

# *SQANTI report*

*Unique Genes: 11709*

*Unique Isoforms: 59523*

### *Gene classification*

<b>Category</b>	<b># Genes</b>
Annotated Genes	11517
Novel Genes	192

### *Splice Junction Classification*

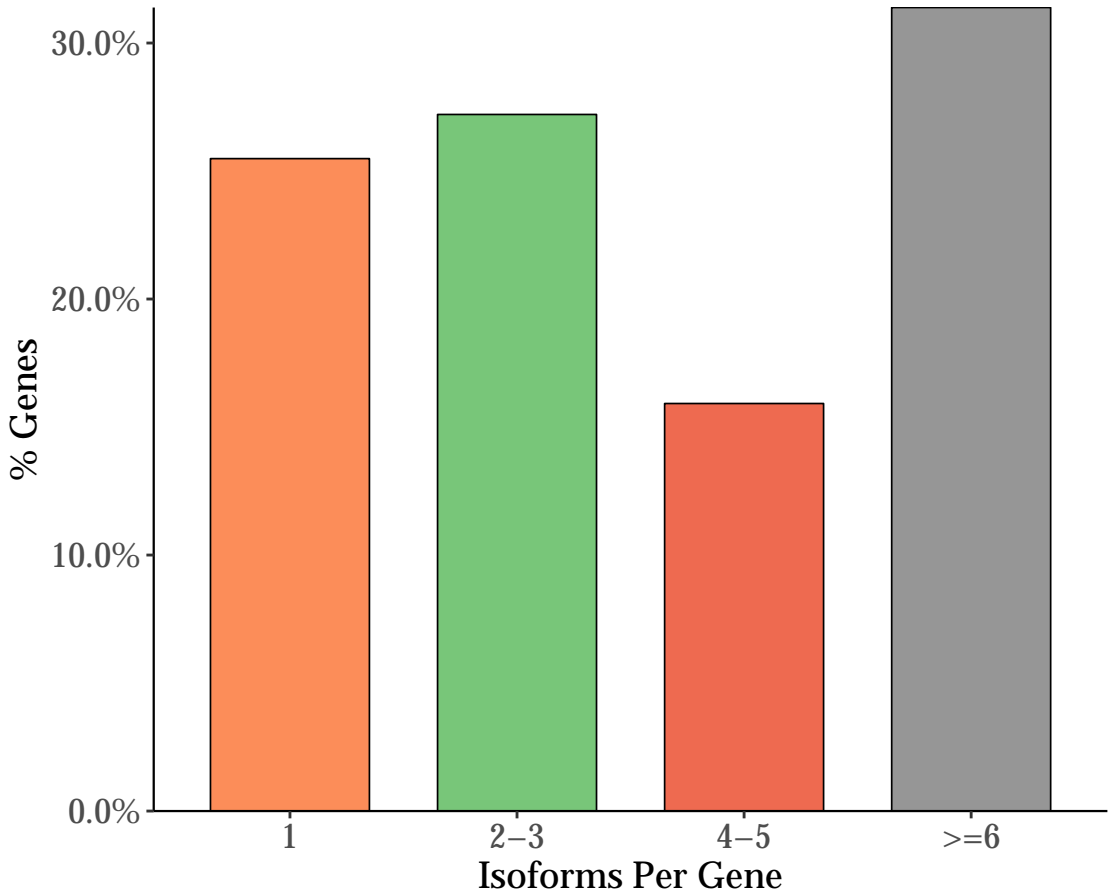
<b>Category</b>	<b># SJs</b>	<b>Percent</b>
Known canonical	123028	90.97
Known Non-canonical	73	0.05
Novel canonical	12133	8.97
Novel Non-canonical	1	0.00

### *Characterization of transcripts based on splice junctions*

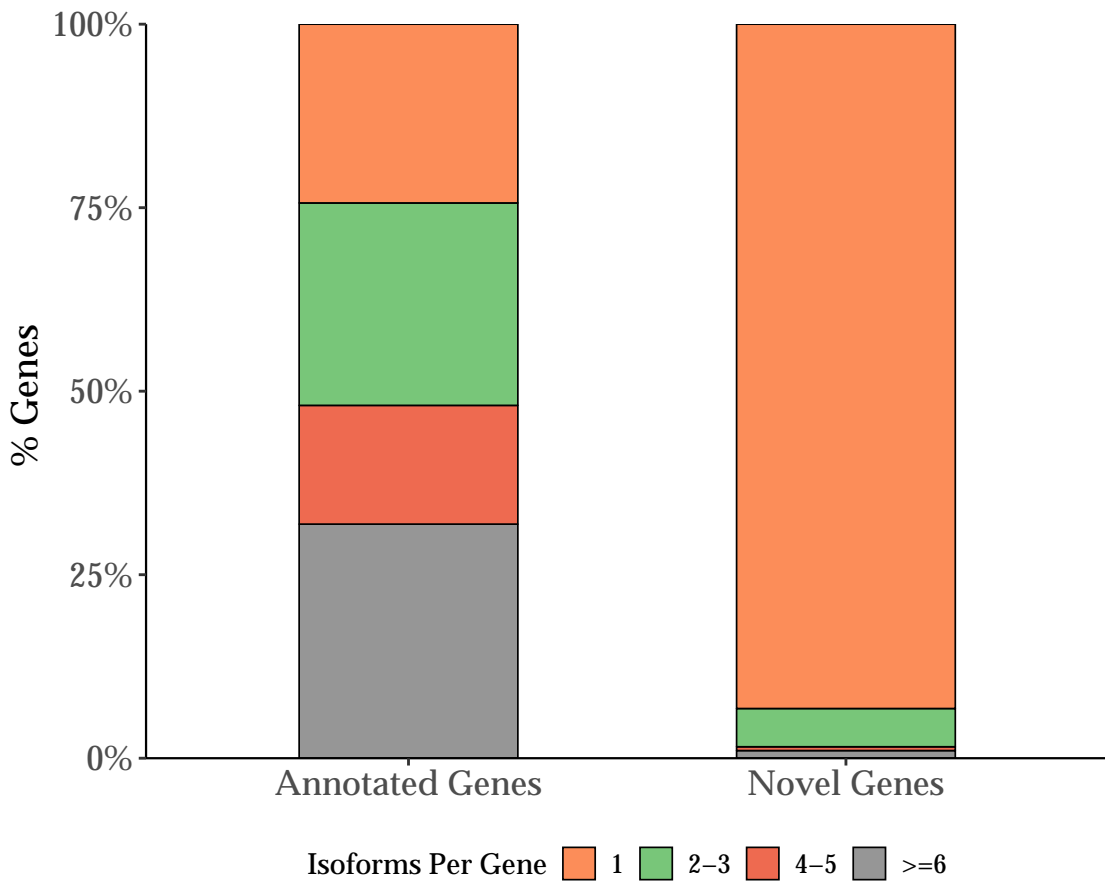
<b>Category</b>	<b># Isoforms</b>
FSM	27772
ISM	5166
NIC	17808
NNC	8555
Genic Genomic	0
Antisense	82
Fusion	0
Intergenic	140
Genic Intron	0

## *Gene Characterization*

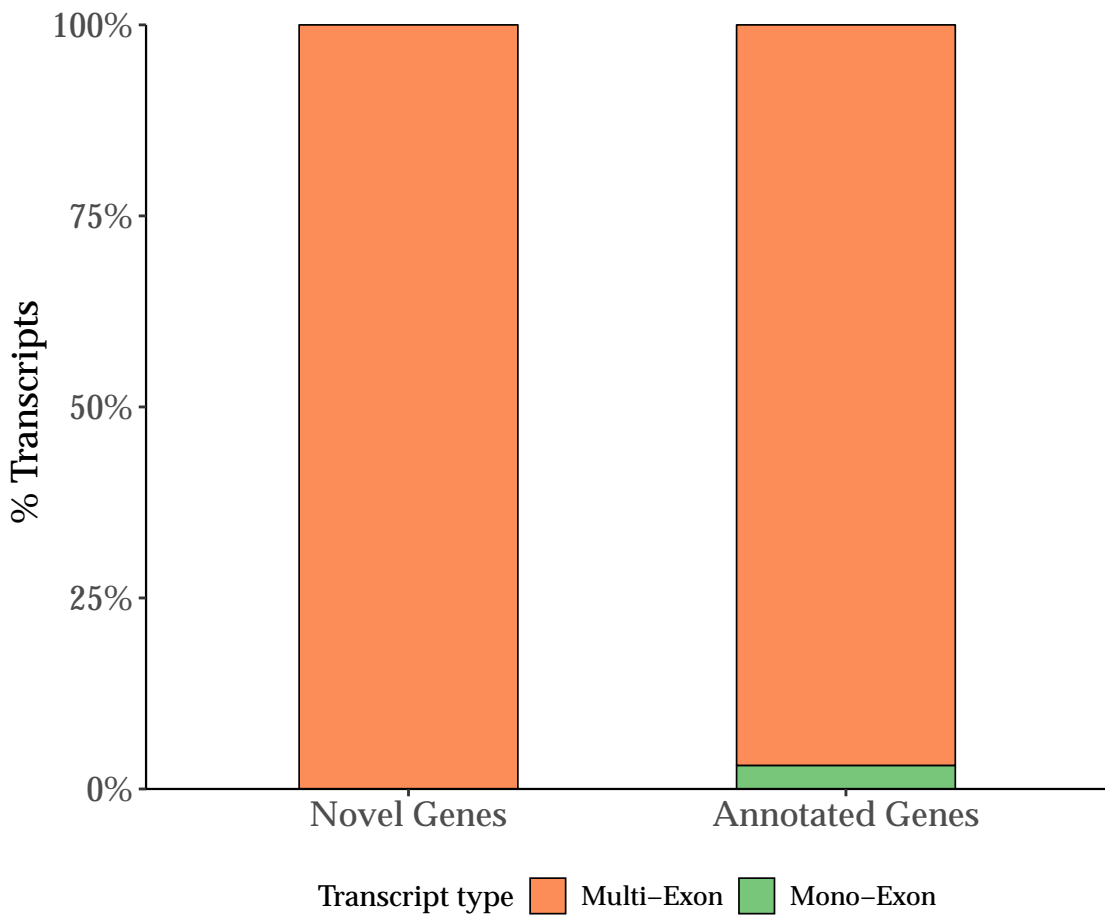
## Number of Isoforms per Gene



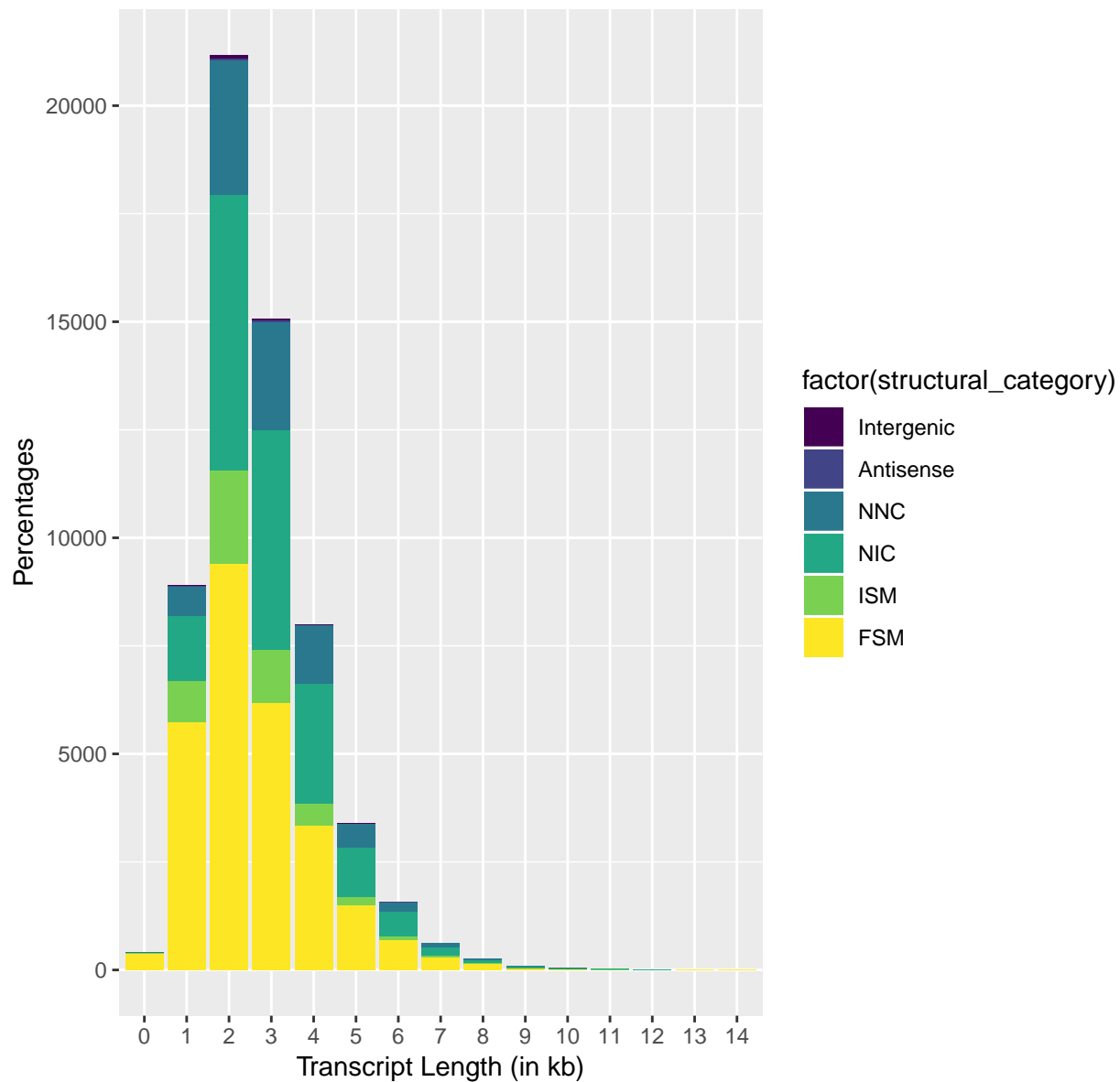
## Number of Isoforms per Gene, Novel vs Known Genes



