Transcriptome analysis reveals extensive alternative splicing coupled with nonsense-mediated mRNA decay in human cells
Courtney E. French1, Gang Wei2, Angela N. Brooks1 and Steven E. Brenner1,2
1Department of Molecular and Cell Biology, 2Department of Plant and Microbial Biology, University of California, Berkeley, CA, 94720-3102, USA.
brenner@compbio.berkeley.edu
*These authors contributed equally to this work

INTRODUCTION:
Alternative splicing plays a major role in the generation of proteomic diversity. However, mistakes in this process can introduce a premature termination codon (PTC) and result in non-functional proteins that are harmful to the cell. Such transcripts are usually degraded by nonsense-mediated mRNA decay (NMD). The coupling of alternative splicing and NMD has also been reported as an important regulatory mechanism for certain sets of genes [1]. Though many NMD targets have been identified in various species, we still lack a comprehensive view of the landscape of those transcripts degraded by NMD. Here, we characterize the transcripts normally degraded by NMD in human HeLa cells by inhibiting NMD through knockdown of the essential NMD factor UPF1 and performing RNA-seq analysis.

GOALS:
How many genes produce isoforms that are targets for NMD in human cells? How highly transcribed are NMD targets before degradation? What is the functional role of NMD-related regulation?

APPROACH:
1. NMD inhibition through knockdown of UPF1.
2. Directional and paired-end RNA-seq library preparation.
3. High-throughput sequencing using Illumina machine.
4. Map reads to genome with TopHat [3].
5. Transcript assembly and quantification with Cufflinks [4].
6. Premature termination codon (PTC) prediction using canonical 50nt rule of transcripts targeted by NMD in mammals [5].

RESULTS:

Almost 2,500 transcripts identified as putative NMD targets

A. No bias in direction of abundance changes for productive isoforms

B. Isoforms with a PTC tend to increase in abundance

C. Many isoforms degraded by NMD are expressed at high levels

Many genes encoding splicing factors were identified as having isoforms targeted by NMD

Splicing factor category

Glia
genes expressed

Number targeted by NMD

Genes with isoforms targeted by NMD

<table>
<thead>
<tr>
<th>Splicing factor category</th>
<th>Glia genes expressed</th>
<th>Number targeted by NMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor UPF1</td>
<td>5,809 PTC50nt</td>
<td>1,618 productive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>isoforms targeted by NMD</td>
</tr>
</tbody>
</table>

Many NMD targets have been identified in various species. NMD targets are significantly enriched for ultraconserved elements. Splicing regulators are significantly enriched for NMD targets.

CONCLUSIONS:
- Almost 2,500 isoforms from 1,900 genes are degraded by NMD.
- Splicing regulators are significantly enriched for NMD targets.
- Exons of transcripts targeted by NMD are significantly enriched for ultraconserved elements.
- Their overall conservation is similar to that of all exons genome-wide.
- NMD-targeted isoforms can be generated by various splicing events.
- Coupling of alternative splicing and NMD appears to regulate the expression of hundreds of genes.

REFERENCES: