SQANTI3 report

Unique Genes: 24242

Unique Isoforms: 534383

Transcript Classification

Category	Isoforms, count
FSM	137383
ISM	213048
NIC	99744
NNC	76762
Genic Genomic	653
Antisense	1493
Fusion	3337
Intergenic	1874
Genic Intron	0

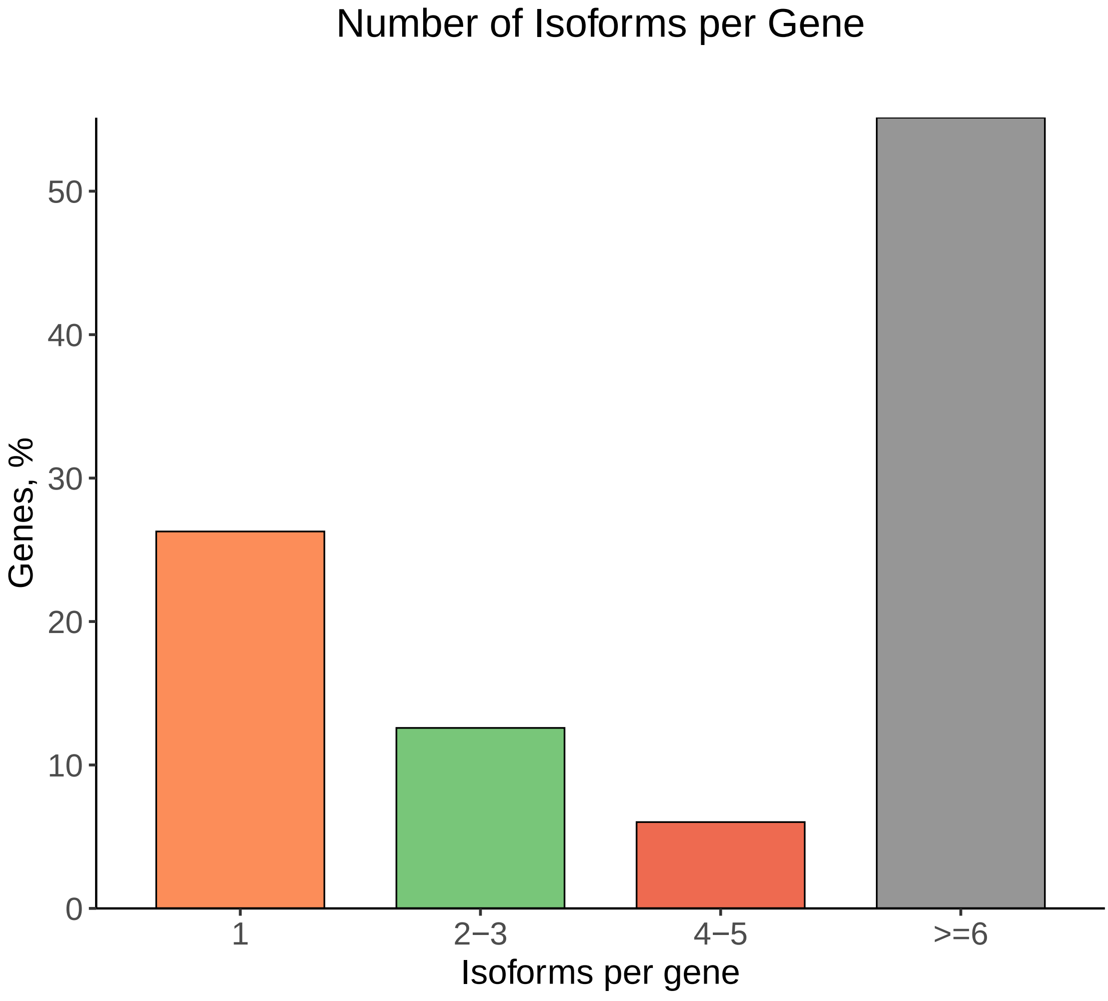
Gene Classification

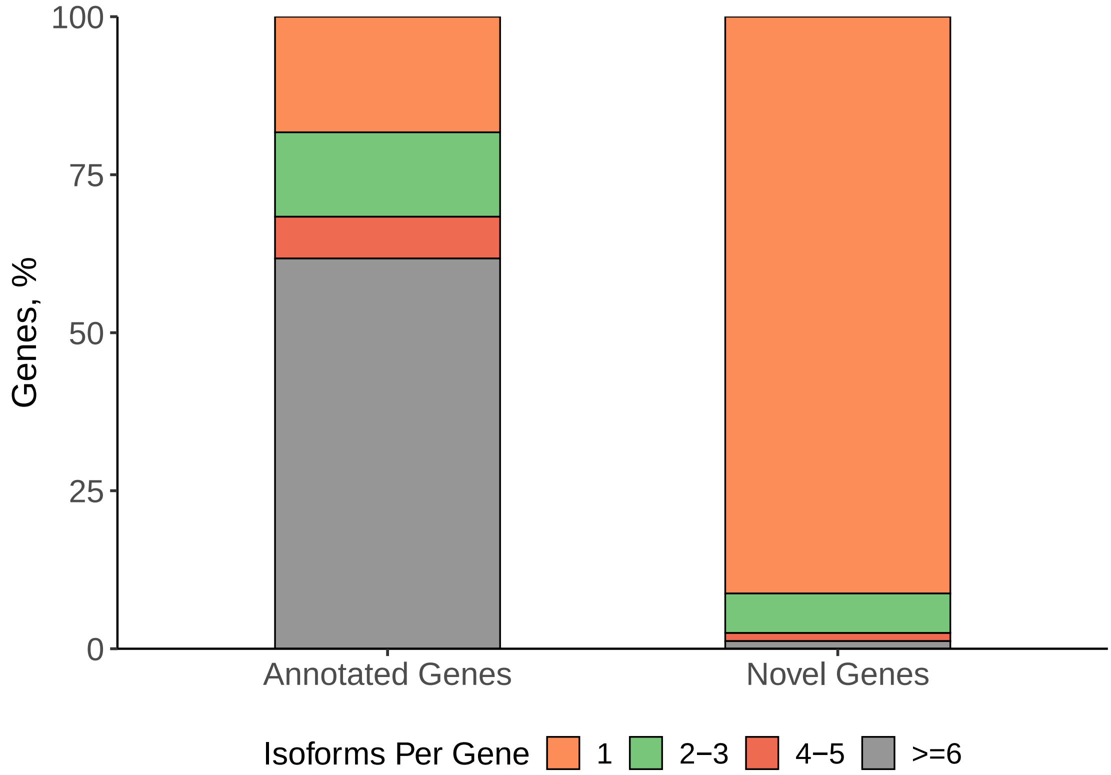
Category	Genes, count
Annotated Genes	21589
Novel Genes	2653

Splice Junction Classification

Category	SJs, count	Percent
Known canonical	194343	69.40
Known Non−canonical	83	0.03
Novel canonical	85619	30.57
Novel Non−canonical	0	0.00

Gene Characterization

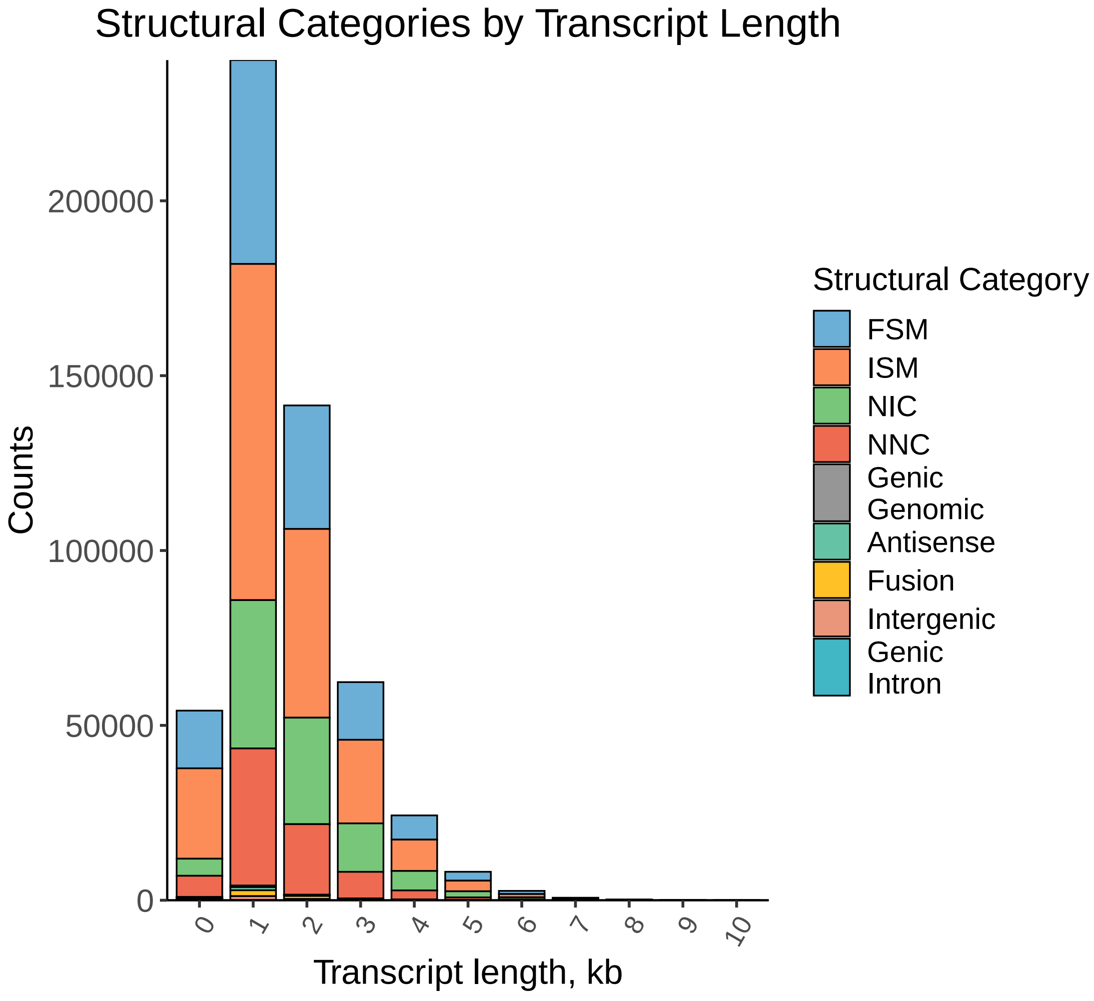

Number of Isoforms per Gene
Genes, %
50
40
30
20
10
0
1
2−3
4−5
>=6
Isoforms per gene



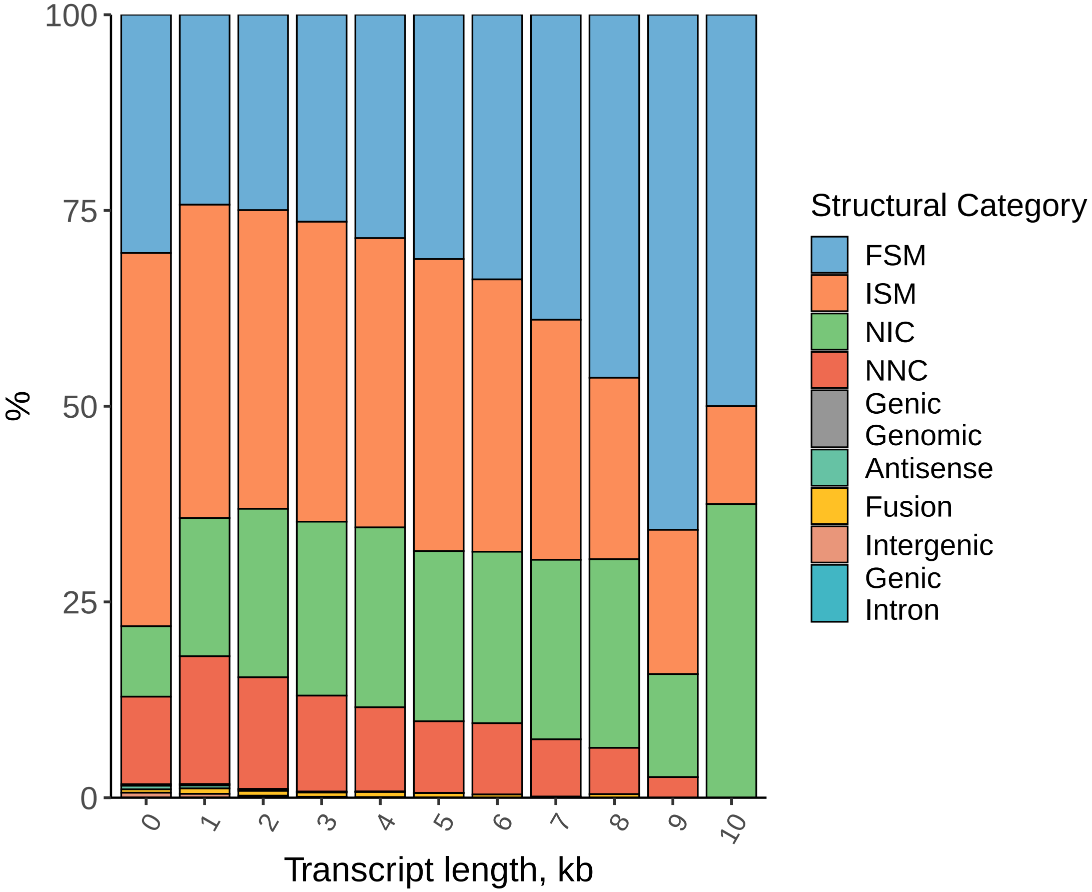
Number of Isoforms per Gene

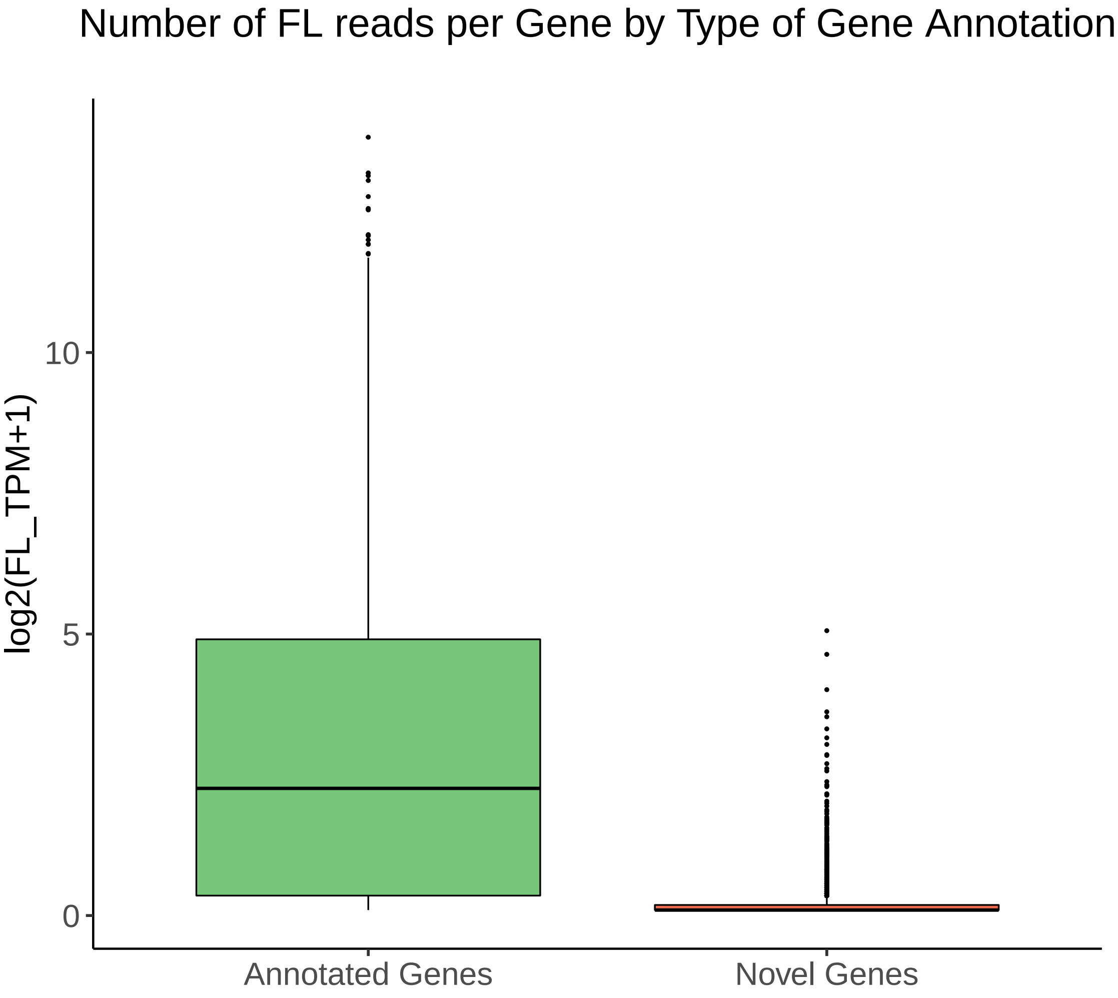
Known vs Novel Genes

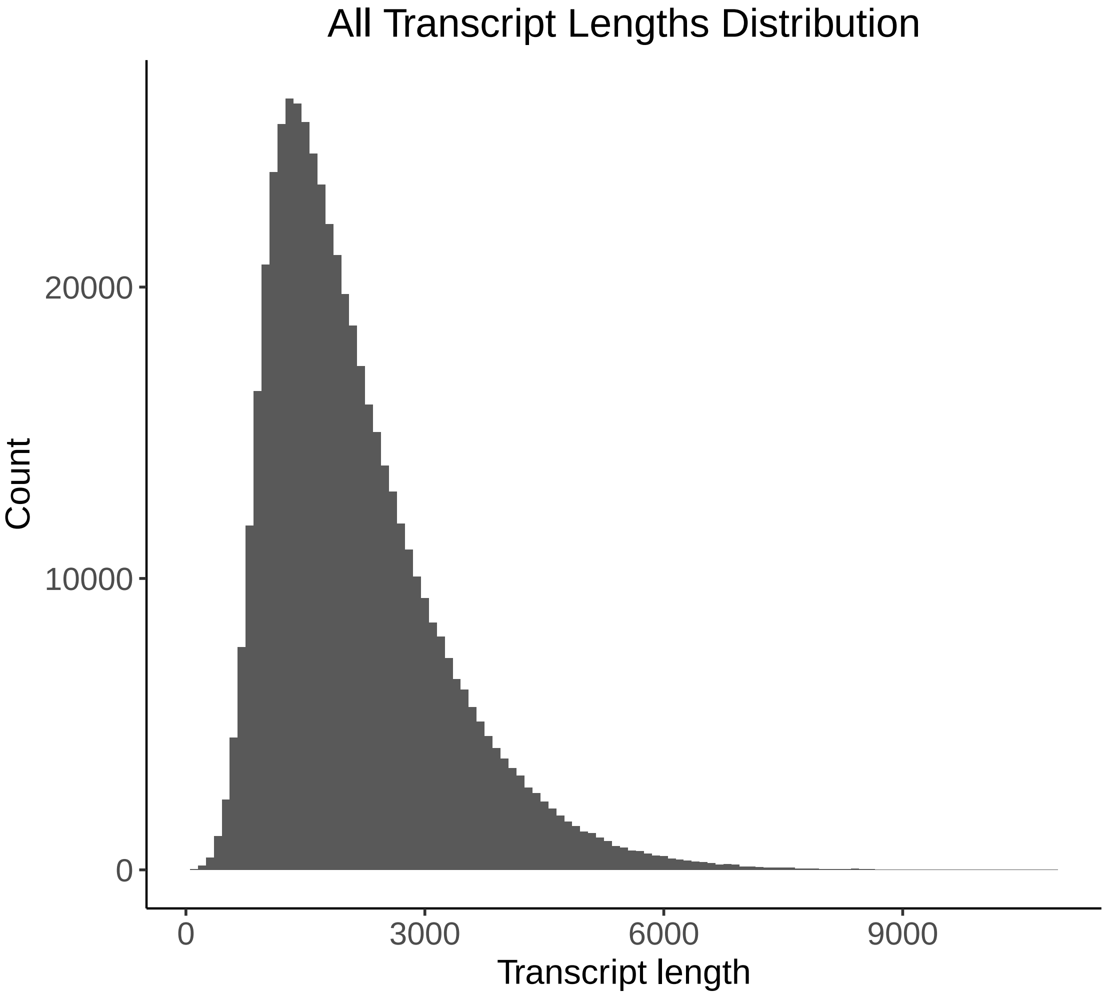

Distribution of Mono− vs Multi−Exon Transcripts
Transcripts, %
100
75
50
25
0
Novel Genes
Annotated Genes
Transcript type
Multi−Exon
Mono−Exon

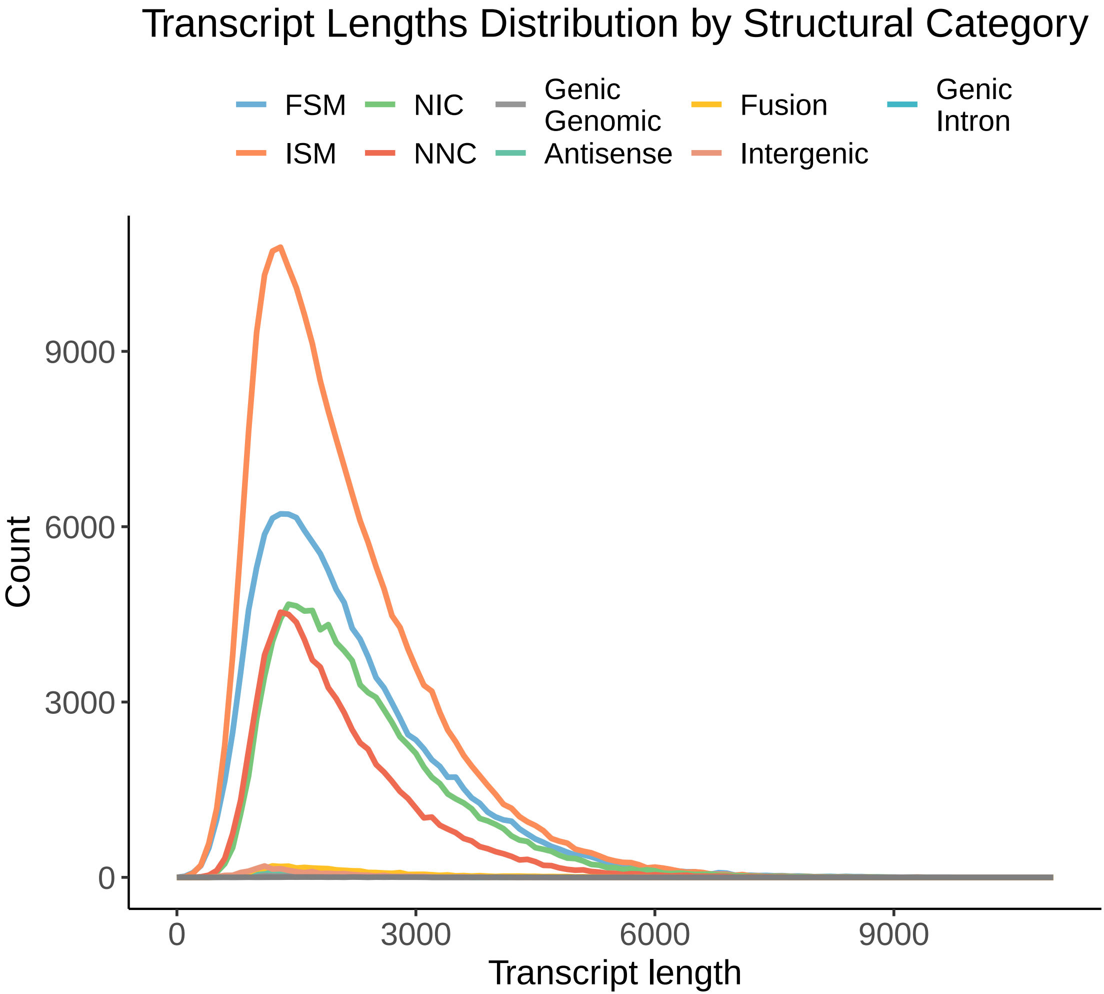

Structural Categories by Transcript Length
Counts
0
50000
100000
150000
200000
Transcript length, kb
0
1
2
3
4
5
6
7
8
9
10
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

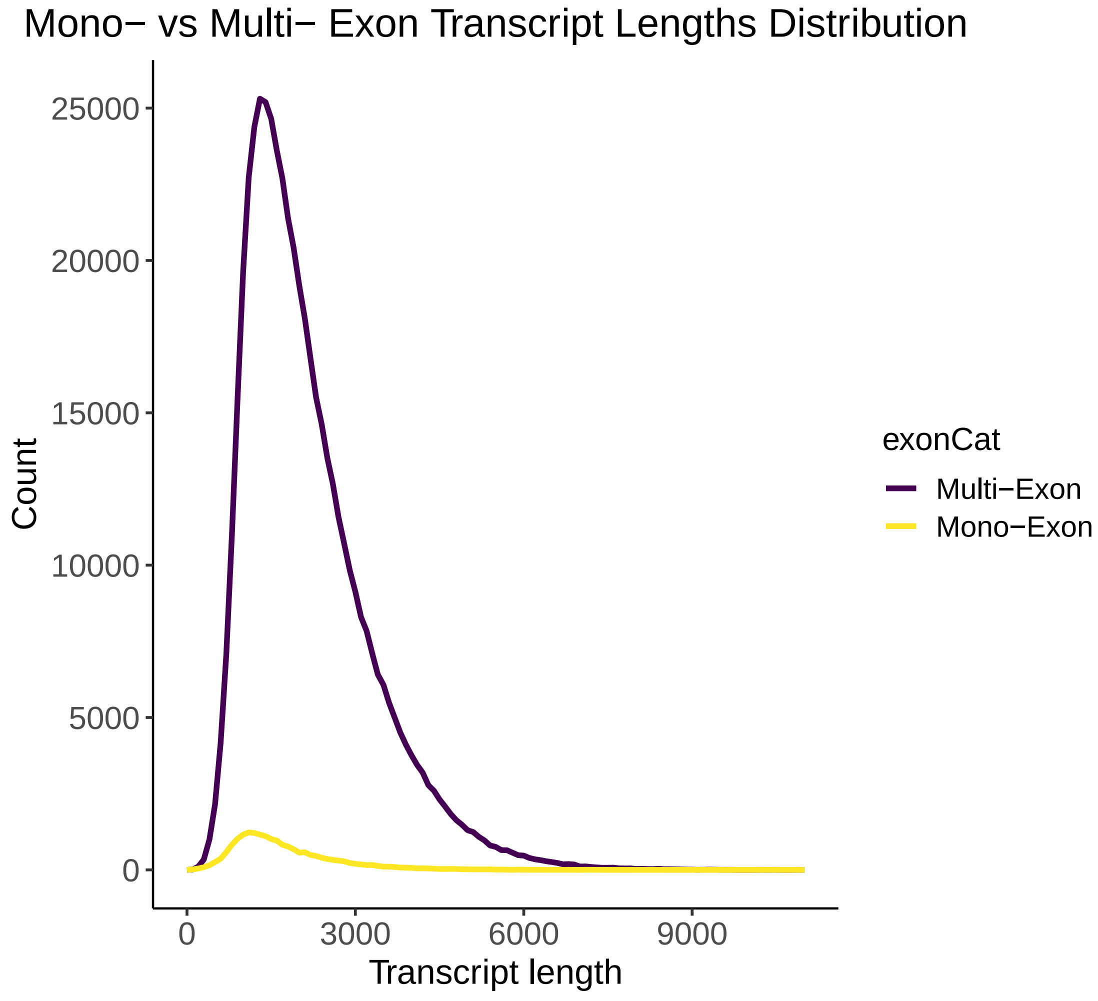
Structural Categories by Transcript Length



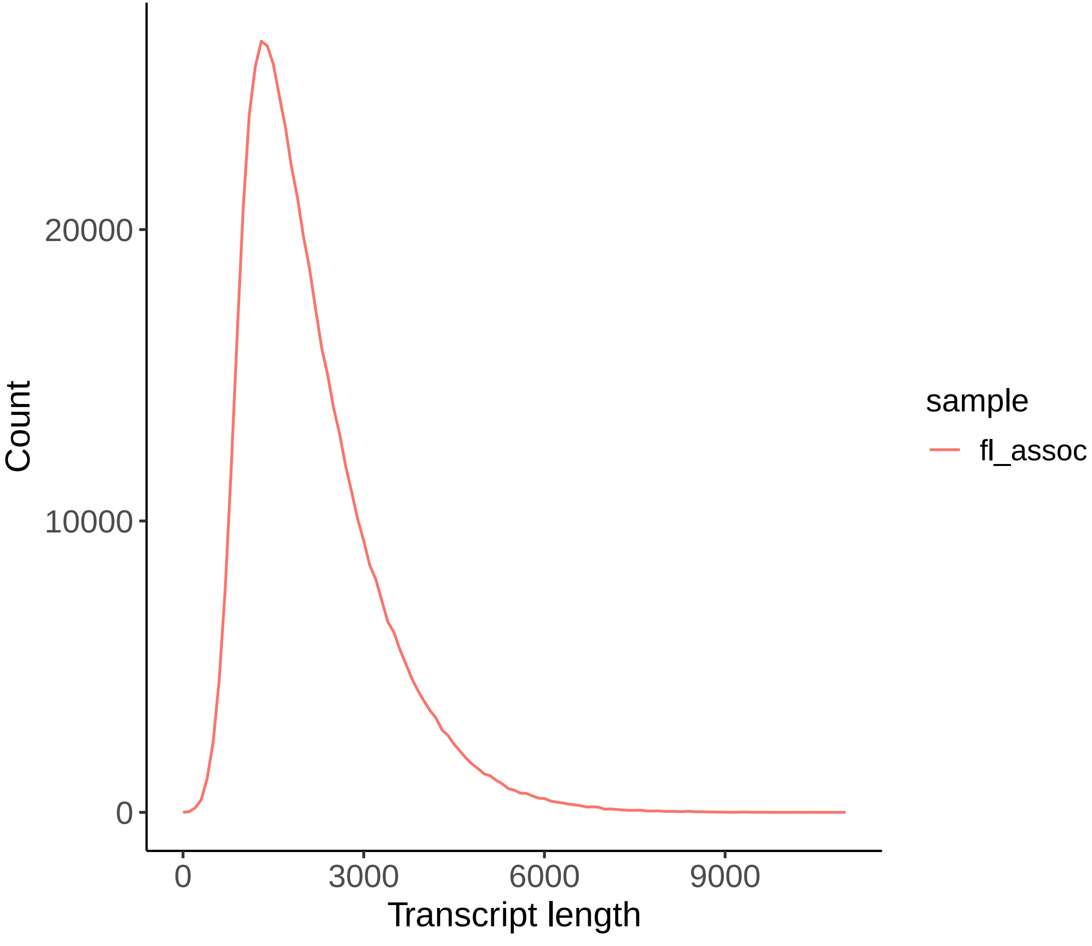

Number of FL reads per Gene by Type of Gene Annotation
log2(FL_TPM+1)
10
5
0
Annotated Genes
Novel Genes

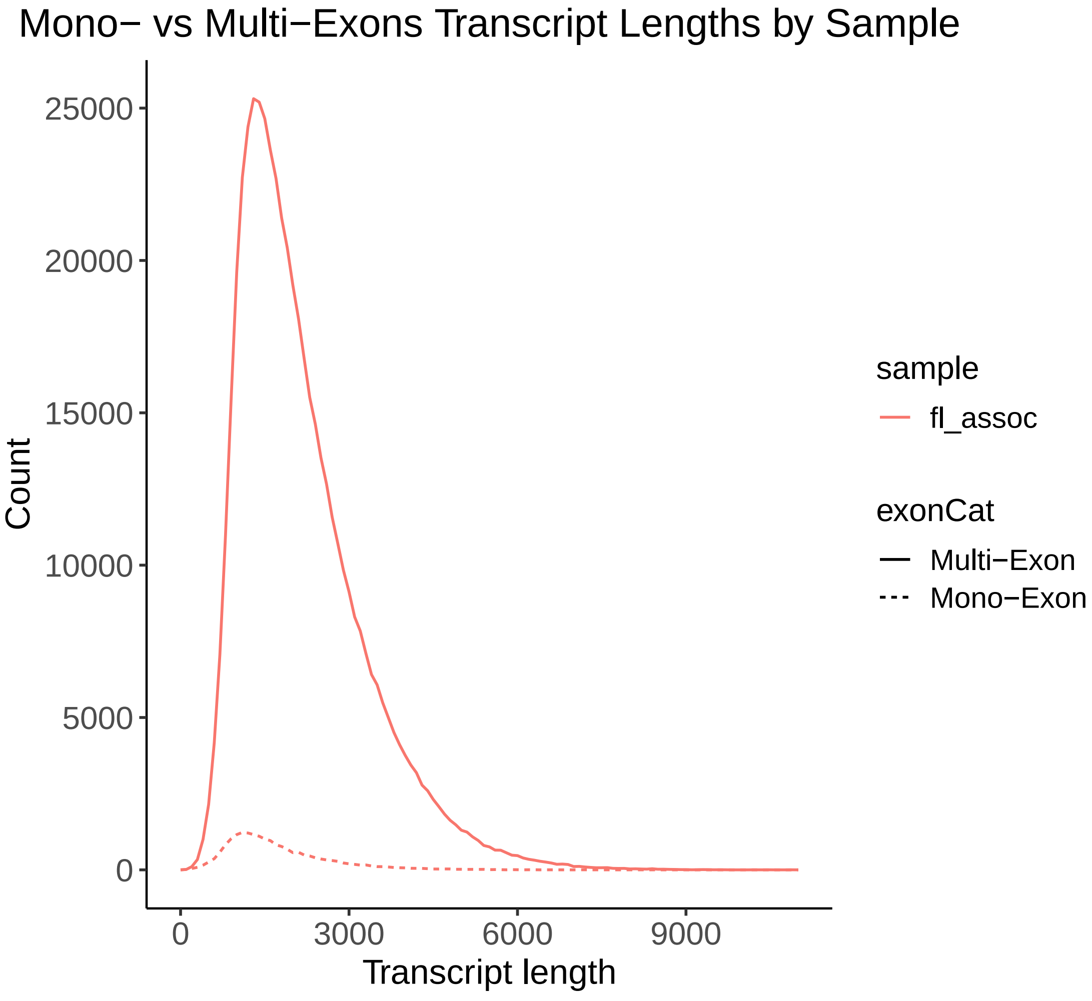

All Transcript Lengths Distribution
Count
20000
10000
0
0
3000
6000
9000
Transcript length


