PCR cycles should be increased to 12–14 cycles to achieve sufficient yield for target enrichment, which requires a minimum of 100 ng.

Results

A) RNA samples.

We used the recommendations provided above to prepare RNA-Seq libraries from two very low quality FFPE-derived RNA samples. Note that for this particular study, the quality of FFPE RNA used is likely the lowest quality that should be considered, especially if targeted enrichment is to be performed. We intentionally used poor quality RNA to illustrate the minimum performance limits of the SureSelect RNA-Seq kit. RNA was isolated and processed in parallel from snap-frozen and FFPE tissue samples. As shown in Figure 2, RIN scores from the snap-frozen tissue samples (7.0, 5.4) were below the RIN value of 8 that Agilent SureSelect recommends. The RIN values for FFPE-derived RNAs were 2.40 and 2.20, and inspection of 2100 Bioanalyzer traces showed the majority of fragments were slightly greater than 100 bp in size. Based on QC metrics, we adjusted the protocol for FFPE RNA as follows: 4 µg input RNA, 5 minute denaturation @65°C (no shearing step), and 1.8:1 SPRI: sample ratio for all clean up steps. The standard kit protocol was followed for the snap-frozen samples, except that 2 µg total RNA went into ribosomal depletion.

The QC traces shown in Figure 3 were generated from one technical replicate for each FFPE sample and its matched snap-frozen control. These traces are from the QC step which follows amplification of the adapter ligated libraries and was run using the D1K* ScreenTape assay on the 2200 TapeStation system. Based on the input amounts used (4 µg of total RNA) for each FFPE sample, the cycling guidelines in the protocol suggest that 9–11 cycles be performed. However, due to the high level of degradation in these FFPE samples, and the likely loss of much of that material due to the size select limits of the SPRI beads used, 14 cycles of PCR were performed for both FFPE and snap-frozen samples during the pre-capture library prep protocol and 12 cycles were run on all samples for the post-capture amplification step.

