SQANTI3 report

Unique Genes: 24764

Unique Isoforms: 626161

Transcript Classification

Category	Isoforms, count
FSM	143351
ISM	244958
NIC	126877
NNC	102353
Genic Genomic	608
Antisense	1812
Fusion	3274
Intergenic	2846
Genic Intron	0

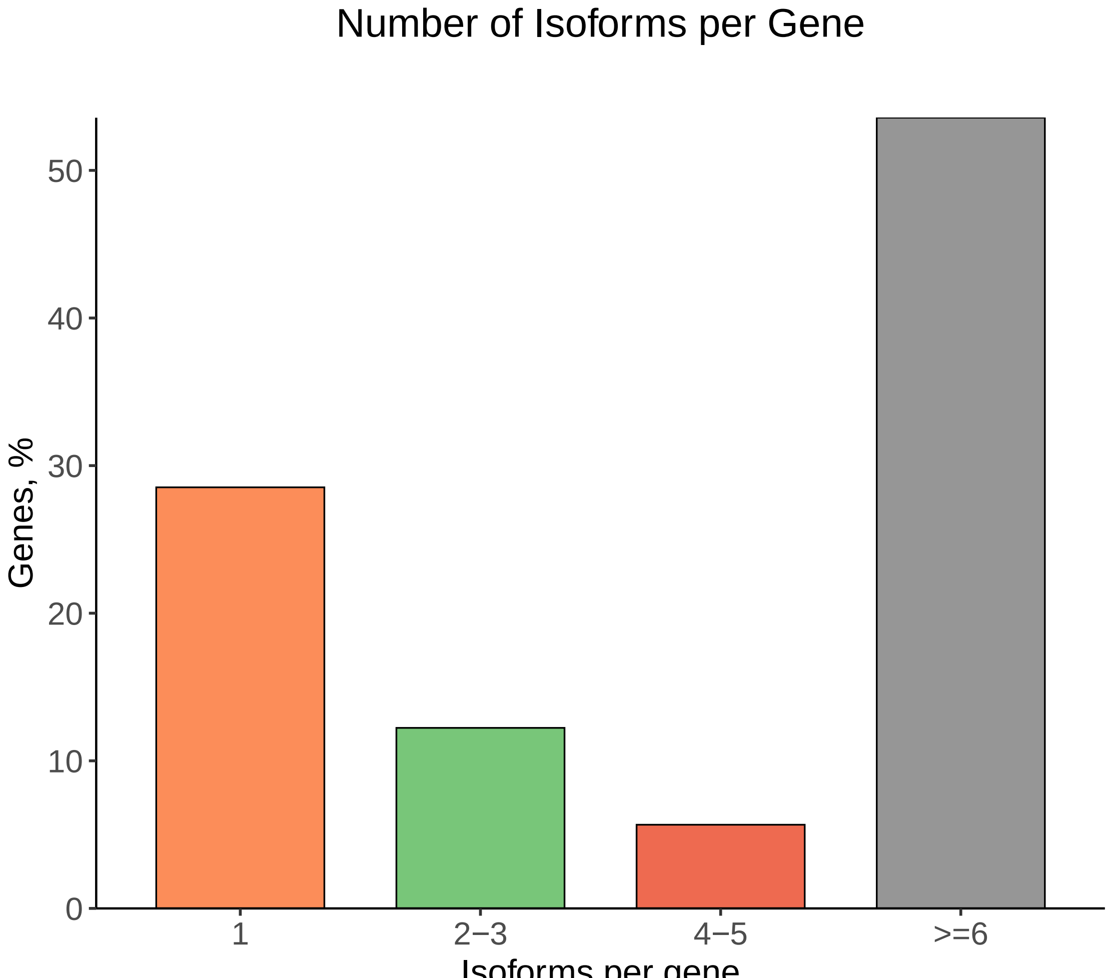
Gene Classification

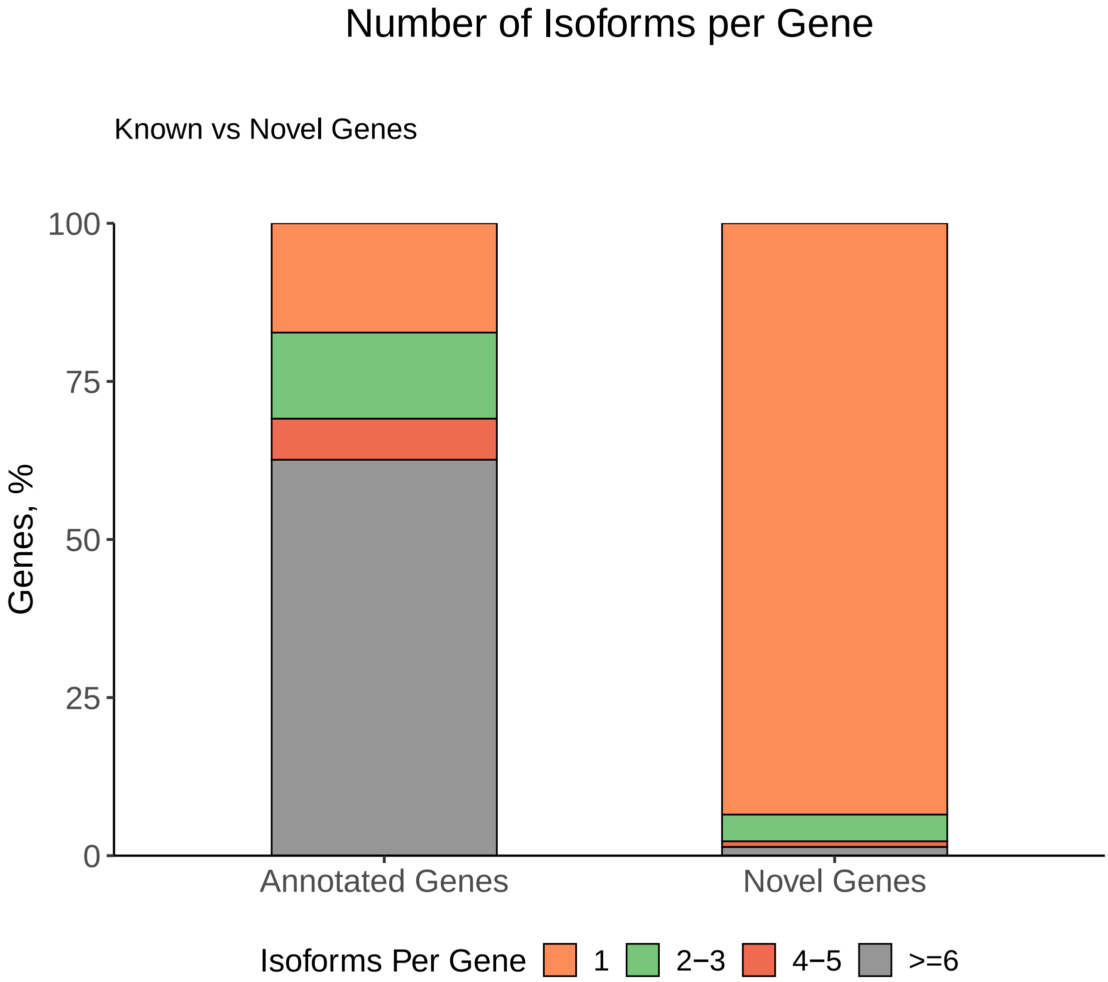
Category	Genes, count
Annotated Genes	21111
Novel Genes	3653

Splice Junction Classification

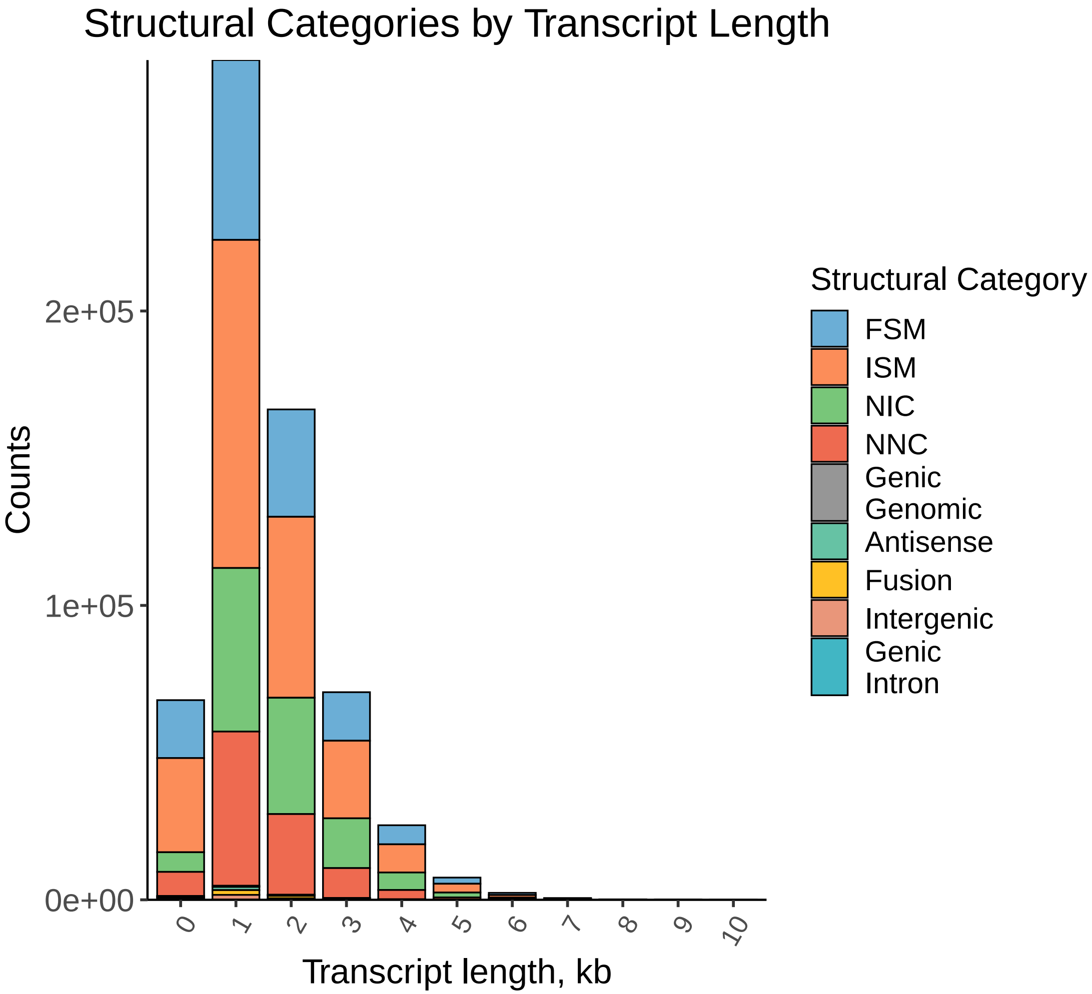
Category	SJs, count	Percent
Known canonical	193511	64.26
Known Non−canonical	85	0.03
Novel canonical	107554	35.71
Novel Non−canonical	0	0.00

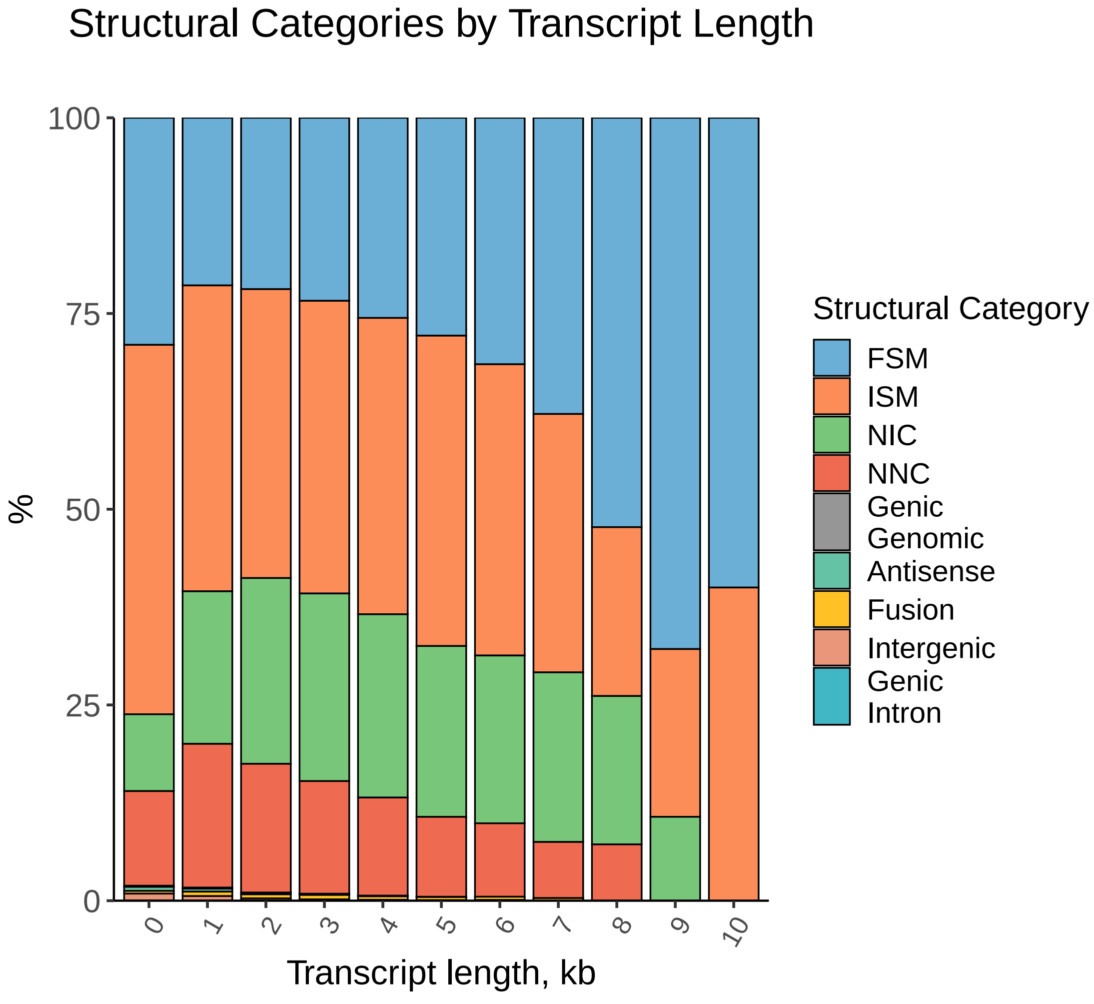
Gene Characterization

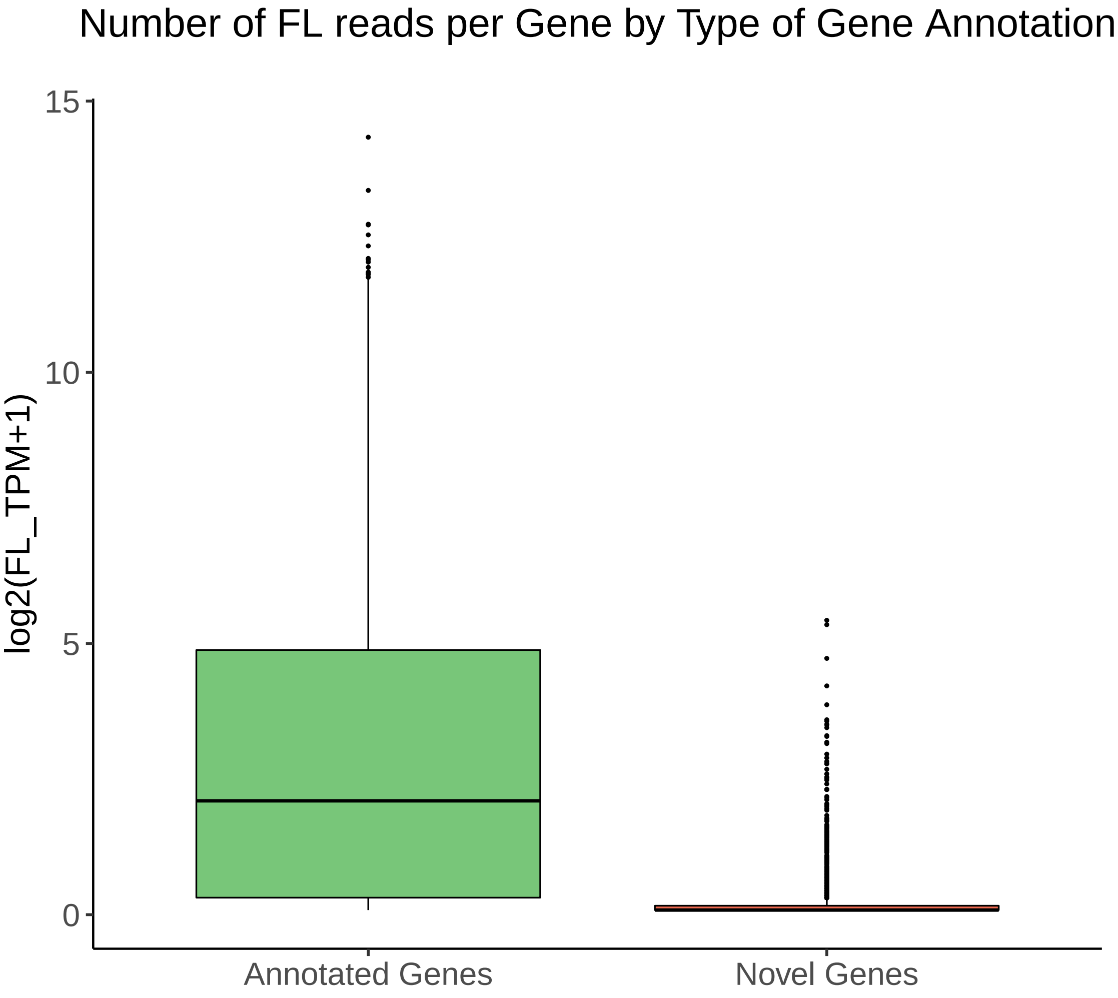

Number of Isoforms per Gene
Genes, %
0
10
20
30
40
50
1
2−3
4−5
>=6
Isoforms per gene

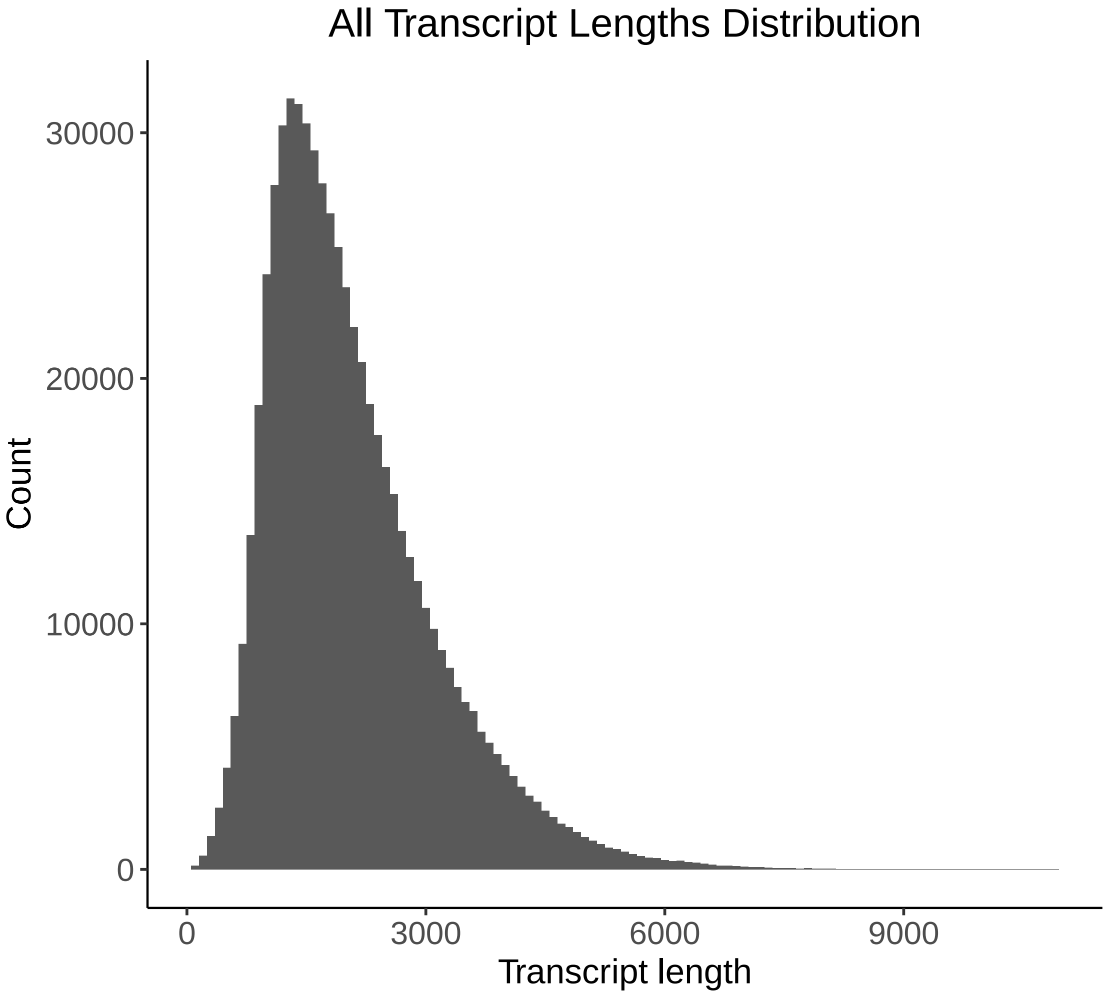

Number of Isoforms per Gene
Known vs Novel Genes
Genes, %
100
75
50
25
0
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3
4−5
>=6

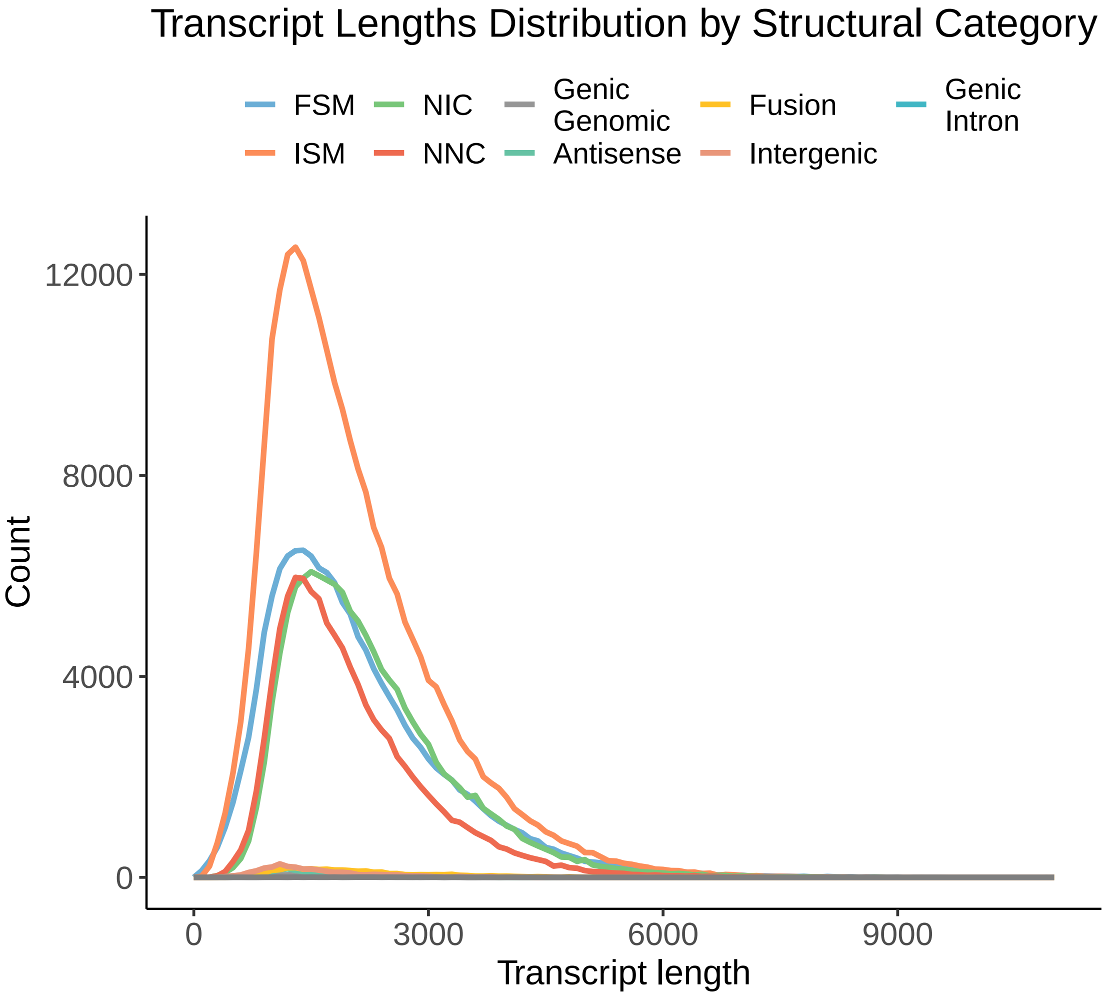

Distribution of Mono− vs Multi−Exon Transcripts
Transcripts, %
100
75
50
25
0
Novel Genes
Annotated Genes
Transcript type
Multi−Exon
Mono−Exon

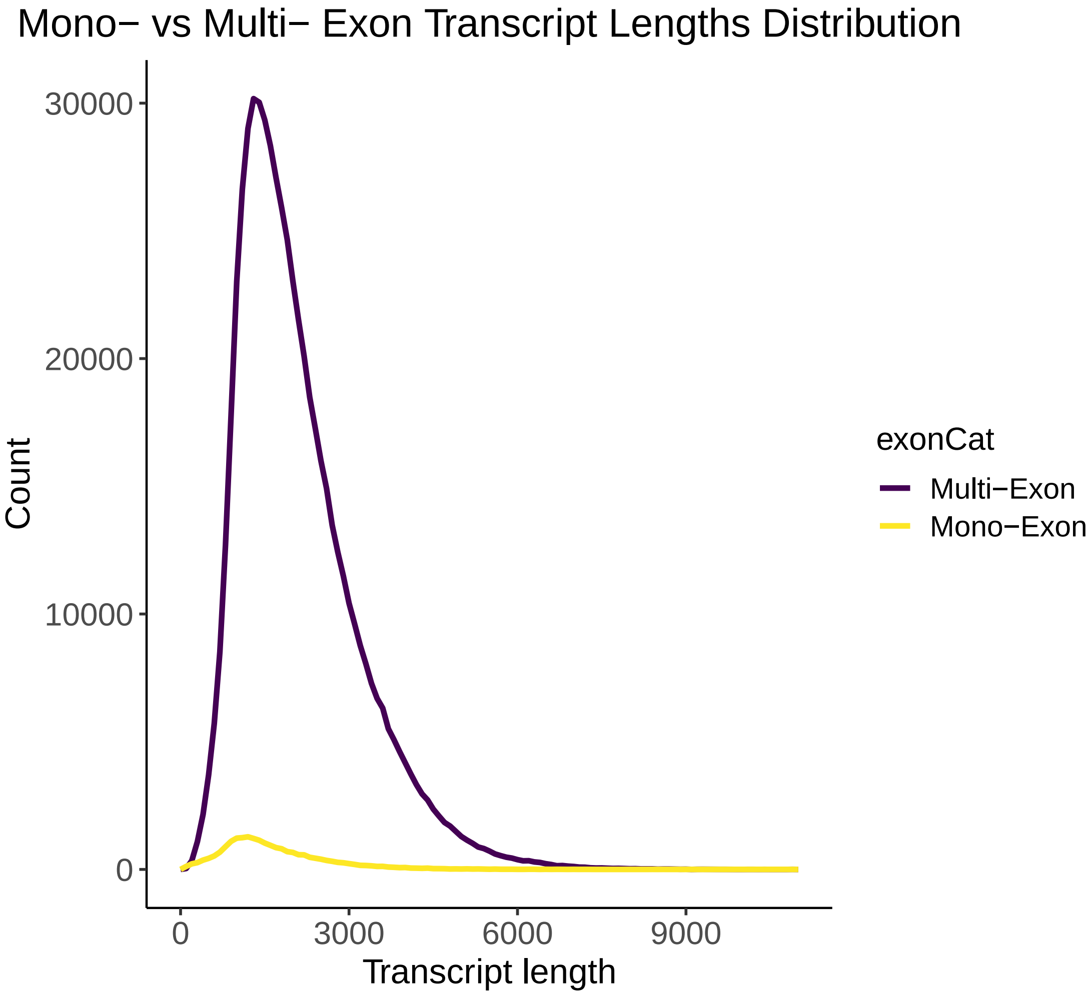

Structural Categories by Transcript Length
Counts
0e+00
1e+05
2e+05
0
1
2
3
4
5
6
7
8
9
10
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

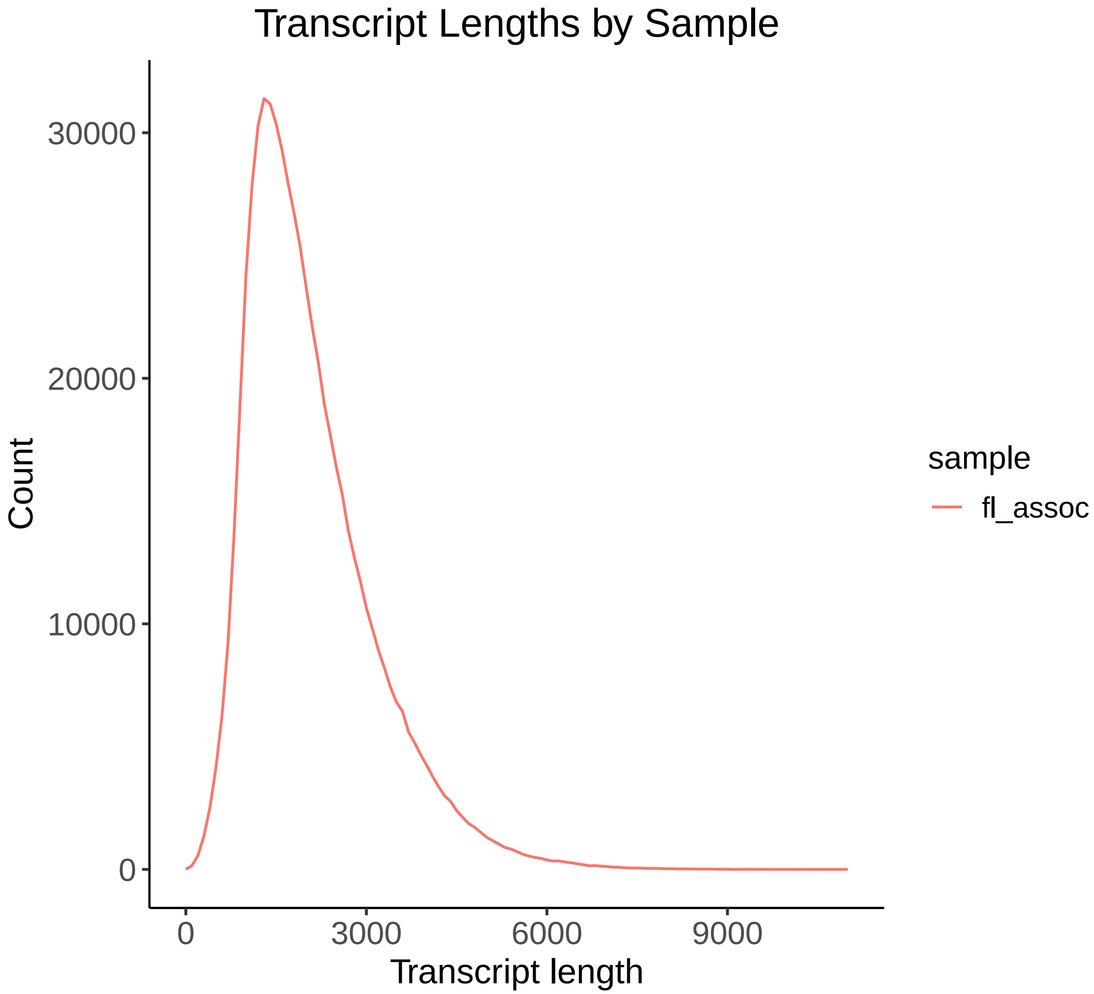

Structural Categories by Transcript Length
%
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
10
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