Transcript Lengths Distribution by Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Count
0
3000
6000
9000
0
3000
6000
9000
Transcript length


Mono− vs Multi− Exon Transcript Lengths Distribution
Count
25000
20000
15000
10000
5000
0
0
3000
6000
9000
Transcript length
exonCat
Multi−Exon
Mono−Exon

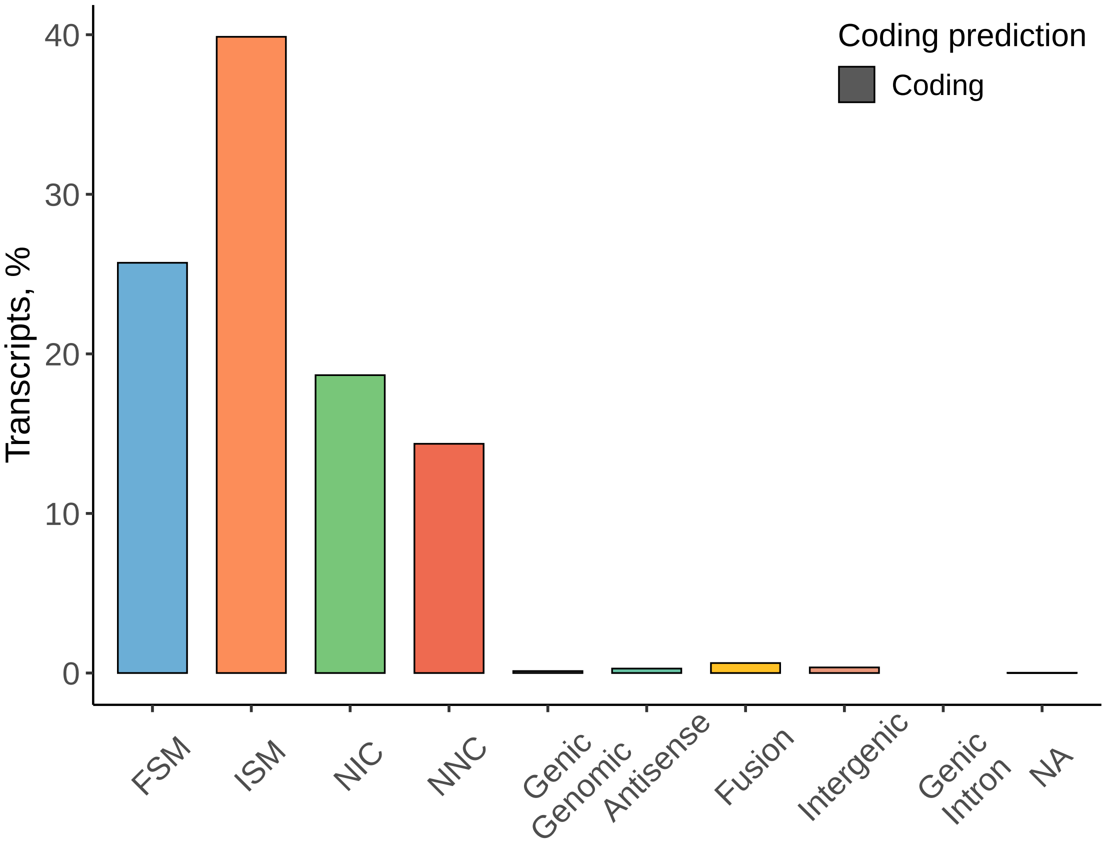
Transcript Lengths by Sample

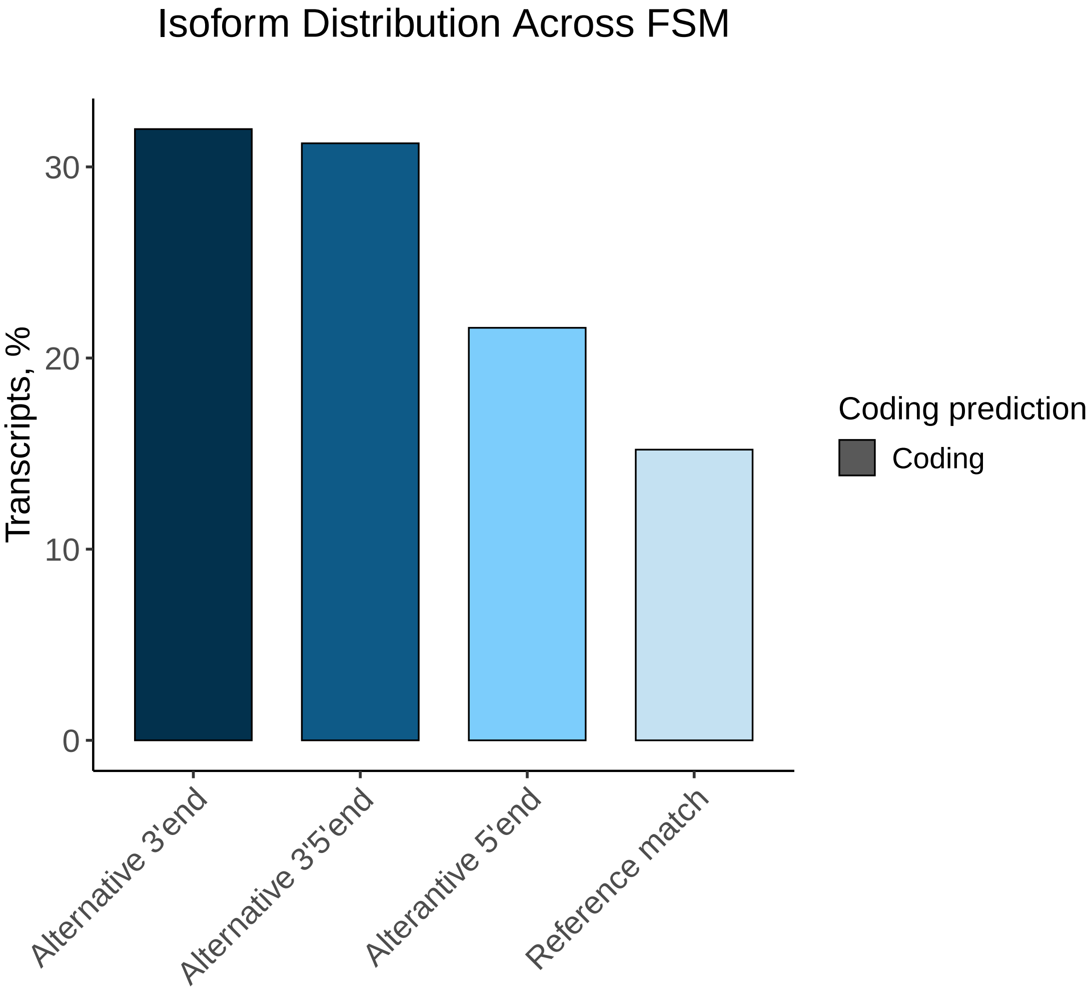


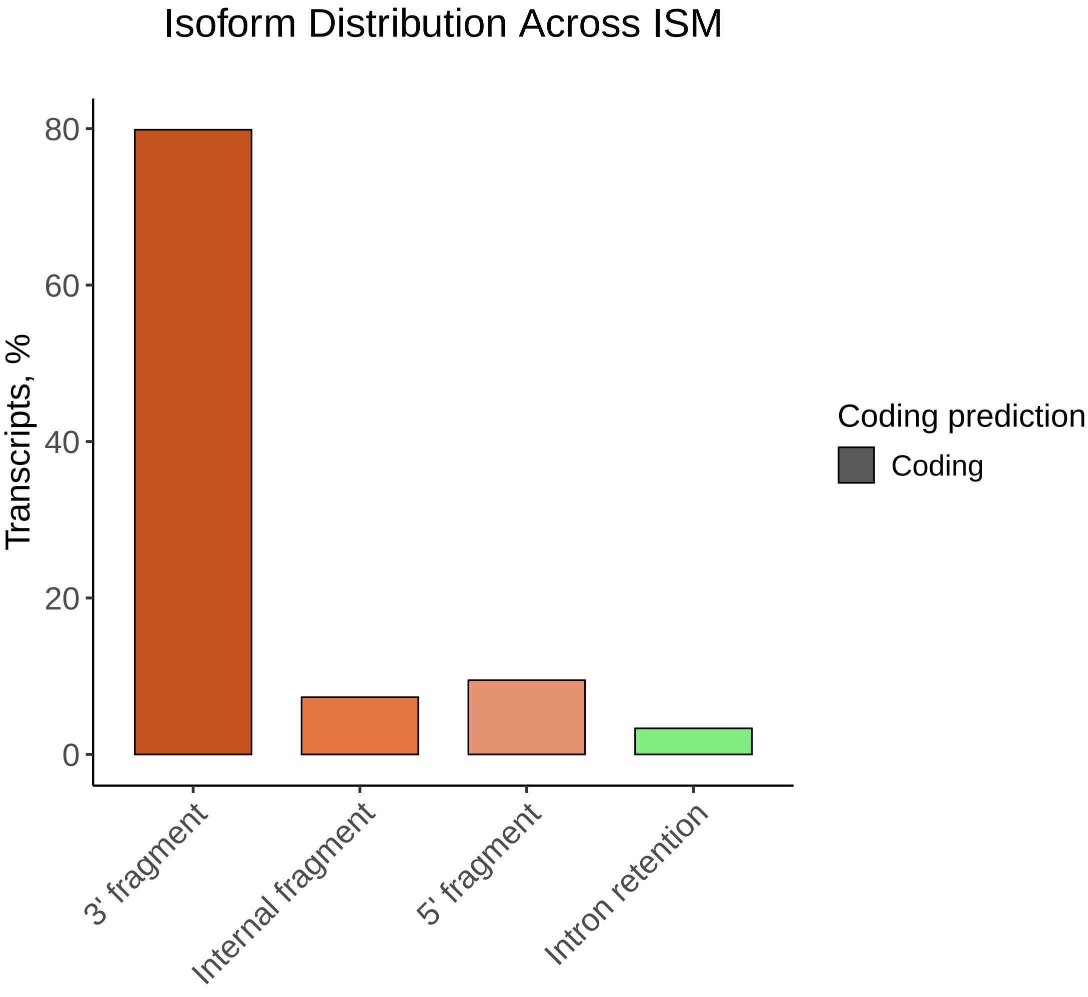

Mono− vs Multi−Exons Transcript Lengths by Sample
Count
25000
20000
15000
10000
5000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc
exonCat
Multi−Exon
Mono−Exon

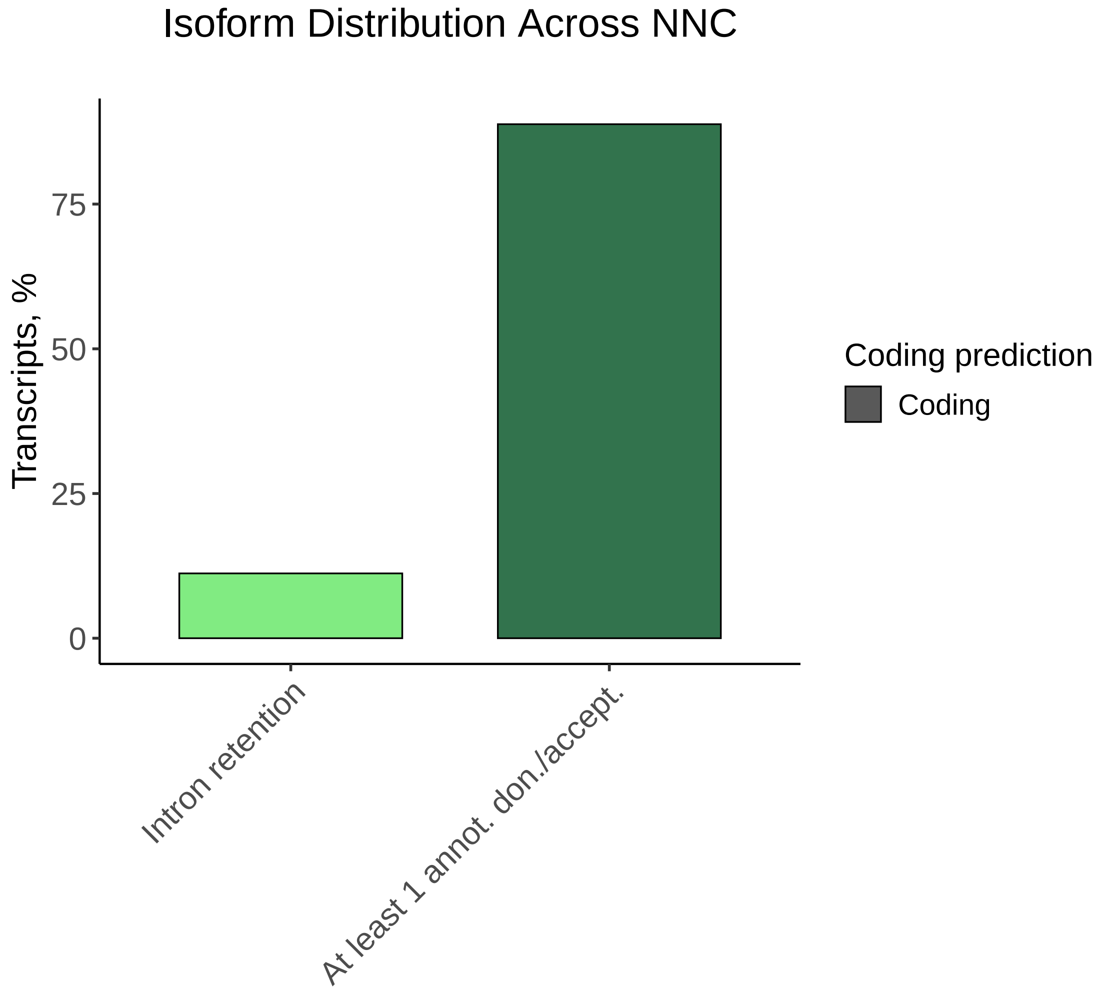
Structural Isoform Characterization

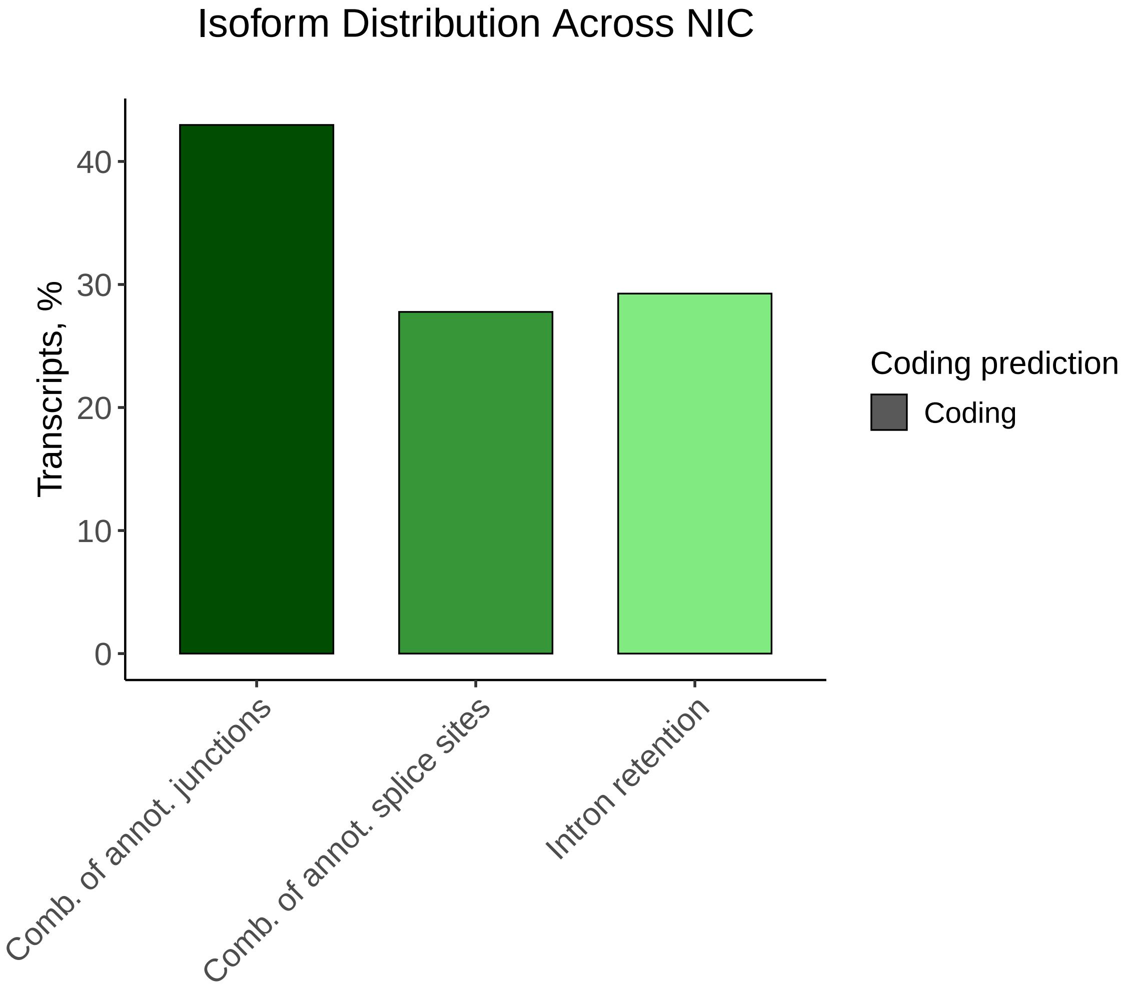
Isoform Distribution Across Structural Categories

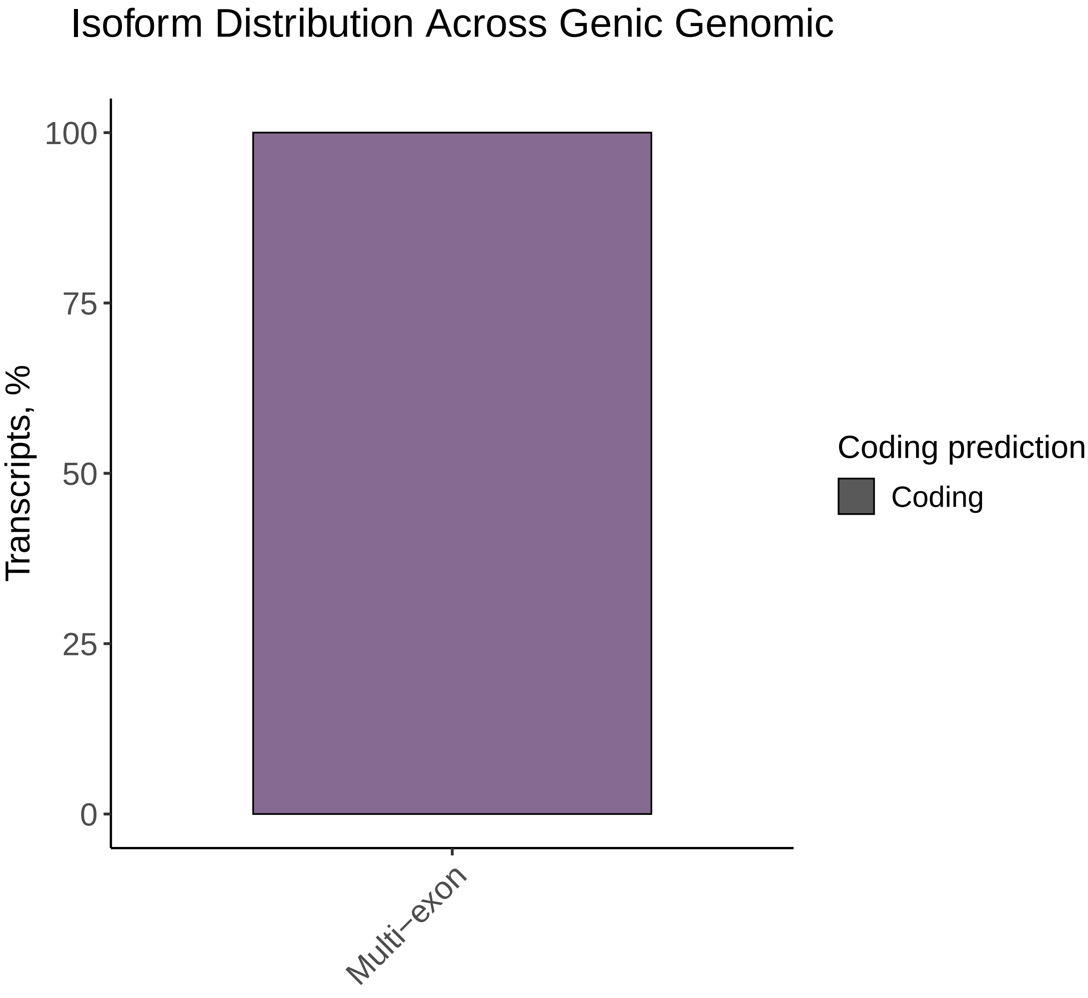


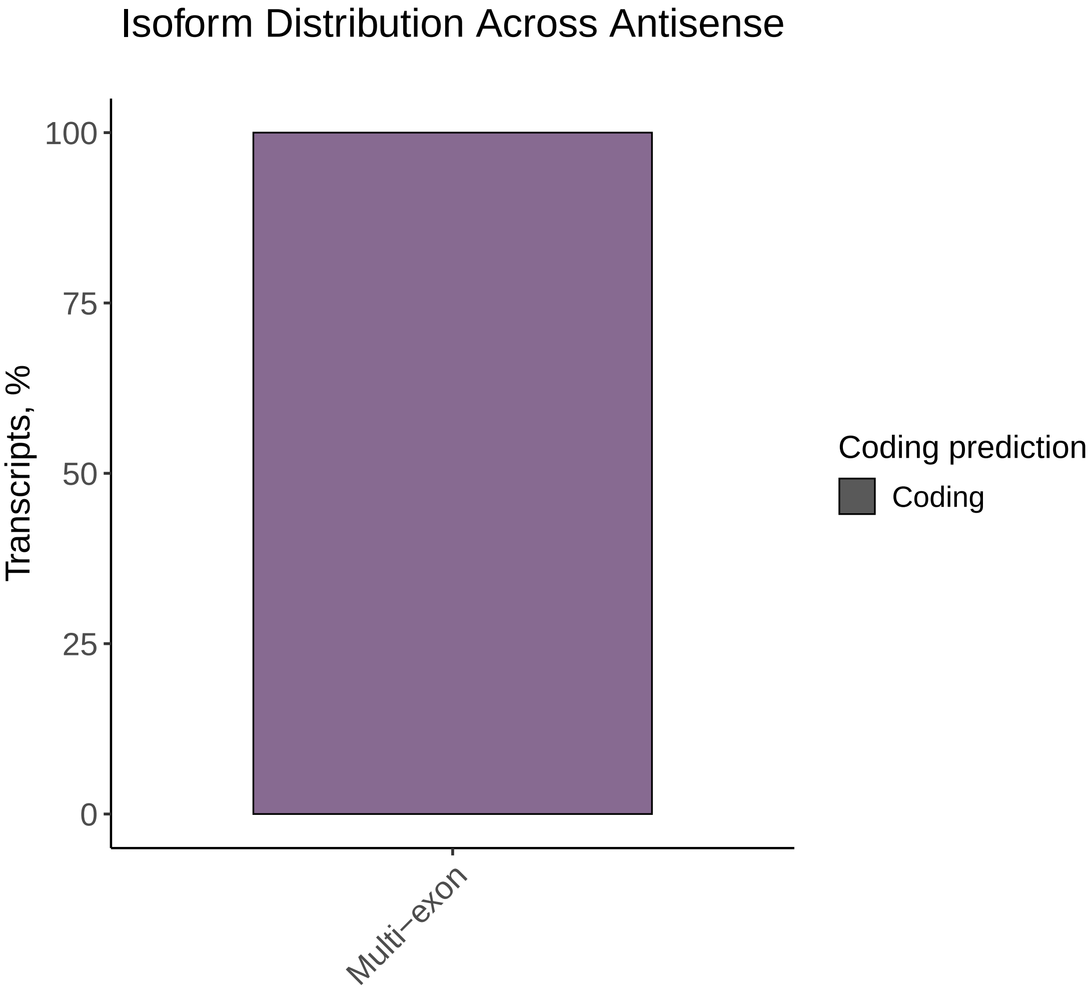

Isoform Distribution Across FSM
Transcripts, %
0
10
20
30
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Coding prediction
Coding

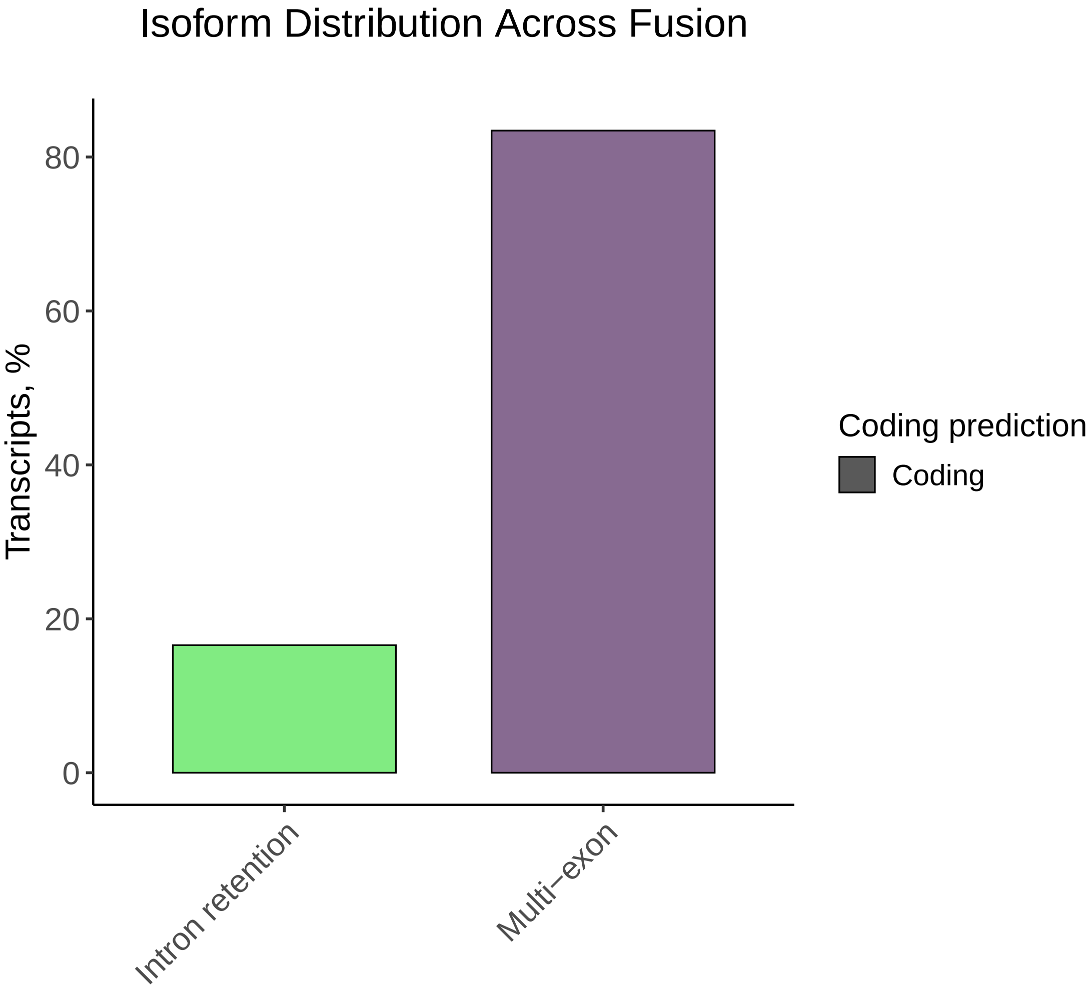

Isoform Distribution Across ISM
Transcripts, %
80
60
40
20
0
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding

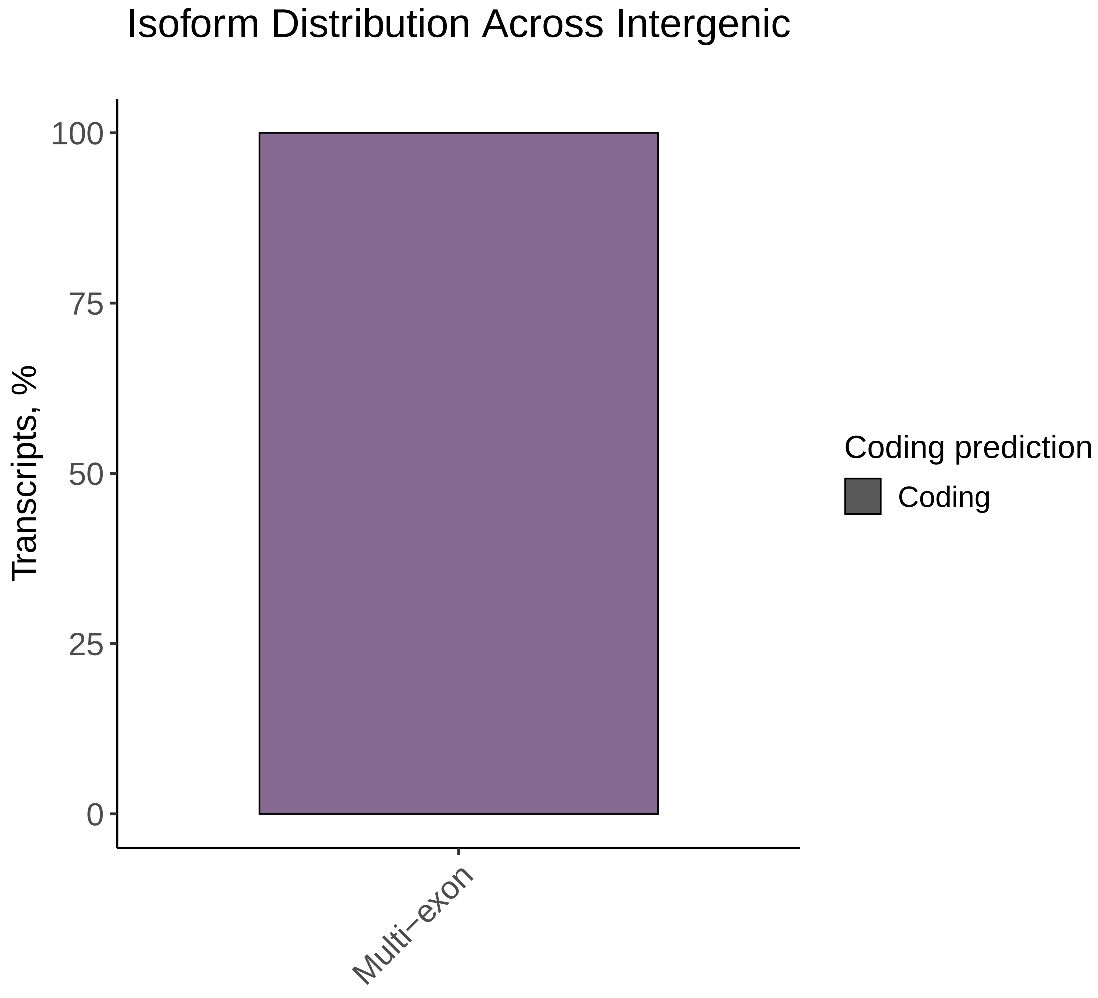

Isoform Distribution Across NNC
Transcripts, %
0
25
50
75
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding

