SQANTI3 report

Unique Genes: 22991

Unique Isoforms: 552028

Transcript Classification

Category	Isoforms, count
FSM	139688
ISM	201592
NIC	111363
NNC	91483
Genic Genomic	623
Antisense	1822
Fusion	2788
Intergenic	2597
Genic Intron	0

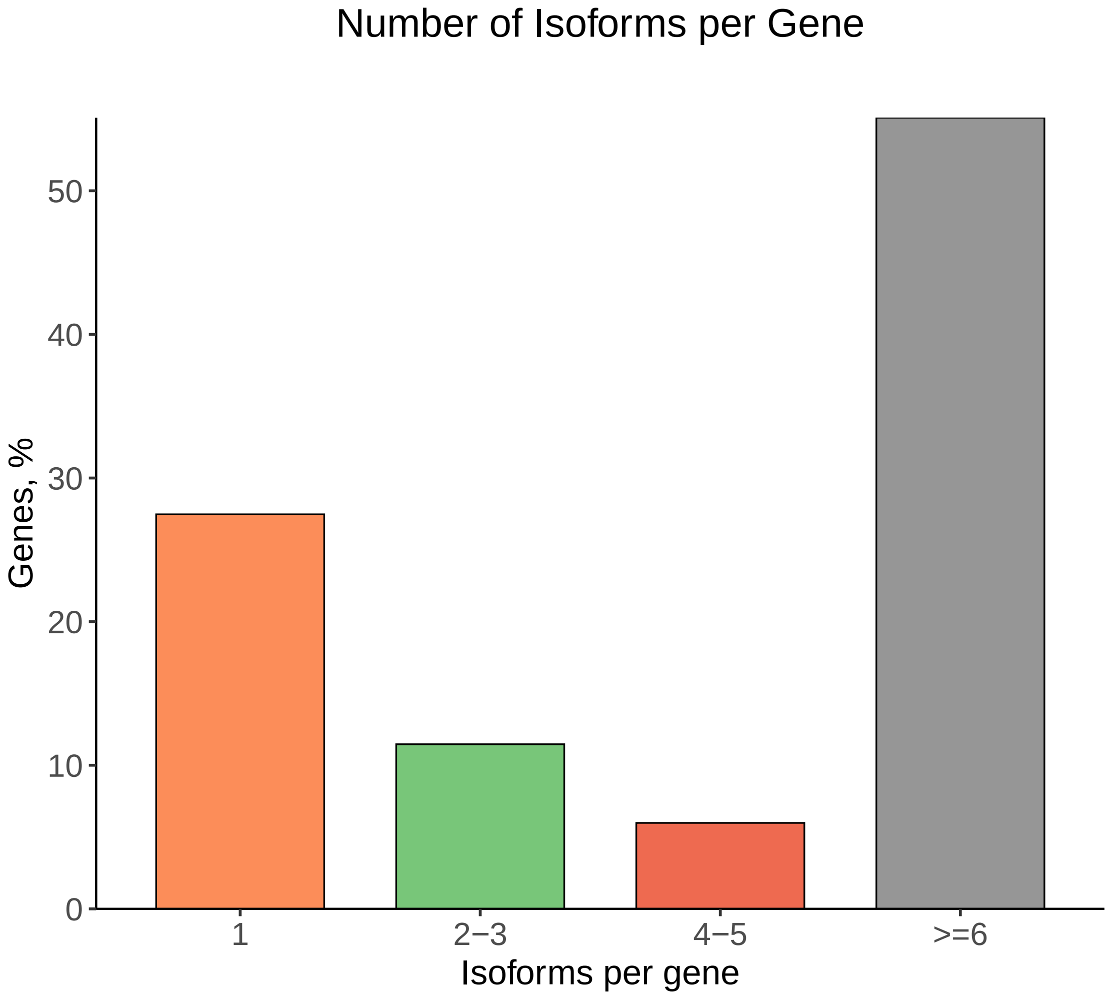
Gene Classification

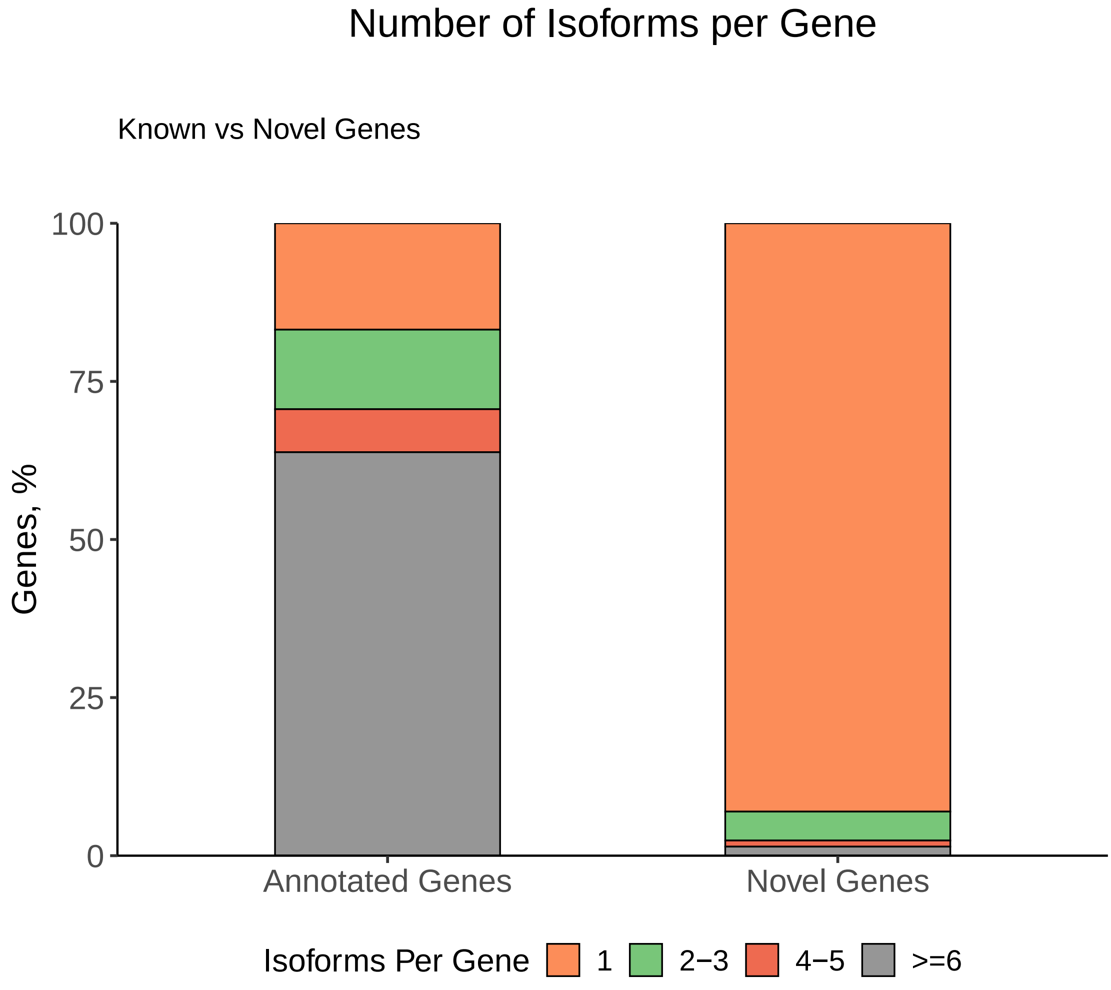
Category	Genes, count
Annotated Genes	19777
Novel Genes	3214

Splice Junction Classification

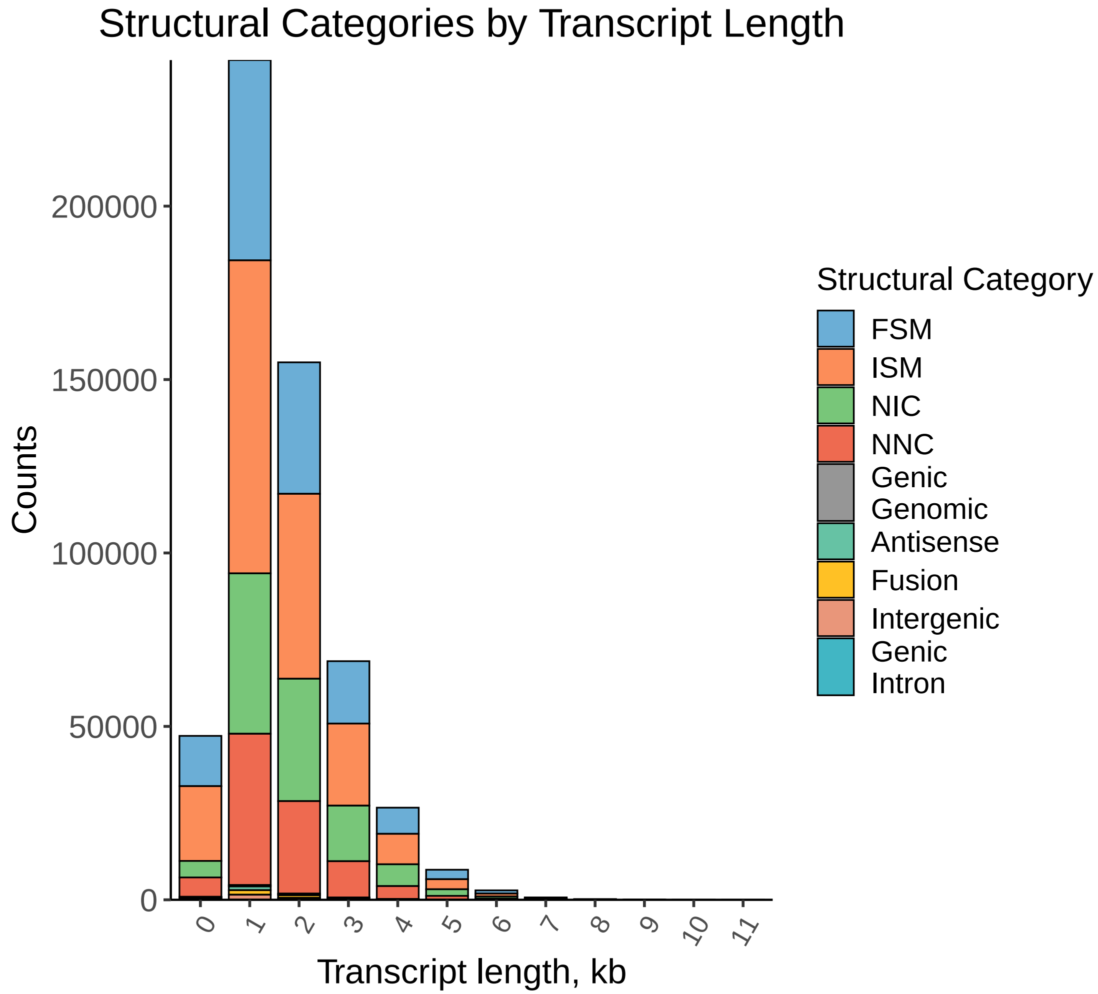
Category	SJs, count	Percent
Known canonical	185974	66.62
Known Non−canonical	90	0.03
Novel canonical	93108	33.35
Novel Non−canonical	0	0.00

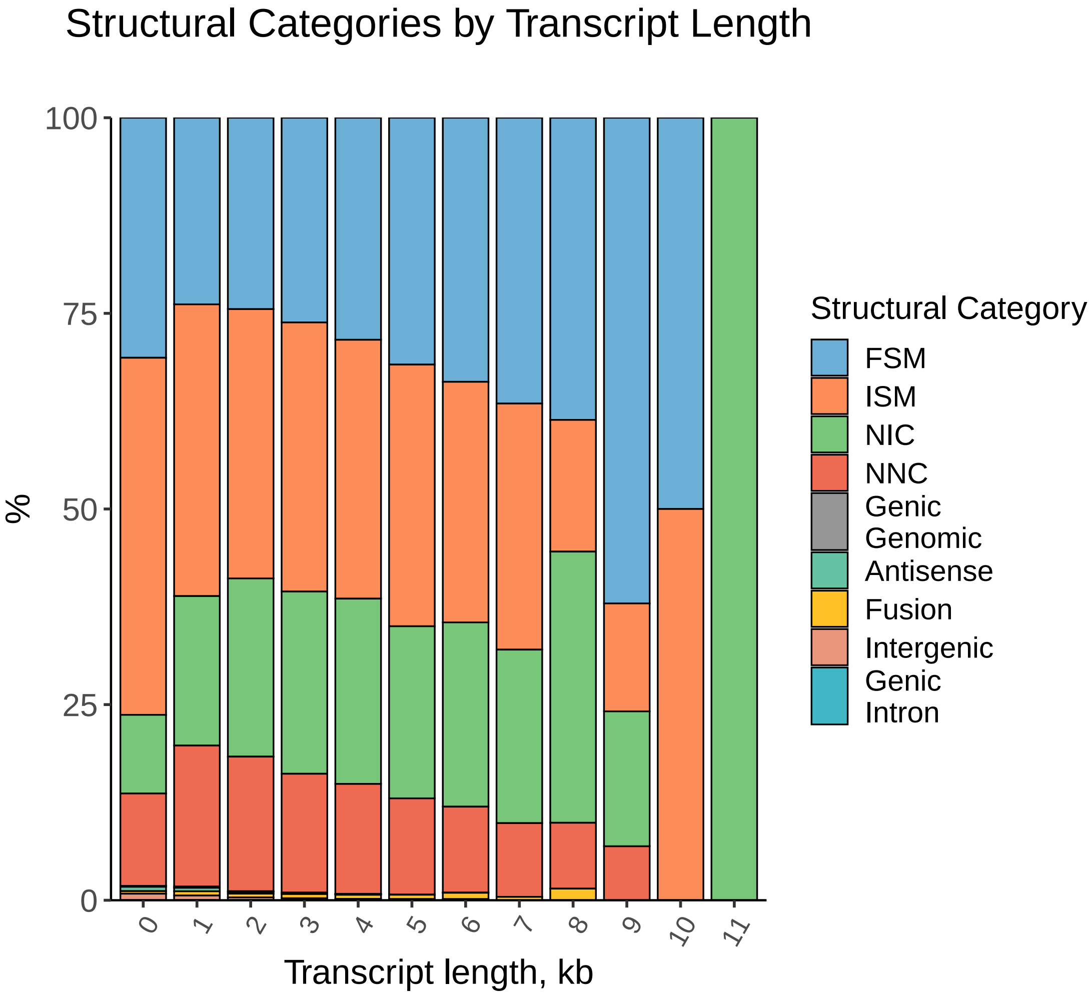
Gene Characterization

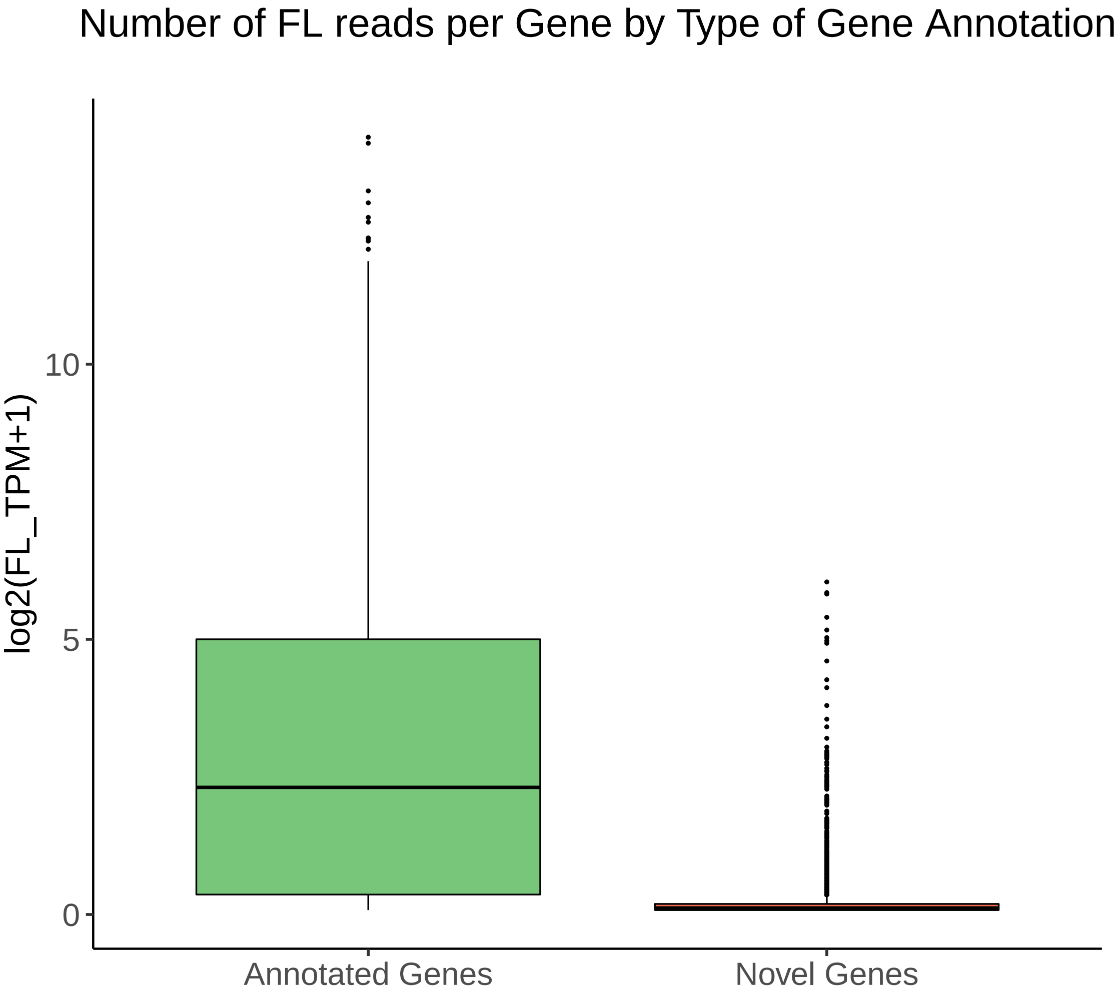

Number of Isoforms per Gene
Genes, %
0
10
20
30
40
50
1
2−3
4−5
>=6
Isoforms per gene

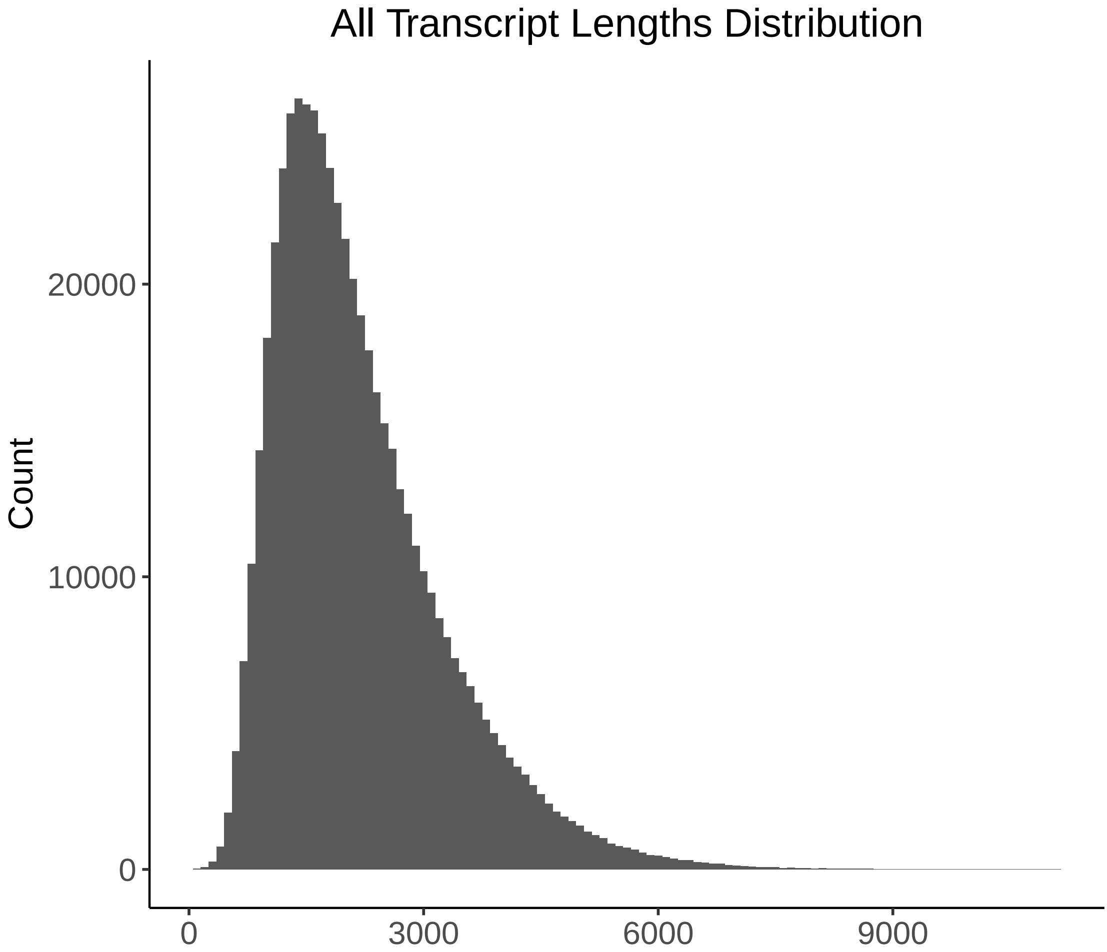

Number of Isoforms per Gene
Known vs Novel Genes
Genes, %
100
75
50
25
0
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3
4−5
>=6

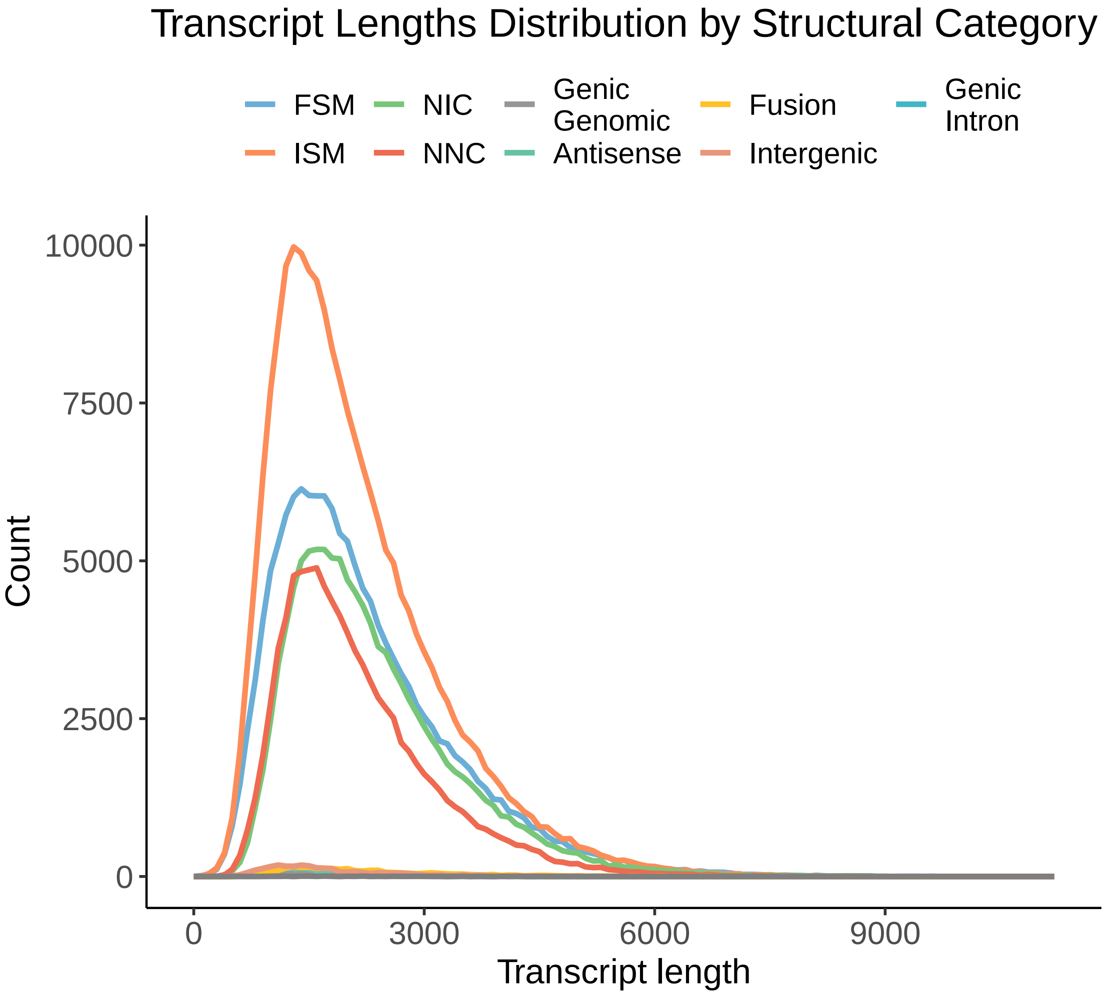

Distribution of Mono− vs Multi−Exon Transcripts
Transcripts, %
100
75
50
25
0
Novel Genes
Annotated Genes
Transcript type
Multi−Exon
Mono−Exon

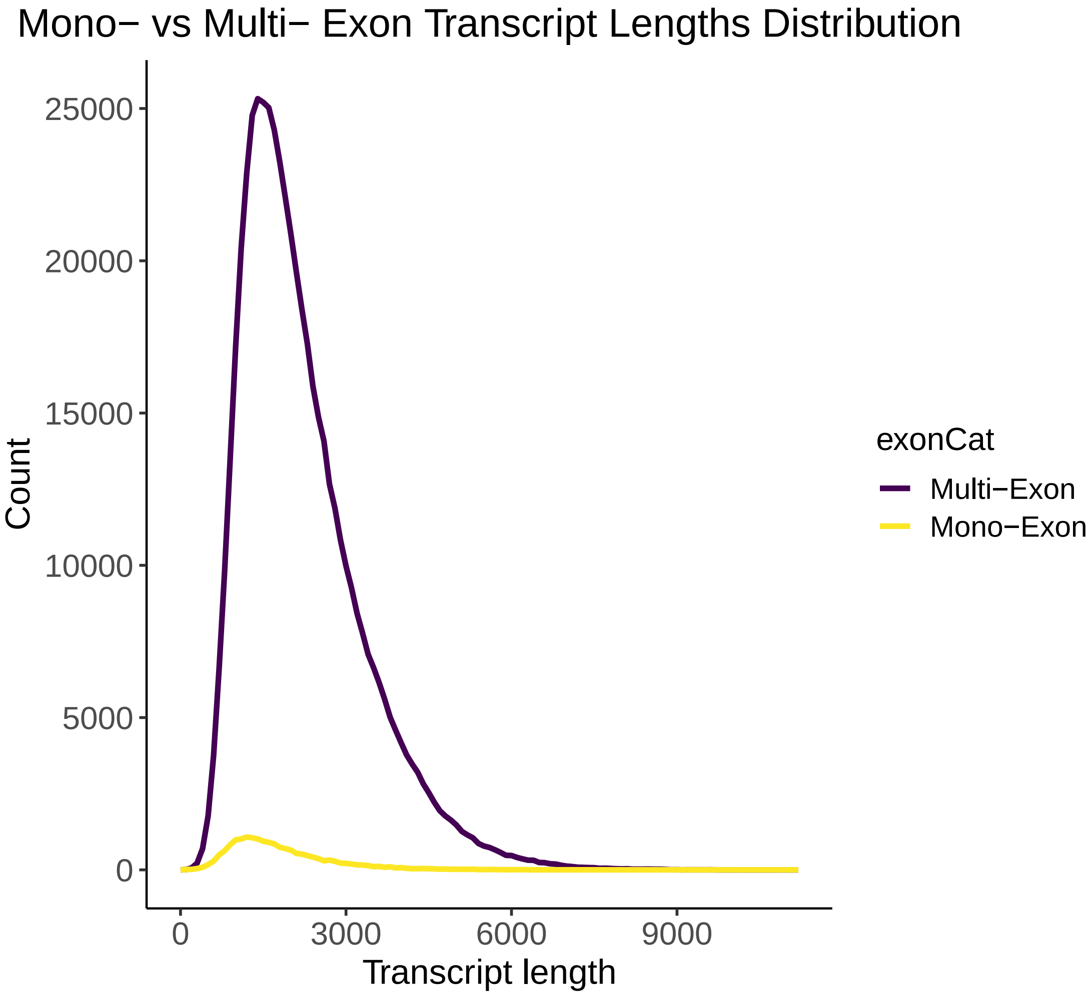

Structural Categories by Transcript Length
Counts
200000
150000
100000
50000
0
0
1
2
3
4
5
6
7
8
9
10
11
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

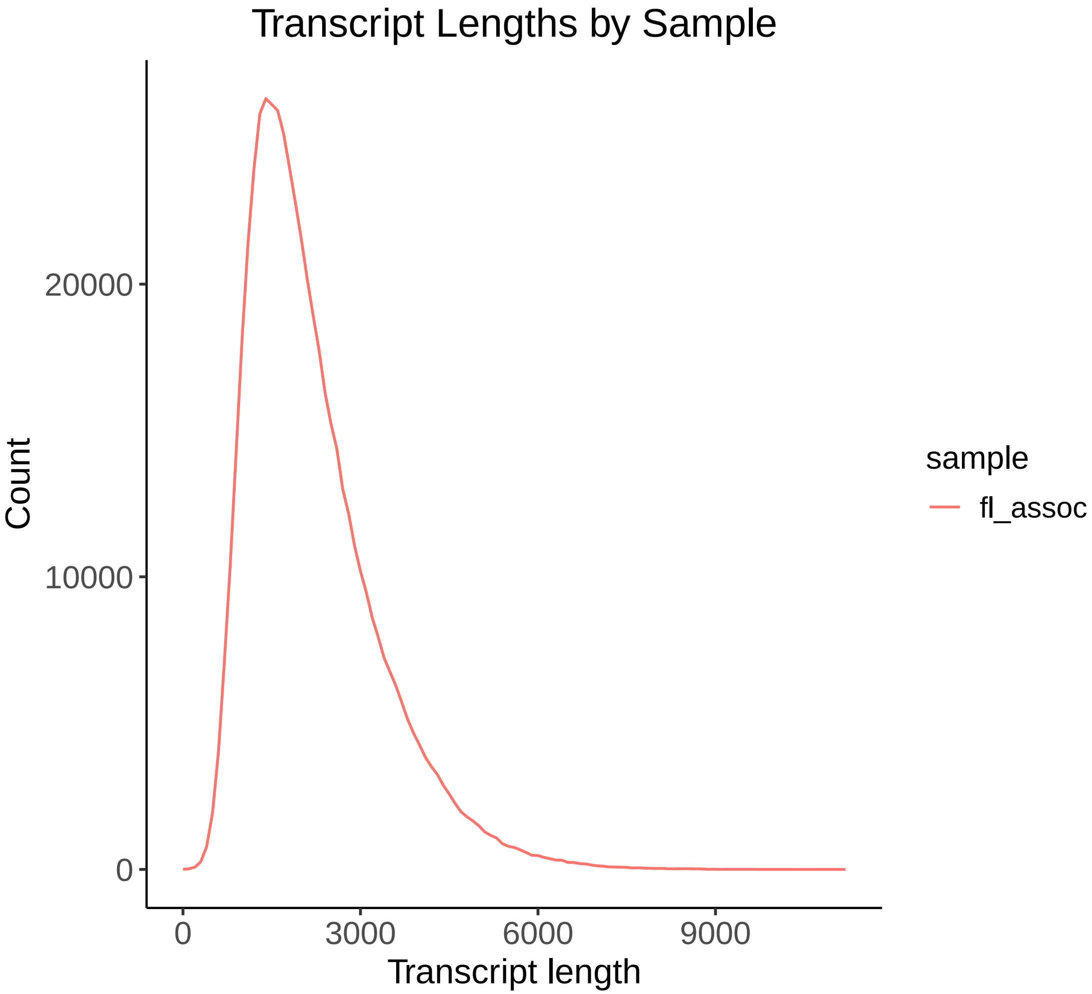

Structural Categories by Transcript Length
100
75
50
25
0
%
0
1
2
3
4
5
6
7
8
9
10
11
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron


