

SQANTI3 report

Unique Genes: 5286
Unique Isoforms: 11821

Gene classification

Category	# Genes
Annotated Genes	5280
Novel Genes	6

Splice Junction Classification

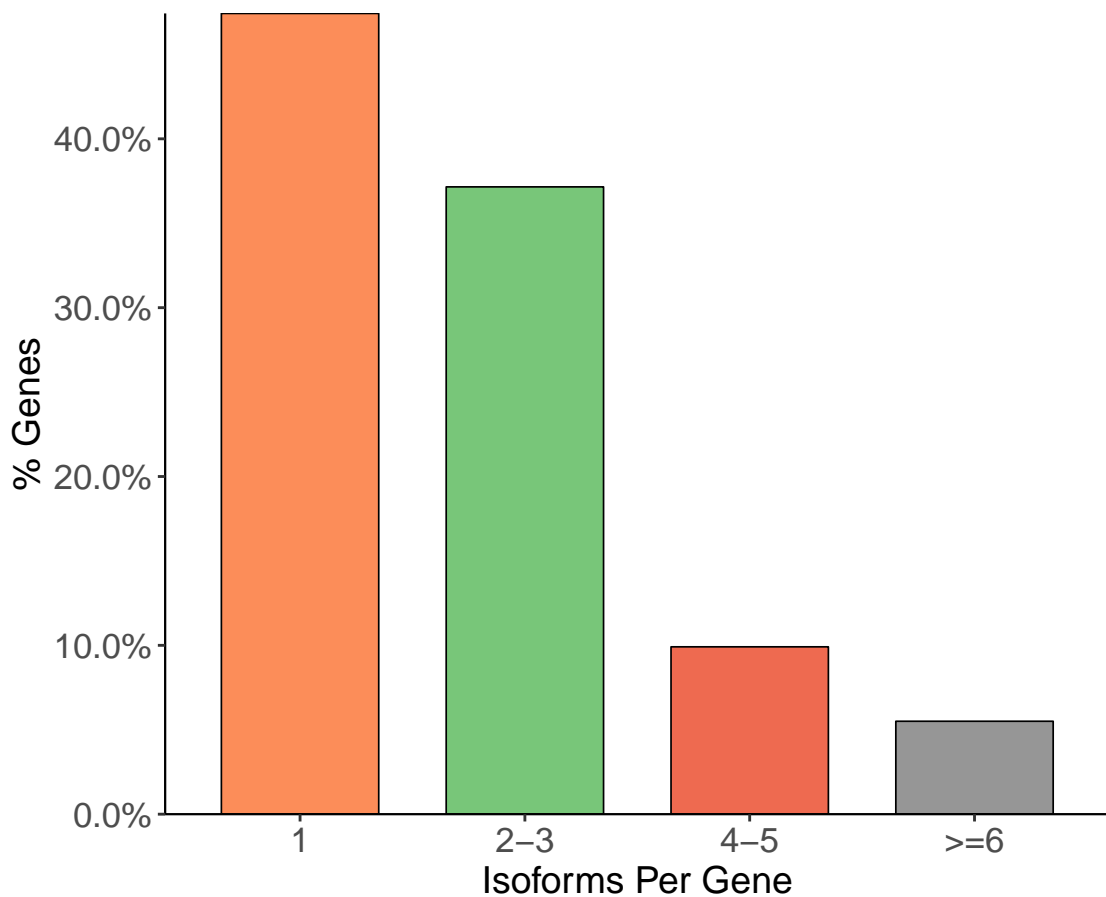
Category	# SJs	Percent
Known canonical	47774	98.68
Known Non-canonical	21	0.04
Novel canonical	617	1.27
Novel Non-canonical	0	0.00

*Characterization of transcripts
based on splice junctions*

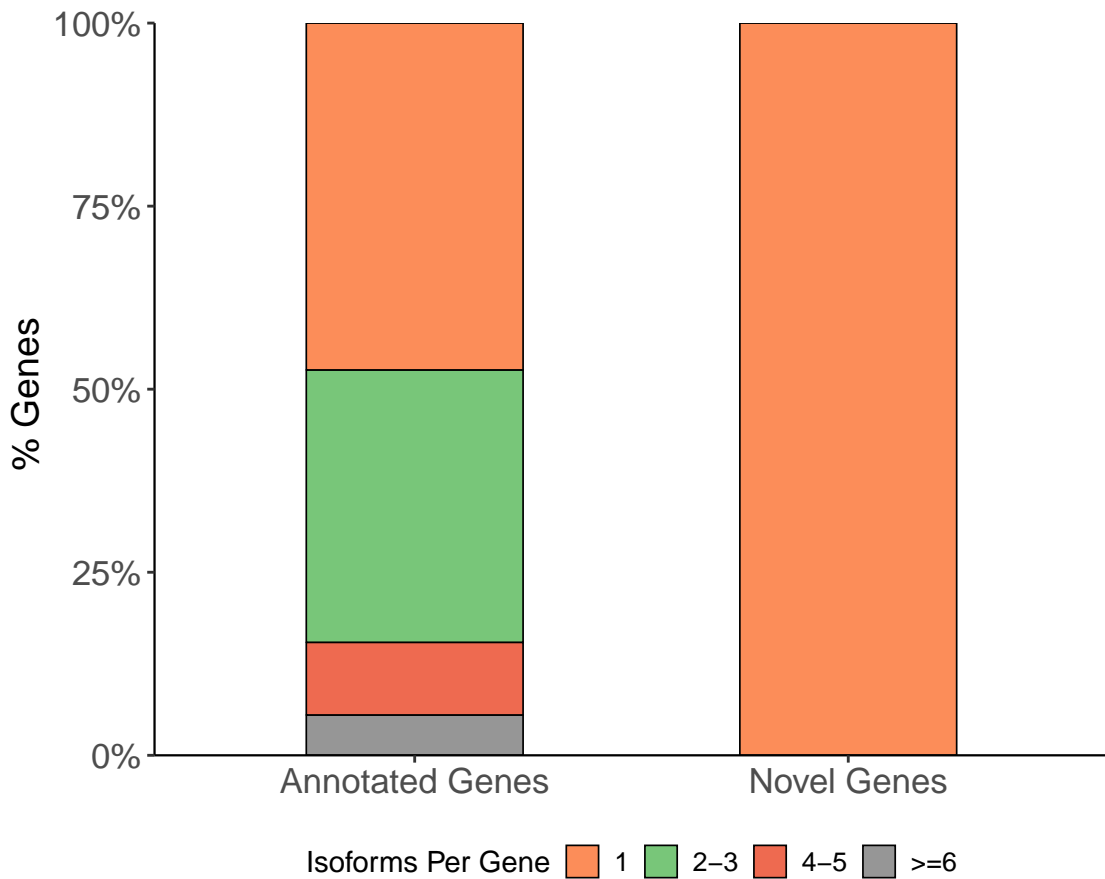
Category	# Isoforms
FSM	8268
ISM	2082
NIC	1071
NNC	381
Genic Genomic	5
Antisense	1
Fusion	8
Intergenic	5
Genic Intron	0

Gene Characterization

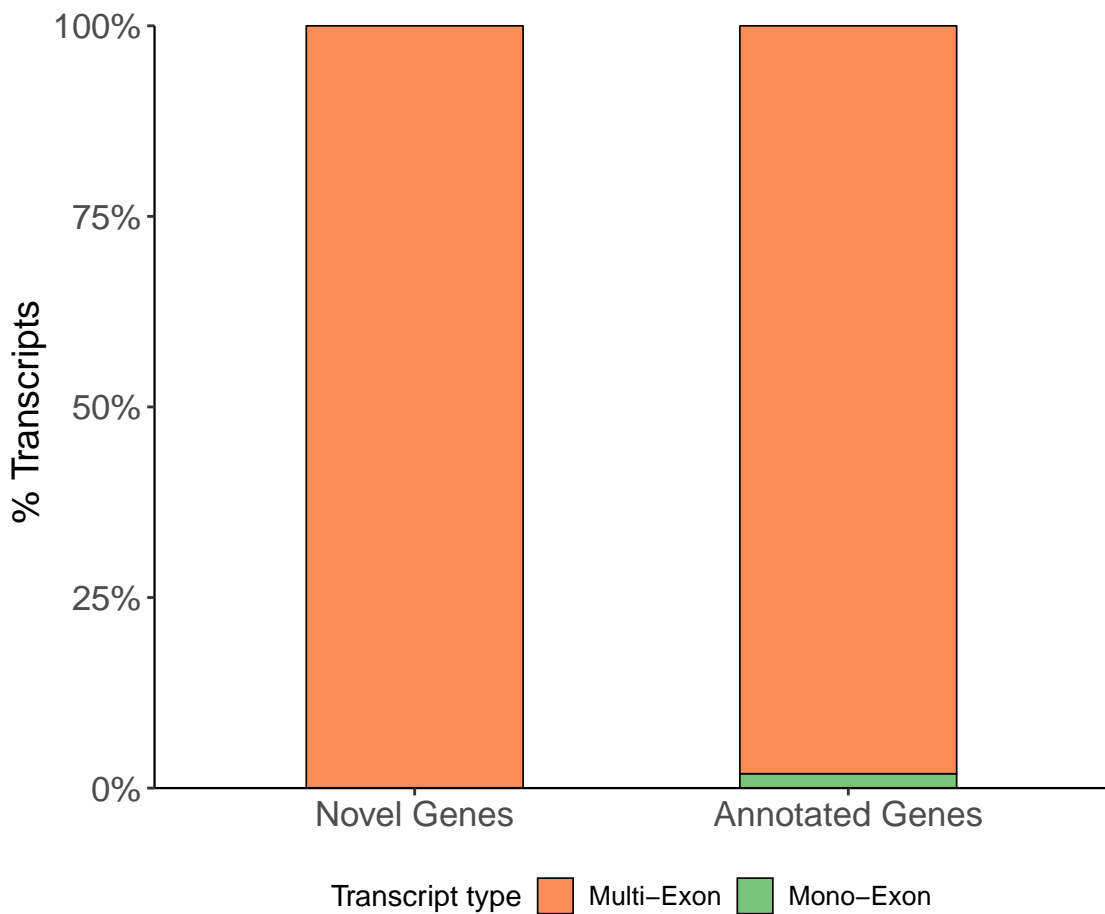
Number of Isoforms per Gene

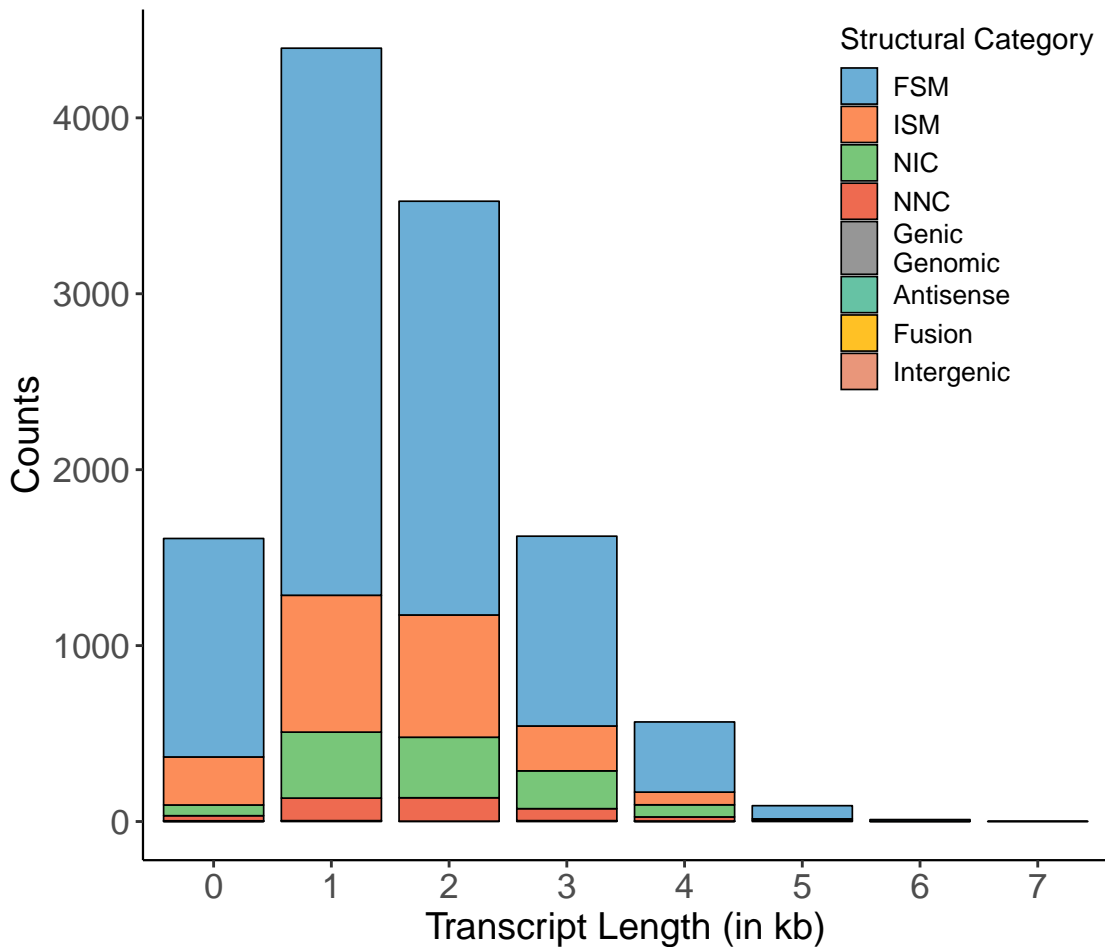


Number of Isoforms per Gene, Known vs Novel Genes

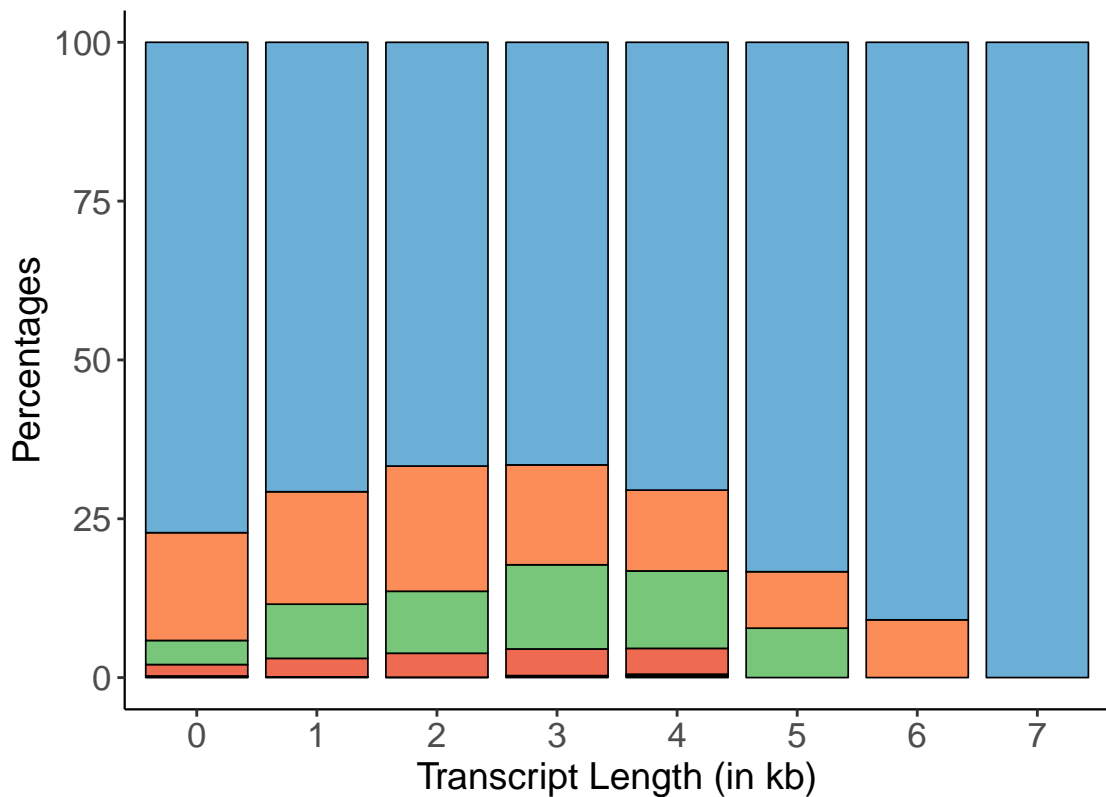


Distribution of Mono- vs Multi-Exon Transcripts



[illegible]

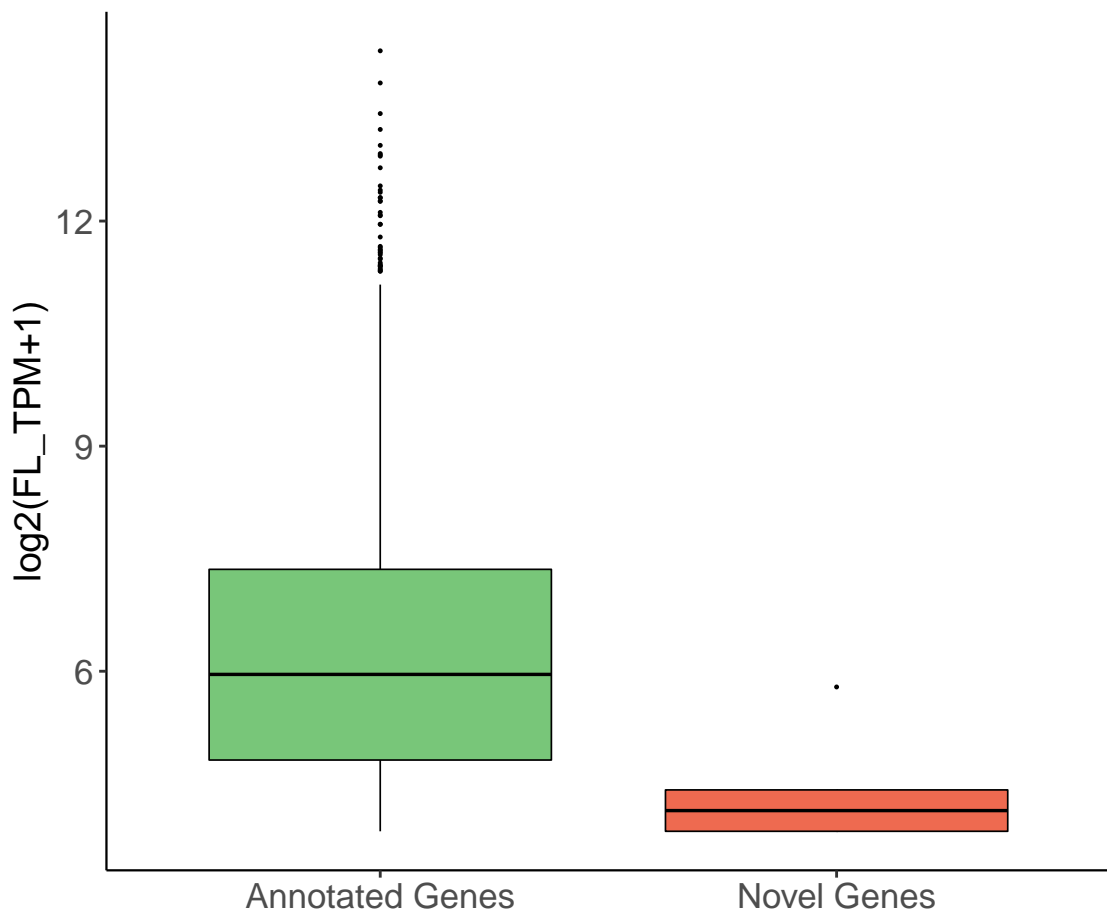
Classifications by Transcript Length, normalized