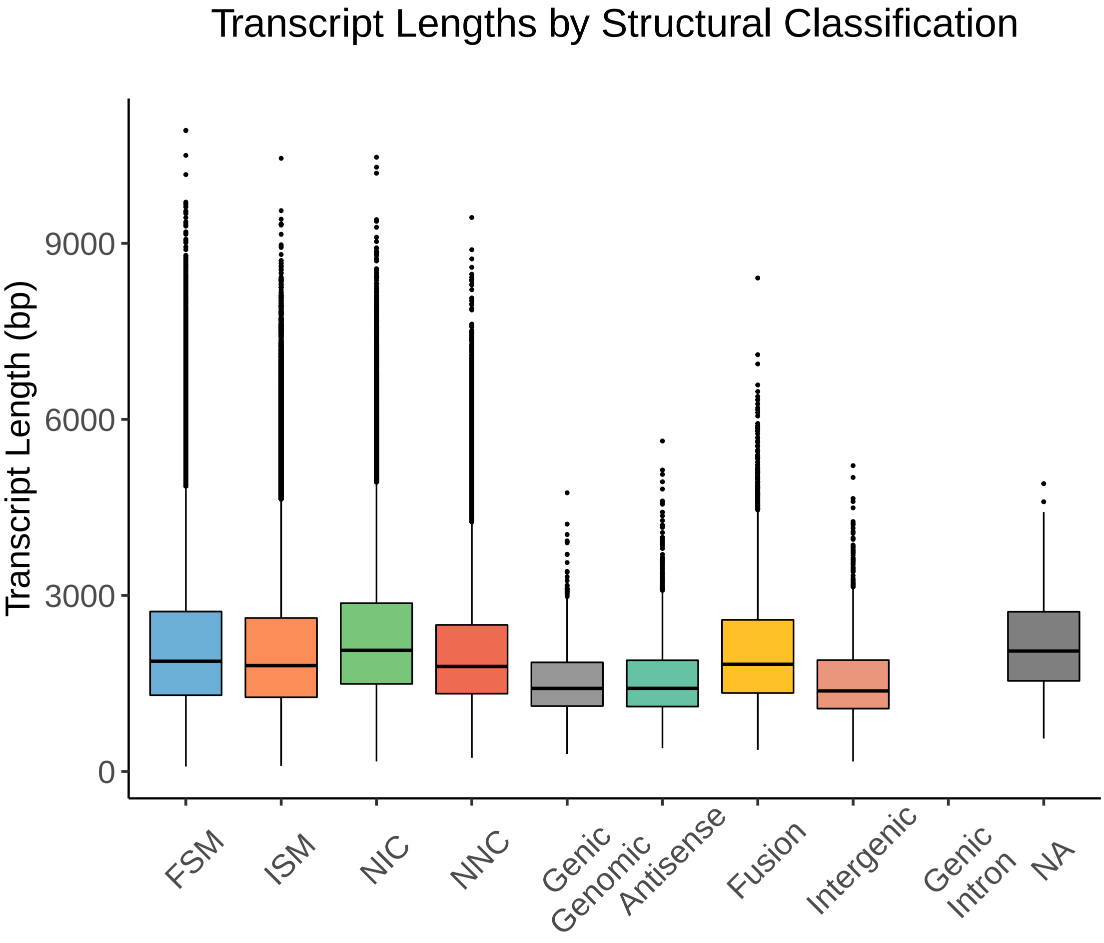

Isoform Distribution Across NIC
Transcripts, %
0
10
20
30
40
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Coding prediction
Coding

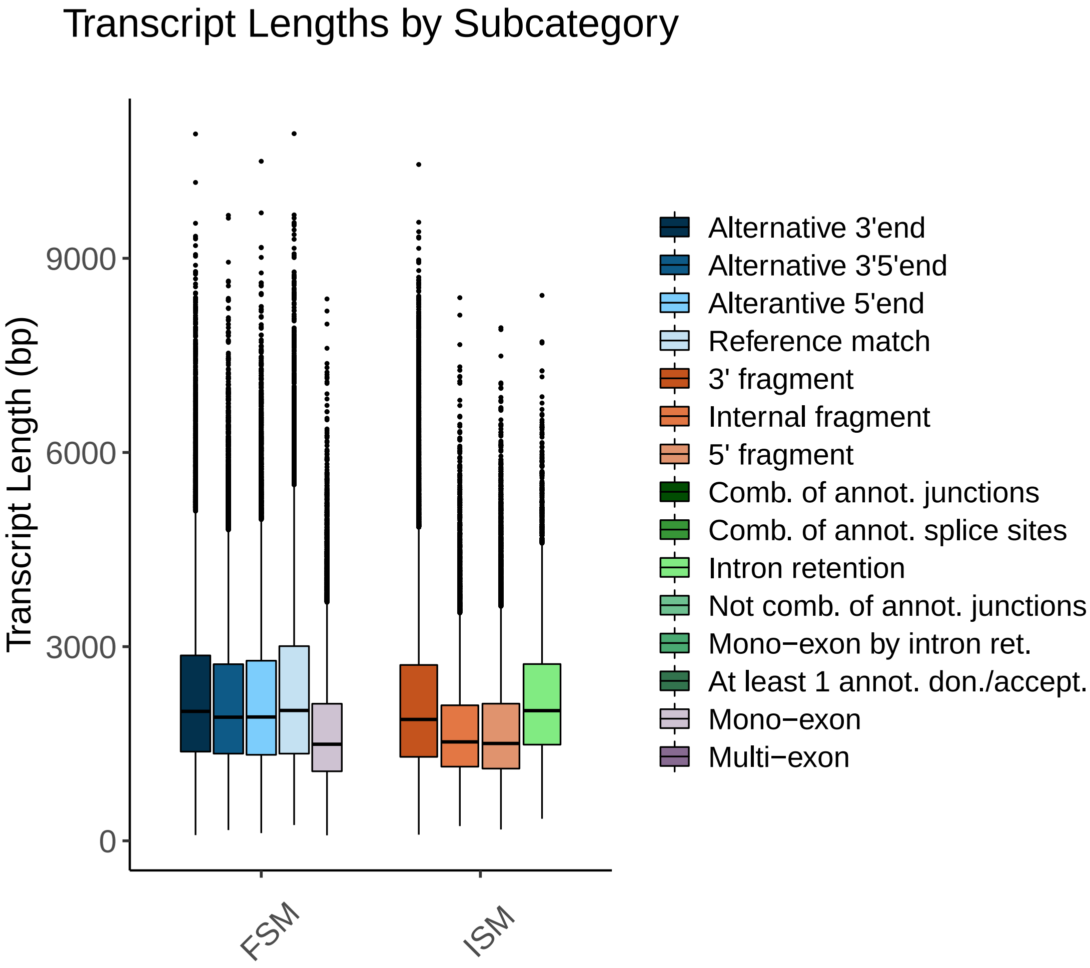

Isoform Distribution Across Genic Genomic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding

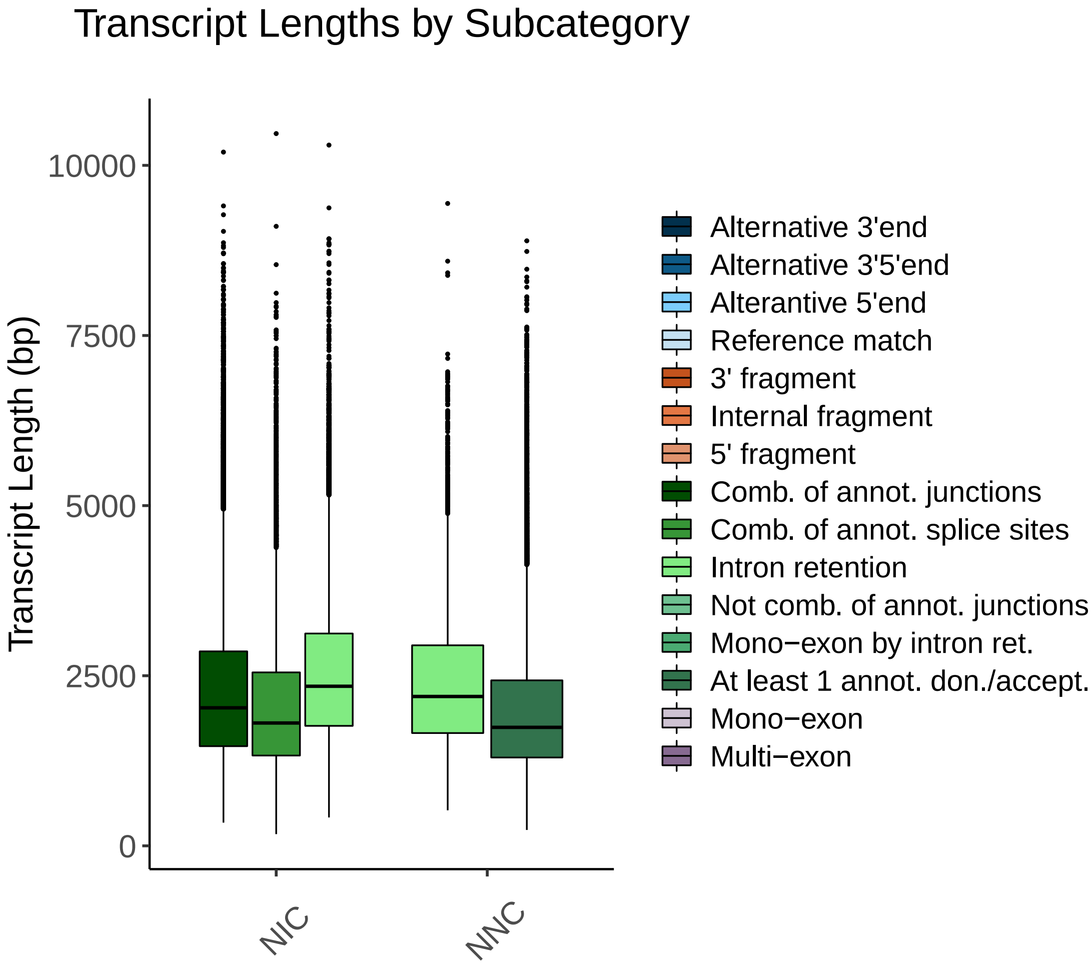

Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding

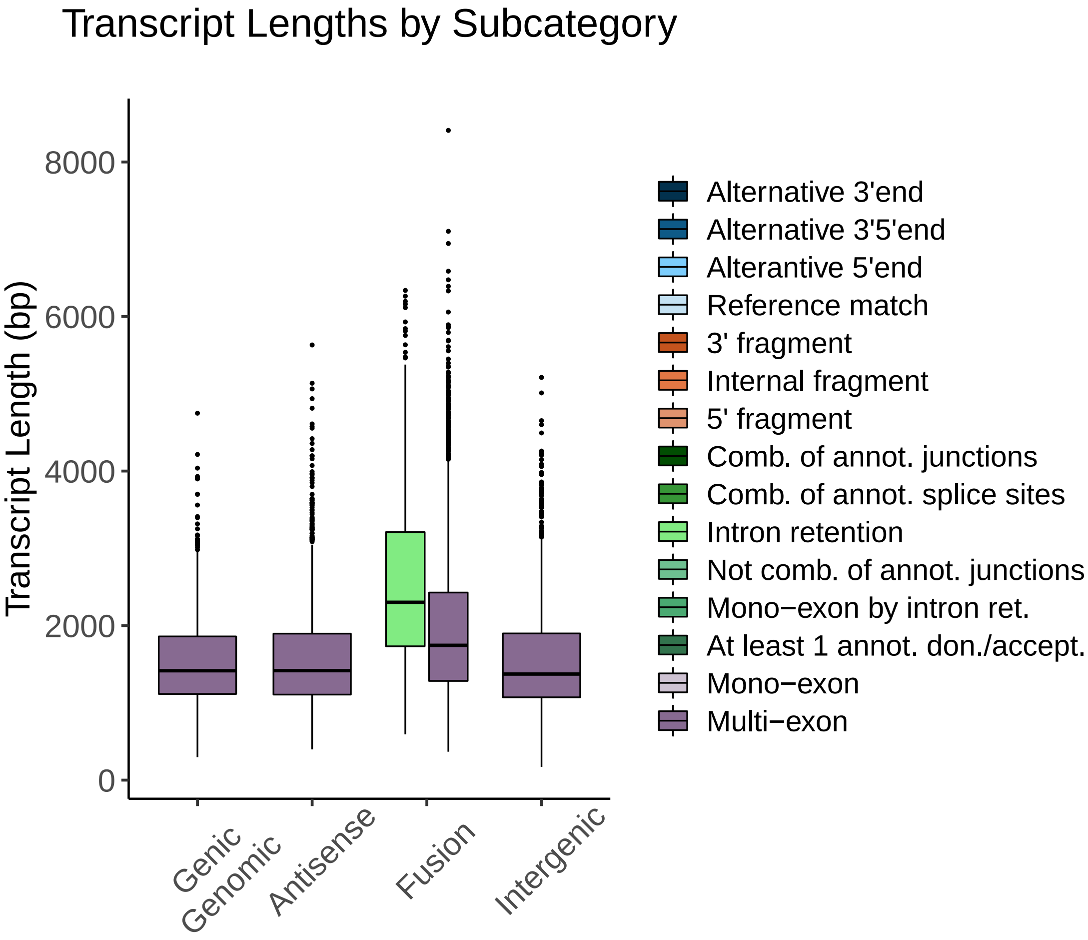

Isoform Distribution Across Fusion
Transcripts, %
0
20
40
60
80
Intron retention
Multi–exon
Coding prediction
Coding

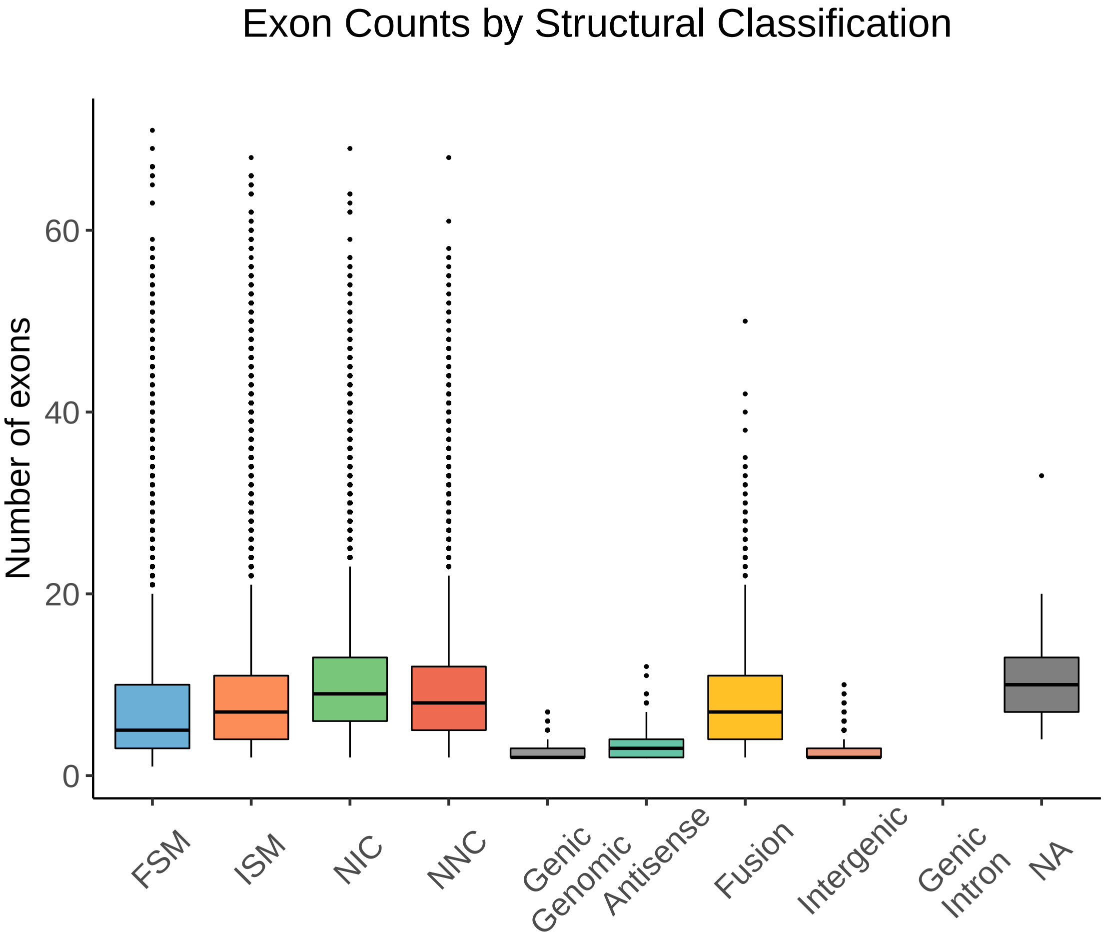

Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding

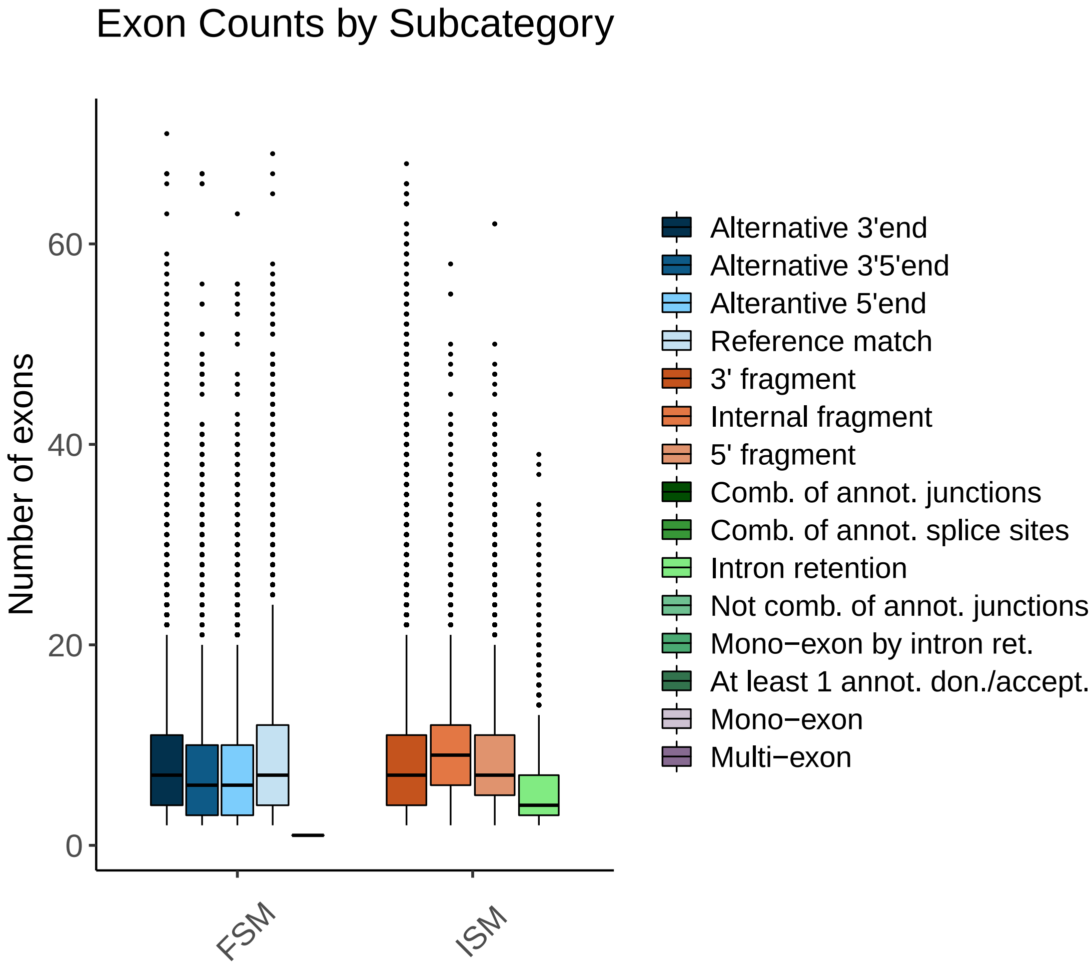

Transcript Lengths by Structural Classification
Transcript Length (bp)
0
3000
6000
9000
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

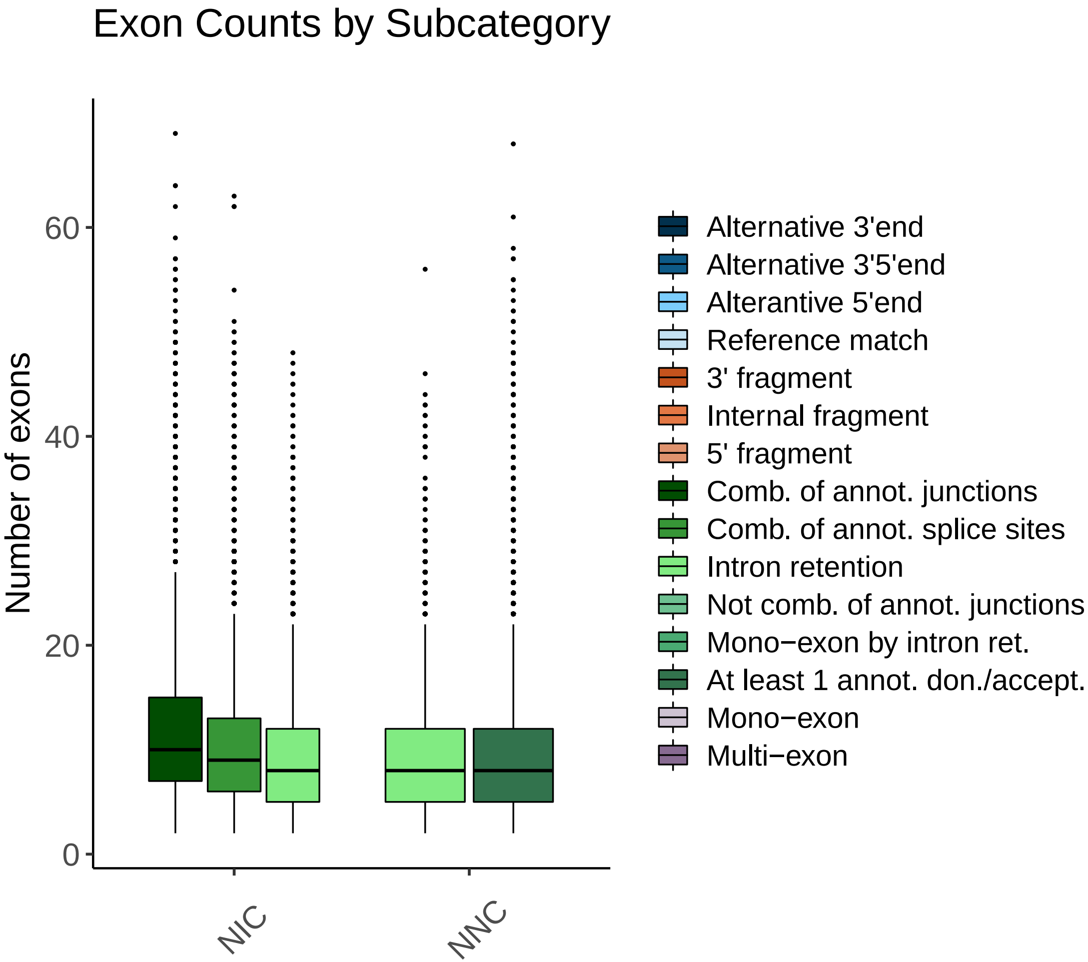

Transcript Lengths by Subcategory
Transcript Length (bp)
9000
6000
3000
0
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

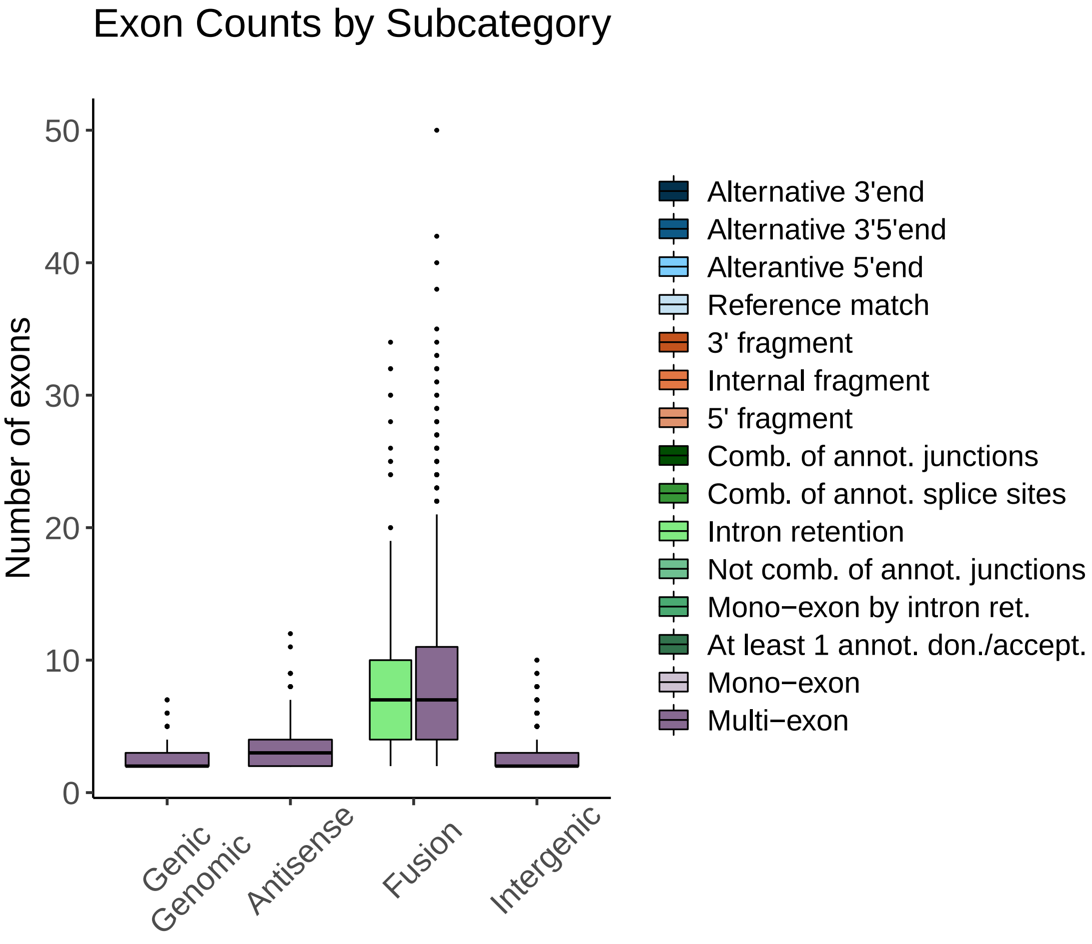

Transcript Lengths by Subcategory
Transcript Length (bp)
10000
7500
5000
2500
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

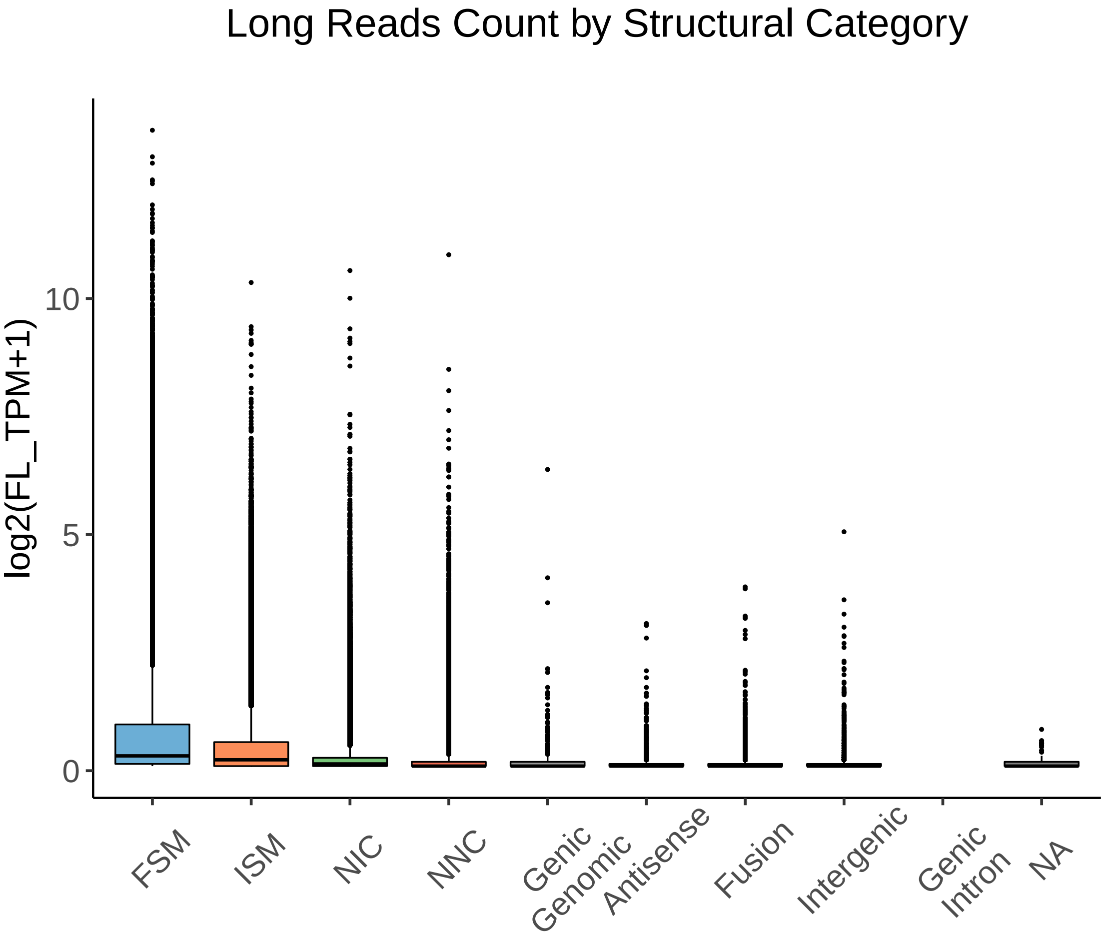

Transcript Lengths by Subcategory
Transcript Length (bp)
0
2000
4000
6000
8000
Genic
Genomic
Antisense
Fusion
Intergenic
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

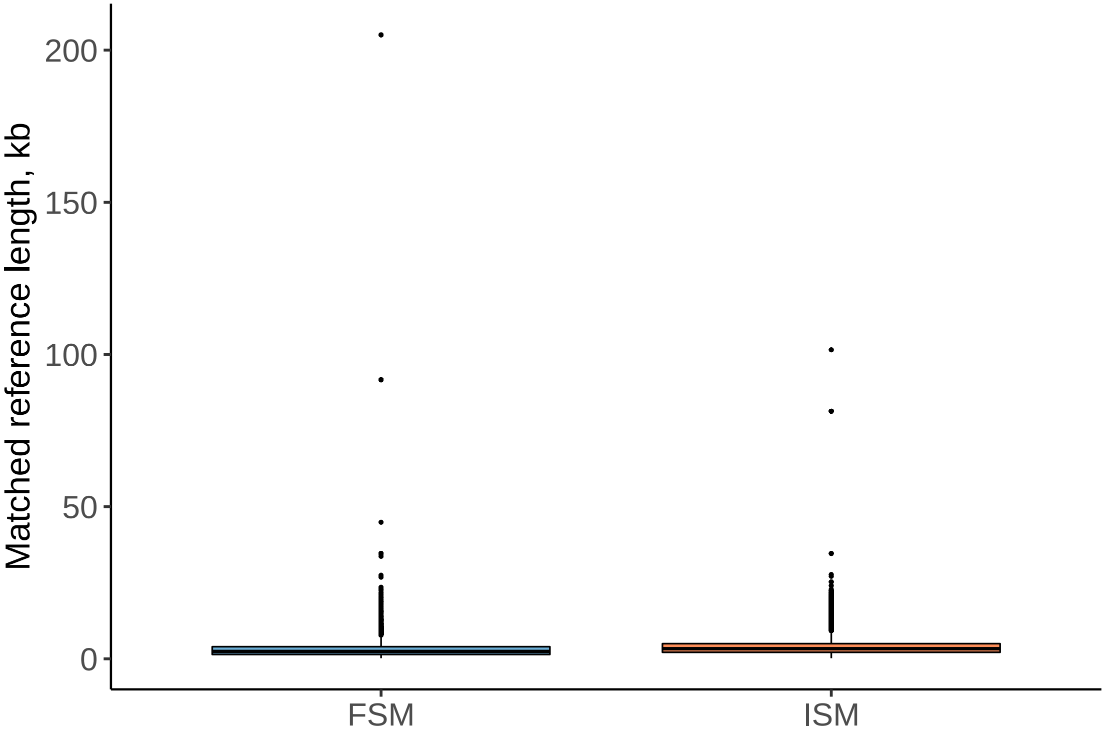

Exon Counts by Structural Classification
Number of exons
0
20
40
60
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA


Exon Counts by Subcategory
Number of exons
0
20
40
60
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon


Exon Counts by Subcategory
Number of exons
0
20
40
60
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon


Exon Counts by Subcategory
Number of exons
0
10
20
30
40
50
Genic Genomic
Antisense
Fusion
Intergenic
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon


Long Reads Count by Structural Category
log2(FL_TPM+1)
10
5
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

