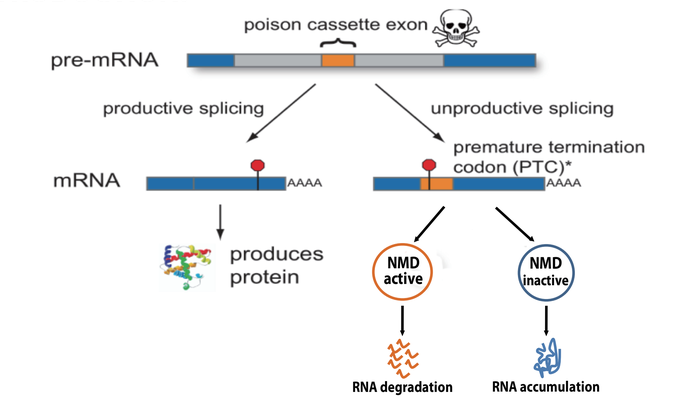
Thousands of targets of nonsense-mediated mRNA decay revealed by transcriptome analysis offer clues to the mechanism in human, fish, and fly

Download PDF

Courtney E. French[1], Gang Wei[2,3], Angela N. Brooks[4], Thomas L. Gallagher[5], Li Yang[6], Brenton R. Gravely[7], Sharon L. Amacher[5] and Steven E. Brenner[1,2*]

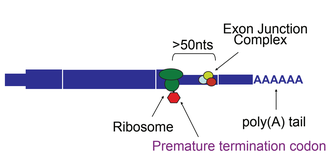
[1]Department of Molecular and Cell Biology, [2]Department of Plant and Microbial Biology, University of California, Berkeley, CA, 94720, USA. [3]Now at Fudan University, Shanghai, China. [4]Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA. [5]Department of Molecular Genetics , Ohio State University, Columbus, OH, 43210, USA. [6]Partner Institute of Computational Biology, Shanghai, China. [7]Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmingtion, CT, 06030, USA. *brenner@compbio.berkeley.edu

INTRODUCTION:

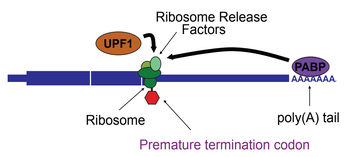


Nonsense-mediated mRNA decay (NMD) is an RNA surveillance system that degrades aberrant isoforms containing a premature termination codon. This pathway is conserved throughout eukaryotes and protects against the production of harmful truncated proteins. Additionally, NMD coupled with alternative splicing is a mechanism of post-transcriptional gene regulation that affects the mRNA levels of hundreds of genes in human [1]. The canonical model of defining a premature termination codon in mammals is the 50nt rule: a termination codon more than 50 nucleotides upstream of an exon-exon junction is premature and triggers degradation by NMD [2]. In other animals, a 3' UTR intron is not required for NMD. There is also evidence that a longer 3' UTR triggers NMD in plants, flies, and mammals [3]. The importance of each mechanism appears to vary between species, and it is currently unclear which is the major mechanism at work in human cells. We used RNA-seq analysis done on cells with inhibted NMD to determine the features associated with degradation in human and in fly.

50nt rule (mammals)



longer 3' UTR (fly)

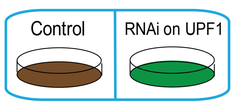


GOALS:

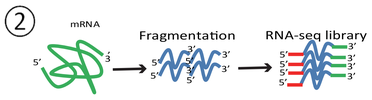
How conserved are the targets of alternative splicing coupled with NMD?

What features define a premature termination codon in different species?

APPROACH:

① 

NMD inhibition through knockdown of UPF1.

② 

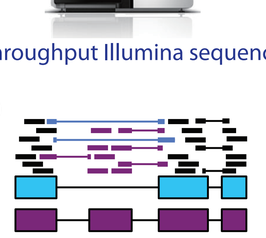
Directional and paired-end RNA-seq library preparation.