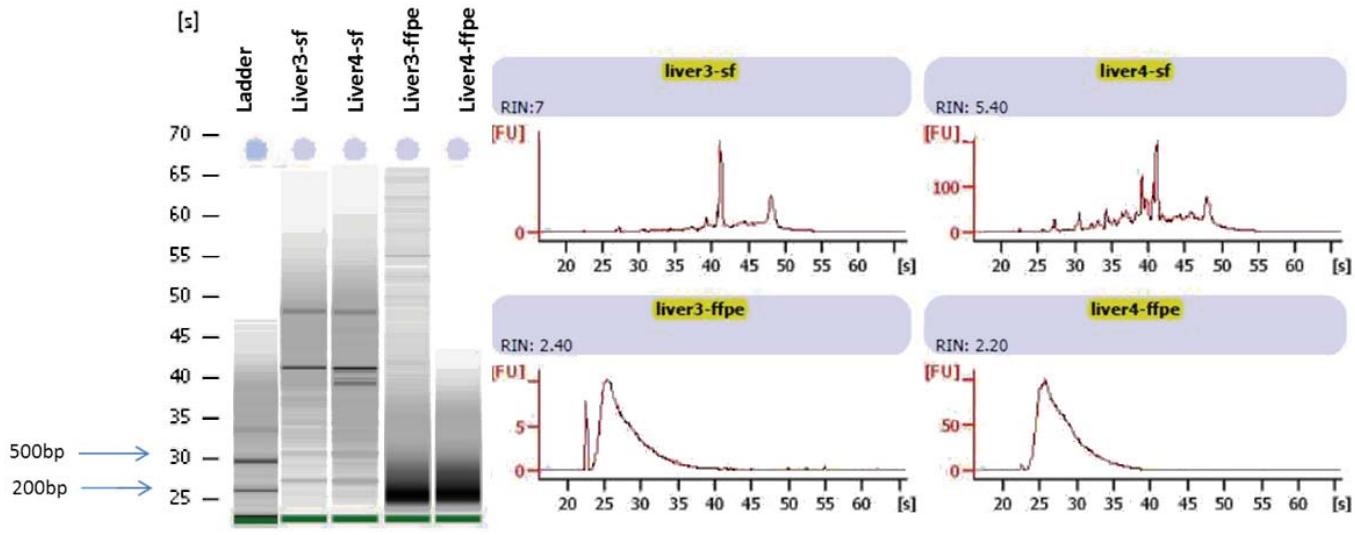

Figure 2. *Total RNA from matched snap-frozen liver and FFPE derived liver tissue were isolated using the Promega Maxwell and analyzed using the Bioanalyzer RNA 6000 Nano assay. Electropherograms and gel images are shown. RIN values of the SF samples were 7 and 5.4. FFPE RIN values were 2.4 and 2.2. The size distribution of the FFPE samples is slightly larger than 100 bp as shown by the gel image.*

*The D1K ScreenTape assay has been obsoleted. Comparable results will be obtained on the D1000 ScreenTape assay (part numbers 5067-5582 and 5067-5583).

Total library yields from FFPE and snap-frozen samples were sufficient for whole transcriptome sequencing. However, to meet the input requirements of 100 ng of pre-capture library per hybridization reaction, the technical replicates from the Liver4 FFPE sample were pooled together. Furthermore, the Liver3-FFPE replicate that yielded 134 ng of pre-capture library was selected for target enrichment. Once this portion of the workflow was completed, the remaining steps in the SureSelect protocol for targeted RNA sequencing were followed (see current protocol available for download from Agilent's Genomics website). Table 2 summarizes the modifications made in the library prep protocol and also displays the yields from the library preps for the snap-frozen and FFPE RNA samples in used in this study.