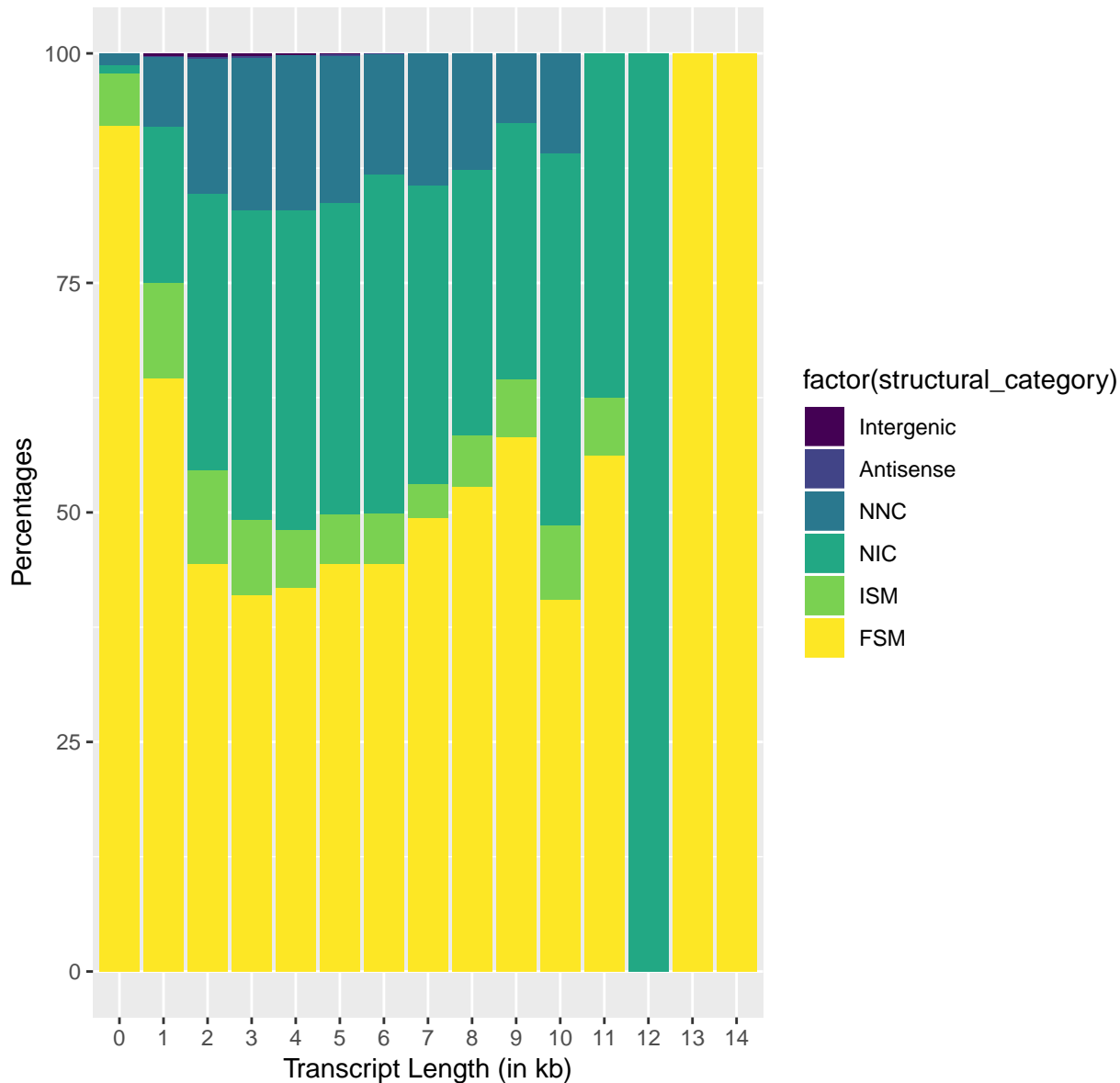
## Distribution of Mono- vs Multi-Exon Transcripts