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

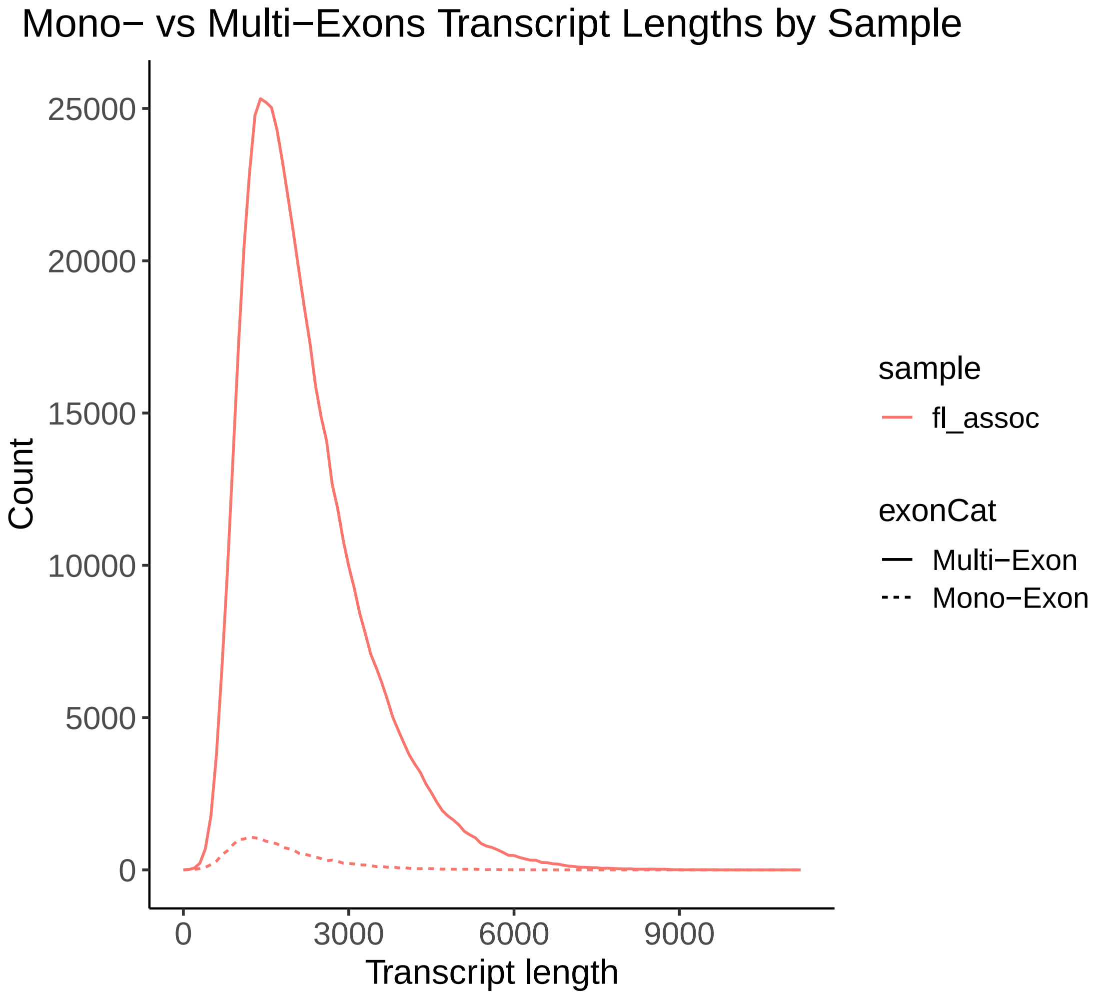

All Transcript Lengths Distribution
Count
20000
10000
0
0
3000
6000
9000
Transcript length


Transcript Lengths Distribution by Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
10000
7500
5000
2500
0
Count
0
3000
6000
9000
Transcript length

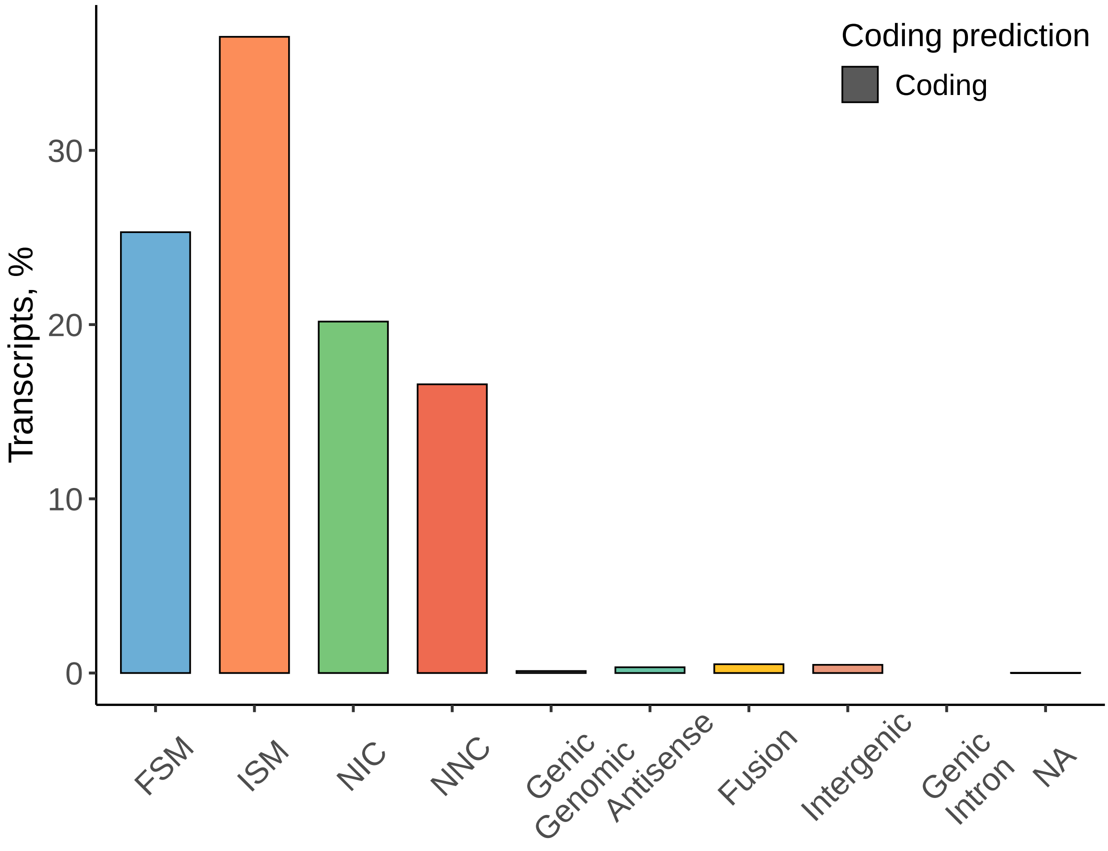

Mono− vs Multi− Exon Transcript Lengths Distribution
Count
25000
20000
15000
10000
5000
0
0
3000
6000
9000
Transcript length
exonCat
Multi−Exon
Mono−Exon



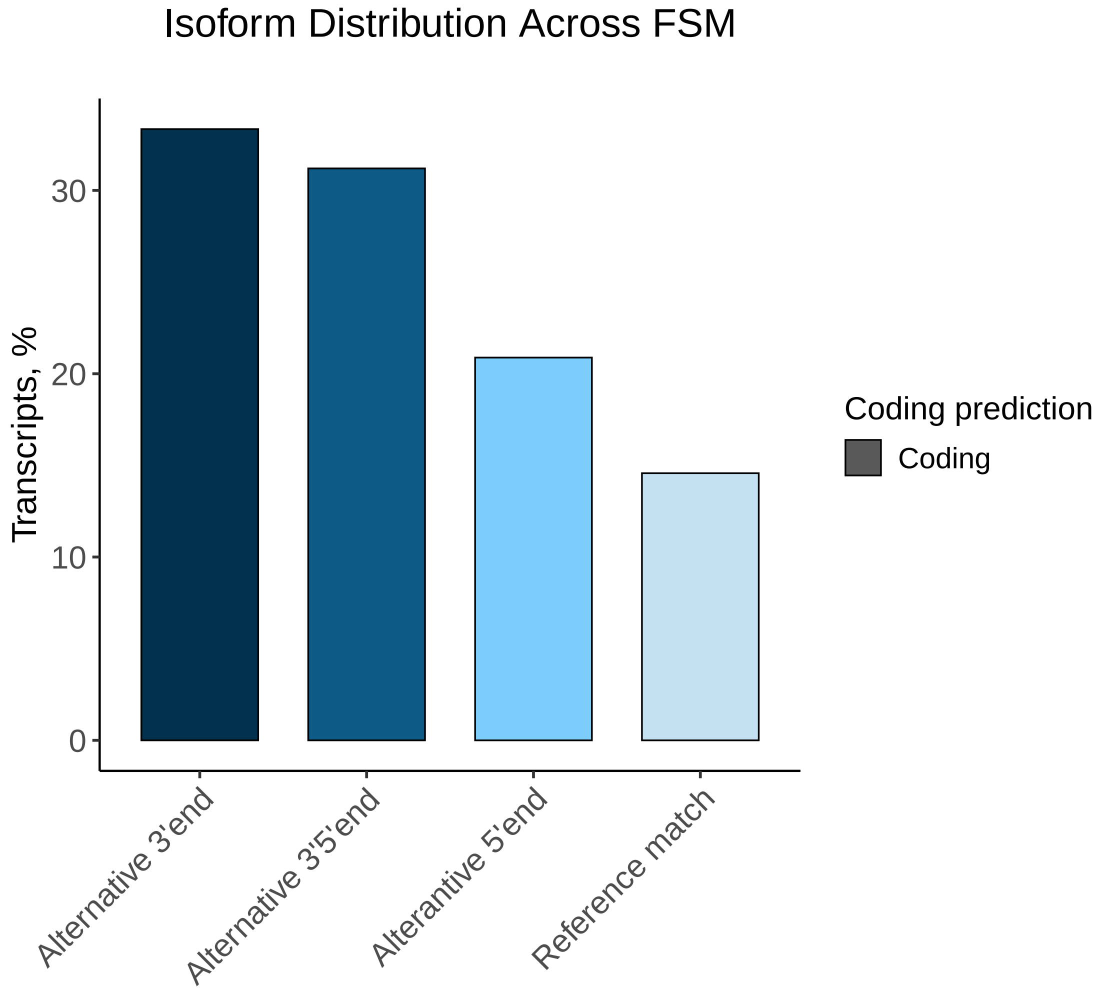
Transcript Lengths by Sample
Count
20000
10000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc

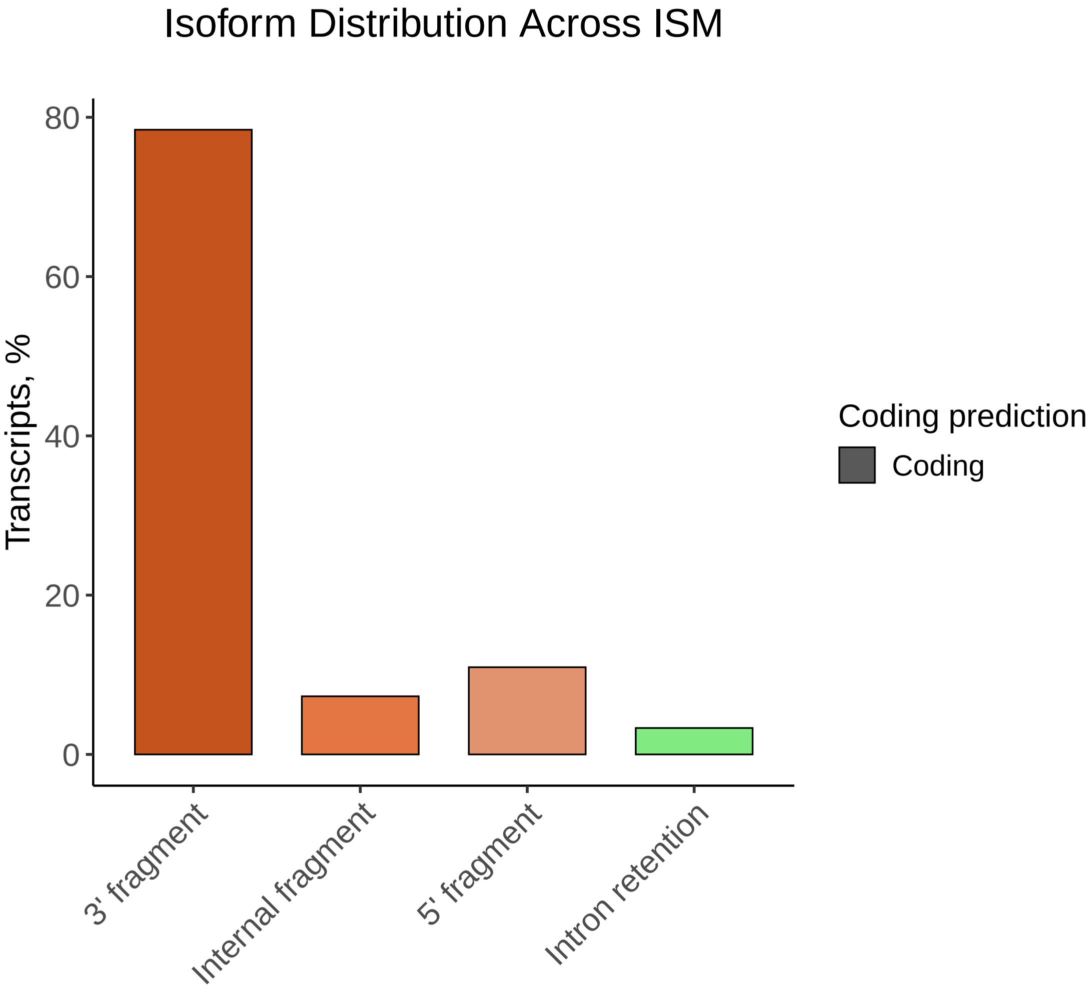

Mono− vs Multi−Exons Transcript Lengths by Sample
Count
25000
20000
15000
10000
5000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc
exonCat
Multi−Exon
Mono−Exon

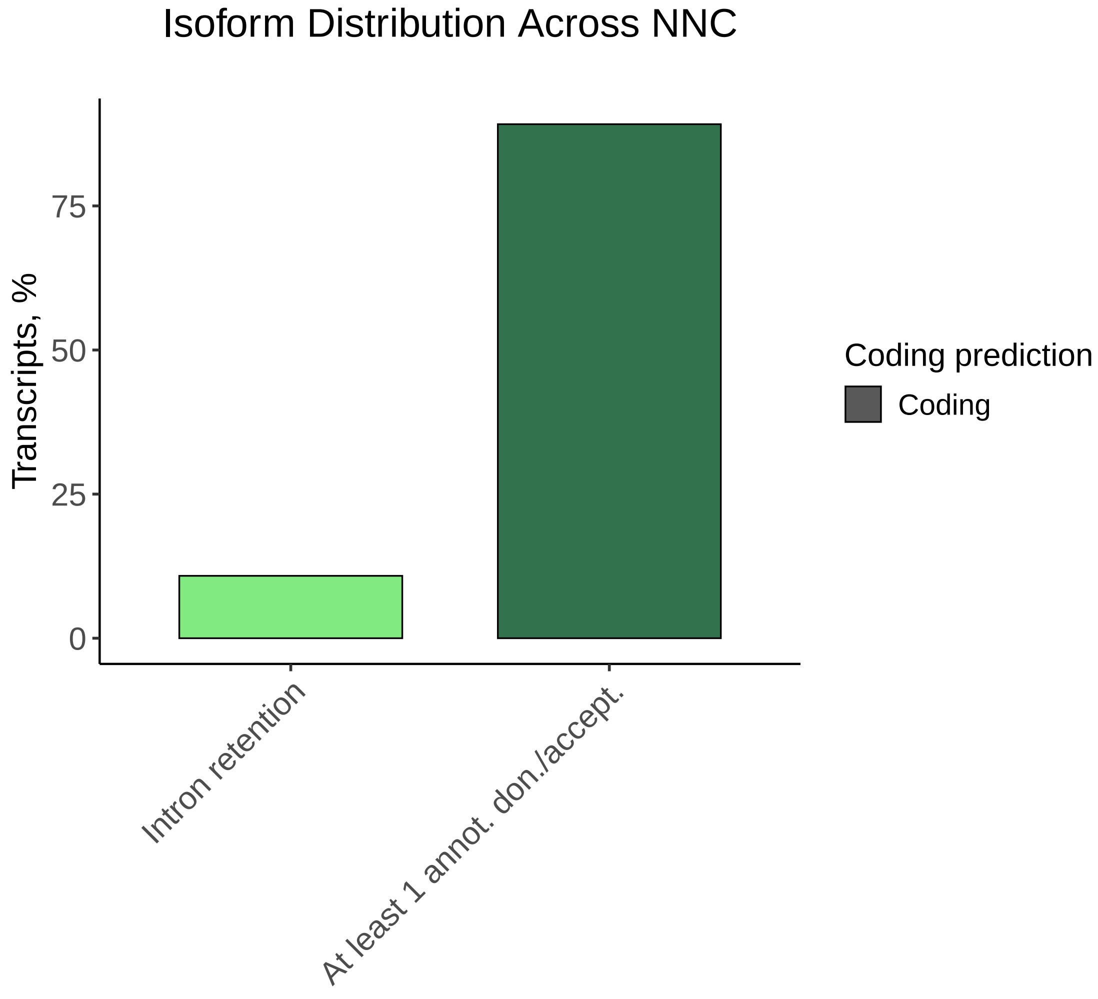
Structural Isoform Characterization

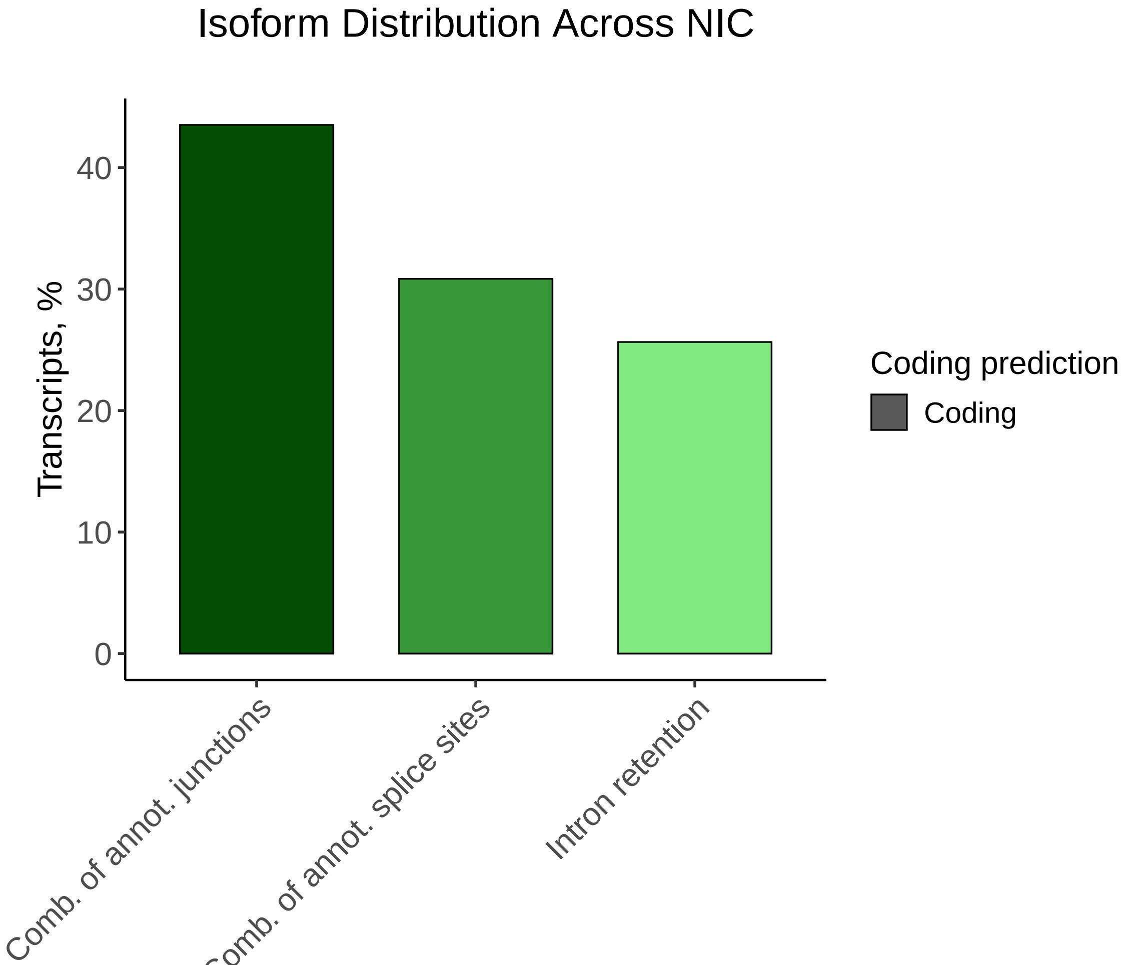


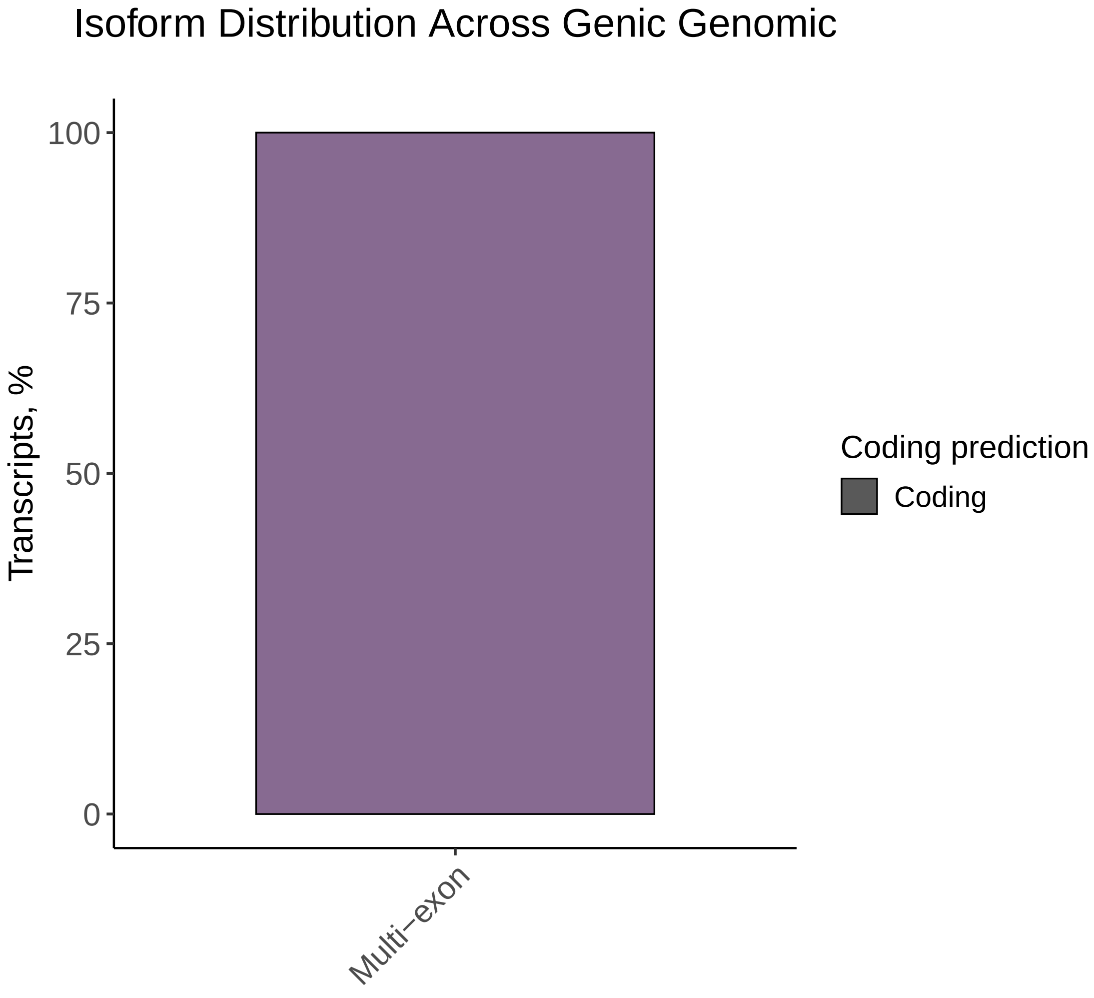
Isoform Distribution Across Structural Categories

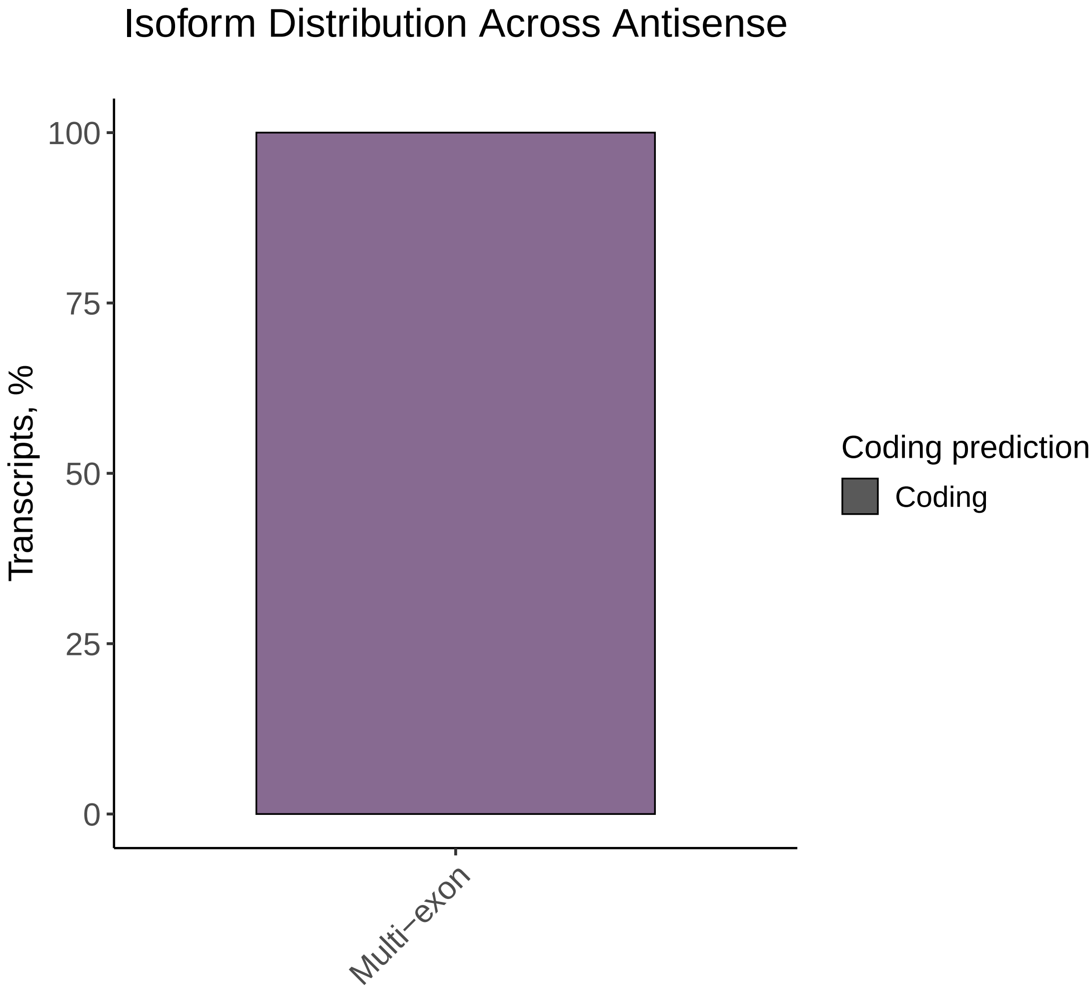

Isoform Distribution Across FSM
Transcripts, %
0
10
20
30
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Coding prediction
Coding

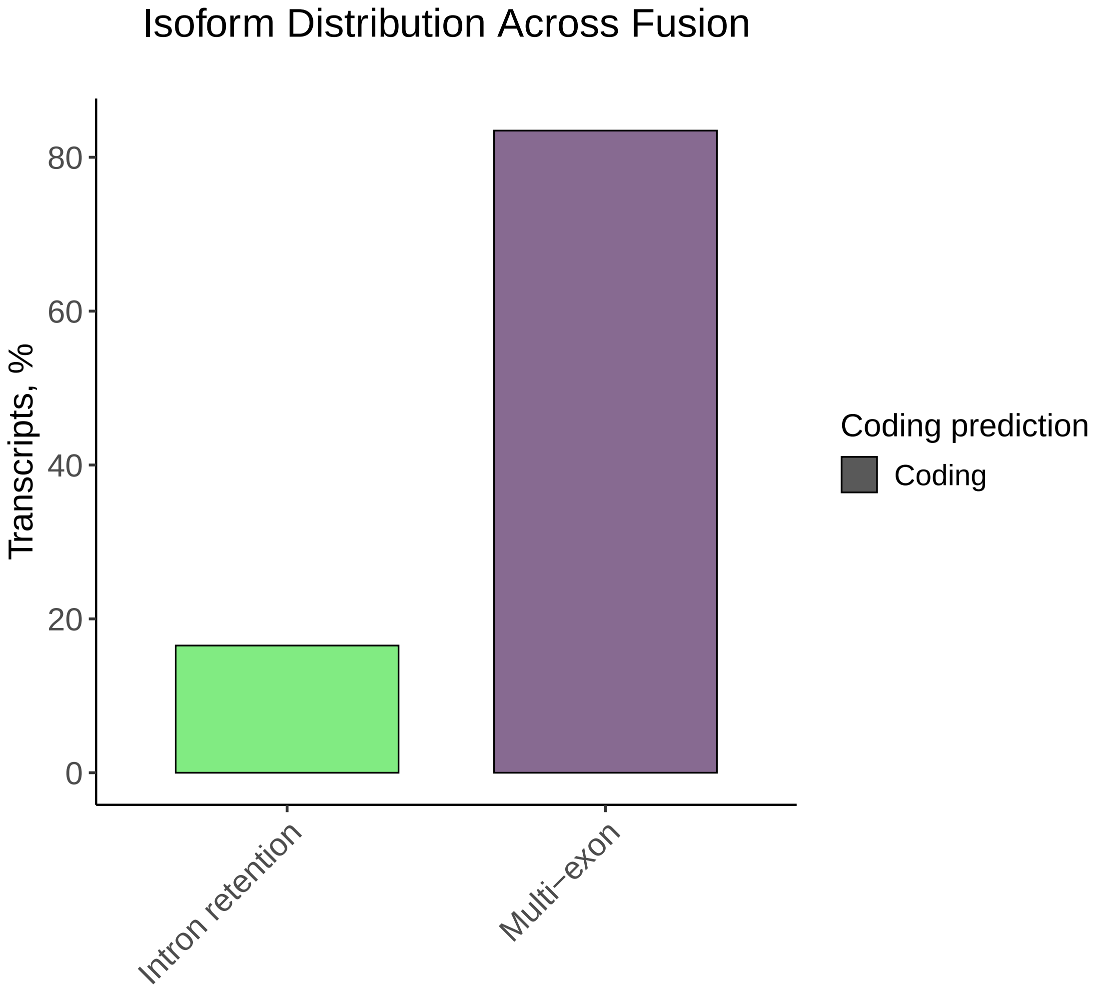

Isoform Distribution Across ISM
Transcripts, %
80
60
40
20
0
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding

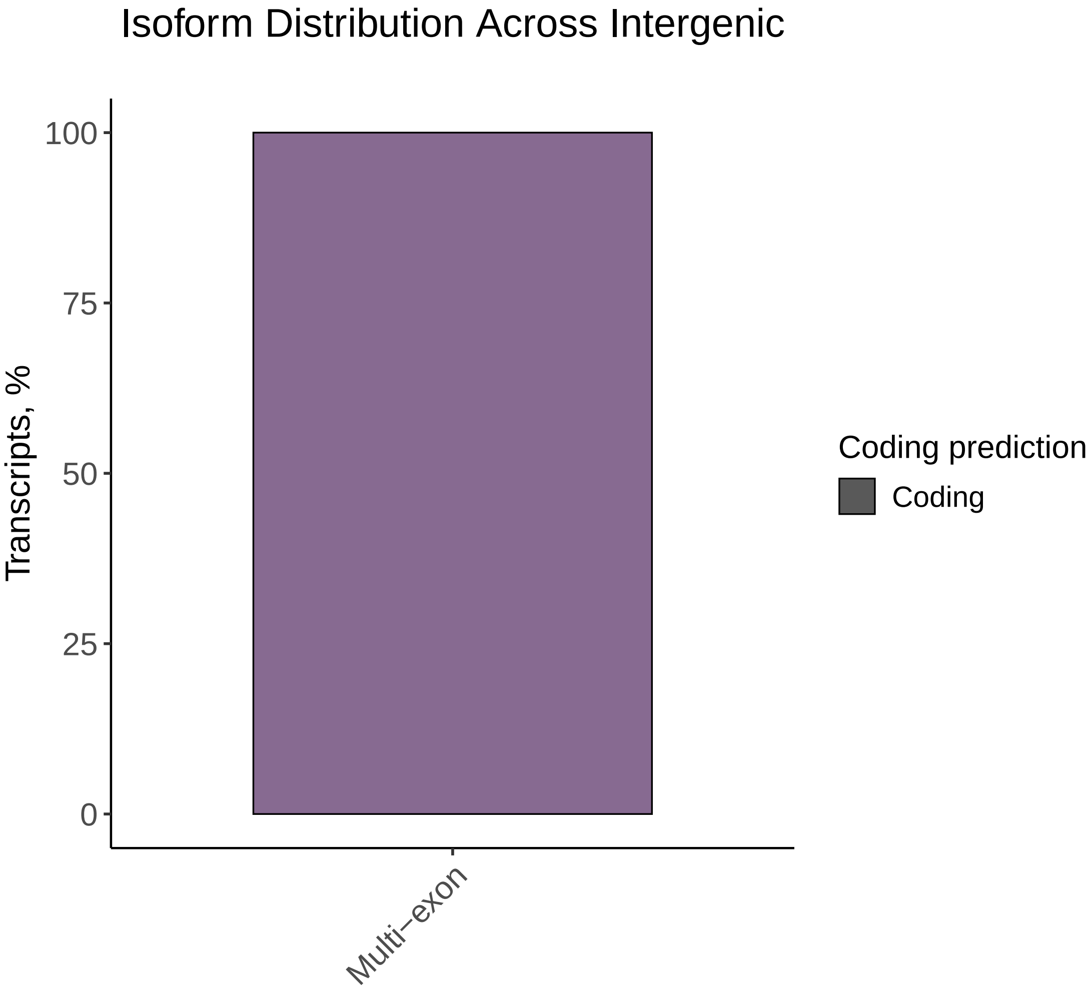

Isoform Distribution Across NNC
Transcripts, %
0
25
50
75
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding

