SQANTI3 report

Unique Genes: 17169

Unique Isoforms: 77017

Gene Classification

Category	Genes, count
Annotated Genes	14017
Novel Genes	3152

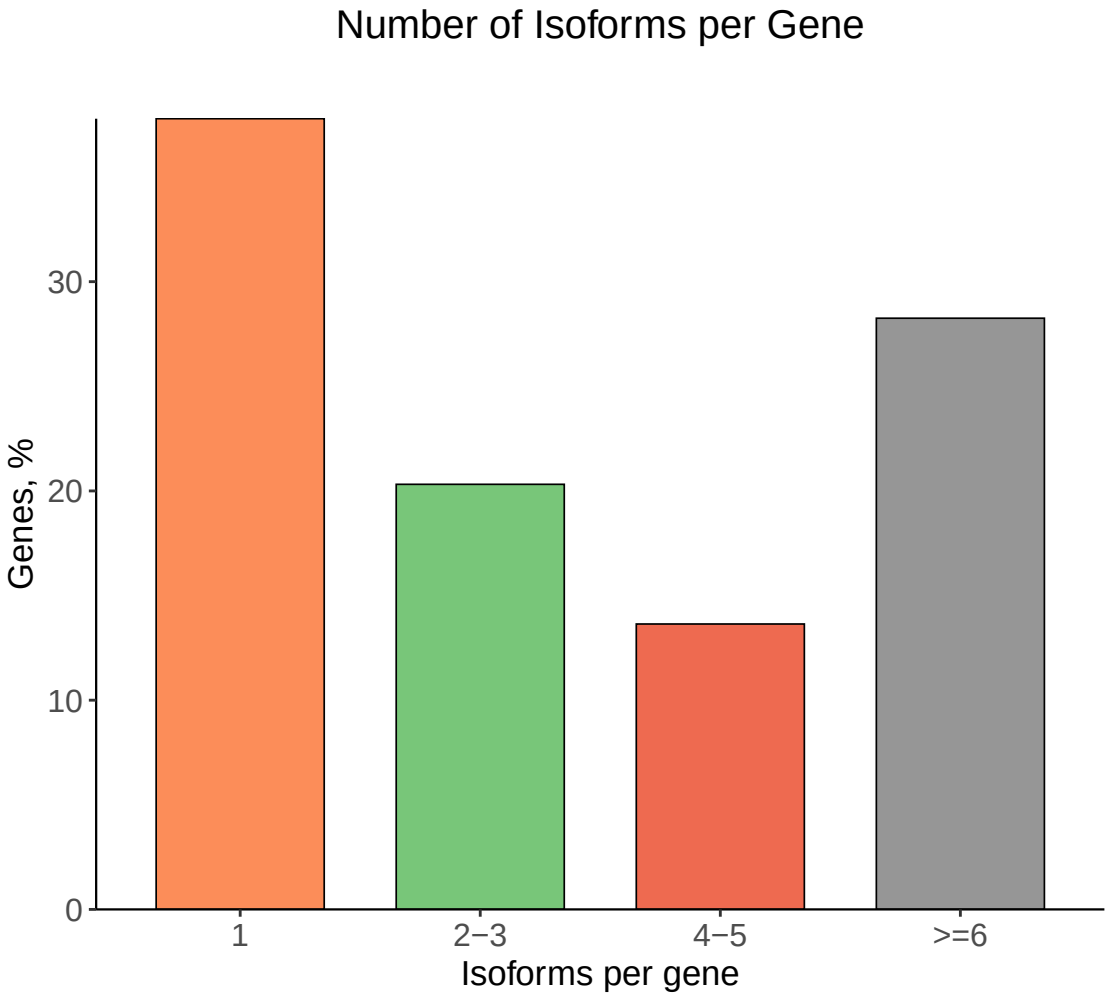
Transcript Classification

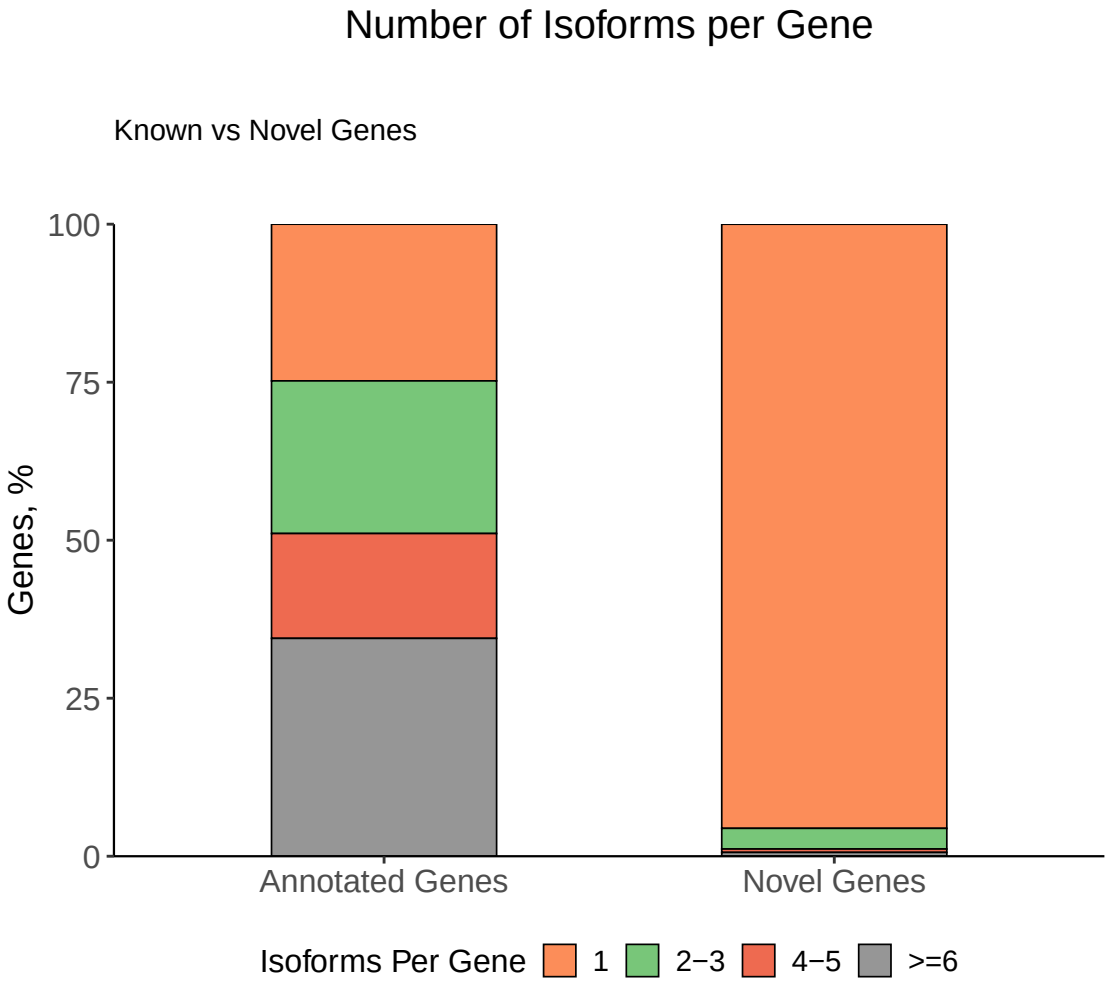
Category	Isoforms, count
FSM	25235
ISM	3781
NIC	22995
NNC	19878
Genic Genomic	1041
Antisense	856
Fusion	549
Intergenic	2131
Genic Intron	551

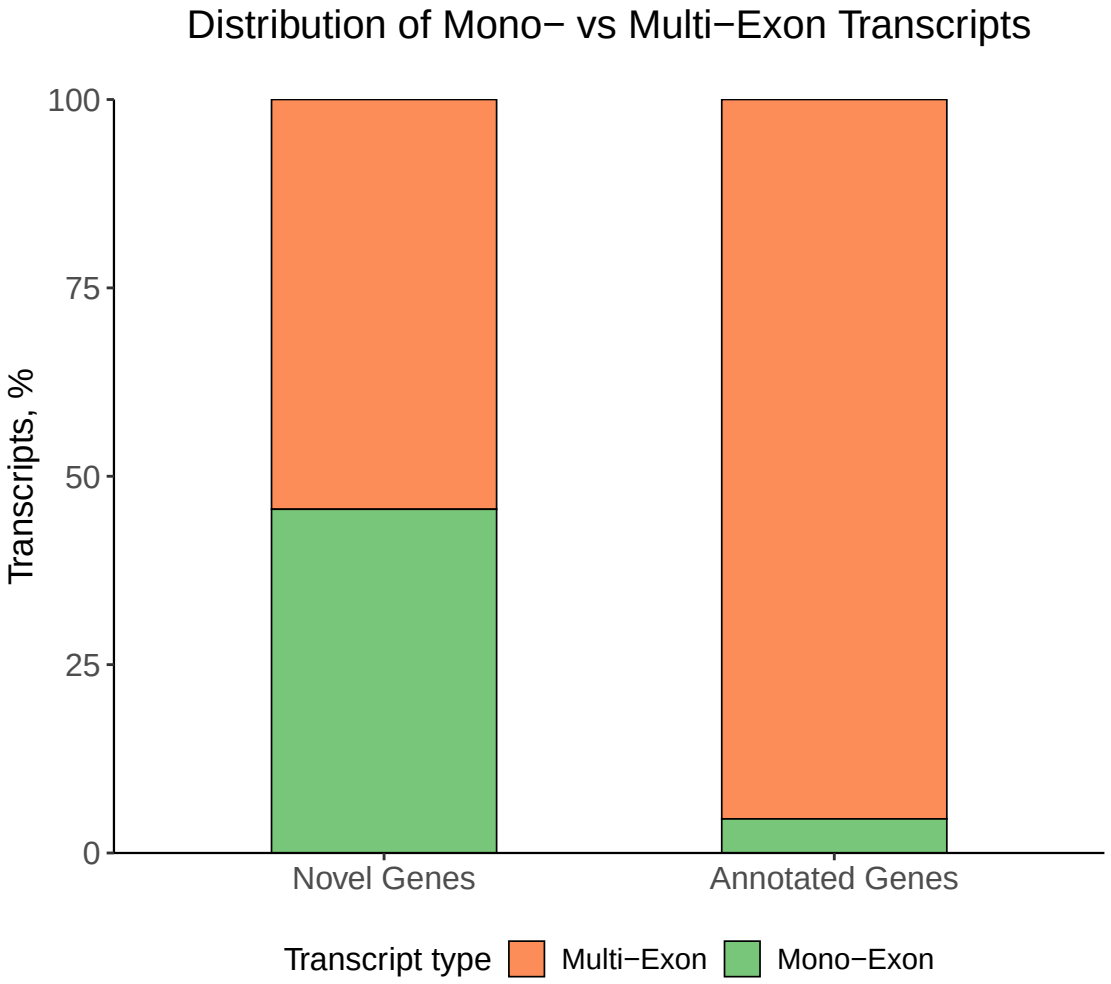
Splice Junction Classification

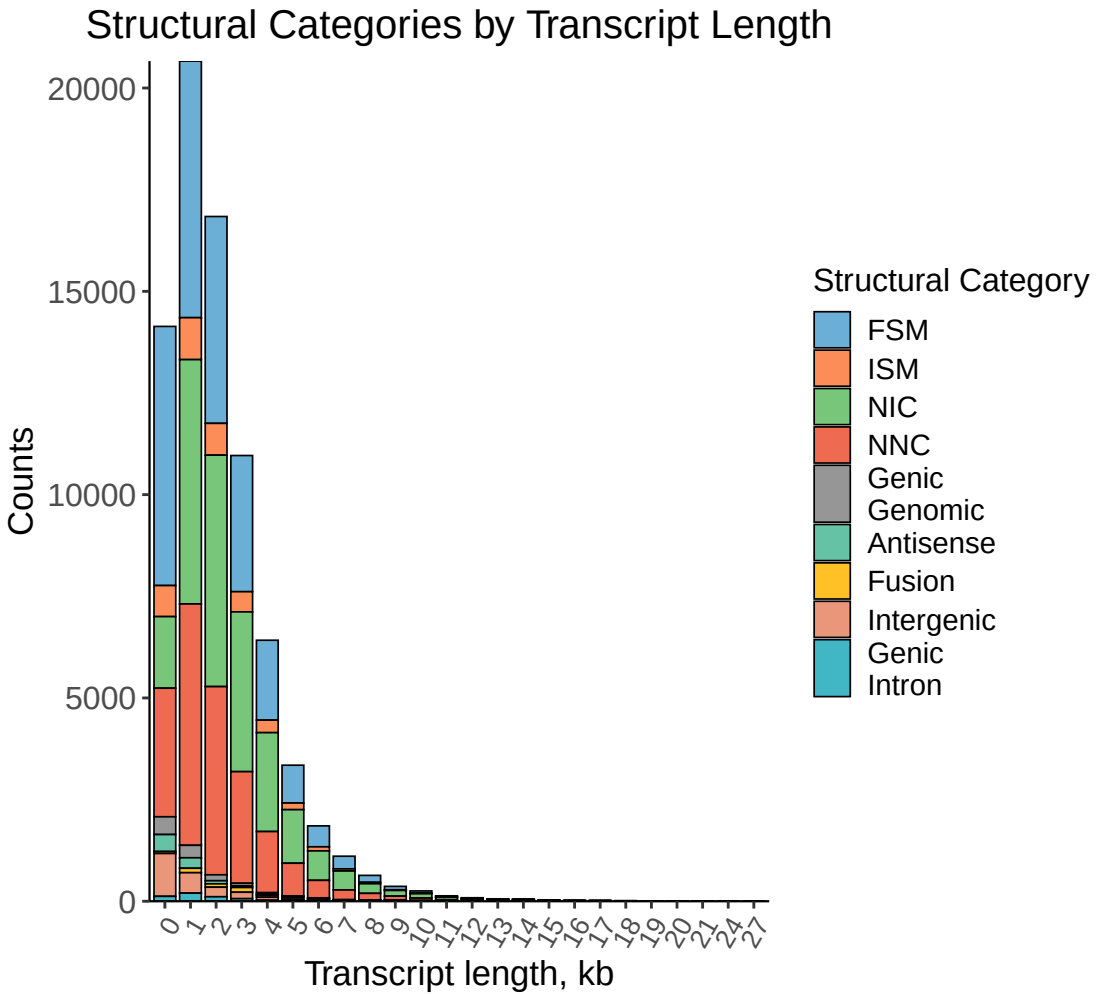
Category	SJs, count	Percent
Known canonical	130781	81.31
Known Non−canonical	164	0.10
Novel canonical	25897	16.10
Novel Non−canonical	3994	2.48

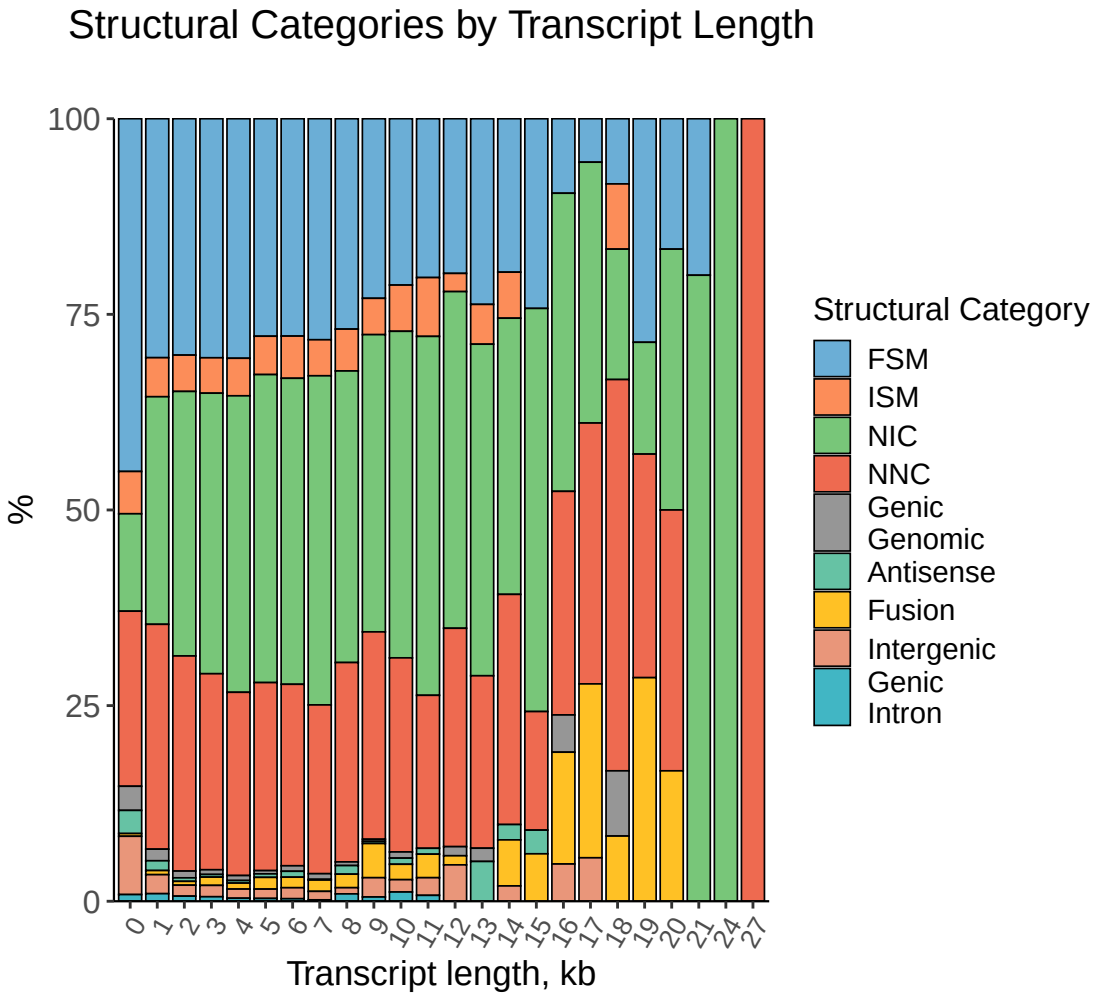
Gene Characterization

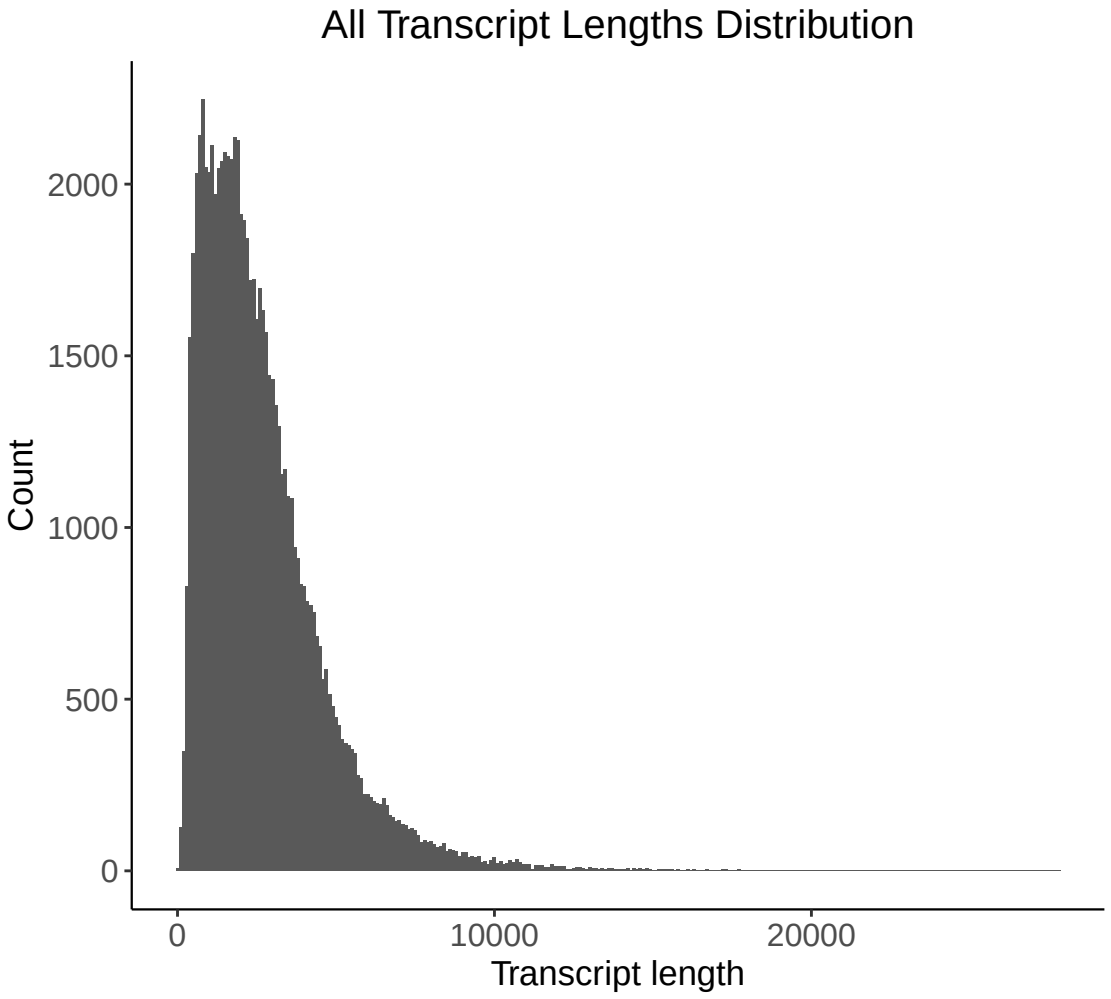

Number of Isoforms per Gene
Genes, %
30
20
10
0
1
2−3
4−5
>=6
Isoforms per gene

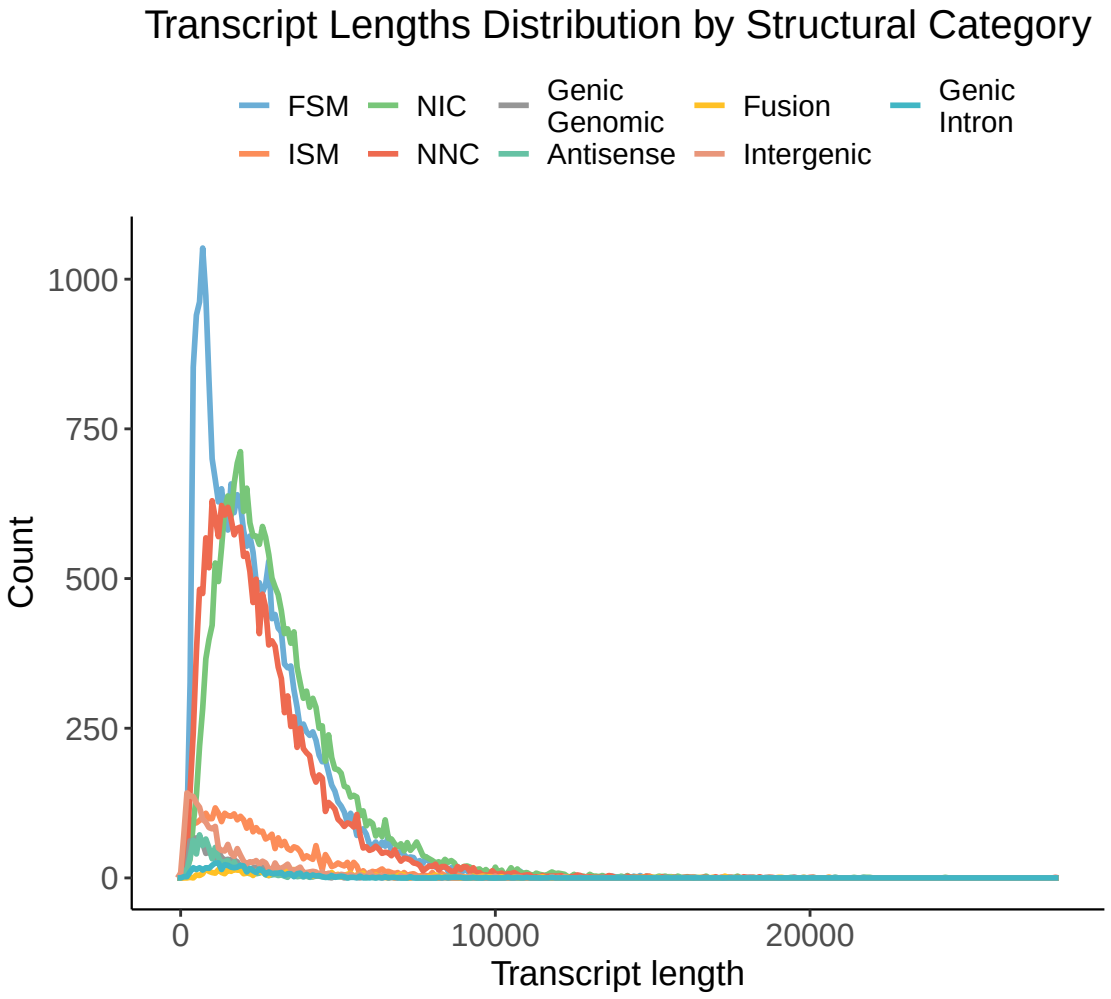

Number of Isoforms per Gene
Known vs Novel Genes
Genes, %
100
75
50
25
0
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3
4−5
>=6

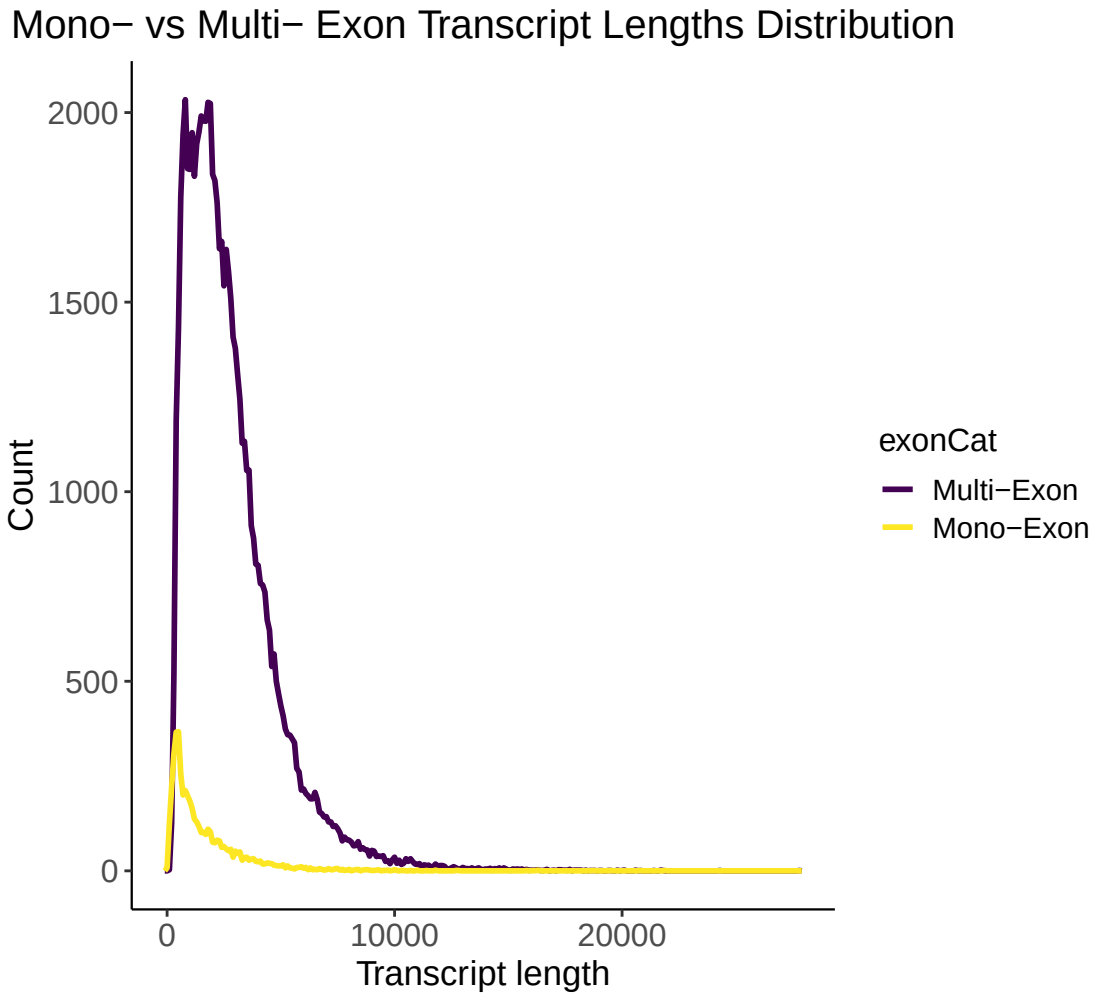

Distribution of Mono− vs Multi−Exon Transcripts
Transcripts, %
100
75
50
25
0
Novel Genes
Annotated Genes
Transcript type
Multi−Exon
Mono−Exon


Structural Categories by Transcript Length
Counts
20000
15000
10000
5000
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 24 27
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

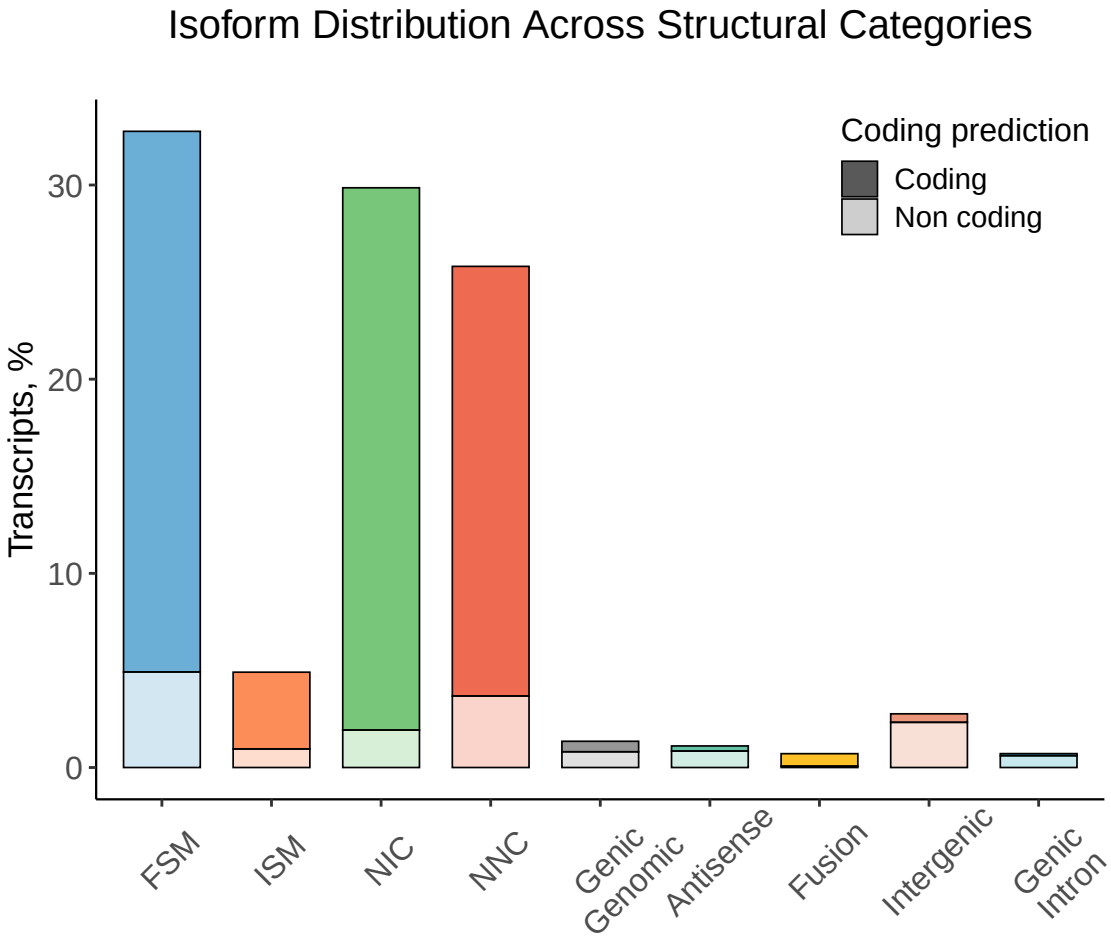

Structural Categories by Transcript Length
100
75
50
25
0
%
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
24
27
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