# Classifications by Transcript Length

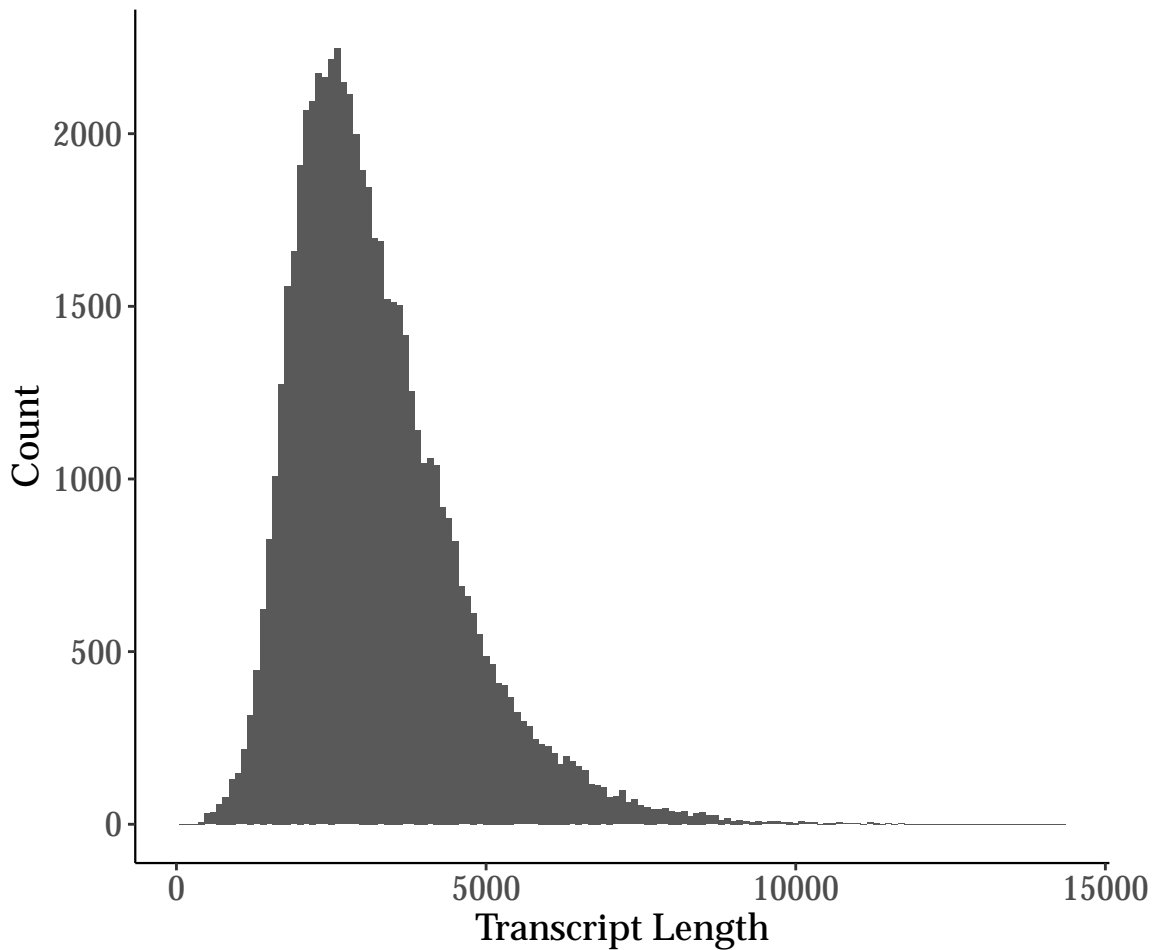


Classifications by Transcript Length, normalized

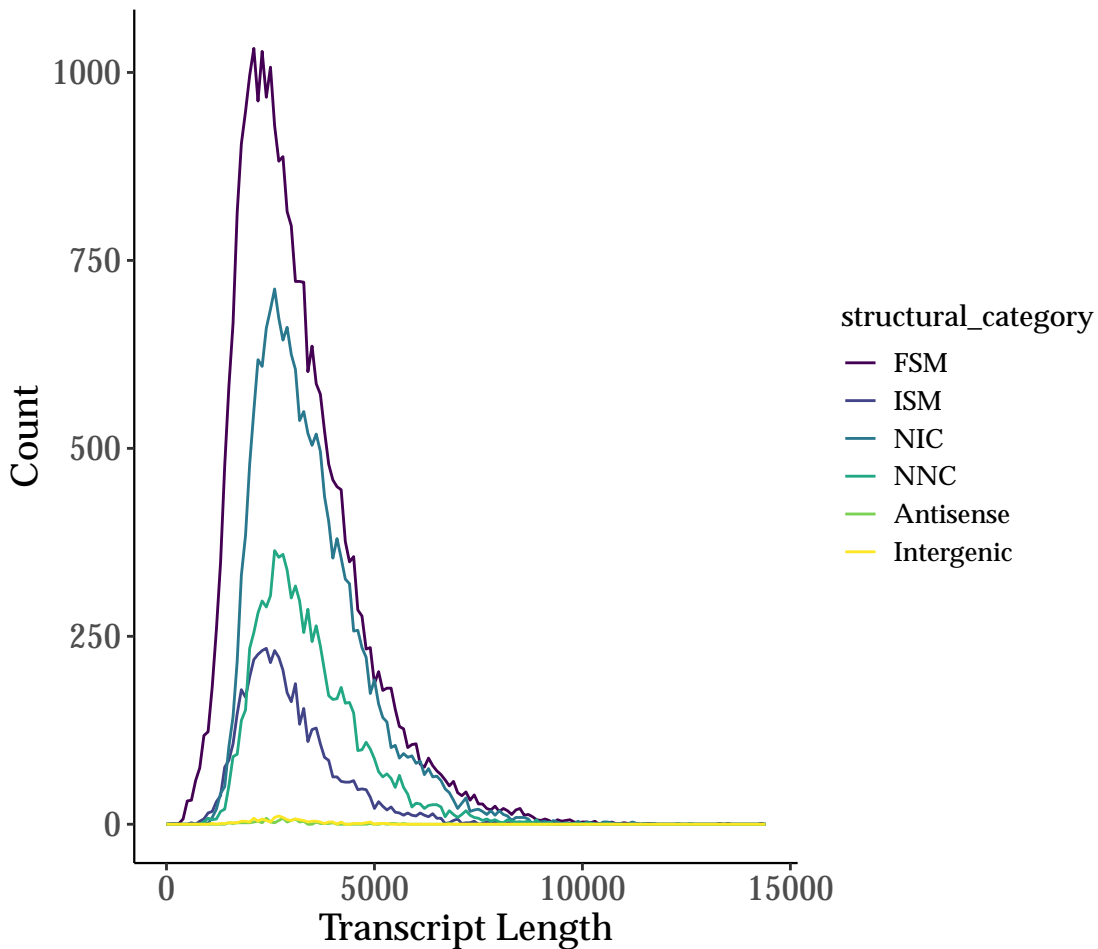




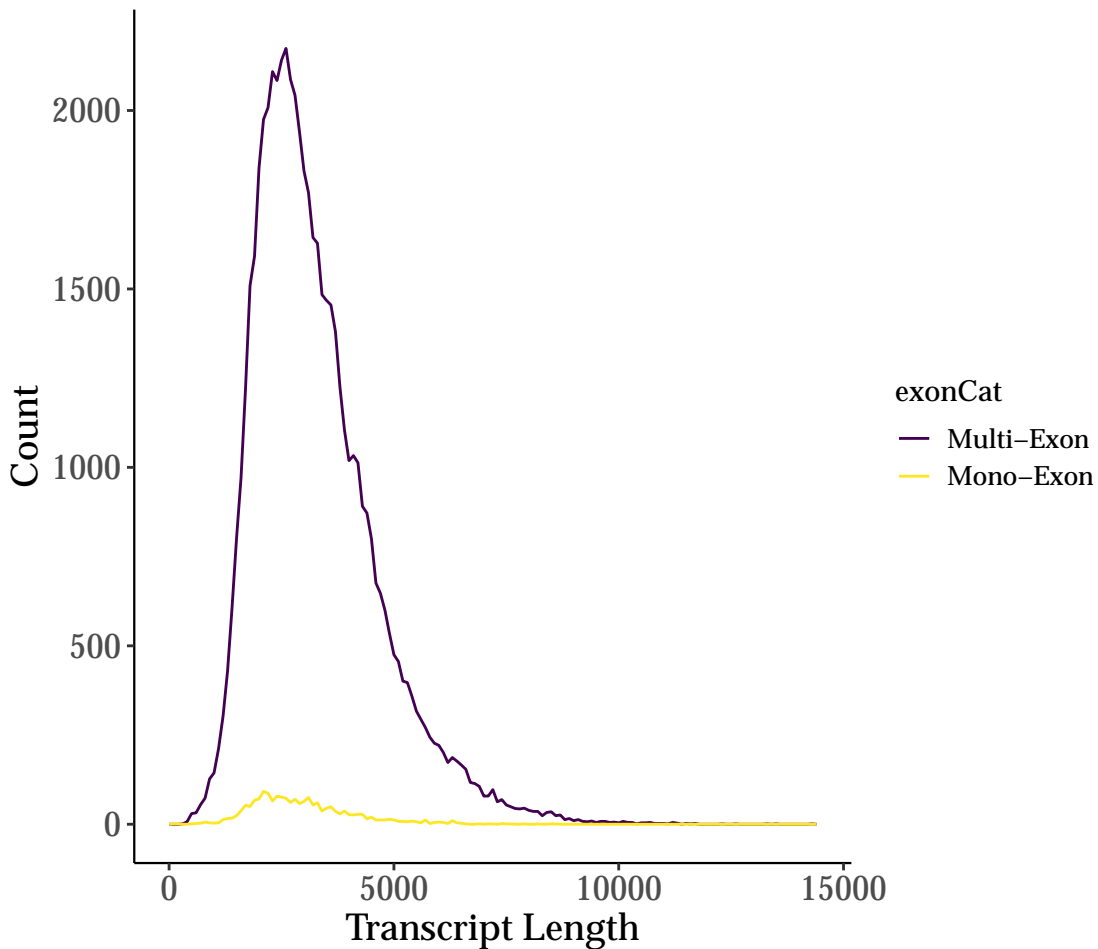
Transcript Lengths, all transcripts



Transcript Lengths, by structural category

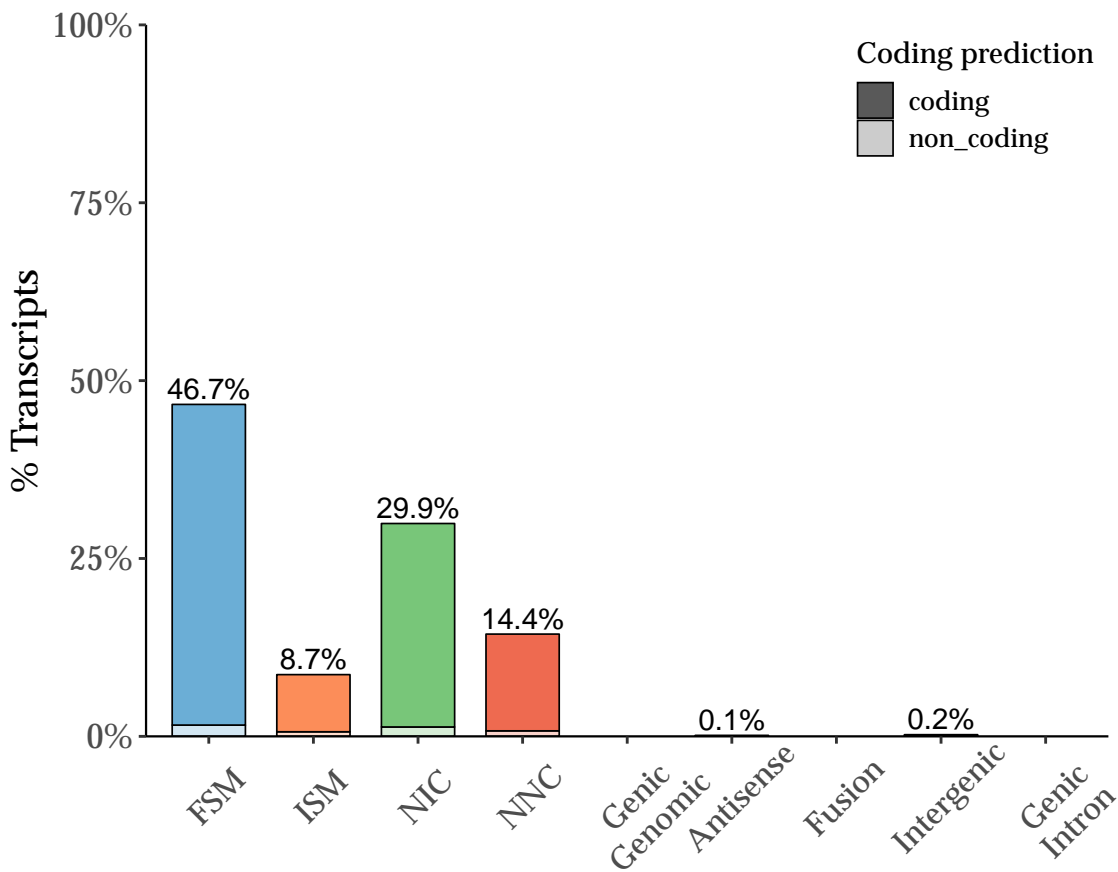


Transcript Lengths, Mono- vs Multi-Exons

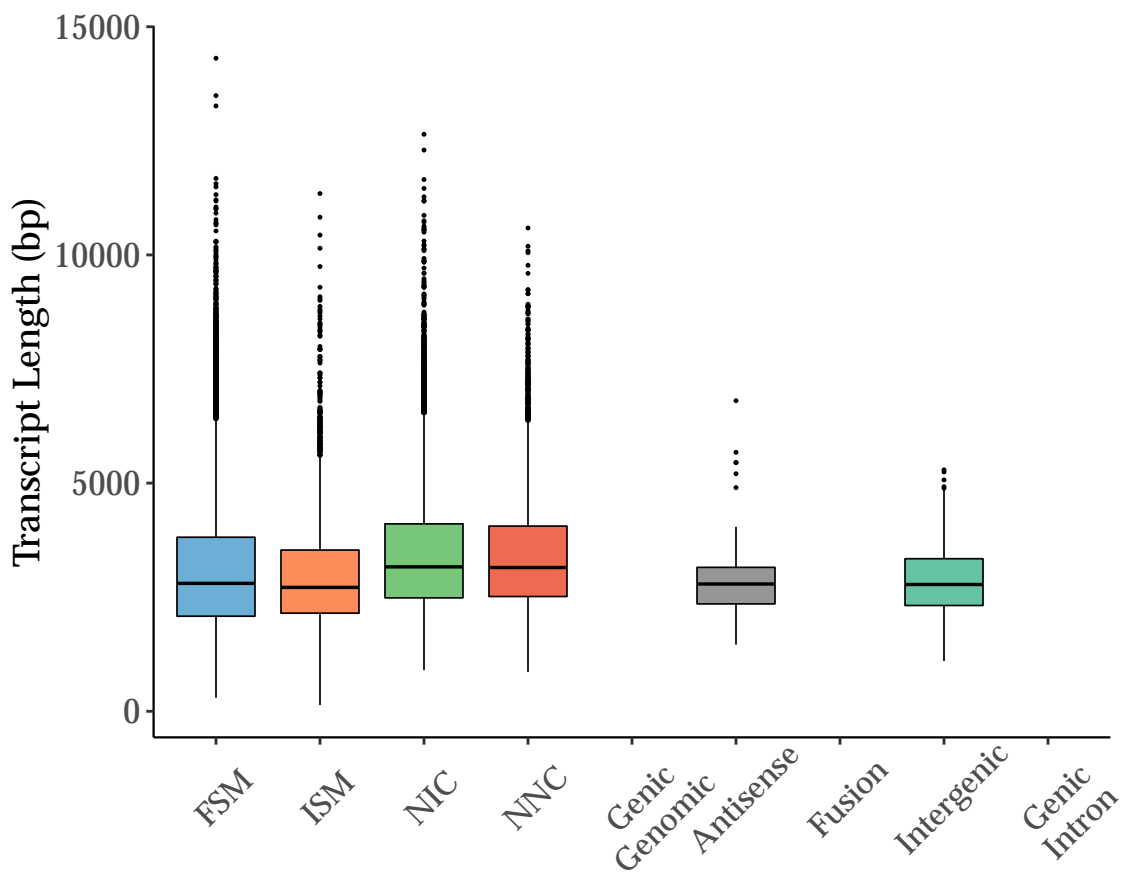


*Structrual Isoform Characterization  
by Splice Junctions*

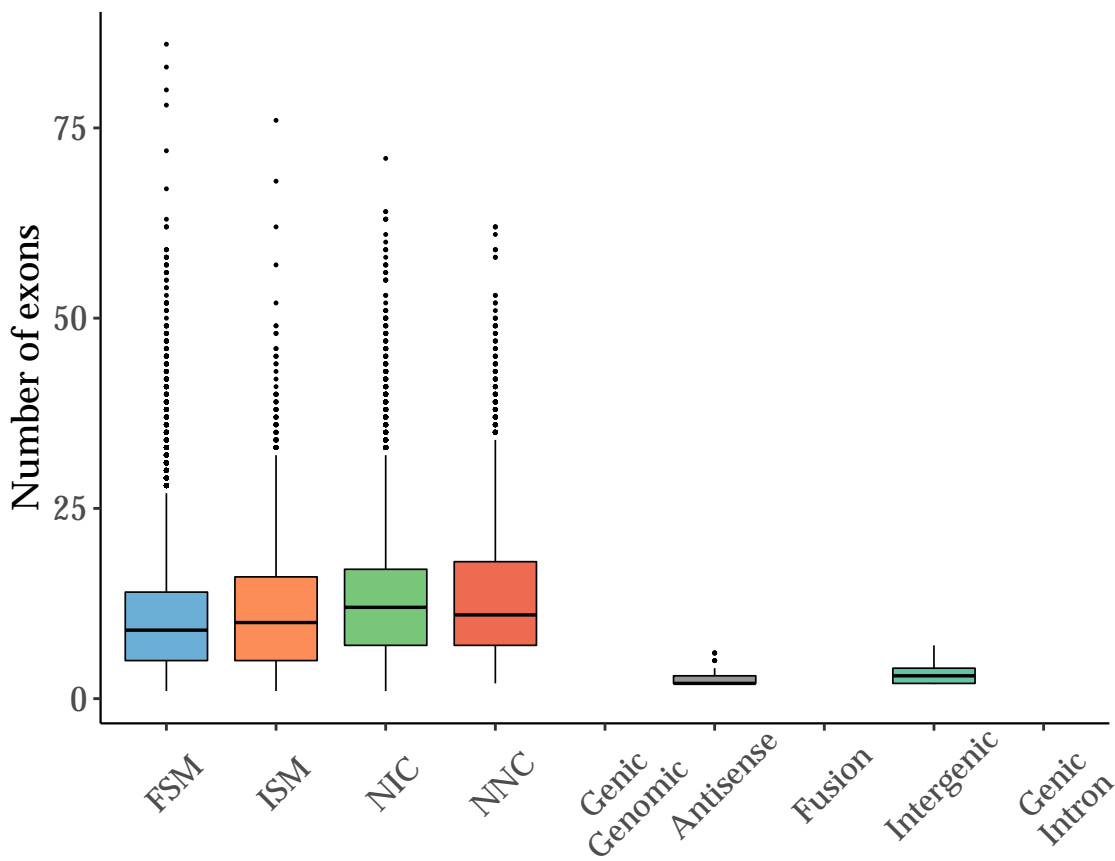
## Isoform distribution across structural categories