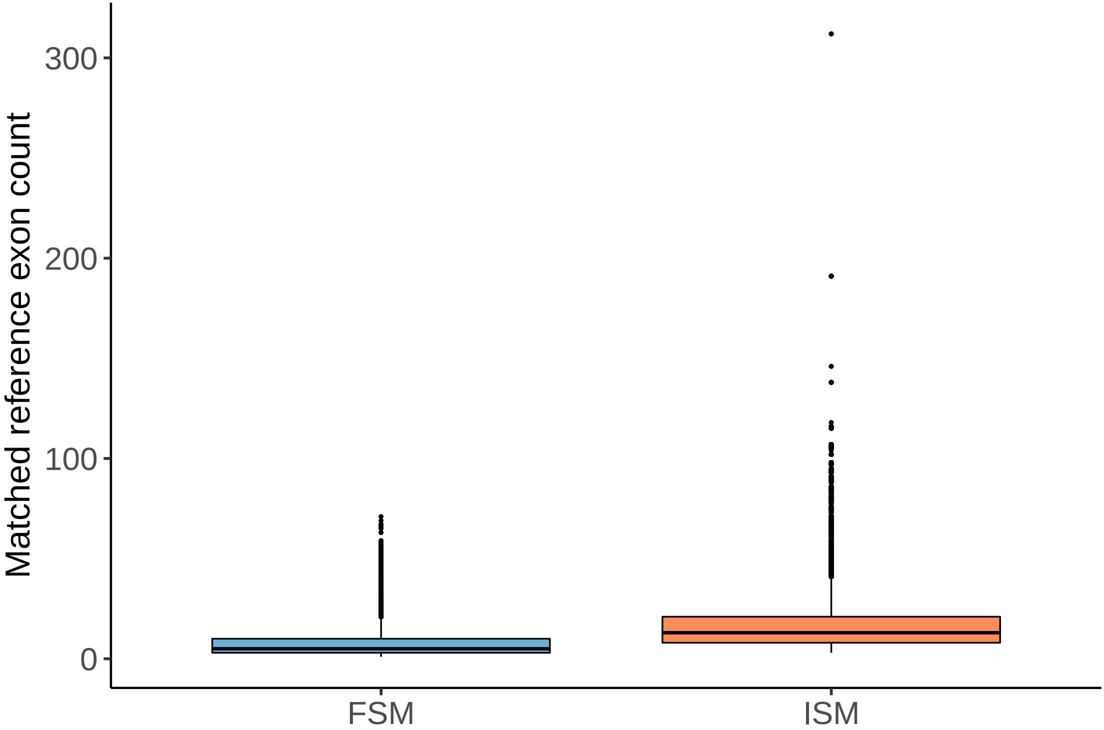


Length Distribution of Matched Reference Transcripts

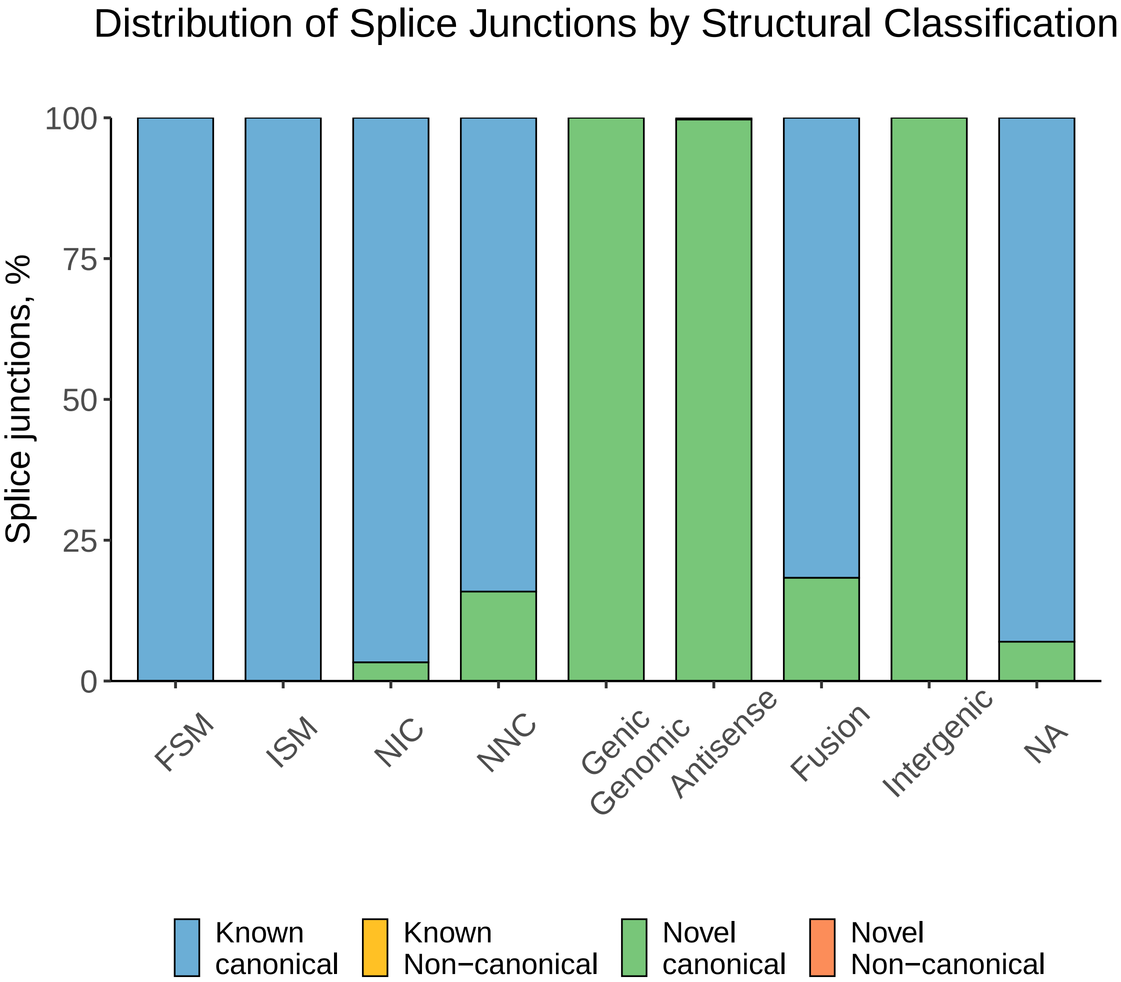
Applicable Only to FSM and ISM Categories

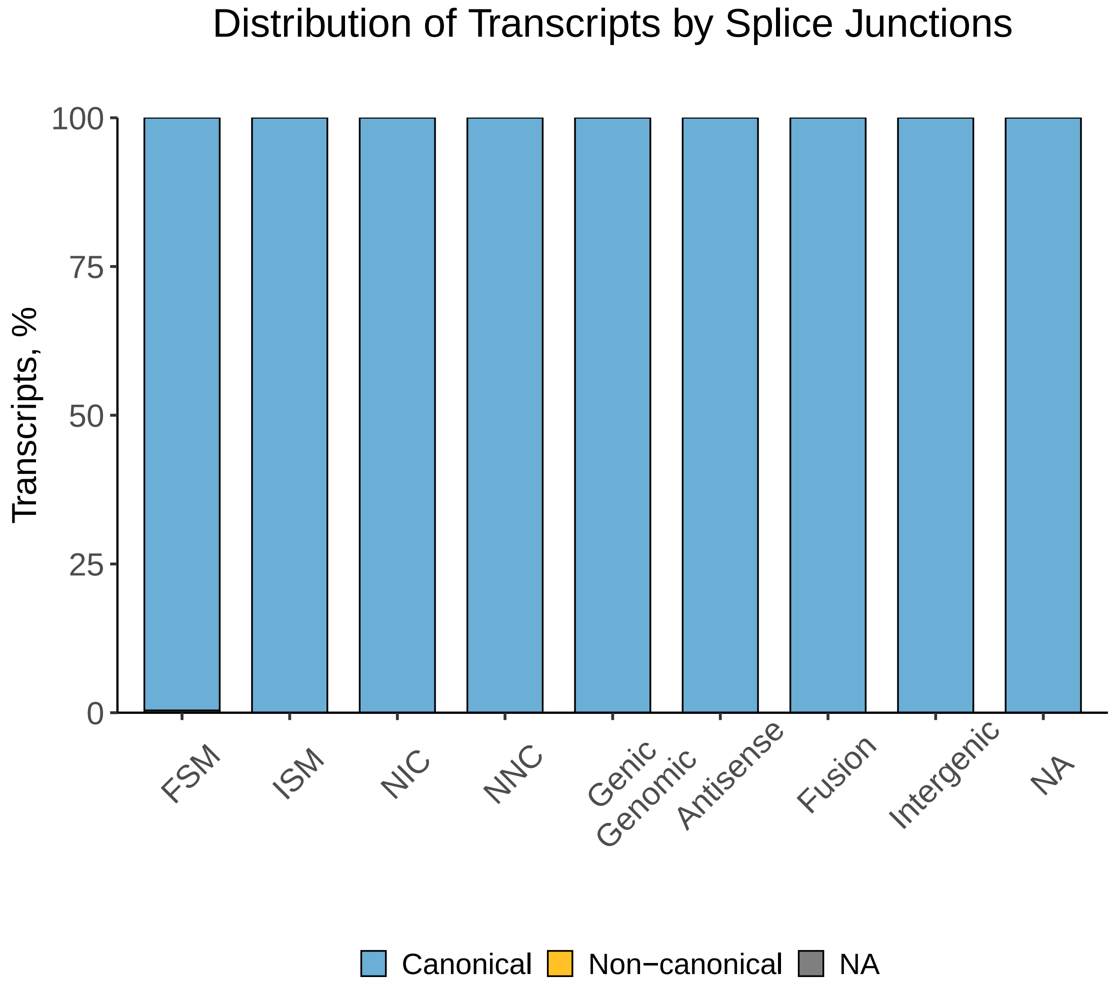
Exon Count Distribution of Matched Reference Transcripts

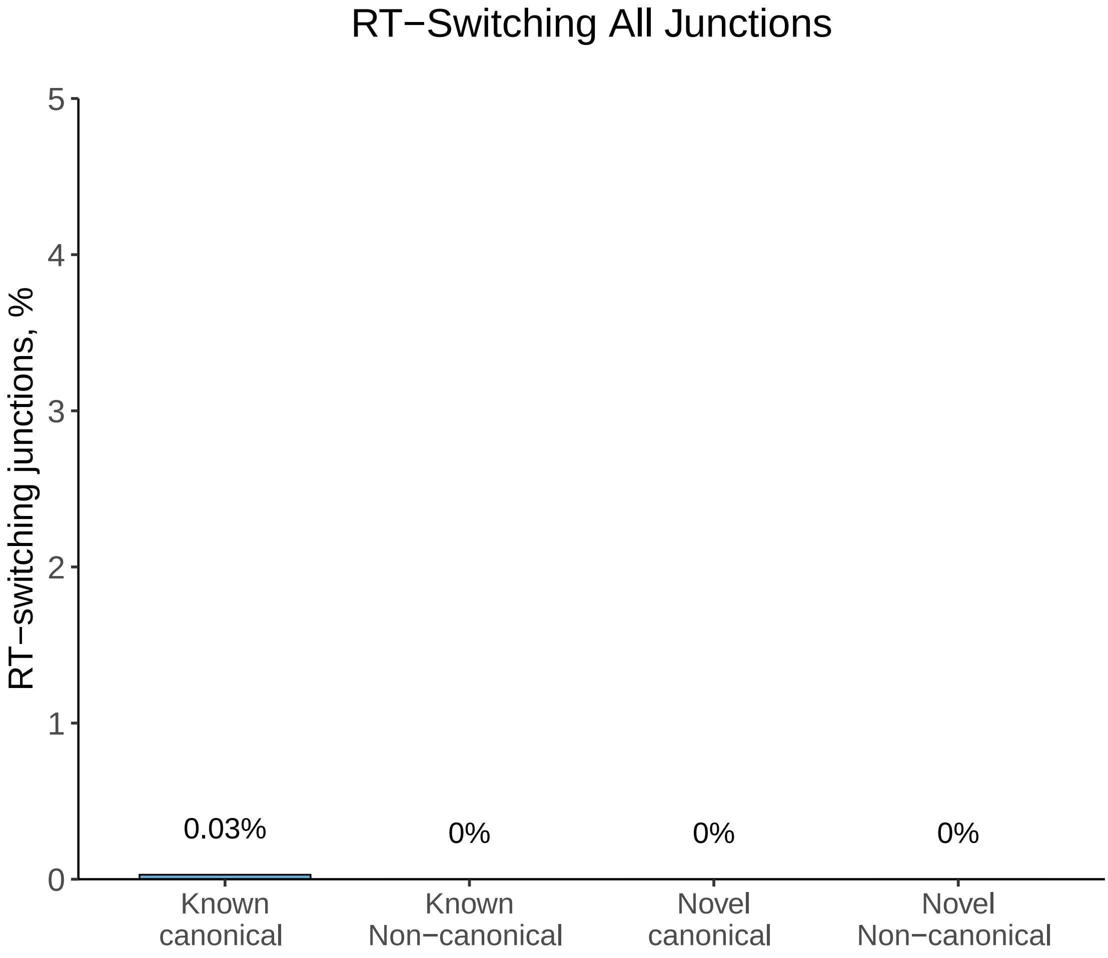
Applicable Only to FSM and ISM Categories

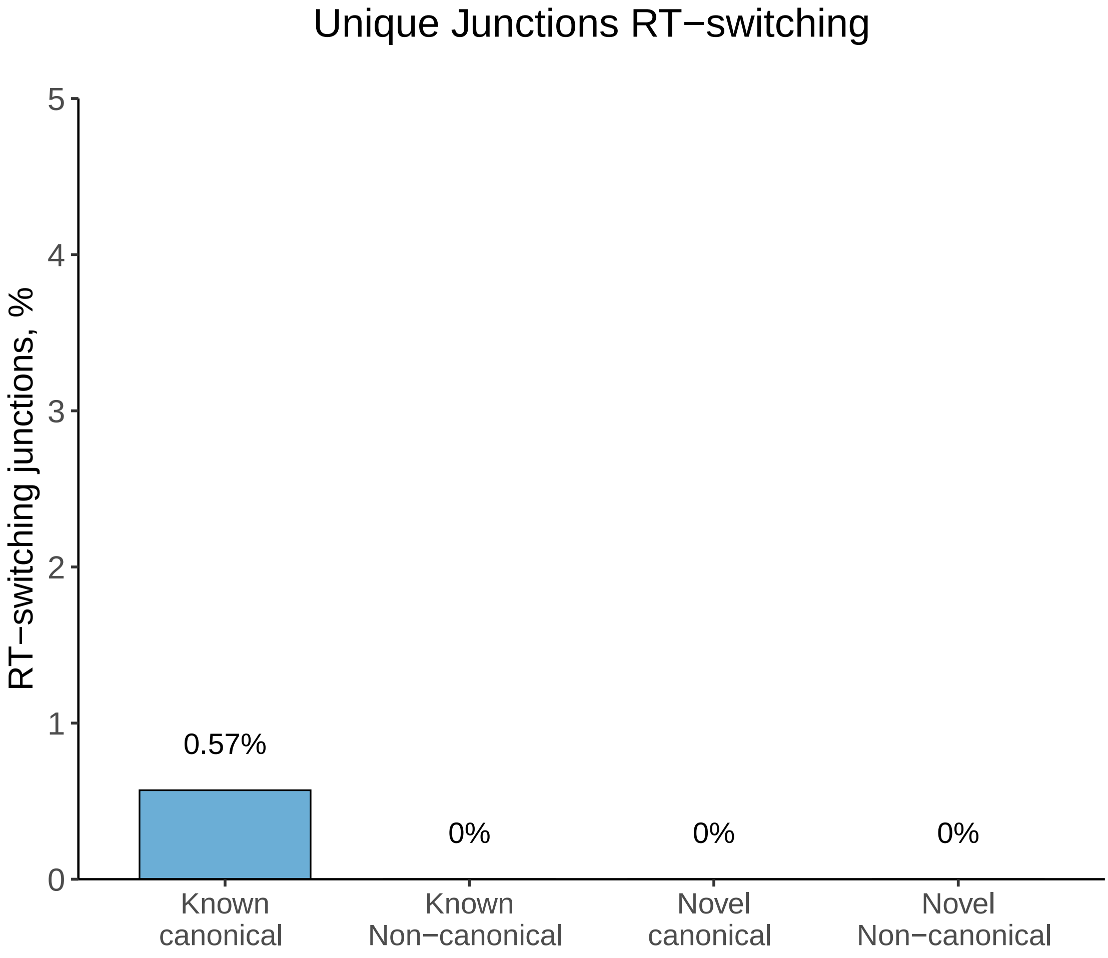


Splice Junction Characterization

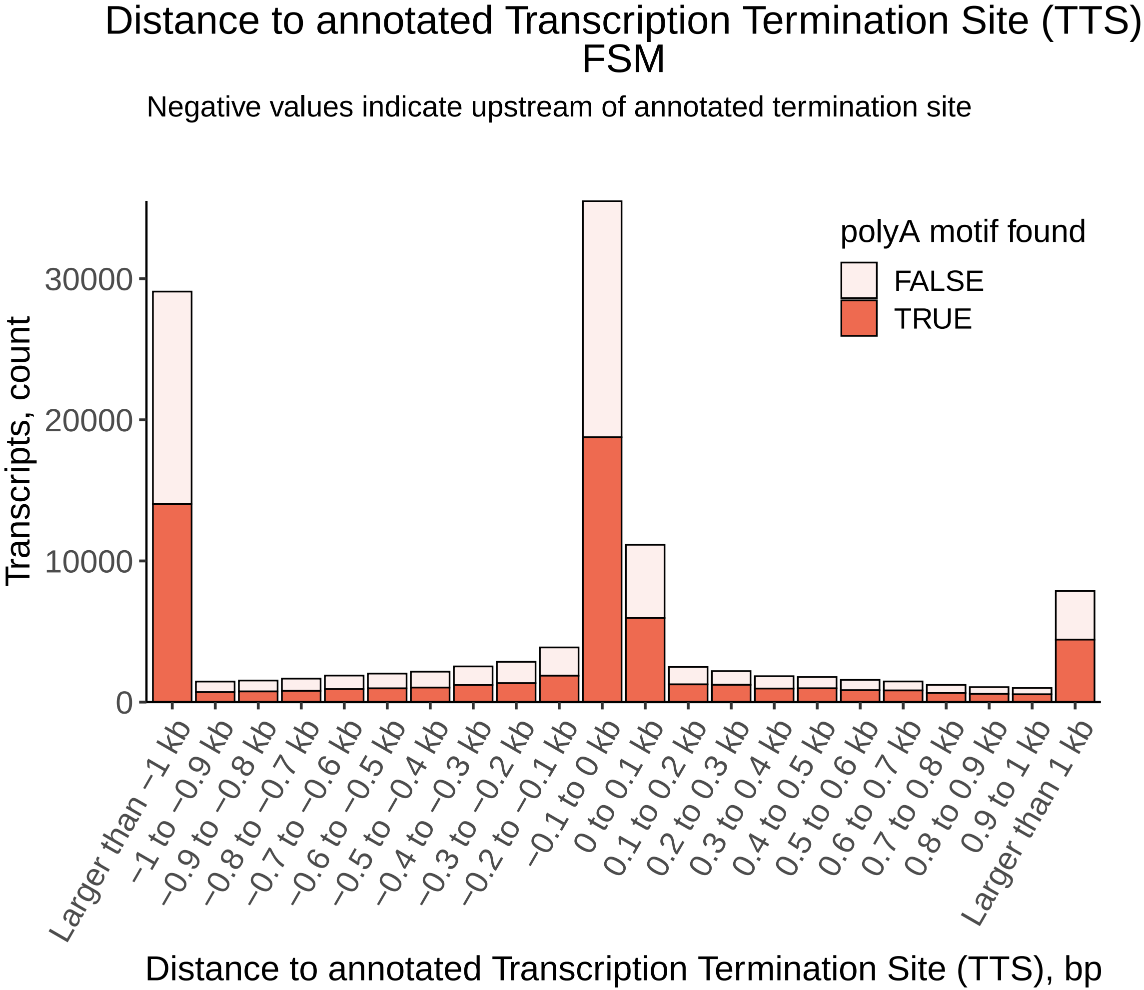

Distribution of Splice Junctions by Structural Classification
Splice junctions, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic
Genomic
Antisense
Fusion
Intergenic
NA
Known
canonical
Known
Non−canonical
Novel
canonical
Novel
Non−canonical


Distribution of Transcripts by Splice Junctions
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic
Genomic
Antisense
Fusion
Intergenic
NA
Canonical
Non−canonical
NA


RT−Switching All Junctions
RT−switching junctions, %
5
4
3
2
1
0
0.03%
0%
0%
0%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

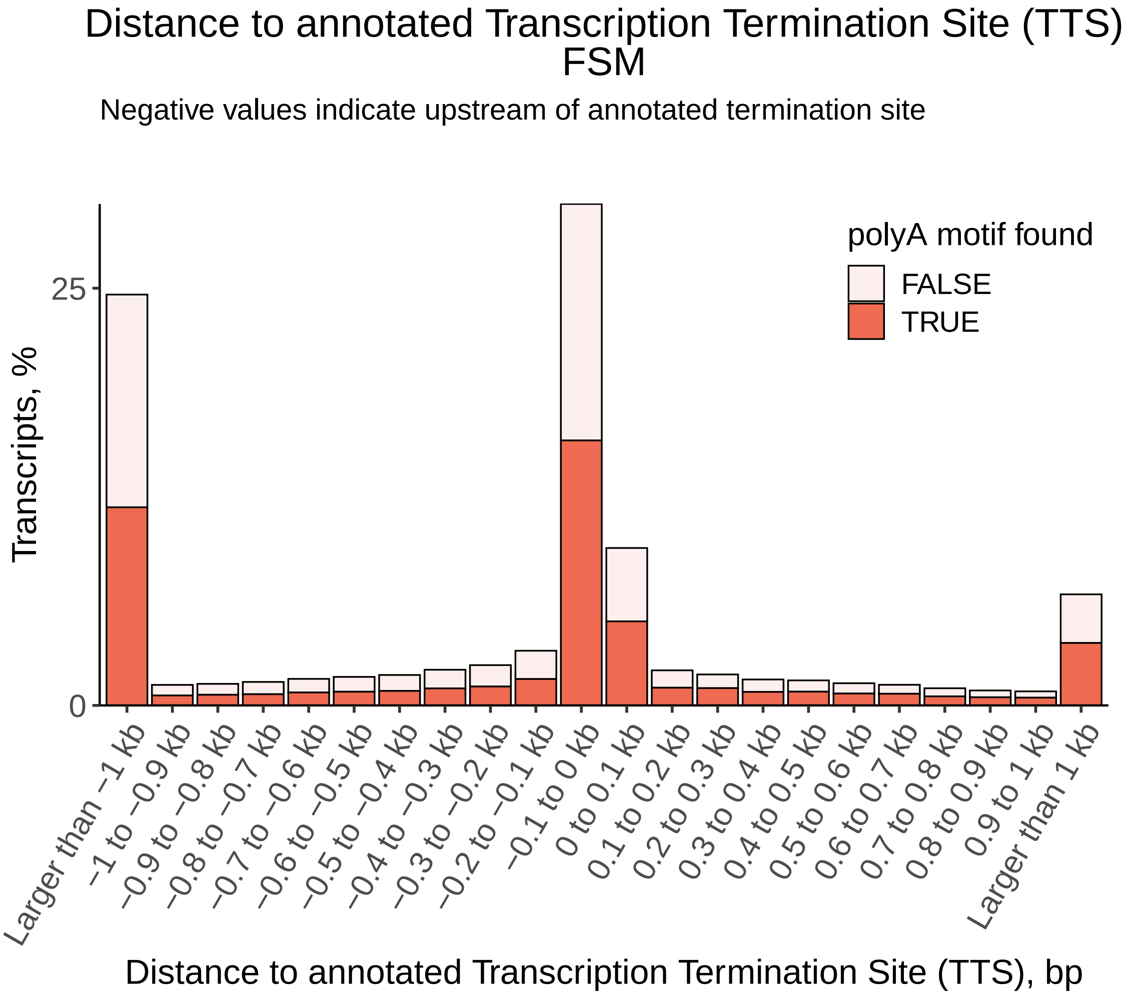

Unique Junctions RT−switching
RT−switching junctions, %
5
4
3
2
1
0
0.57%
0%
0%
0%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

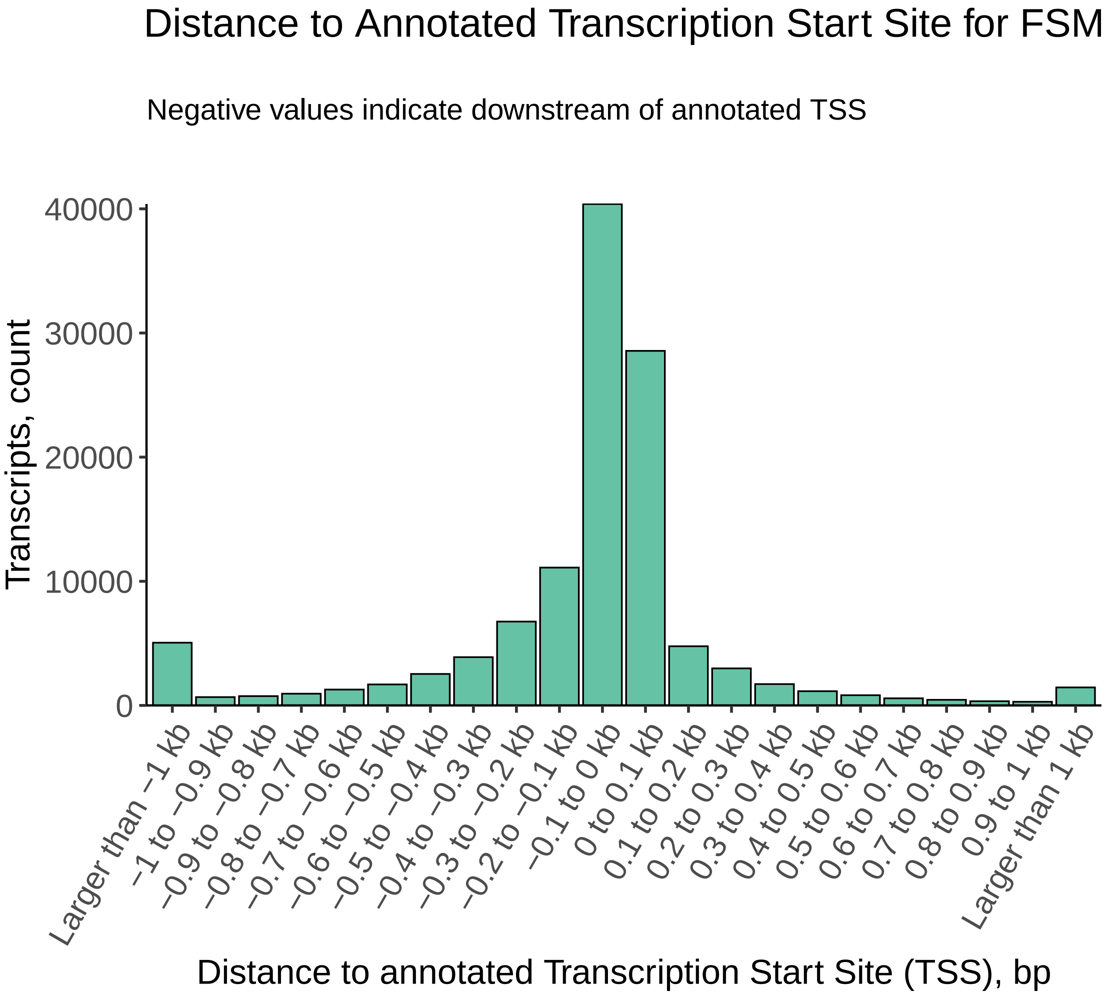
Comparison With Annotated TSS and TTS

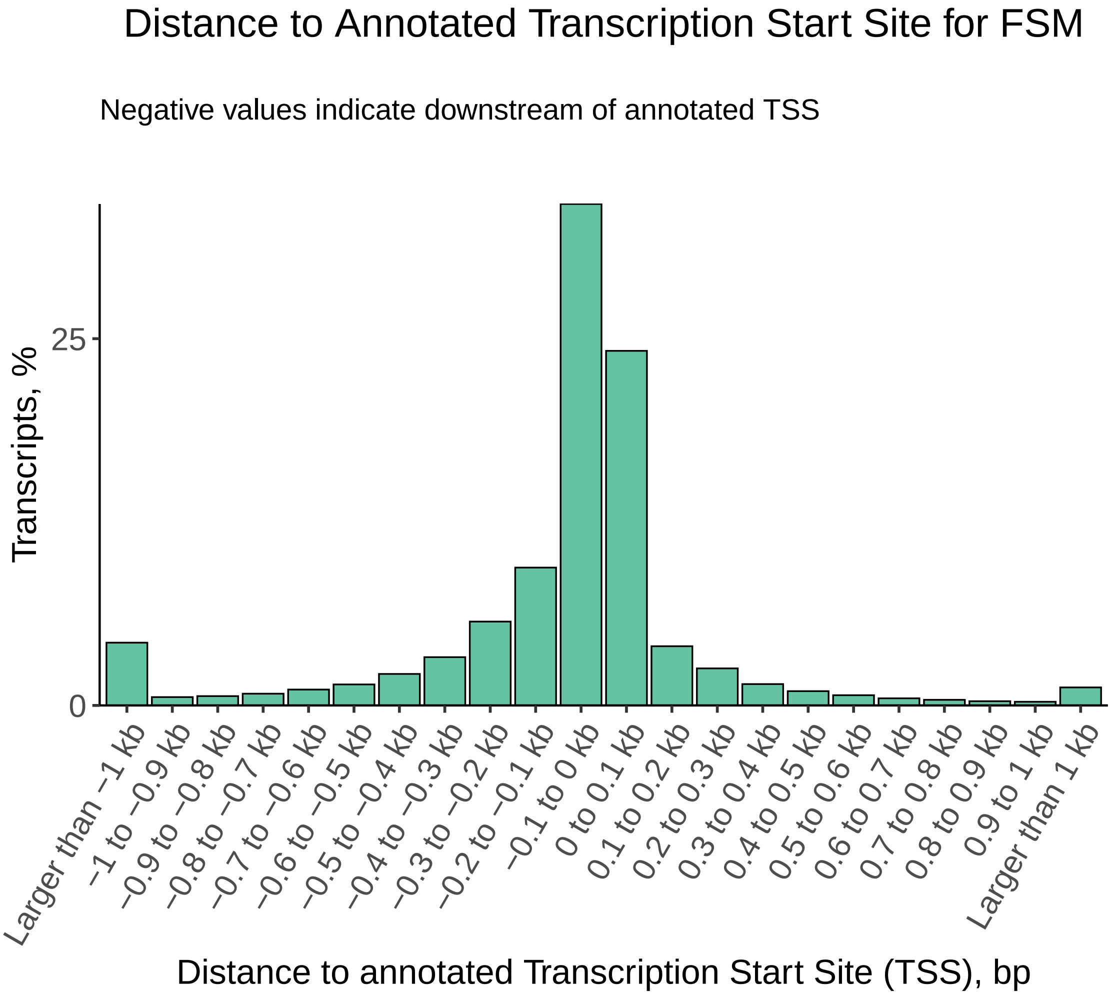


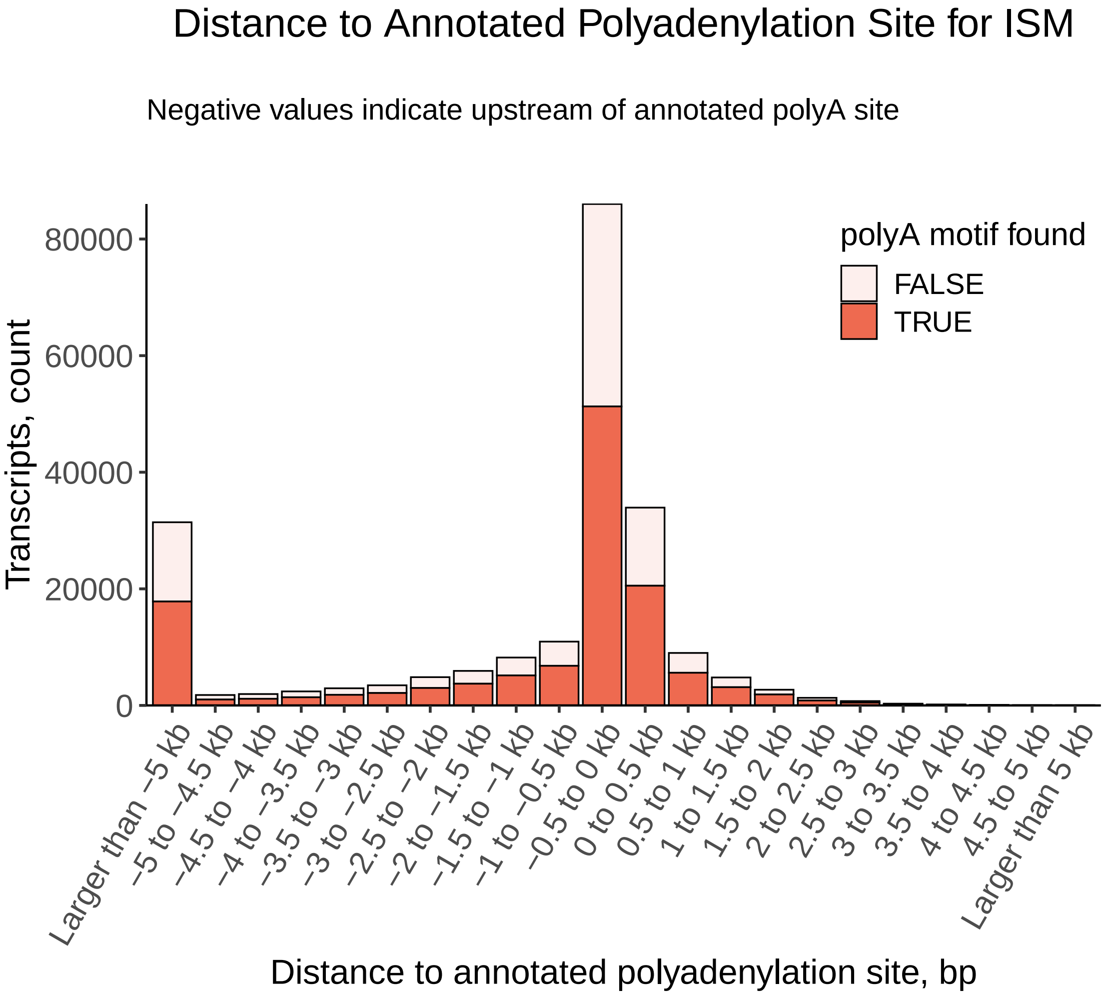
Distance to annotated Transcription Termination Site (TTS)
FSM