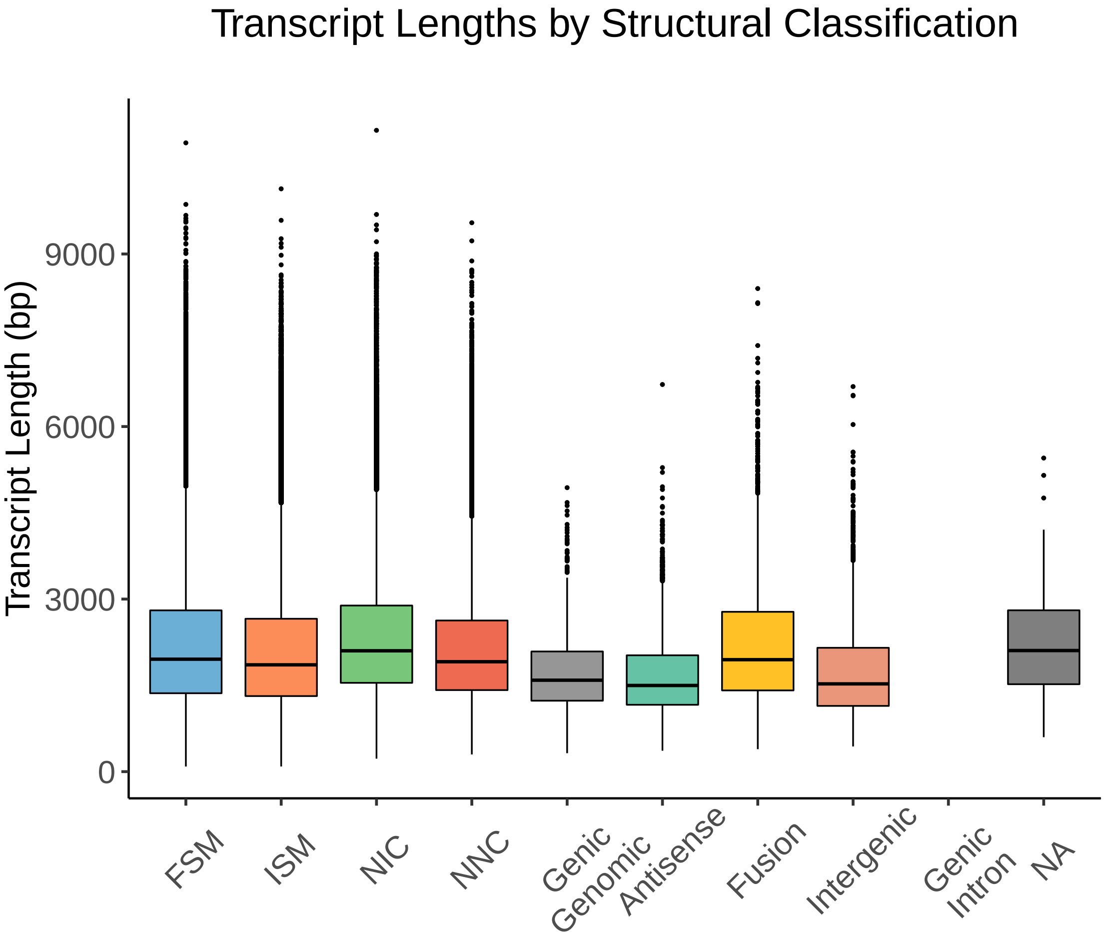

Isoform Distribution Across NIC
Transcripts, %
0
10
20
30
40
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Coding prediction
Coding


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

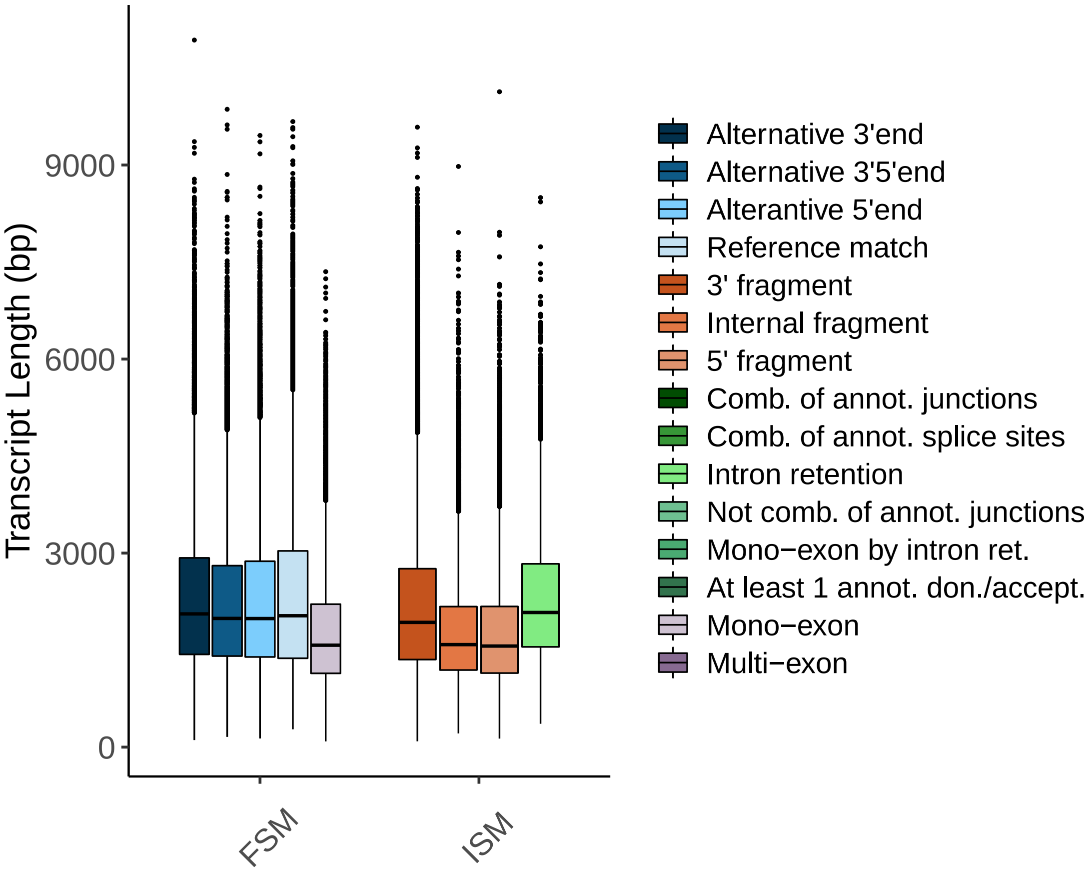

Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding

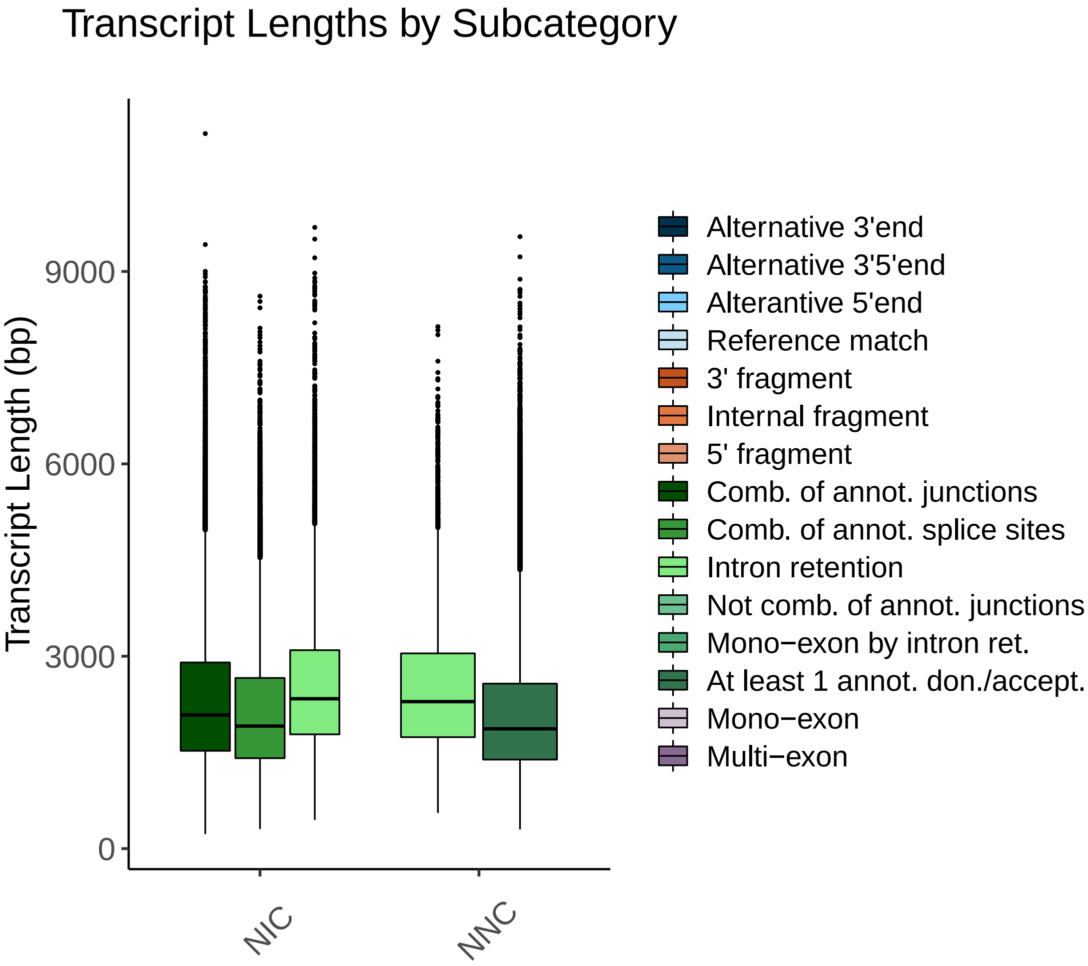

Isoform Distribution Across Fusion
Transcripts, %
0
20
40
60
80
Intron retention
Multi−exon
Coding prediction
Coding

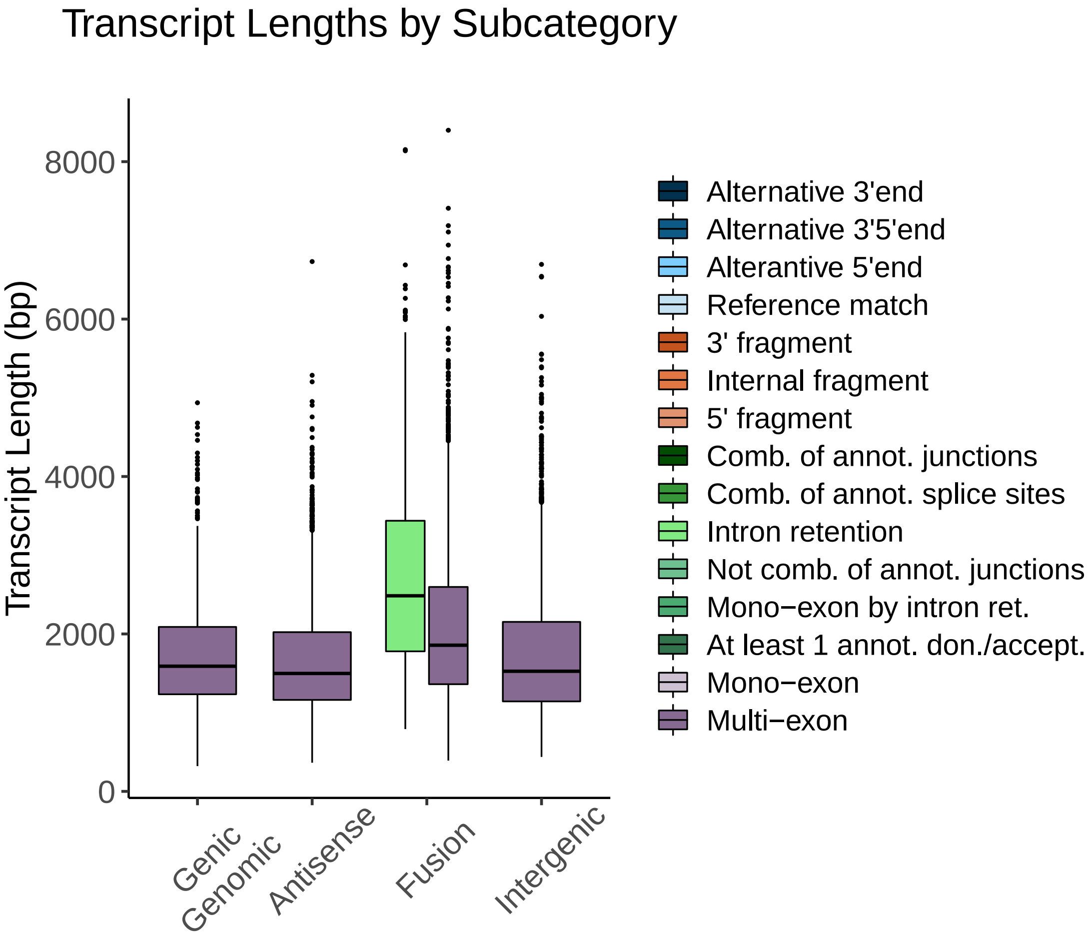

Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding

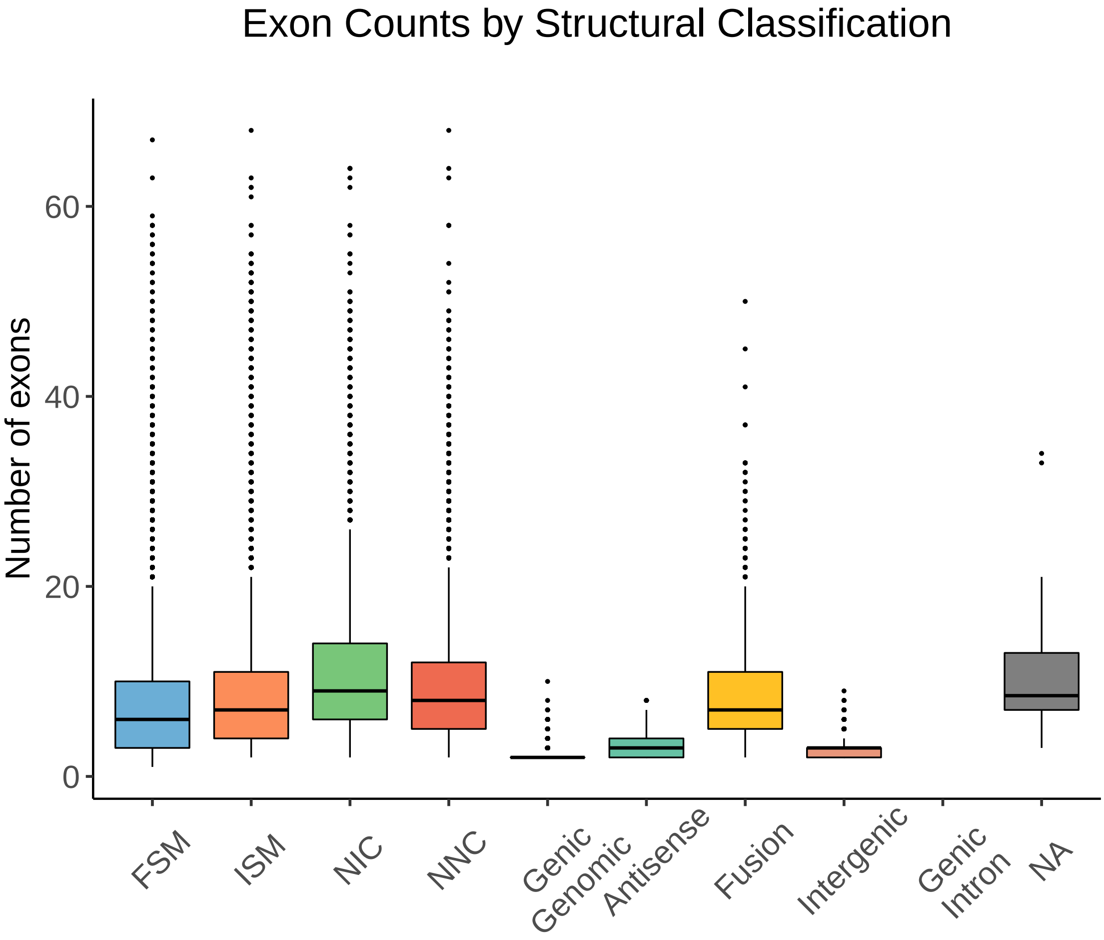

Transcript Lengths by Structural Classification
Transcript Length (bp)
9000
6000
3000
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

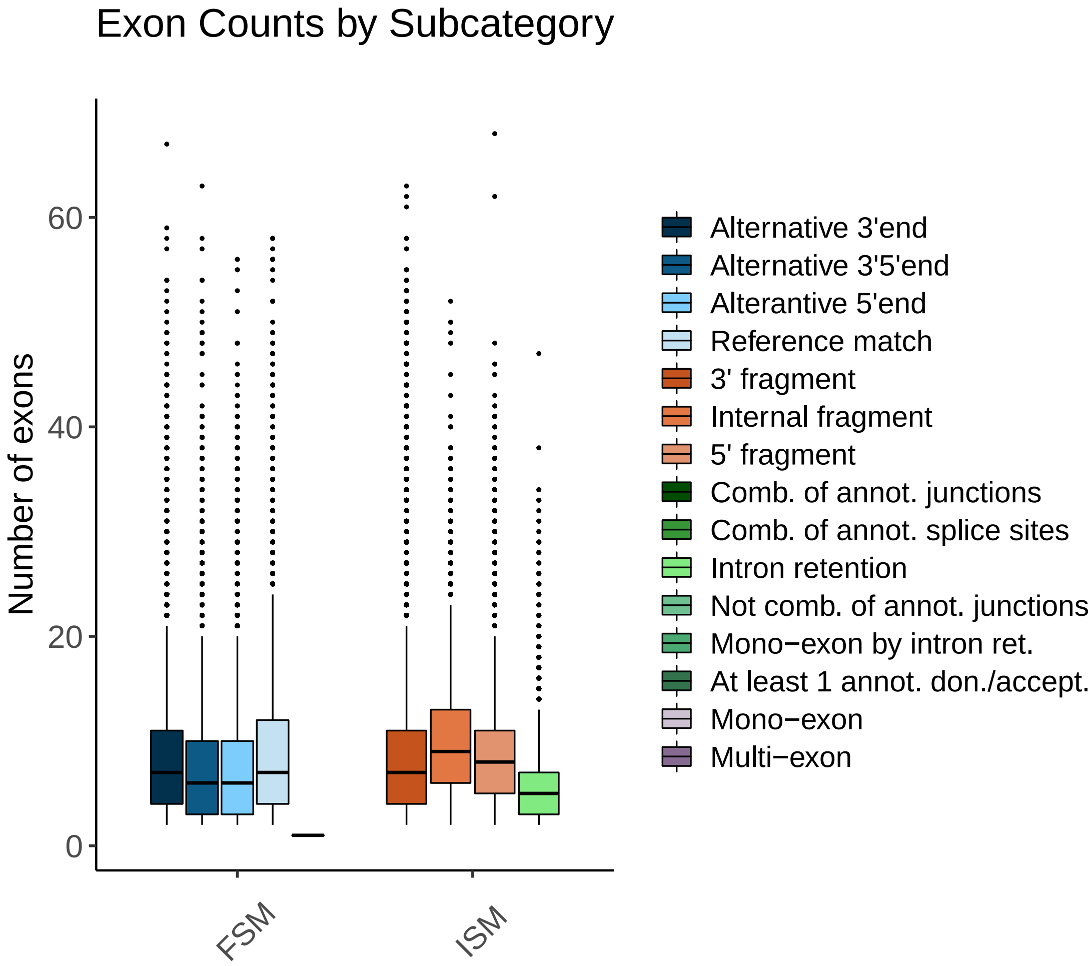
Transcript Lengths by Subcategory

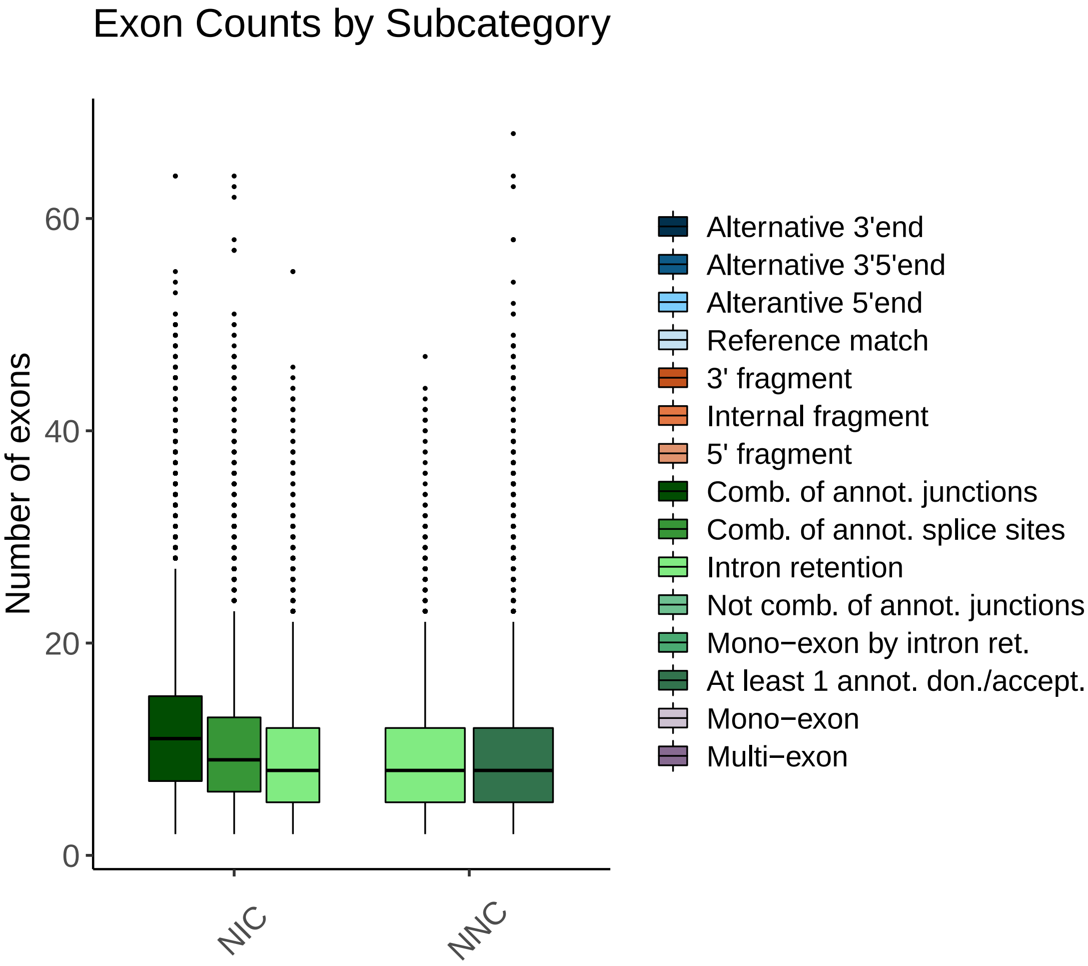


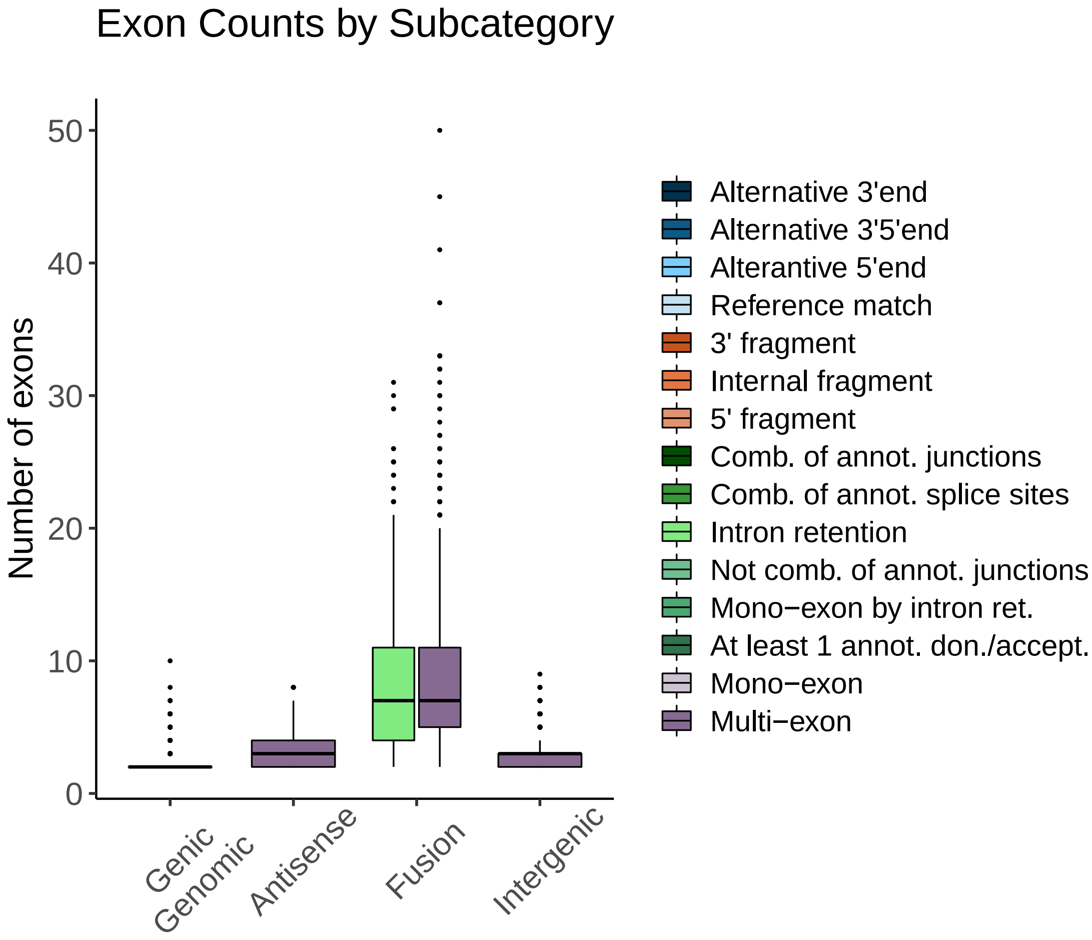

Transcript Lengths by Subcategory
Transcript Length (bp)
0
3000
6000
9000
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon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

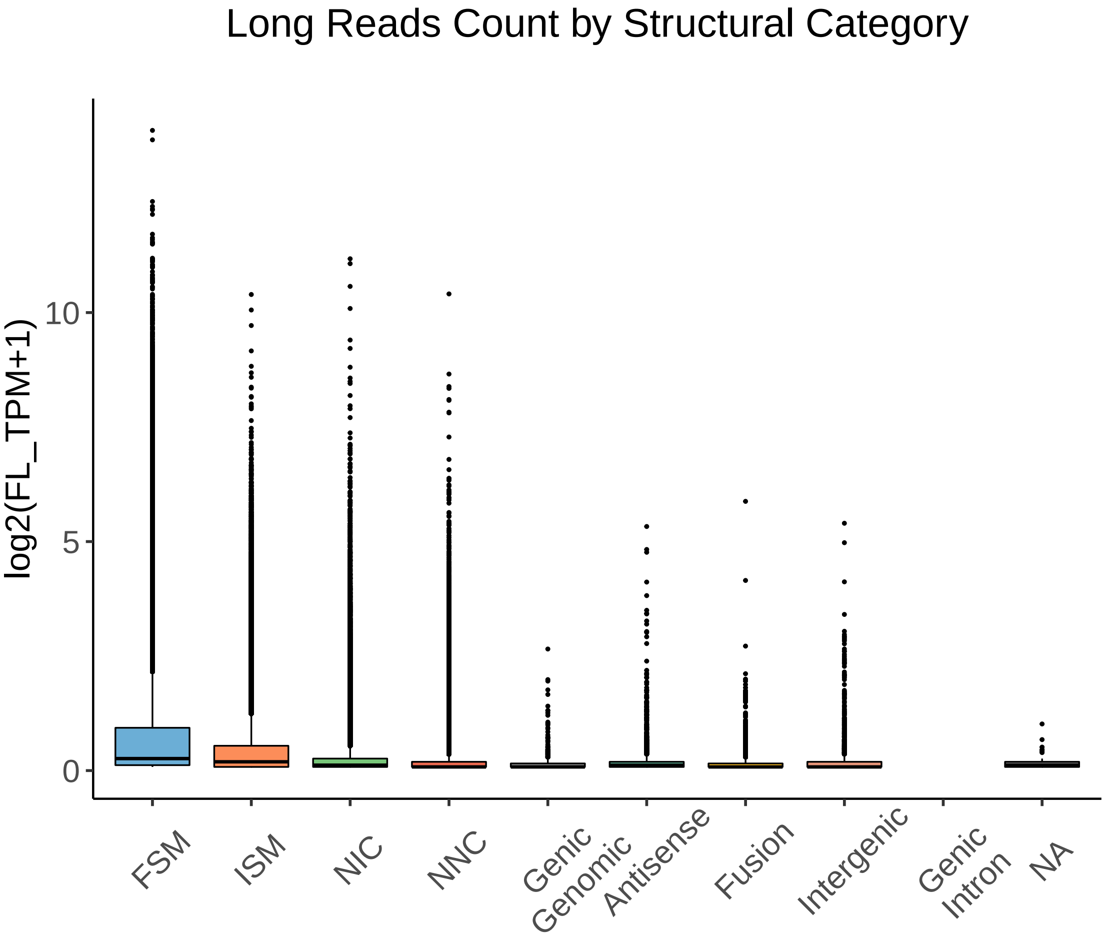

Transcript Lengths by Subcategory
Transcript Length (bp)
8000
6000
4000
2000
0
Genic
Genomic
Antisense
Fusion
Intergenic
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention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

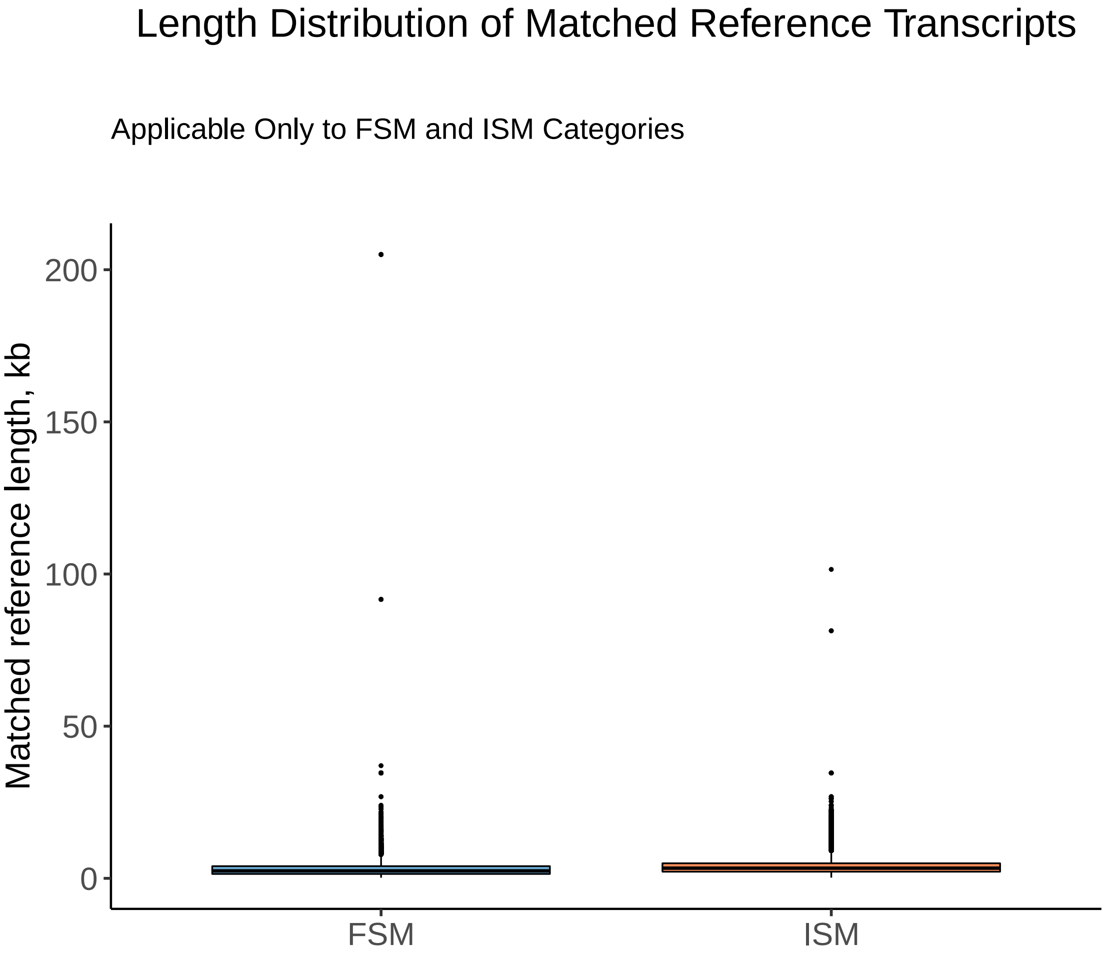

Exon Counts by Structural Classification
Number of exons
0
20
40
60
FSM
ISM
NIC
NNC
Genic
Genomic
Antisense
Fusion
Intergenic
Genic
Intron
NA


