

SQANTI3 report

Unique Genes: 6305
Unique Isoforms: 14341

Gene classification

Category	# Genes
Annotated Genes	5763
Novel Genes	542

Splice Junction Classification

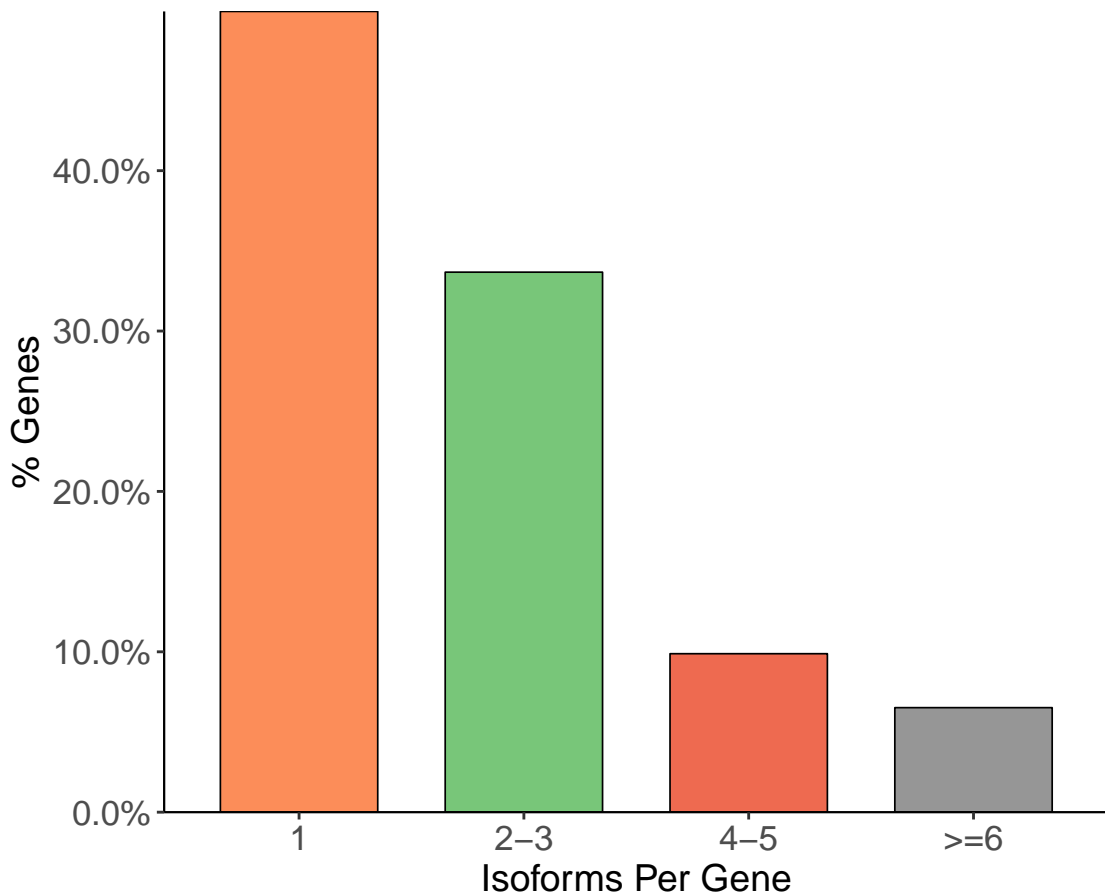
Category	# SJs	Percent
Known canonical	51388	97.80
Known Non-canonical	28	0.05
Novel canonical	726	1.38
Novel Non-canonical	402	0.77

*Characterization of transcripts
based on splice junctions*

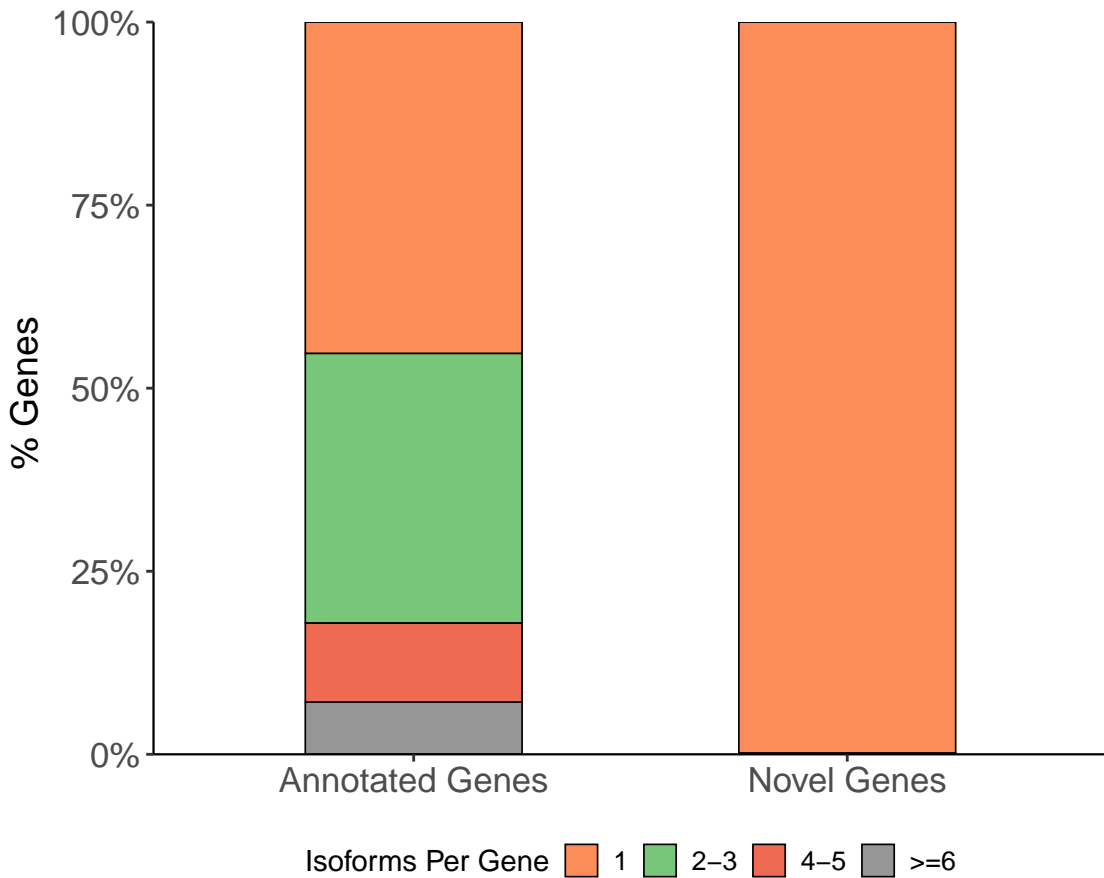
Category	# Isoforms
FSM	8627
ISM	2853
NIC	1294
NNC	965
Genic Genomic	51
Antisense	13
Fusion	8
Intergenic	530
Genic Intron	0

Gene Characterization

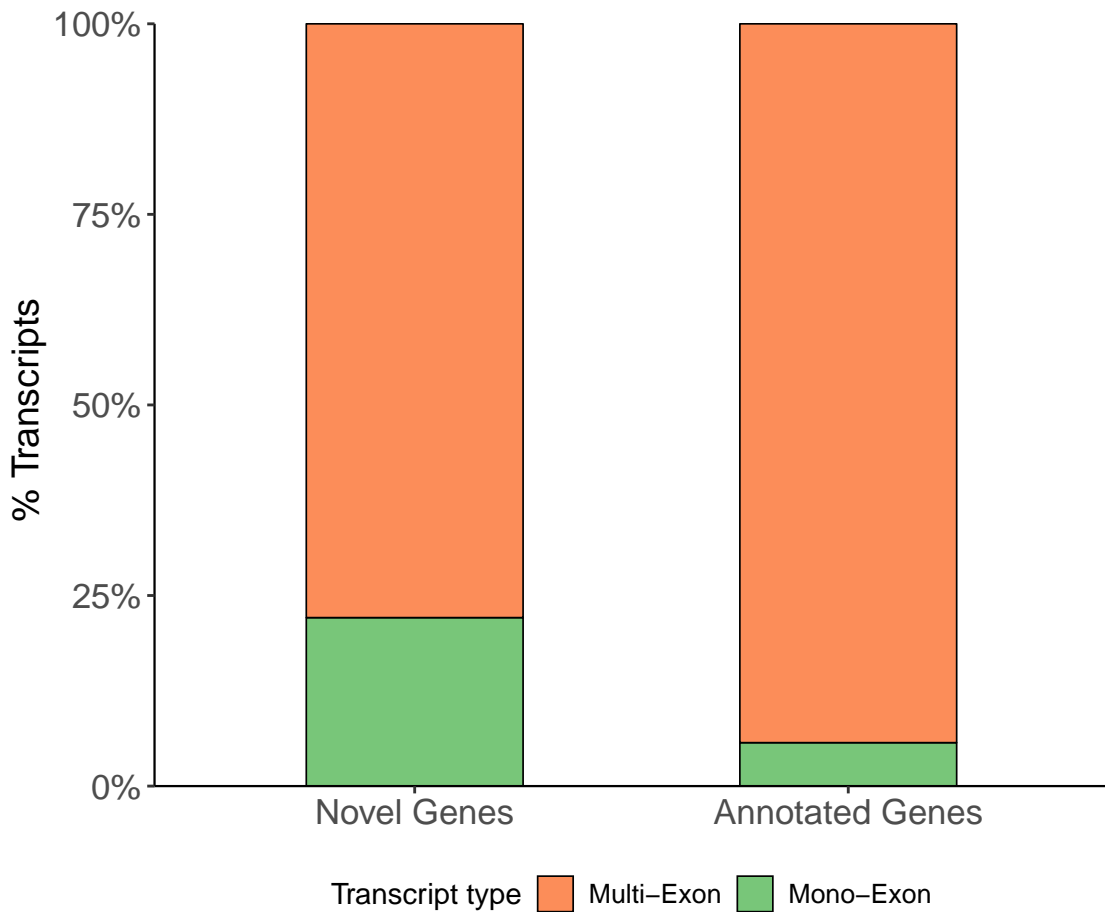
Number of Isoforms per Gene



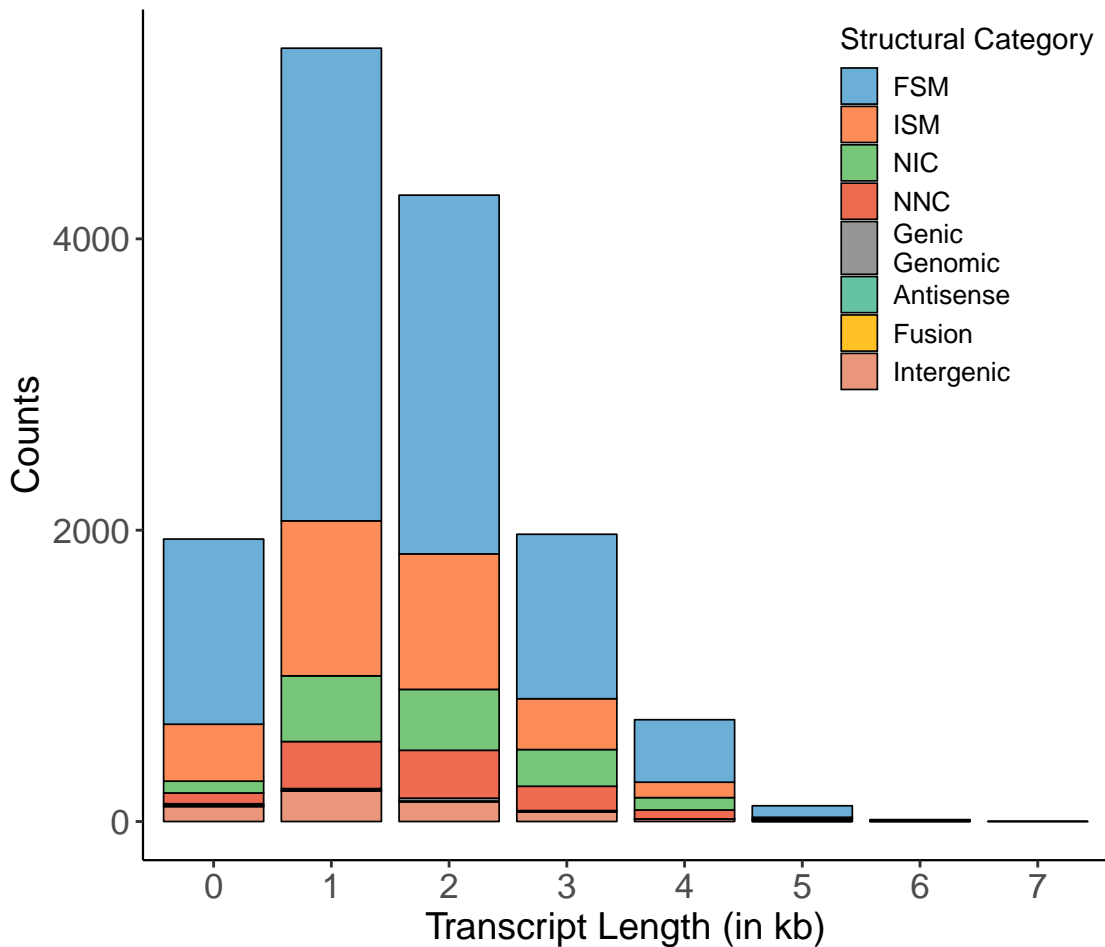
Number of Isoforms per Gene, Known vs Novel Genes



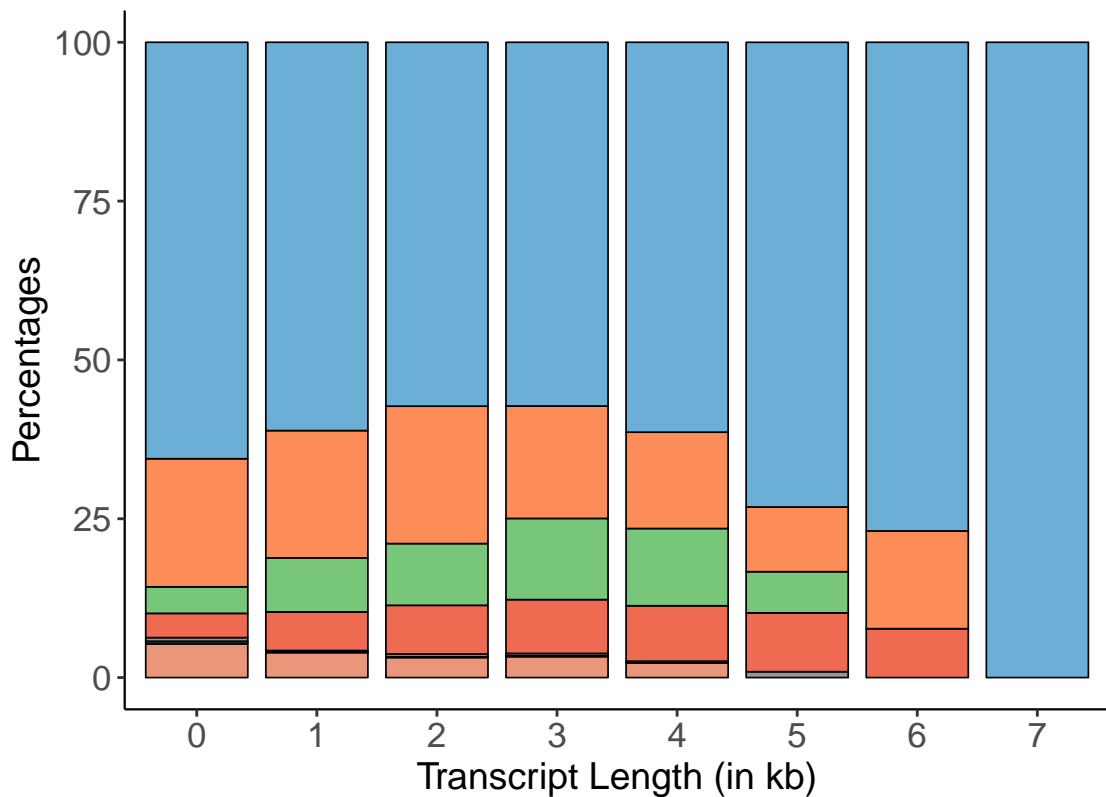
Distribution of Mono- vs Multi-Exon Transcripts



Stacked bar chart showing the distribution of transcript lengths (in kb) across different structural categories. The x-axis represents Transcript Length (in kb) from 0 to 7. The y-axis represents the count of transcripts. The legend includes: FSM (blue), ISM (orange), NIC (green), NNC (red), Genic (grey), Genomic (teal), Antisense (yellow), Fusion (light orange), and Intergenic (light grey).



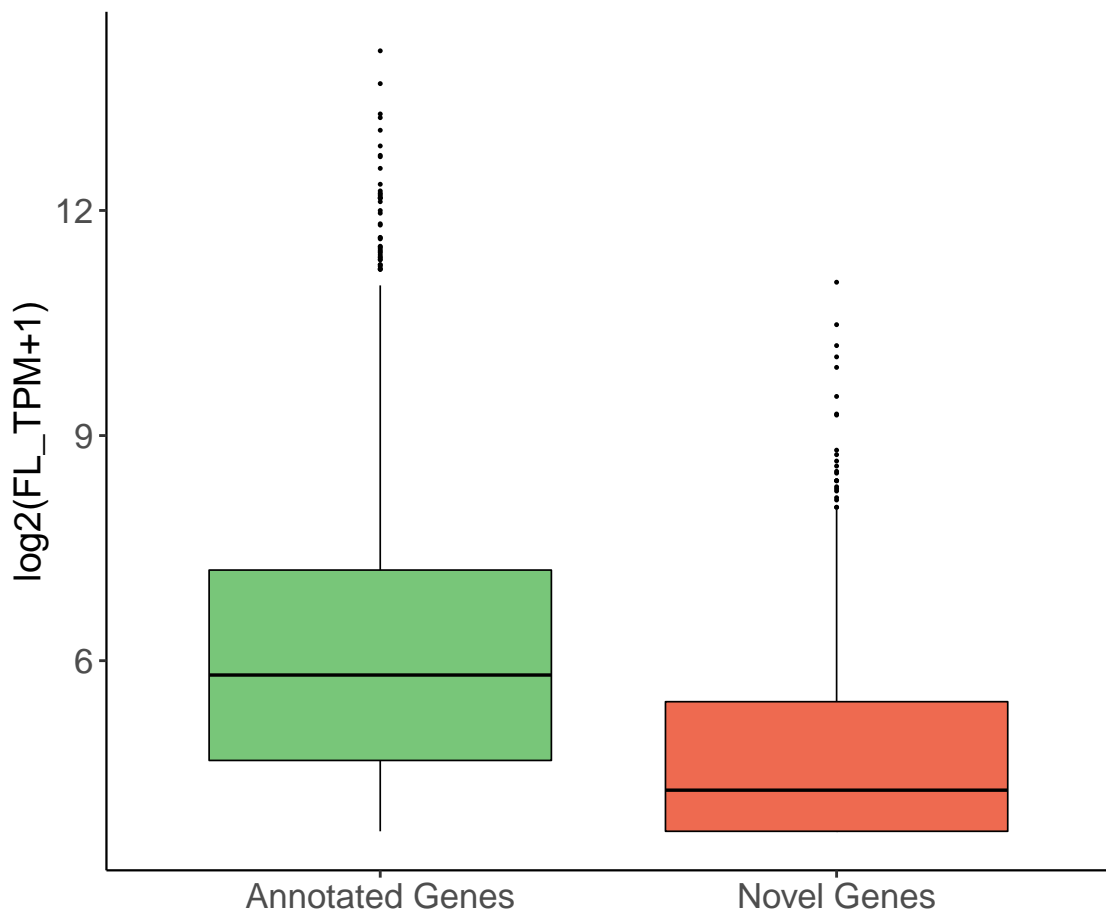
Classifications by Transcript Length, normalized



