QuantSeq yields good correlation between high and low quality (FFPE) samples

The suitability of QuantSeq when using highly degraded samples (e.g., formalin-fixed, paraffin-embedded (FFPE) material) was evaluated by comparing two samples derived from one source but of different RNA qualities.

A xenograft of the MOLP-8 human tumor cell line was split into two pieces, which were subsequently processed either as fresh frozen cryo-block or embedded FFPE material, leading to different RNA qualities from the same original sample. To determine RNA quality often the RIN (RNA Integrity Number) is used, with an RIN of >8 indicating high RNA quality. For heavily degraded samples this is not a sensitive measure, and hence the DV_{200} value (distribution value of RNA fragments >200 nucleotides) should rather be used. Low RNA integrity correlates with low DV_{200} values.

After RNA extraction, the FFPE sample had a DV_{200} of 87 % (RIN of 2.8), while the cryo sample yielded a RIN of 8.3. The libraries were generated with the QuantSeq 3′ mRNA-Seq FWD kit using 50 ng total RNA input. QuantSeq libraries were also successfully generated of FFPE-derived RNA with a DV_{200} of down to 52 % (data not shown). For the FFPE sample the protocol recommendations for low quality RNA input were followed, for the cryo sample the standard protocol was applied. The libraries were sequenced on a HiSeq 2500 instrument at 1x50 bp read length.

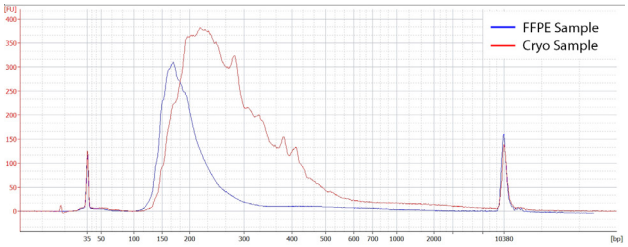

Figure 5 | Bioanalyzer 2100 HS DNA chip traces of QuantSeq 3′ mRNA-Seq FWD prepared libraries using FFPE (blue) or cryo RNA input (red). For the degraded input, the resulting library shows a smooth size distributing with no visible linker-linker by-products but only a shift towards shorter fragments. Average library size is 204 bp (FFPE) and 286 bp, respectively.

Plotting the relative coverage across the normalized transcript length shows that coverage is focused on the transcripts' 3′ end, independent of the input RNA quality (Fig. 6). However, as QuantSeq FWD libraries are sequenced towards the poly(A) site coverage is dependent on library size and sequencing length.

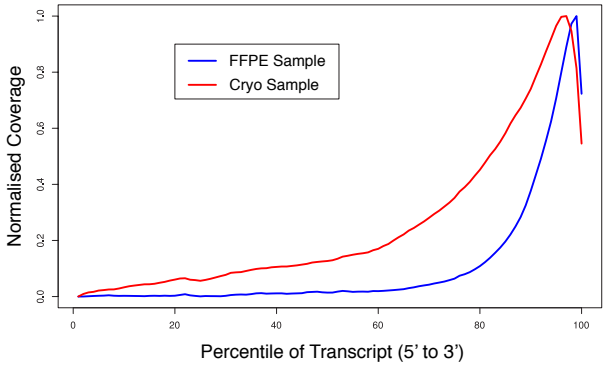

Figure 6 | QuantSeq read coverage versus normalized transcript length of NGS libraries derived from FFPE RNA (blue) and cryo-preserved RNA (red).

The FFPE libraries were significantly shorter than the cryo sample, therefore the ends of the transcripts were reached more frequently, which is reflected in the coverage plots (Fig. 6).

Gene expression correlation between libraries derived from FFPE and cryo-preserved RNA is high (R^2 0.86) and indicates that QuantSeq performs consistently well on samples of different RNA quality (Fig. 7).

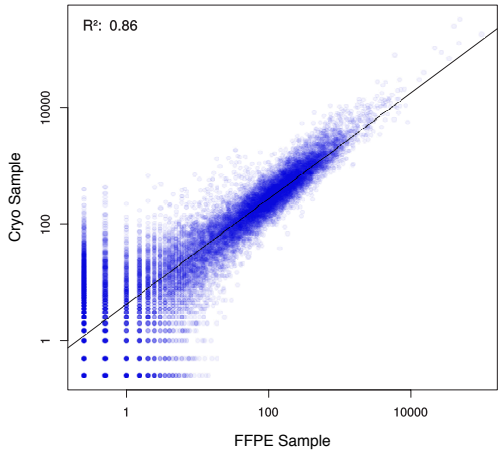

Figure 7 | Correlation of gene counts of FFPE and cryo samples.

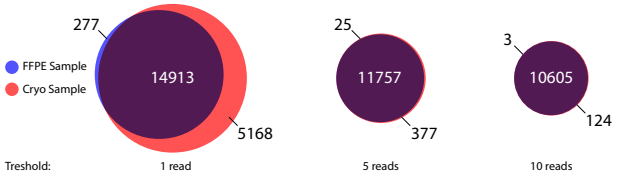

Figure 8 | Venn diagrams of genes detected by QuantSeq at a uniform read depth of 2.5 M reads in FFPE and cryo samples with 1, 5, and 10 reads/gene thresholds.

QuantSeq has a very high sensitivity, at 26.5 M reads, 25842 genes were detected in the intact Cryo RNA sample (data not shown). At uniformly 2.5 M reads, 20081 genes were detected in the Cryo sample with at least 1 read, compared to 15190 genes in the FFPE samples, which represents a 24 % difference (Fig. 8). However, increasing the detection level to 5 or 10 reads/gene reduces the difference to 3 % and 1 %, respectively. These alignments show that QuantSeq reliably detects gene expression in both cryo and FFPE samples; the difference in detection of lowly expressed genes is due to the higher susceptibility of their low copy transcripts to degradation during FFPE treatment, storage, and recovery.

QuantSeq is based on oligodT priming of the reverse transcription and only generates one fragment per transcript. This enables accurate gene expression quantification independent of the RNA quality (including FFPE samples). Standard mRNA-Seq protocols aim to cover the whole transcript, but will result in a heavy 3′ bias when used on degraded RNA input. Therefore, QuantSeq 3′ mRNA-Seq is an efficient tool to generate NGS libraries from low quality samples compared to other mRNA-Seq protocols using poly(A) selection.