LEXOGEN
The RNA Experts

219P00015V0100

# Identifying hallmarks of kidney tumors from FFPE samples using whole transcriptome or 3' mRNA-Seq library preps

Alice Wenger, Pamela Moll, Nora Nikoloska, Yvonne Goepel, Kathrina Radl, Cynthia Wollner, Musashi Tsujita, Mira Schaupper, and Torsten Reda

Lexogen GmbH, Campus Vienna Biocenter 5, 1030 Wien, Austria

## ABSTRACT

Formalin-fixed paraffin-embedded (FFPE) material presents a readily available resource for studying biomarkers. However, RNA isolated from FFPE samples is often heavily degraded and chemically modified. Here we compare RNA-Seq data from FFPE tissues generated with two different NGS sample prep methods: **3' mRNA-Seq** vs. **whole transcriptome sequencing (WTS)**. We used our new CORALL FFPE RNA-Seq Whole Transcriptome Library Prep Kit together with RiboCop HMR V2 rRNA depletion as well as our novel QuantSeq FFPE 3' mRNA-Seq Library Prep Kit to compare the performance on matched human normal and tumor FFPE RNA.
Both kits show a high degree of overlap between normal and tumor samples, but also signifcant detection of differentially expressed genes. Despite using rRNA depletion WTS vs. 3' mRNA-Seq, we observed a significant overlap between the two sample prep methods. The area of interest (gene or transcript expression, mRNA or all non-rRNA, screening or in-depth analysis) determines the ideal method of choice.

## Workflows

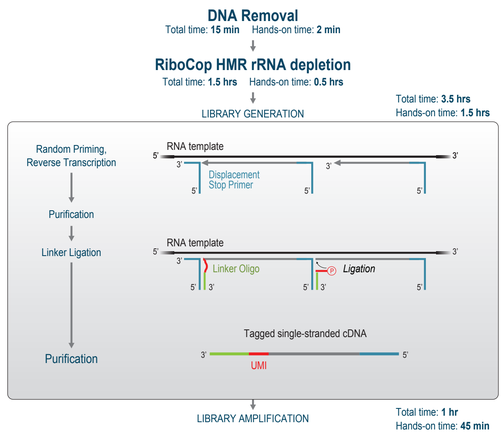


Figure 1 | Schematic overview of the CORALL FFPE library preparation workflow.

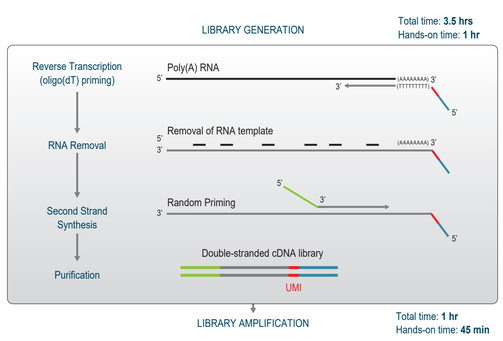


Figure 2 | Schematic overview of the QuantSeq FFPE library preparation workflow.

### CORALL FFPE Whole Transcriptome

**CORALL FFPE** is validated on various FFPE samples and covers the whole length of the transcripts. CORALL uses Lexogen's proprietary displacement-stop technology to generate NGS library inserts **omitting any RNA fragmentation steps** (Fig. 1). Paired with **RiboCop for enzyme-free rRNA depletion**, CORALL FFPE is ideal for any application requiring coverage uniformity, including coverage analysis, alternative splicing or fusion gene detection and analysis of non-coding RNA (e.g., lncRNA biomarkers).

### QuantSeq FFPE 3' mRNA-Seq

**QuantSeq FFPE** focuses on the 3' ends of all polyadenylated transcripts and is thus ideal for cost-efficient expression profiling from degraded and FFPE RNA. **No depletion or prior enrichment is required**, allowing preparation of ready-to-sequence libraries in only 4.5 hours total workflow time. Starting from total RNA, an oligo(dT) primer initiates reverse transcription (first strand synthesis) and random priming (second strand synthesis) completes cDNA library generation (Fig. 2). **UMIs are seamlessly introduced** in the first reaction step for highly accurate expression analysis. Lexogen's FFPE expression profiling workflow builds on the established **QuantSeq technology** and is now **validated for FFPE samples from varying tissues and species**.

## Experiment and Results

### Comparsion of gene body coverage

CORALL FFPE RNA-Seq is a whole transcriptome (WTS) sample prep method which was used on rRNA depleted fresh frozen (FF) tumor and normal FFPE RNA, hence coverage profiles consistently show uniformity across the transcript body (Fig.3). QuantSeq FFPE is a 3' mRNA-Seq sample prep method that uses oligo(dT) priming and captures only the most 3' end. When using degraded RNA such as FFPE material, an even more prominent shift towards the 3' end can be observed.