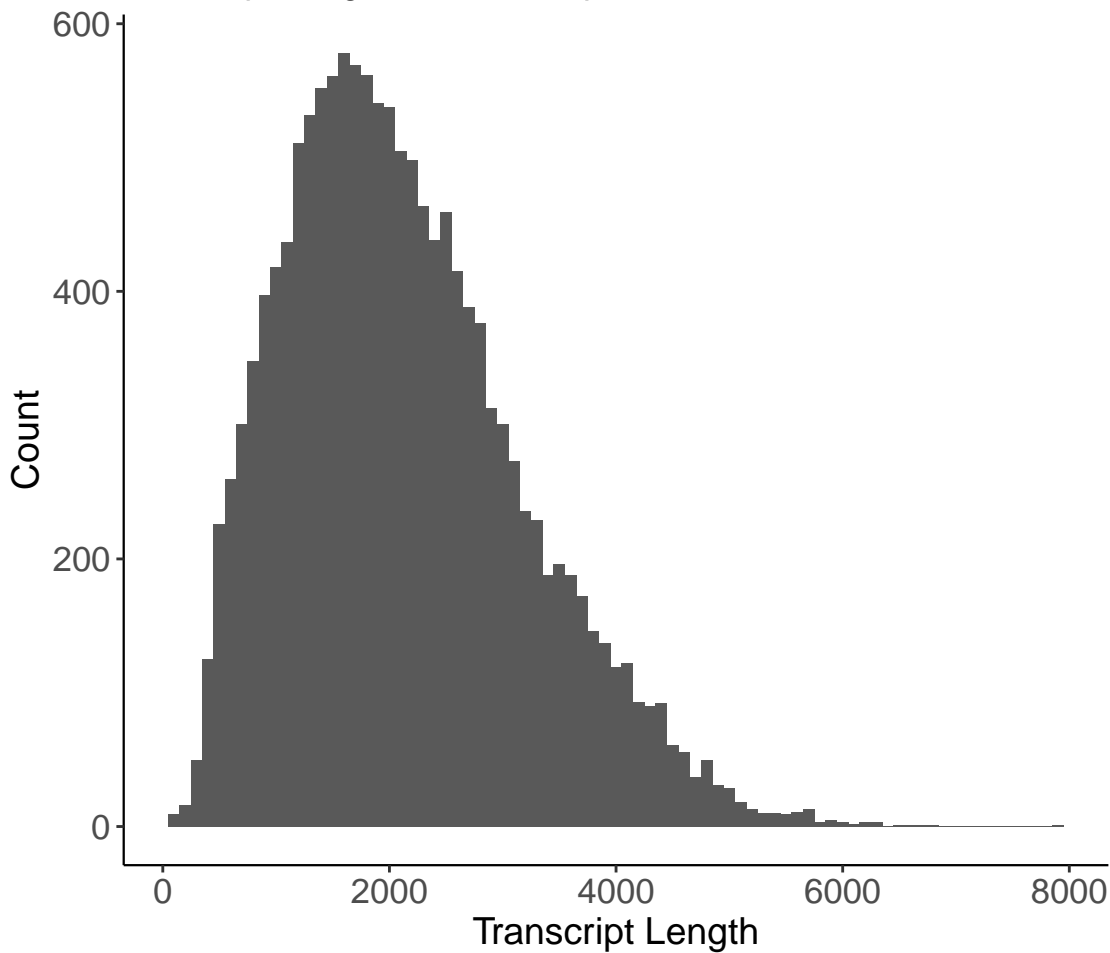
Structural Category

FSM	NIC	Genic	Fusion
ISM	NNC	Genomic	Intergenic
		Antisense	

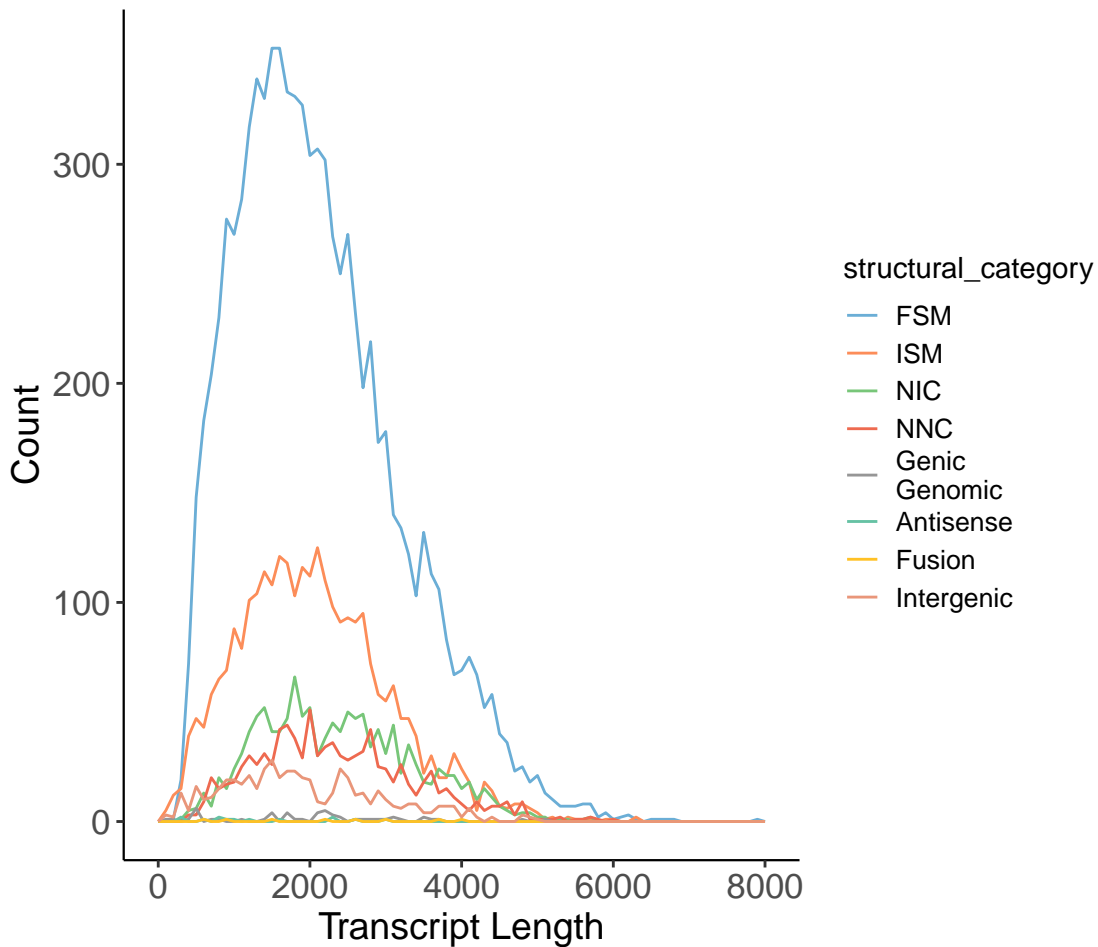
Number of FL reads per Gene by type of gene annotation



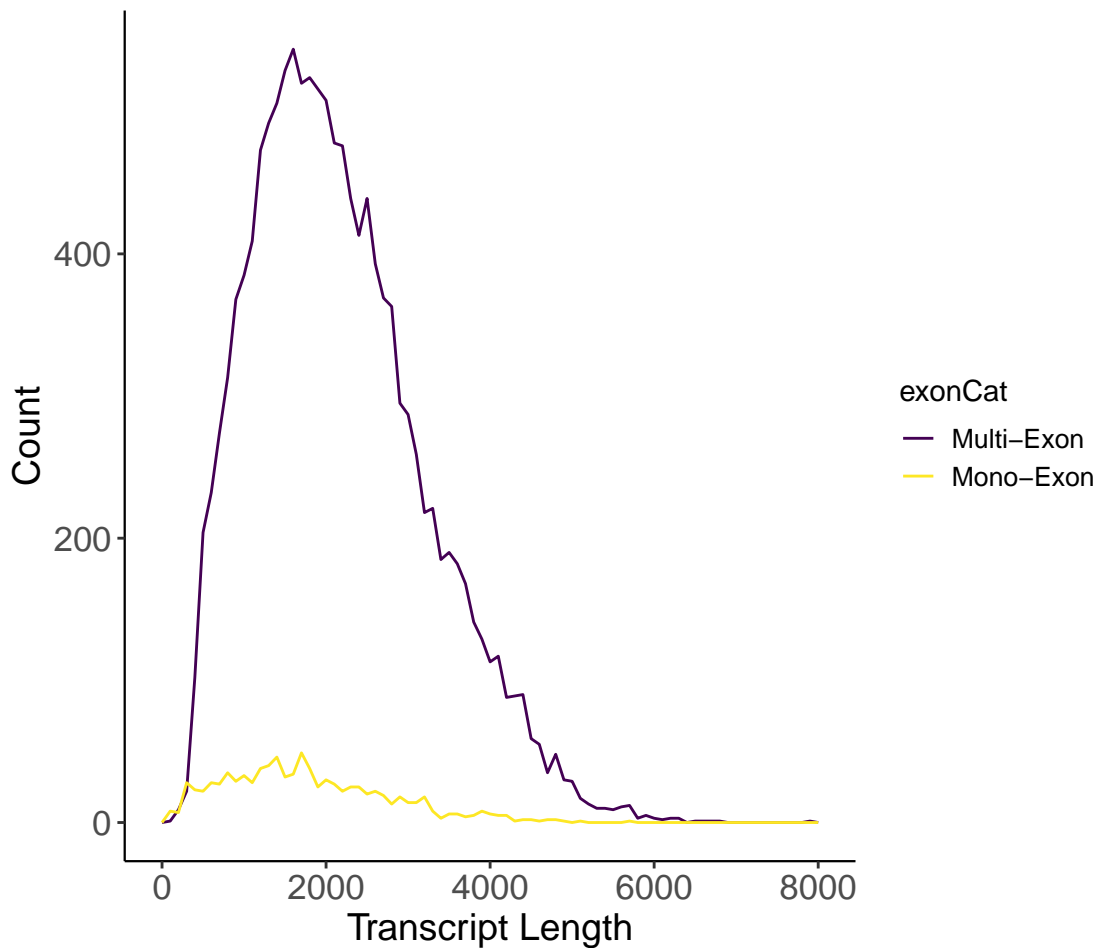
Transcript Lengths, all transcripts



Transcript Lengths, by structural category

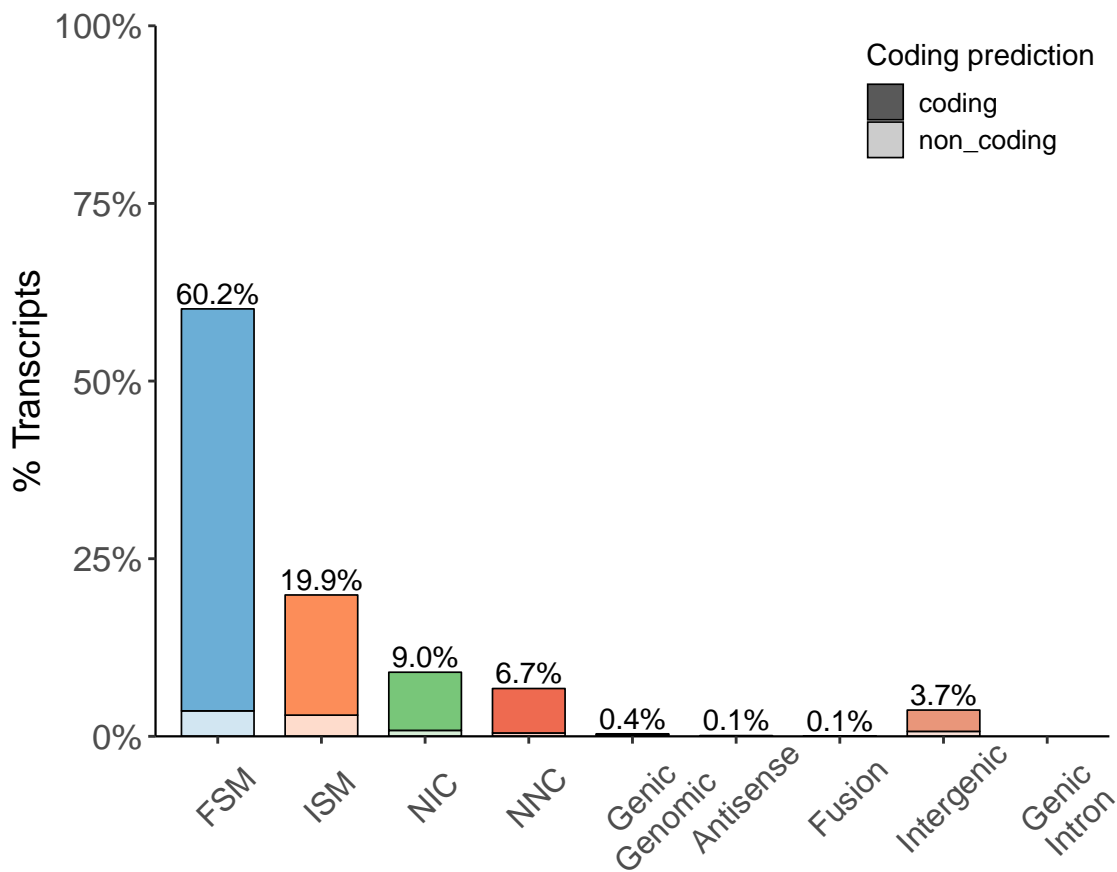


Transcript Lengths, Mono- vs Multi-Exons

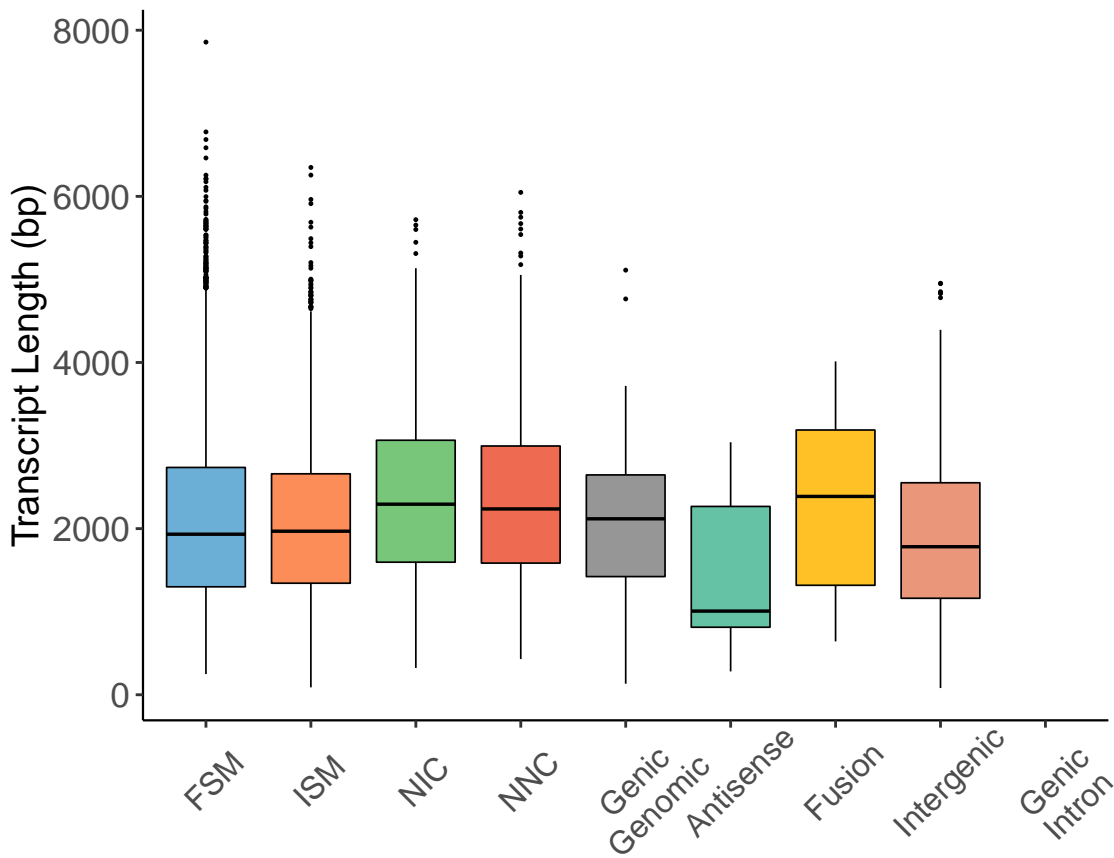


*Structural Isoform Characterization
by Splice Junctions*

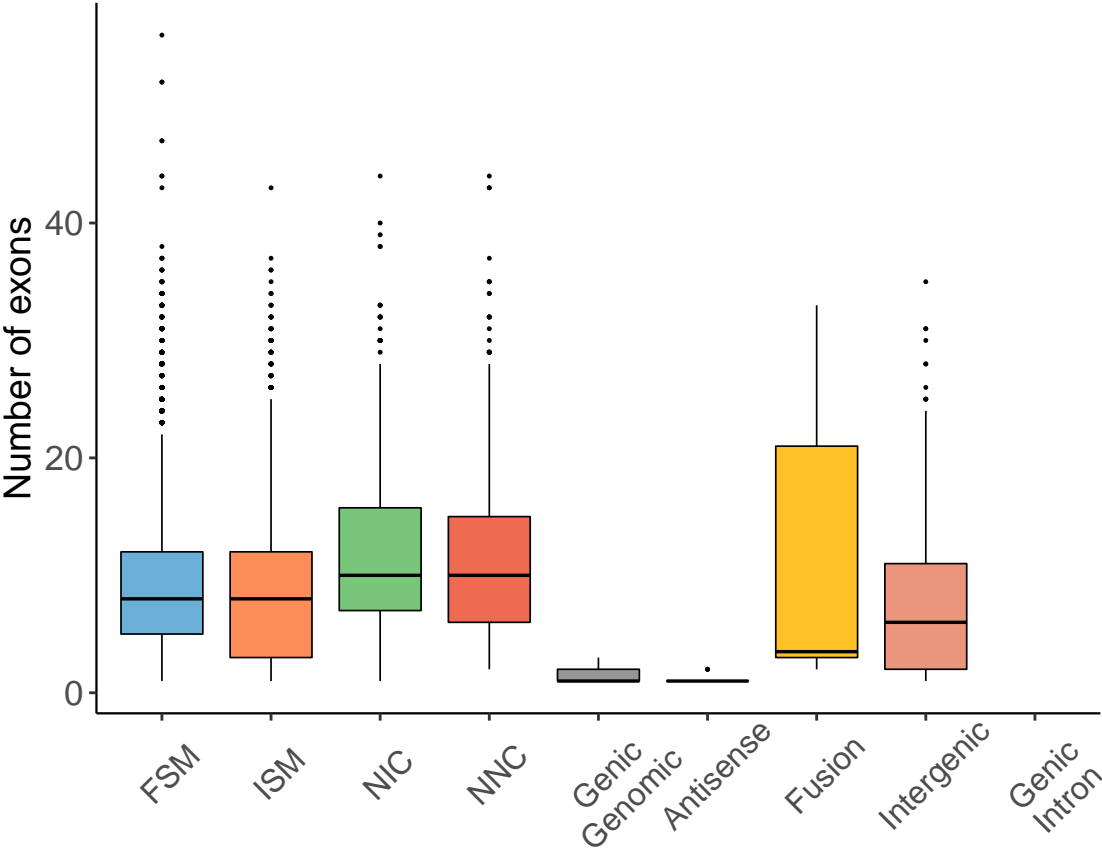
Isoform distribution across structural categories



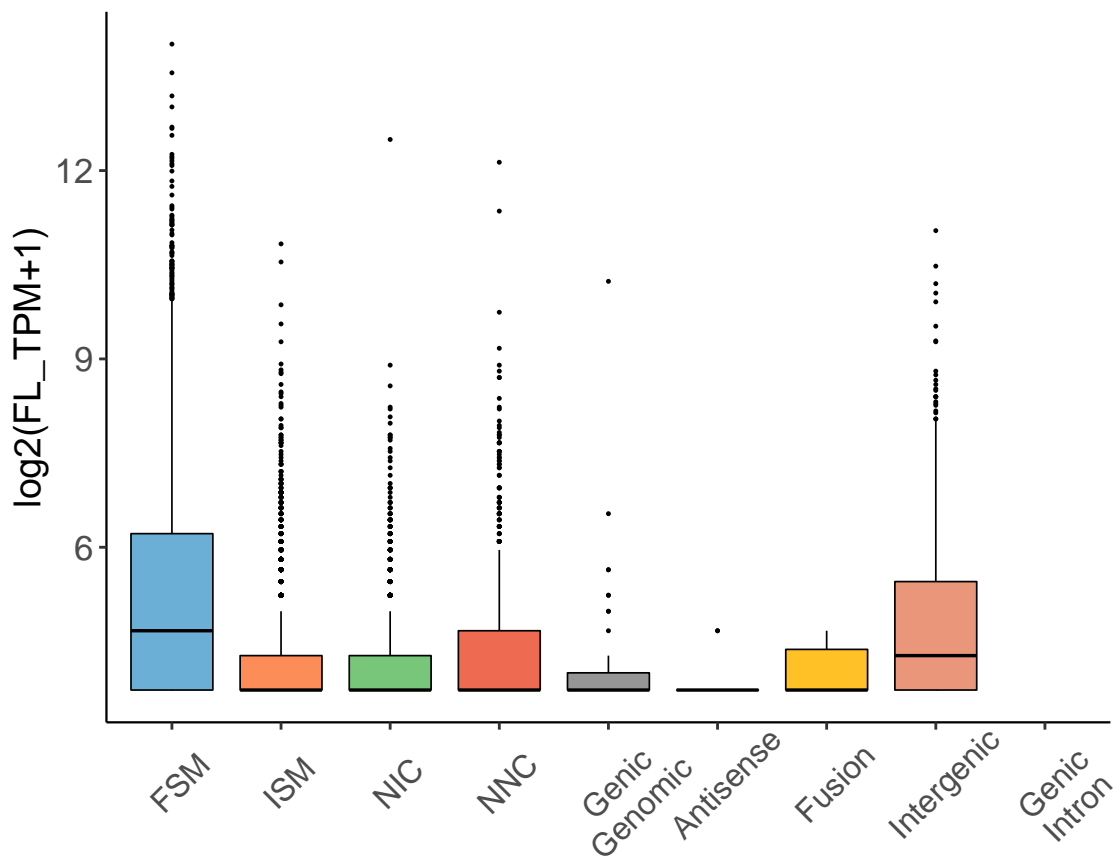
Transcript Lengths by Structural Classification



Exon Counts by Structural Classification

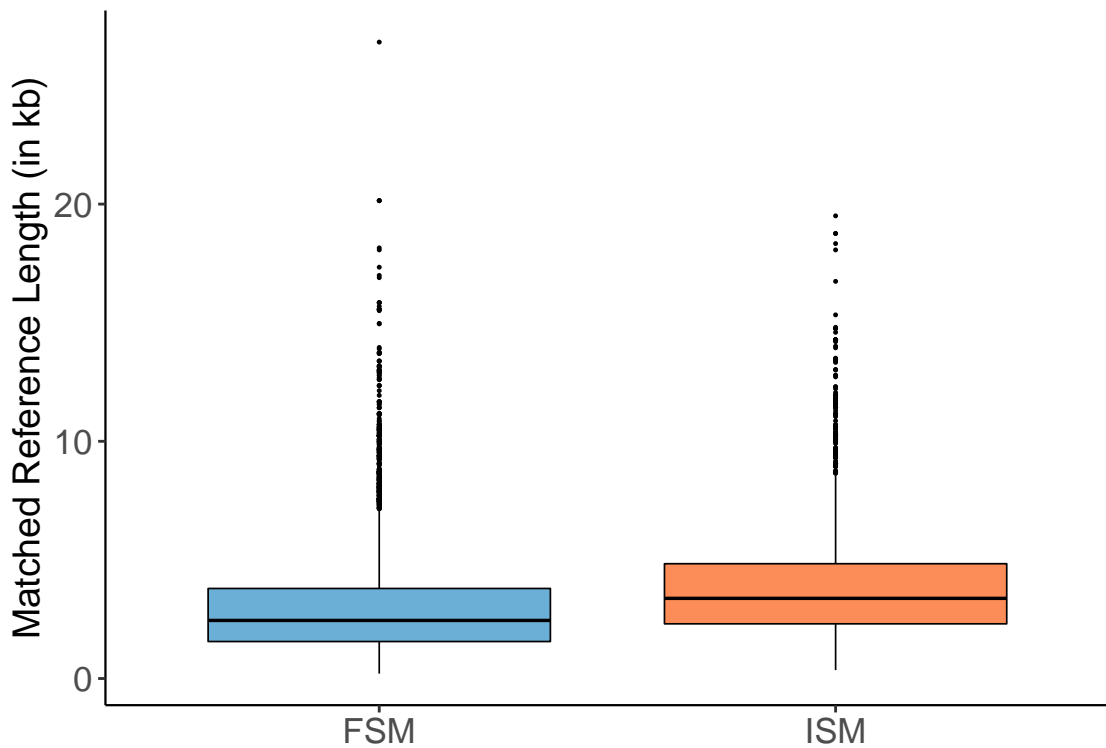


FL Count (normalized) by Structural Category



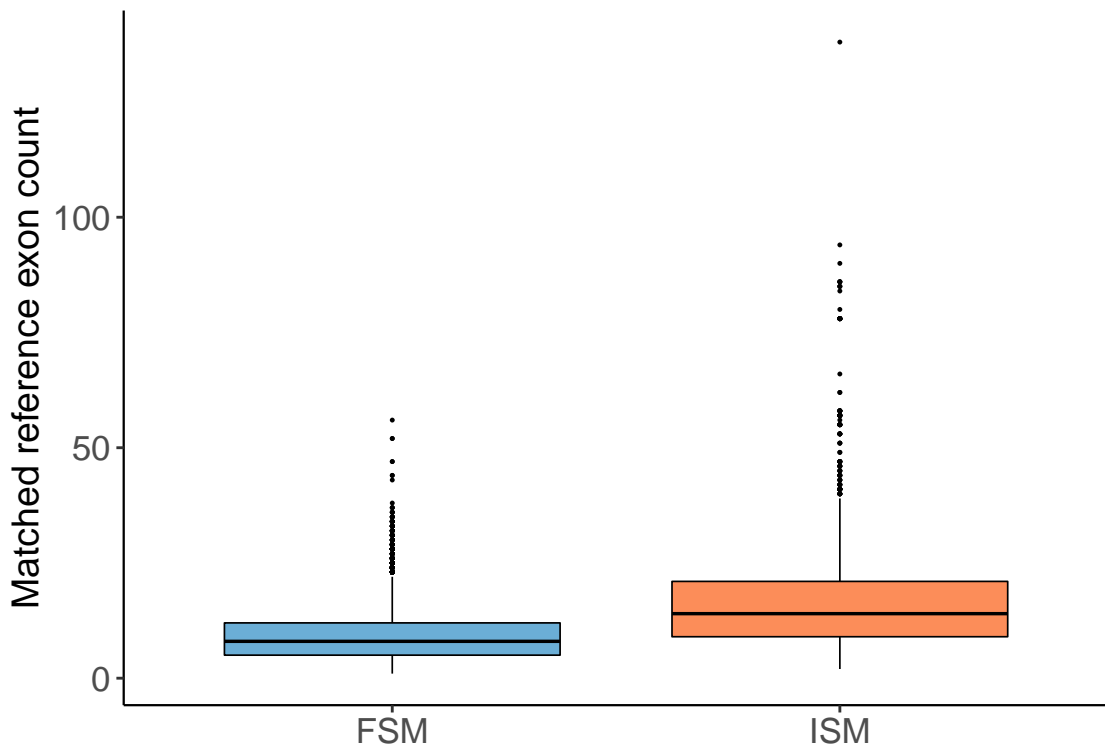
Length Distribution of Matched Reference Transcripts

