

SQANTI3 report

Unique Genes: 24581
Unique Isoforms: 579611

Transcript Classification

Category	Isoforms, count
FSM	145277
ISM	213164
NIC	118186
NNC	94898
Genic Genomic	708
Antisense	1682
Fusion	3498
Intergenic	2111
Genic Intron	0

Gene Classification

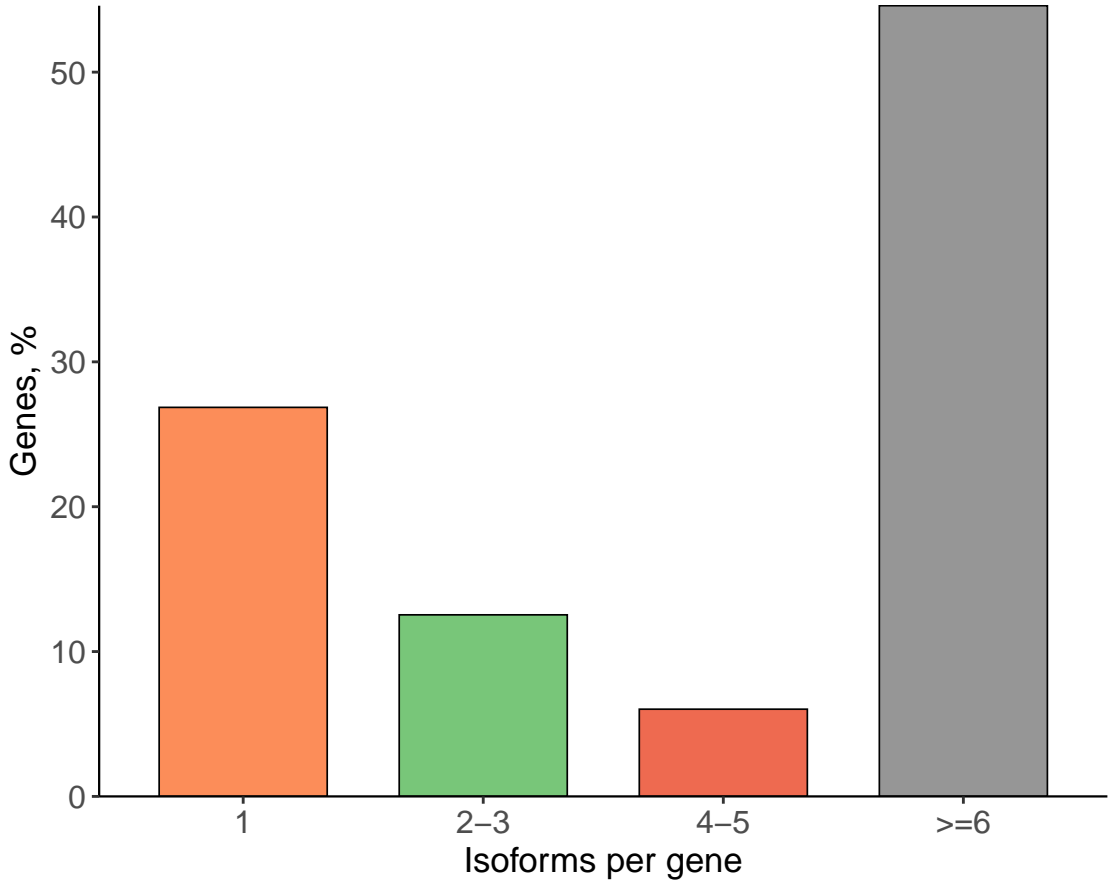
Category	Genes, count
Annotated Genes	21612
Novel Genes	2969

Splice Junction Classification

Category	SJs, count	Percent
Known canonical	195813	65.84
Known Non-canonical	90	0.03
Novel canonical	101494	34.13
Novel Non-canonical	0	0.00

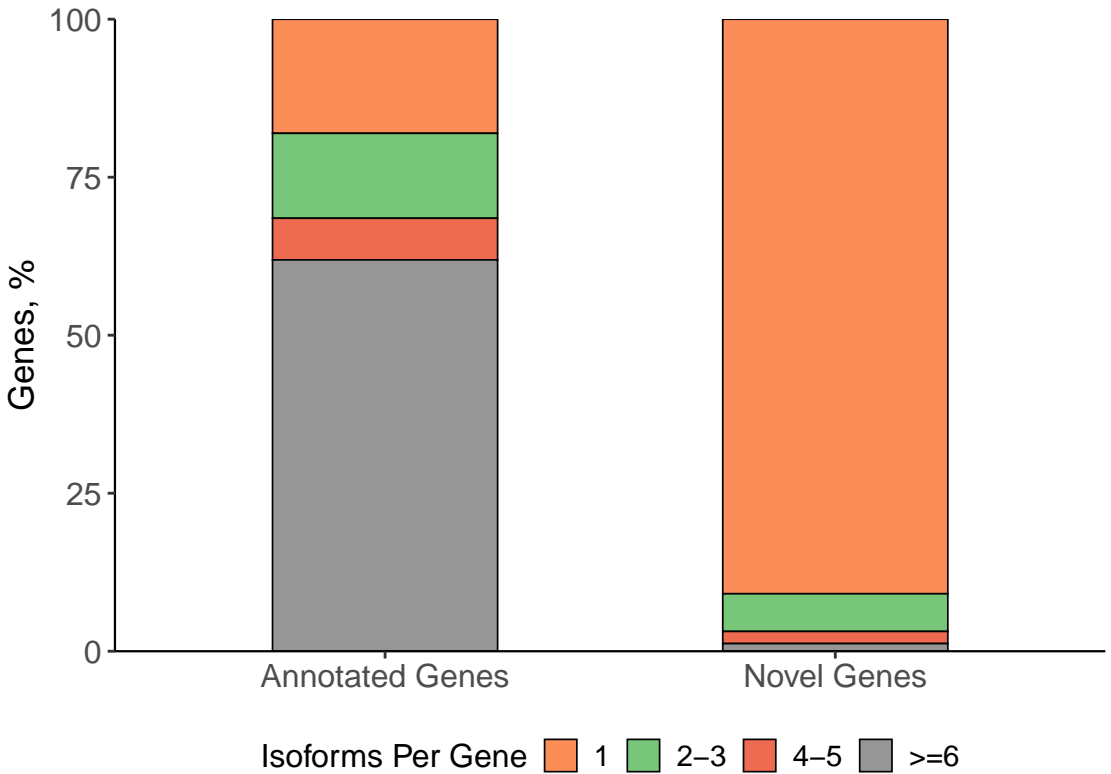
Gene Characterization

Number of Isoforms per Gene

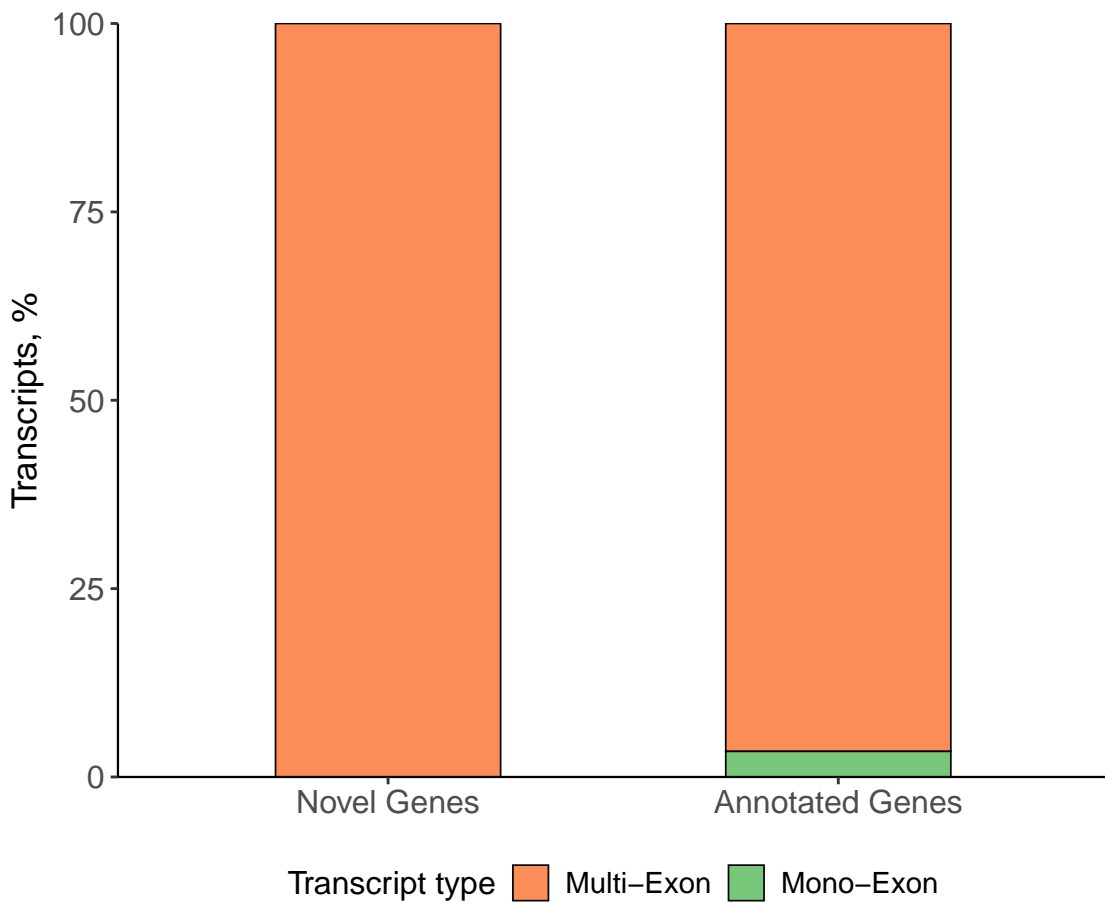


Number of Isoforms per Gene

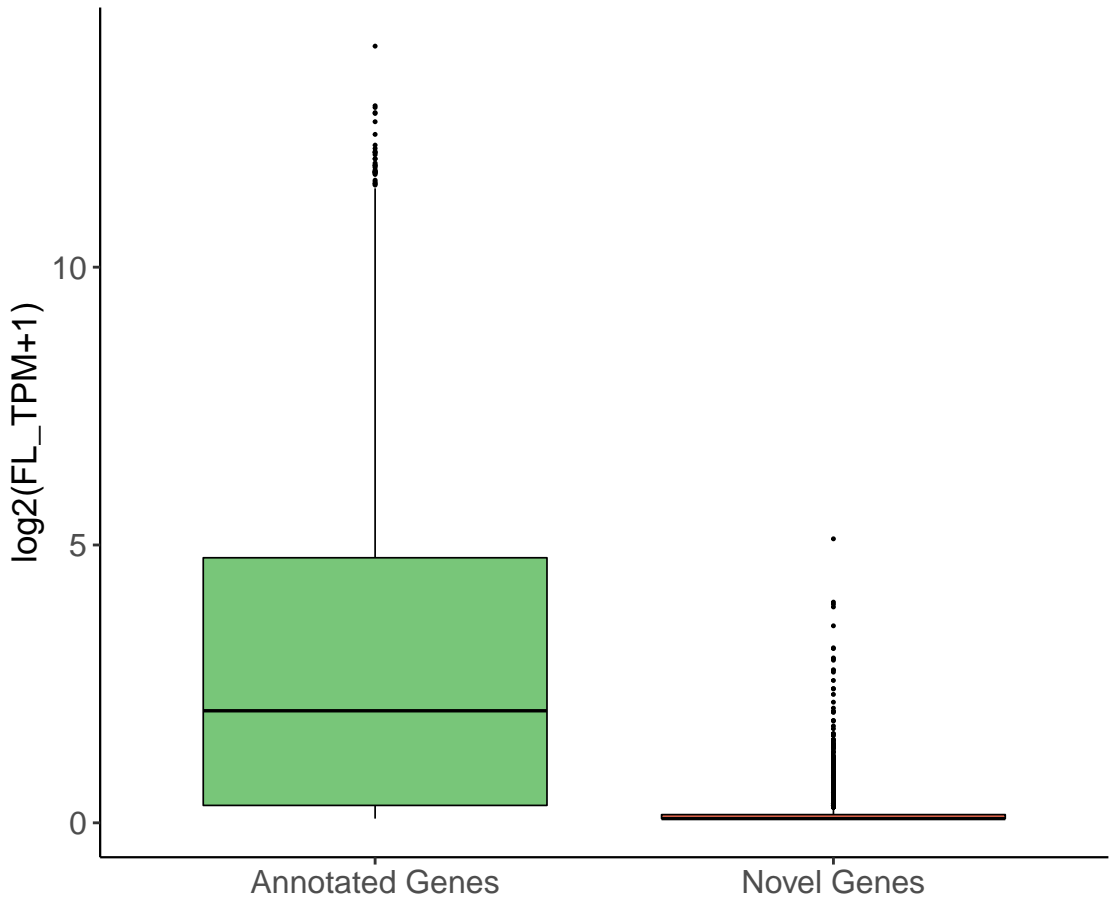
Known vs Novel Genes



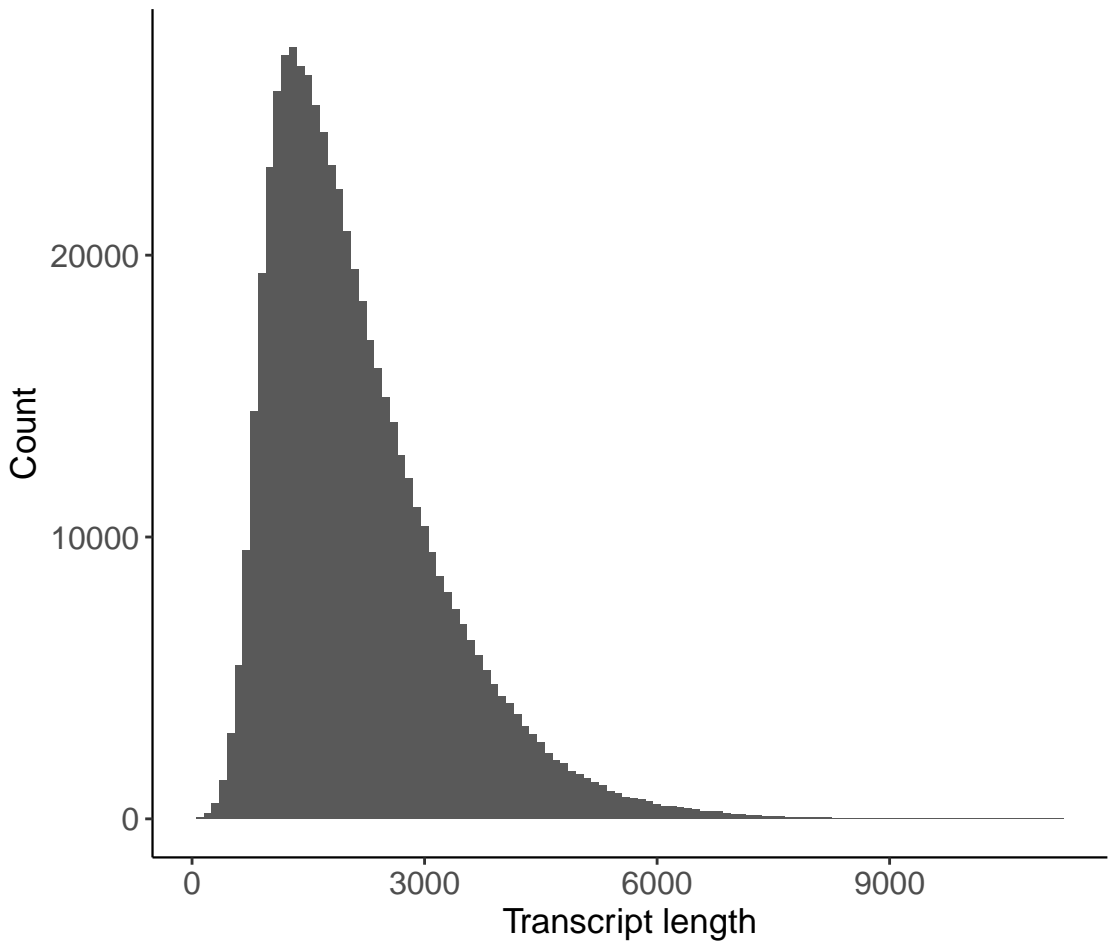
Distribution of Mono- vs Multi-Exon Transcripts



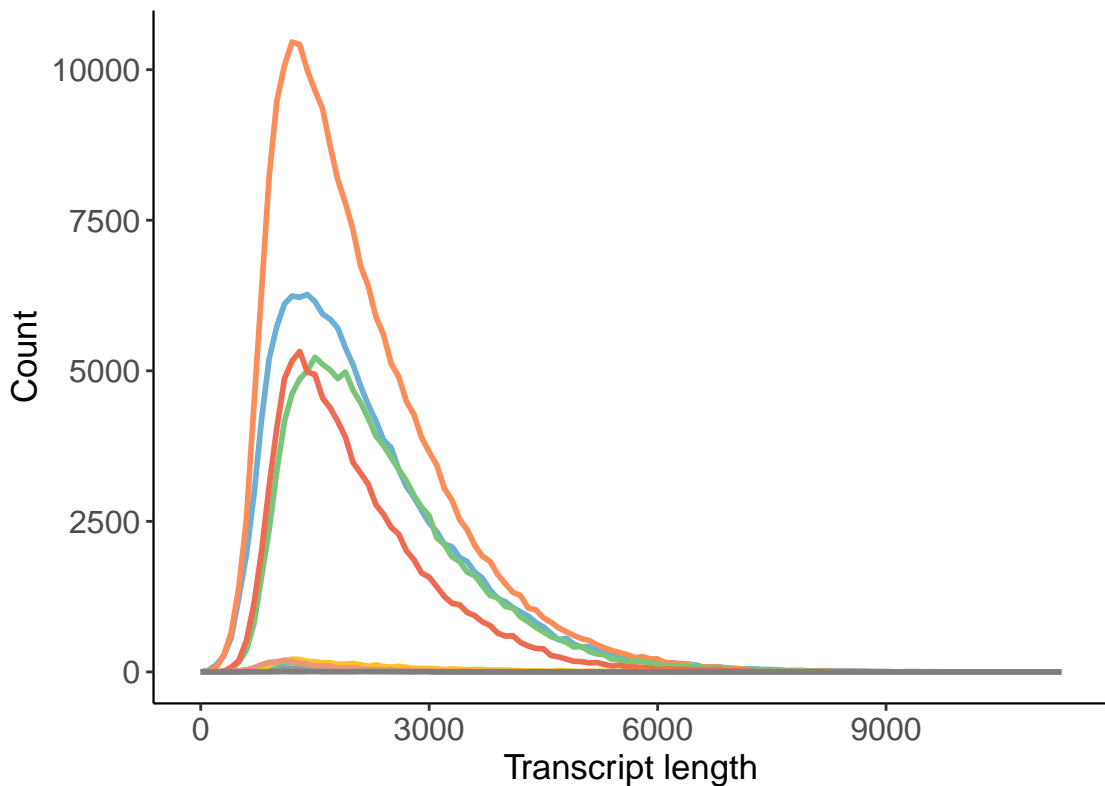
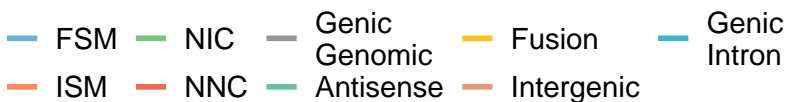
Number of FL reads per Gene by Type of Gene Annotation



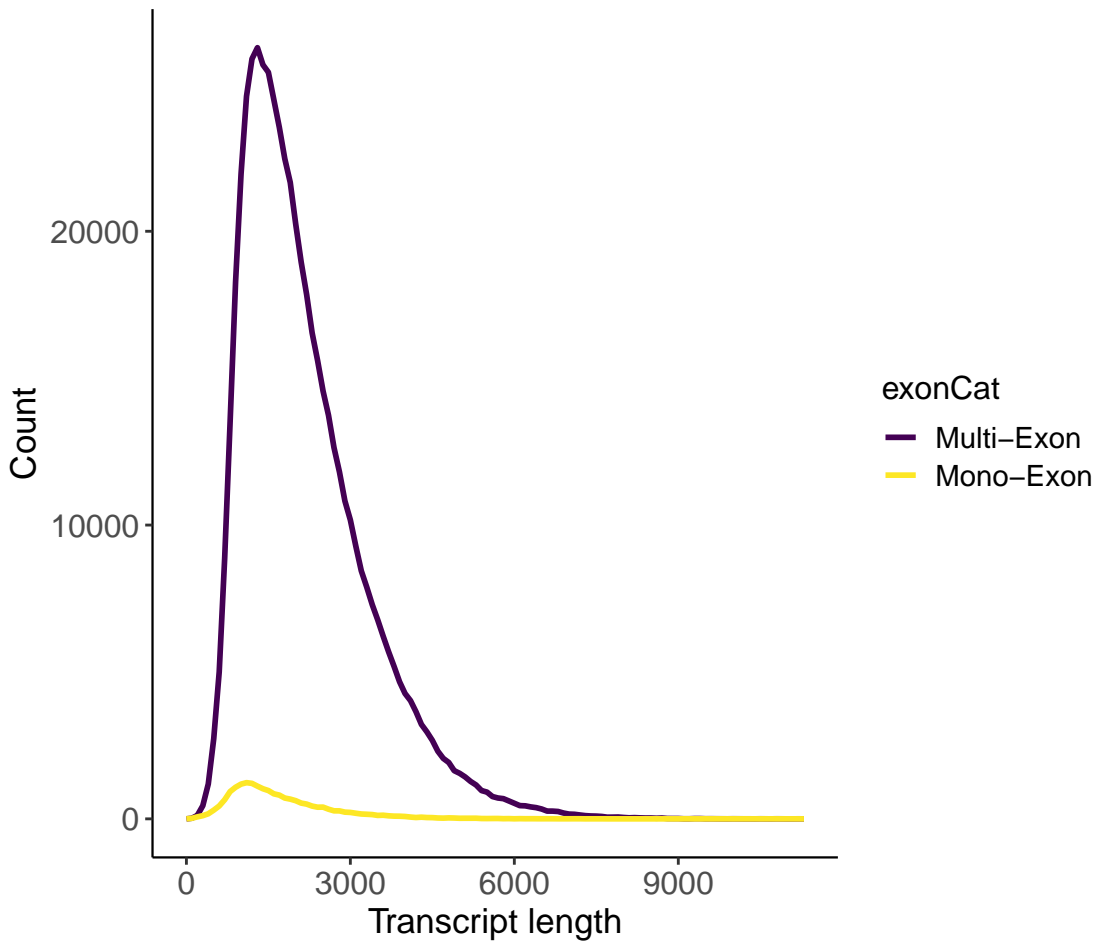
All Transcript Lengths Distribution



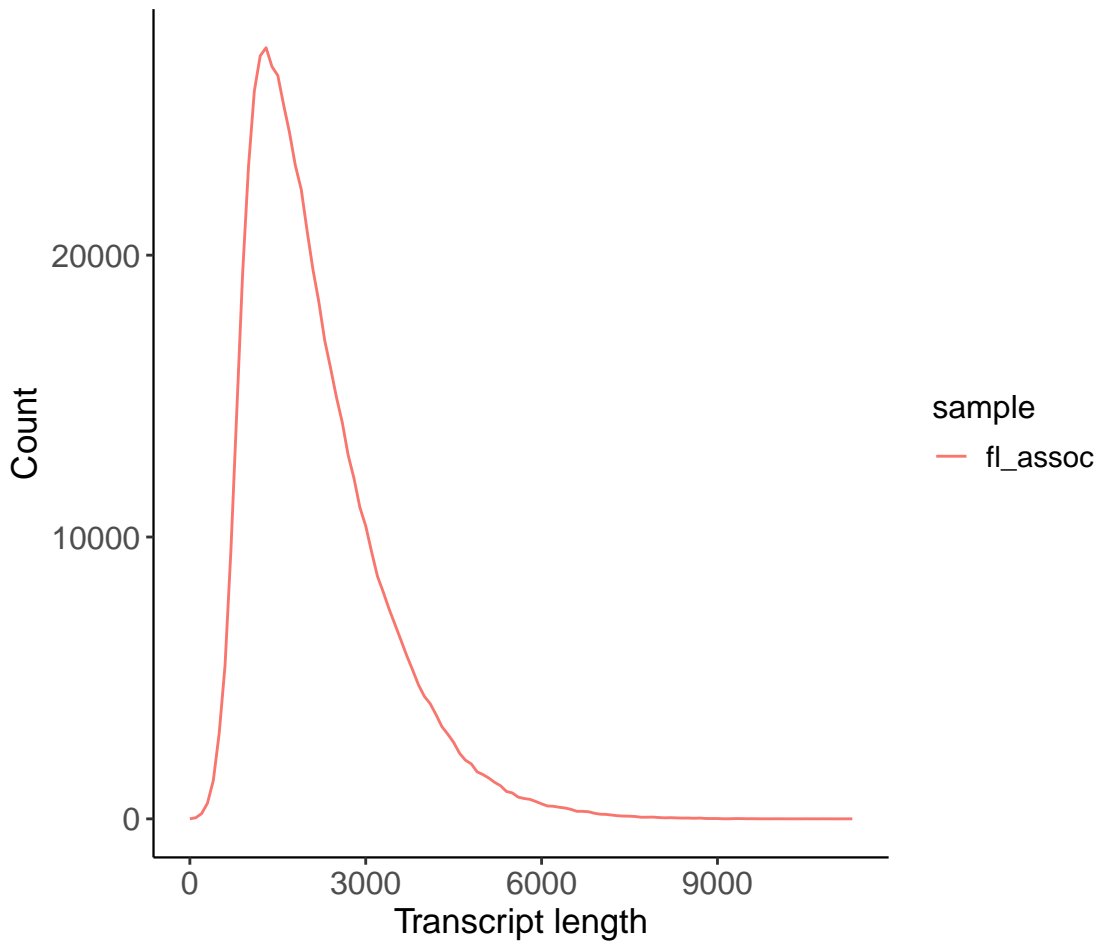
Transcript Lengths Distribution by Structural Category



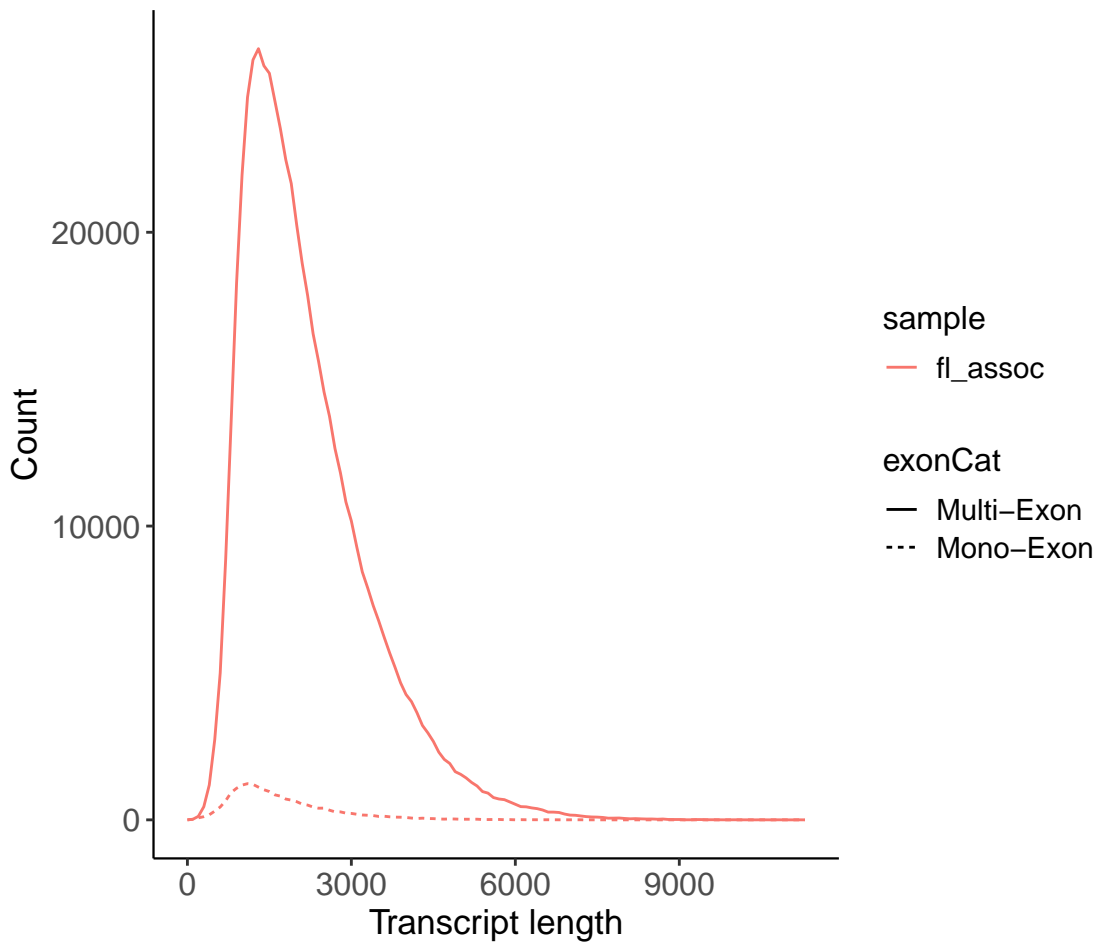
Mono- vs Multi- Exon Transcript Lengths Distribution