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
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

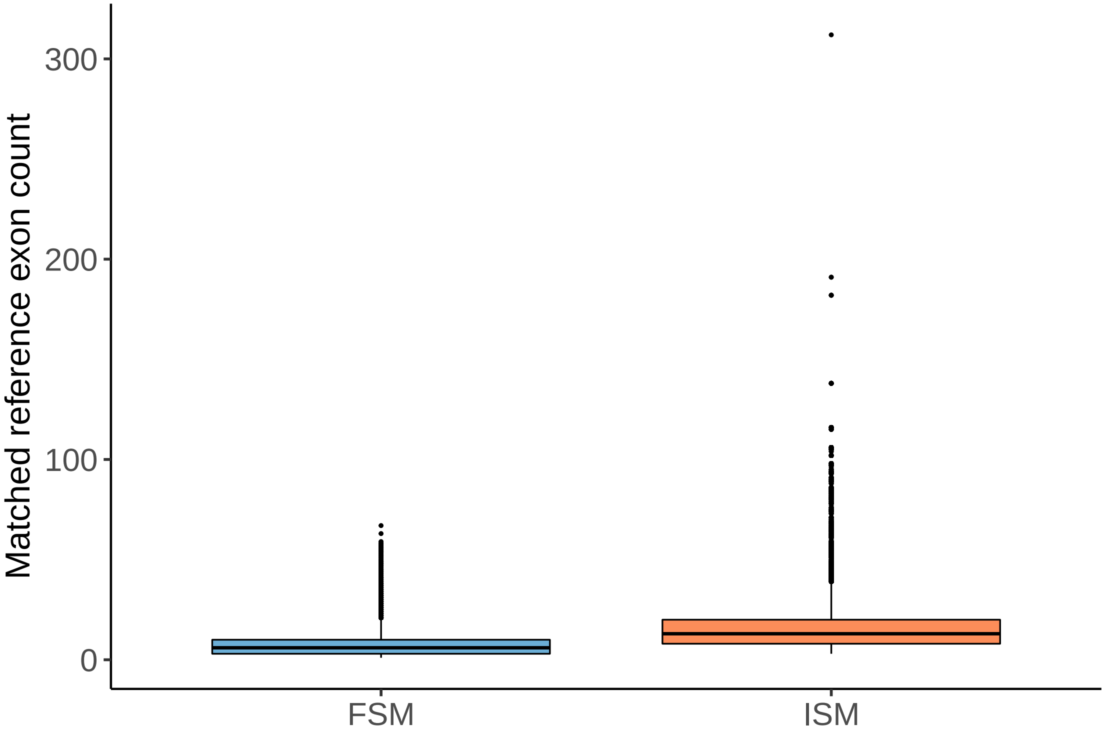


Length Distribution of Matched Reference Transcripts

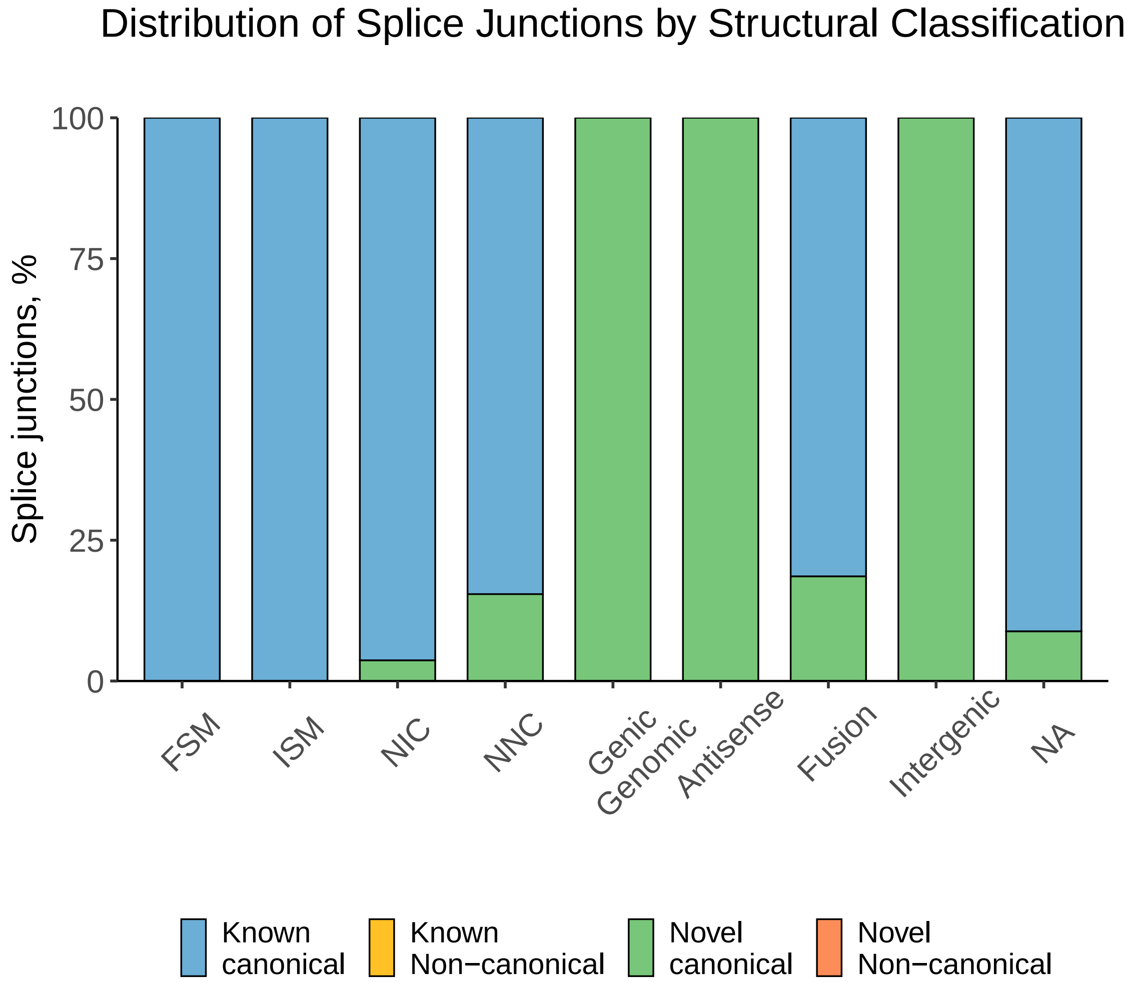
Applicable Only to FSM and ISM Categories

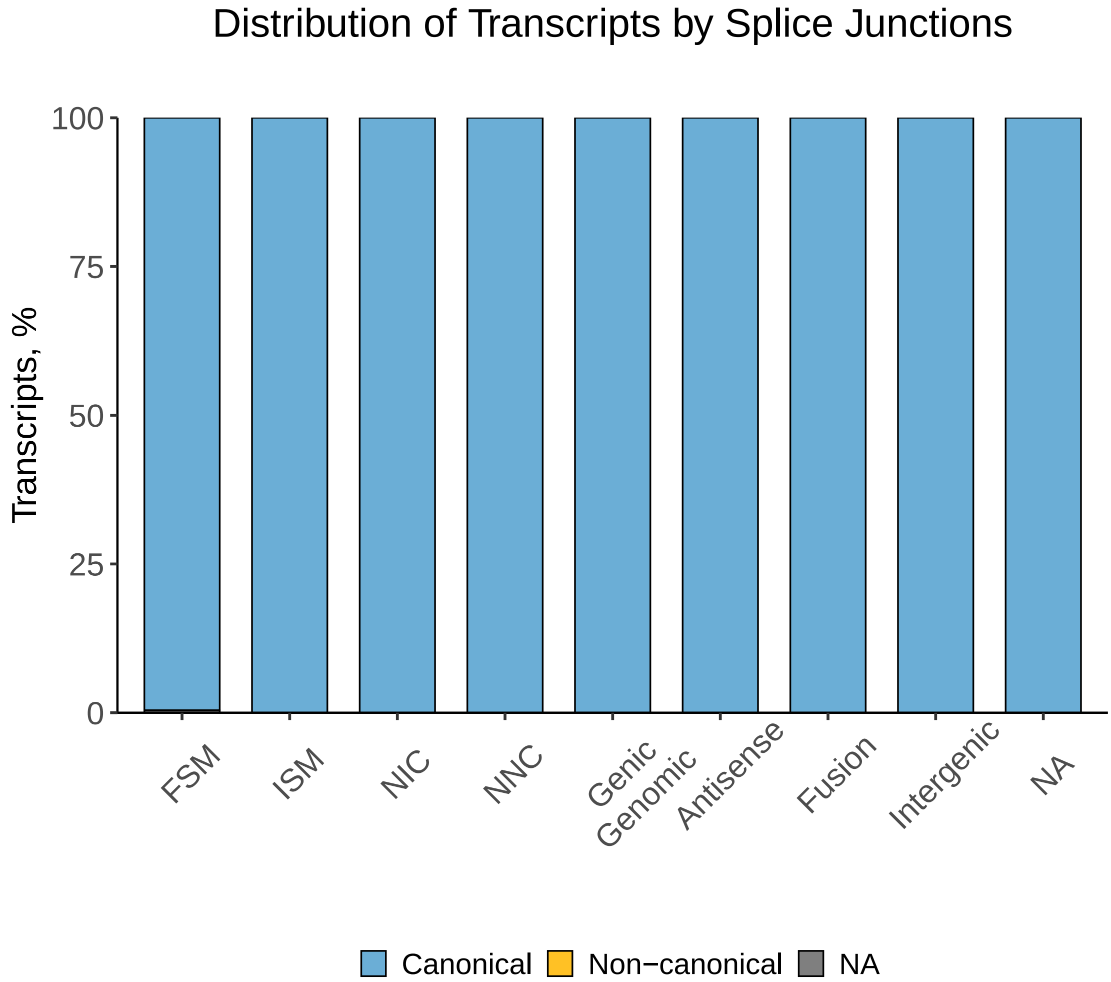
Exon Count Distribution of Matched Reference Transcripts

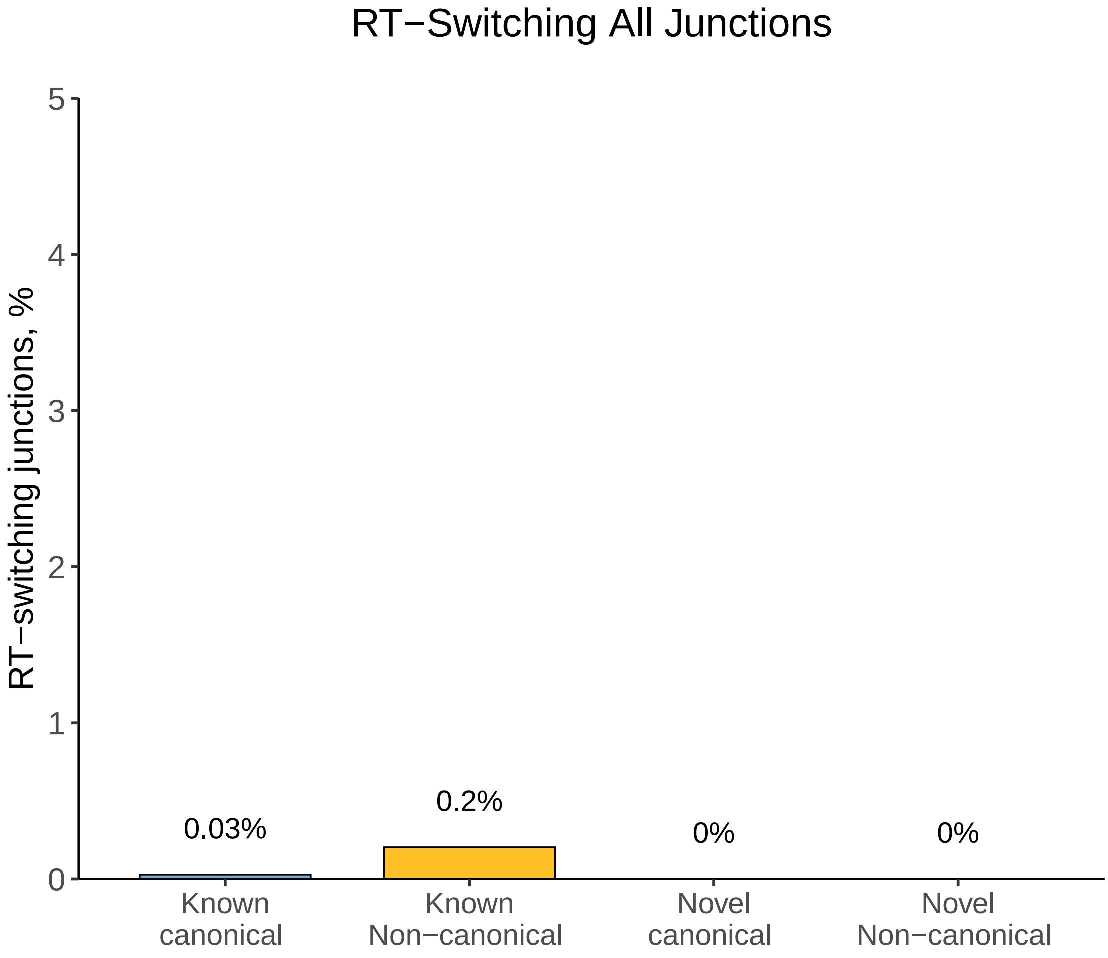
Applicable Only to FSM and ISM Categories

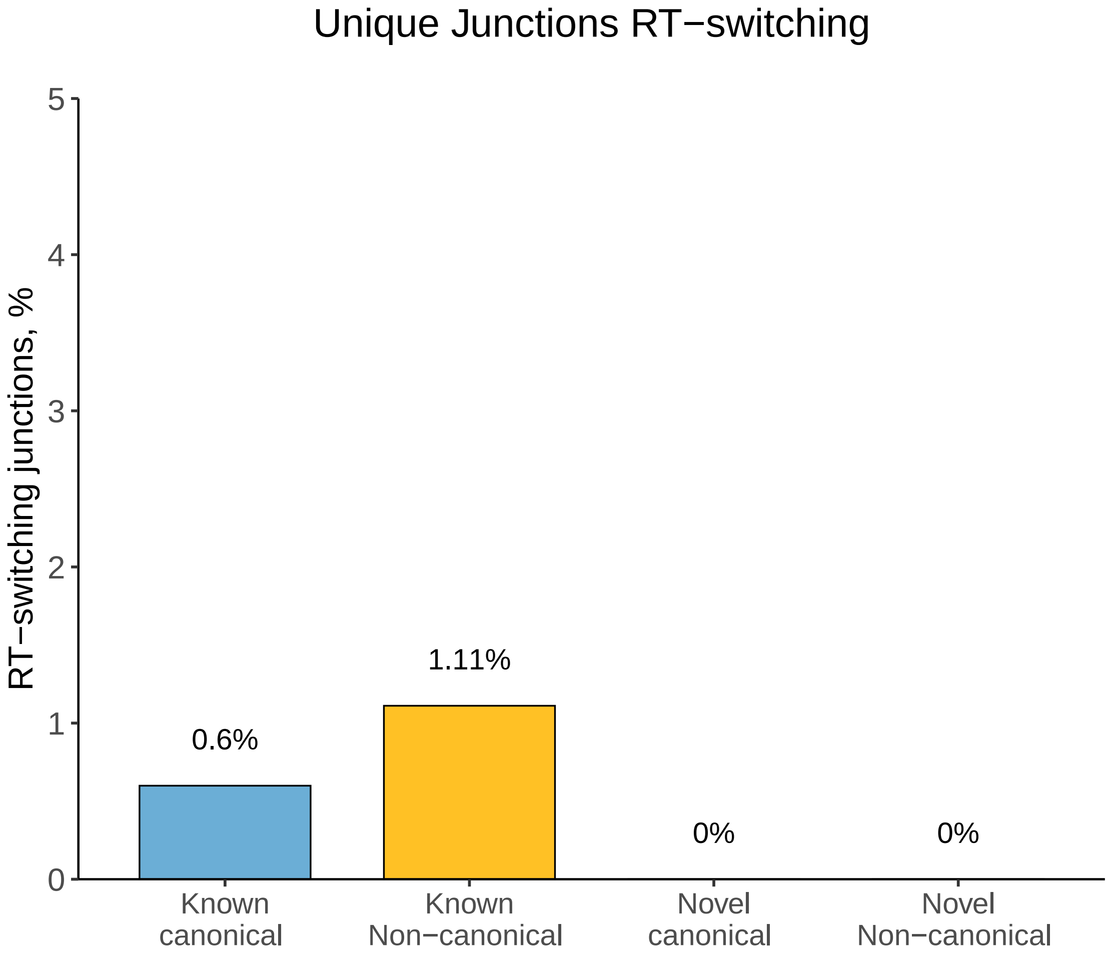


Splice Junction Characterization


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

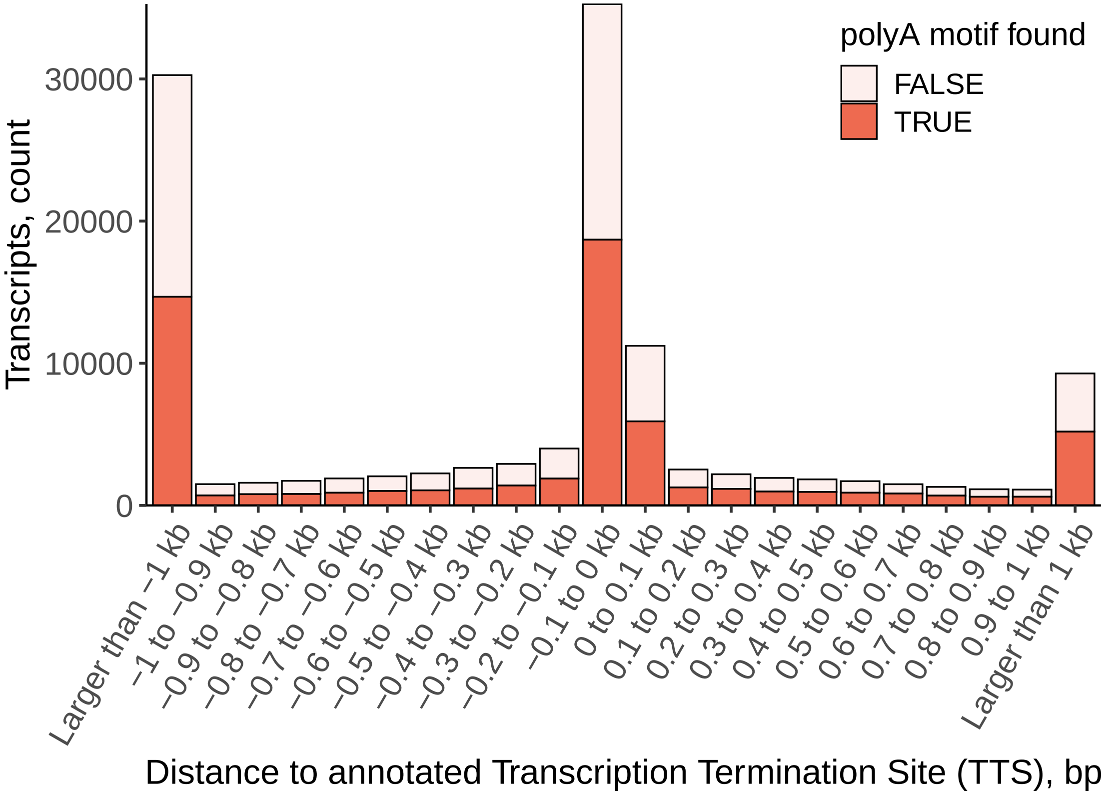

RT−Switching All Junctions
RT−switching junctions, %
5
4
3
2
1
0
0.03%
0.2%
0%
0%
Known
canonical
Known
Non−canonical
Novel
canonical
Novel
Non−canonical

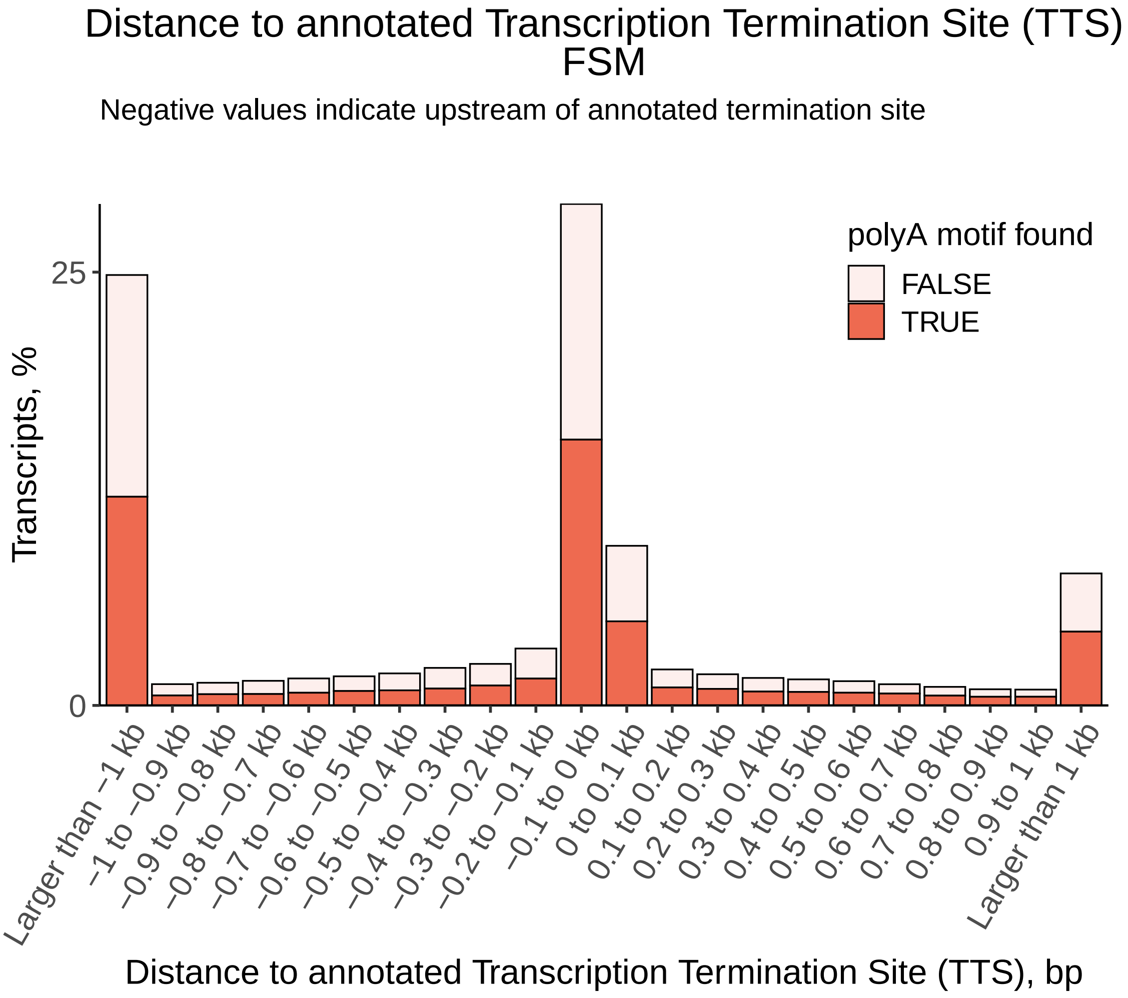

Unique Junctions RT−switching
RT−switching junctions, %
0
1
2
3
4
5
0.6%
1.11%
0%
0%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

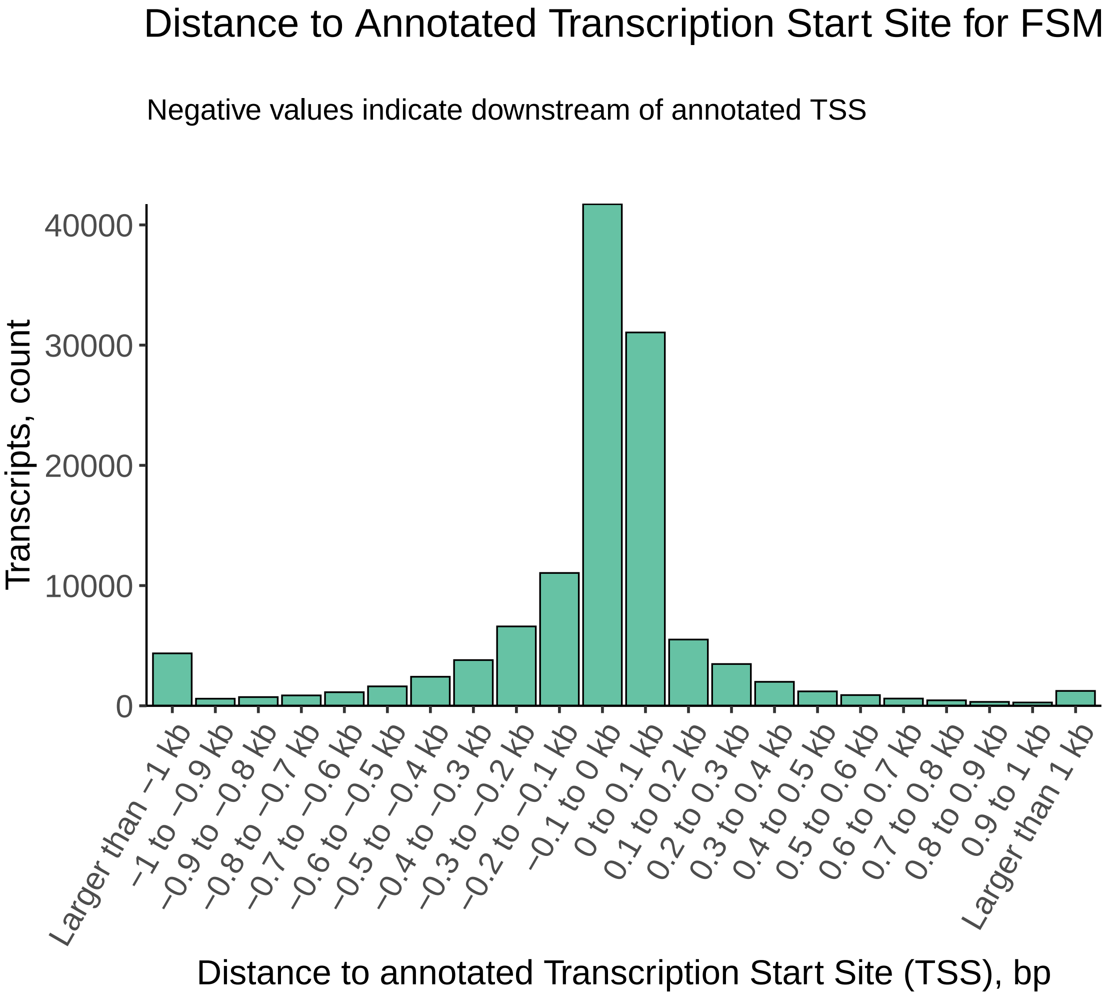
Comparison With Annotated TSS and TTS

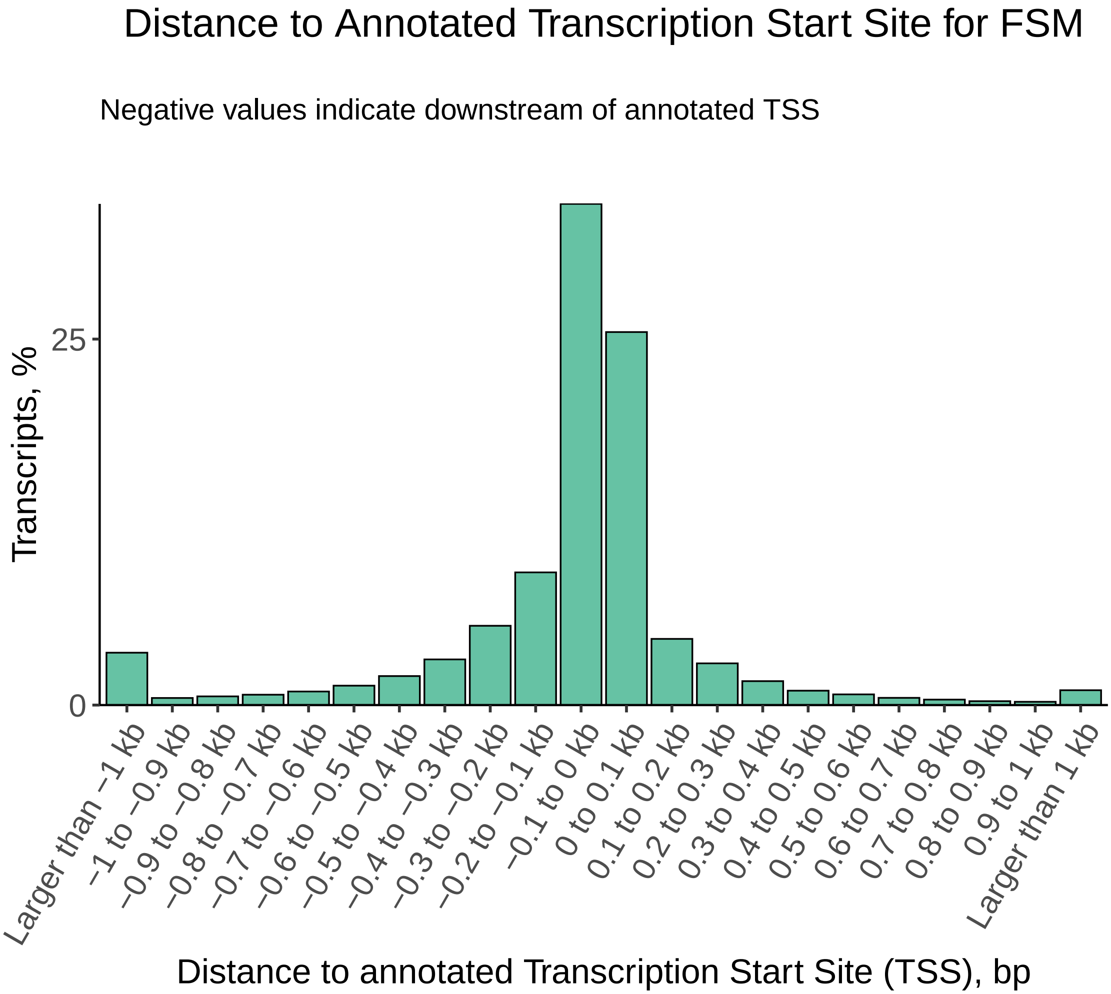
Distance to annotated Transcription Termination Site (TTS)
FSM

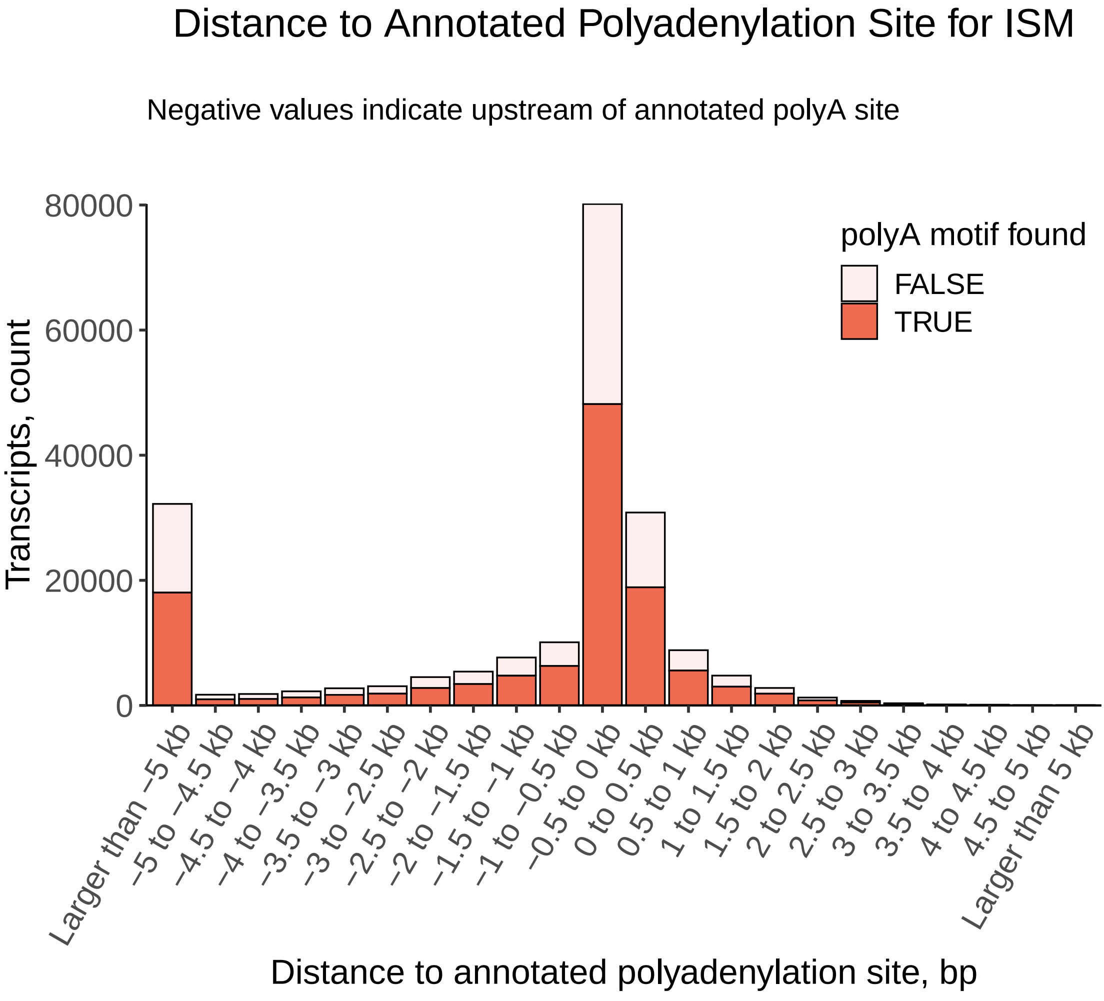
Negative values indicate upstream of annotated termination site