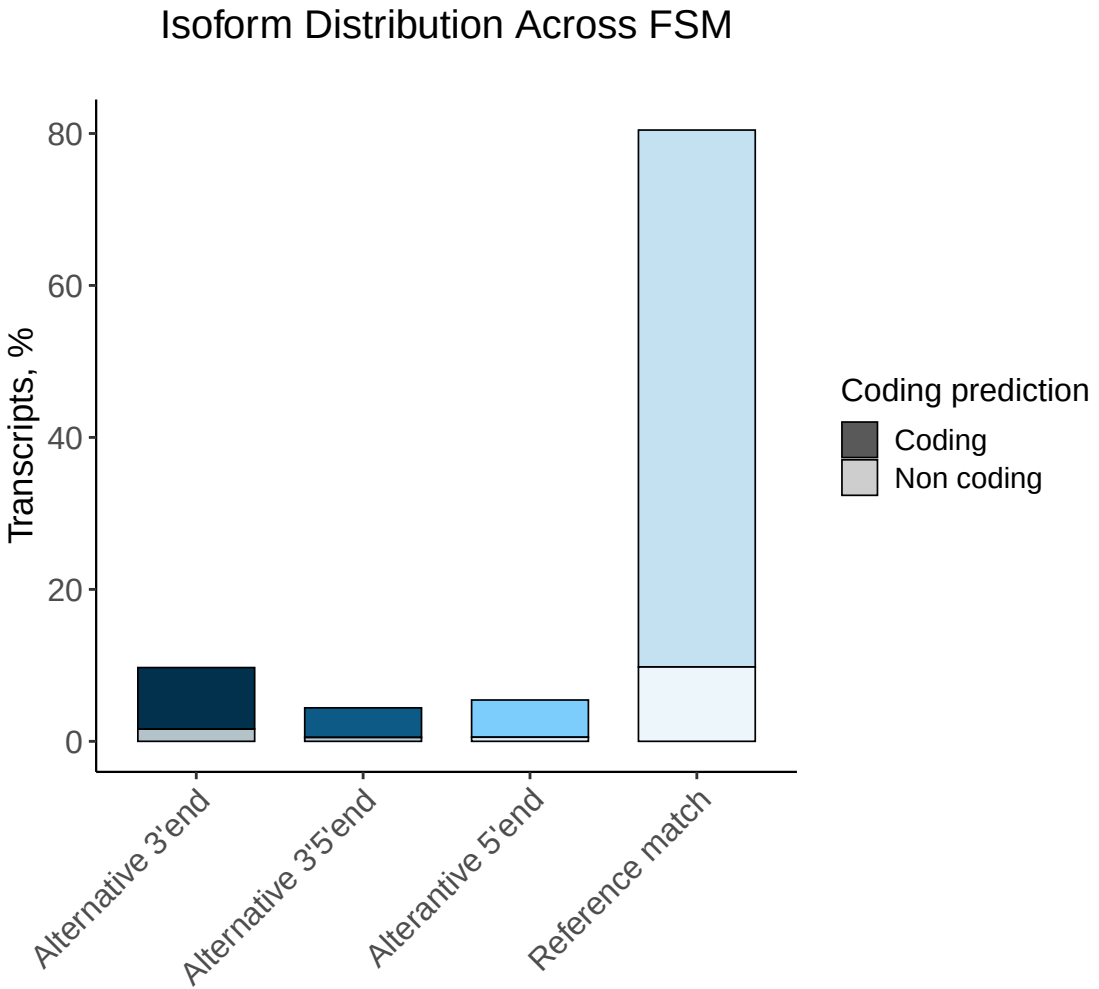

All Transcript Lengths Distribution
Count
2000
1500
1000
500
0
0
10000
20000
Transcript length

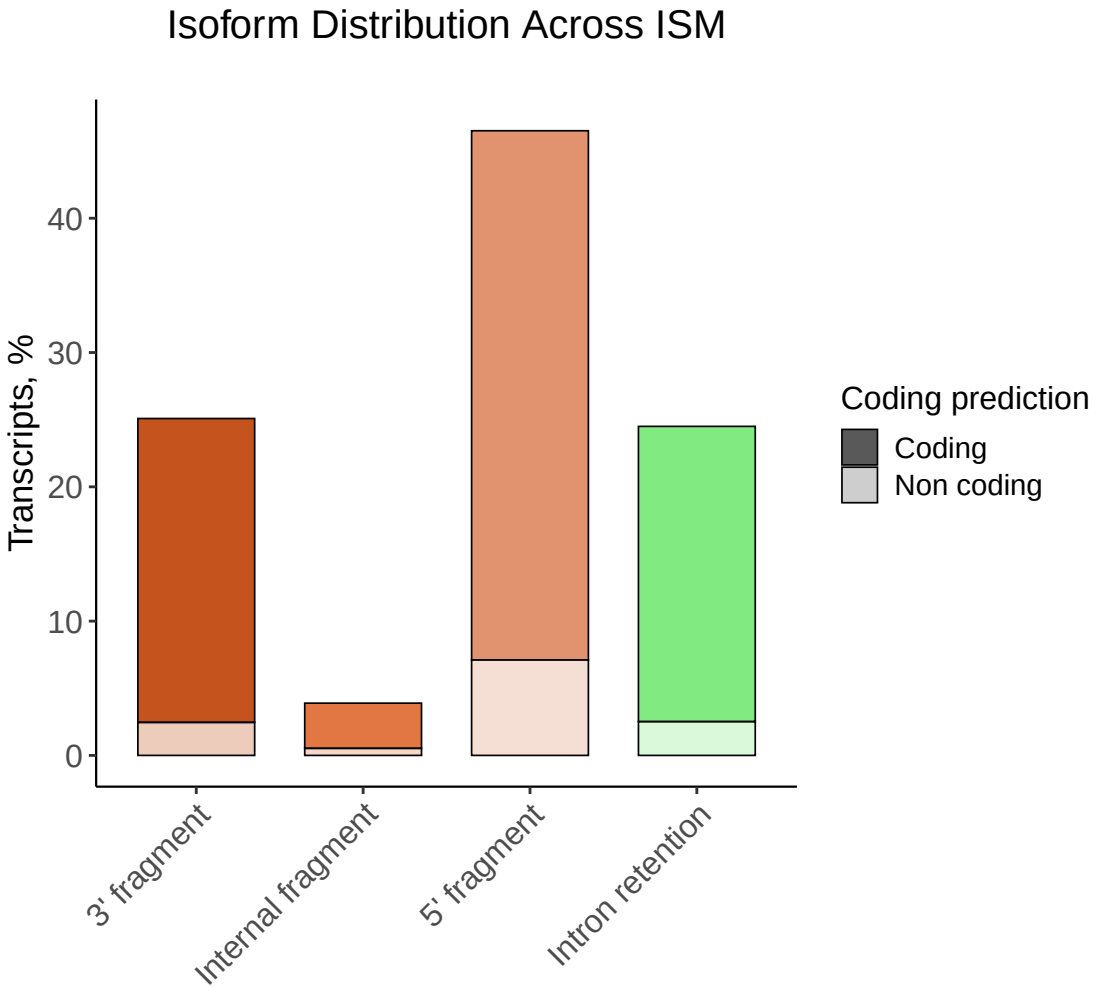

Transcript Lengths Distribution by Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Count
1000
750
500
250
0
0
10000
20000
Transcript length

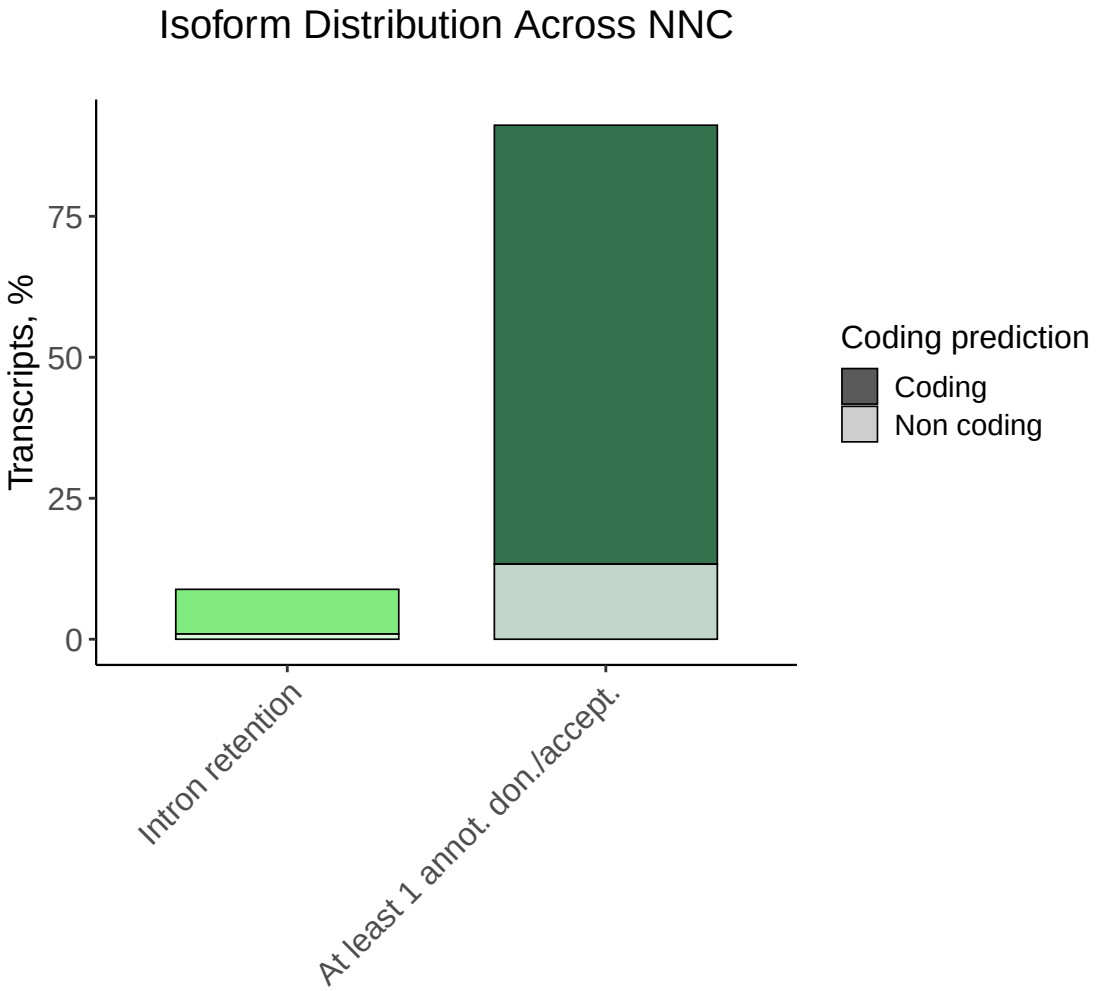

Mono− vs Multi− Exon Transcript Lengths Distribution
Count
2000
1500
1000
500
0
0
10000
20000
Transcript length
exonCat
Multi−Exon
Mono−Exon

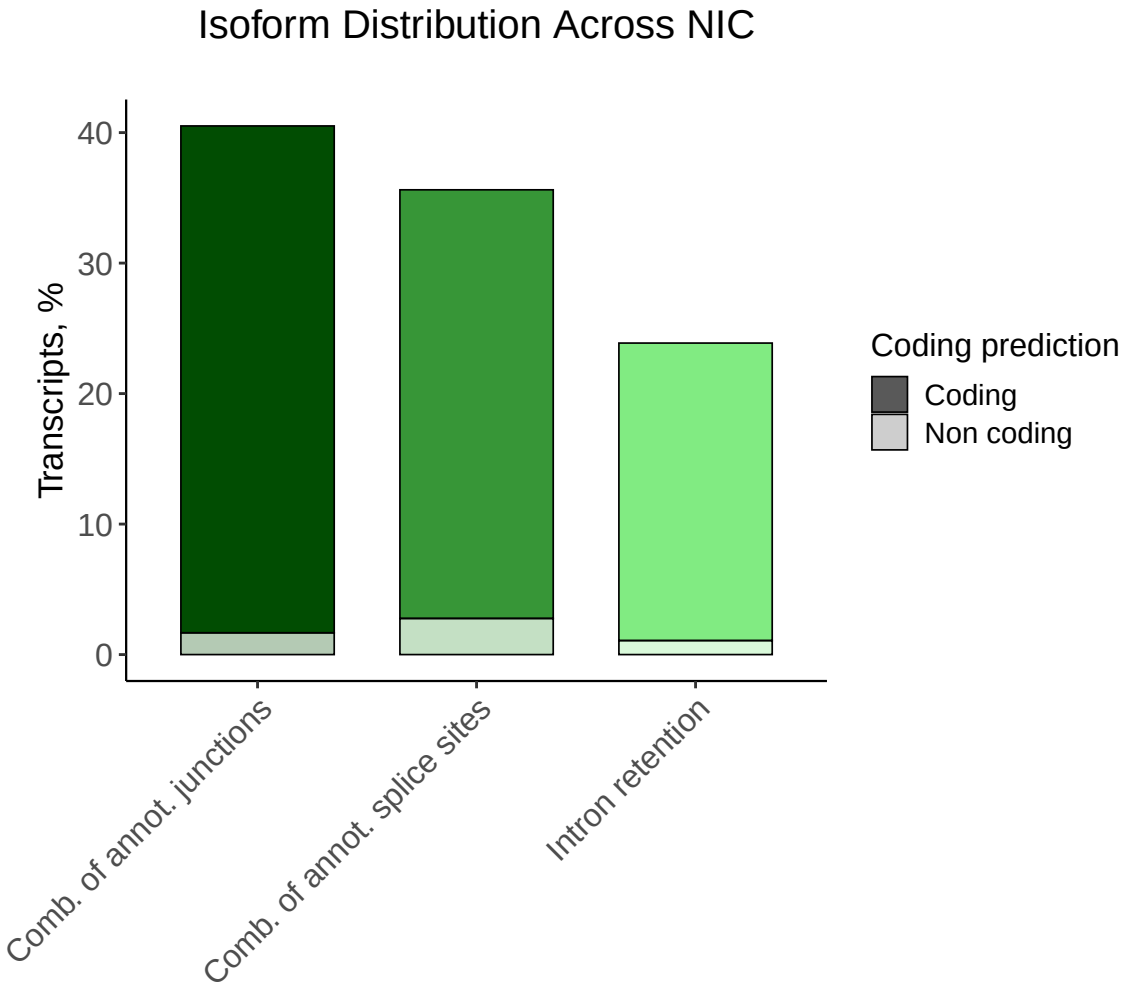
Structural Isoform Characterization

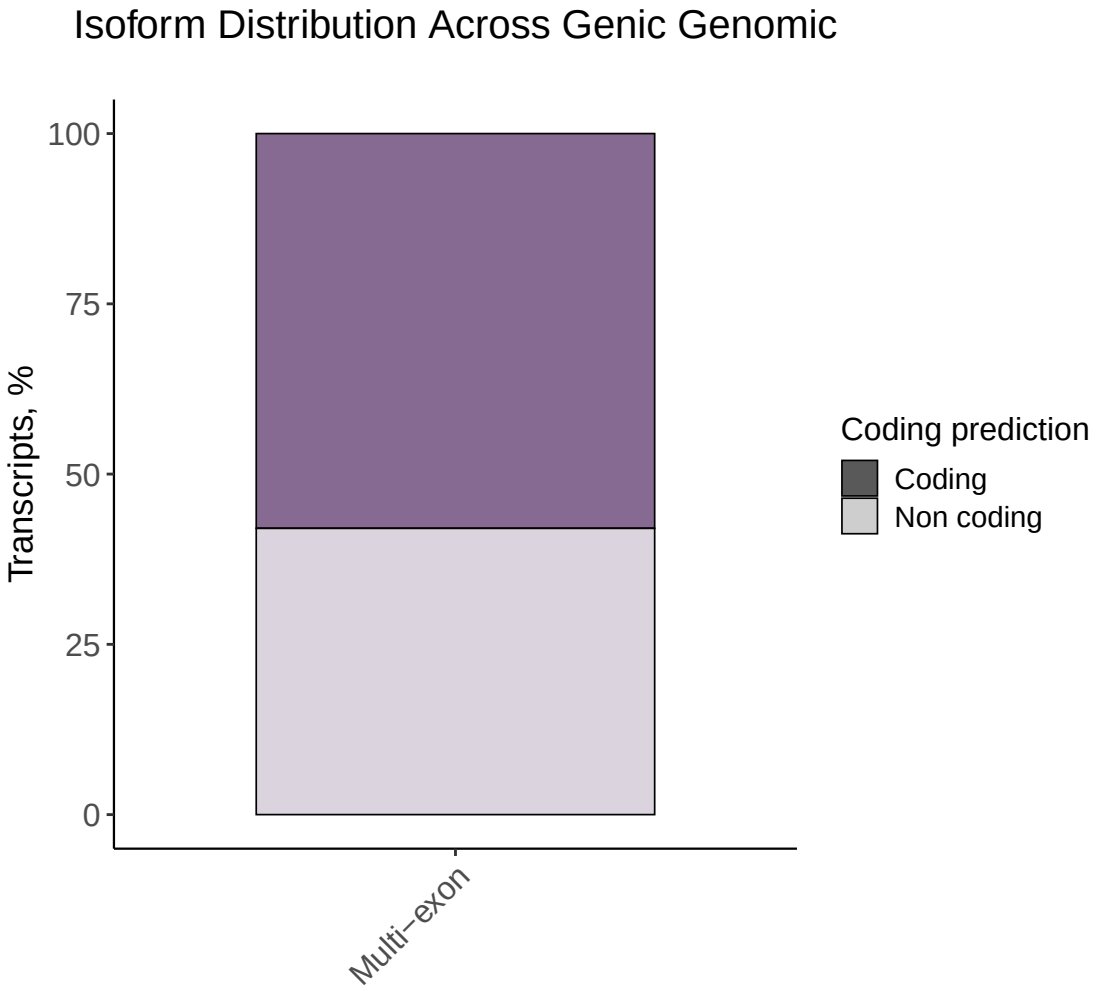

Isoform Distribution Across Structural Categories
Coding prediction
Coding
Non coding
Transcripts, %
30
20
10
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

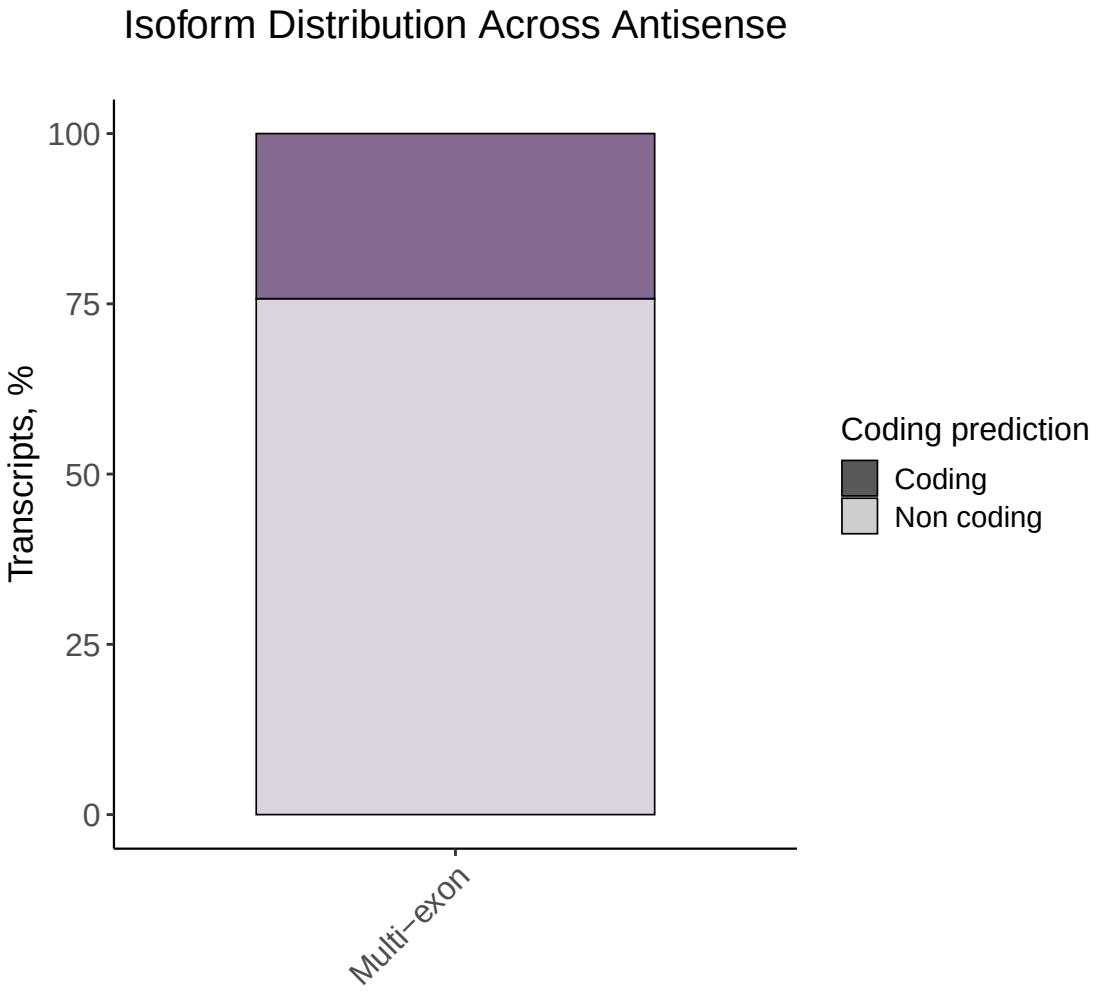

Isoform Distribution Across FSM
Transcripts, %
80
60
40
20
0
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Coding prediction
Coding
Non coding

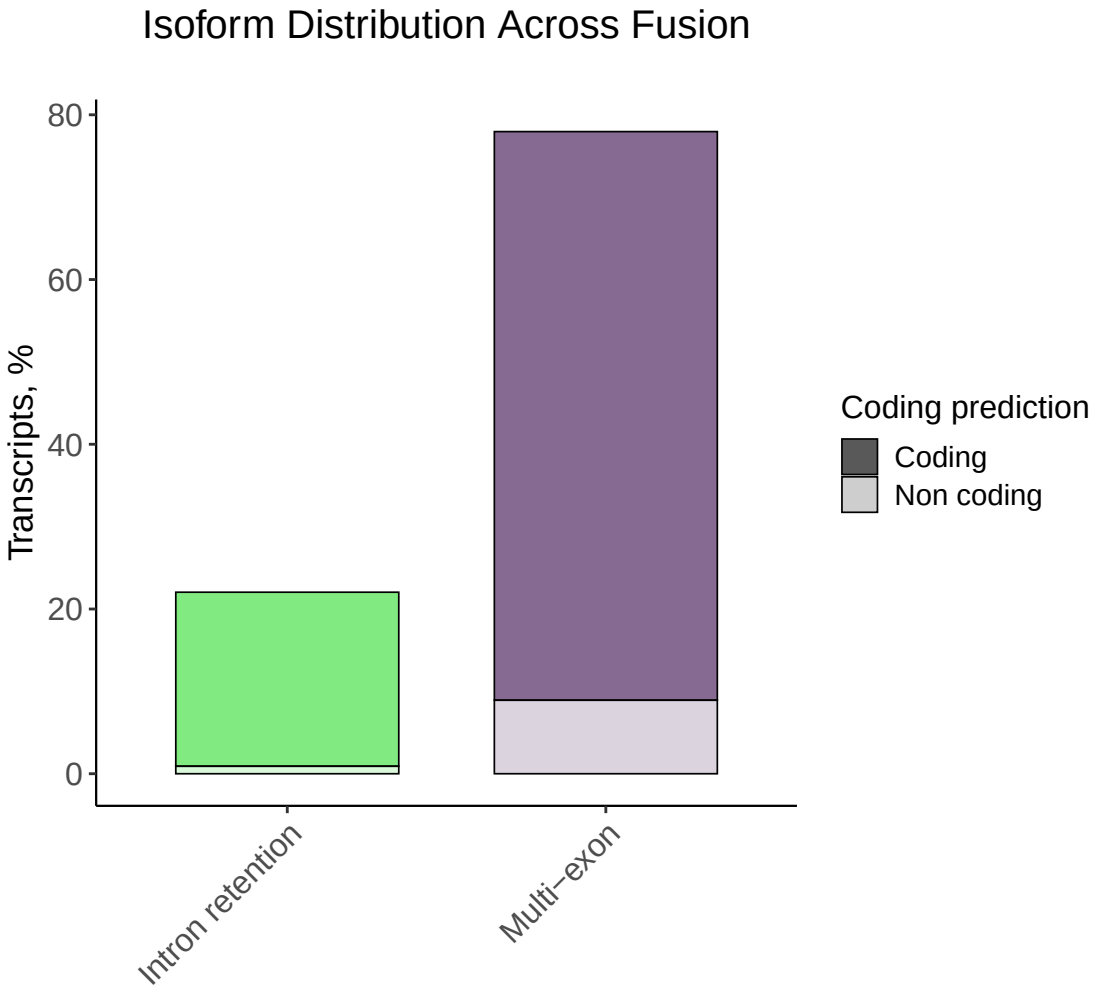

Isoform Distribution Across ISM
Transcripts, %
0
10
20
30
40
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding
Non coding

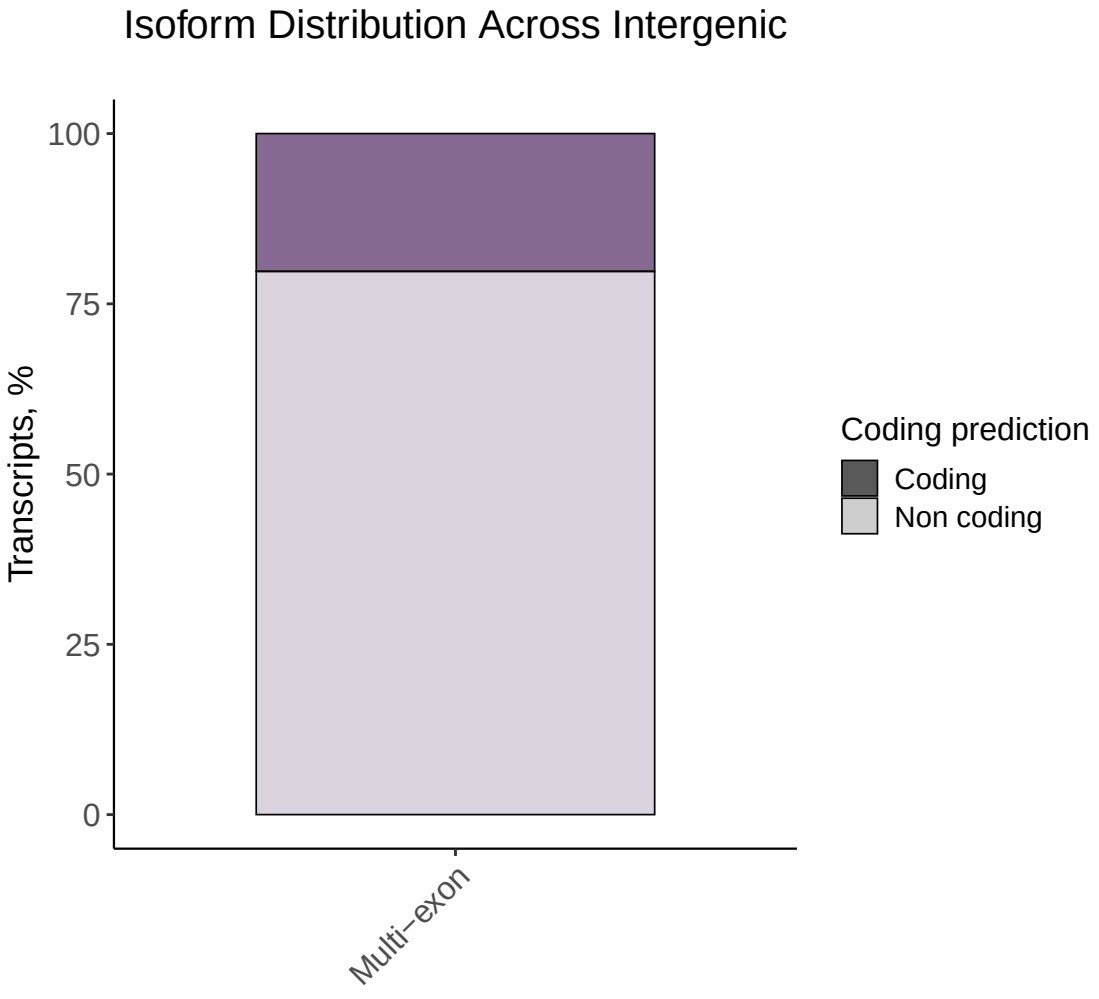

Isoform Distribution Across NNC
Transcripts, %
75
50
25
0
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding
Non coding

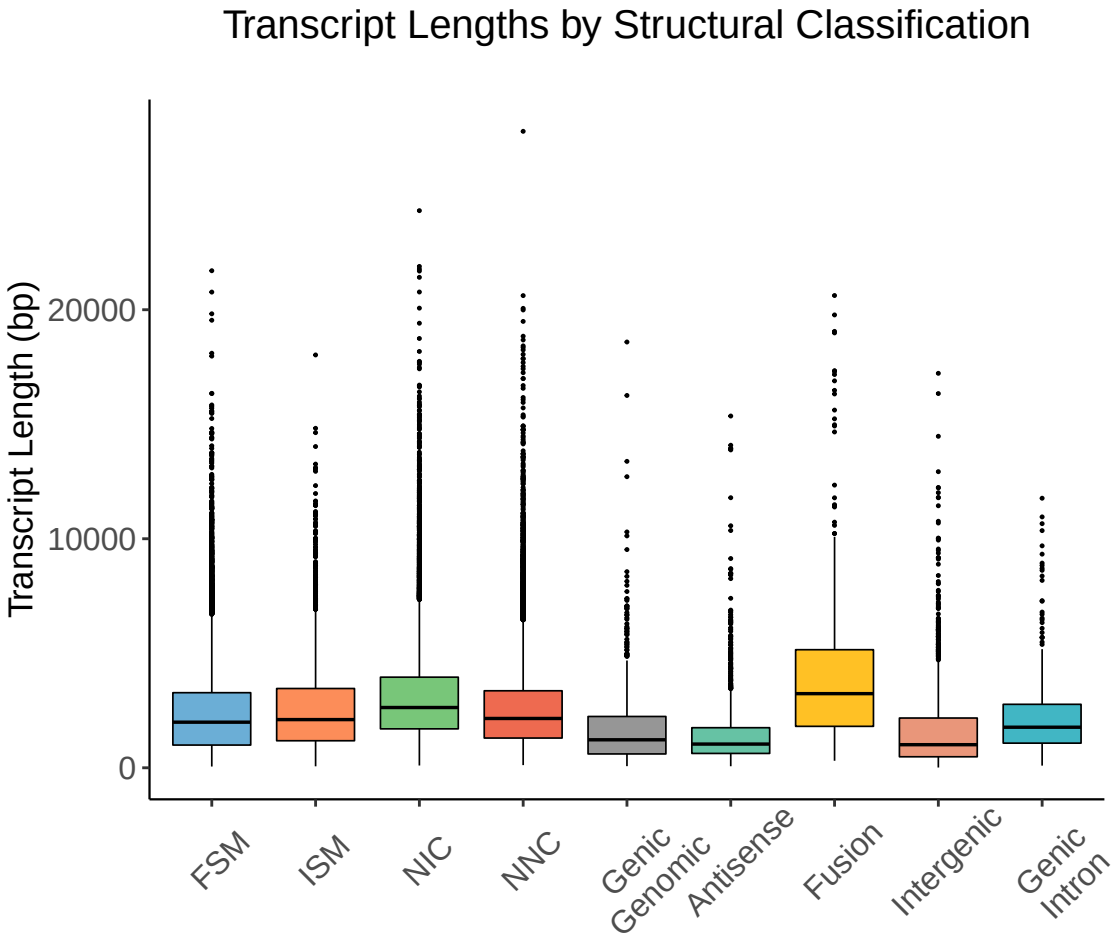

Isoform Distribution Across NIC
Transcripts, %
40
30
20
10
0
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Coding prediction
Coding
Non coding

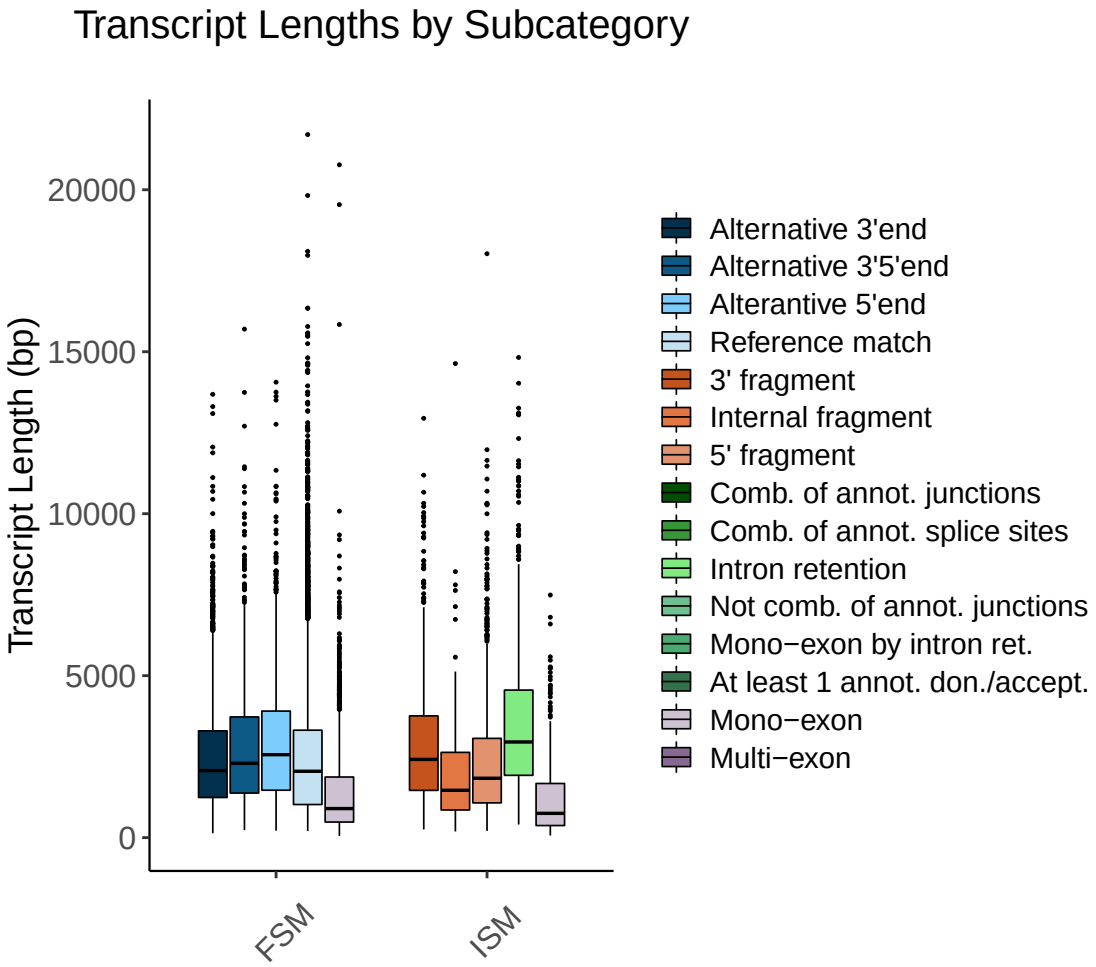

Isoform Distribution Across Genic Genomic
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding
Non coding