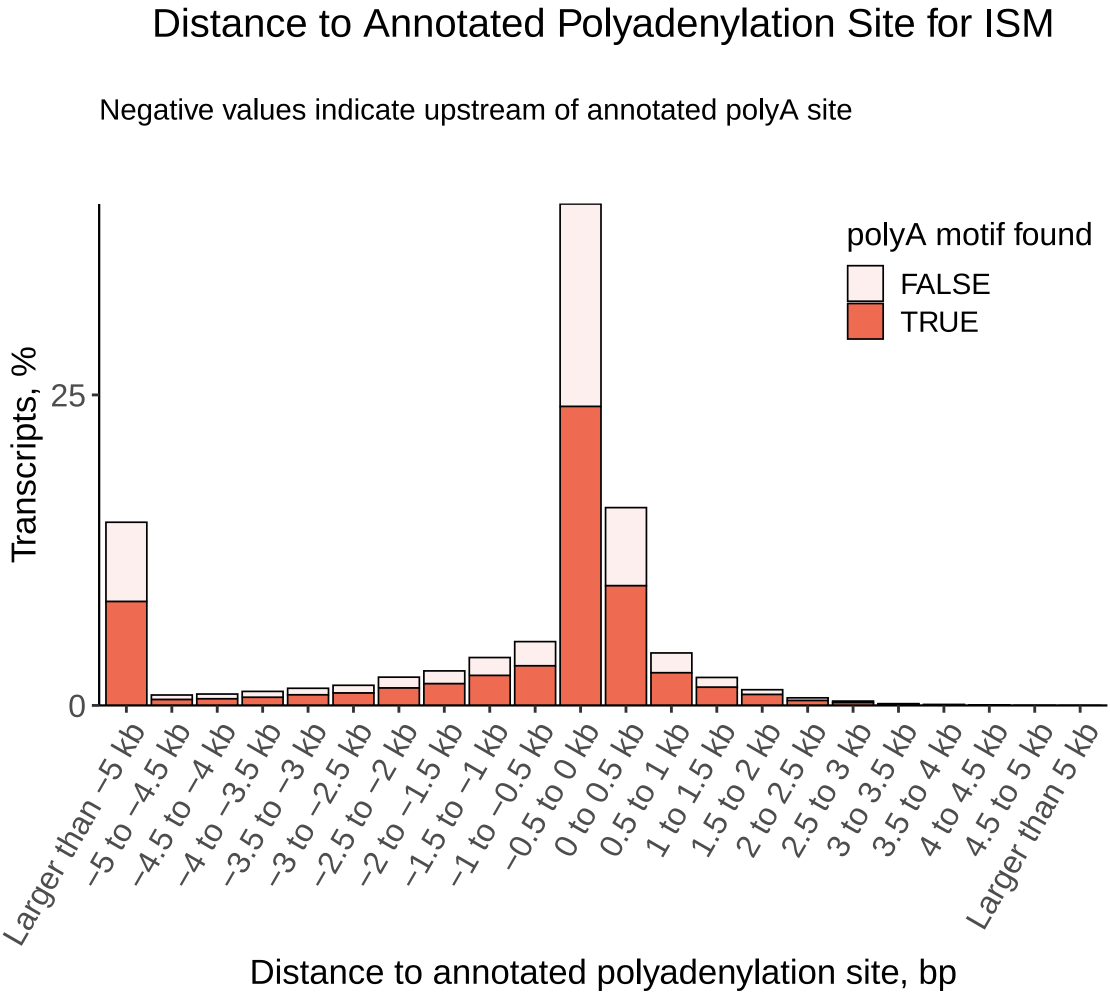
Negative values indicate upstream of annotated termination site


Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Termination Site (TTS), bp


Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
Transcripts, count
40000
30000
20000
10000
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Start Site (TSS), bp


Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Start Site (TSS), bp

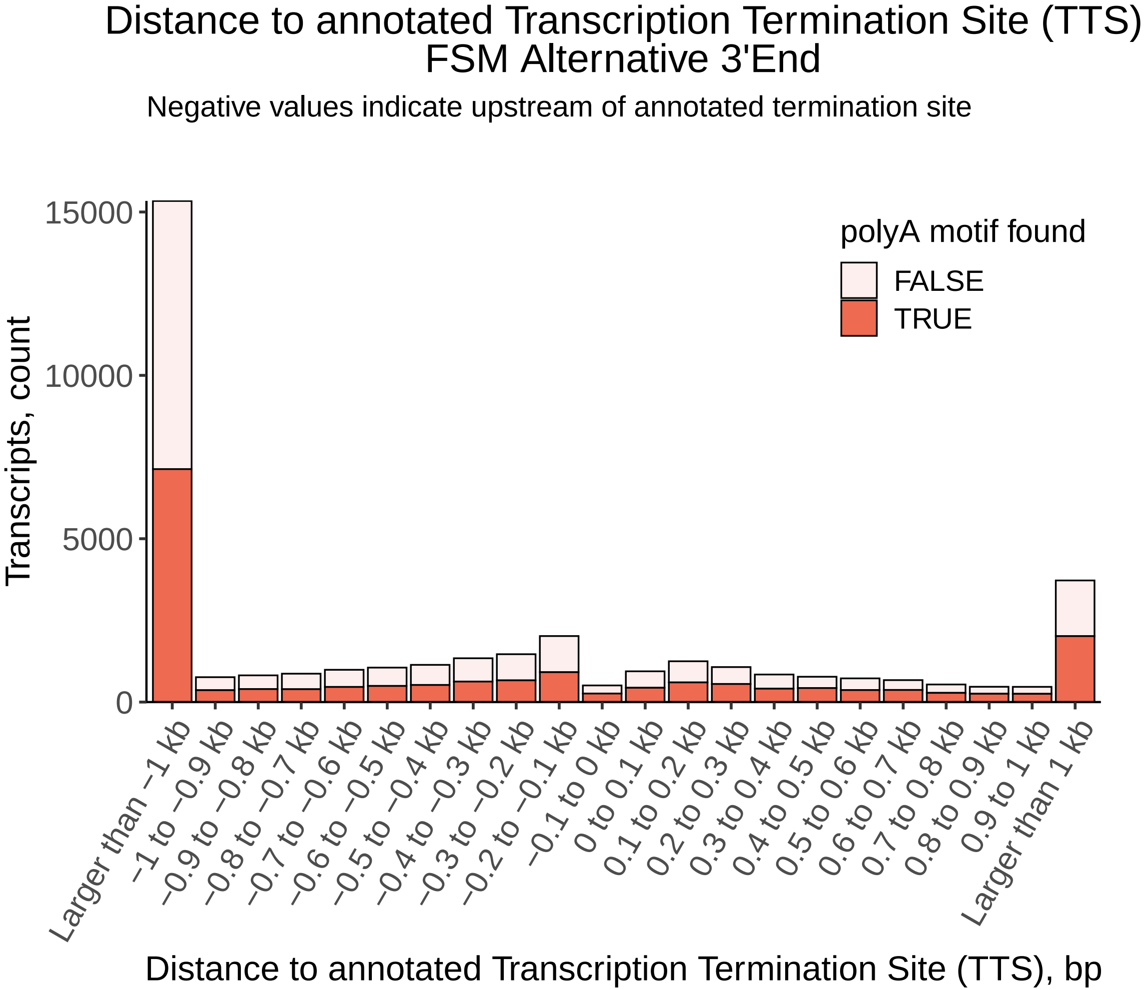

Distance to Annotated Polyadenylation Site for ISM
Negative values indicate upstream of annotated polyA site
Transcripts, count
80000
60000
40000
20000
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp

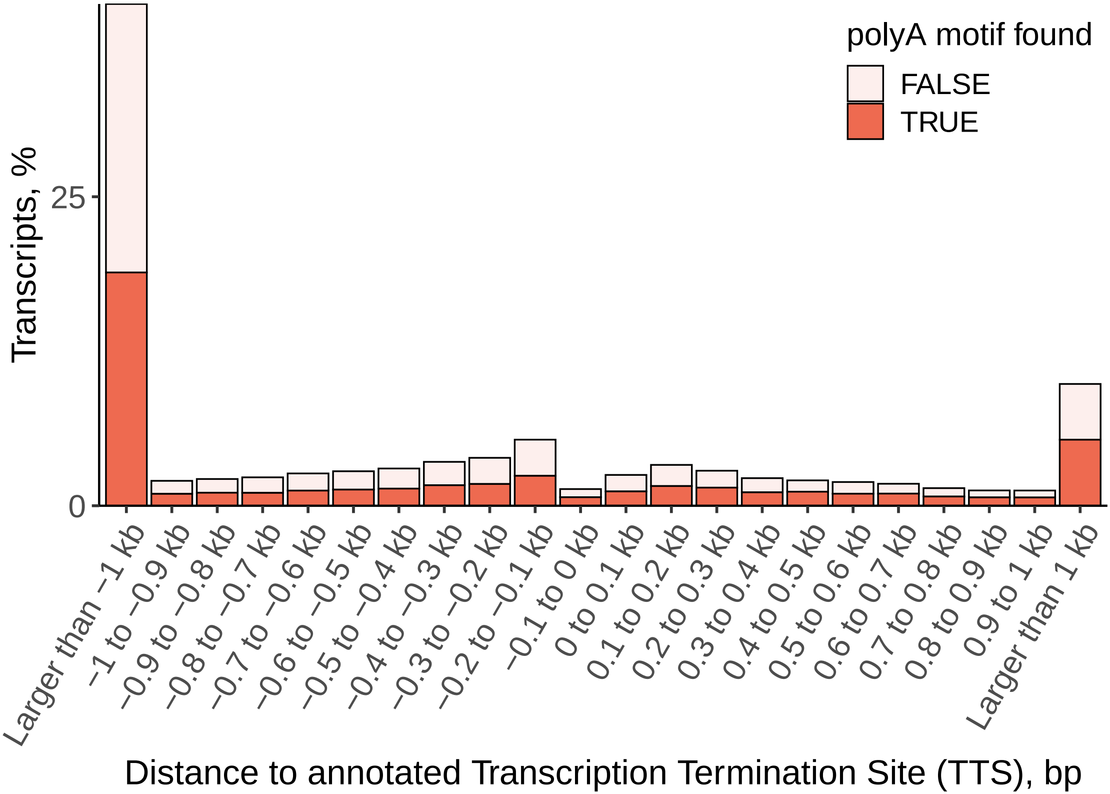


Distance to Annotated Polyadenylation Site for ISM

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
Transcripts, count
15000
10000
5000
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Termination Site (TTS), bp

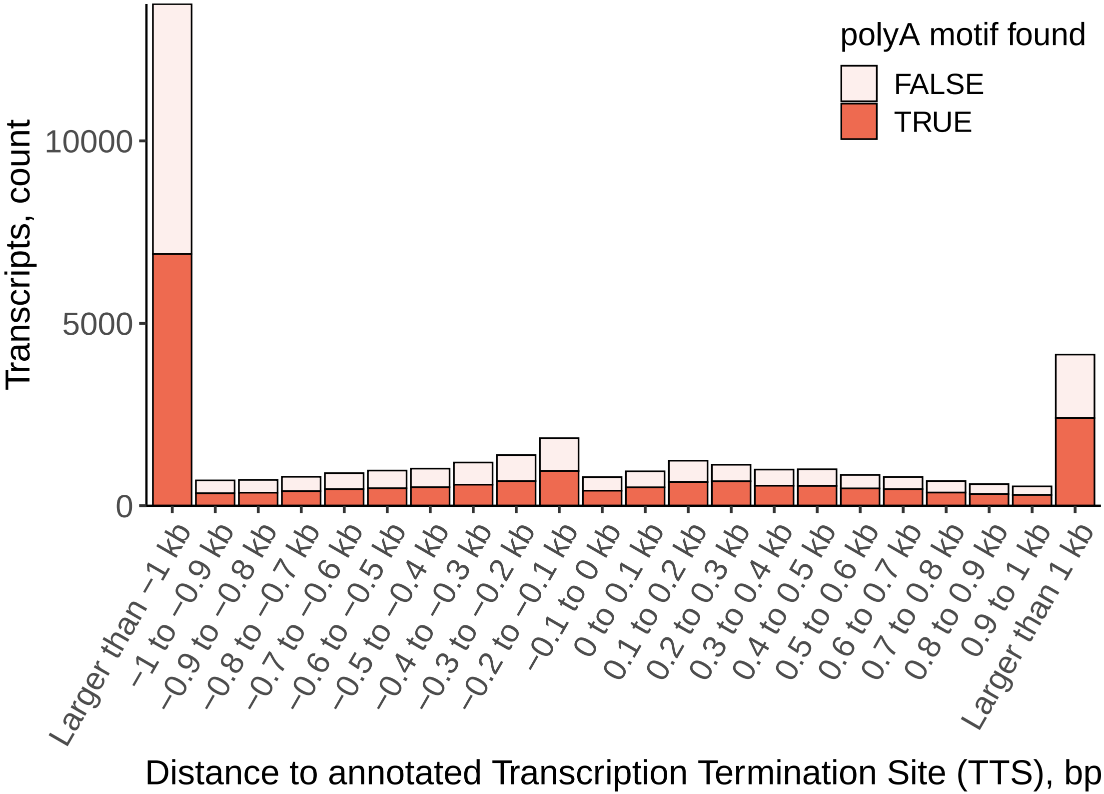


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End

Negative values indicate upstream of annotated termination site

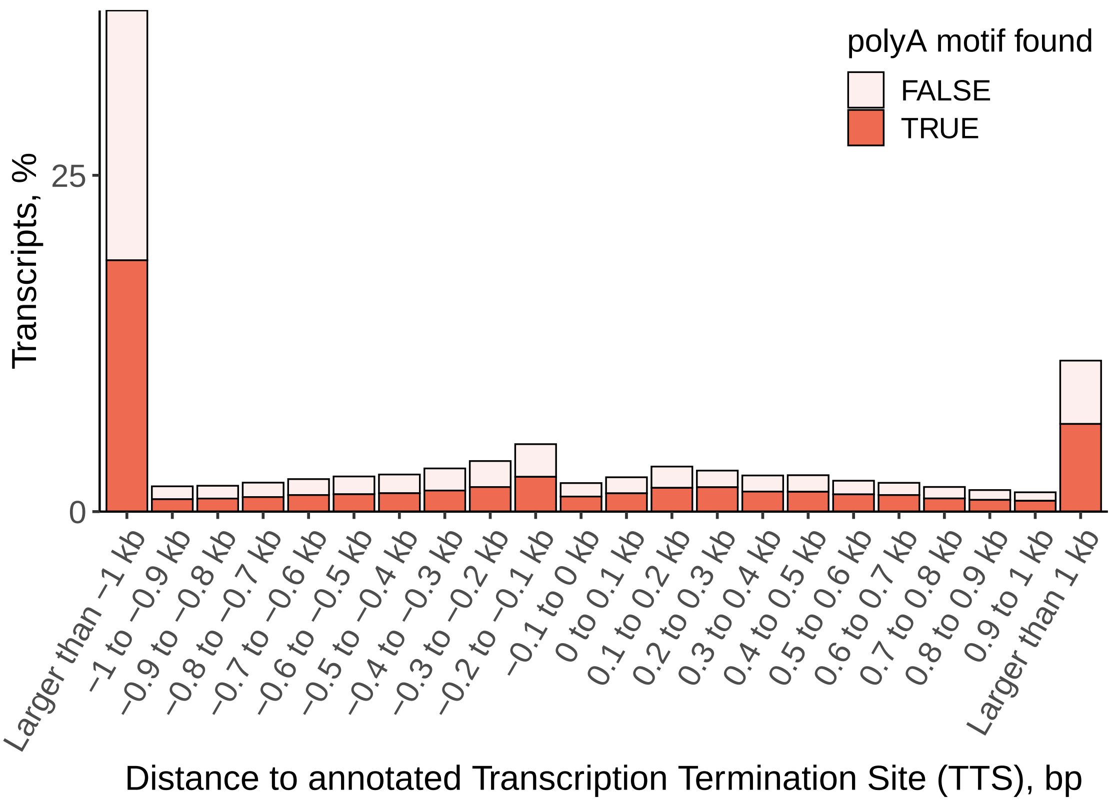
Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

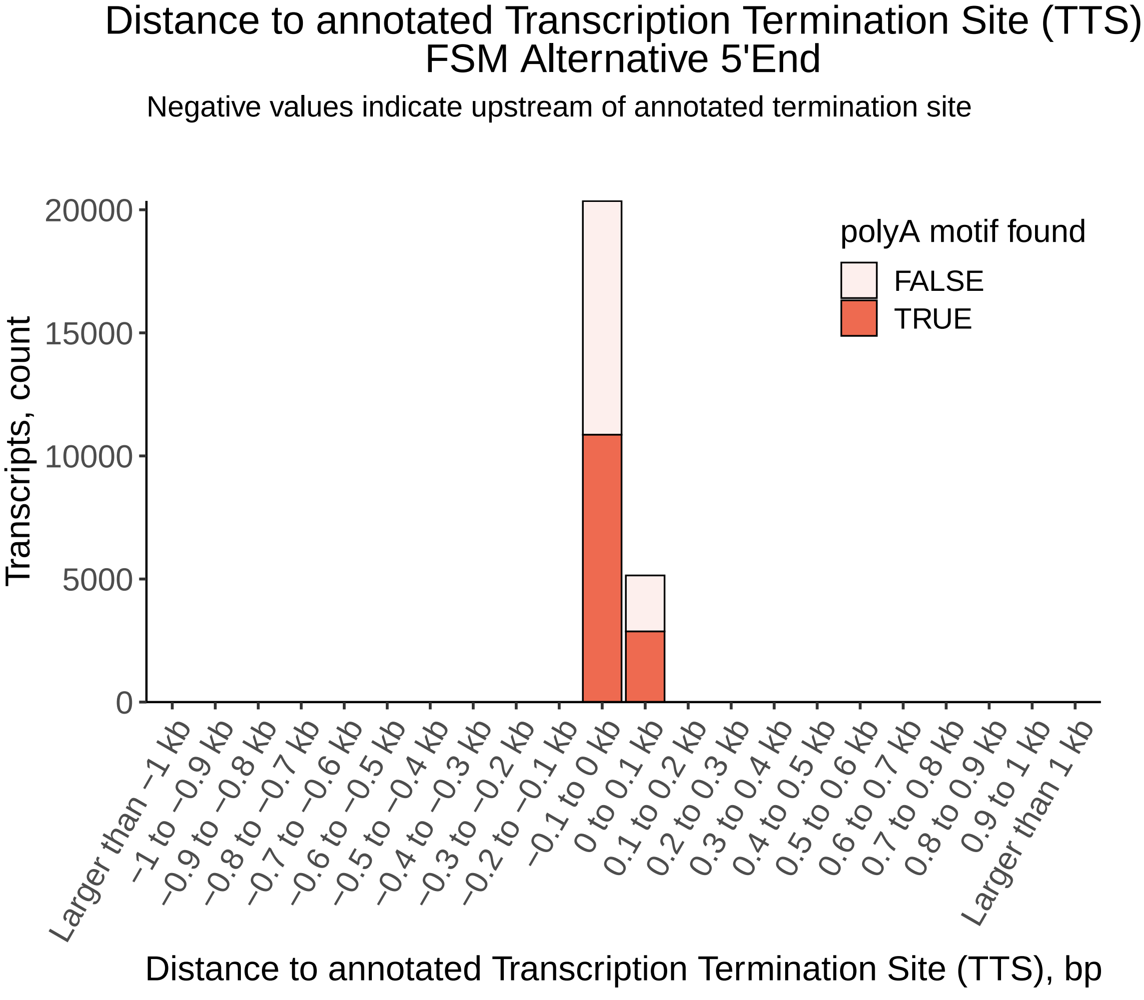
Negative values indicate upstream of annotated termination site

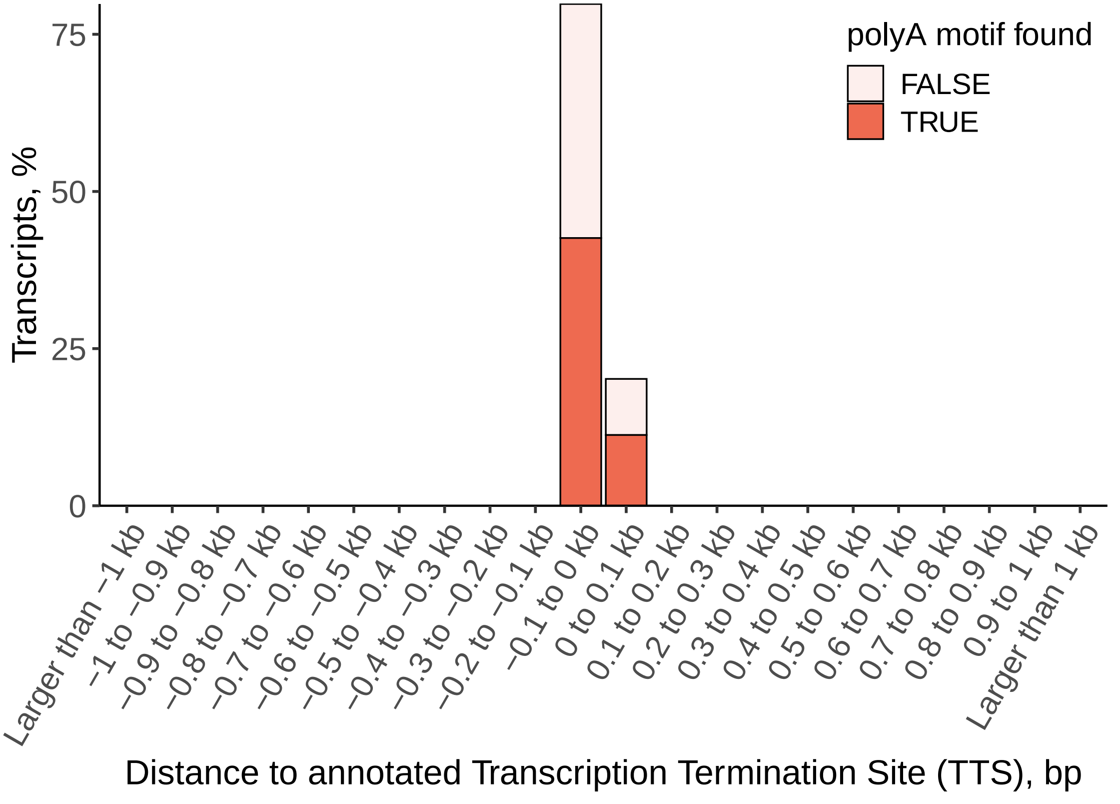


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

Negative values indicate upstream of annotated termination site

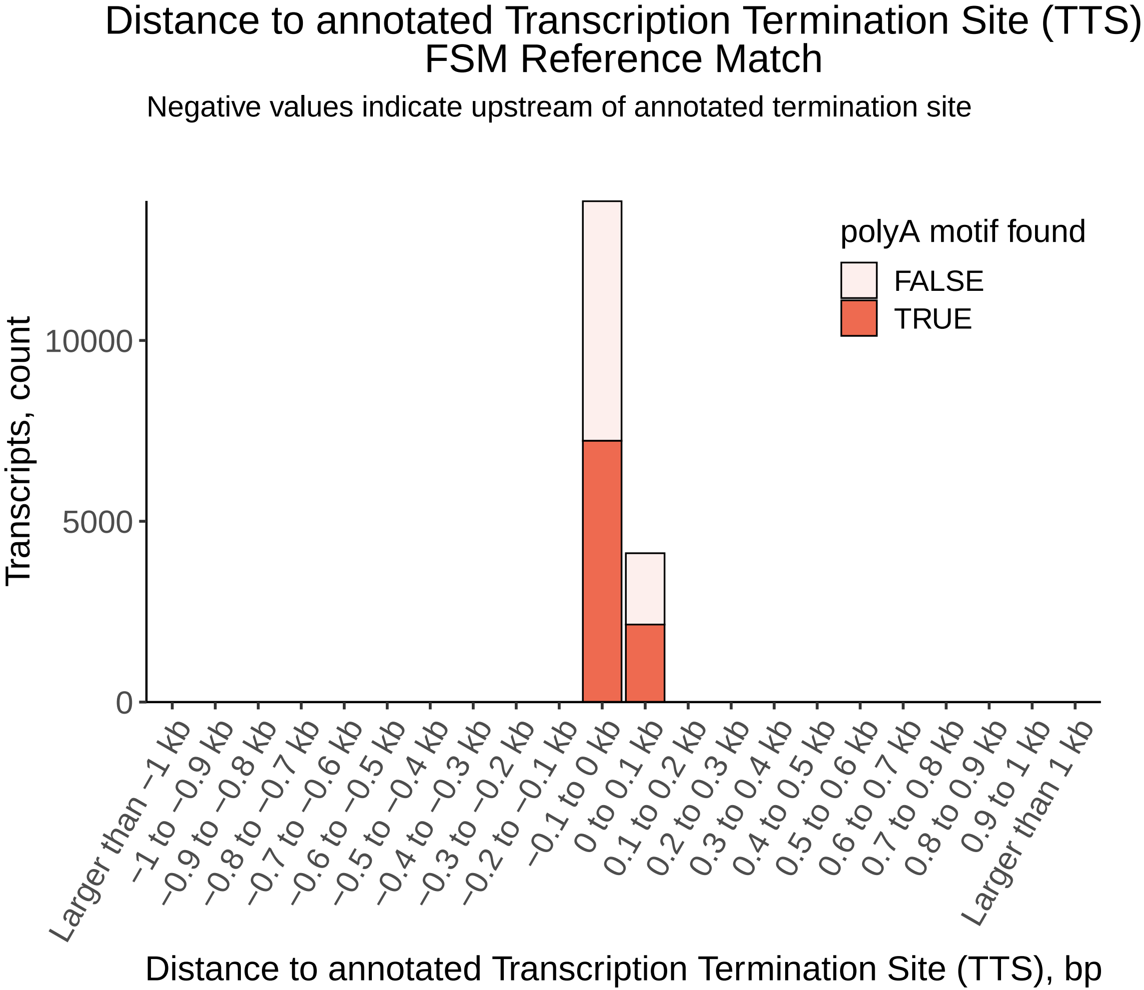


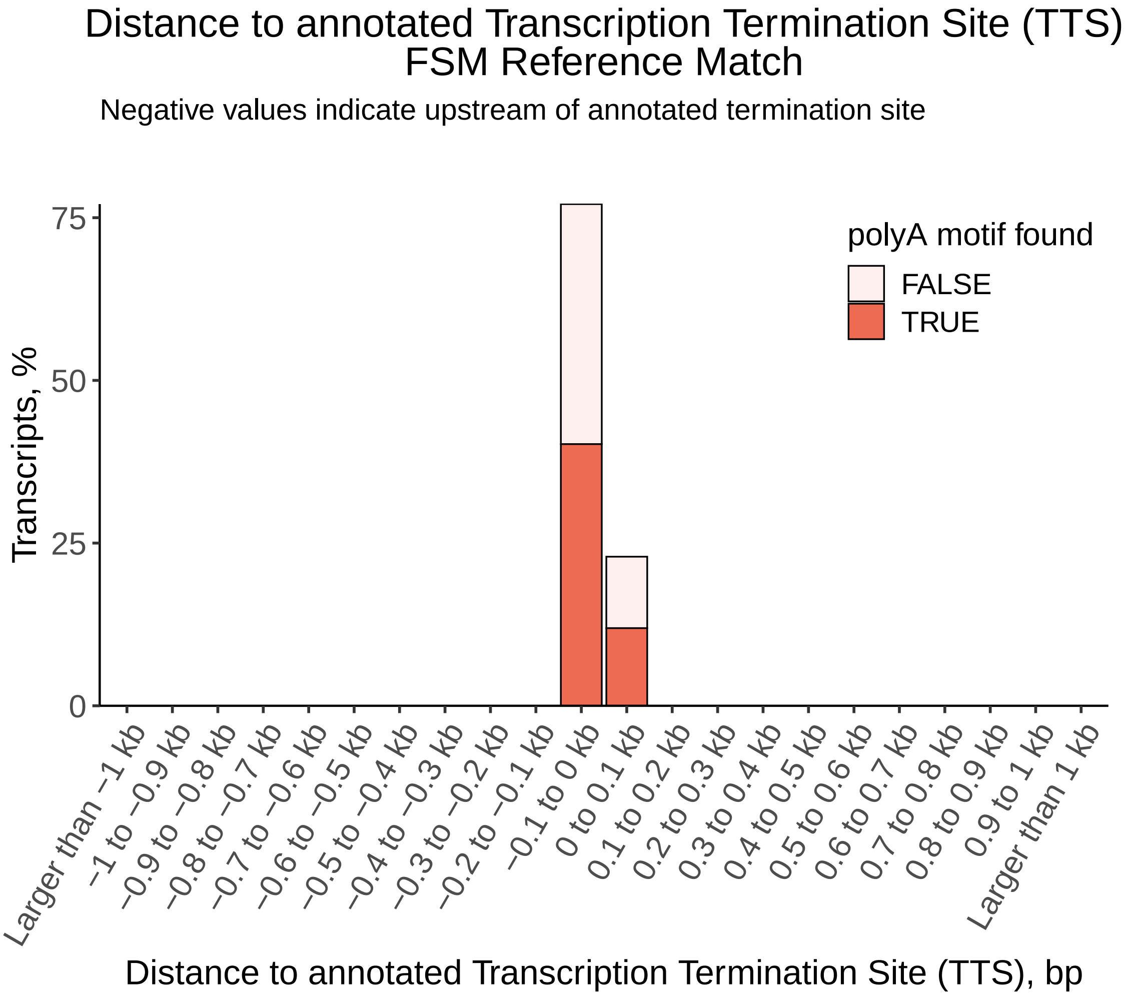

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
Transcripts, count
0
5000
10000
15000
20000
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Termination Site (TTS), bp