③ 

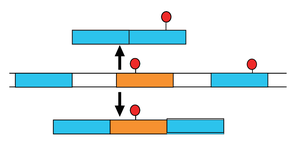
High throughput Illumina sequencing.

④ 

Map reads to genome with TopHat [4].

⑤ 

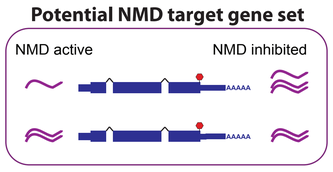
Transcript assembly and quantification with Cufflinks [5] or JuncBASE [6].

⑥ 

Premature termination codon prediction.

Thousands of genes produce alternative isoforms degraded by NMD in human. Fish and fly have hundreds.

Potential NMD target gene set

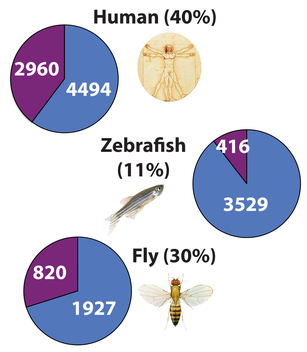


We infer that a gene produces an alternative isoform that may be degraded by NMD if the gene has at least one isoform that increases >2x when NMD is inhibited and at least one isoform that does not increase.

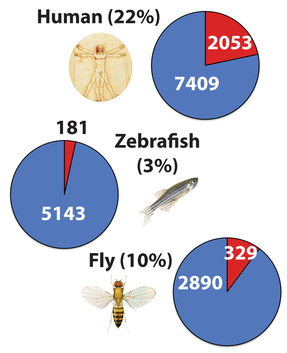
Thousands of genes fall into this category for human. Zebrafish and fly have hundreds (**PURPLE**).

Ribosomal and translation genes are enriched in these NMD targeted genes for fly, and intracellular signaling genes are enriched in zebrafish (Fisher's exact test, FDR<0.05).

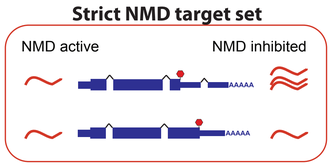
Fraction of genes with >1 isoform (FPKM>1) that are **potential NMD targets**



Fraction of alternatively spliced genes that are in our **strict set of NMD targets**



Strict NMD target set

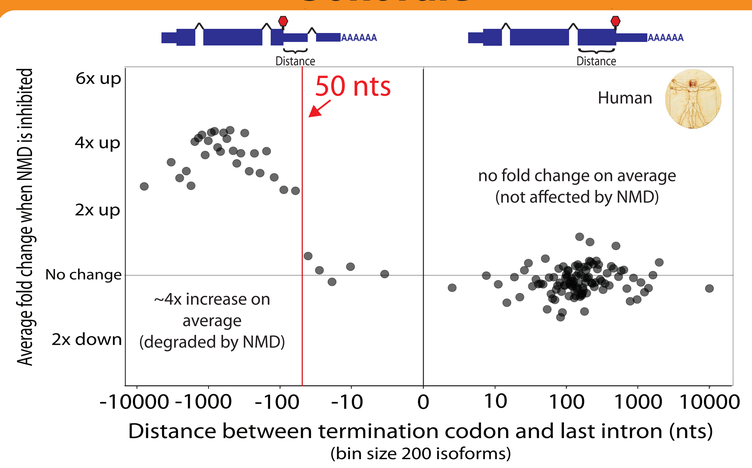


Since an intron sufficiently downstream of the termination codon is known to trigger NMD in human (50nt rule), we defined a strict set of probable NMD targeted genes that are required to have an isoform that follows the 50nt rule (contains a 3' UTR intron) and increases >2x (after controlling for transcriptional changes) when NMD is inhibited.