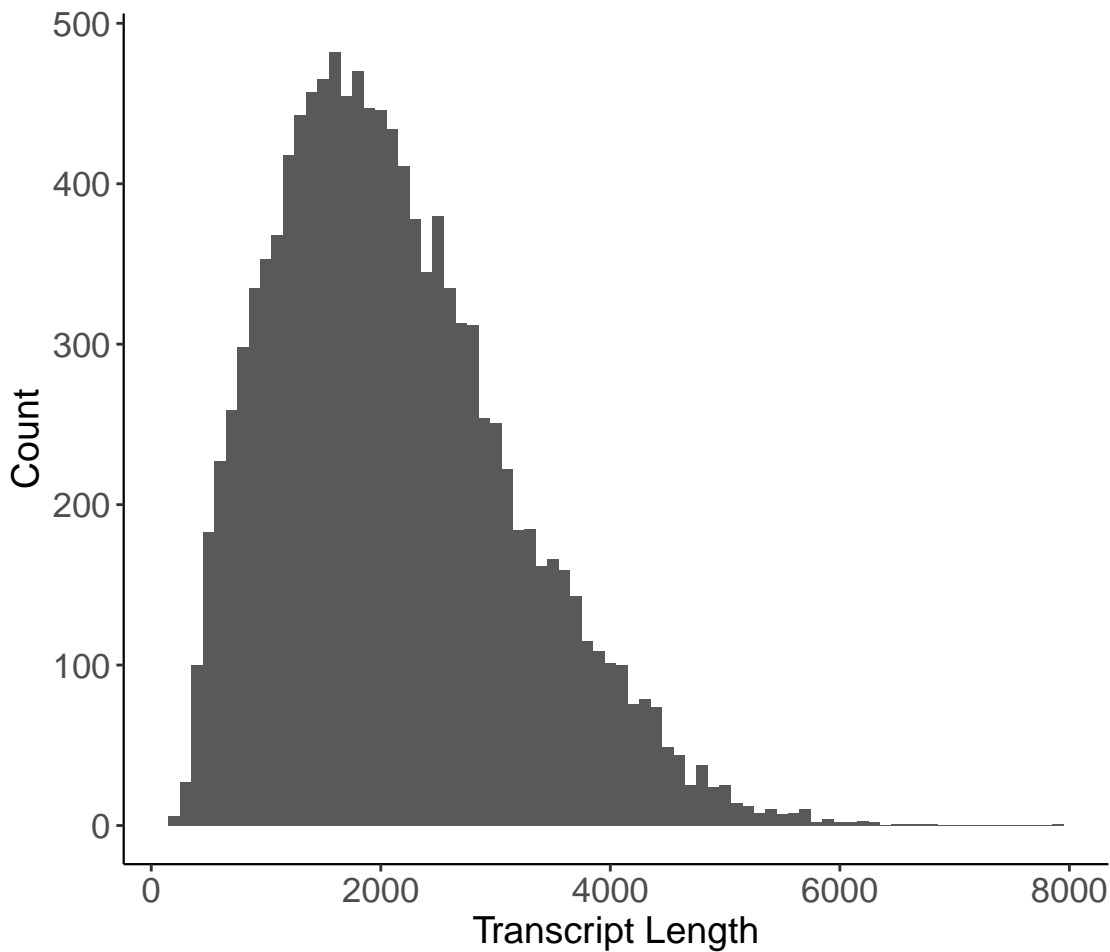
Structural Category

■ FSM	■ NIC	■ Genic	■ Fusion
■ ISM	■ NNC	■ Antisense	■ Intergenic

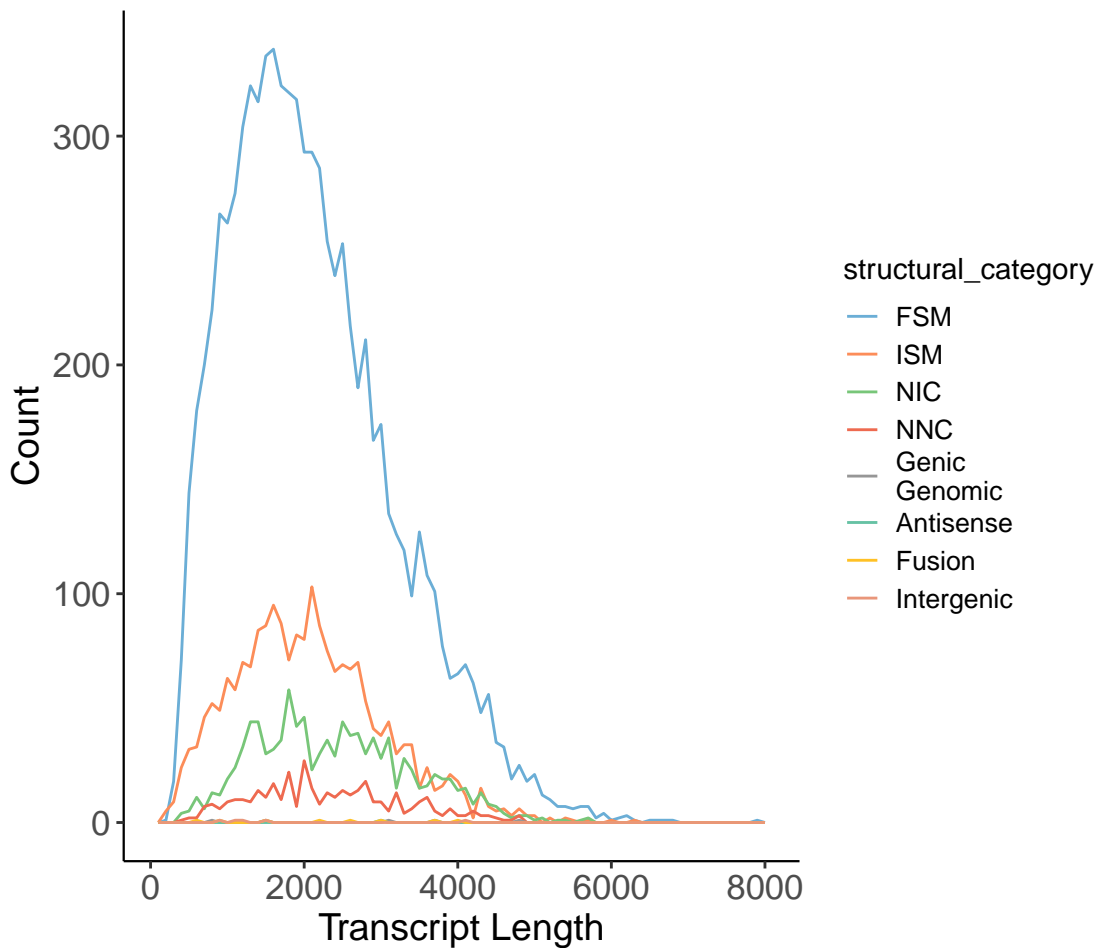
Number of FL reads per Gene by type of gene annotation



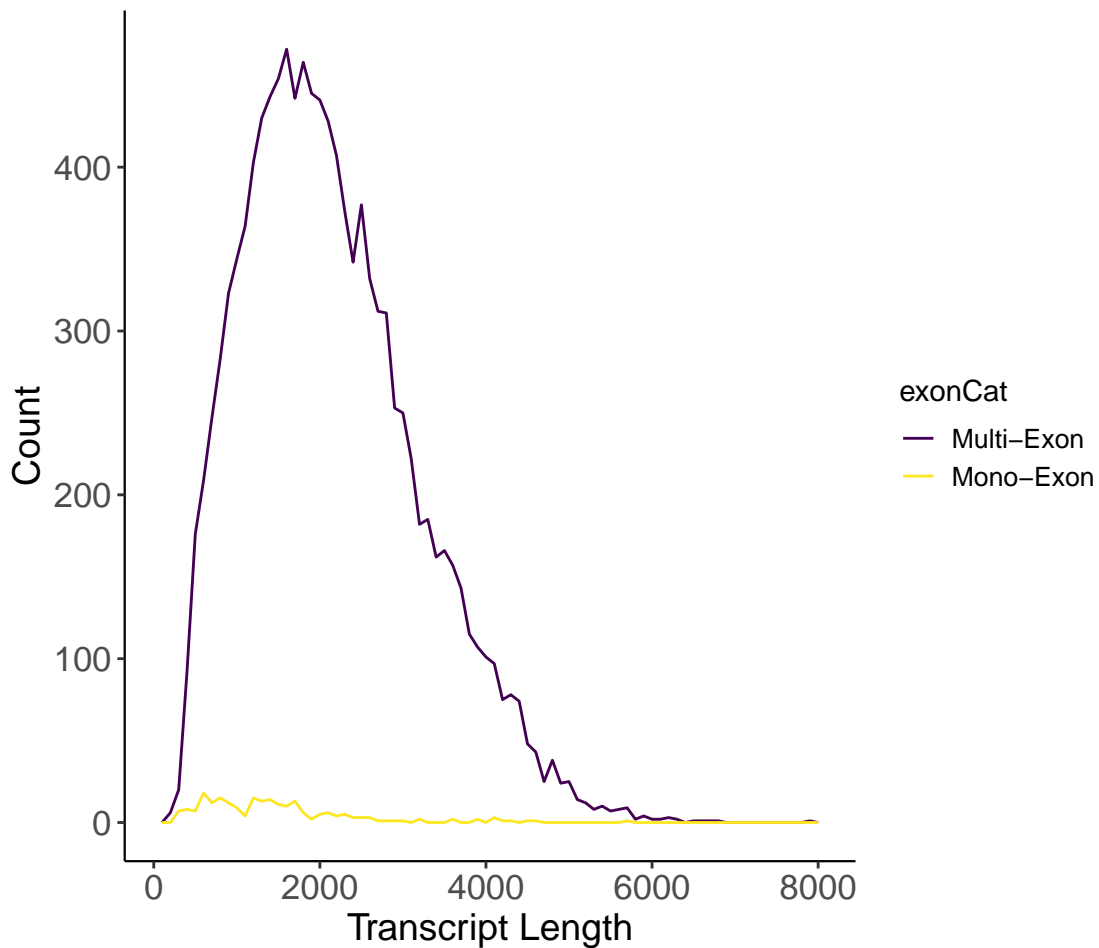
Transcript Lengths, all transcripts



Transcript Lengths, by structural category

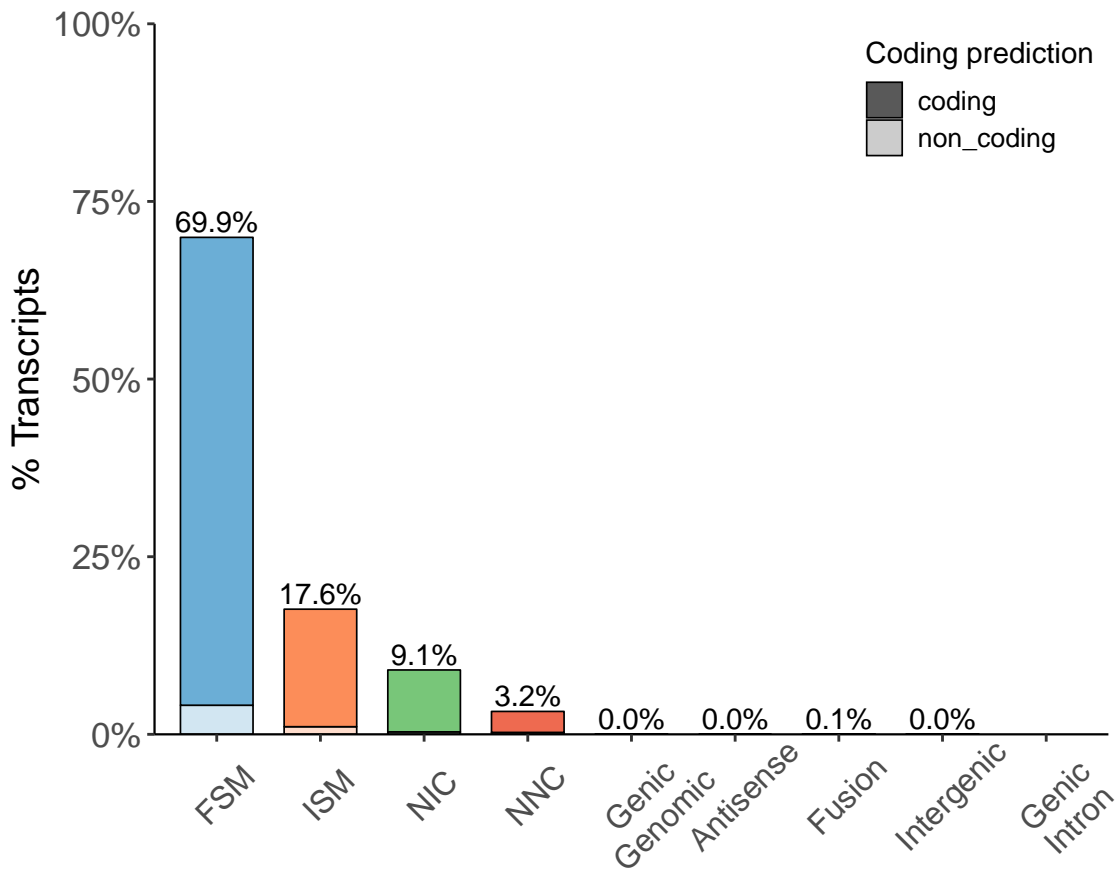


Transcript Lengths, Mono- vs Multi-Exons

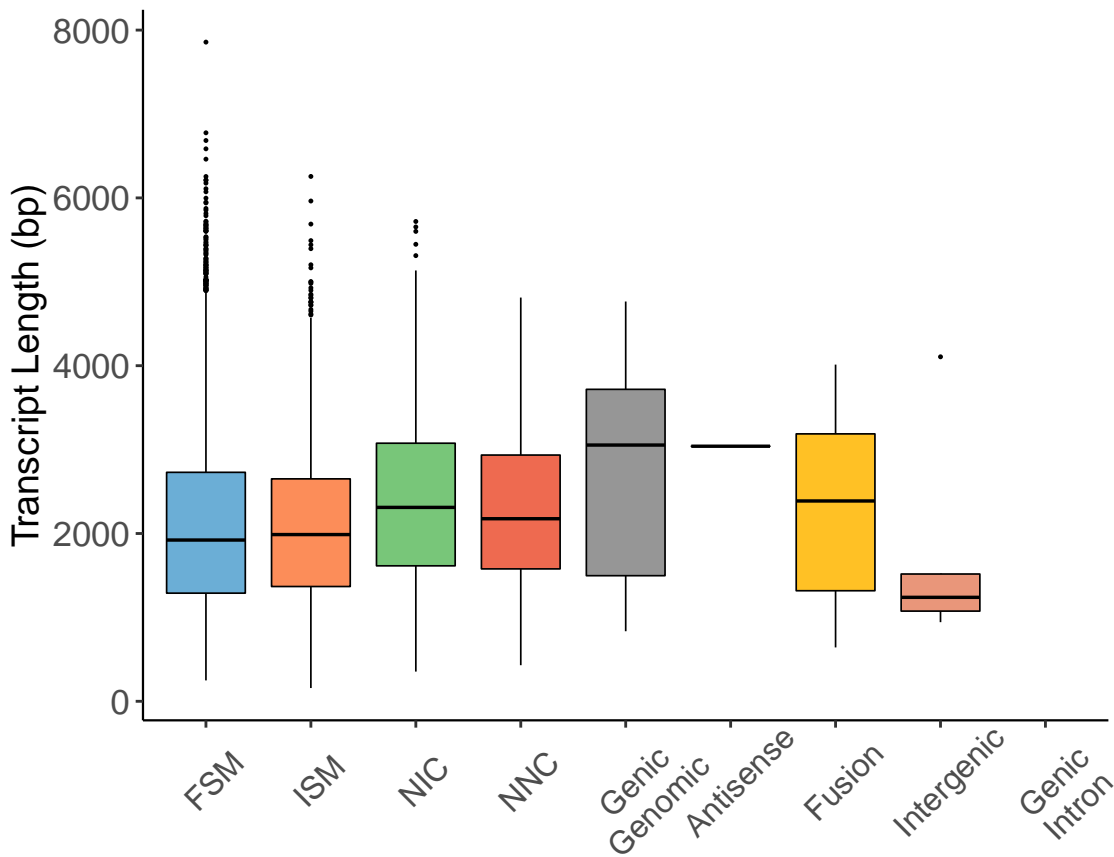


*Structural Isoform Characterization
by Splice Junctions*

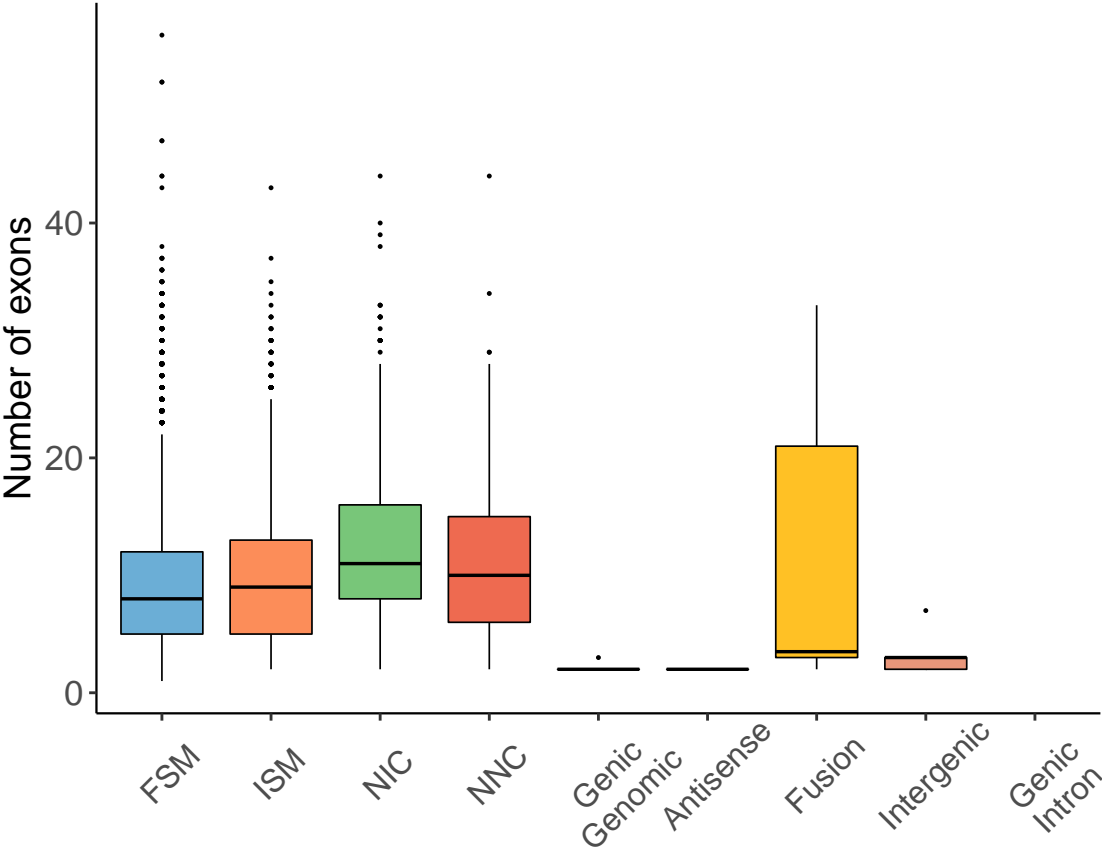
Isoform distribution across structural categories



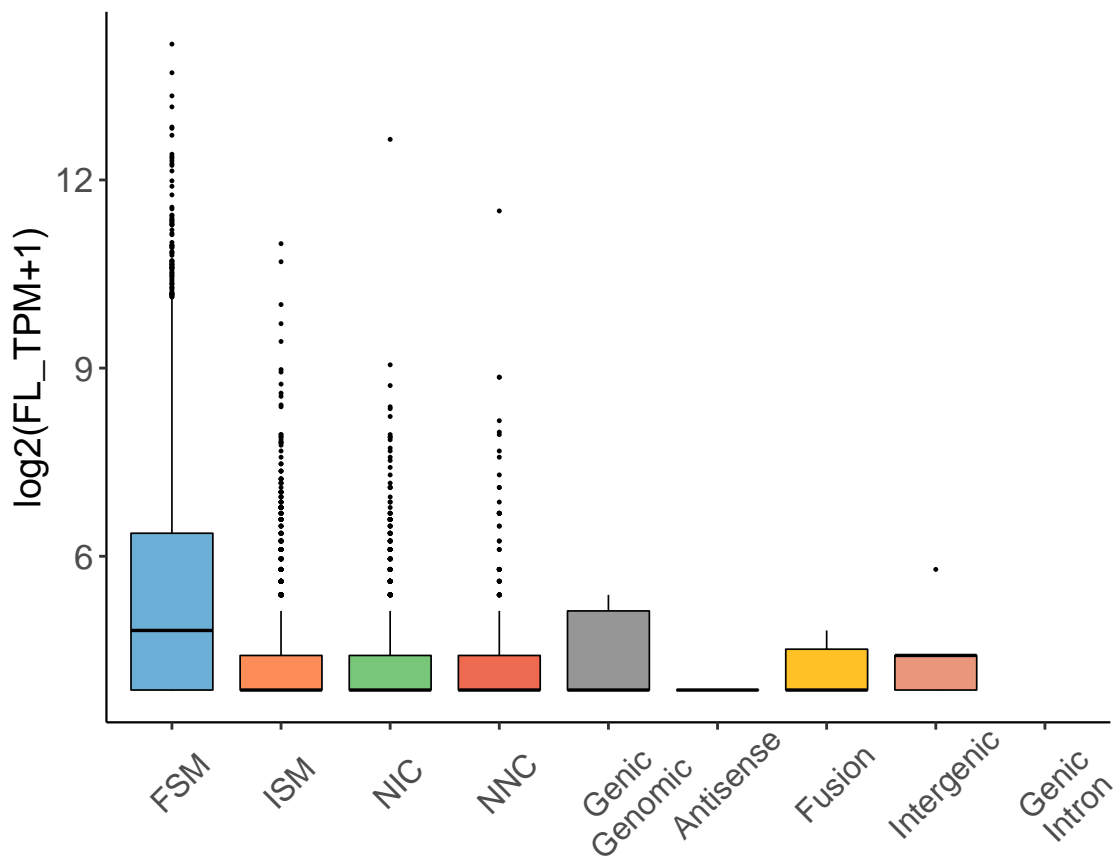
Transcript Lengths by Structural Classification



Exon Counts by Structural Classification

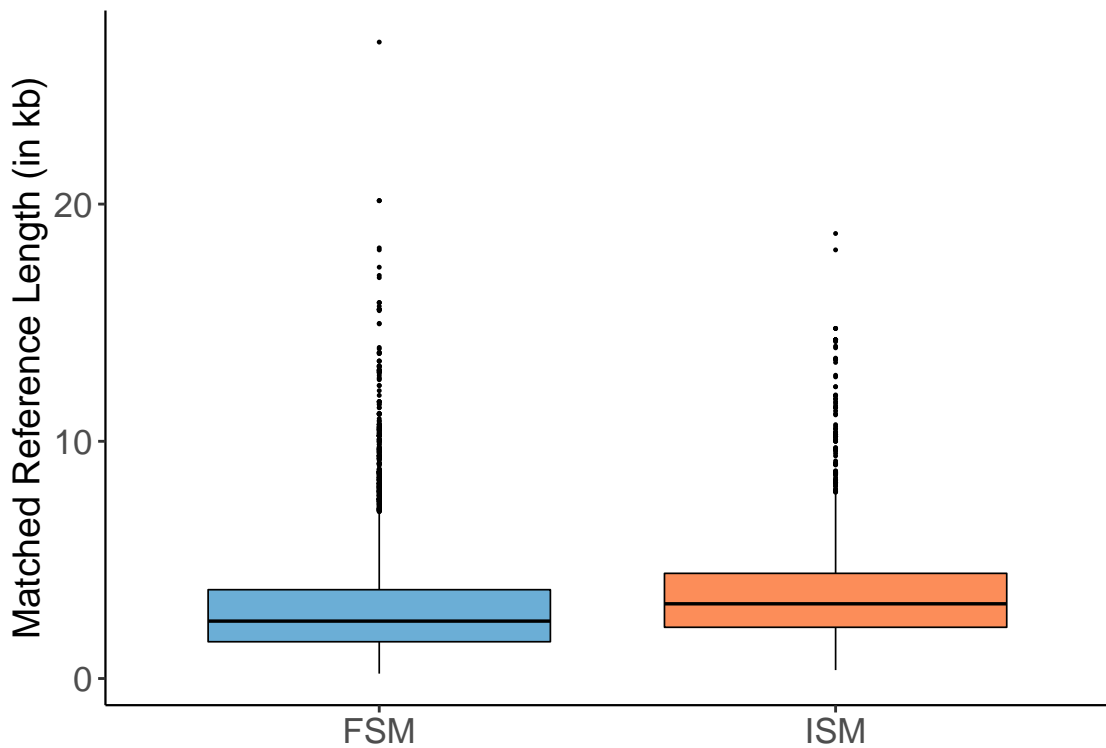


FL Count (normalized) by Structural Category



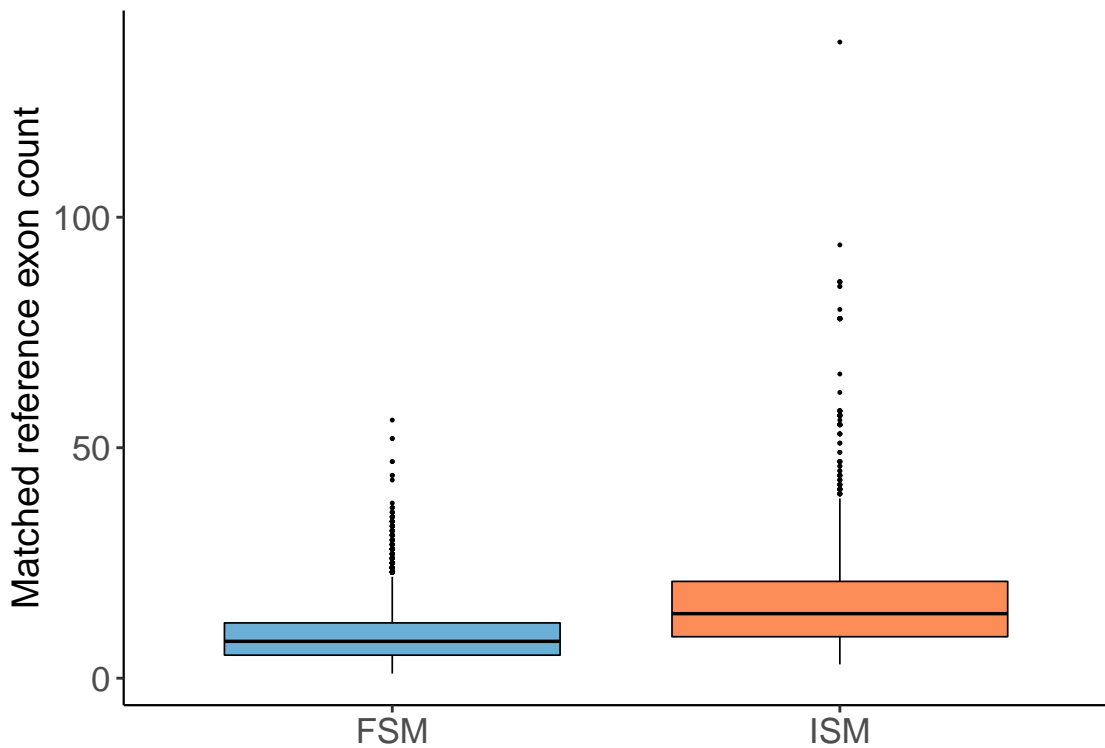
Length Distribution of Matched Reference Transcripts