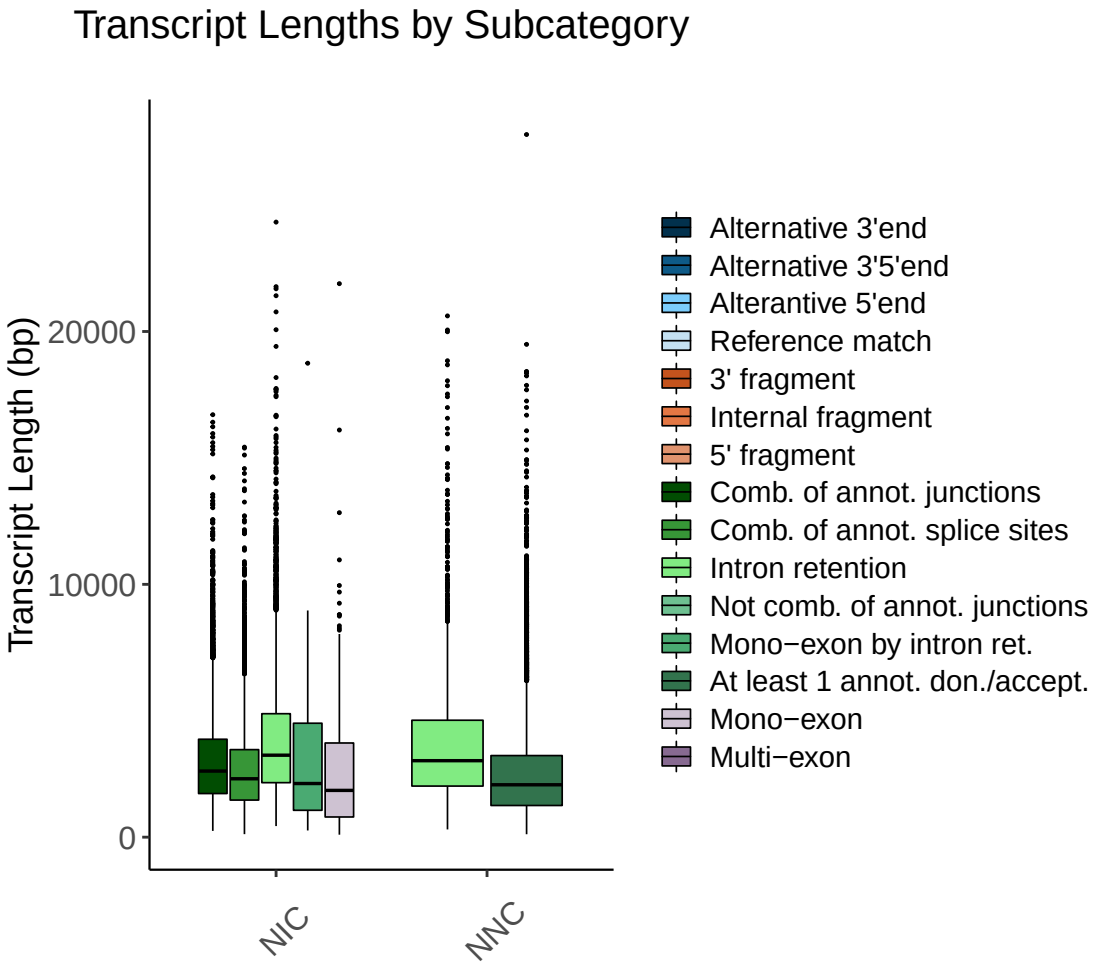

Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding
Non coding

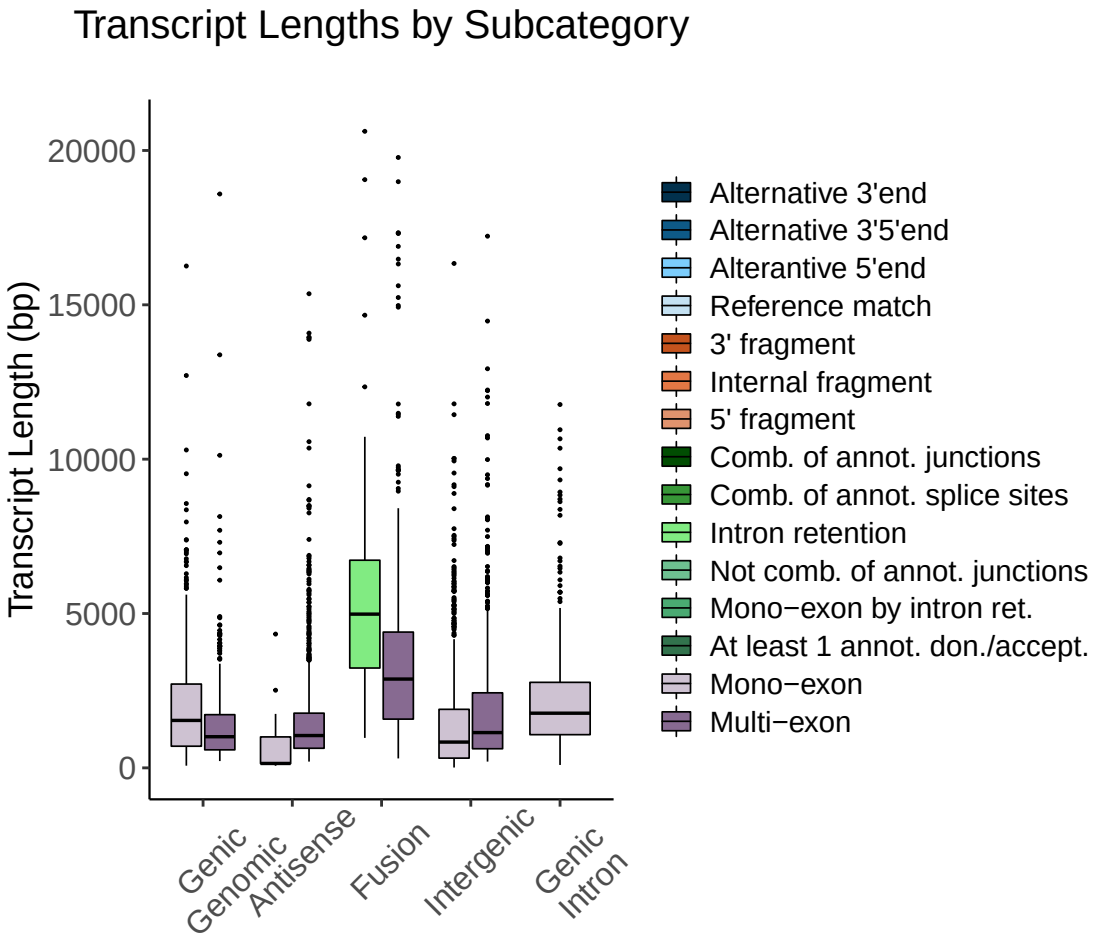

Isoform Distribution Across Fusion
Transcripts, %
80
60
40
20
0
Intron retention
Multi−exon
Coding prediction
Coding
Non coding

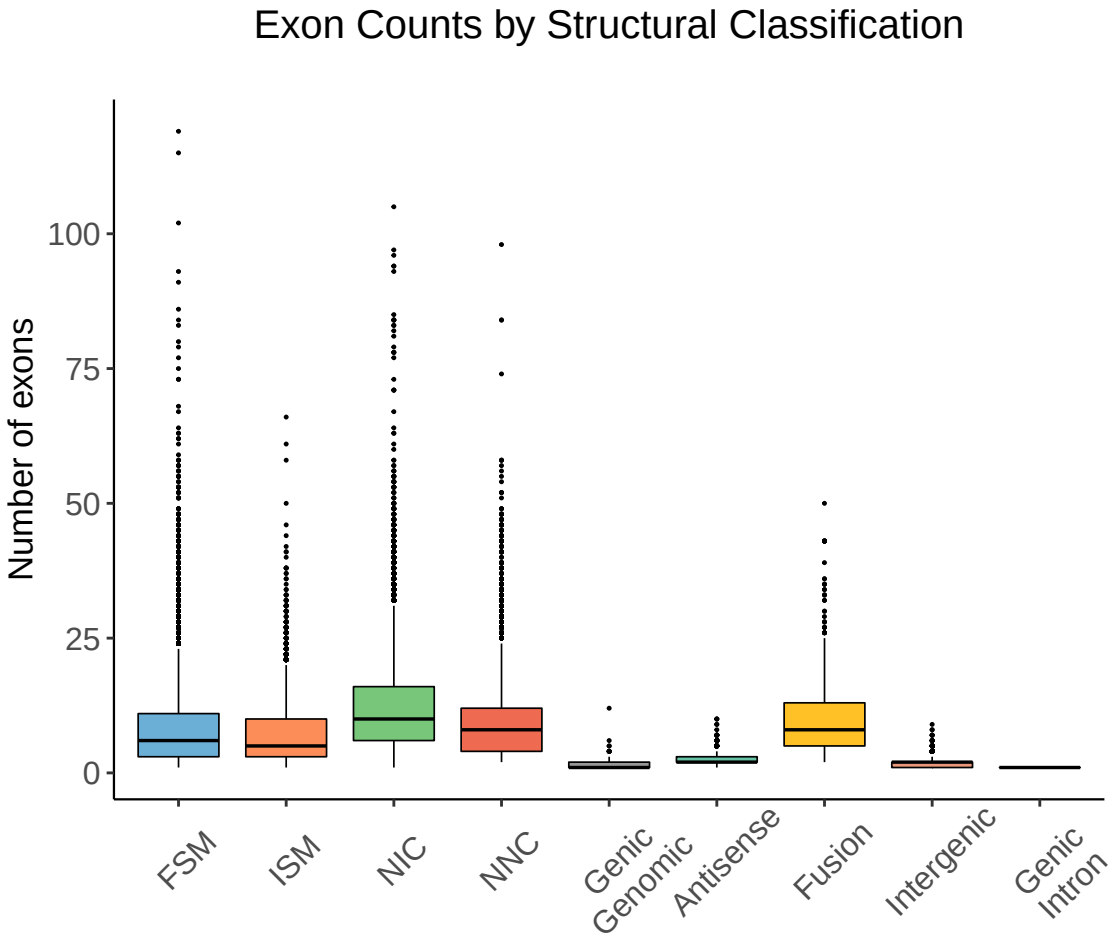

Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding
Non coding

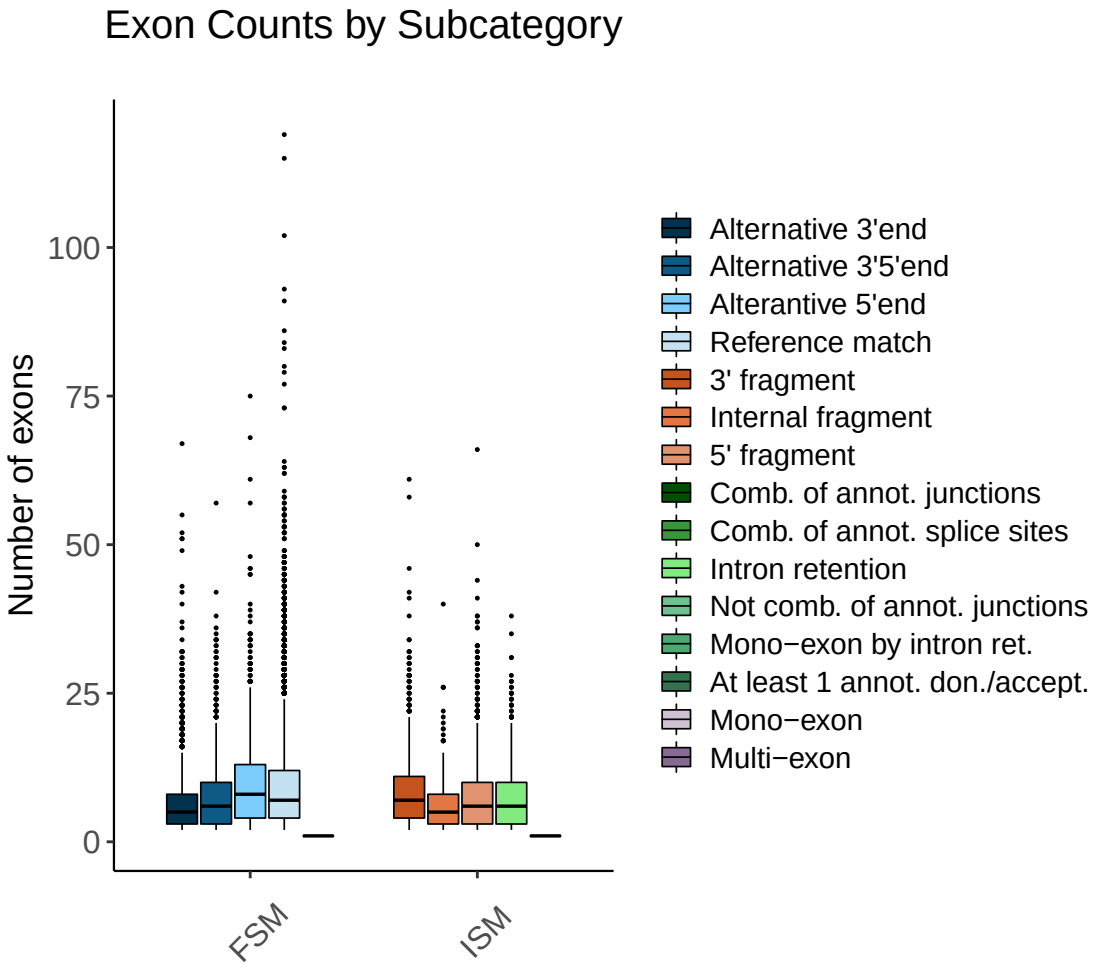

Transcript Lengths by Structural Classification
Transcript Length (bp)
20000
10000
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

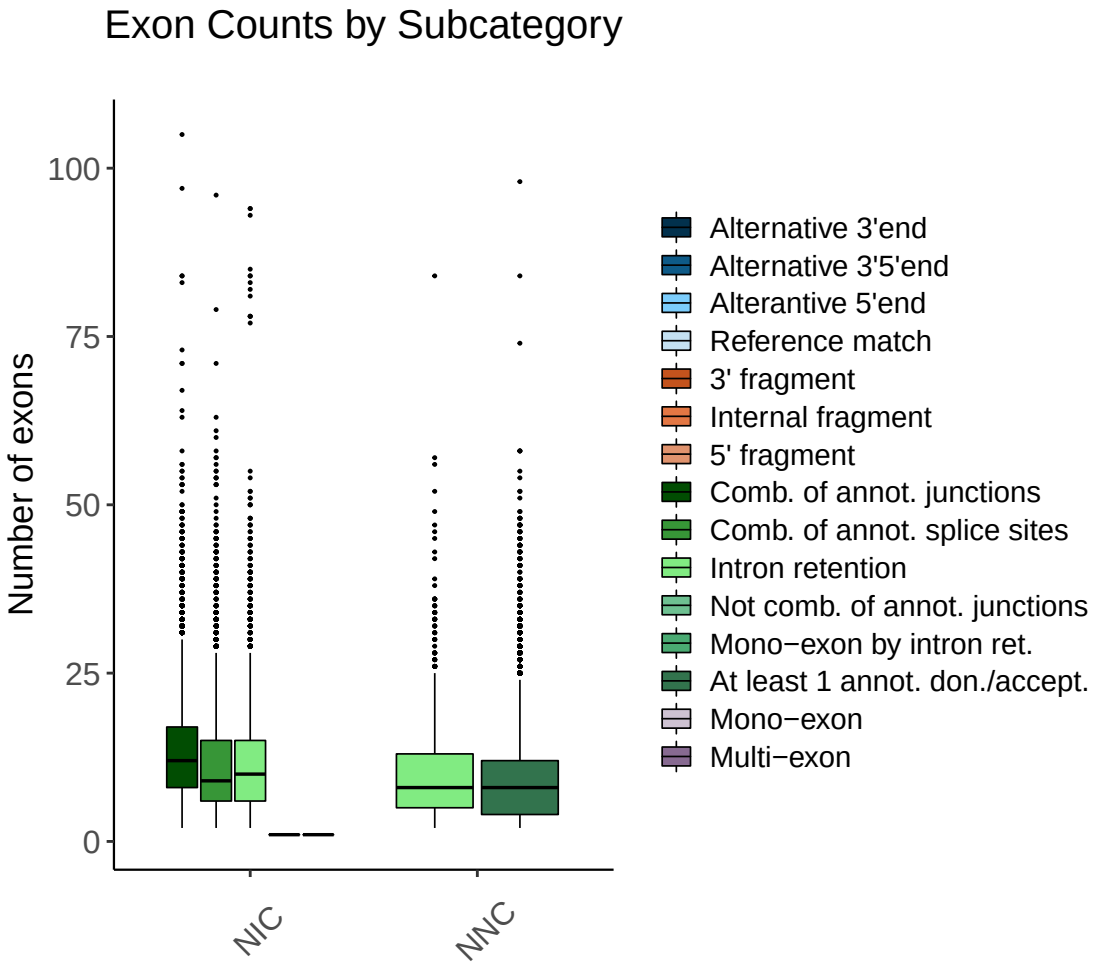

Transcript Lengths by Subcategory
Transcript Length (bp)
20000
15000
10000
5000
0
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

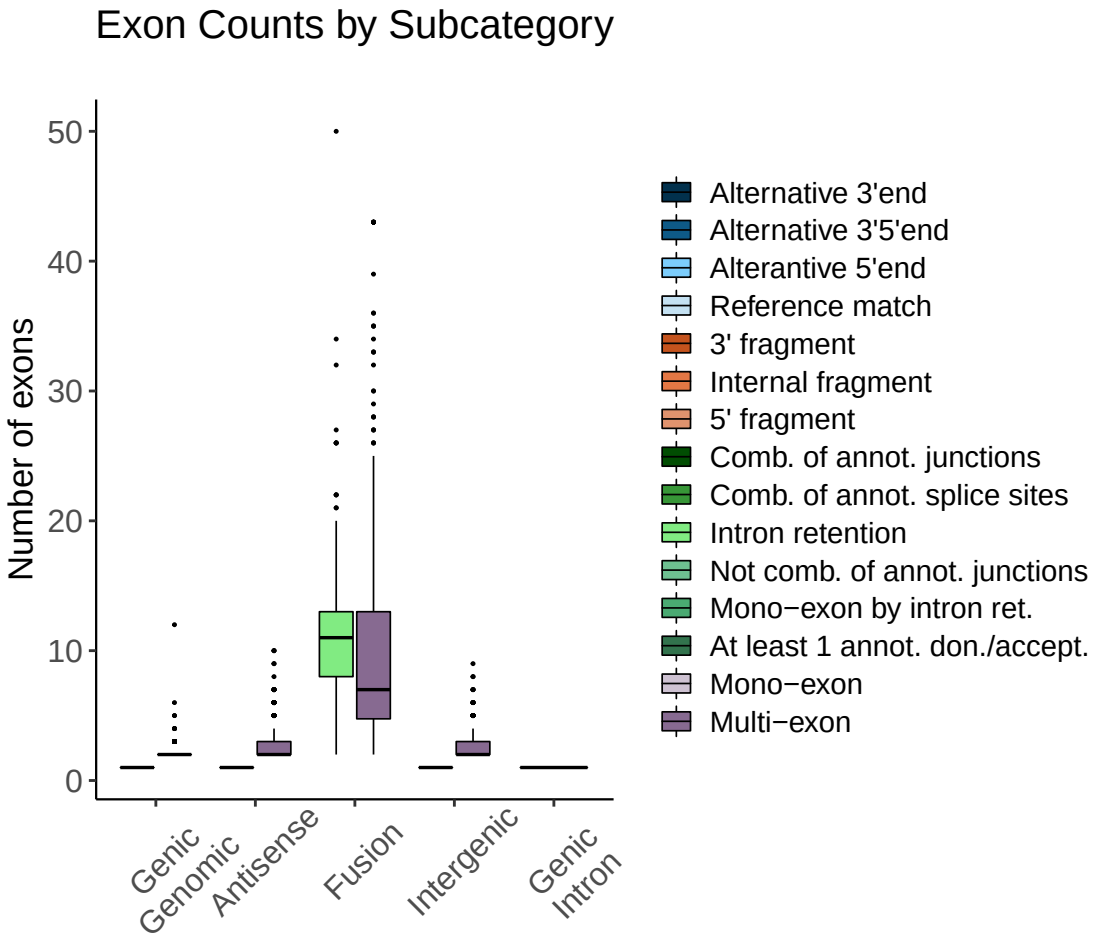

Transcript Lengths by Subcategory
Transcript Length (bp)
20000
10000
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

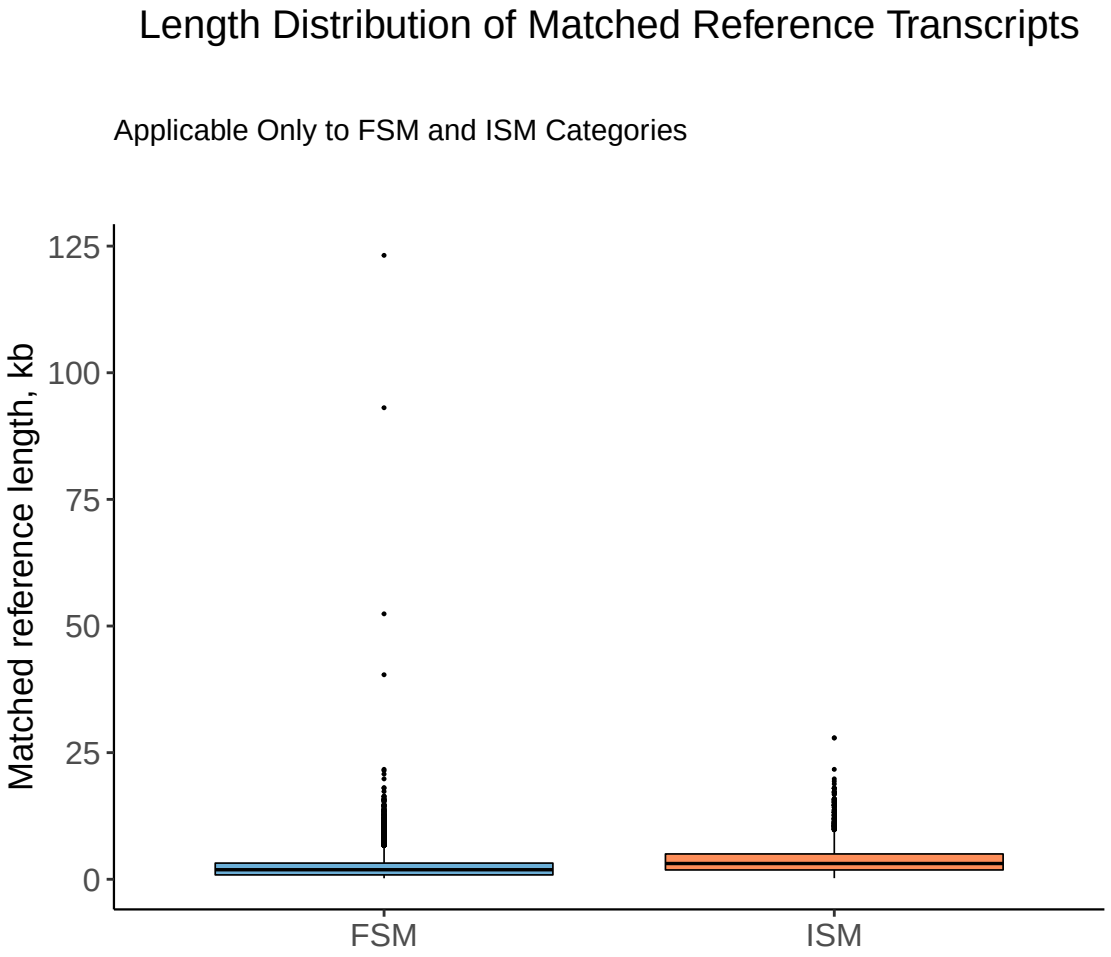

Transcript Lengths by Subcategory
Transcript Length (bp)
20000
15000
10000
5000
0
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
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
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron


Exon Counts by Subcategory
Number of exons
100
75
50
25
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

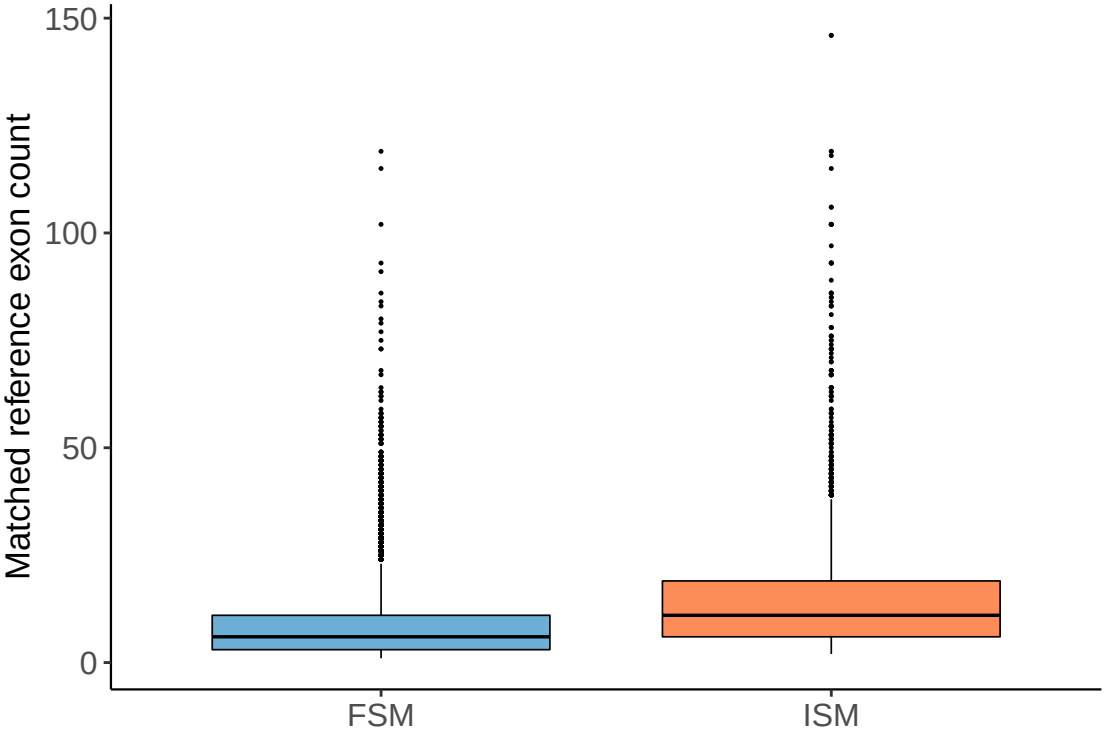

Exon Counts by Subcategory
Number of exons
100
75
50
25
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon


Exon Counts by Subcategory
Number of exons
50
40
30
20
10
0
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
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

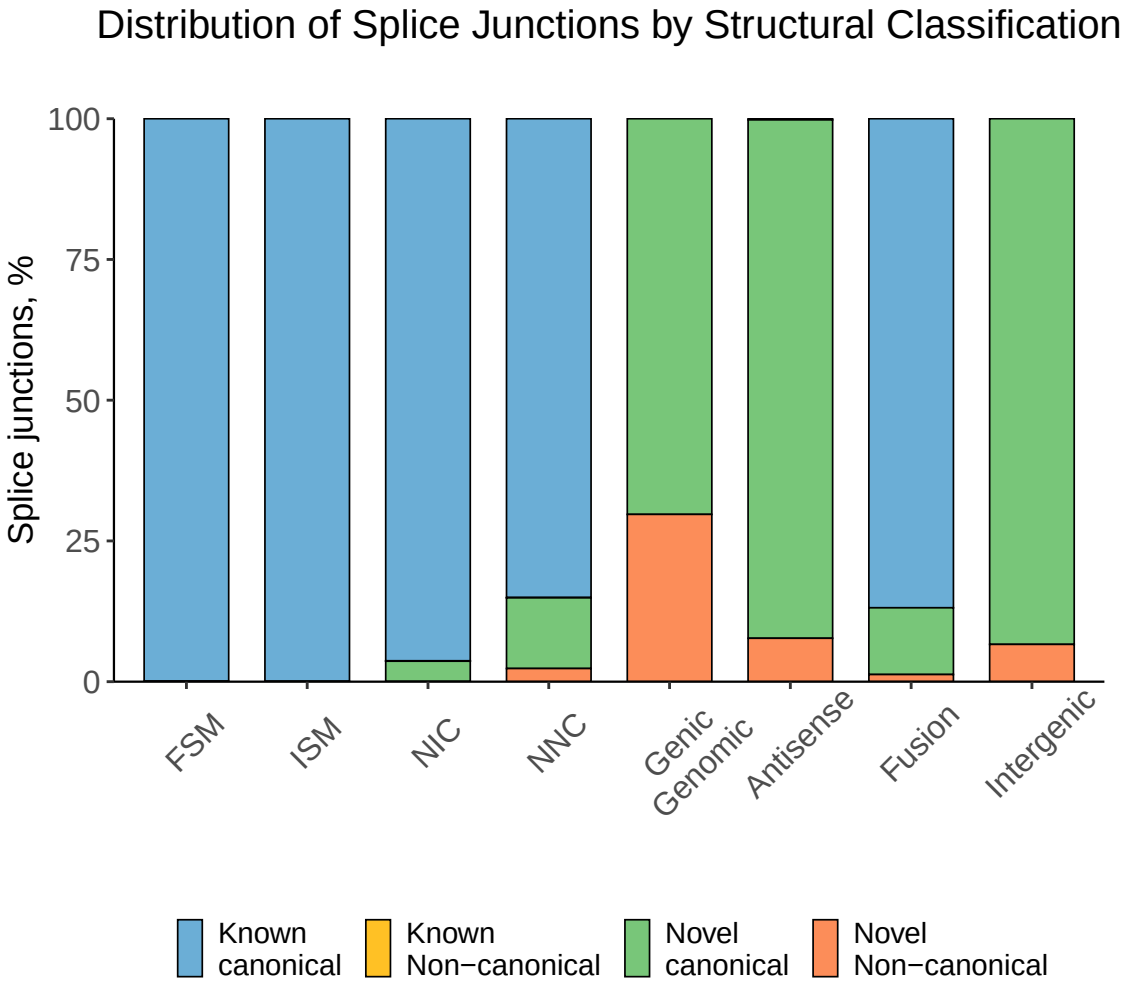

Length Distribution of Matched Reference Transcripts
Applicable Only to FSM and ISM Categories
Matched reference length, kb
125
100
75
50
25
0
FSM
ISM

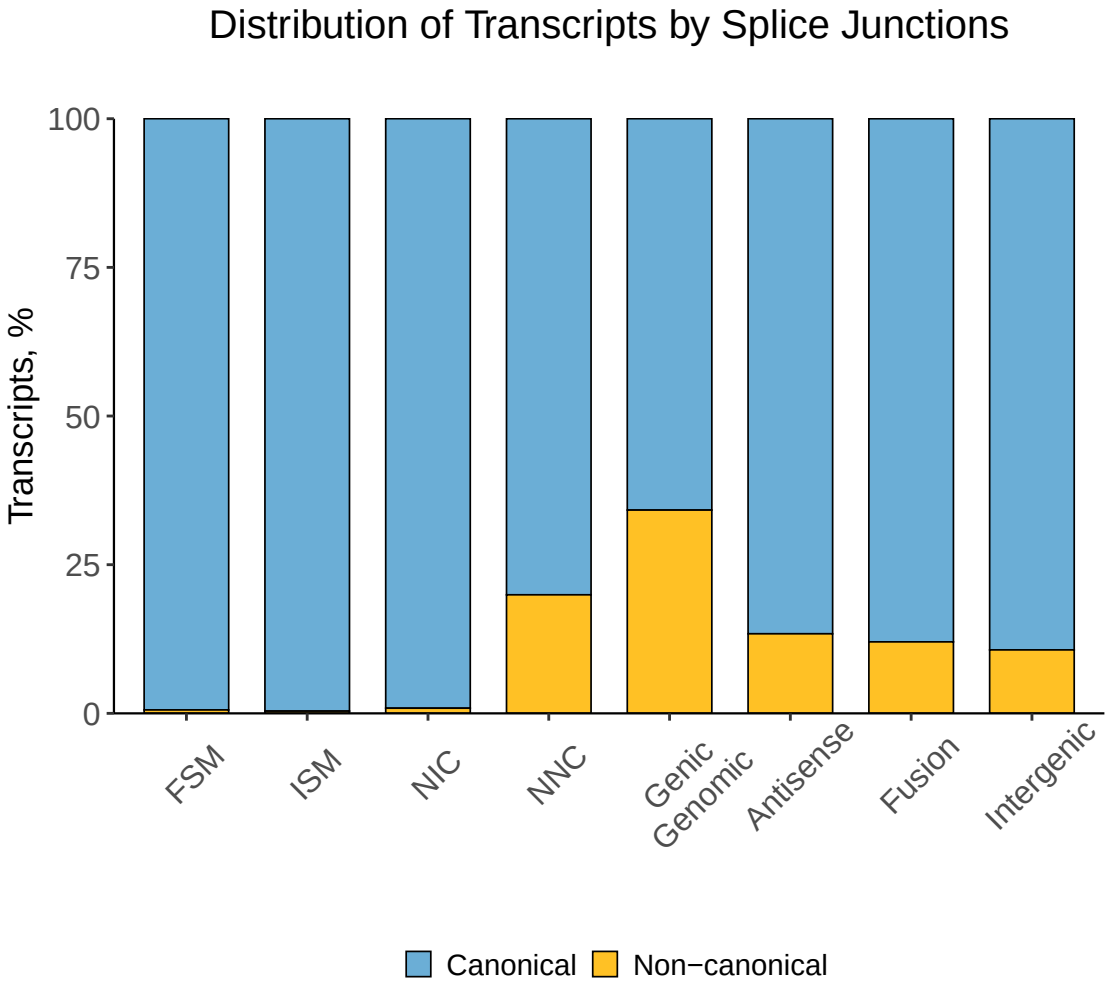
Exon Count Distribution of Matched Reference Transcripts