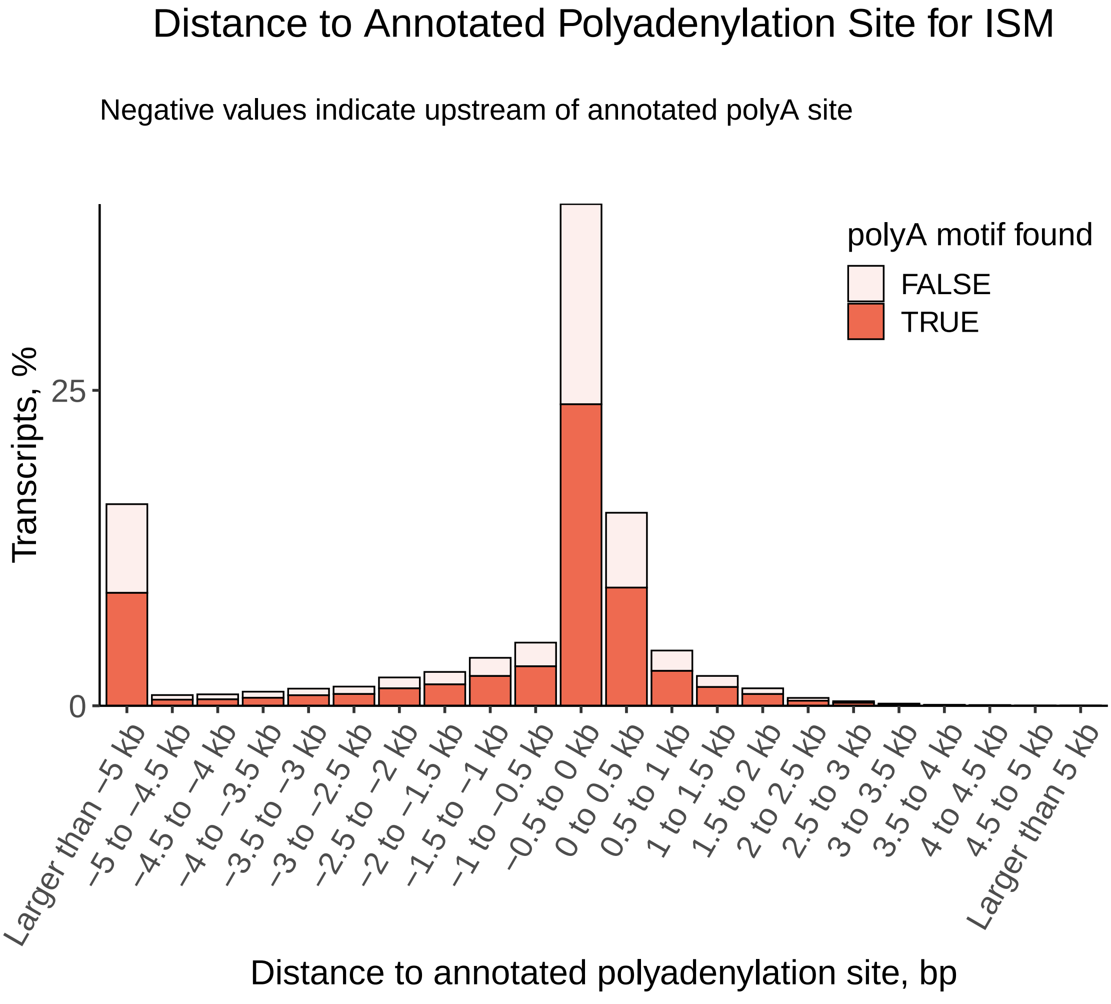



Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3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

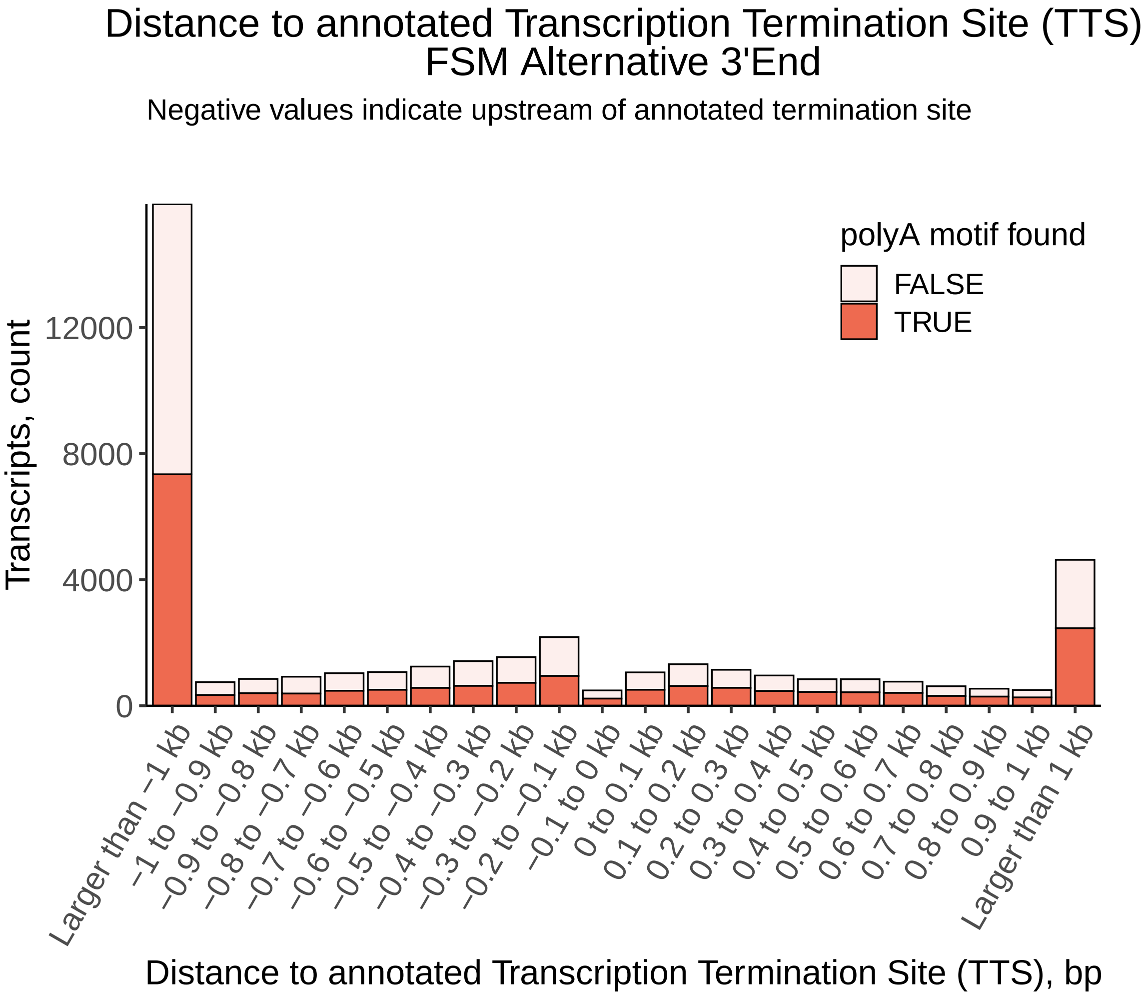

Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
Transcripts, count
40000
30000
20000
10000
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2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

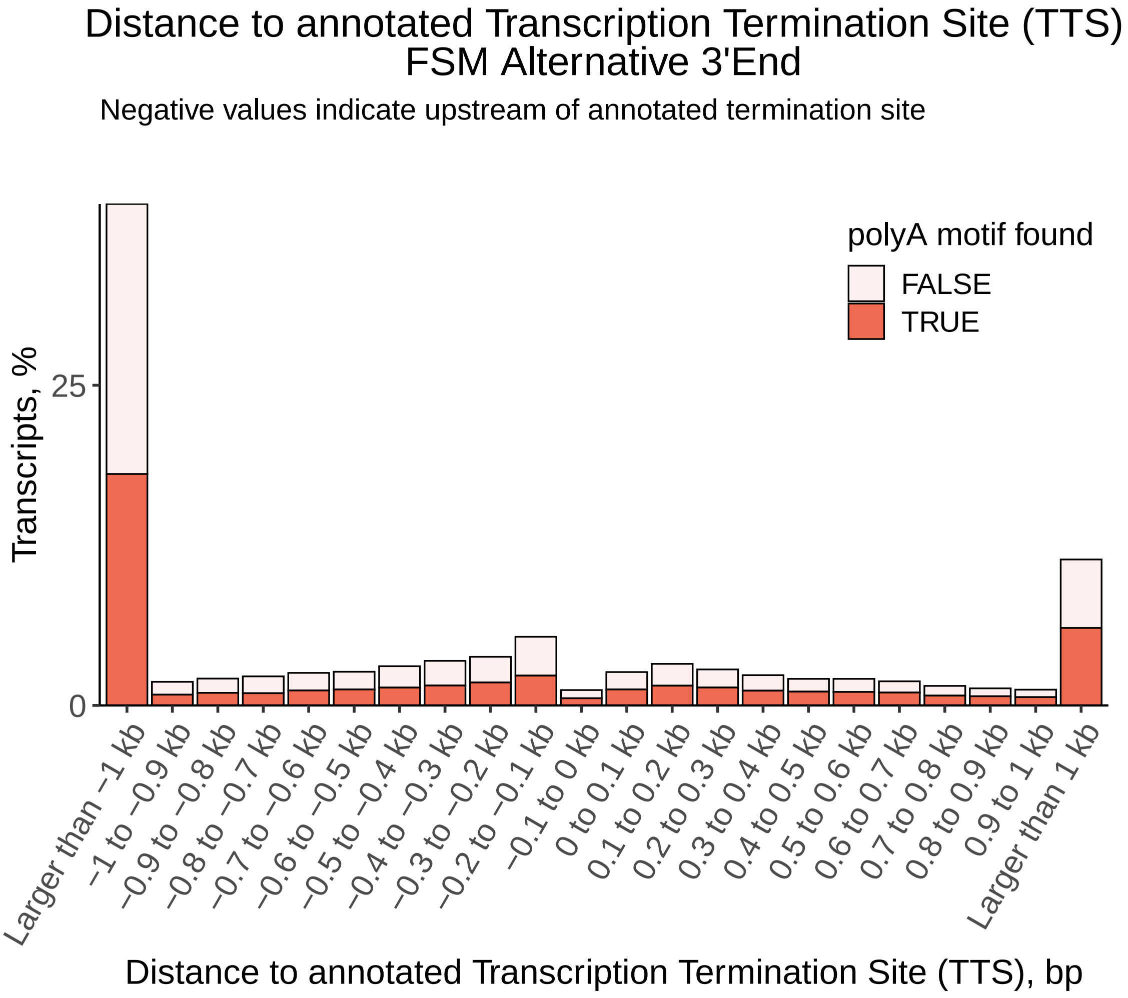

Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb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

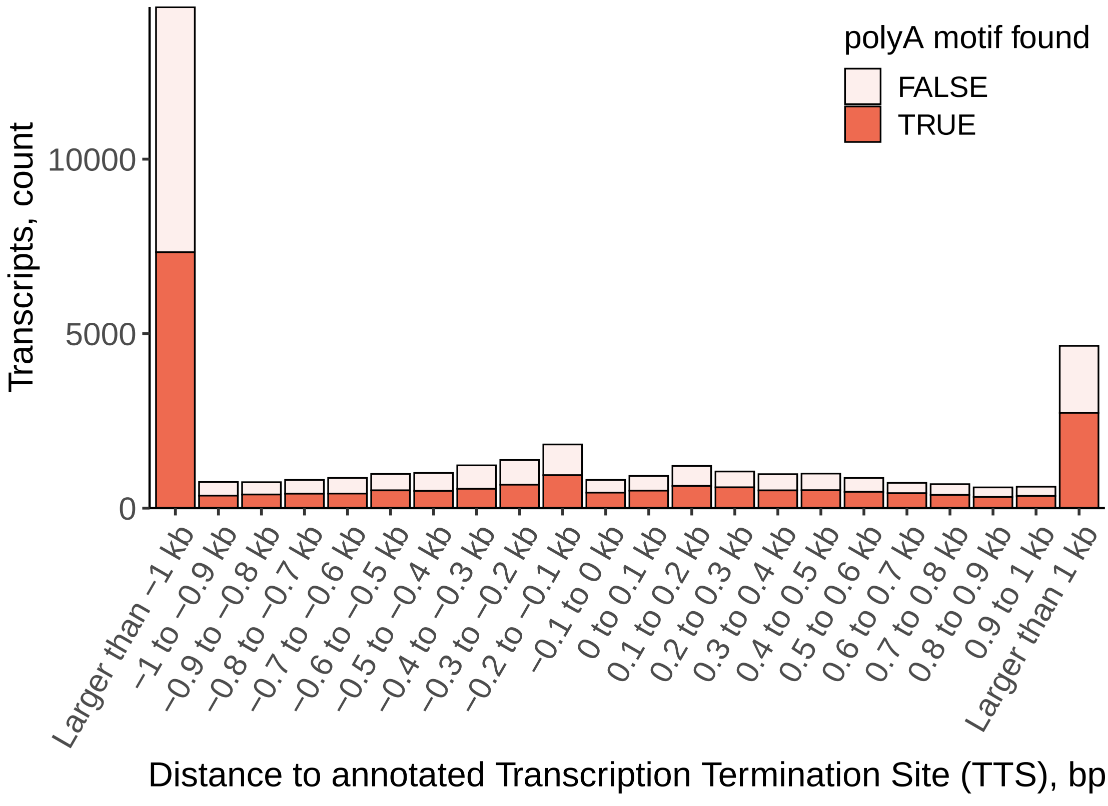
Comparison With Annotated TSS and TTS by Subcategories

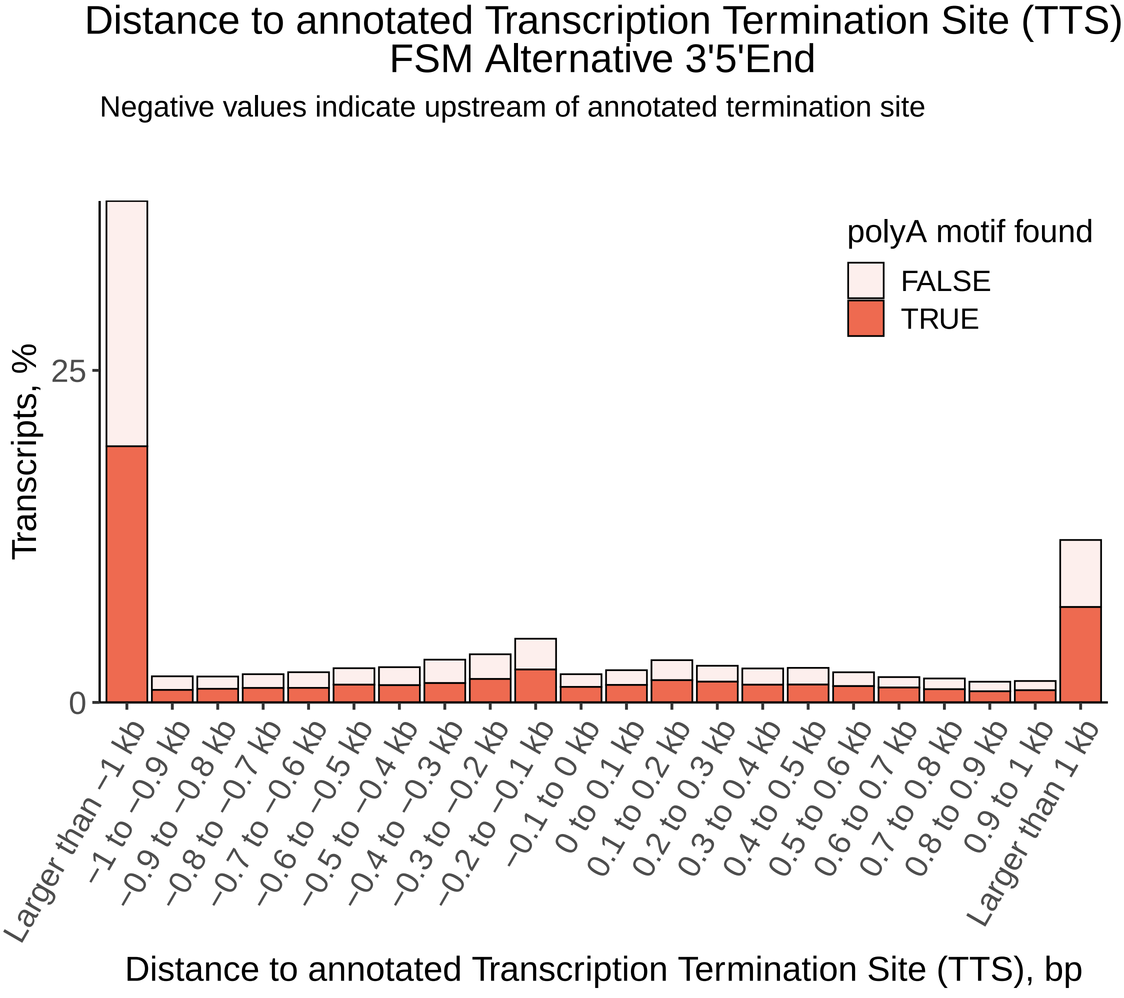

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
4000
8000
12000
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 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
Transcripts, %
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 Transcription Termination Site (TTS)
FSM Alternative 3'5'End

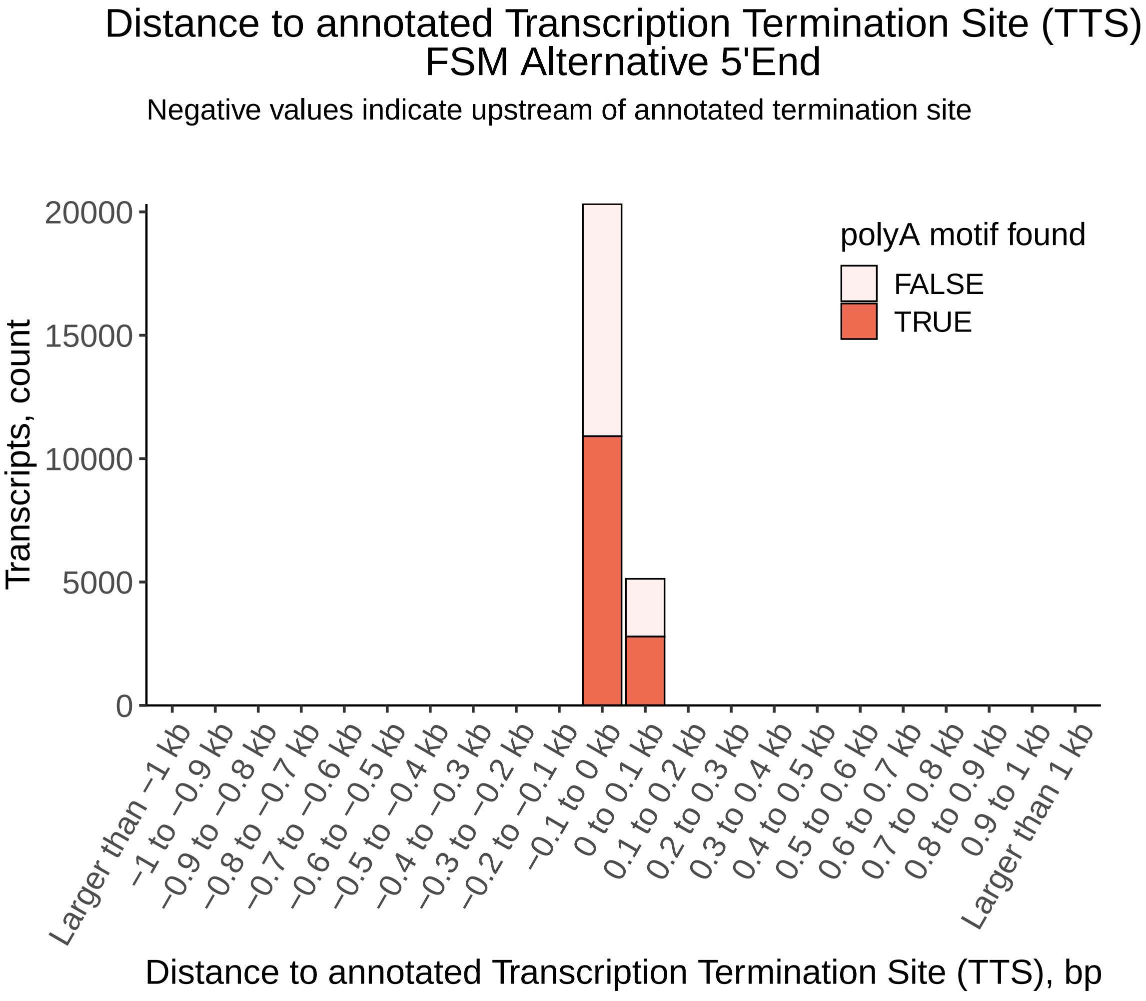
Negative values indicate upstream of annotated termination site

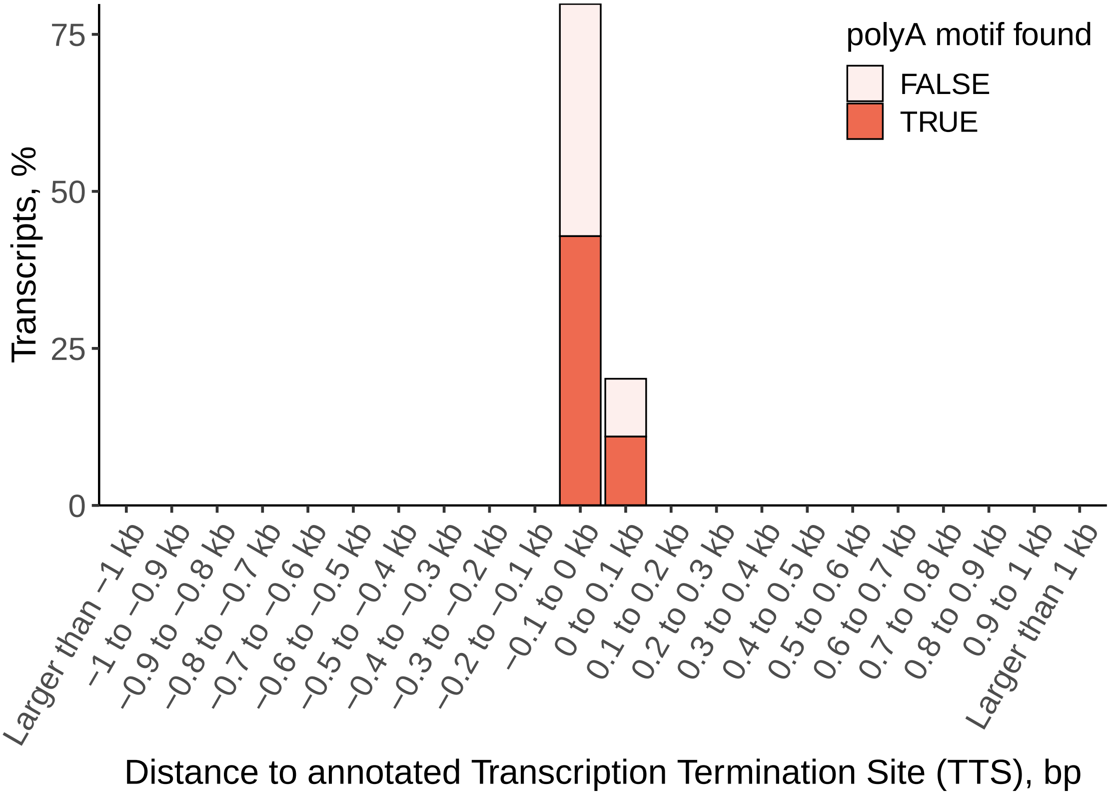




Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

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
20000
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

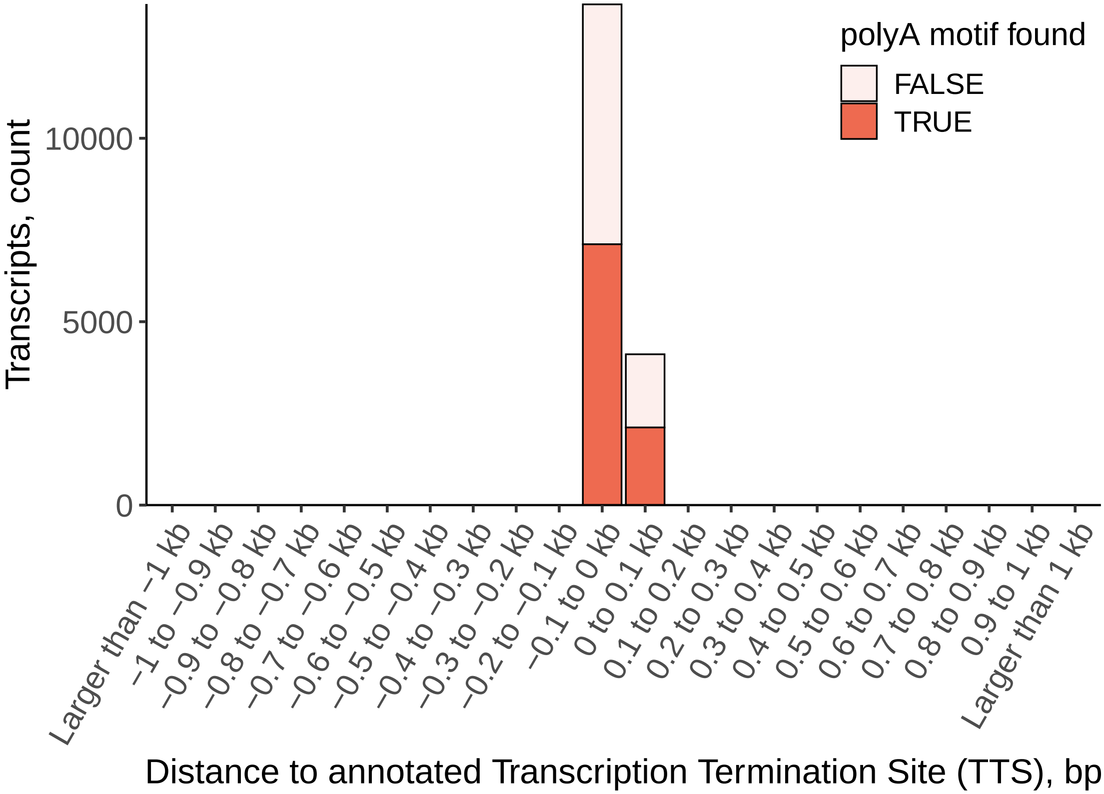


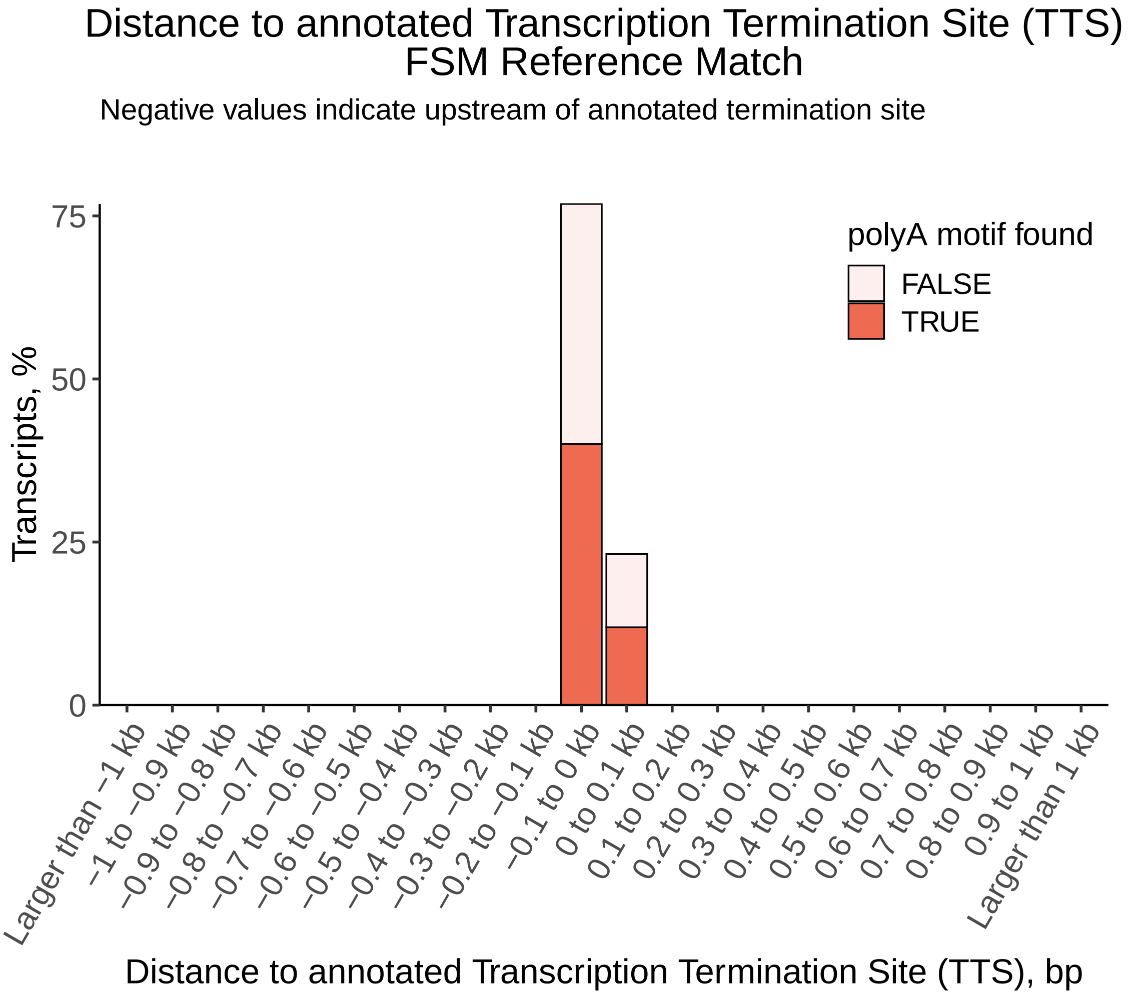
Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End

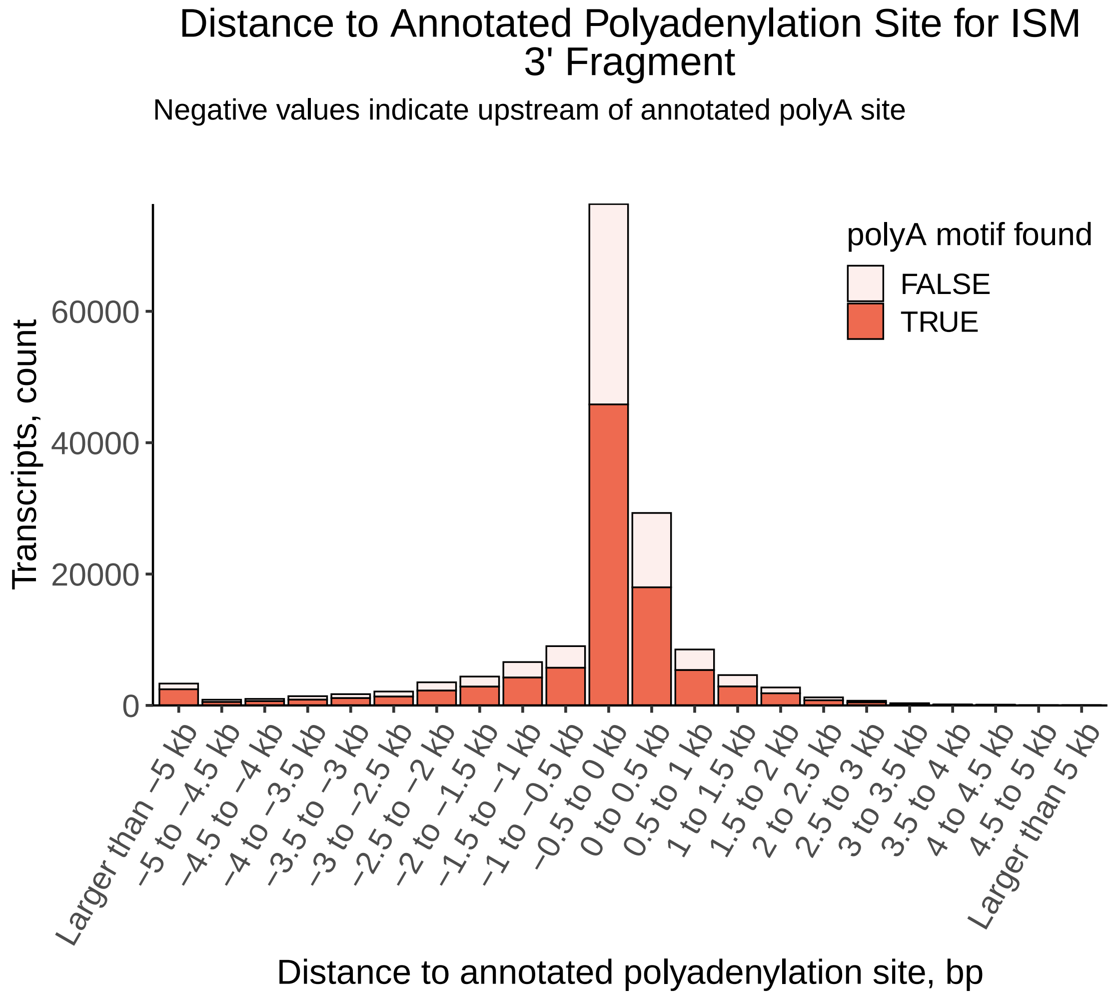
Negative values indicate upstream of annotated termination site