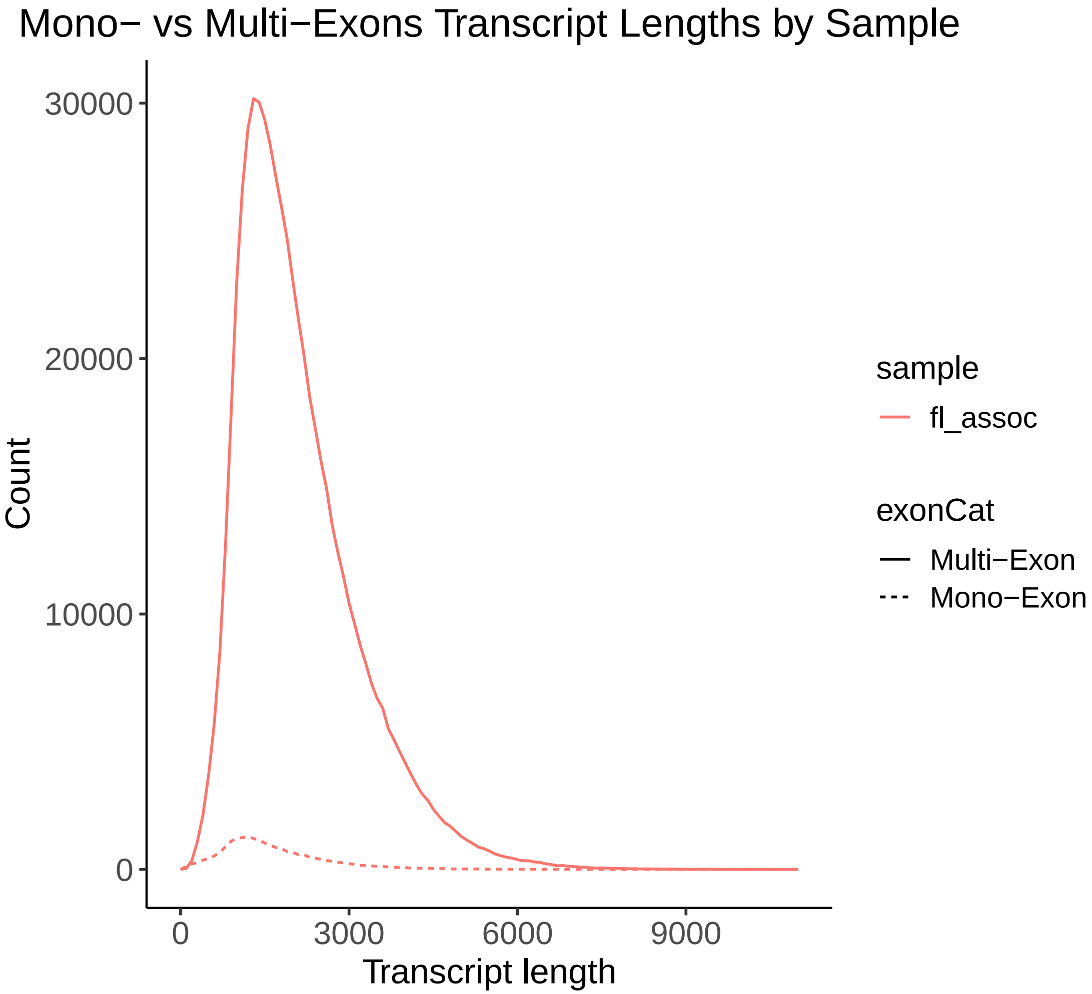

Number of FL reads per Gene by Type of Gene Annotation
log2(FL_TPM+1)
15
10
5
0
Annotated Genes
Novel Genes


All Transcript Lengths Distribution
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length

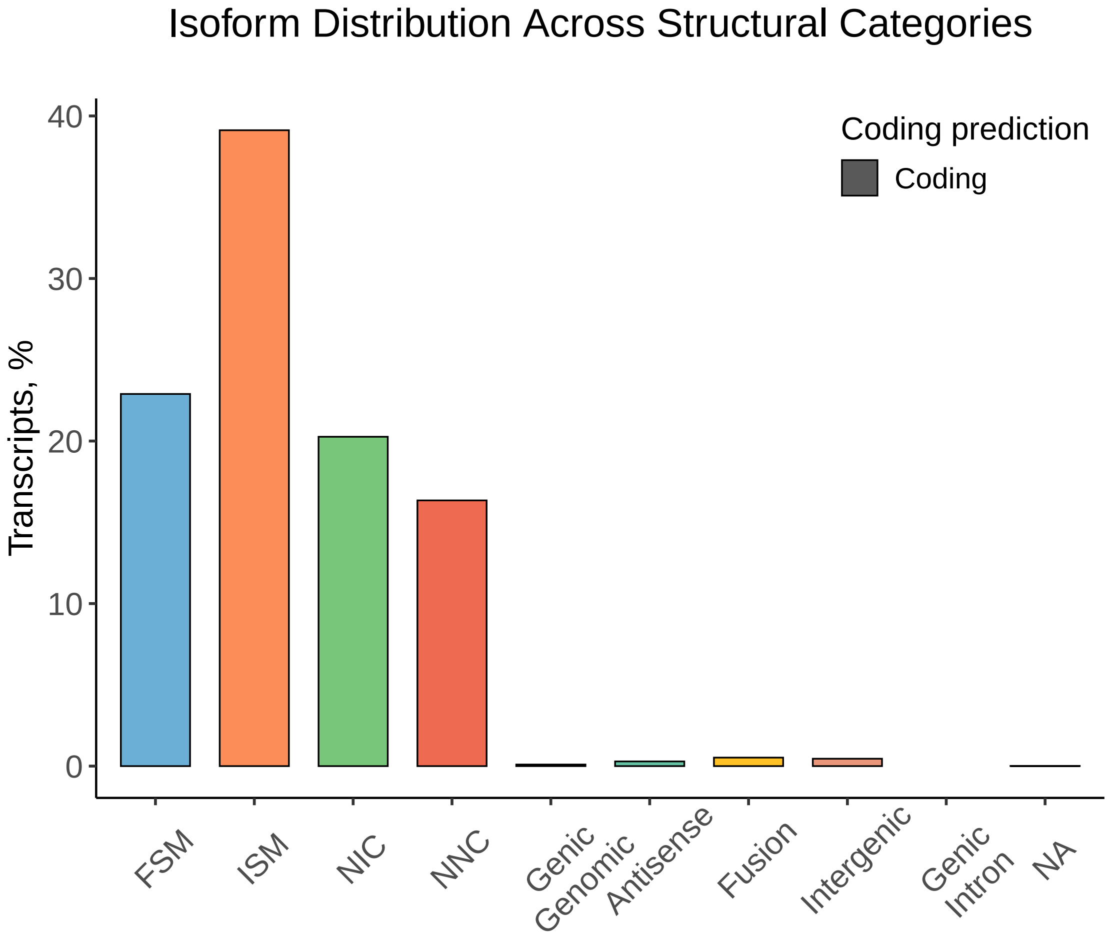

Transcript Lengths Distribution by Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Count
12000
8000
4000
0
0
3000
6000
9000
Transcript length

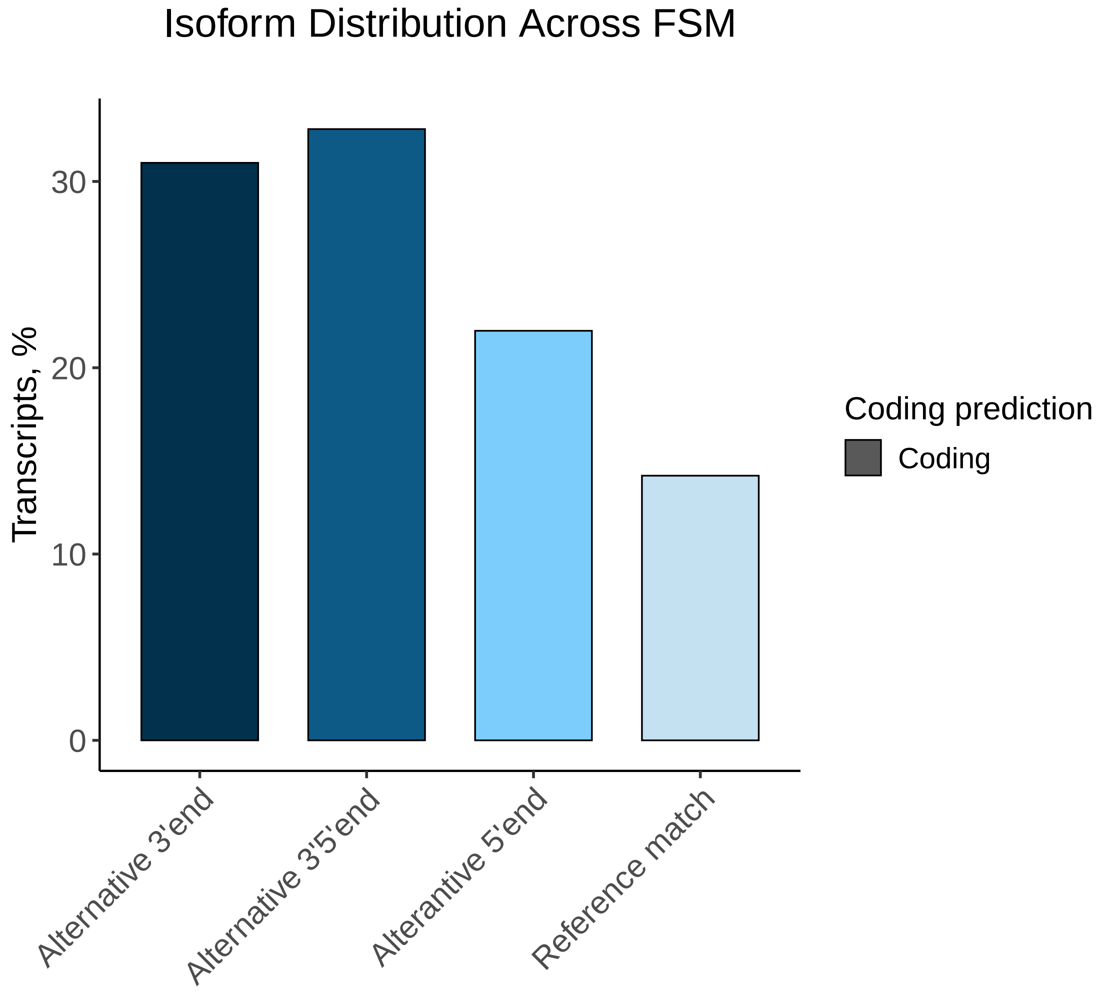

Mono− vs Multi− Exon Transcript Lengths Distribution
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length
exonCat
Multi−Exon
Mono−Exon

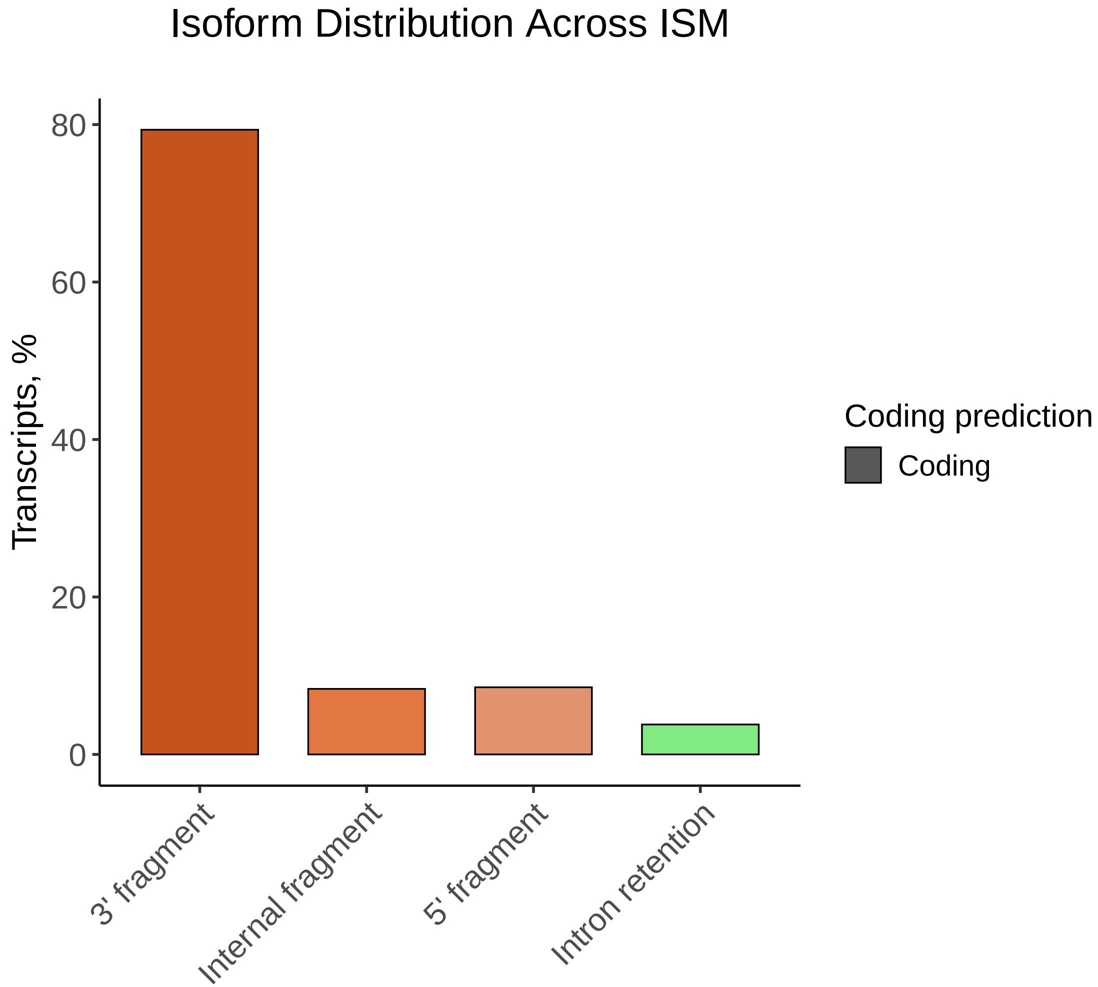

Transcript Lengths by Sample
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc

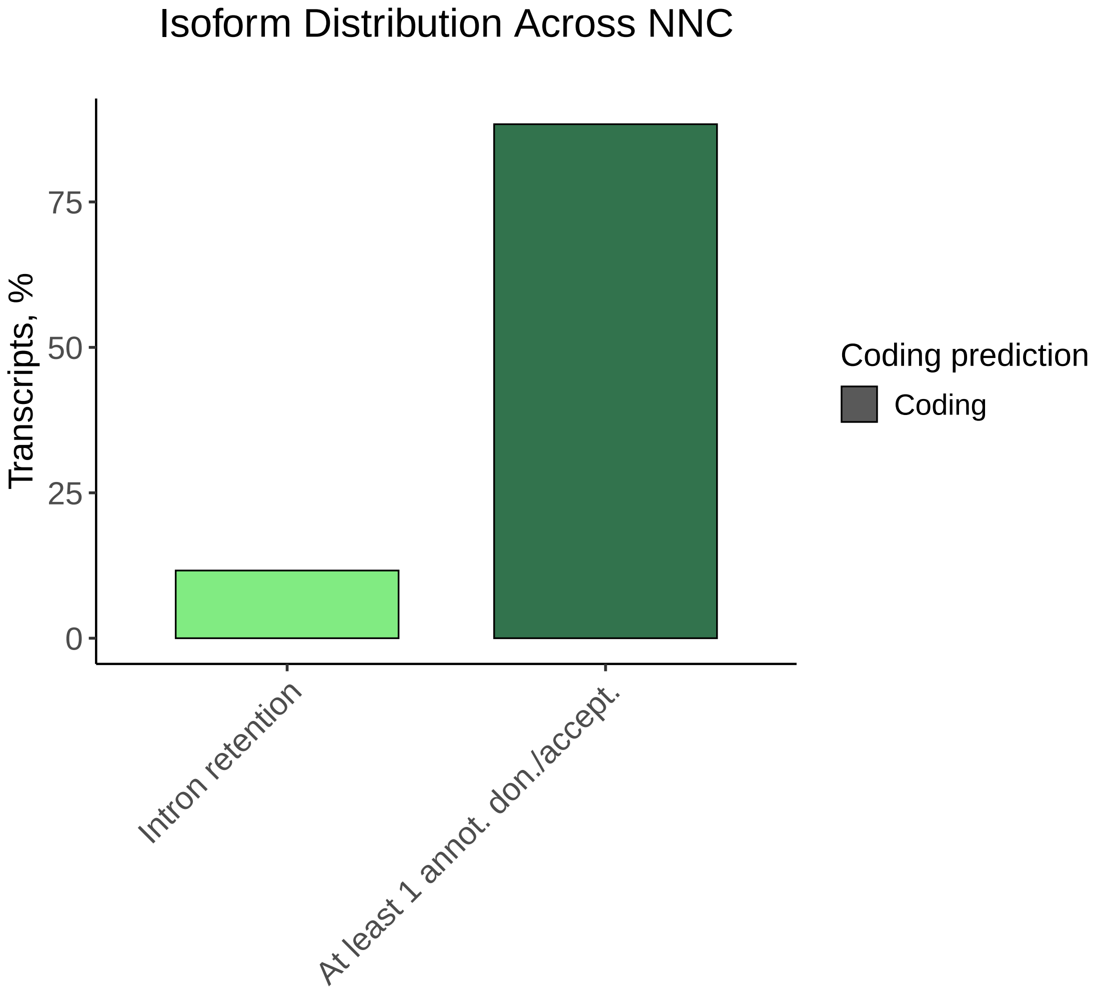

Mono− vs Multi−Exons Transcript Lengths by Sample
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc
exonCat
Multi−Exon
Mono−Exon

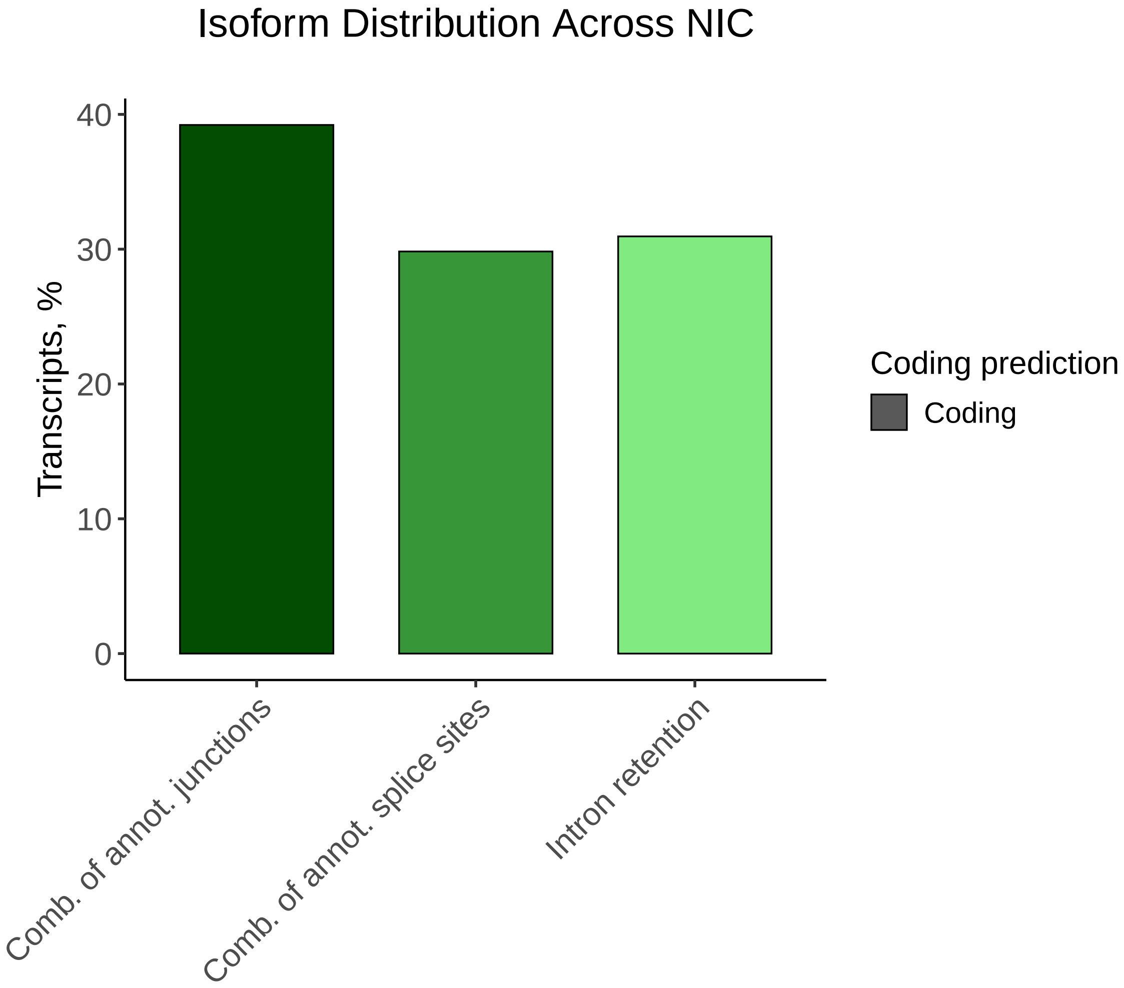
Structural Isoform Characterization

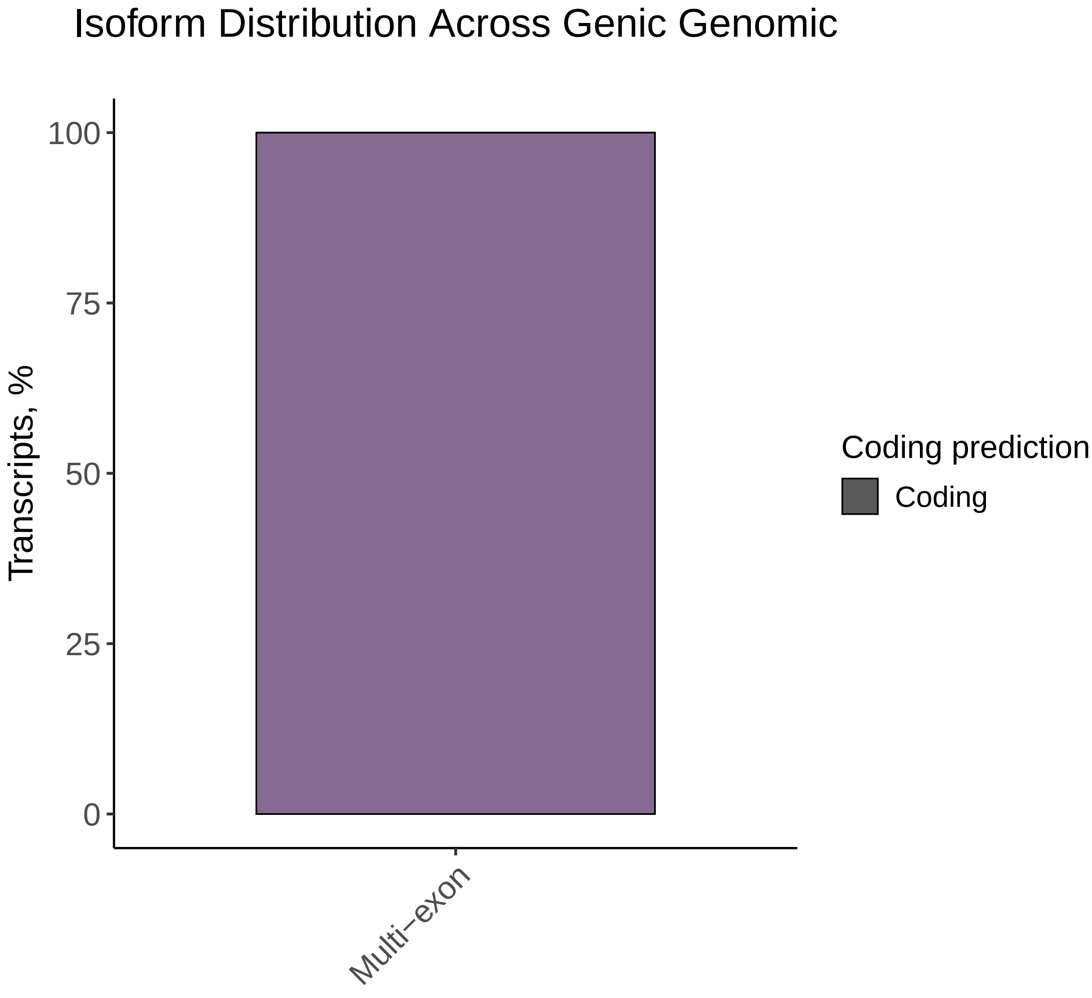

Isoform Distribution Across Structural Categories
Transcripts, %
0
10
20
30
40
Coding prediction
Coding
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

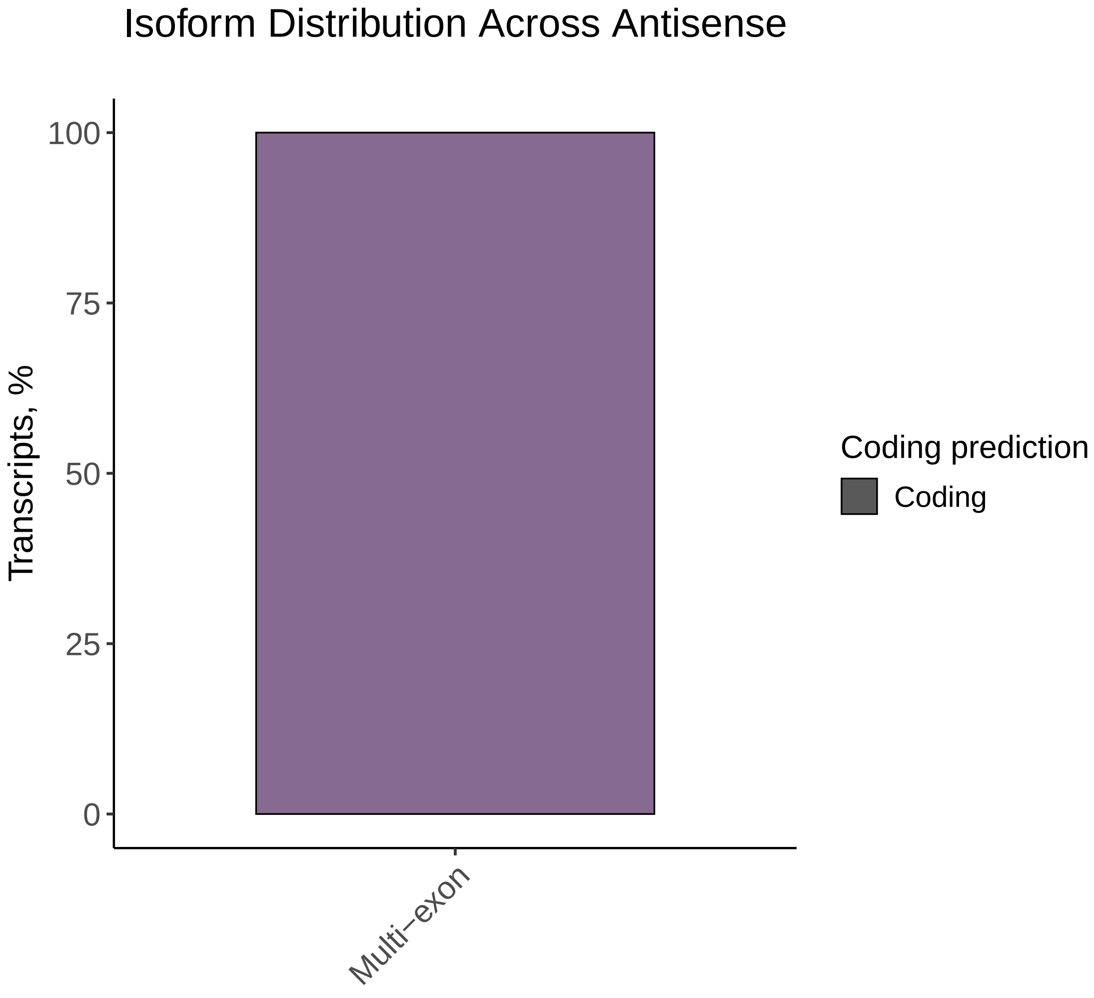

Isoform Distribution Across FSM
Transcripts, %
0
10
20
30
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Coding prediction
Coding

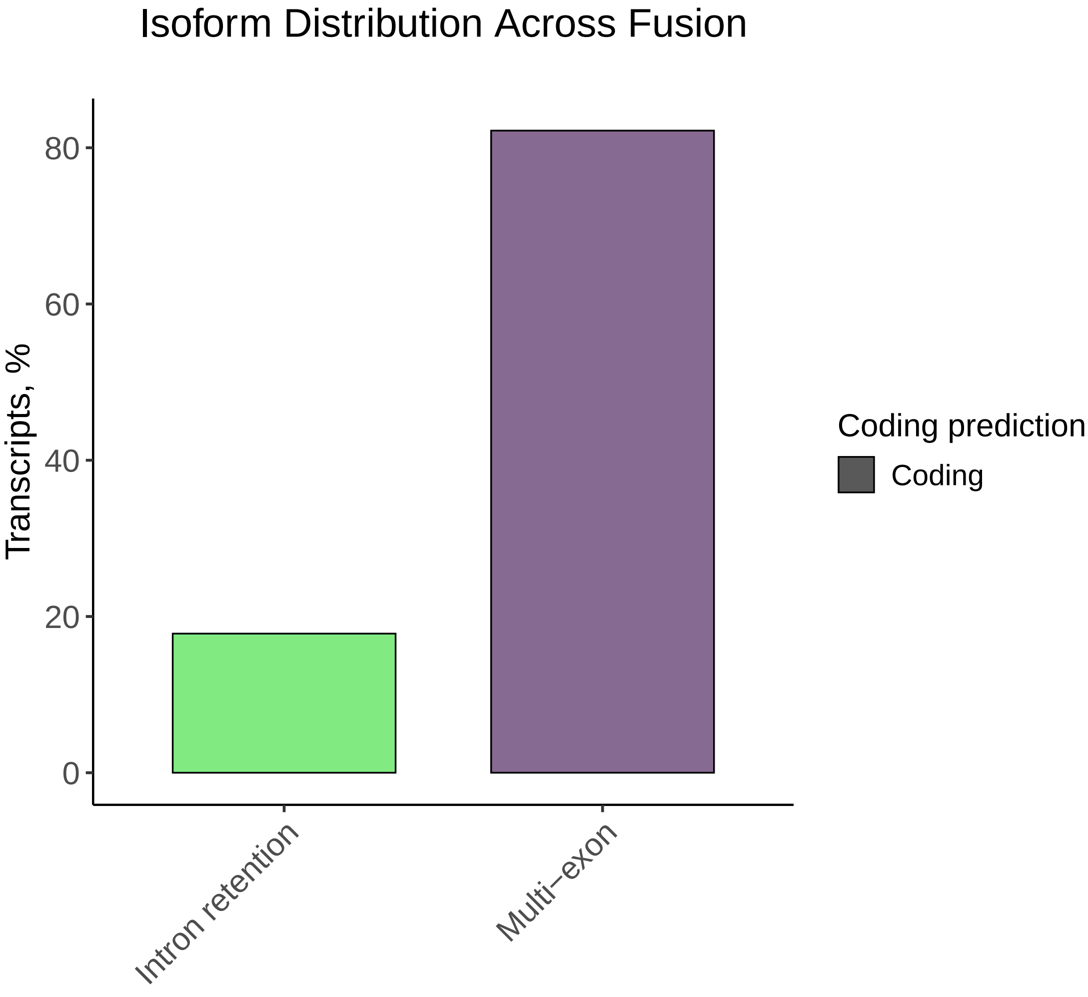

Isoform Distribution Across ISM
Transcripts, %
80
60
40
20
0
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding

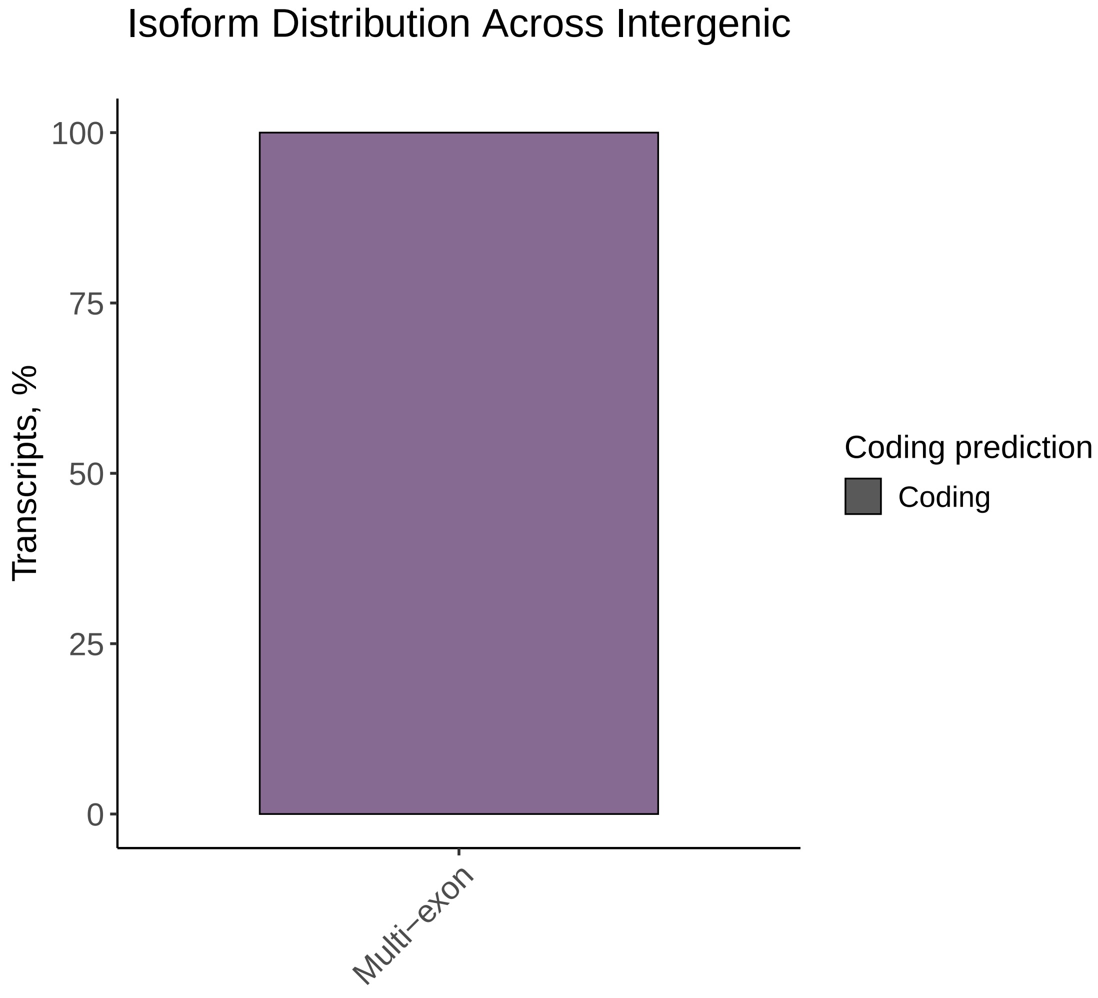

Isoform Distribution Across NNC
Transcripts, %
0
25
50
75
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding

