# SQANTI3 report

Unique Genes: 31829

Unique Isoforms: 107649

# Gene classification

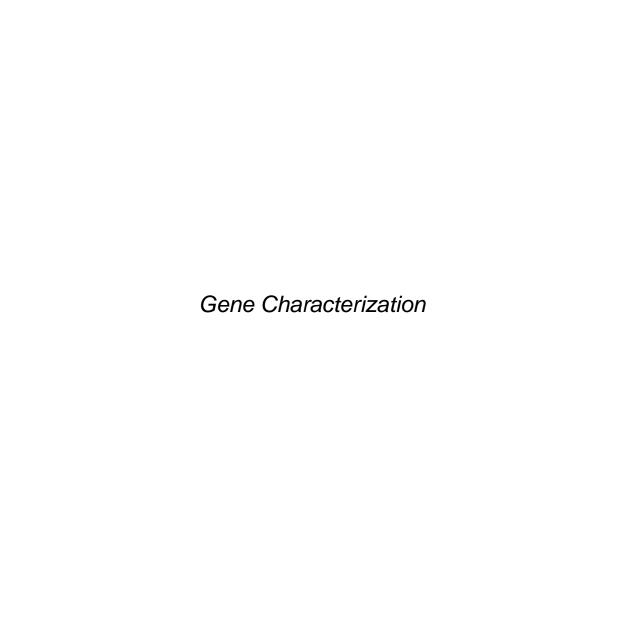
Category	# Genes
Annotated Genes	28124
Novel Genes	3705

# Splice Junction Classification

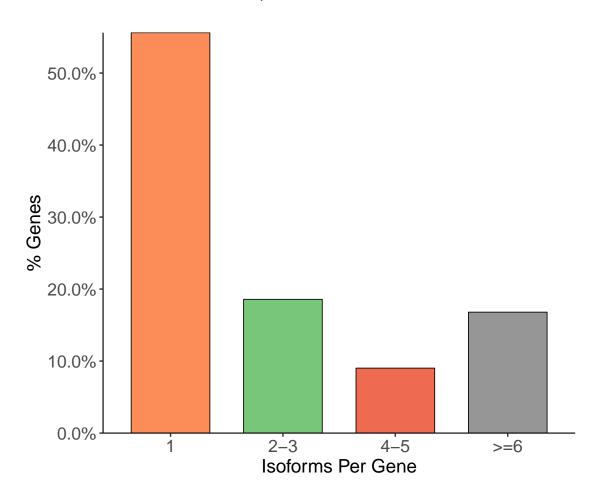
Category	# SJs	Percent
Known canonical	220195	88.47
Known Non-canonical	860	0.35
Novel canonical	27810	11.17
Novel Non-canonical	37	0.01

# Characterization of transcripts based on splice junctions

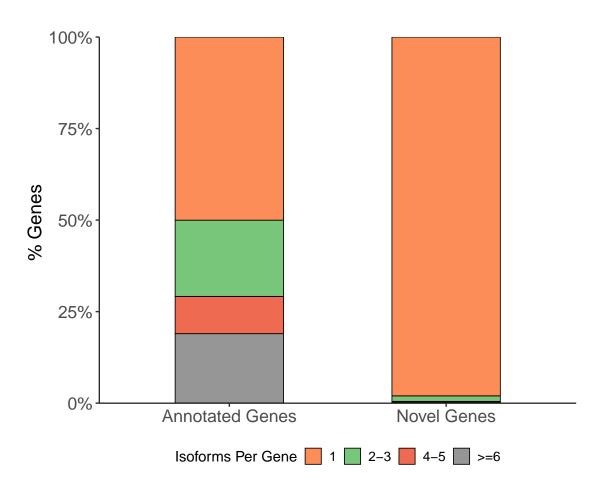
Category	# Isoforms
FSM	55043
ISM	7585
NIC	19082
NNC	21055
Genic	72
Genomic	12
Antisense	575
Fusion	947
Intergenic	3290
Genic	0
Intron	