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

Negative values indicate upstream of annotated termination site

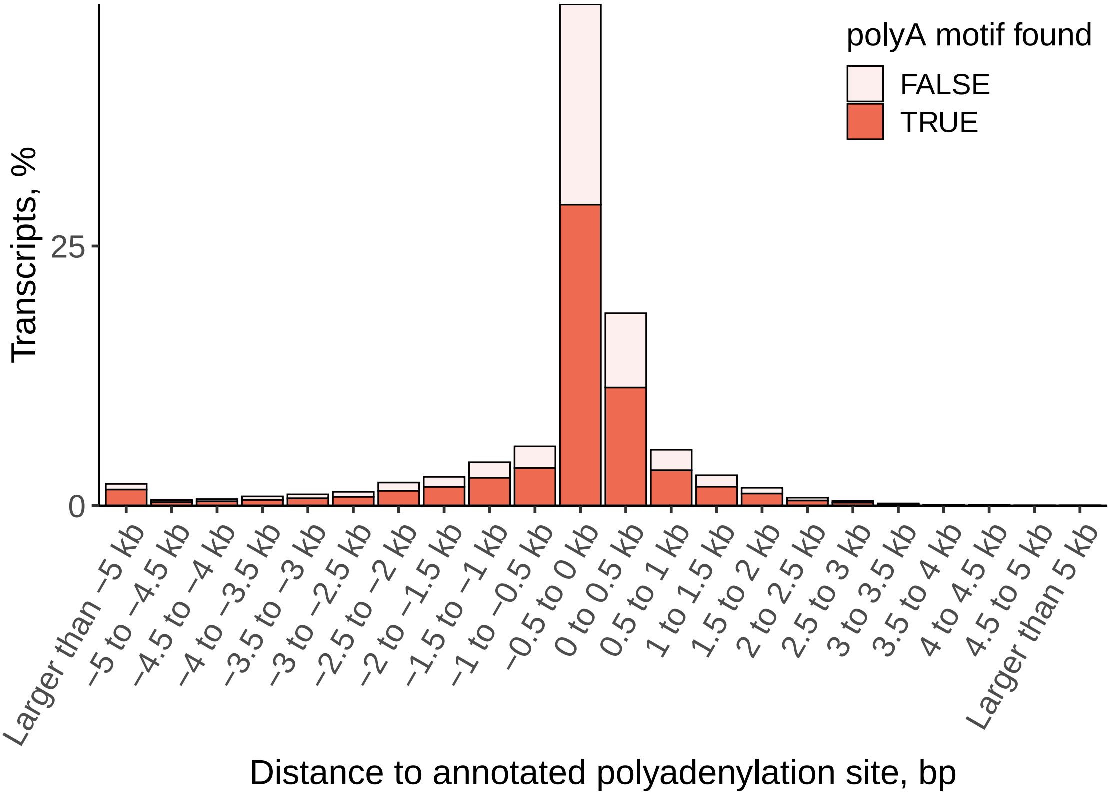



Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, %
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to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

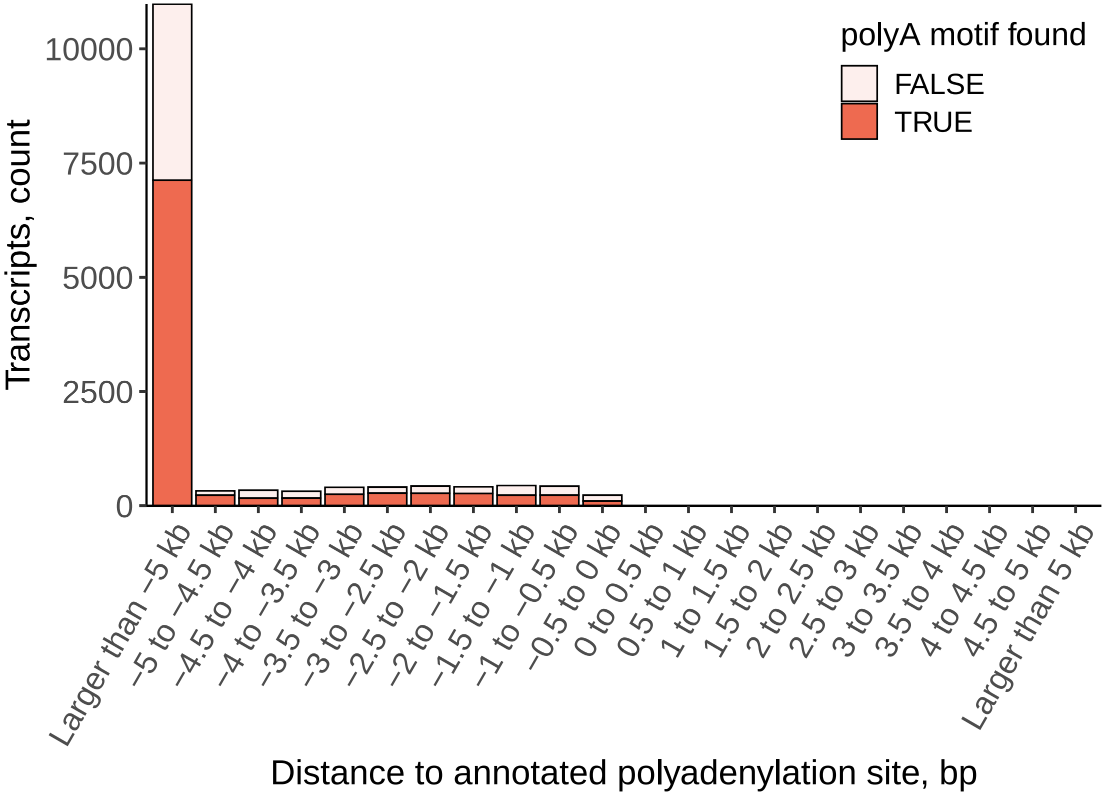
Distance to Annotated Polyadenylation Site for ISM 3' Fragment

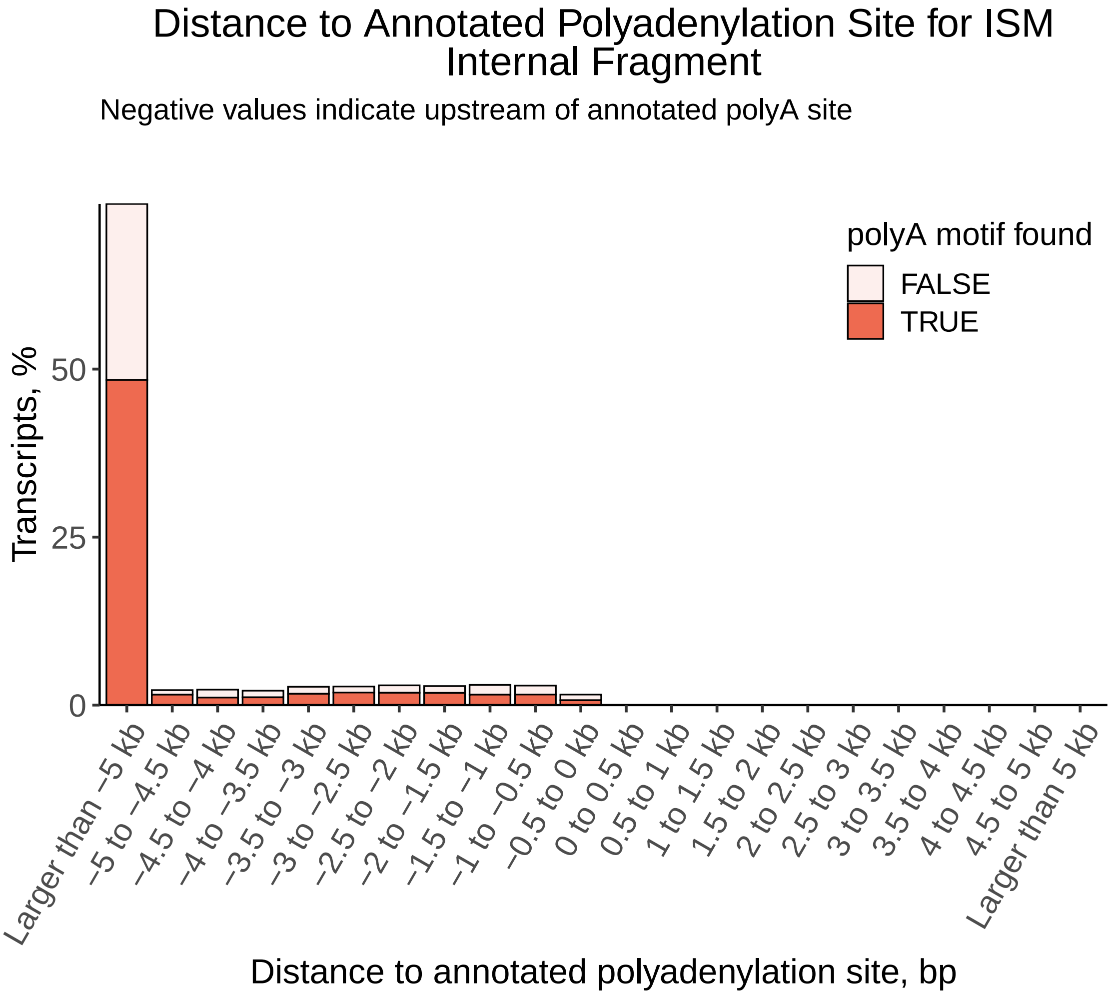
Negative values indicate upstream of annotated polyA site

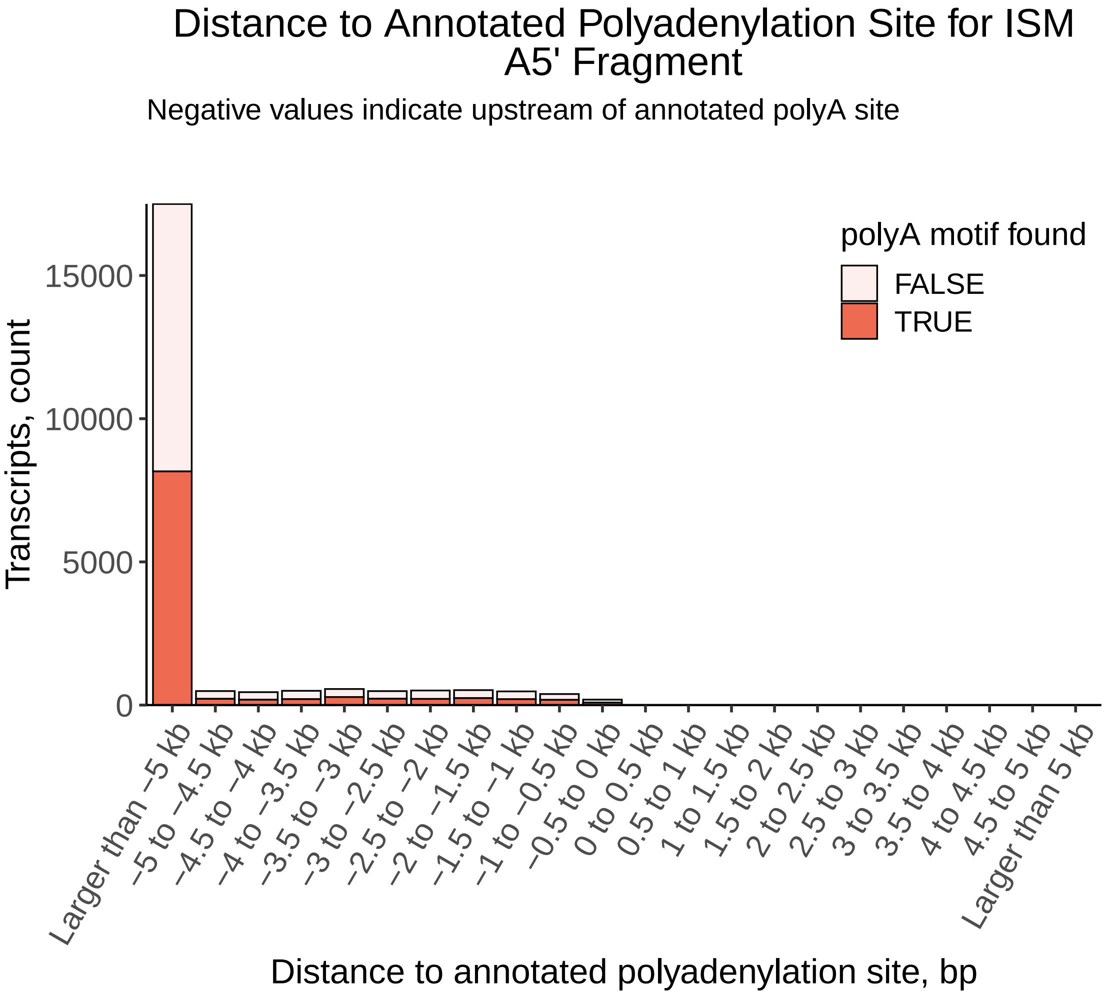


Distance to Annotated Polyadenylation Site for ISM Internal Fragment

Negative values indicate upstream of annotated polyA site

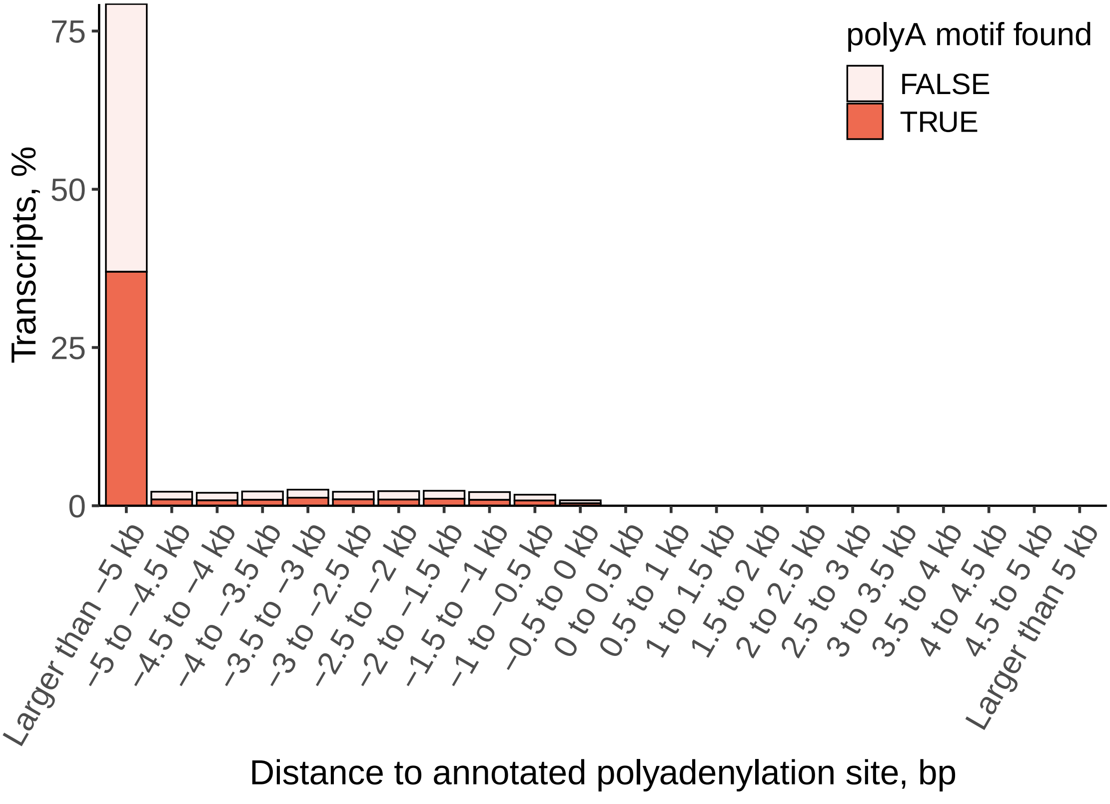


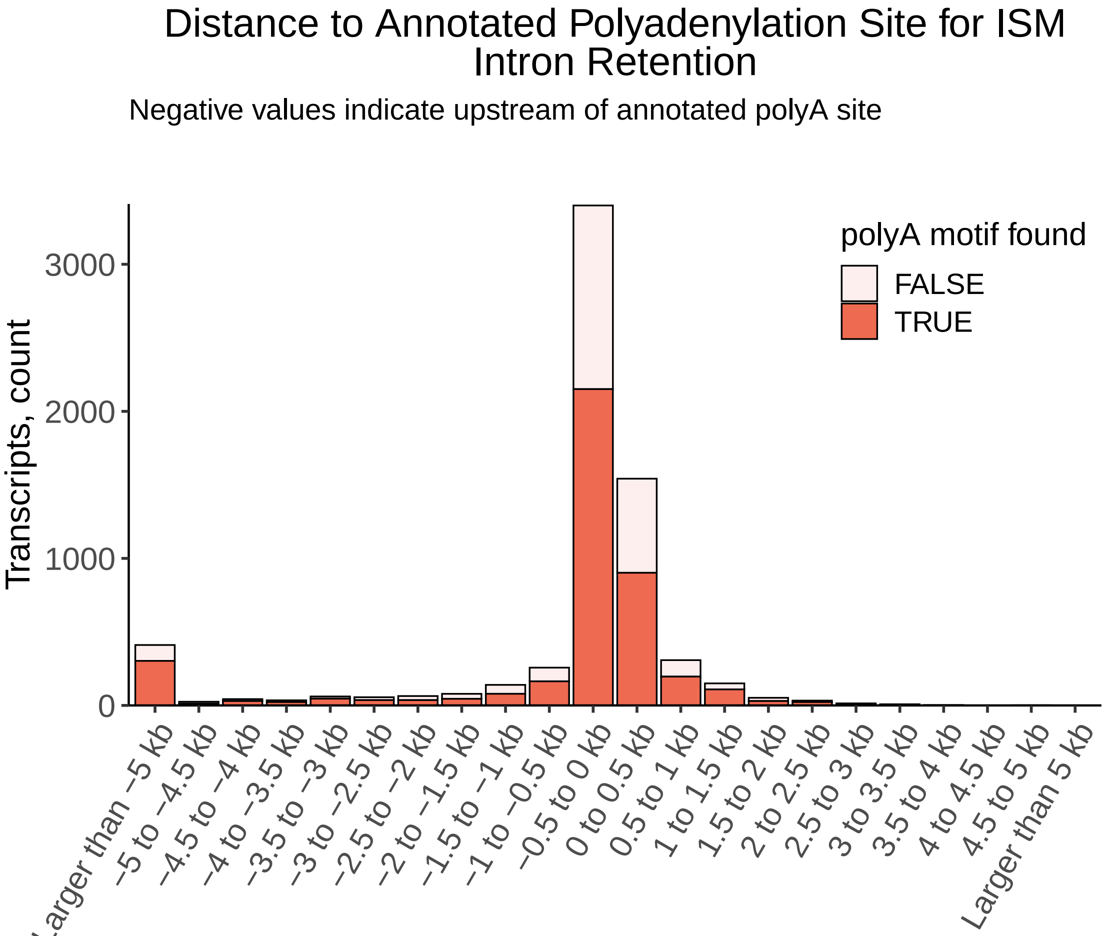

Distance to Annotated Polyadenylation Site for ISM Internal Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
50
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to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

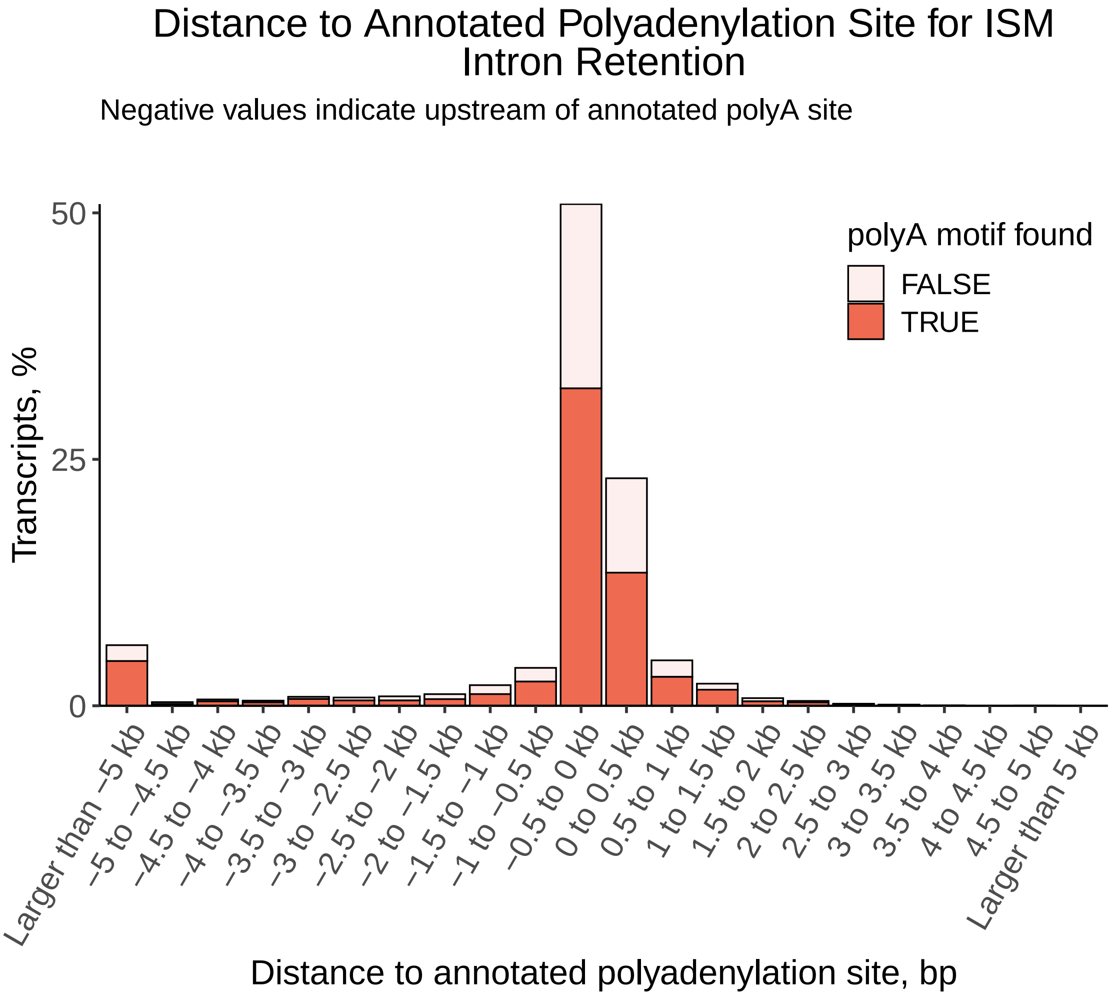

Distance to Annotated Polyadenylation Site for ISM A5' Fragment
Negative values indicate upstream of annotated polyA site
Transcripts, count
15000
10000
5000
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 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 A5' Fragment

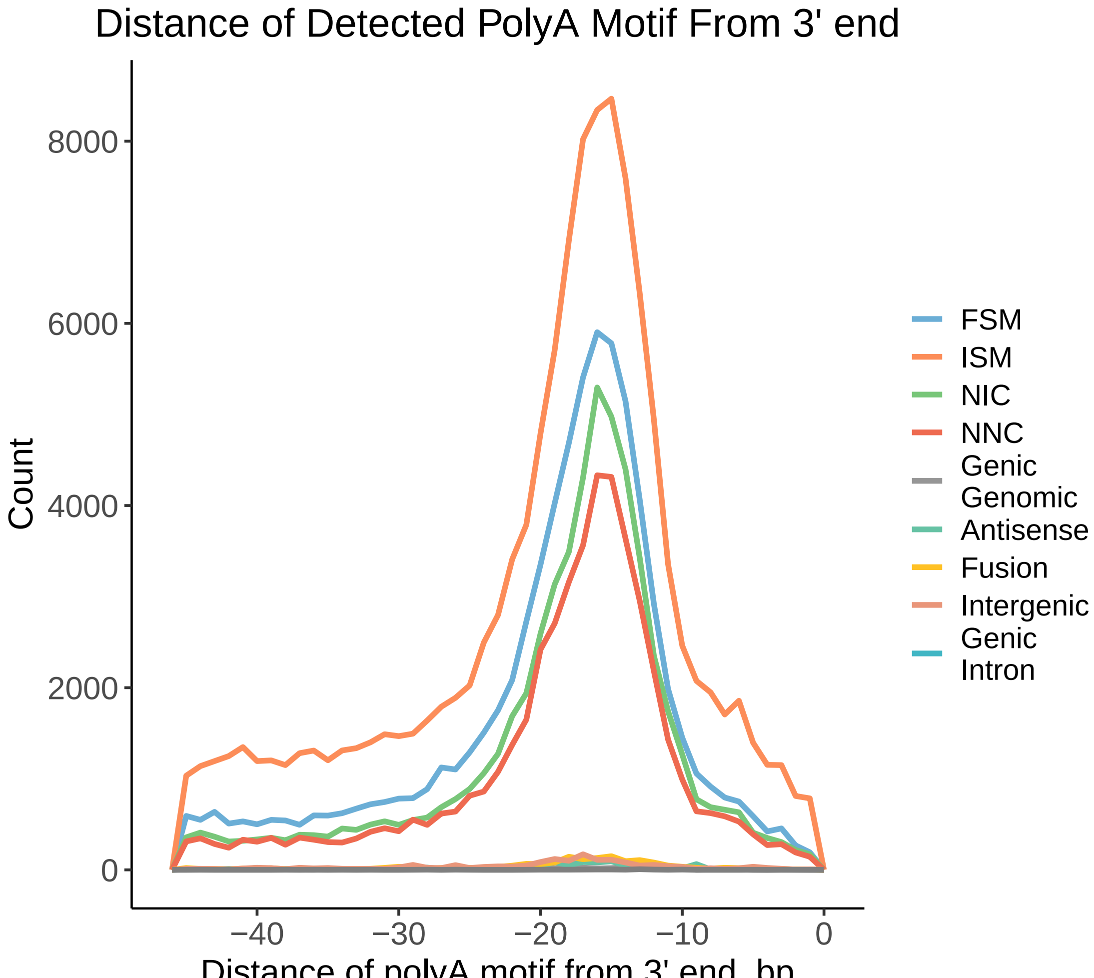
Negative values indicate upstream of annotated polyA site




Distance to Annotated Polyadenylation Site for ISM Intron Retention
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
0
1000
2000
3000
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

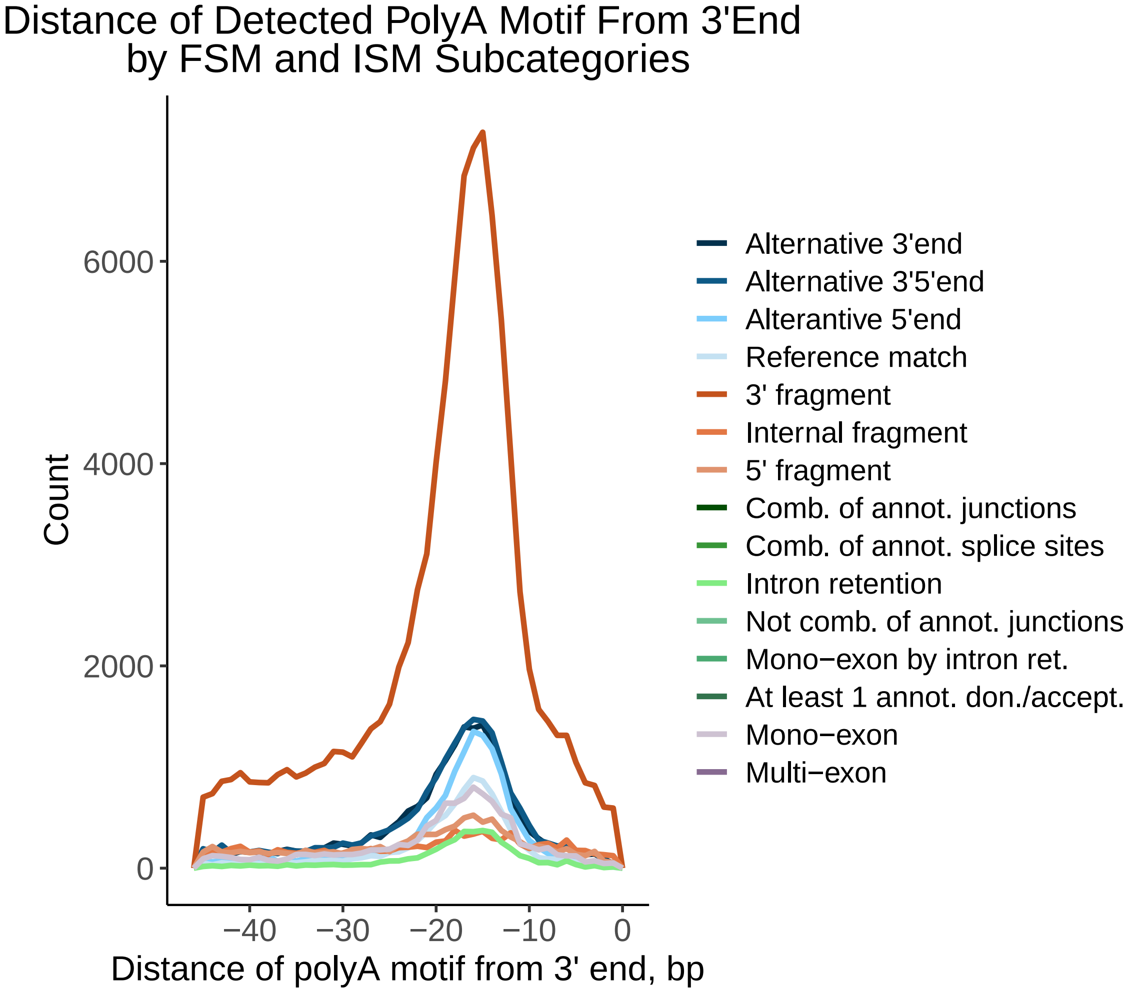

Distance of Detected PolyA Motif From 3' end
Count
8000
6000
4000
2000
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

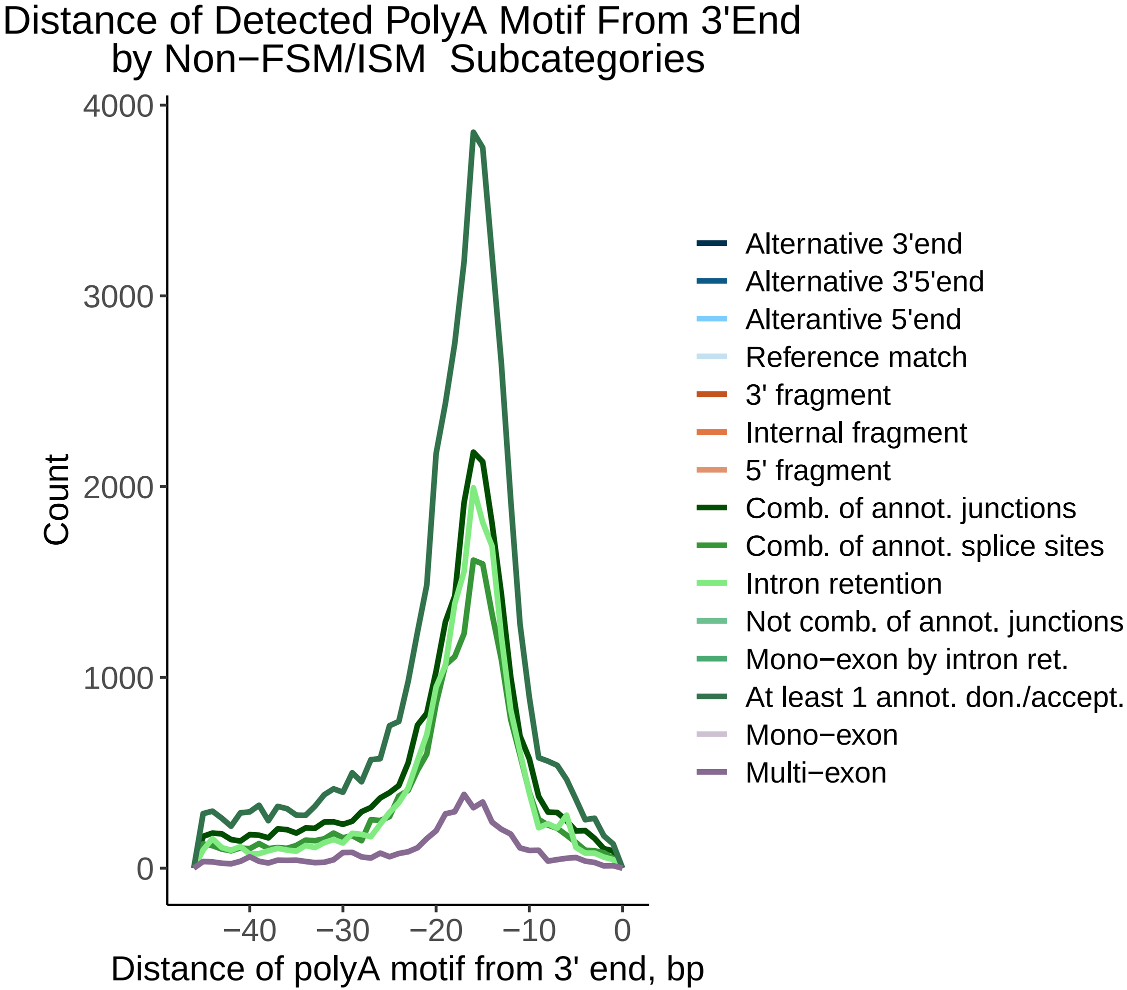
Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	139688	73058	52
ISM	201592	121443	60
NIC	111363	56938	51
NNC	91483	48812	53
Genic Genomic	623	320	51
Antisense	1822	1050	58
Fusion	2788	1640	59
Intergenic	2597	1637	63
NA	72	47	65