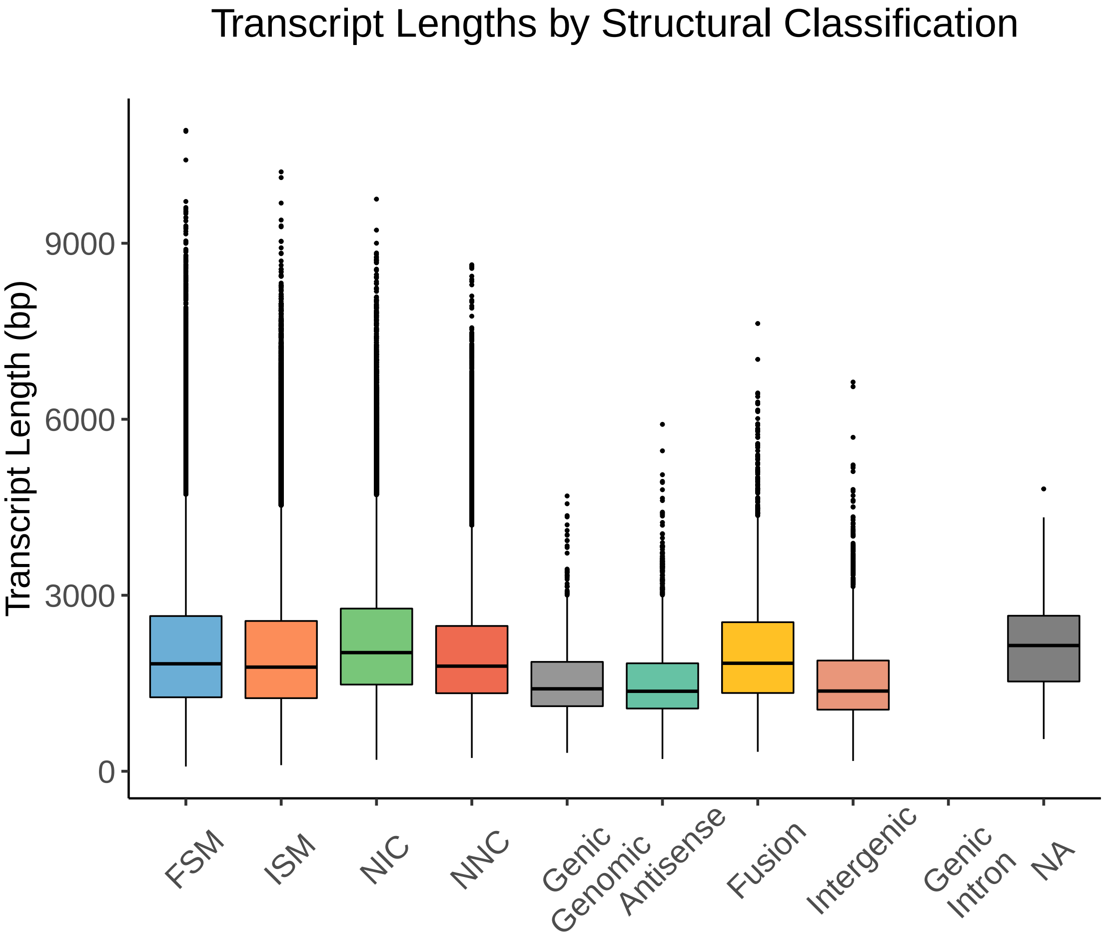

Isoform Distribution Across NIC
Transcripts, %
40
30
20
10
0
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Coding prediction
Coding

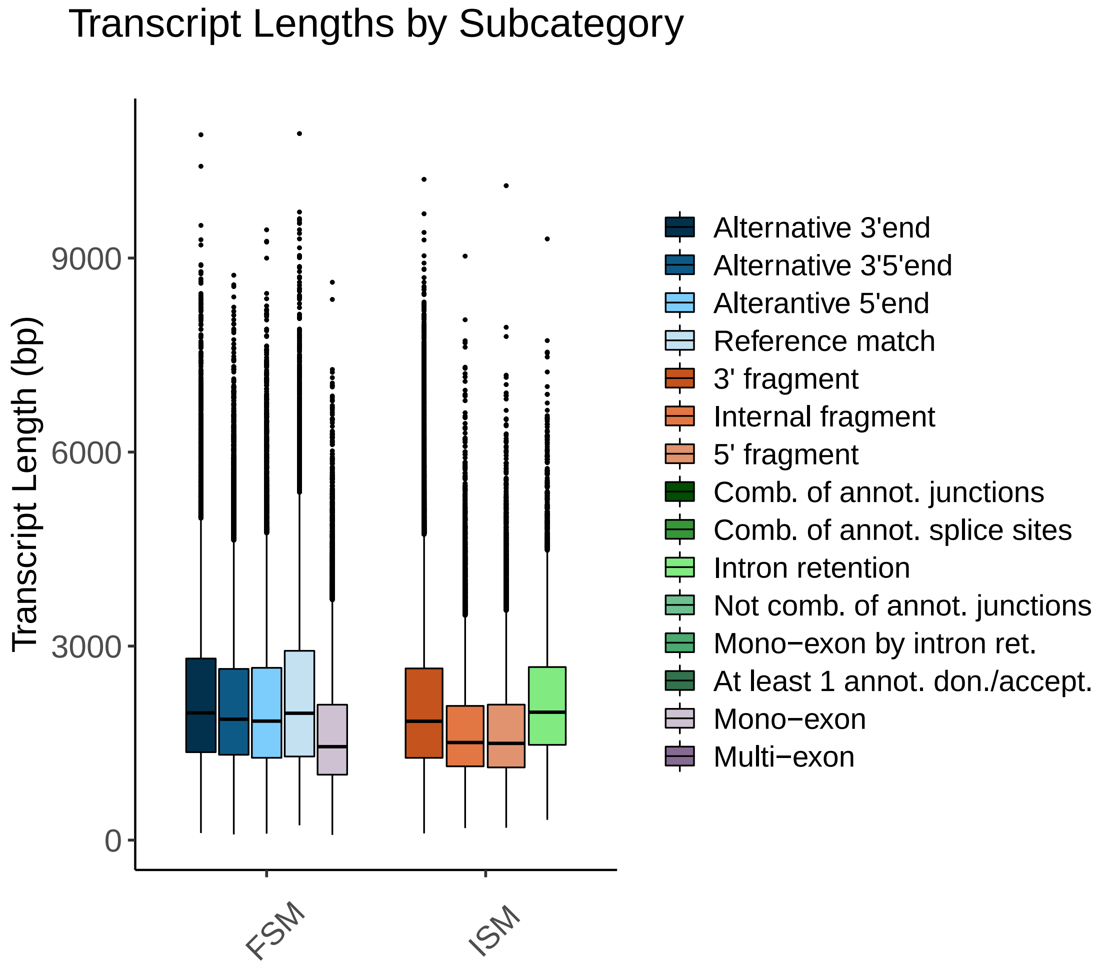

Isoform Distribution Across Genic Genomic
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding

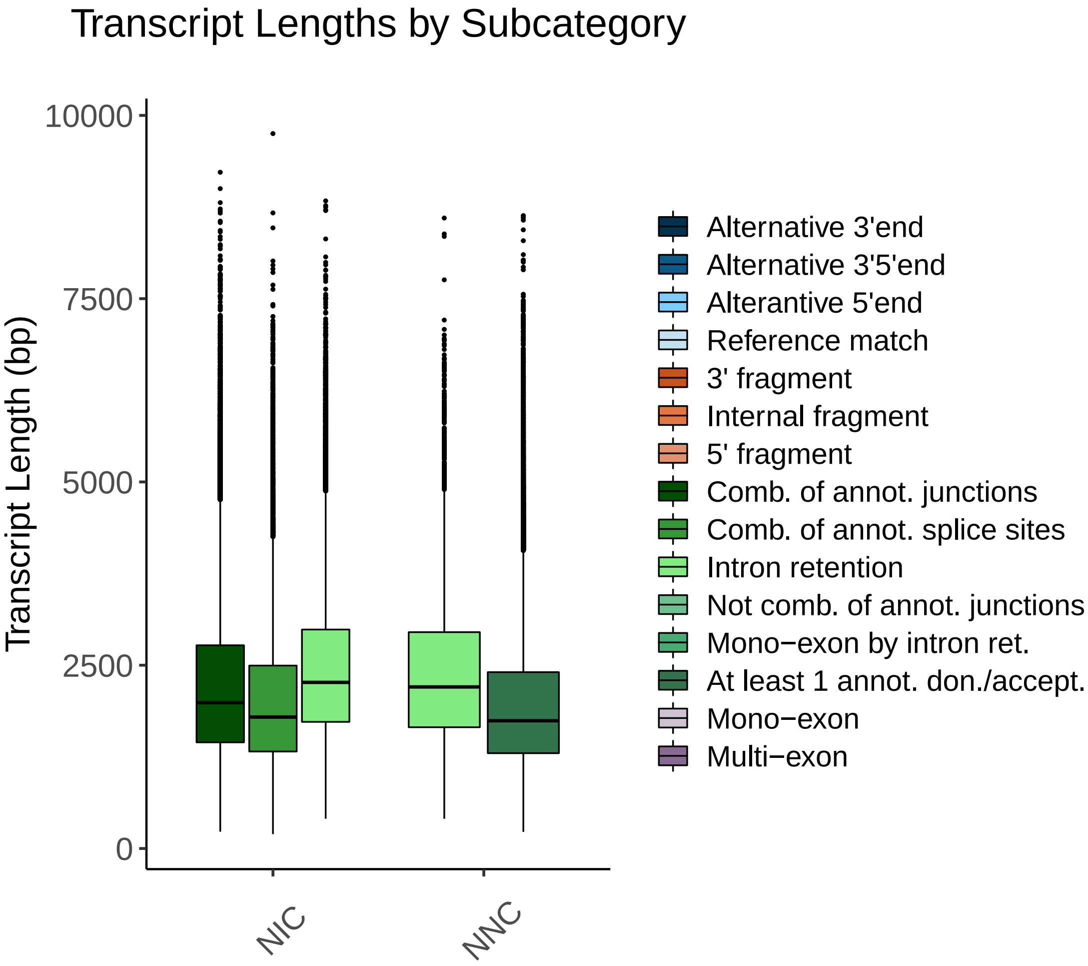

Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding

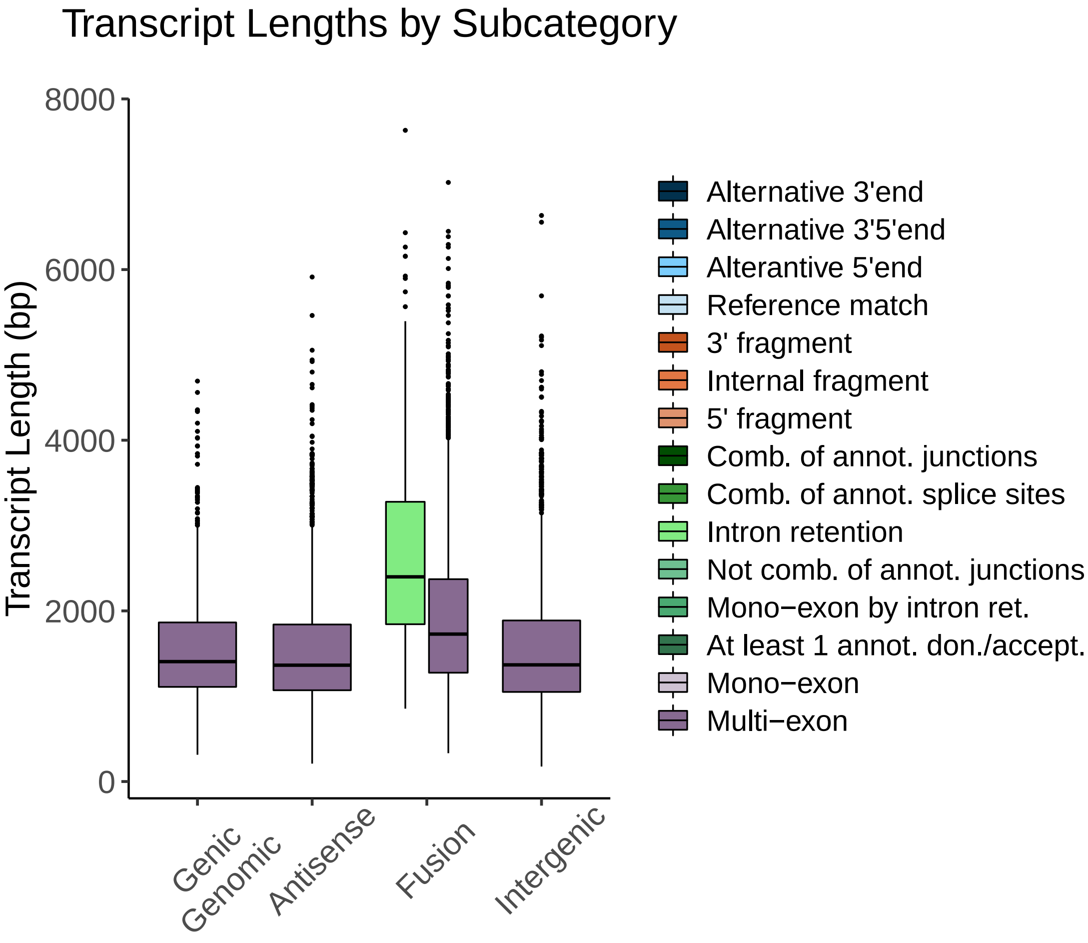

Isoform Distribution Across Fusion
Transcripts, %
0
20
40
60
80
Intron retention
Multi–exon
Coding prediction
Coding

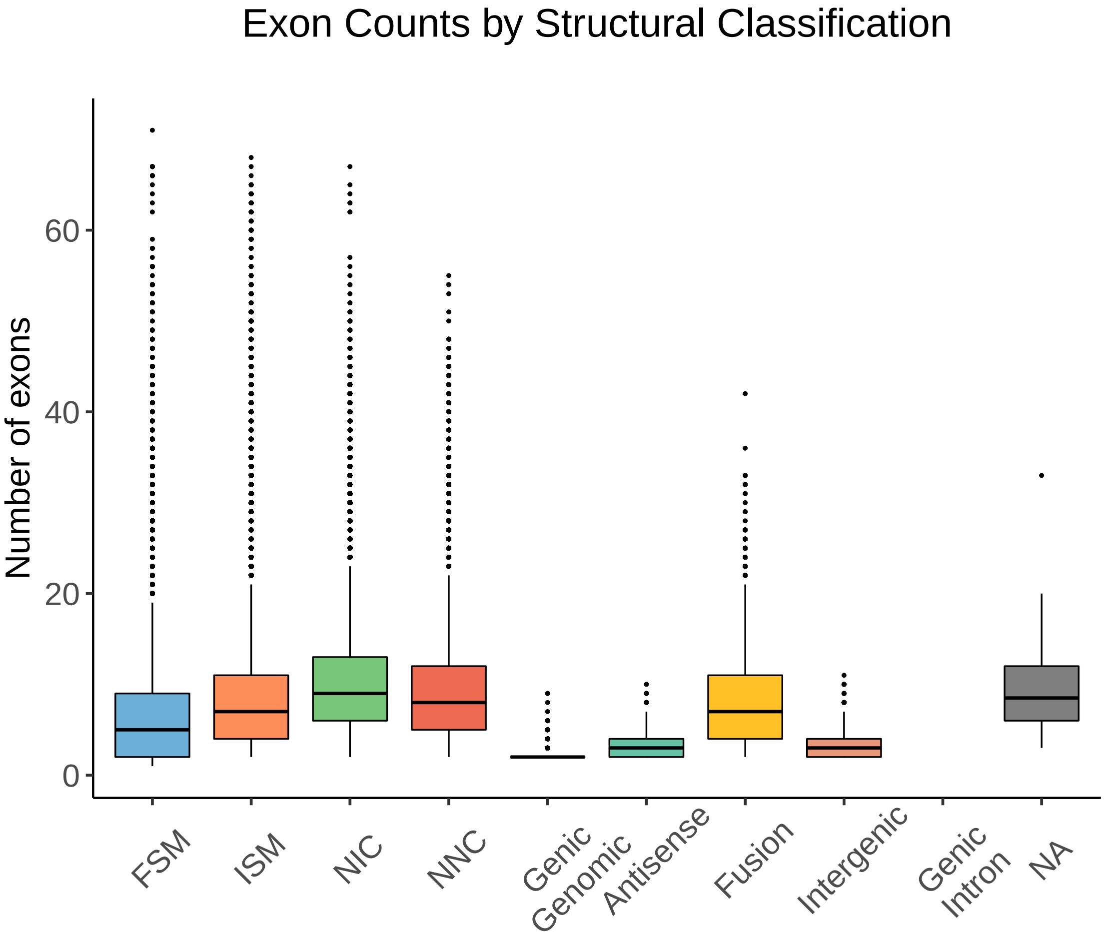

Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding

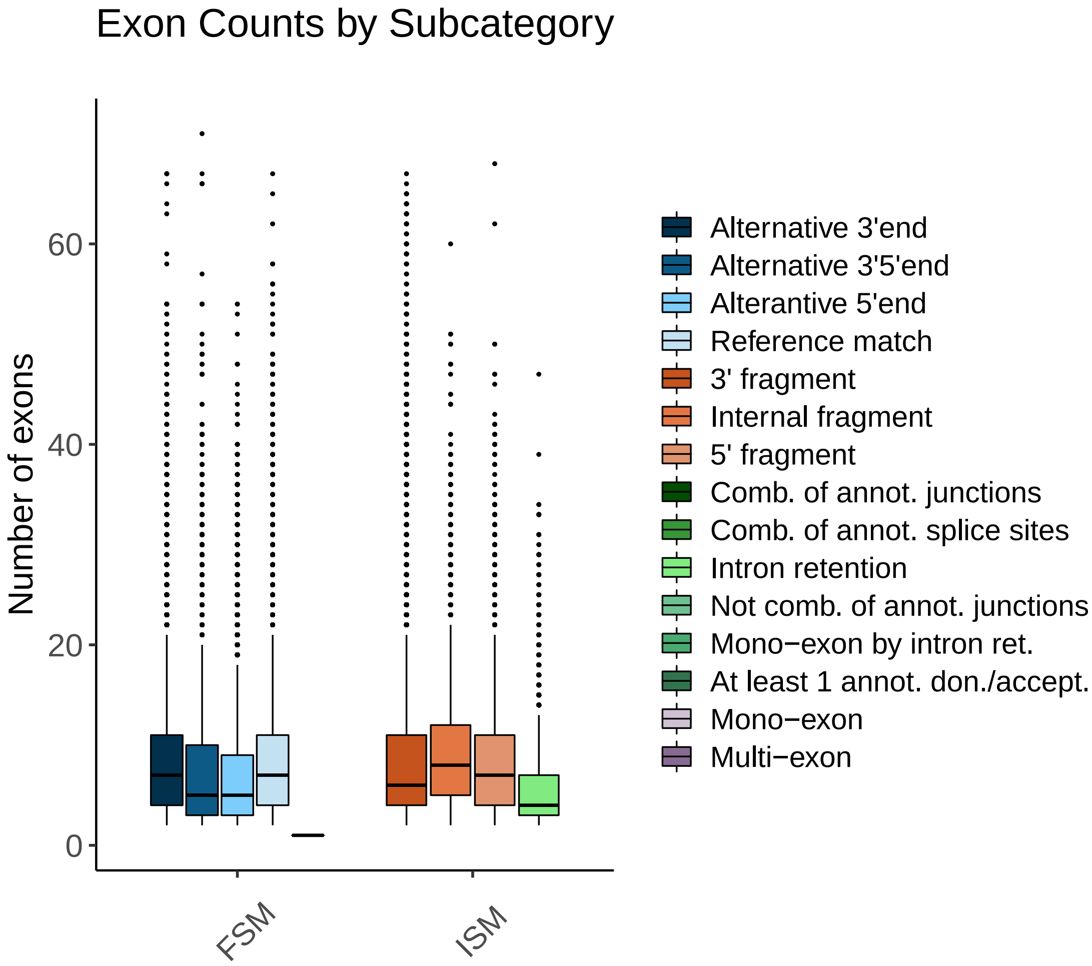

Transcript Lengths by Structural Classification
Transcript Length (bp)
9000
6000
3000
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

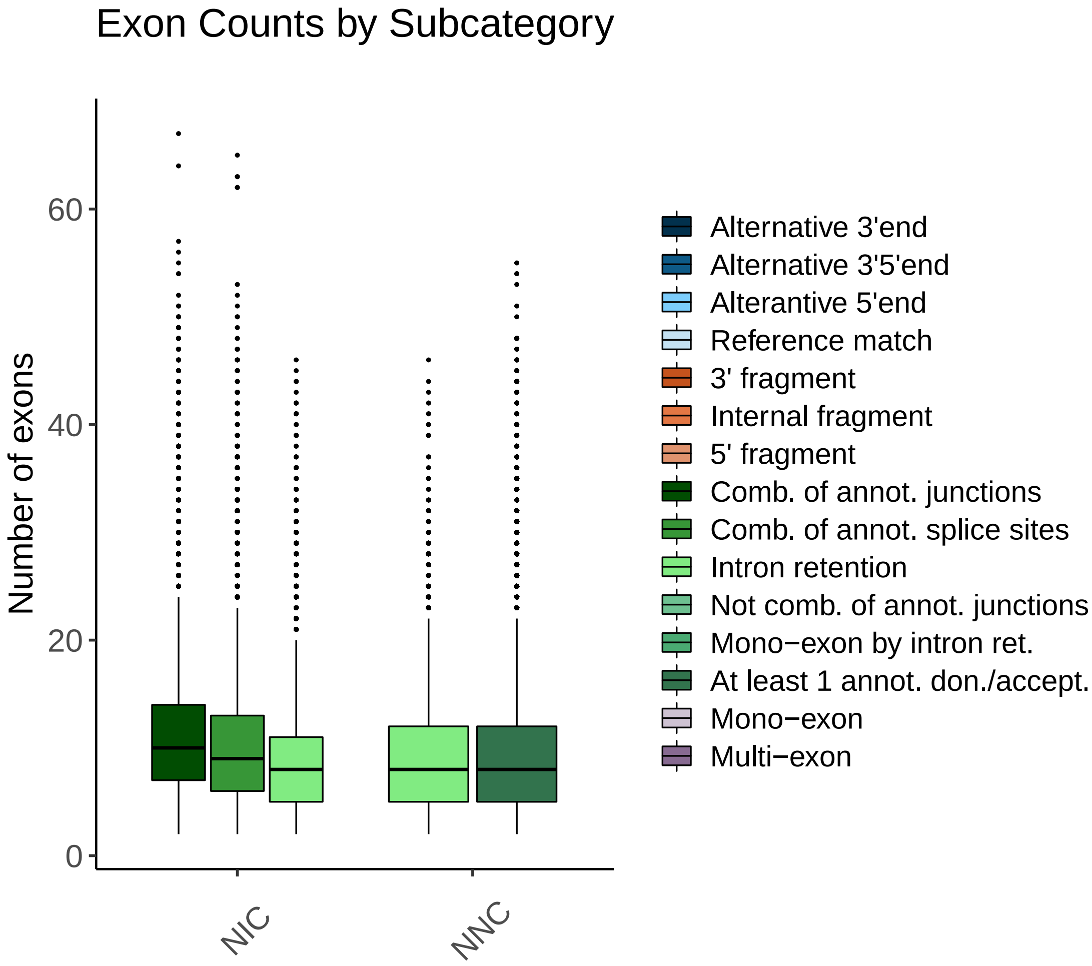

Transcript Lengths by Subcategory
Transcript Length (bp)
0
3000
6000
9000
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon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

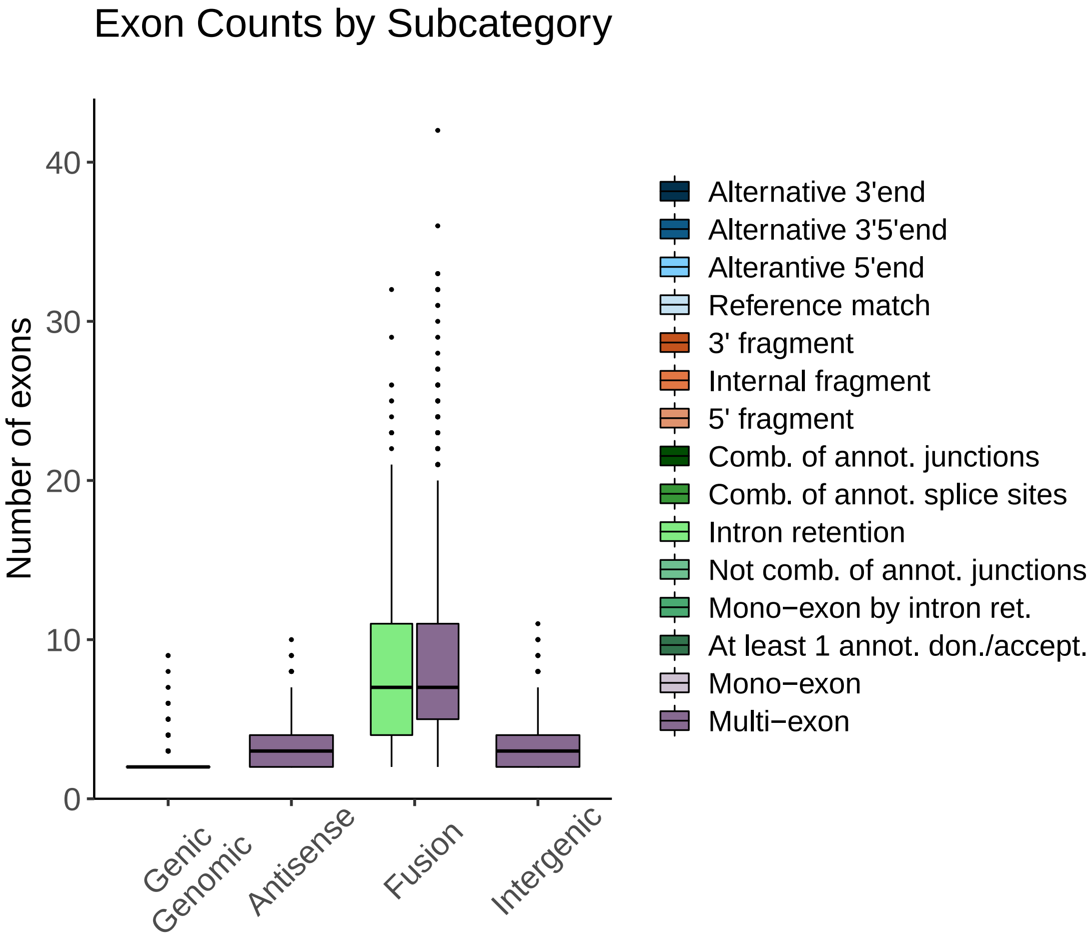

Transcript Lengths by Subcategory
Transcript Length (bp)
10000
7500
5000
2500
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions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

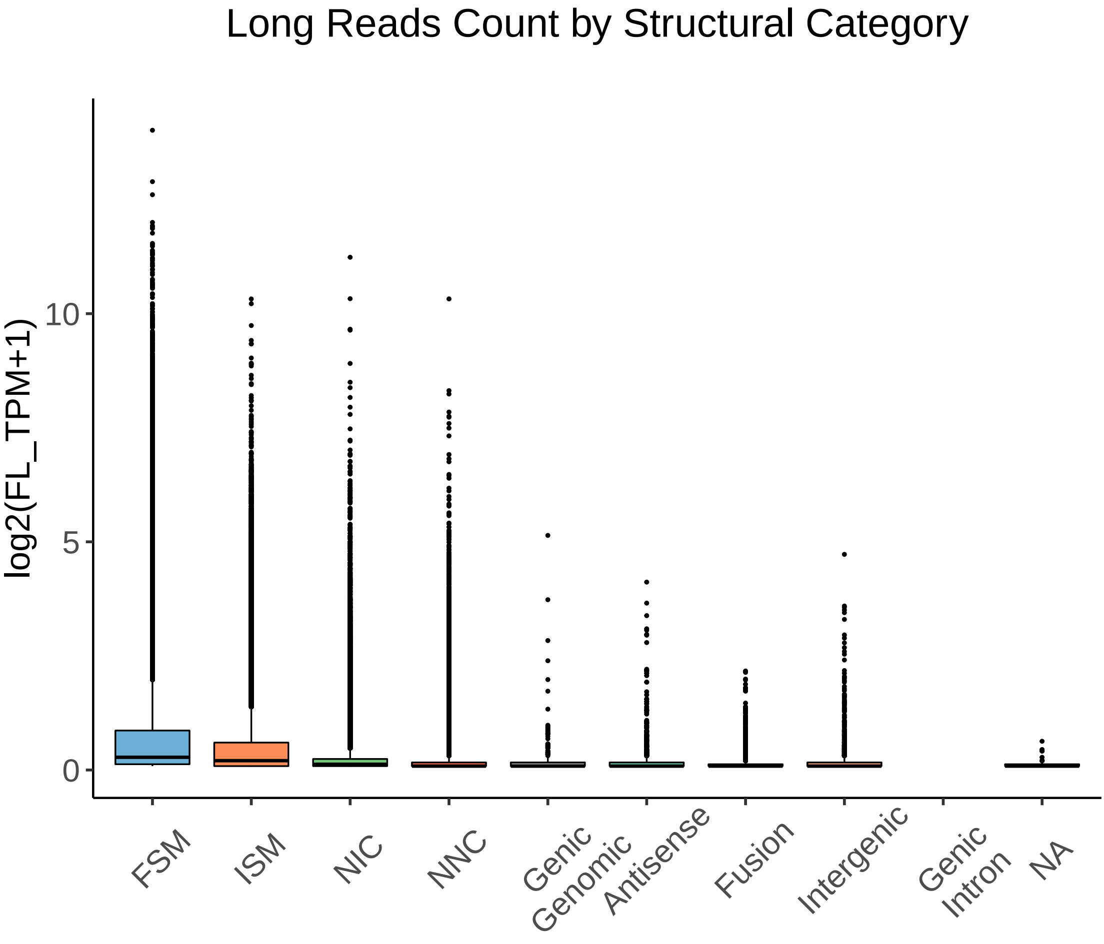

Transcript Lengths by Subcategory
Transcript Length (bp)
8000
6000
4000
2000
0
Genic Genomic
Antisense
Fusion
Intergenic
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention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


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

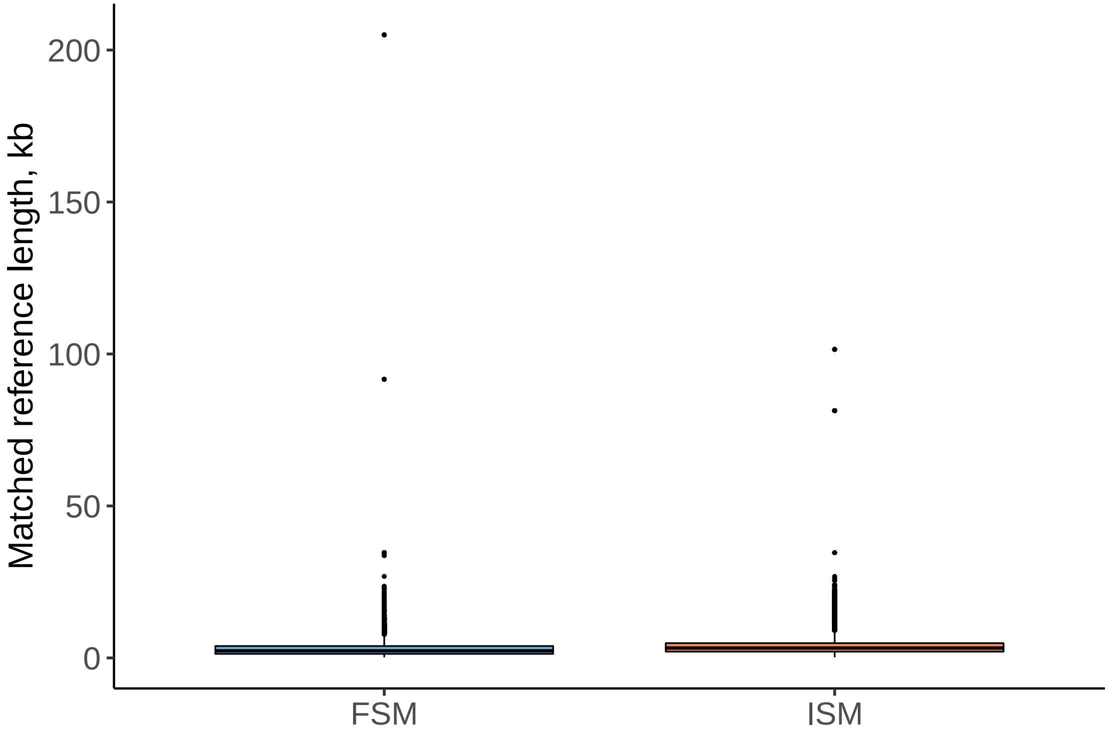

Exon Counts by Subcategory
Number of exons
60
40
20
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.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