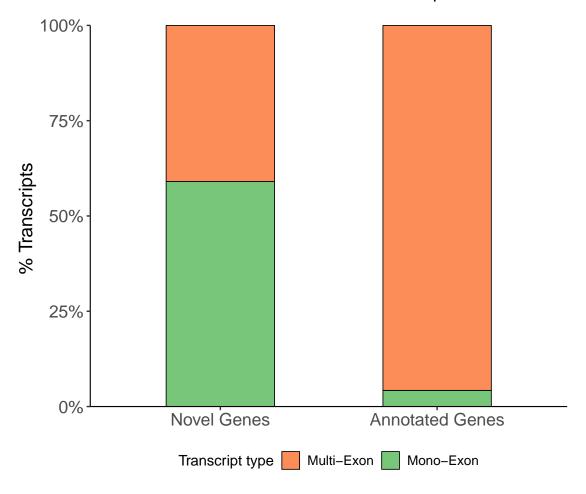
# Number of Isoforms per Gene

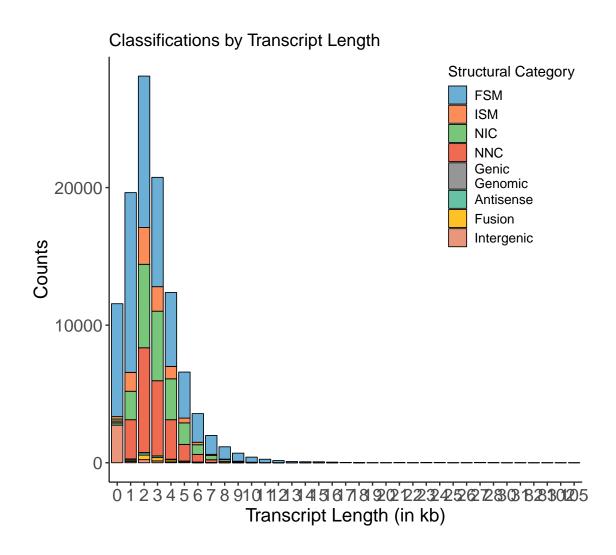


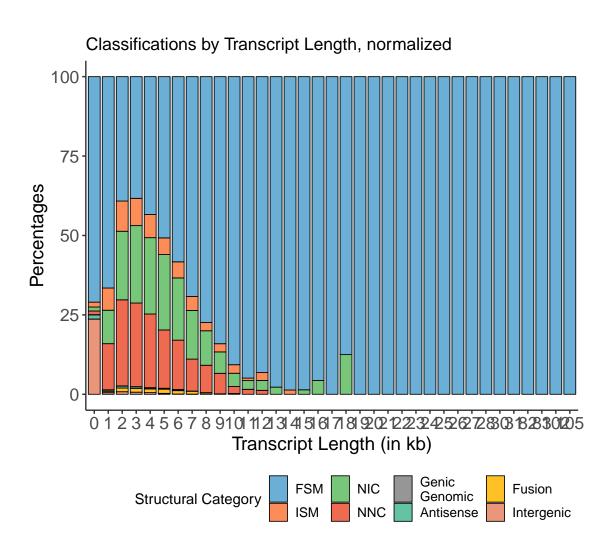
#### Number of Isoforms per Gene, Known vs Novel Genes



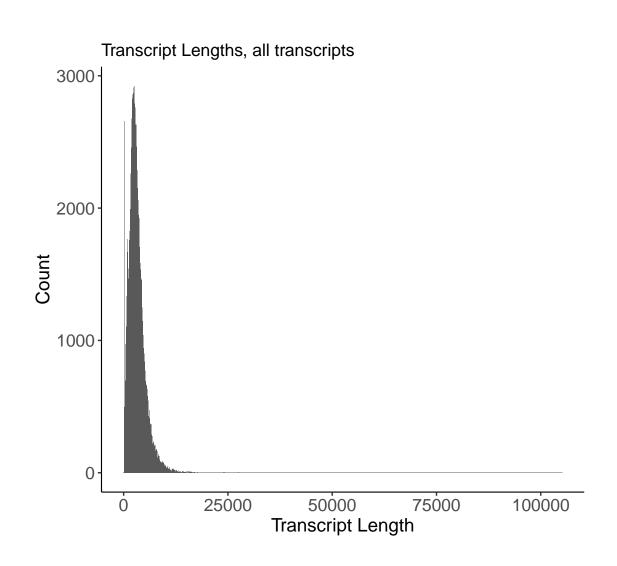
#### Distribution of Mono- vs Multi-Exon Transcripts