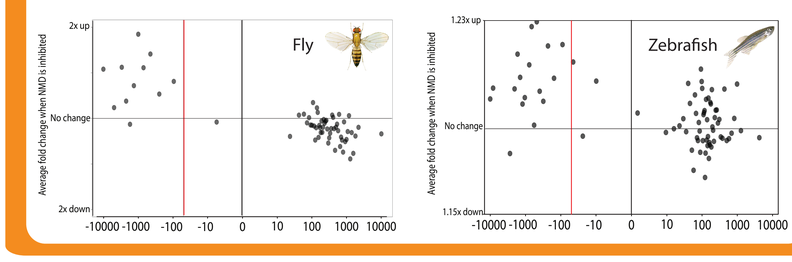
Over 20% of alternatively spliced genes are targeted by NMD in human (**RED**). These genes are enriched for splicing genes in human and in fly (Fisher's exact test, FDR<0.05).

The 50nt rule is a strong predictor of NMD in all three species while a longer 3' UTR has a limited effect

50nt rule

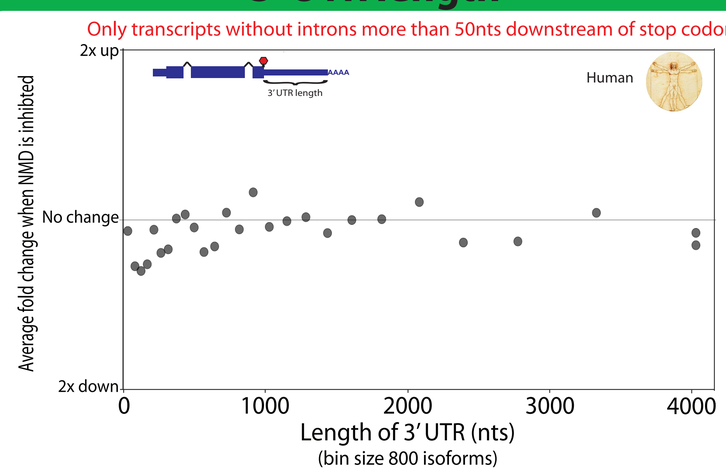


Transcripts with an exon-exon junction over 50nts downstream of a stop codon (left) are significantly more likely to increase when NMD is inhibited in human, fly, and zebrafish (K-S test: $p < 2x10^{-308}$, $p = 2x10^{-21}$, $p = 1x10^{-79}$, respectively).
Each point is the mean distance and fold change of 200 isoforms.