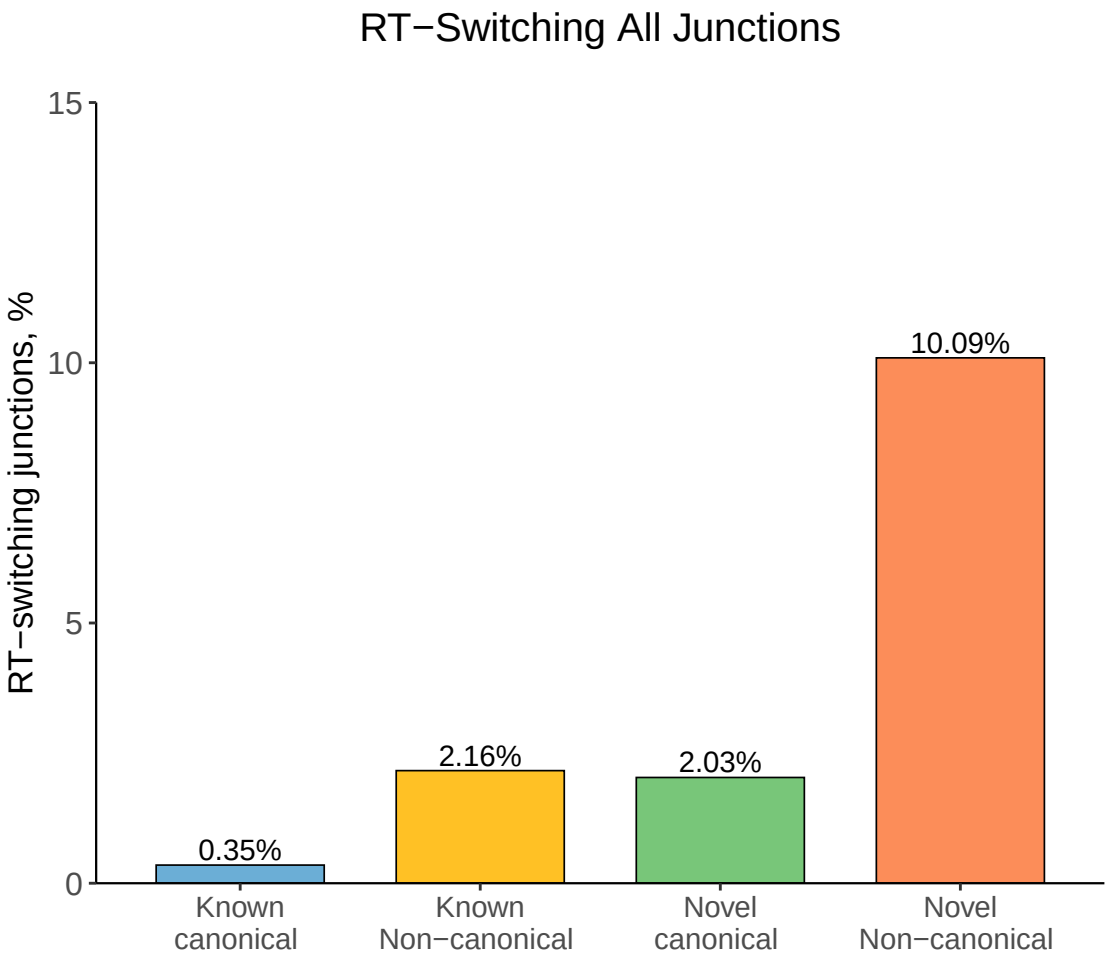
Applicable Only to FSM and ISM Categories

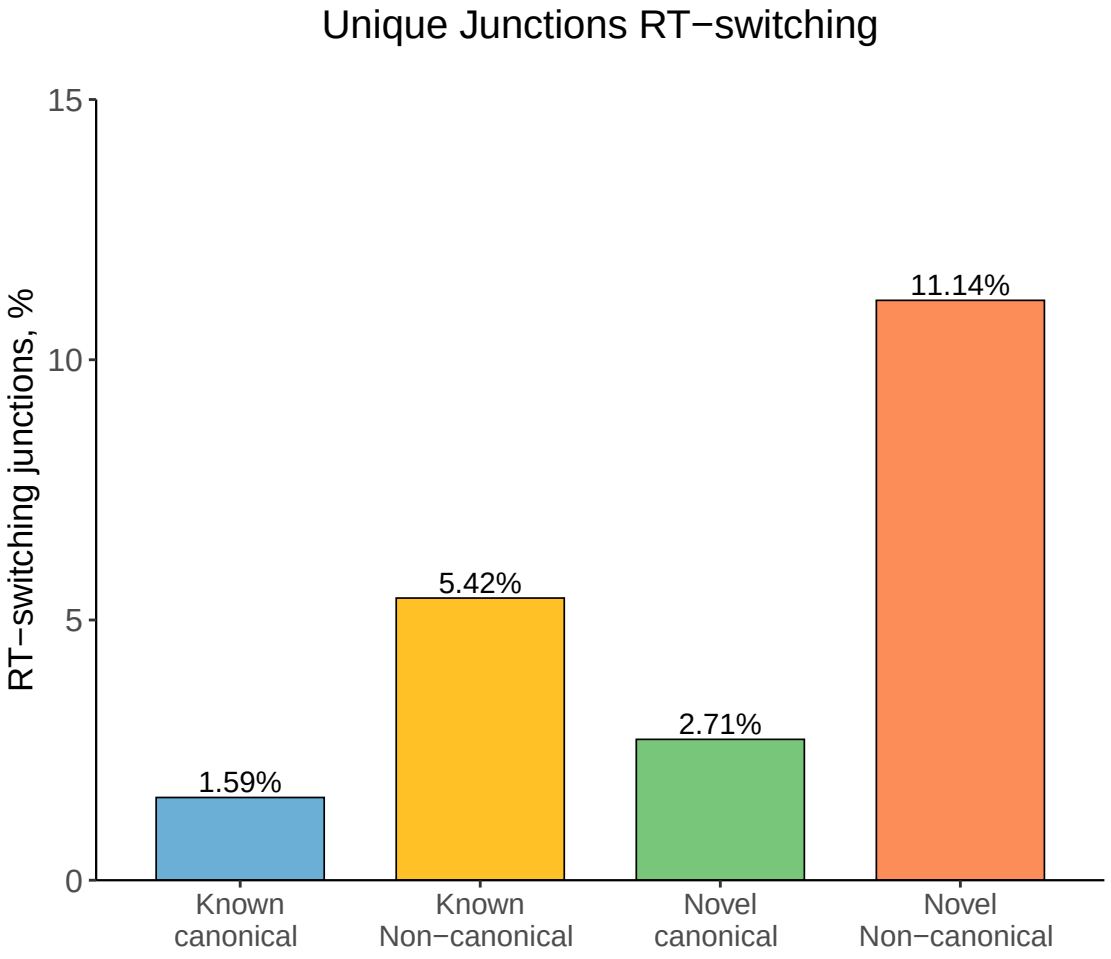


Splice Junction Characterization

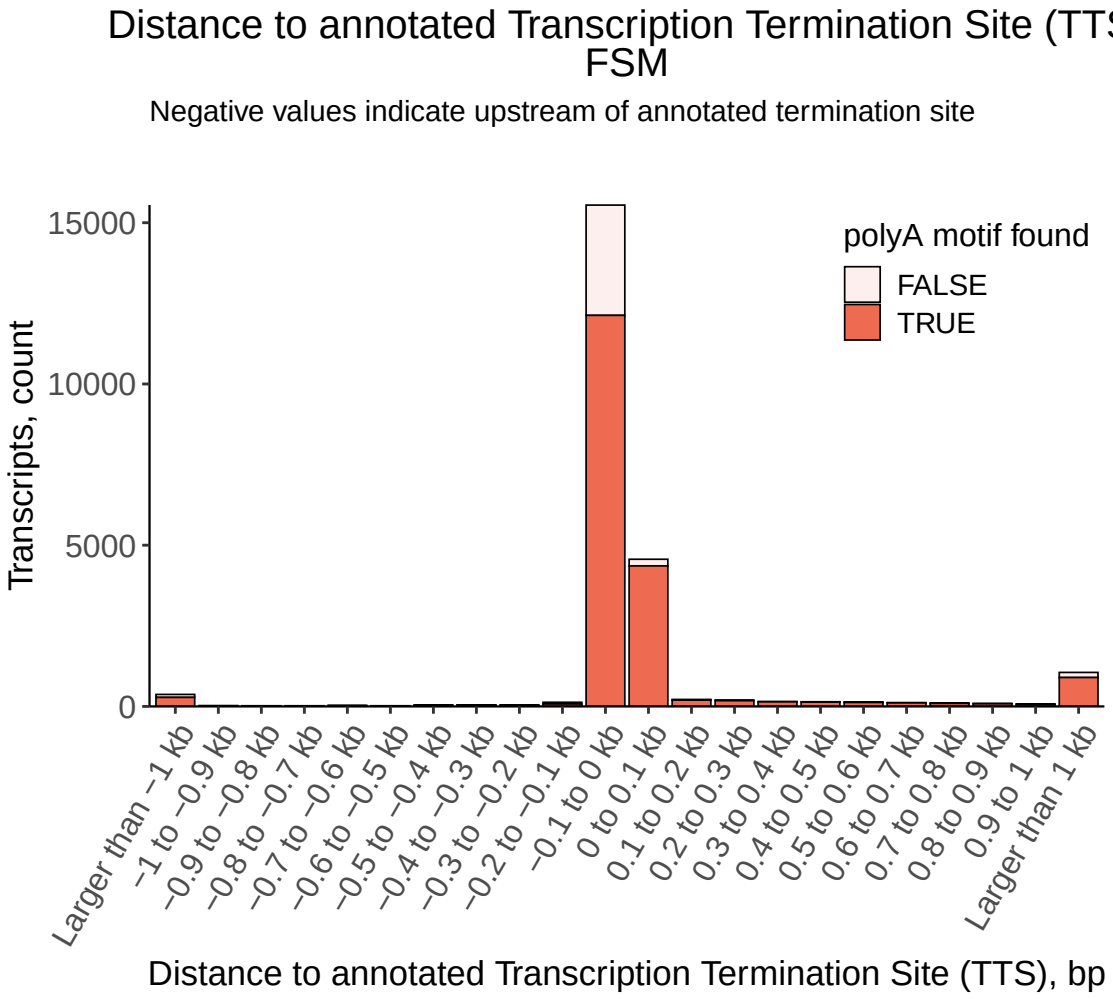

Distribution of Splice Junctions by Structural Classification
Splice junctions, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

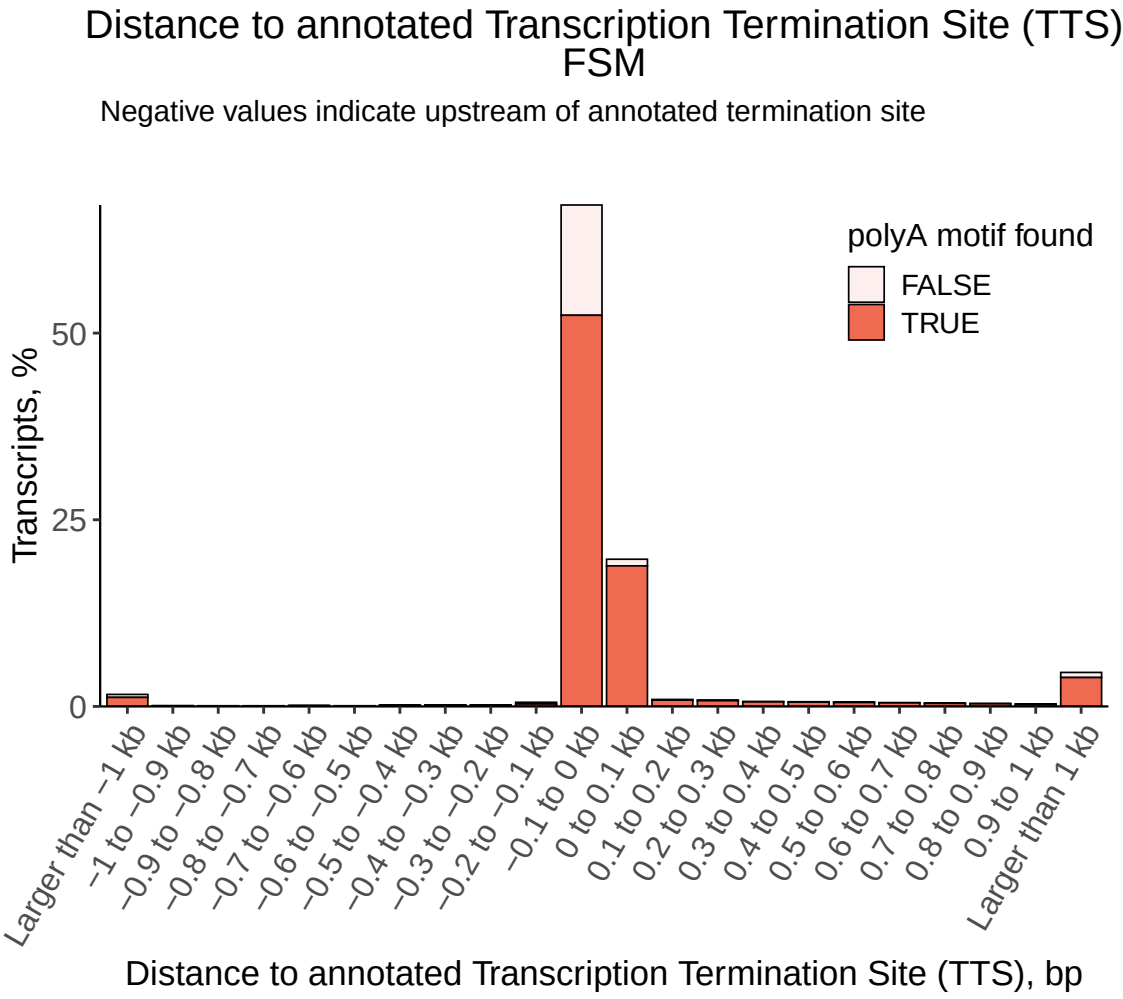

Distribution of Transcripts by Splice Junctions
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Canonical
Non−canonical

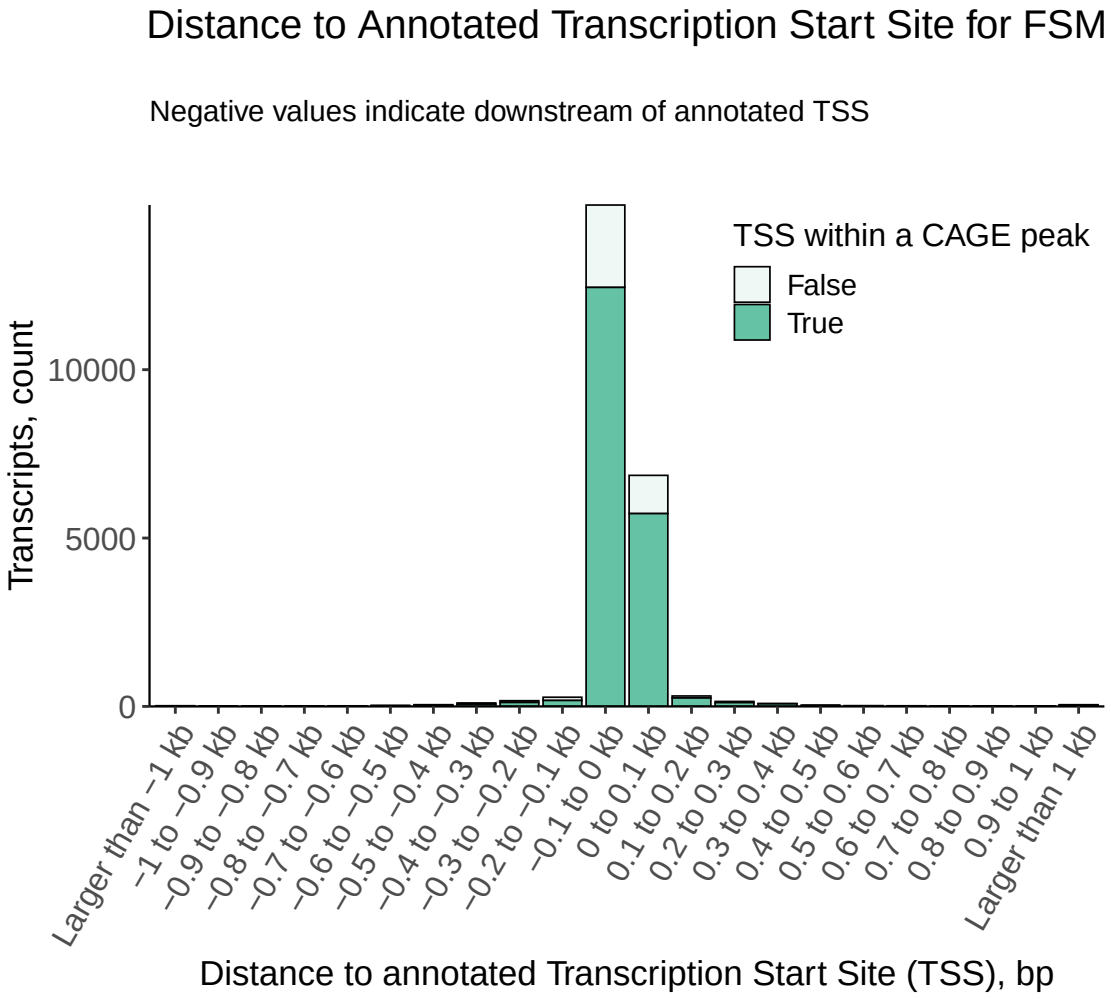

RT−Switching All Junctions
RT−switching junctions, %
15
10
5
0
0.35%
2.16%
2.03%
10.09%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

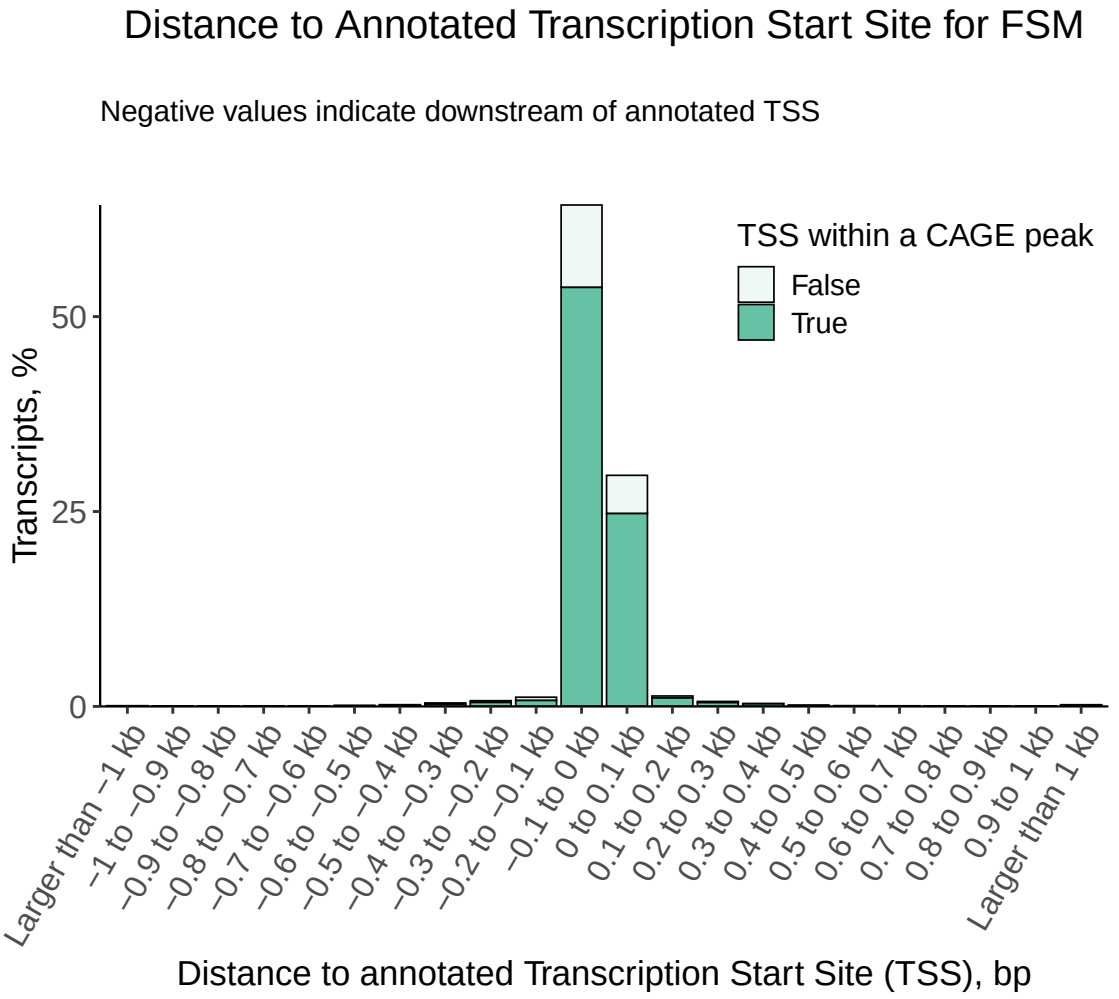

Unique Junctions RT−switching
RT−switching junctions, %
15
10
5
0
1.59%
5.42%
2.71%
11.14%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

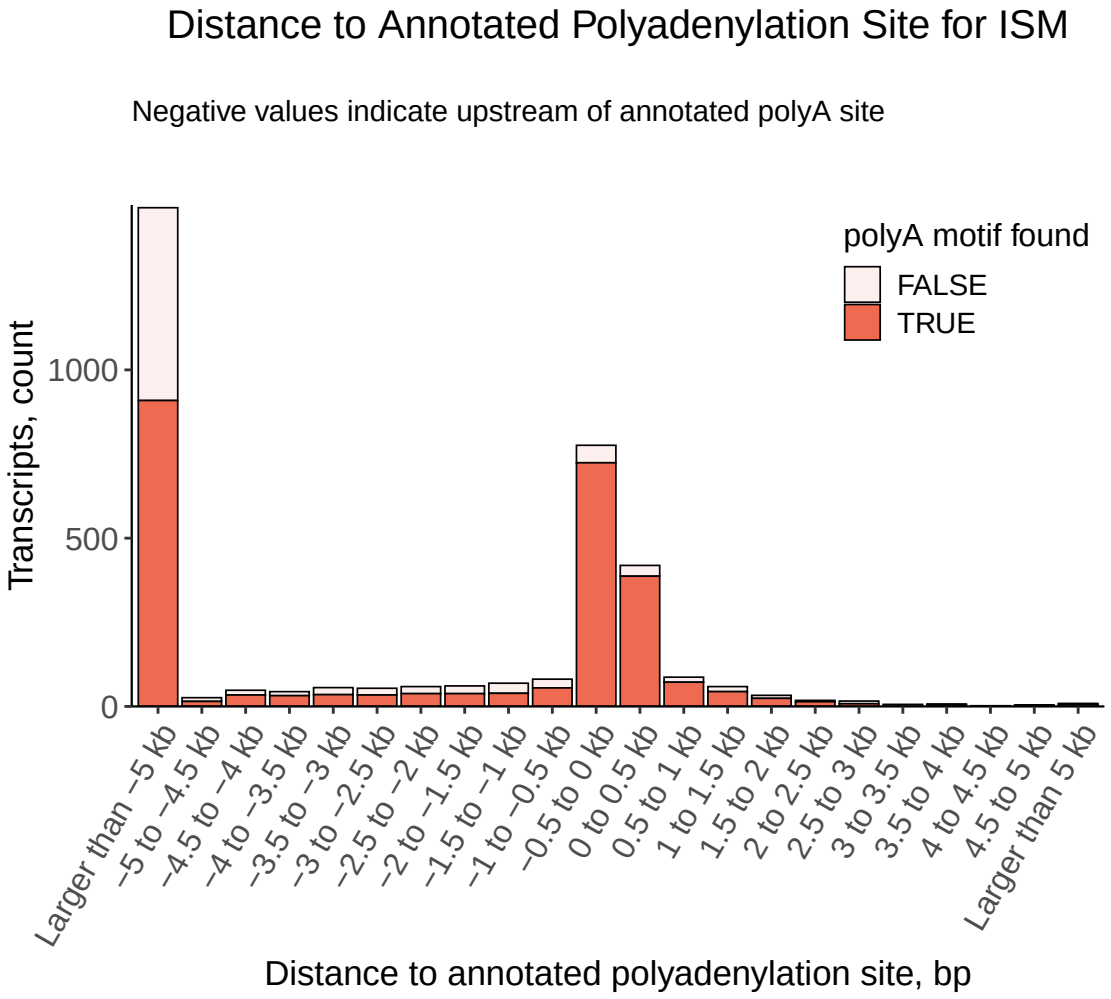
Comparison With Annotated TSS and TTS

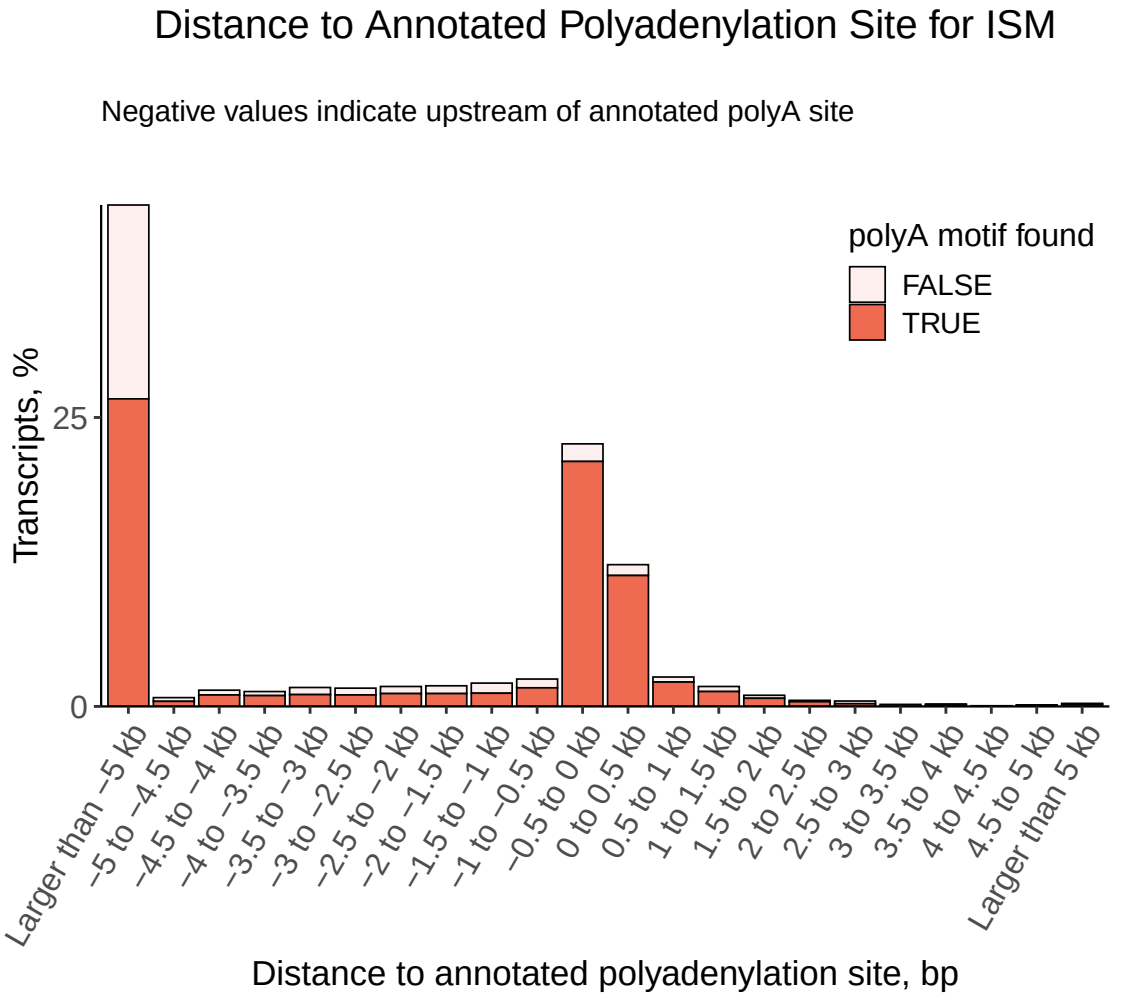

Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
Transcripts, count
15000
10000
5000
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

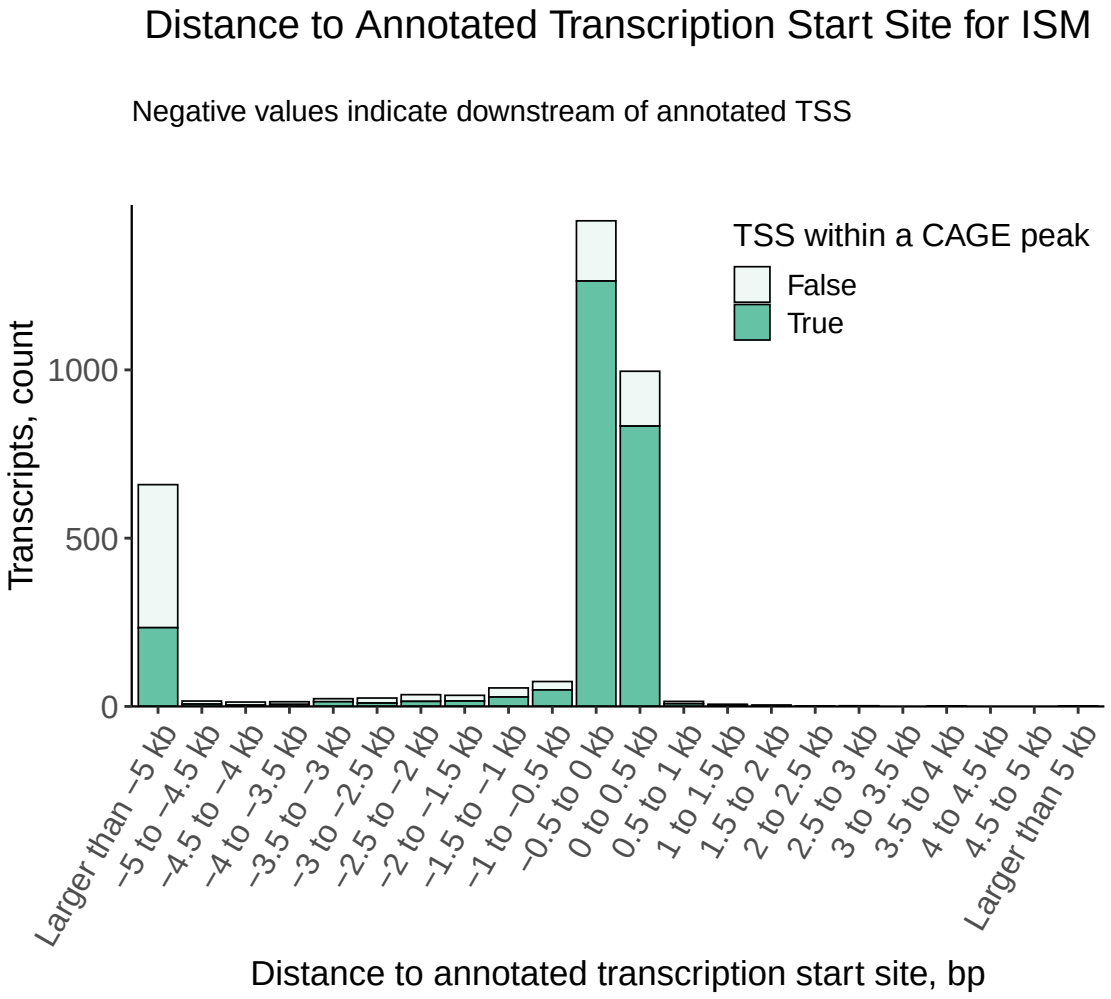

Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3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