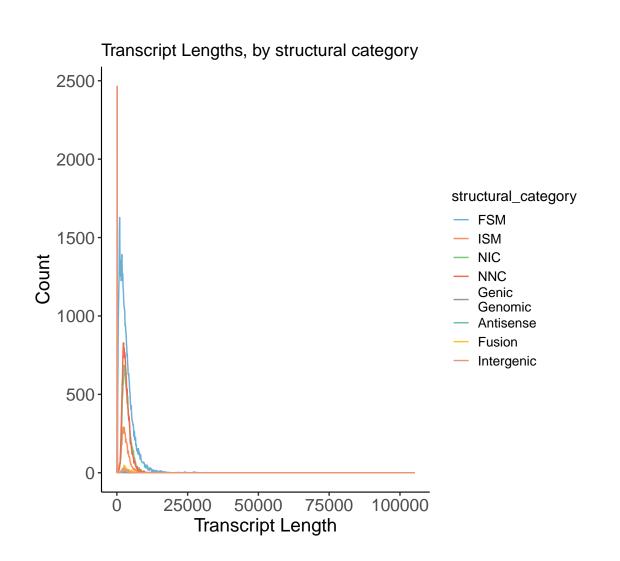


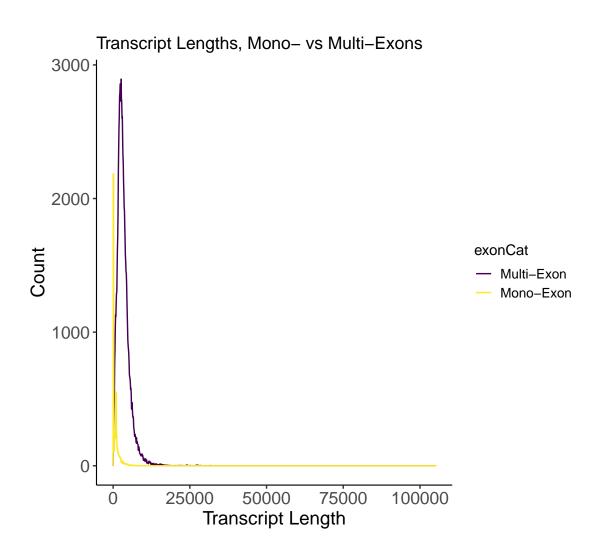


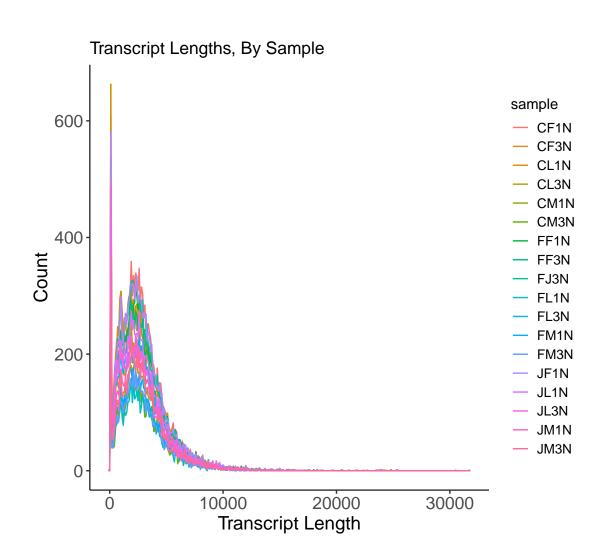


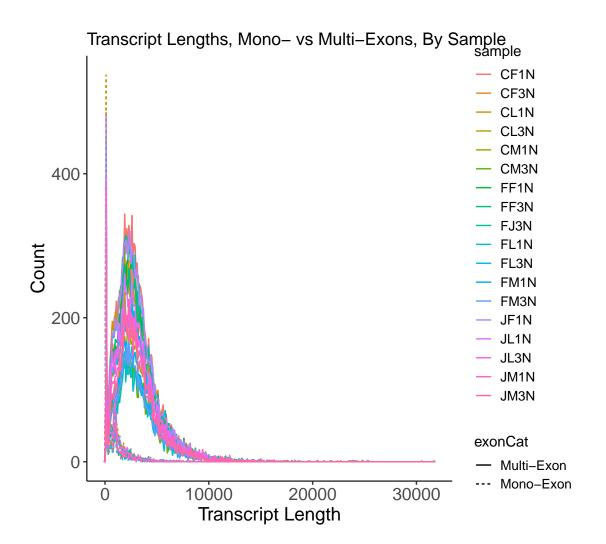
Number of FL reads per Gene by type of gene annotation





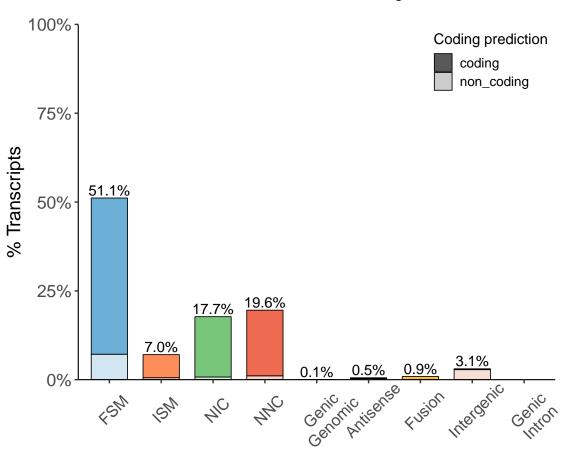




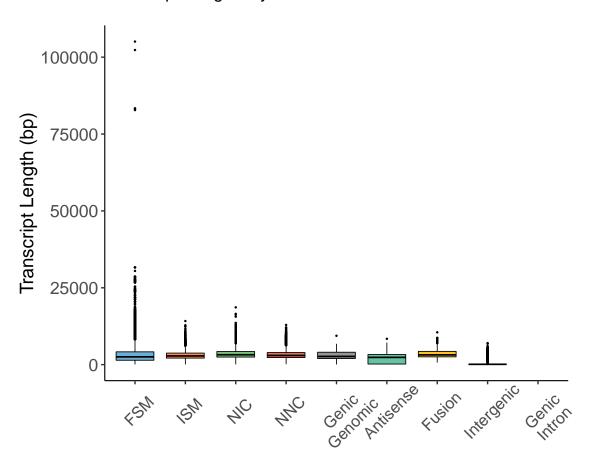


# Structural Isoform Characterization by Splice Junctions

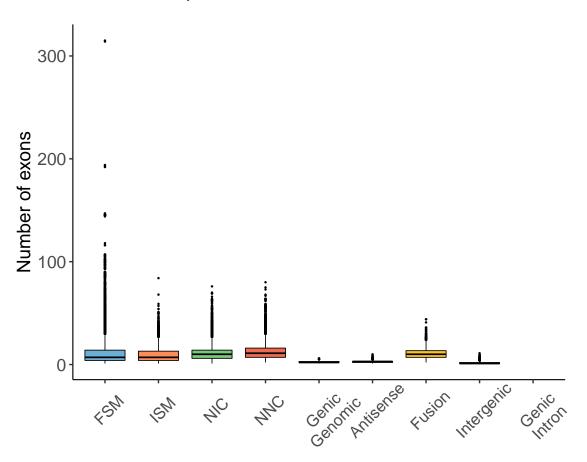
#### Isoform distribution across structural categories