Transcript Lengths by Structural Classification

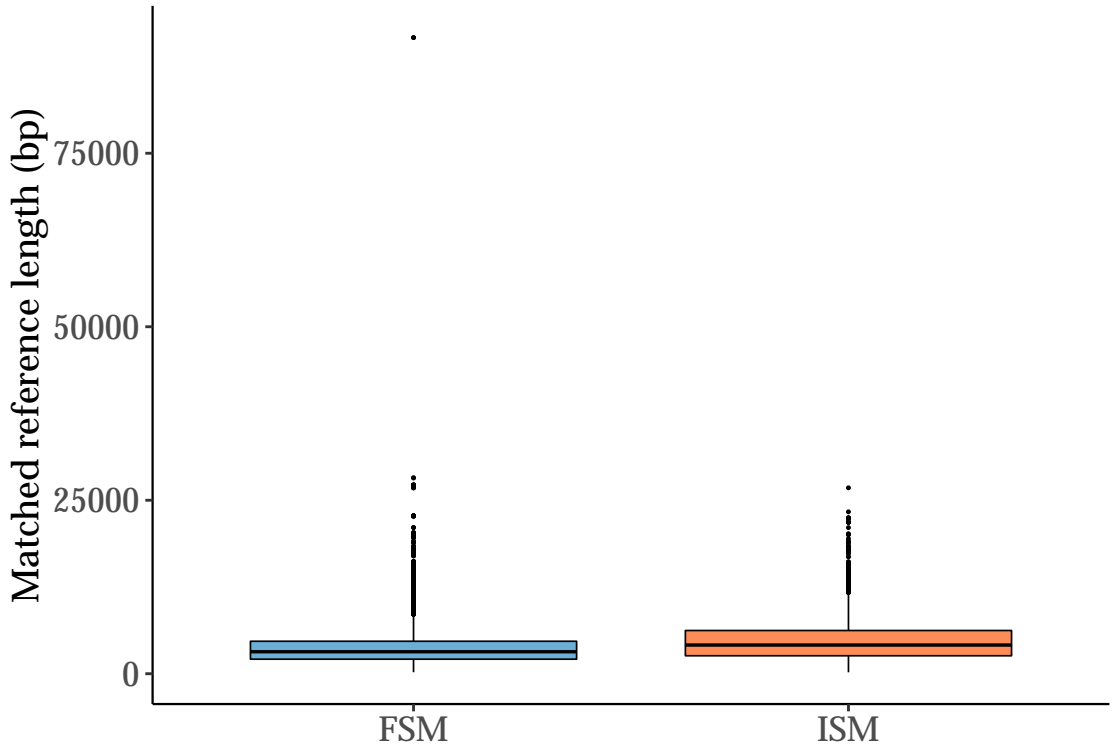


## Exon Counts by Structural Classification



# Length distribution of matched reference transcripts

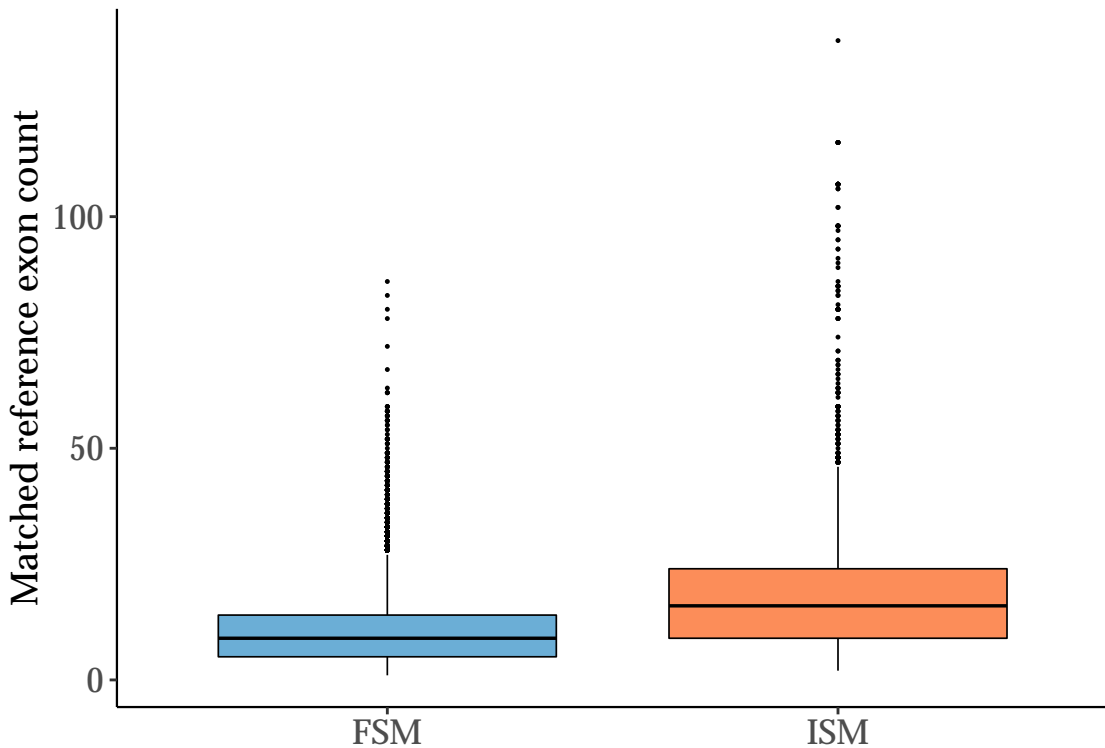
Applicable only to FSM and ISM categories





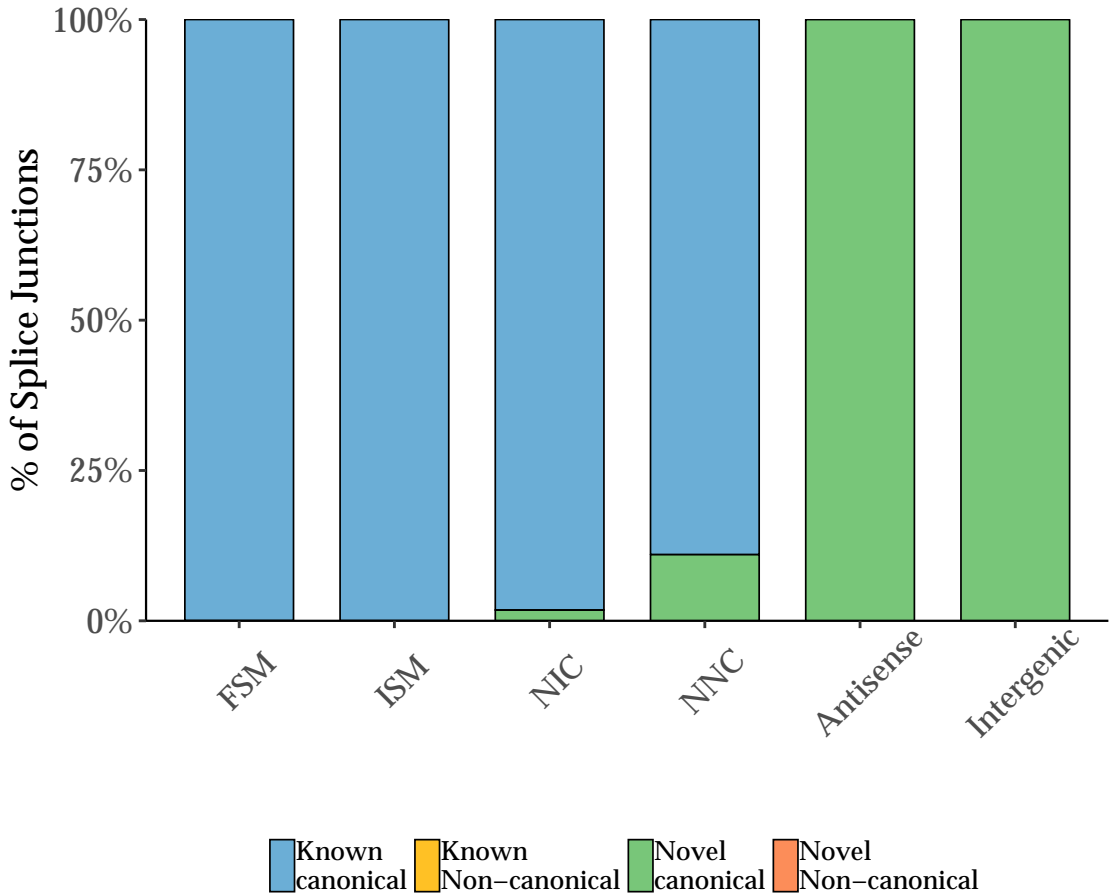
## Exon number distribution of matched reference transcripts