Applicable only to FSM and ISM categories



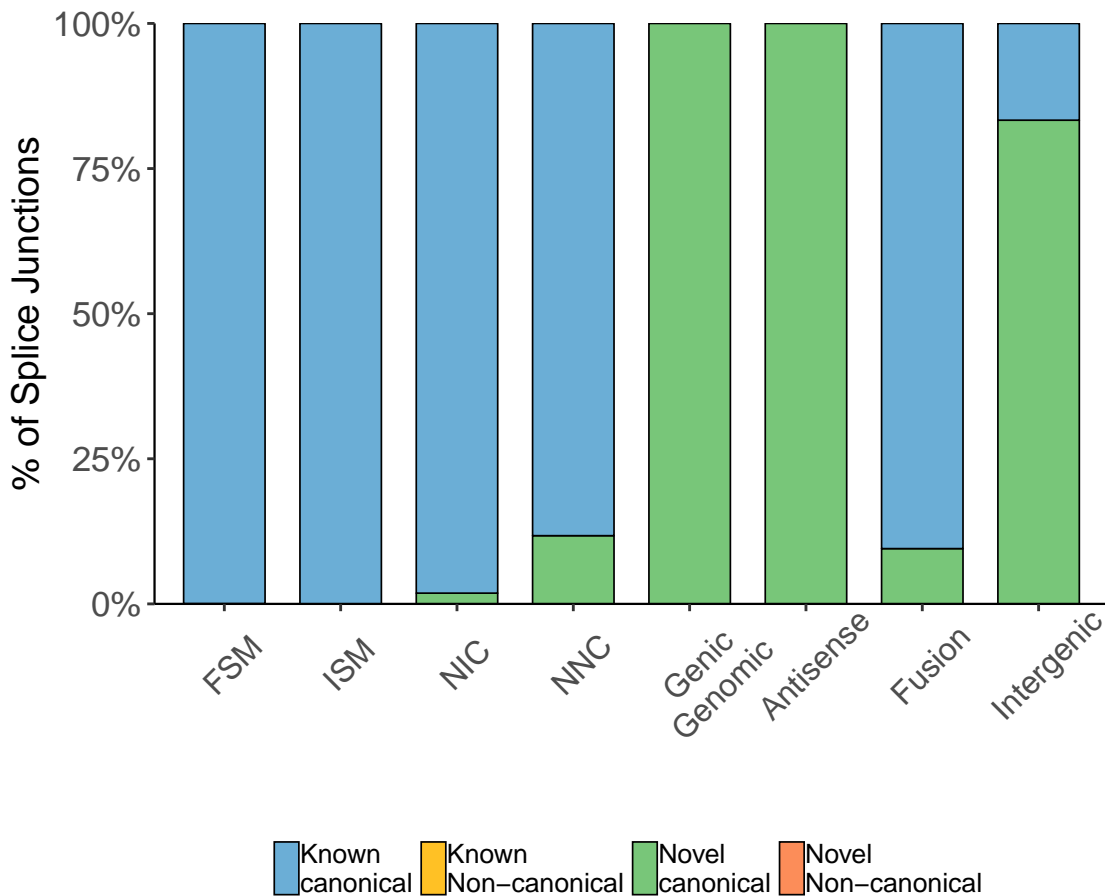
Exon Count Distribution of Matched Reference Transcripts

Applicable only to FSM and ISM categories

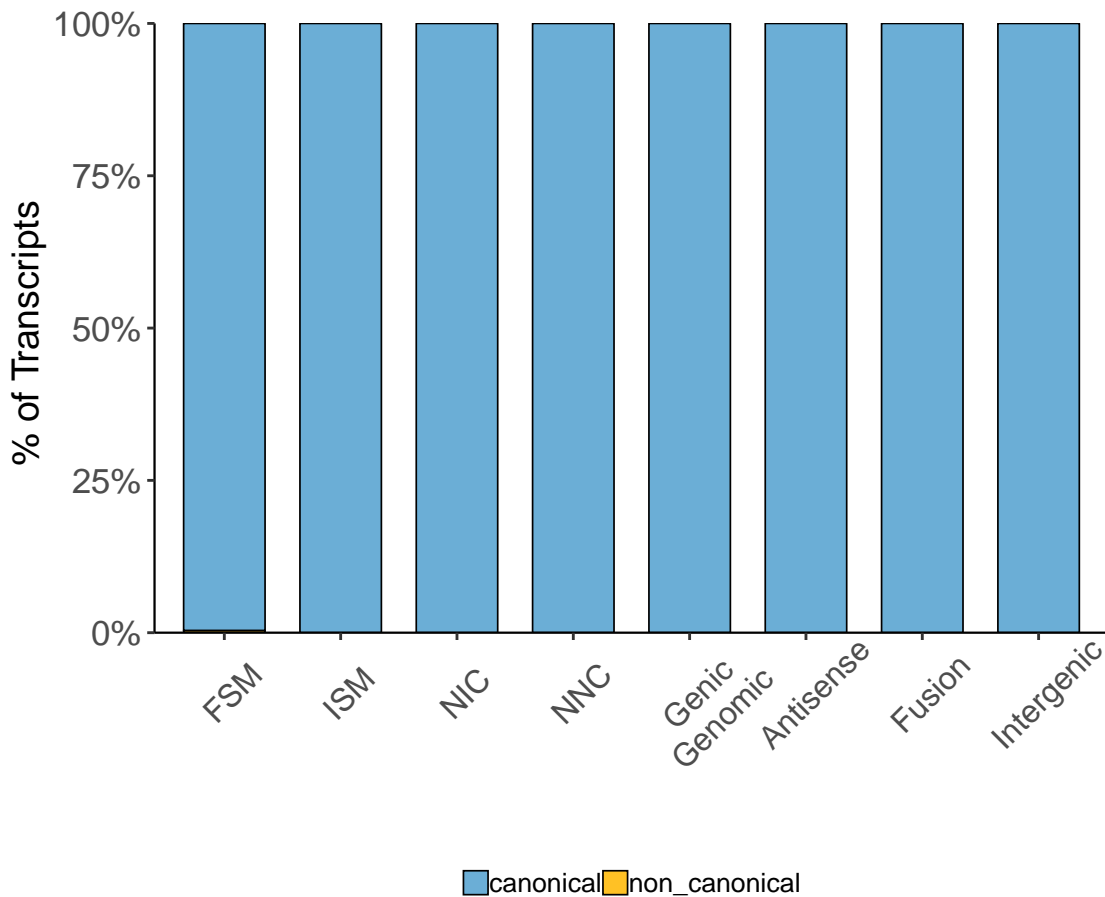


Splice Junction Characterization

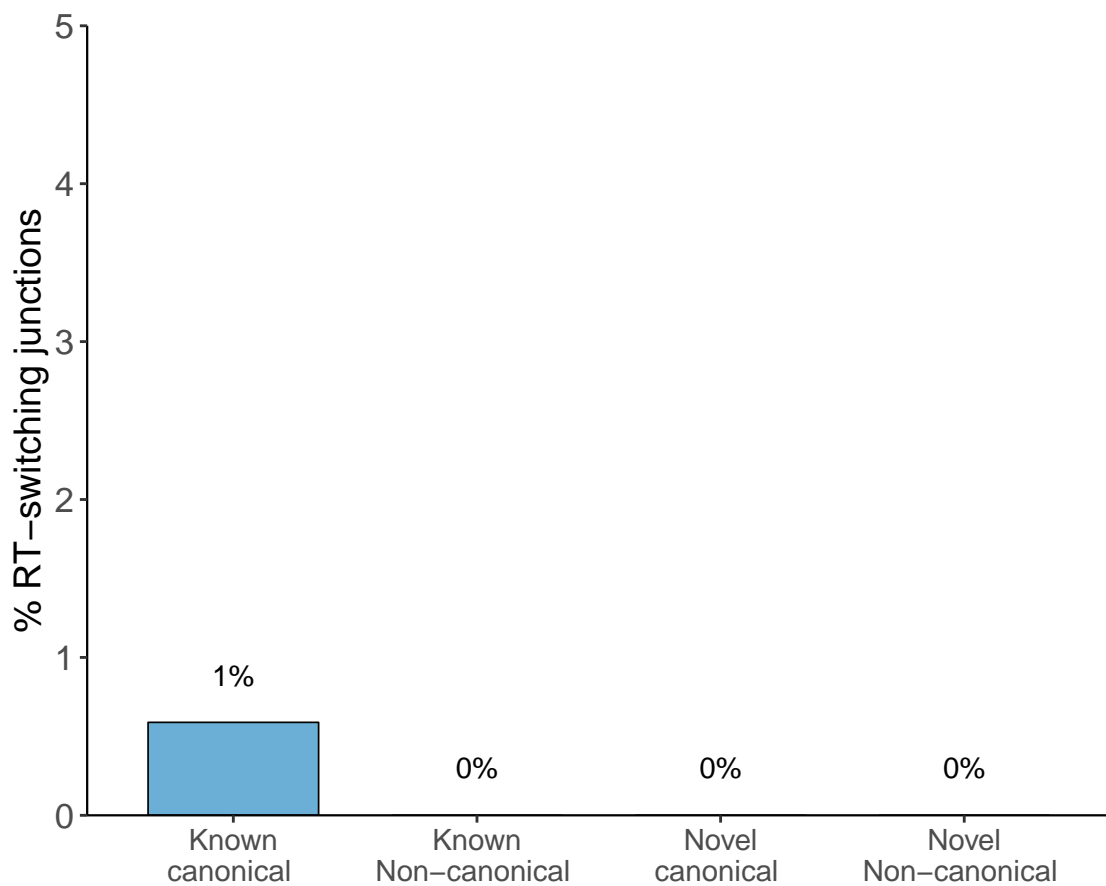
Distribution of Splice Junctions by Structural Classification



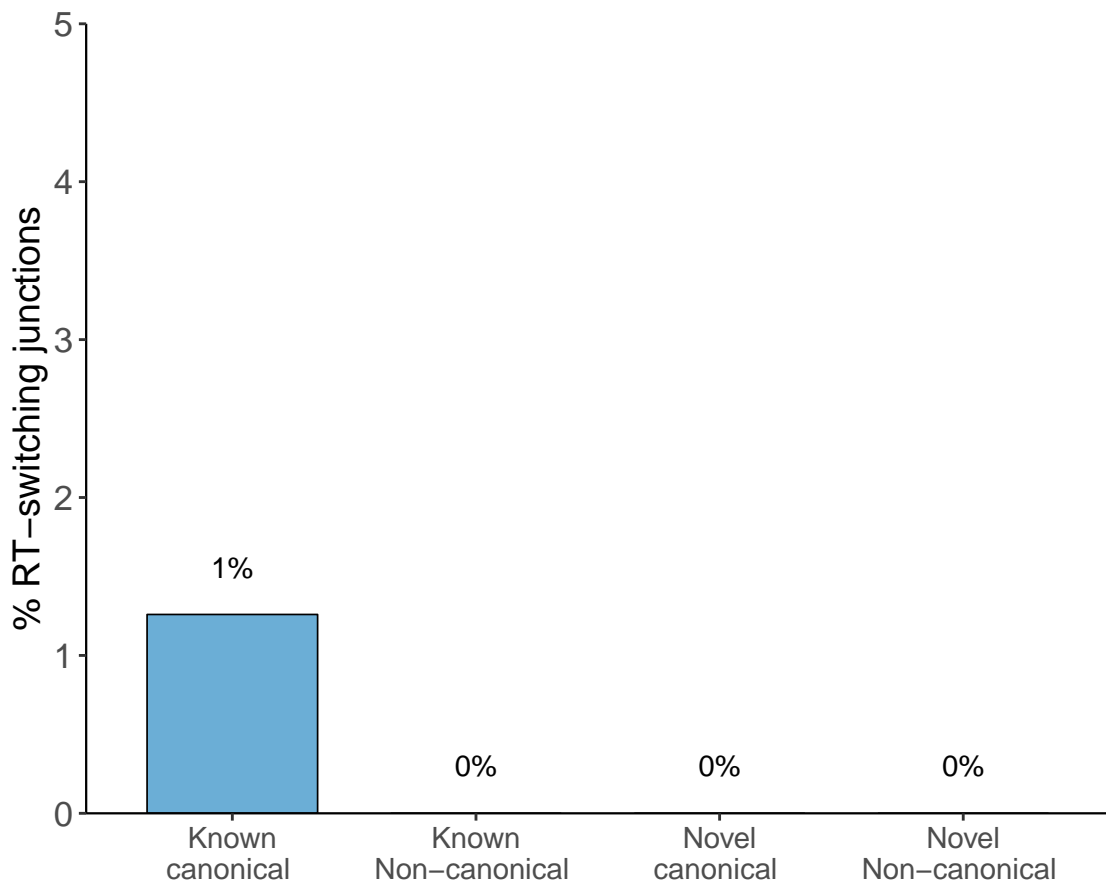
Distribution of Transcripts by Splice Junctions



RT-switching, all junctions



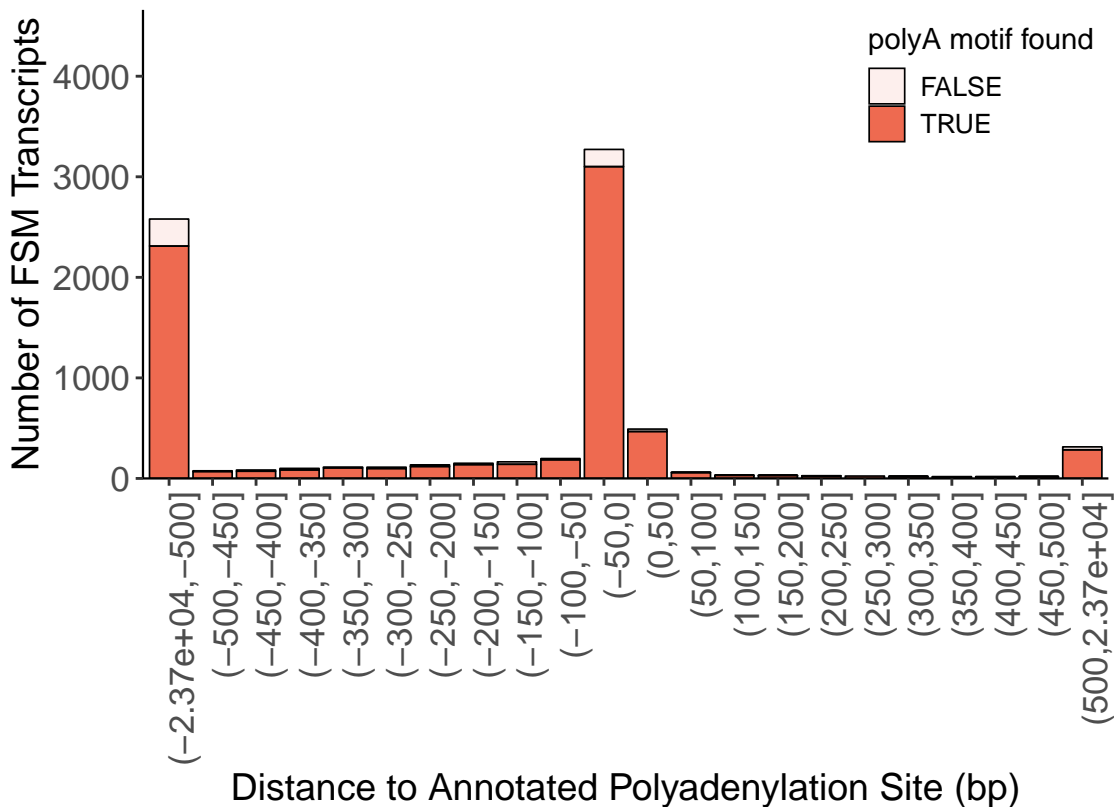
RT-switching, unique junctions



Comparison with Annotated TSS and PolyA Sites

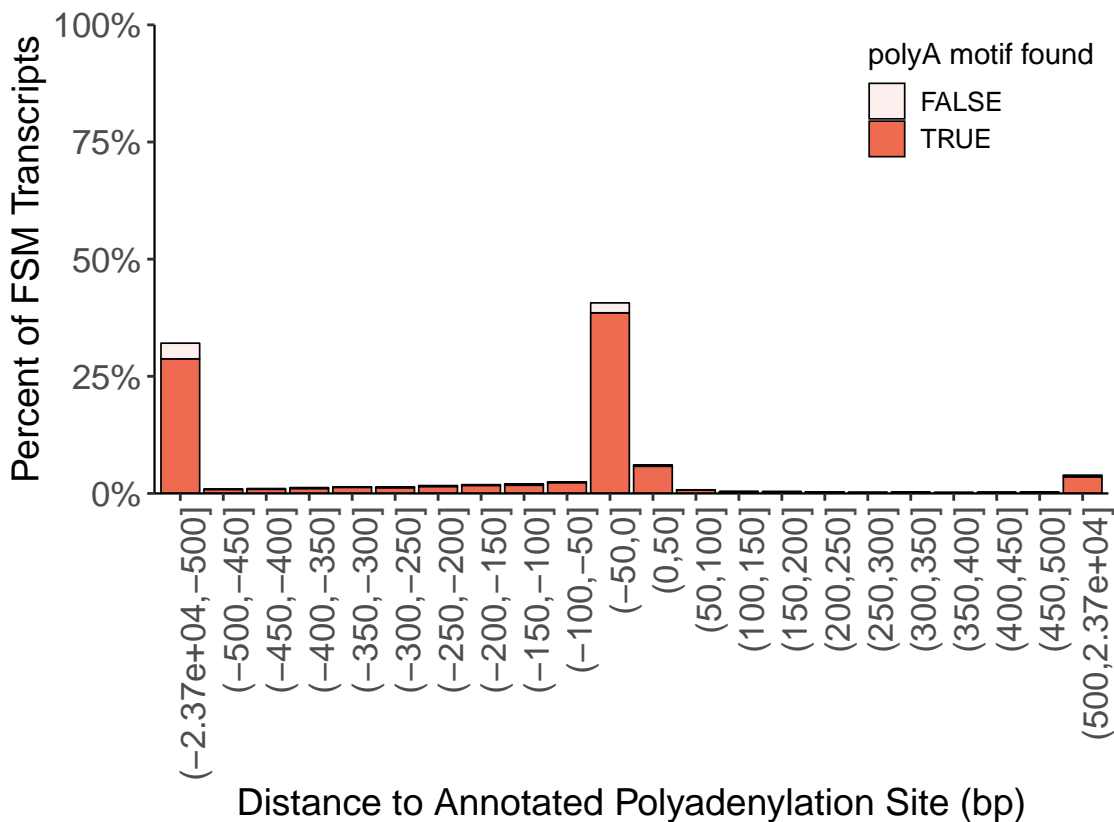
Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



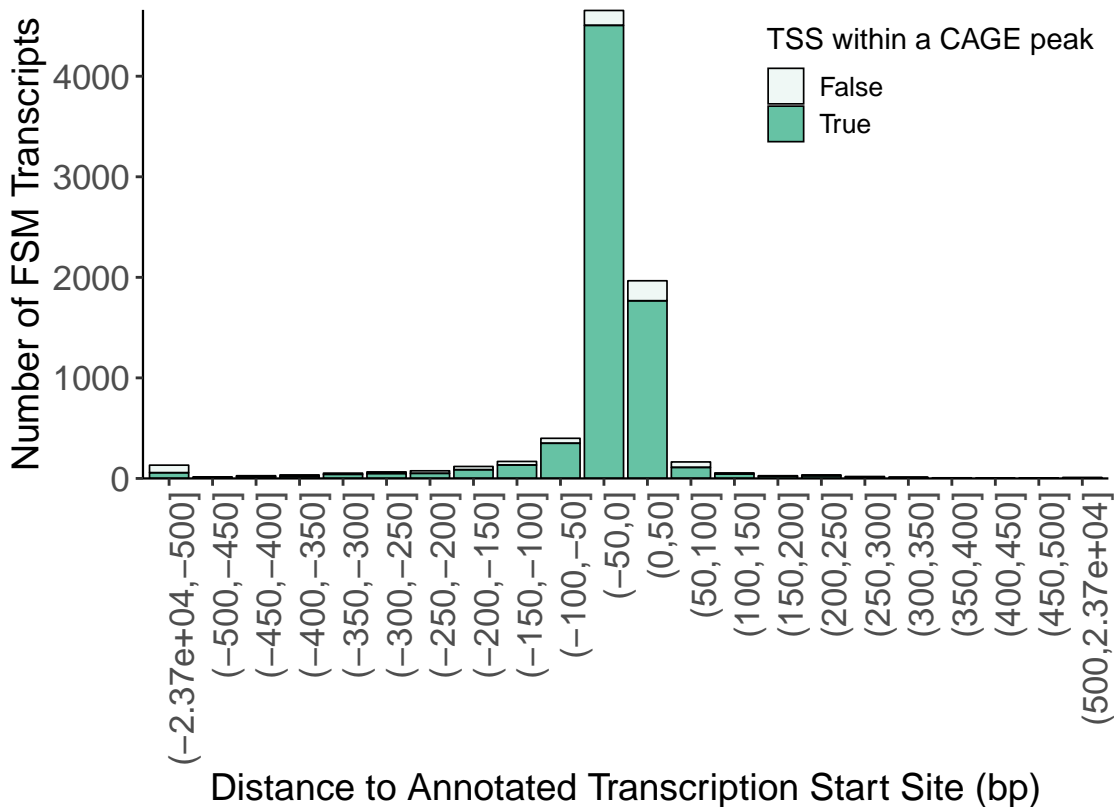
Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