Transcript Lengths by Sample

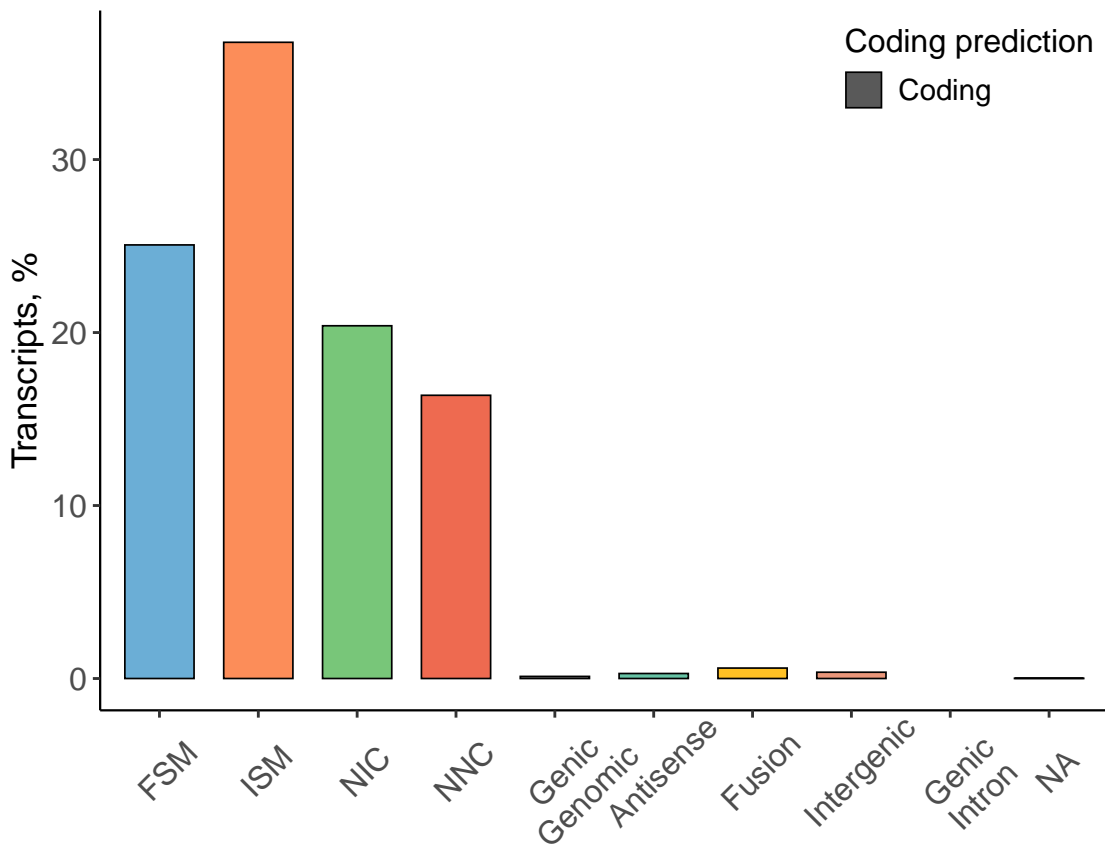


Mono- vs Multi-Exons Transcript Lengths by Sample

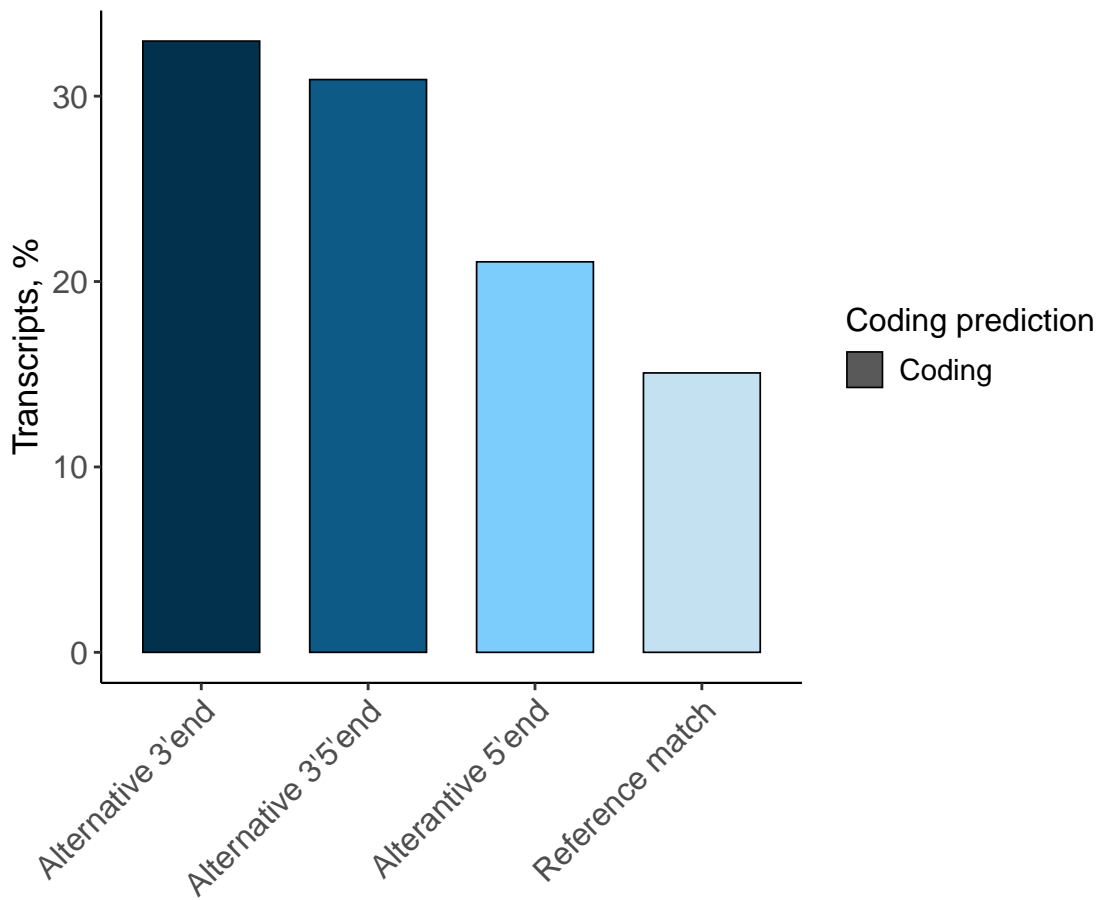


Structural Isoform Characterization

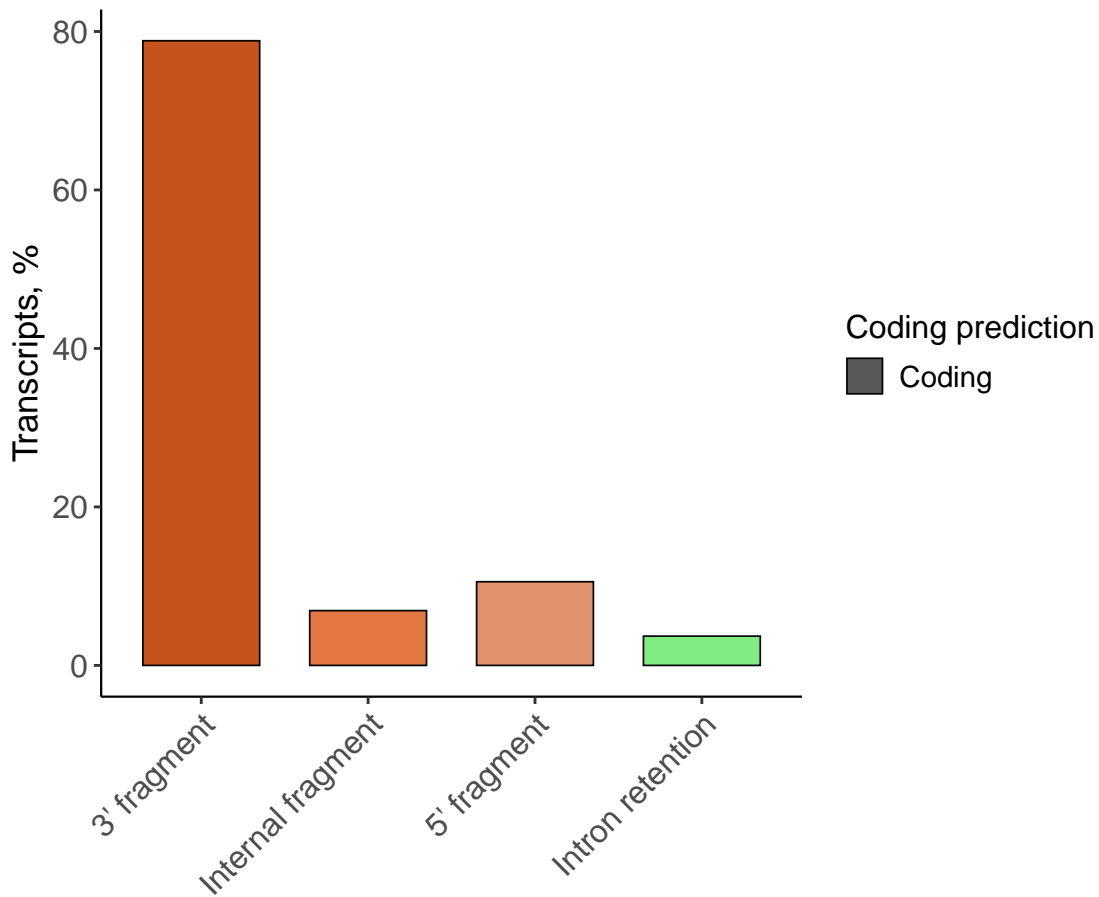
Isoform Distribution Across Structural Categories



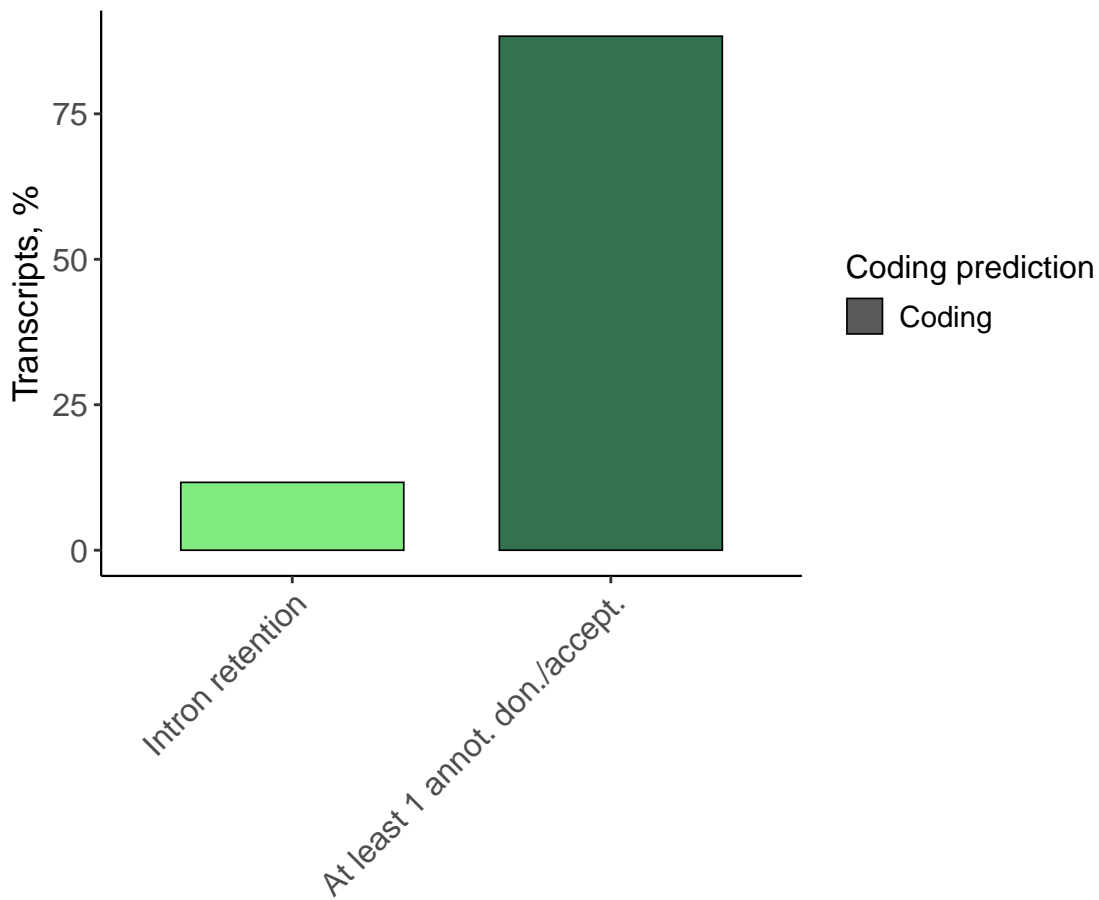
Isoform Distribution Across FSM



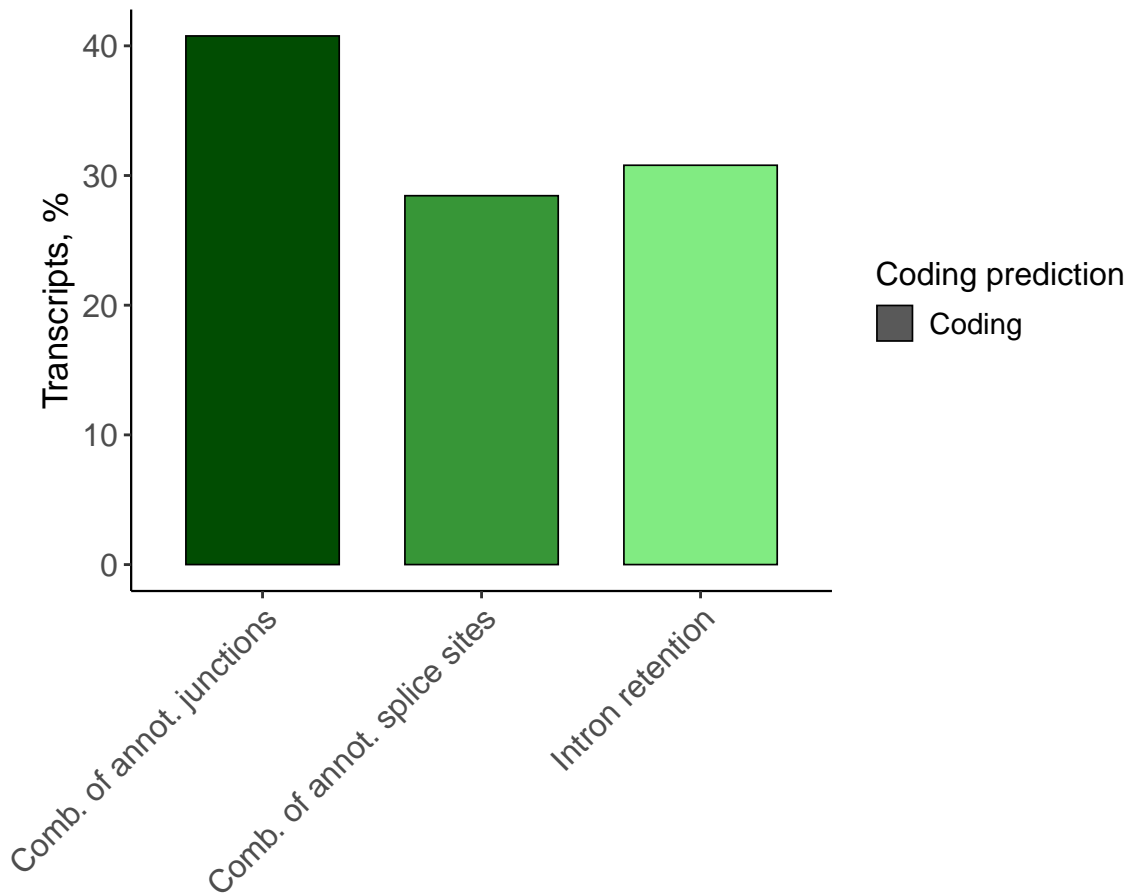
Isoform Distribution Across ISM



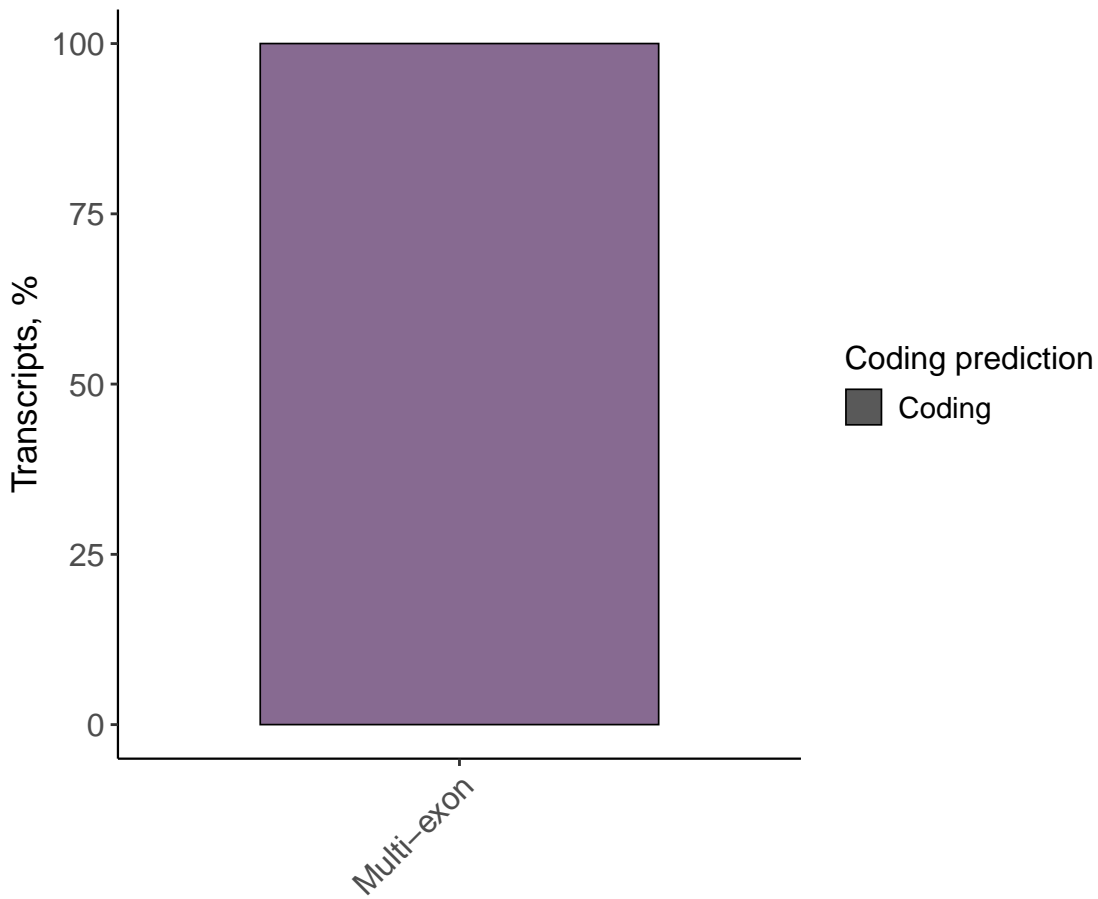
Isoform Distribution Across NNC



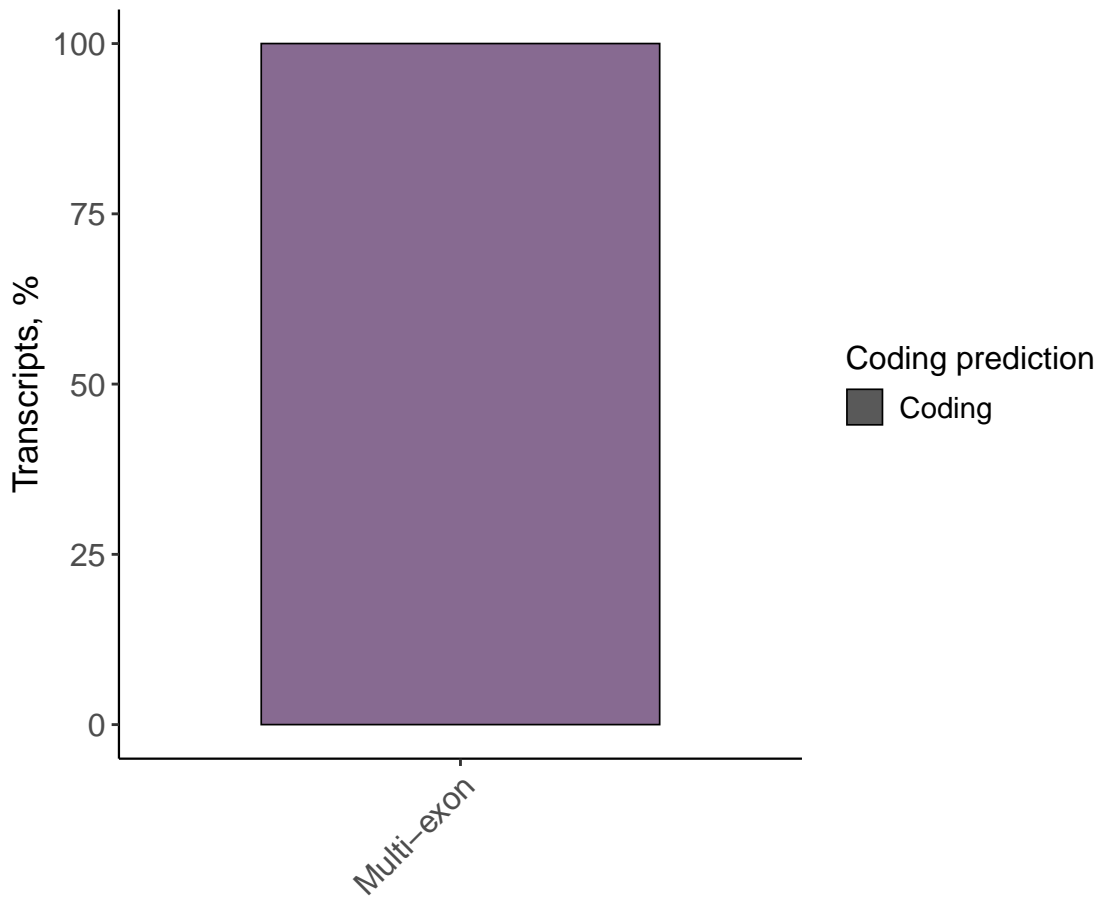
Isoform Distribution Across NIC



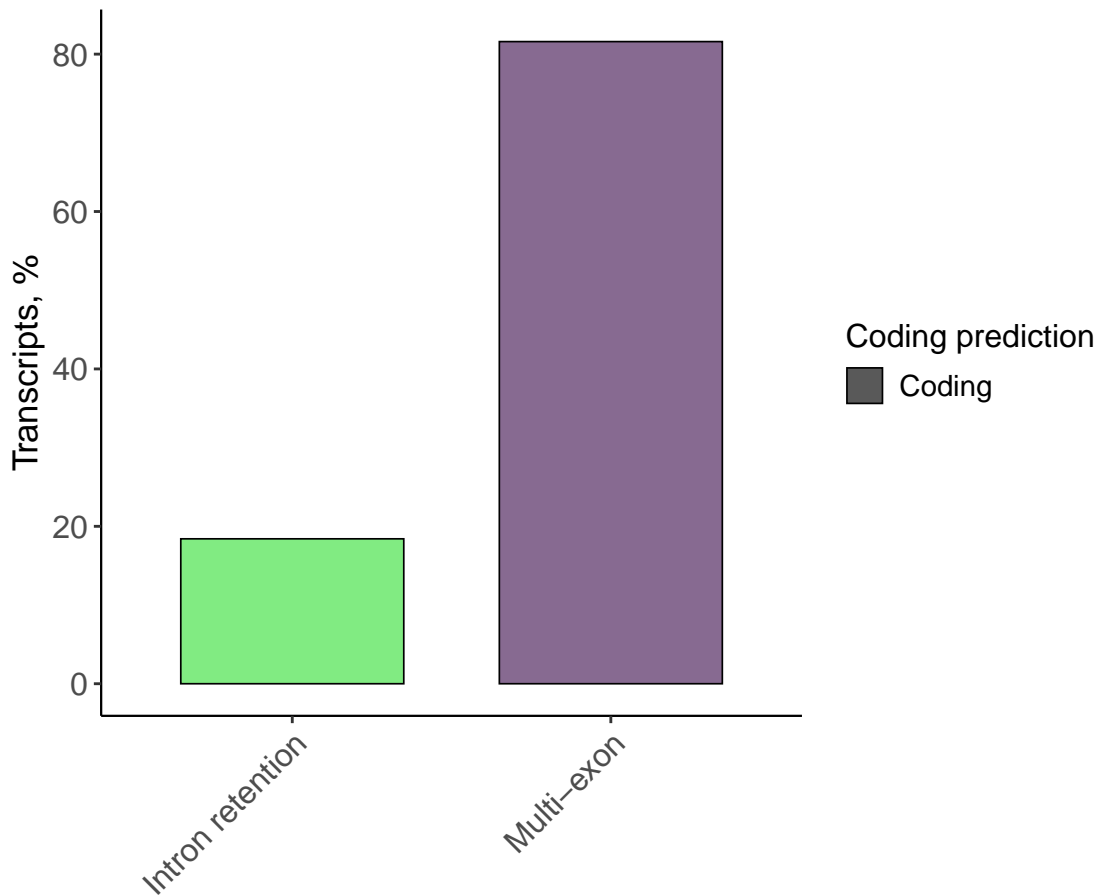
Isoform Distribution Across Genic Genomic