Applicable only to FSM and ISM categories



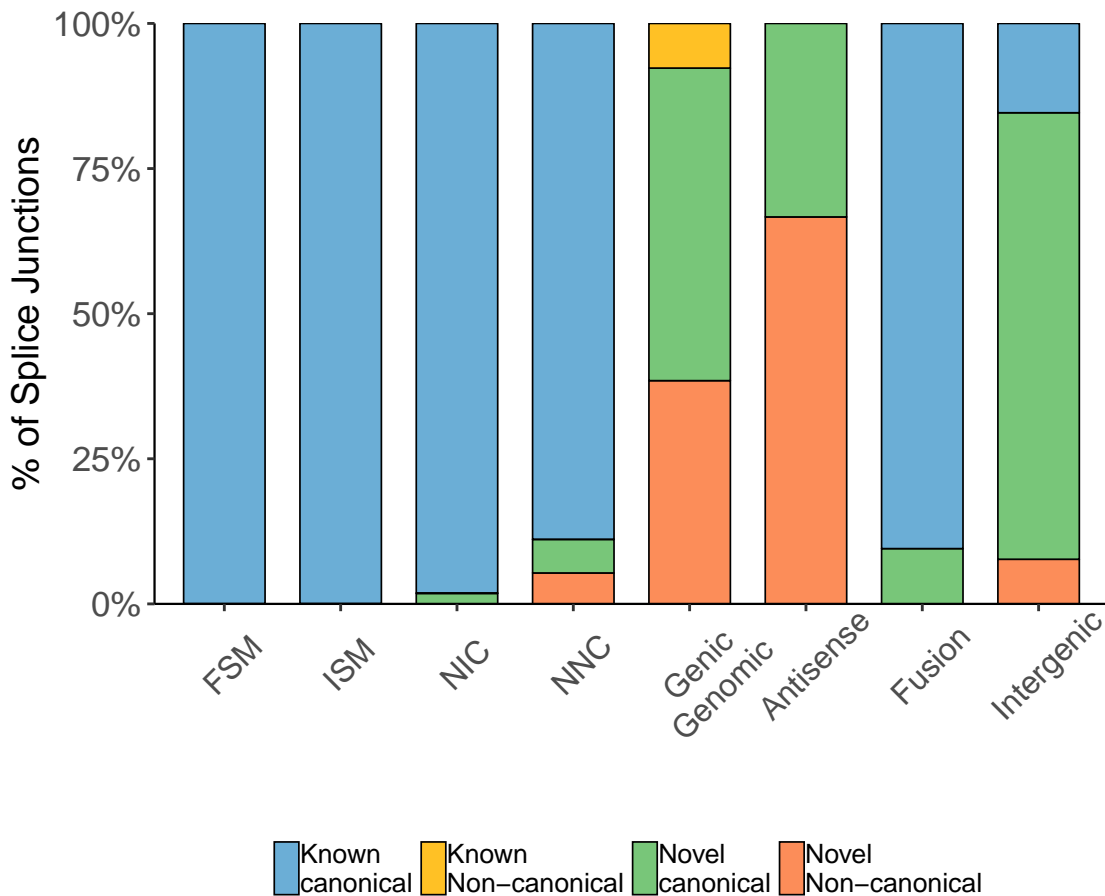
Exon Count Distribution of Matched Reference Transcripts

Applicable only to FSM and ISM categories

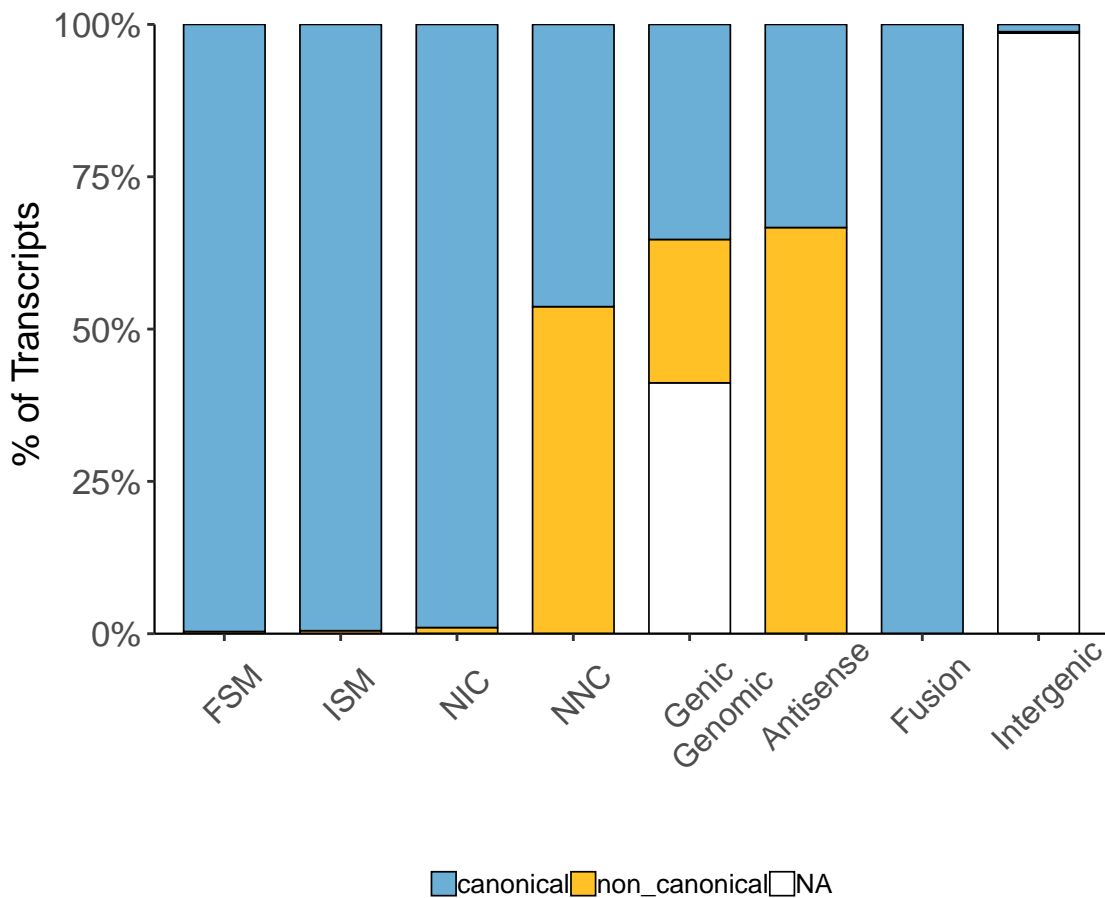


Splice Junction Characterization

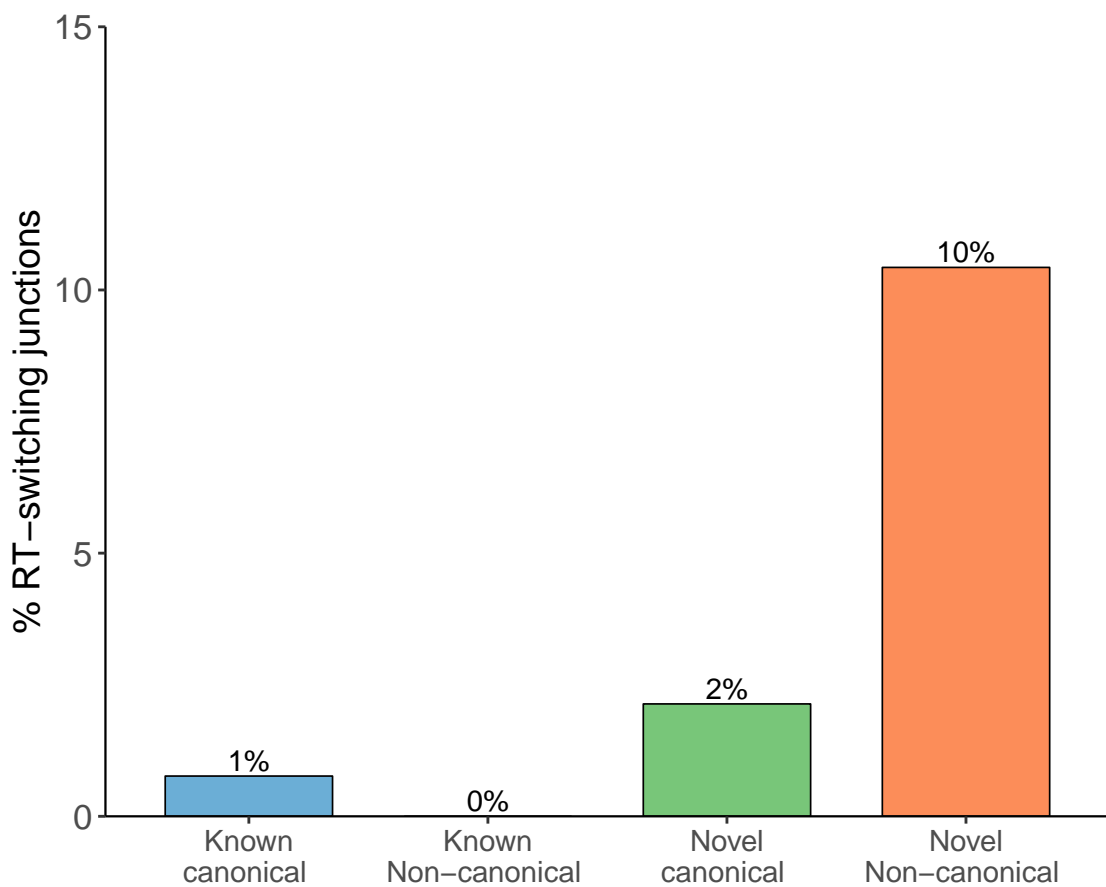
Distribution of Splice Junctions by Structural Classification



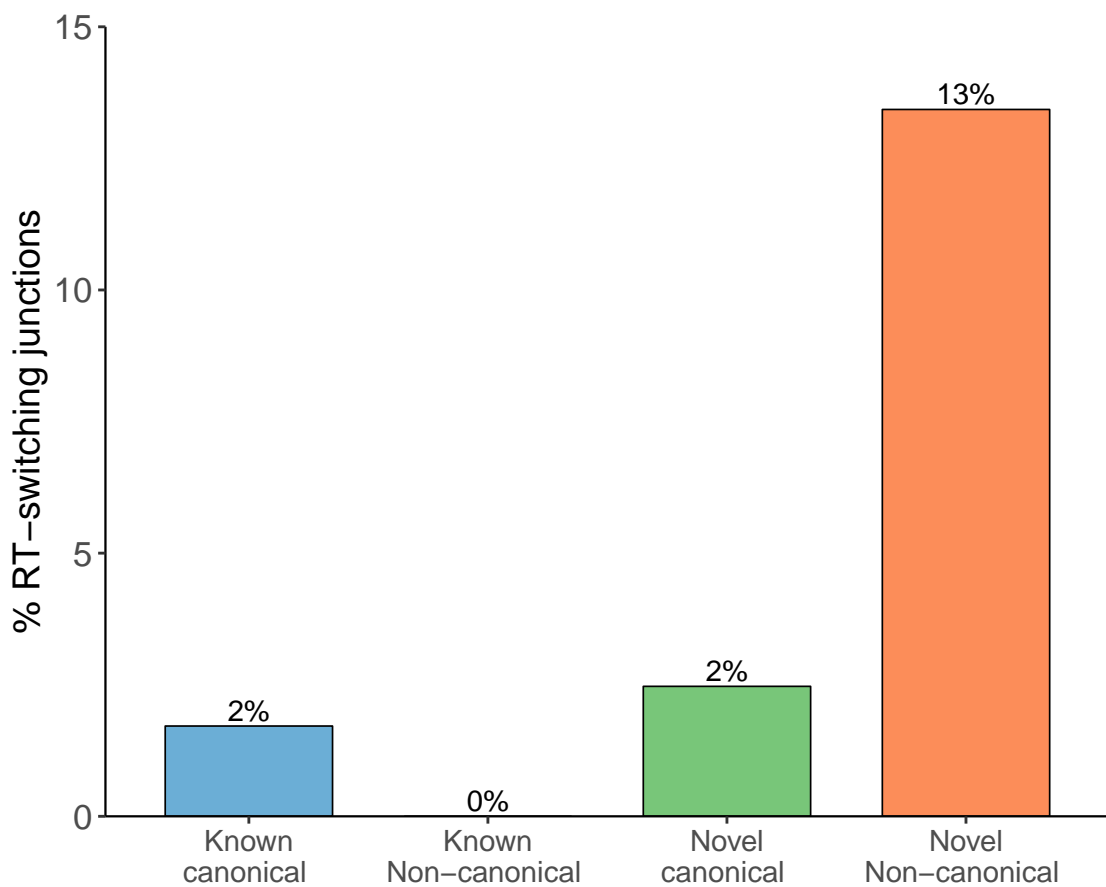
Distribution of Transcripts by Splice Junctions



RT-switching, all junctions



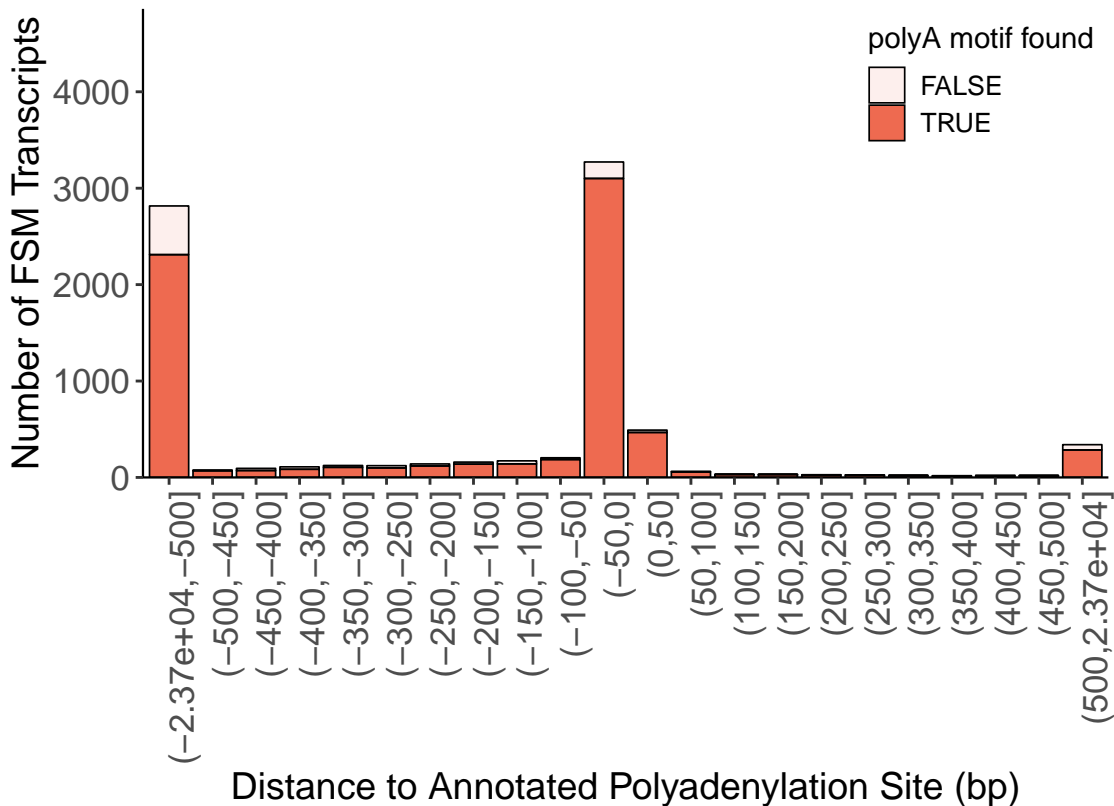
RT-switching, unique junctions



Comparison with Annotated TSS and PolyA Sites

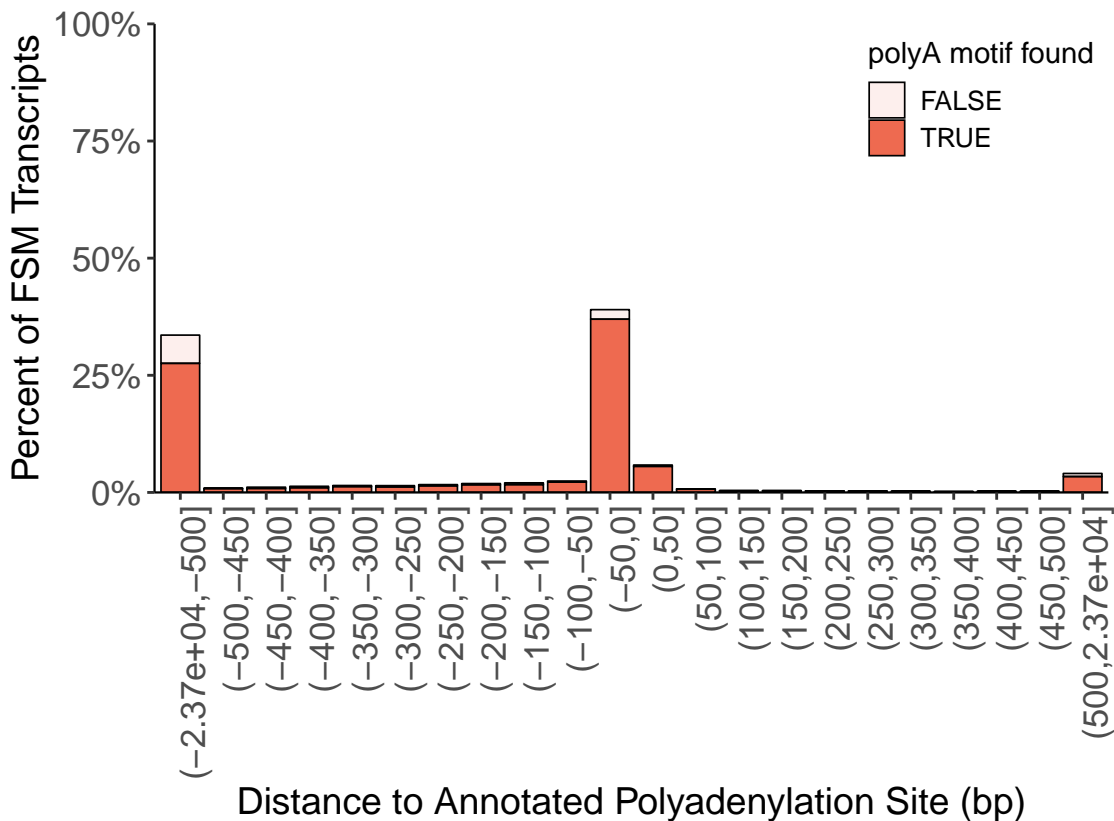
Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



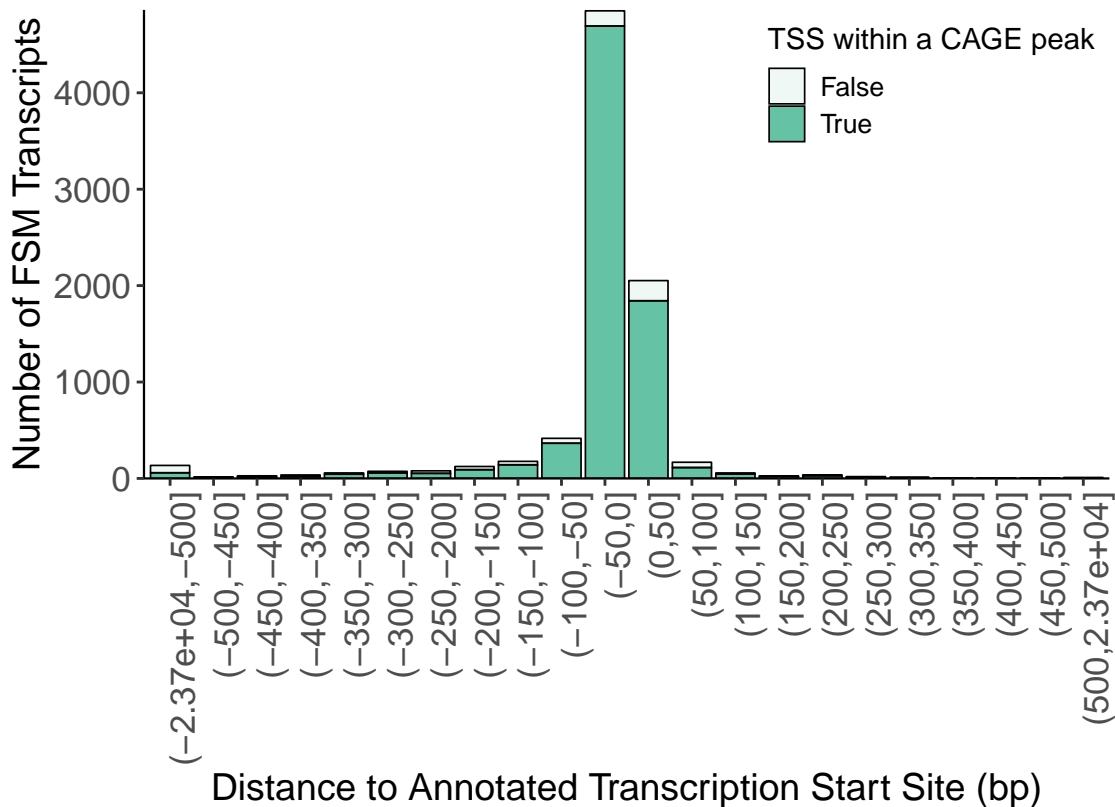
Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