Exon Counts by Subcategory
Number of exons
0
10
20
30
40
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

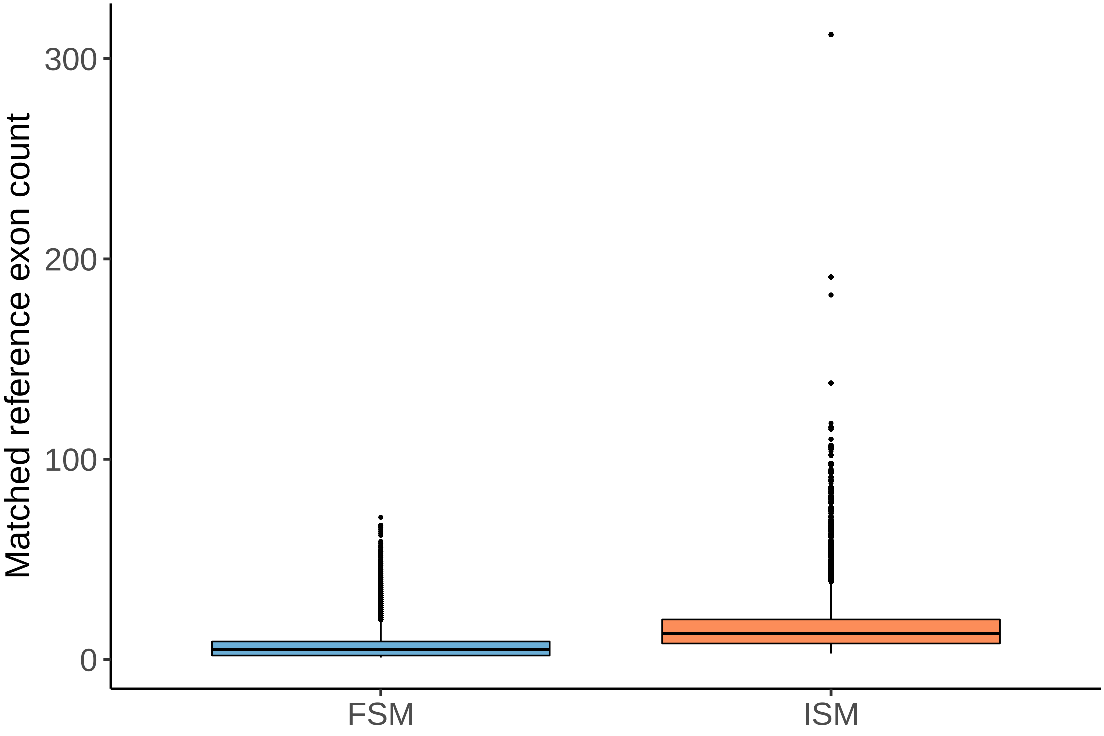
Length Distribution of Matched Reference Transcripts

Applicable Only to FSM and ISM Categories

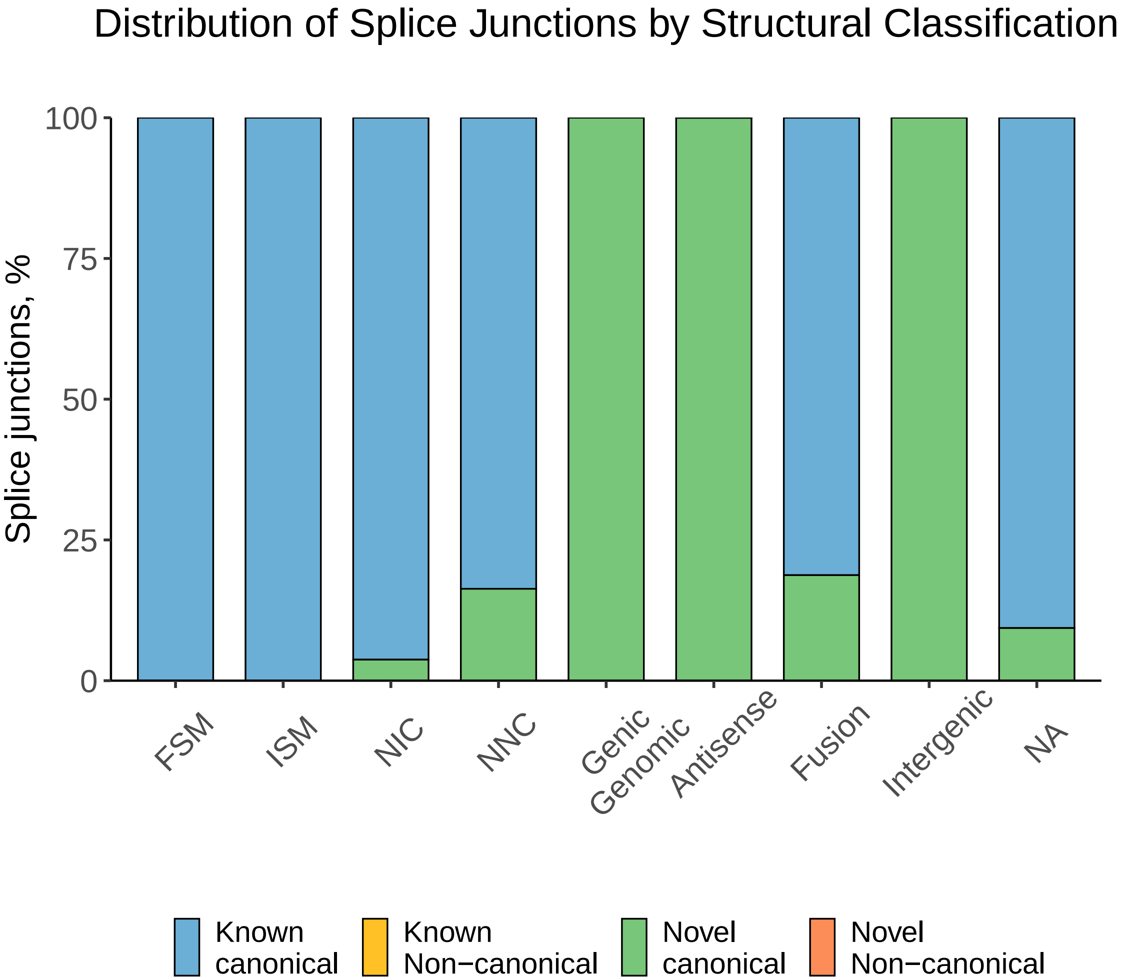


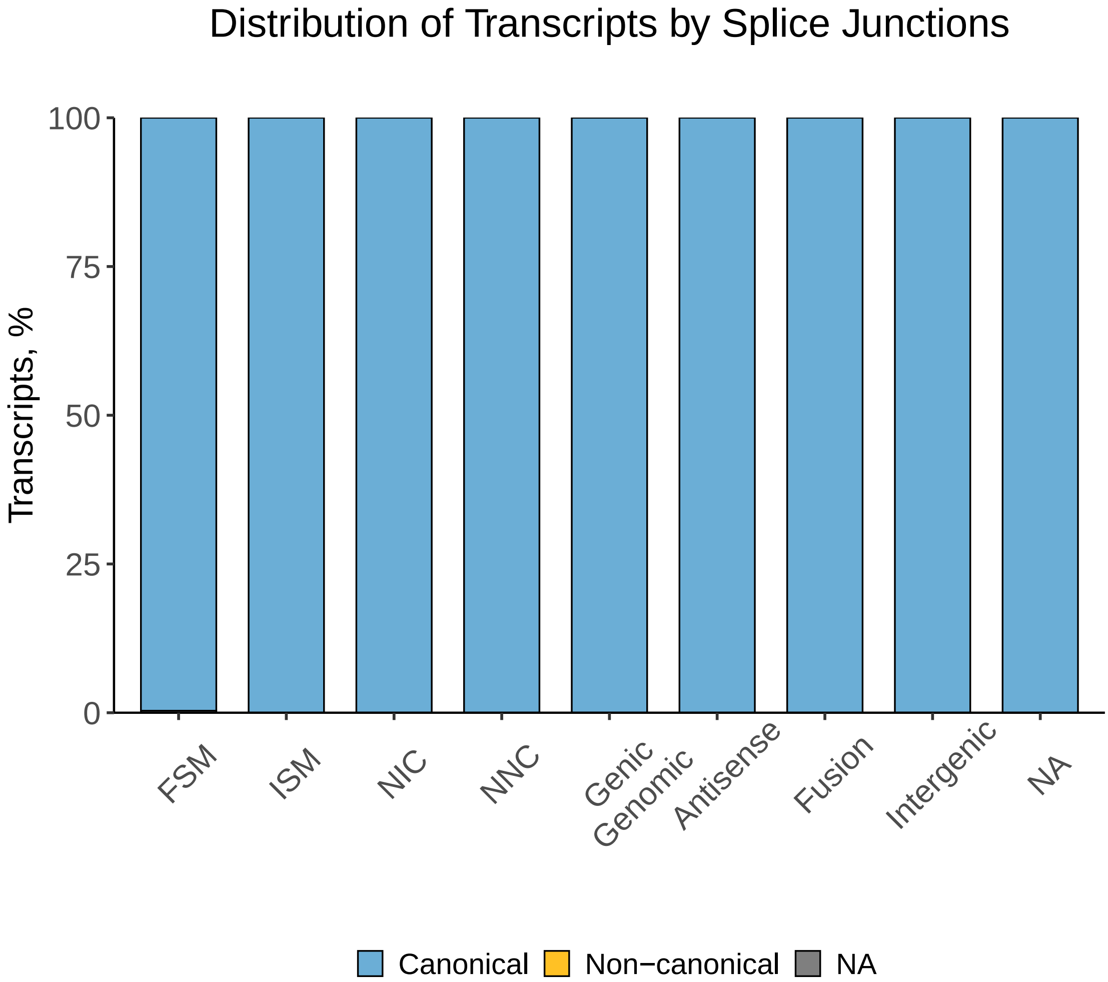
Exon Count Distribution of Matched Reference Transcripts

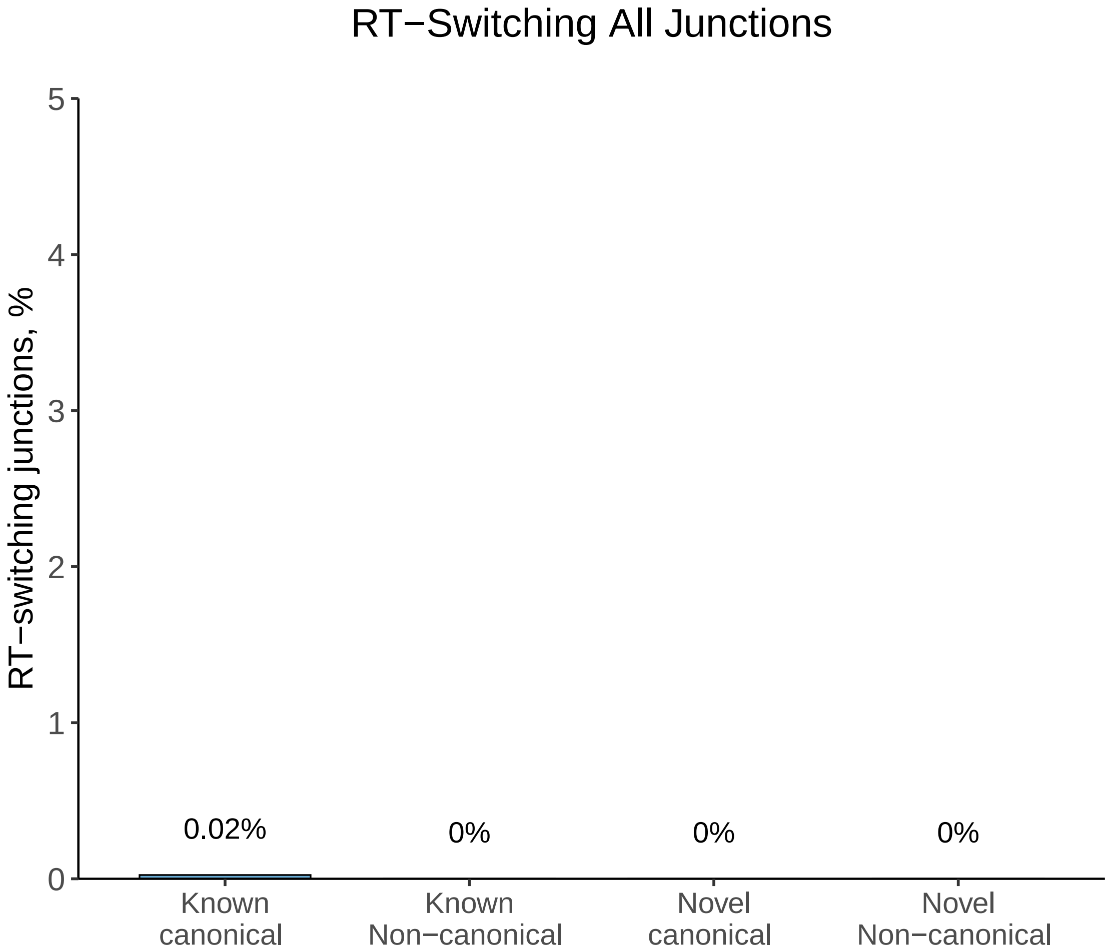
Applicable Only to FSM and ISM Categories

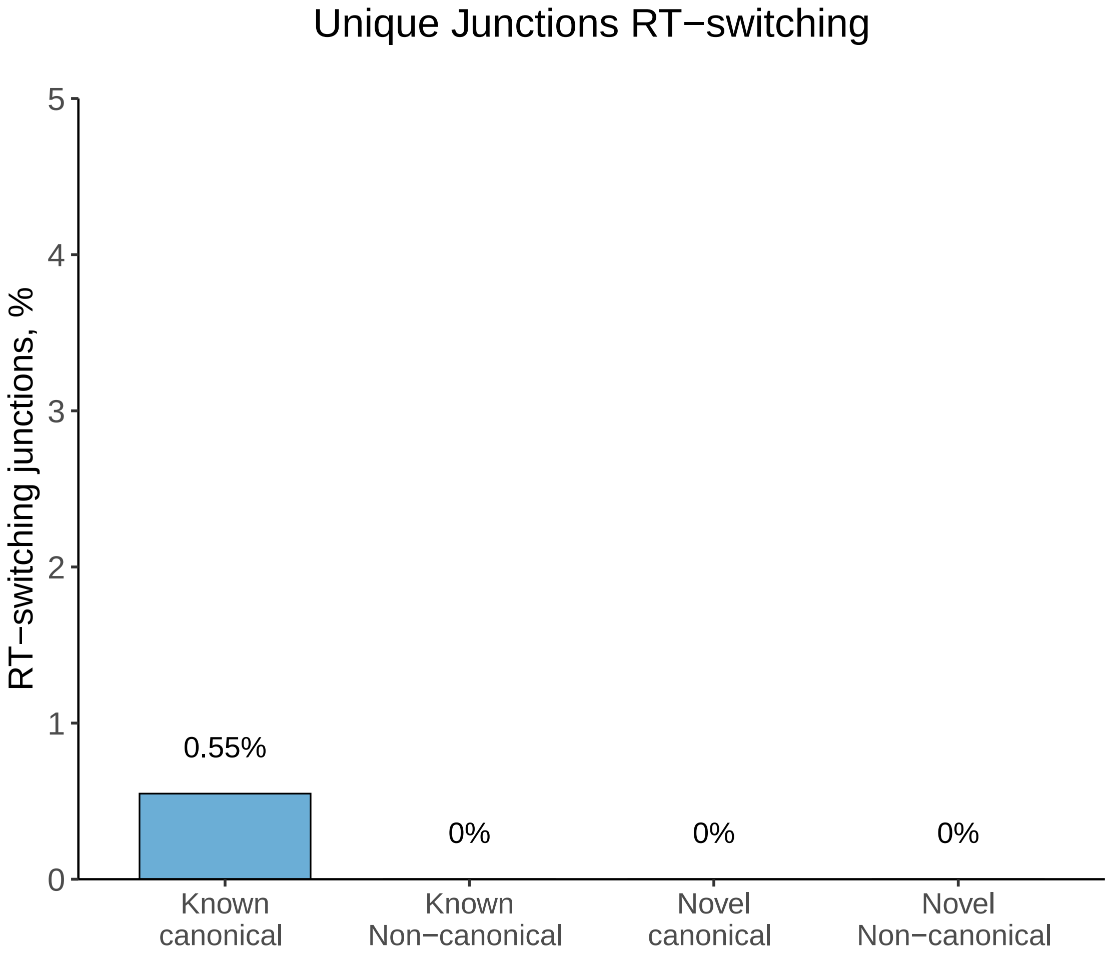


Splice Junction Characterization

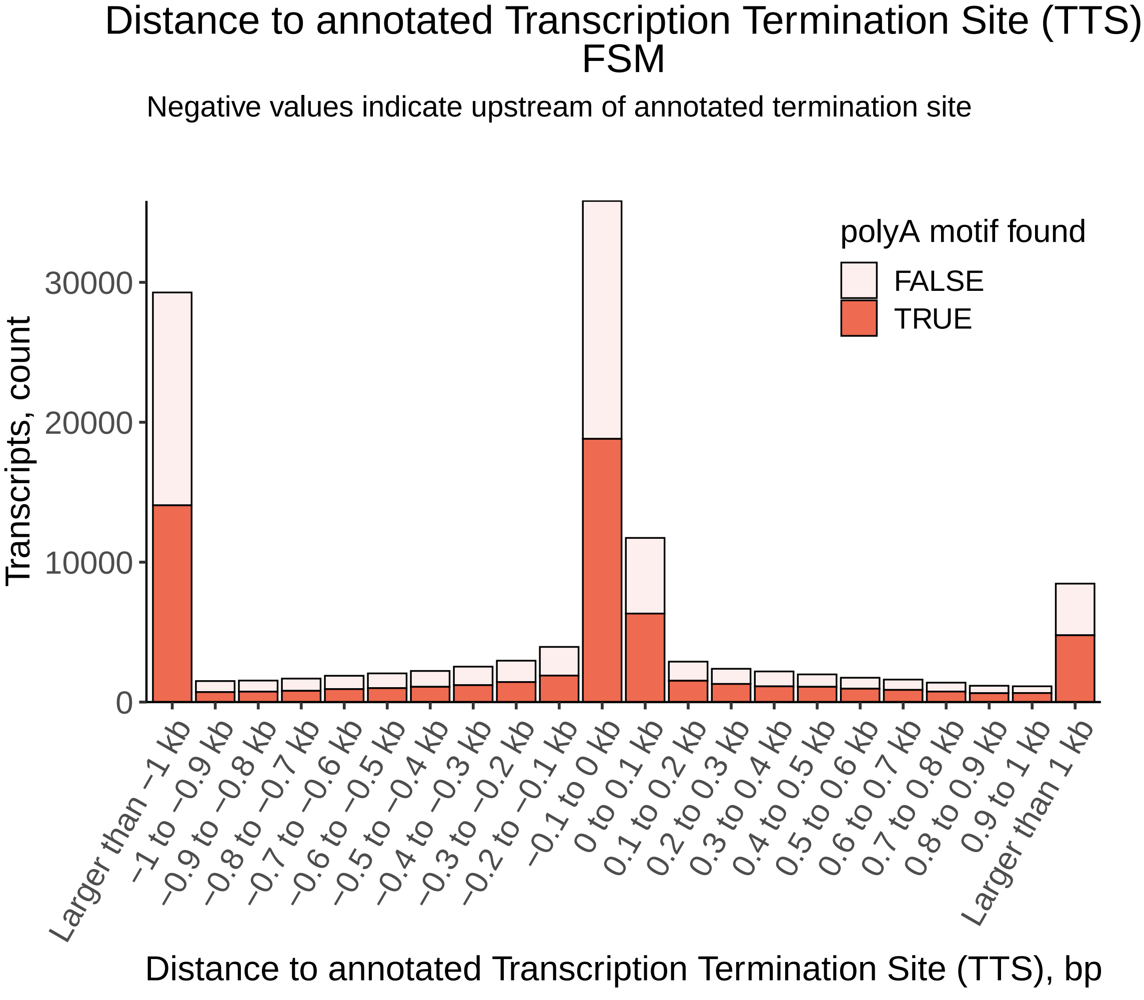
Distribution of Splice Junctions by Structural Classification
Splice junctions, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
NA
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

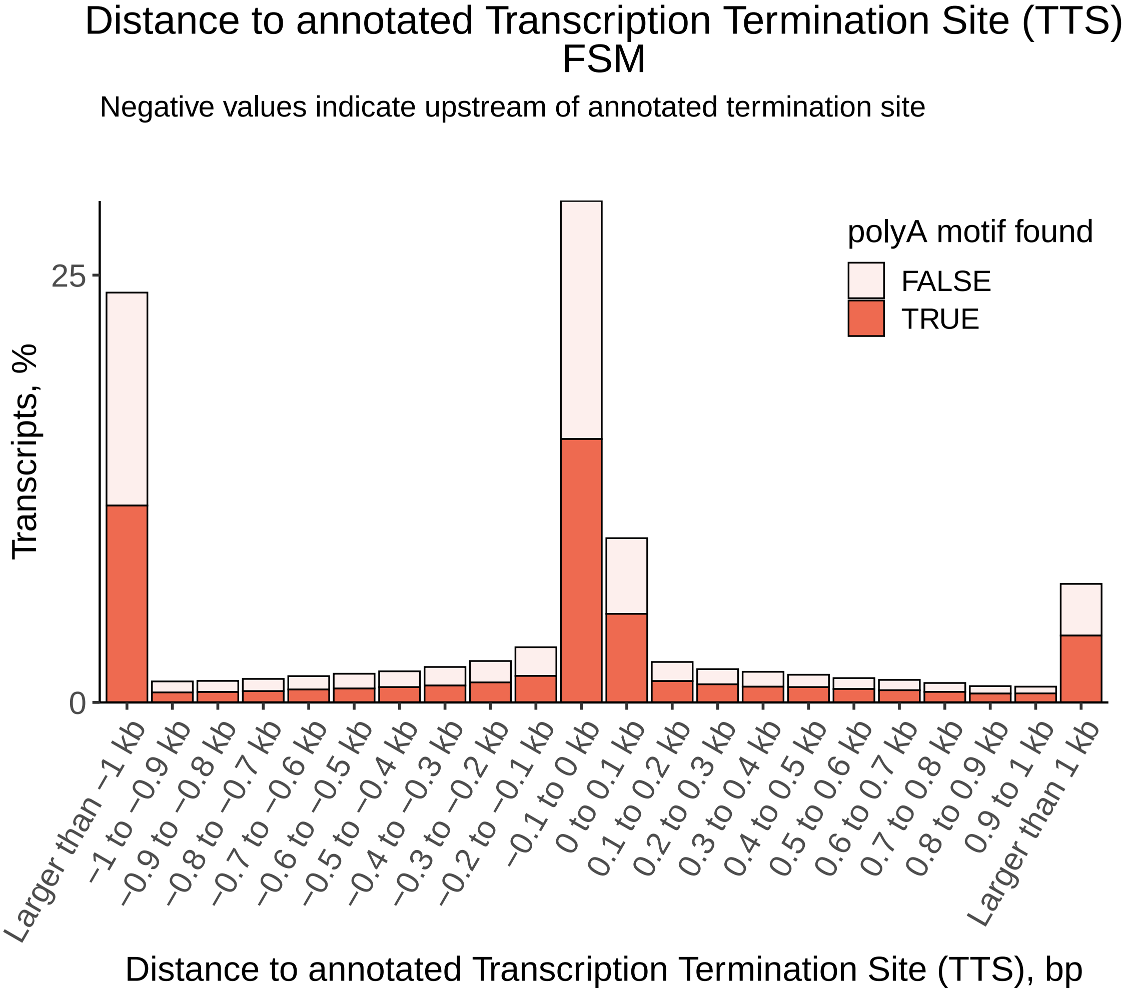

Distribution of Transcripts by Splice Junctions
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
NA
Canonical
Non−canonical
NA


RT−Switching All Junctions
RT−switching junctions, %
5
4
3
2
1
0
0.02%
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
0.55%
0%
0%
0%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

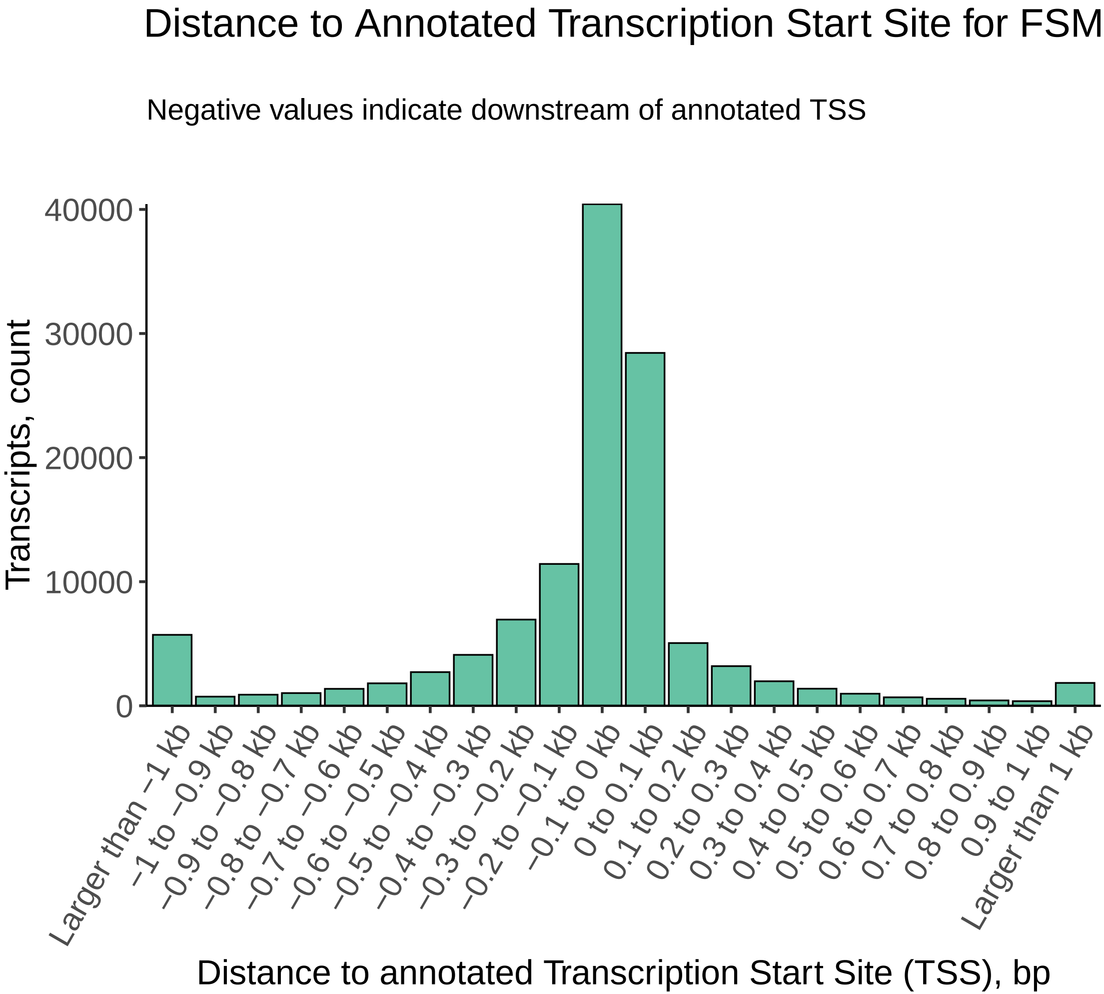
Comparison With Annotated TSS and TTS


Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
10000
20000
30000
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



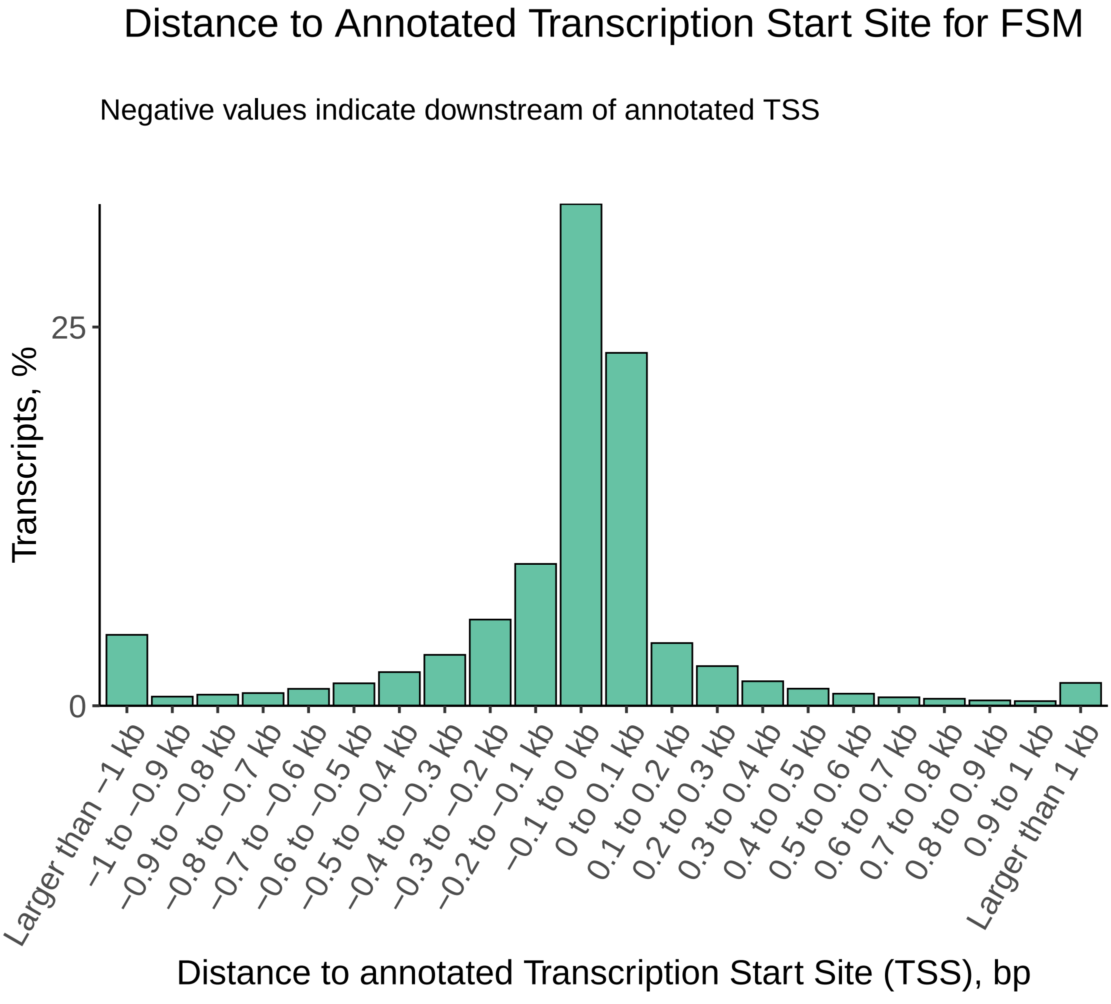
Distance to annotated Transcription Termination Site (TTS)
FSM

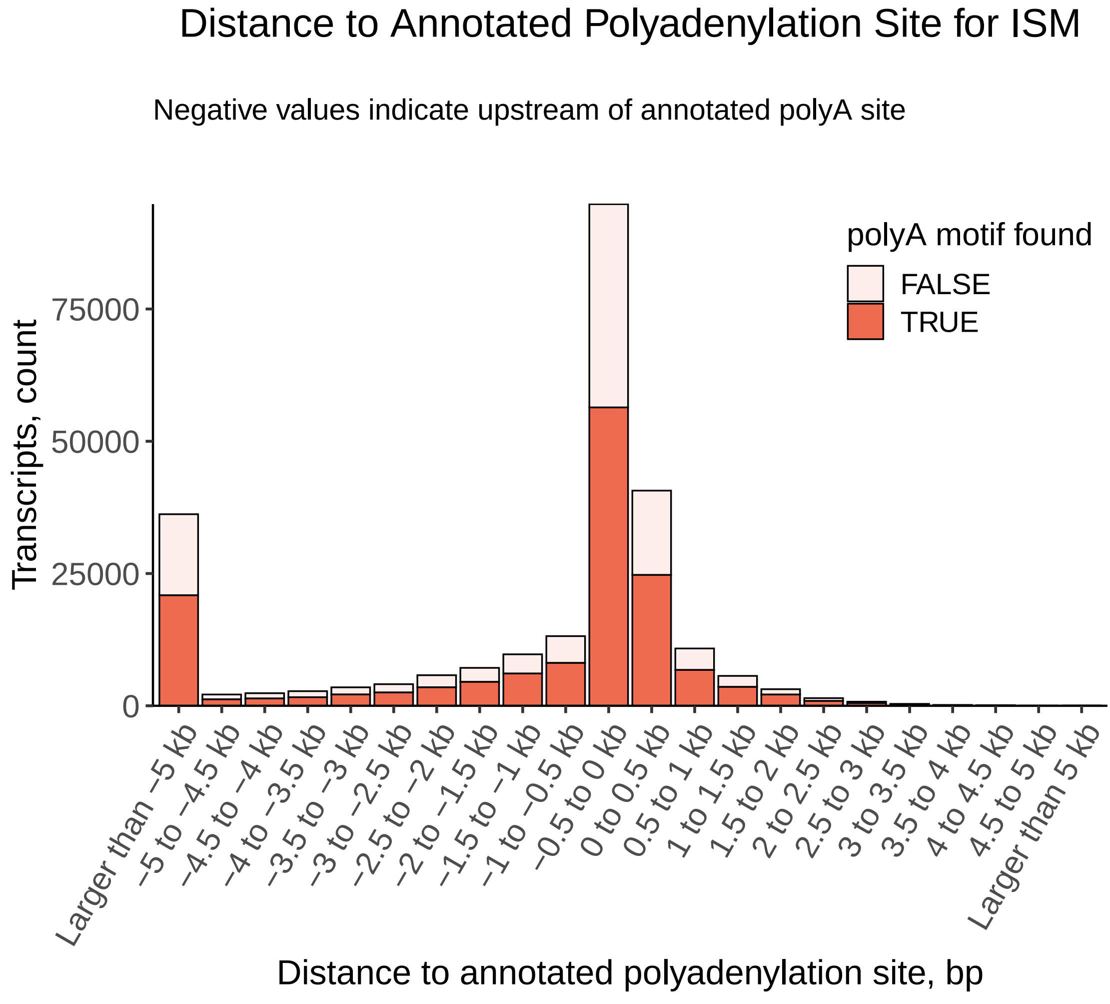
Negative values indicate upstream of annotated termination site