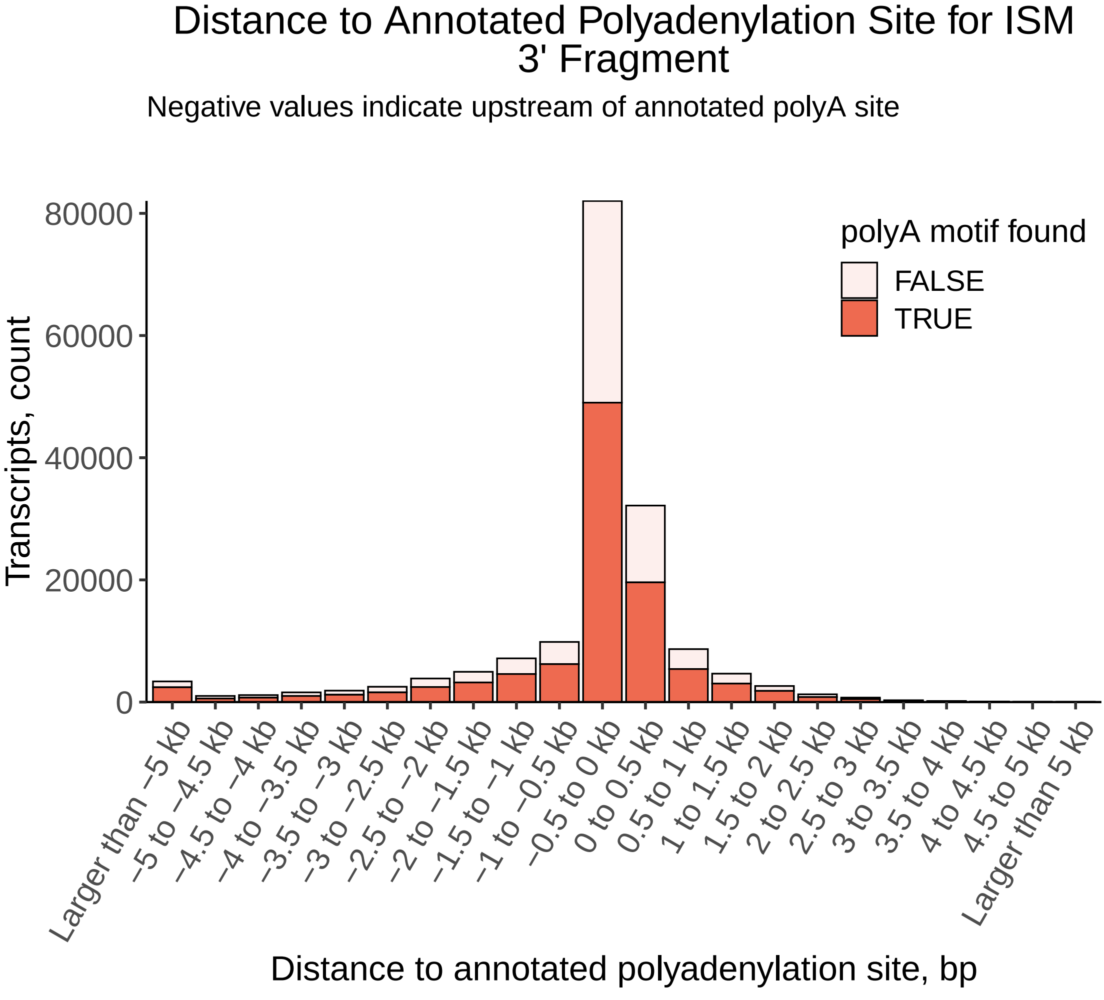


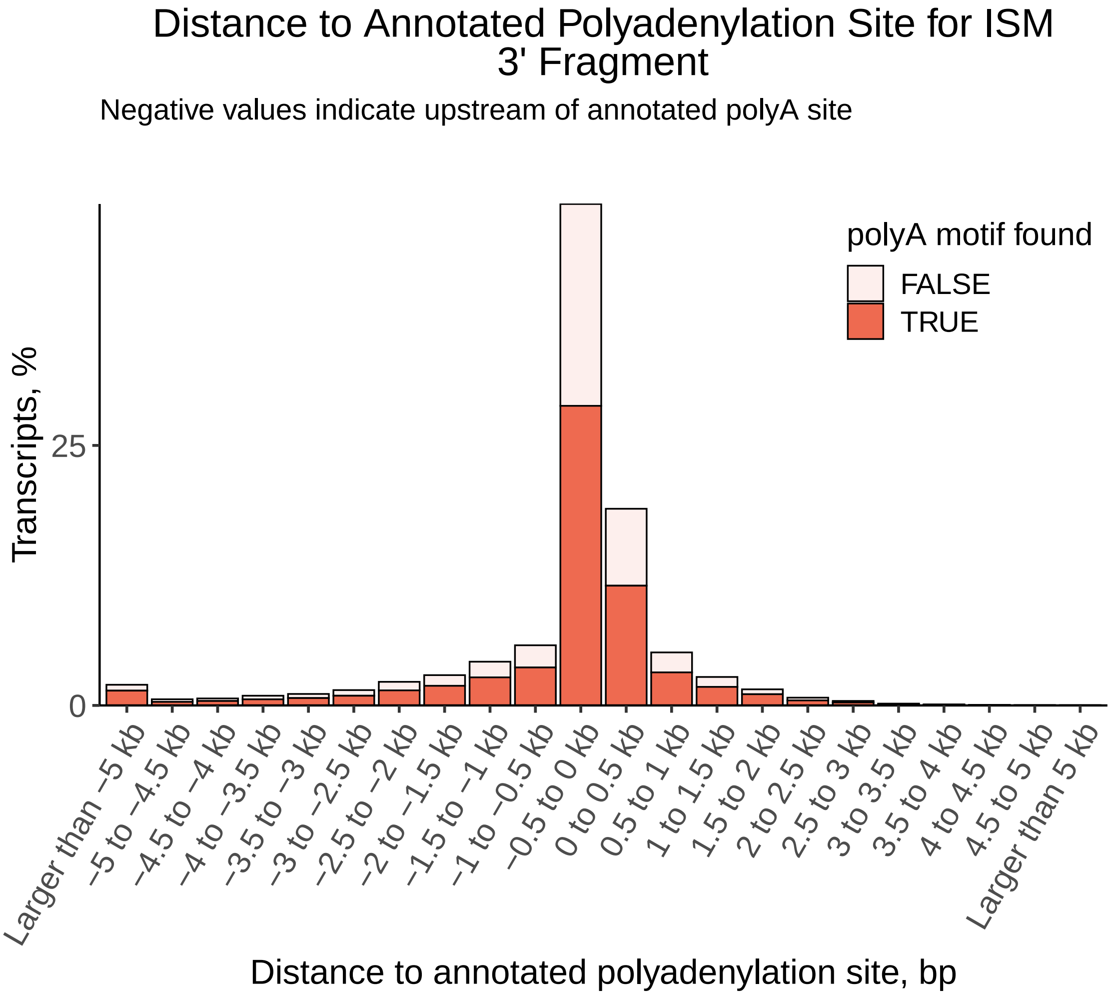
Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End

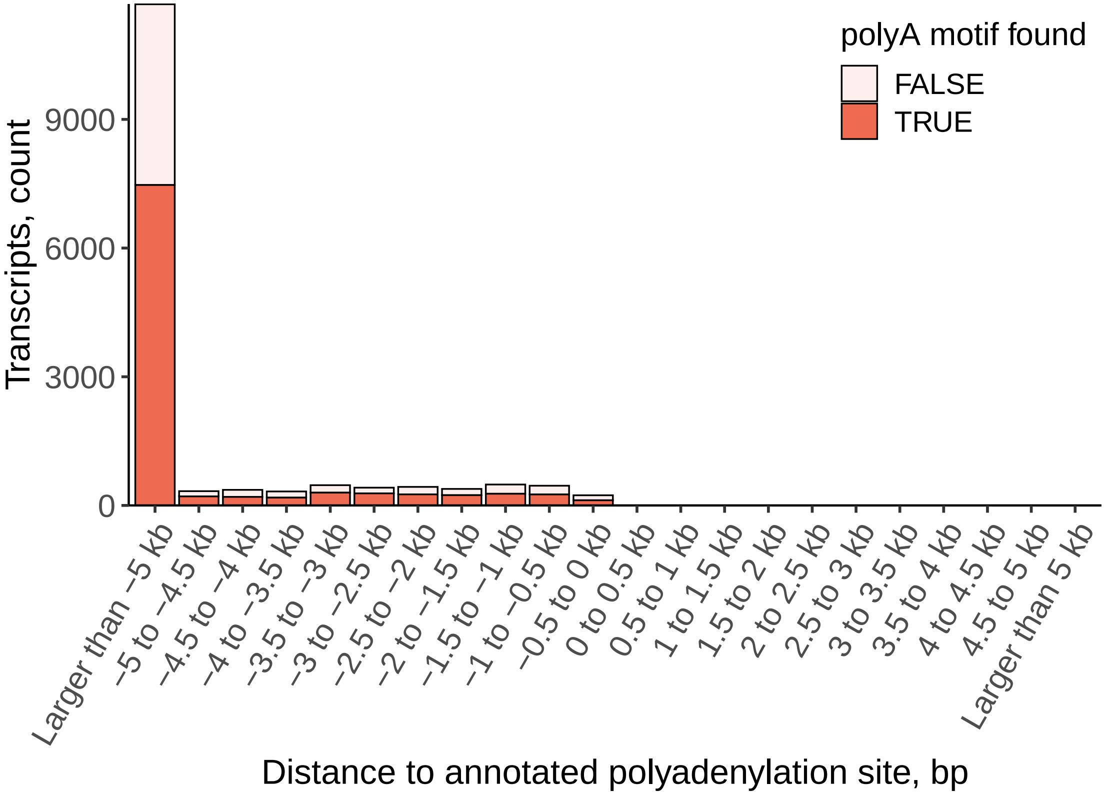
Negative values indicate upstream of annotated termination site


Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, count
10000
5000
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Termination Site (TTS), bp


Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, %
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to annotated Transcription Termination Site (TTS), bp

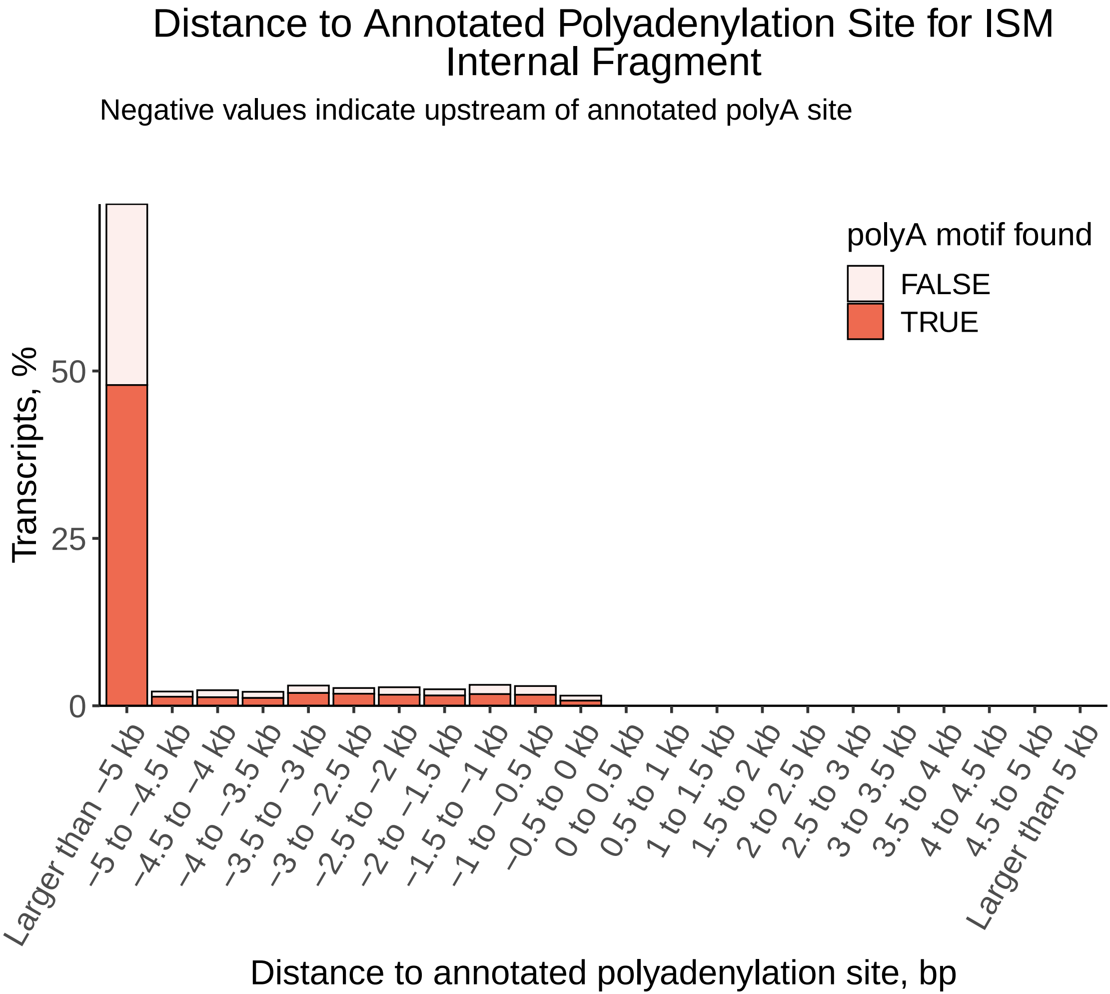

Distance to Annotated Polyadenylation Site for ISM 3' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
80000
60000
40000
20000
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp


Distance to Annotated Polyadenylation Site for ISM 3' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp



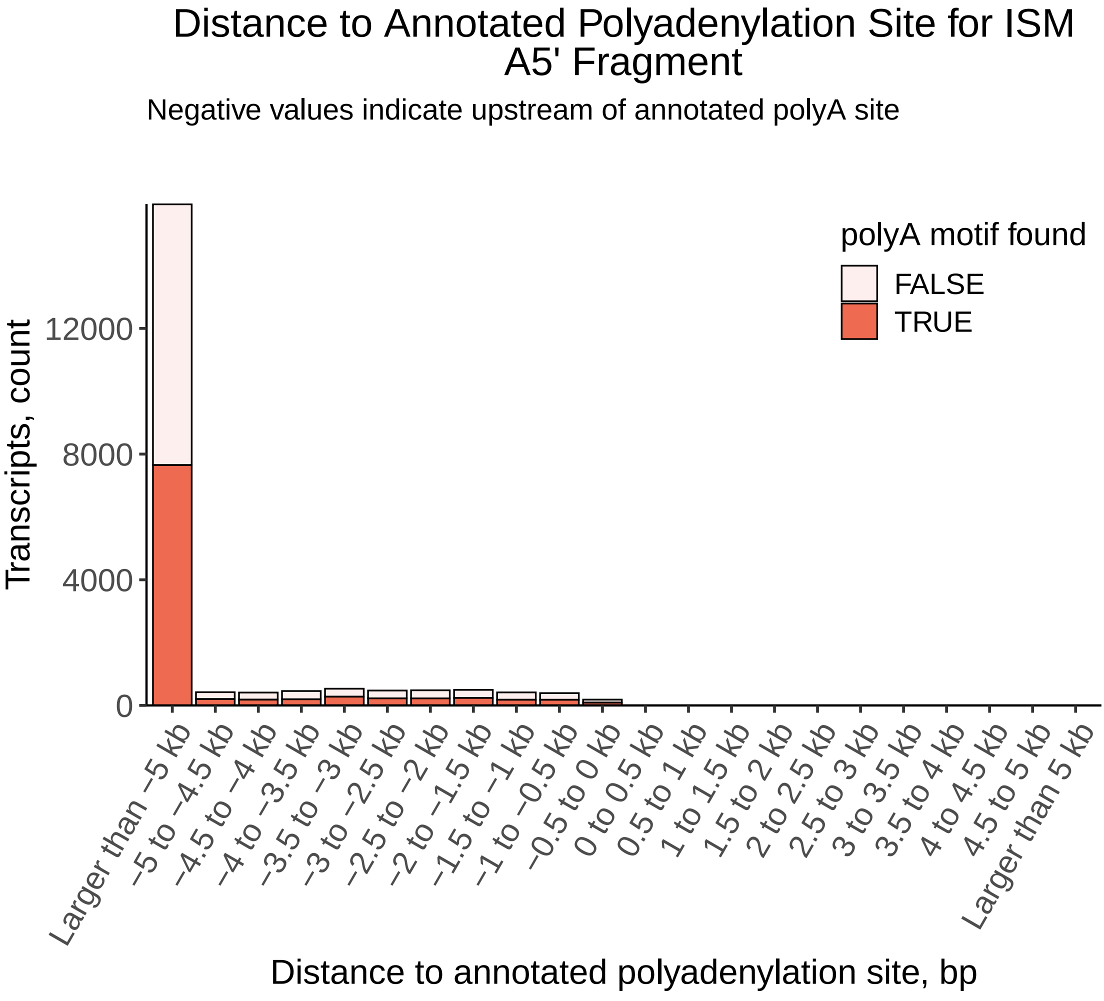
Distance to Annotated Polyadenylation Site for ISM Internal Fragment

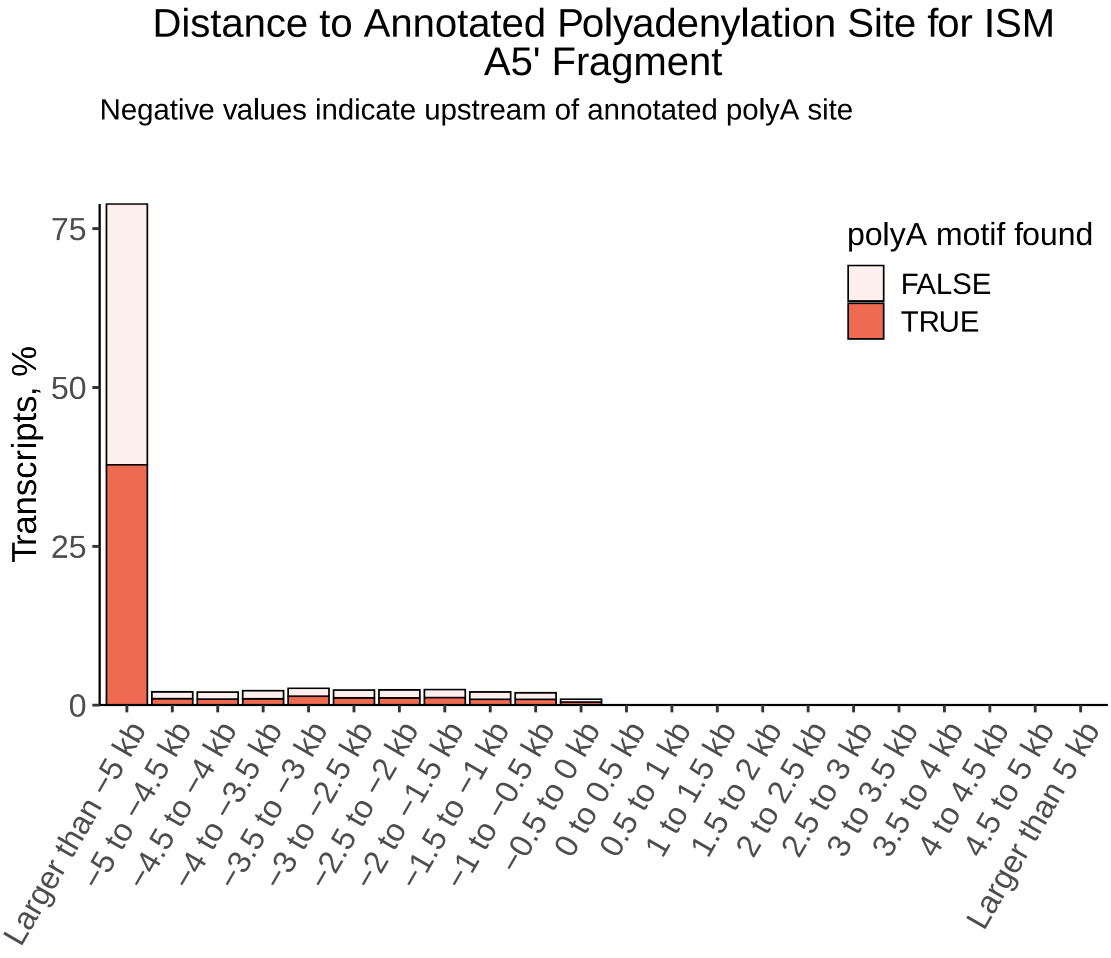
Negative values indicate upstream of annotated polyA site



Distance to Annotated Polyadenylation Site for ISM Internal Fragment

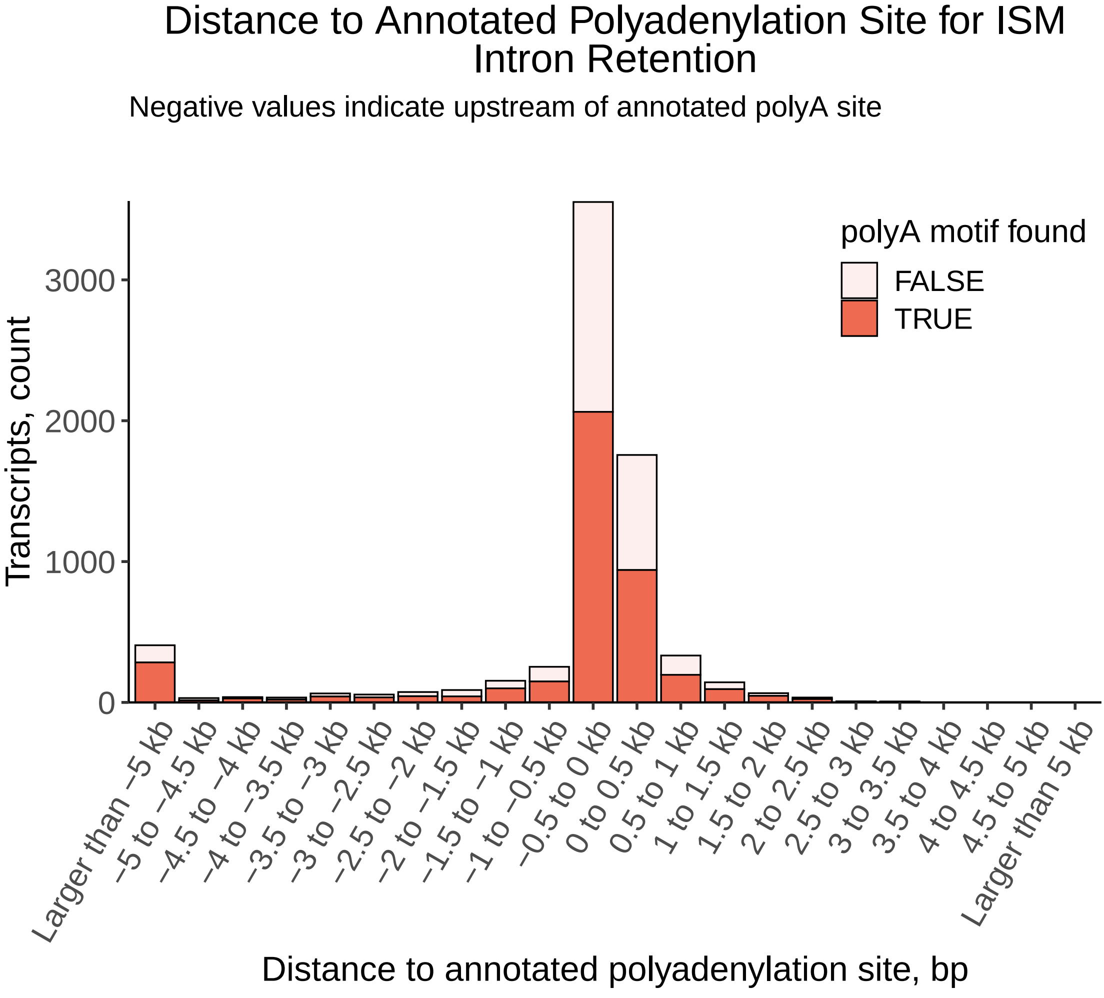
Negative values indicate upstream of annotated polyA site

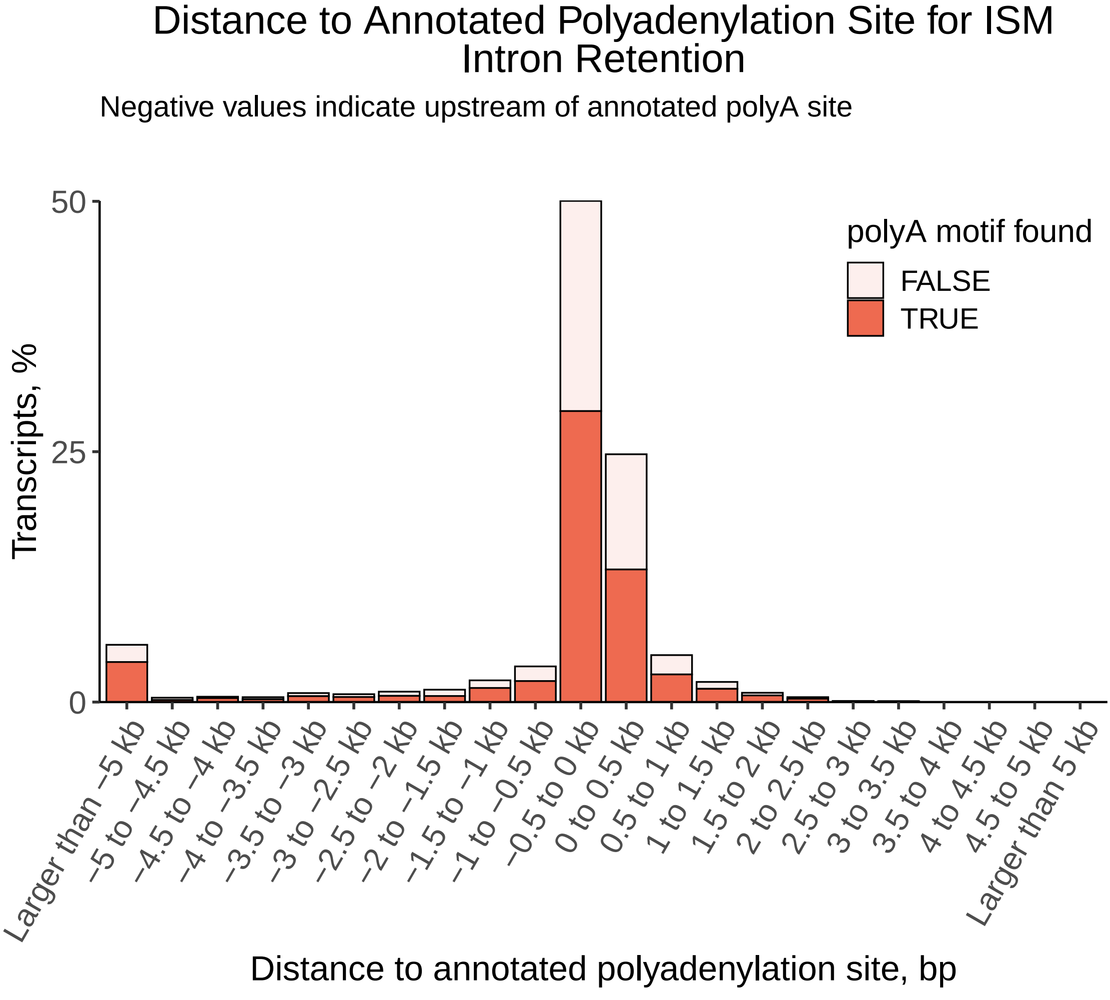

Distance to Annotated Polyadenylation Site for ISM A5' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
0
4000
8000
12000
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp



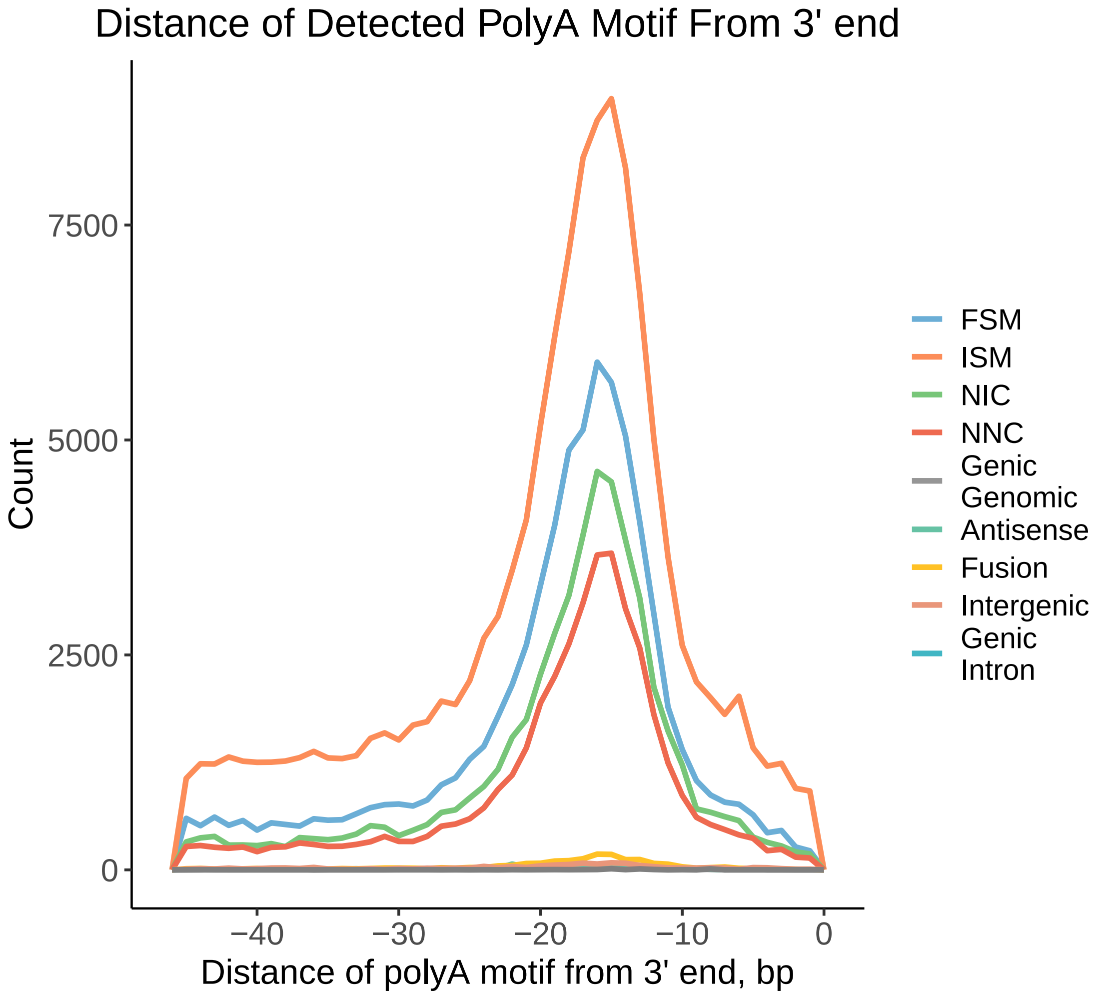
Distance to Annotated Polyadenylation Site for ISM A5' Fragment

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
Transcripts, count
3000
2000
1000
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp


Distance to Annotated Polyadenylation Site for ISM Intron Retention
Negative values indicate upstream of annotated polyA site
Transcripts, %
50
25
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp

PolyA Distance Analysis


Distance of Detected PolyA Motif From 3' end
Count
7500
5000
2500
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

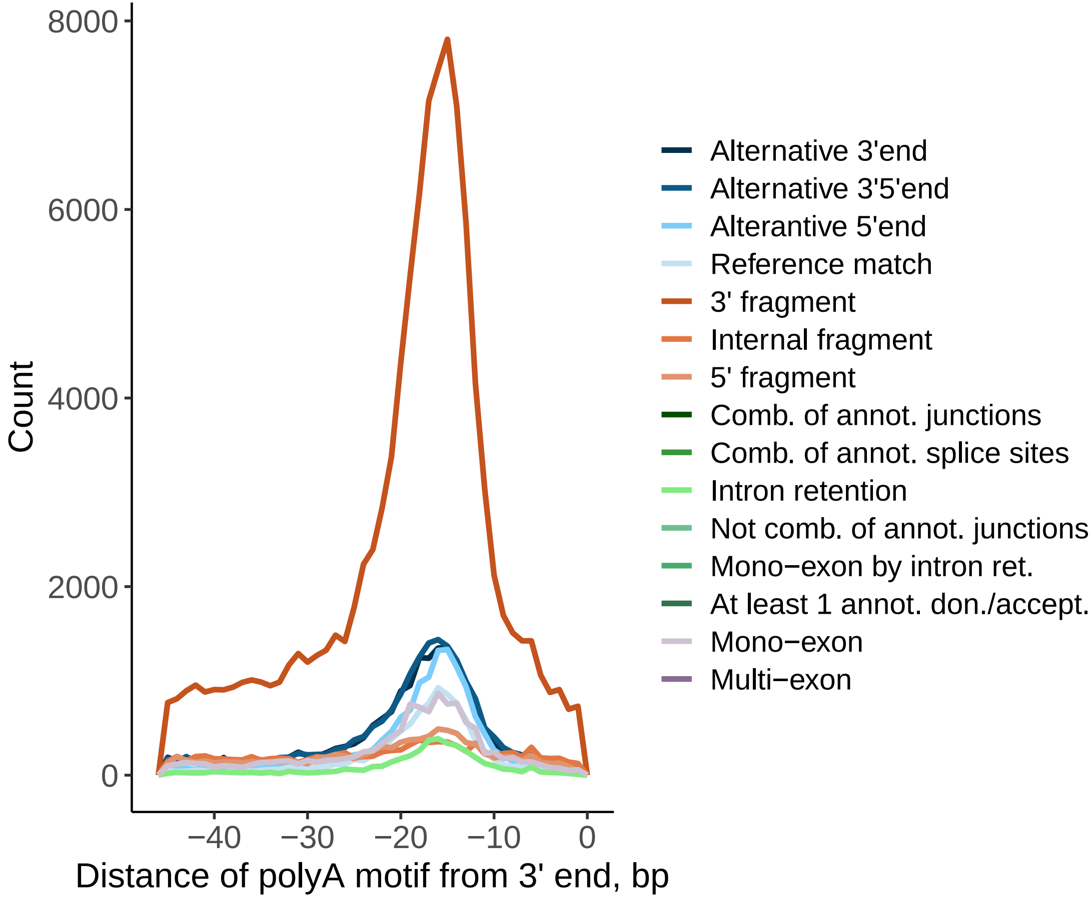
Number of polyA Motifs Detected

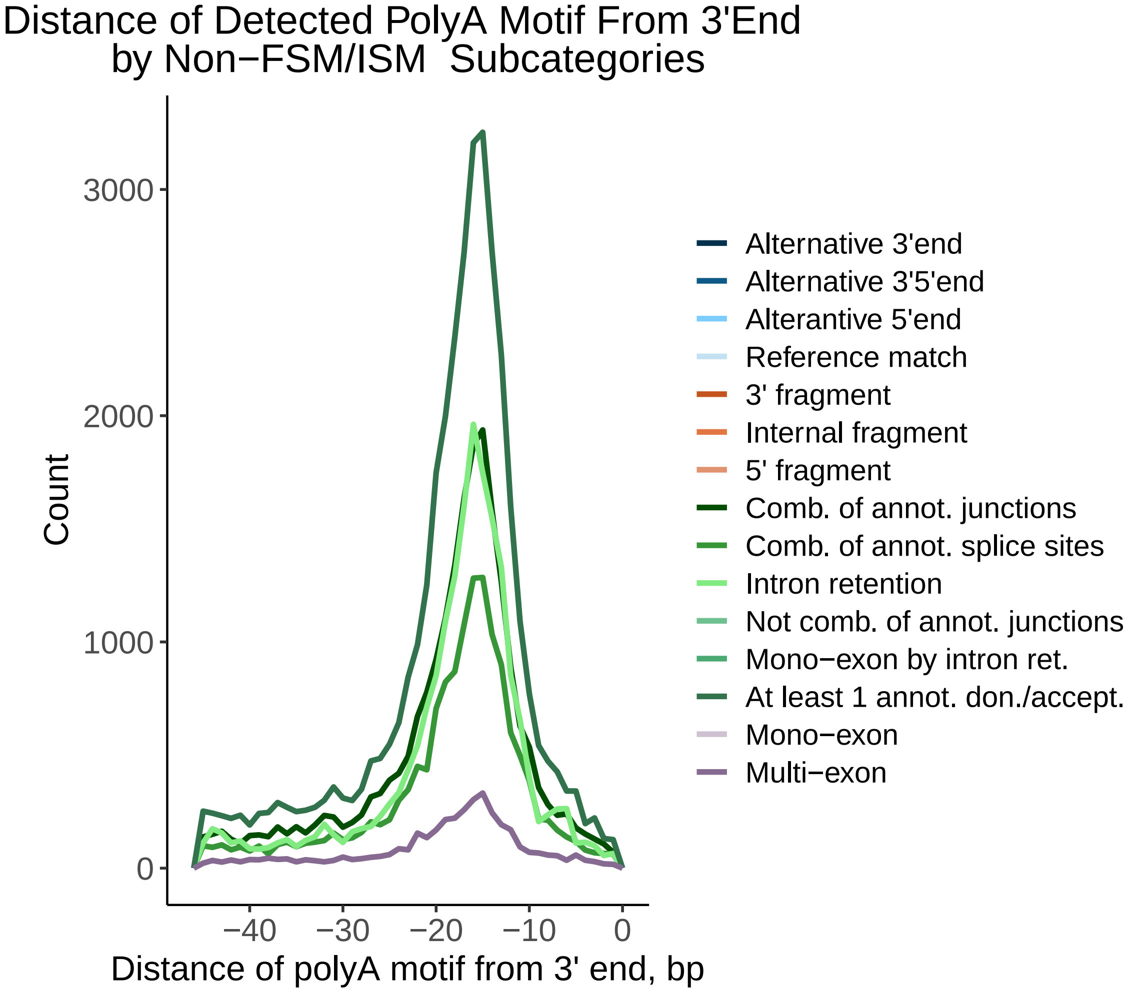
Category	Count	polyA Detected	%
FSM	137383	72142	53
ISM	213048	128212	60
NIC	99744	51581	52
NNC	76762	41039	53
Genic Genomic	653	346	53
Antisense	1493	820	55
Fusion	3337	1844	55
Intergenic	1874	1195	64
NA	89	67	75