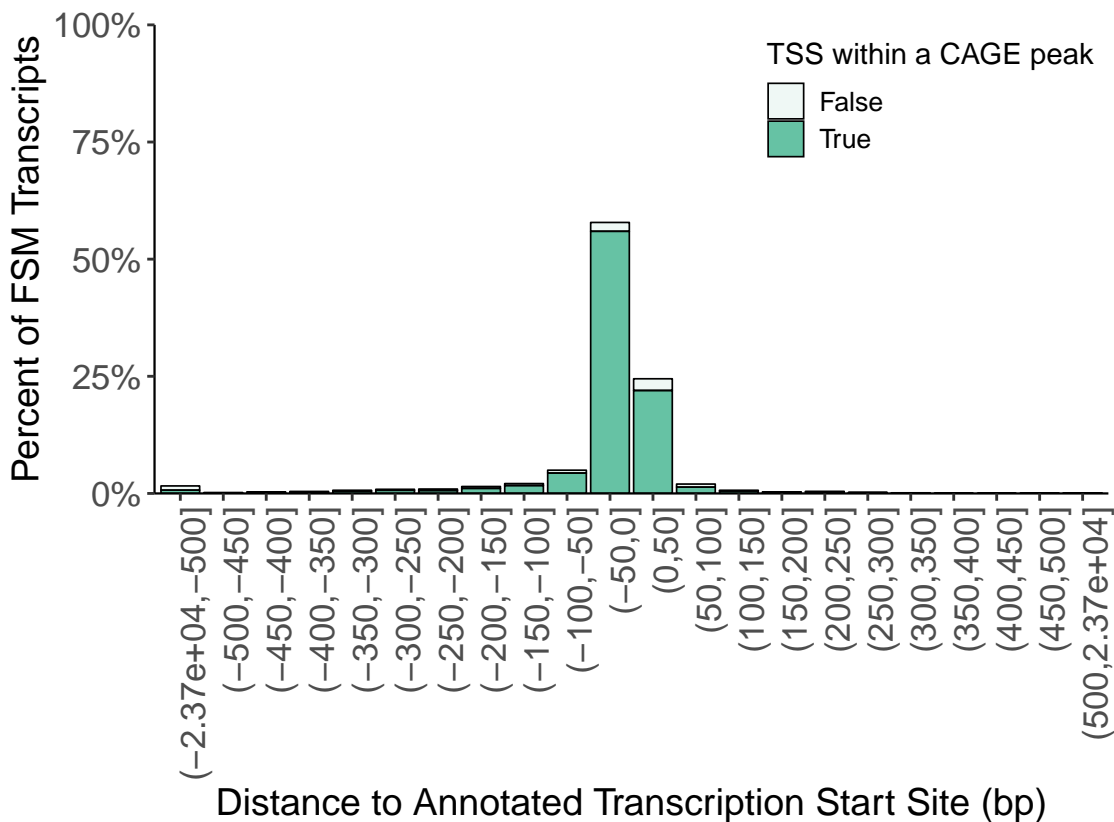
Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



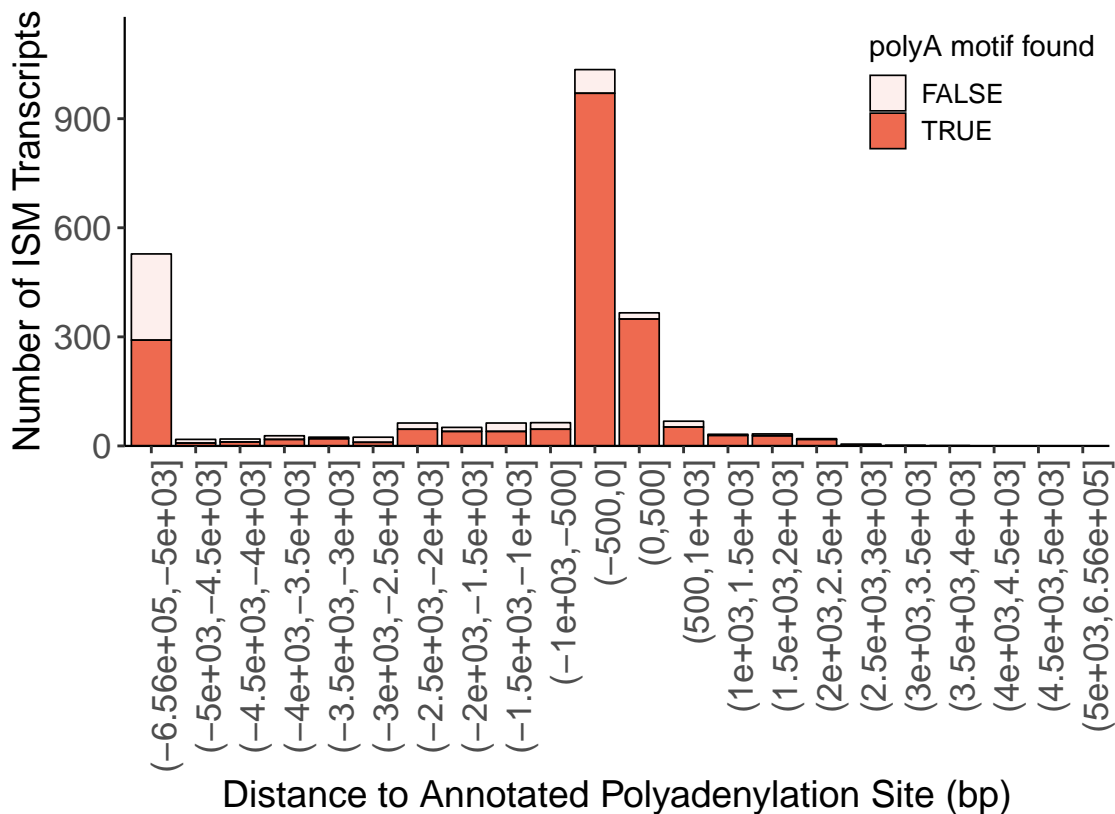
Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



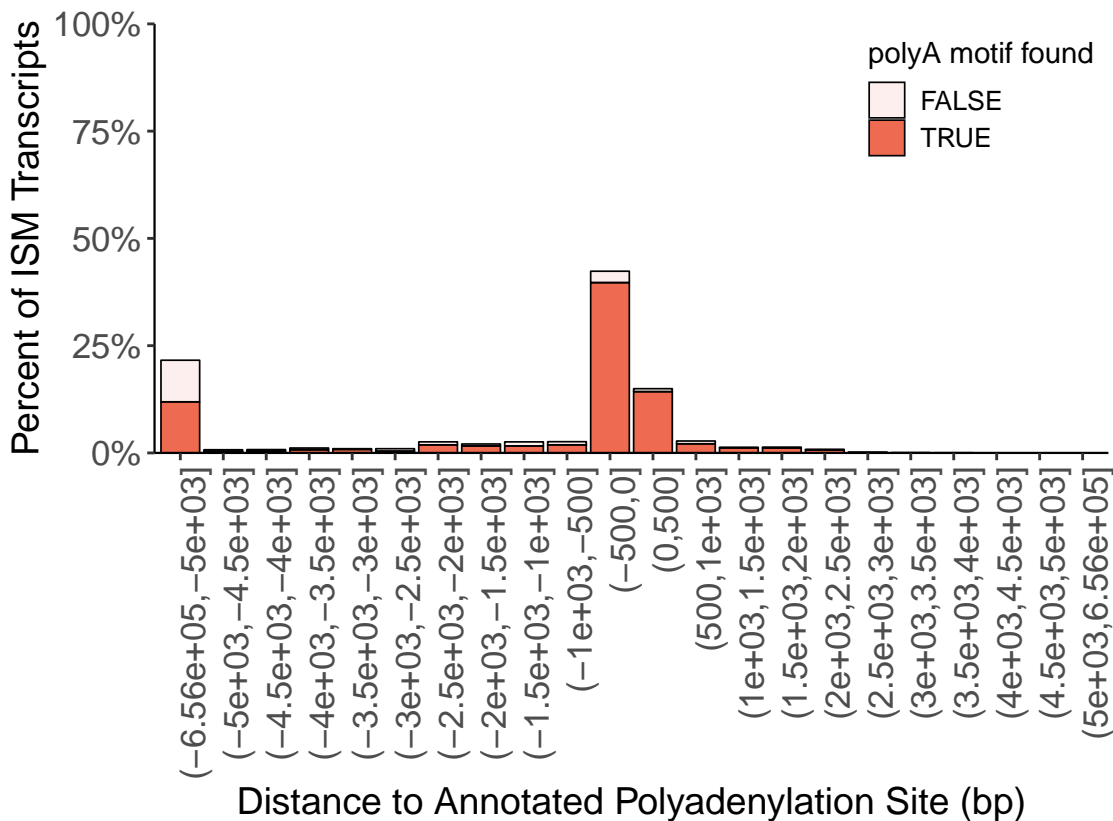
Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



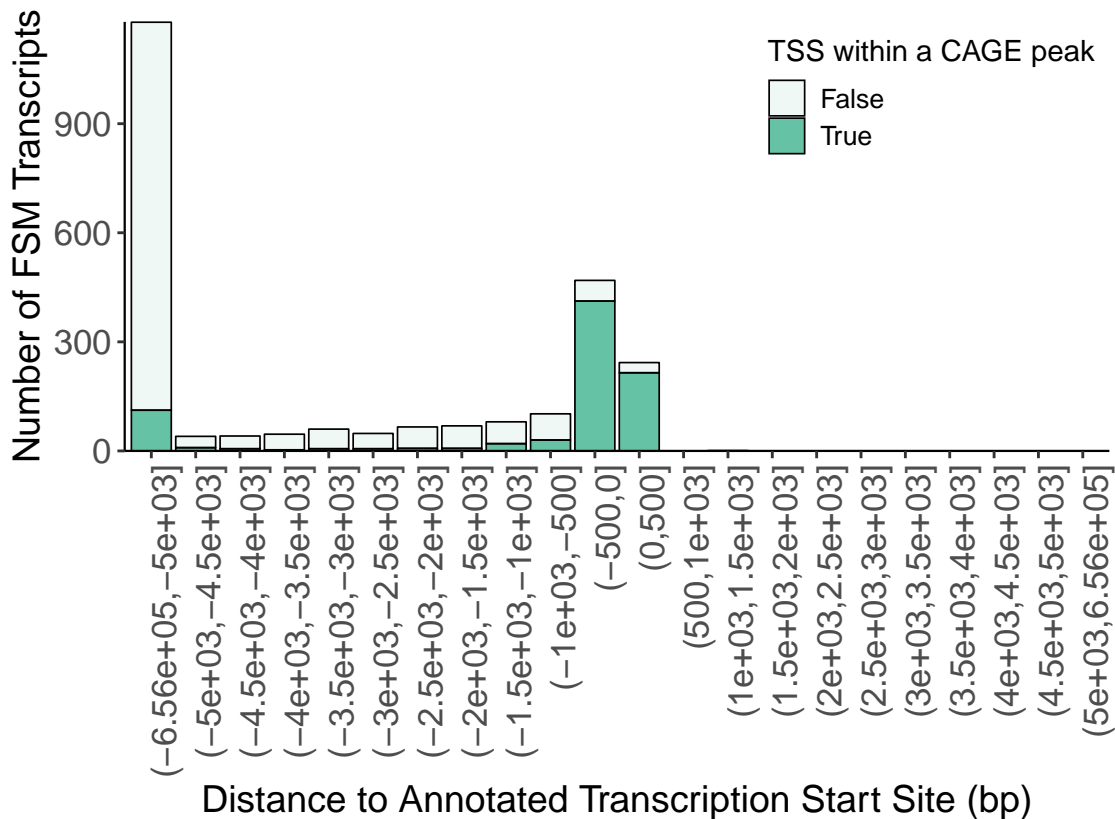
Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



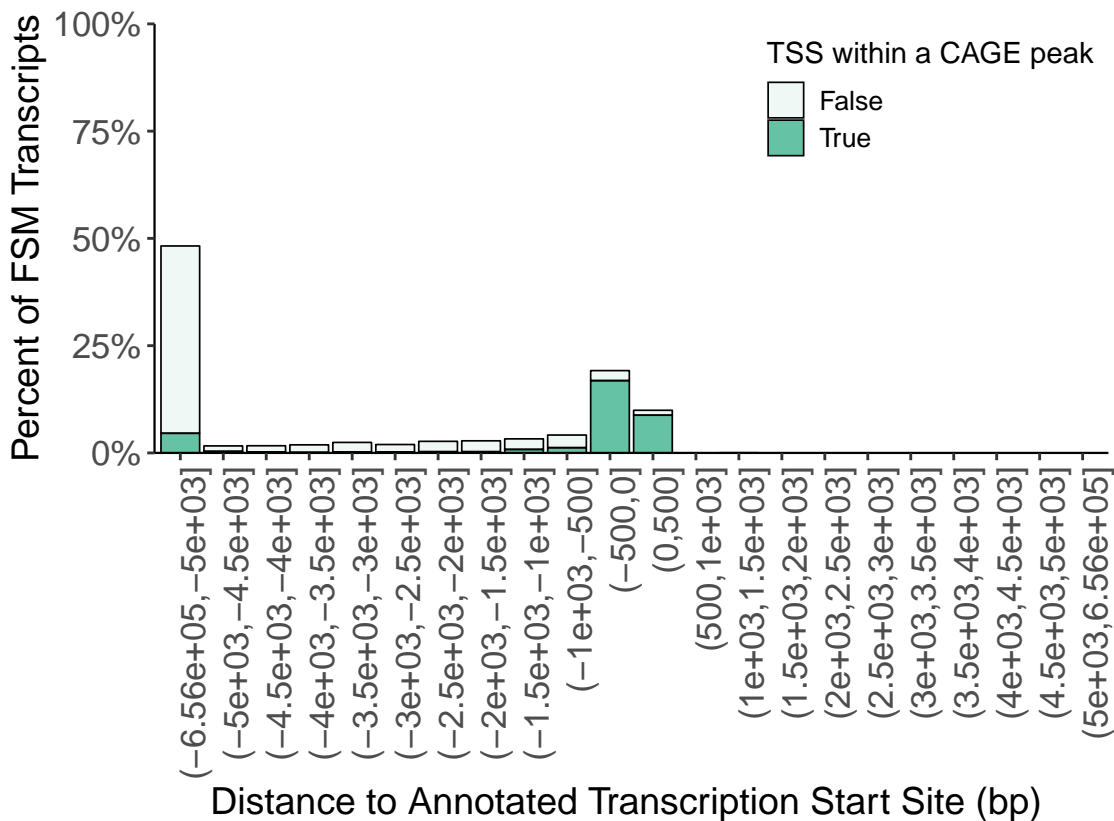
Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS



Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS



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

Frequency of polyA motifs

Number of polyA Motifs Detected

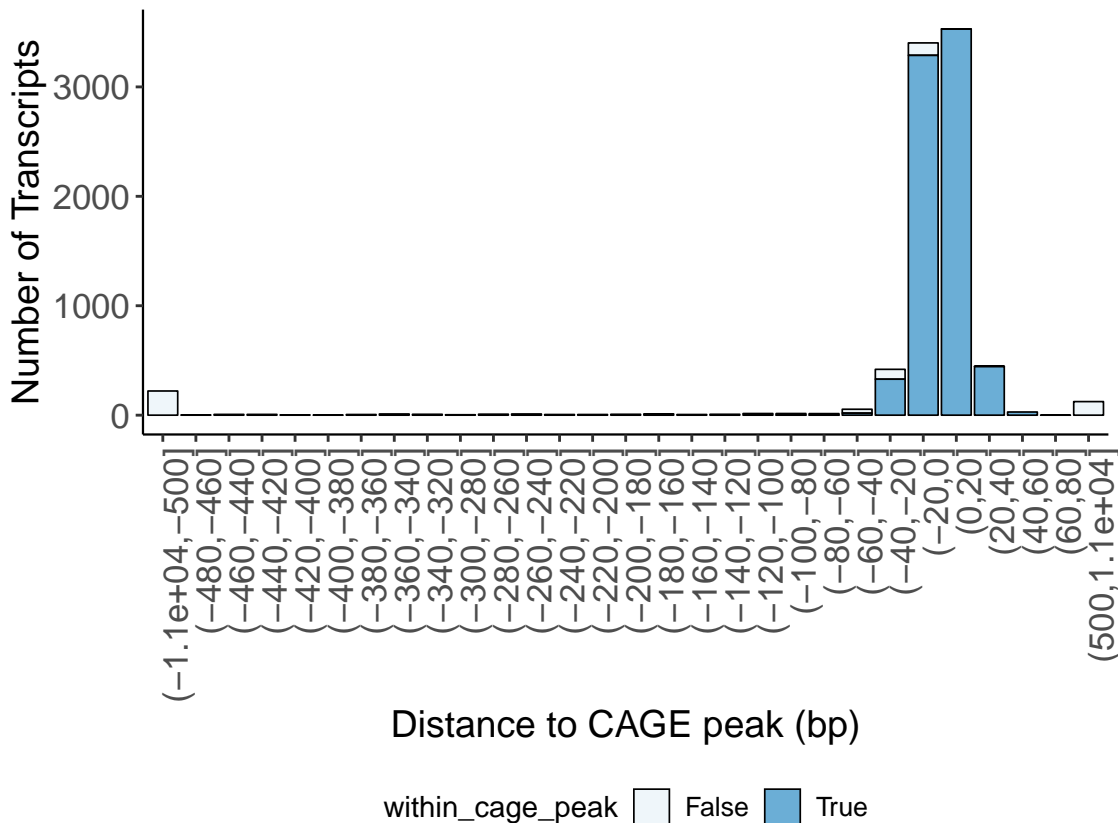
Category	Count	polyA Detected	%
FSM	8627	7541	87
ISM	2853	2365	83
NIC	1294	1109	86
NNC	965	848	88
Genic Genomic	51	37	73
Antisense	13	4	31
Fusion	8	7	88
Intergenic	530	410	77

Motif	Count	%
AATAAA	8160	66.2
ATTAAA	2015	16.4
TATAAA	347	2.8
AGTAAA	344	2.8
AAGAAA	203	1.6
AATATA	171	1.4
CATAAA	165	1.3
AAAAAG	163	1.3
AATACA	152	1.2
GATAAA	141	1.1
TTTAAA	116	0.9
AATGAA	115	0.9
AAAACA	79	0.6
ACTAAA	74	0.6
AATAGA	67	0.5
GGGGCT	9	0.1