Frequency of PolyA Motifs

Motif	Count	%
AATAAA	152200	49.9
ATTAAA	44265	14.5
TATAAA	12075	4.0
AGTAAA	11289	3.7
AAAAAG	11141	3.7
AAGAAA	10987	3.6
GGGGCT	9225	3.0
TTTAAA	8022	2.6
CATAAA	6849	2.2
AATATA	6738	2.2
GATAAA	6524	2.1
AATACA	6476	2.1
AAAACA	6253	2.1
AATGAA	5903	1.9
AATAGA	3649	1.2
ACTAAA	3349	1.1


Distance of Detected PolyA Motif From 3'End
by FSM and ISM Subcategories
Count
6000
4000
2000
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

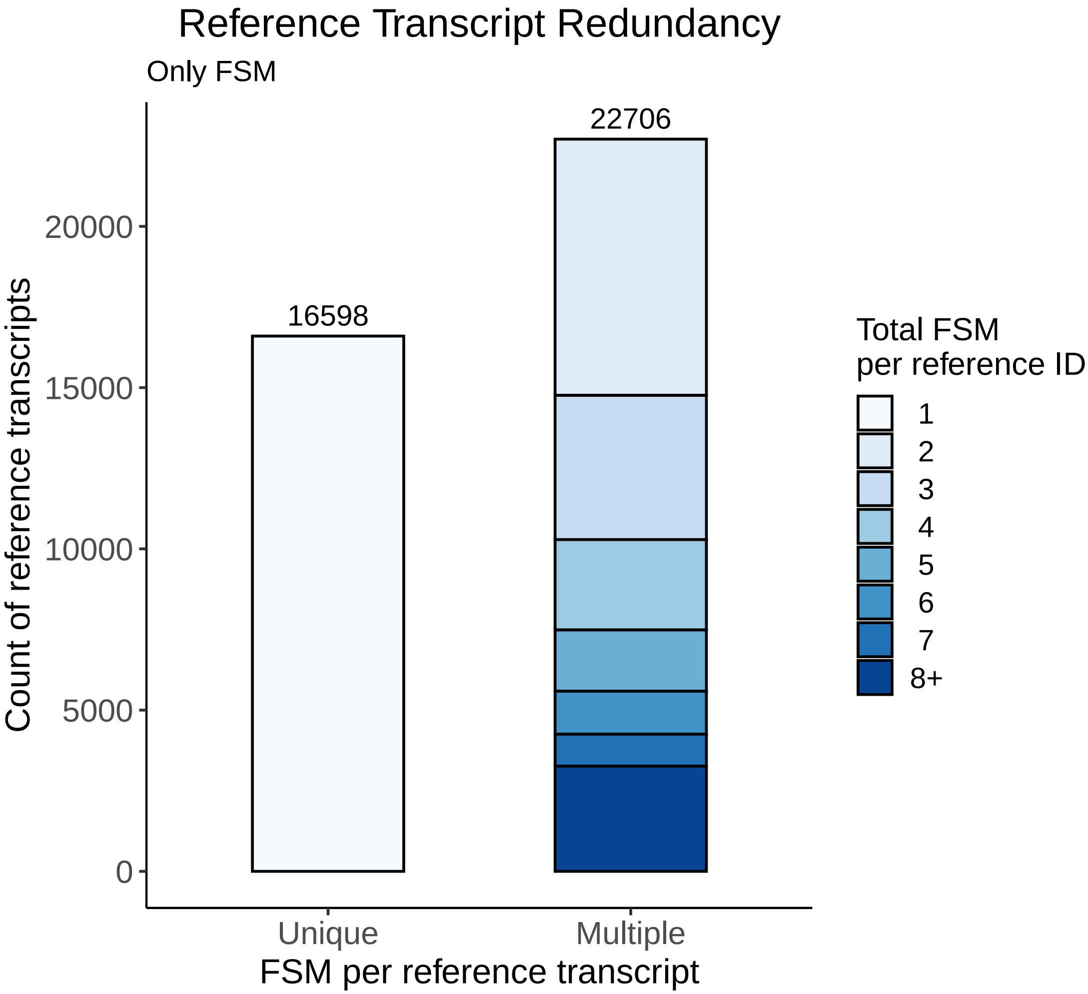
Number of polyA Motifs Detected

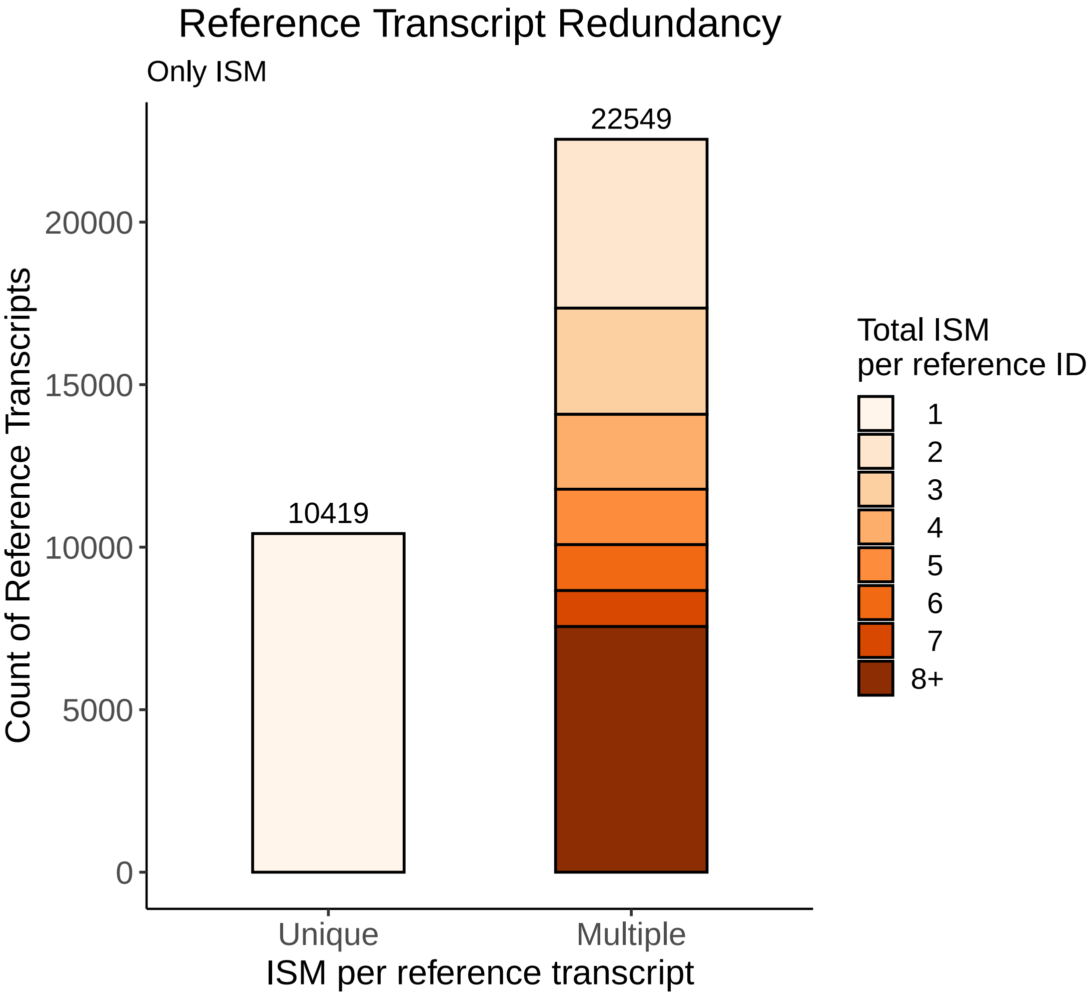
Subcategory	Count	polyA Detected	%
Alternative 3'end	40643	19327	48
Alternative 3'5'end	38026	19945	52
Alterantive 5'end	25448	13698	54
Reference match	17762	9226	52
3' fragment	158123	97730	62
Internal fragment	14717	9310	63
5' fragment	22071	10201	46
Comb. of annot. junctions	48453	24678	51
Comb. of annot. splice sites	34351	17974	52
Intron retention	45613	23823	52
At least 1 annot. don./accept.	81580	43722	54
Mono−exon	17809	10862	61
Multi−exon	7369	4408	60
NA	63	41	65

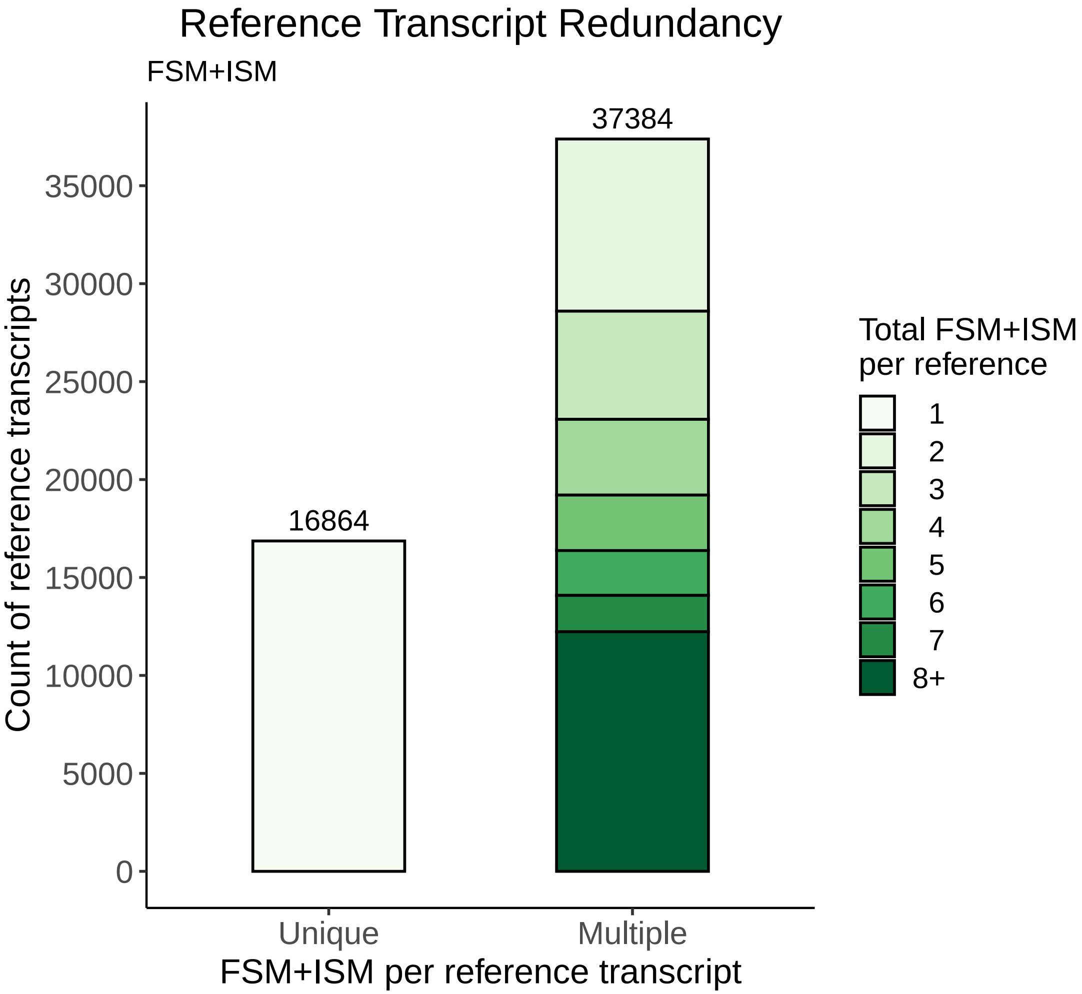
Frequency of PolyA Motifs

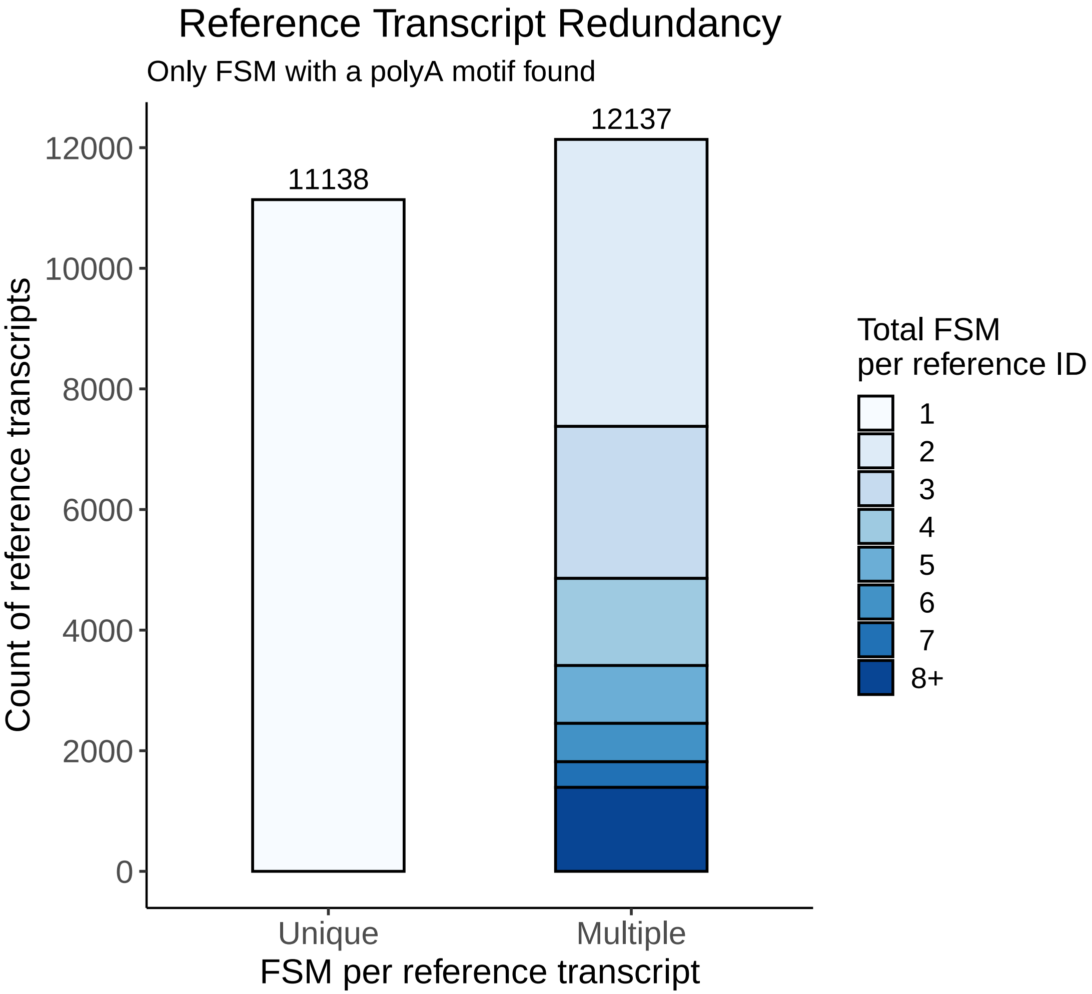
Motif	Count	%
AATAAA	152200	49.9
ATTAAA	44265	14.5
TATAAA	12075	4.0
AGTAAA	11289	3.7
AAAAAG	11141	3.7
AAGAAA	10987	3.6
GGGGCT	9225	3.0
TTTAAA	8022	2.6
CATAAA	6849	2.2
AATATA	6738	2.2
GATAAA	6524	2.1
AATACA	6476	2.1
AAAACA	6253	2.1
AATGAA	5903	1.9
AATAGA	3649	1.2
ACTAAA	3349	1.1

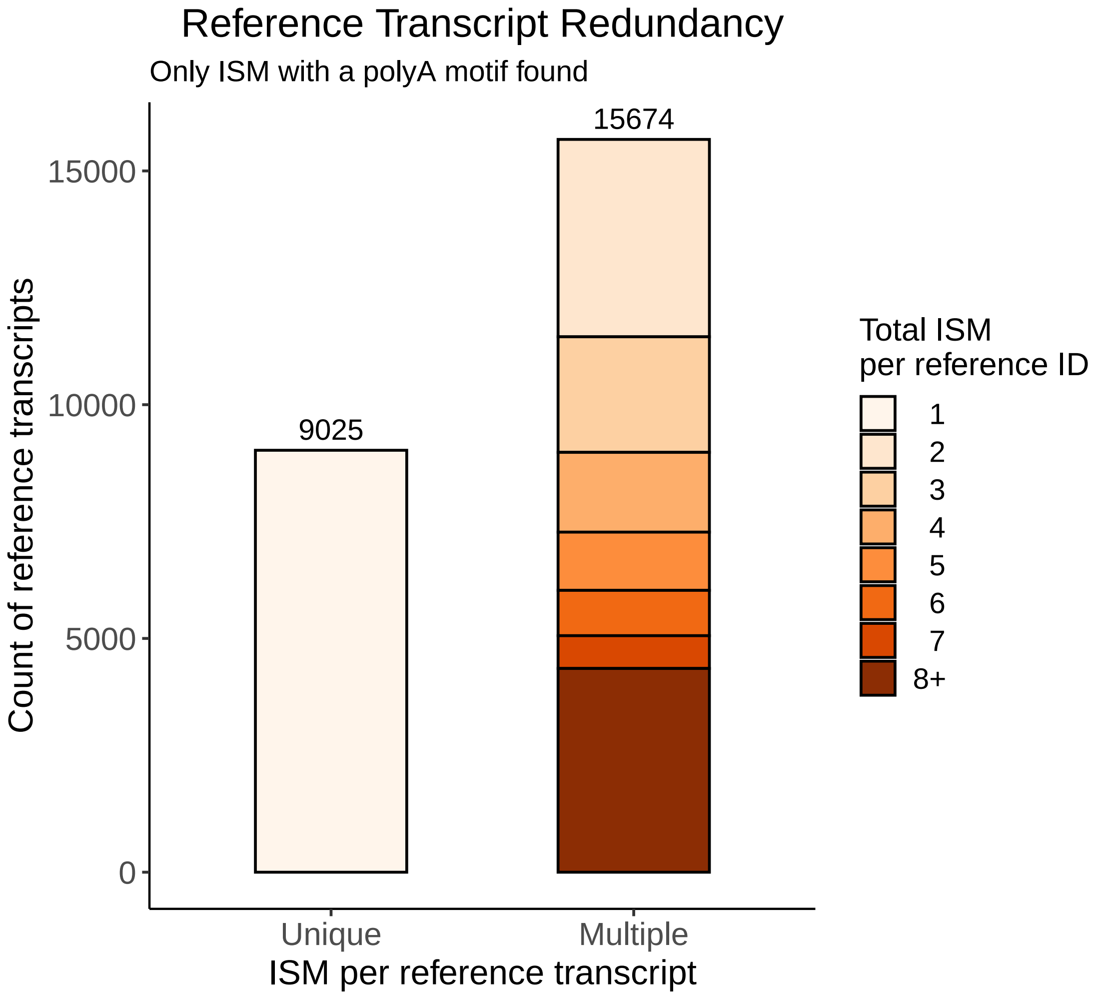
Redundancy Analysis

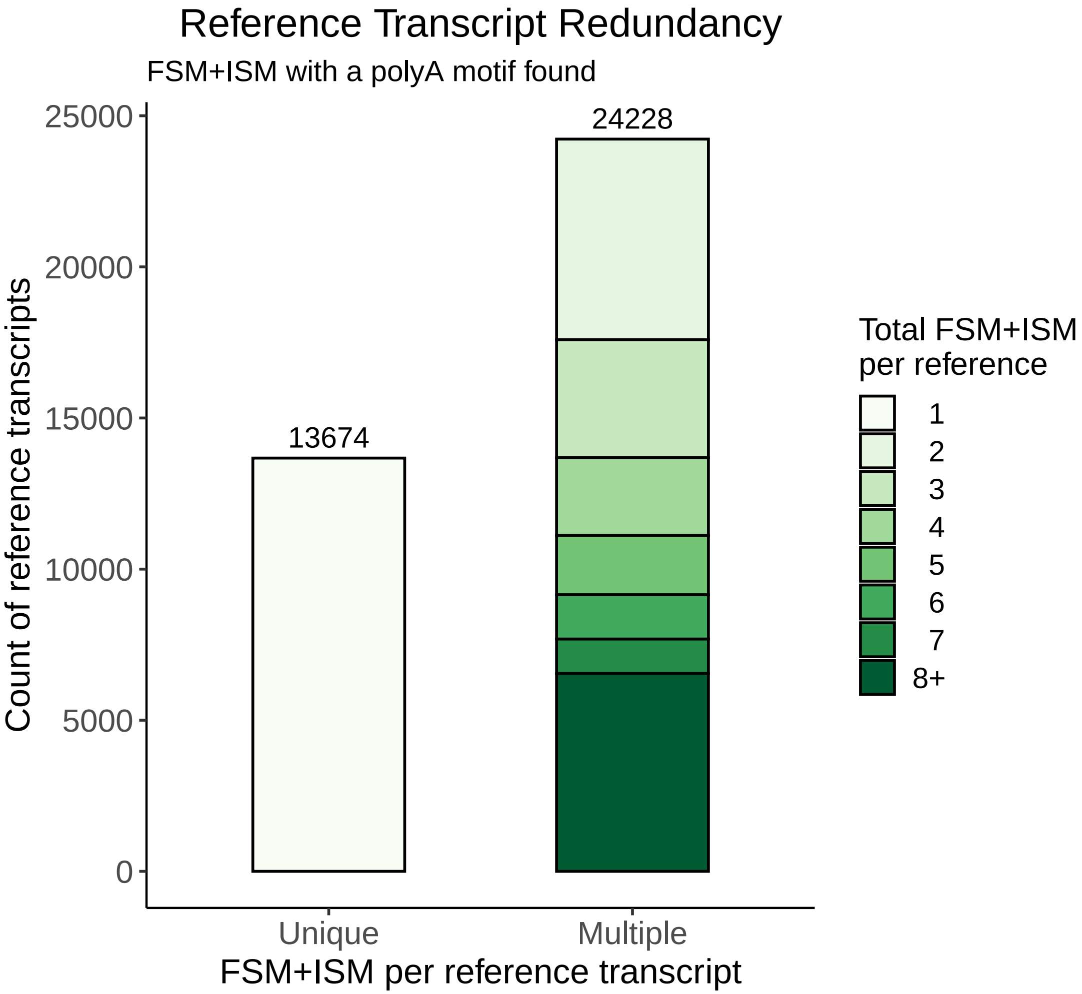

Reference Transcript Redundancy
Only FSM
Count of reference transcripts
0
5000
10000
15000
20000
16598
22706
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only ISM
Count of Reference Transcripts
0
5000
10000
15000
20000
10419
22549
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

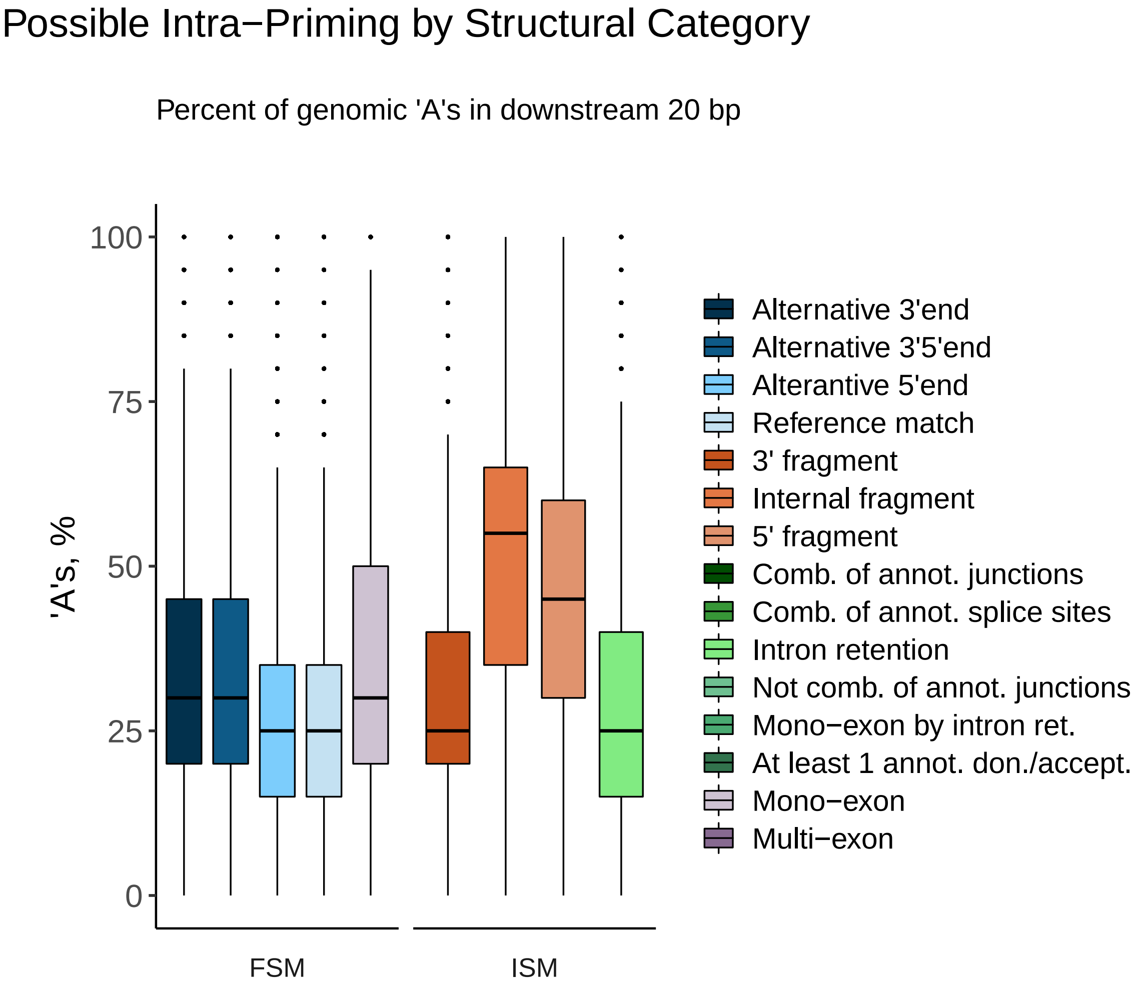

Reference Transcript Redundancy
FSM+ISM
37384
16864
35000
30000
25000
20000
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only FSM with a polyA motif found
Count of reference transcripts
12000
10000
8000
6000
4000
2000
0
11138
12137
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
15674
9025
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

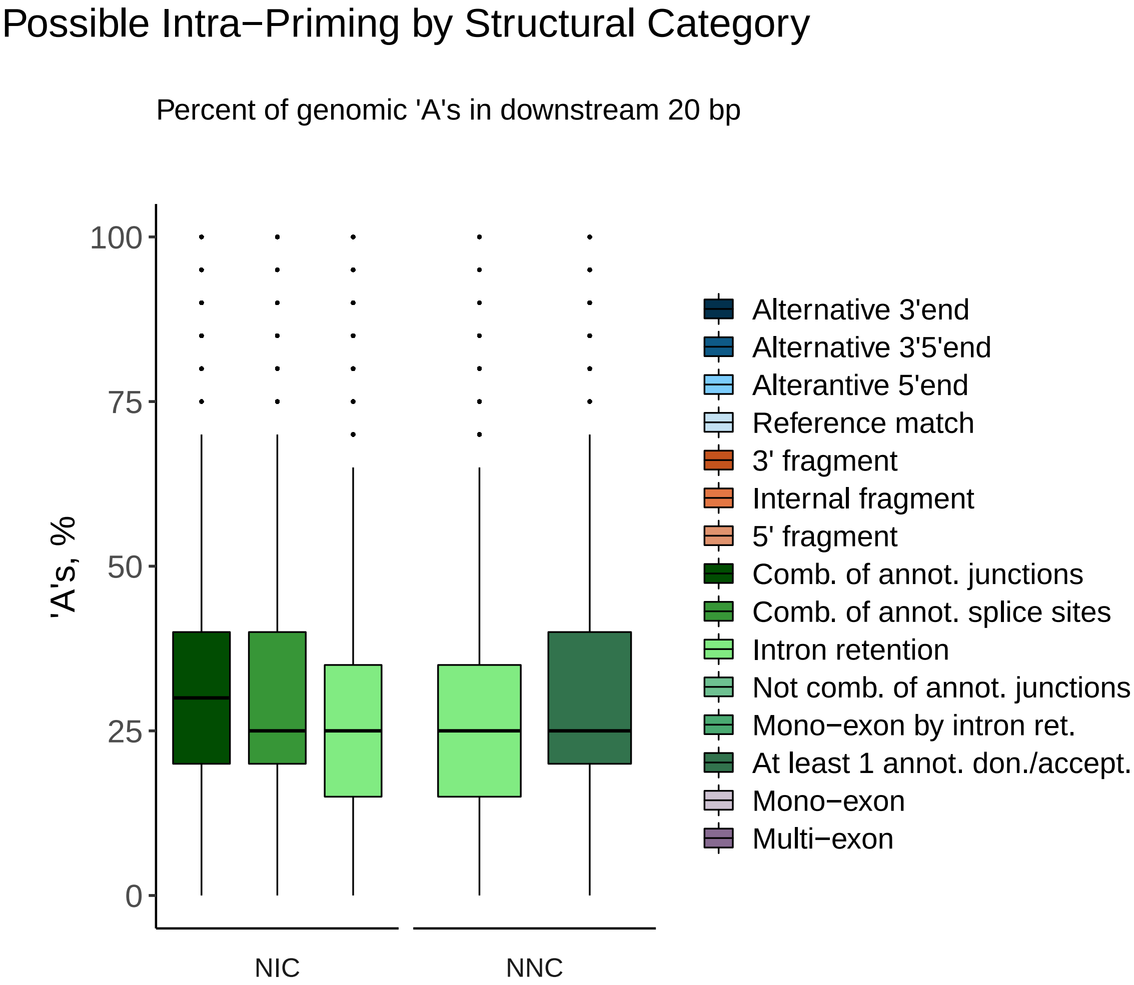

Reference Transcript Redundancy
FSM+ISM with a polyA motif found
Count of reference transcripts
25000
20000
15000
10000
5000
0
13674
24228
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+