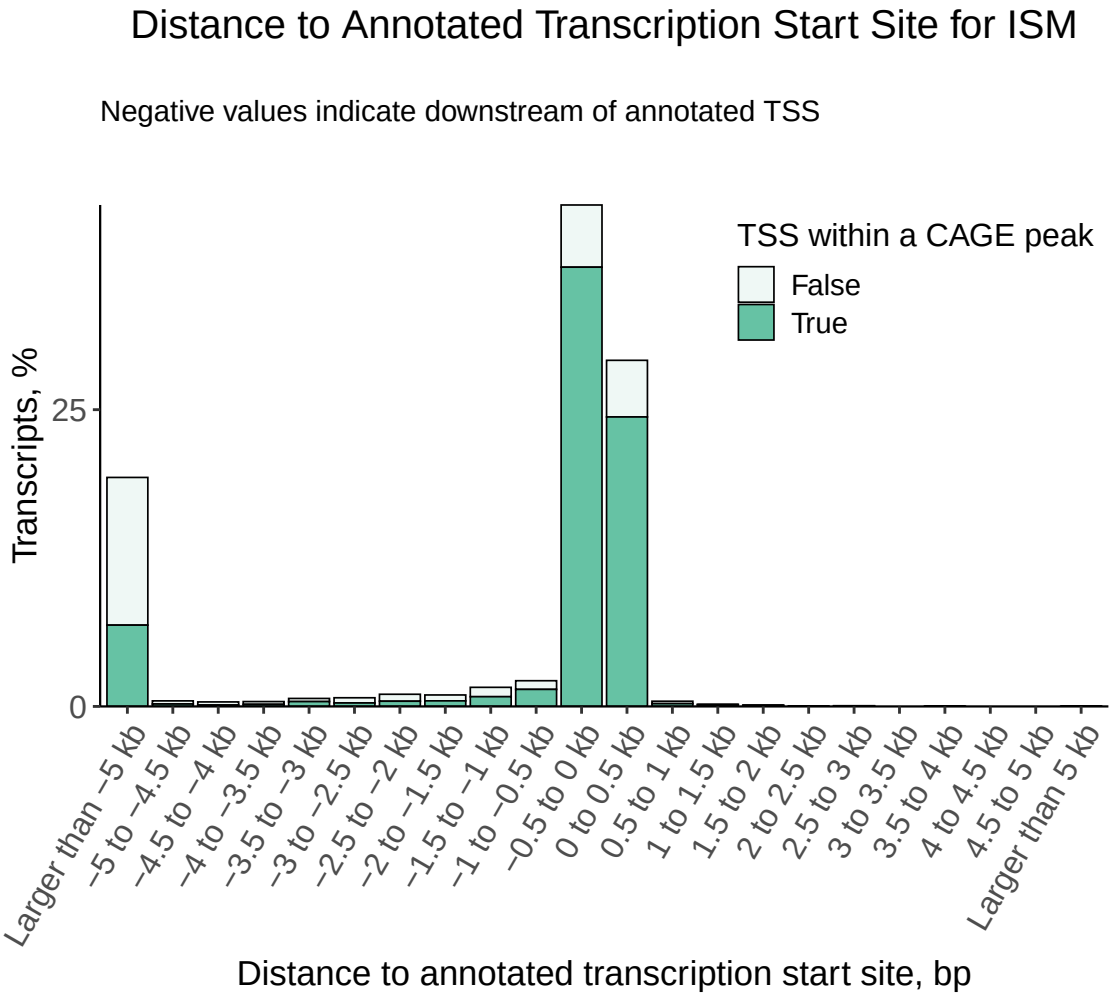

Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
10000
5000
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−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 Start Site (TSS), bp


Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
50
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

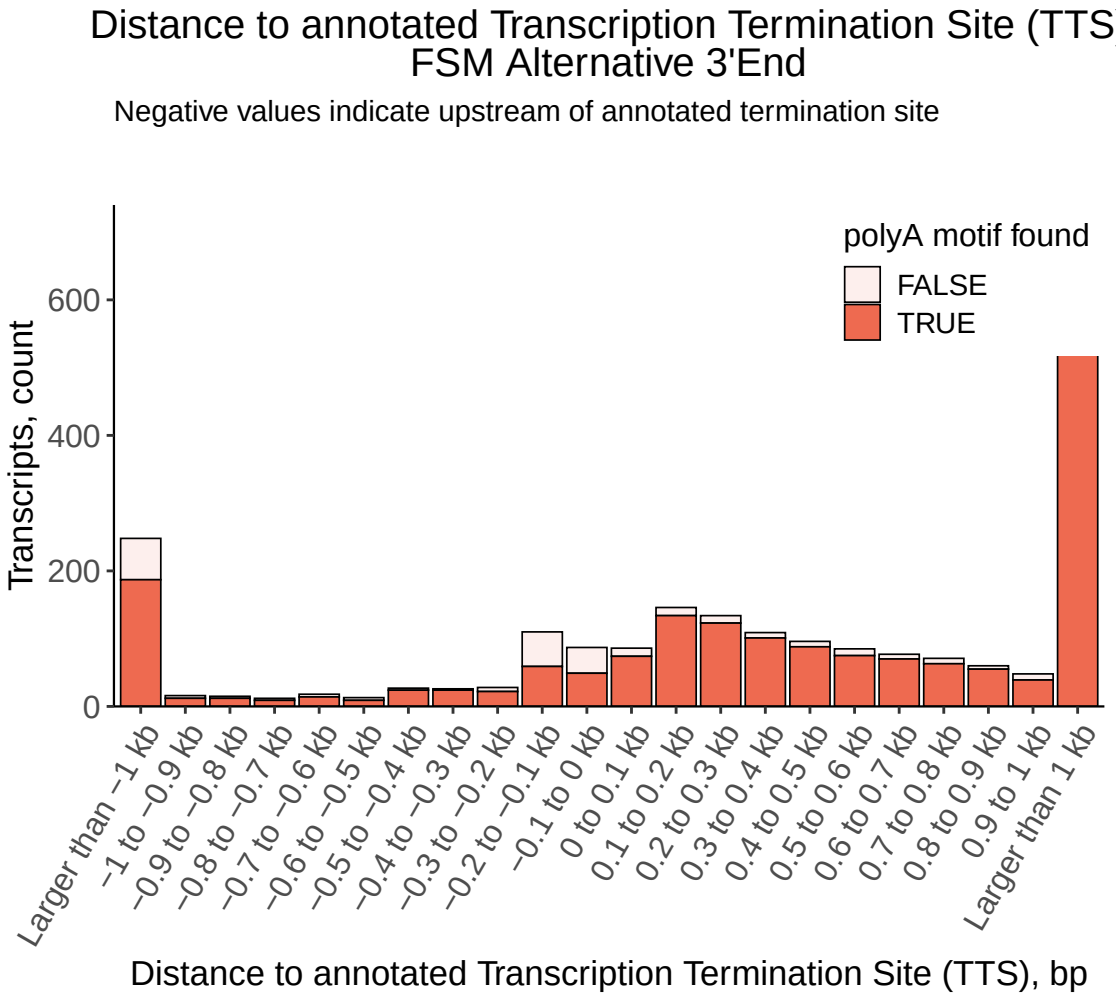

Distance to Annotated Polyadenylation Site for ISM
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
0
500
1000
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 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

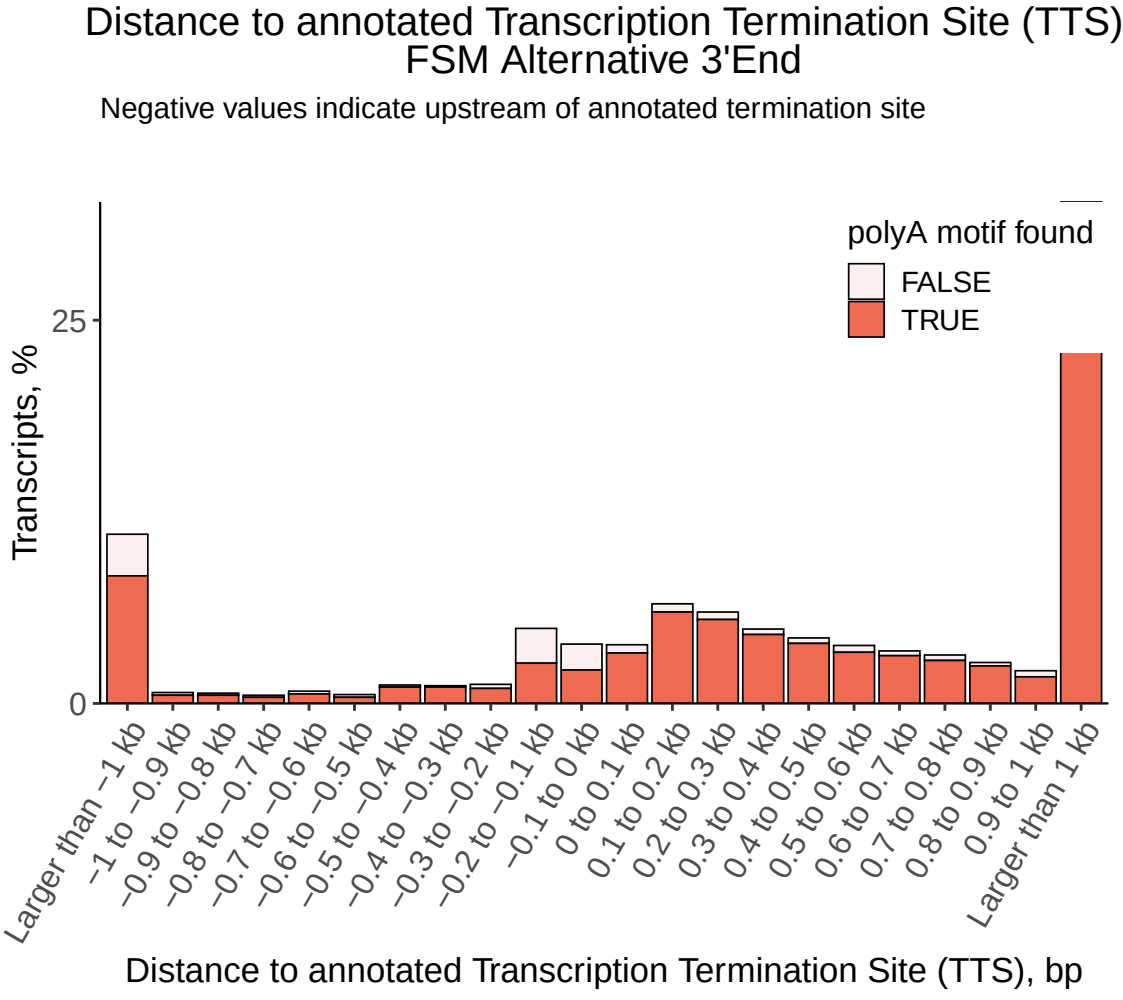

Distance to Annotated Polyadenylation Site for ISM
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb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

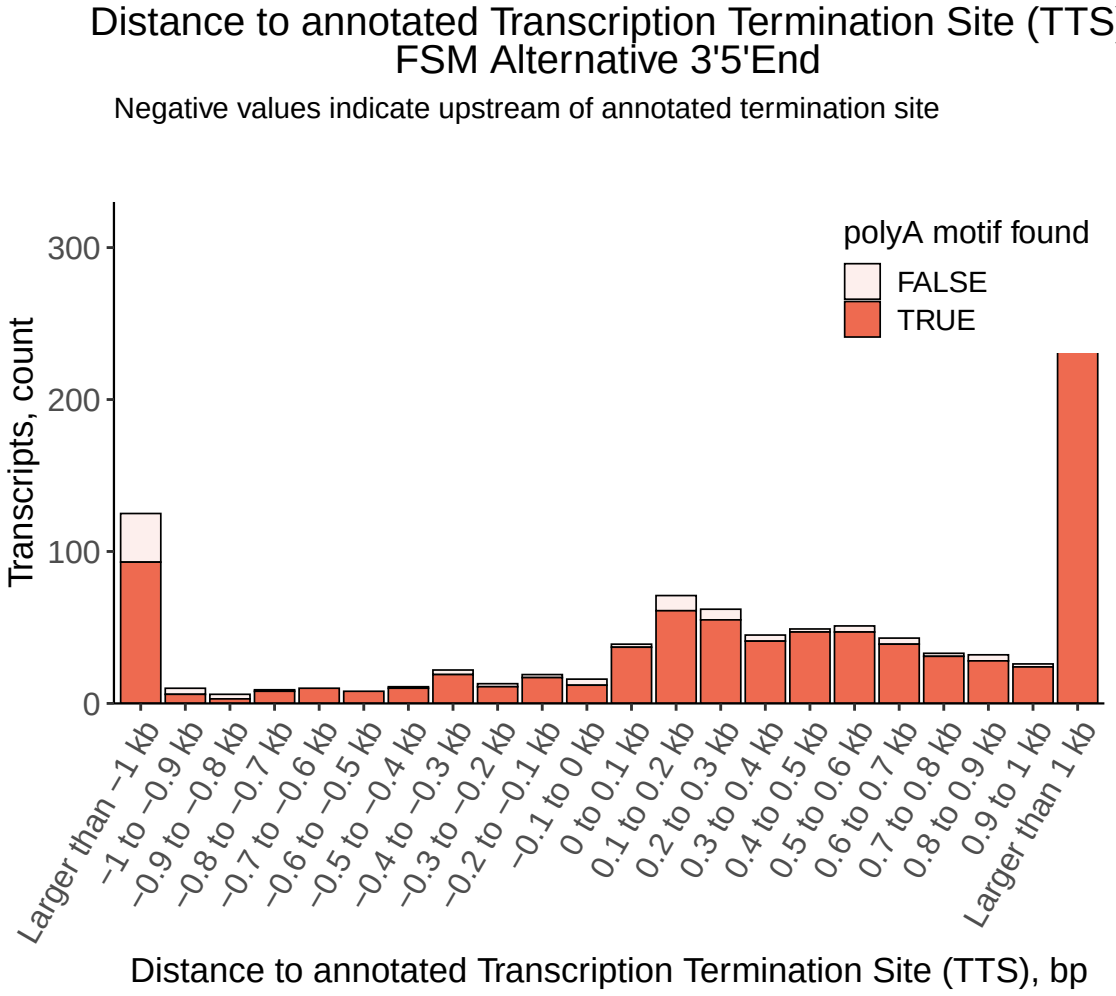

Distance to Annotated Transcription Start Site for ISM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
0
500
1000
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−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 transcription start site, bp

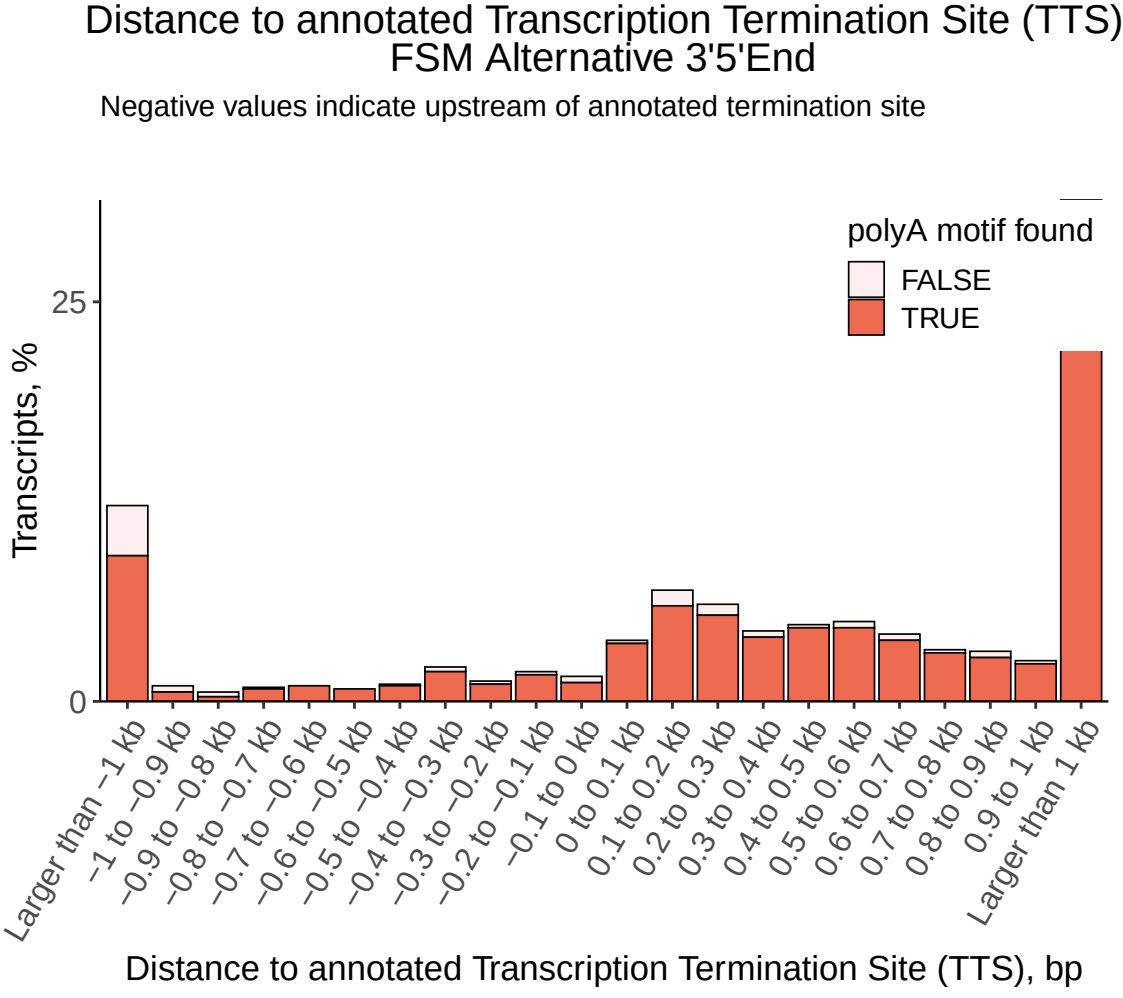

Distance to Annotated Transcription Start Site for ISM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 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 transcription start site, bp

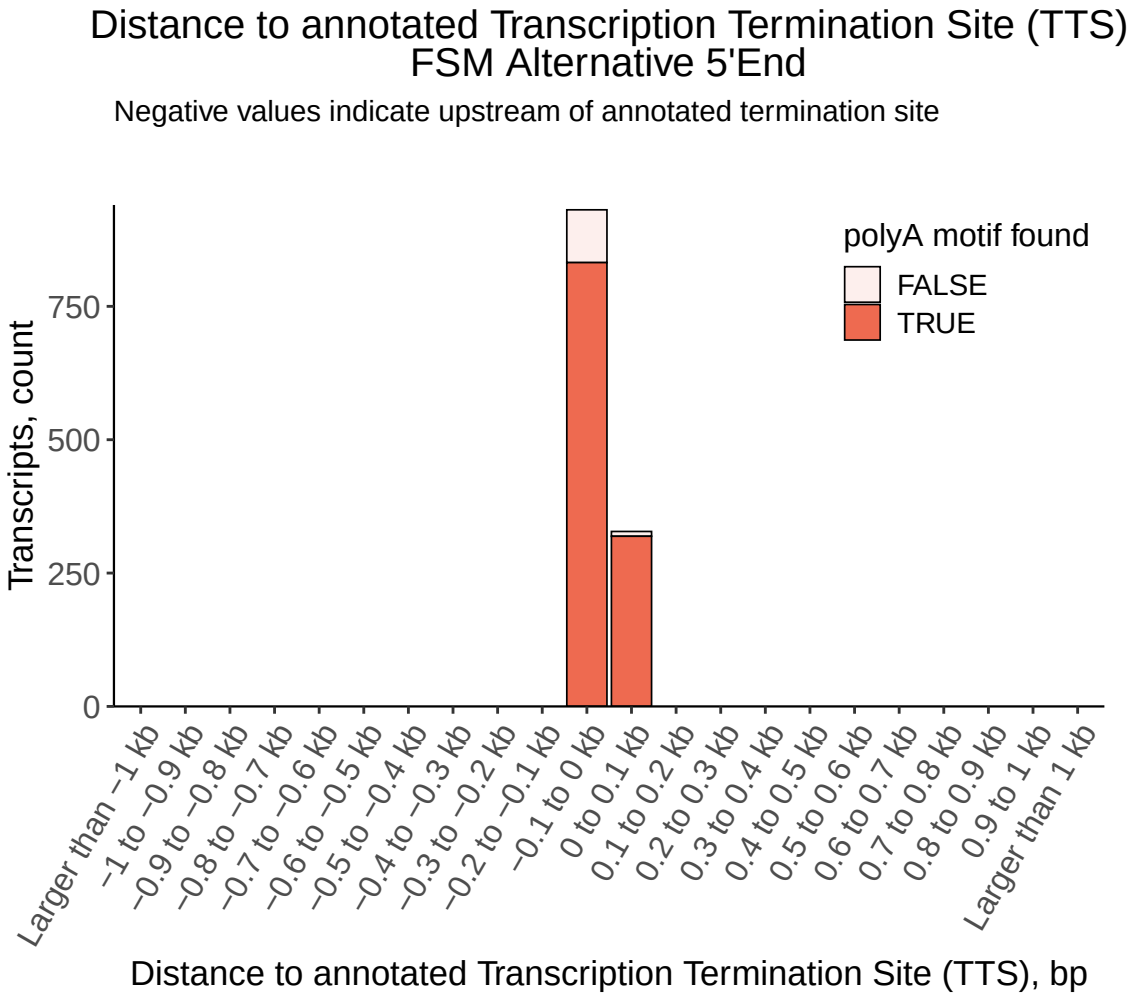
Comparison With Annotated TSS and TTS by Subcategories

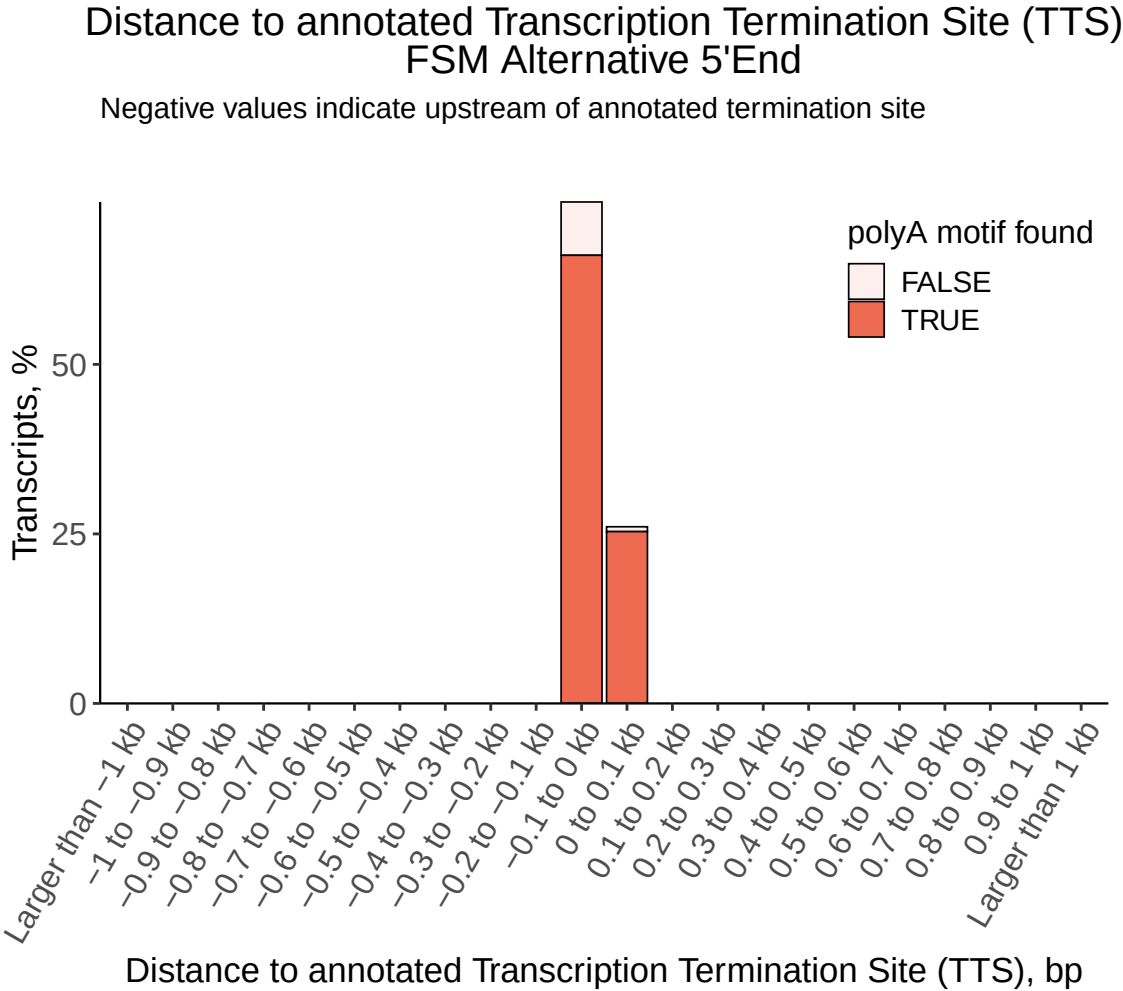

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
200
400
600
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−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

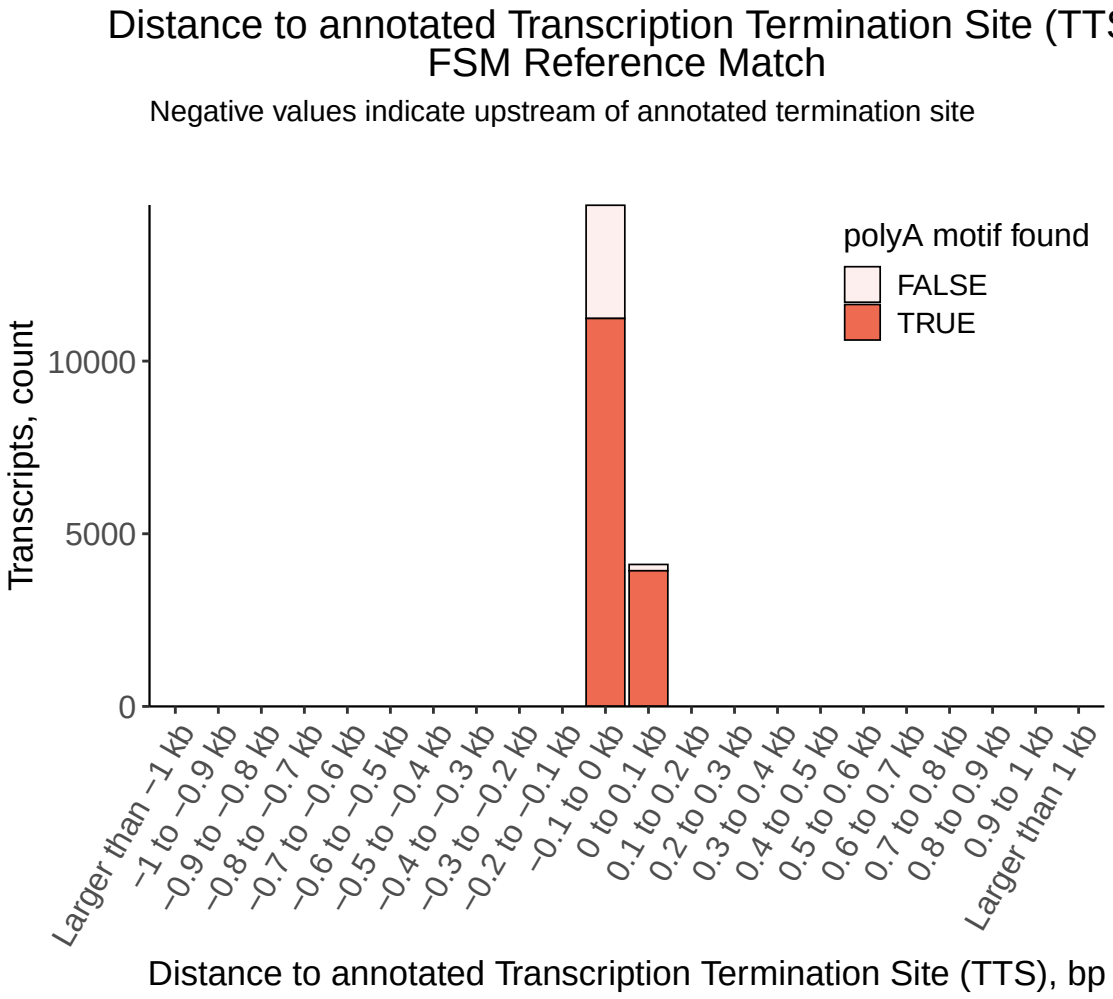

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

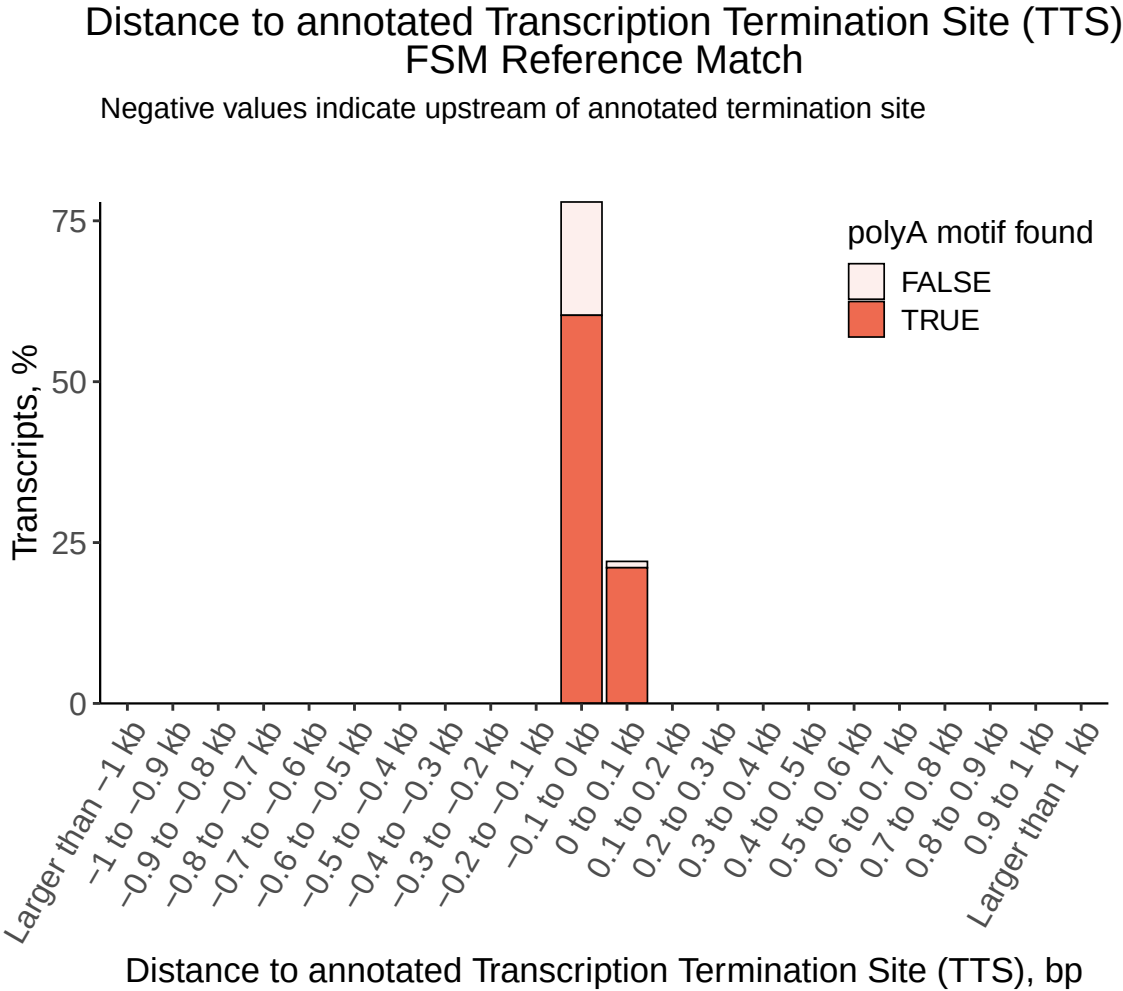

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
300
200
100
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 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