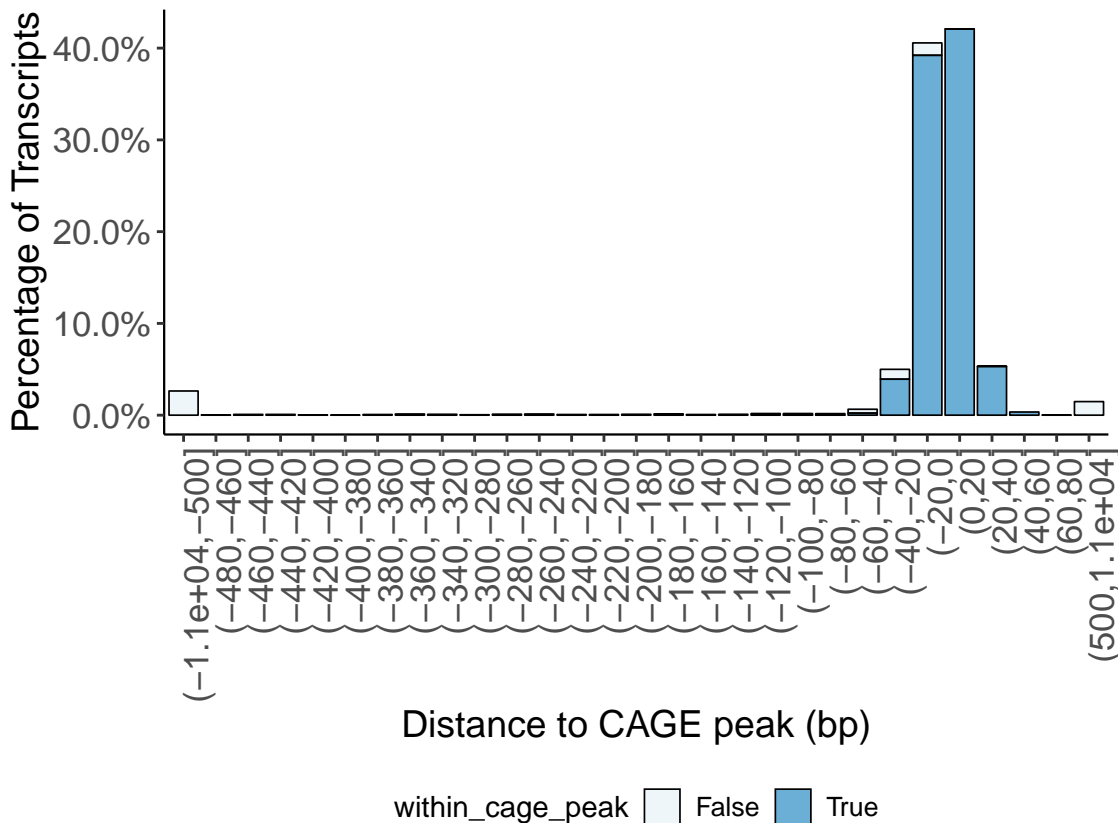
CAGE distances analysis

Distance to CAGE peak of multi-exonic FSM

Negative values indicate downstream of annotated CAGE peak

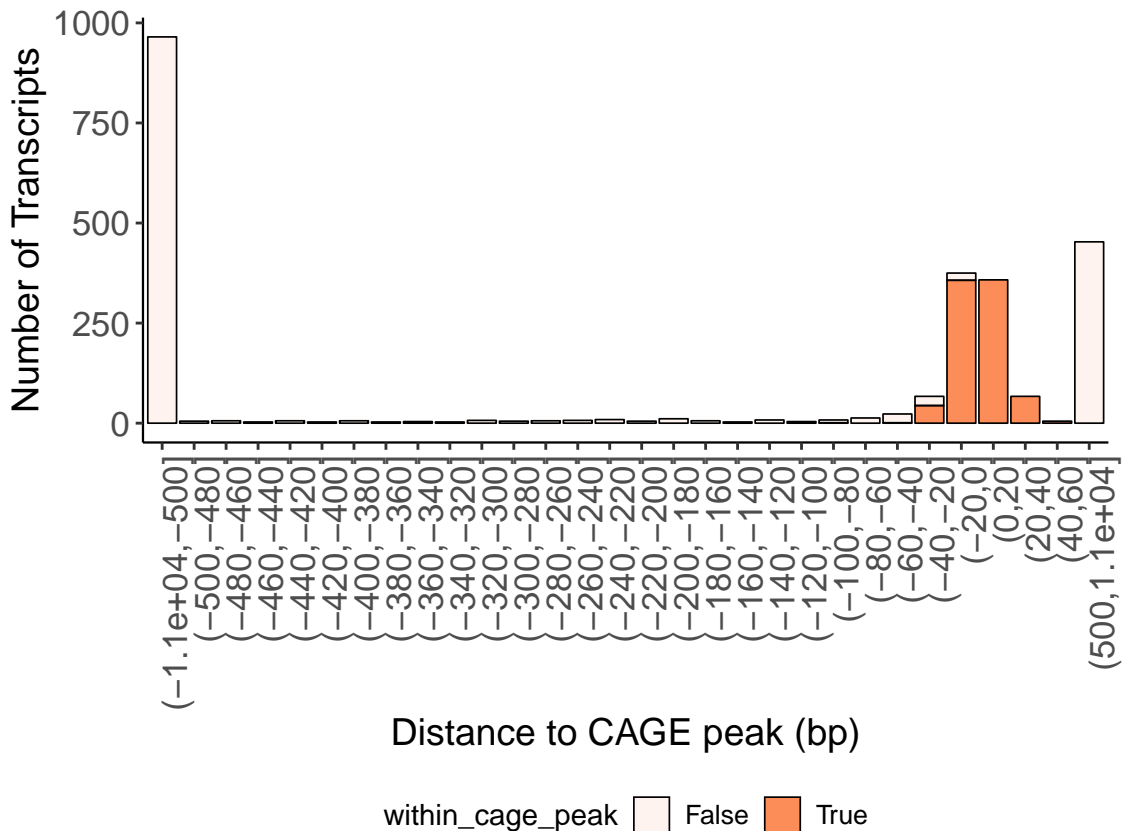


Negative values indicate downstream of annotated CAGE peak



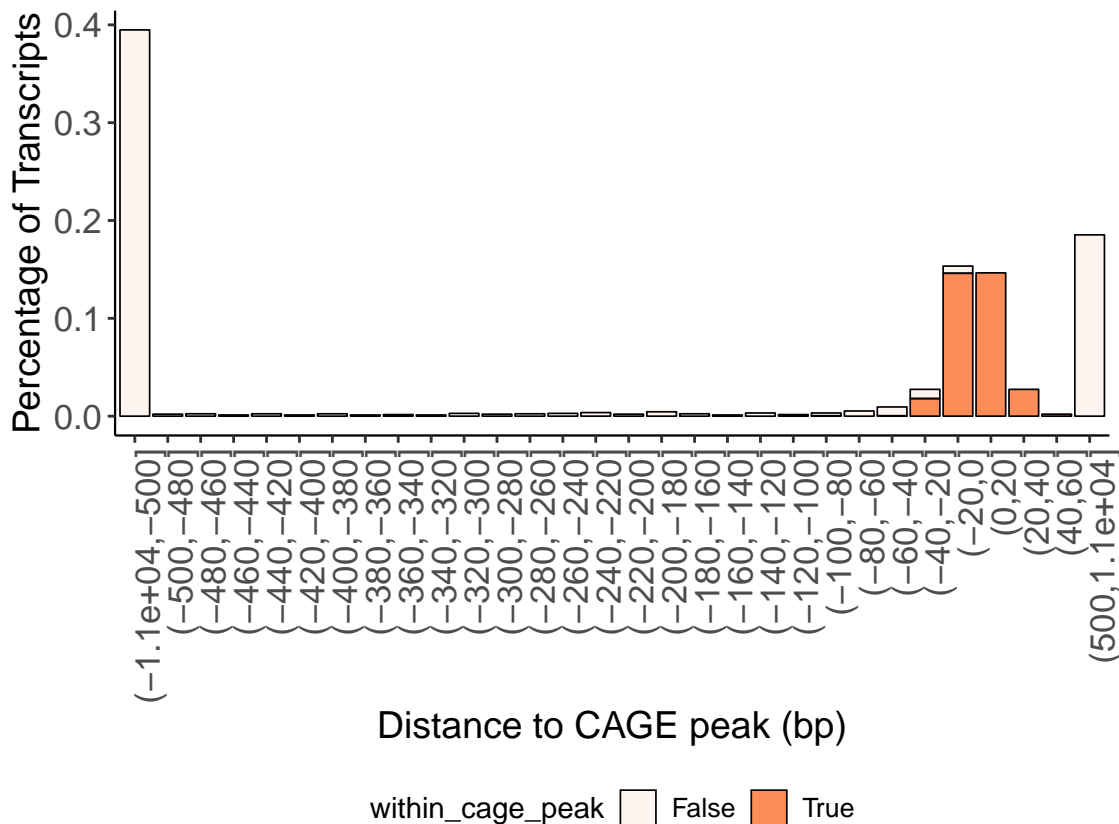
Distance to CAGE peak of multi-exonic ISM

Negative values indicate downstream of annotated CAGE peak

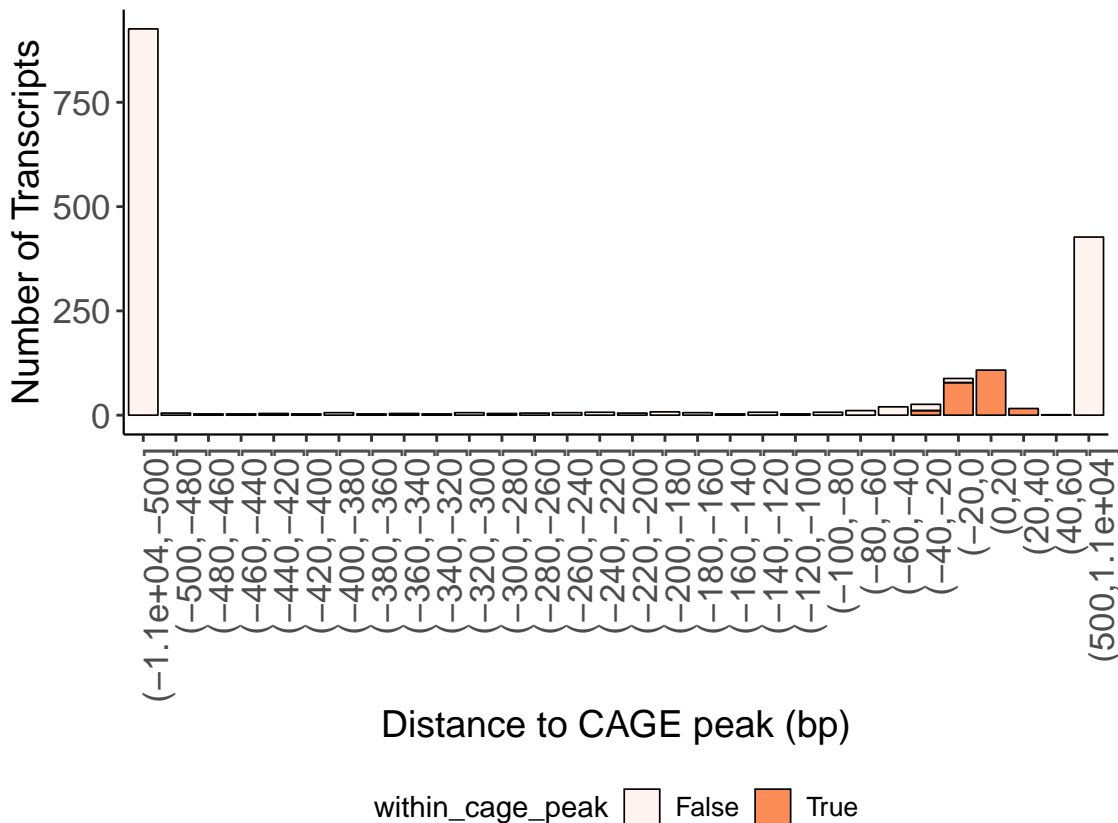


Distance to CAGE peak of multi-exonic ISM

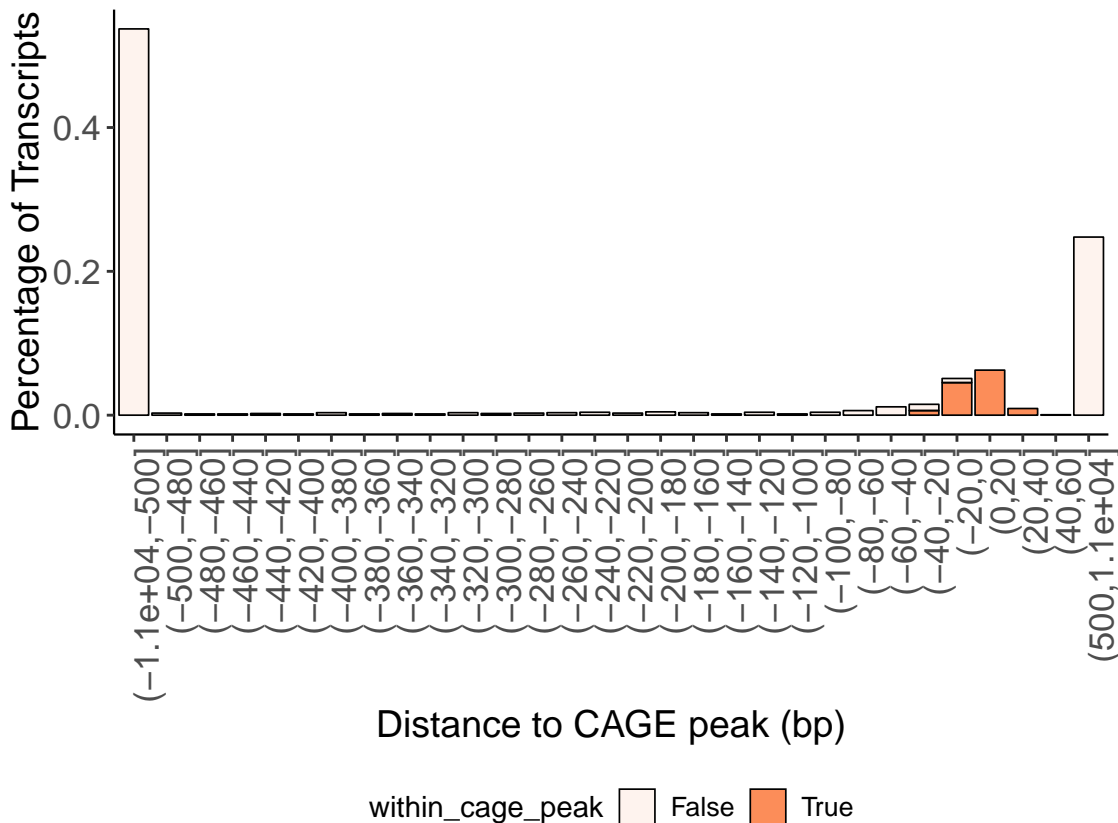
Negative values indicate downstream of annotated CAGE peak



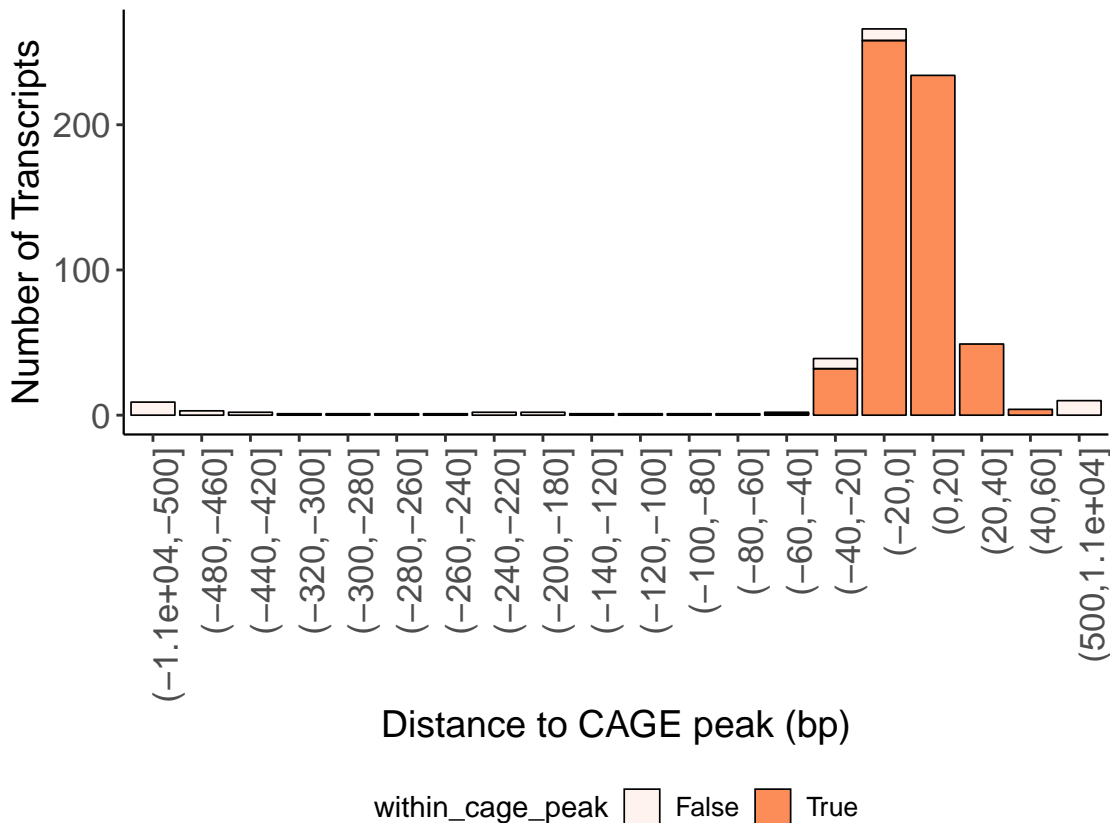
Distance to CAGE peak of multi-exonic ISM 3prime fragments
Negative values indicate downstream of annotated CAGE peak



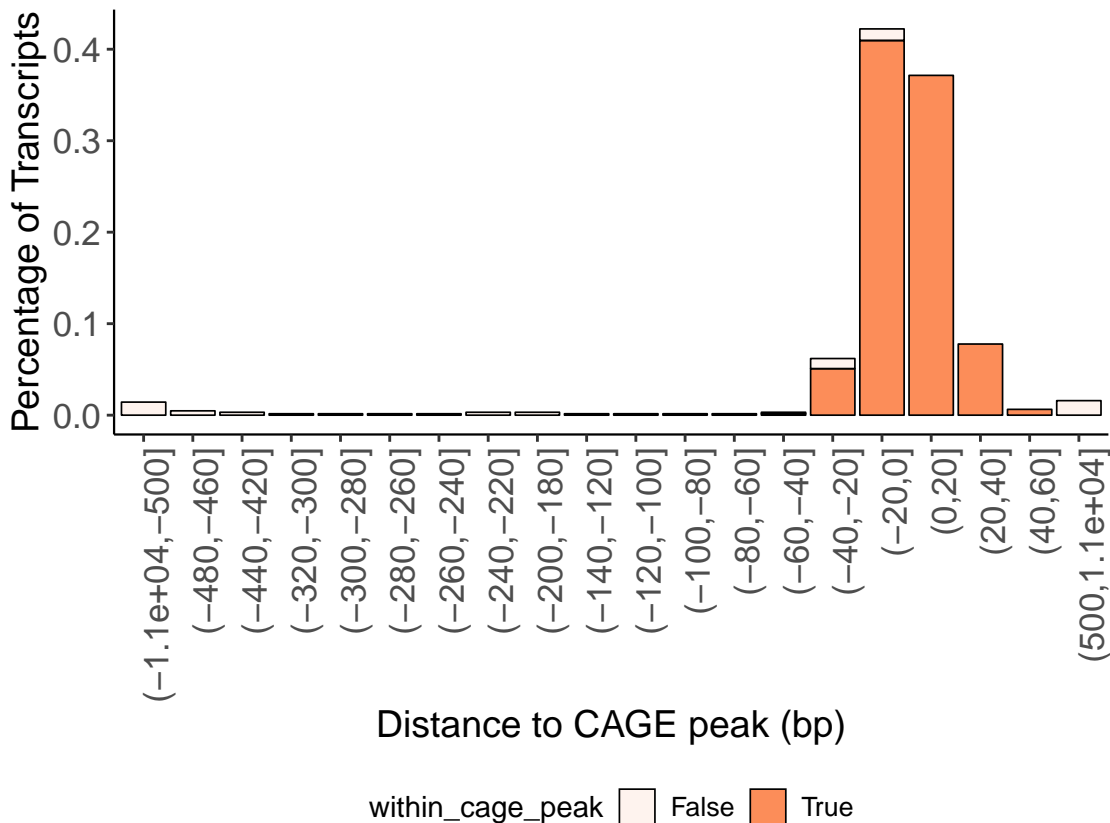
Distance to CAGE peak of multi-exonic ISM 3prime fragments
Negative values indicate downstream of annotated CAGE peak



Distance to CAGE peak of multi-exonic ISM 5prime fragments
Negative values indicate downstream of annotated CAGE peak

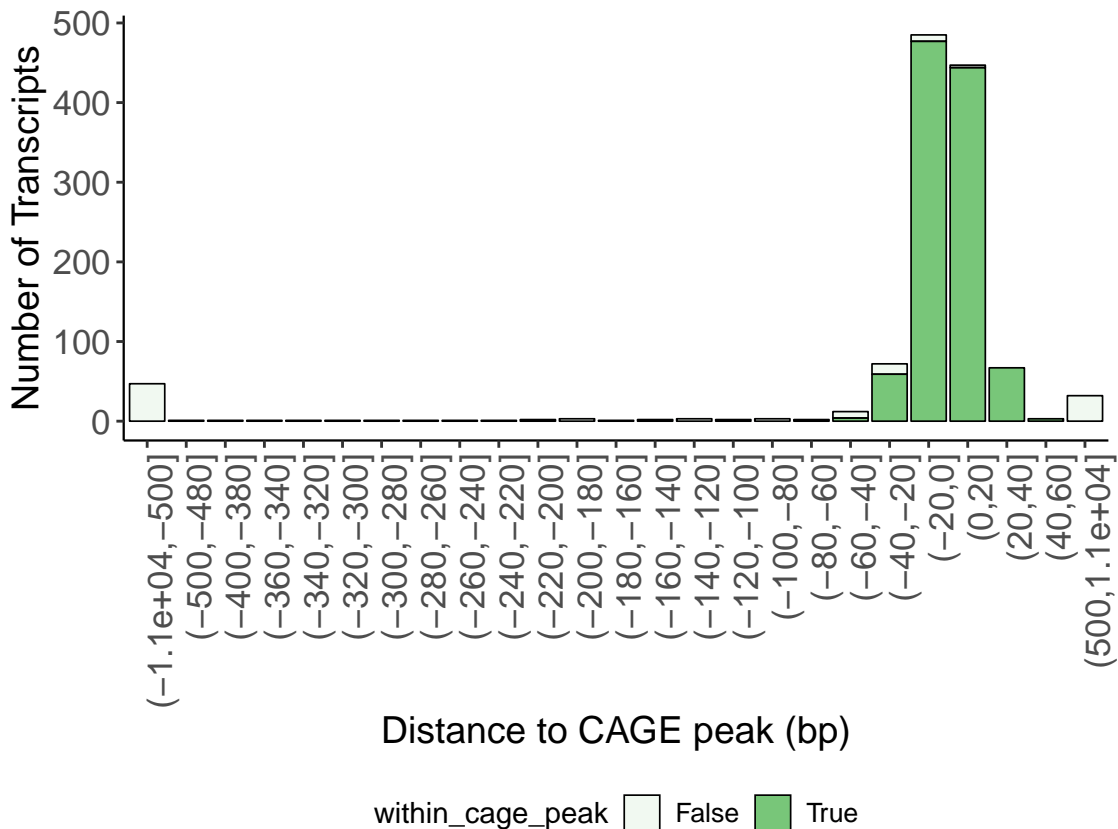


Distance to CAGE peak of multi-exonic ISM 5prime fragments
Negative values indicate downstream of annotated CAGE peak



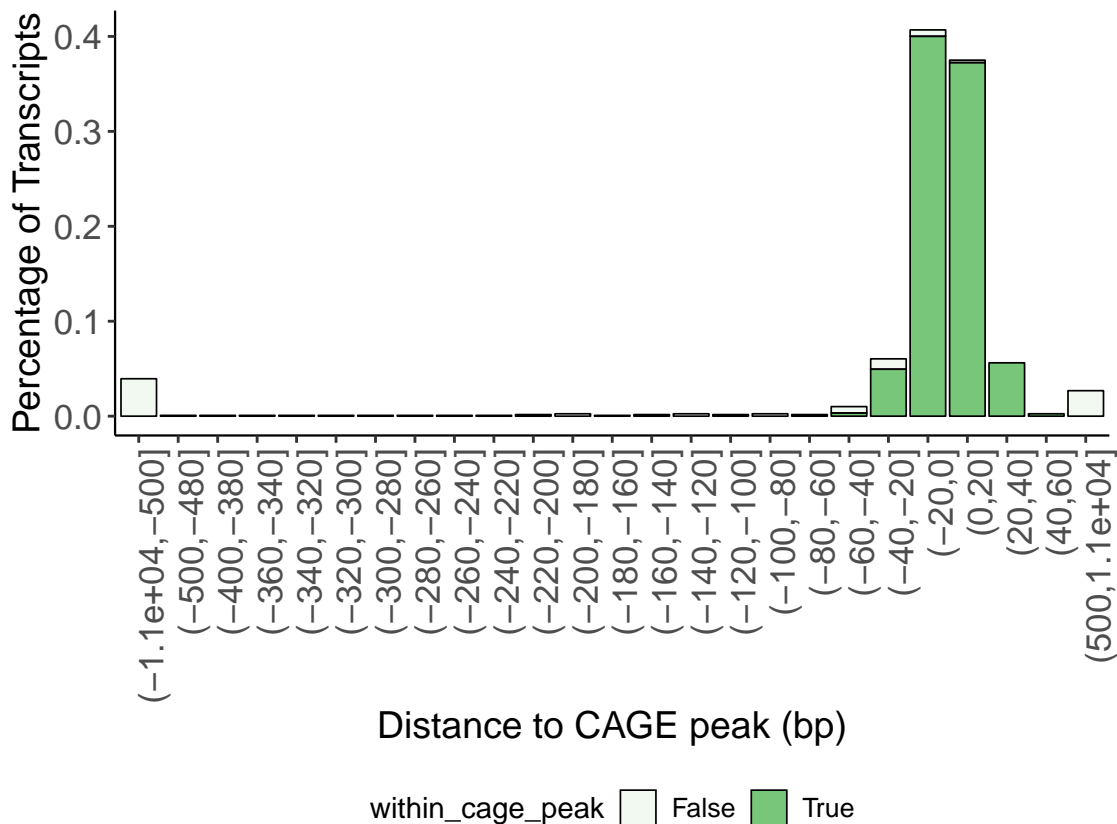
Distance to CAGE peak of multi-exonic NIC