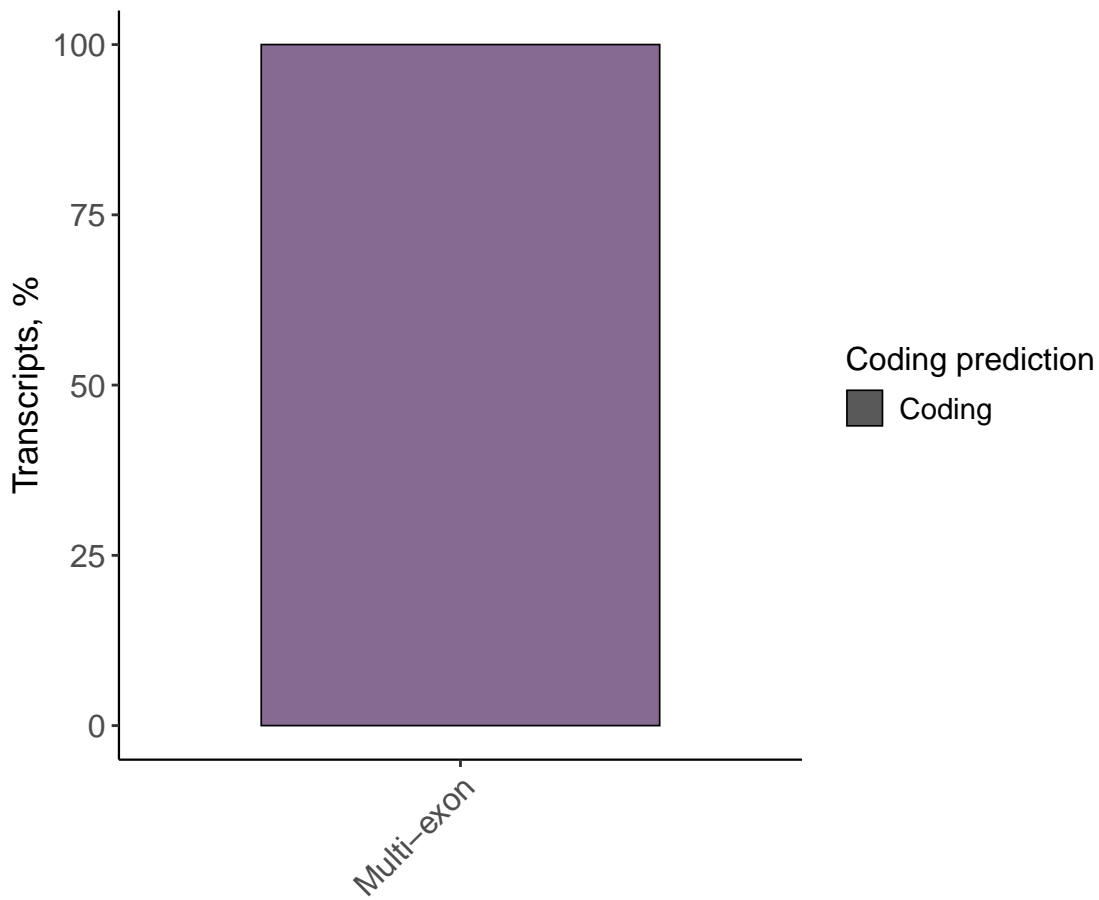
Isoform Distribution Across Antisense



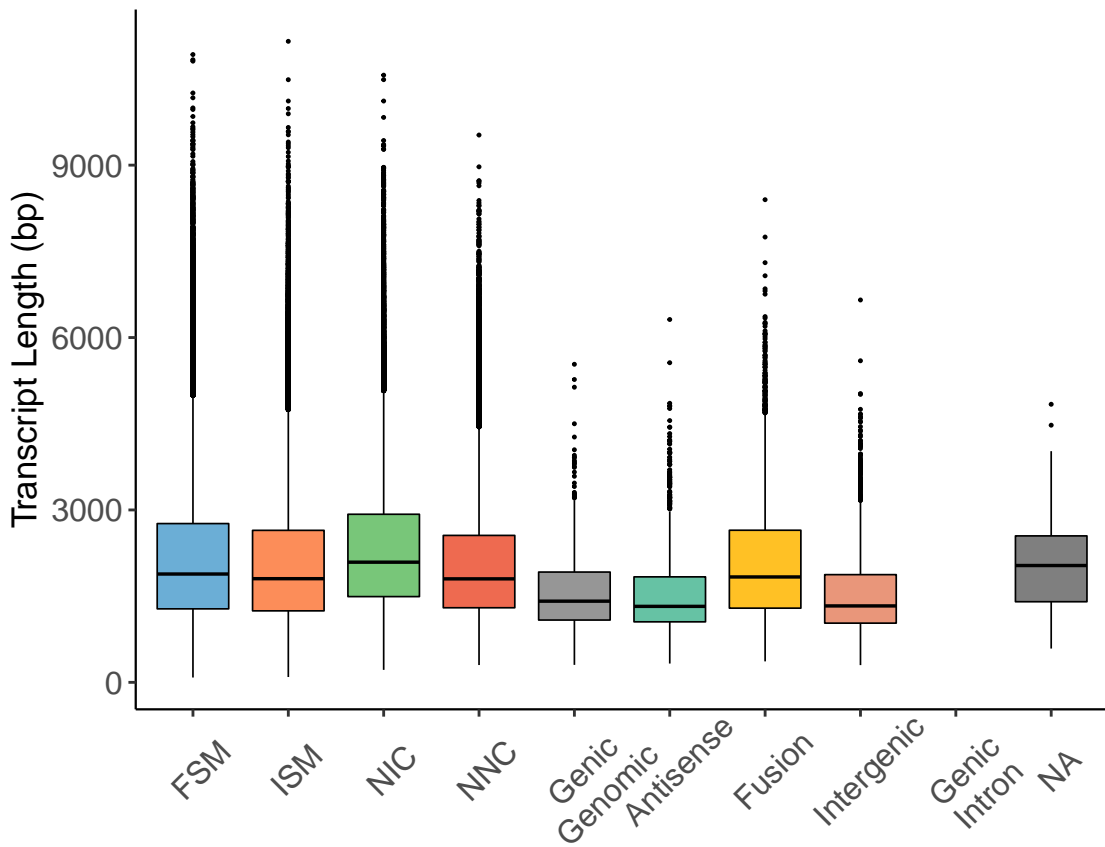
Isoform Distribution Across Fusion



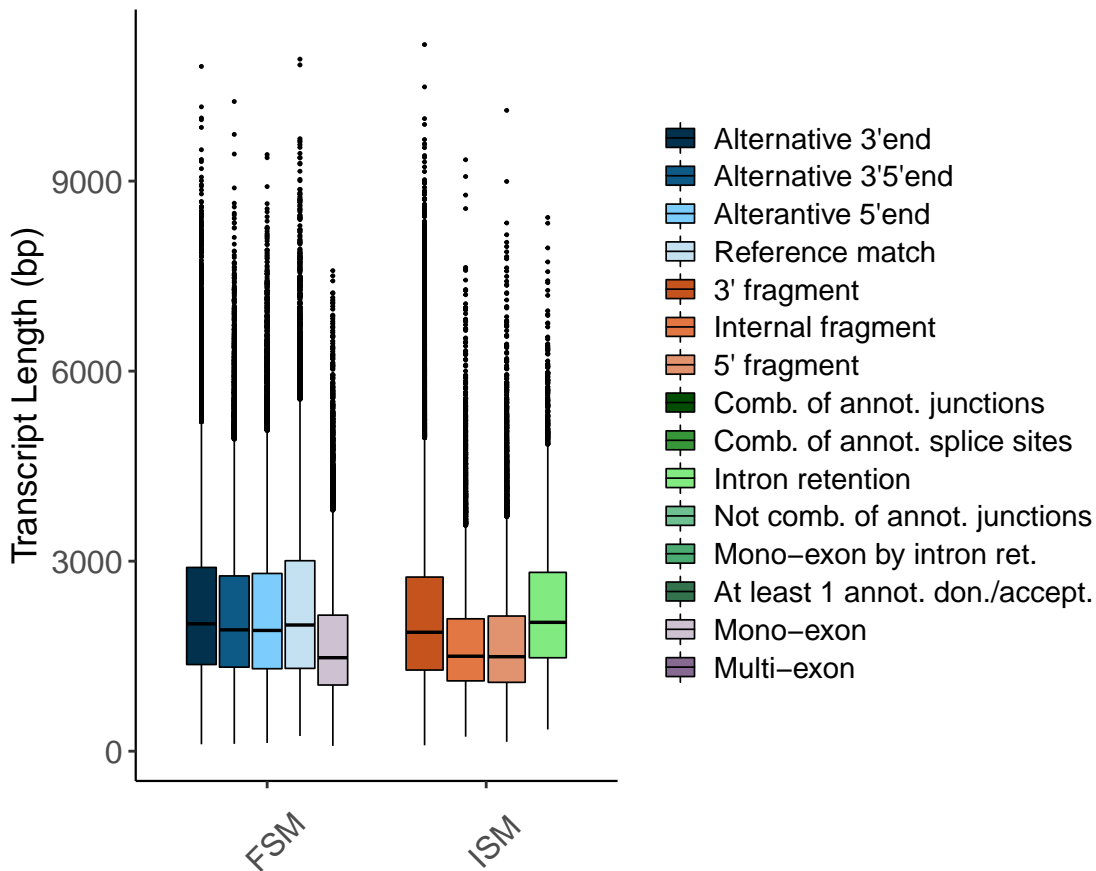
Isoform Distribution Across Intergenic



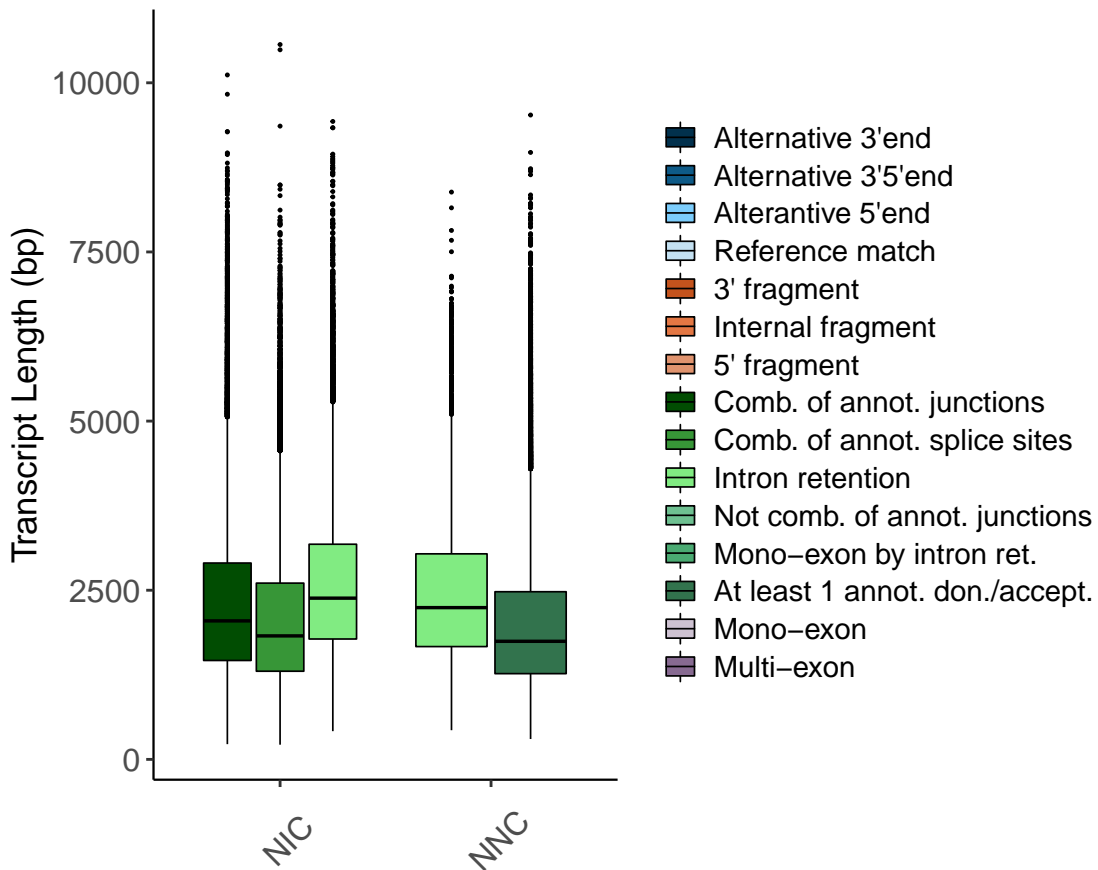
Transcript Lengths by Structural Classification



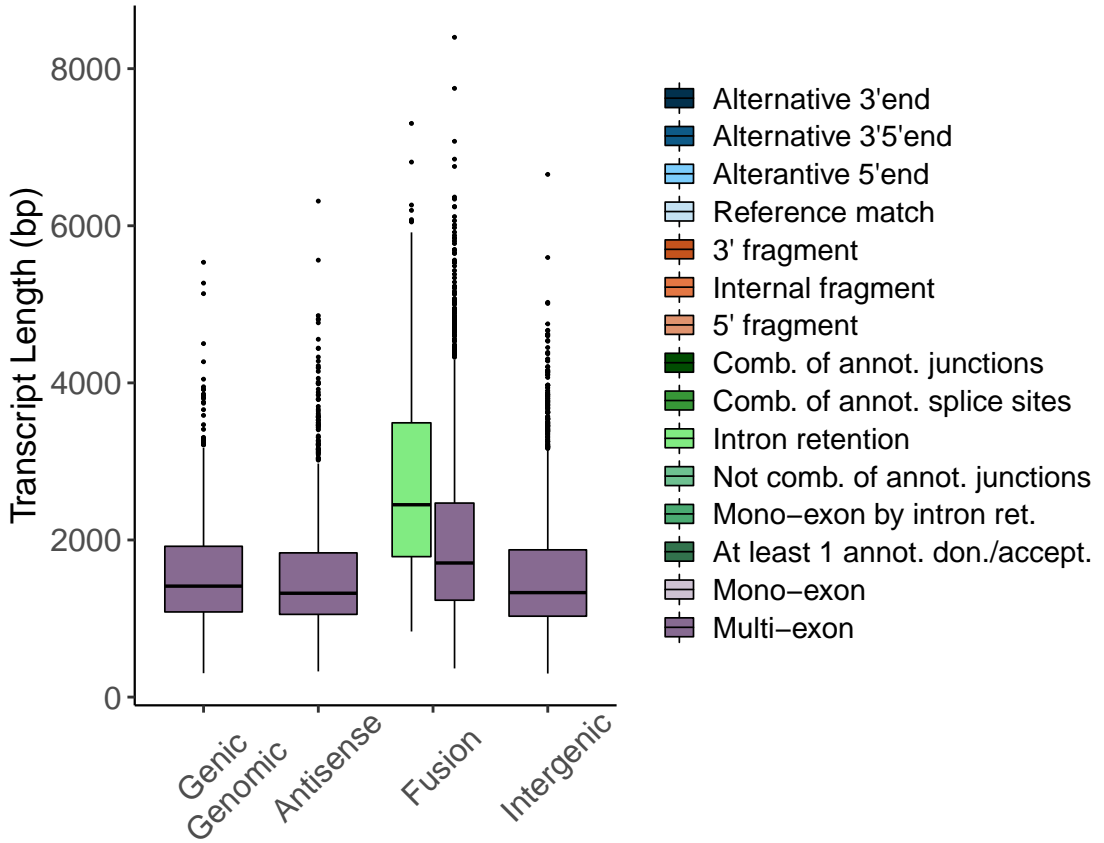
Transcript Lengths by Subcategory



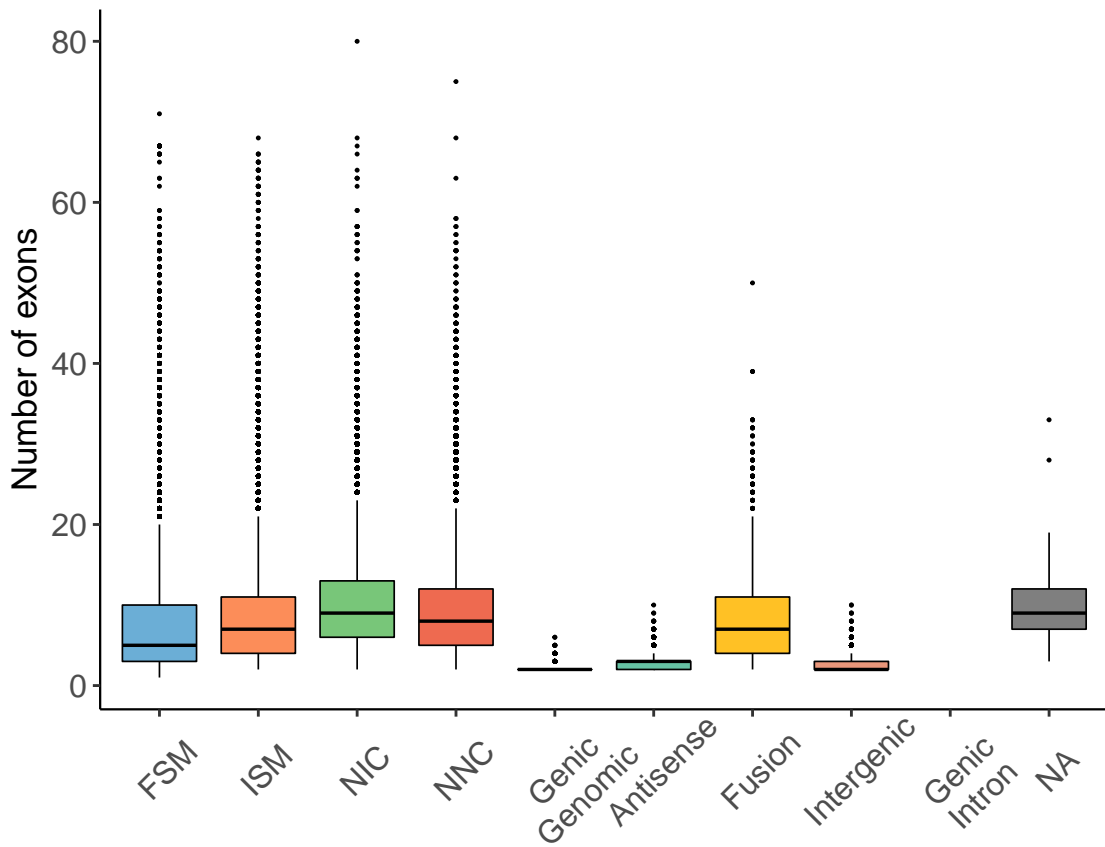
Transcript Lengths by Subcategory



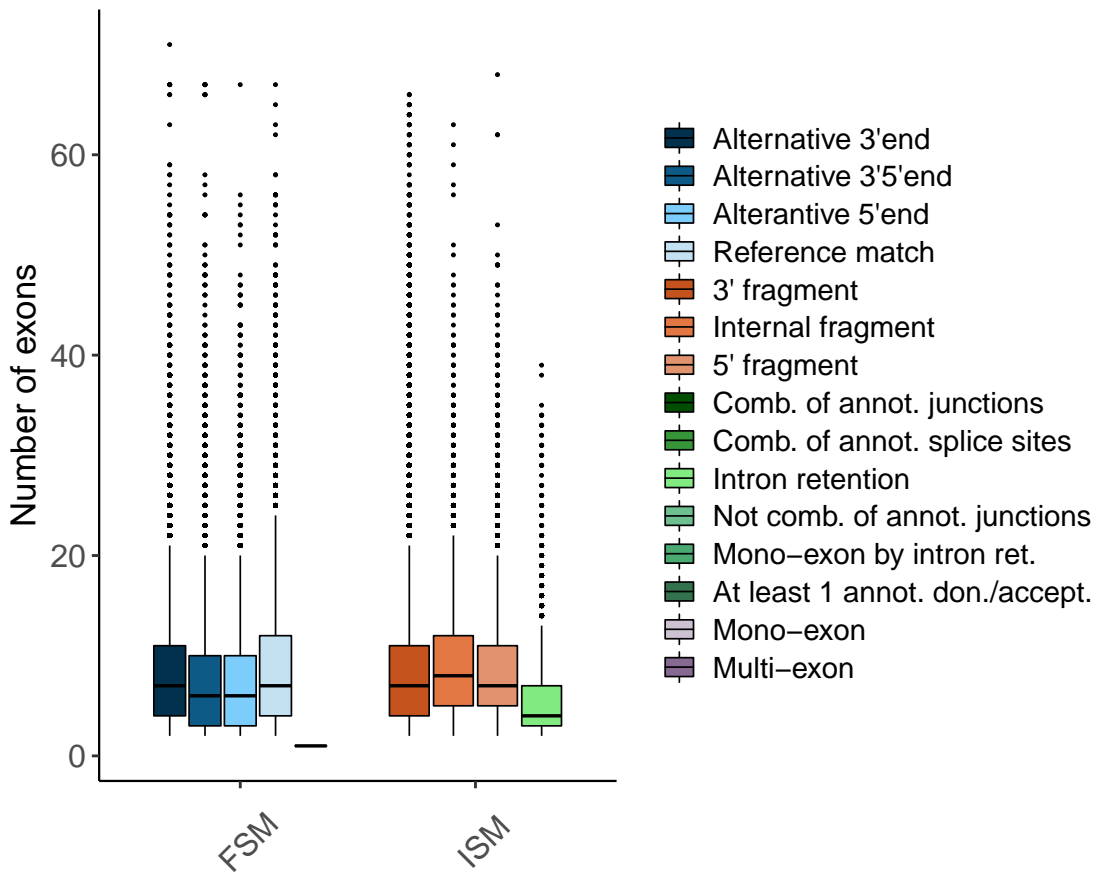
Transcript Lengths by Subcategory



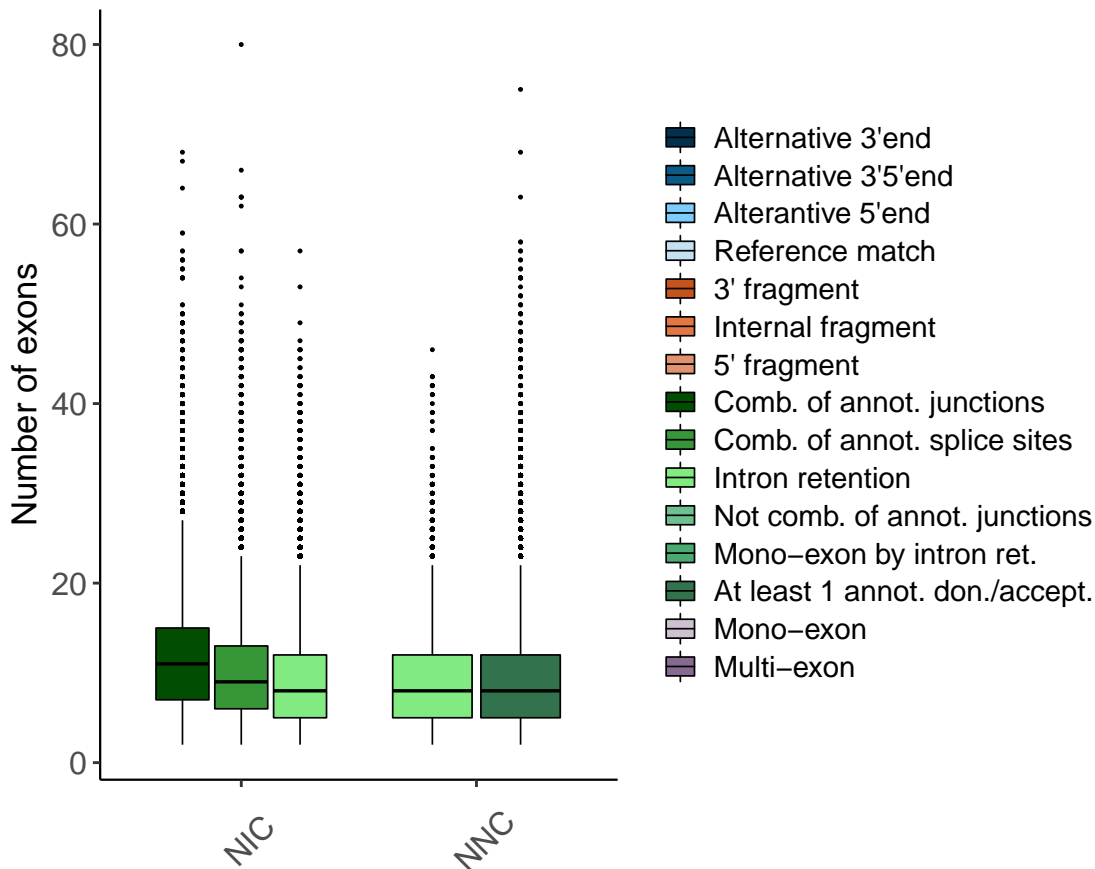
Exon Counts by Structural Classification