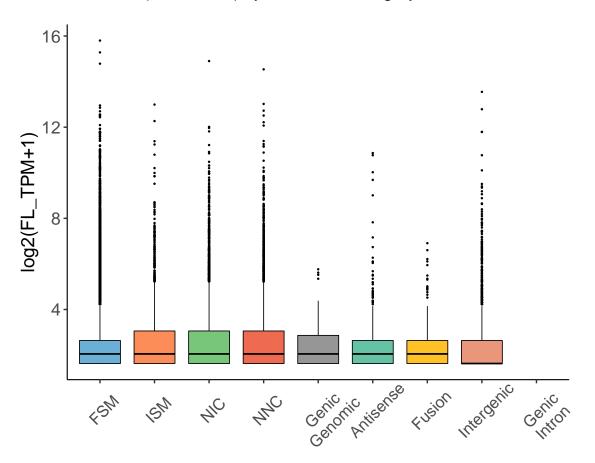
# Transcript Lengths by Structural Classification



### **Exon Counts by Structural Classification**

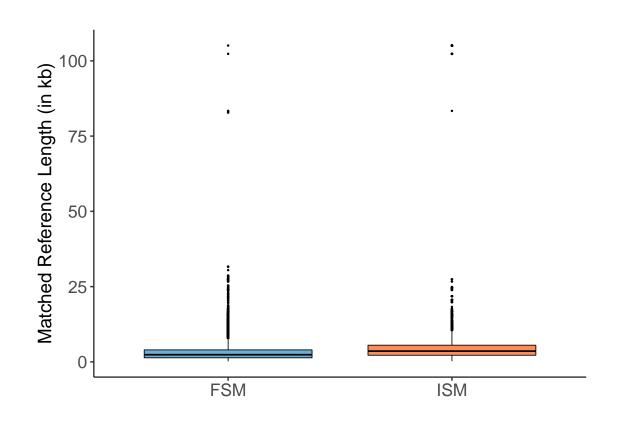


FL Count (normalized) by Structural Category



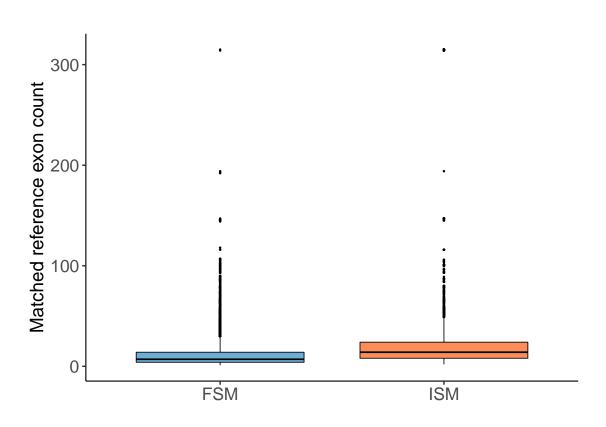
### Length Distribution of Matched Reference Transcripts

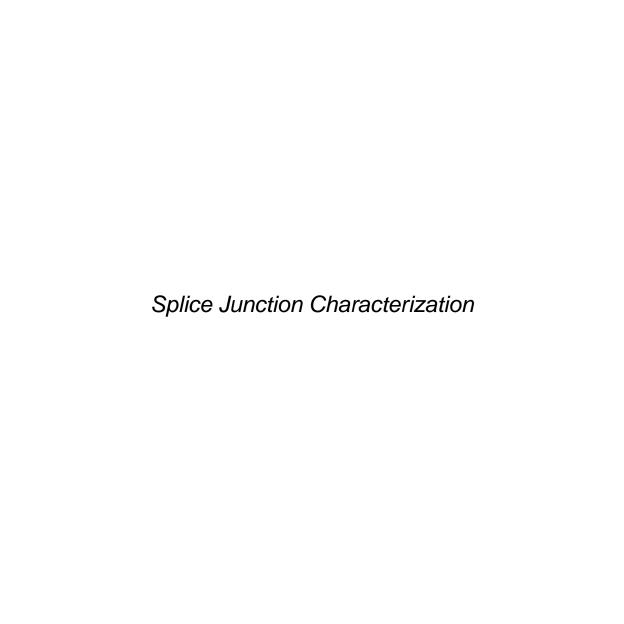
Applicable only to FSM and ISM categories



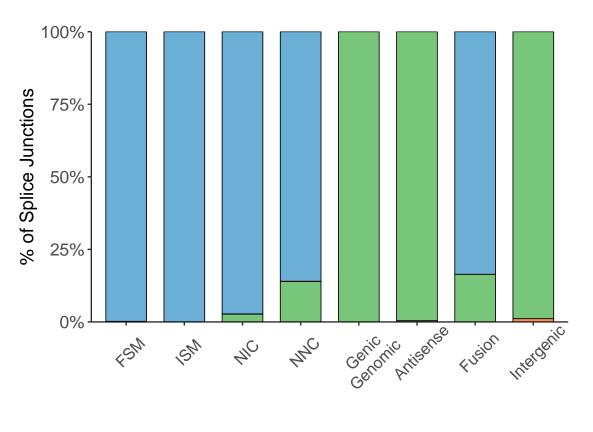
### **Exon Count Distribution of Matched Reference Transcripts**