Applicable only to FSM and ISM categories

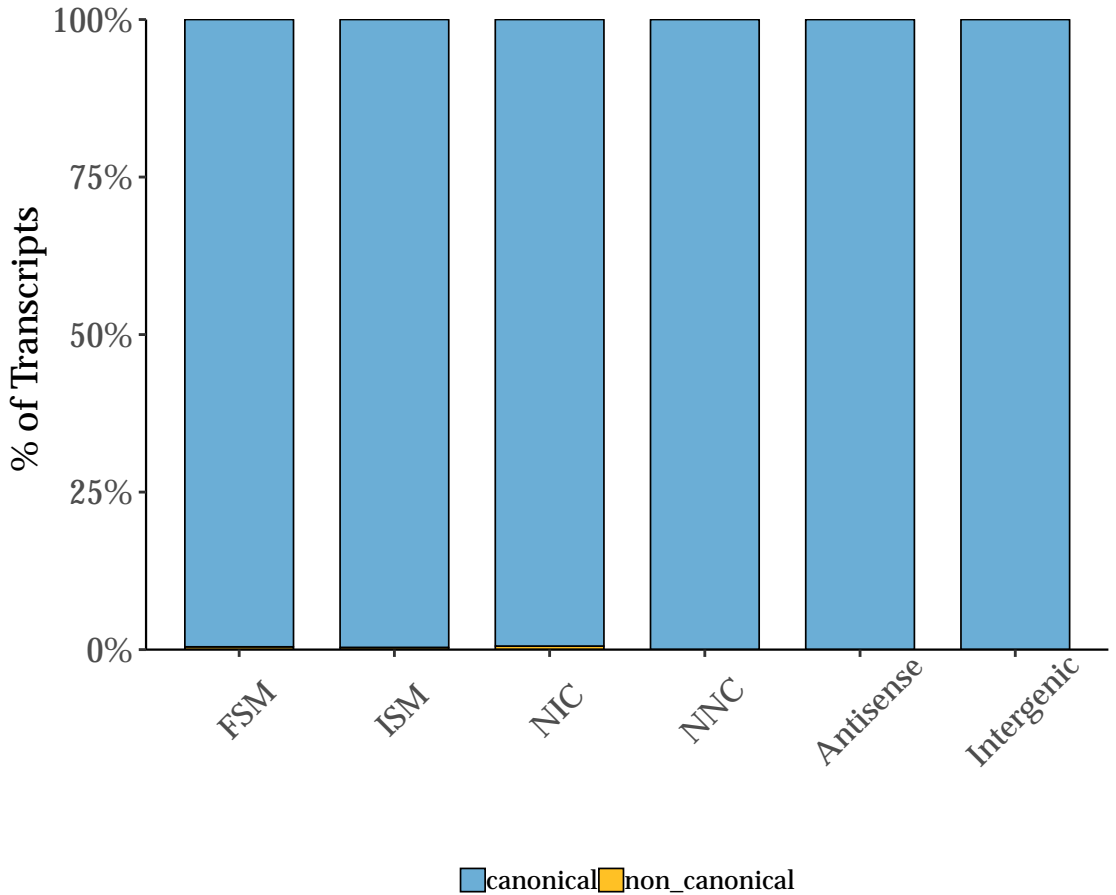


## *Splice Junction Characterization*

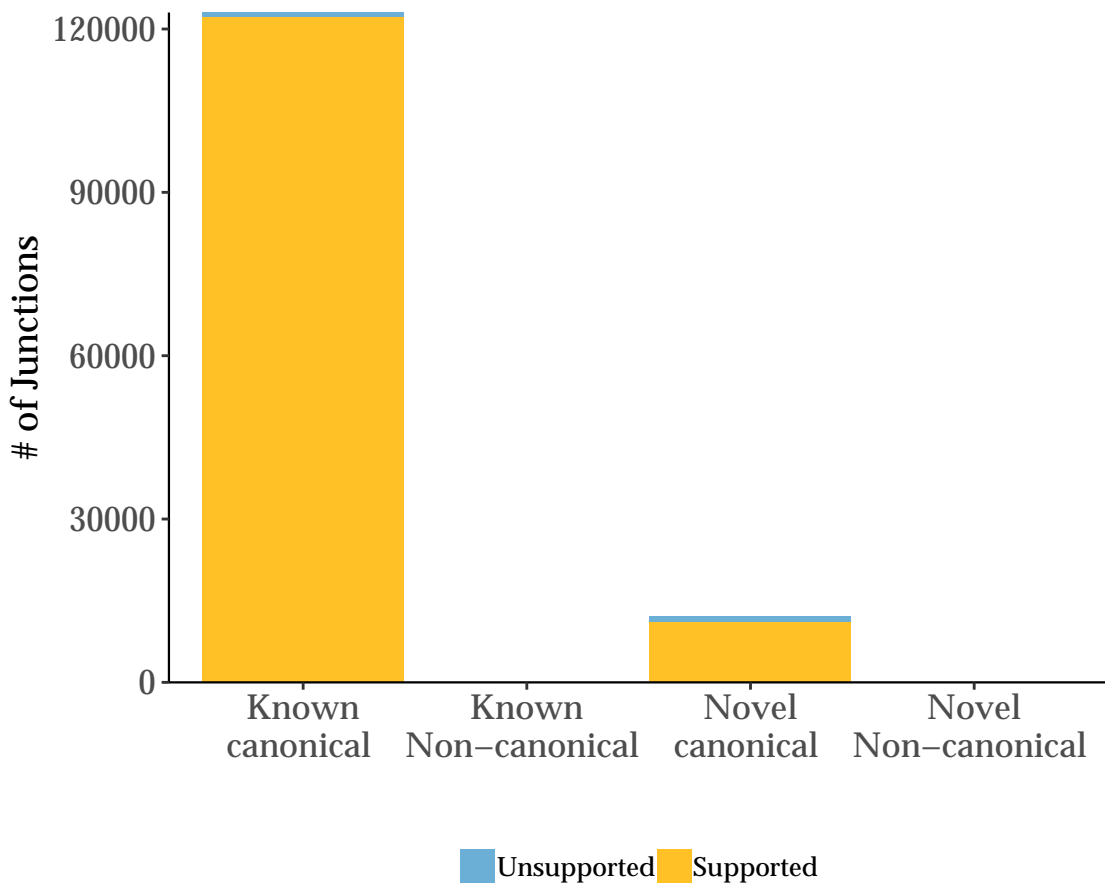
## Distribution of Splice Junctions by Structural Classification



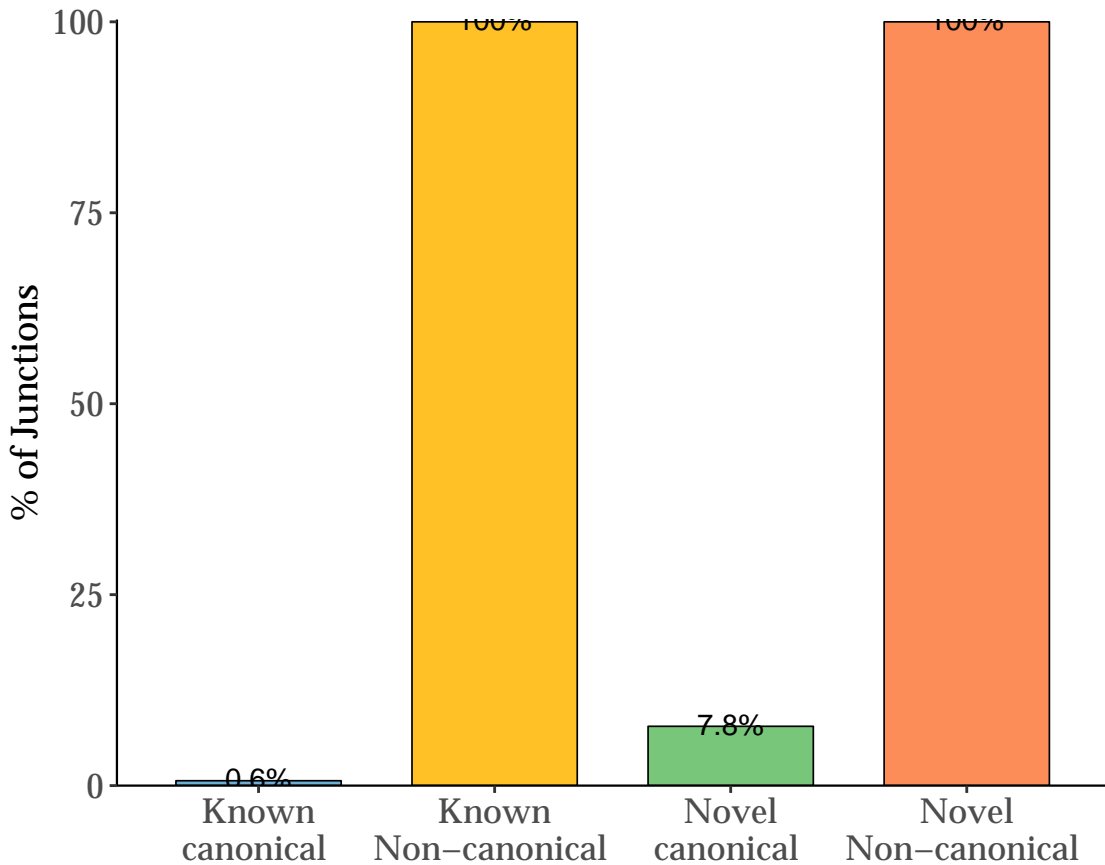
## Distribution of Transcripts by Splice Junctions



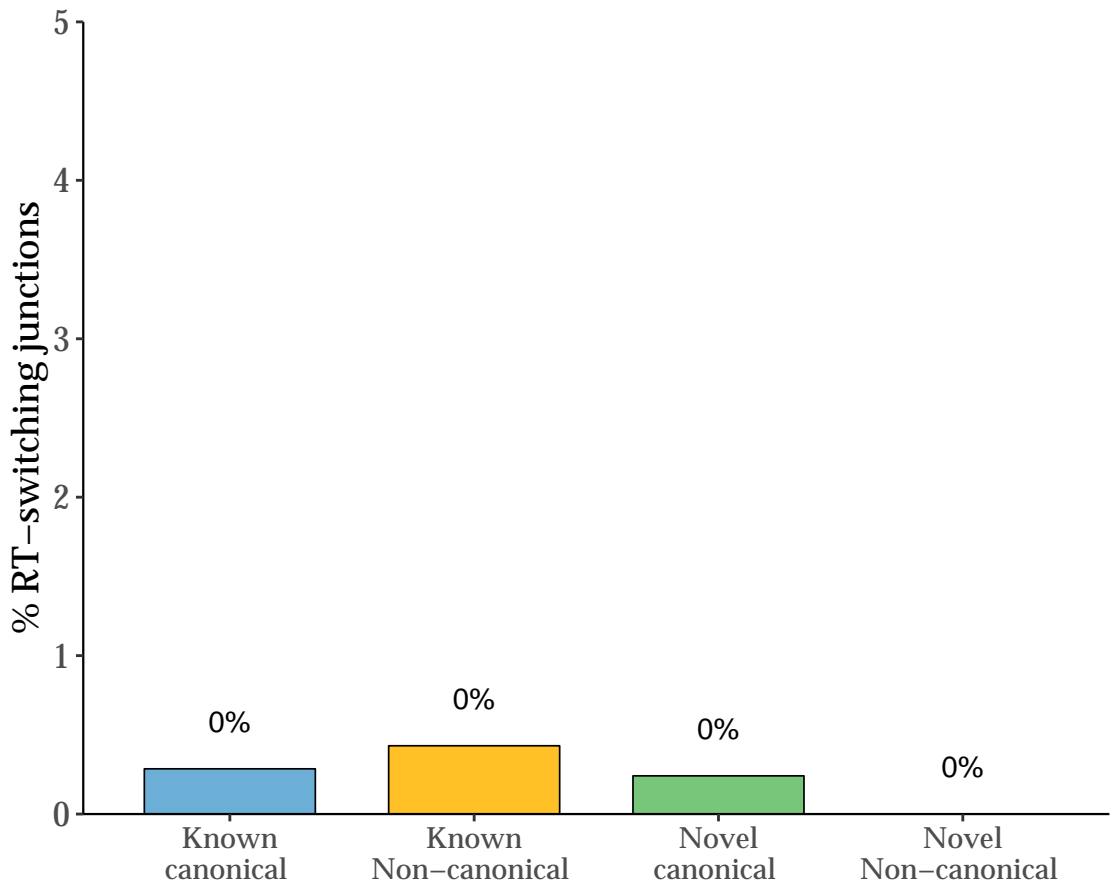
## Unique junctions w/ or w/out short read coverage



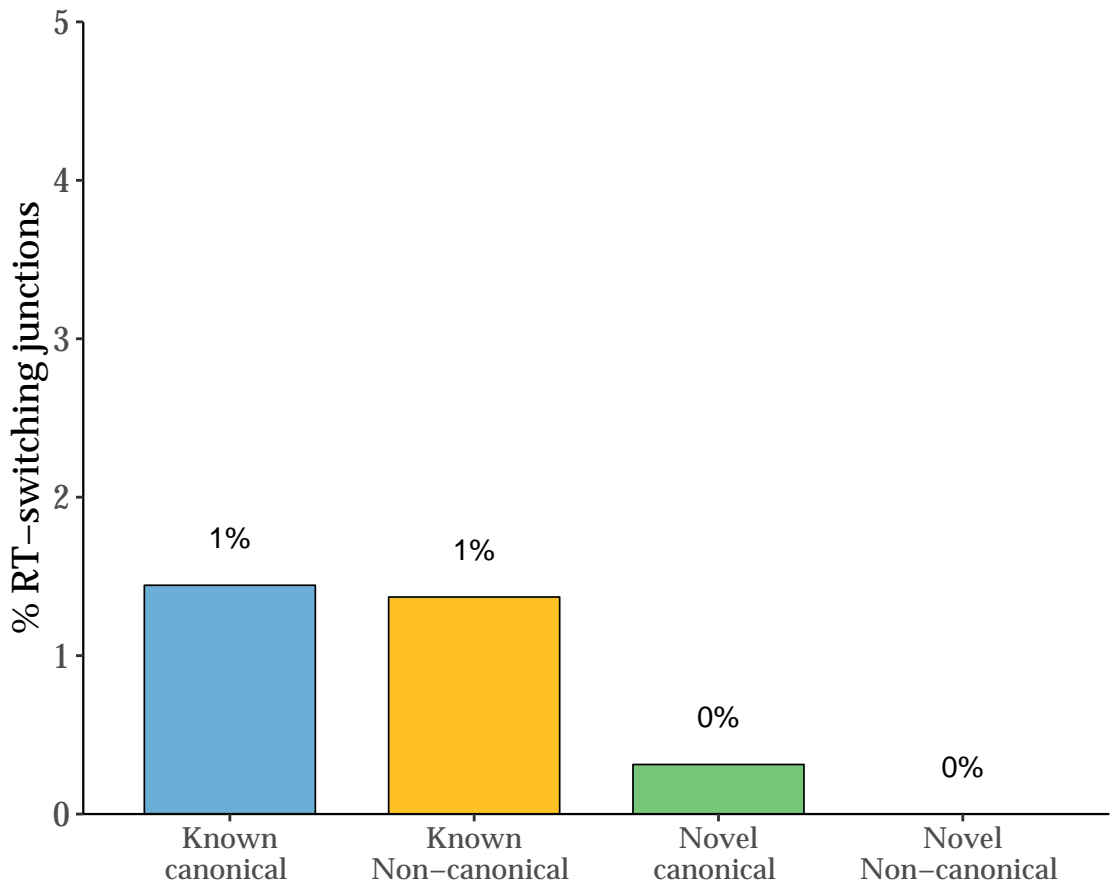
### Unique junctions w/out short read coverage (percentage)