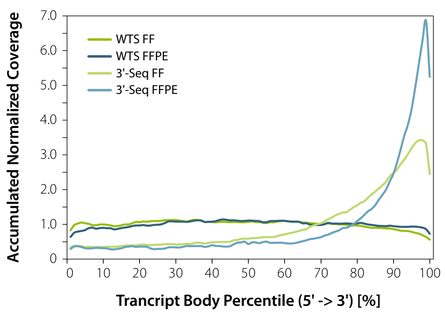


**Figure 3 | Accumulated transcript body coverage (whole transcriptome (WTS), dark blue and dark green vs. 3' mRNA-Seq, light blue and light green).** Libraries were prepared from 25 ng FF and FFPE human liver RNA and either subjected to rRNA depletion before CORALL WTS sample prep or directly inserted into oligo(dT) primed 3' mRNA-Seq library preps. Libraries were sequenced on NextSeq500 in an asymmetric paired end run (51+24). Reads were aligned to the human reference genome GRCh38.94 and quantified with featureCounts for 3' mRNA-Seq and Mix$^2$ version 1.4.0.12 (Lexogen) for CORALL WTS[1, 2]. Coverage across all transcripts was generated using the gbc-tool provided by RSeQC (transcripts length normalized to 100 %).

### Human kidney normal vs. tumor samples analyzed with CORALL and QS FFPE

20 ng human kidney normal and tumor FFPE RNA were inserted into QuantSeq FFPE 3' mRNA-Seq library preps or used for rRNA depletion with RiboCop HMR V2 and subsequent library preps with CORALL FFPE (Fig. 4). Libraries were sequenced on NextSeq2000 (2x75bp). Reads were downsampled to 8 M reads per condition.

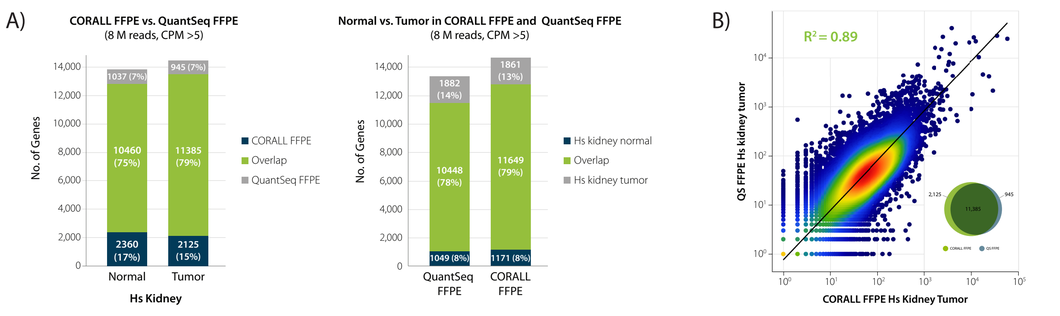


**Figure 4 | QuantSeq FFPE and CORALL FFPE on human kidney normal vs. tumor FFPE RNA.** A) Overlap in gene detection in CORALL FFPE library preps from RiboCop HMR V2 rRNA depleted samples vs. QuantSeq FFPE 3' mRNA-Seq preps for normal and tumor kidney FFPE samples as well as overlaps in gene detection for normal vs. tumor (8 M reads per condition, uniquely mapping reads, CPM >5). B) Correlation plots for human tumor FFPE RNA in CORALL FFPE (x-axis, rRNA depleted RNA) vs. oligo(dT) primed QuantSeq FFPE preps. Venn Diagram depicts overlaps in gene detection with both preps for human kidney tumor samples (uniquely mapping reads, CPM >5).

As CORALL FFPE uses rRNA depletion additional genes can be detected (long non-coding RNAs, non-polyadenylated RNAs) compared to QuantSeq FFPE, which selectively detects the polyadenylated mRNA fraction.

### Differential expression in human kidney normal vs. tumor FFPE samples

Matched tissue samples from human kindey normal and tumor FFPE material were analyzed for differential gene expression using QuantSeq FFPE (Fig. 5).