Applicable only to FSM and ISM categories





#### Distribution of Splice Junctions by Structural Classification



Known

canonical

Known

Non-canonical

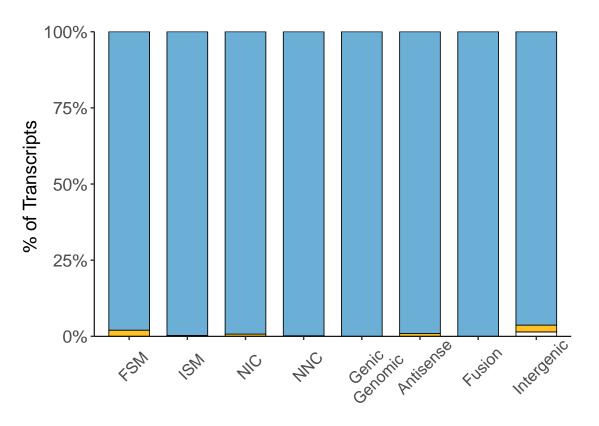
Novel

lcanonical

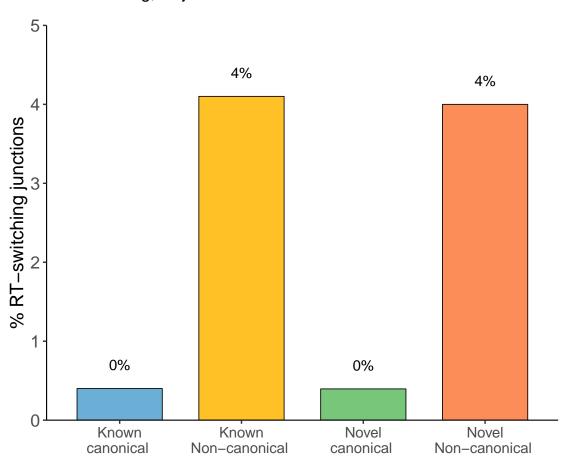
Novel

Non-canonical

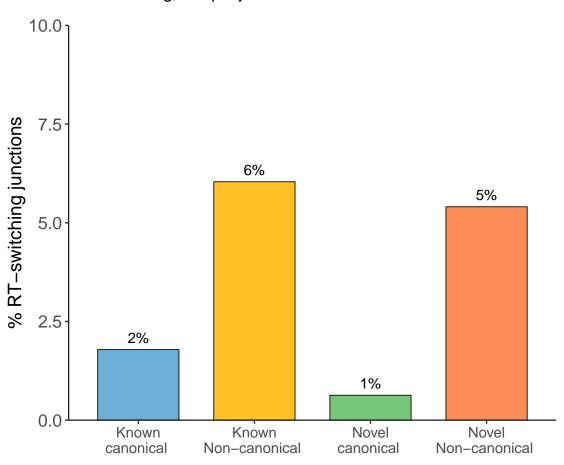
#### Distribution of Transcripts by Splice Junctions

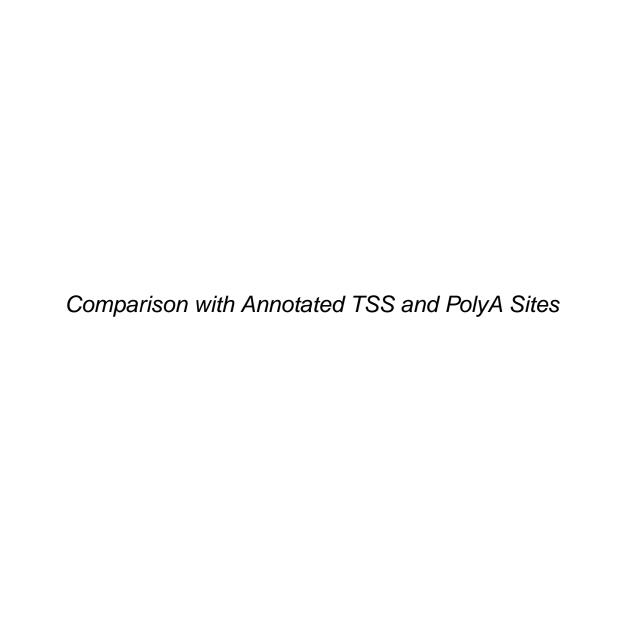


# RT-switching, all junctions



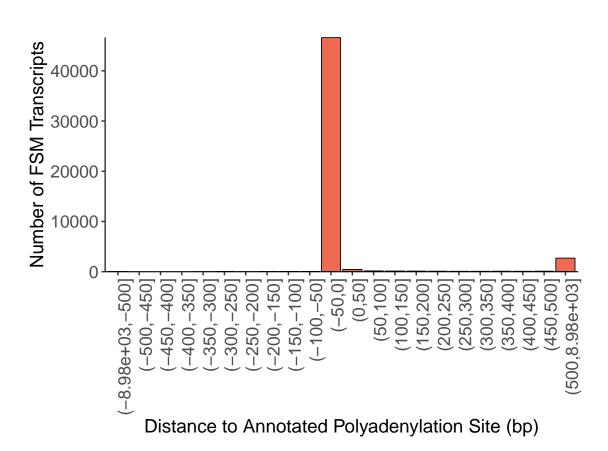
# RT-switching, unique junctions





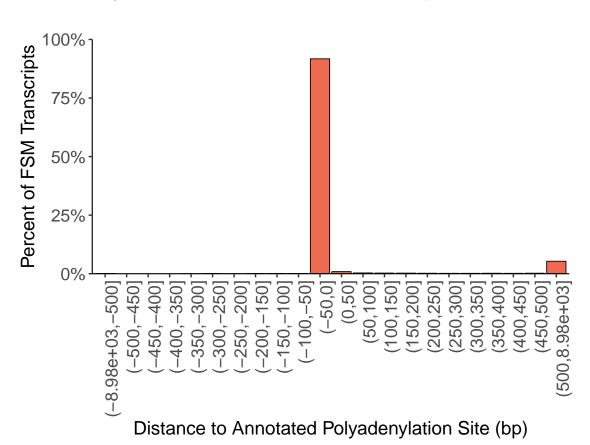
#### Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