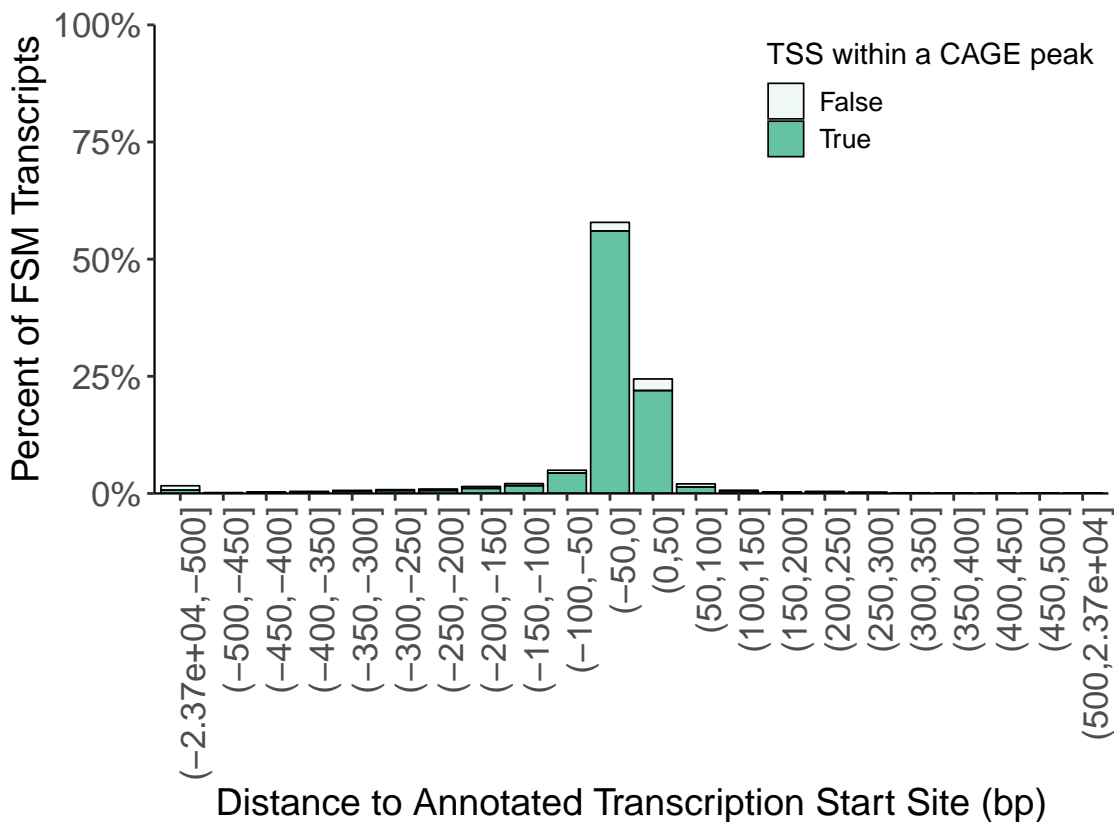
Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



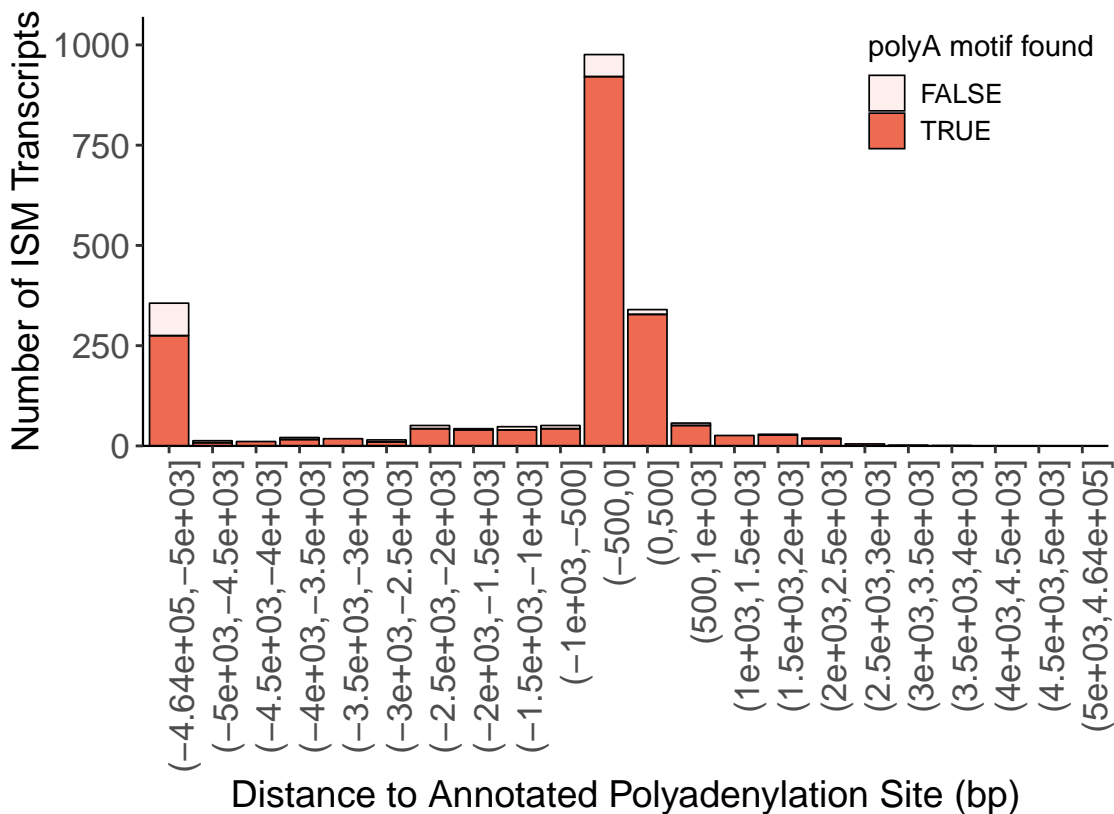
Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



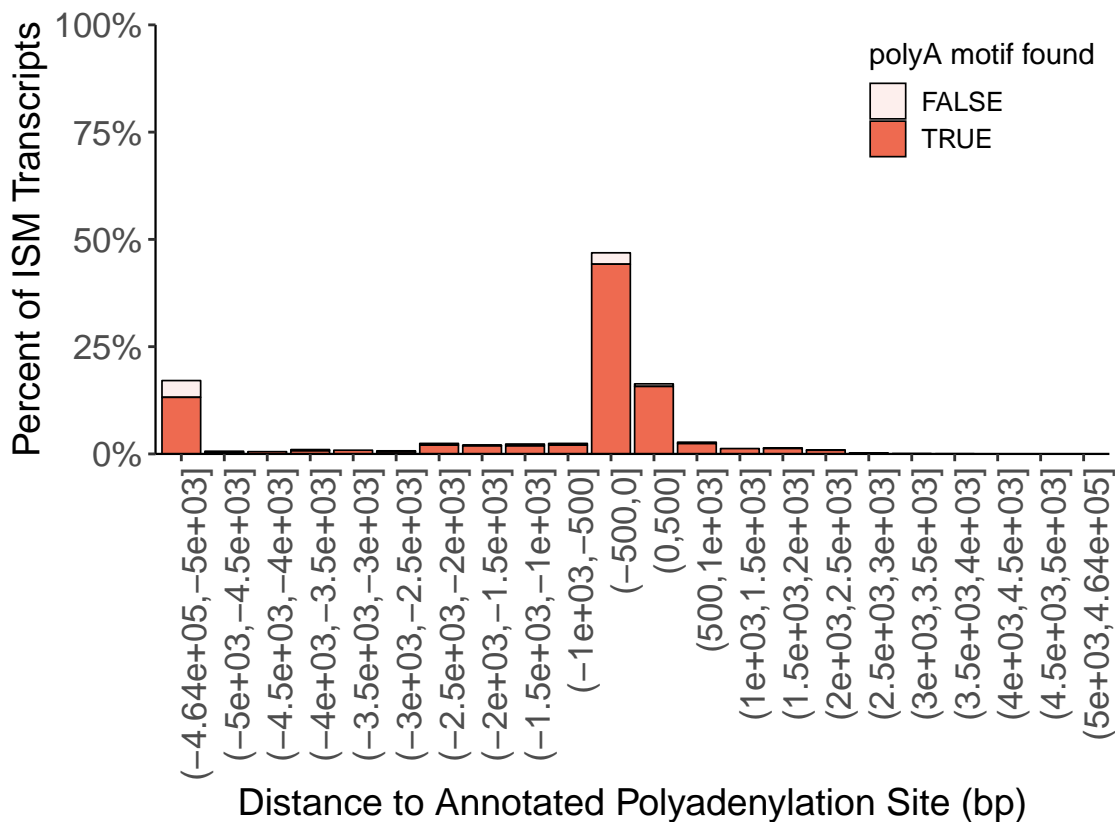
Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



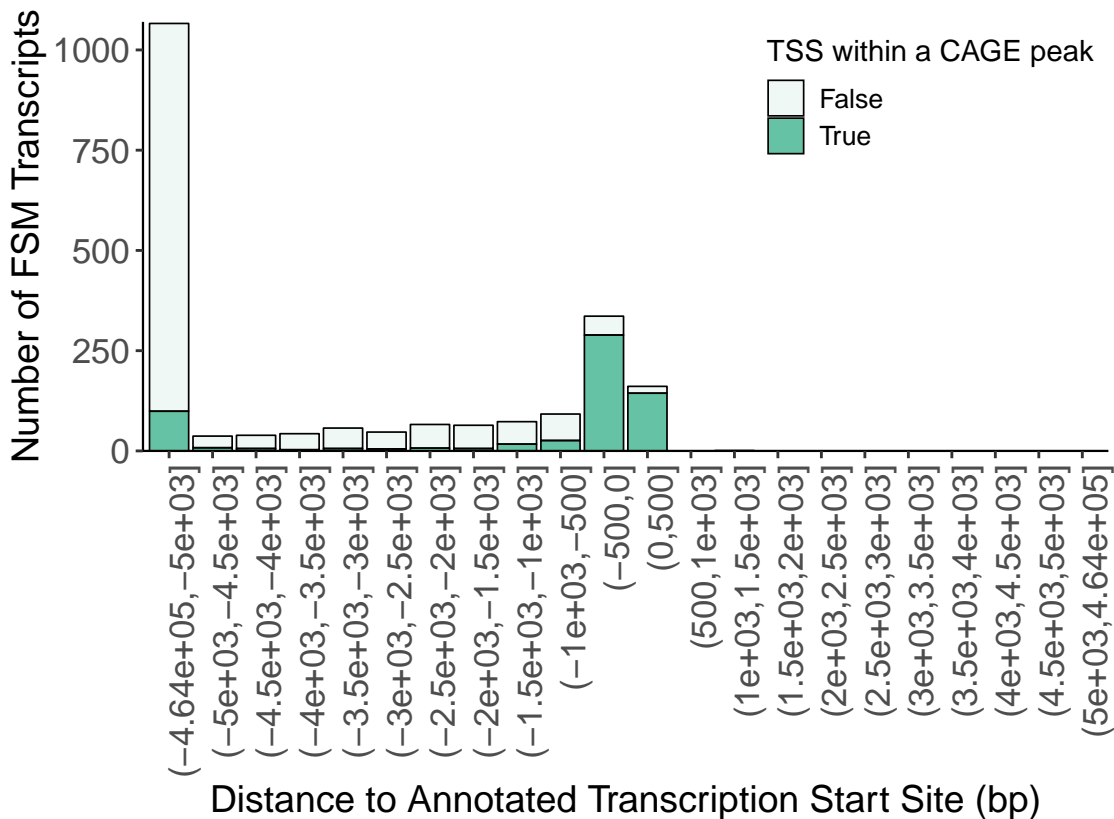
Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



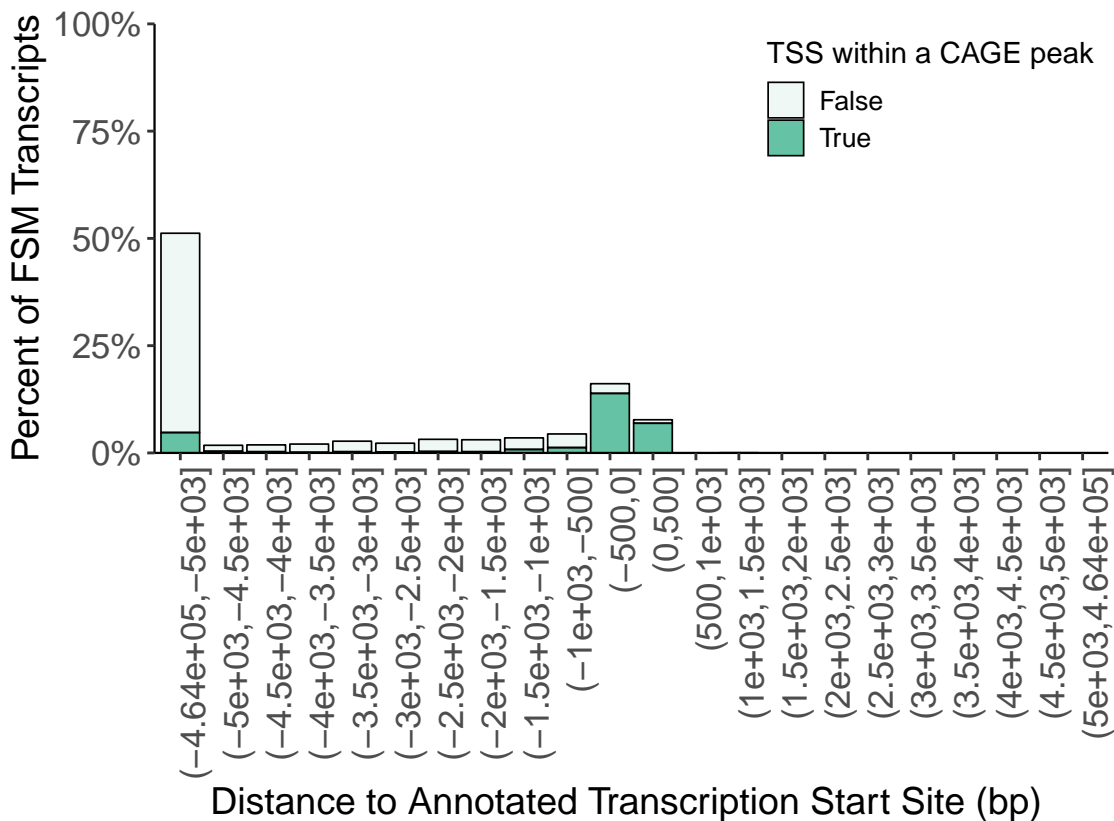
Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS



Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS



The graph displays the distribution of polyA motifs from the 3' end (bp) for four different conditions. The x-axis represents the distance of the polyA motif from the 3' end in base pairs (bp), ranging from -45 to 0. The y-axis represents the frequency or count of motifs. The purple line shows a sharp peak at approximately -15 bp. The blue line shows a broader peak at approximately -15 bp. The green line shows a small peak at approximately -15 bp. The yellow line is flat at 0.

- FSM
- ISM
- NIC
- NNC
- Genic
- Genomic
- Antisense
- Fusion
- Intergenic

Frequency of polyA motifs

Number of polyA Motifs Detected

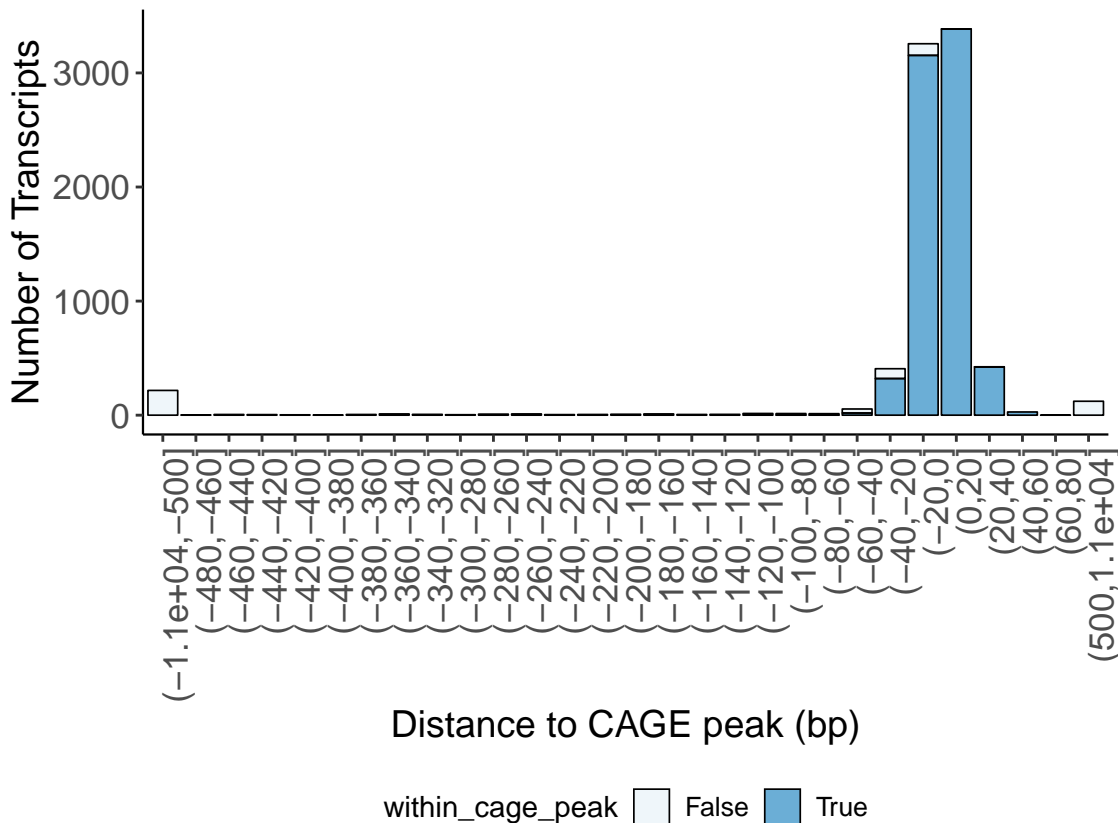
Category	Count	polyA Detected	%
FSM	8268	7541	91
ISM	2082	1883	90
NIC	1071	979	91
NNC	381	357	94
Genic Genomic	5	5	100
Antisense	1	1	100
Fusion	8	7	88
Intergenic	5	5	100

Motif	Count	%
AATAAA	7093	65.8
ATTAAA	1774	16.5
TATAAA	307	2.8
AGTAAA	305	2.8
AAGAAA	177	1.6
AATATA	154	1.4
AAAAAG	149	1.4
CATAAA	143	1.3
AATACA	142	1.3
GATAAA	123	1.1
TTTAAA	104	1.0
AATGAA	99	0.9
AAAACA	72	0.7
ACTAAA	66	0.6
AATAGA	62	0.6
GGGGCT	8	0.1

CAGE distances analysis

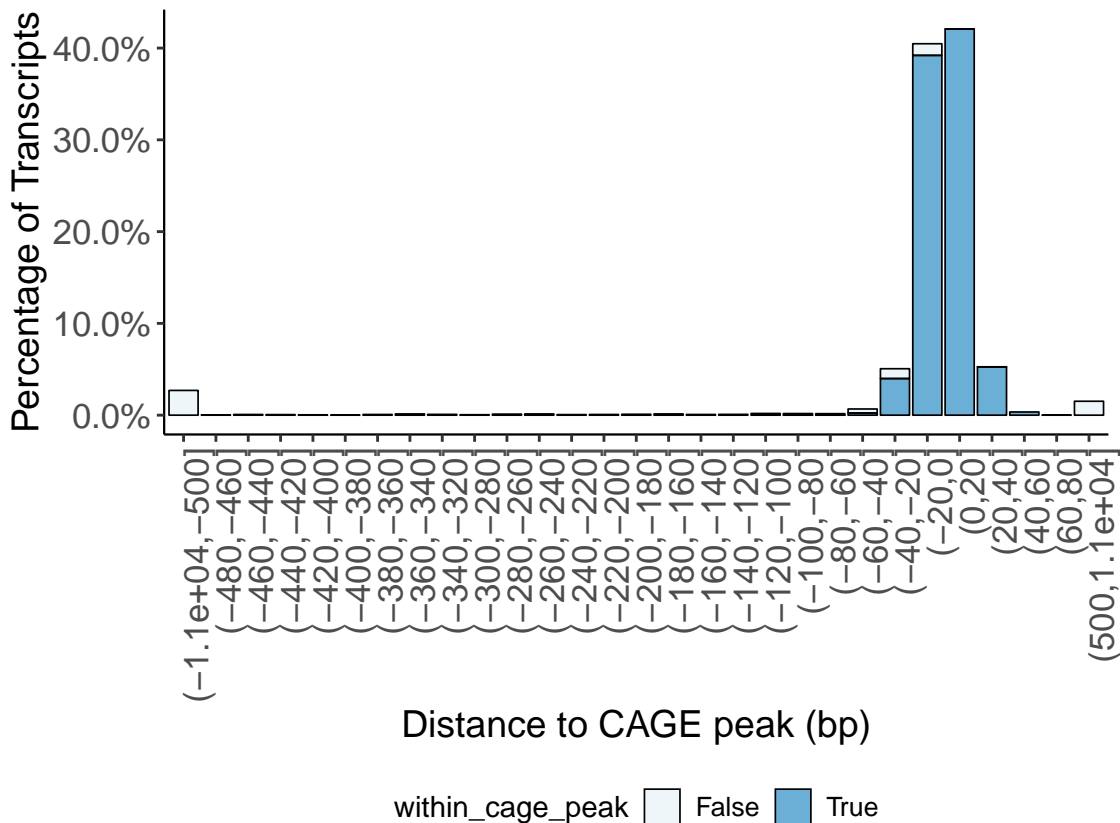
Distance to CAGE peak of multi-exonic FSM