## RT-switching, all junctions



### RT-switching, unique junctions

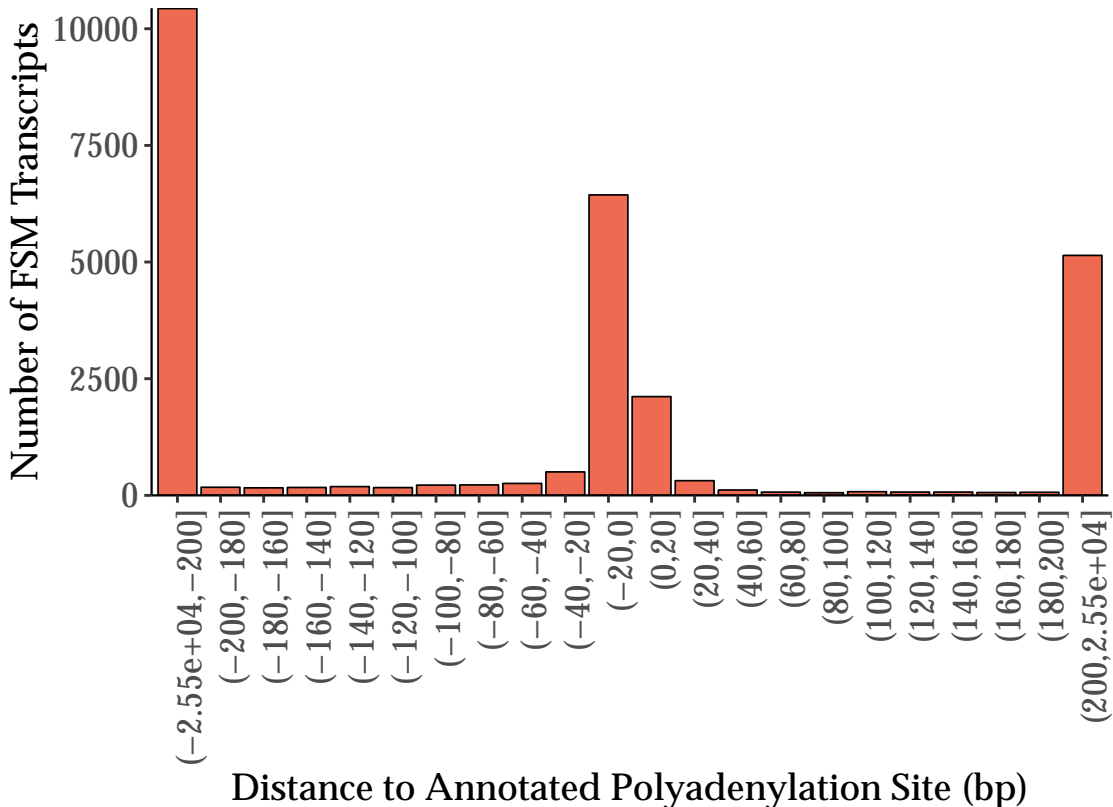




*Comparison with Annotated TSS and PolyA Sites*

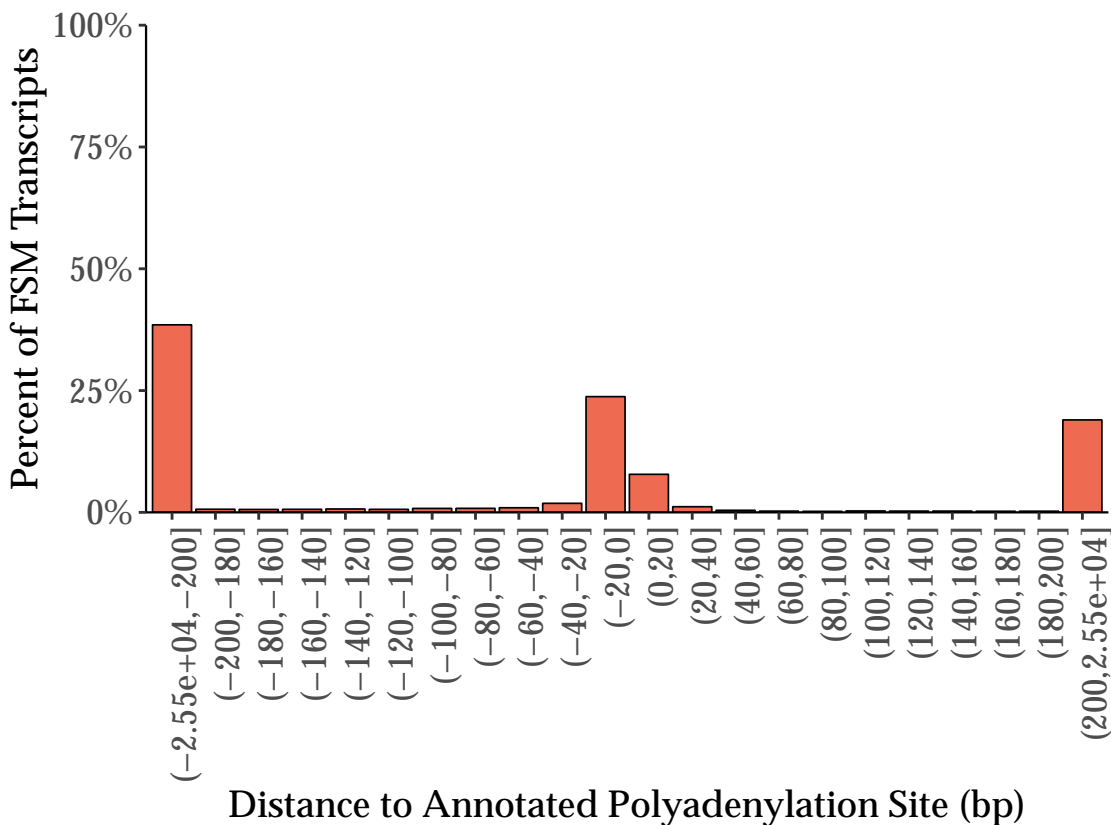
## Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



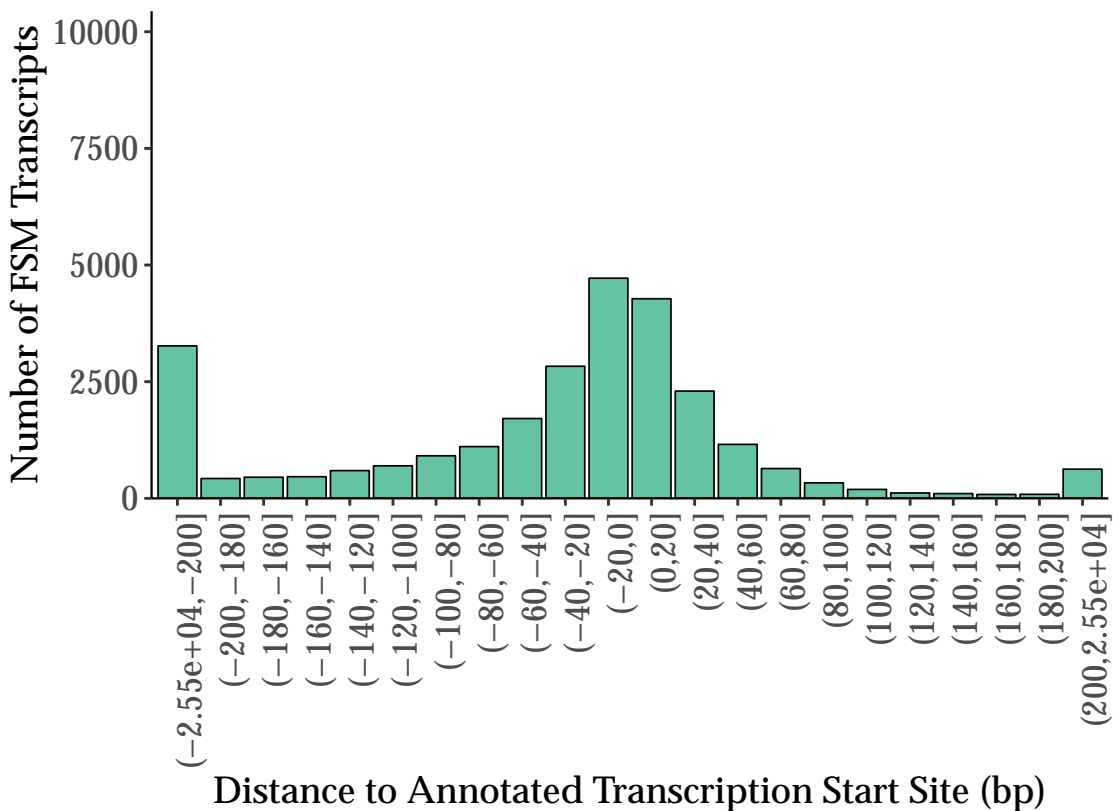
## Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



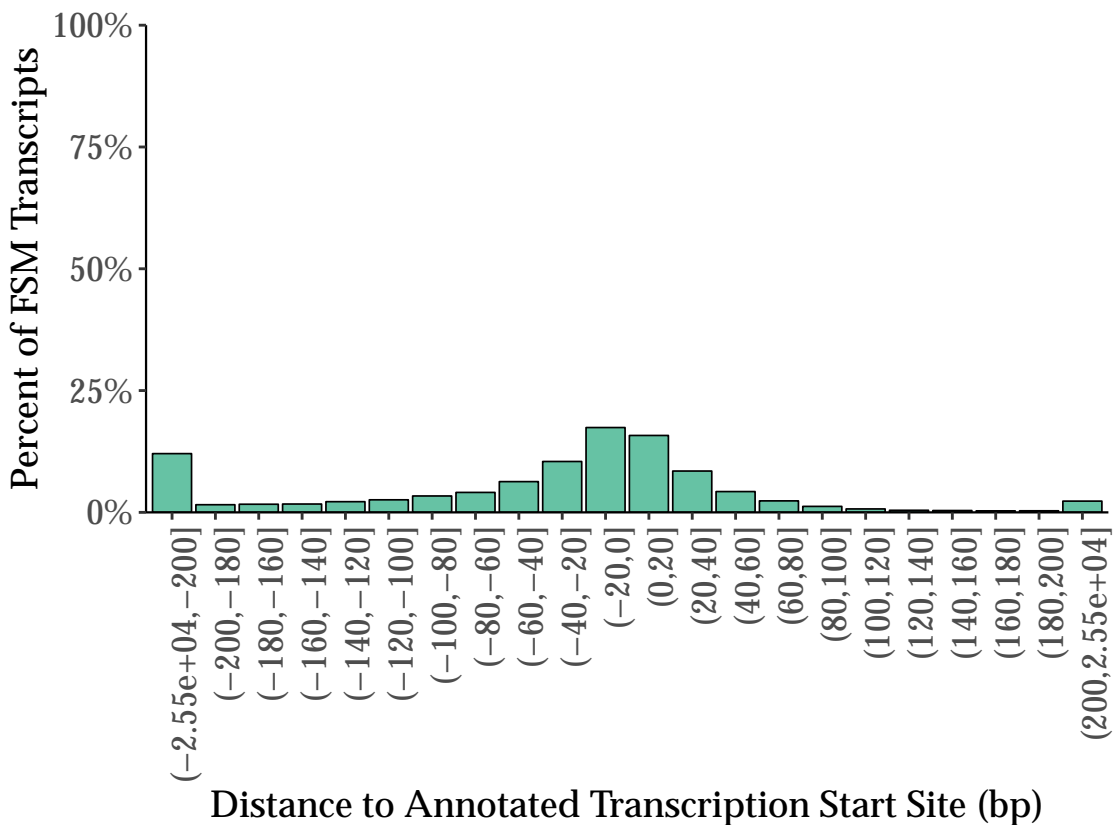
## Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