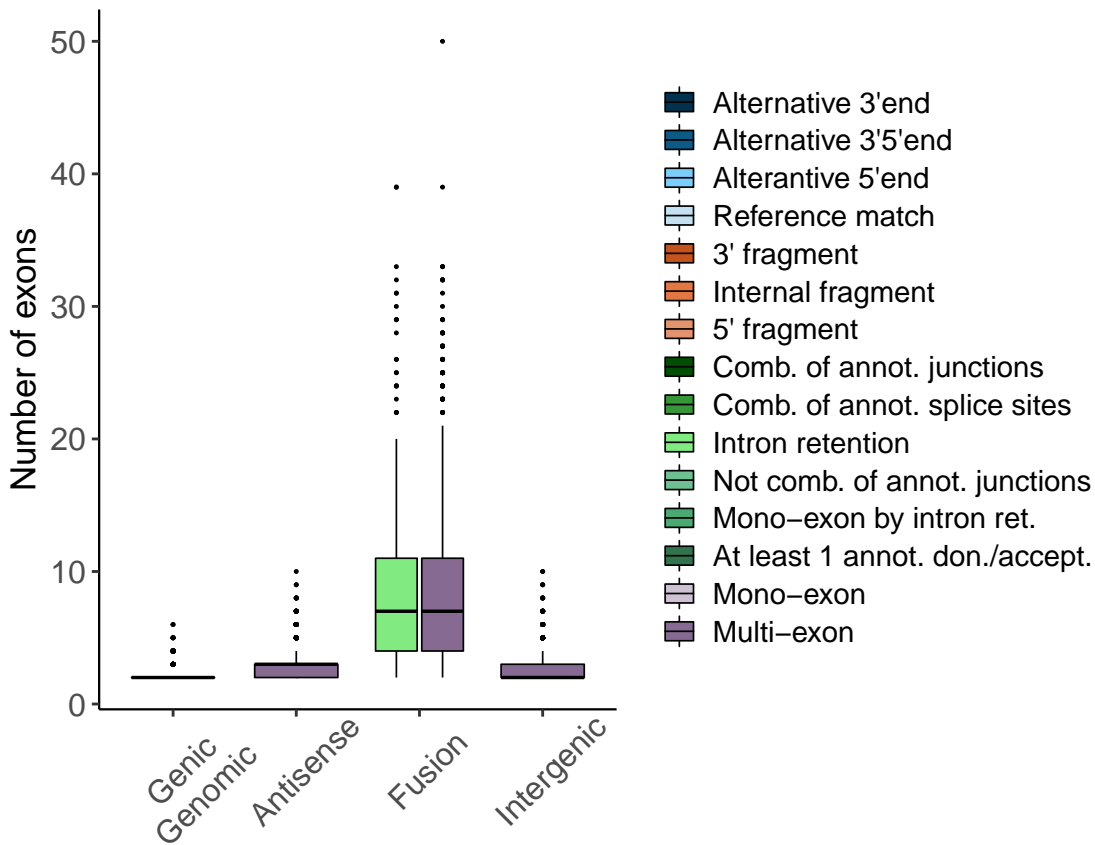
Exon Counts by Subcategory



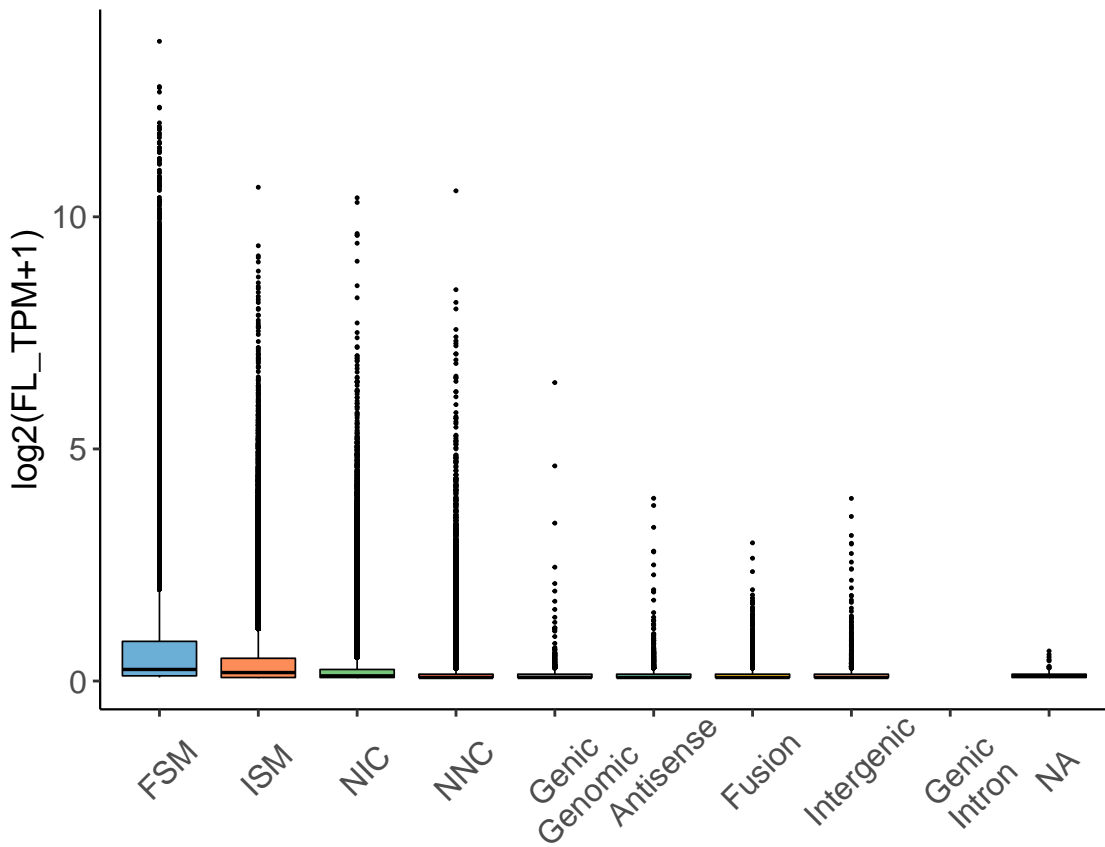
Exon Counts by Subcategory



Exon Counts by Subcategory

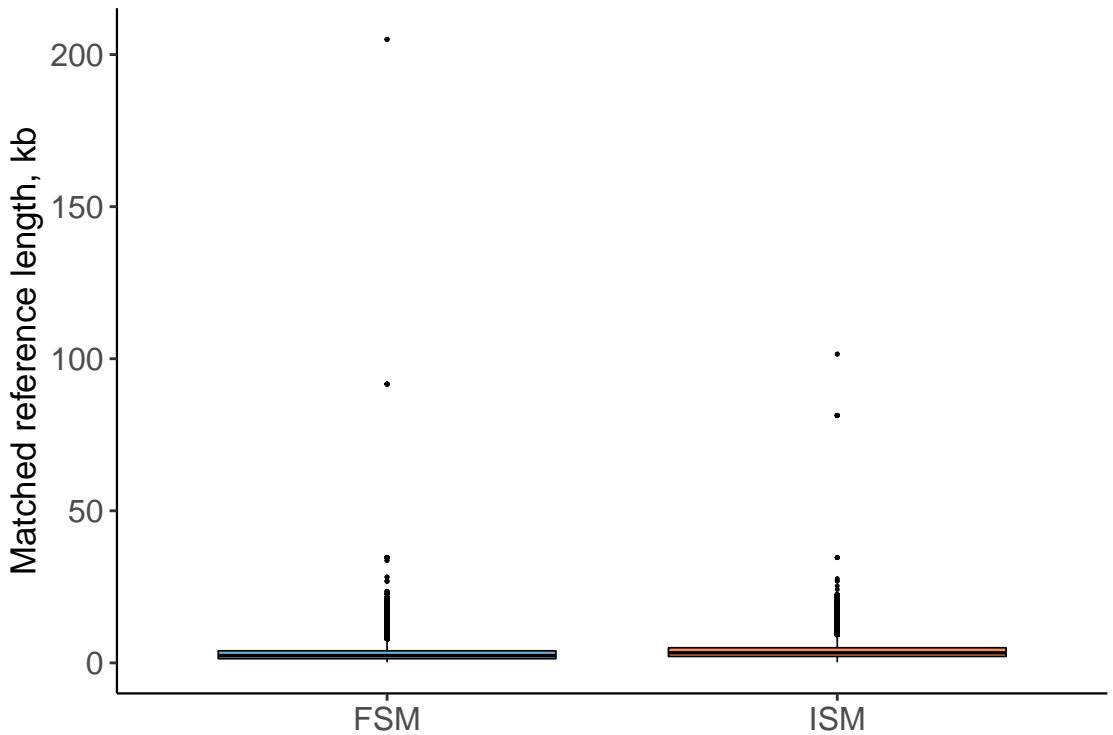


Long Reads Count by Structural Category



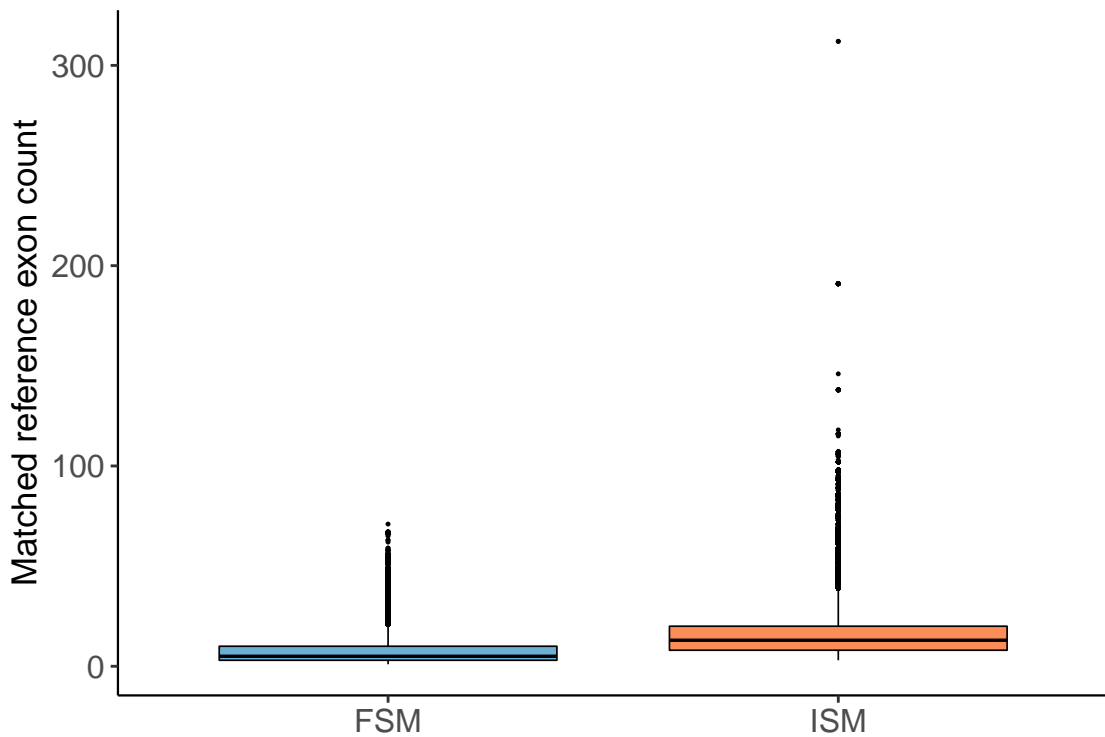
Length Distribution of Matched Reference Transcripts

Applicable Only to FSM and ISM Categories



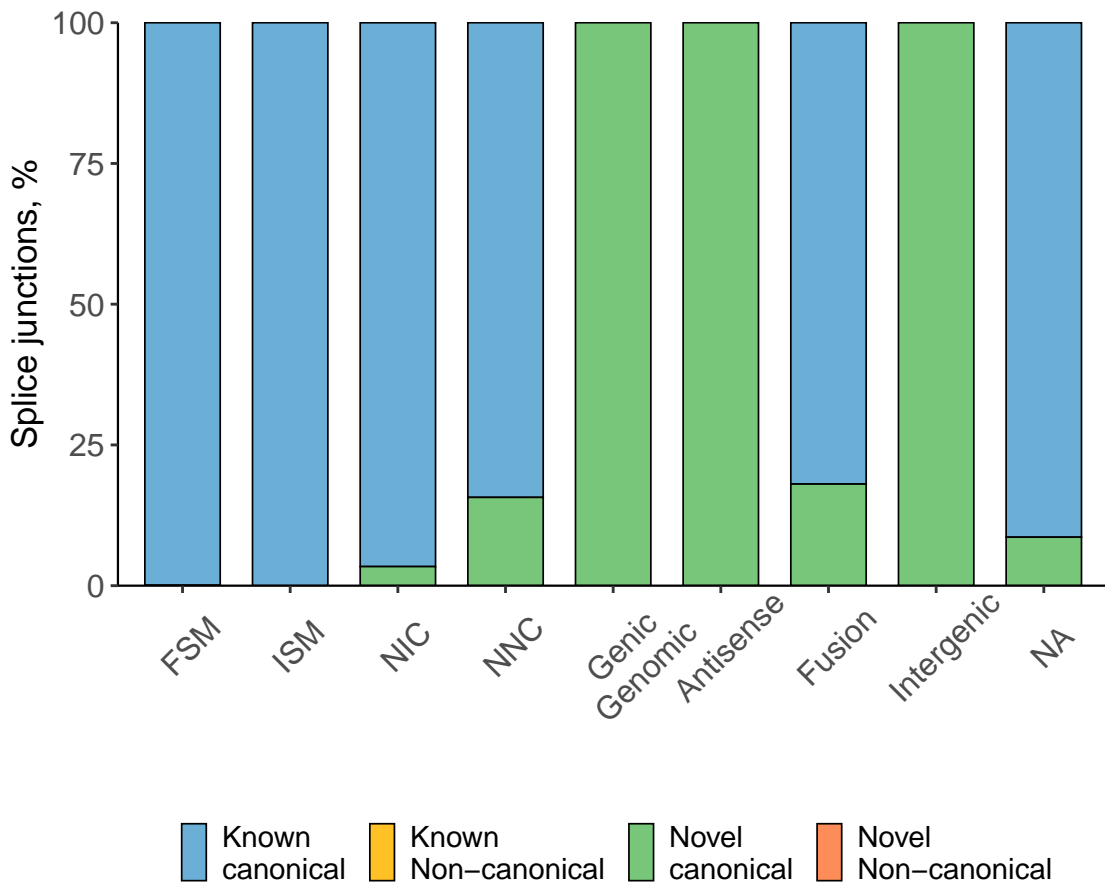
Exon Count Distribution of Matched Reference Transcripts

Applicable Only to FSM and ISM Categories

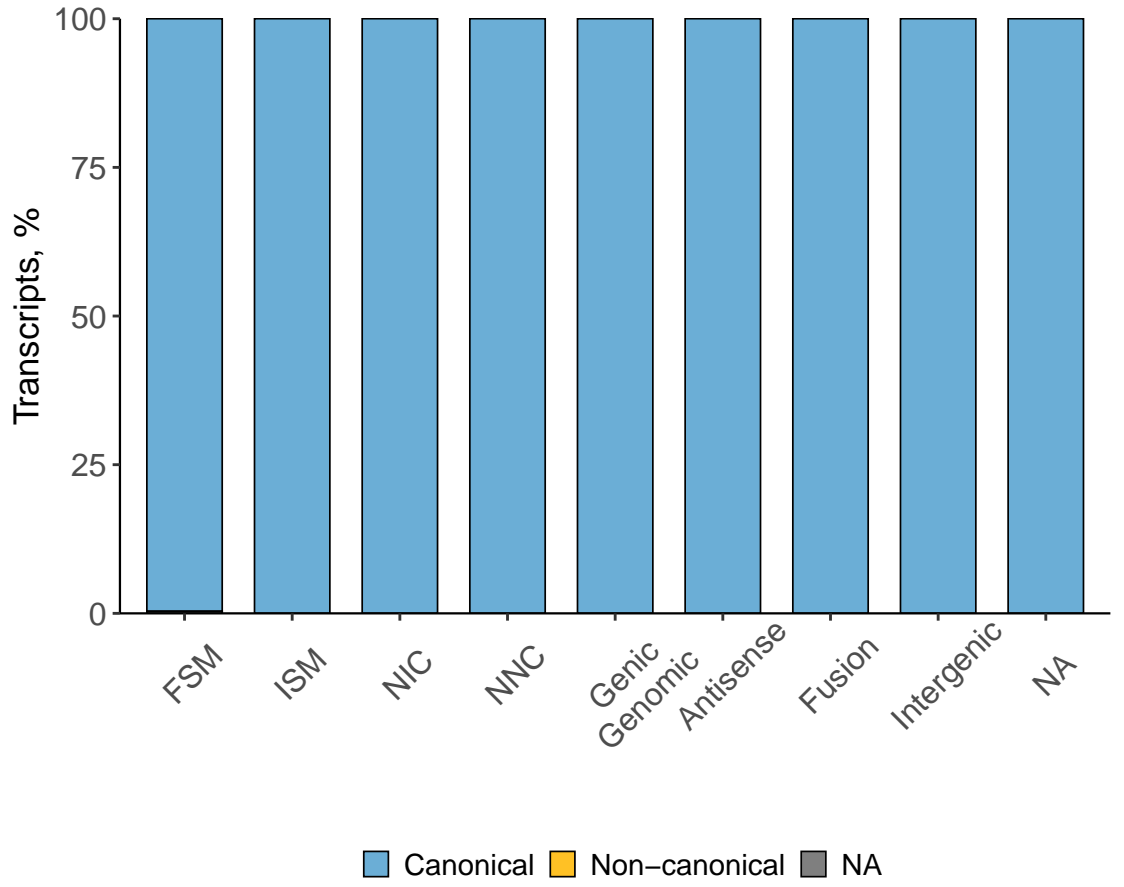


Splice Junction Characterization

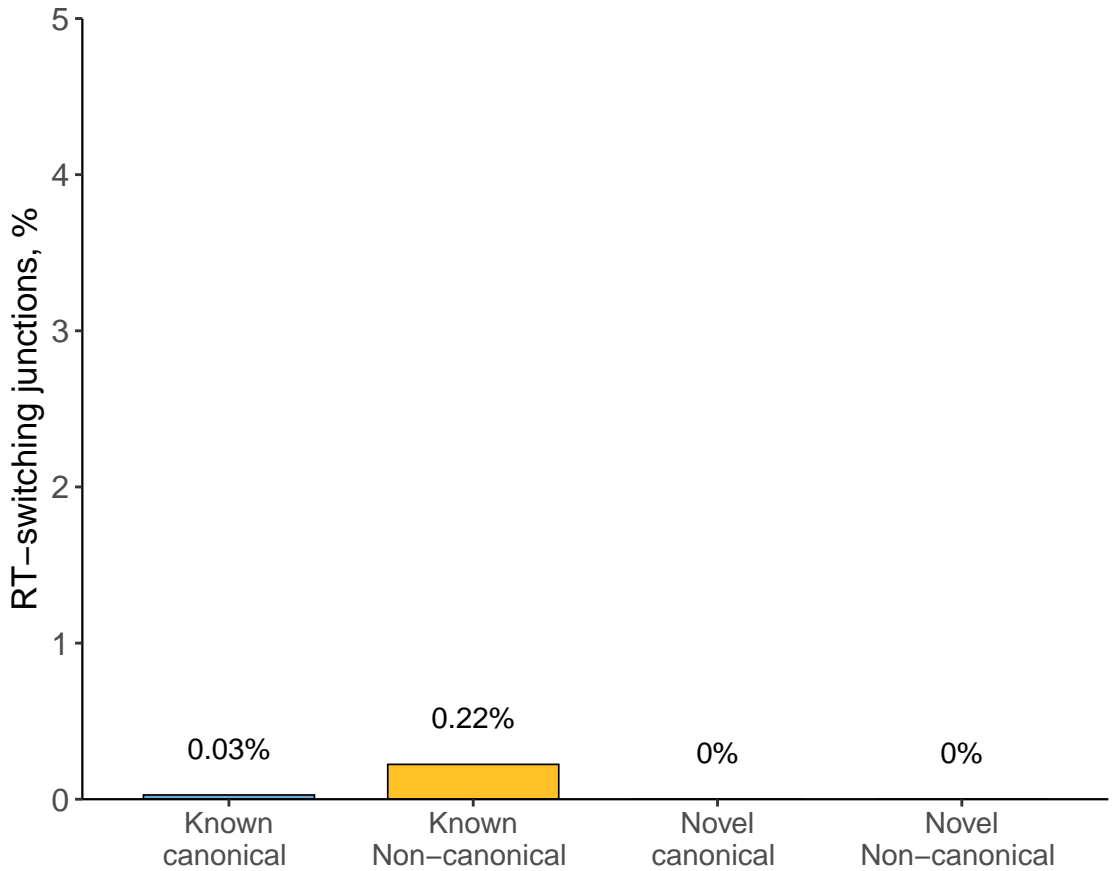
Distribution of Splice Junctions by Structural Classification



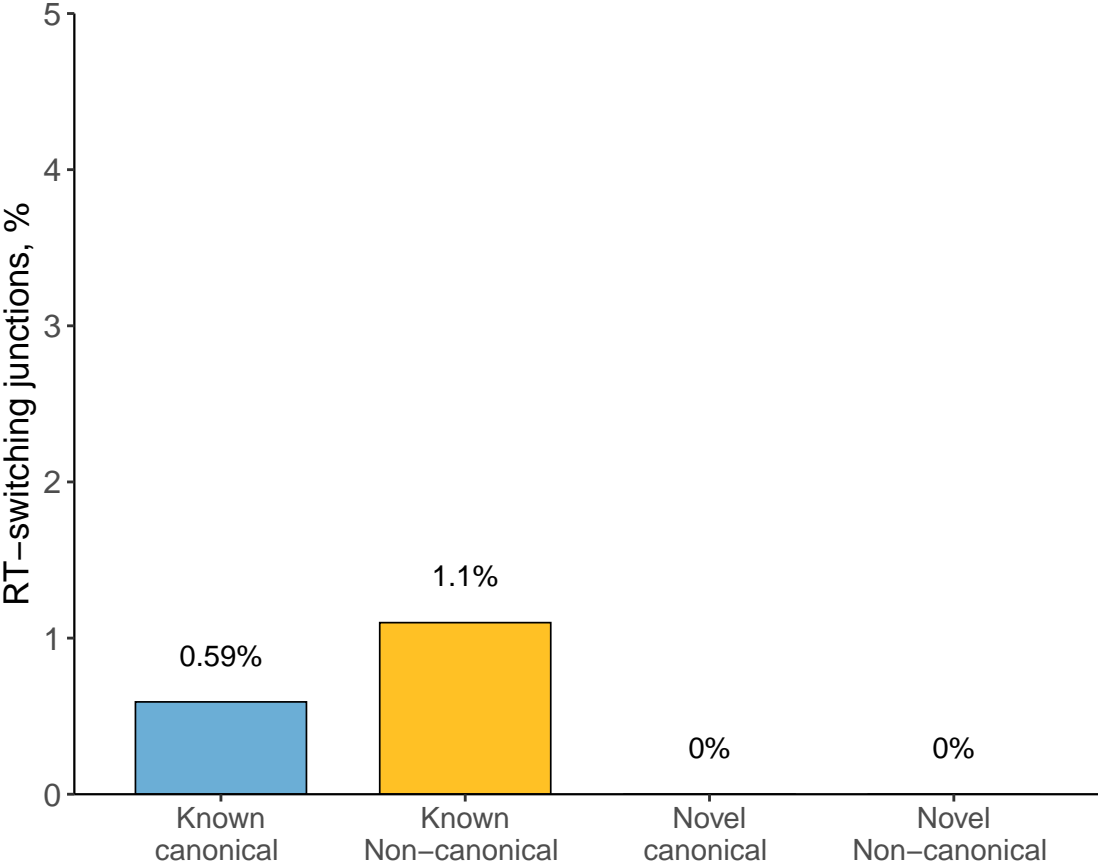
Distribution of Transcripts by Splice Junctions



RT-Switching All Junctions



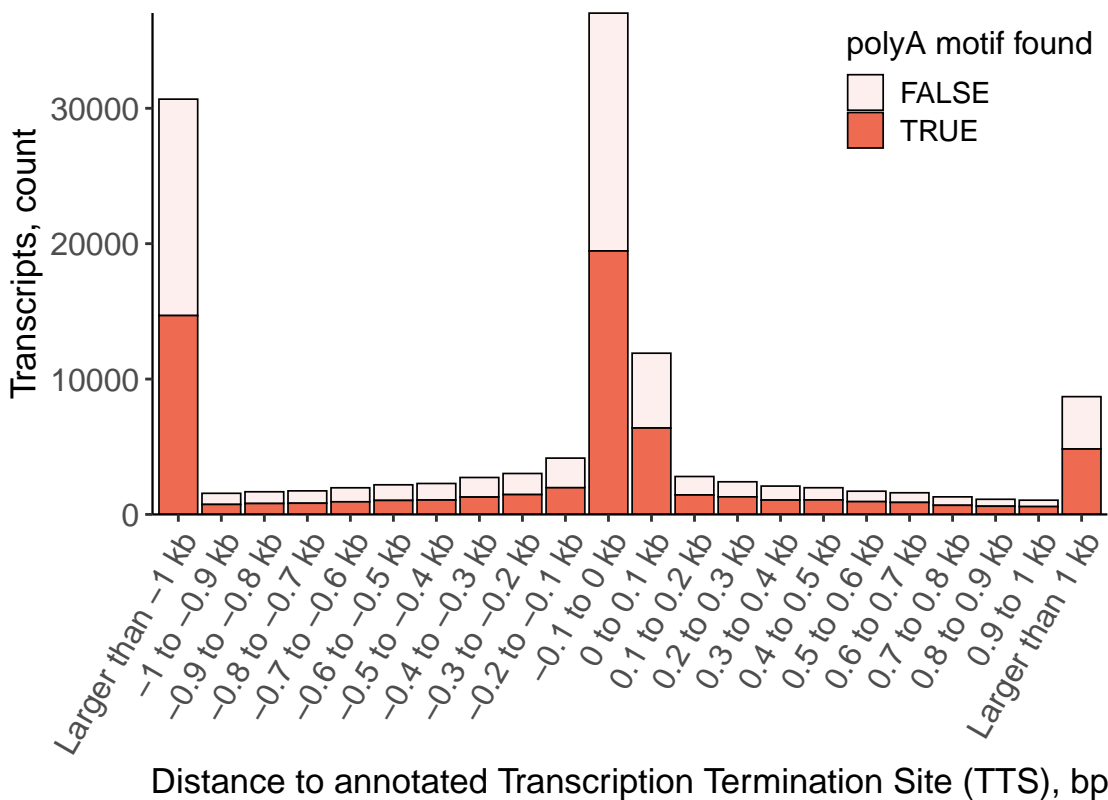
Unique Junctions RT-switching



Comparison With Annotated TSS and TTS

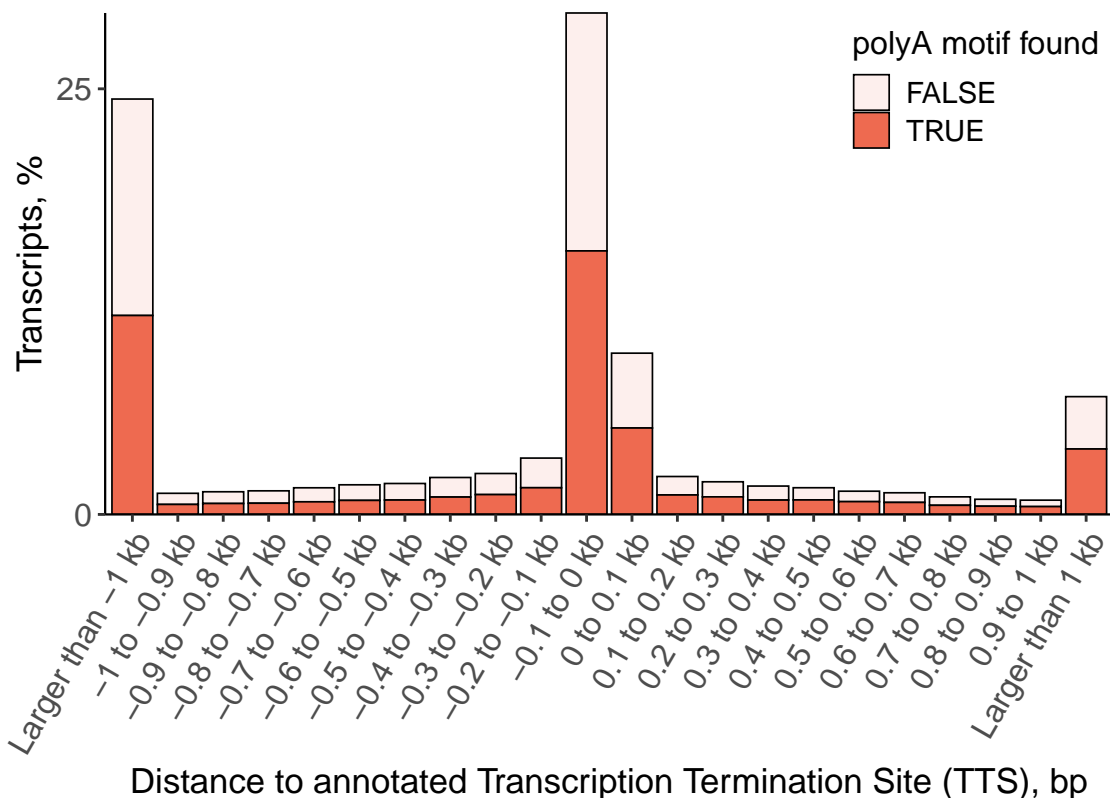
Distance to annotated Transcription Termination Site (TTS) FSM

Negative values indicate upstream of annotated termination site



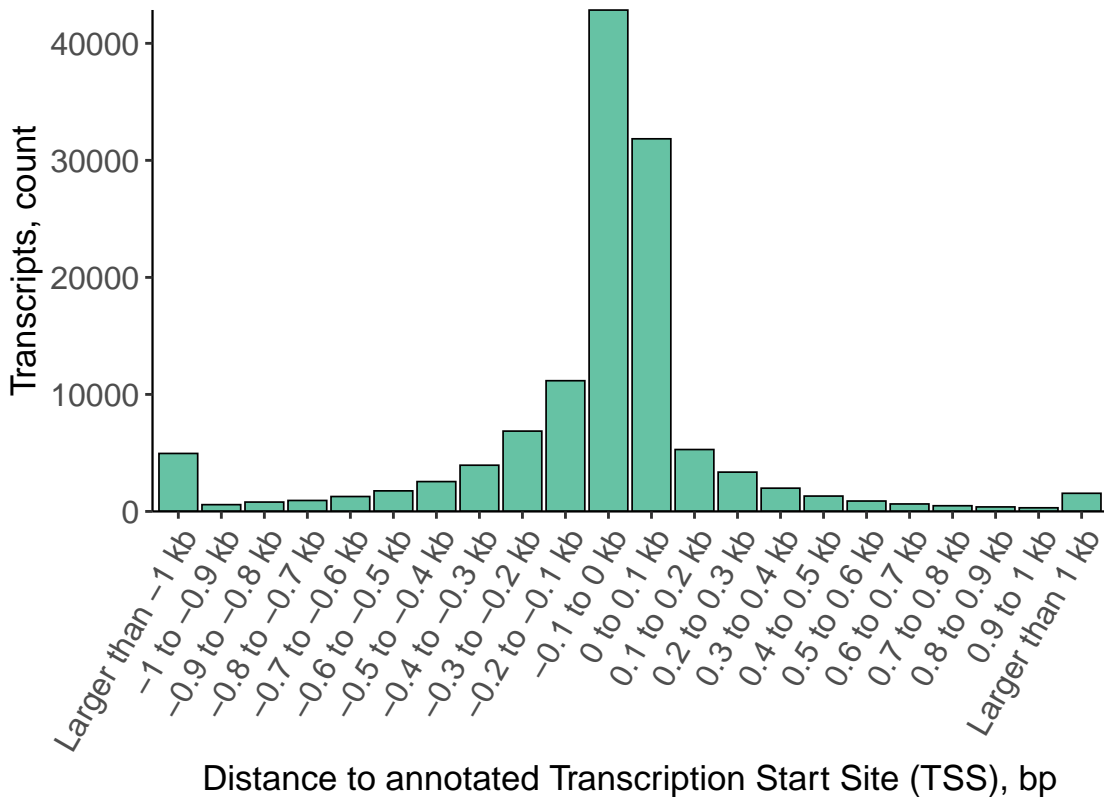
Distance to annotated Transcription Termination Site (TTS) FSM