Negative values indicate downstream of annotated CAGE peak



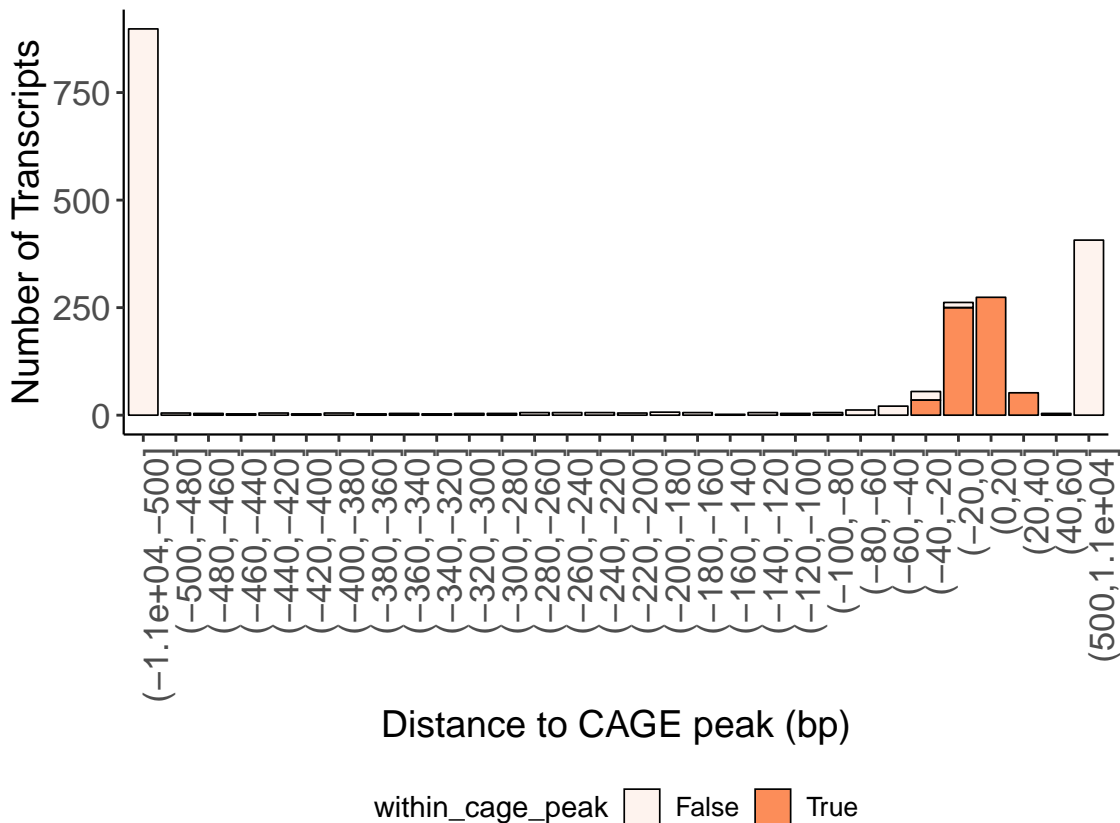
Negative values indicate downstream of annotated CAGE peak

Negative values indicate downstream of annotated CAGE peak



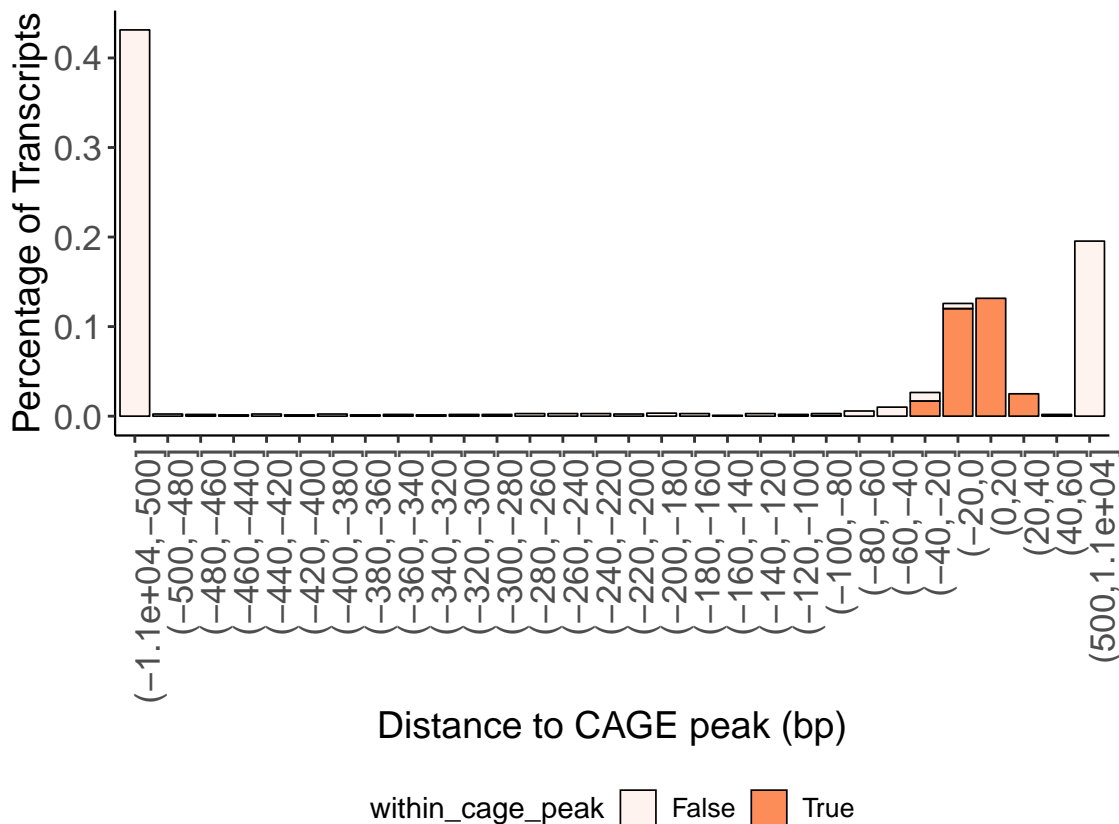
Distance to CAGE peak of multi-exonic ISM

Negative values indicate downstream of annotated CAGE peak

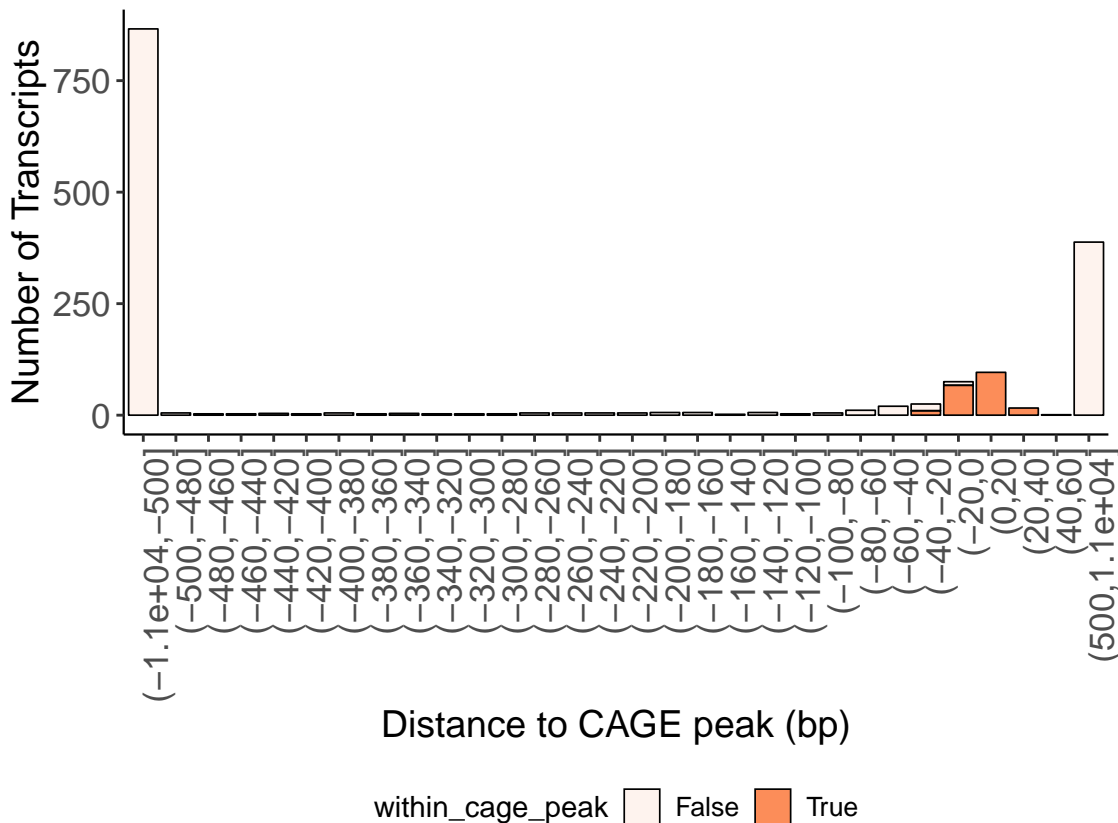


Distance to CAGE peak of multi-exonic ISM

Negative values indicate downstream of annotated CAGE peak

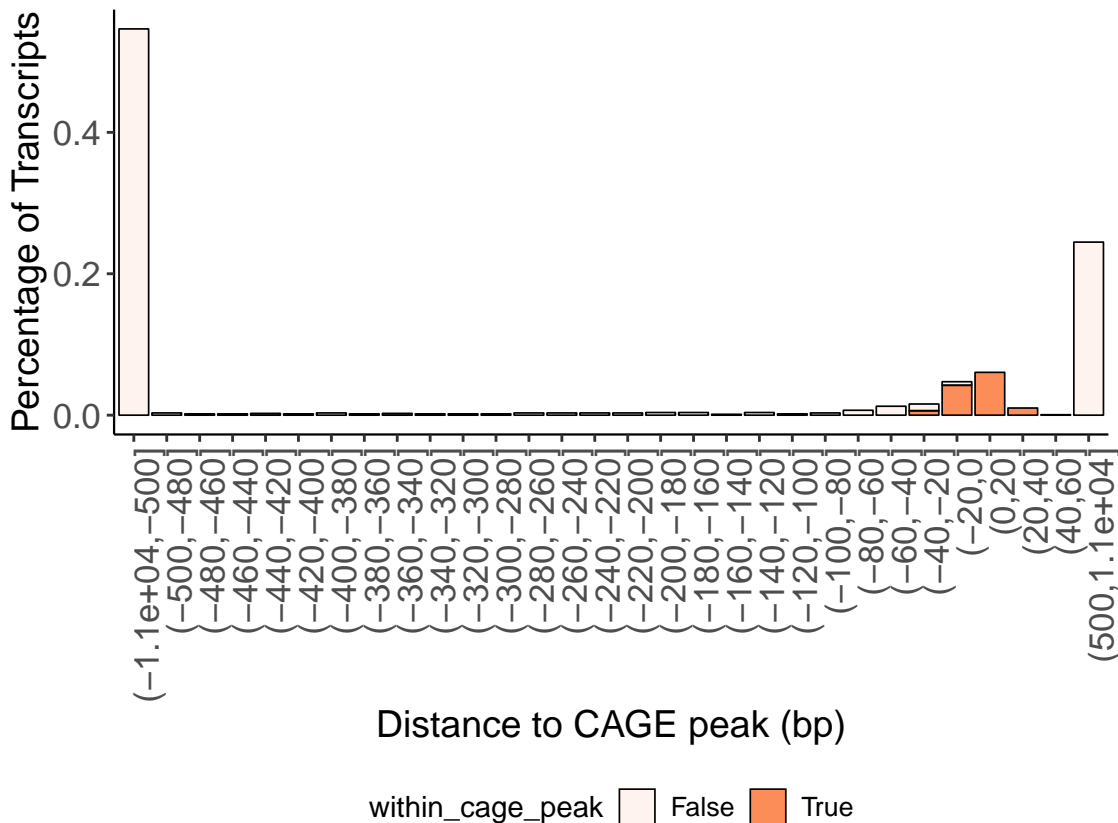


Distance to CAGE peak of multi-exonic ISM 3prime fragments
Negative values indicate downstream of annotated CAGE peak



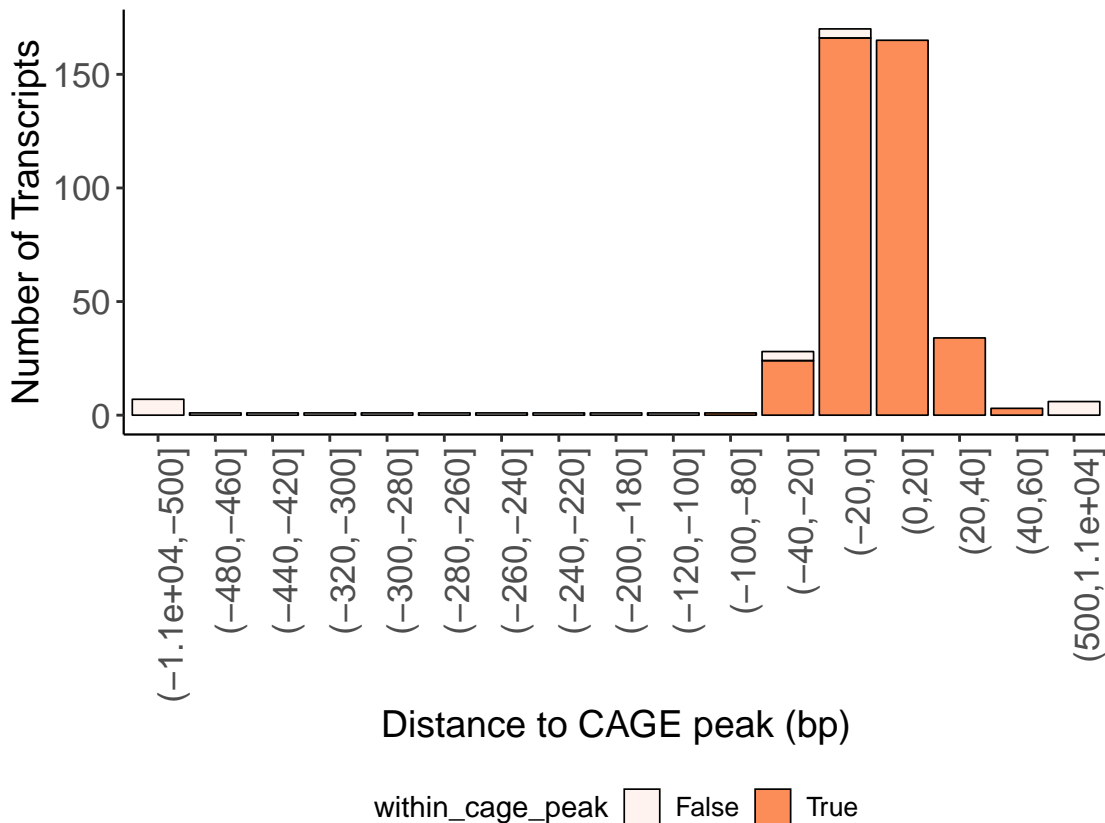
Distance to CAGE peak of multi-exonic ISM 3prime fragments

Negative values indicate downstream of annotated CAGE peak



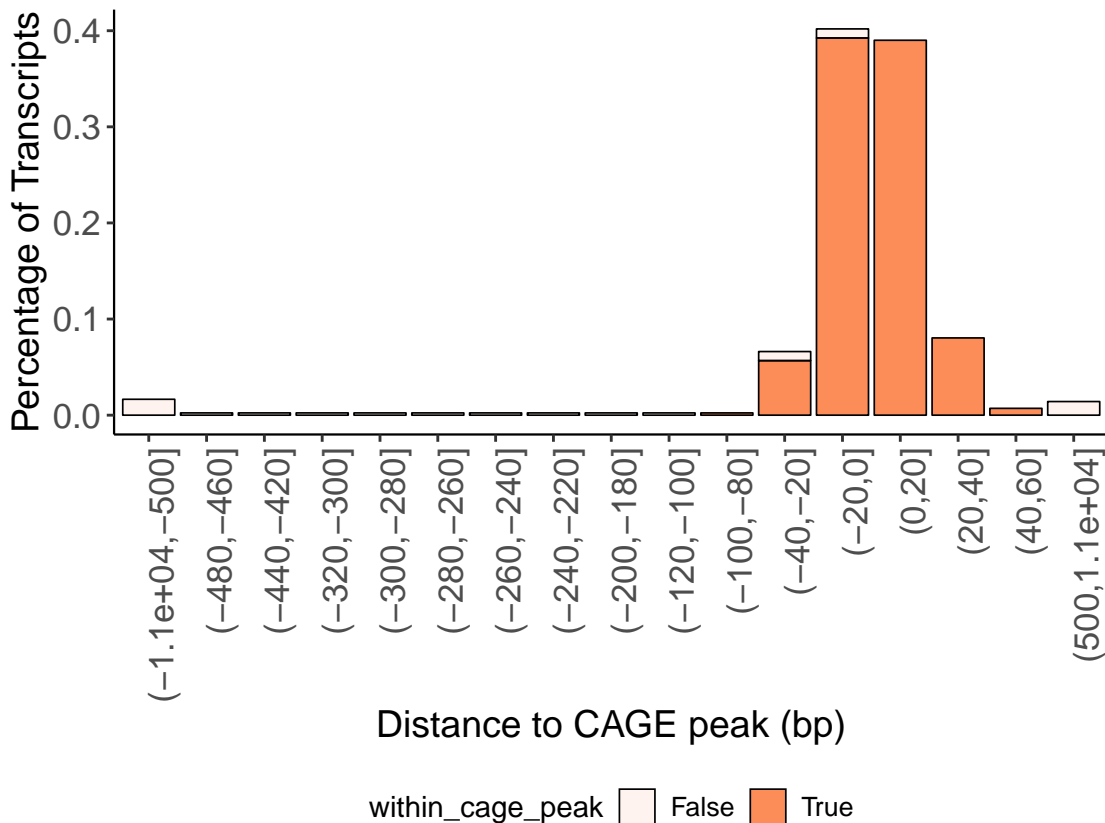
Distance to CAGE peak of multi-exonic ISM 5prime fragments

Negative values indicate downstream of annotated CAGE peak



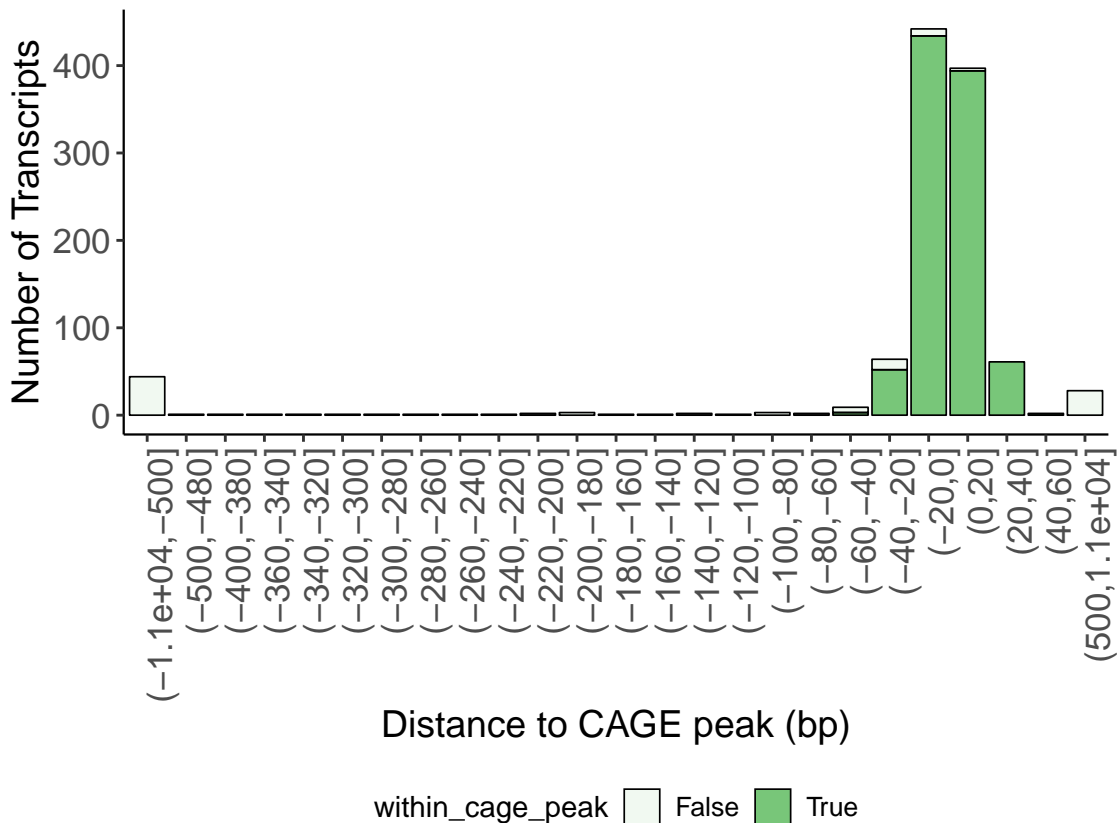
Distance to CAGE peak of multi-exonic ISM 5prime fragments

Negative values indicate downstream of annotated CAGE peak



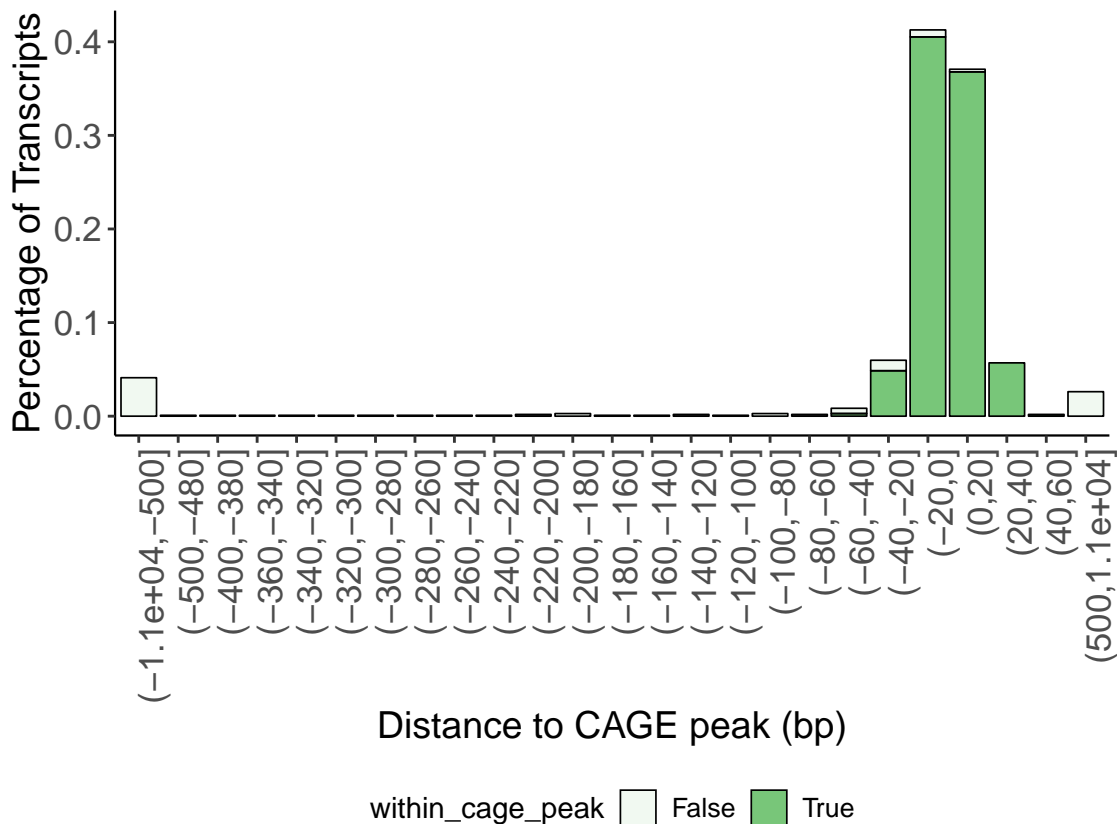
Distance to CAGE peak of multi-exonic NIC