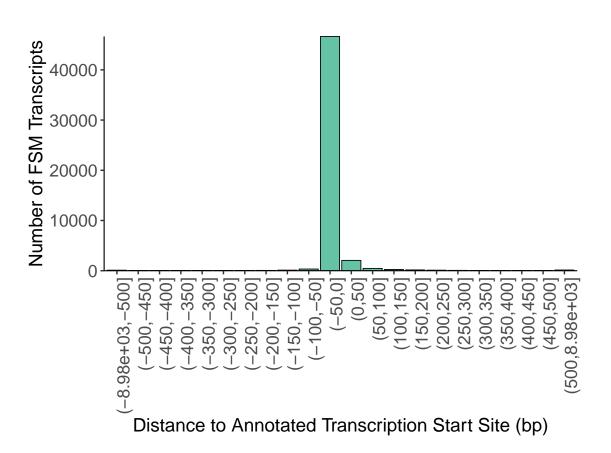
#### Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



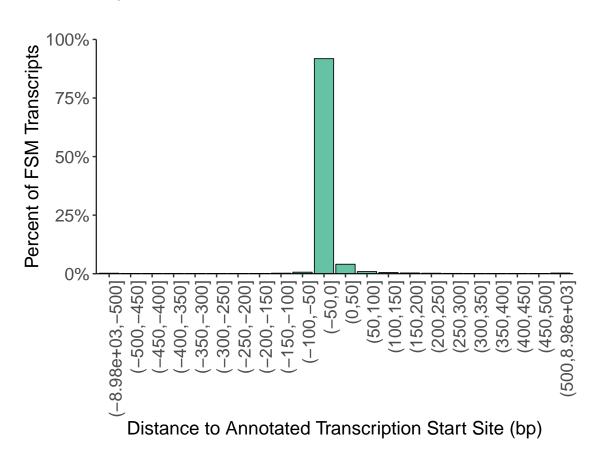
#### Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



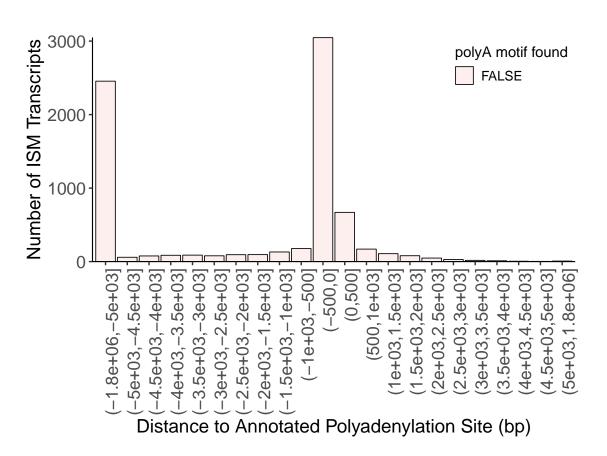
#### Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



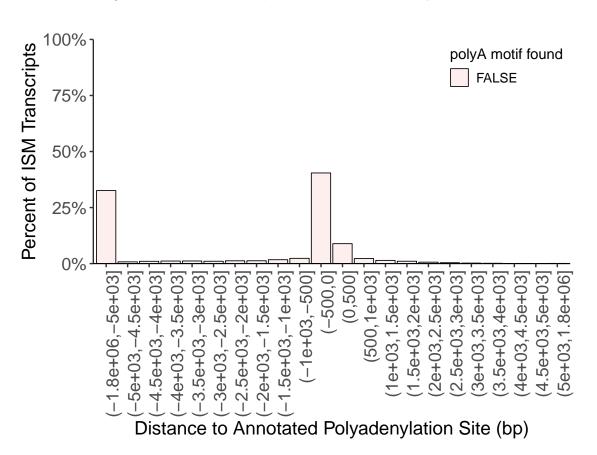
#### Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



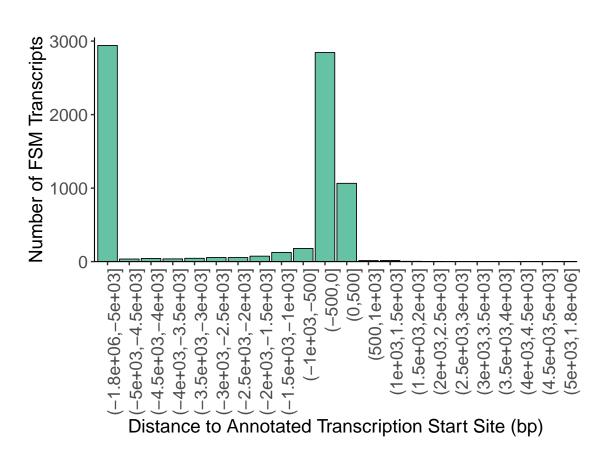
#### Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