Negative values indicate upstream of annotated termination site



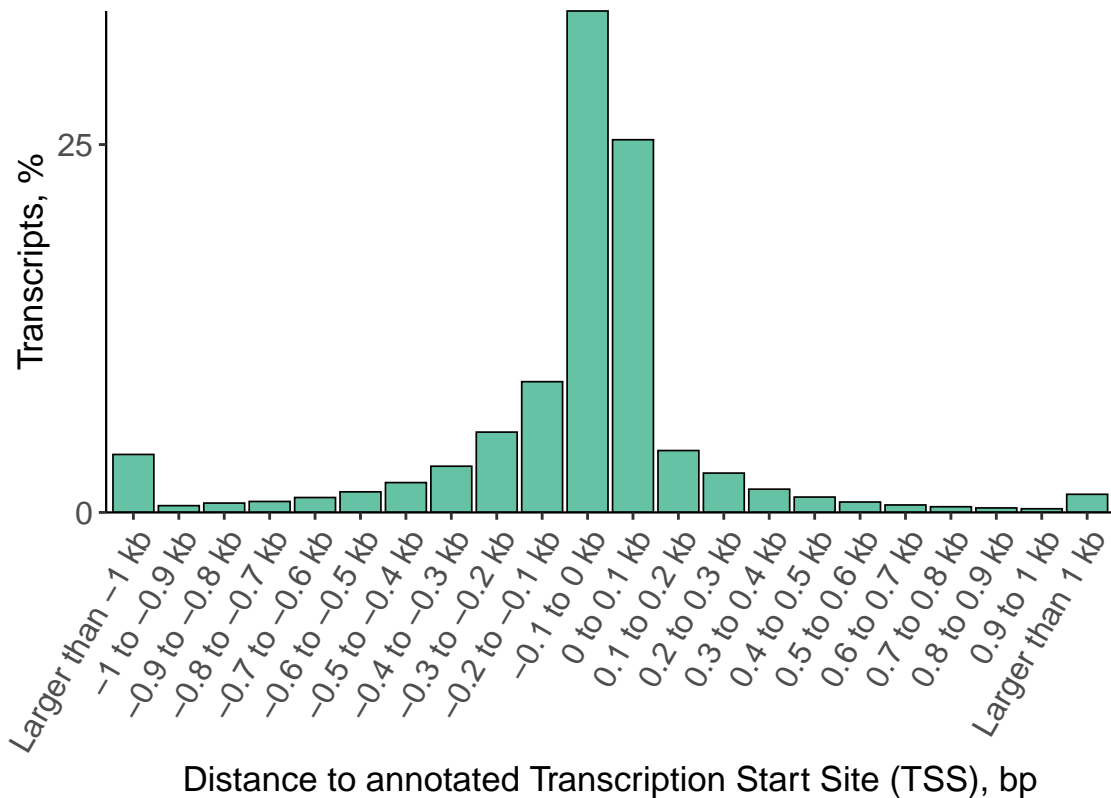
Distance to Annotated Transcription Start Site for FSM

Negative values indicate downstream of annotated TSS



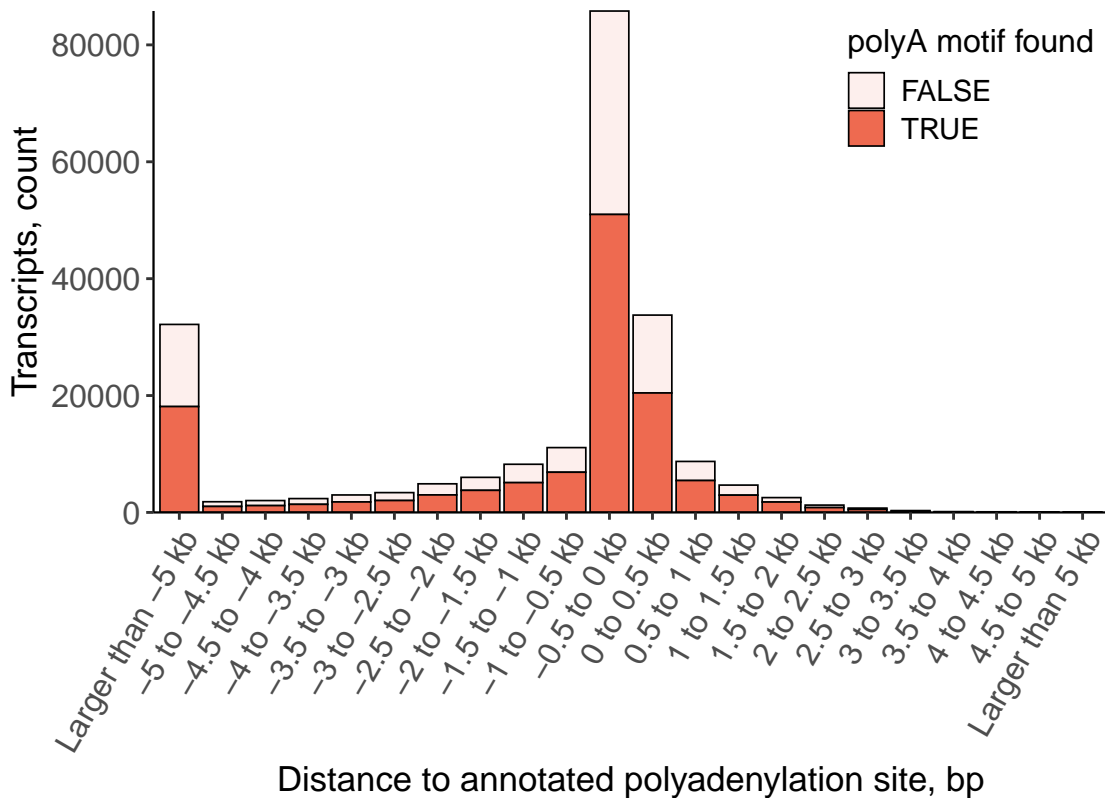
Distance to Annotated Transcription Start Site for FSM

Negative values indicate downstream of annotated TSS



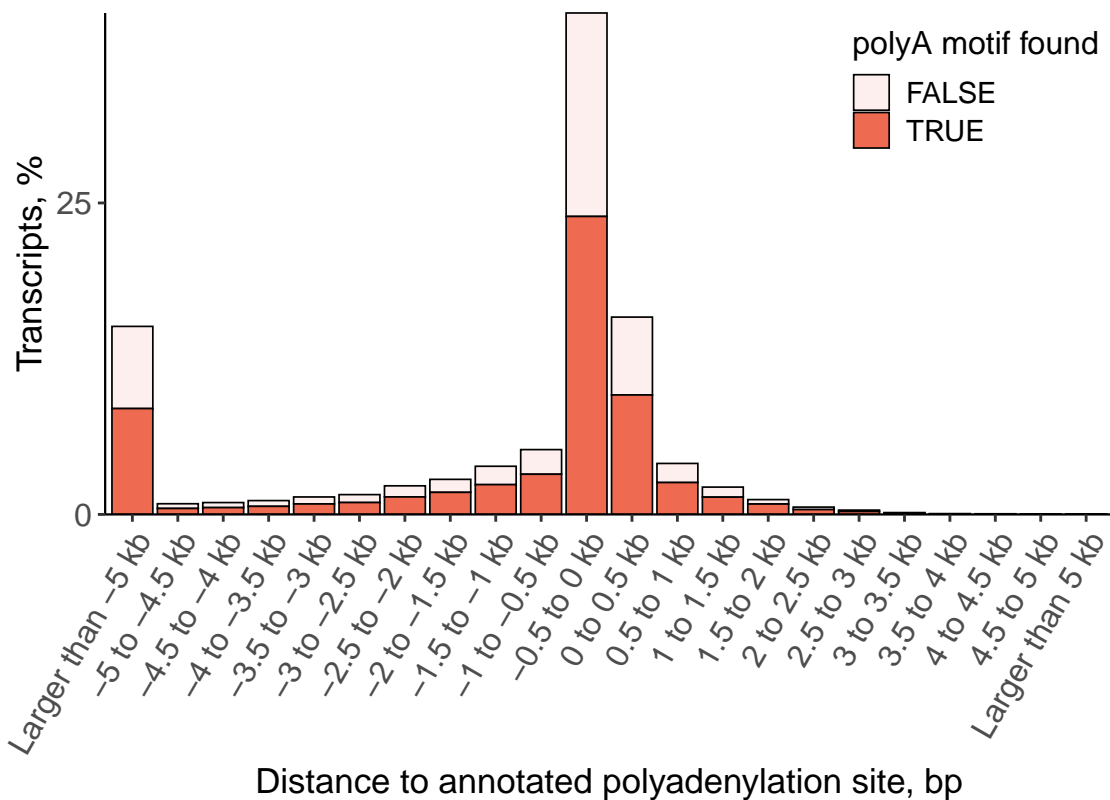
Distance to Annotated Polyadenylation Site for ISM

Negative values indicate upstream of annotated polyA site



Distance to Annotated Polyadenylation Site for ISM

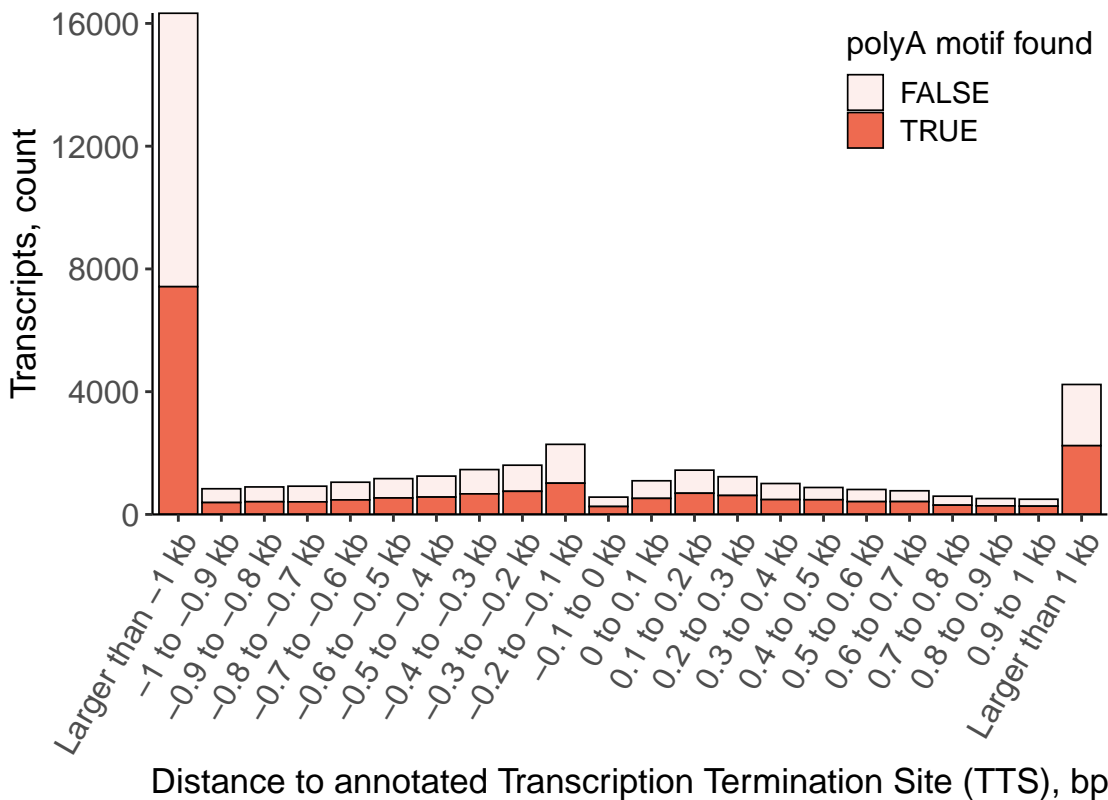
Negative values indicate upstream of annotated polyA site



*Comparison With Annotated TSS and TTS
by Subcategories*

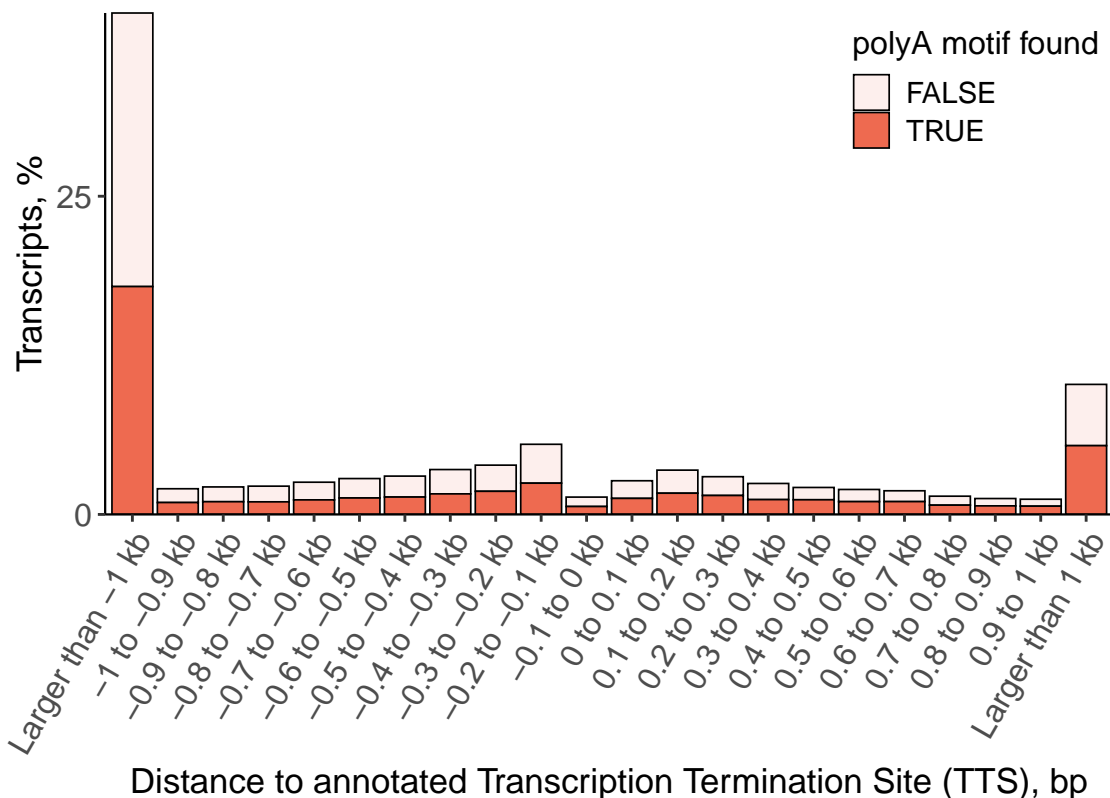
Distance to annotated Transcription Termination Site (TTS) FSM Alternative 3'End

Negative values indicate upstream of annotated termination site



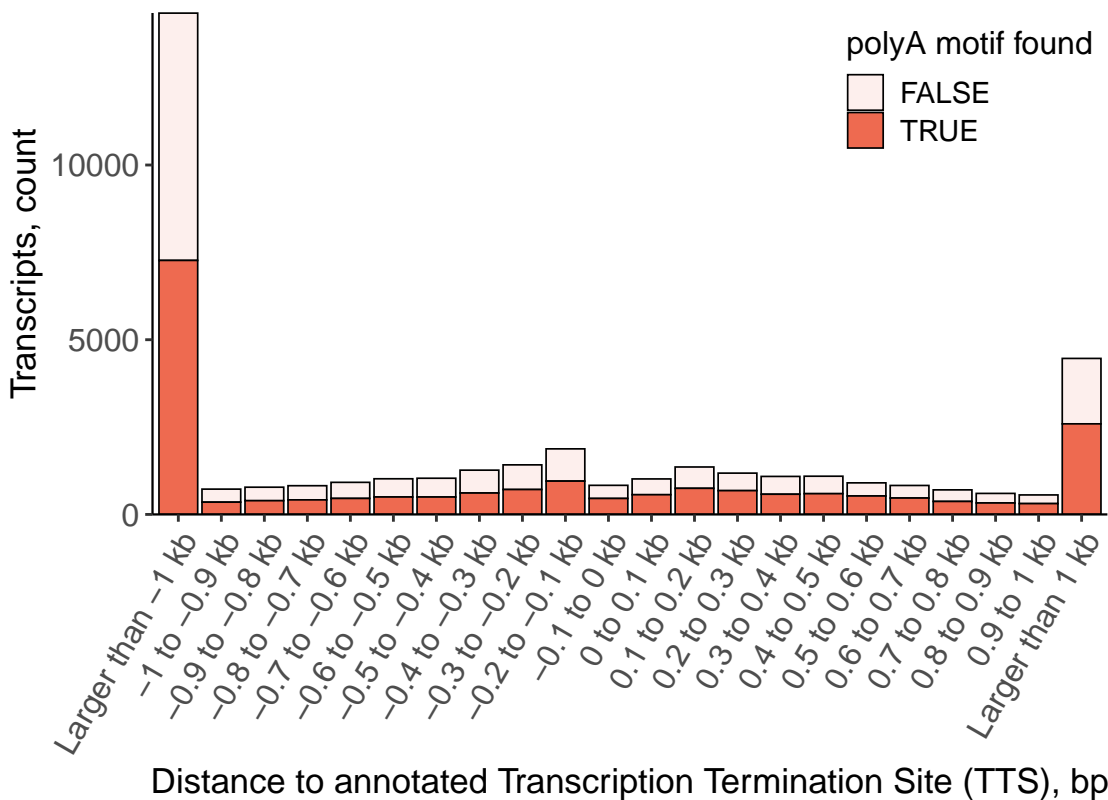
Distance to annotated Transcription Termination Site (TTS) FSM Alternative 3'End

Negative values indicate upstream of annotated termination site



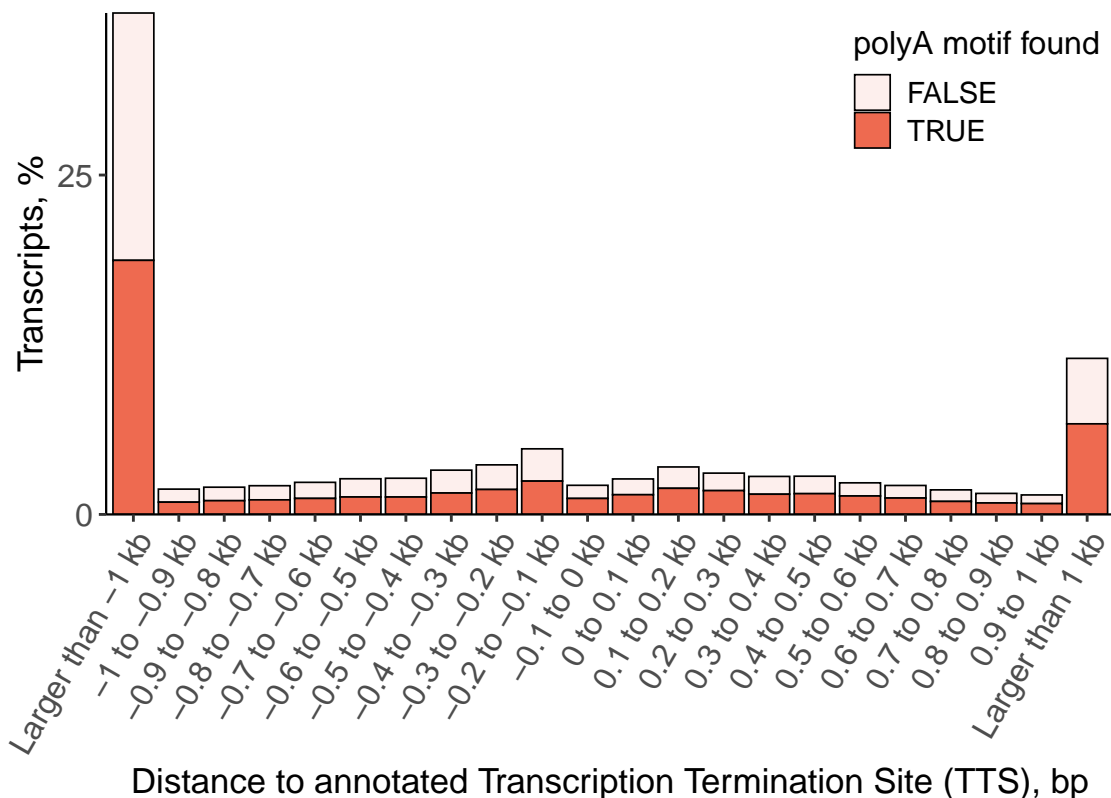
Distance to annotated Transcription Termination Site (TTS) FSM Alternative 3'5'End

Negative values indicate upstream of annotated termination site



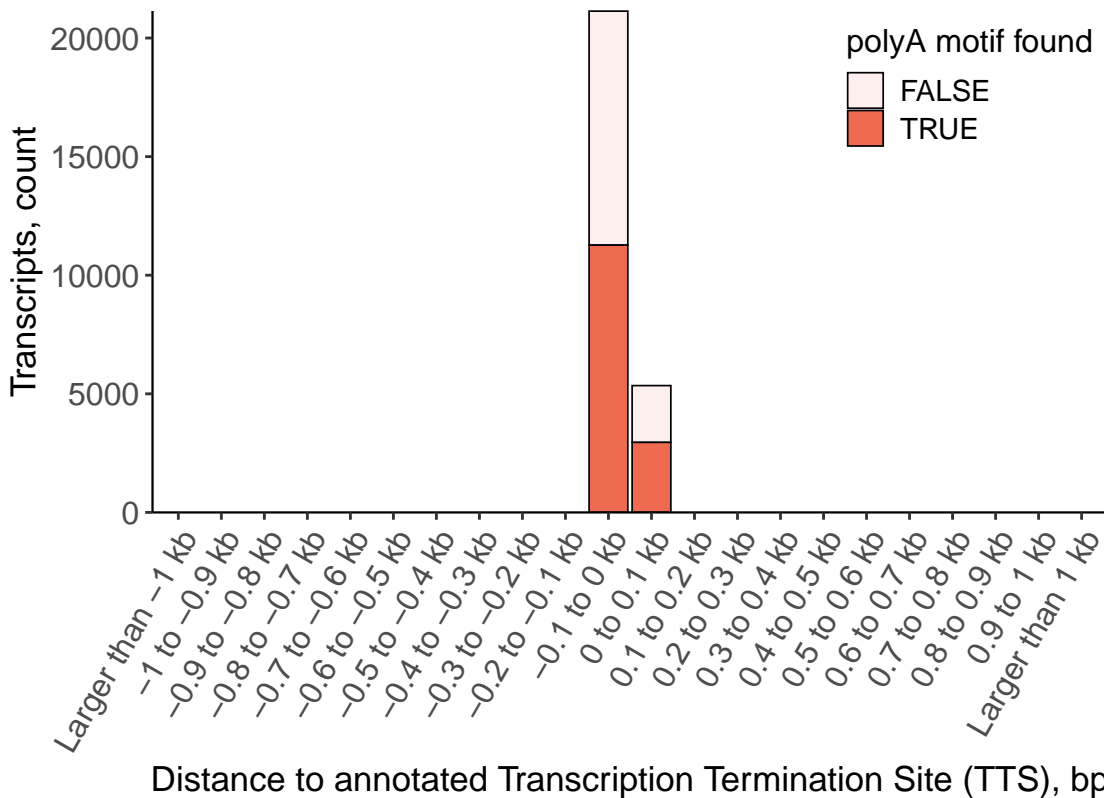
Distance to annotated Transcription Termination Site (TTS) FSM Alternative 3'5'End

Negative values indicate upstream of annotated termination site



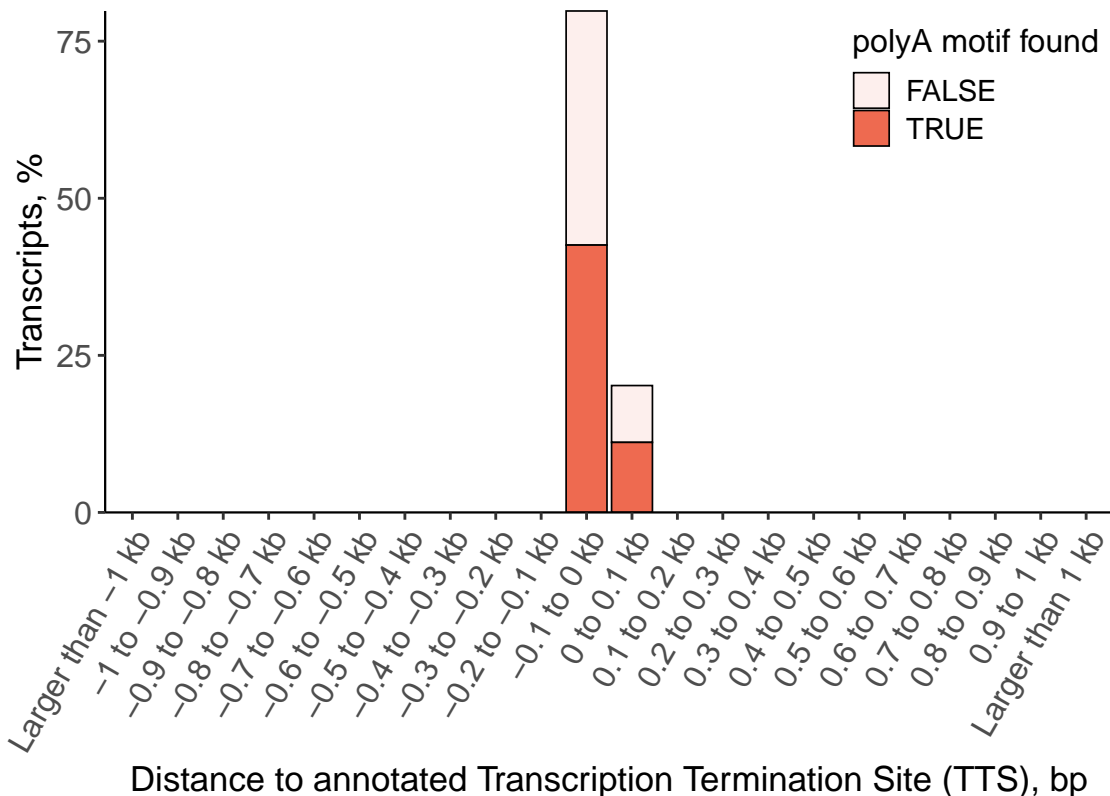
Distance to annotated Transcription Termination Site (TTS) FSM Alternative 5'End

Negative values indicate upstream of annotated termination site



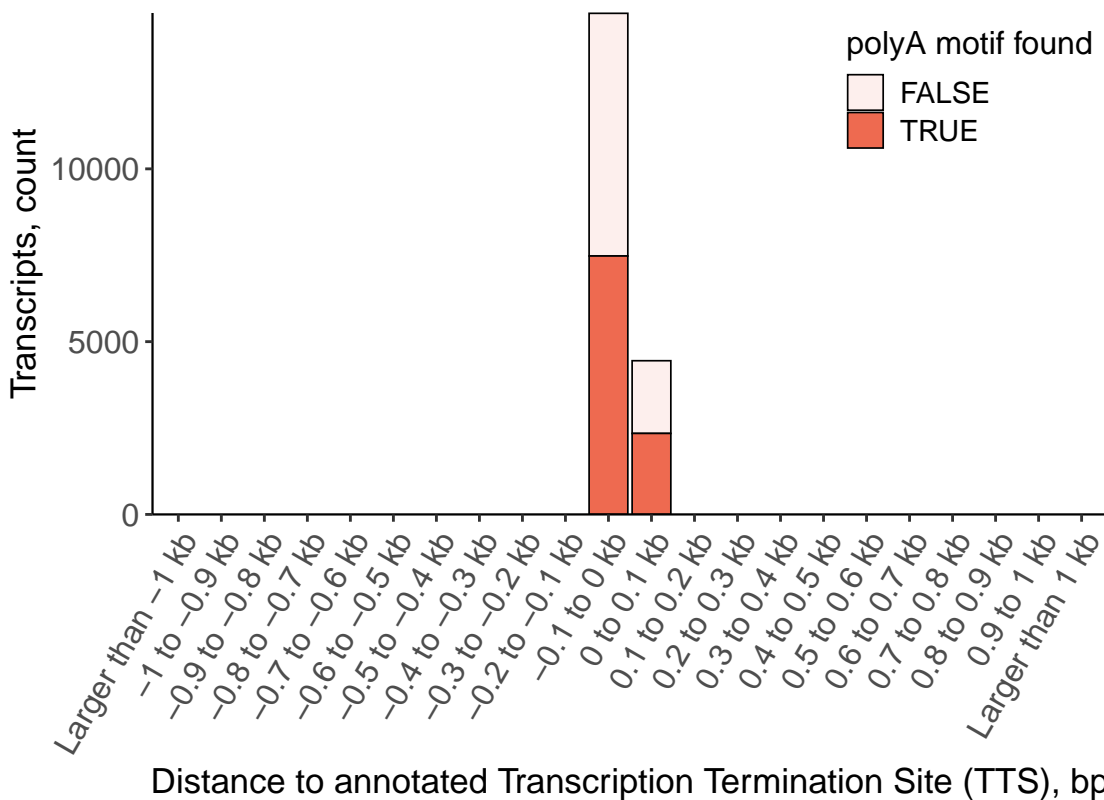
Distance to annotated Transcription Termination Site (TTS) FSM Alternative 5'End