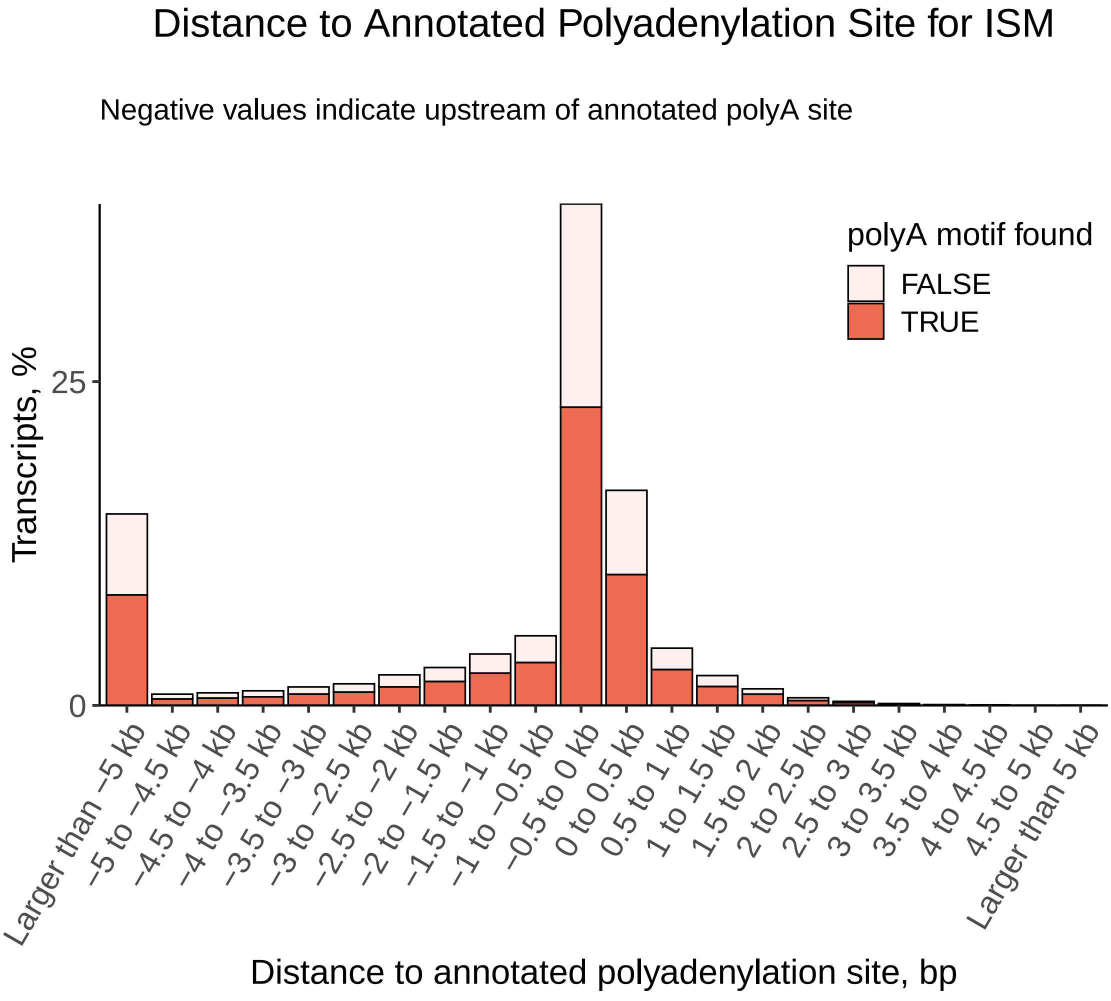


Distance to Annotated Transcription Start Site for FSM

Negative values indicate downstream of annotated TSS


Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
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

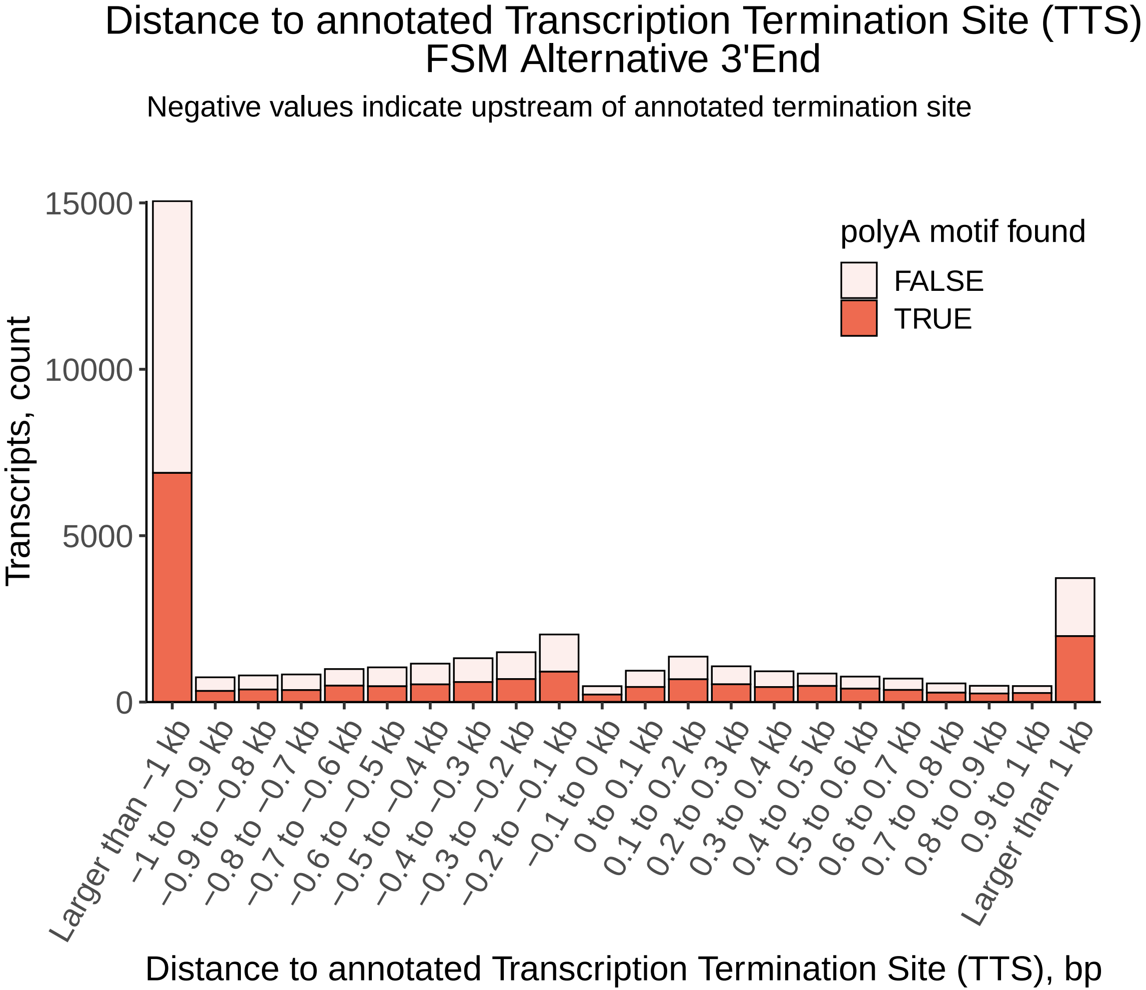

Distance to Annotated Polyadenylation Site for ISM
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
0
25000
50000
75000
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −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

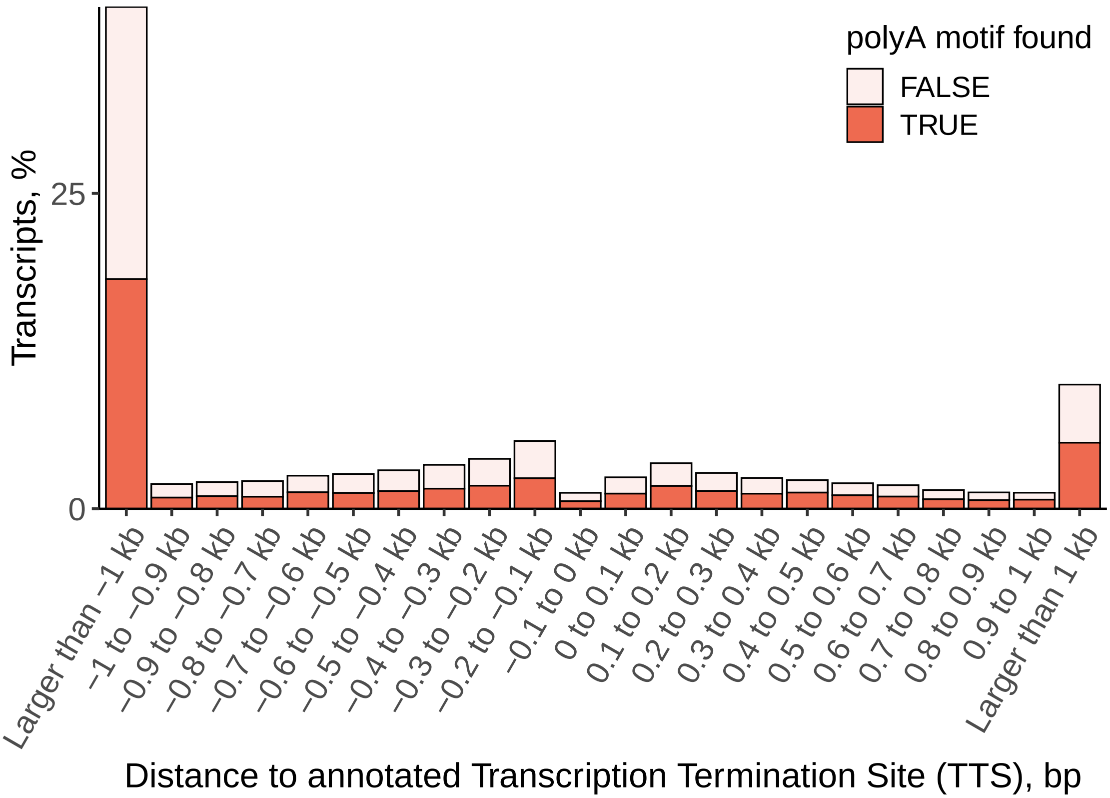
Negative values indicate upstream of annotated polyA site

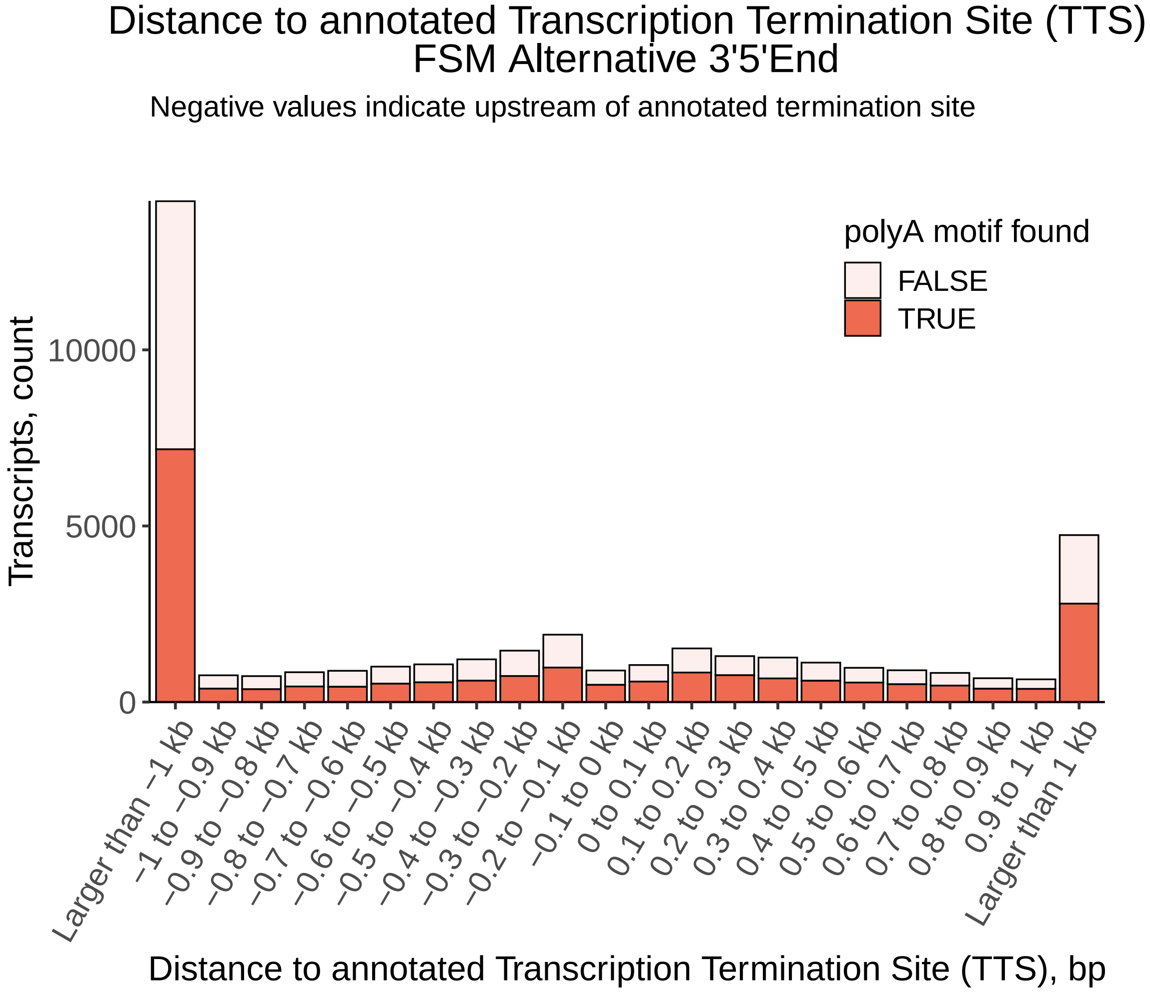
Comparison With Annotated TSS and TTS by Subcategories


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
15000
10000
5000
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

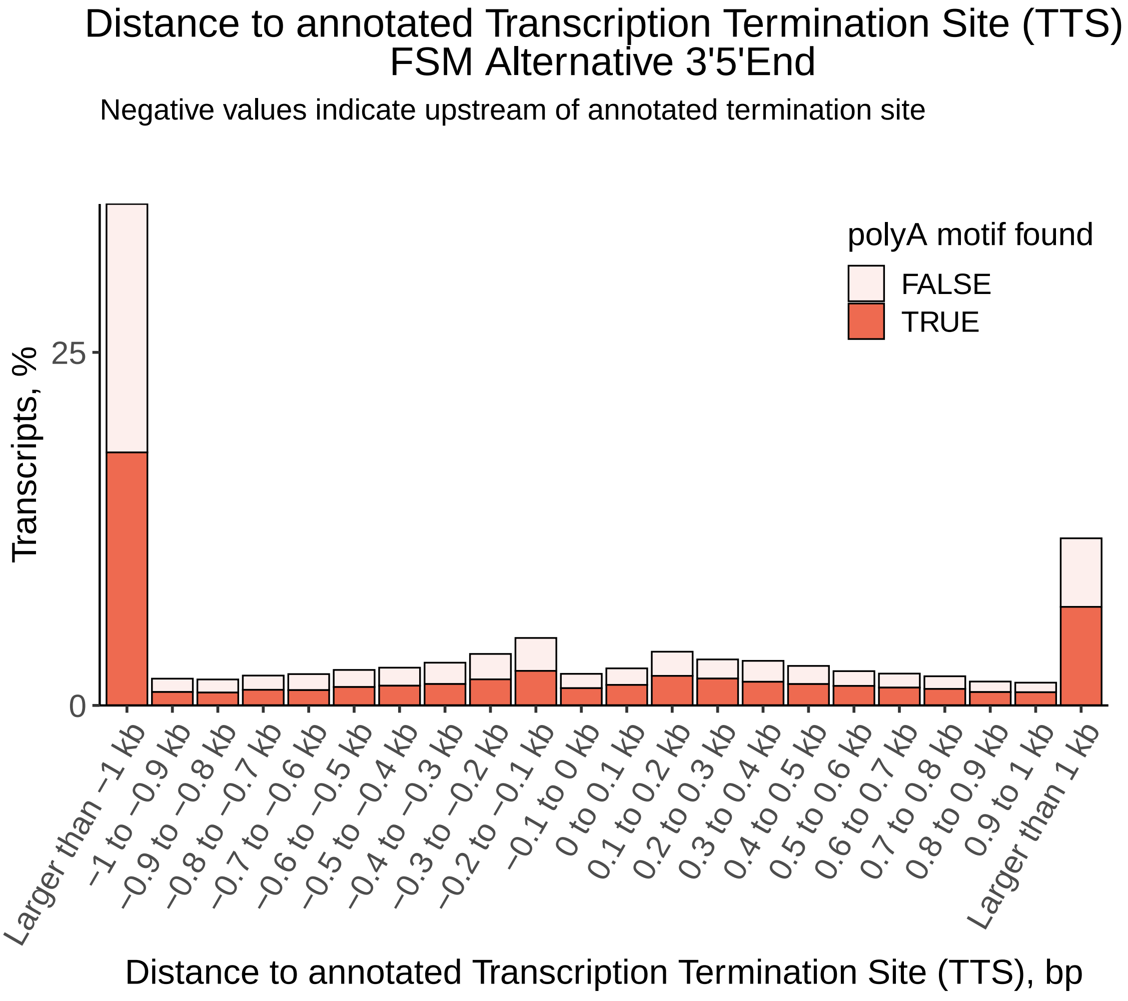
Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End

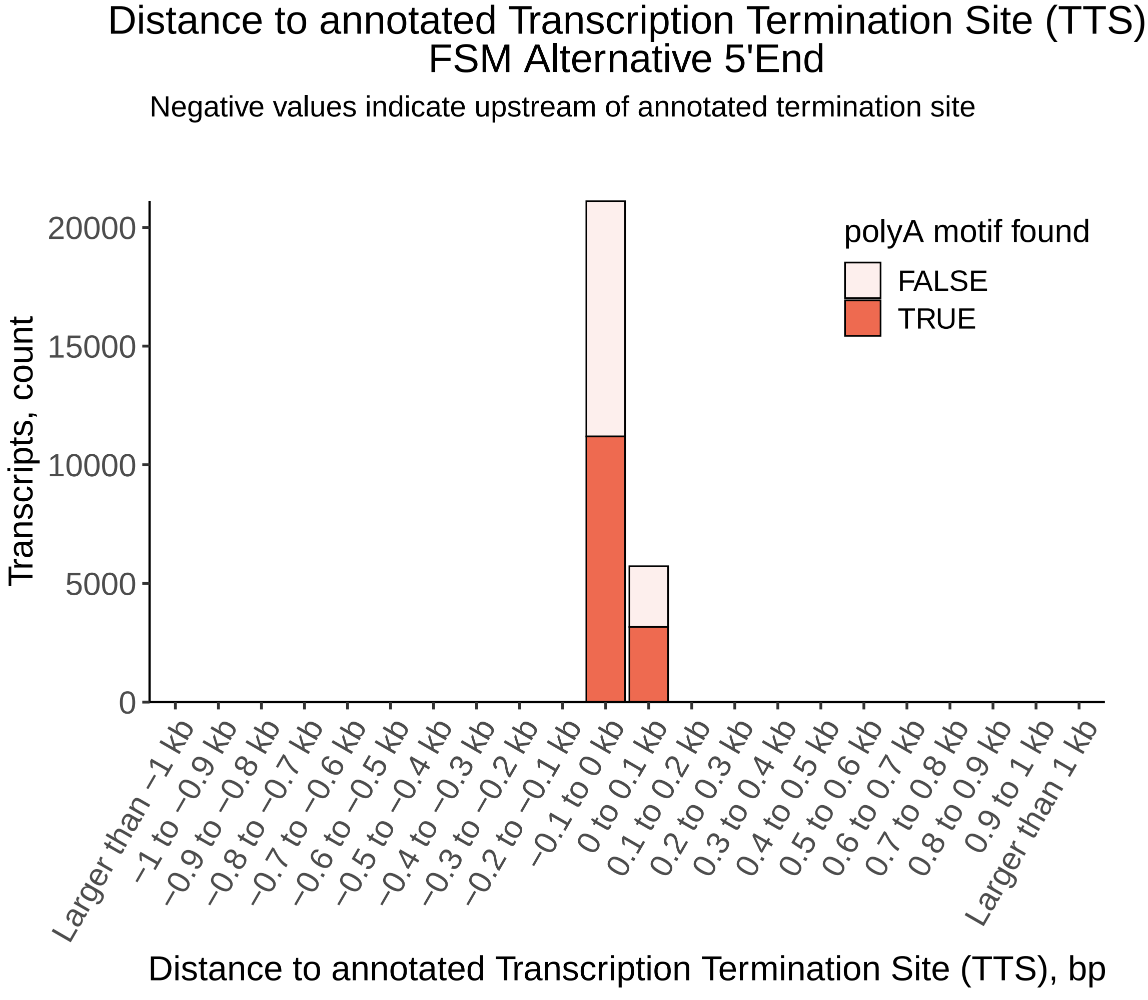
Negative values indicate upstream of annotated termination site

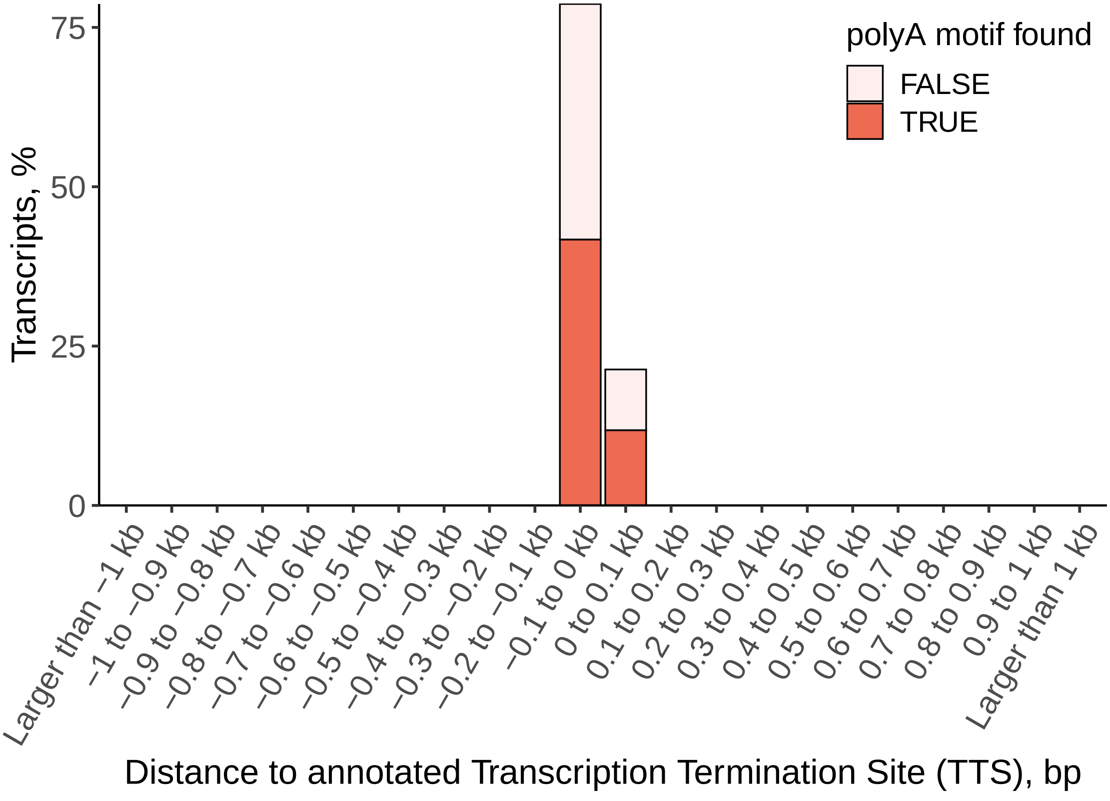



Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

Negative values indicate upstream of annotated termination site

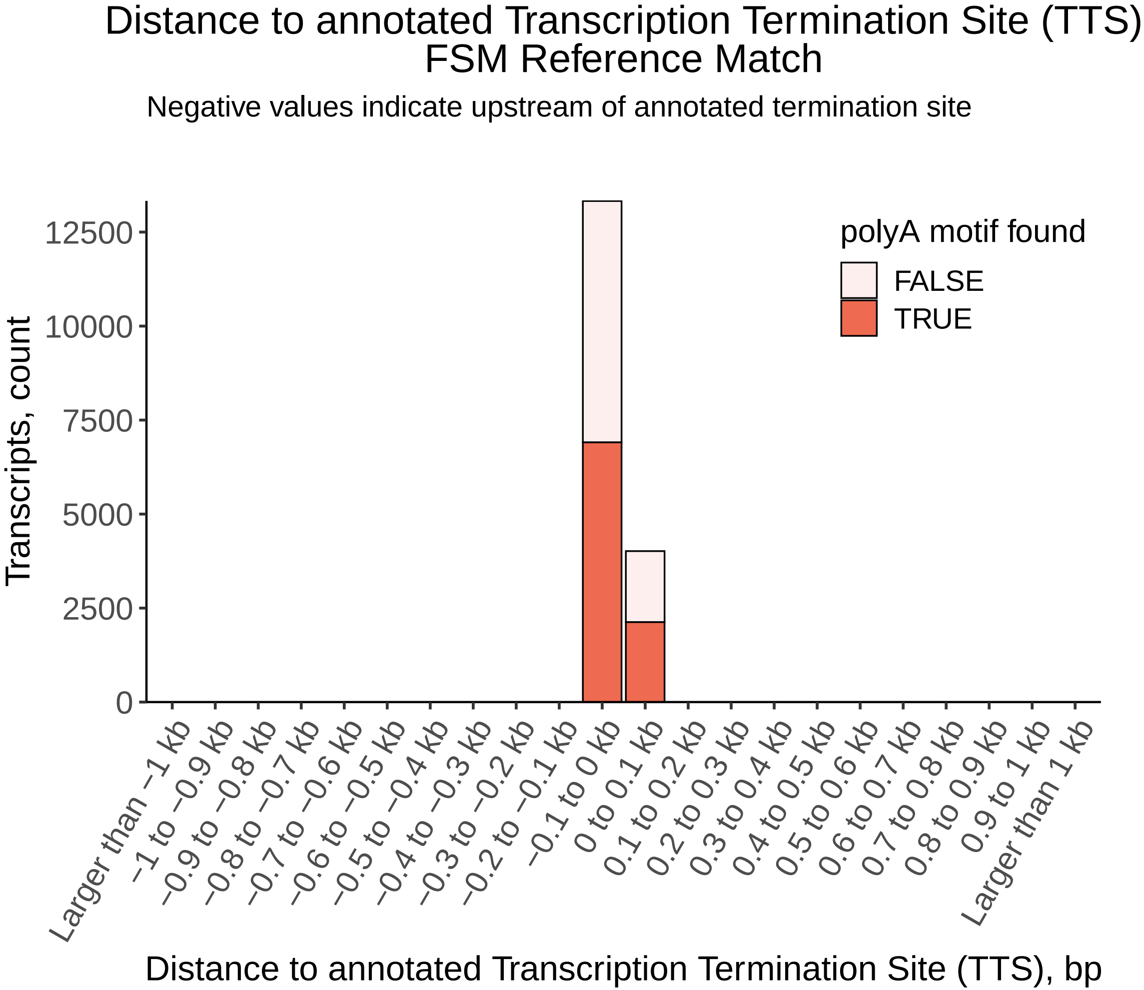

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 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

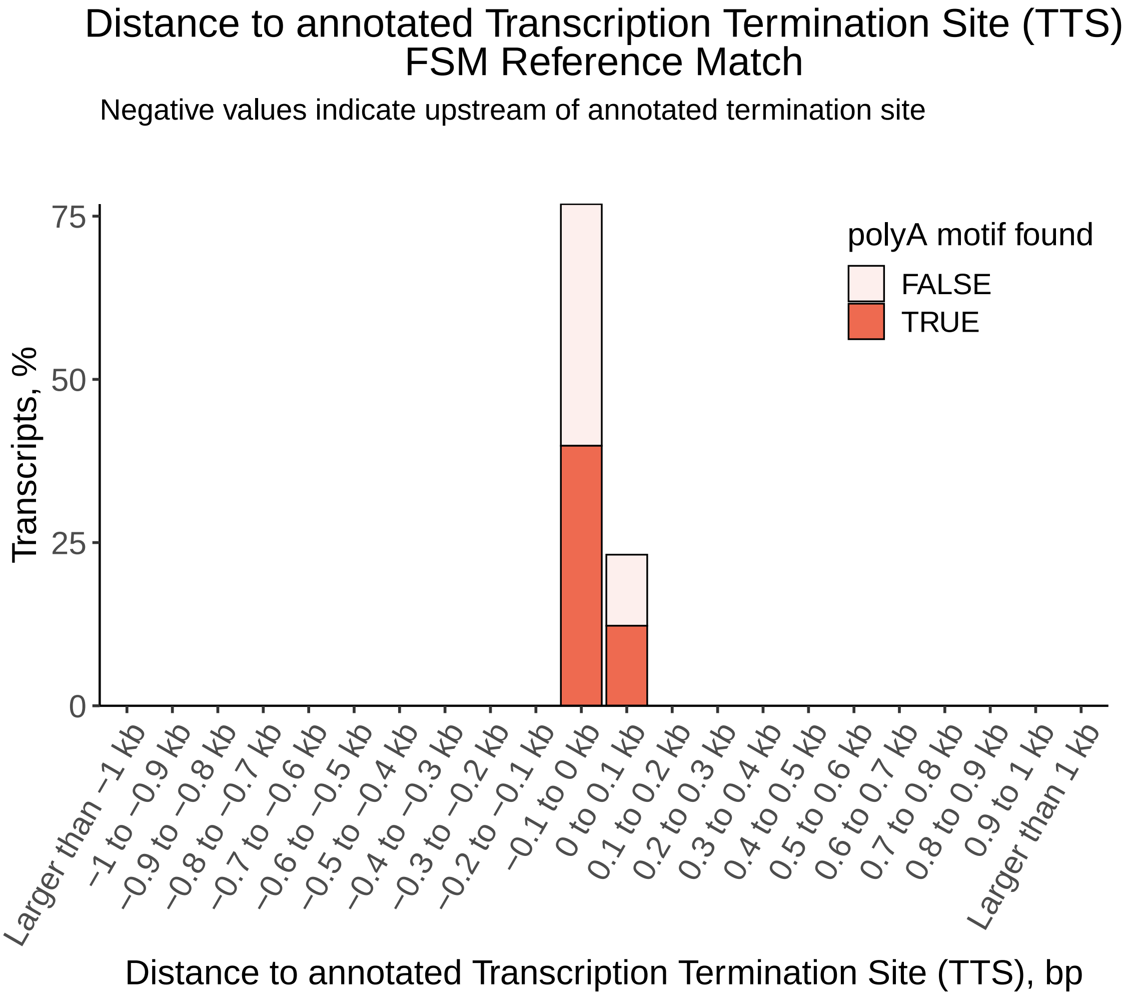

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
5000
10000
15000
20000
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6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