Negative values indicate downstream of annotated CAGE peak



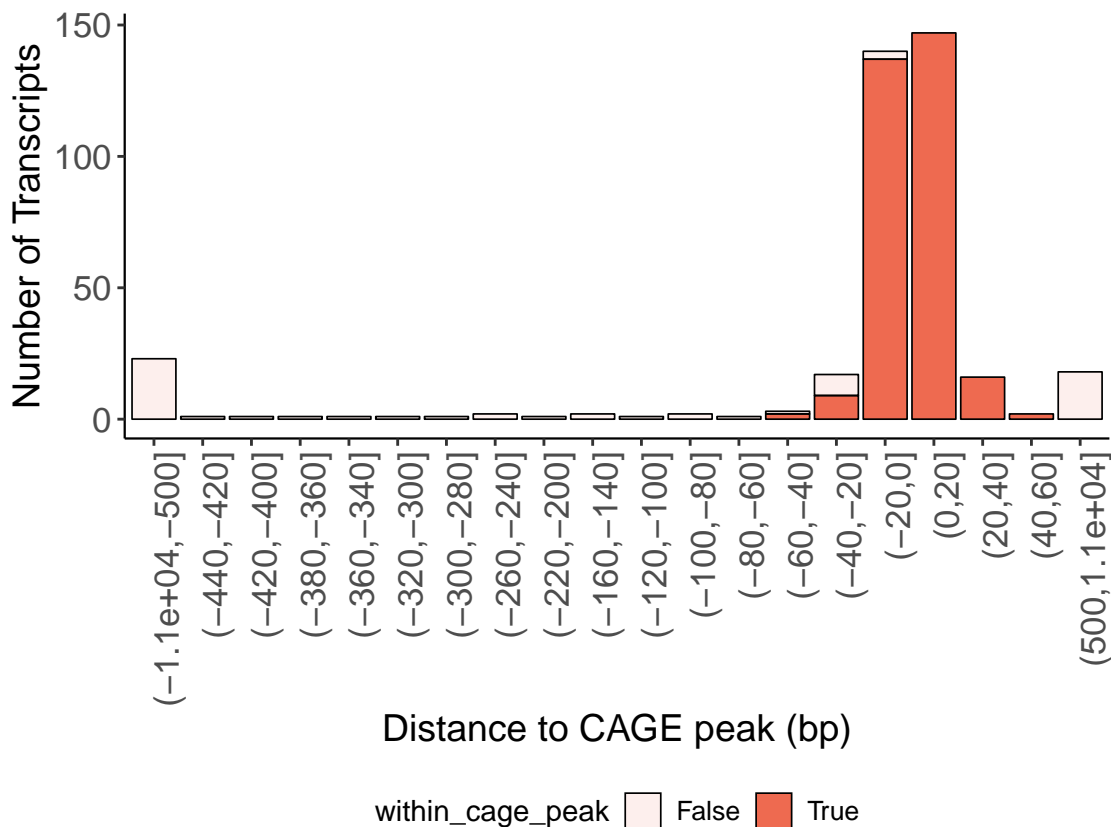
Distance to CAGE peak of multi-exonic NIC

Negative values indicate downstream of annotated CAGE peak



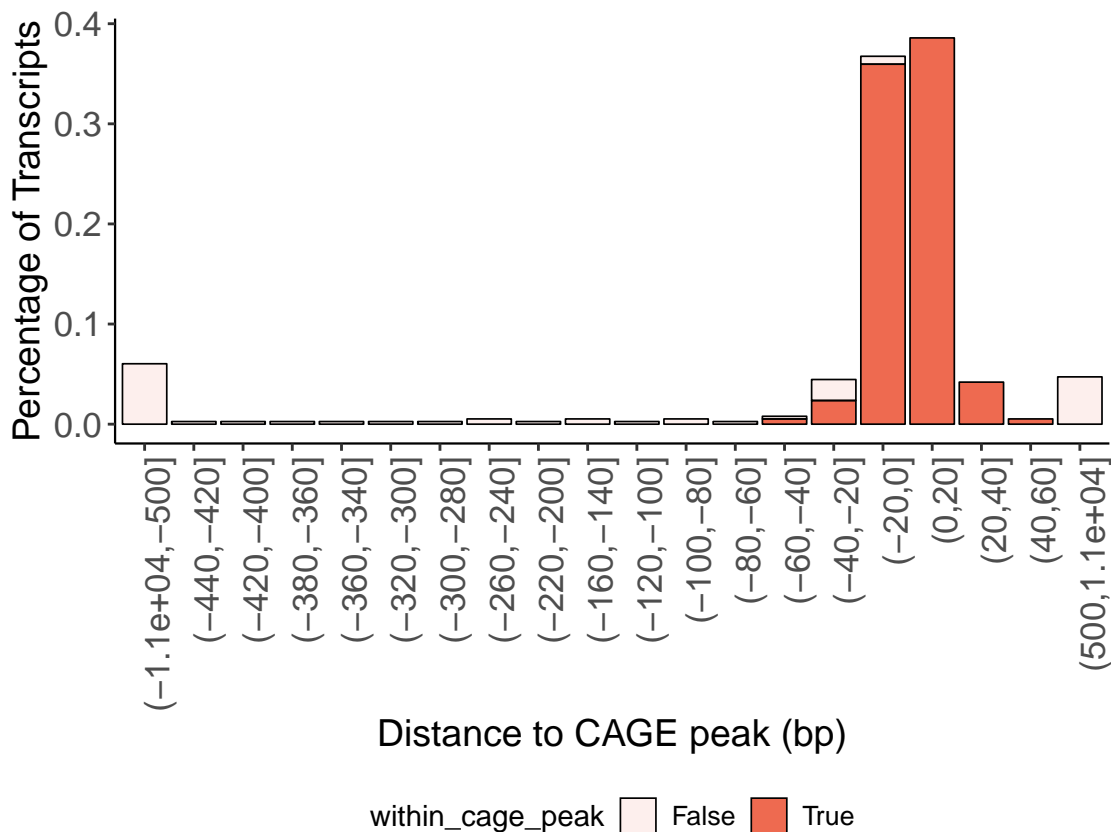
Distance to CAGE peak of multi-exonic NNC

Negative values indicate downstream of annotated CAGE peak



Distance to CAGE peak of multi-exonic NNC

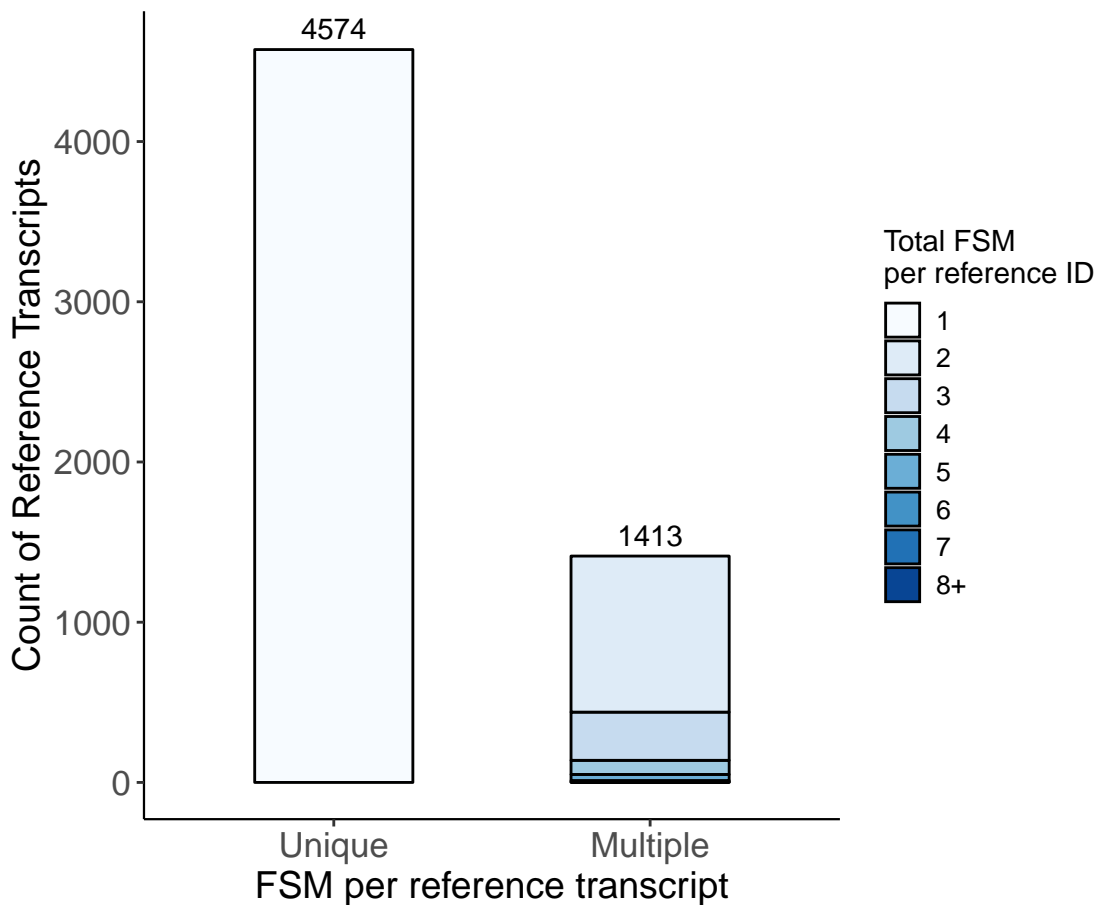
Negative values indicate downstream of annotated CAGE peak