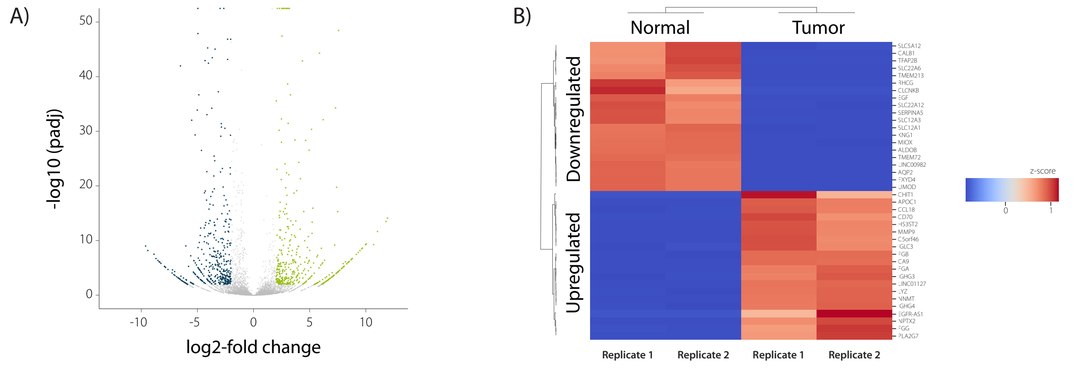


**Figure 5 | Differential gene expression in human kindey normal vs. tumor FFPE samples.** A) Volcanco plot for QuantSeq FFPE samples. B) Hierachical heat map analyzing differential gene expression in normal vs. tumor samples with QuantSeq FFPE (top 50 DE genes).

Significant dfferential expression can be detected between normal and tumor kidney samples. The affected genes are associated with described pathways differentially regulated during renal cancer formation.

### Coverage profiles of hallmarks of renal carcinoma

CORALL FFPE is a whole transcriptome library prep enabling analysis of coverage along the entire transcript for differentialy expressed genes in tumor vs. normal human kindey samples (Fig. 6).

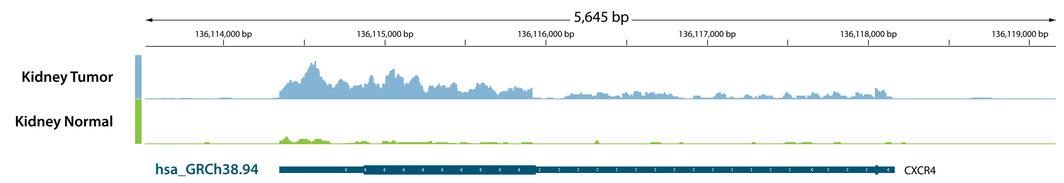


Figure 6 | CXCR4 coverage in tumor vs. normal kidney FFPE samples with CORALL FFPE.

Chemokine receptor CXCR4 is associated with an increased risk, the incidence of renal cell carcinoma, and its progression.

### Exon skipping in human kidney tumor samples

CORALL FFPE samples were further used to analyse alternative splicing in tumor samples. Differential splicing, exon skipping or retention has been previously shown to be associated with tumor formation (Fig. 7).


**Figure 7 | Exon skipping in human kidney tumor detected by CORALL FFPE.** Alternative splicing events were computed with SplAdder[3] using aligned BAM files with default parameters. The exon skipping event on exon 6 of the VEGF-A gene is plotted (adjusted p-value 0.011). Tracks show sum of coverage for two replicates. The role of alternative splicing of VEGF-A in pathological processes among which liver and kidney disease and cancers has been described in [4].

## Conclusion

CORALL FFPE and QuantSeq FFPE can both be used for **efficient detection of differential gene expression identifying drivers of e.g., tumorigensis**. Despite using rRNA depleted RNA vs. mRNA there is a significant overlap between the two sample prep methods. The area of interest (gene or transcript expression, mRNA or all non-rRNA, splice variant analysis, screening or in-depth analysis) determines the method of choice. **CORALL FFPE is the method of choice when being interessed in alternative splicing, long non-coding RNA** while **QuantSeq FFPE is ideal to detect differential gene expression**. The lower read-depth requirement for 3' mRNA-Seq technologies, such as QuantSeq allows **upscaling of experiments to easily process hundreds or thousands of samples**. Both library preps have been optimized and validated with FFPE samples derived from various tissues and species and can be used even with low quality FFPE RNA (down to DV200 <10%) or low quantity down to 5 ng FFPE RNA.

**References:**

[1] Dobin A., et al. (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics, DOI: 10.1093/bioinformatics/bts635
[2] Tuerk A., et al. (2017) Mixture models reveal multiple positional bias types in RNA-Seq data and lead to accurate transcript concentration estimates. PLoS Computational Biology, DOI: 10.1371/journal.pcbi.1005515
[3] Kahles A., et al. (2016) SplAdder: identification, quantification and testing of alternative splicing events from RNA-Seq data. Bioinformatics, DOI: 10.1093/bioinformatics/btw076
[4] Dehghanian F., et al., (2014) New Insights into VEGF-A Alternative Splicing: Key Regulatory Switching in the Pathological Process. Avicenna J Med Biotechnol. PMID: 25414781; PMCID: PMC4224658.

For more information please visit our website:
www.lexogen.com/applications/FFPE-transcriptomics