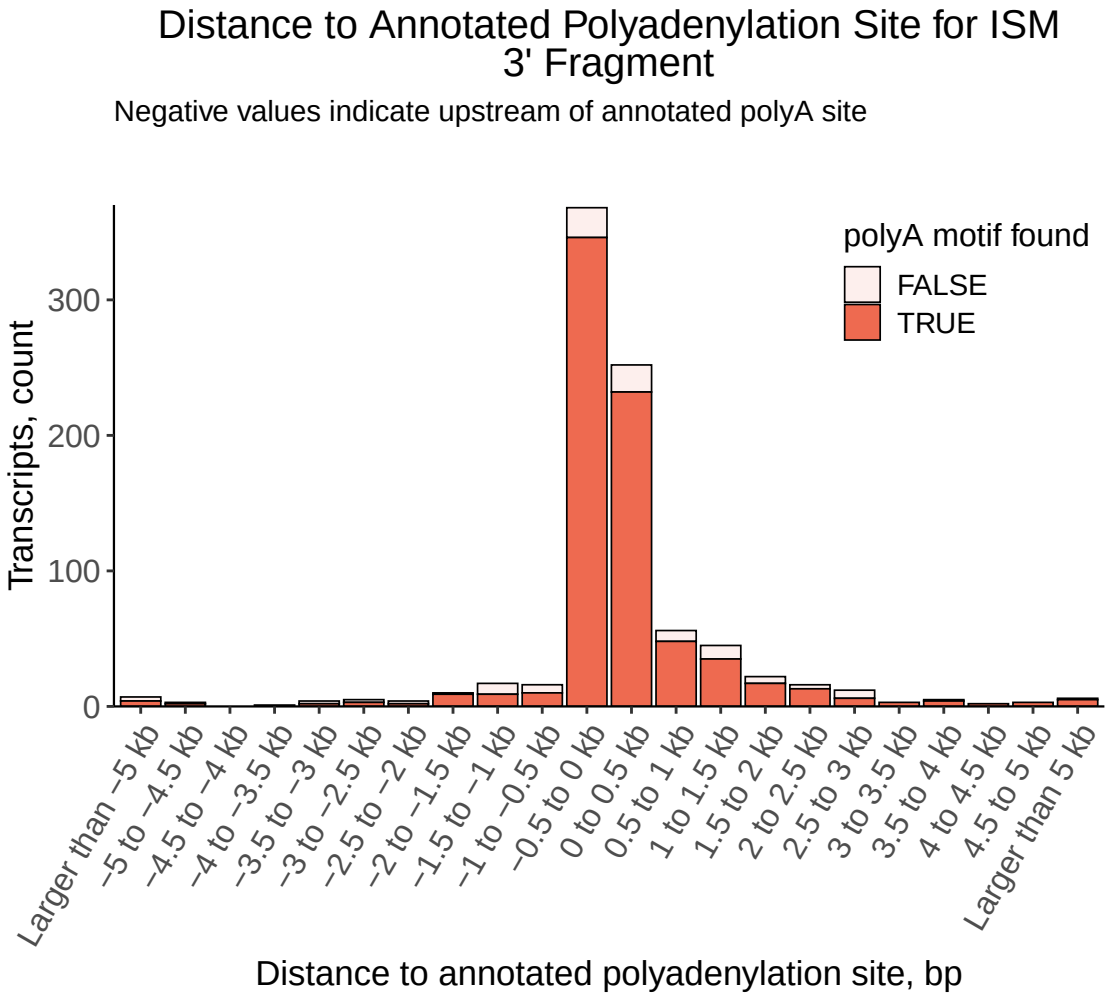

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

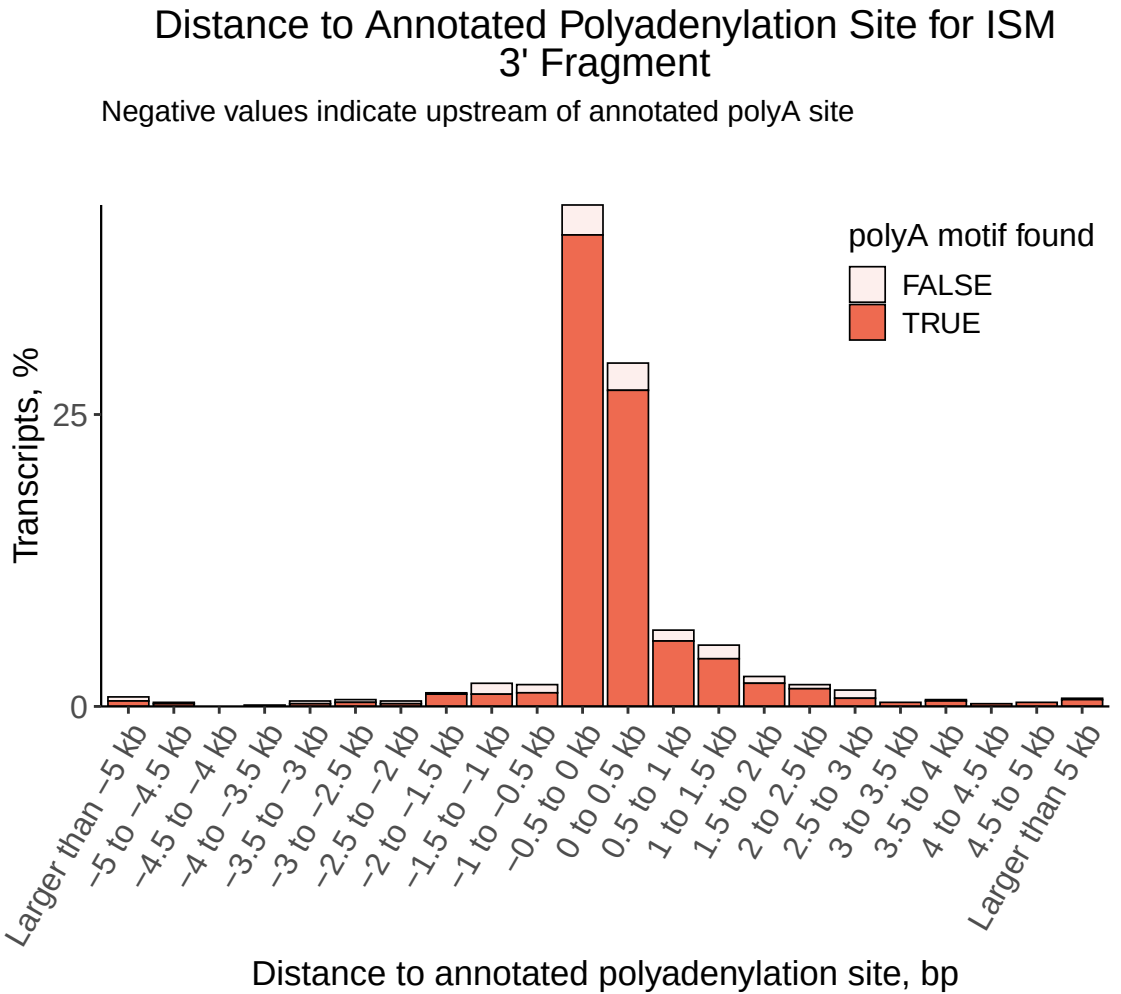

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
250
500
750
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−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

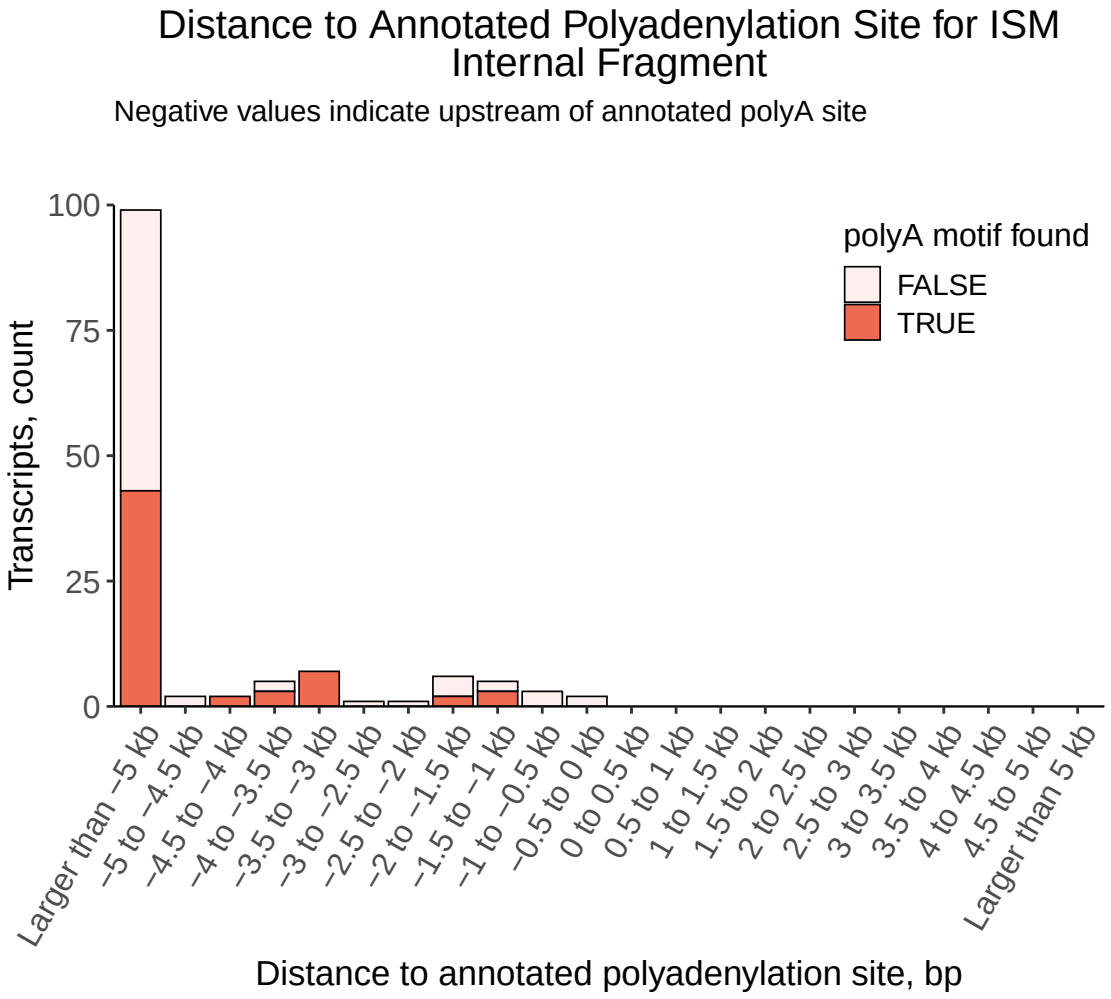

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −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

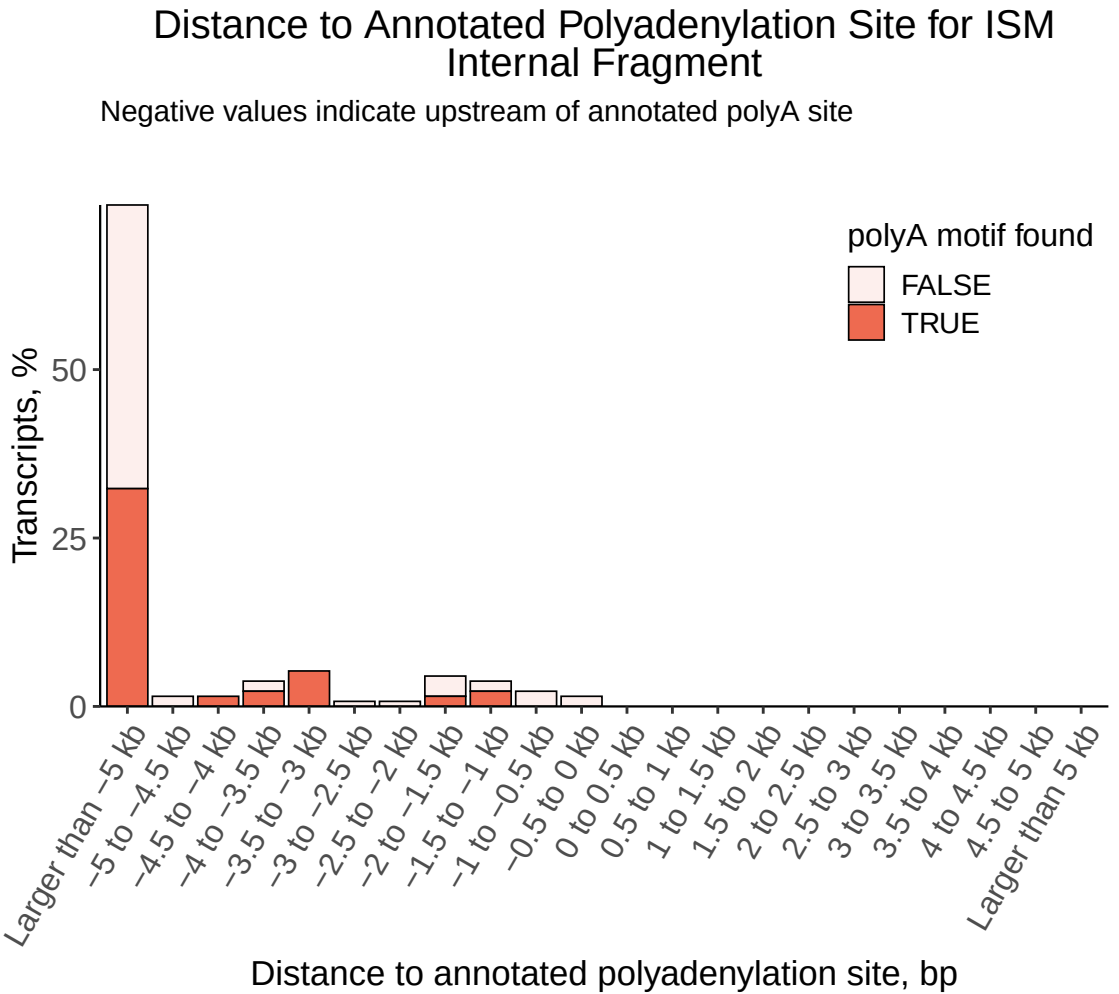

Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
5000
10000
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5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

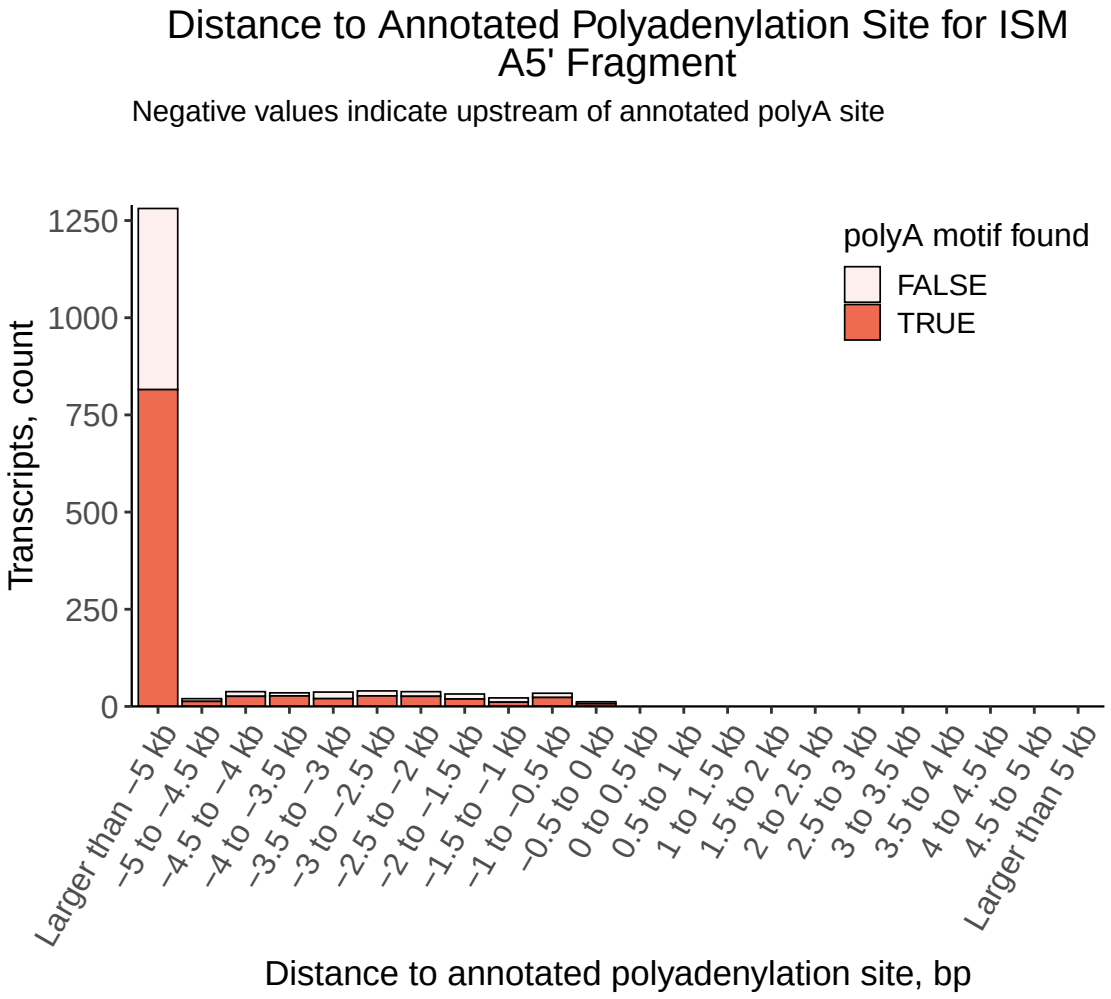

Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, %
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to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

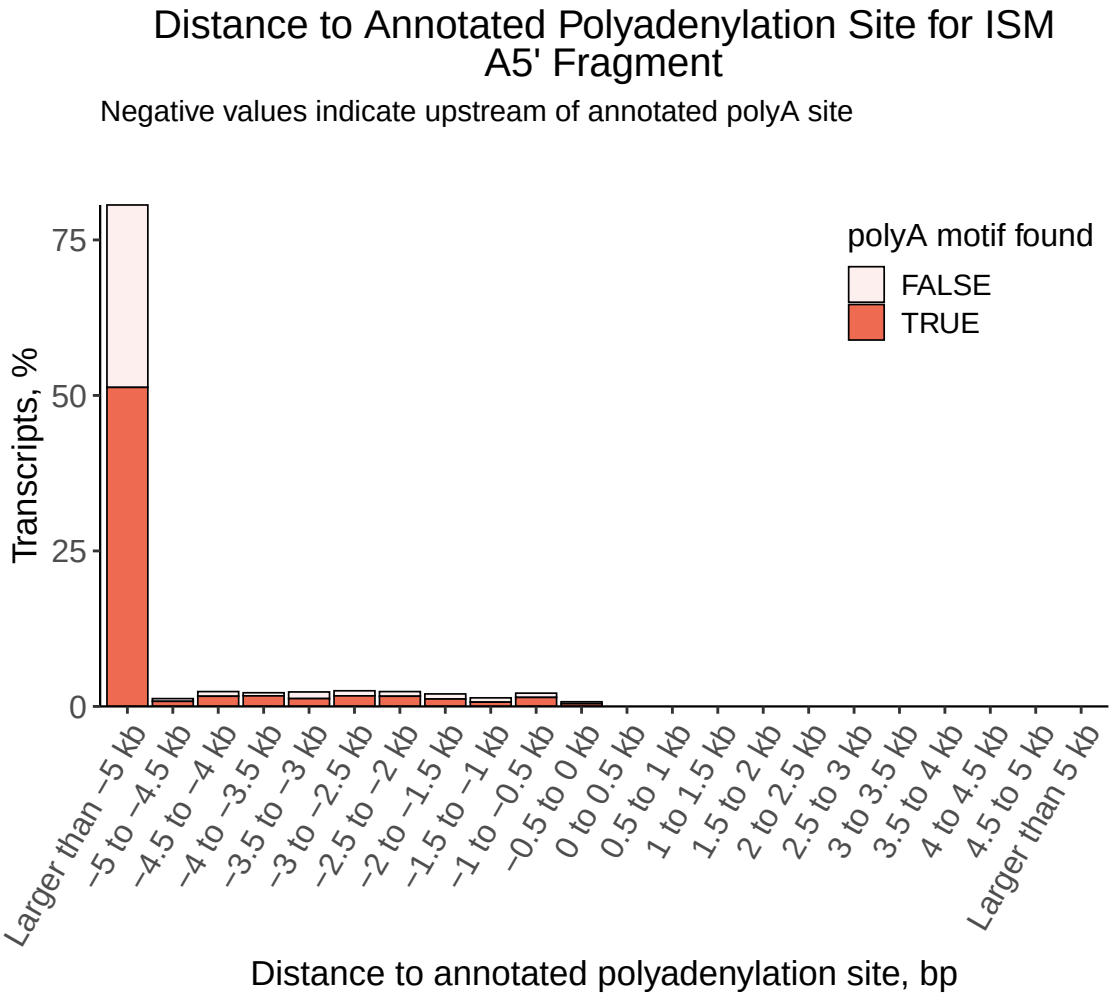

Distance to Annotated Polyadenylation Site for ISM 3' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
0
100
200
300
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 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

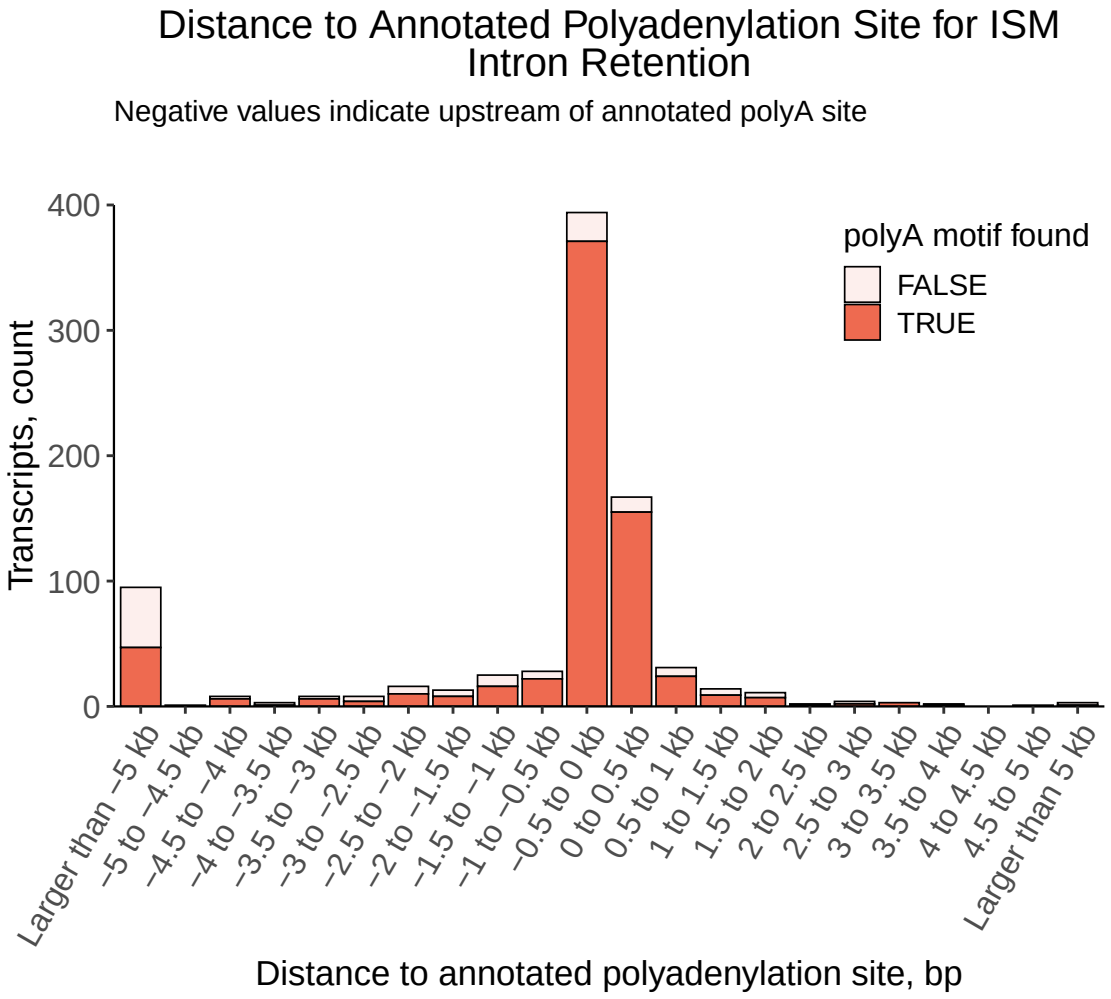

Distance to Annotated Polyadenylation Site for ISM 3' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1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

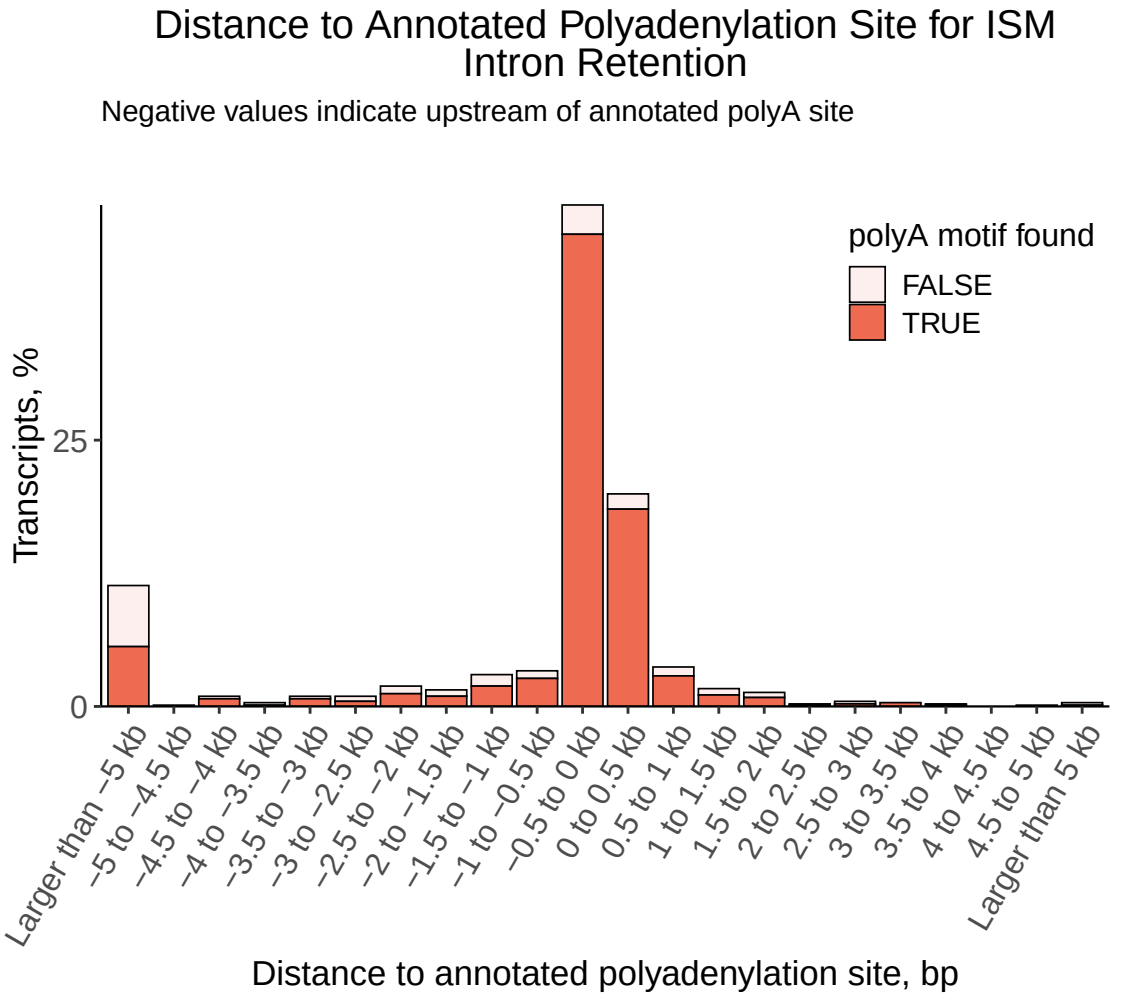

Distance to Annotated Polyadenylation Site for ISM Internal Fragment
Negative values indicate upstream of annotated polyA site
Transcripts, count
100
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 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

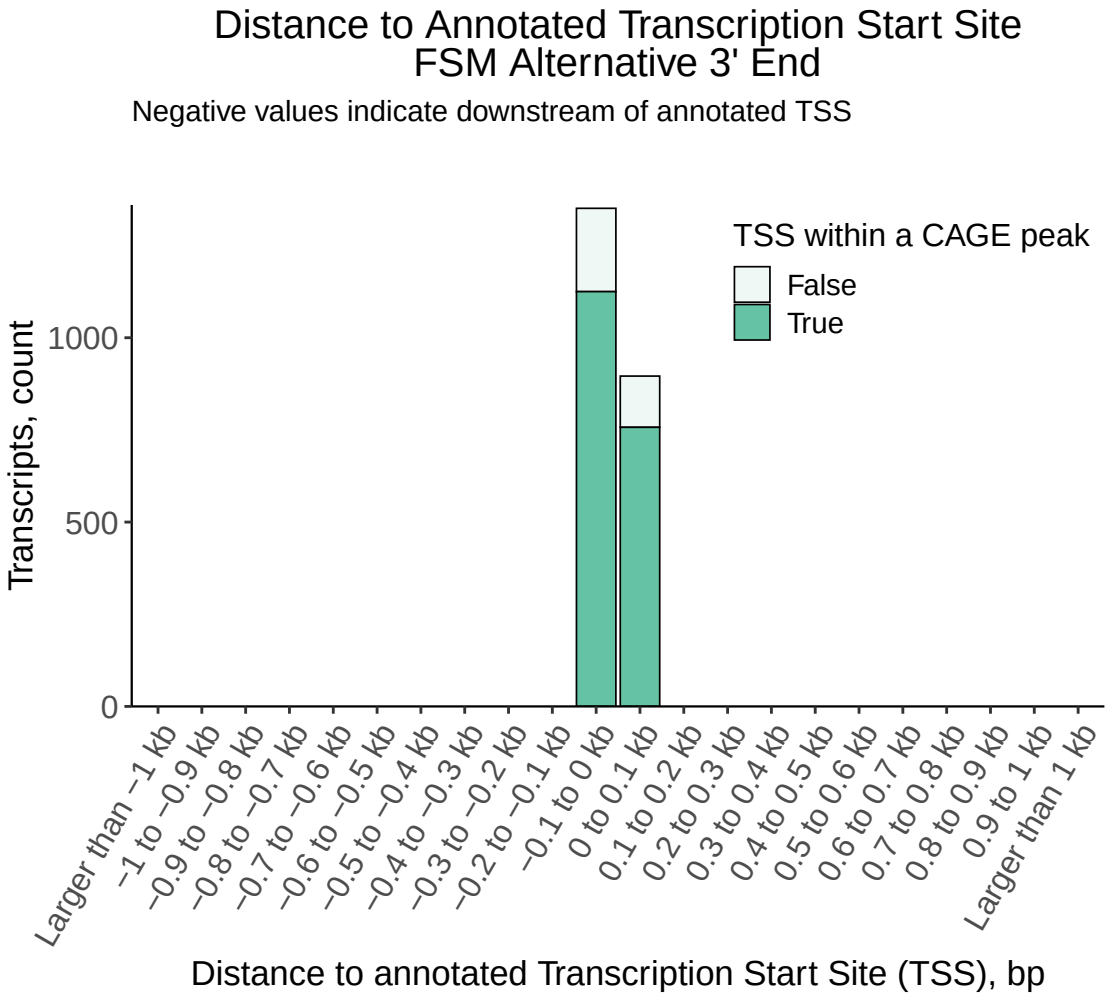

Distance to Annotated Polyadenylation Site for ISM Internal Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to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

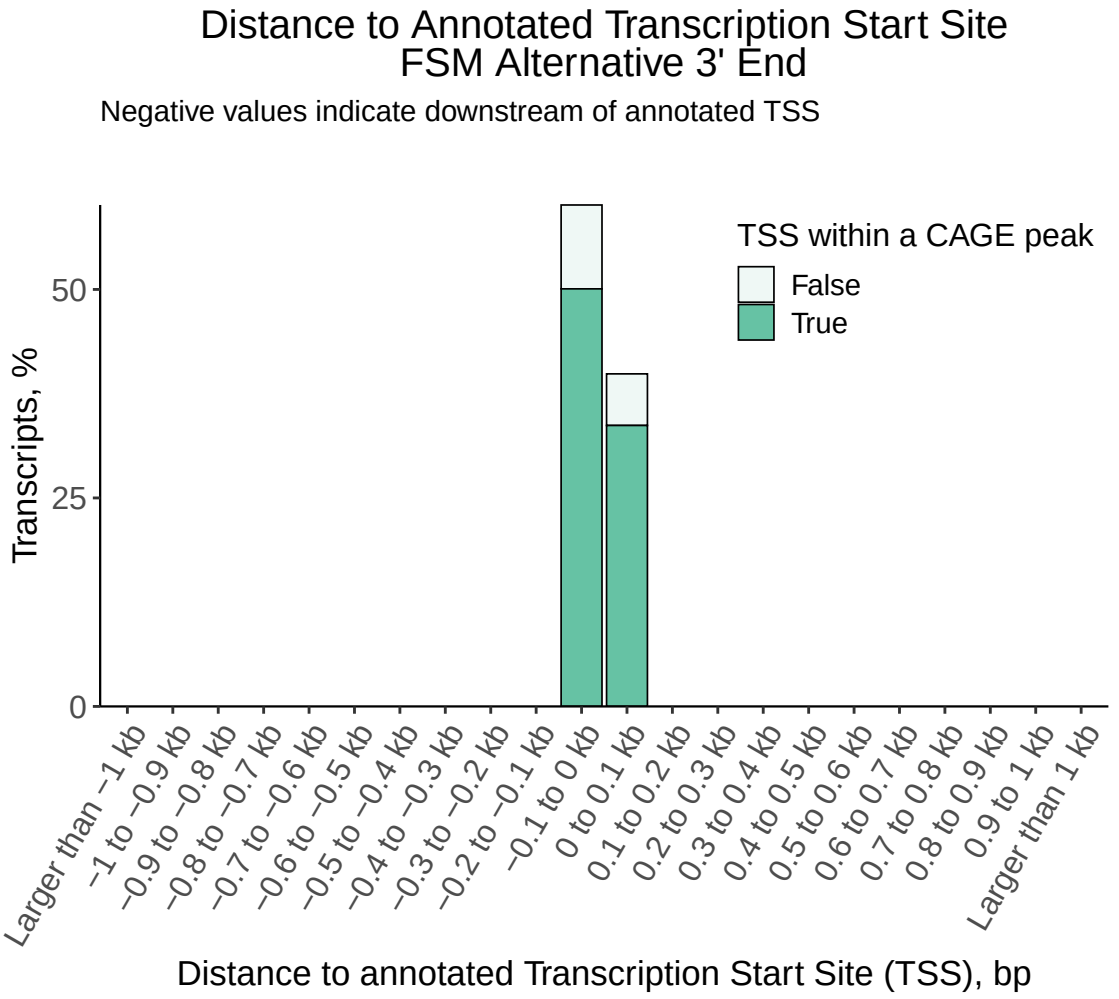

Distance to Annotated Polyadenylation Site for ISM A5' Fragment
Negative values indicate upstream of annotated polyA site
Transcripts, count
1250
1000
750
500
250
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −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