Negative values indicate upstream of annotated termination site



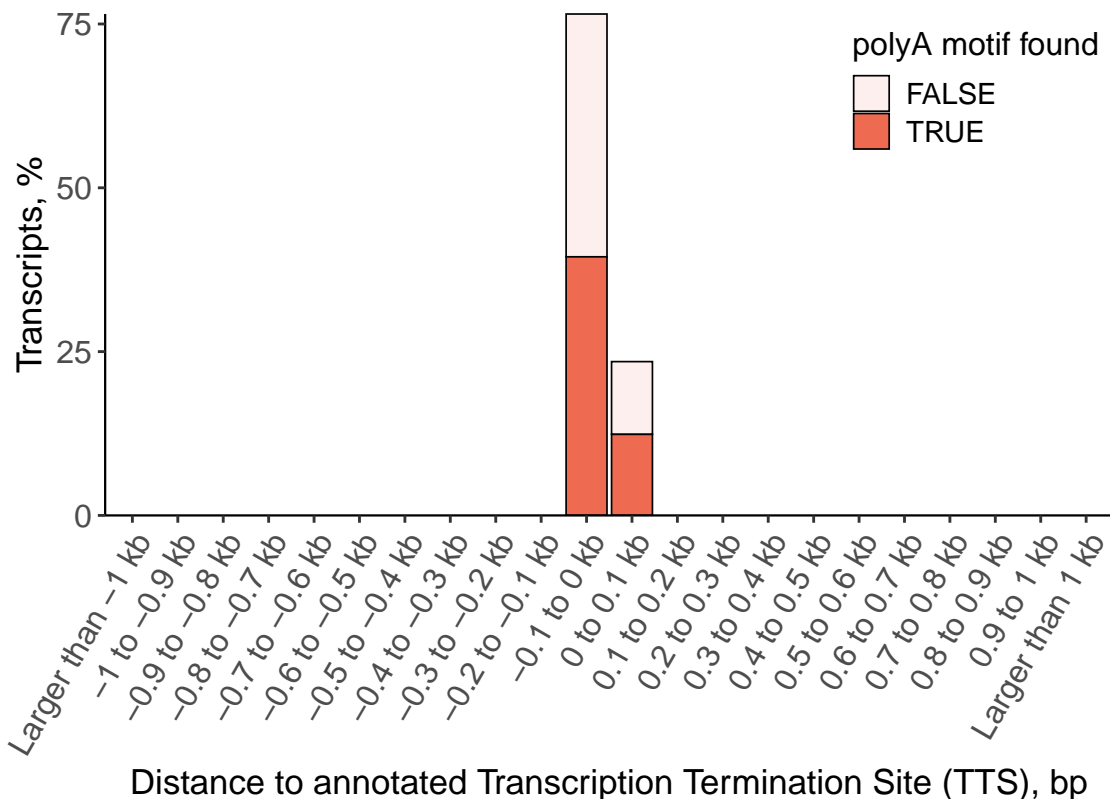
Distance to annotated Transcription Termination Site (TTS) FSM Reference Match

Negative values indicate upstream of annotated termination site



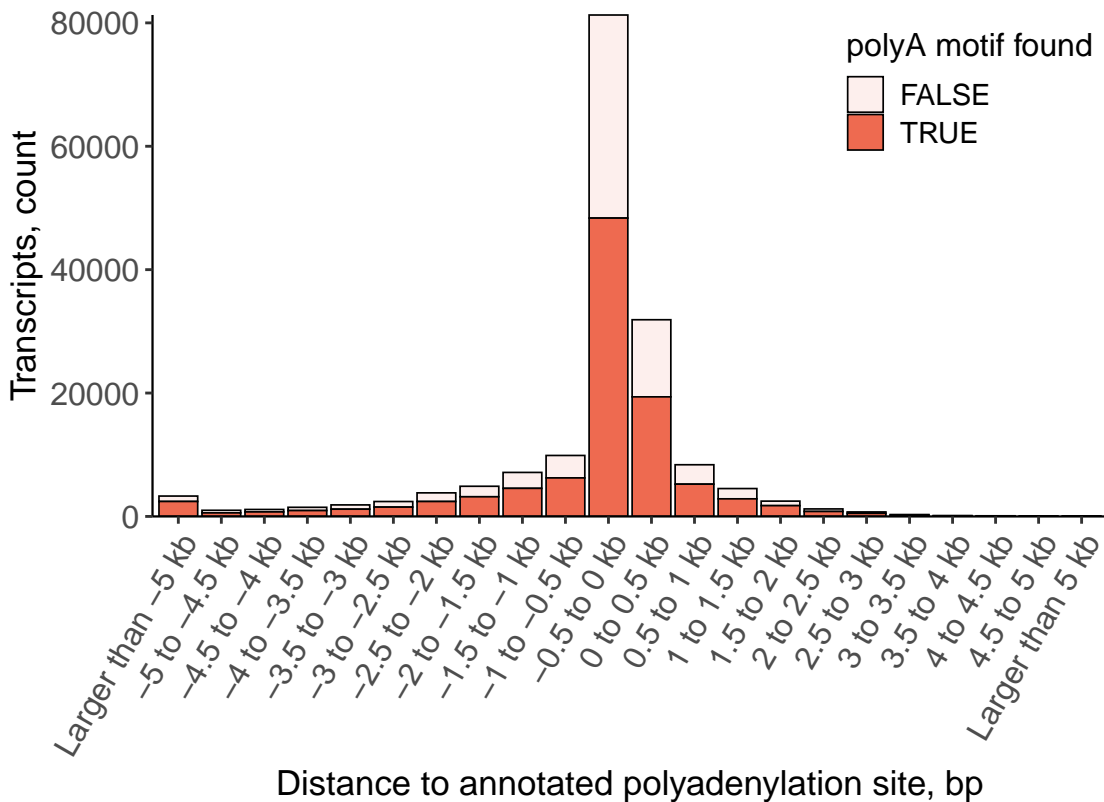
Distance to annotated Transcription Termination Site (TTS) FSM Reference Match

Negative values indicate upstream of annotated termination site



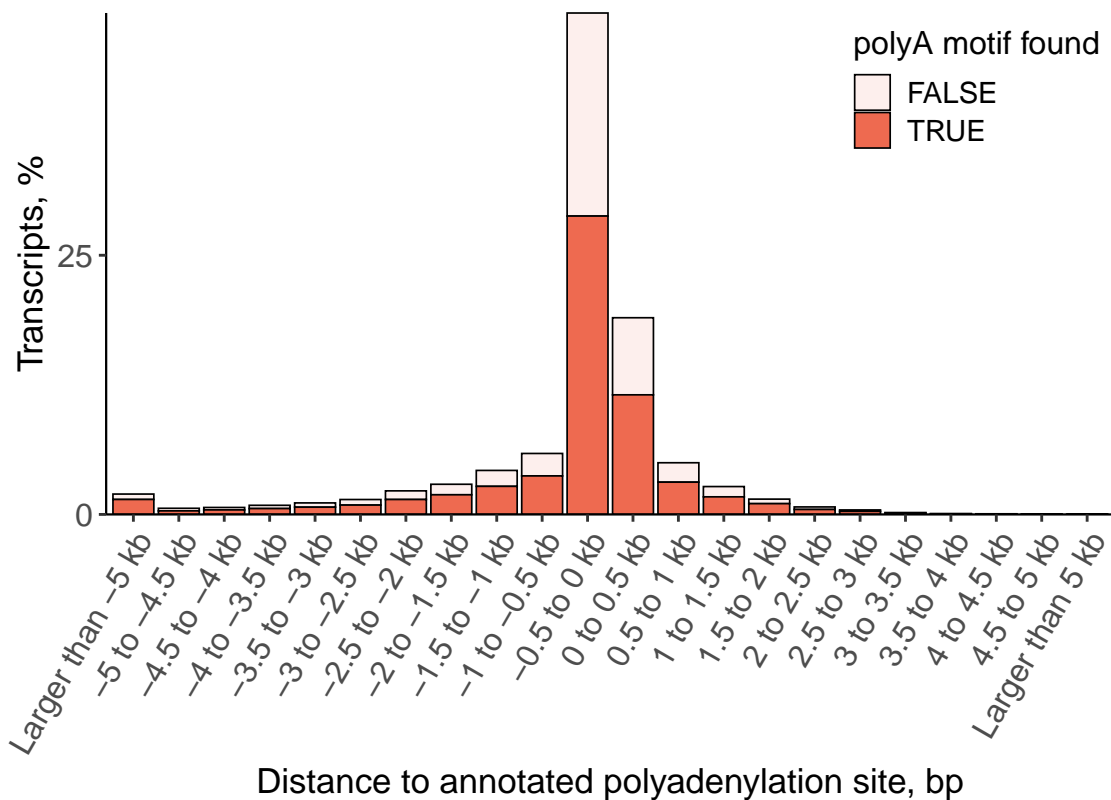
Distance to Annotated Polyadenylation Site for ISM 3' Fragment

Negative values indicate upstream of annotated polyA site



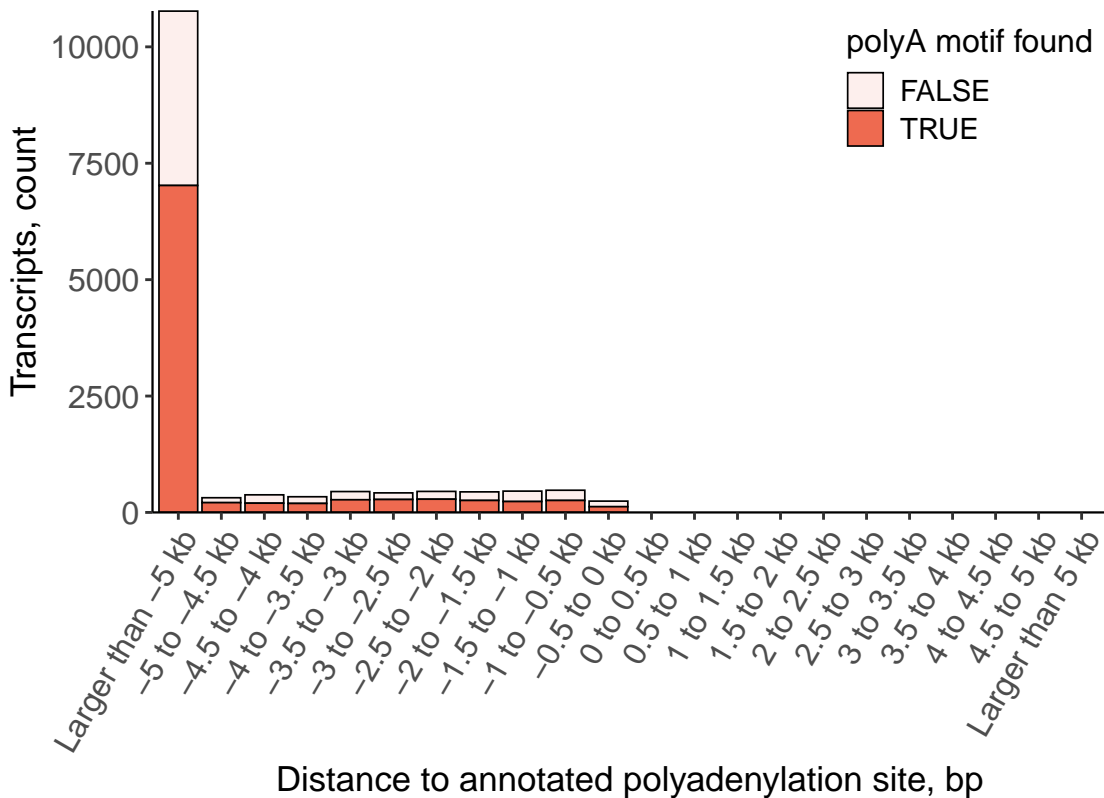
Distance to Annotated Polyadenylation Site for ISM 3' Fragment

Negative values indicate upstream of annotated polyA site



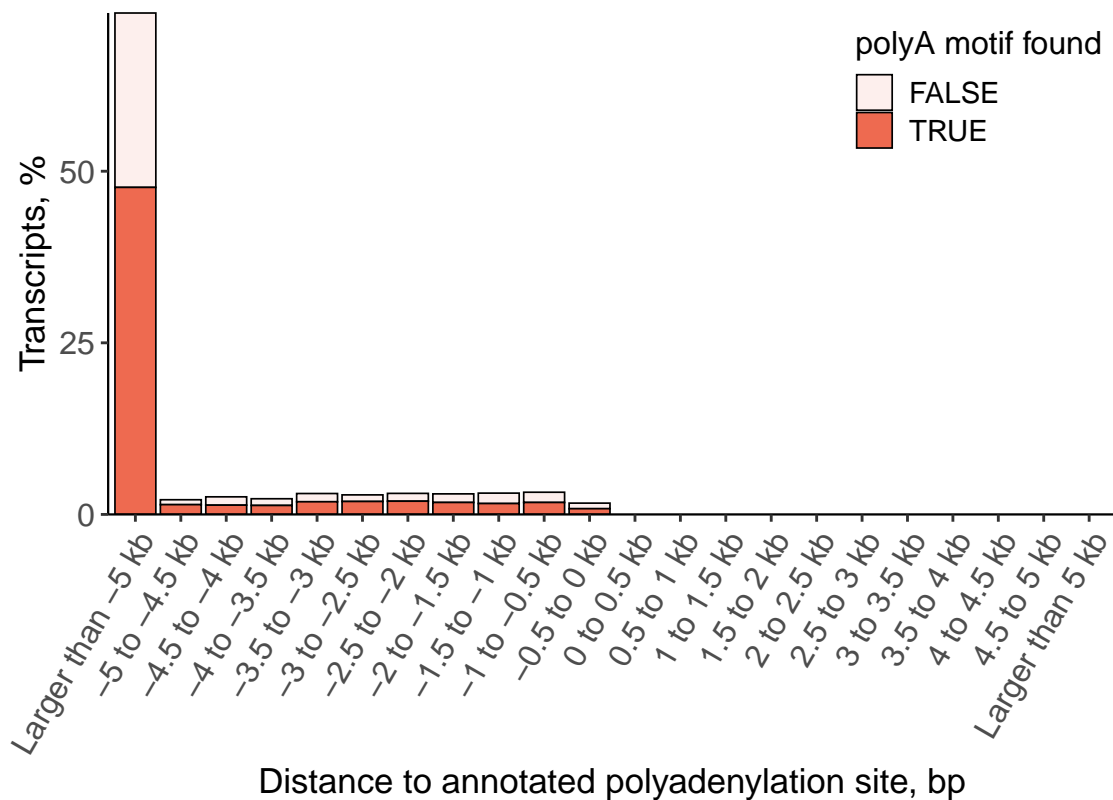
Distance to Annotated Polyadenylation Site for ISM Internal Fragment

Negative values indicate upstream of annotated polyA site



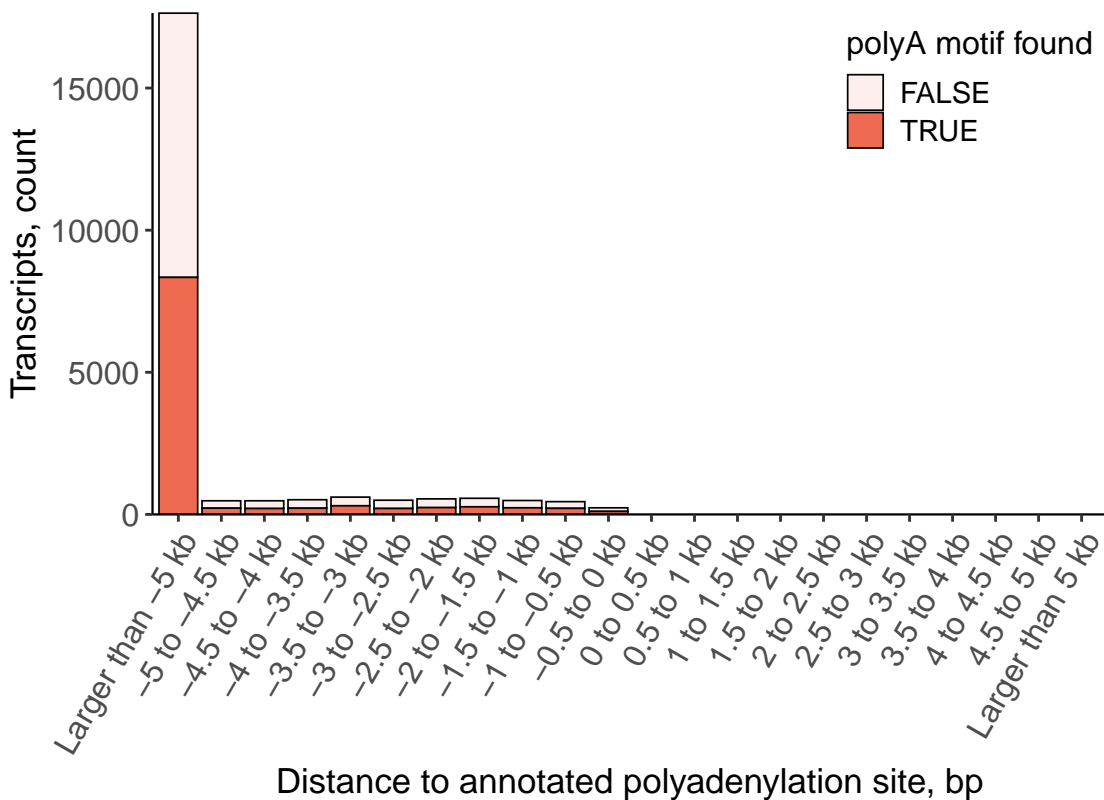
Distance to Annotated Polyadenylation Site for ISM Internal Fragment

Negative values indicate upstream of annotated polyA site



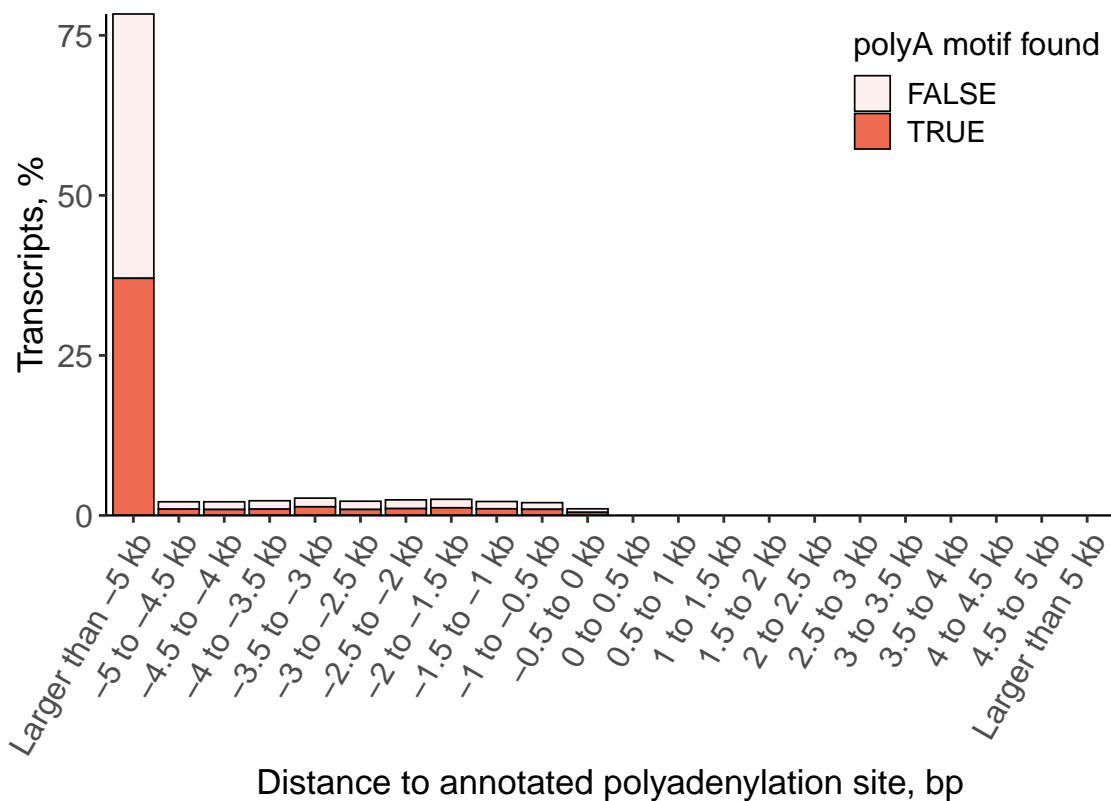
Distance to Annotated Polyadenylation Site for ISM A5' Fragment

Negative values indicate upstream of annotated polyA site



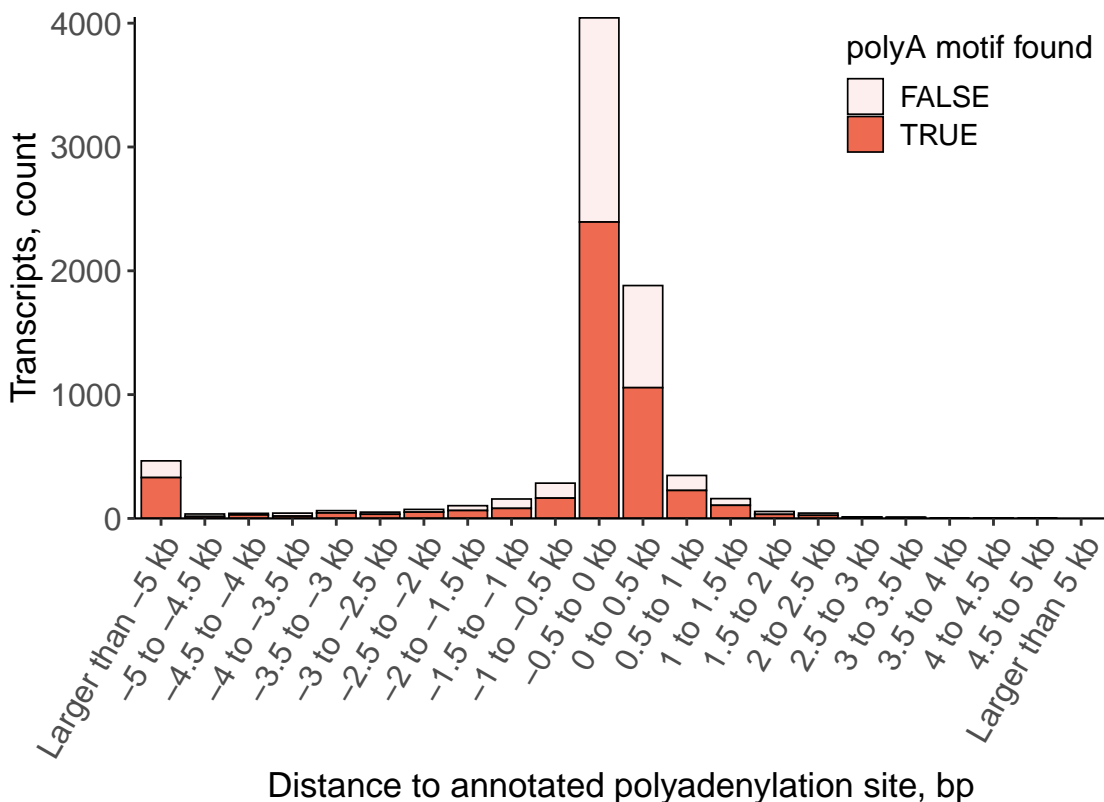
Distance to Annotated Polyadenylation Site for ISM A5' Fragment

Negative values indicate upstream of annotated polyA site



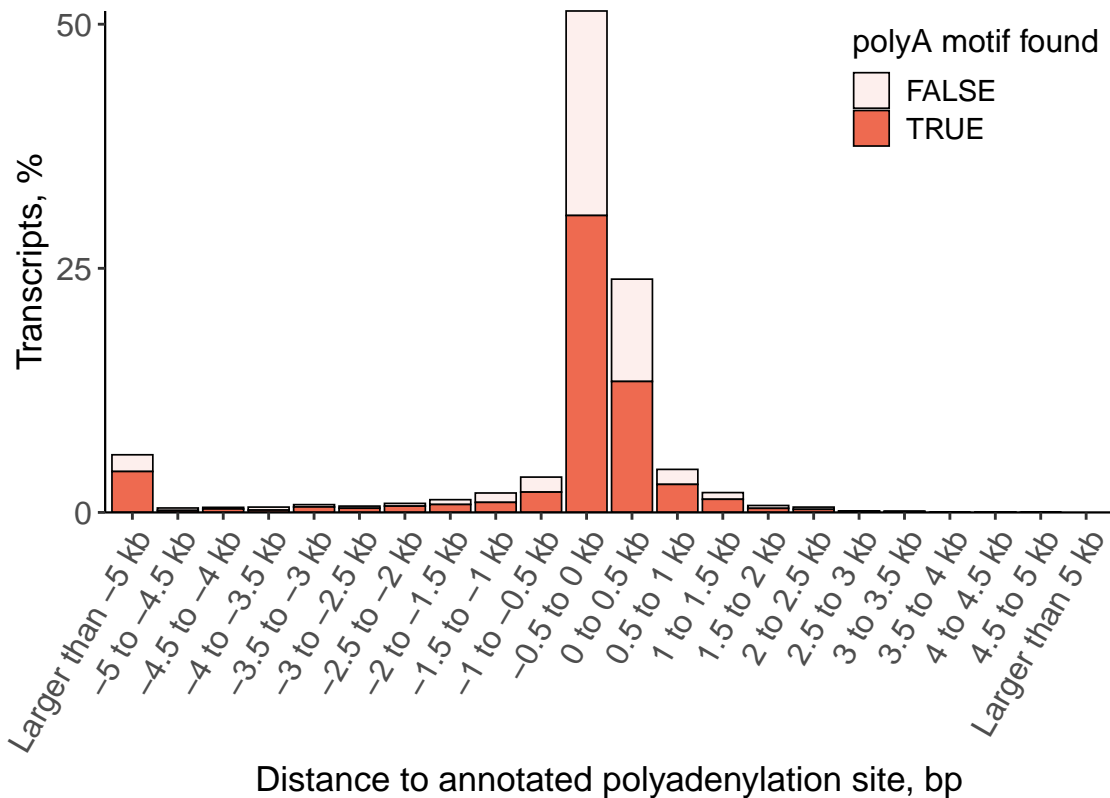
Distance to Annotated Polyadenylation Site for ISM Intron Retention