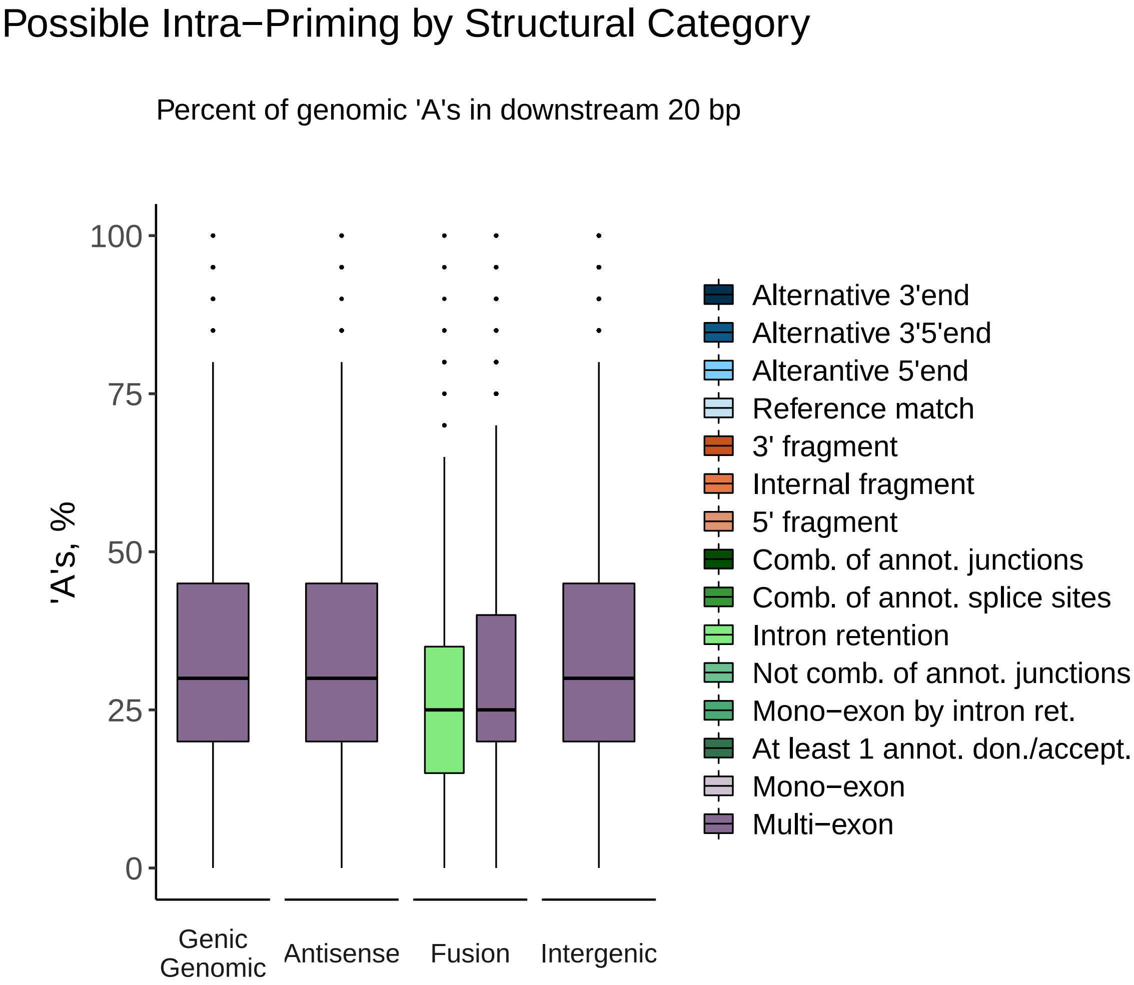
Intra−Priming Quality Check

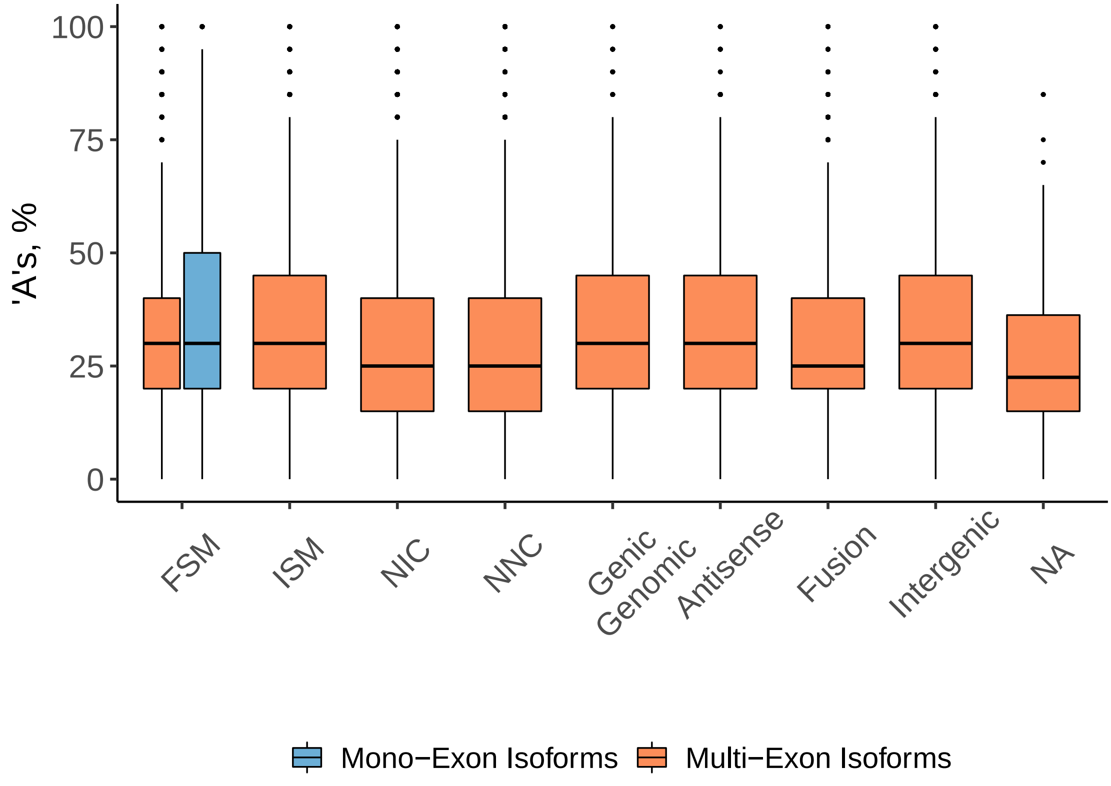


Possible Intra−Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp

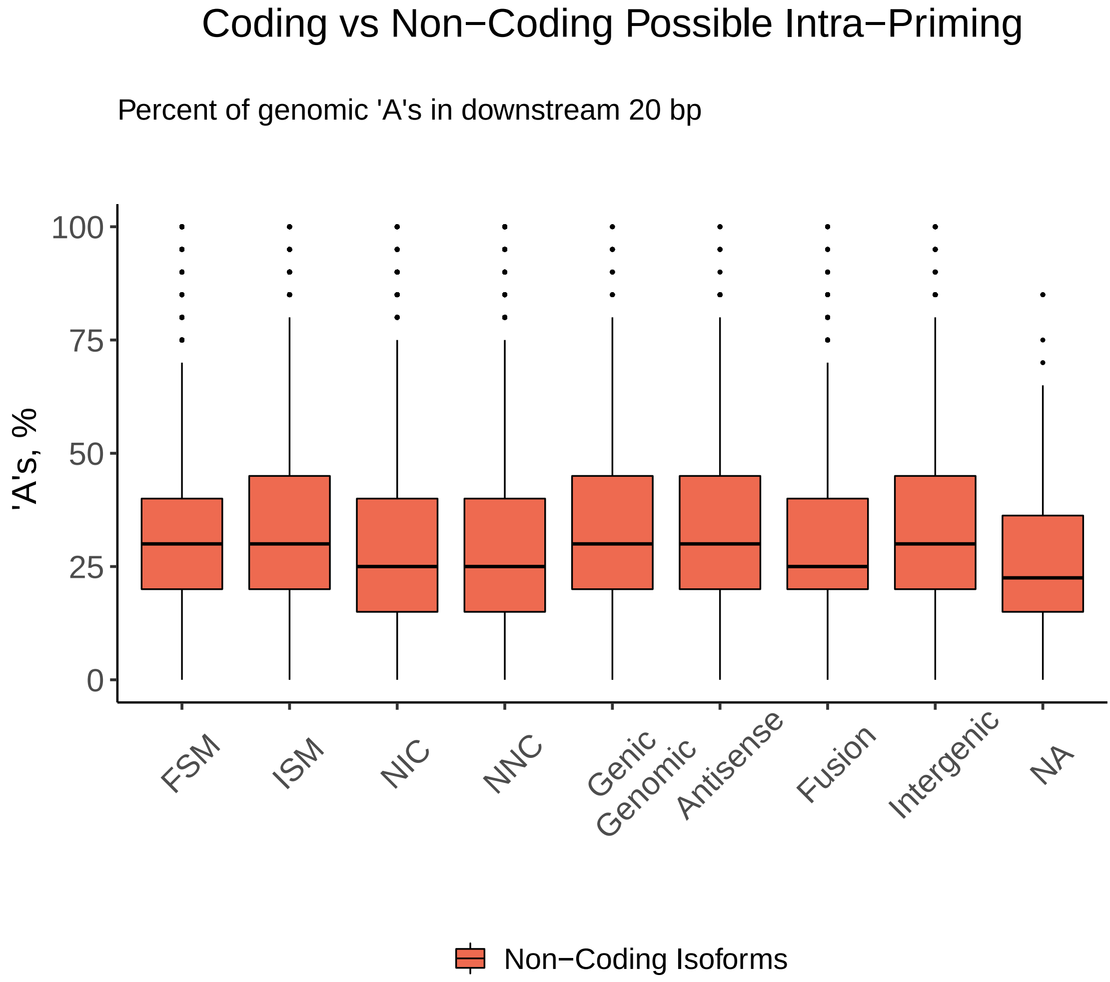

Possible Intra−Priming by Structural Category
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention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


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

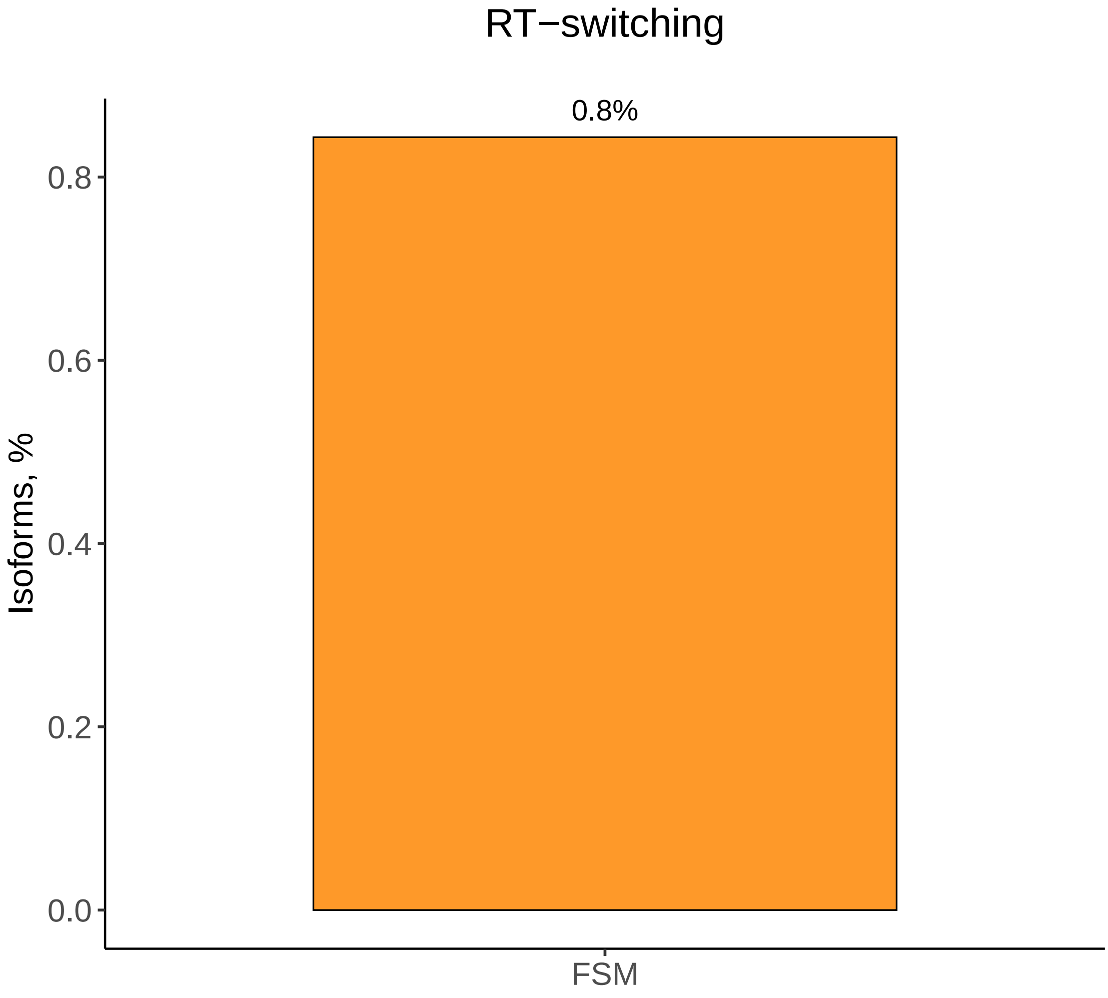


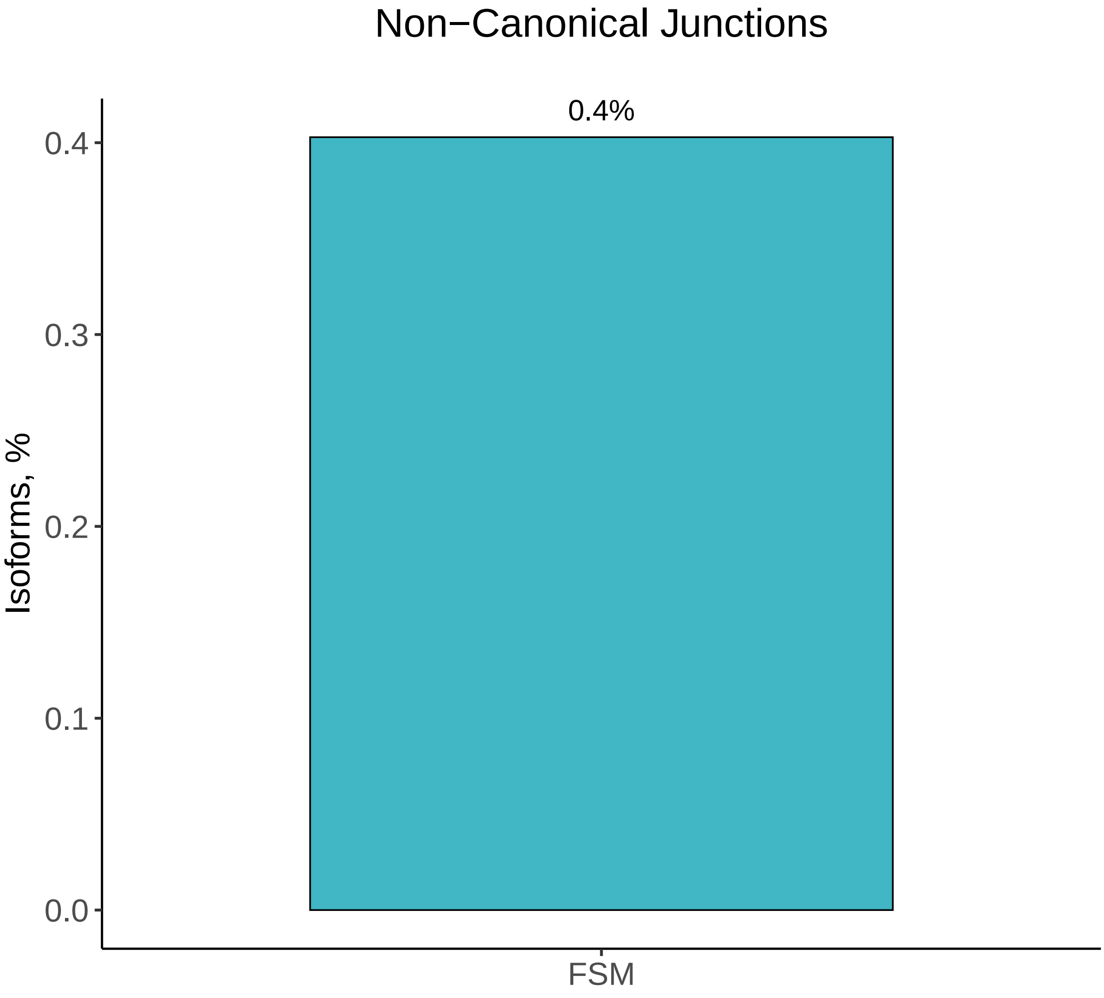
Mono− vs Multi−Exon Possible Intra−Priming

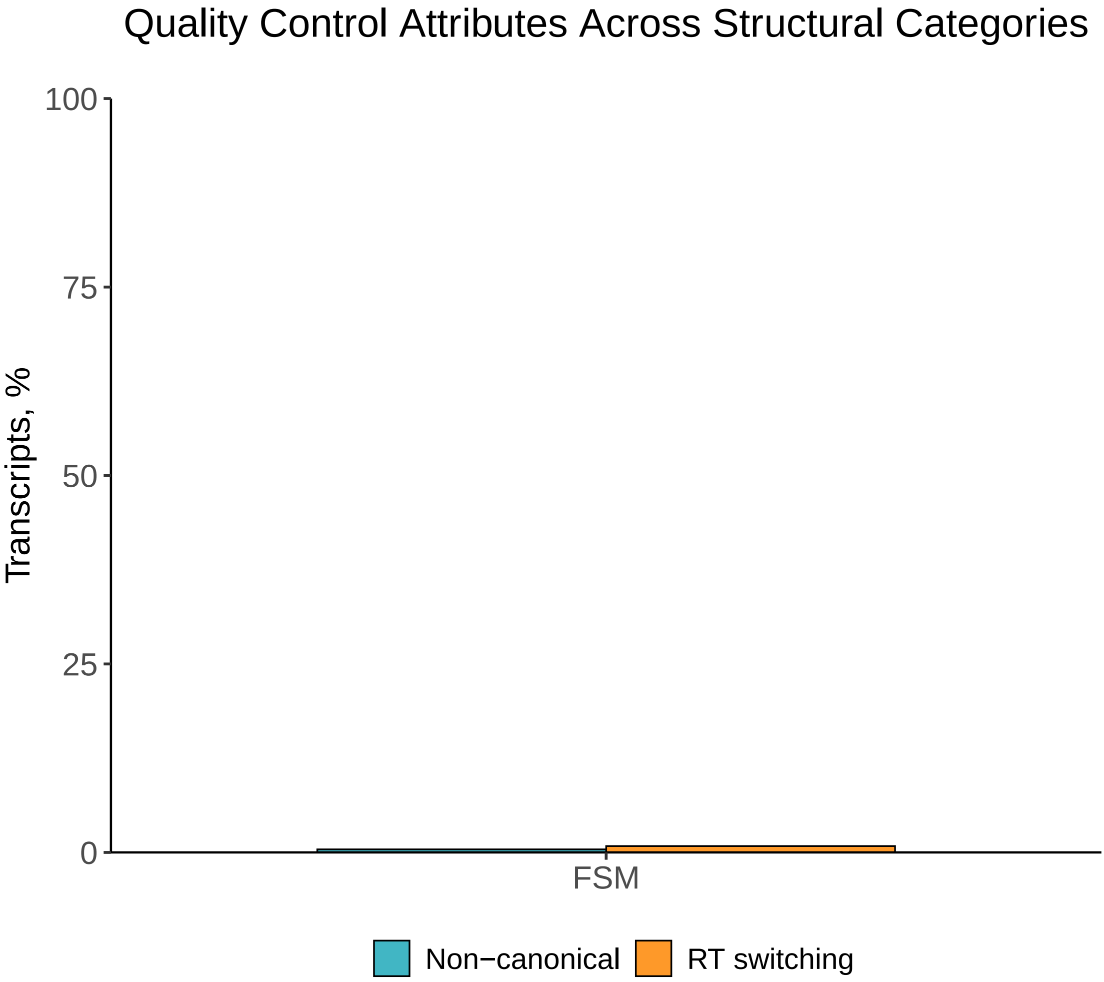
Percent of genomic 'A's in downstream 20 bp


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

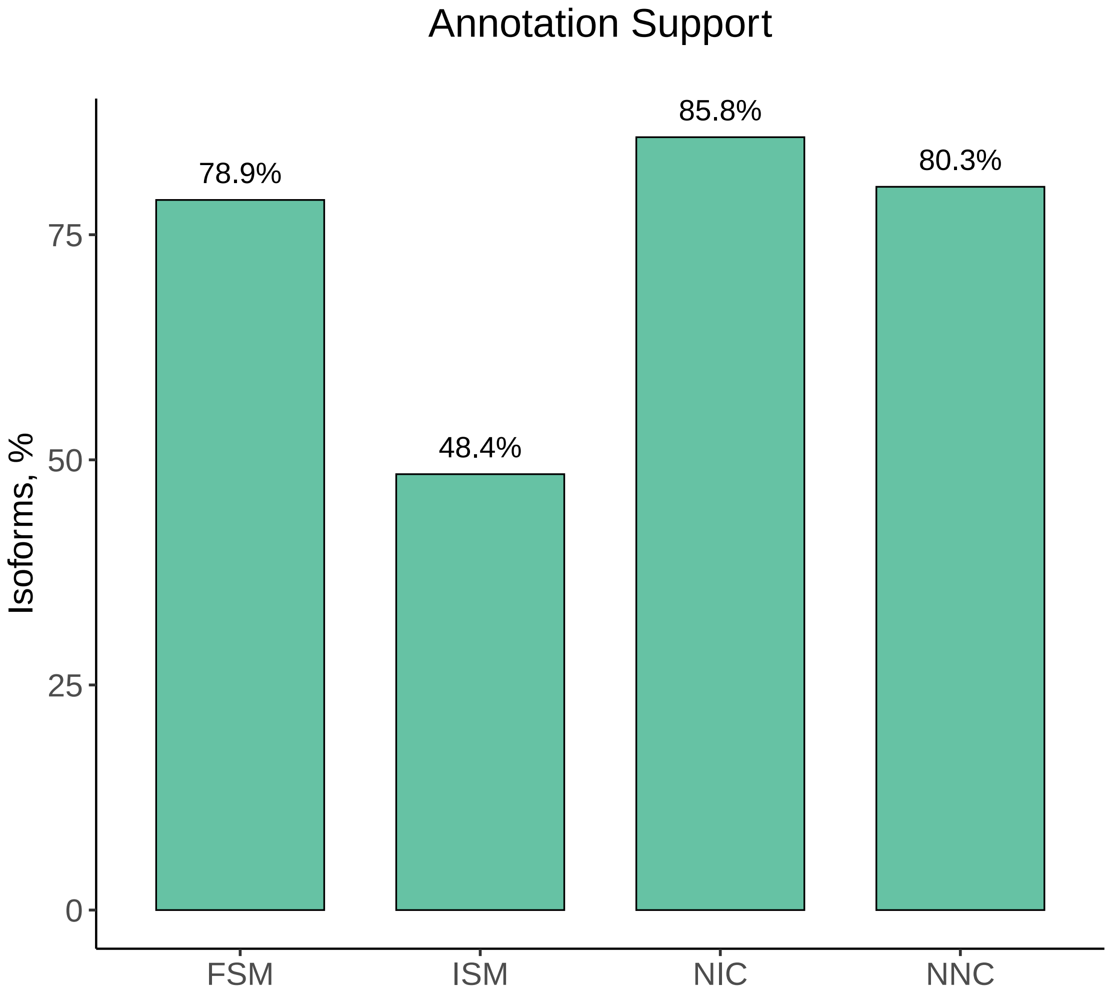
Features of Bad Quality

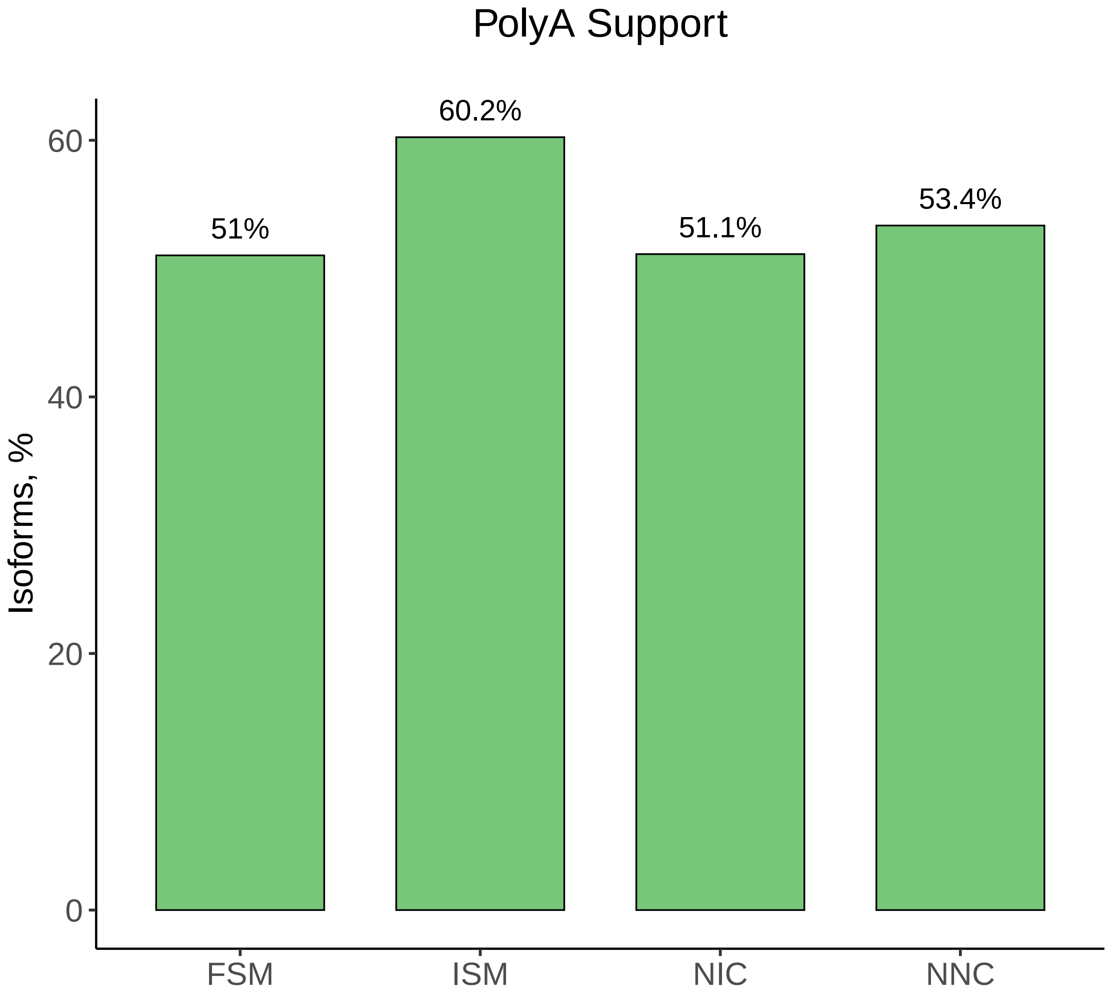

RT−switching
0.8%
0.8
0.6
0.4
0.2
0.0
Isoforms, %
FSM

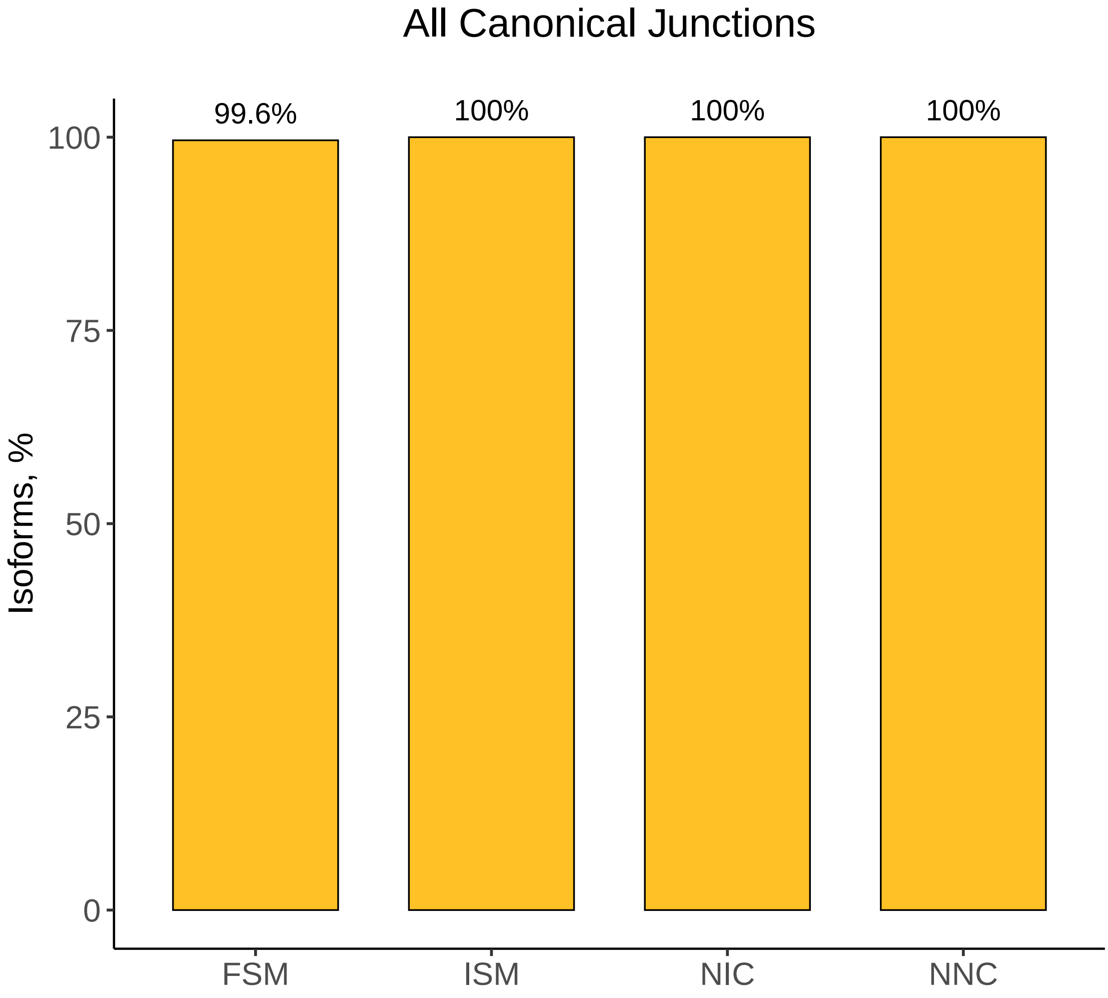

Non−Canonical Junctions
0.4%
Isoforms, %
0.4
0.3
0.2
0.1
0.0
FSM

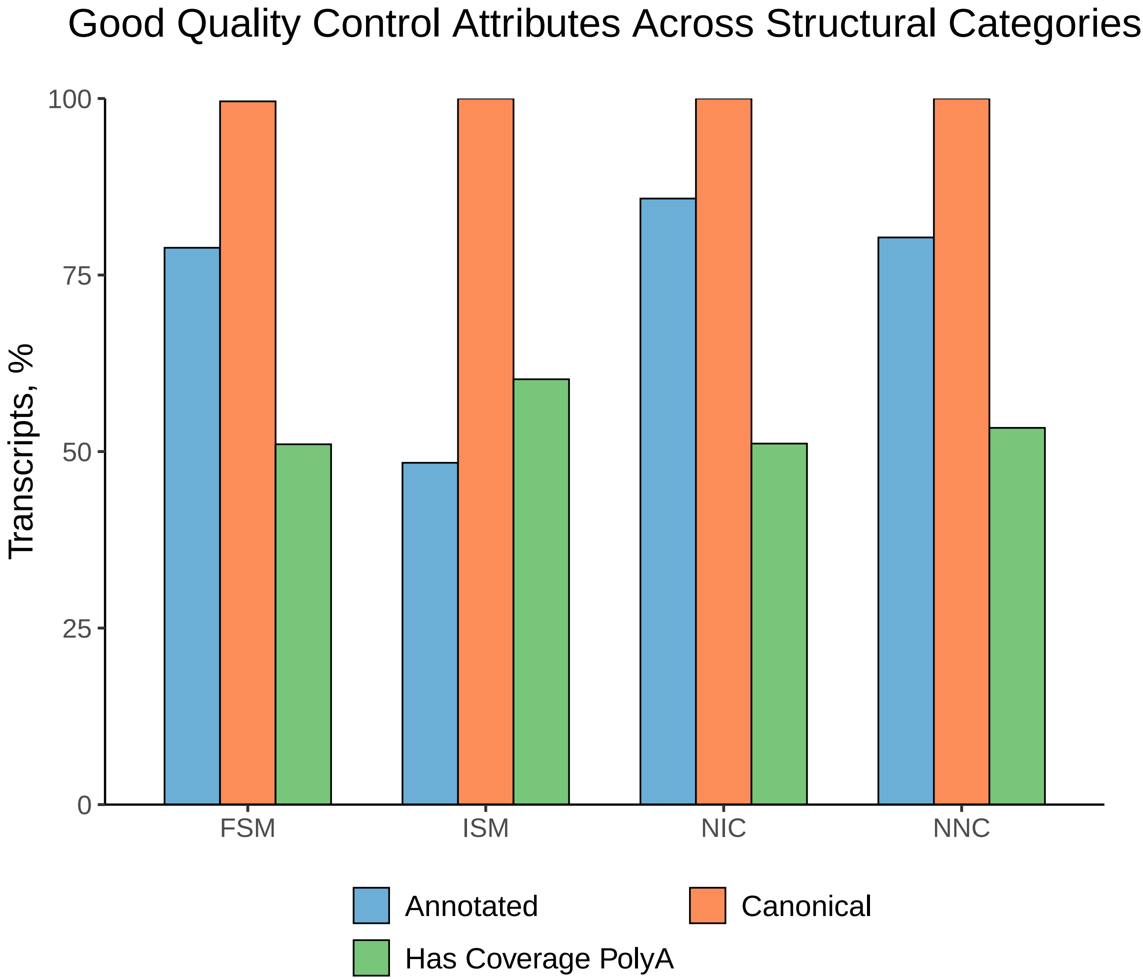

Quality Control Attributes Across Structural Categories
100
75
50
25
0
Transcripts, %
FSM
Non−canonical
RT switching

Features of Good Quality


Annotation Support
Isoforms, %
0
25
50
75
78.9%
48.4%
85.8%
80.3%
FSM
ISM
NIC
NNC


PolyA Support
Isoforms, %
0
20
40
60
51%
60.2%
51.1%
53.4%
FSM
ISM
NIC
NNC


All Canonical Junctions
Isoforms, %
0
25
50
75
100
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