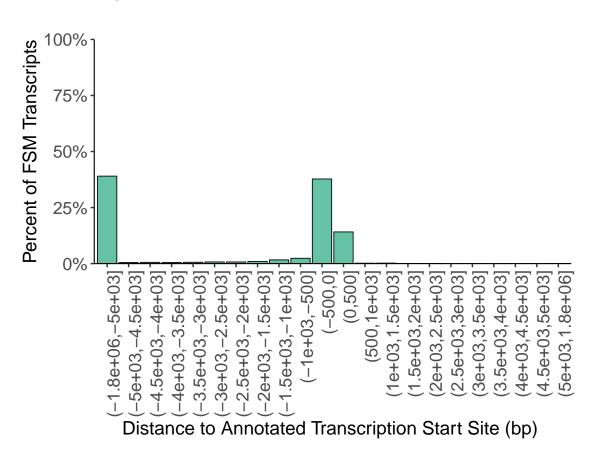
#### Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS

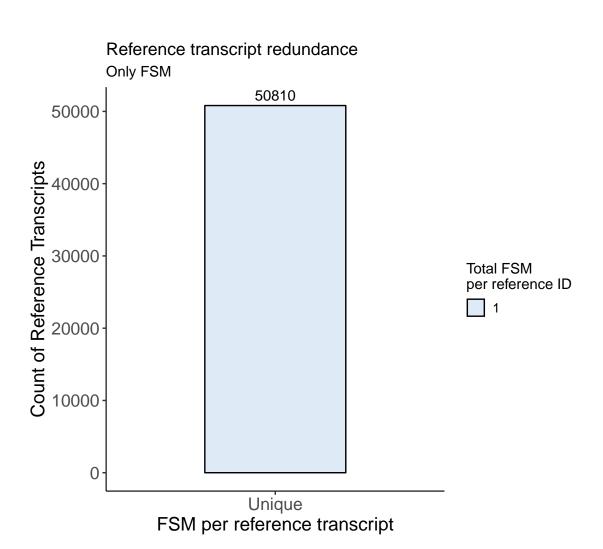


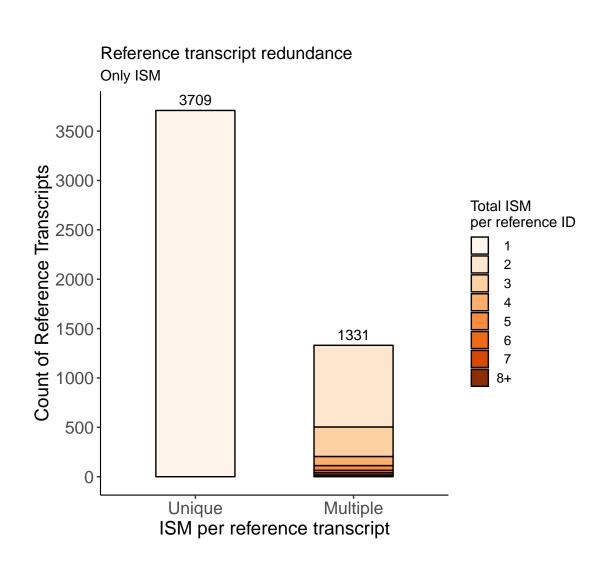
#### Distance to Annotated Transcription Start Site, ISM only

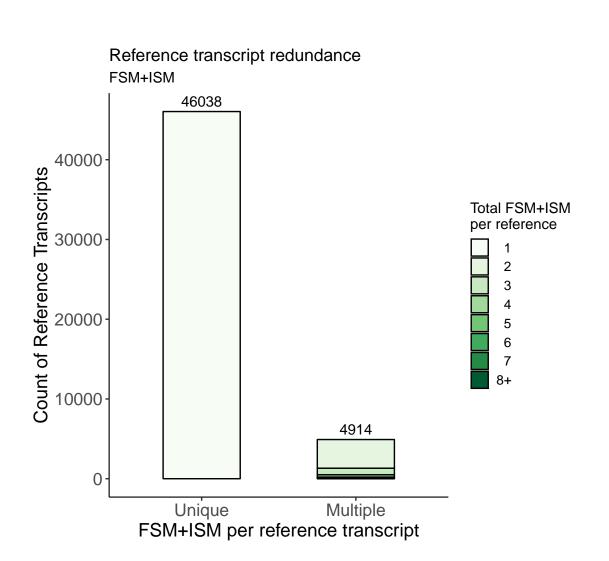
Negative values indicate downstream of annotated TSS



Accumulation of FSM and ISM to the same reference transcript



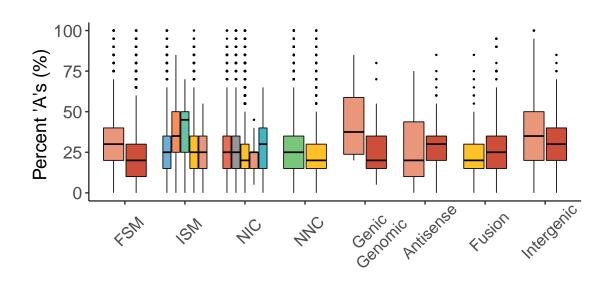


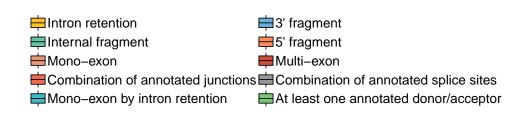




#### Possible Intra-Priming by Structural Category

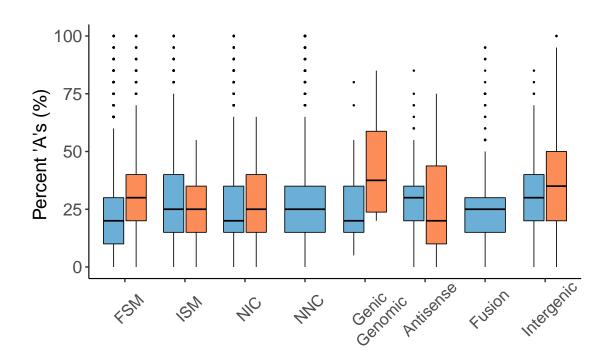
Percent of genomic 'A's in downstream 20 bp