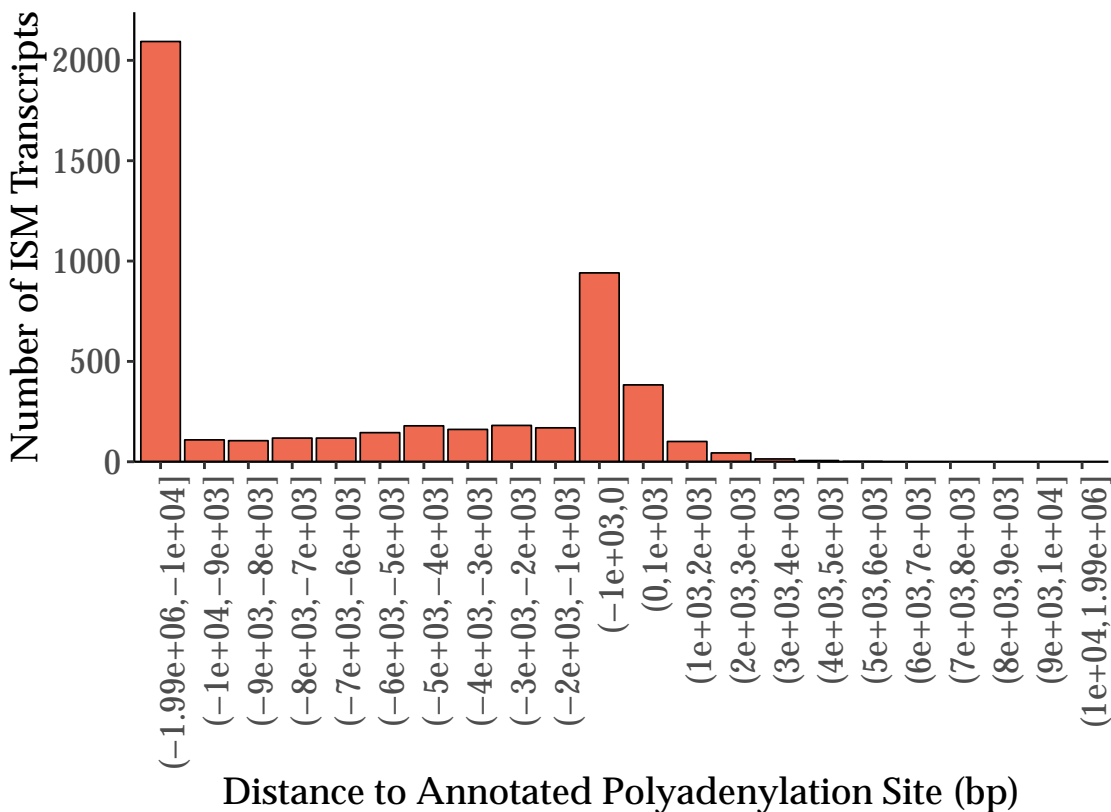
# Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



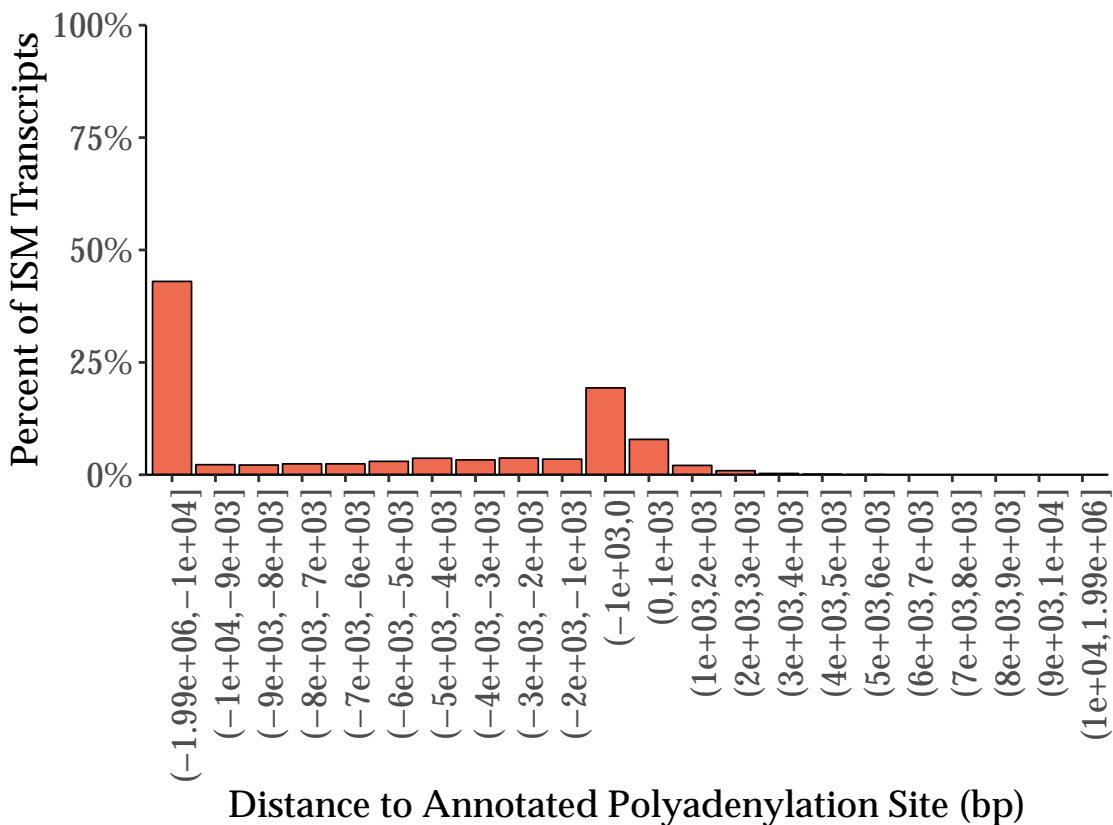
## Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



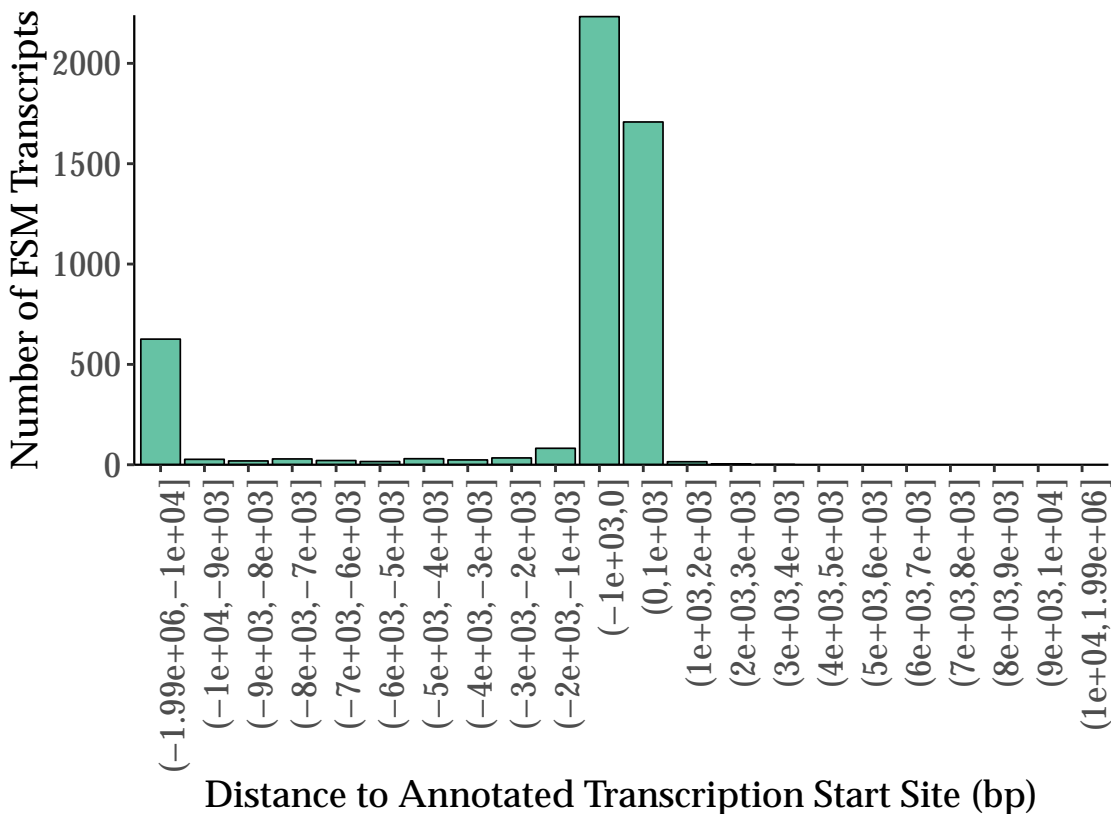
## Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



# Distance to Annotated Transcription Start Site, ISM only

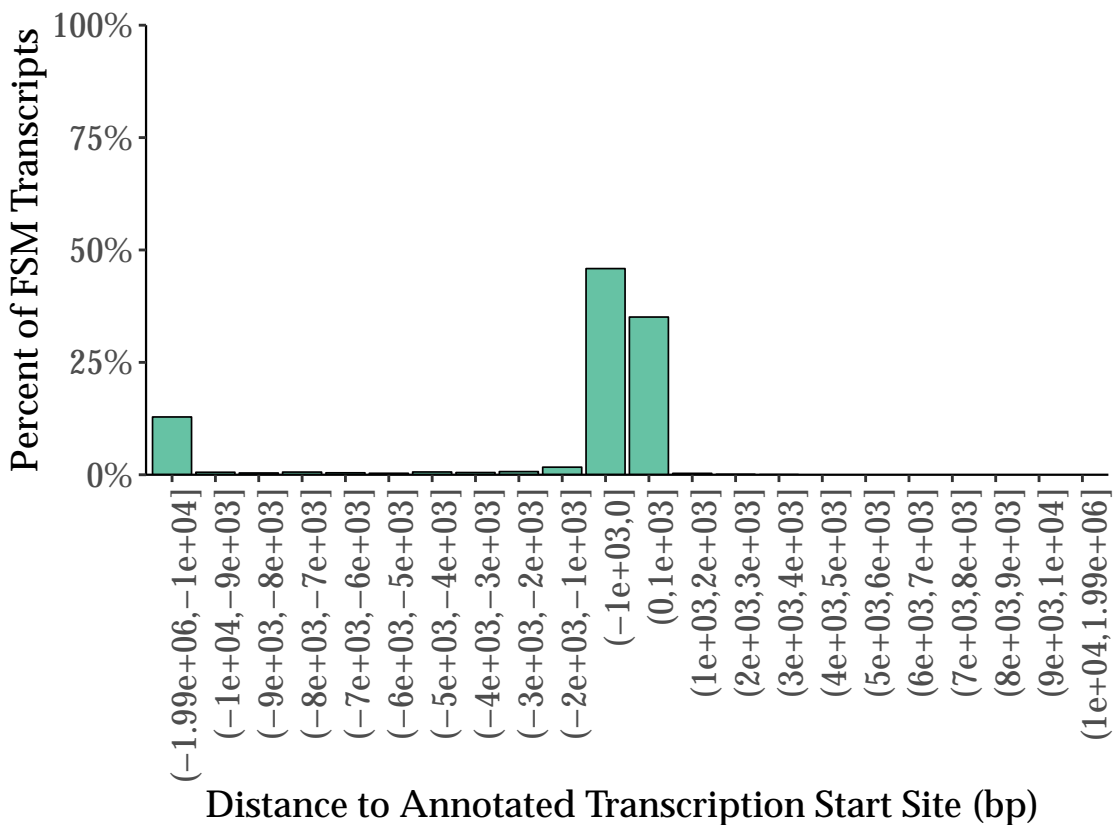
Negative values indicate downstream of annotated TSS