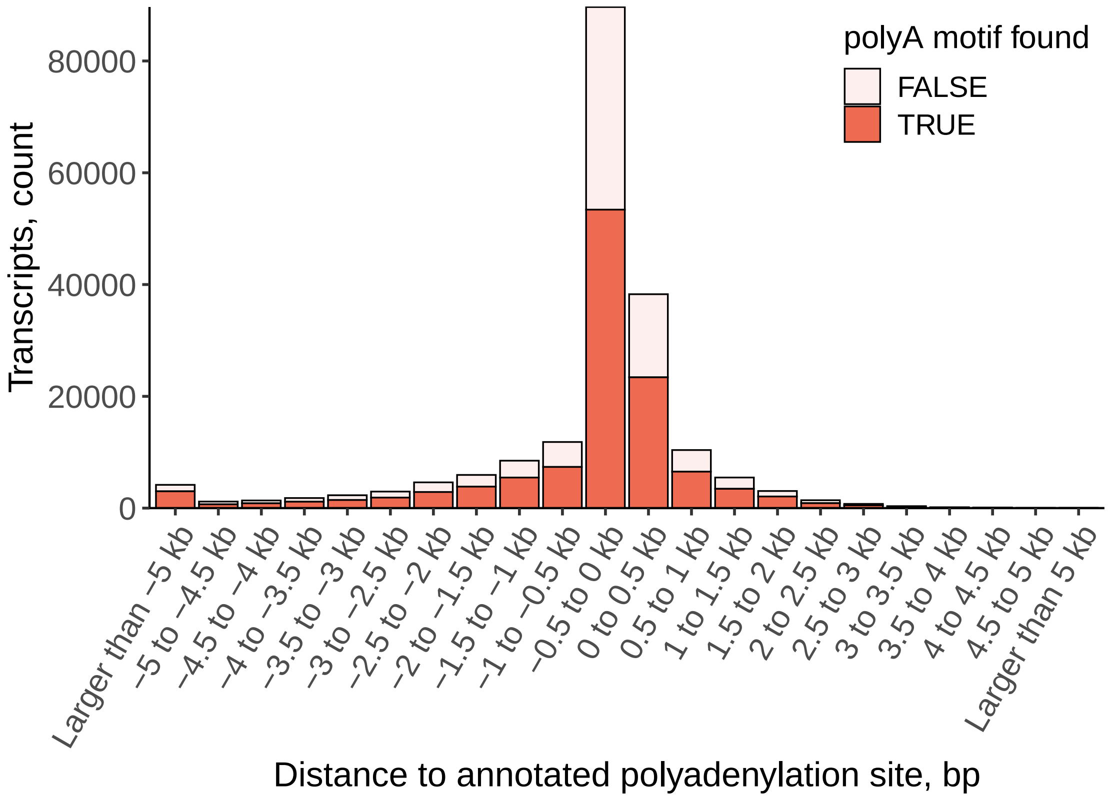


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End

Negative values indicate upstream of annotated termination site

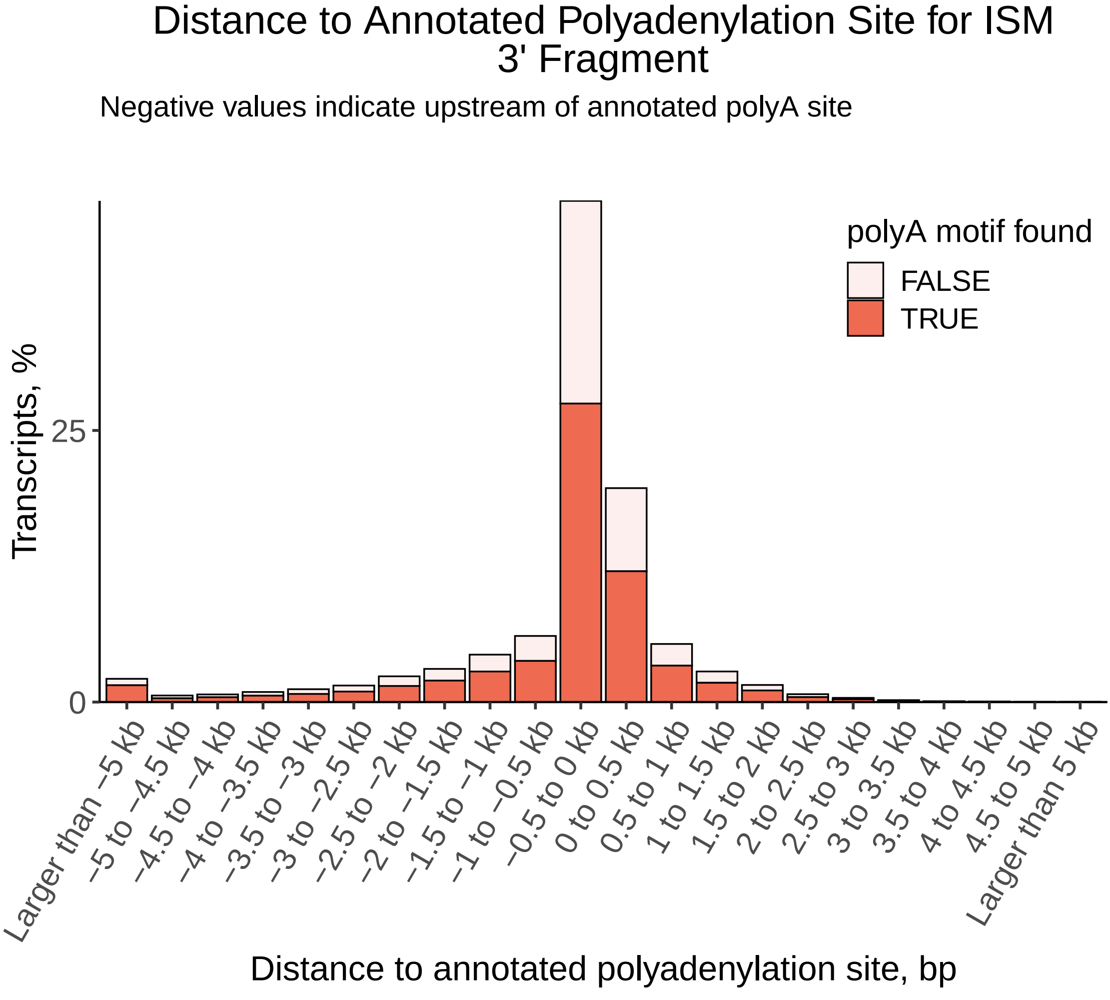

Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, count
12500
10000
7500
5000
2500
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 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

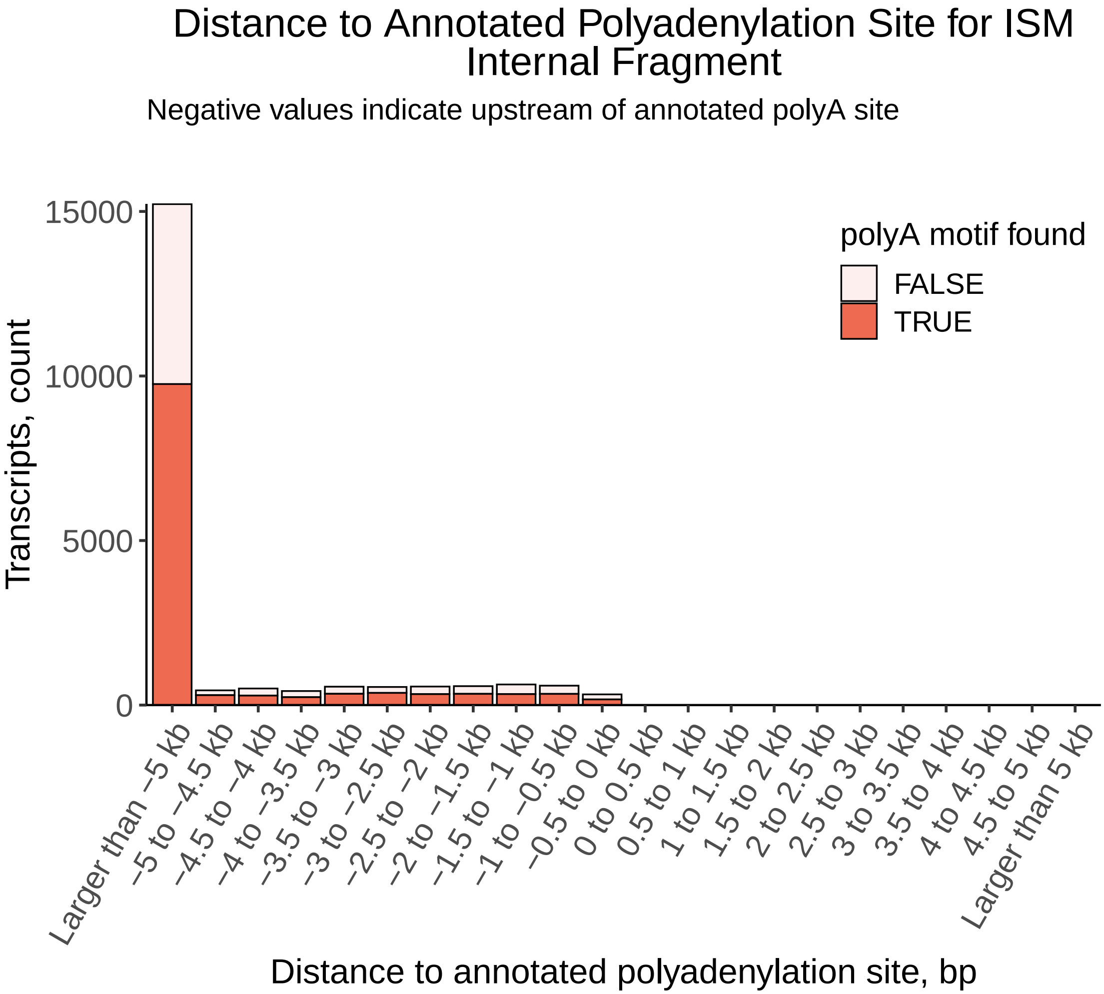

Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, %
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to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

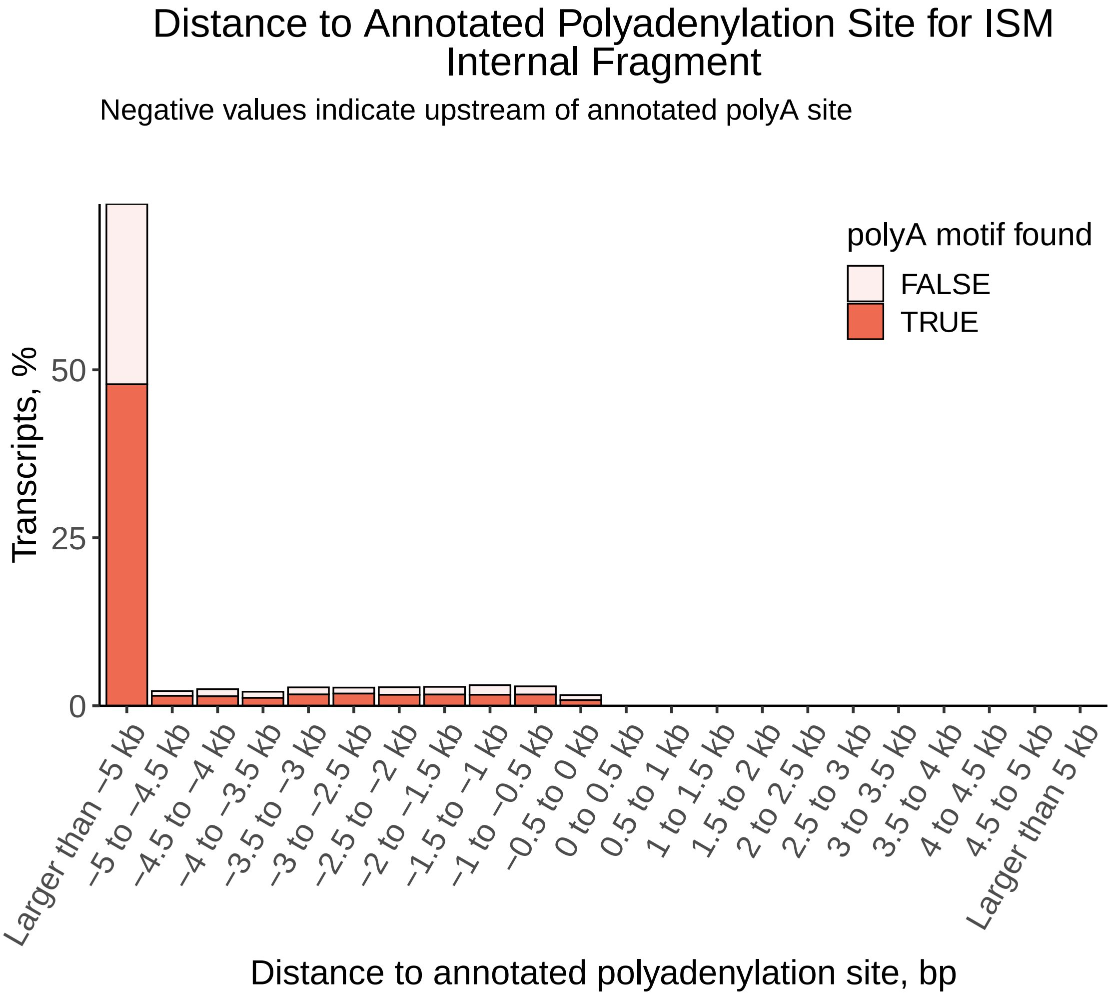
Negative values indicate upstream of annotated polyA site

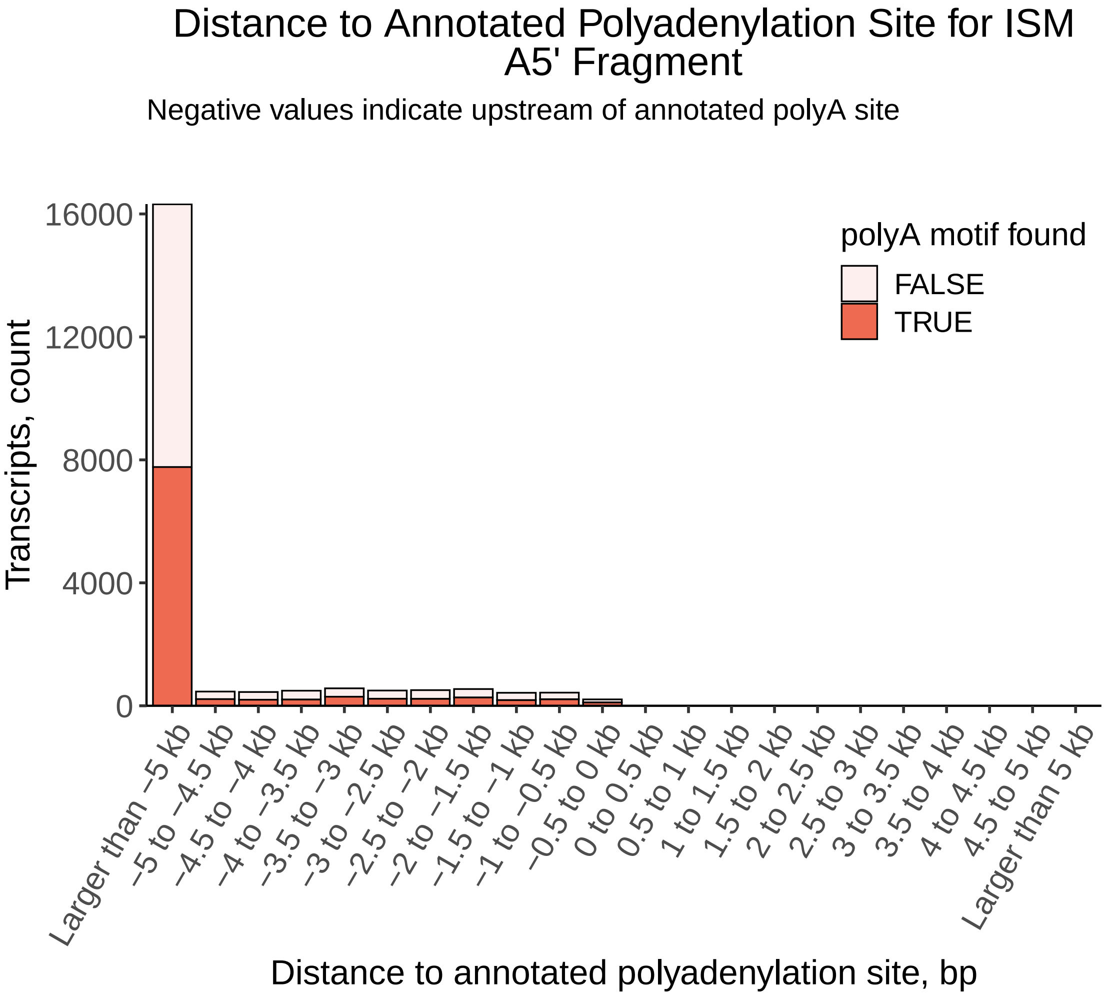

Distance to Annotated Polyadenylation Site for ISM 3' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1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

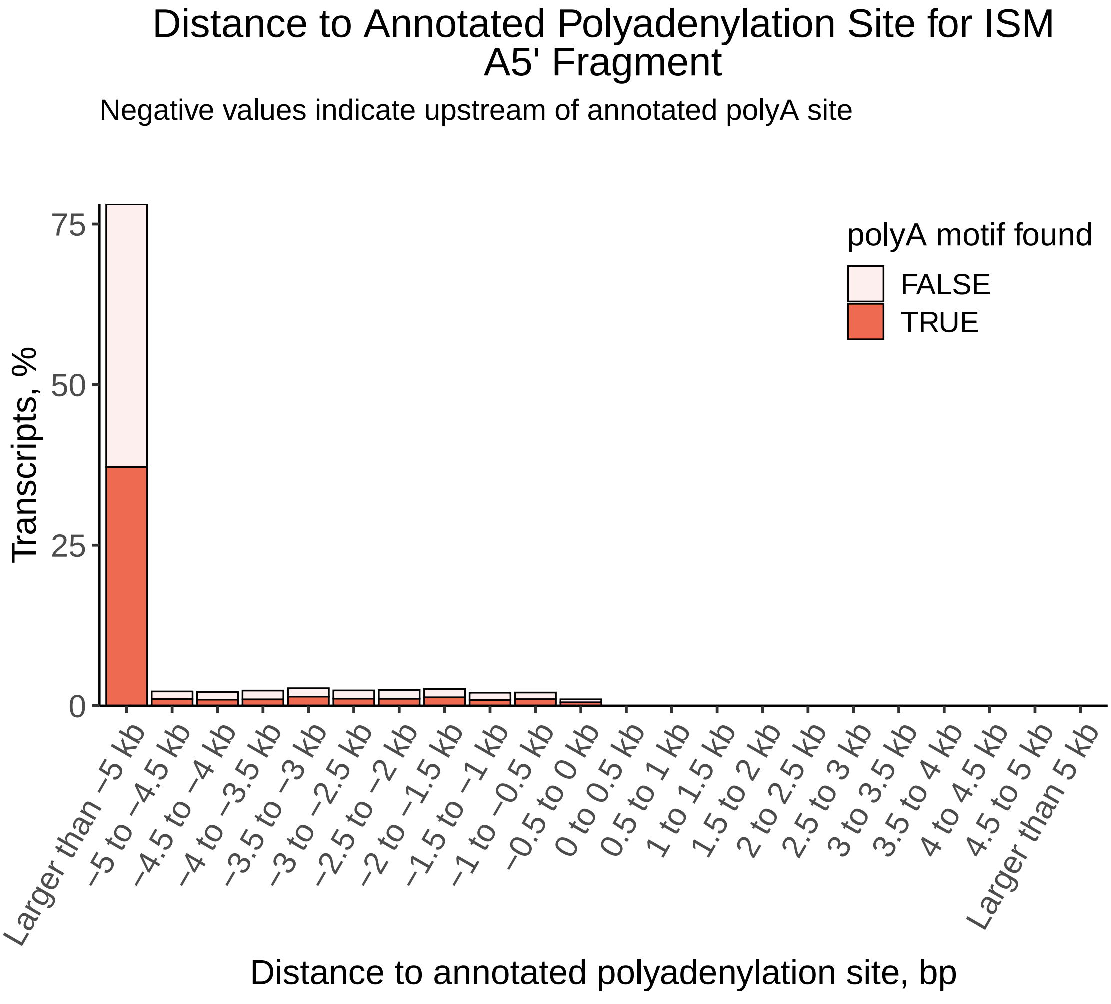


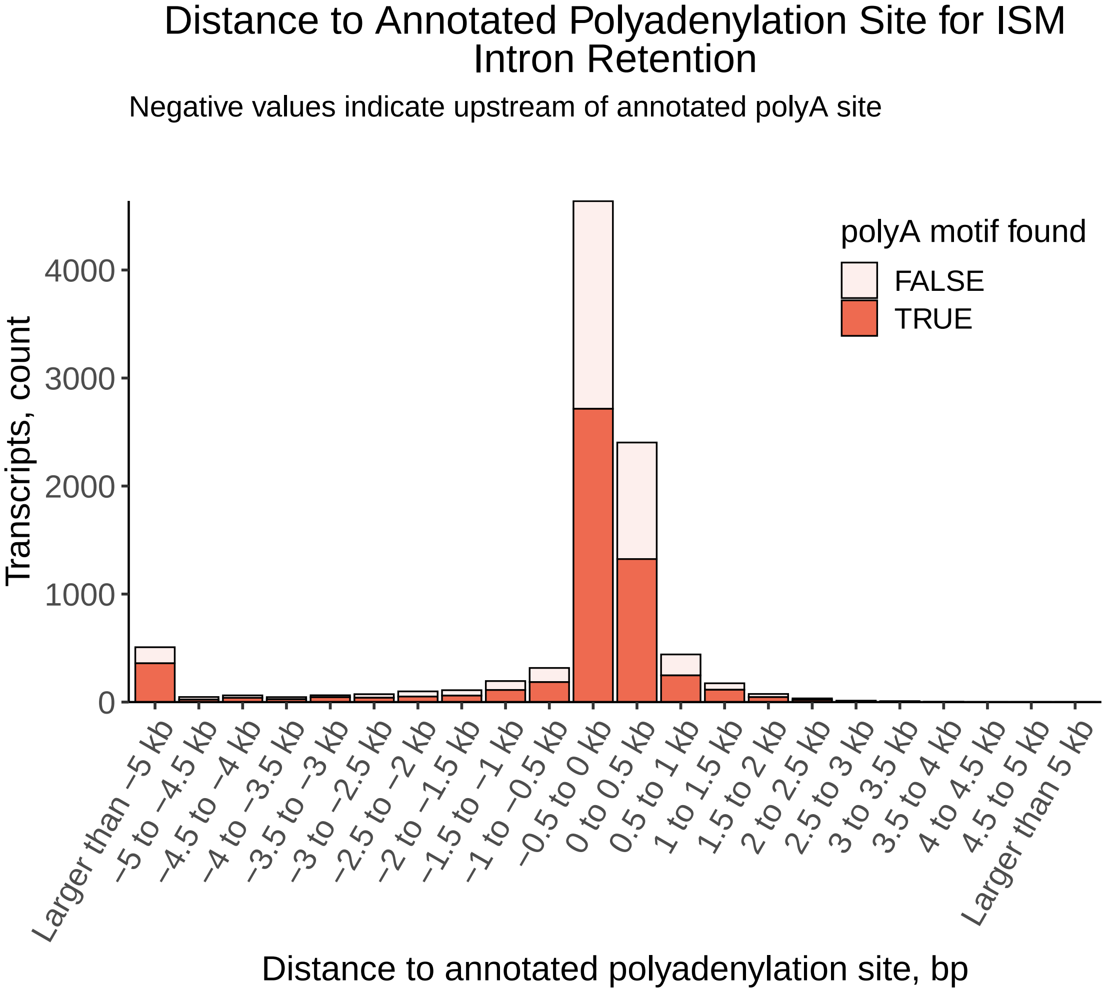
Distance to Annotated Polyadenylation Site for ISM Internal Fragment

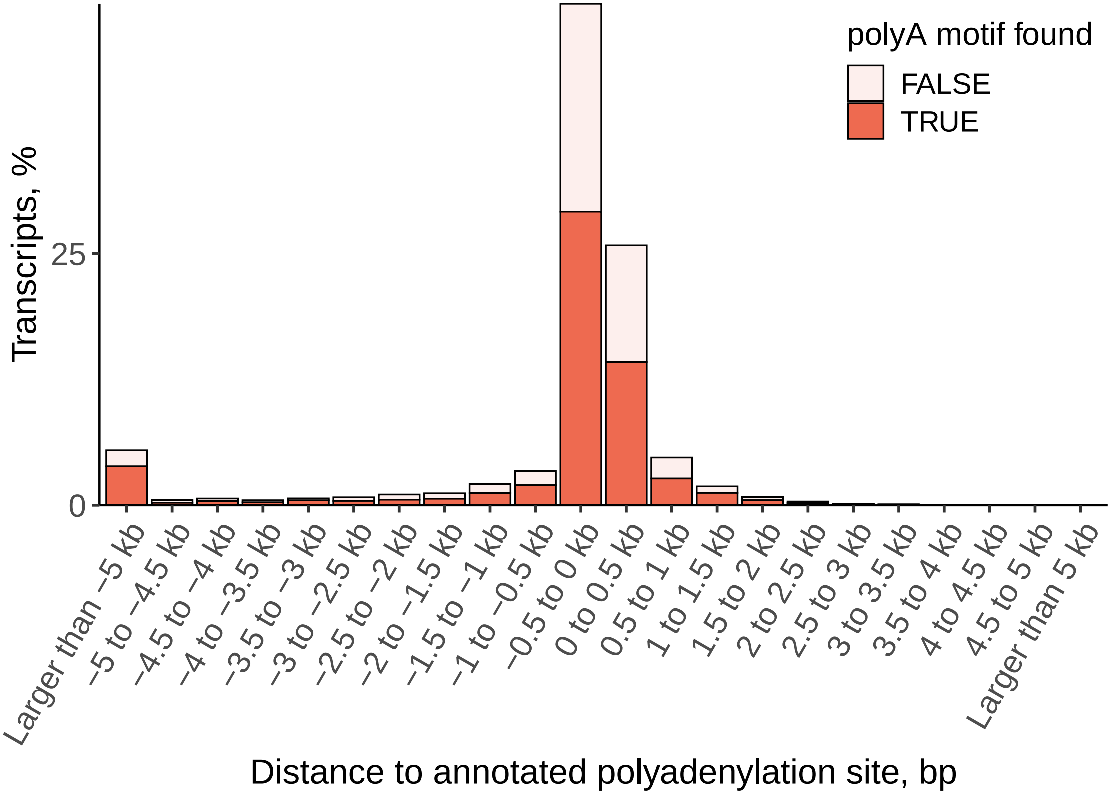
Negative values indicate upstream of annotated polyA site


Distance to Annotated Polyadenylation Site for ISM Internal Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
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
polyA motif found
FALSE
TRUE
Transcripts, count
16000
12000
8000
4000
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 A5' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
75
50
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

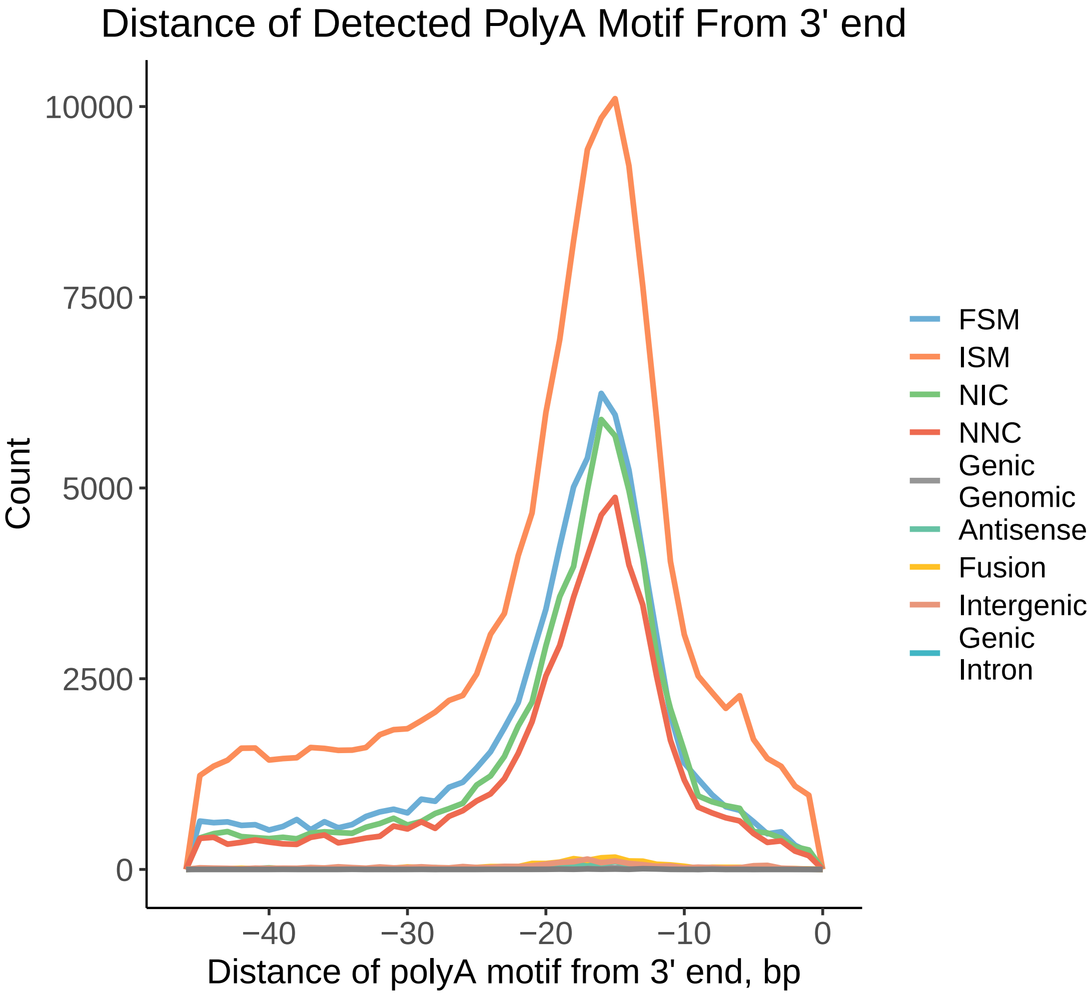

Distance to Annotated Polyadenylation Site for ISM Intron Retention
Negative values indicate upstream of annotated polyA site
Transcripts, count
0
1000
2000
3000
4000
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 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

PolyA Distance Analysis

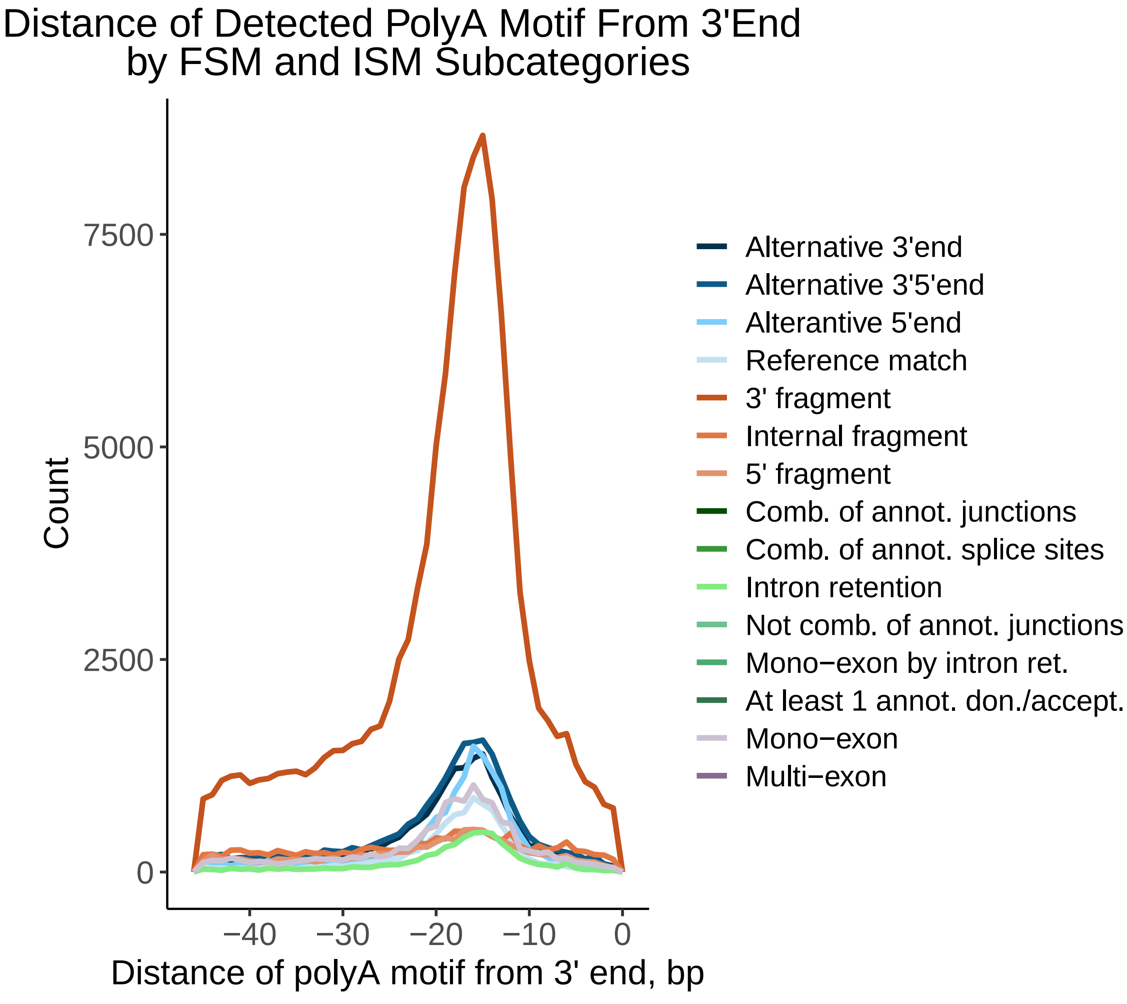

Distance of Detected PolyA Motif From 3' end
Count
10000
7500
5000
2500
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