Negative values indicate upstream of annotated polyA site



Distance to Annotated Polyadenylation Site for ISM Intron Retention

Negative values indicate upstream of annotated polyA site



PolyA Distance Analysis

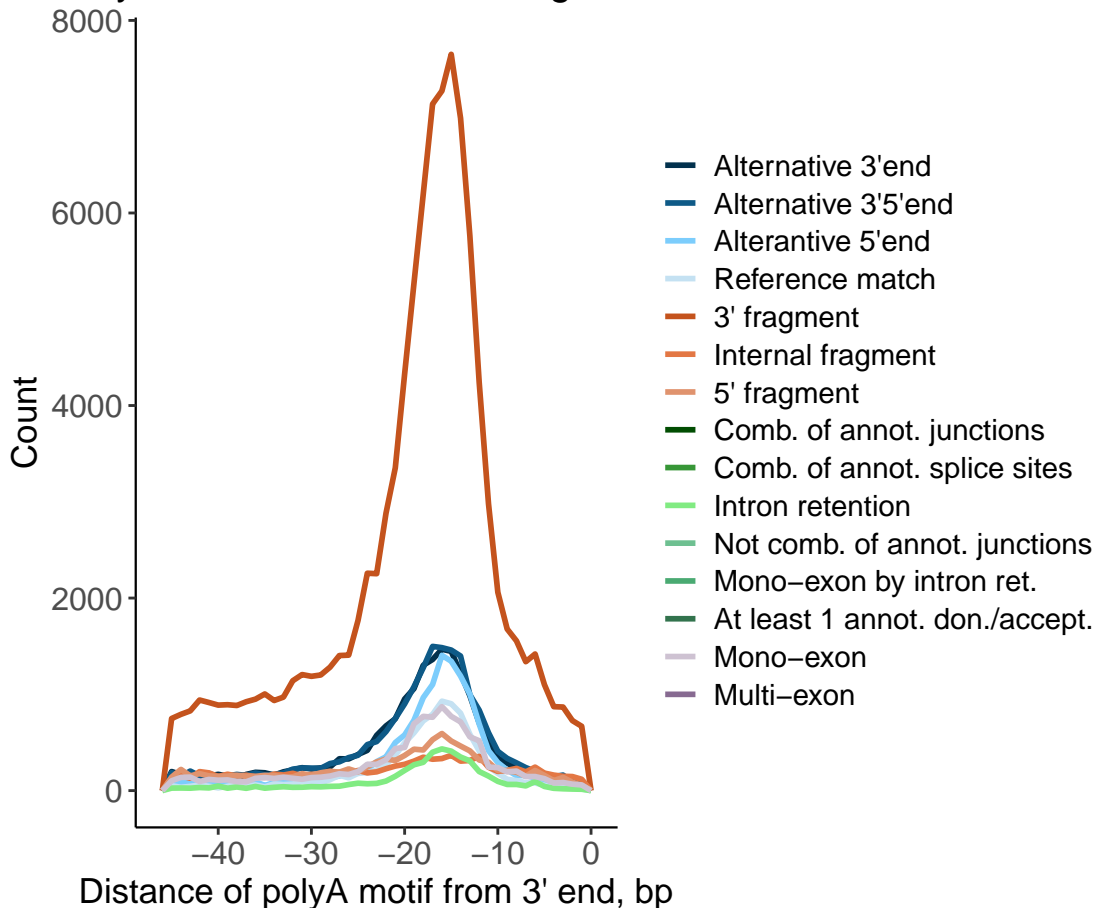
Frequency of PolyA Motifs

Motif	Count	%
AATAAA	158777	49.7
ATTA AA	46236	14.5
TATAAA	13486	4.2
AGTAAA	11868	3.7
AAAAAG	11441	3.6
AAGAAA	11412	3.6
GGGGCT	9727	3.0
TTTAAA	8629	2.7
AATATA	7262	2.3
CATAAA	6980	2.2
AATACA	6952	2.2
GATAAA	6820	2.1
AAAACA	6536	2.0
AATGAA	5955	1.9
AATAGA	3790	1.2
ACTAAA	3439	1.1

Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	145277	75776	52
ISM	213164	127689	60
NIC	118186	60589	51
NNC	94898	50628	53
Genic Genomic	708	375	53
Antisense	1682	980	58
Fusion	3498	1891	54
Intergenic	2111	1318	62
NA	87	64	74

Distance of Detected PolyA Motif From 3'End by FSM and ISM Subcategories



Number of polyA Motifs Detected

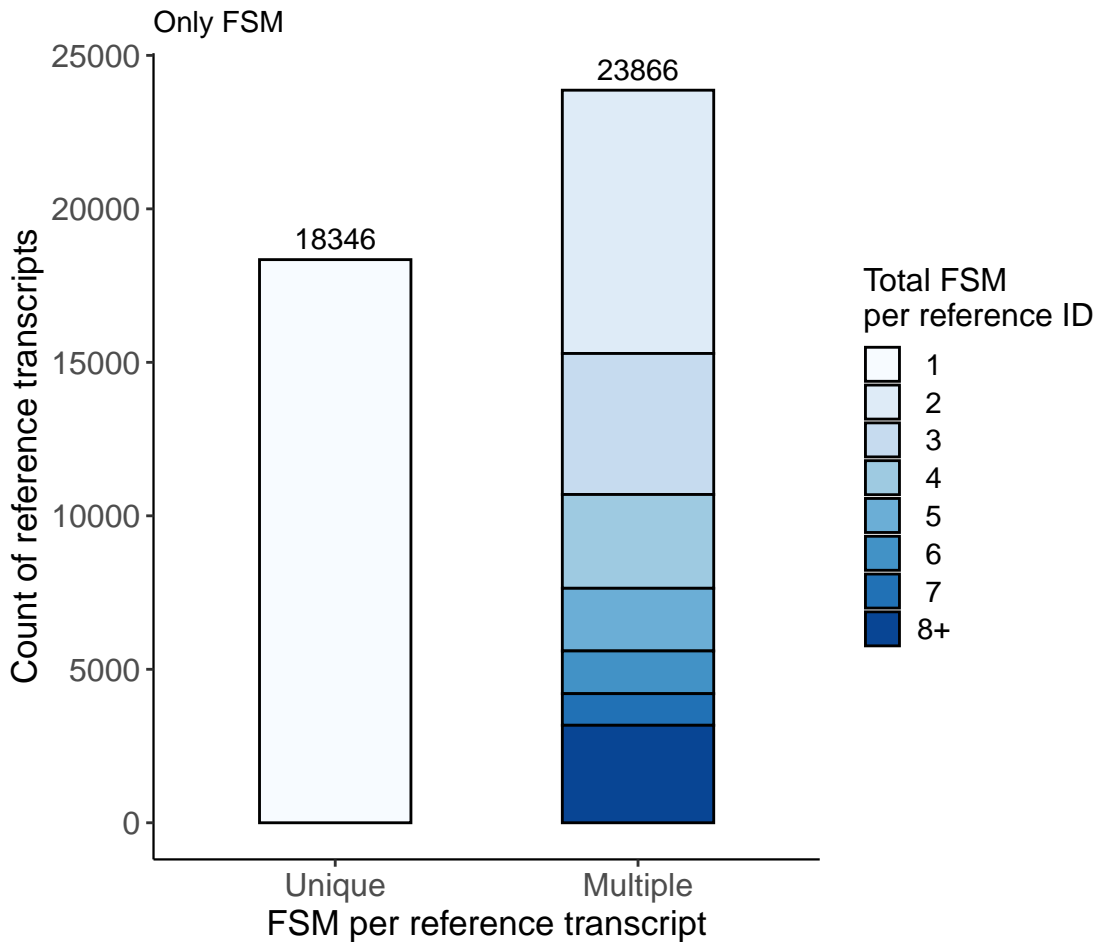
Subcategory	Count	polyA Detected	%
Alternative 3'end	41451	19623	47
Alternative 3'5'end	38841	20385	52
Alterantive 5'end	26481	14220	54
Reference match	18948	9822	52
3' fragment	168041	103108	61
Internal fragment	14739	9329	63
5' fragment	22511	10561	47
Comb. of annot. junctions	48176	24274	50
Comb. of annot. splice sites	33617	17776	53
Intron retention	55989	29338	52
At least 1 annot. don./accept.	83831	44863	54
Mono-exon	19556	11726	60
Multi-exon	7355	4231	58
NA	75	54	72

Frequency of PolyA Motifs

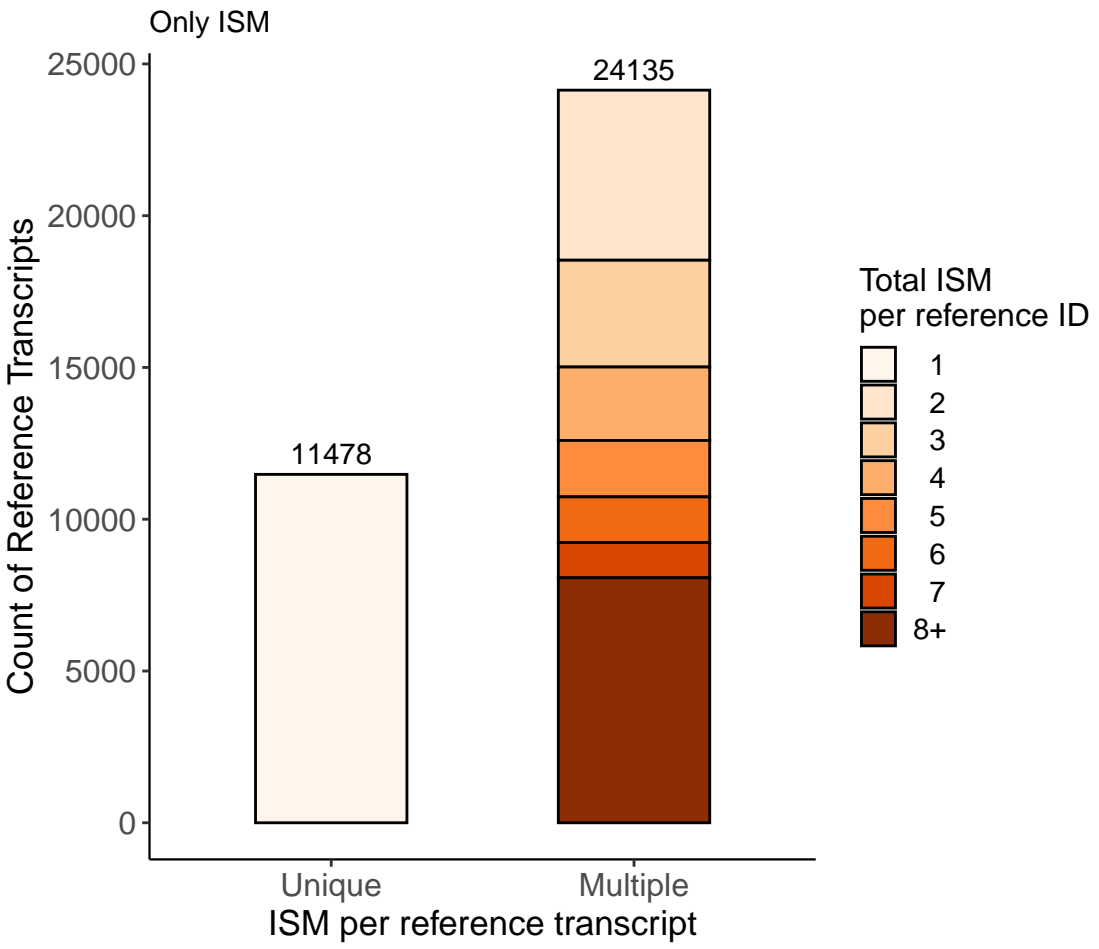
Motif	Count	%
AATAAA	158777	49.7
ATTAAA	46236	14.5
TATAAA	13486	4.2
AGTAAA	11868	3.7
AAAAAG	11441	3.6
AAGAAA	11412	3.6
GGGGCT	9727	3.0
TTTAAA	8629	2.7
AATATA	7262	2.3
CATAAA	6980	2.2
AATACA	6952	2.2
GATAAA	6820	2.1
AAAACA	6536	2.0
AATGAA	5955	1.9
AATAGA	3790	1.2
ACTAAA	3439	1.1

Redundancy Analysis

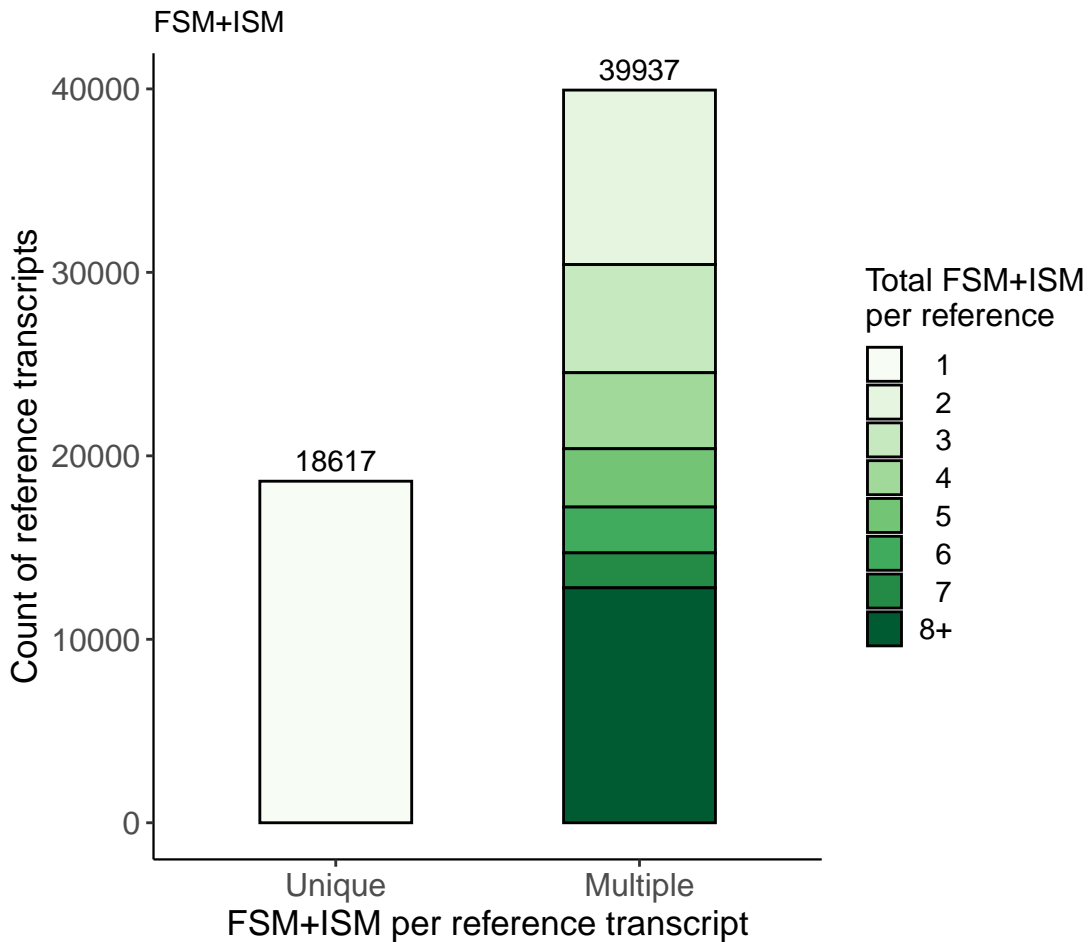
Reference Transcript Redundancy



Reference Transcript Redundancy

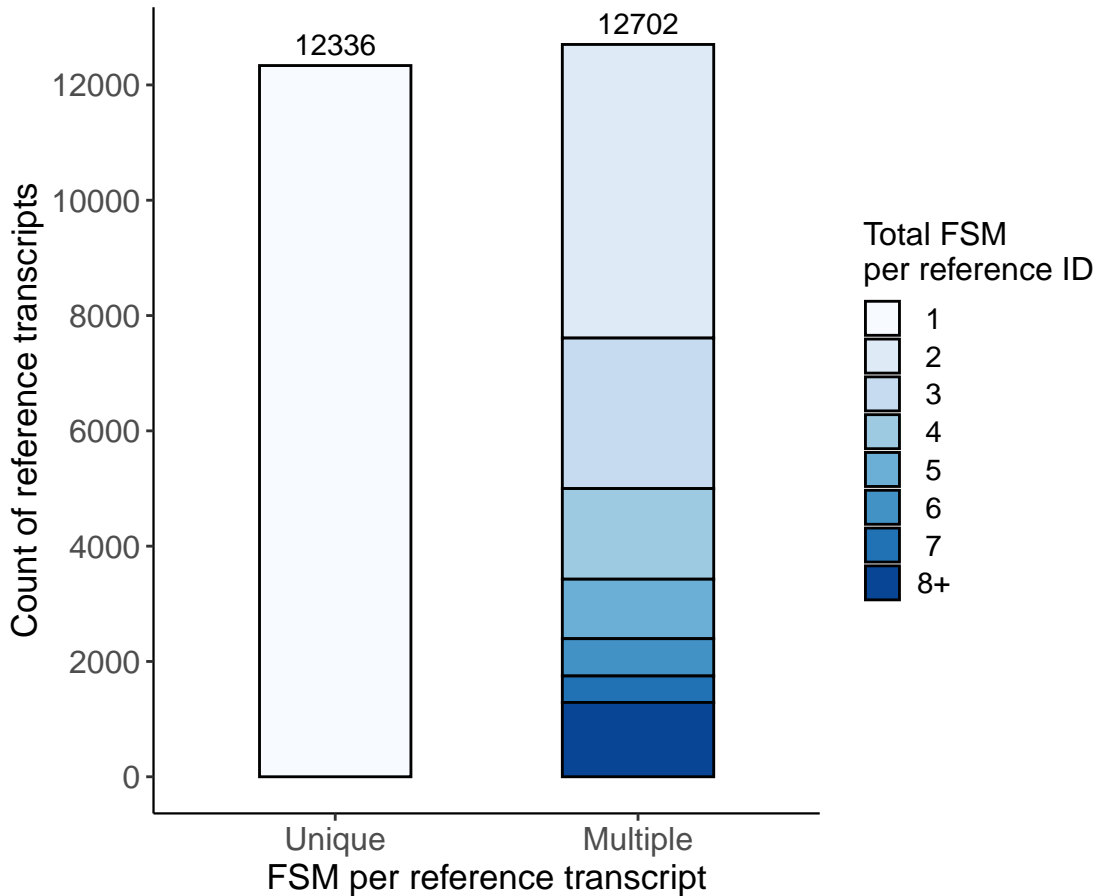


Reference Transcript Redundancy



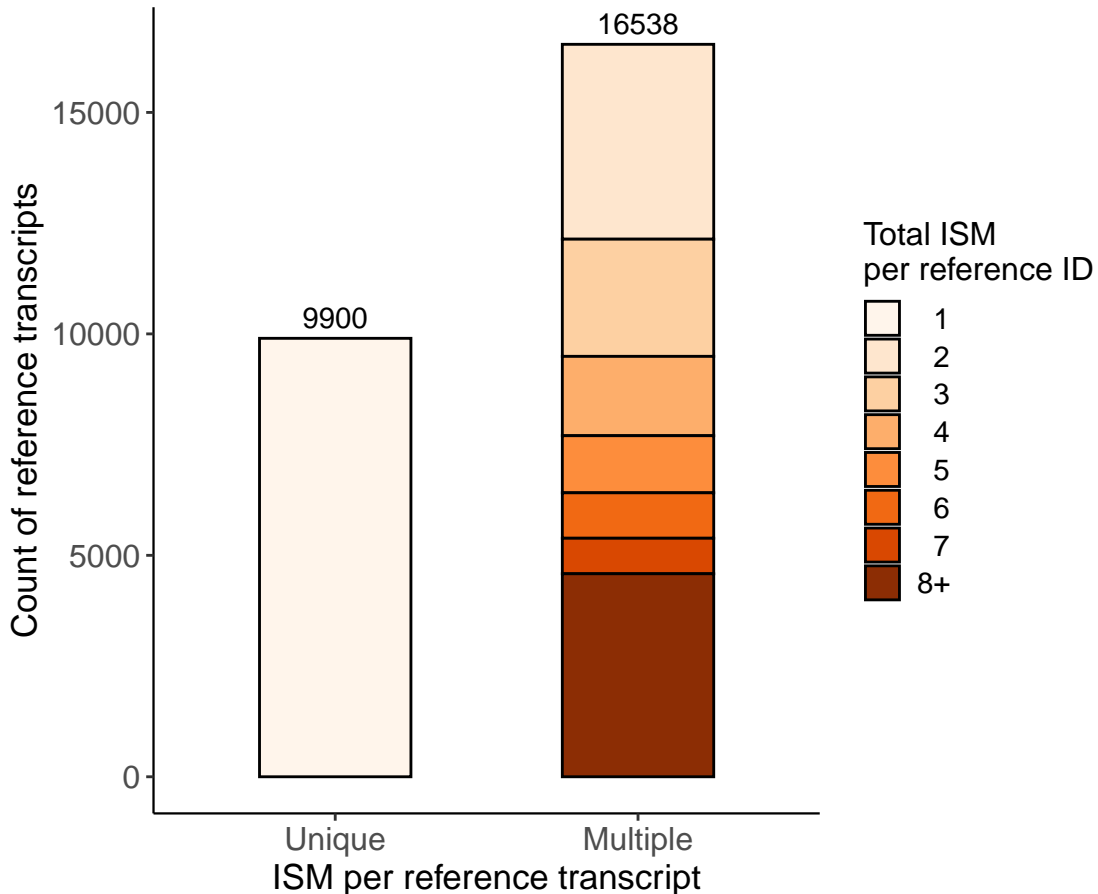
Reference Transcript Redundancy

Only FSM with a polyA motif found



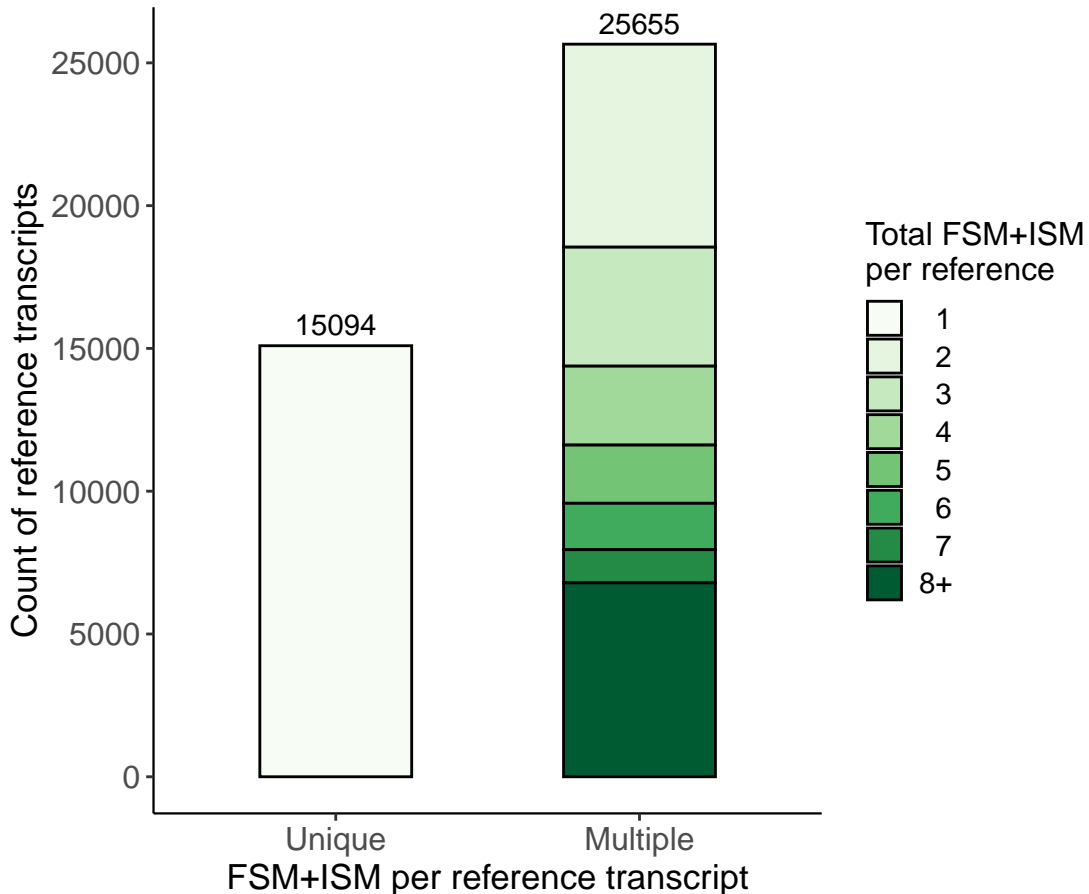
Reference Transcript Redundancy

Only ISM with a polyA motif found



Reference Transcript Redundancy

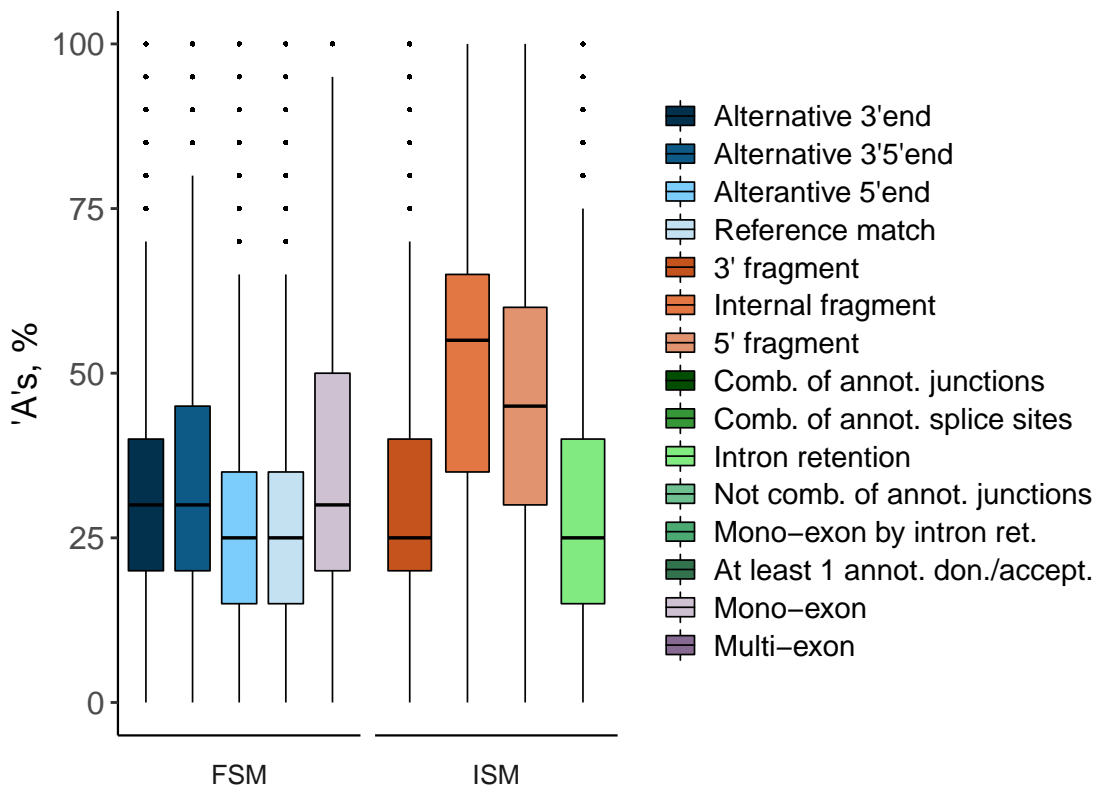
FSM+ISM with a polyA motif found



Intra-Priming Quality Check

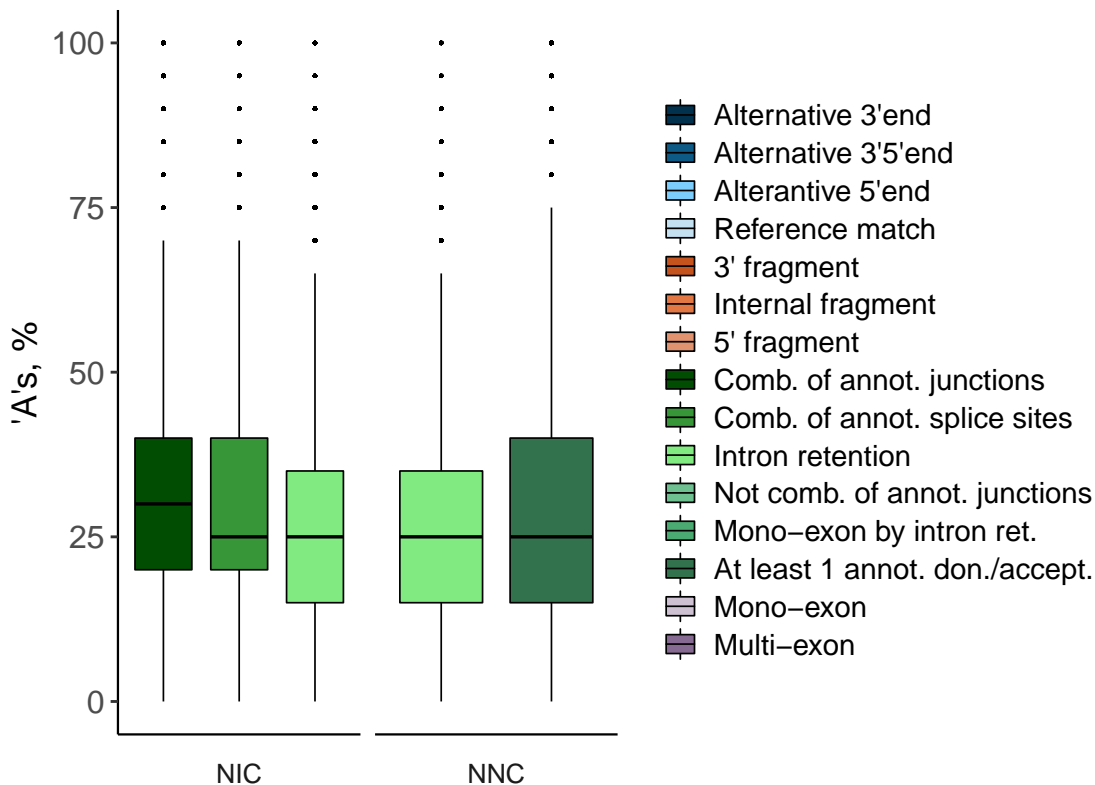
Possible Intra-Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



