Searching For Alternative Splice Sites In FFPE Samples

Does Tissue Preservation Method Matter?

Ryan Videgar-Laird, Mikayla J. Feldbauer, Peyton Kuhlers, Katherine A. Hoadley

UNC Biological & Biomedical Sciences Program

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Introduction

In addition to normal cellular function, alternative splicing (AS) is known to be associated with various cancers and metastatic processes. Robust identification of AS events can therefore be used to better understand mechanisms of disease, establish diagnostic biomarkers, and drive development of novel cancer treatments.

A large proportion of available tumor samples are preserved as formalin-fixed, paraffin-embedded samples due to the relative ease and affordability of doing so (no need for long-term cryogenic storage). However, this preservation method introduces random artifacts and fragmentation into the RNA sequence, potentially confounding analysis.

To date, much of the work done to compare FFPE samples to Fresh Frozen (FF) samples has only considered differential gene expression. We began an exploratory analysis comparing AS events between FF and FFPE samples, utilizing a unique dataset from the Trans Cancer Genome Atlas (TCGA) involving 35 patients with paired, time-matched samples.

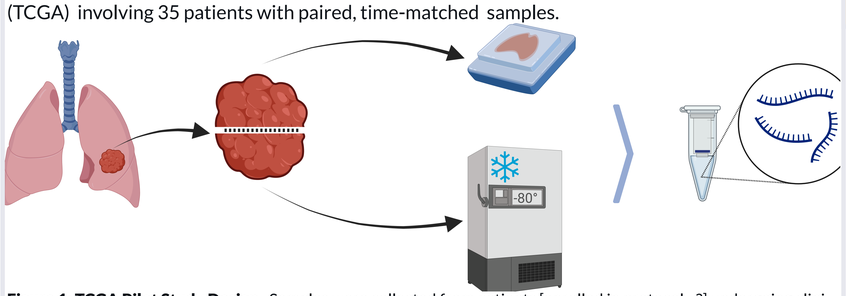


Figure 1: TCGA Pilot Study Design. Samples were collected from patients [enrolled in protocol…?] undergoing clinical resection of tumors . Tumors were immediately bisected , with one half preserved FF and the other as FFPE. RiboZero Total RNA sequencing was performed on all paired sets.

Table 1: Count of paired samples by tissue type

Tissue	N Paired FF/FFPE Samples
LUAD	10
COAD	9
BRCA	5
KIRC	4
UCEC	4
BLCA	3
Total	**35**

Methods

rMATS turbo (4.1.1) was used to identify AS events. Ironically, we chose not to use it for differential analysis between replicate groups. Instead it was ran independently for all 70 samples, as we were mostly concerned with bulk sensitivity rather than significant differences.

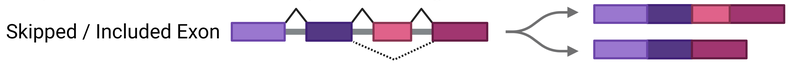


Figure 2: Example of AS Junction Counts. Solid lines represent Inclusion Junction Count (IJC) reads, dashed lines Skipped Junction Counts (SJC). Junction Read Support (JRS) is the sum of IJC and SJC. JRS > 20 is considered a good rule-of-thumb for an evidence cut-off.

Picard (2.7.1) was used for RNAseq QC metrics, chiefly nucleotide mapping.

BAM files were indexed with **Samtools**.

Gene expression levels obtained via <**Ask Mikayla**>.

R (4.1.2) was used for all data aggregation, analysis, and plotting.

Key Questions

1. **How do AS events between FF and FFPE samples compare?**
   - Raw counts
   - Shared Events (intersection)
2. **What is driving any differences between the two?**
   - RNA quality
     - Read mapping
   - Gene expression
3. **How does event support vary with preservation method?**
   - SJC, IJC, Percent Spliced In (PSI)
     - JRS

Results

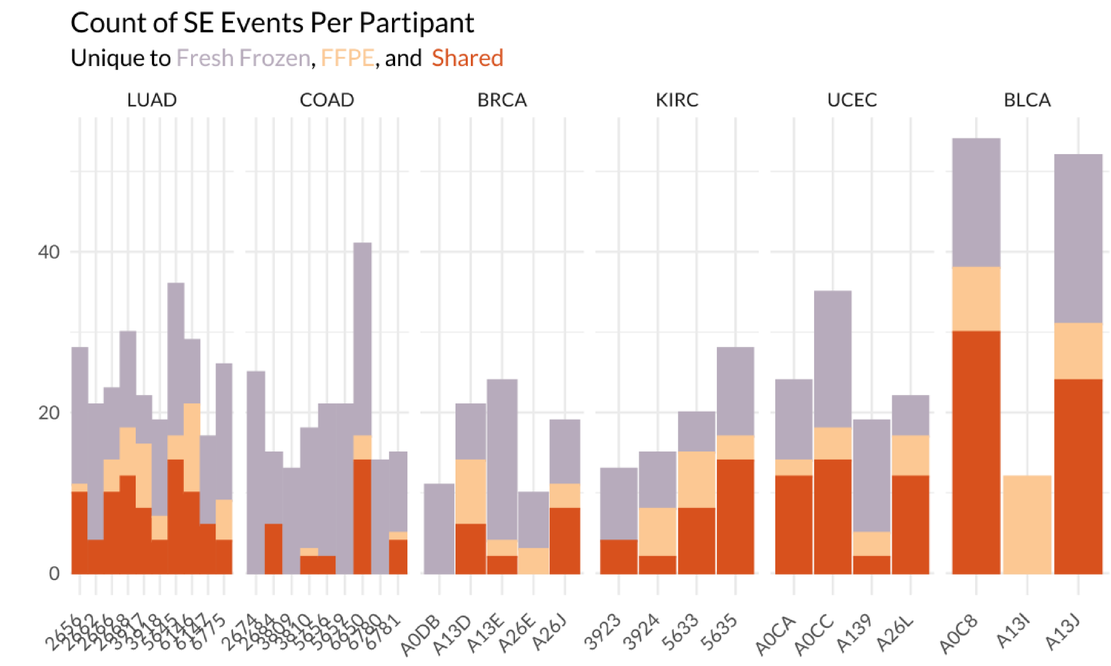
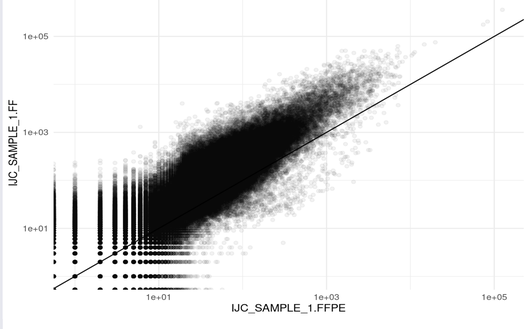


Figure 3: Comparing Proportion of Shared Skipped Exon (SE) Events in Highly Expressed Genes. Coding genes with an expression level in the upper quartiles for at least 5 samples were examined for shared AS events.



Discussion

Acknowledgments