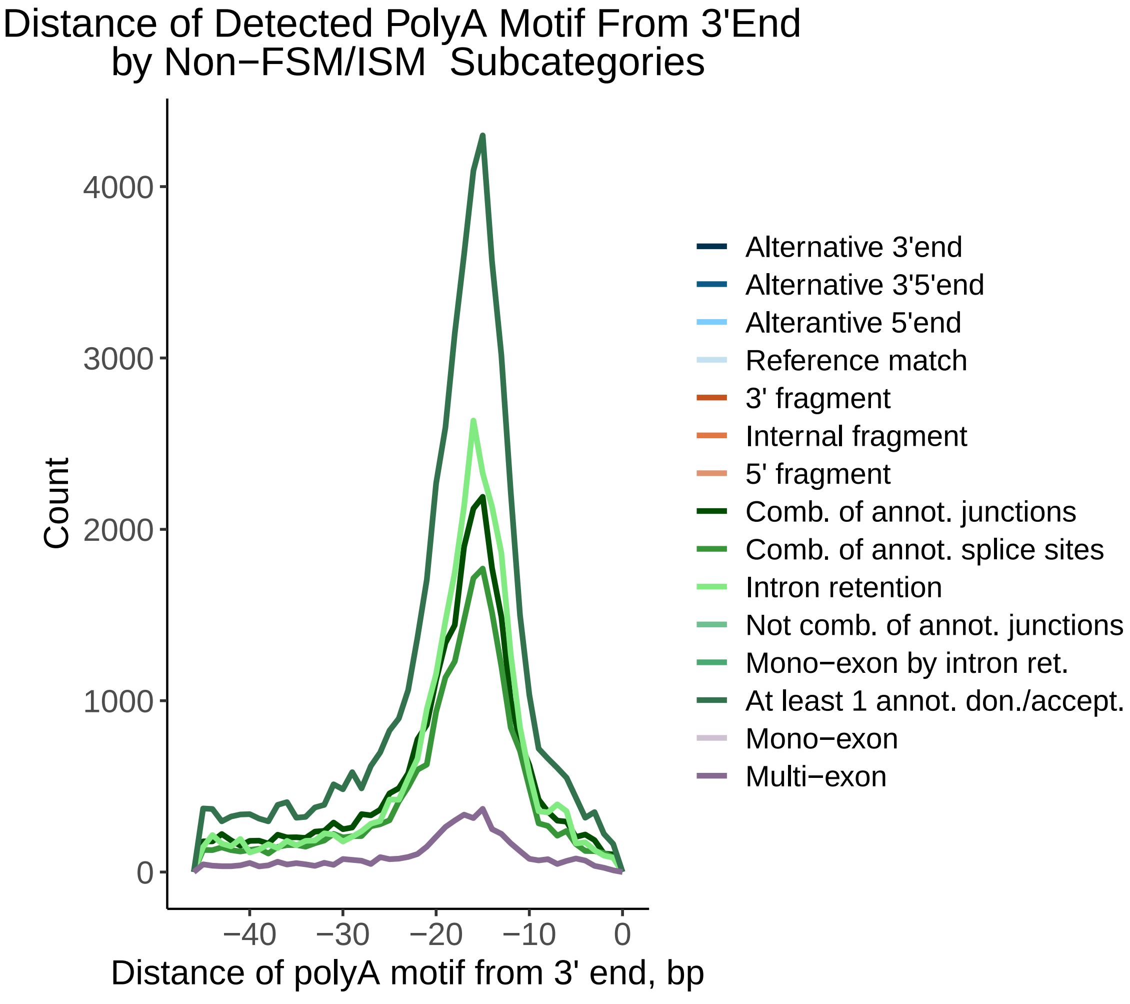
Frequency of PolyA Motifs

Motif	Count	%
AATAAA	169445	48.3
ATTAAA	50694	14.5
TATAAA	15253	4.4
AGTAAA	13528	3.9
AAGAAA	13496	3.9
AAAAAG	13448	3.8
TTTAAA	9974	2.8
GGGGCT	9612	2.7
AATATA	8184	2.3
AATACA	8076	2.3
CATAAA	7983	2.3
GATAAA	7817	2.2
AAAACA	7649	2.2
AATGAA	7042	2.0
AATAGA	4275	1.2
ACTAAA	4007	1.1

Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	143351	75785	53
ISM	244958	147457	60
NIC	126877	66656	53
NNC	102353	55639	54
Genic Genomic	608	354	58
Antisense	1812	1051	58
Fusion	3274	1823	56
Intergenic	2846	1667	59
NA	82	51	62


Distance of Detected PolyA Motif From 3'End
by FSM and ISM Subcategories
Count
7500
5000
2500
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


Distance of Detected PolyA Motif From 3'End
by Non−FSM/ISM Subcategories
Count
4000
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

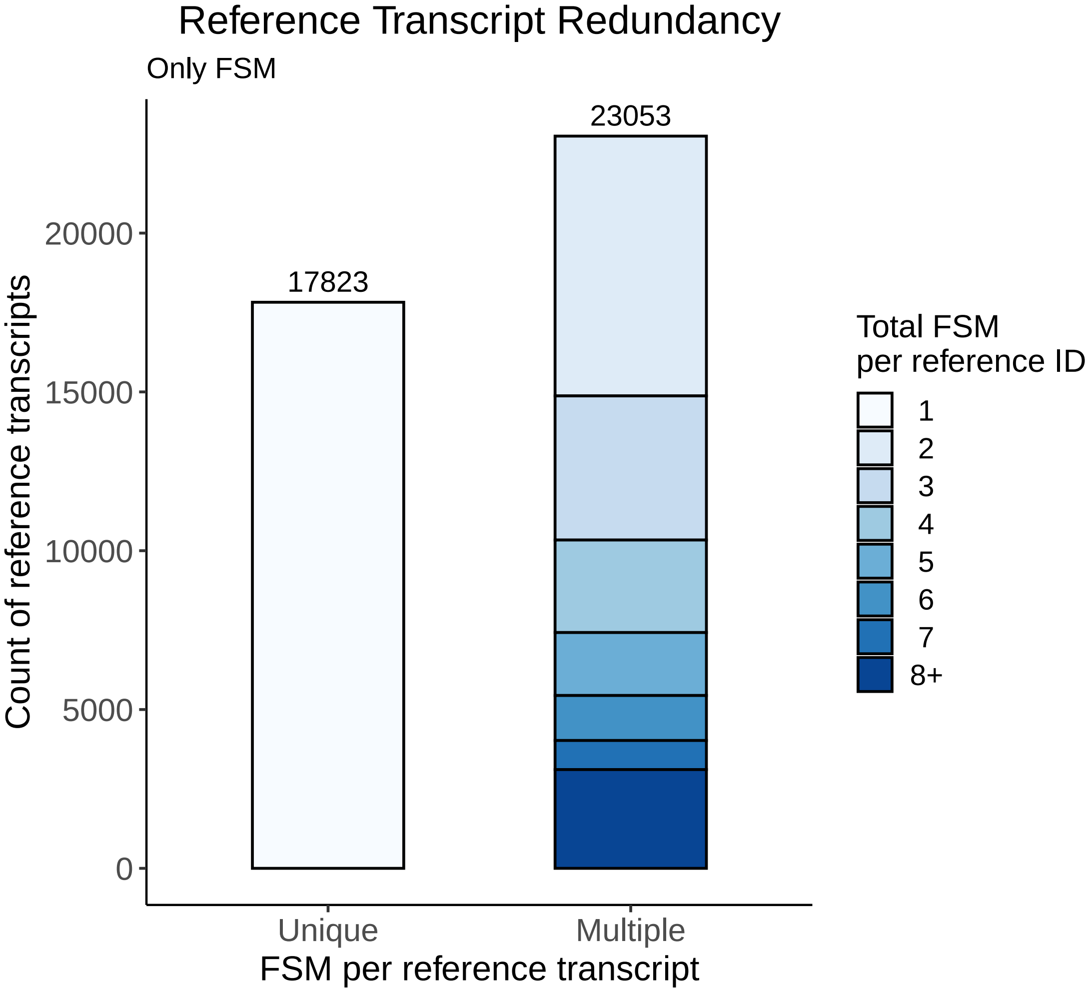
Number of polyA Motifs Detected

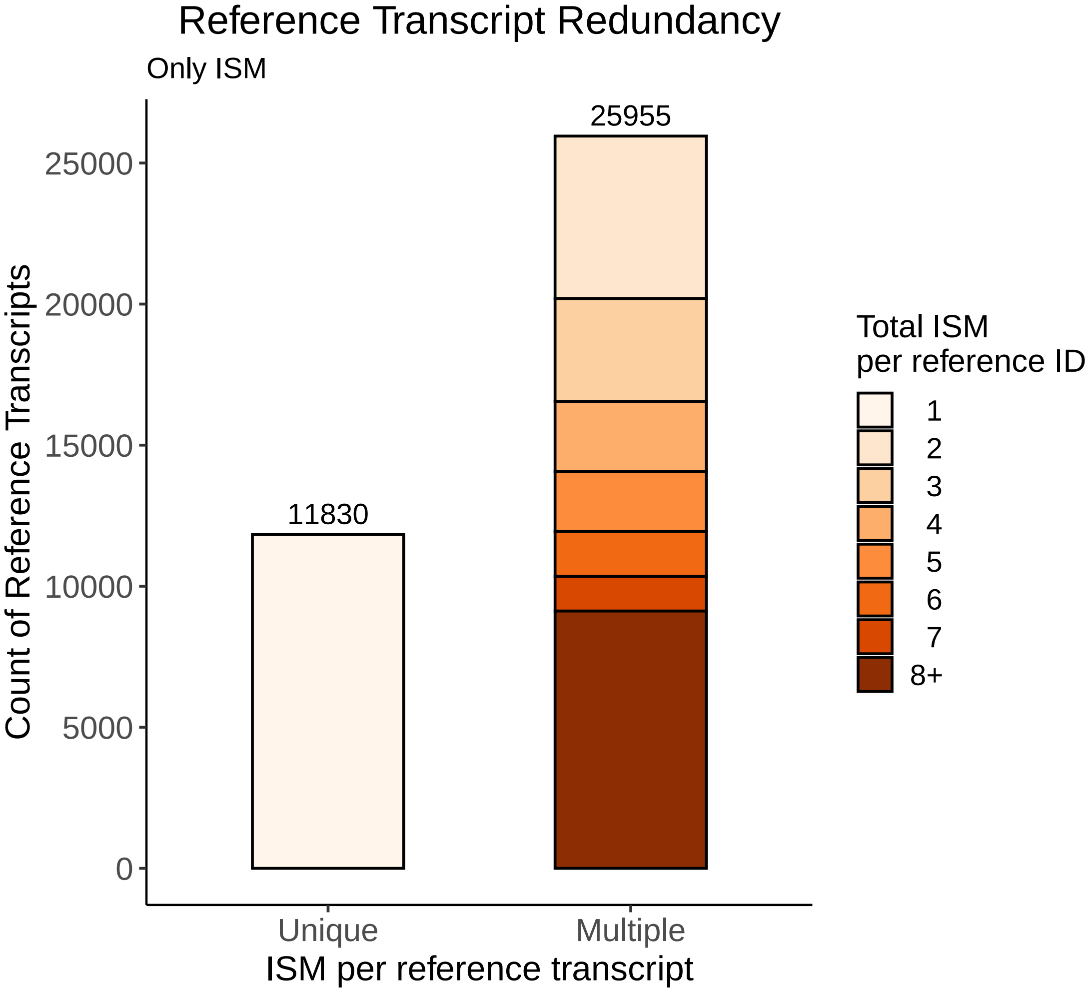
Subcategory	Count	polyA Detected	%
Alternative 3'end	37842	18066	48
Alternative 3'5'end	40055	21233	53
Alterantive 5'end	26837	14356	53
Reference match	17342	9033	52
3' fragment	194362	119314	61
Internal fragment	20391	12816	63
5' fragment	20894	9900	47
Comb. of annot. junctions	49758	25688	52
Comb. of annot. splice sites	37849	20434	54
Intron retention	61087	32460	53
At least 1 annot. don./accept.	90442	49457	55
Mono−exon	21275	13097	62
Multi−exon	7957	4589	58
NA	70	40	57

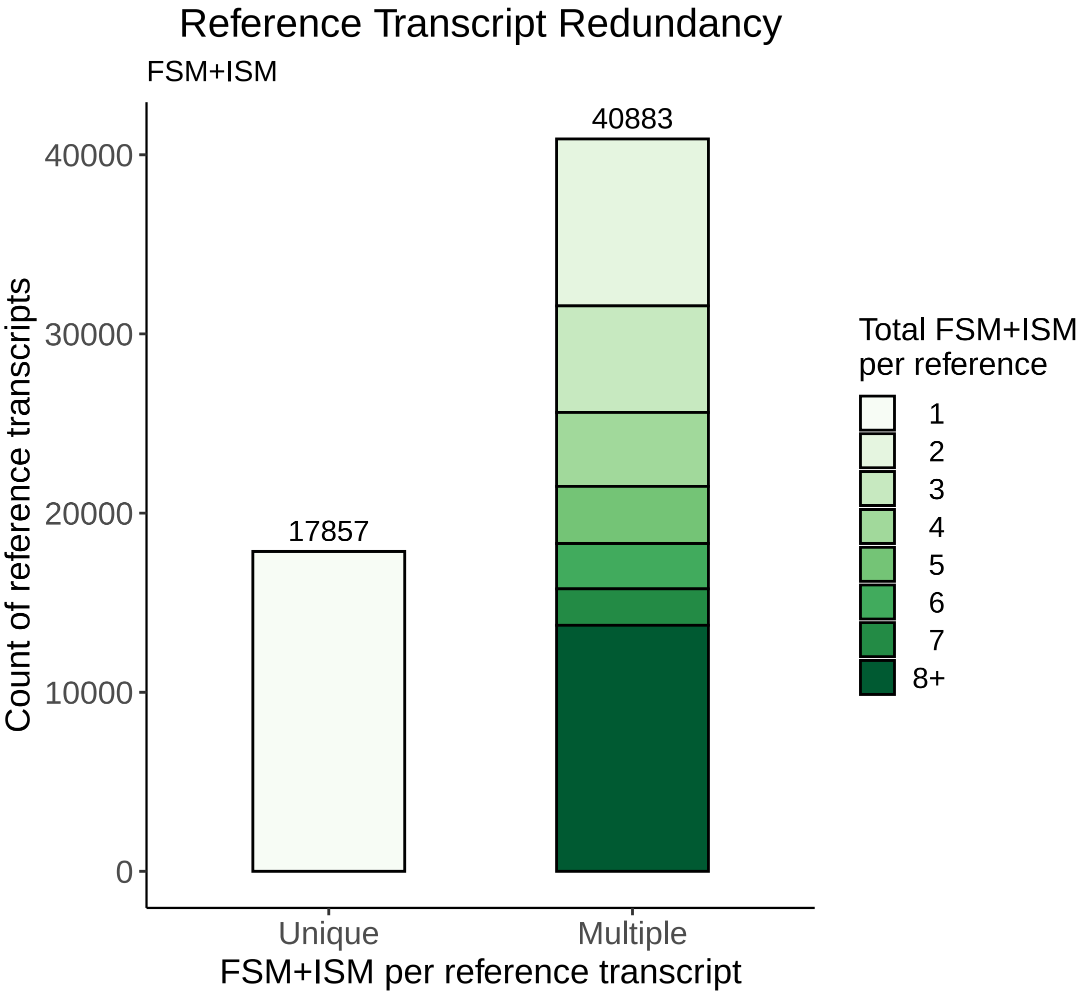
Frequency of PolyA Motifs

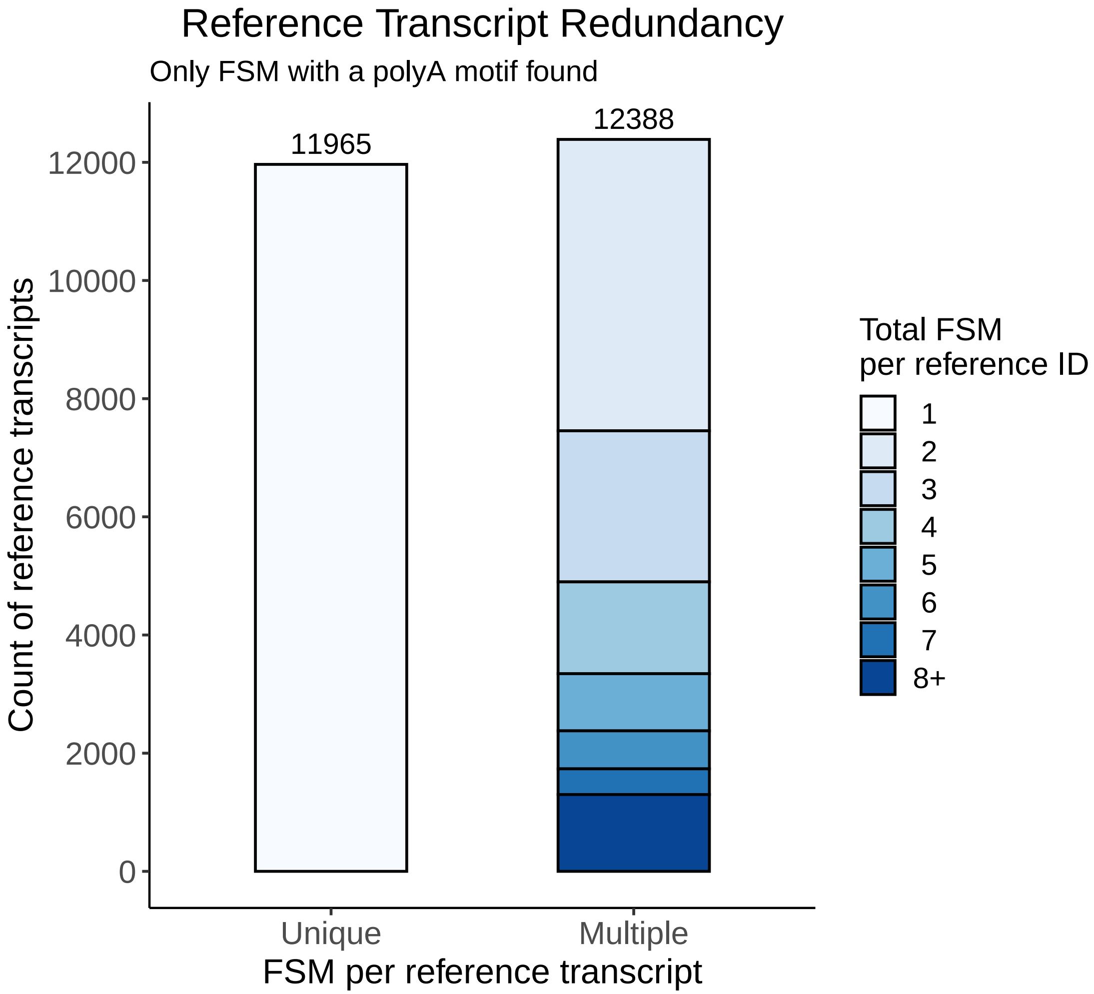
Motif	Count	%
AATAAA	169445	48.3
ATTAAA	50694	14.5
TATAAA	15253	4.4
AGTAAA	13528	3.9
AAGAAA	13496	3.9
AAAAAG	13448	3.8
TTTAAA	9974	2.8
GGGGCT	9612	2.7
AATATA	8184	2.3
AATACA	8076	2.3
CATAAA	7983	2.3
GATAAA	7817	2.2
AAAACA	7649	2.2
AATGAA	7042	2.0
AATAGA	4275	1.2
ACTAAA	4007	1.1

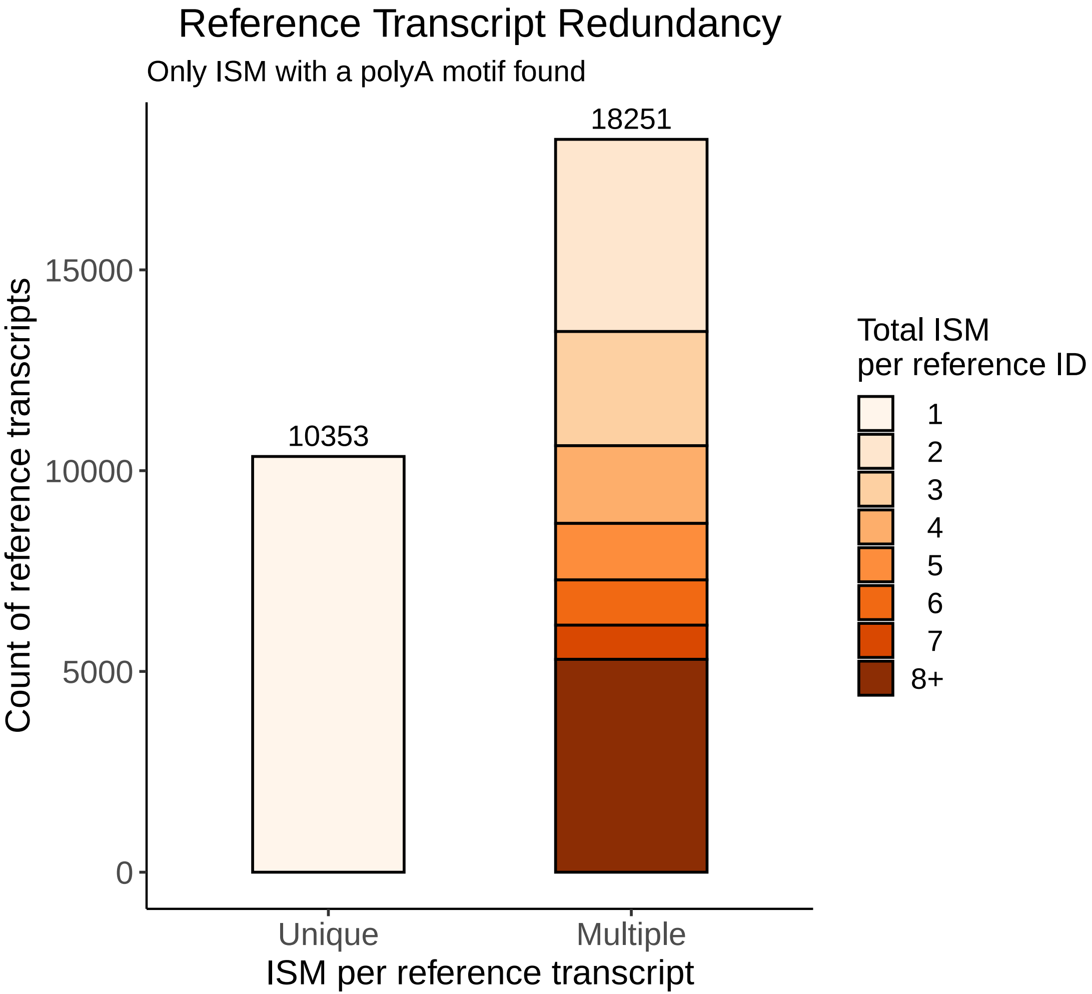
Redundancy Analysis

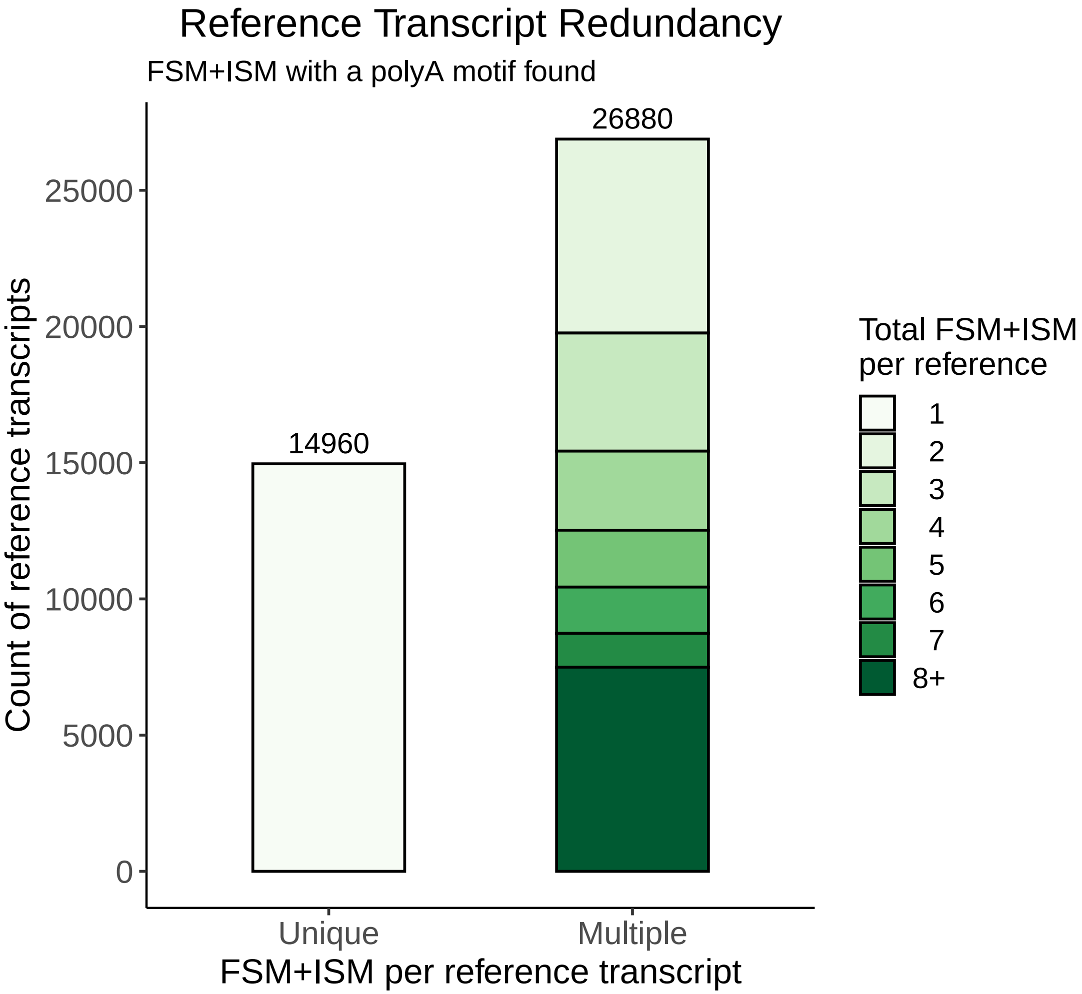

Reference Transcript Redundancy
Only FSM
23053
17823
20000
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only ISM
Count of Reference Transcripts
0
5000
10000
15000
20000
25000
11830
25955
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
FSM+ISM
40883
17857
40000
30000
20000
10000
0
Count of reference transcripts
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
11965
12388
0
2000
4000
6000
8000
10000
12000
Count of reference transcripts
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+

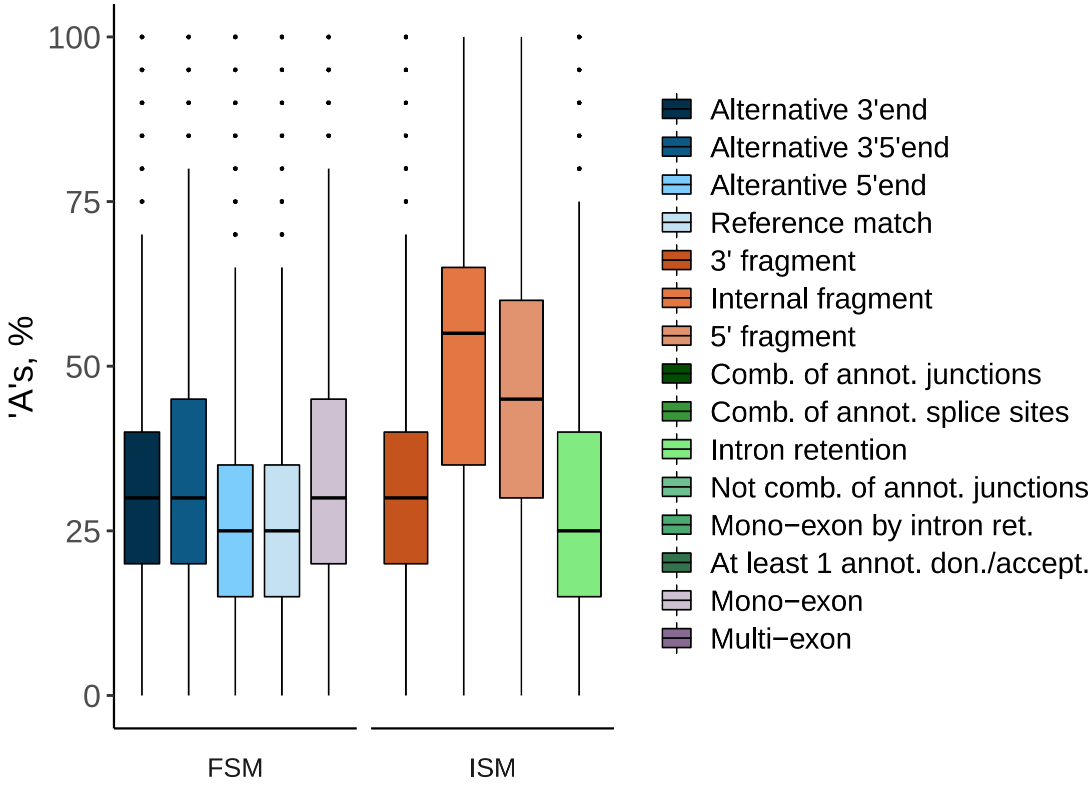

Reference Transcript Redundancy
Only ISM with a polyA motif found
18251
10353
Count of reference transcripts
15000
10000
5000
0
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

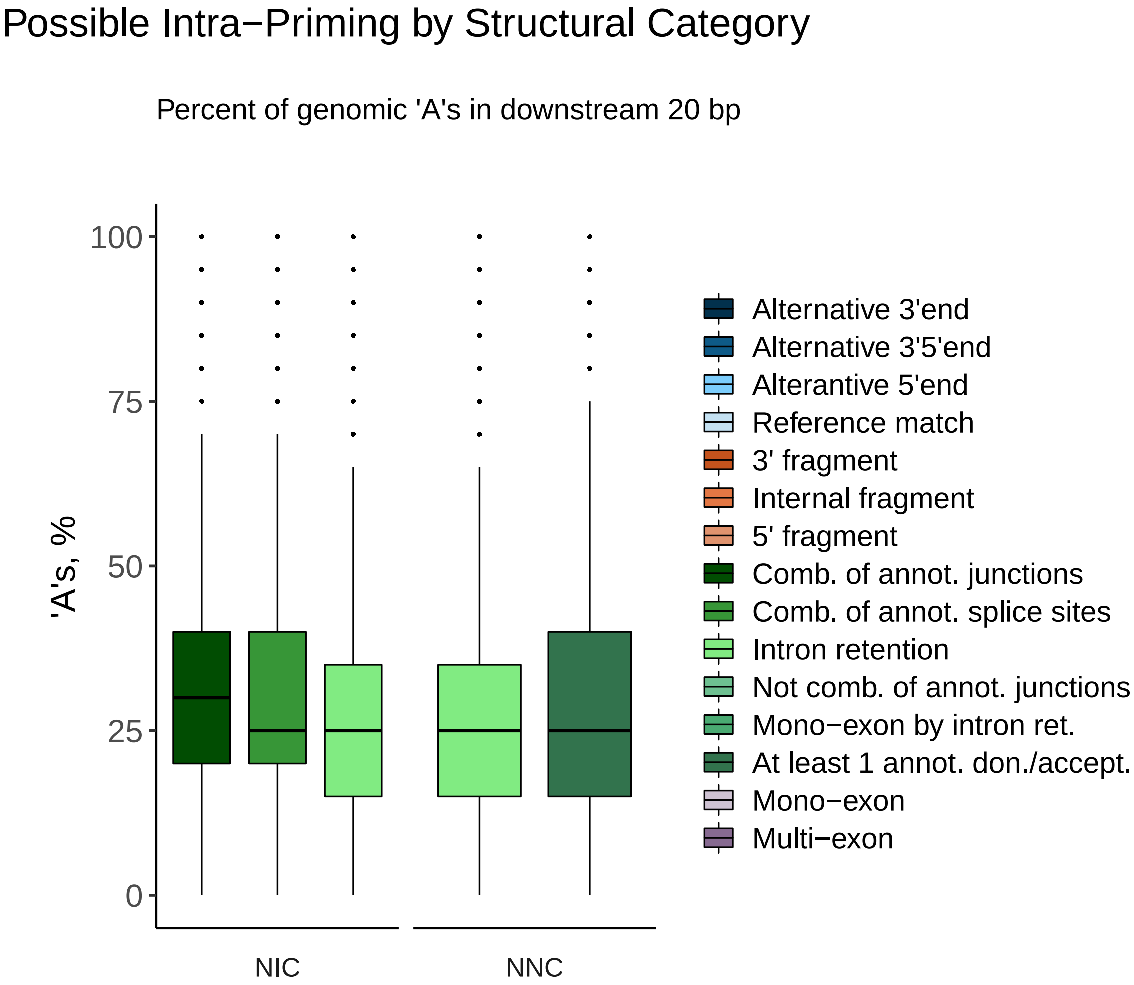

Reference Transcript Redundancy
FSM+ISM with a polyA motif found
Count of reference transcripts
25000
20000
15000
10000
5000
0
14960
26880
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+