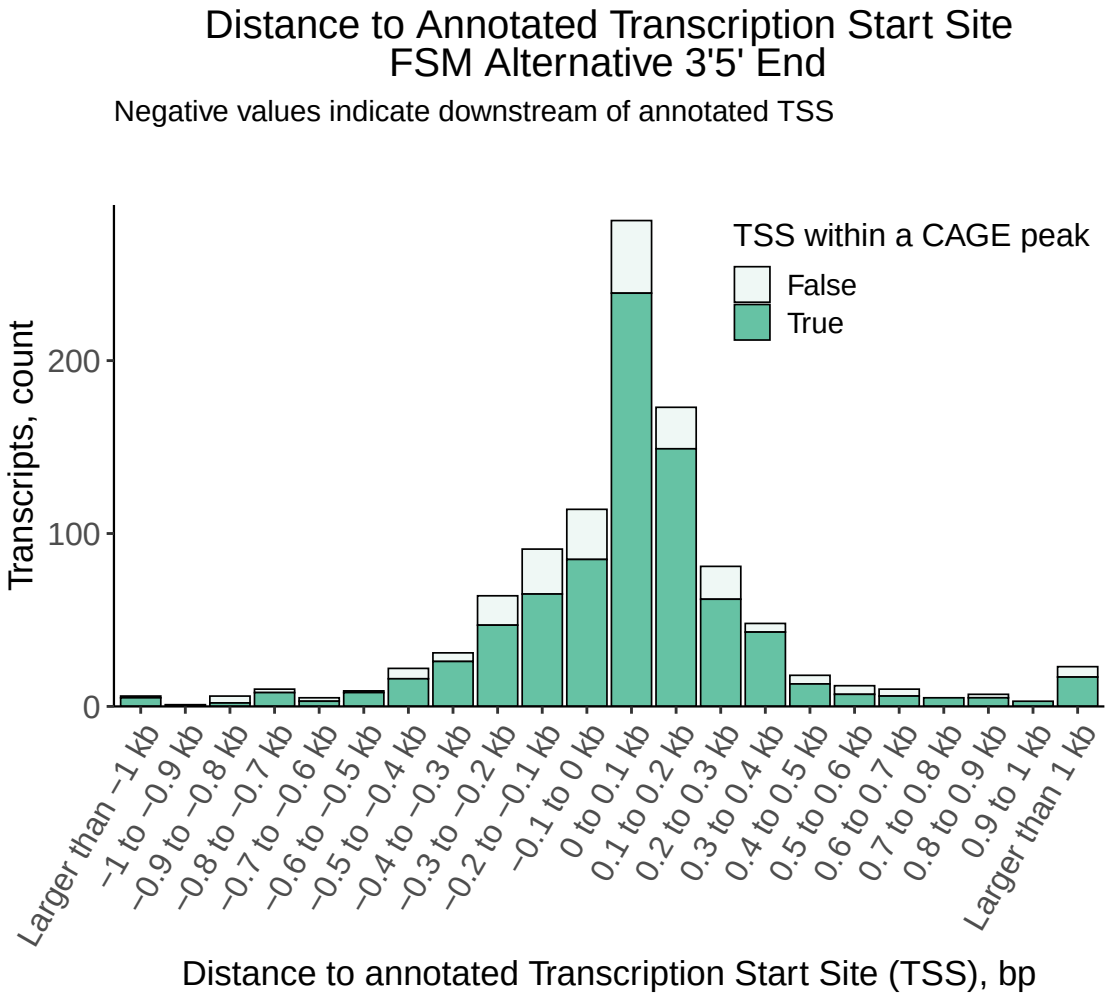

Distance to Annotated Polyadenylation Site for ISM A5' Fragment
Negative values indicate upstream of annotated polyA site
Transcripts, %
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −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

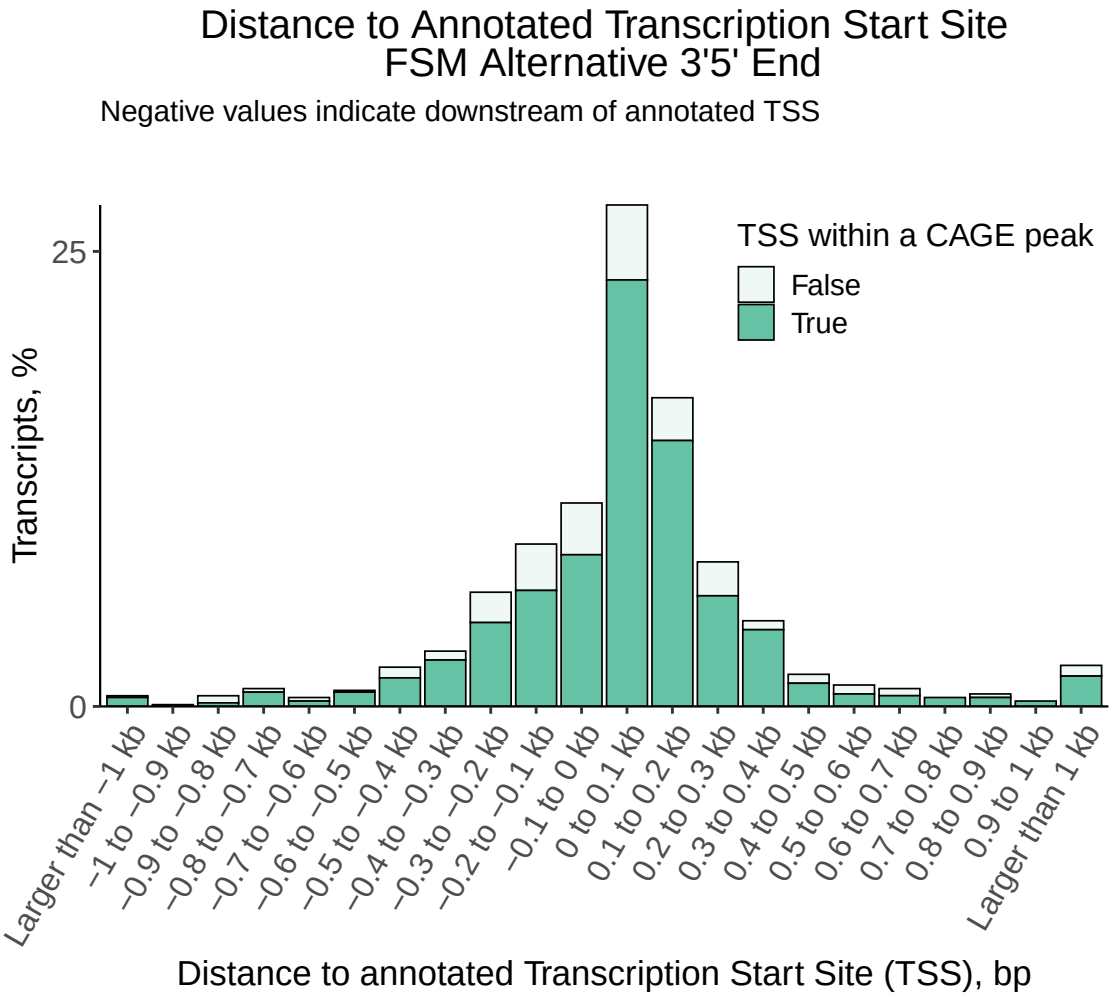

Distance to Annotated Polyadenylation Site for ISM Intron Retention
Negative values indicate upstream of annotated polyA site
Transcripts, count
400
300
200
100
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 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

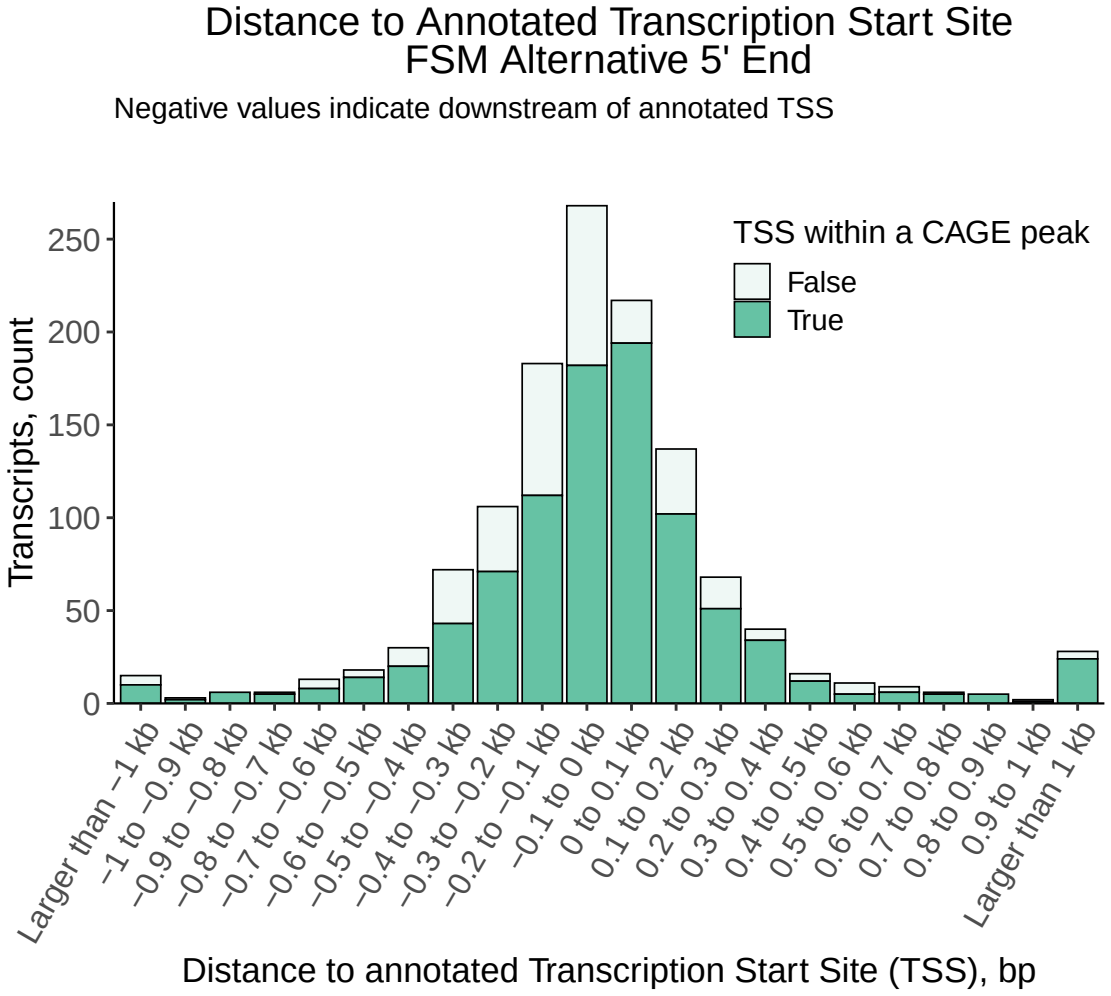

Distance to Annotated Polyadenylation Site for ISM Intron Retention
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1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

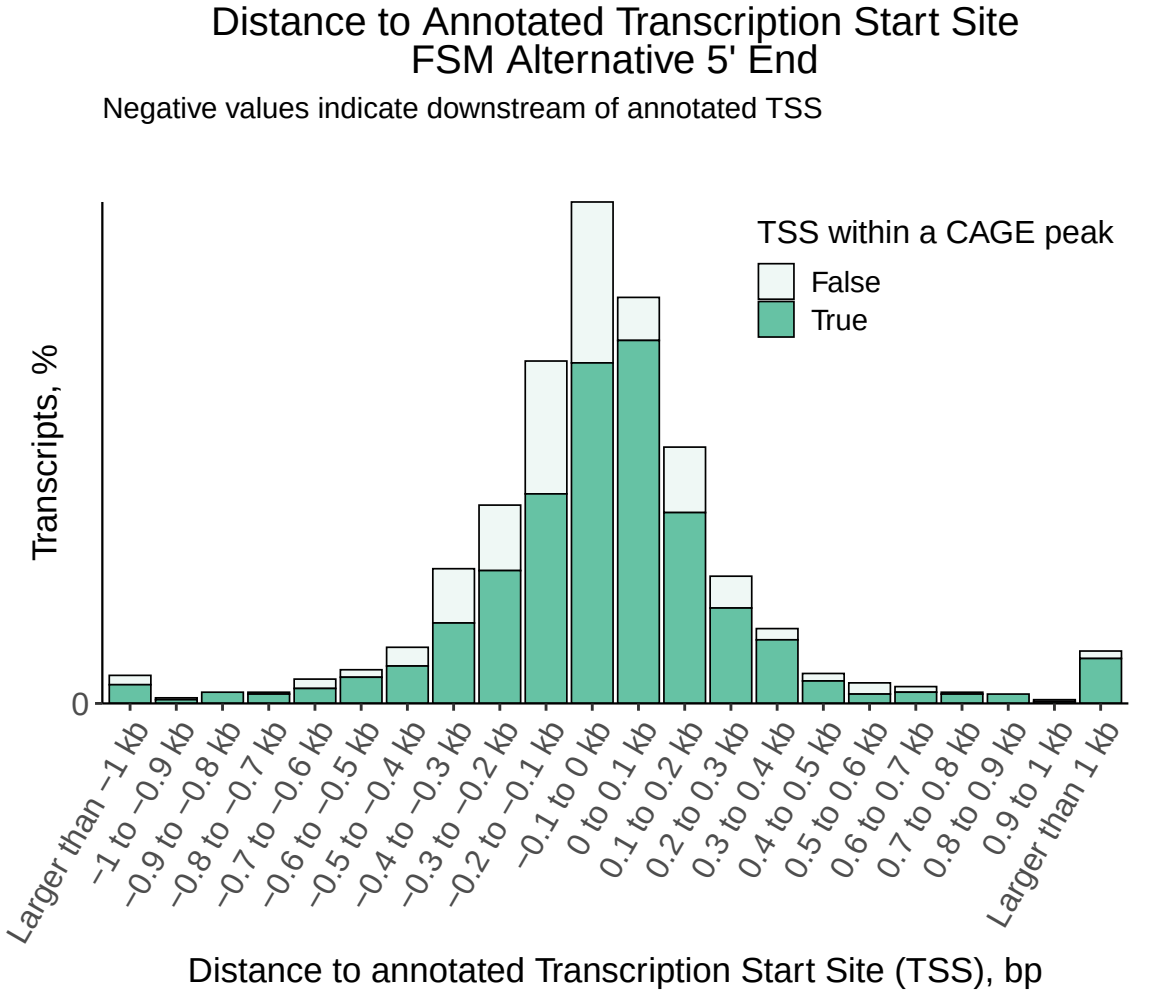

Distance to Annotated Transcription Start Site
FSM Alternative 3' End
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
0
500
1000
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 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

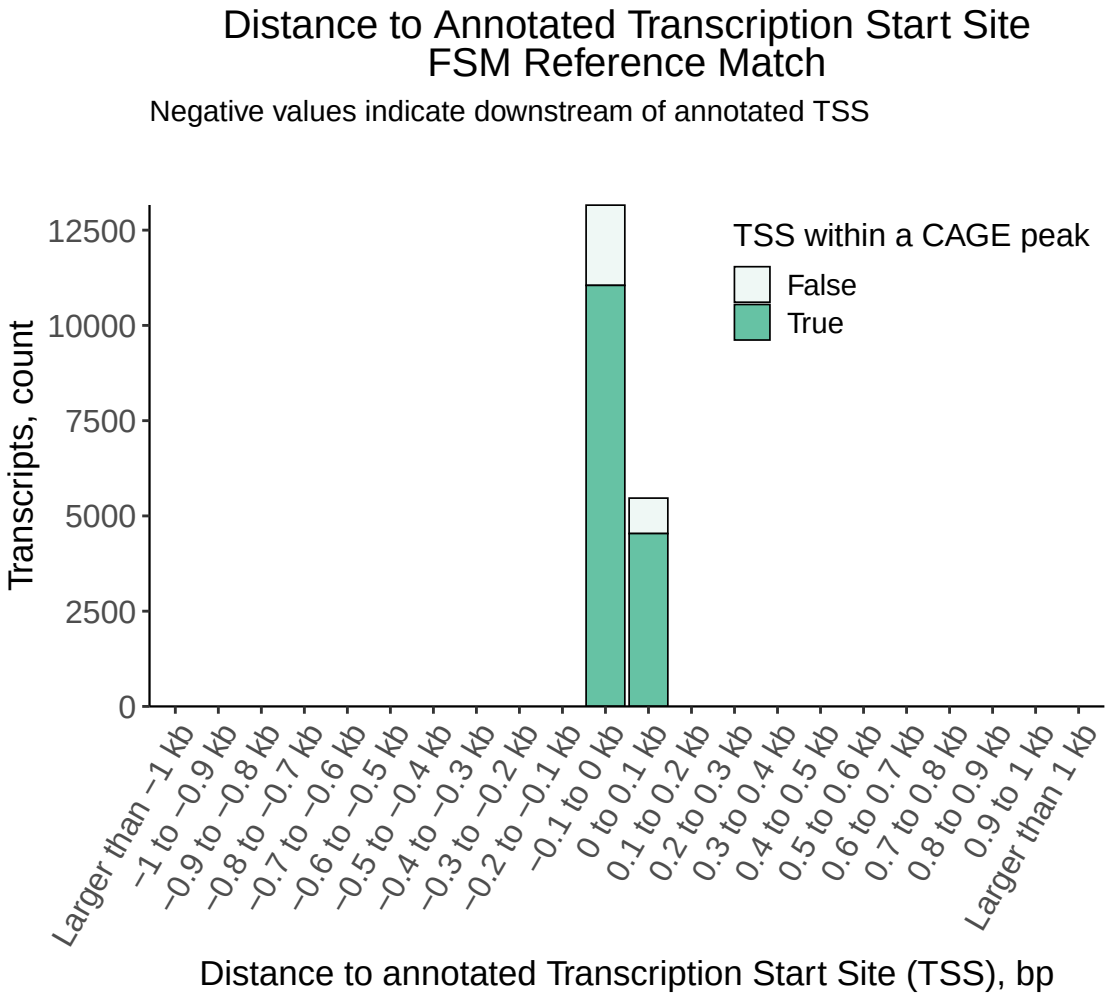

Distance to Annotated Transcription Start Site
FSM Alternative 3' End
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
50
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 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

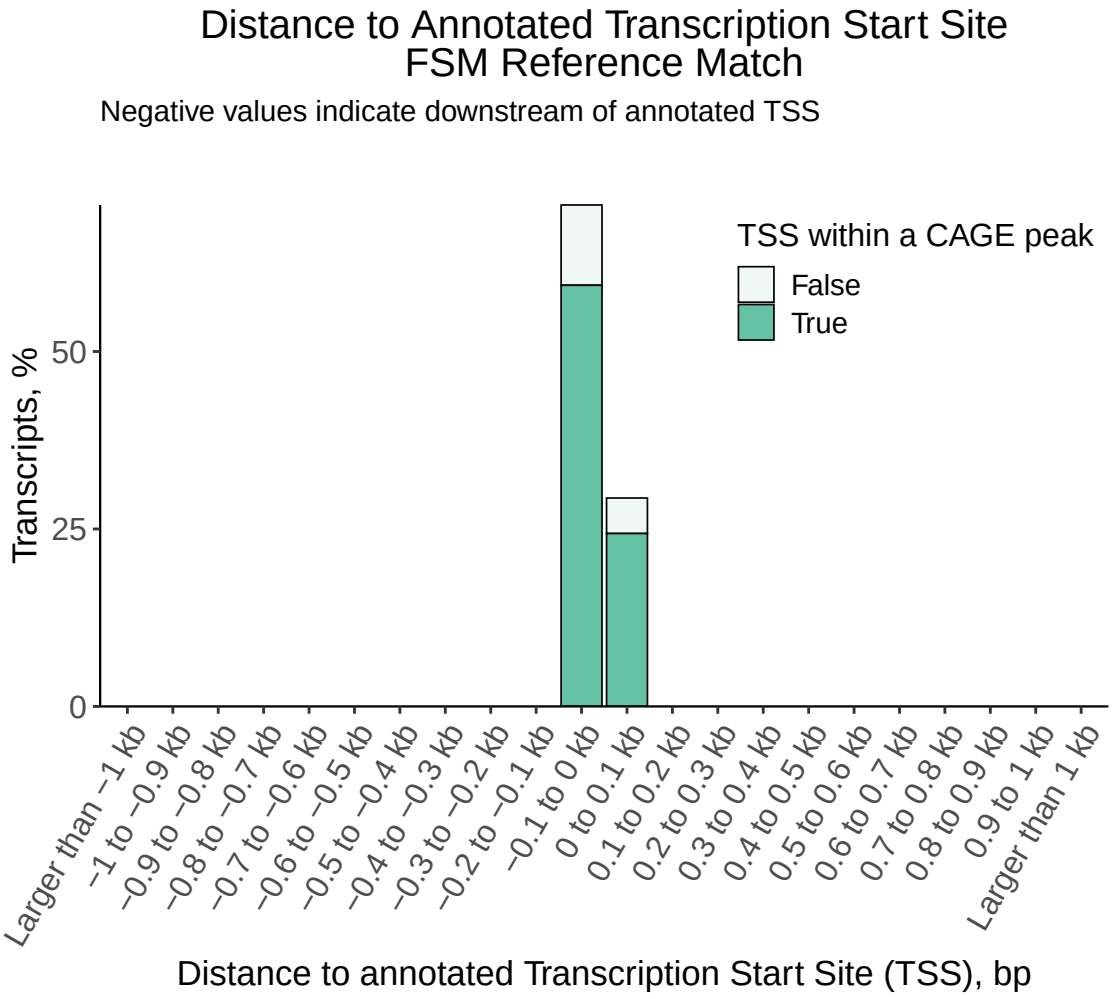

Distance to Annotated Transcription Start Site
FSM Alternative 3'5' End
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
0
100
200
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 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

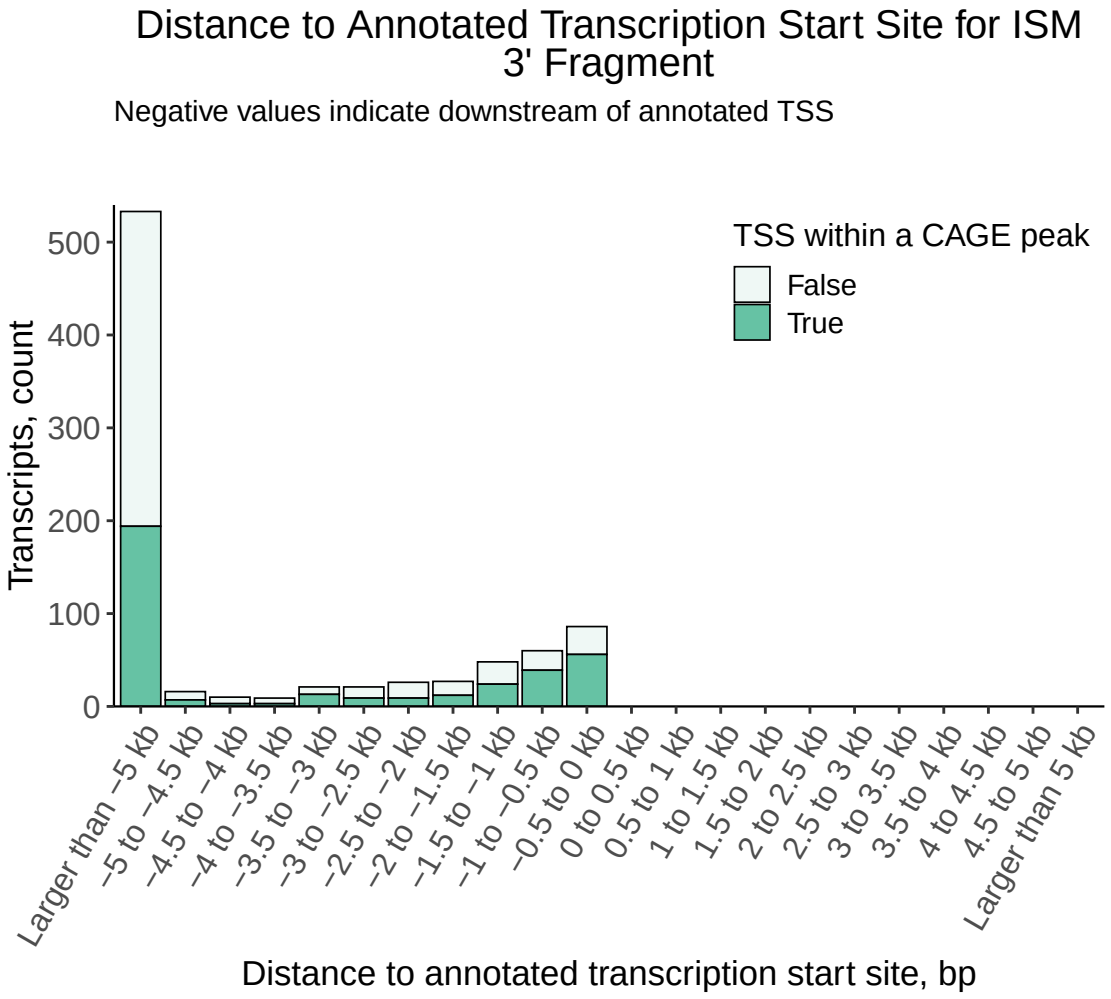

Distance to Annotated Transcription Start Site
FSM Alternative 3'5' End
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 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 Start Site (TSS), bp

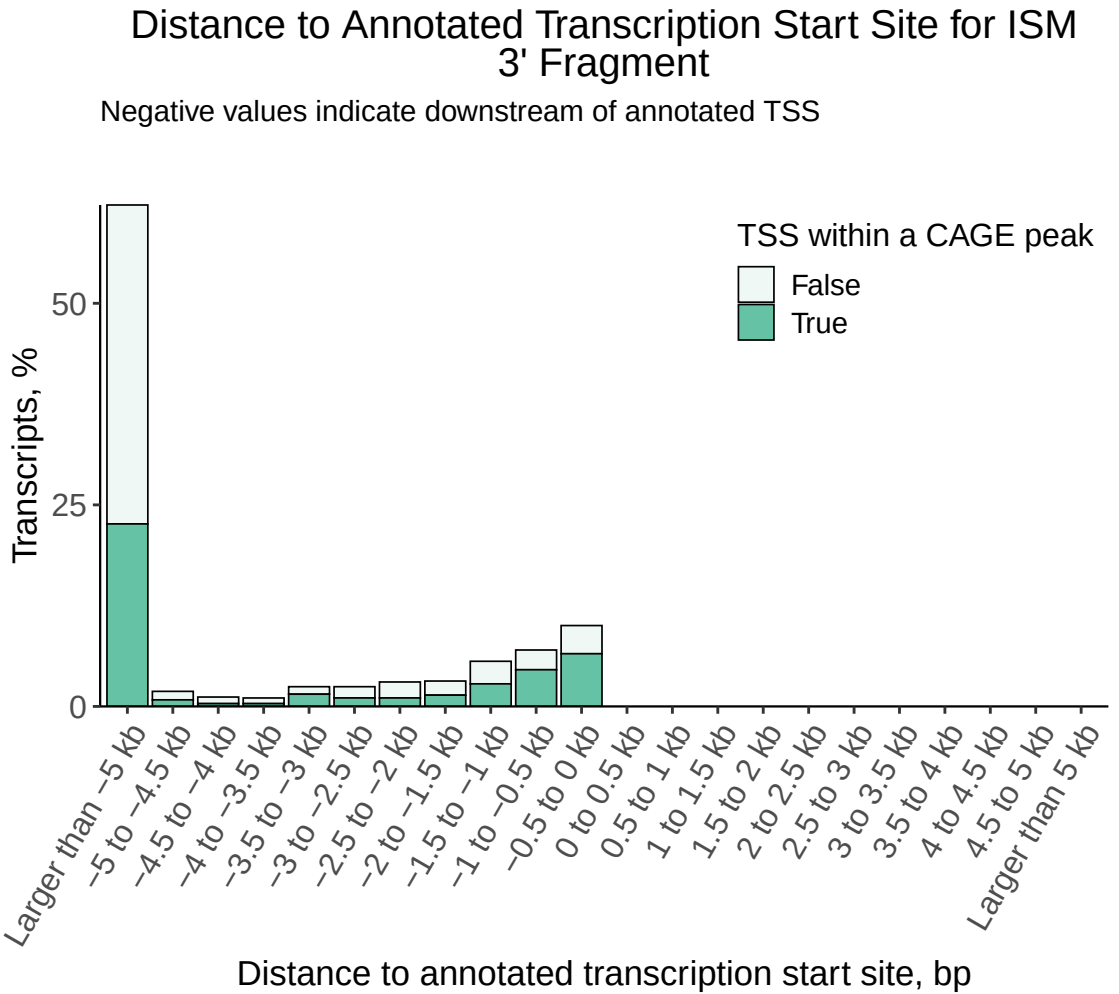

Distance to Annotated Transcription Start Site
FSM Alternative 5' End
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
0
50
100
150
200
250
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −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

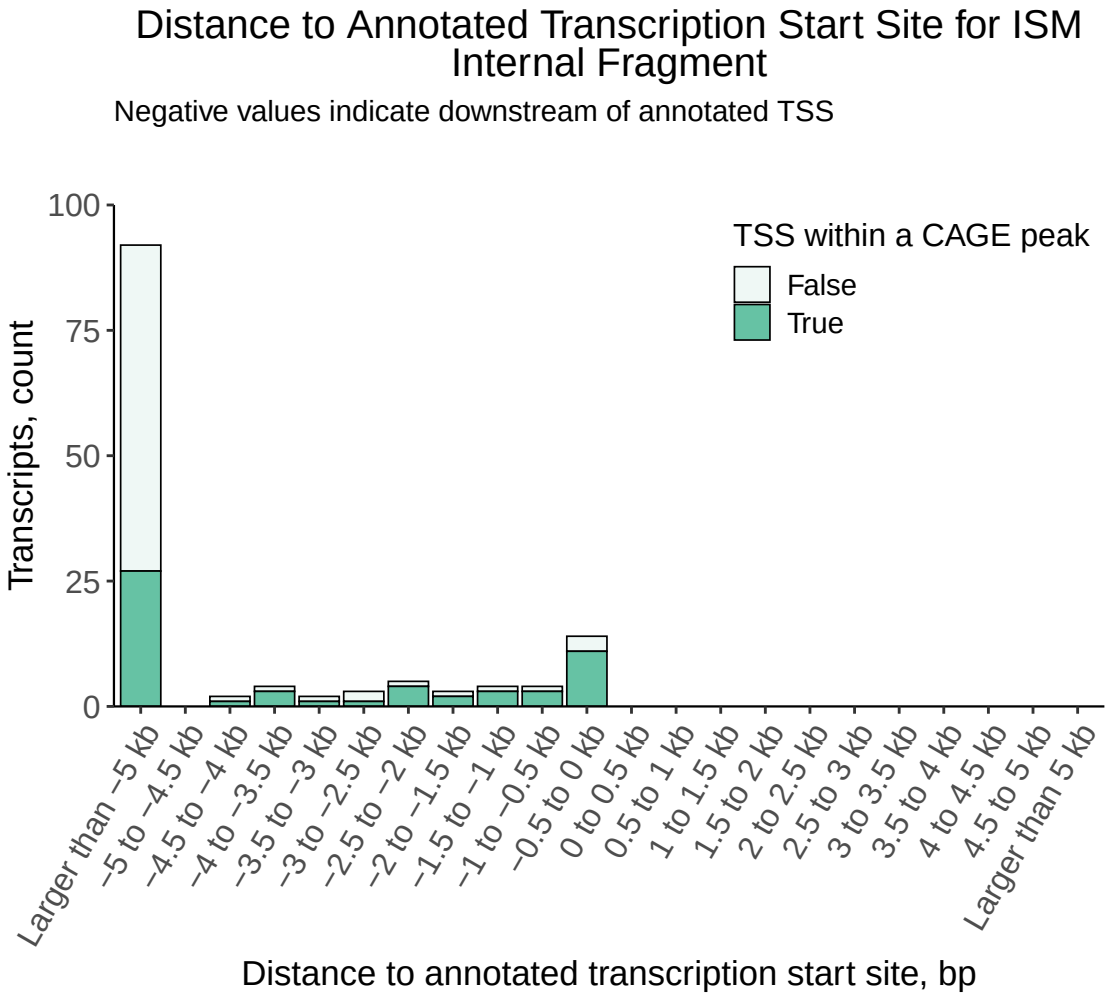

Distance to Annotated Transcription Start Site
FSM Alternative 5' End
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−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