Negative values indicate downstream of annotated CAGE peak



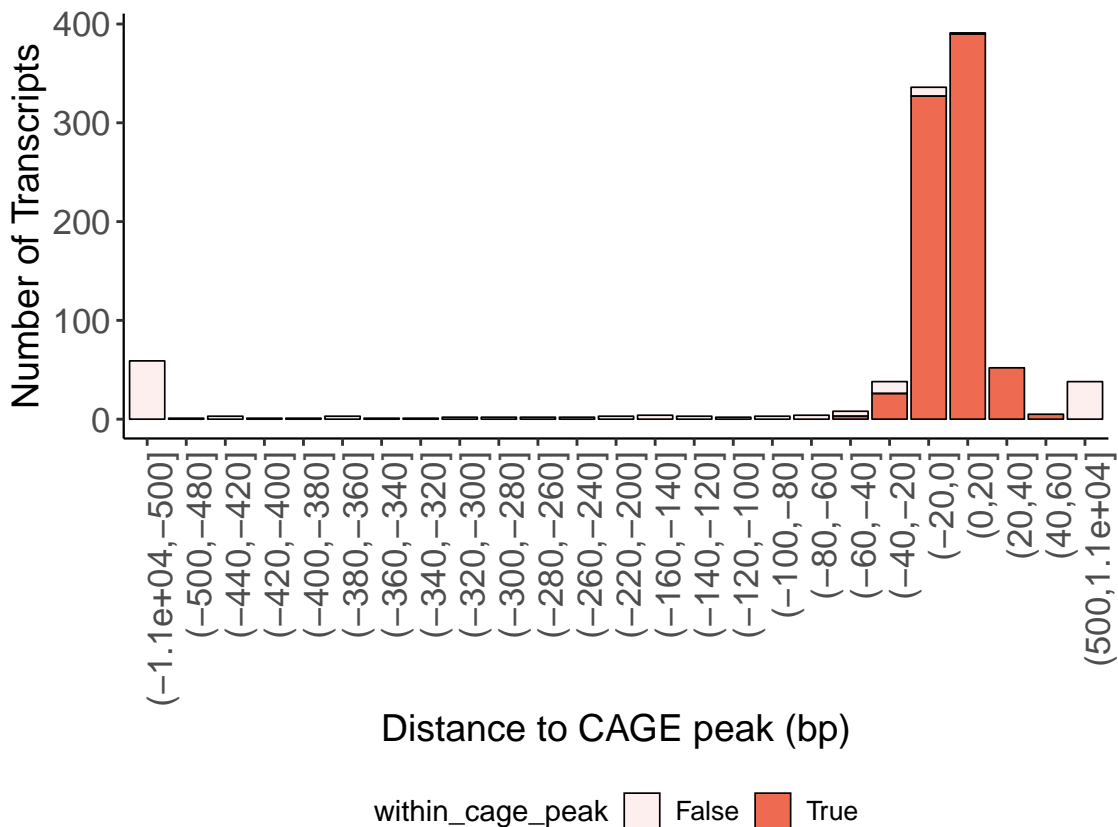
Distance to CAGE peak of multi-exonic NIC

Negative values indicate downstream of annotated CAGE peak



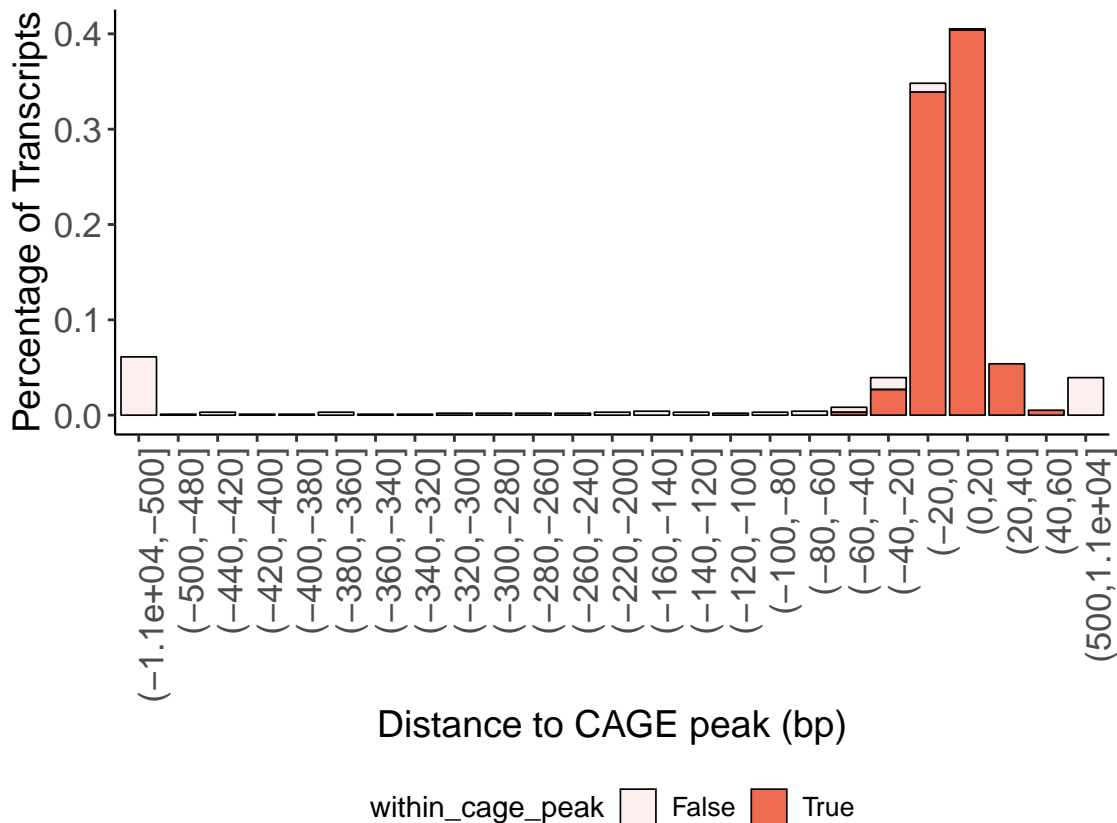
Distance to CAGE peak of multi-exonic NNC

Negative values indicate downstream of annotated CAGE peak



Distance to CAGE peak of multi-exonic NNC

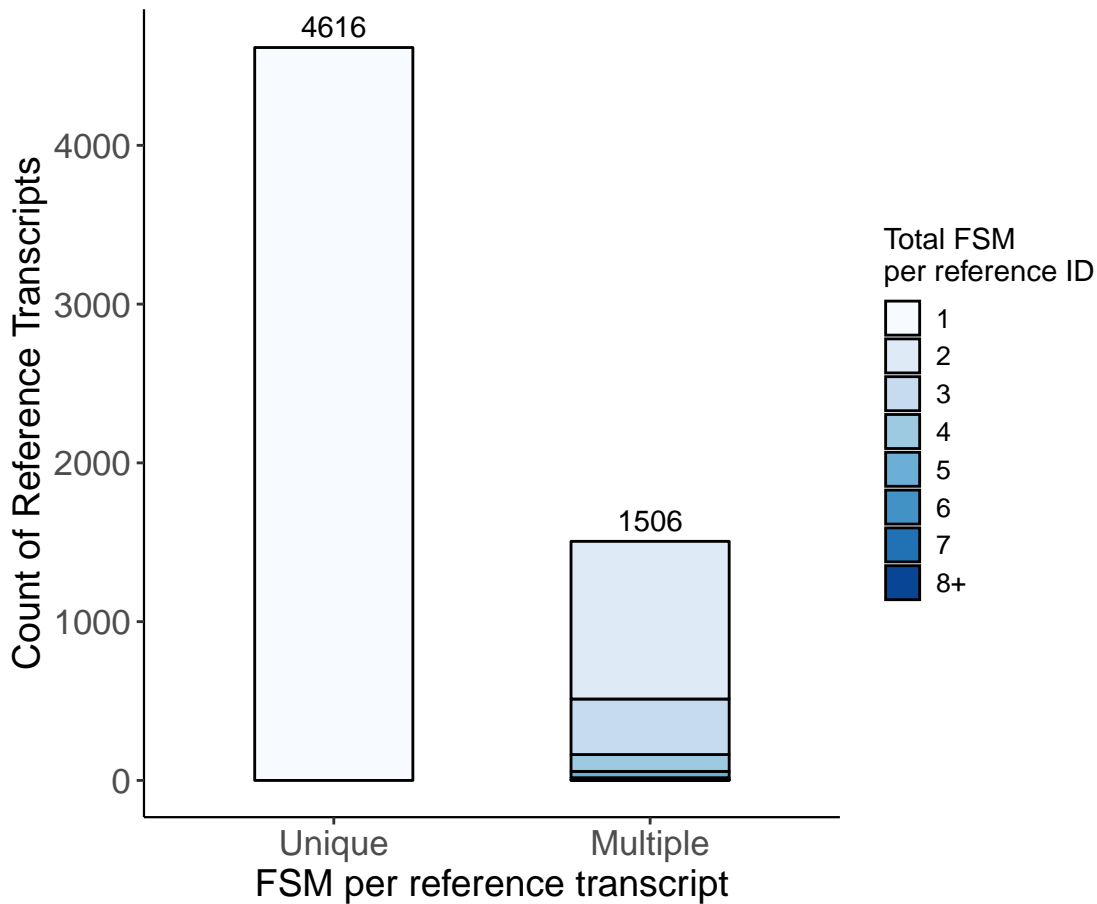
Negative values indicate downstream of annotated CAGE peak