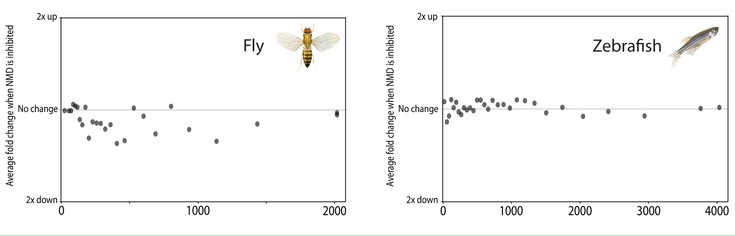


3' UTR length

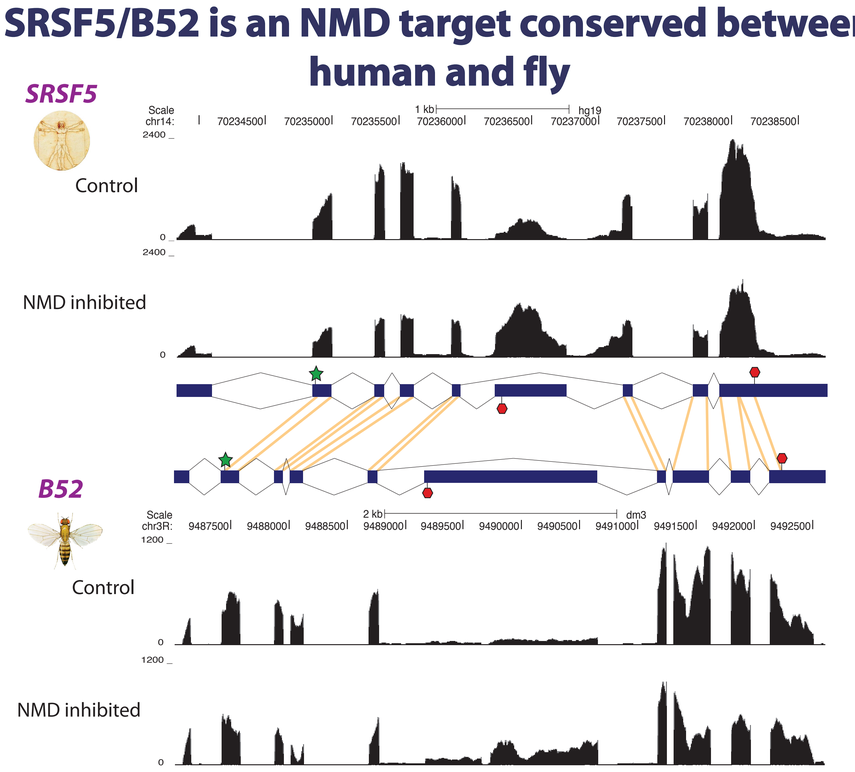
Only transcripts without introns more than 50nts downstream of stop codon



When only looking at transcripts that cannot be affected by 50nt rule, we see only a slight correlation between 3' UTR length and an increase when NMD is inhibited in human and no correlation in fish or fly (K-S test: human $p = 3x10^{-11}$, fish $p = 0.13$, fly $p = 0.81$).
Each point is the mean length and fold change of 800 isoforms.



SRSF5/B52 is an NMD target conserved between human and fly



Both the human gene SRSF5 and its fly ortholog B52 produce NMD-targeted alternative isoforms when the sixth exon is included. All SR genes have NMD-targeted isoforms in human and mouse [7,8].

CONCLUSIONS:

Thousands of alternatively spliced genes (>20%) produce transcripts that fall into our strict set of NMD targets in human.

Hundreds of alternatively spliced genes (10-30%) produce transcripts possibly degraded by NMD in fly and zebrafish.

Splicing genes are significantly enriched in NMD targeted genes in human and in fly.

The 50nt rule is a strong predictor of NMD in human and also appears to have a role in fly and zebrafish.

3' UTR length has little correlation with NMD in human, fly, and zebrafish.

REFERENCES:

1. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated decay in humans. *PNAS.* 2003. 100:189-192
2. Nagy E and Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends in Biochemical Science.* 1998. 23:198-9
3. Kerenyi Z, Merai Z, Hiripi L, Benkovics A, Gyula P, Lacomme C, Barta E, Nagy F, and Silhavt D. Inter-kingdom conservation of mechanisms of nonsense-mediated mRNA decay. *EMBO Journal.* 2008. 27:1585-95
4. Trapnell C, Pachter L, Salzberg SL. Tophat: discovering splice junctions with RNA-seq. *Bioinformatics.* 2009. 25:1105-11
5. Trapnell, et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology.* 2010. 28:511-5
6. Brooks AN, et al. Conservation of an RNA regulatory map between Drosphila and mammals. *Genome Research.* 2011. 21:193-202
7. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature.* 2007. 446:926-9
8. Ni JZ, et al. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes and Development.* 2007. 21:708-18

ACKNOWLEDGEMENTS:

We thank Adam Roberts and Lior Pachter of UC Berkeley for help with the optimization of Cufflinks.
This work was funded by an NIH Grant R01-GM071655 to S.E.B, G.W was supported by a Tang Distinguished Scholarship from QB3 at UC Berkeley, and C.E.F. was supported by an a NDSEG Fellowship.