Intra−Priming Quality Check

Possible Intra−Priming by Structural Category

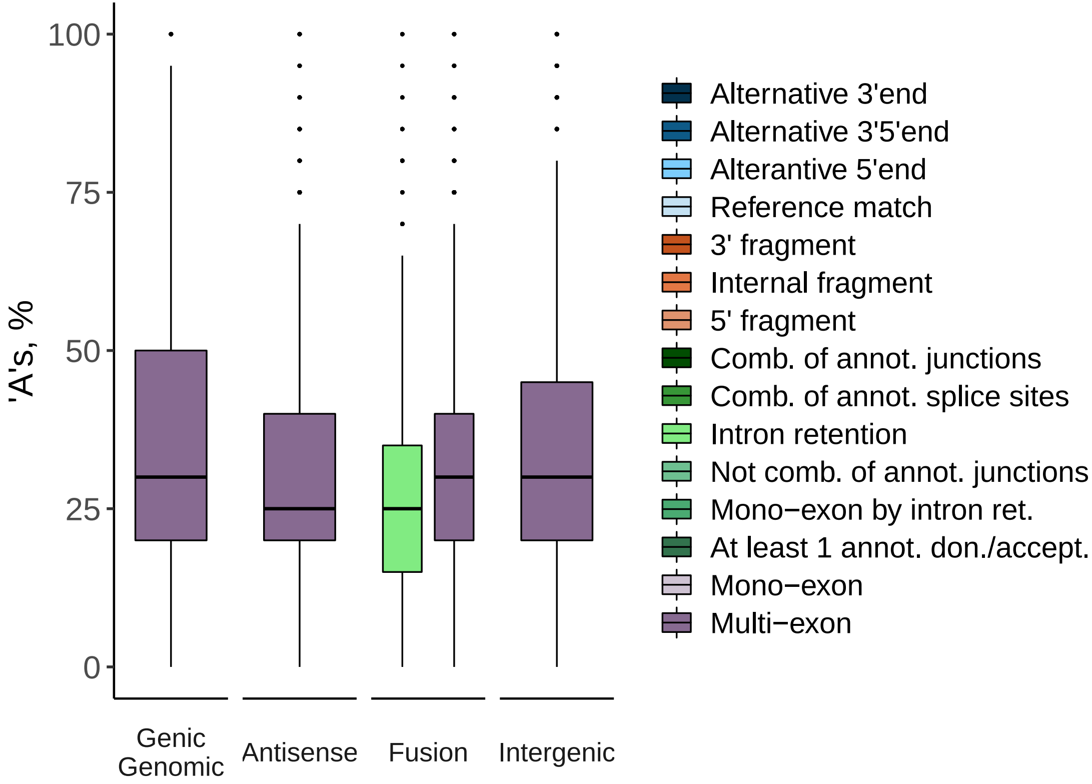
Percent of genomic 'A's in downstream 20 bp





Possible Intra−Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp

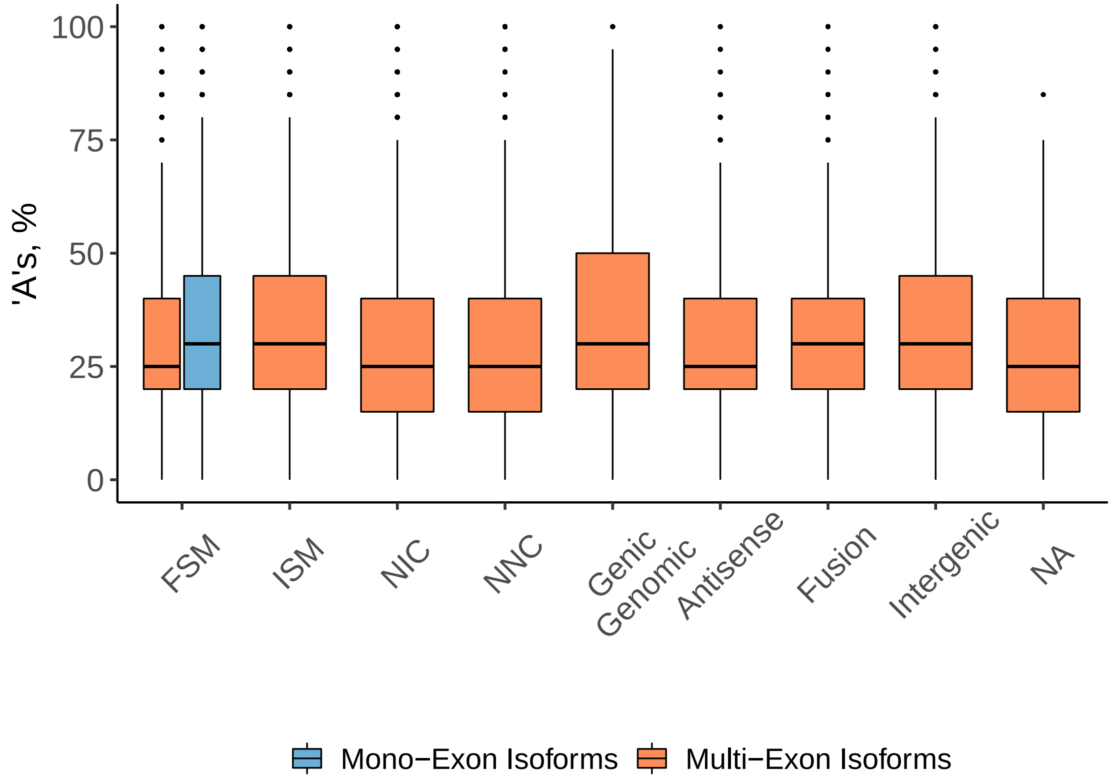


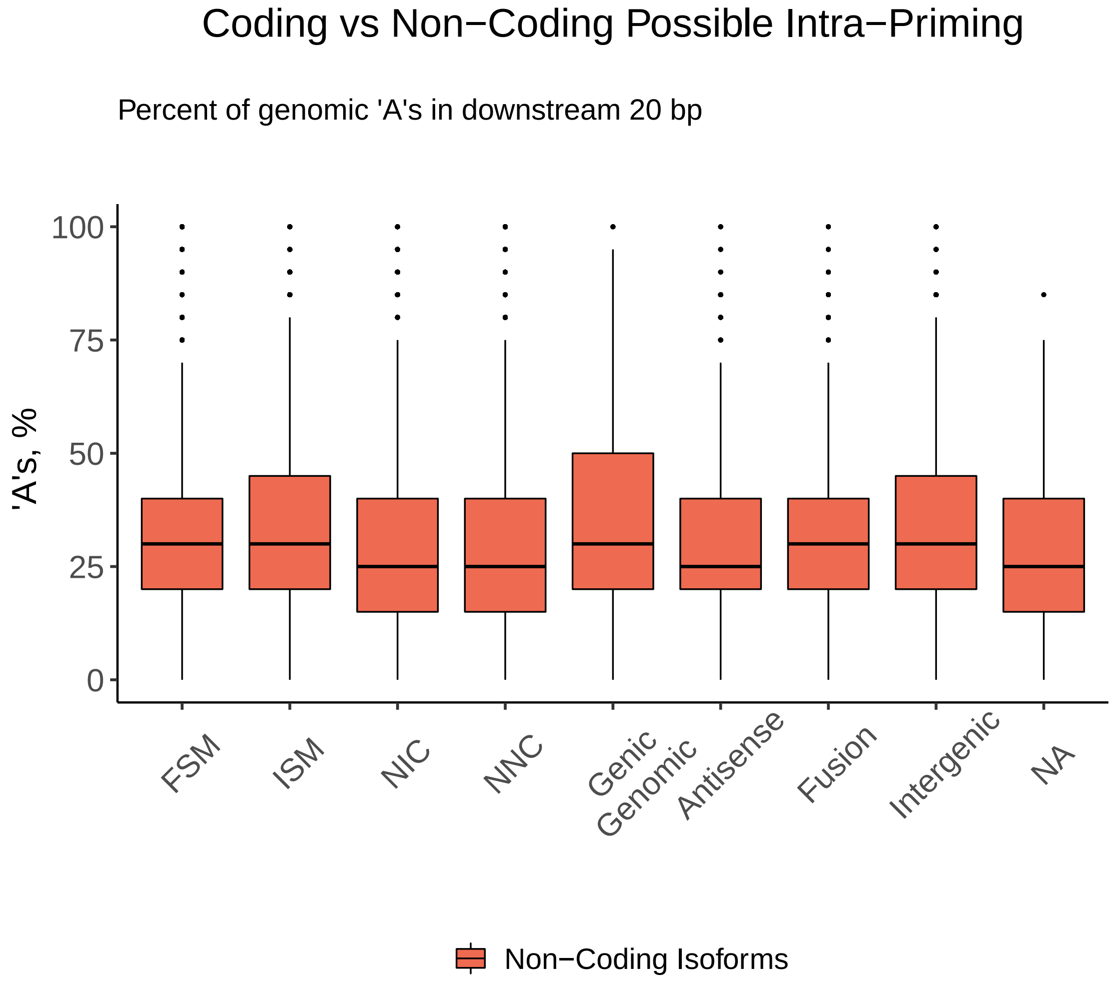
Possible Intra−Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp

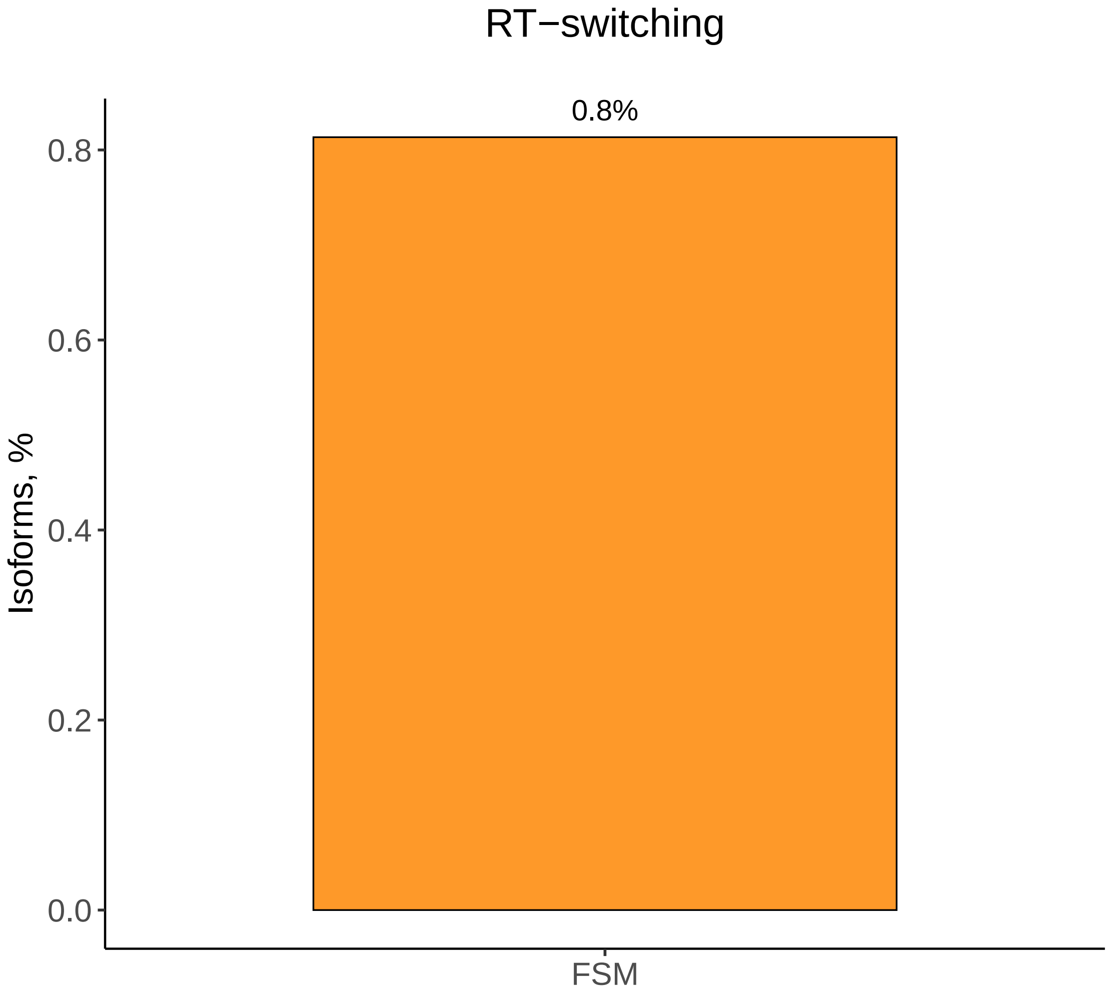
Mono− vs Multi−Exon Possible Intra−Priming

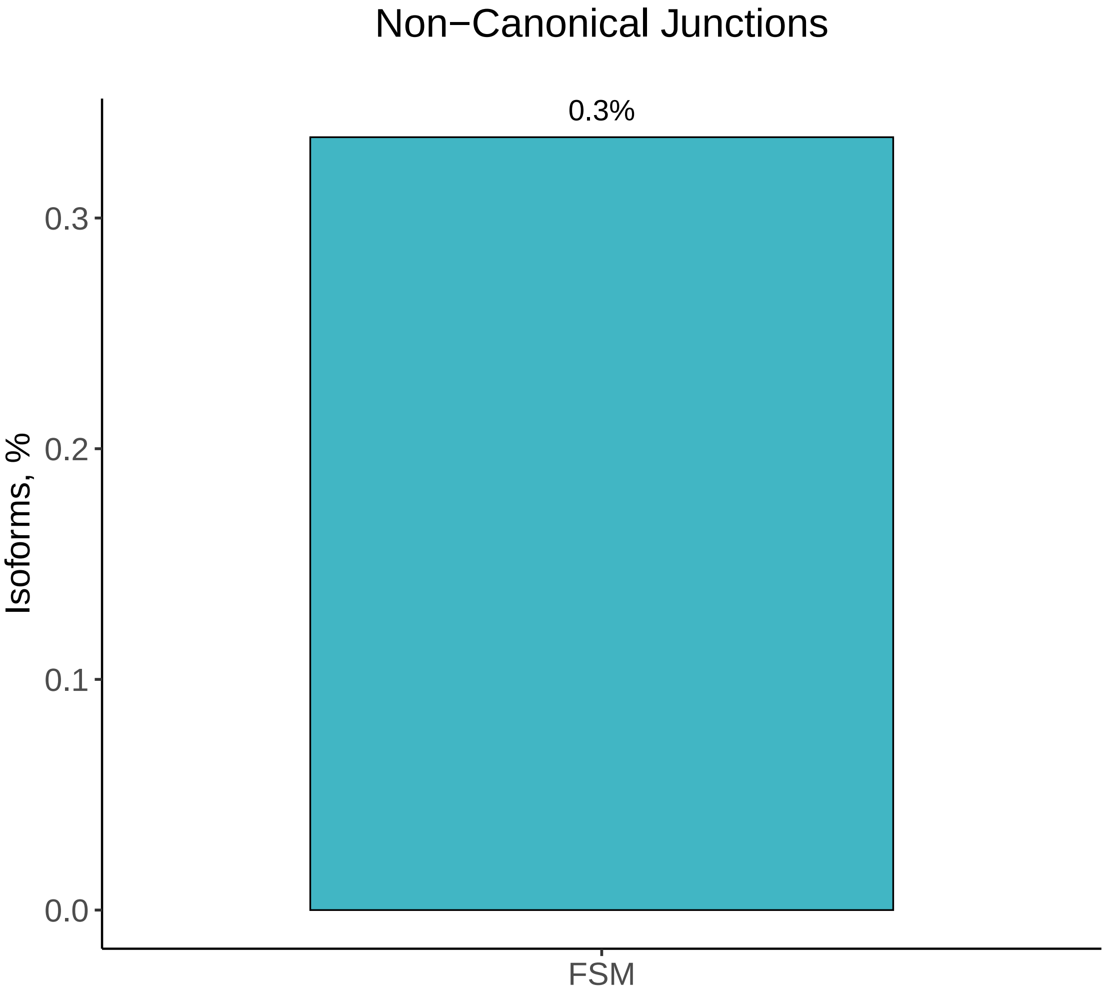
Percent of genomic 'A's in downstream 20 bp

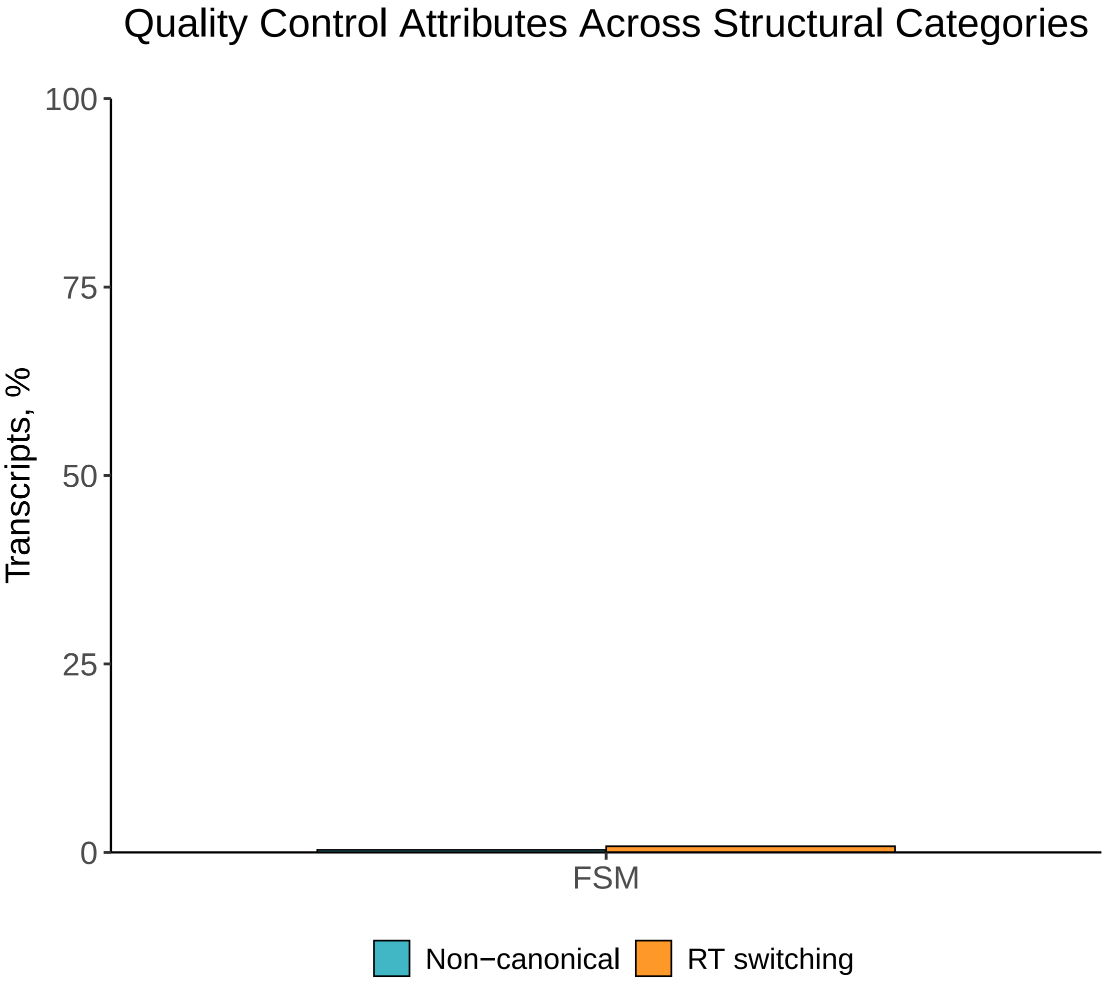



Coding vs Non−Coding Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic
Genomic
Antisense
Fusion
Intergenic
NA
Non−Coding Isoforms

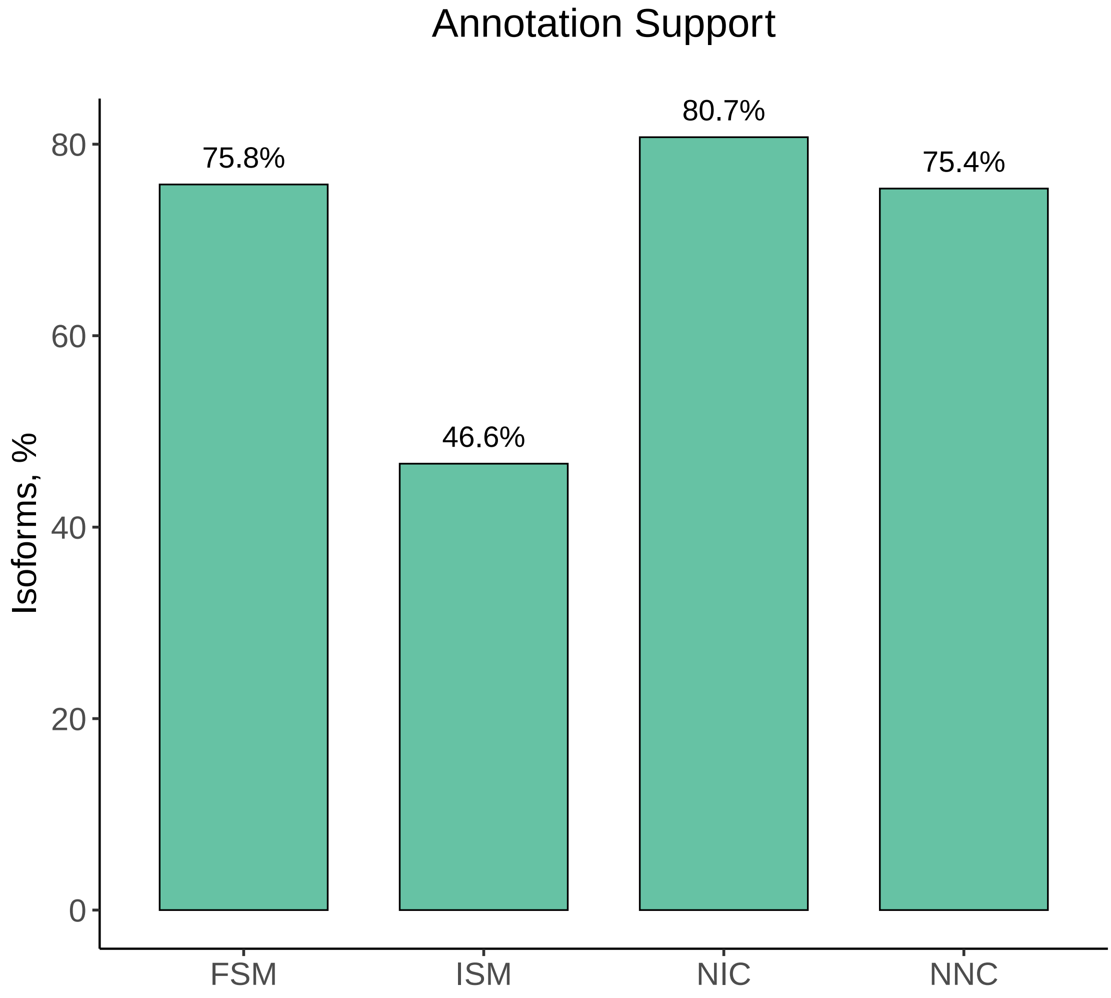
Features of Bad Quality

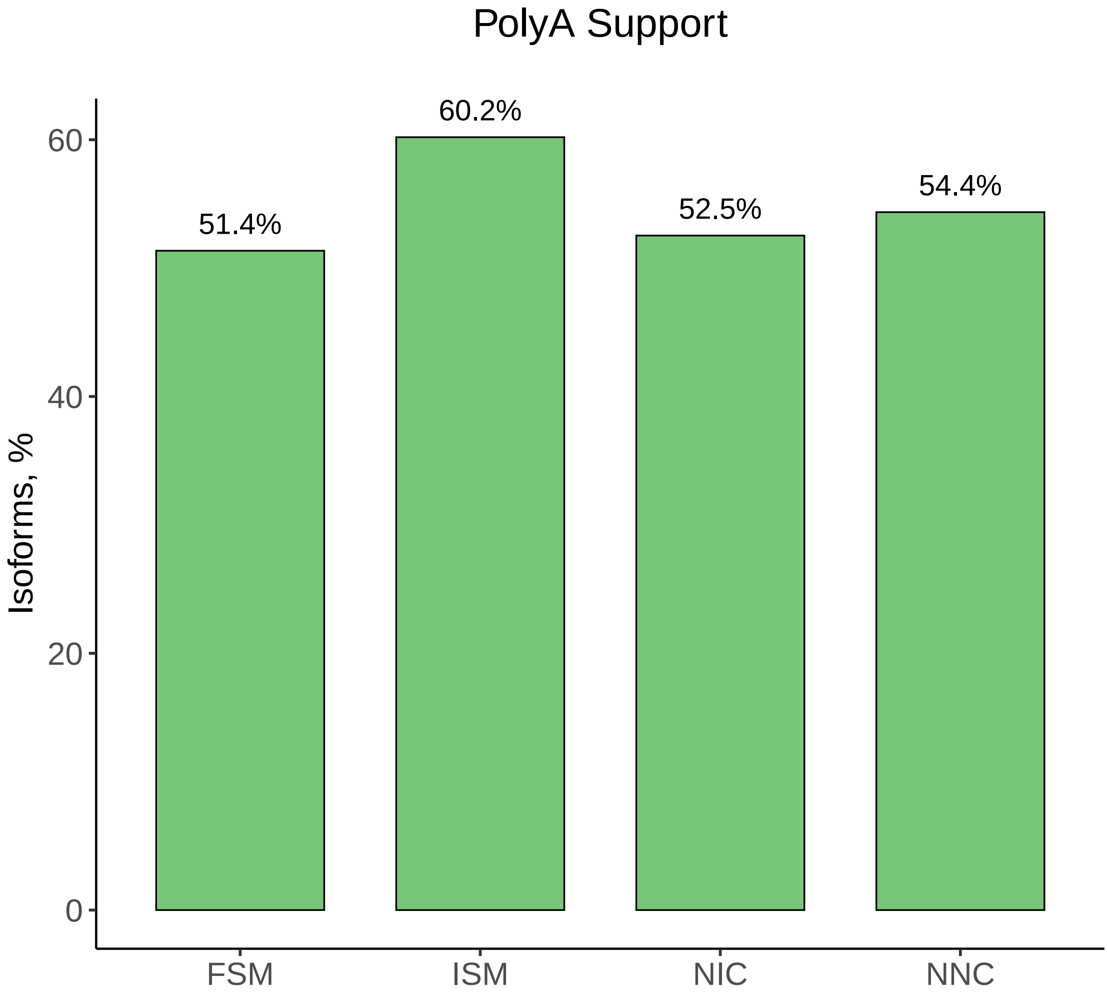

RT−switching
0.8%
Isoforms, %
0.0
0.2
0.4
0.6
0.8
FSM

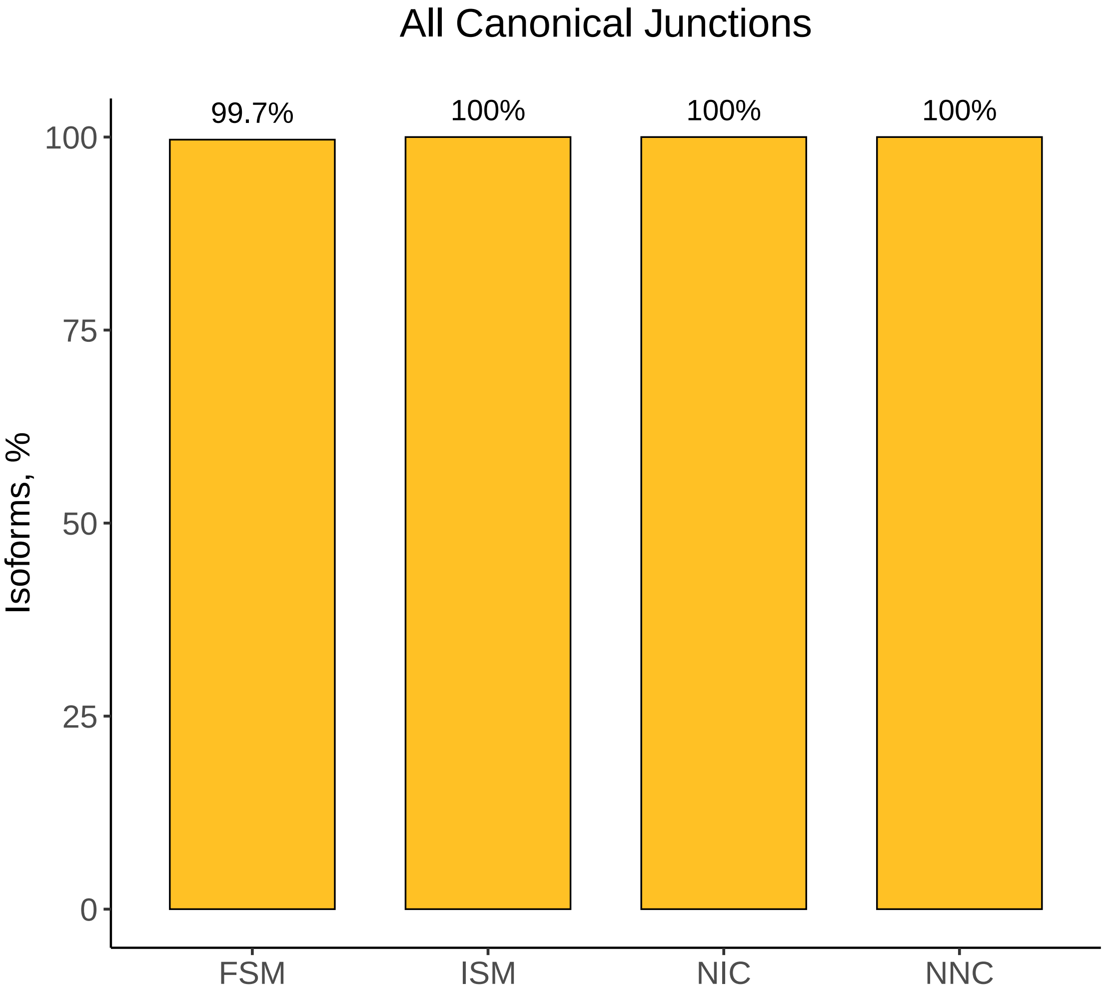

Non−Canonical Junctions
0.3%
Isoforms, %
0.3
0.2
0.1
0.0
FSM

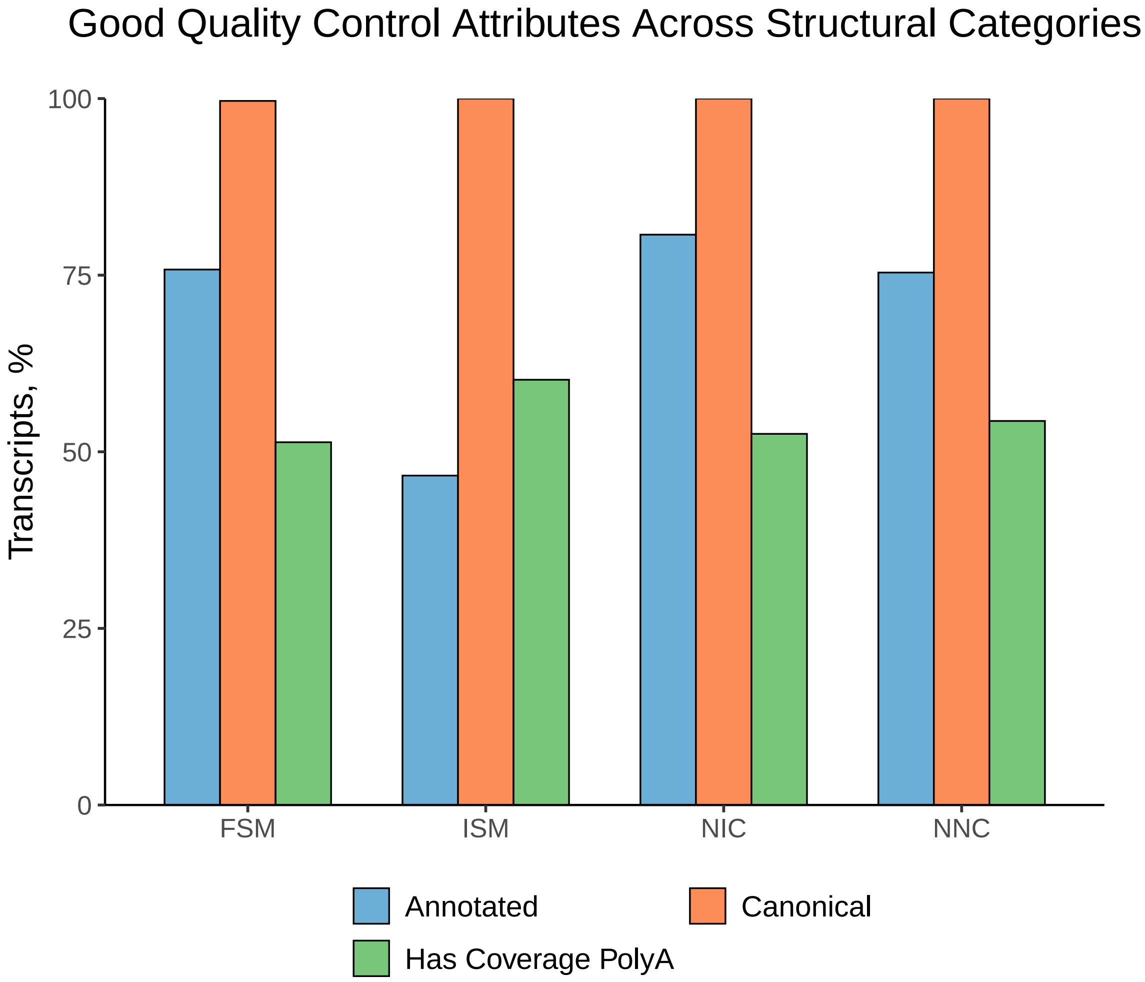

Quality Control Attributes Across Structural Categories
Transcripts, %
100
75
50
25
0
FSM
Non−canonical
RT switching

Features of Good Quality


Annotation Support
Isoforms, %
0
20
40
60
80
75.8%
46.6%
80.7%
75.4%
FSM
ISM
NIC
NNC

PolyA Support
Isoforms, %
0
20
40
60
51.4%
60.2%
52.5%
54.4%
FSM
ISM
NIC
NNC


All Canonical Junctions
Isoforms, %
99.7%
100%
100%
100%
100
75
50
25
0
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