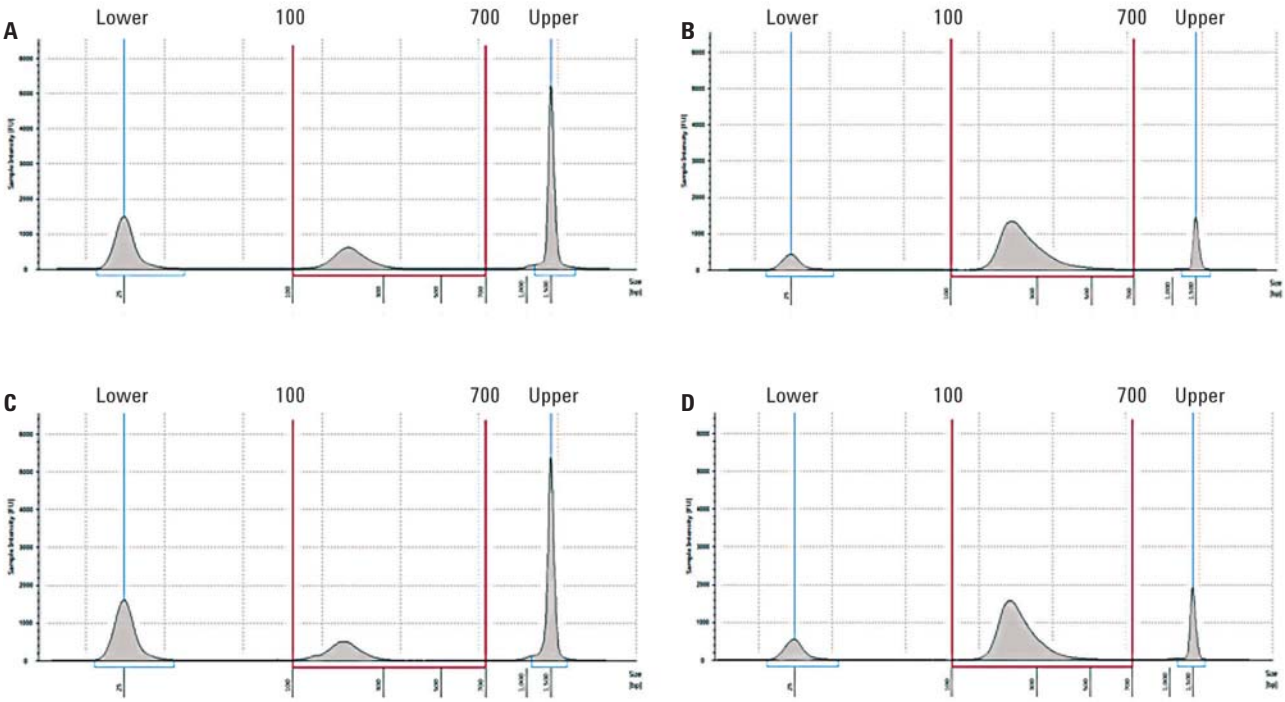

Figure 3. *2200 TapeStation traces from the D1K ScreenTape assay of amplified adapter-ligated FFPE sequencing libraries. Panels A and C are libraries derived from the FFPE liver 3 and liver 4 samples. Panels B and D are the libraries derived from the matched snap-frozen matched (non-FFPE) liver 3 and liver 4 controls. The traces display the amplified adapter-ligated libraries following SPRI bead cleanup after 14 cycles of PCR were run. Total library yields from FFPE derived RNA samples ranged from 59.5 to 134 ng. Yields from the two snap-frozen matched controls were 990 ng for liver 3 and 777.5 ng for liver 4.*

Table 2. *Quick guide to protocol modifications made to prepare libraries from FFPE samples and PCR yield data.*

Sample	RIN	Input into Ribo-Zero (µg)	Fragmentation time/temp	SPRI:DNA ratio after ligation and PCR	Pre-cap yield (ng) replicates show when applicable	Pre-cap PCR cycle#	Post-cap yield (nM)
Liver3-SF	7	2	8min @ 94°C	1.2:1	990	14	12.9
Liver4-SF	5.4	2	8min @ 94°C	1.2:1	777.5	14	9.16
Liver3-FFPE	2.4	4	5min @ 65°C	1.8:1	98, 134	14	N/A, 4.62
Liver4-FFPE*	2.2	4	5min @ 65°C	1.8:1	59.5, 77	14	1.71

*Pooled technical replicates of this sample to yield >100 ng of pre-capture library, required for target enrichment.

B) Sequencing performance metrics.

1- Strand specificity, target enrichment efficiency, and % duplicates (library complexity). A SureSelect bait library of 51229 probes at 2.88 Mb was used for all captures. Number of reads from all four libraries were normalized to 2.88 million/library (2 x 100 bp sequencing) for comparison purposes. As Figure 4 shows, strand specificity (i.e., percent reads on the expected strand) at > 98% is similar between snap-frozen and FFPE libraries. Percent reads in targeted regions are also very similar among all libraries, ranging from 76.4%-76.8% for the liver 3 and 4 FFPE libraries, to 81.2%-81.5% for liver 3 and 4 snap-frozen libraries, respectively. Percent duplicates, however, are significantly higher for FFPE libraries at 74–76% for liver 3 and 4 compared to 36–41% for the snap-frozen counterparts. These results are not unexpected though, since the FFPE samples that were processed were highly fragmented. Furthermore, we determined the estimated library size (Figure 5) for all libraries, which is the number of expected unique fragments based upon the total number of reads and duplication rate assuming a Poisson distribution. Not surprisingly, and consistent with duplication rates, the estimated library size for liver 3 and 4 FFPE samples are ~5-fold less than their snap-frozen counterparts. Overall, we were able to detect ~4300 transcripts for liver 3 and 4 FFPE and ~5600 for snap-frozen libraries.