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

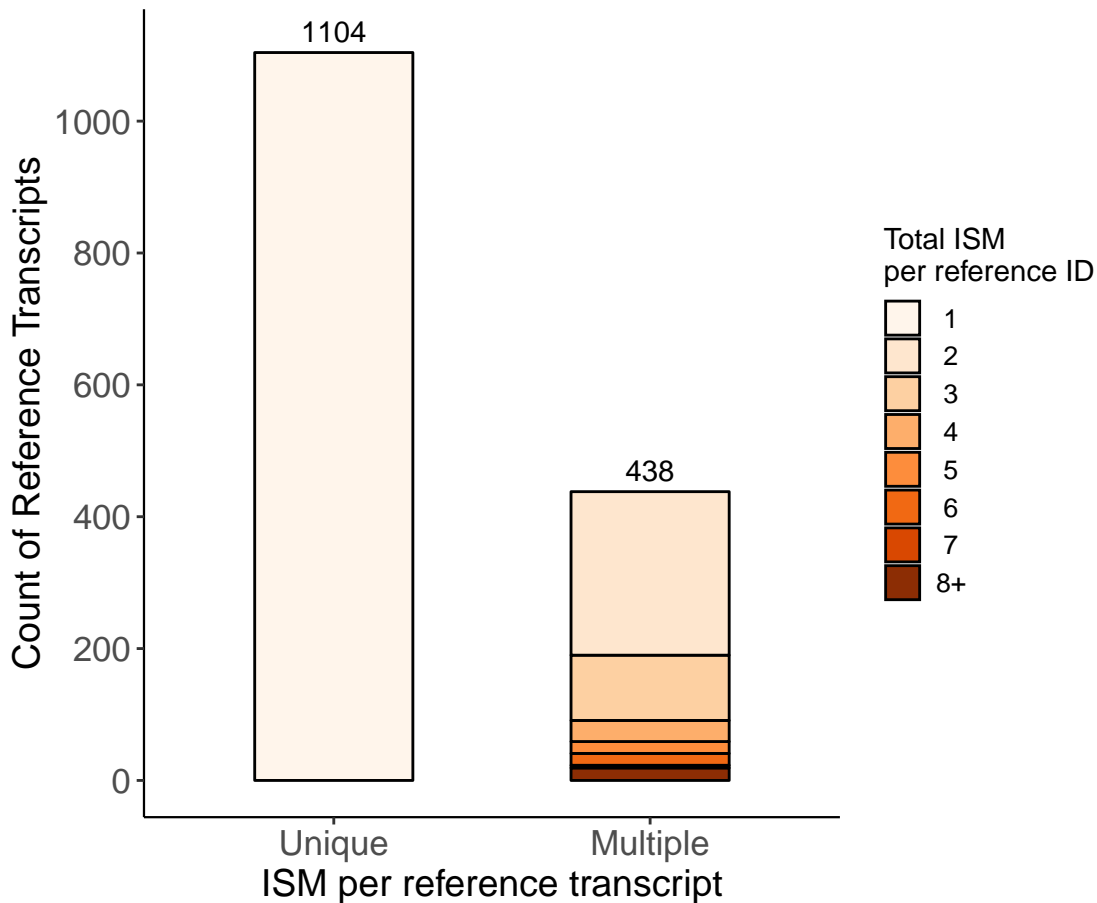
Reference transcript redundance

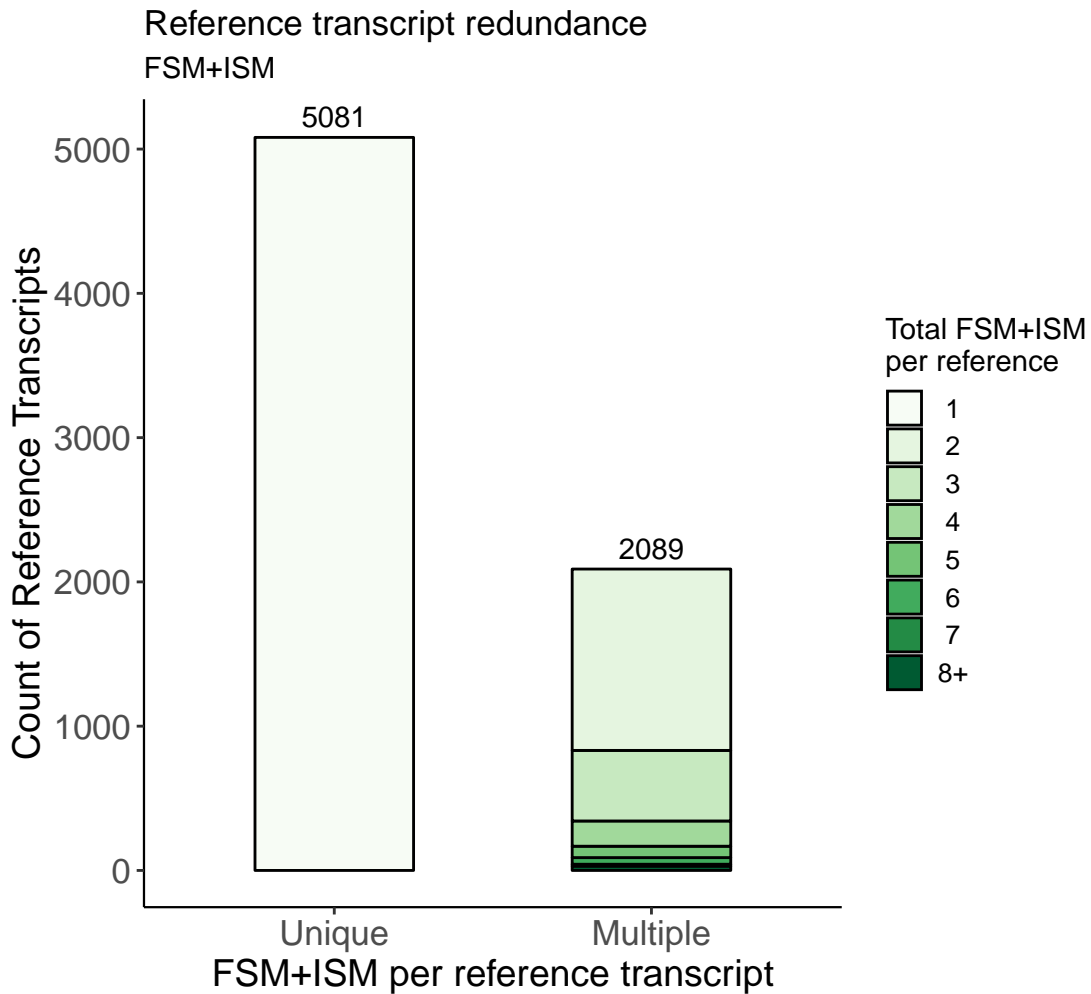
Only FSM



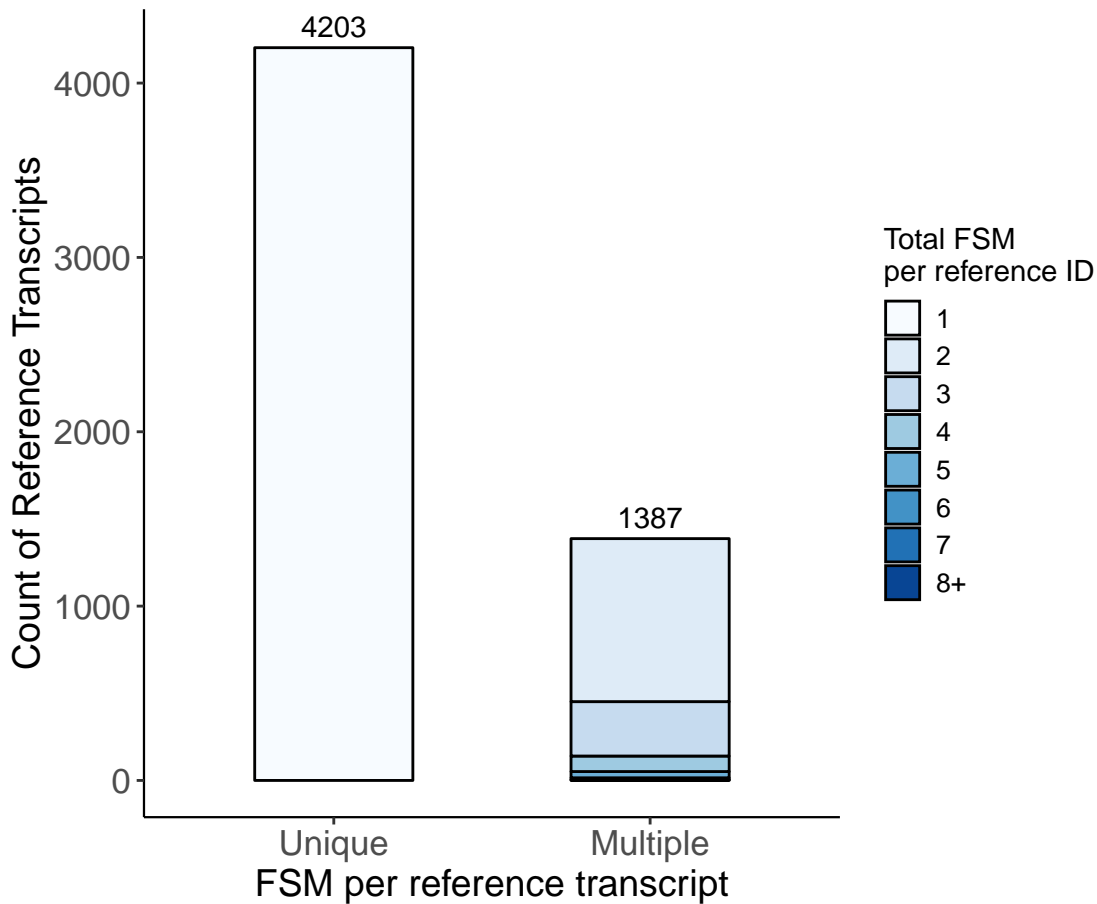
Reference transcript redundance

Only ISM



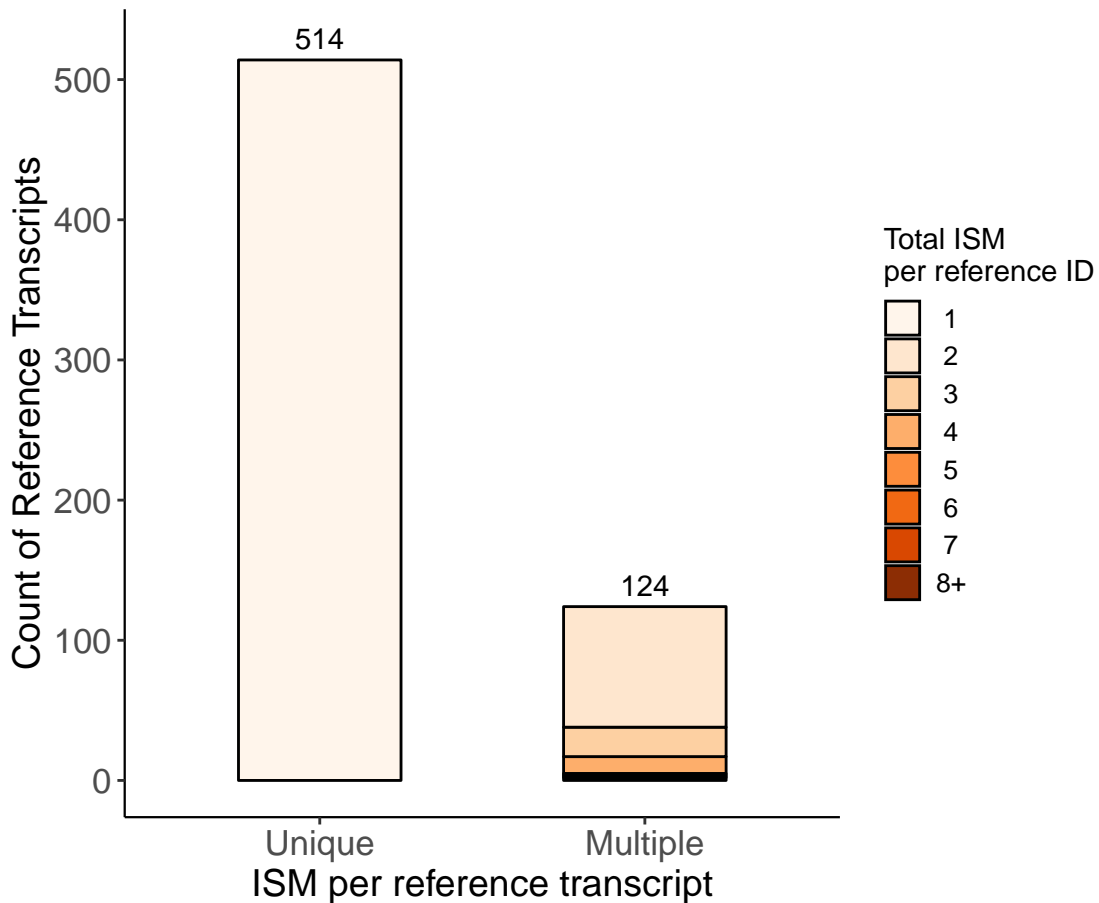


Only FSM with CAGE support

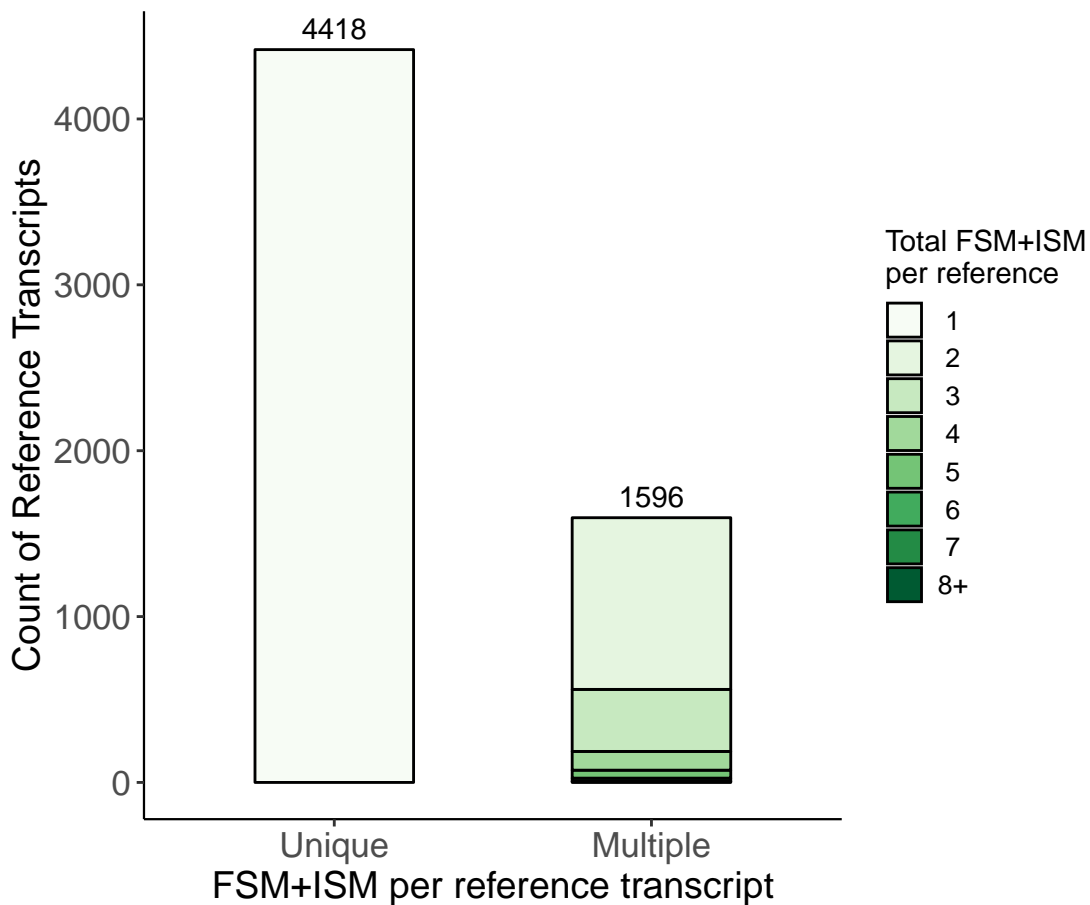


Reference transcript redundance

Only ISM with CAGE support

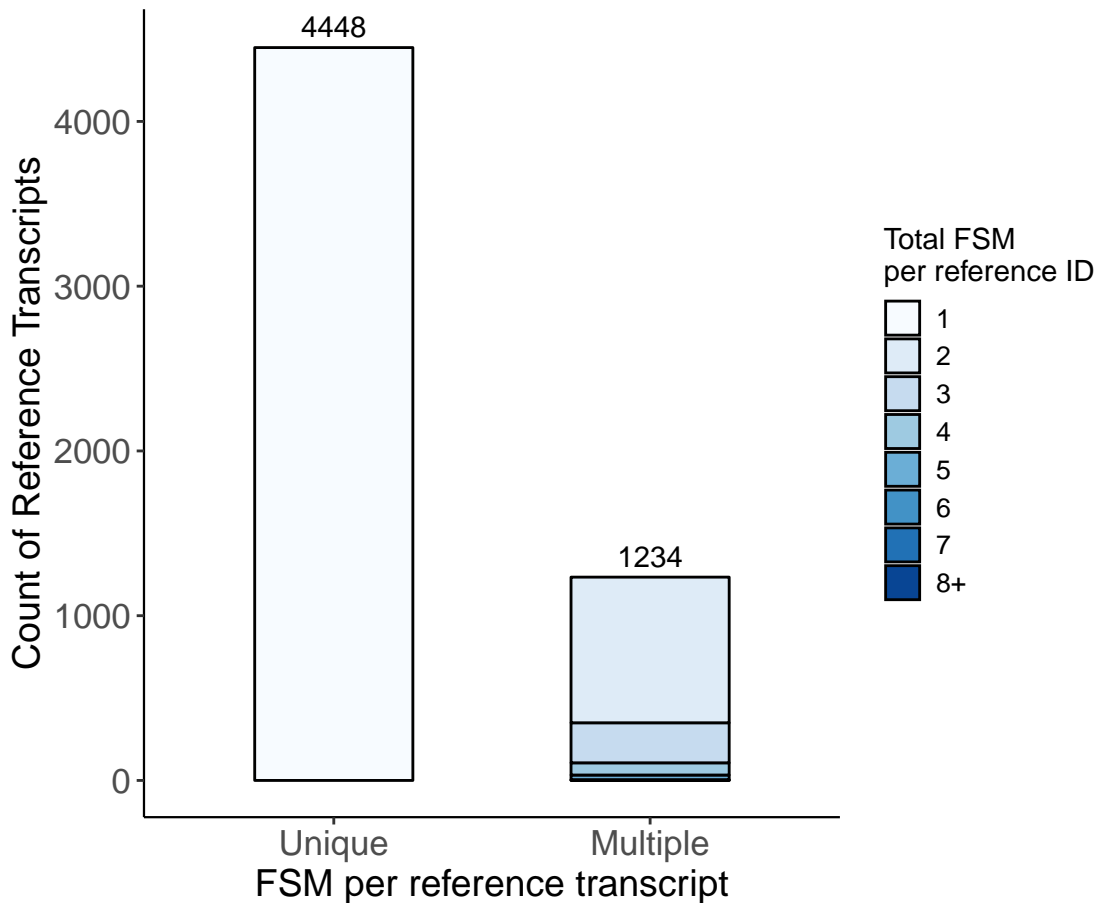


Reference transcript redundance
FSM+ISM with CAGE support



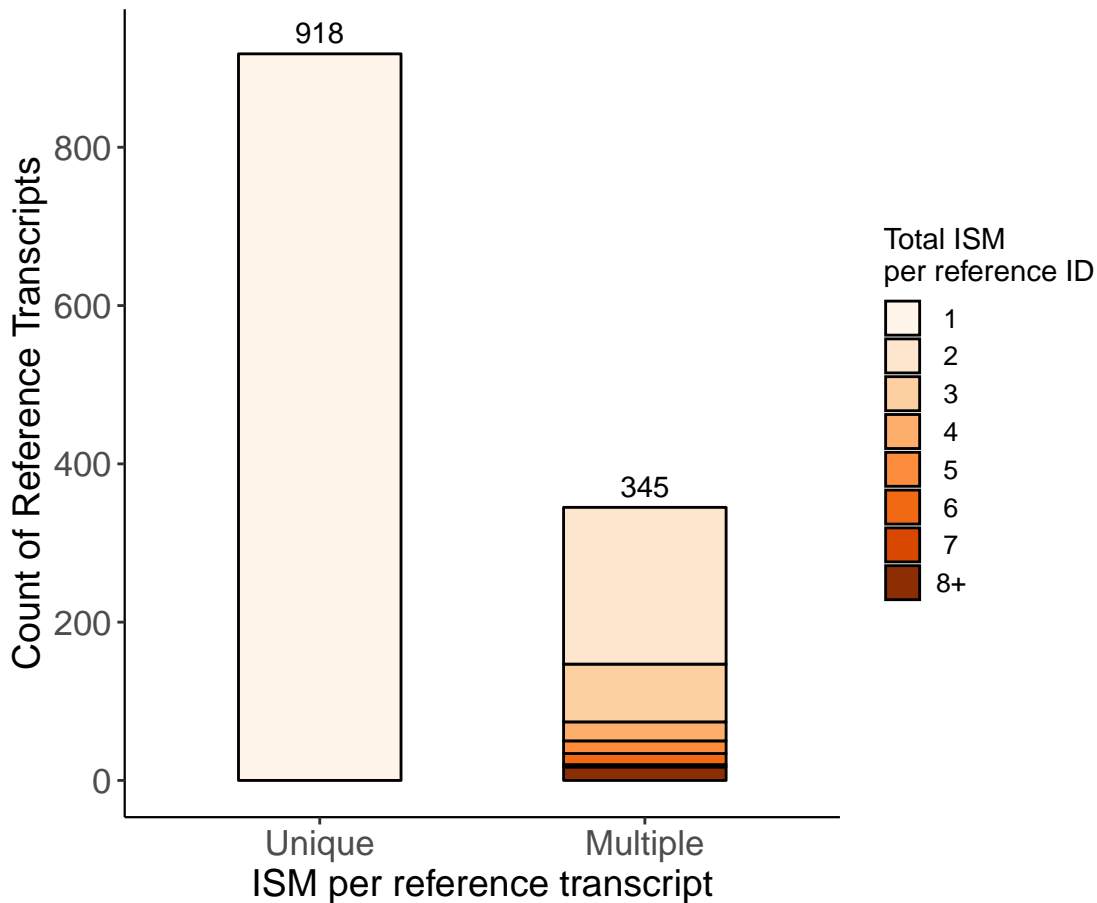
Reference transcript redundance

Only FSM with a polyA motif found



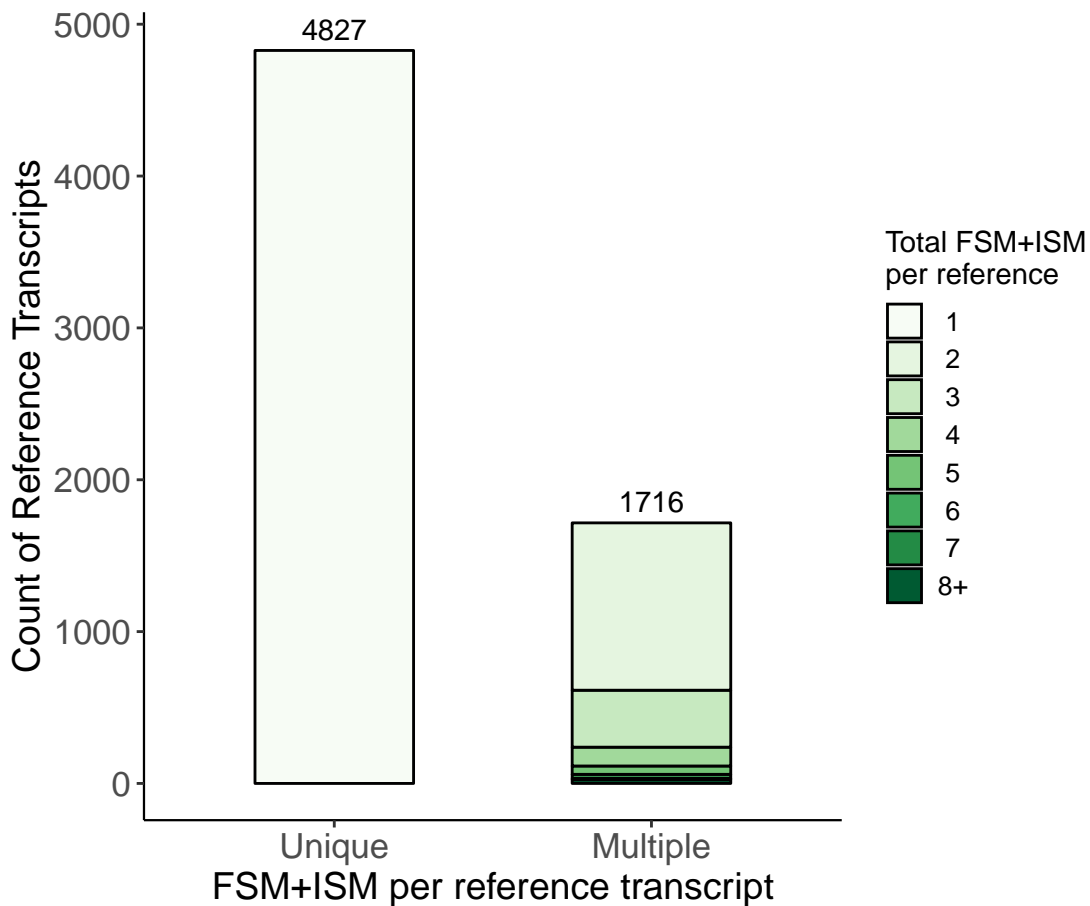
Reference transcript redundancy

Only ISM with a polyA motif found



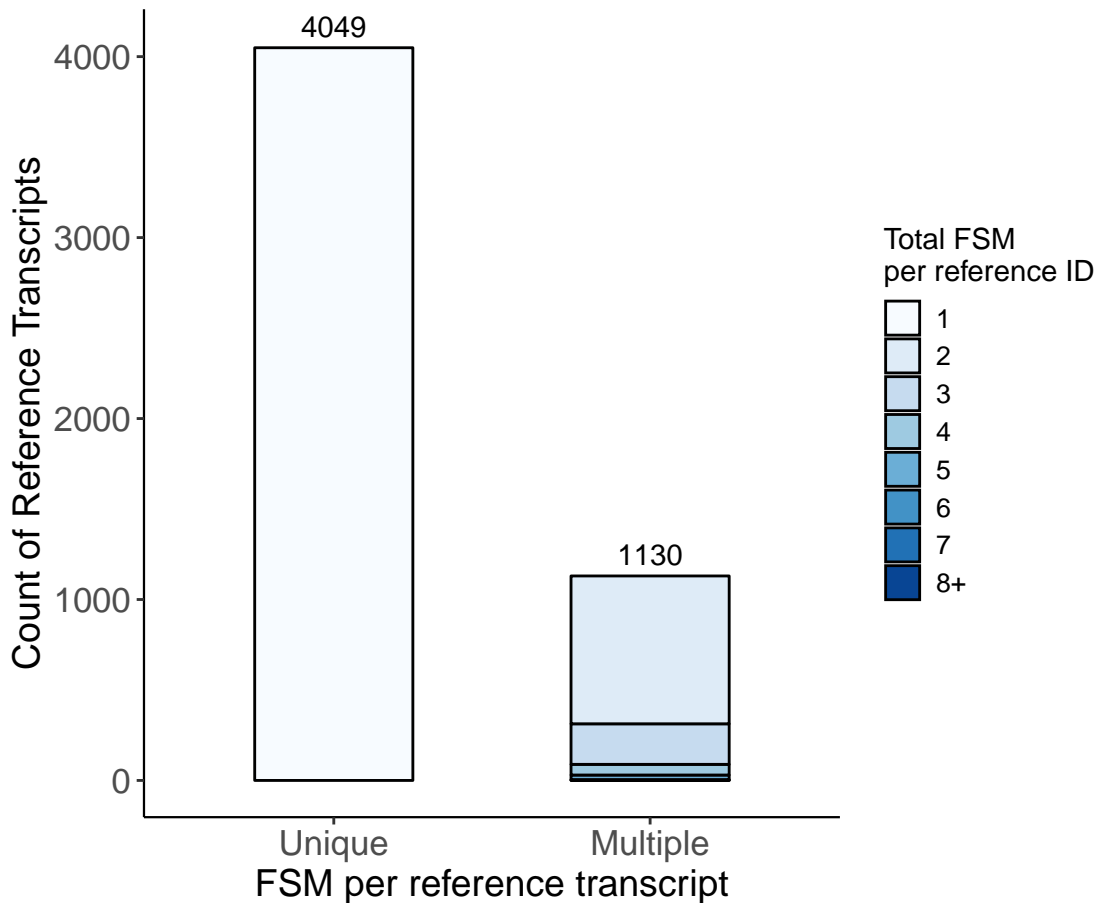
Reference transcript redundance

FSM+ISM with a polyA motif found



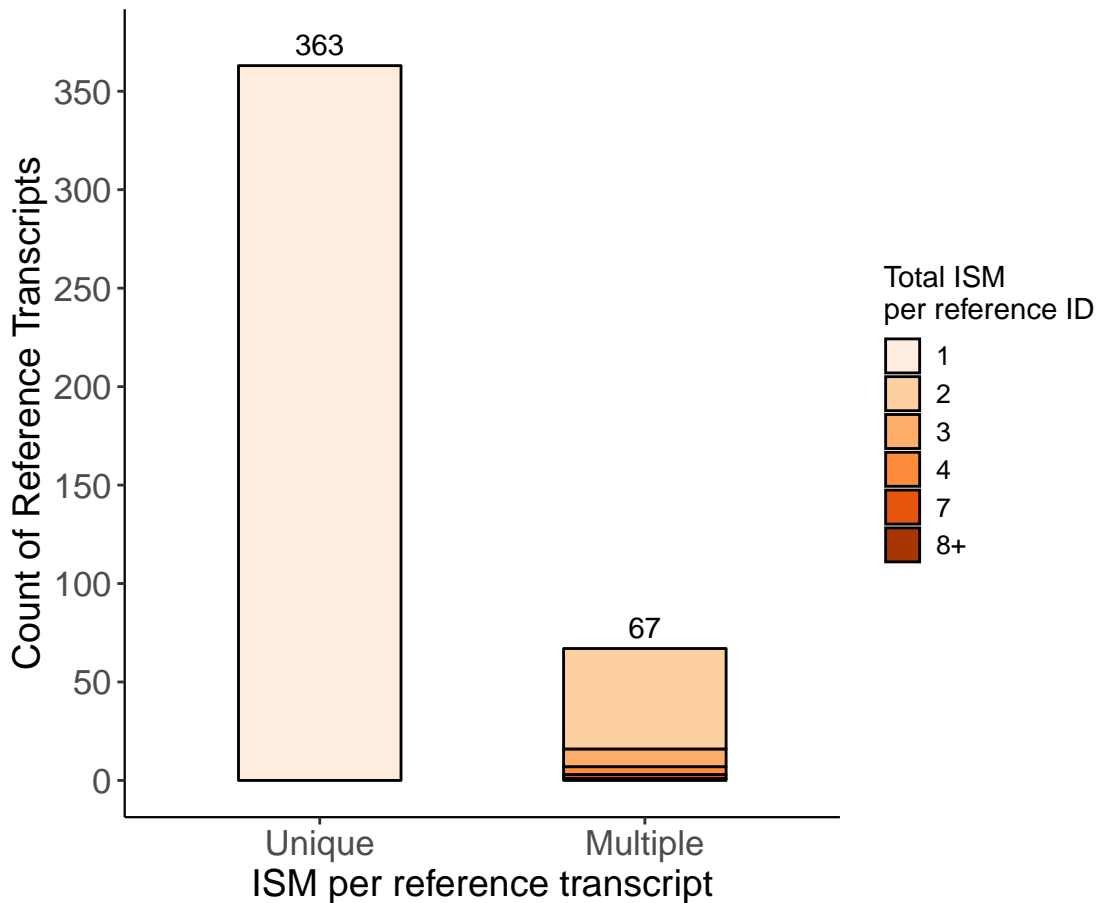
Reference transcript redundance

Only FSM with CAGE support and polyA motif