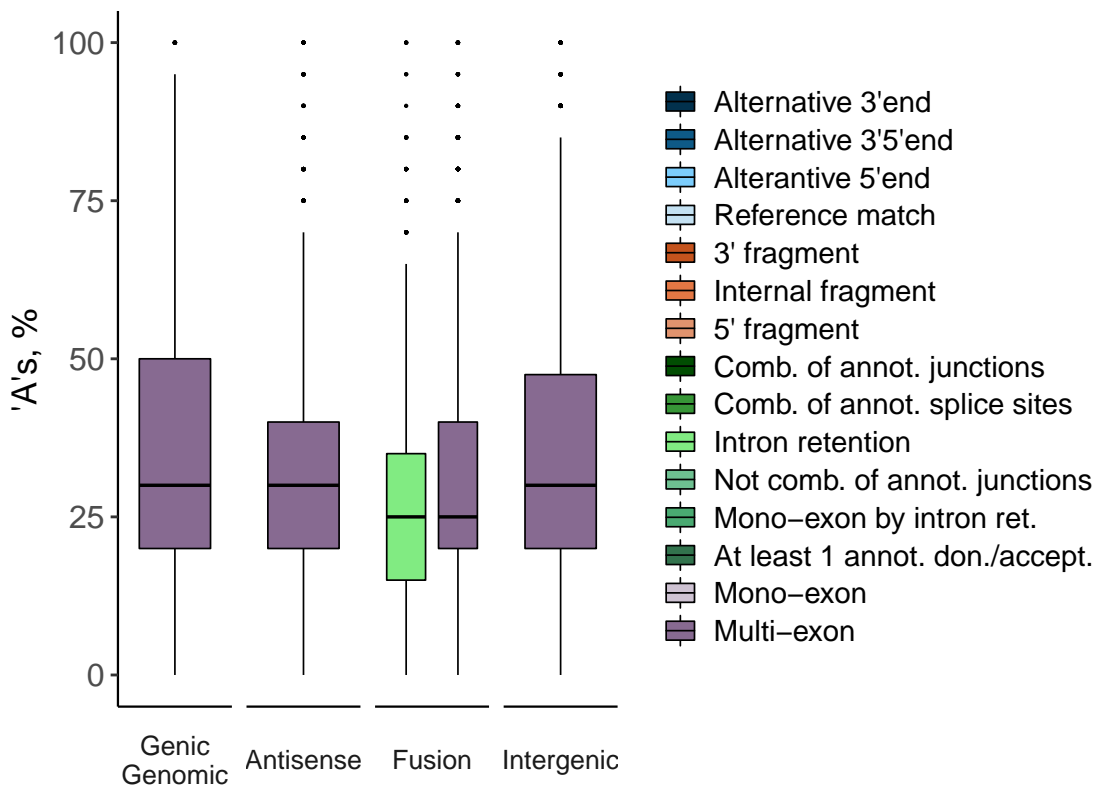
Possible Intra-Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



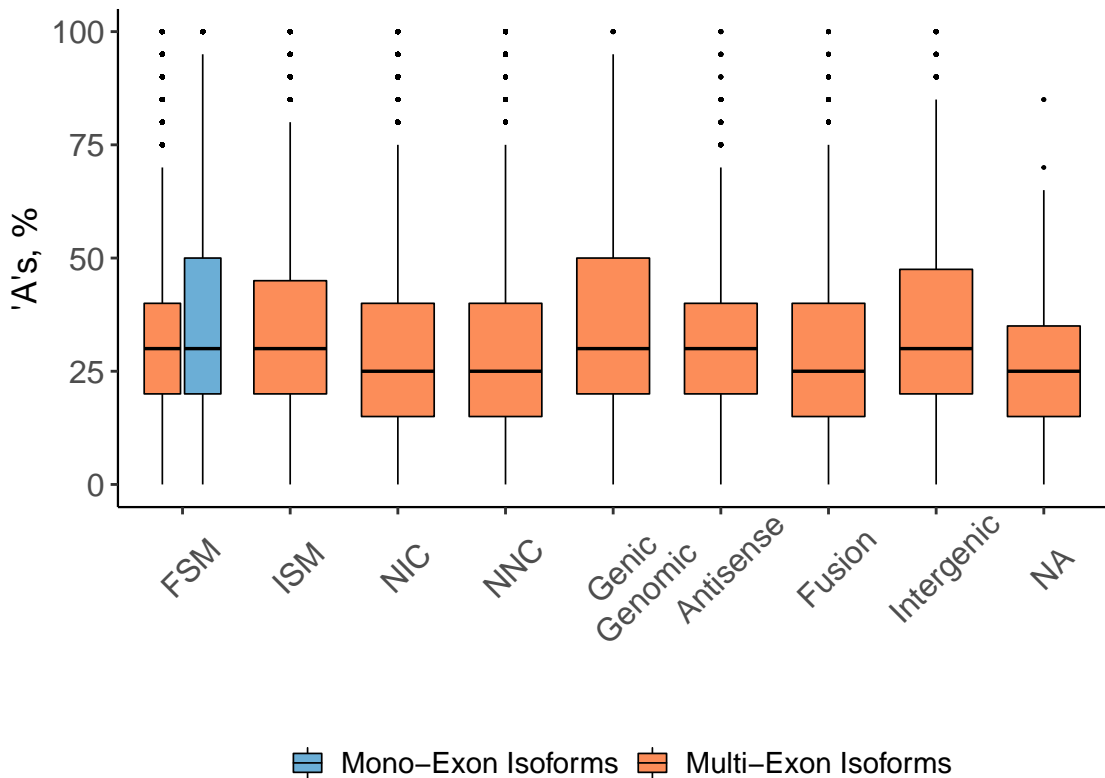
Possible Intra-Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



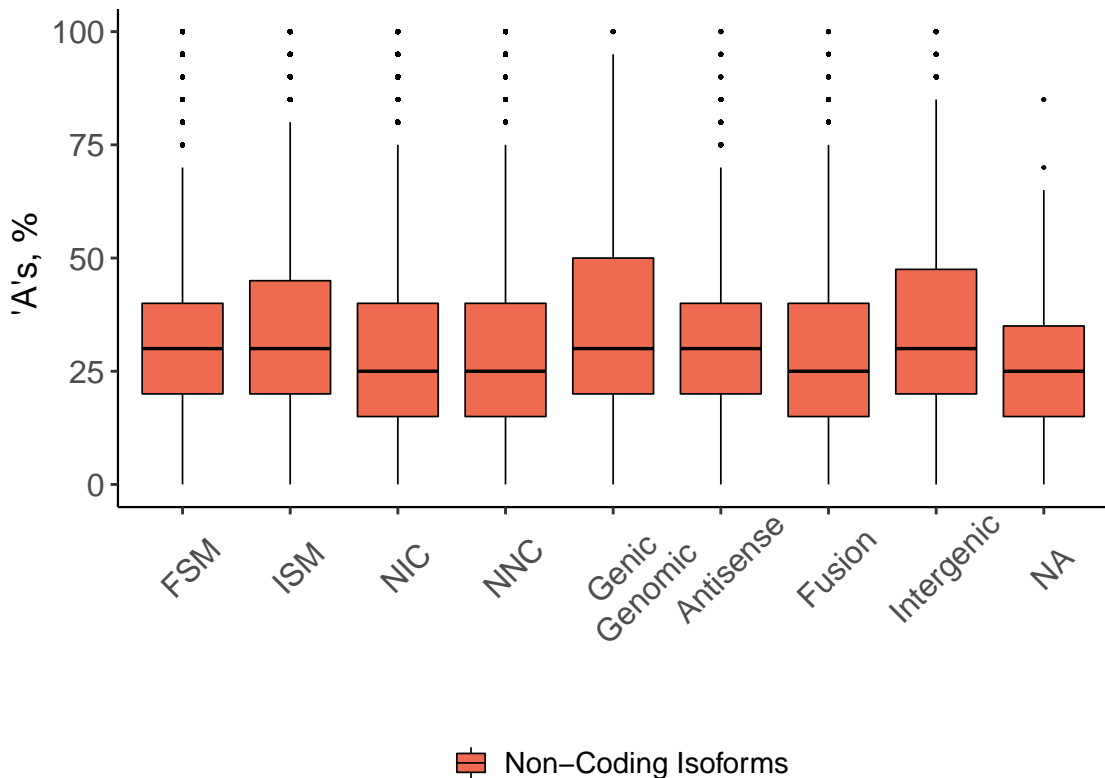
Mono- vs Multi-Exon Possible Intra-Priming

Percent of genomic 'A's in downstream 20 bp



Coding vs Non-Coding Possible Intra-Priming

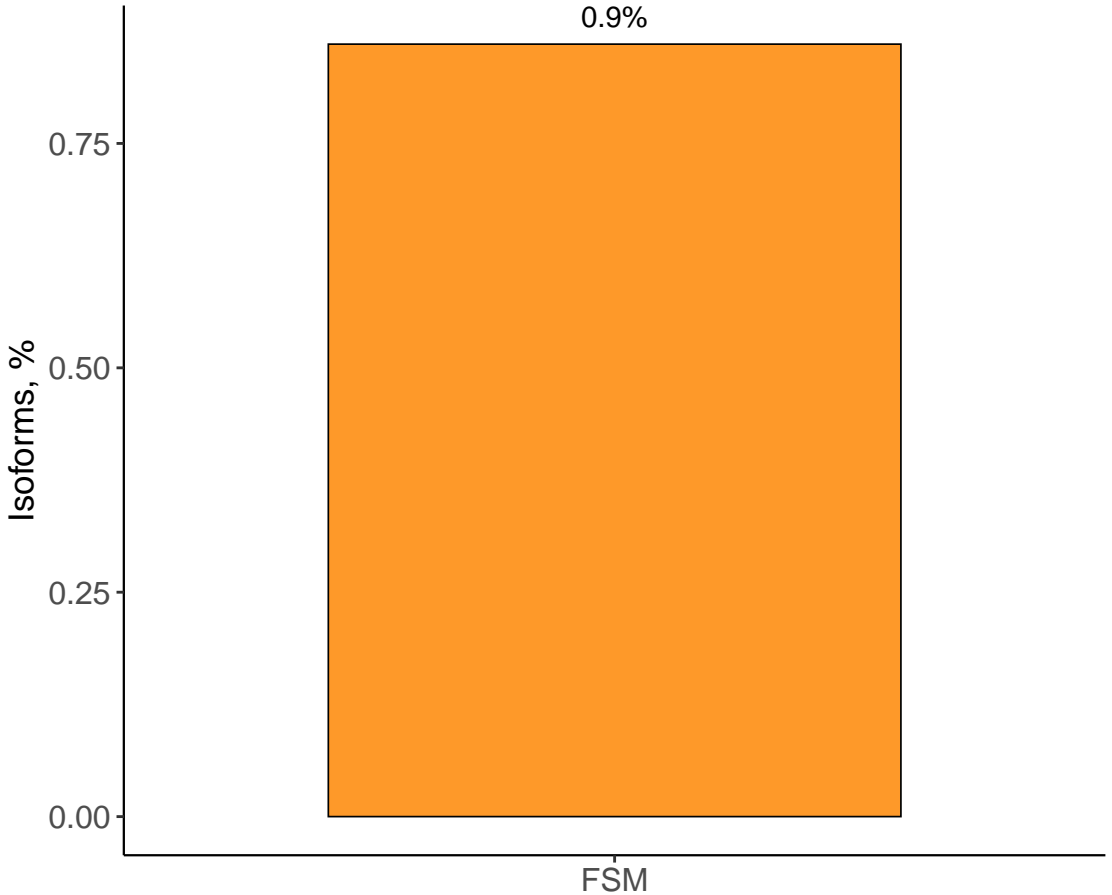
Percent of genomic 'A's in downstream 20 bp



Features of Bad Quality

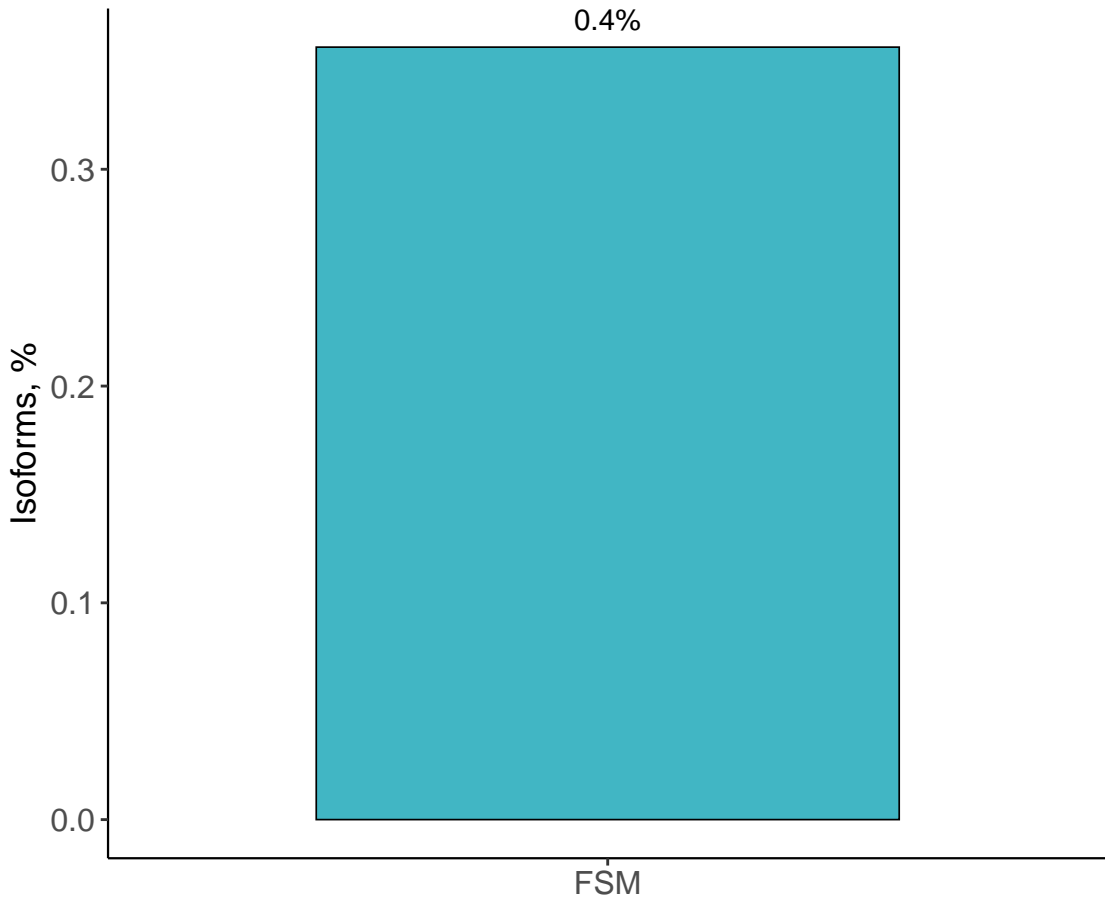
RT-switching

0.9%

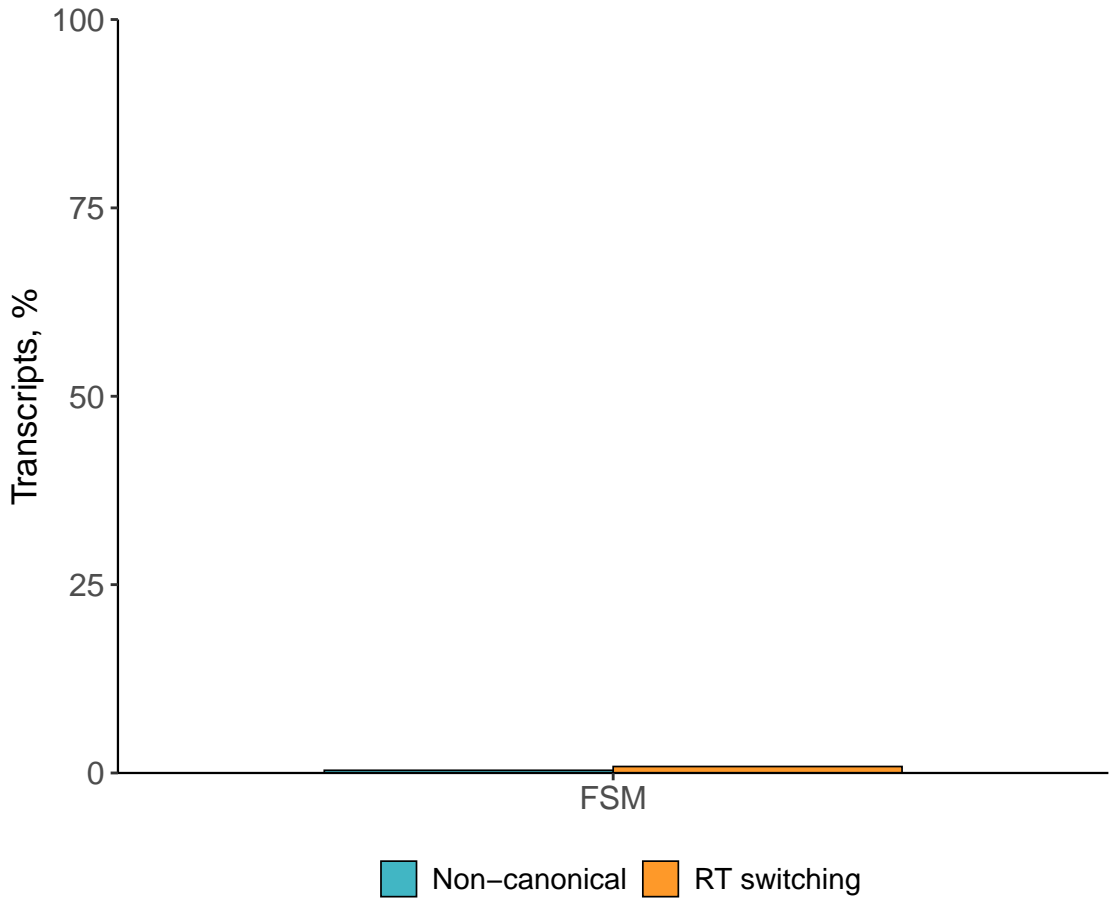


Non-Canonical Junctions

0.4%

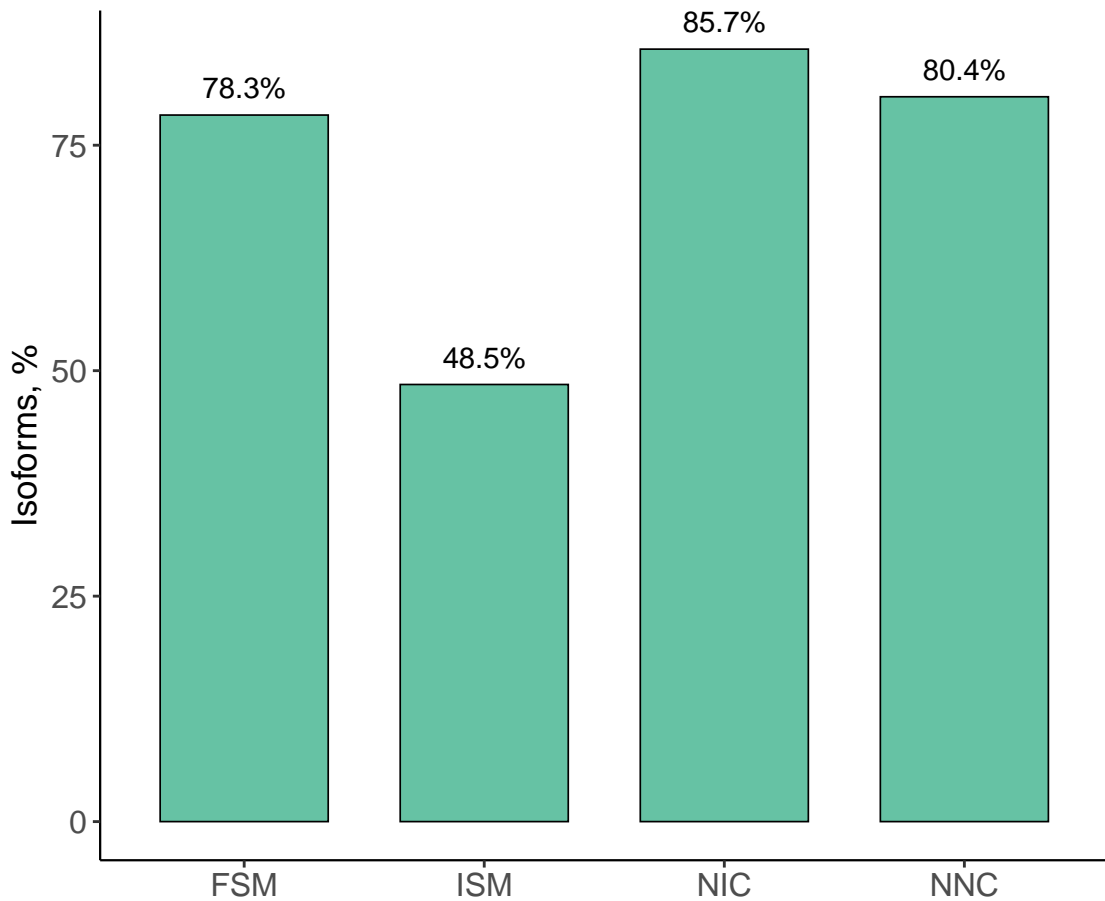


Quality Control Attributes Across Structural Categories

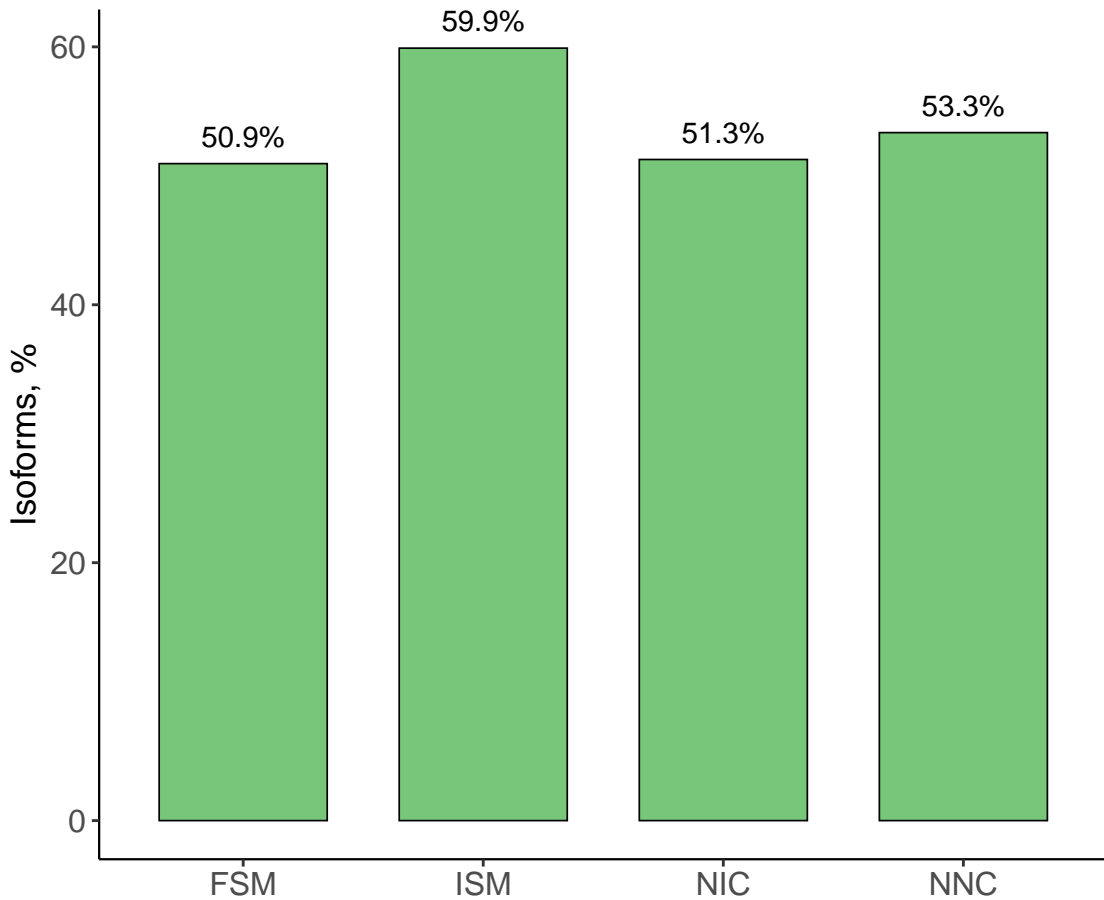


Features of Good Quality

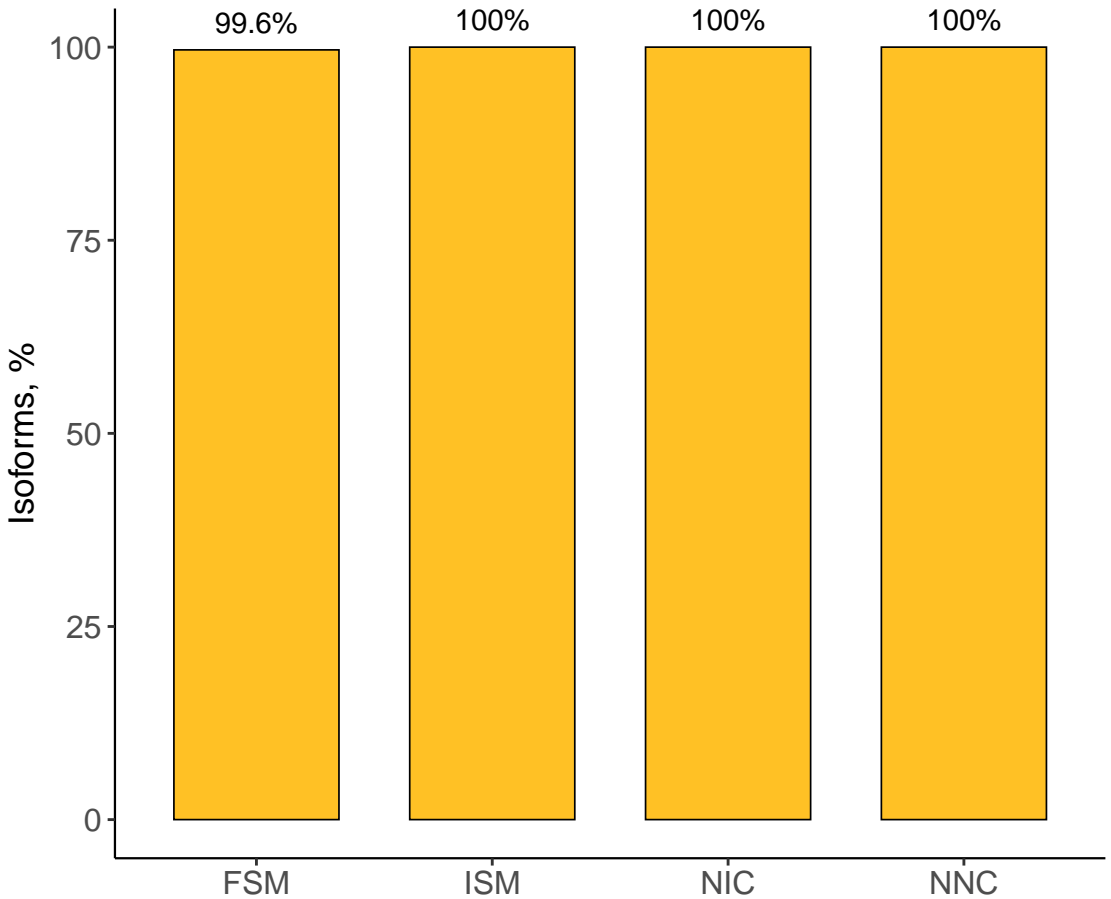
Annotation Support



PolyA Support



All Canonical Junctions



Good Quality Control Attributes Across Structural Categories