*Accumulation of FSM and ISM
to the same reference transcript*

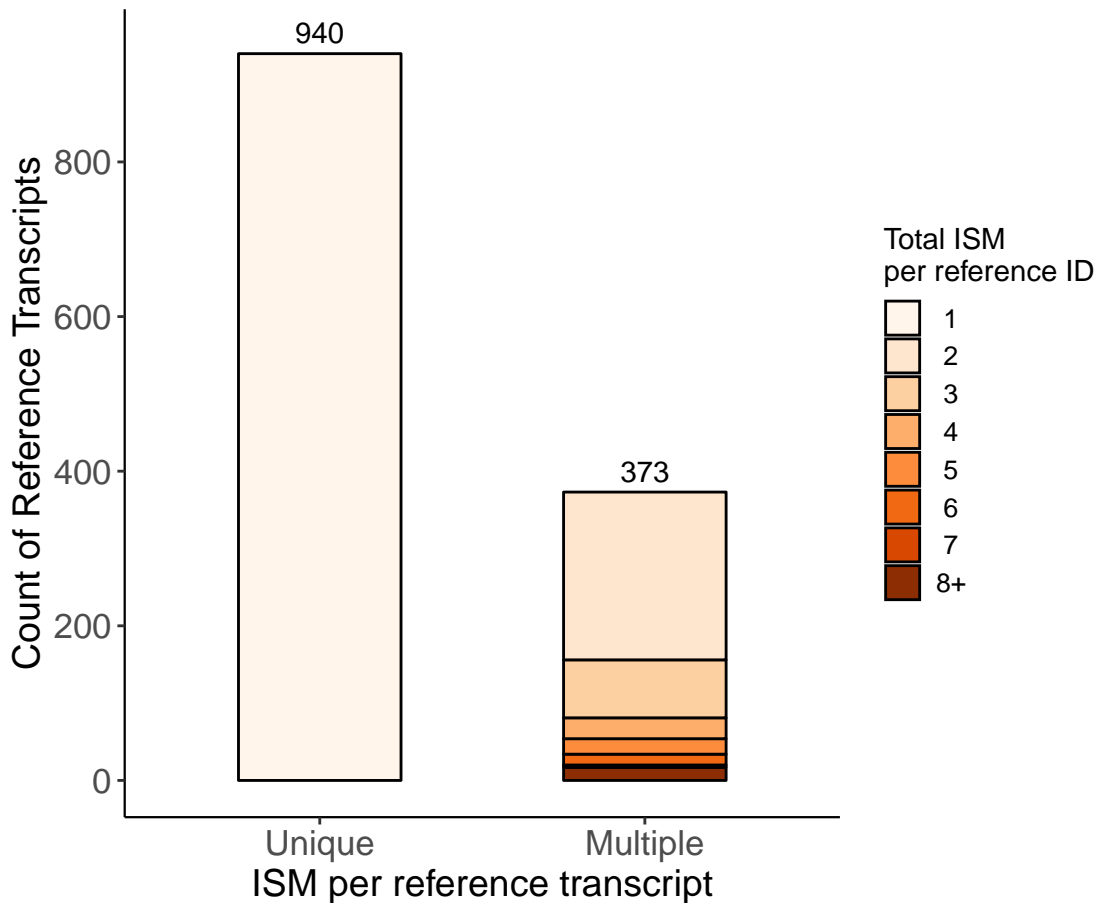
Reference transcript redundance

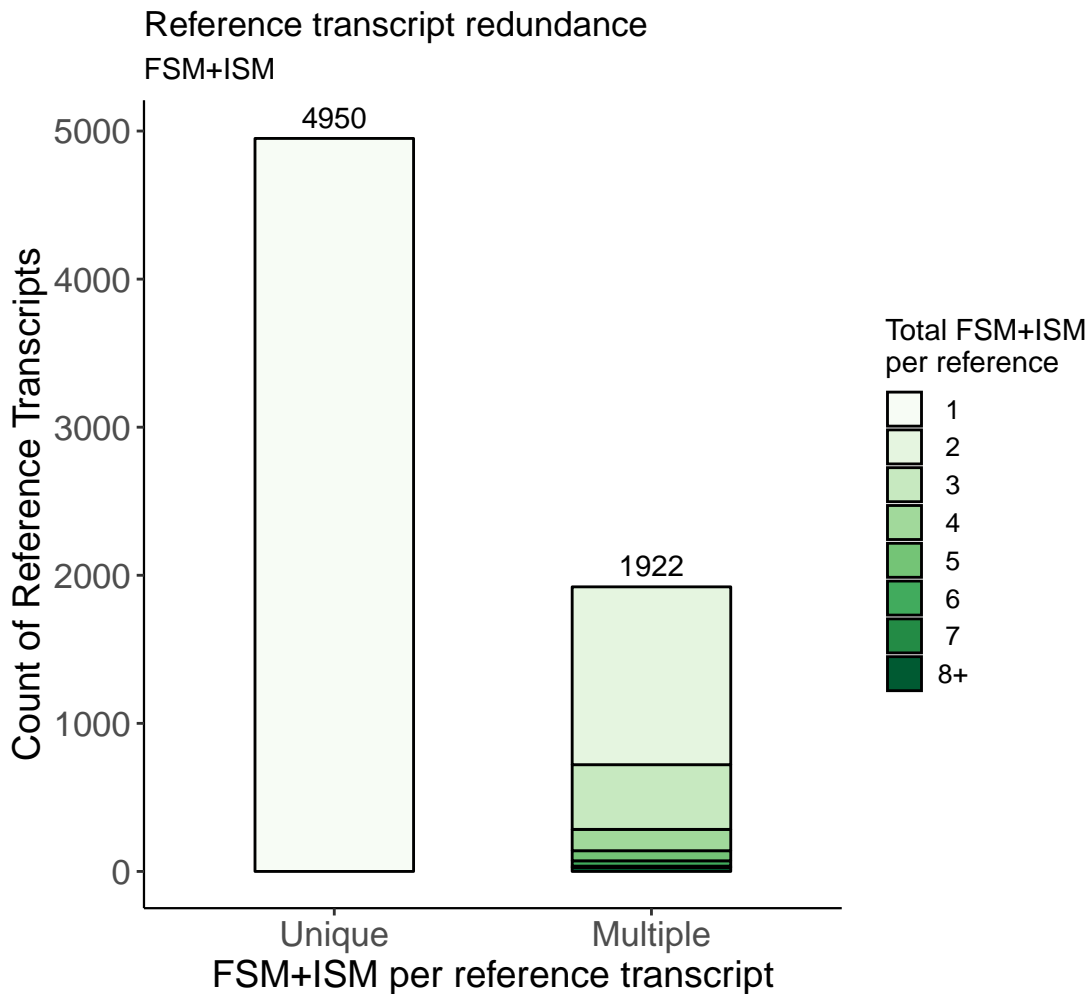
Only FSM



Reference transcript redundance

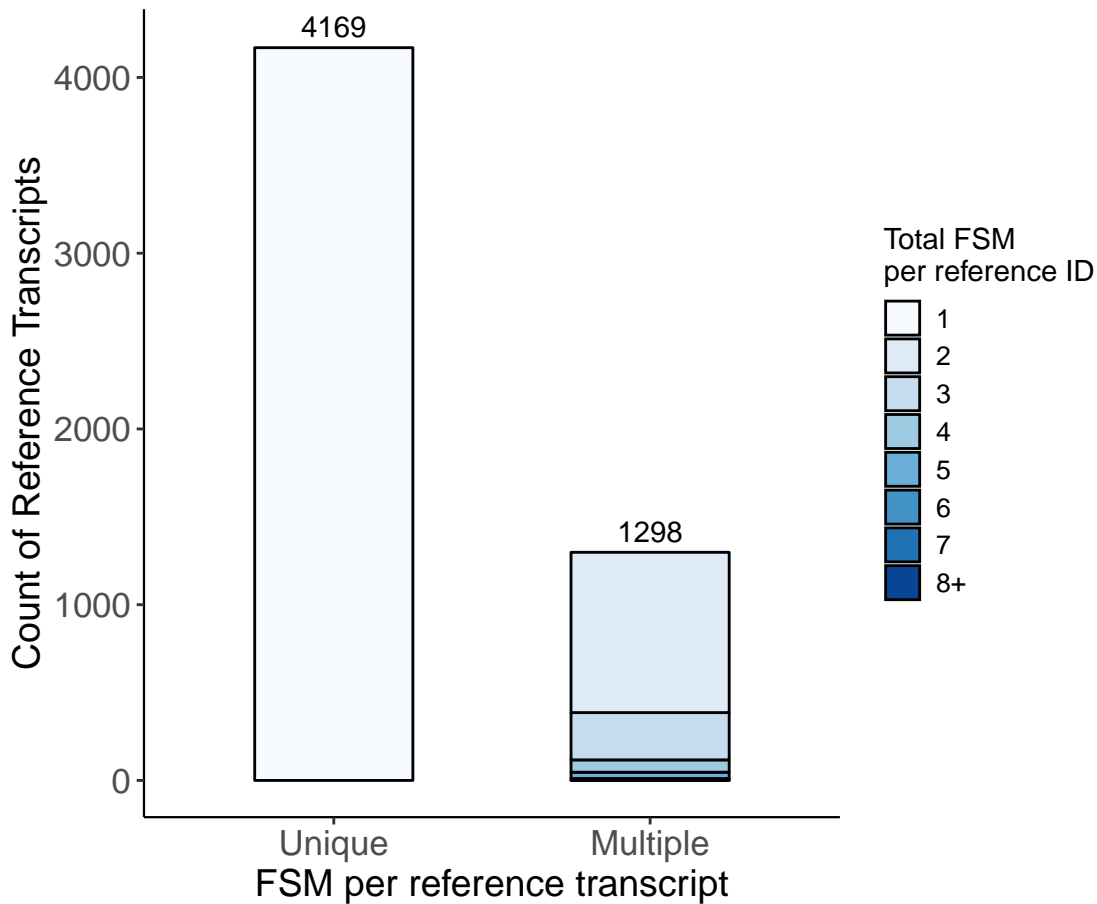
Only ISM





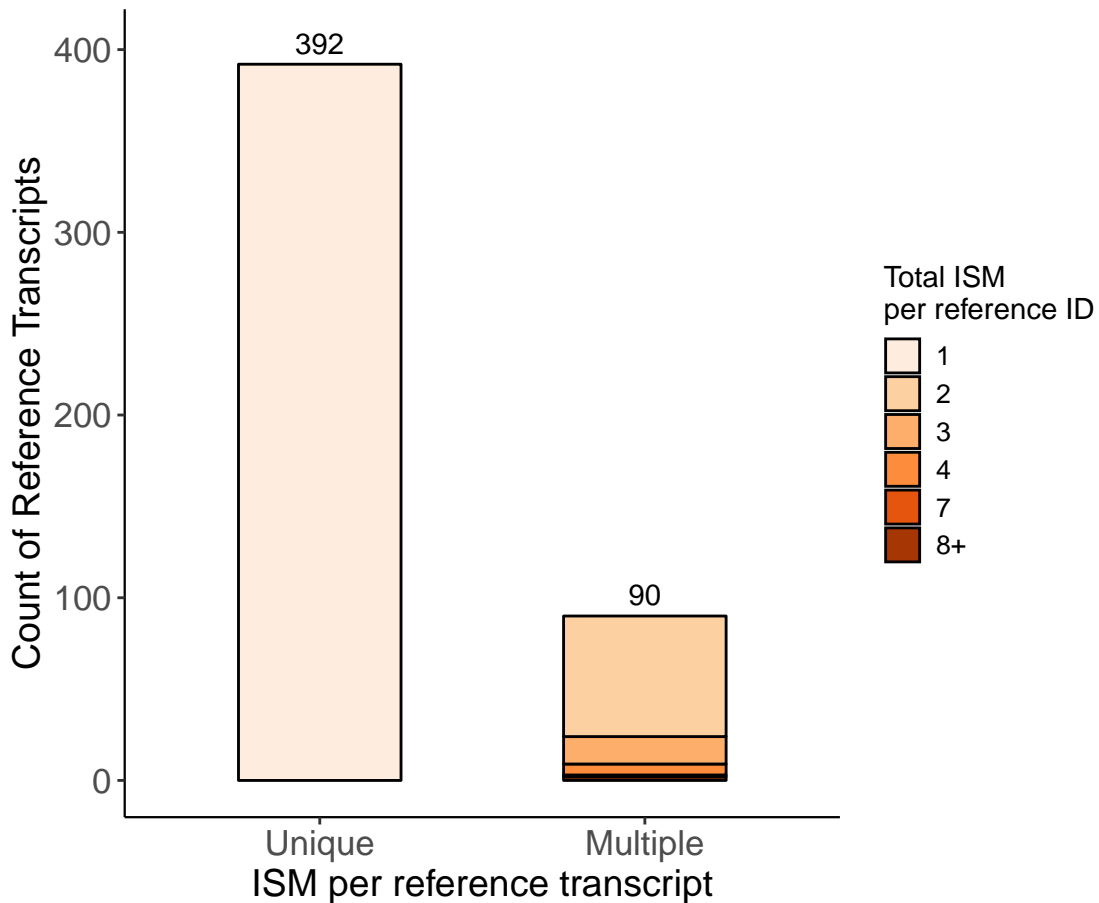
Reference transcript redundance

Only FSM with CAGE support



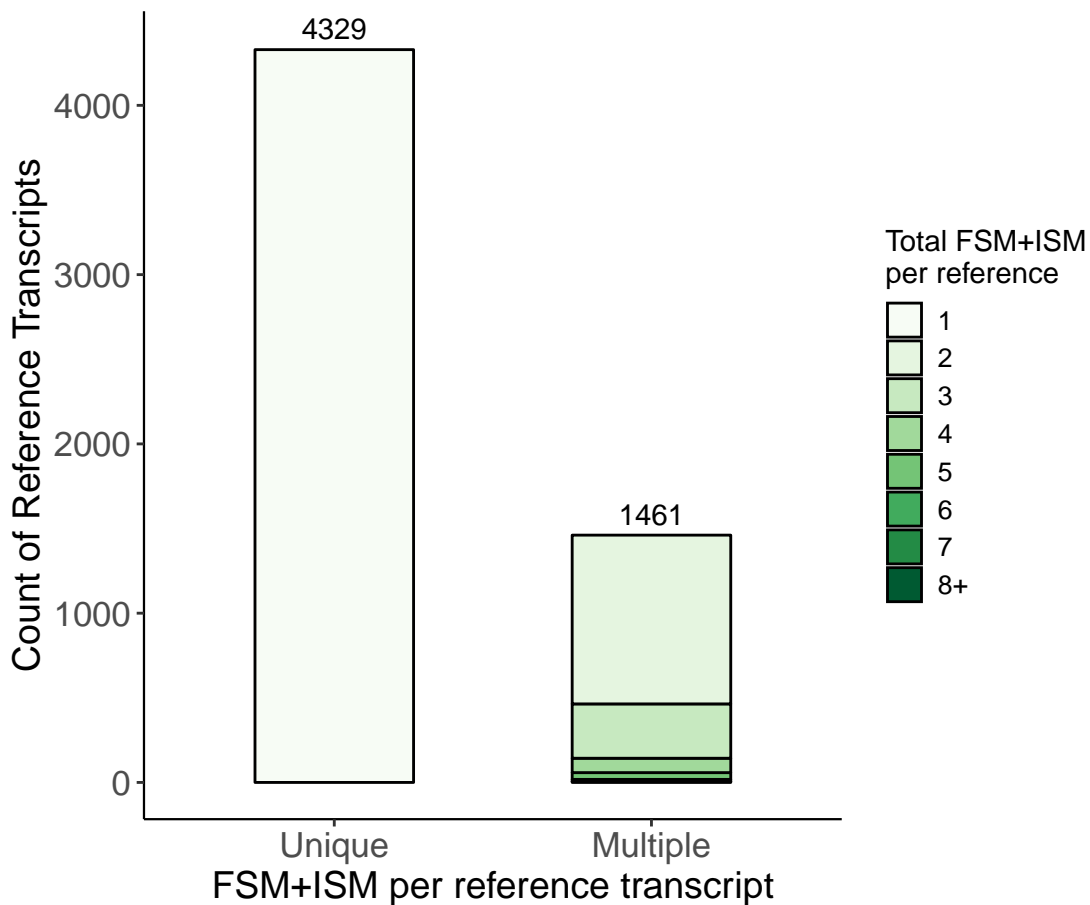
Reference transcript redundance

Only ISM with CAGE support



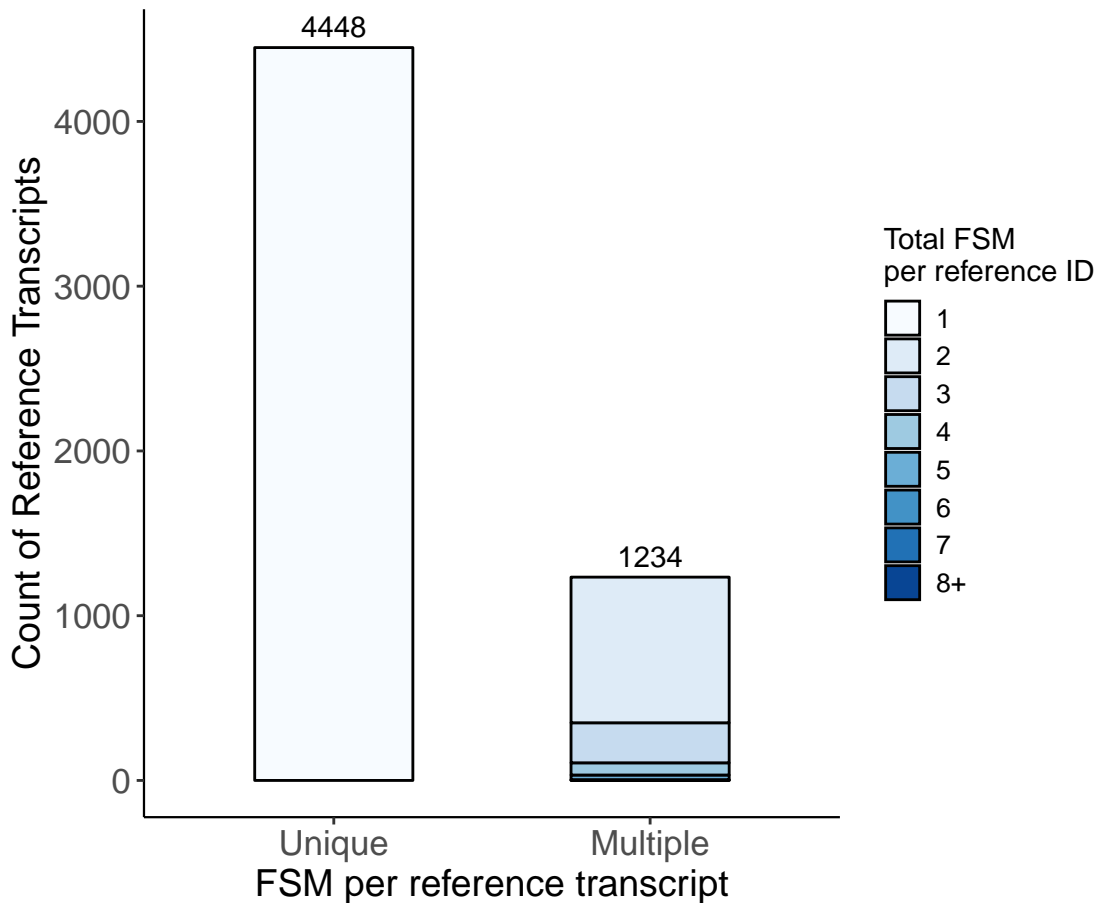
Reference transcript redundance

FSM+ISM with CAGE support



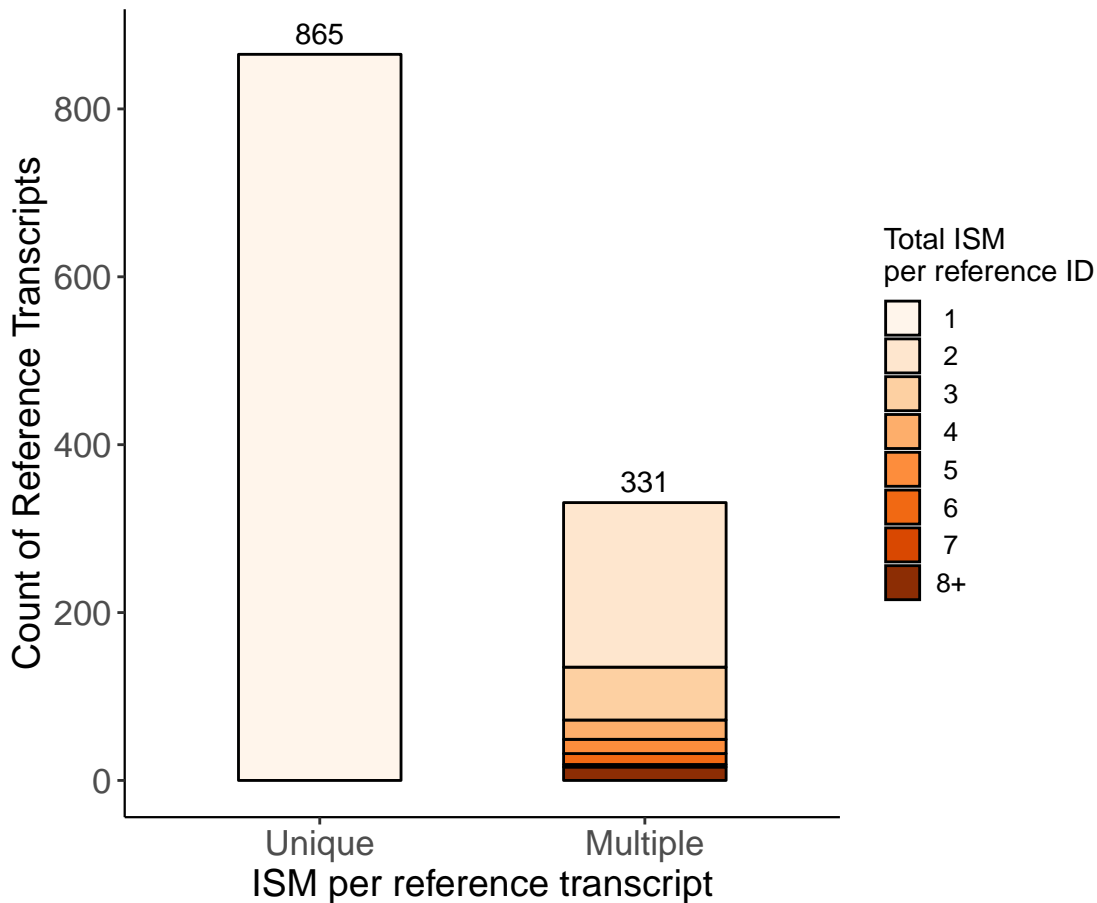
Reference transcript redundance

Only FSM with a polyA motif found



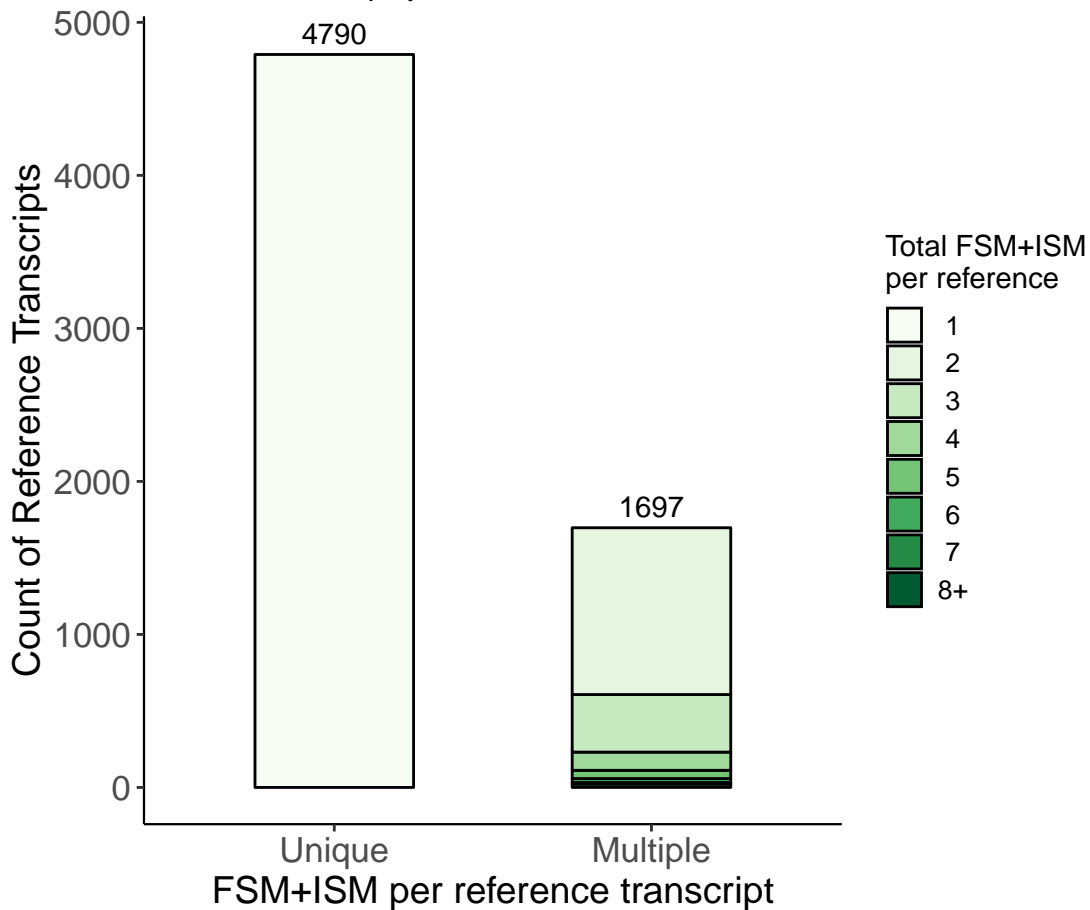
Reference transcript redundancy

Only ISM with a polyA motif found



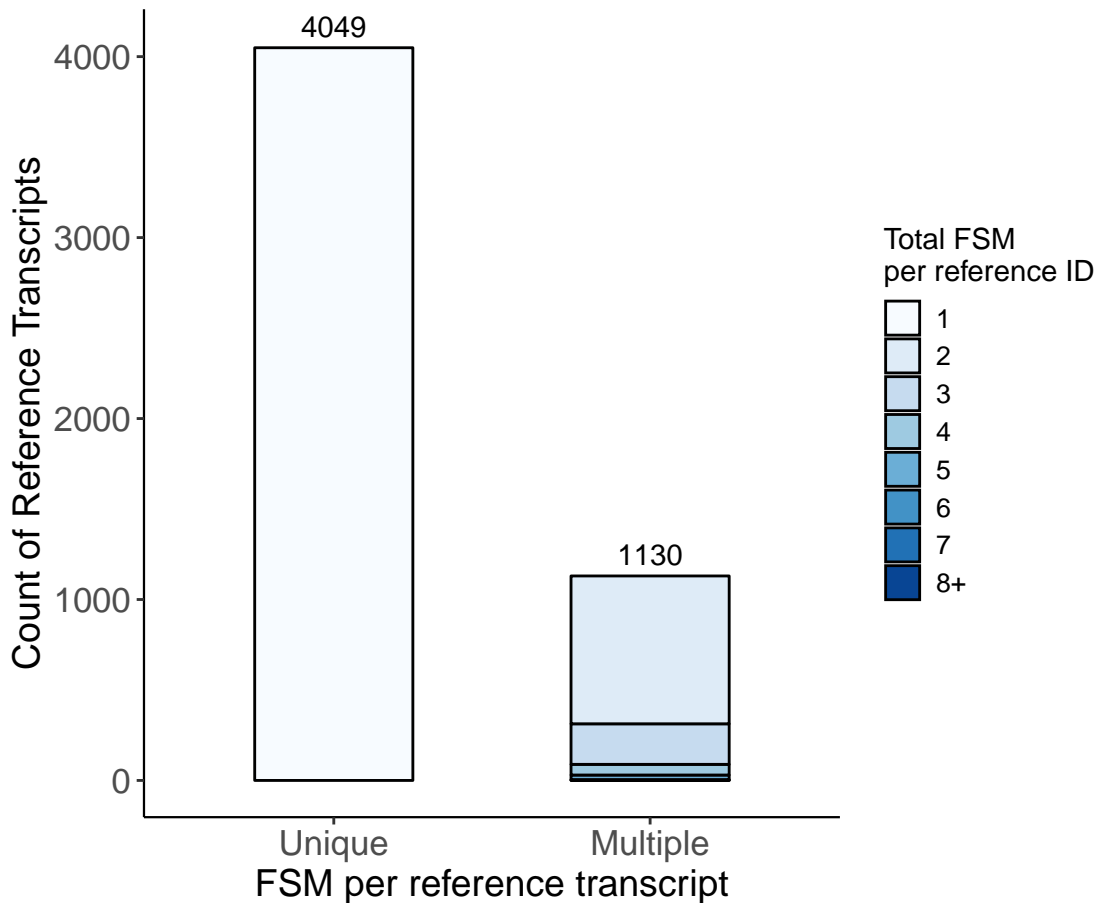
Reference transcript redundance

FSM+ISM with a polyA motif found



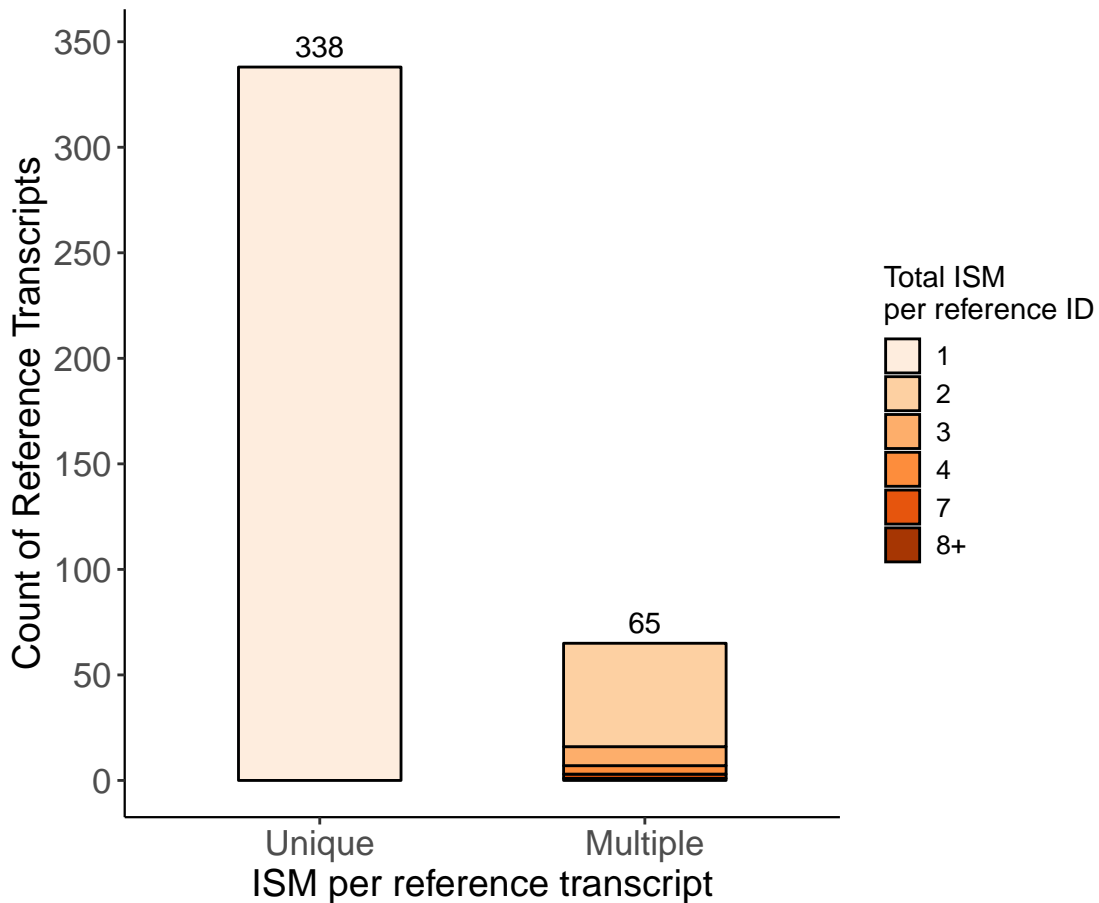
Reference transcript redundance

Only FSM with CAGE support and polyA motif