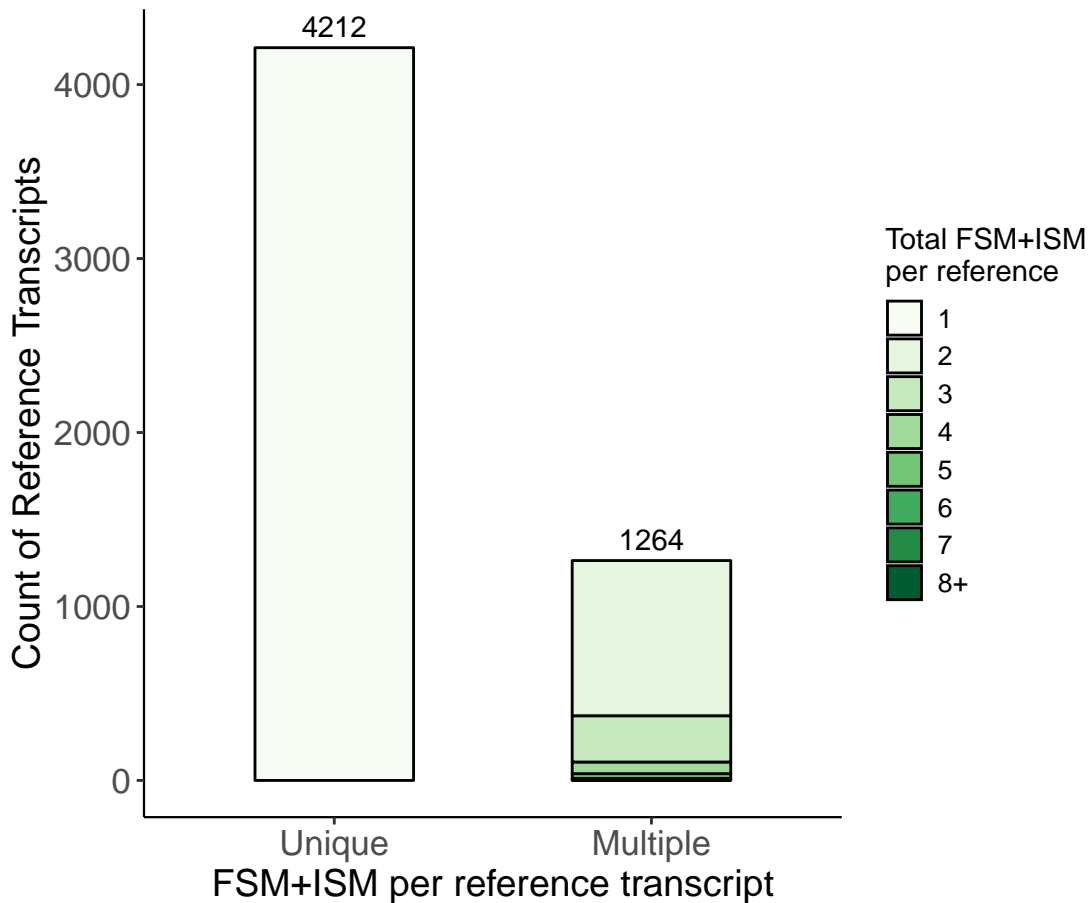


Reference transcript redundance

Only ISM with CAGE support and polyA motif



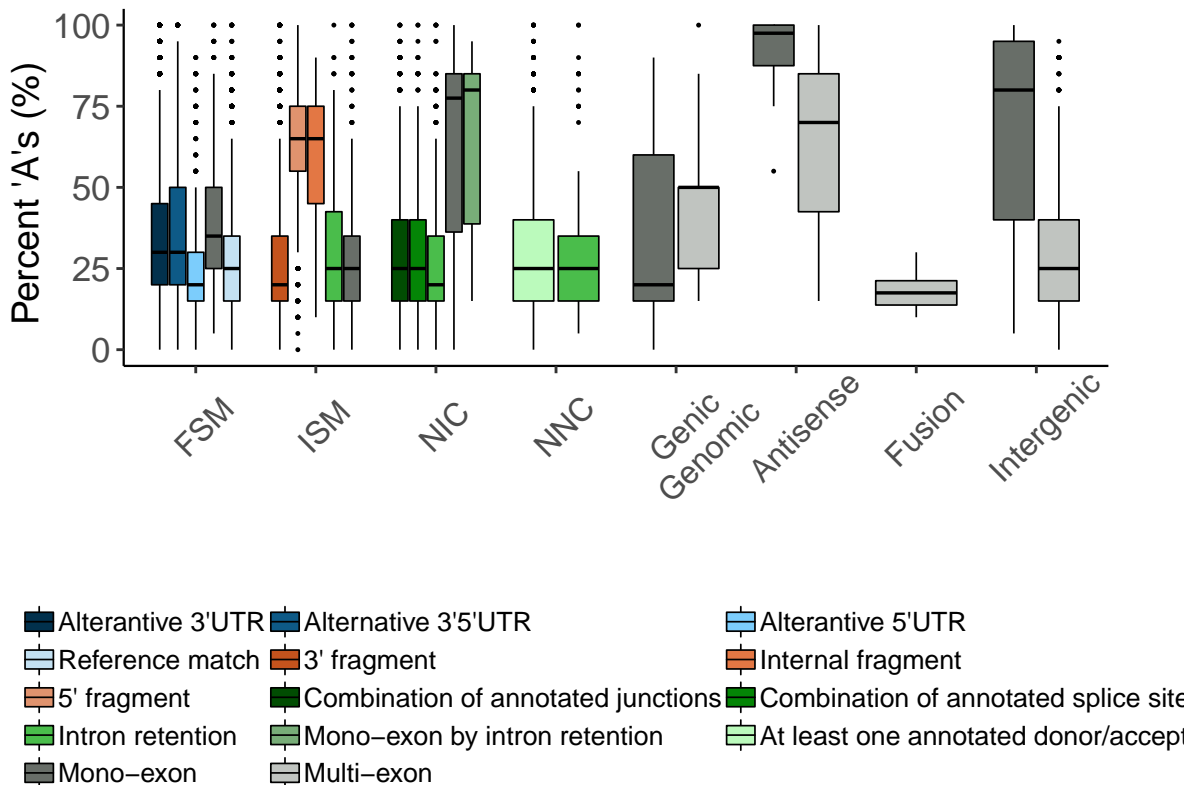
FSM+ISM with CAGE support and polyA motif



Intra-Priming Quality Check

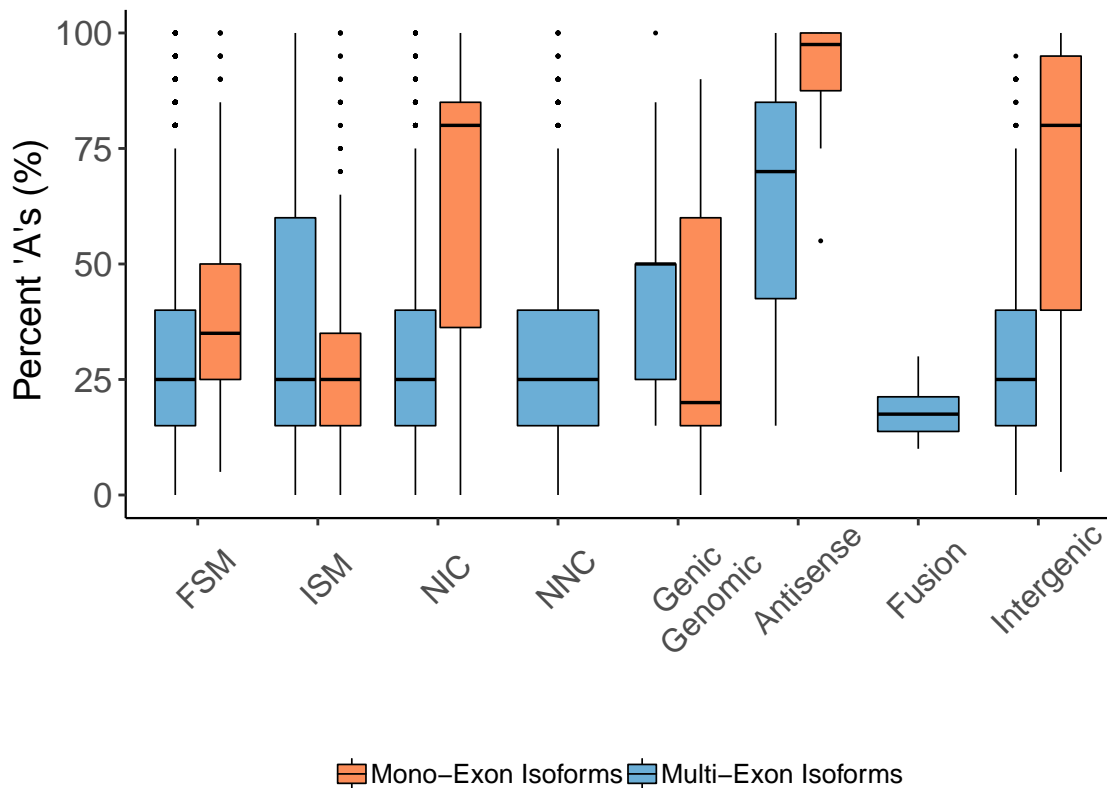
Possible Intra-Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



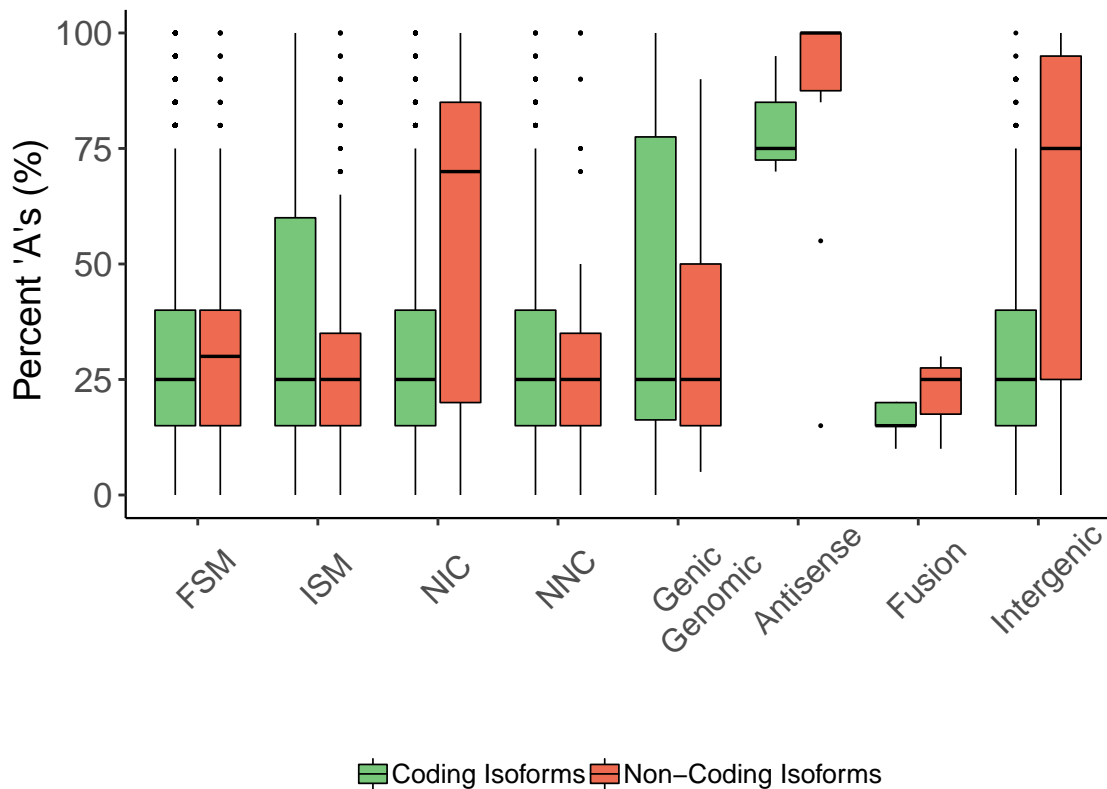
Possible Intra-Priming, Mono- vs Multi-Exon

Percent of genomic 'A's in downstream 20 bp



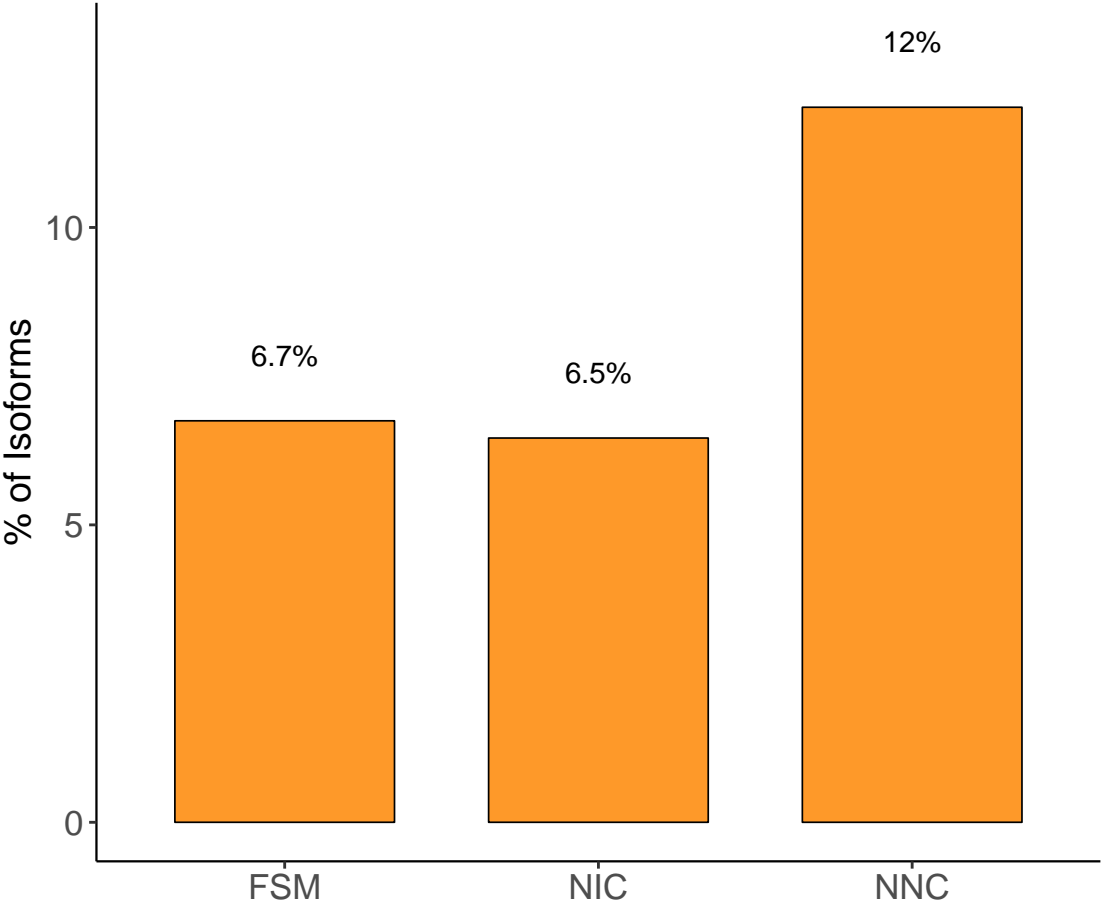
Possible Intra-Priming, Coding vs Non-Coding

Percent of genomic 'A's in downstream 20 bp

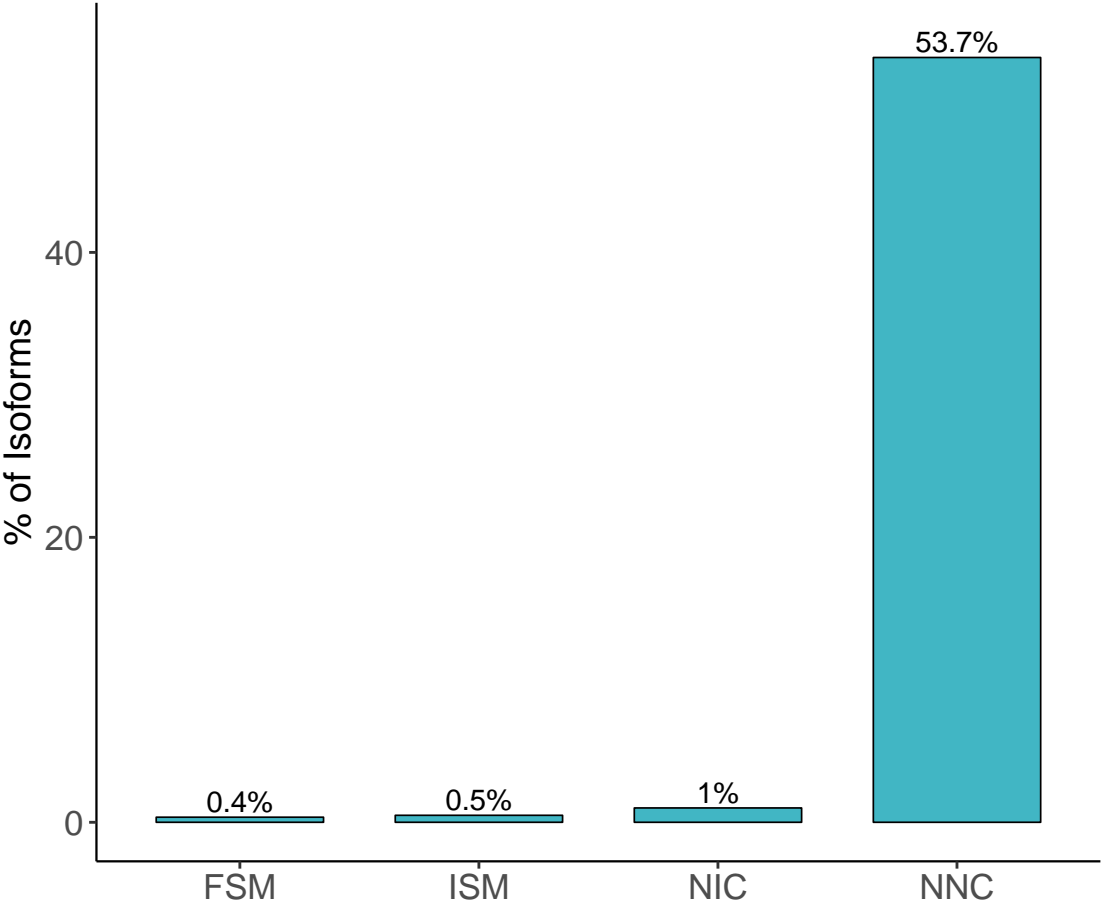


Quality Controls

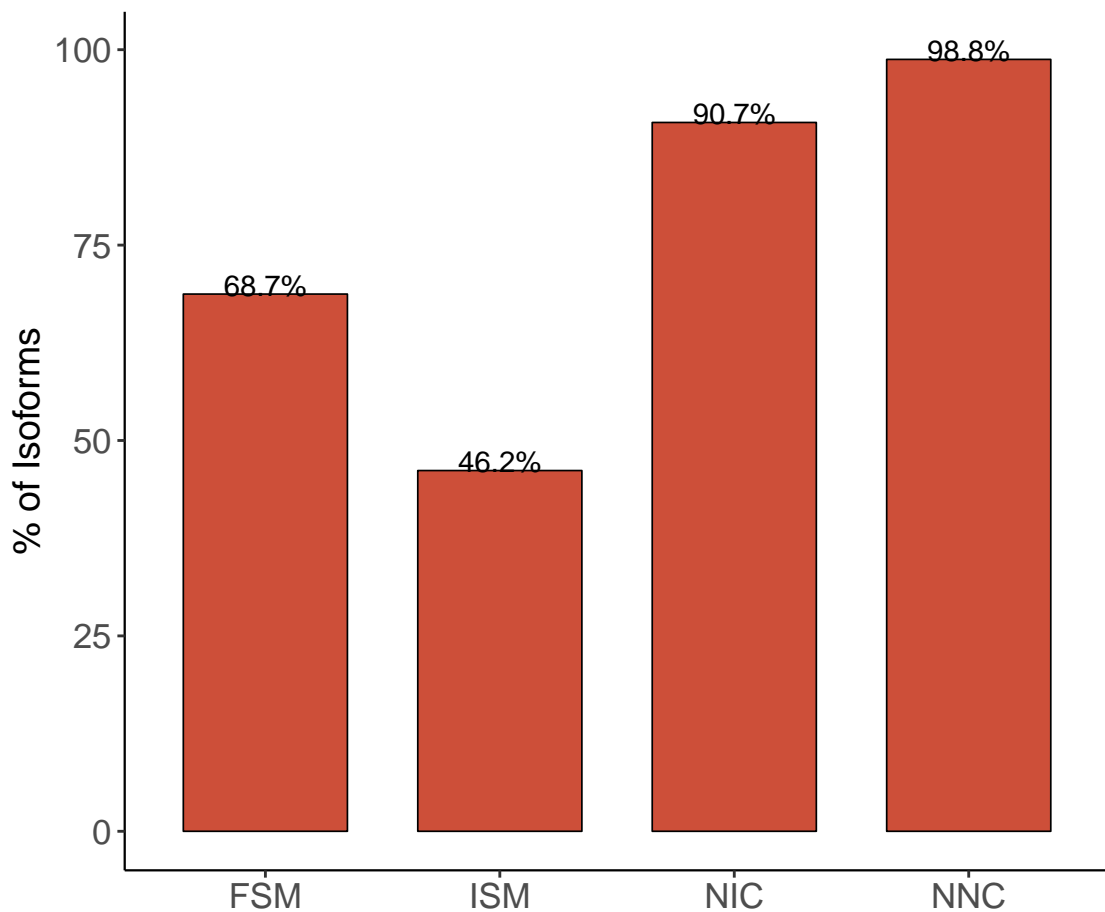
Incidence of RT-switching



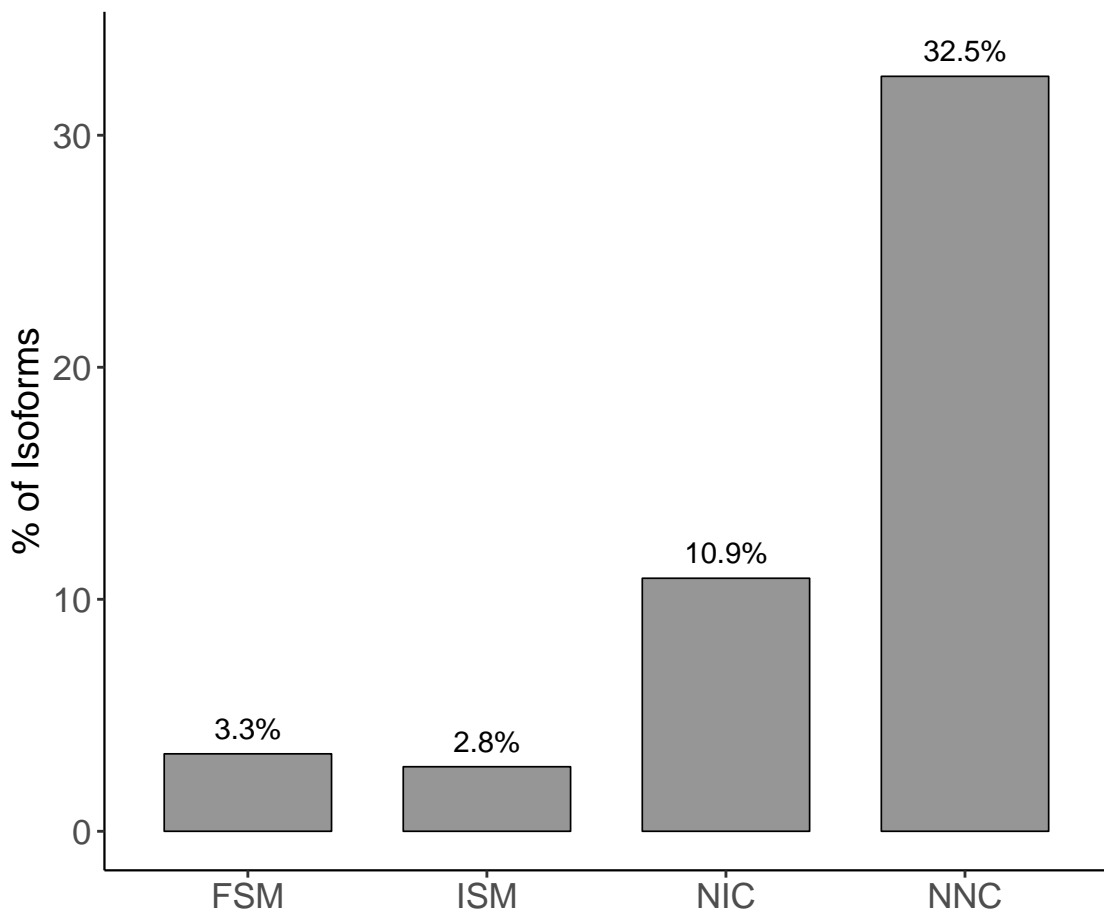
Incidence of Non-Canonical Junctions



Incidence of SJ without SR coverage Junctions



Incidence of NMD by structural category



Quality control attributes across structural categories