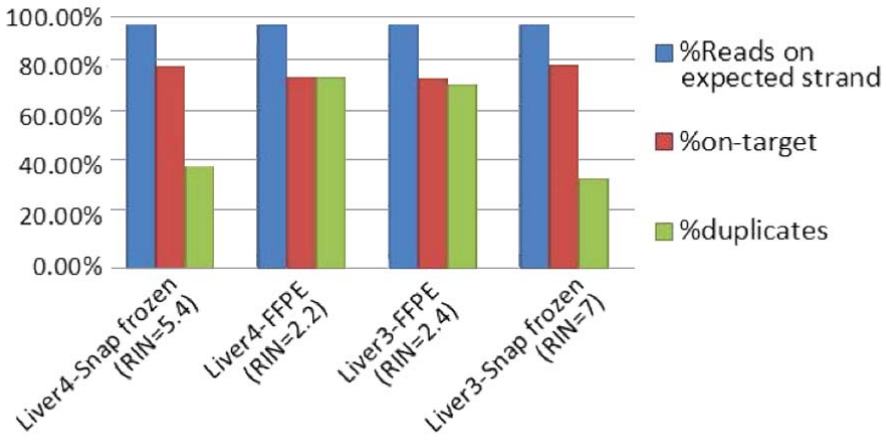

Figure 4. *Strand specificity and target enrichment efficiency of the snap-frozen and FFPE derived libraries that were prepared and sequenced.*

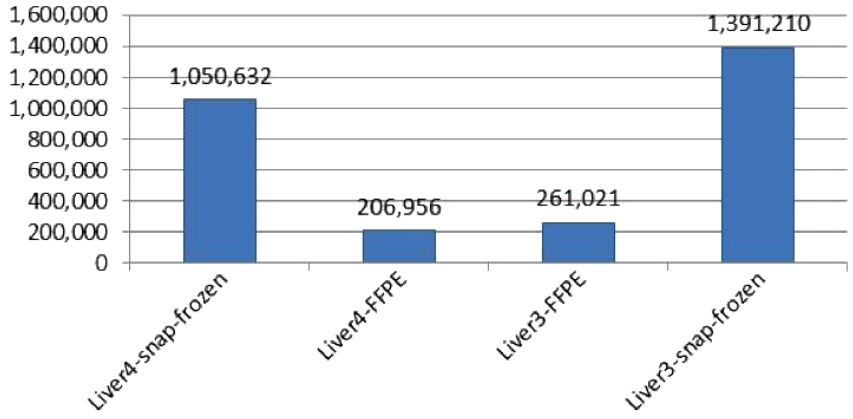

Figure 5. *Estimated library size of the snap-frozen and FFPE sample derived strand specific libraries, which depicts the number of unique fragments present in each library. The FFPE samples showed a ~5 fold lower number of unique fragments, consistent with the high % duplication rate in the FFPE libraries.*

2- Mapped read rate. We then compared average mapping statistics of the human transcriptome content between the FFPE and snap-frozen counterparts. Intergenic rate refers to the fraction of reads that map in between genes, while exonic rate is the fraction mapping within exons, and intronic rate is the fraction mapping within introns. Figure 6 exhibits very similar exonic (~75% for FFPE and ~74% for snap-frozen), intronic (~23% for FFPE and ~24% for snap-frozen) and intergenic (~1.2% for FFPE and ~1.6-2% for snap-frozen) rate for all libraries. Furthermore, the rRNA rate was < 0.1% for all libraries.