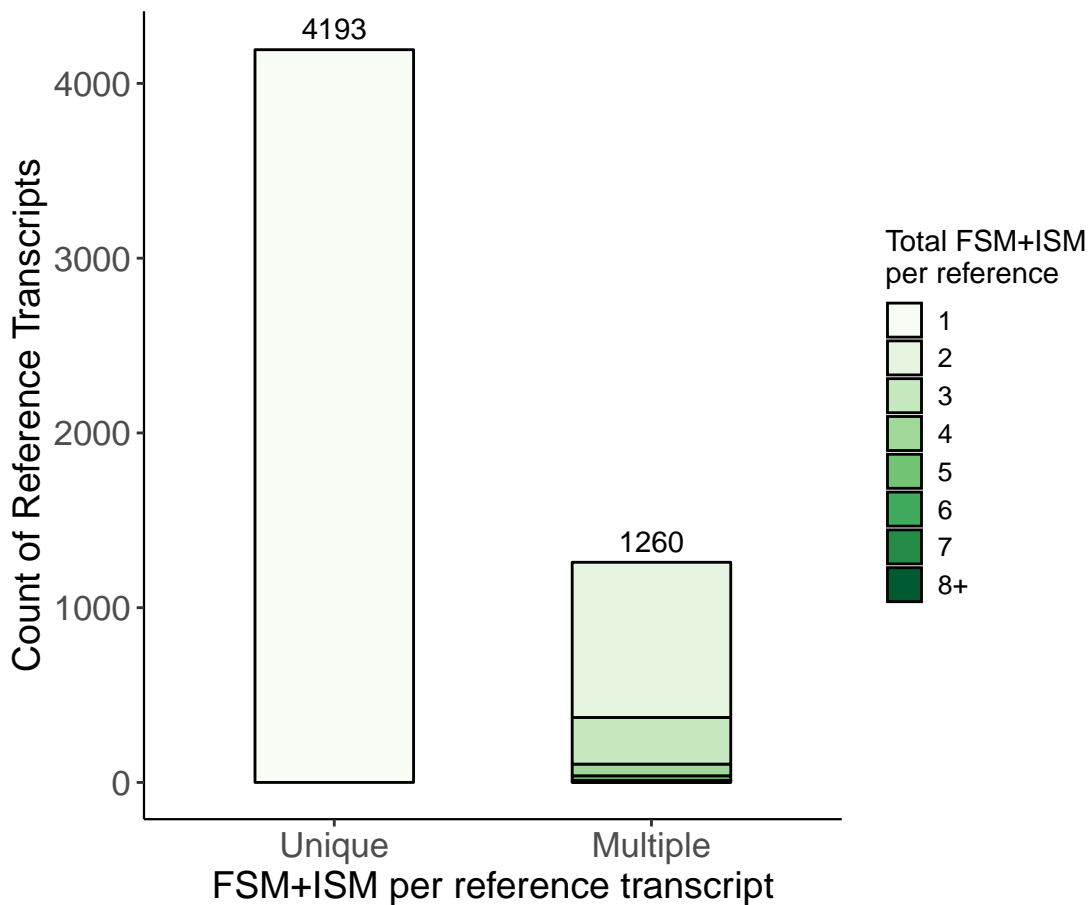


Reference transcript redundance

Only ISM with CAGE support and polyA motif



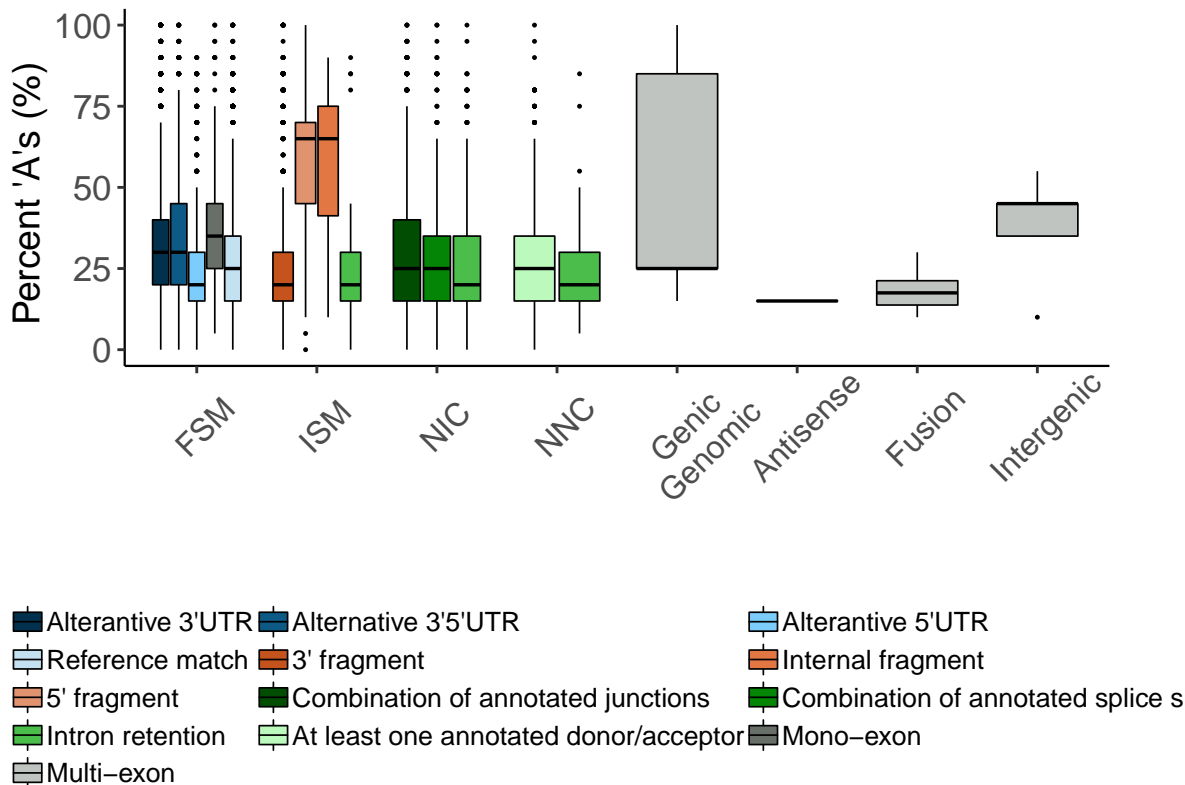
Reference transcript redundance
FSM+ISM with CAGE support and polyA motif



Intra-Priming Quality Check

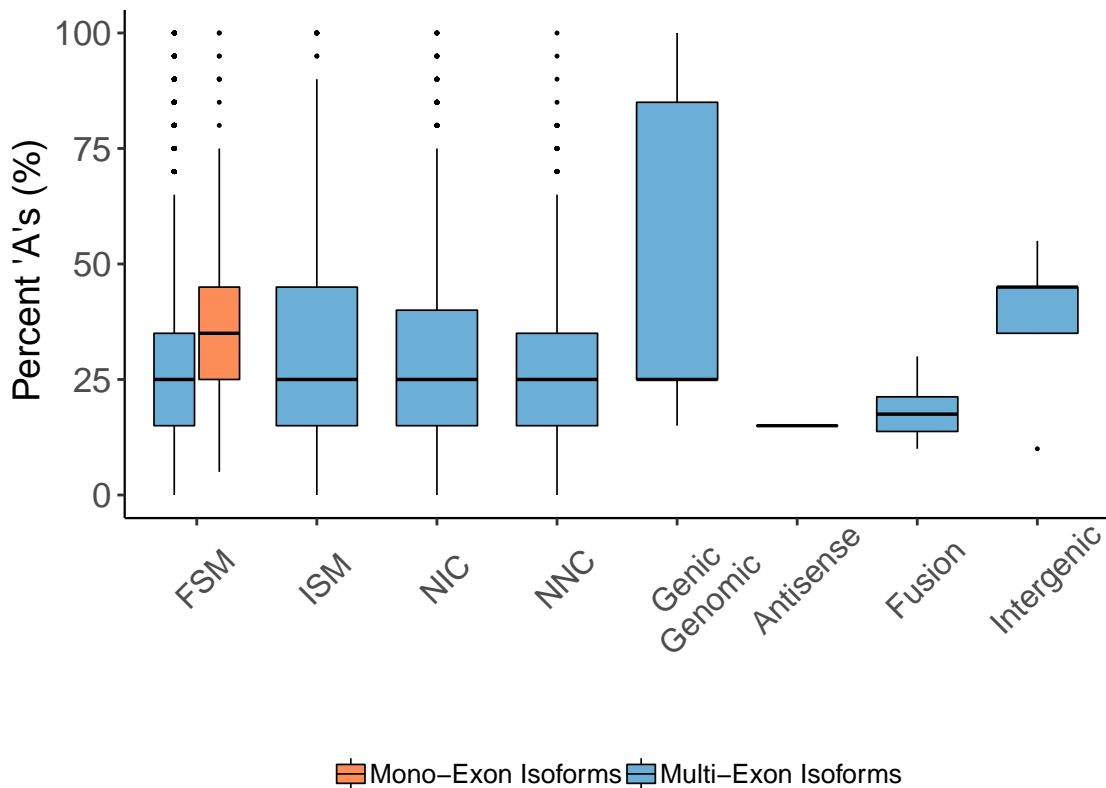
Possible Intra-Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



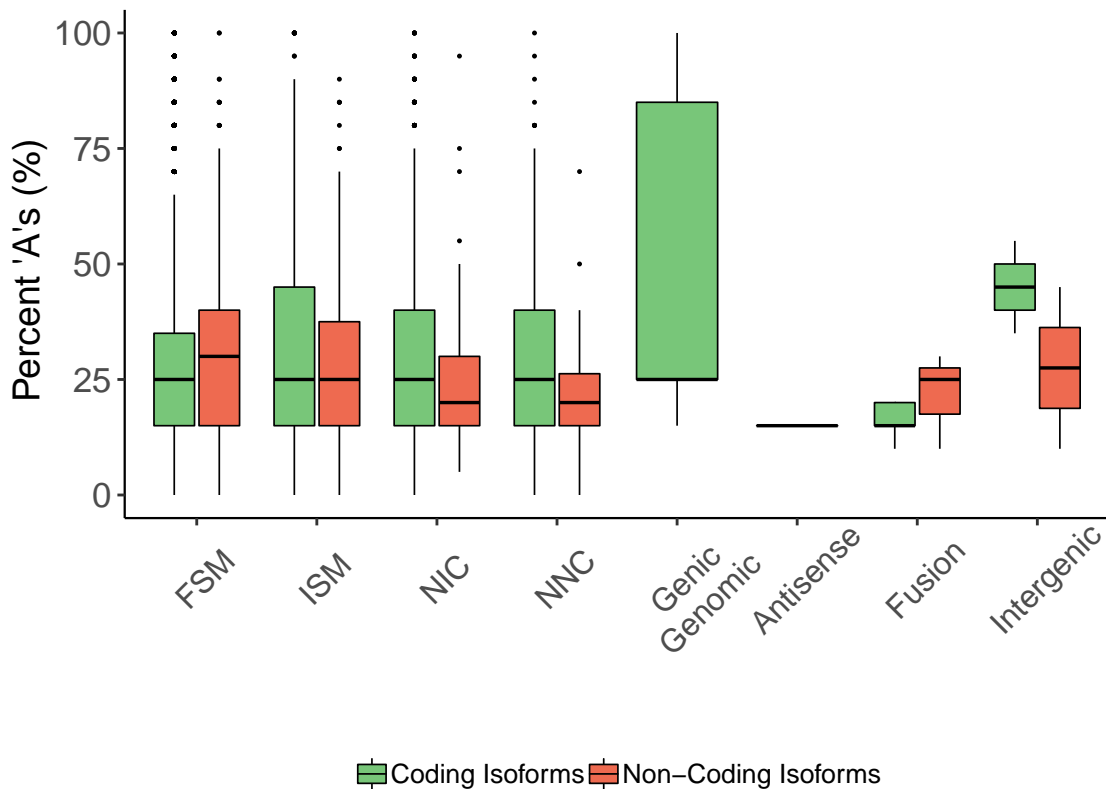
Possible Intra-Priming, Mono- vs Multi-Exon

Percent of genomic 'A's in downstream 20 bp



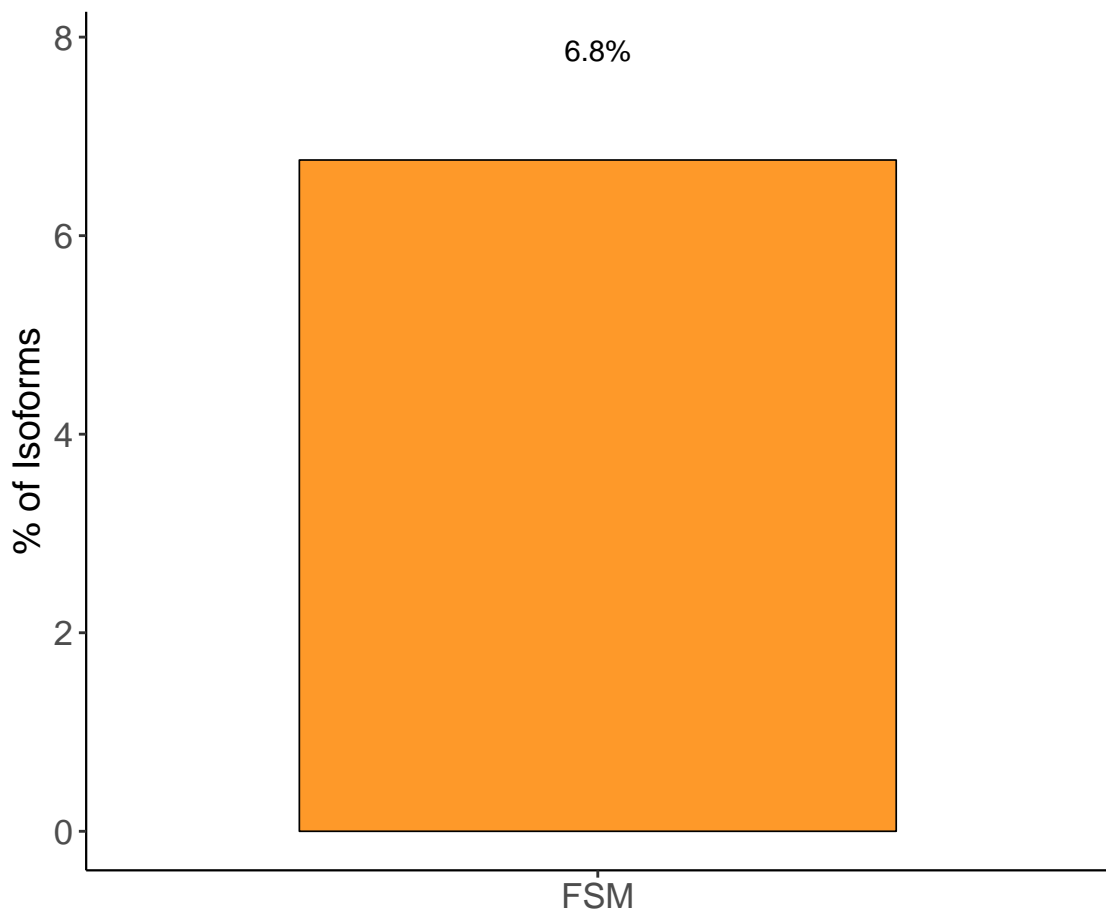
Possible Intra-Priming, Coding vs Non-Coding

Percent of genomic 'A's in downstream 20 bp

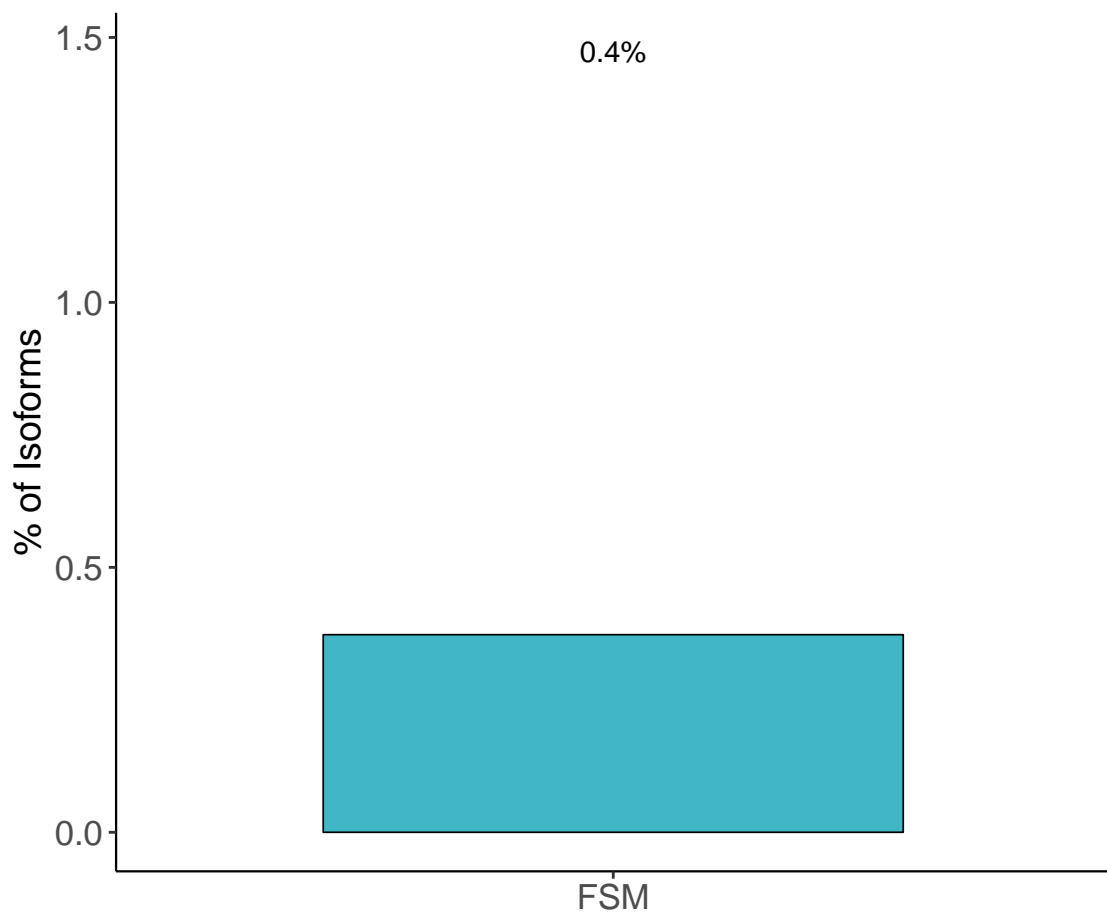


Quality Controls

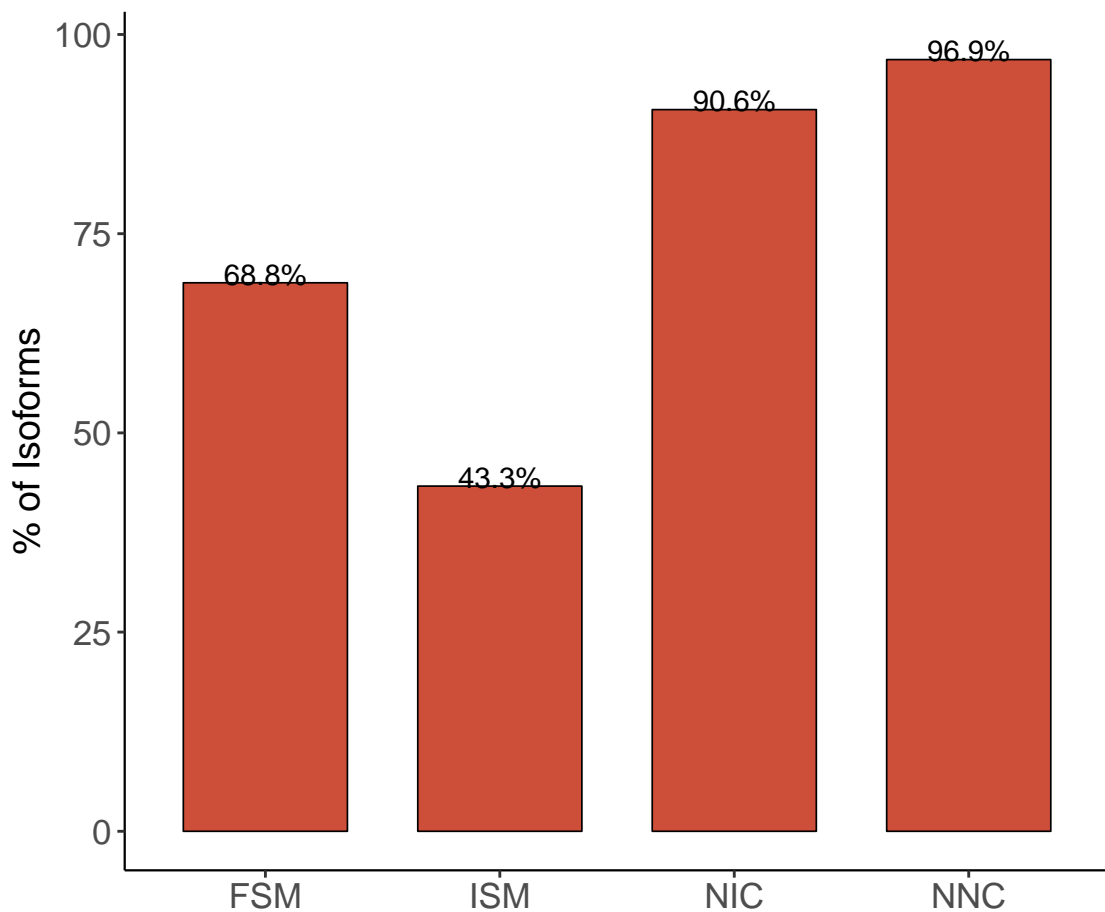
Incidence of RT-switching



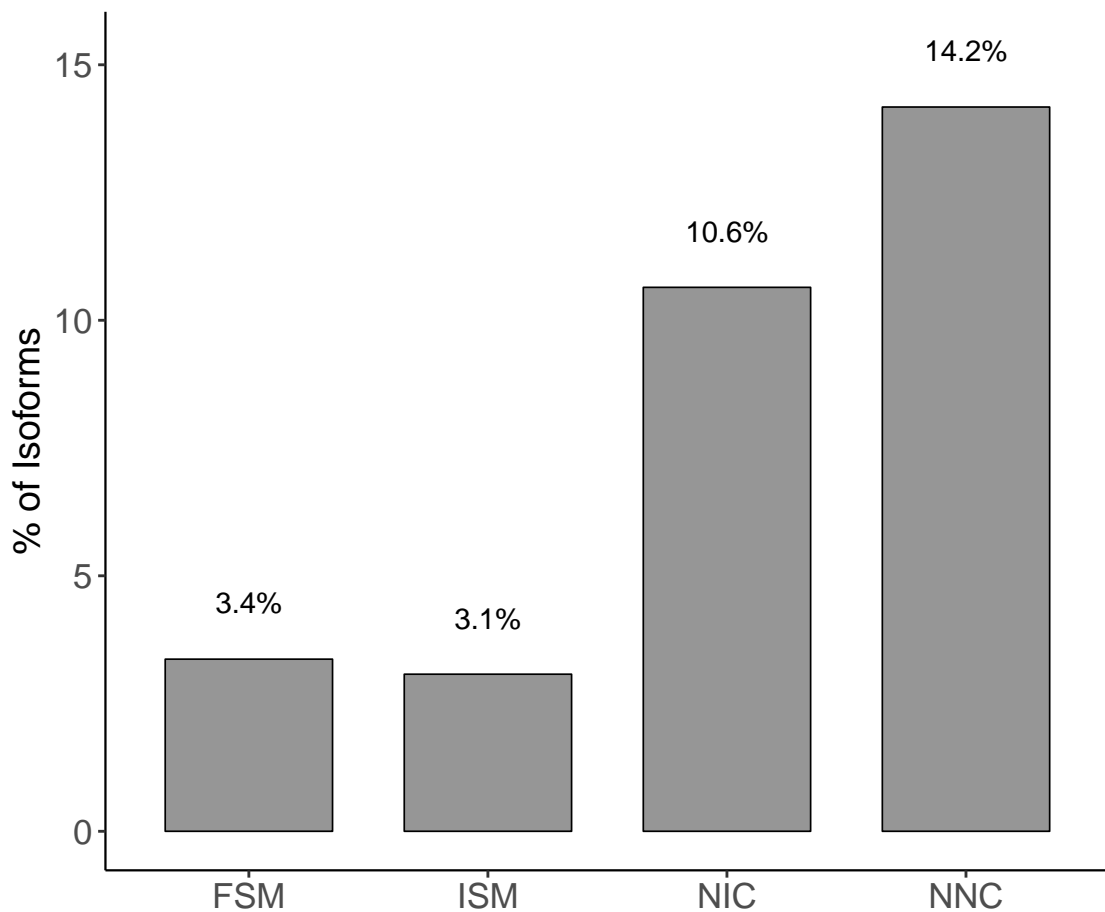
Incidence of Non-Canonical Junctions



Incidence of SJ without SR coverage Junctions



Incidence of NMD by structural category



Quality control attributes across structural categories