Frequency of PolyA Motifs

Motif	Count	%
AATAAA	146485	49.3
ATTAAA	42618	14.3
TATAAA	12561	4.2
AAGAAA	11070	3.7
AAAAAG	10997	3.7
AGTAAA	10988	3.7
GGGGCT	8783	3.0
TTTAAA	8281	2.8
AATATA	6904	2.3
CATAAA	6649	2.2
GATAAA	6588	2.2
AATACA	6480	2.2
AAAACA	6407	2.2
AATGAA	5565	1.9
AATAGA	3544	1.2
ACTAAA	3326	1.1

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




Distance of Detected PolyA Motif From 3'End
by Non−FSM/ISM Subcategories
Count
3000
2000
1000
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

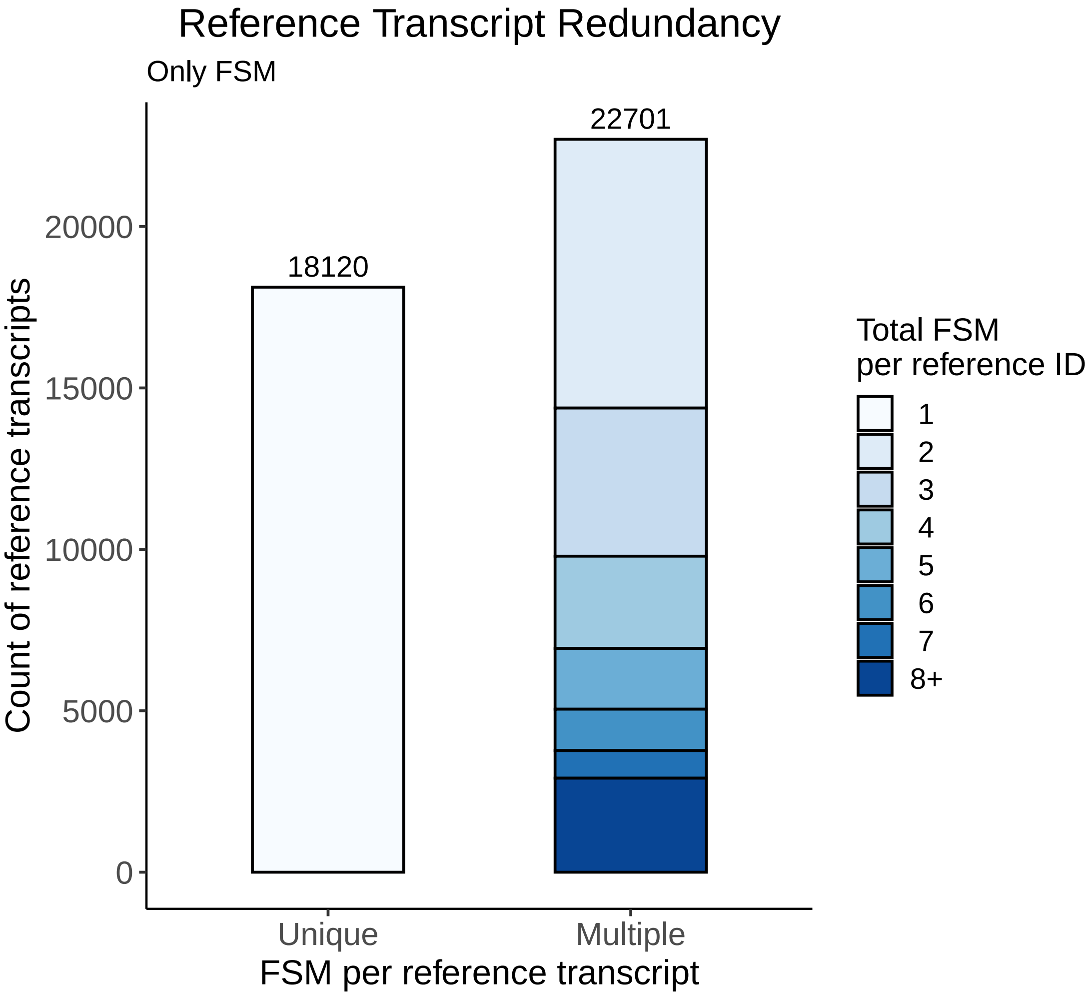
Number of polyA Motifs Detected

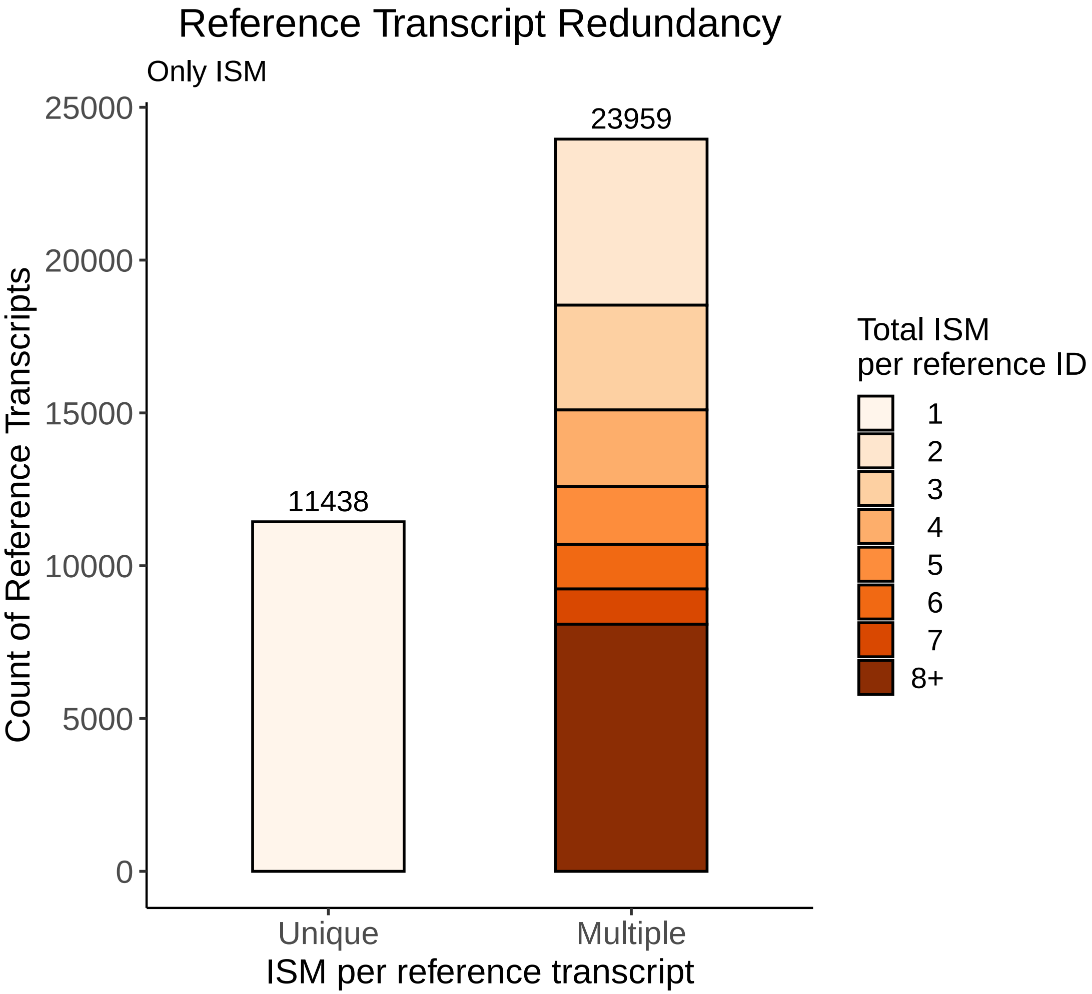
Subcategory	Count	polyA Detected	%
Alternative 3'end	37784	18217	48
Alternative 3'5'end	36904	19279	52
Alterantive 5'end	25499	13729	54
Reference match	17969	9369	52
3' fragment	170134	104641	62
Internal fragment	15594	9786	63
5' fragment	20220	9649	48
Comb. of annot. junctions	42858	21910	51
Comb. of annot. splice sites	27700	14666	53
Intron retention	45454	23934	53
At least 1 annot. don./accept.	68164	36571	54
Mono−exon	19227	11548	60
Multi−exon	6804	3894	57
NA	72	53	74

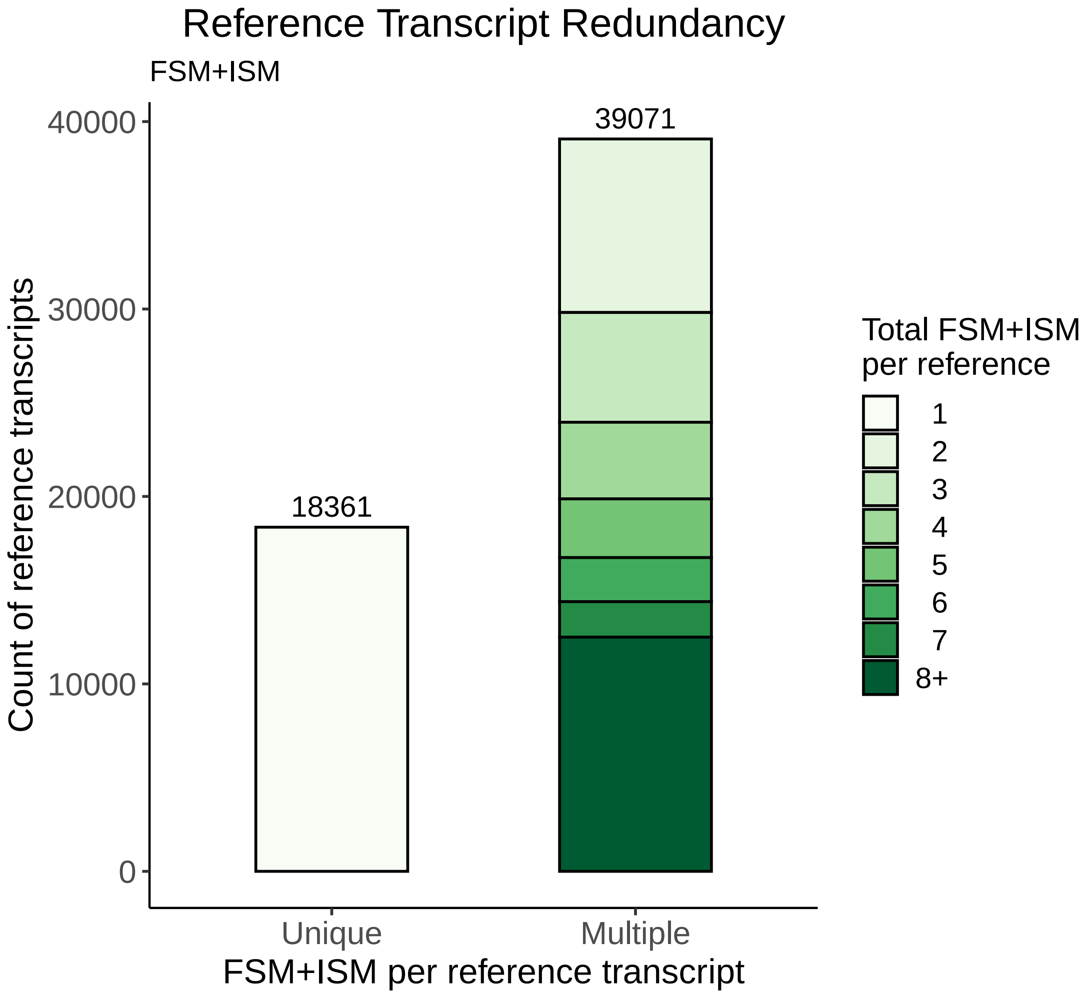
Frequency of PolyA Motifs

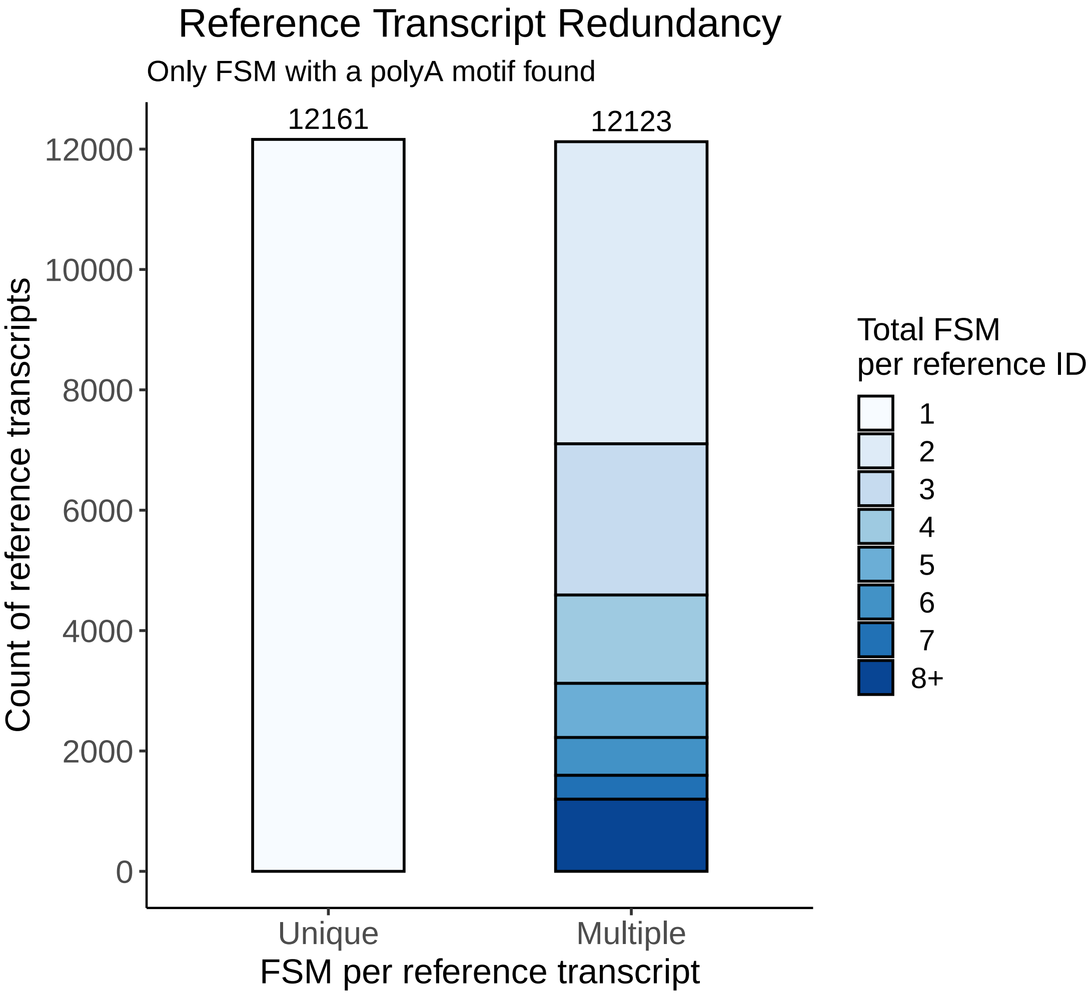
Motif	Count	%
AATAAA	146485	49.3
ATTAAA	42618	14.3
TATAAA	12561	4.2
AAGAAA	11070	3.7
AAAAAG	10997	3.7
AGTAAA	10988	3.7
GGGGCT	8783	3.0
TTTAAA	8281	2.8
AATATA	6904	2.3
CATAAA	6649	2.2
GATAAA	6588	2.2
AATACA	6480	2.2
AAAACA	6407	2.2
AATGAA	5565	1.9
AATAGA	3544	1.2
ACTAAA	3326	1.1

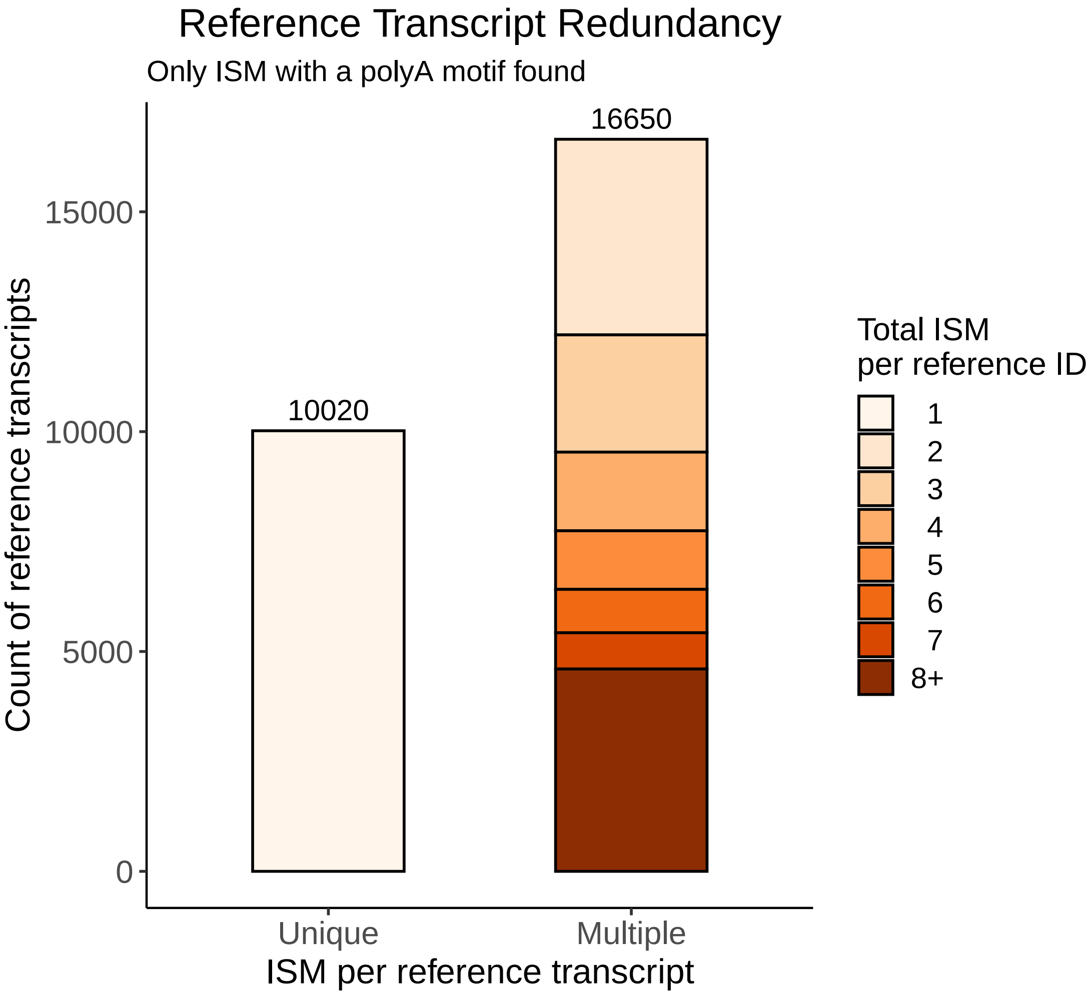
Redundancy Analysis

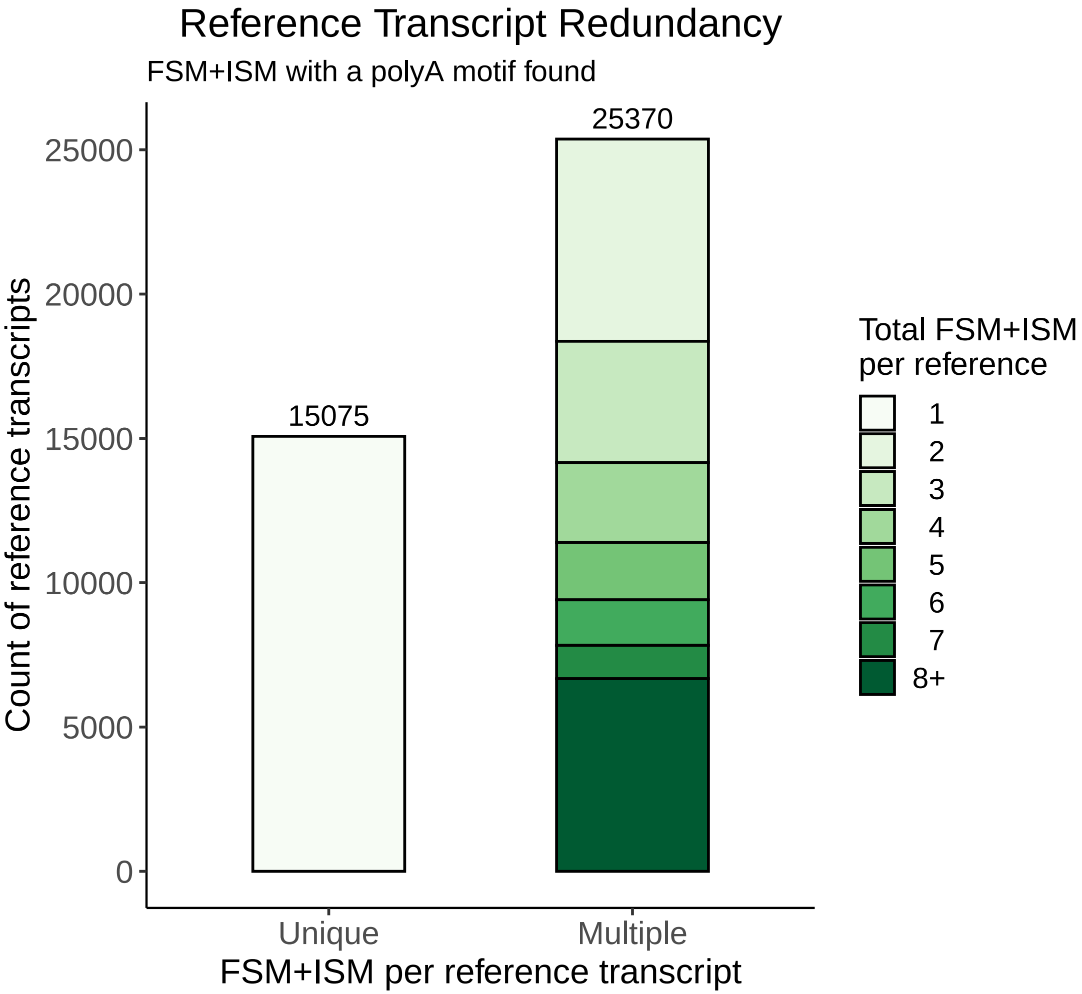

Reference Transcript Redundancy
Only FSM
Count of reference transcripts
0
5000
10000
15000
20000
18120
22701
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only ISM
Count of Reference Transcripts
25000
20000
15000
10000
5000
0
11438
23959
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
FSM+ISM
Count of reference transcripts
40000
30000
20000
10000
0
18361
39071
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only FSM with a polyA motif found
12161
12123
12000
10000
8000
6000
4000
2000
0
Count of reference transcripts
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+

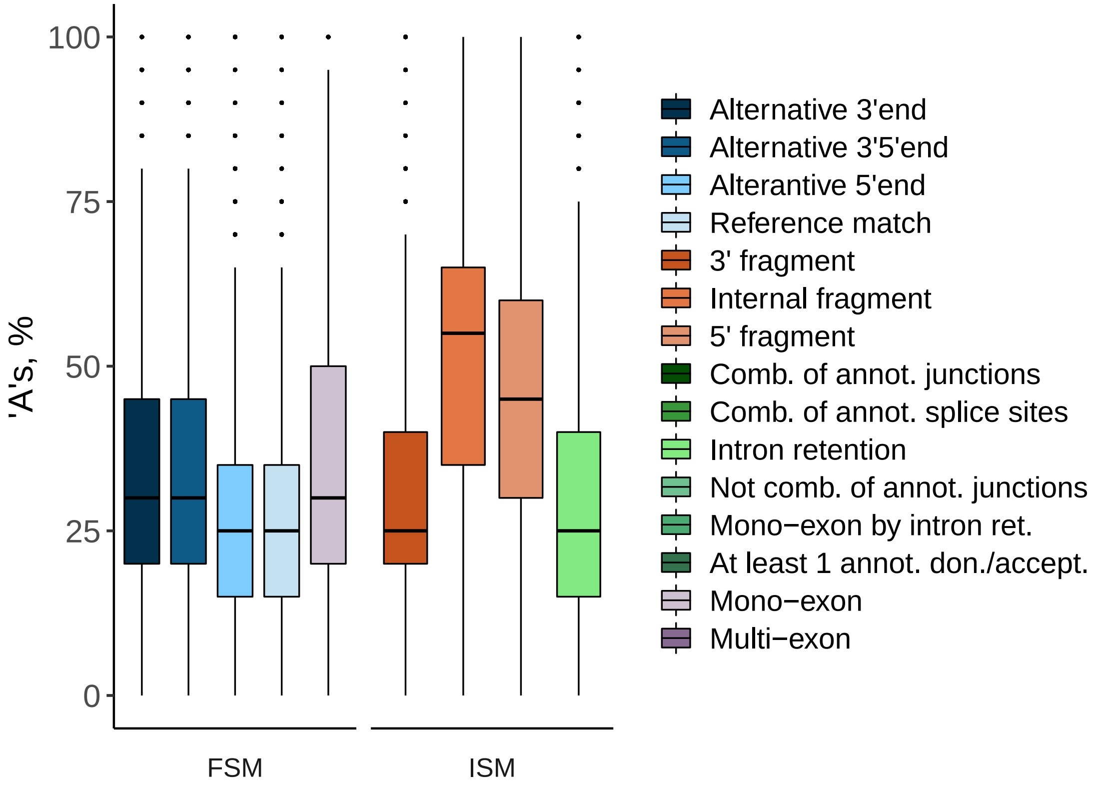

Reference Transcript Redundancy
Only ISM with a polyA motif found
Count of reference transcripts
0
5000
10000
15000
10020
16650
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
FSM+ISM with a polyA motif found
Count of reference transcripts
25000
20000
15000
10000
5000
0
15075
25370
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+

Intra−Priming Quality Check

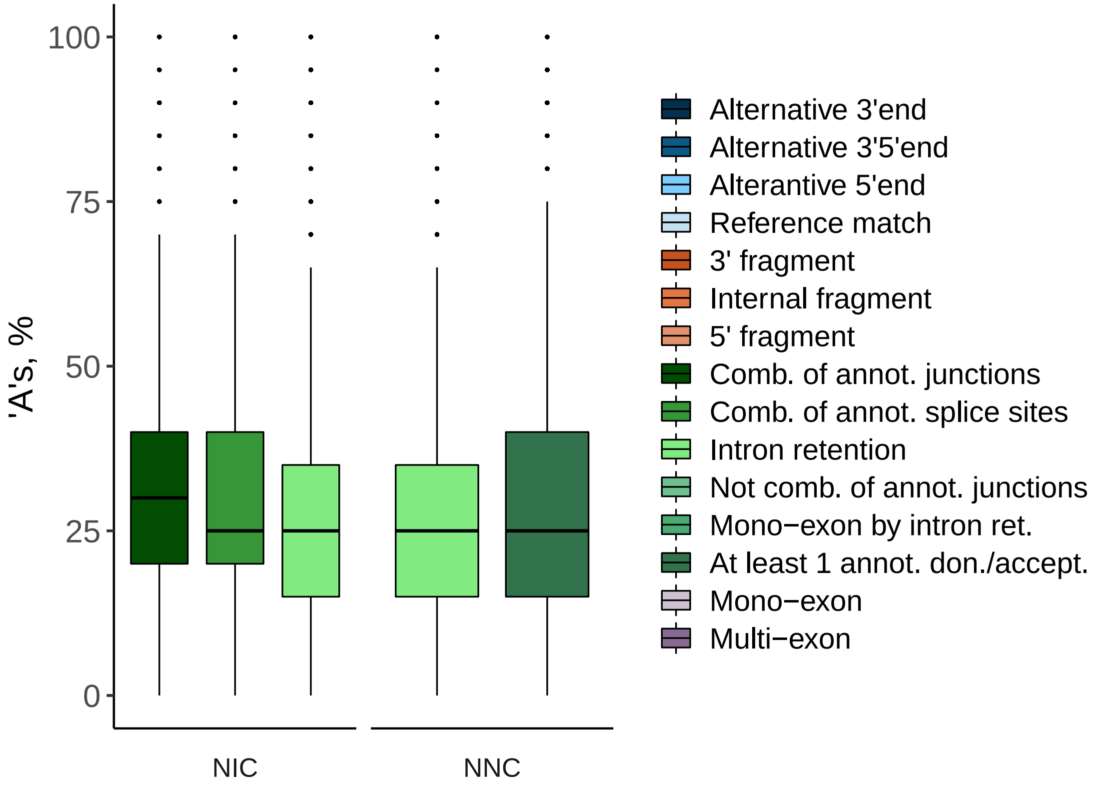
Possible Intra−Priming by Structural Category

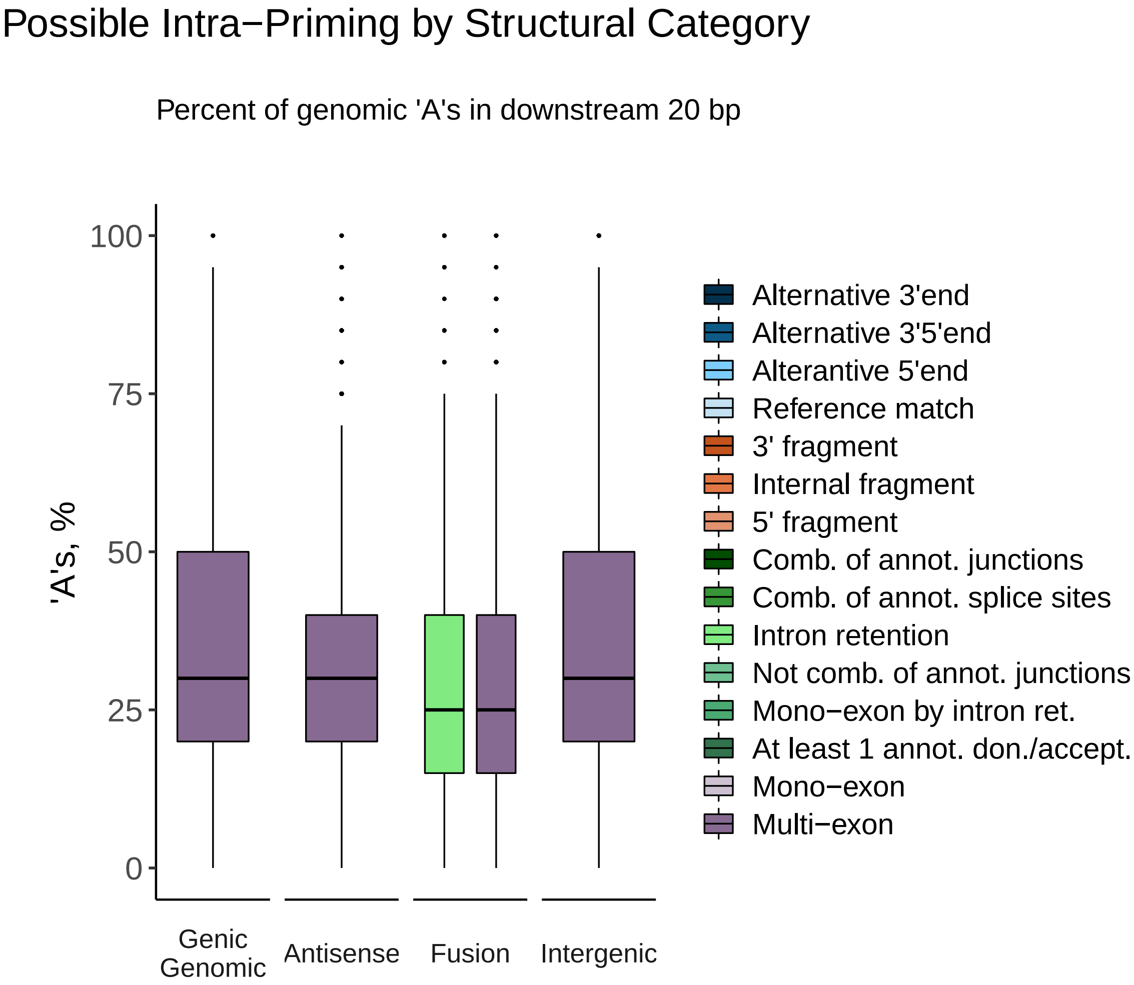
Percent of genomic 'A's in downstream 20 bp