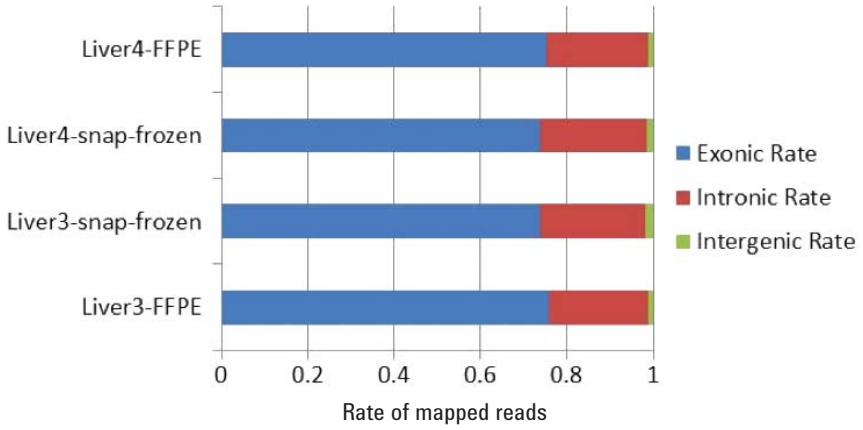

Figure 6. *The rate of mapped reads.*

3- Coverage at 5' and 3' ends of the transcripts. Coverage at 5′ and 3′ ends is important for correctly identifying transcripts. We determined average coverage at each percentile of the length from 5′- to 3′-end of the known transcripts to determine bias for FFPE versus snap-frozen libraries (Figure 7). Mean coverage for snap-frozen libraries is significantly higher than FFPE libraries due lower higher complexity (lower duplicates). There is also slightly higher coverage at the 3′-end of the snap-frozen libraries compared to the 5′-end. However, with FFPE libraries, there is minimal difference in coverage between the 5′- and 3′-ends of the transcripts, indicating minimal bias in library prep/sequencing.

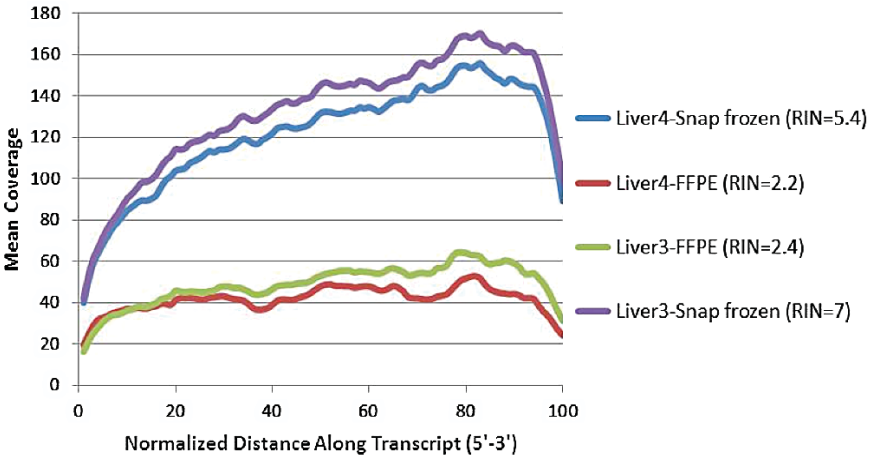

Figure 7. *5′-3′ transcript coverage.*

Conclusions

The data presented here reveal that FFPE and matched snap-frozen samples generate similar % on target, strand specificity, and read mapping statistics. The snap-frozen samples show a slight bias at the 3'-end of the transcript, while their FFPE counterparts exhibit minimal bias. The complexity of the FFPE derived libraries was significantly lower than the snap-frozen controls. This latter observation is to be expected given the low quality and integrity of the FFPE-derived RNA used, as measured by the RNA 6000 Nano assay on the Agilent 2100 Bioanalyzer system. Taken together, these data provide two key conclusions. First, Agilent's strand specific RNA library preparation kit can indeed be utilized to prepare sequencing libraries from FFPE RNA and second, the ability to generate data from the low quality FFPE RNA samples used in this study provide strong evidence that a broad range of FFPE RNA sample qualities can be successfully used with this kit for both whole transcriptome and targeted RNA-seq applications.

References

1. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M (2007) Determinants of RNA Quality from FFPE Samples. PLoS ONE 2(12): e1261. doi:10.1371/journal.pone.0001261
2. Agilent recommends Agencourt AMPure XP Beads from Beckman Coulter, Inc. for all SPRI bead based sample clean up steps. Refer to the following Beckman Coulter website for technical information related to the features of this product: (https://www.beckmancoulter.com/wsrportal/wsr/research-and-discovery/products-and-services/nucleic-acid-sample-preparation/agencourt-ampure-xp-pcr-purification/index.htm)
3. NanoDrop Technical Support Bulletin T009 (2007) 260/280 and 260/230 Ratios NanoDrop® ND-1000 and ND-8000 8-Sample Spectrophotometers.
4. Epicenter Ribo-Zero™ Gold Kit Protocol for (Human/Mouse/Rat)

www.agilent.com/genomics/NGS

For Research Use Only.
Not for use in diagnostic procedures.
This information is subject to change without notice.

PR7000-0416

5991-4794EN

Agilent Technologies