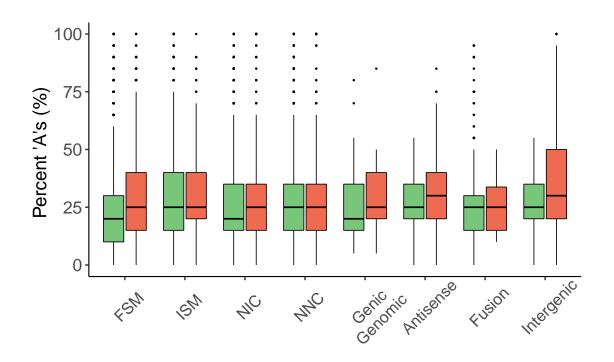
### Possible Intra-Priming, Mono- vs Multi-Exon

Percent of genomic 'A's in downstream 20 bp



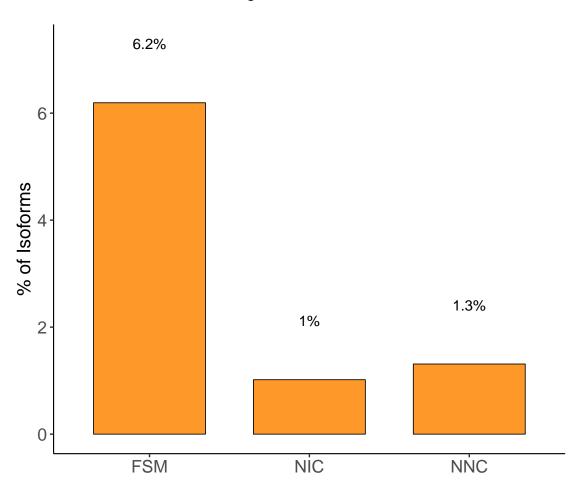
#### Possible Intra-Priming, Coding vs Non-Coding

Percent of genomic 'A's in downstream 20 bp

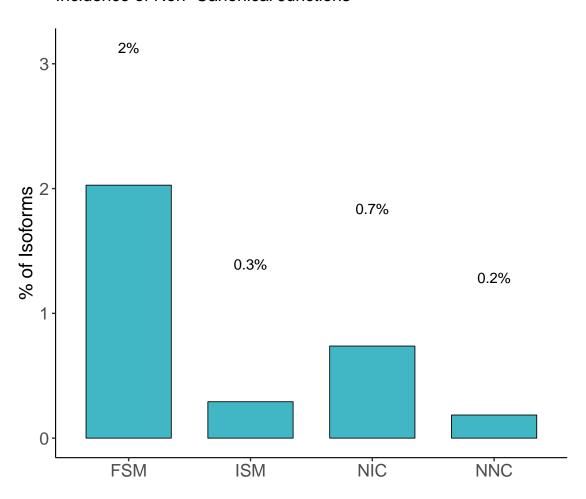




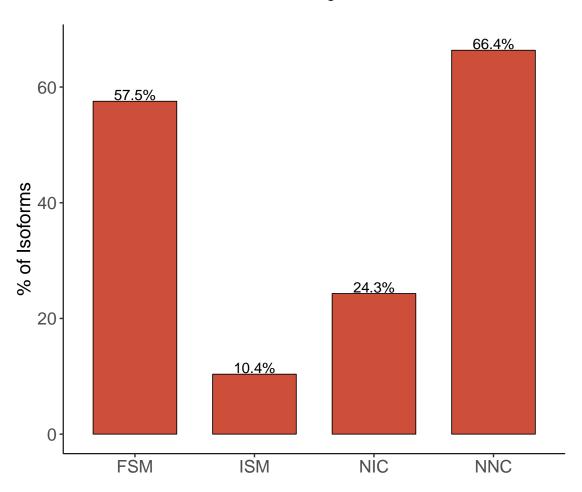
# Incidence of RT-switching



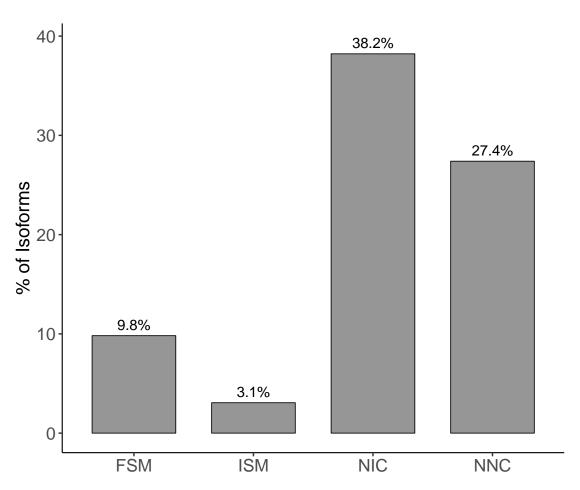
#### Incidence of Non-Canonical Junctions



## Incidence of SJ without SR coverage Junctions



## Incidence of NMD by structural category



#### Quality control attributes across structural categories