## Distance to Annotated Transcription Start Site, ISM only

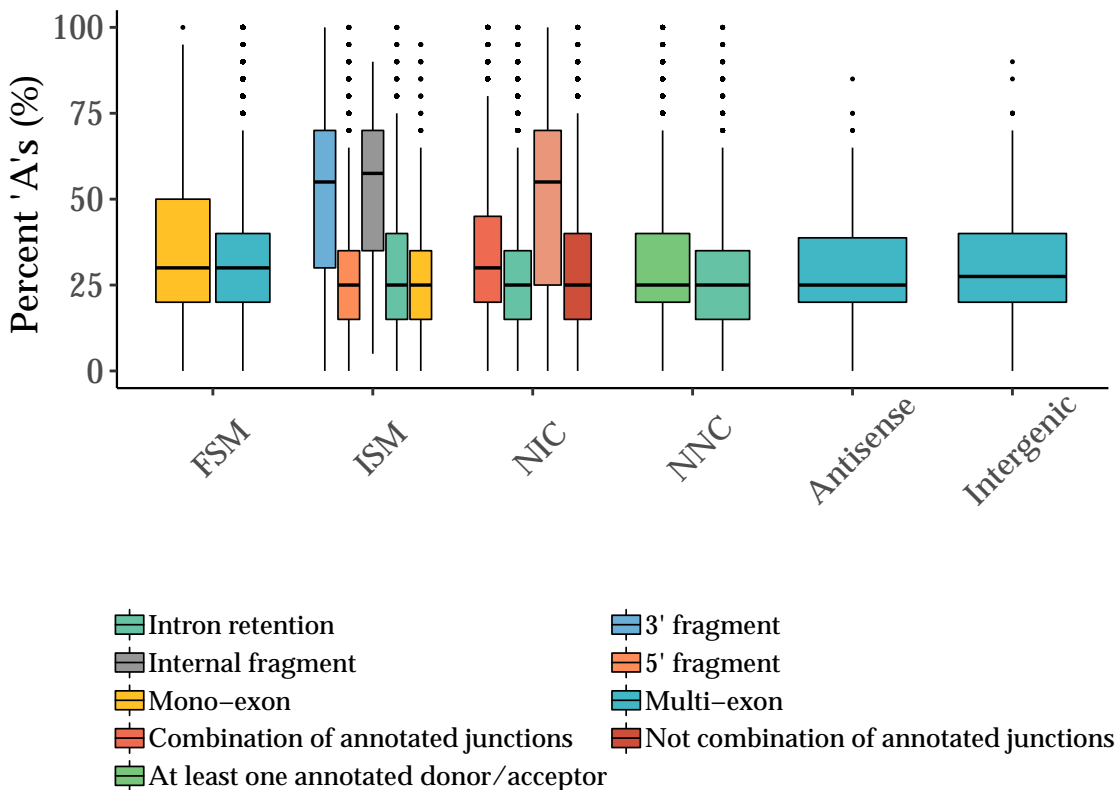
Negative values indicate downstream of annotated TSS



*Intra-Priming Quality Check*

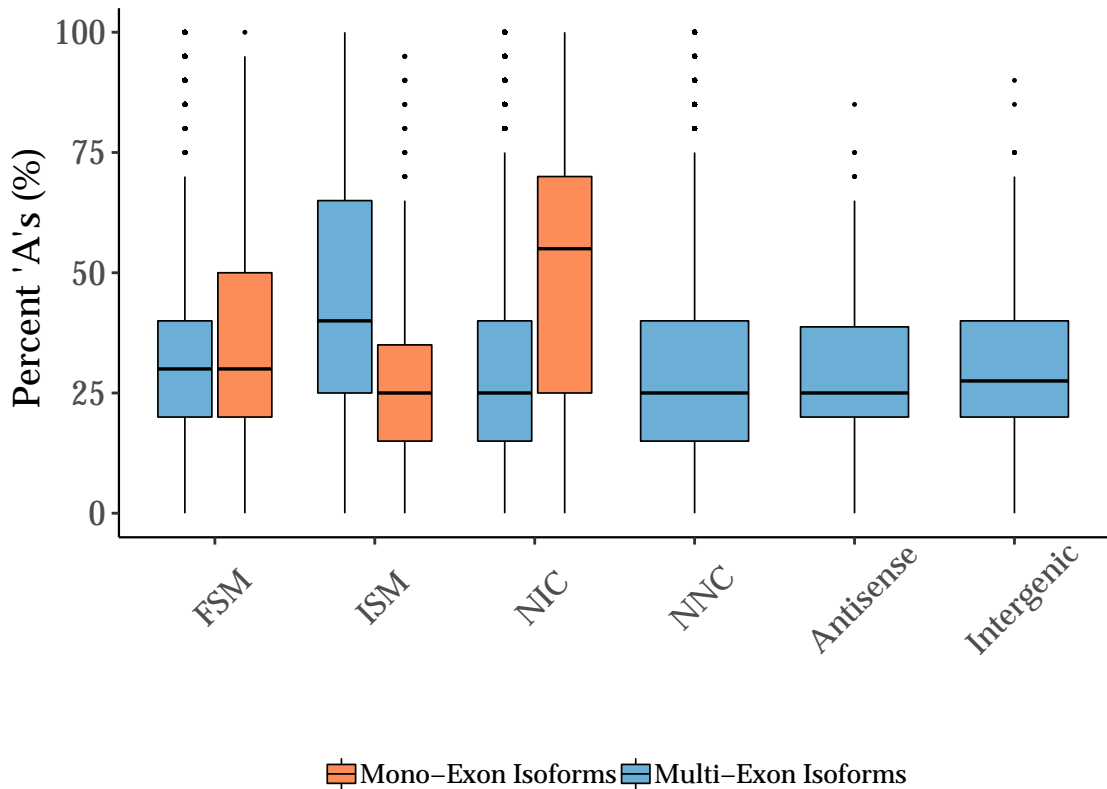
## Possible Intra-Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



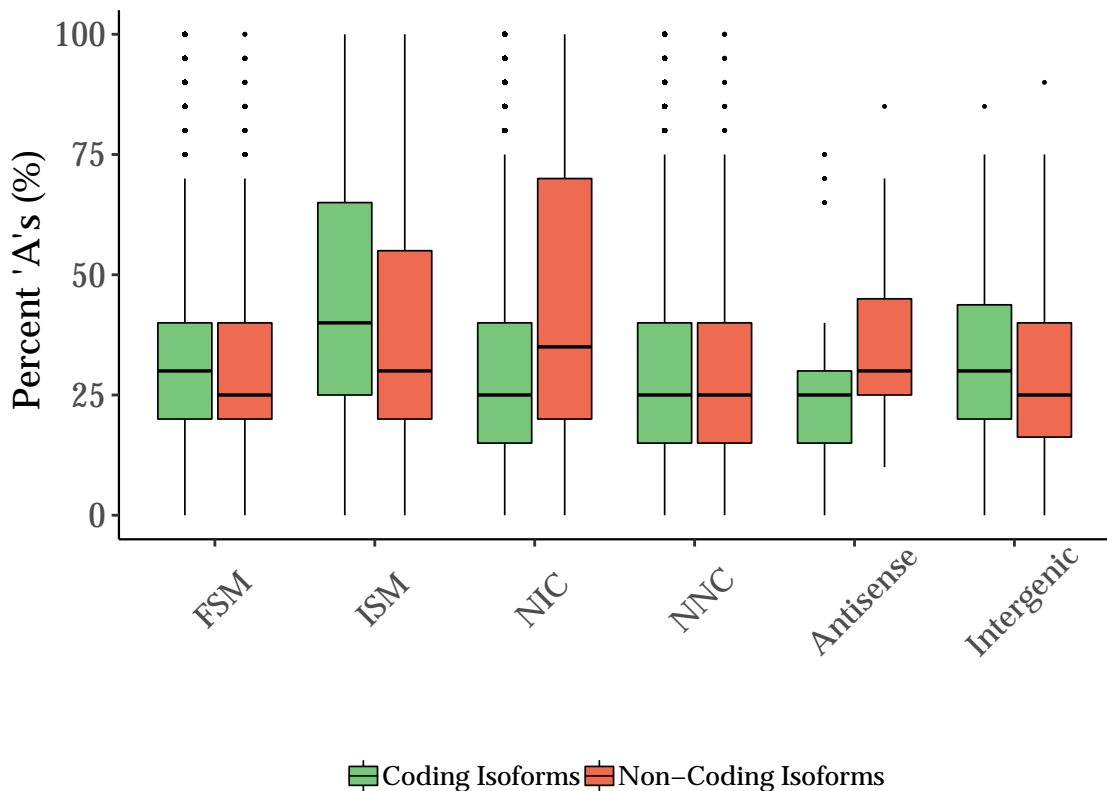
## Possible Intra-Priming, Mono- vs Multi-Exon

Percent of genomic 'A's in downstream 20 bp



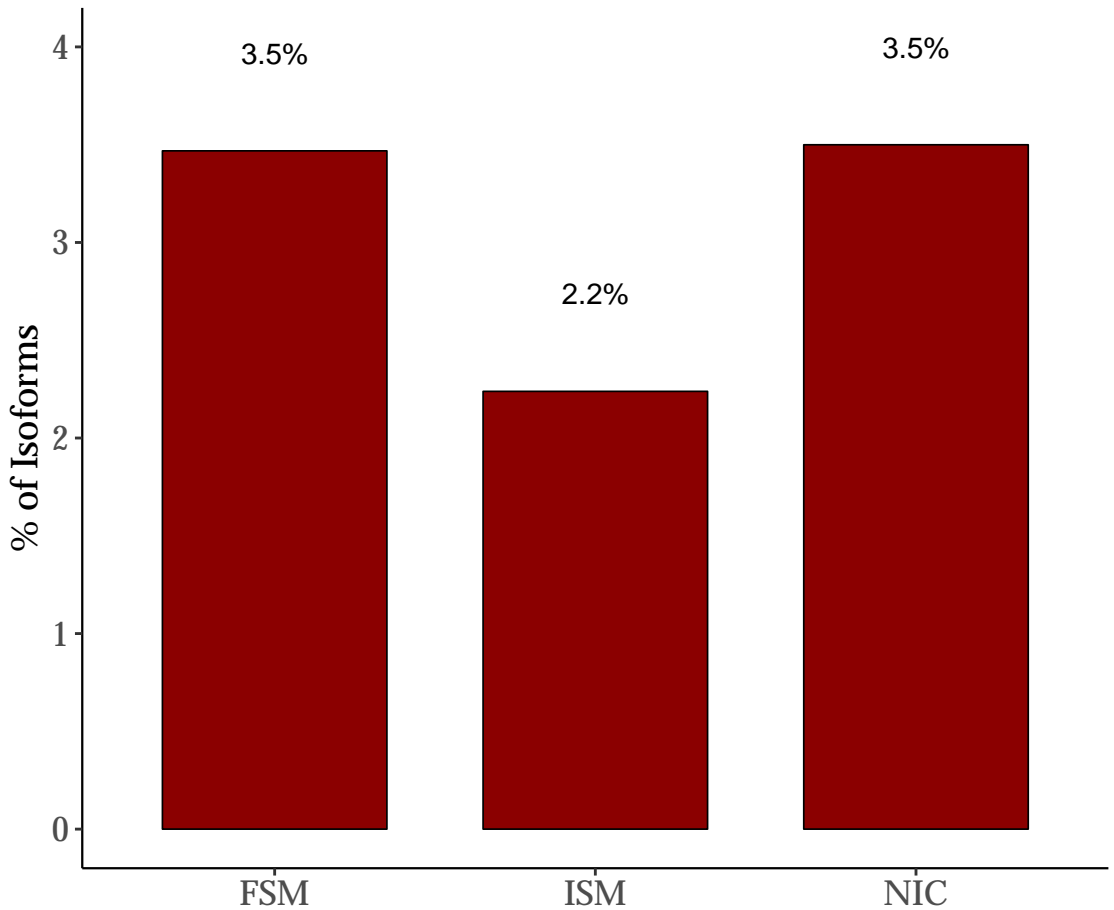
## Possible Intra-Priming, Coding vs Non-Coding

Percent of genomic 'A's in downstream 20 bp



*Quality Controls*

## Incidence of RT-switching



## Incidence of Non-Canonical Junctions