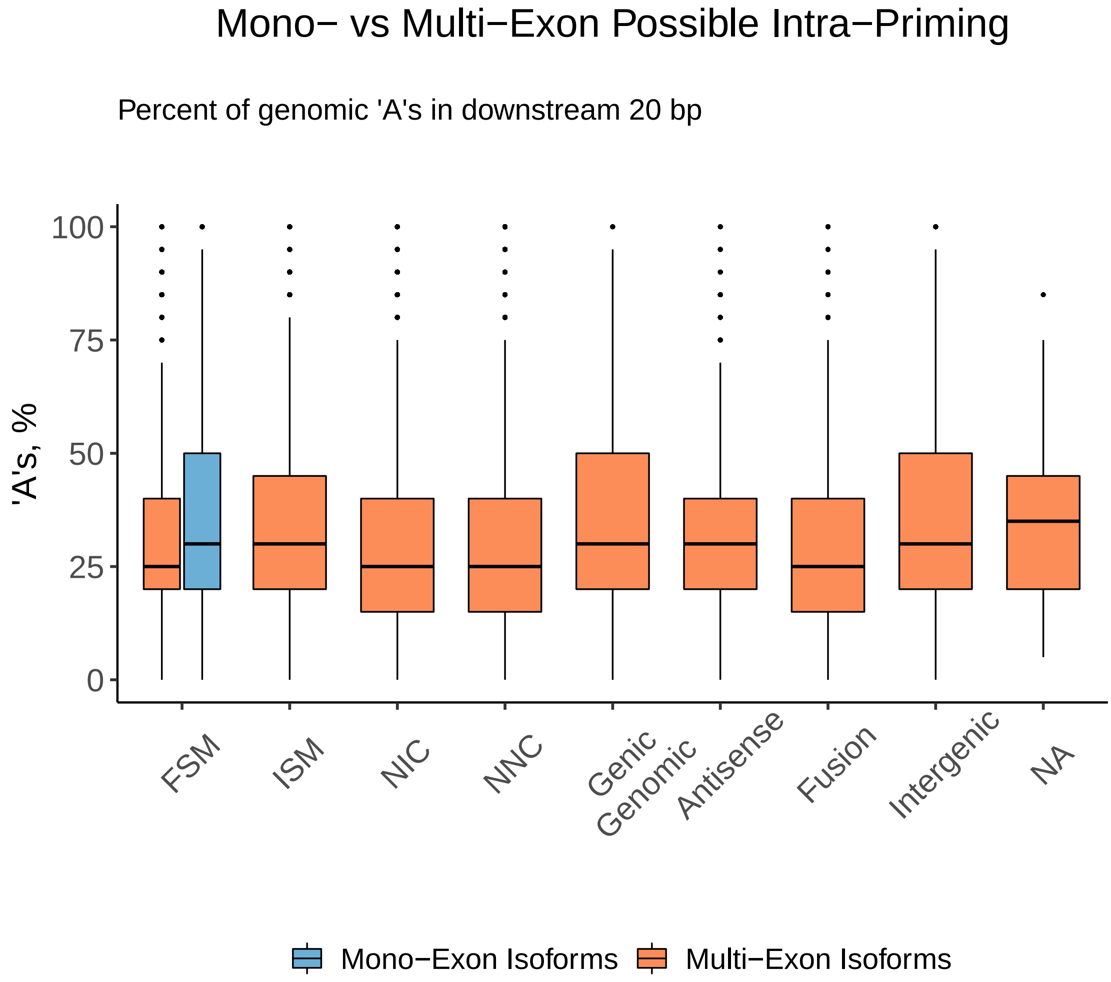


Possible Intra−Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp

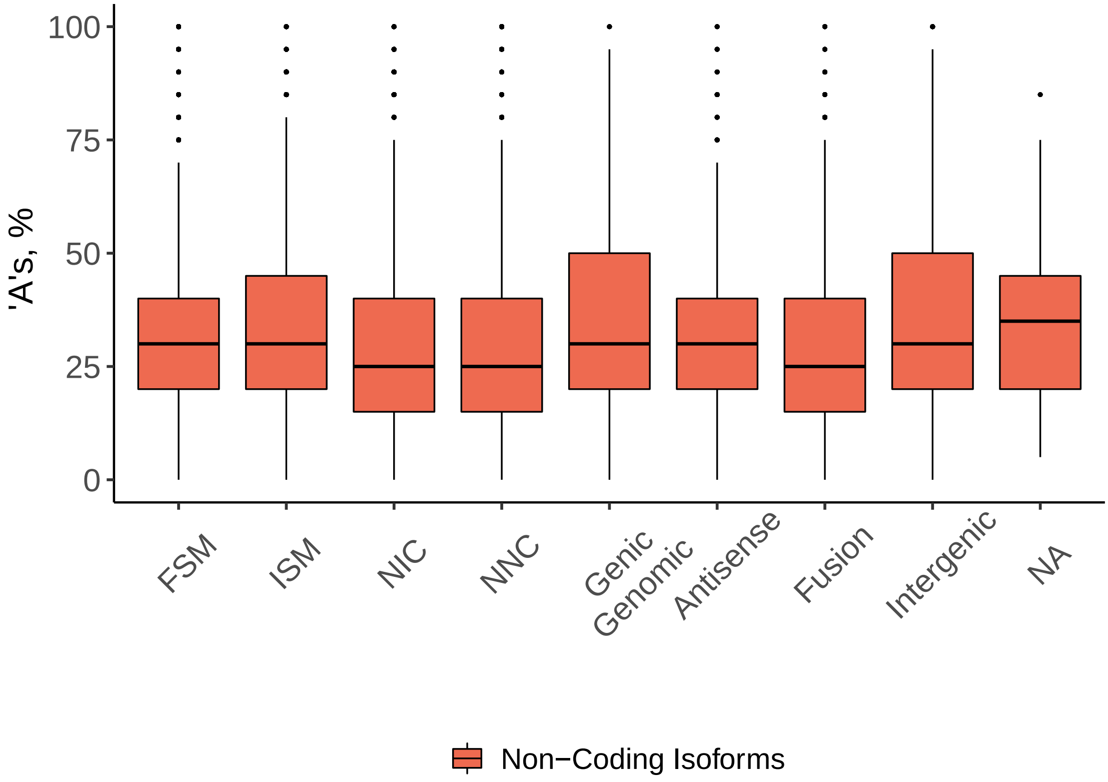



Possible Intra−Priming by Structural Category
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
Genic Genomic
Antisense
Fusion
Intergenic
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon



Mono− vs Multi−Exon Possible Intra−Priming

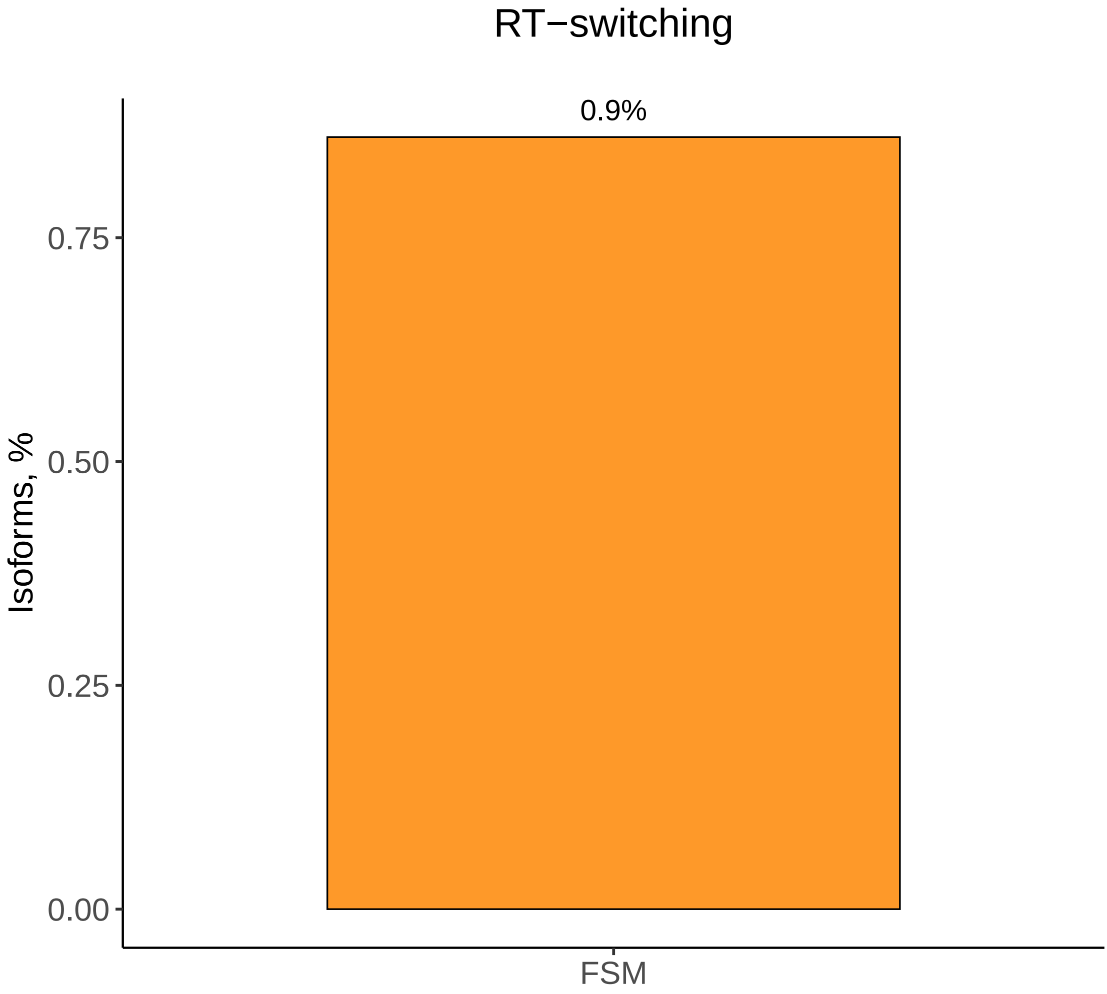
Percent of genomic 'A's in downstream 20 bp

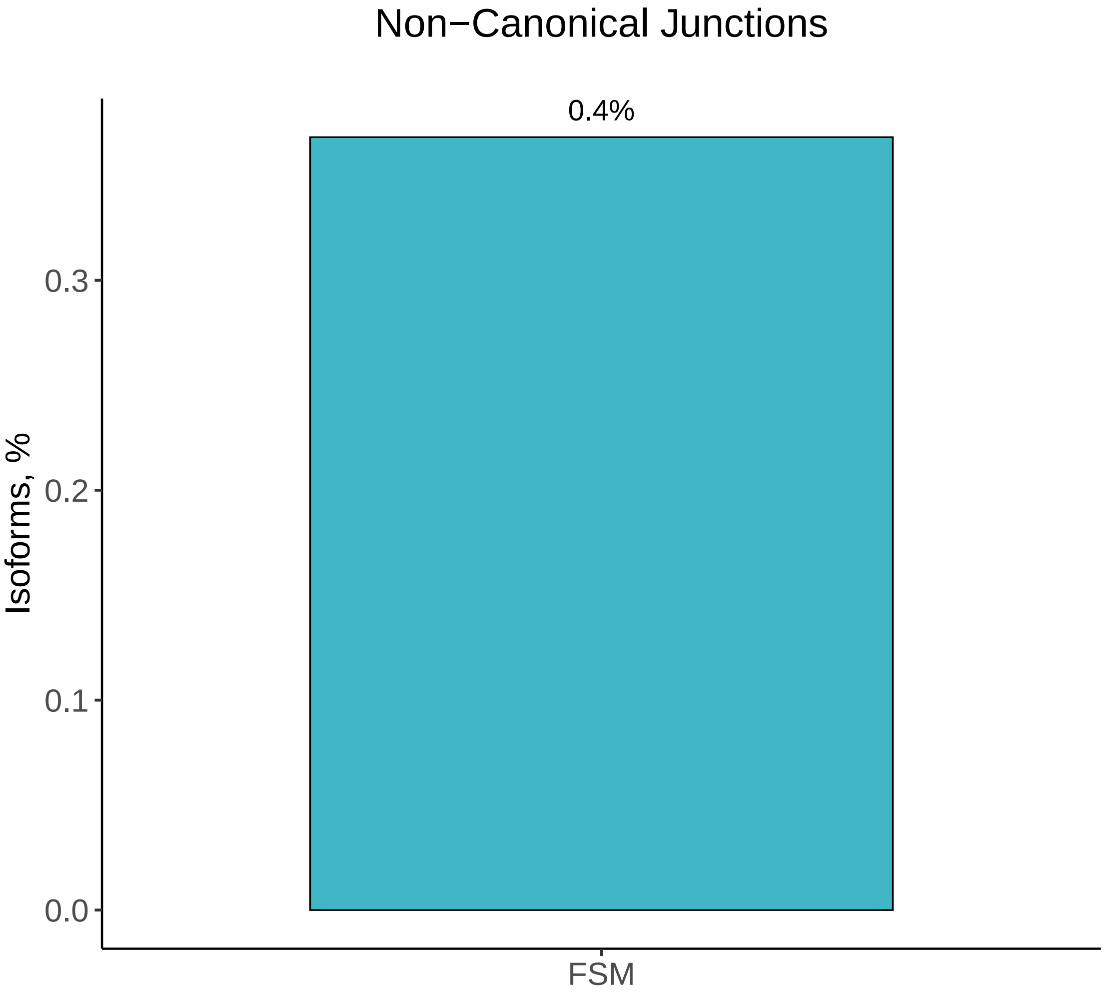


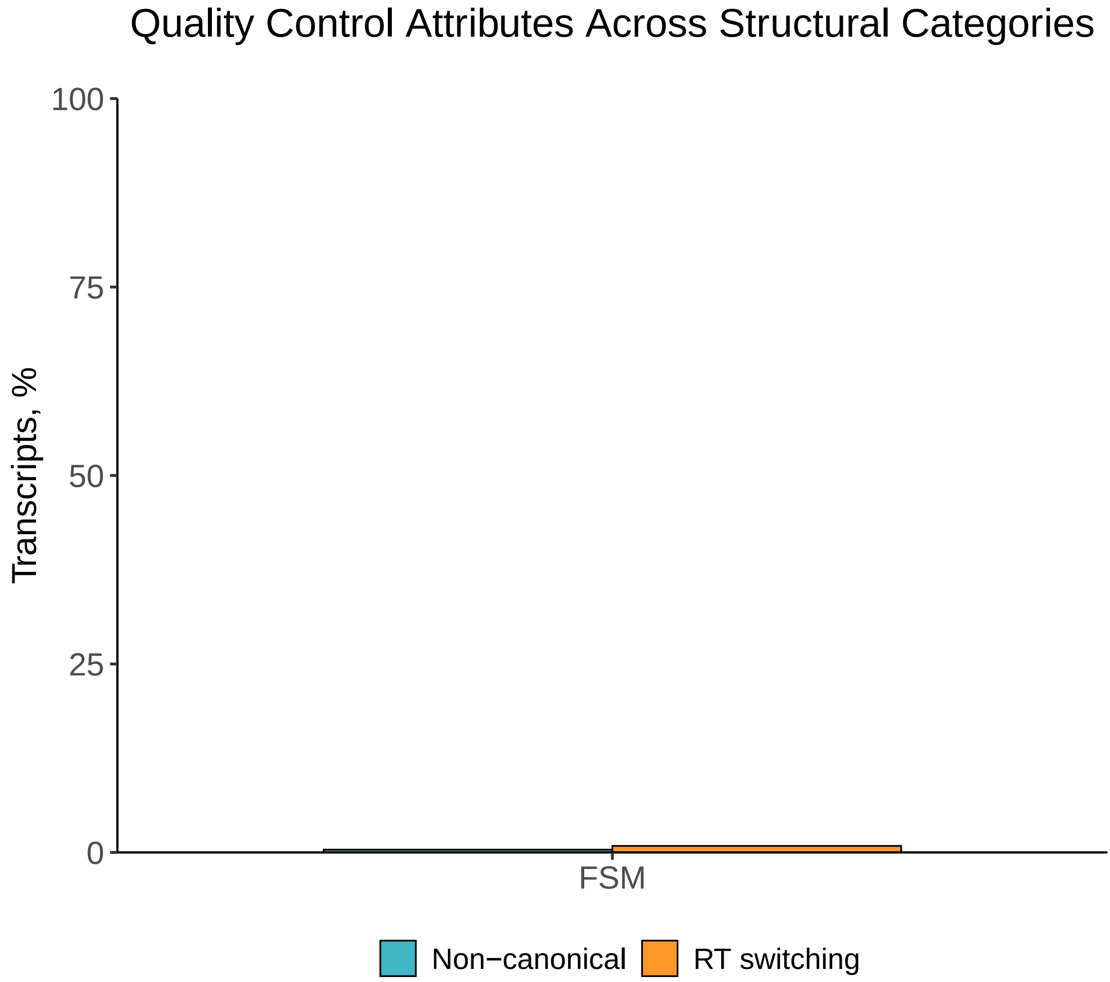
Coding vs Non−Coding Possible Intra−Priming

Percent of genomic 'A's in downstream 20 bp

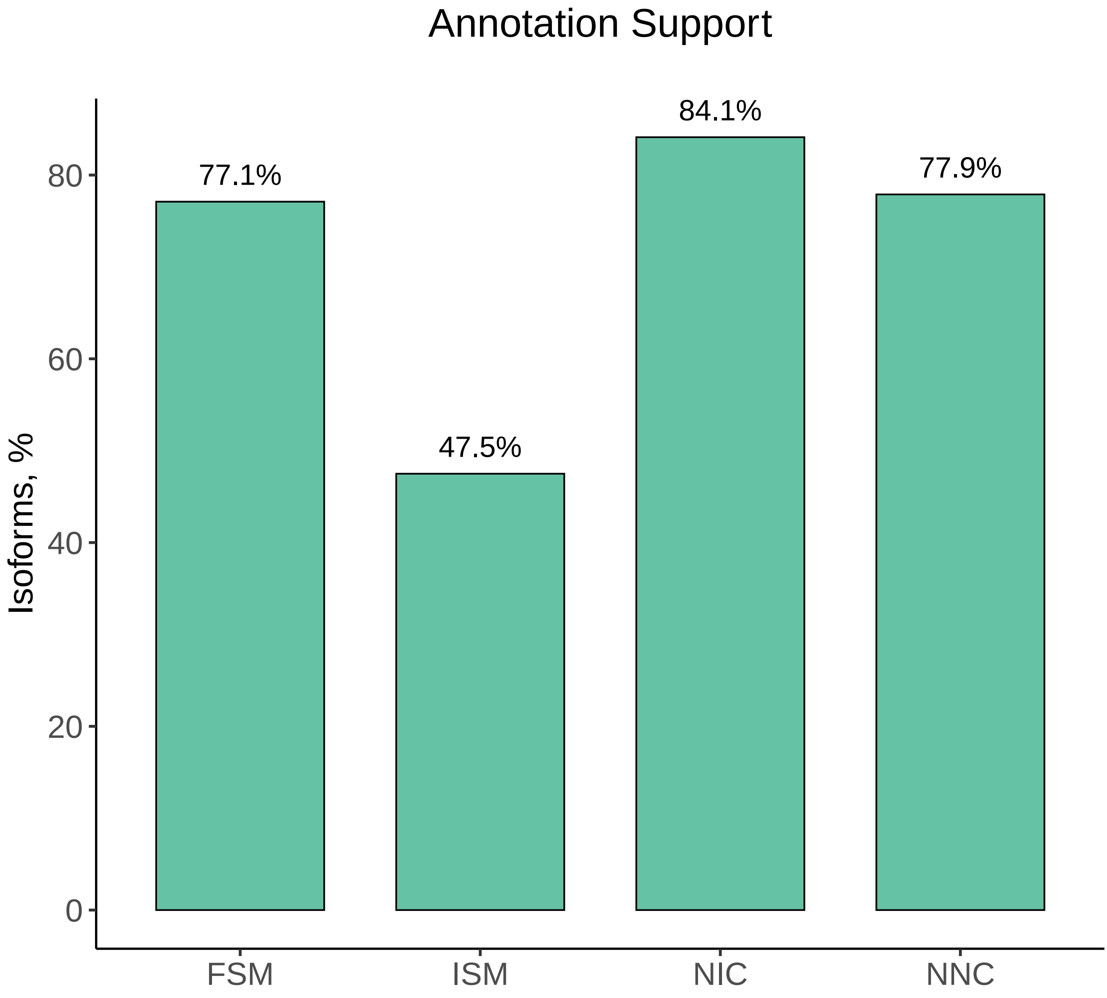
Features of Bad Quality

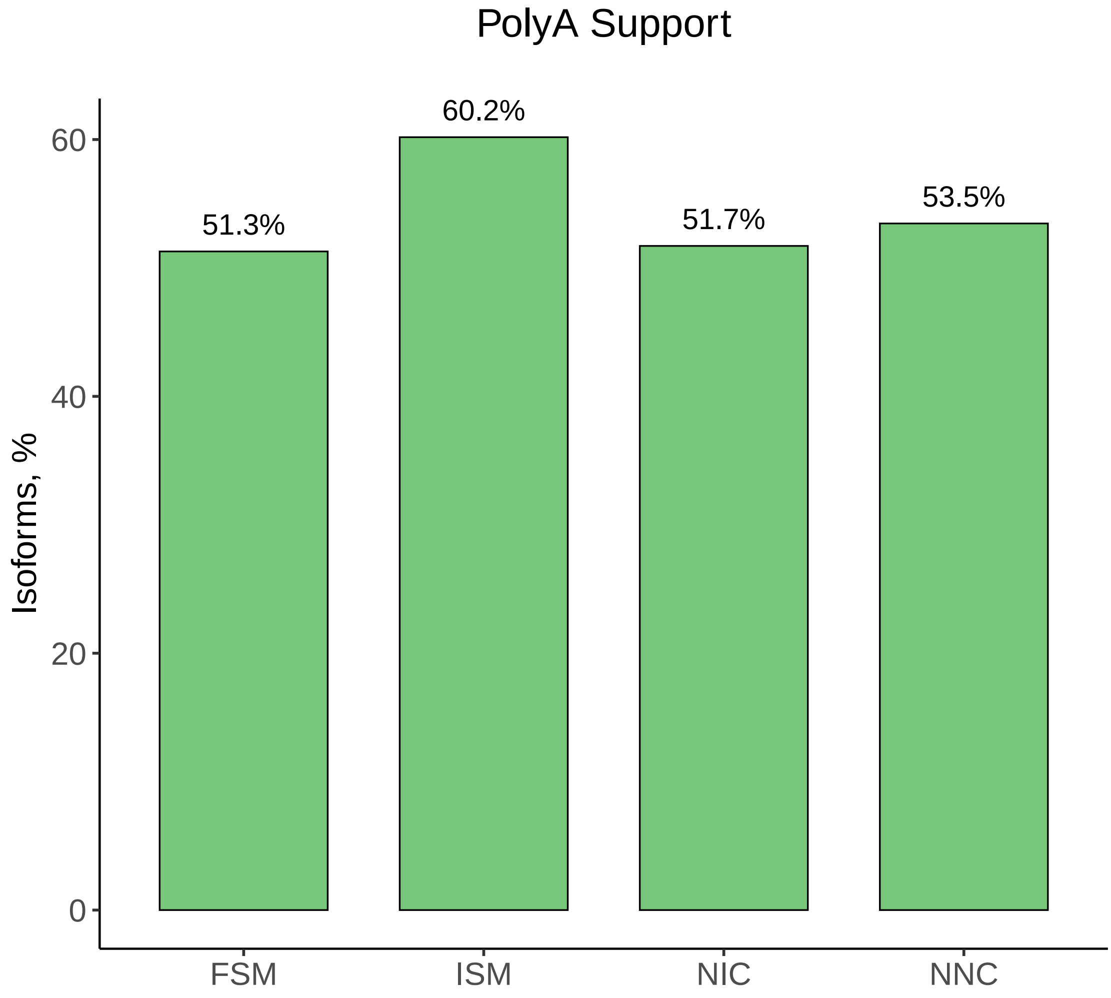

RT−switching
0.9%
Isoforms, %
0.75
0.50
0.25
0.00
FSM

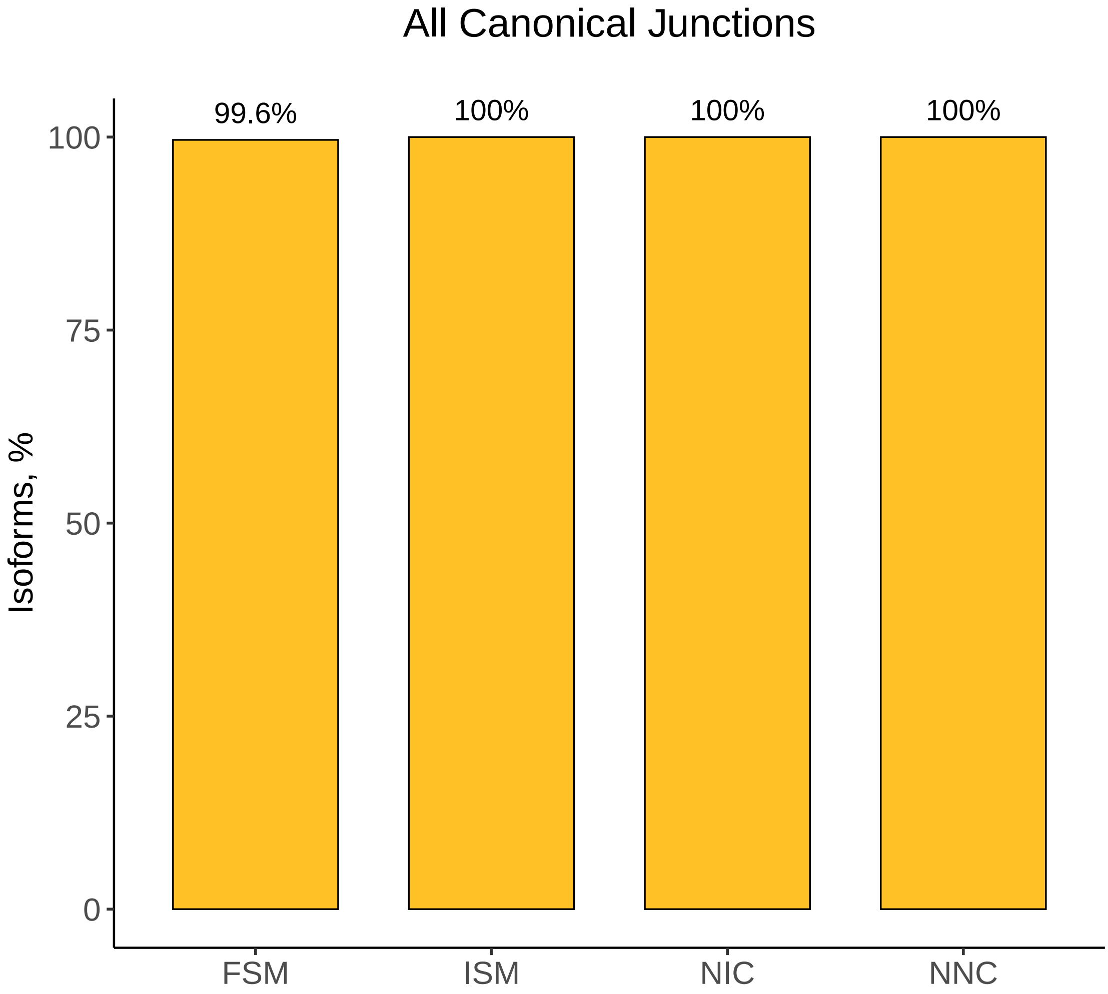

Non−Canonical Junctions
0.4%
Isoforms, %
0.0
0.1
0.2
0.3
FSM

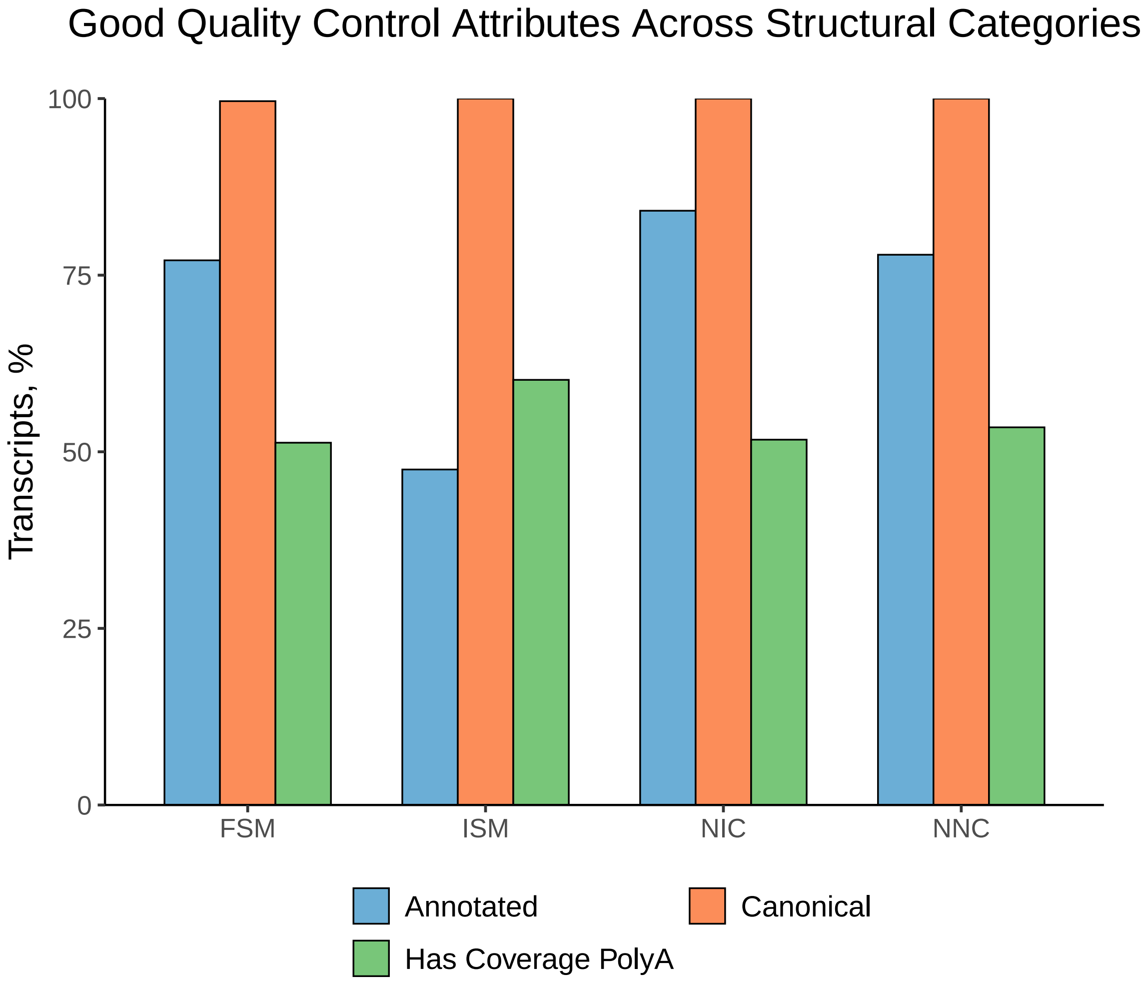

Quality Control Attributes Across Structural Categories
Transcripts, %
100
75
50
25
0
FSM
Non−canonical
RT switching

Features of Good Quality


Annotation Support
Isoforms, %
0
20
40
60
80
77.1%
47.5%
84.1%
77.9%
FSM
ISM
NIC
NNC


PolyA Support
Isoforms, %
0
20
40
60
51.3%
60.2%
51.7%
53.5%
FSM
ISM
NIC
NNC


All Canonical Junctions
Isoforms, %
100
75
50
25
0
99.6%
100%
100%
100%
FSM
ISM
NIC
NNC


Good Quality Control Attributes Across Structural Categories
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Annotated
Canonical
Has Coverage PolyA