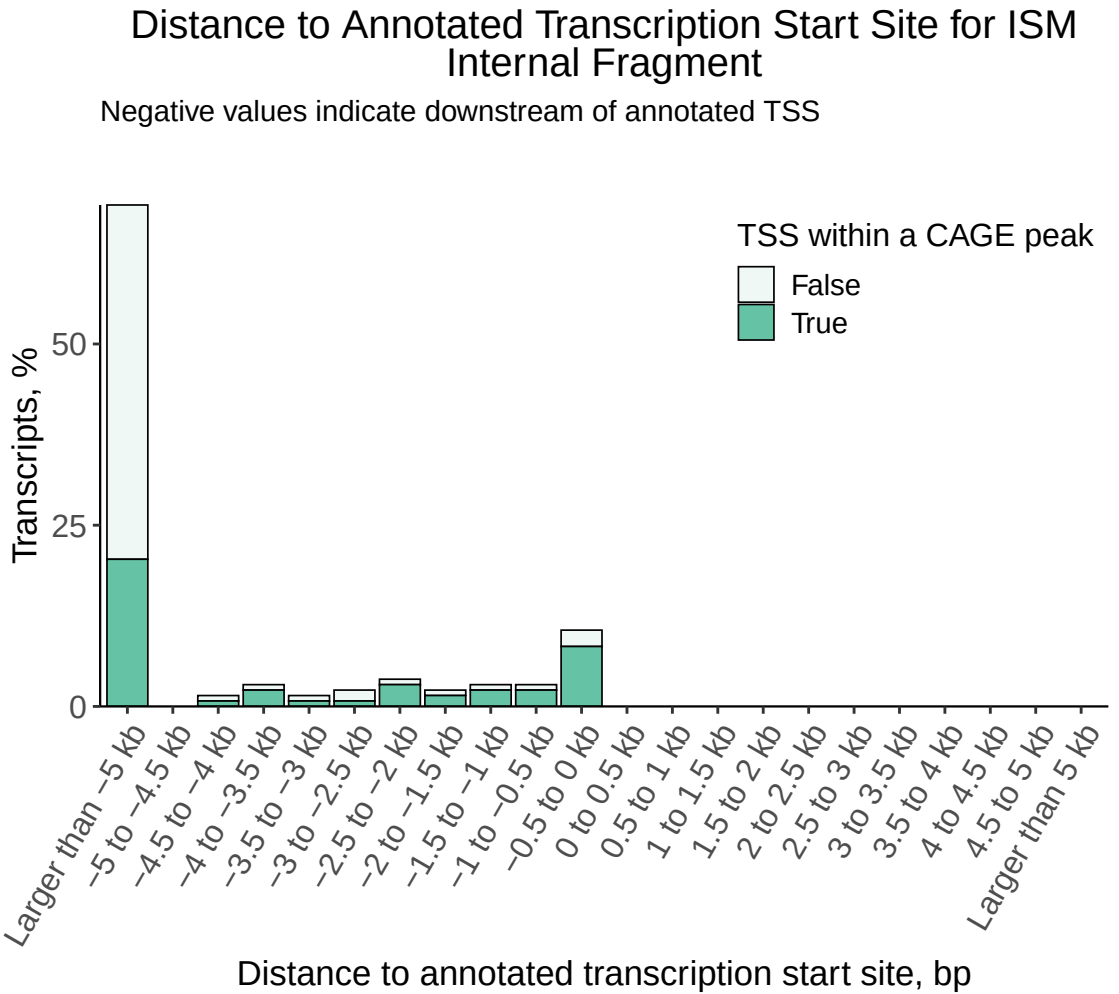

Distance to Annotated Transcription Start Site
FSM Reference Match
Negative values indicate downstream of annotated TSS
Transcripts, count
12500
10000
7500
5000
2500
0
TSS within a CAGE peak
False
True
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 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 Start Site (TSS), bp

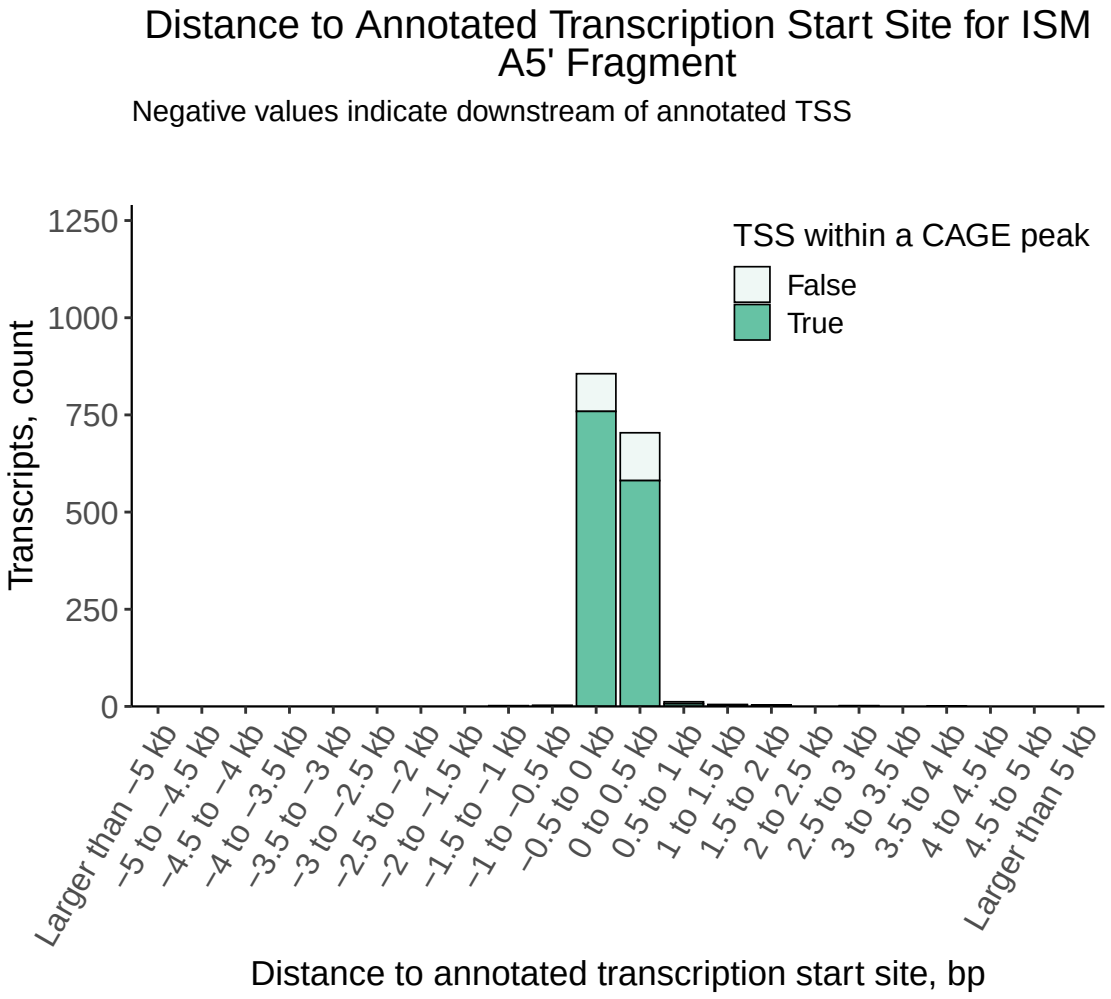

Distance to Annotated Transcription Start Site
FSM Reference Match
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
0
25
50
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 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

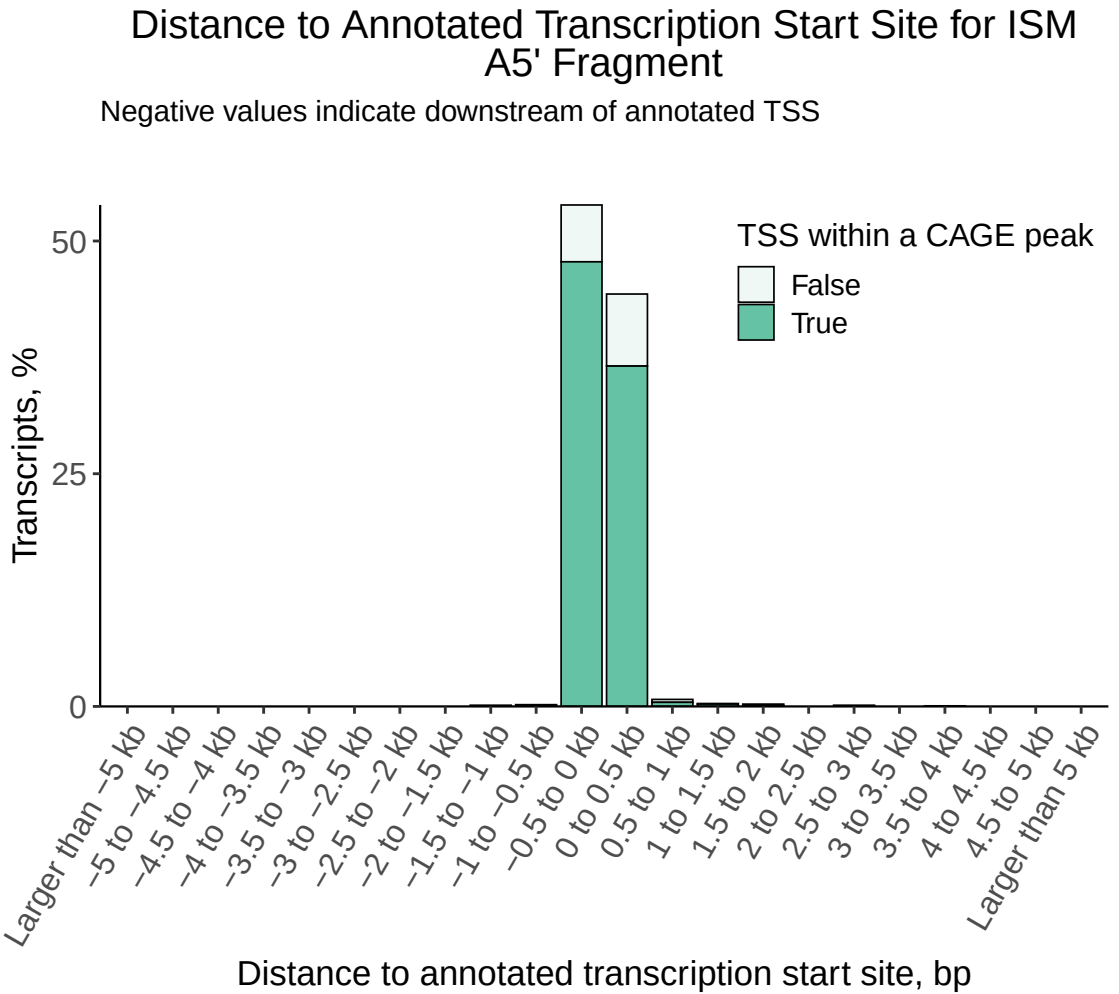

Distance to Annotated Transcription Start Site for ISM 3' Fragment
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
0
100
200
300
400
500
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −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 transcription start site, bp

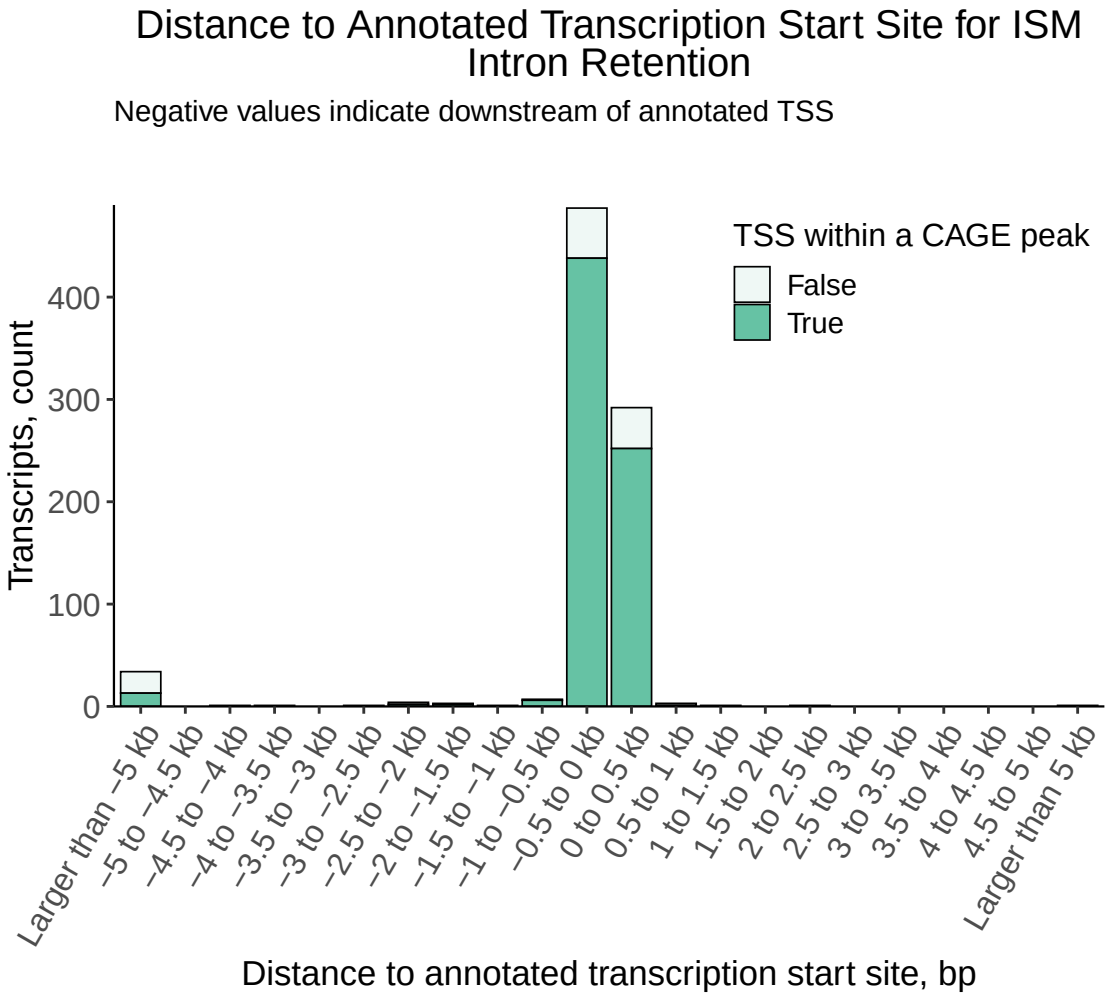

Distance to Annotated Transcription Start Site for ISM 3' Fragment
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, %
50
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −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 transcription start site, bp

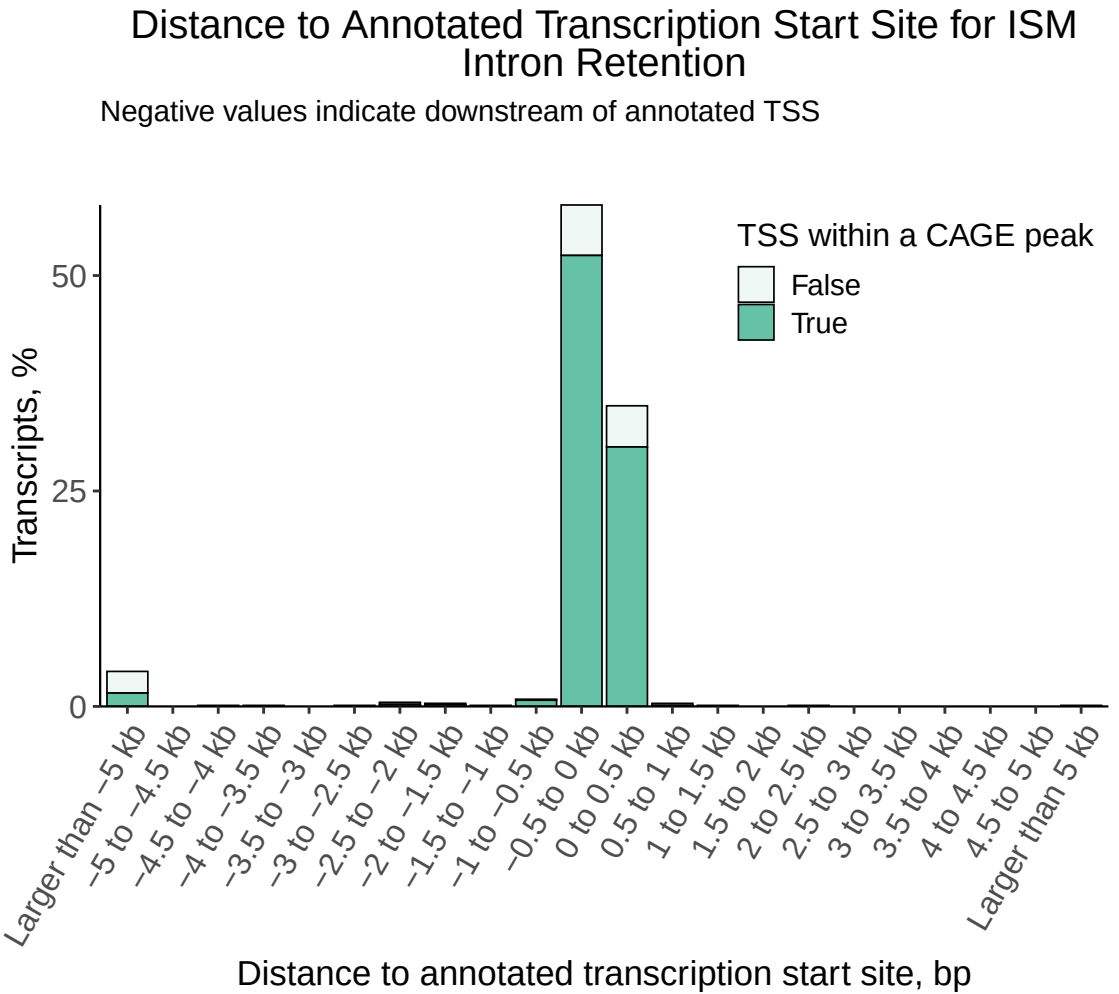

Distance to Annotated Transcription Start Site for ISM Internal Fragment
Negative values indicate downstream of annotated TSS
Transcripts, count
100
75
50
25
0
TSS within a CAGE peak
False
True
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −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 transcription start site, bp


Distance to Annotated Transcription Start Site for ISM Internal Fragment
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
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 transcription start site, bp

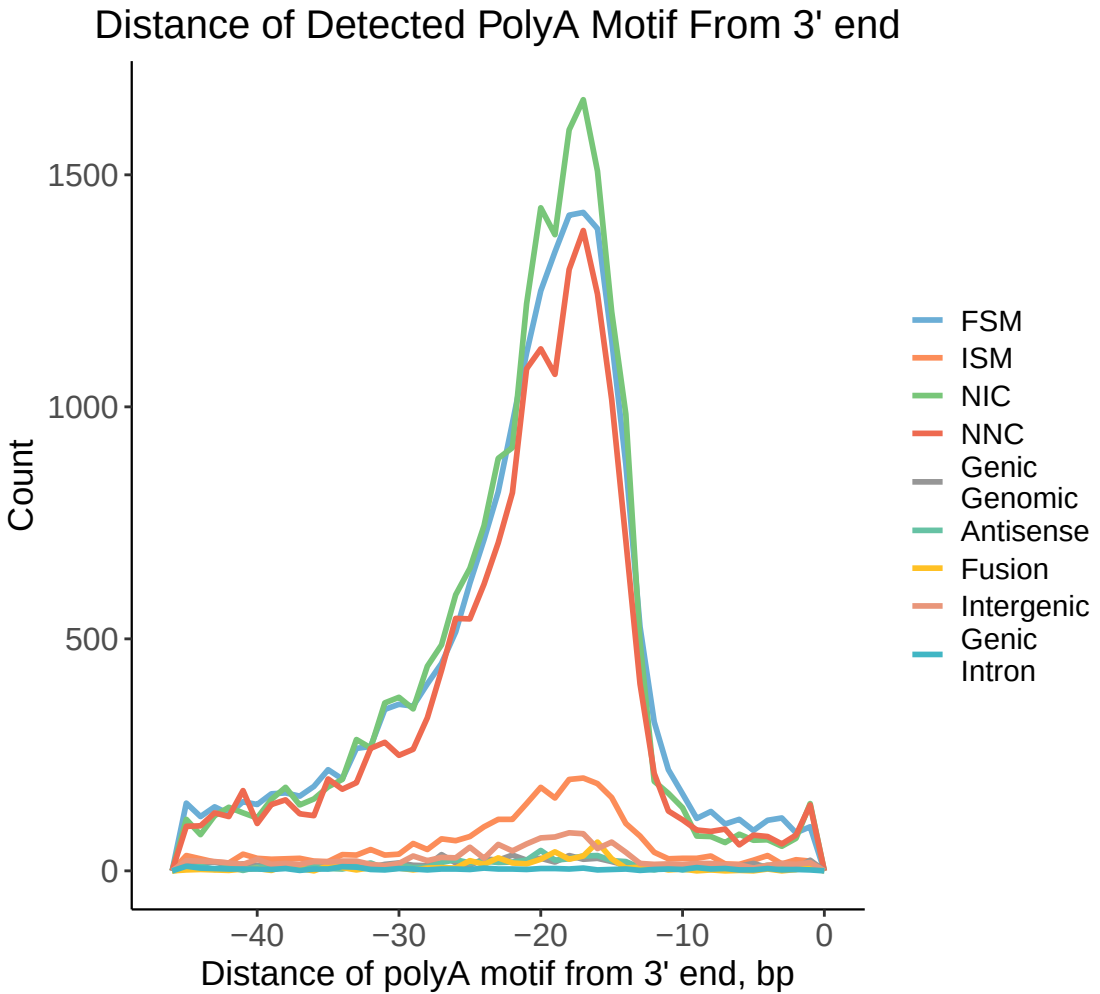

Distance to Annotated Transcription Start Site for ISM A5' Fragment
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
1250
1000
750
500
250
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 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 transcription start site, bp


Distance to Annotated Transcription Start Site for ISM A5' Fragment
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
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 transcription start site, bp


Distance to Annotated Transcription Start Site for ISM Intron Retention
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
False
True
Transcripts, count
0
100
200
300
400
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 transcription start site, bp


Distance to Annotated Transcription Start Site for ISM Intron Retention
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
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 transcription start site, bp

PolyA Distance Analysis

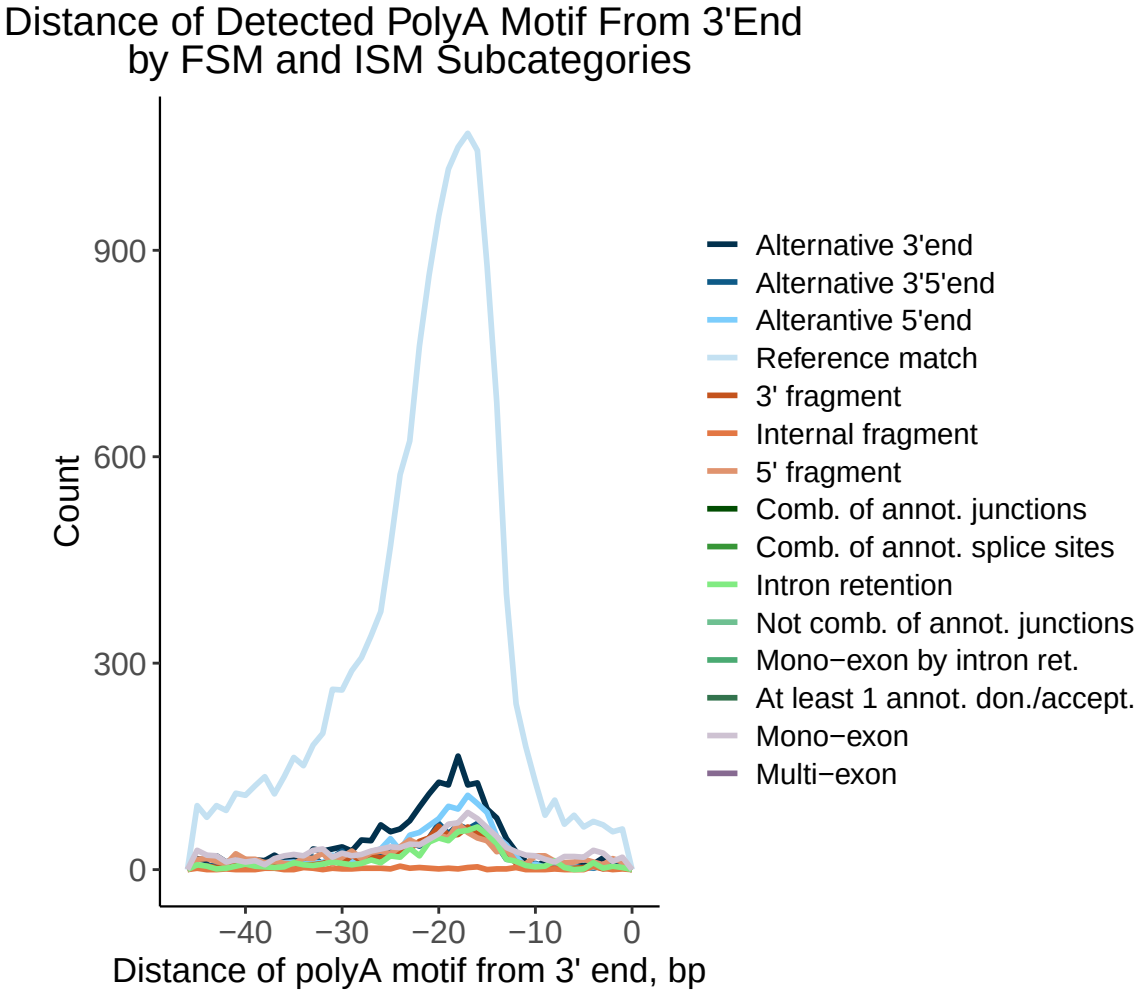

Distance of Detected PolyA Motif From 3' end
Count
1500
1000
500
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

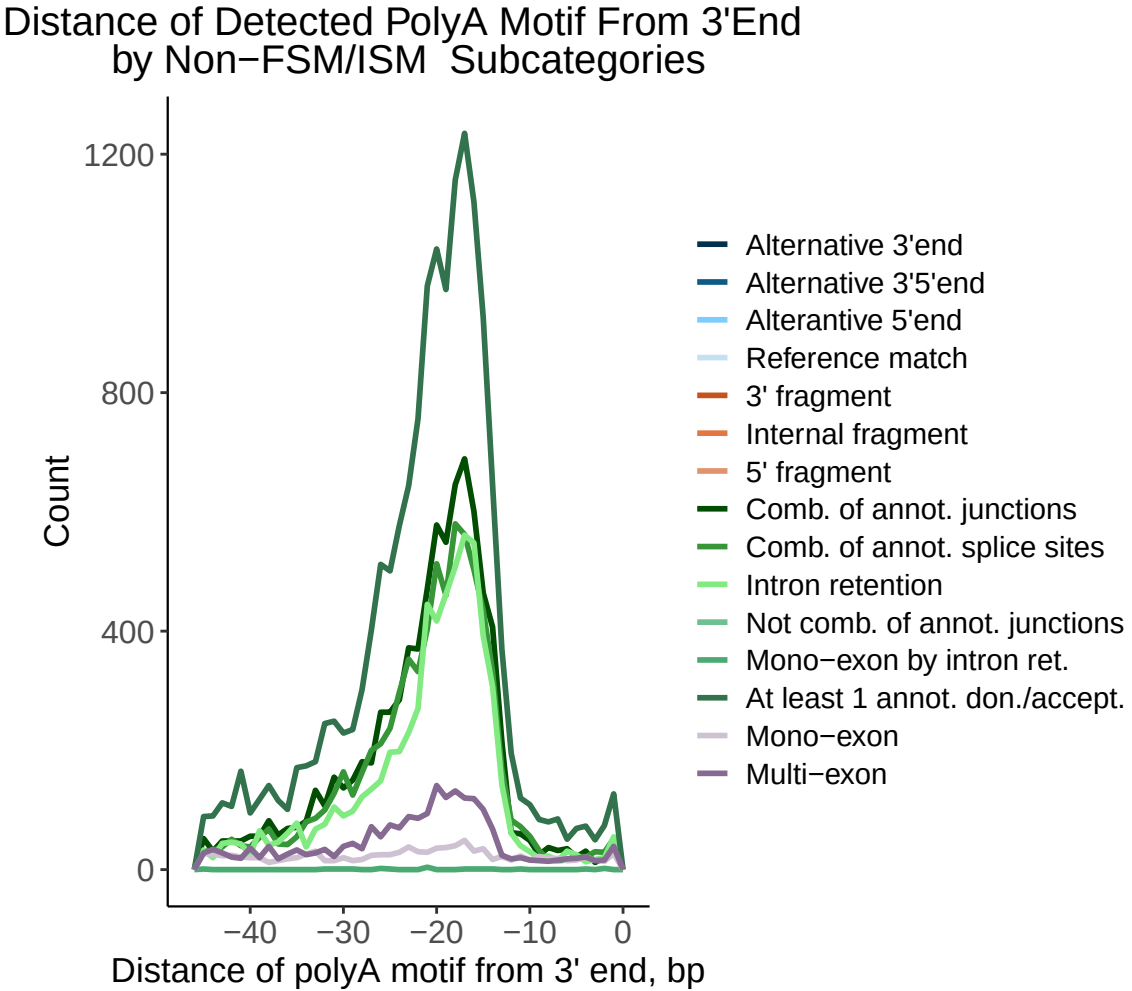
Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	25235	20091	80
ISM	3781	2791	74
NIC	22995	20704	90
NNC	19878	17376	87
Genic Genomic	1041	673	65
Antisense	856	552	64
Fusion	549	443	81
Intergenic	2131	1271	60
Genic Intron	551	183	33

Frequency of PolyA Motifs

Motif	Count	%
AATAAA	40859	63.8
ATTAAA	9814	15.3
AAAAAG	2077	3.2
TATAAA	1999	3.1
AGTAAA	1846	2.9
AAGAAA	1135	1.8
CATAAA	1035	1.6
TTTAAA	934	1.5
AATACA	901	1.4
AATATA	786	1.2
GATAAA	740	1.2
AAAACA	607	0.9
AATGAA	475	0.7
AATAGA	365	0.6
ACTAAA	357	0.6
GGGGCT	154	0.2


Distance of Detected PolyA Motif From 3'End
by FSM and ISM Subcategories
Count
900
600
300
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
1200
800
400
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

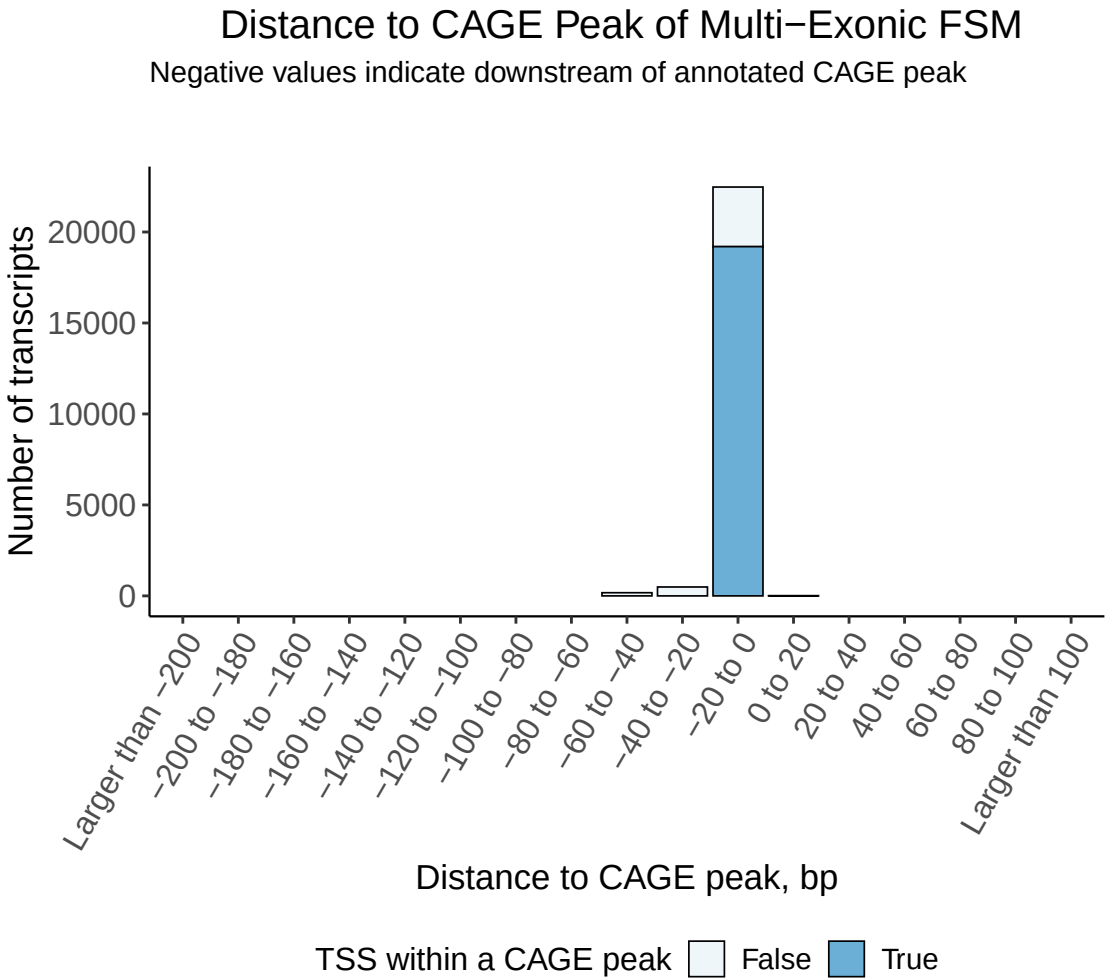
Number of polyA Motifs Detected

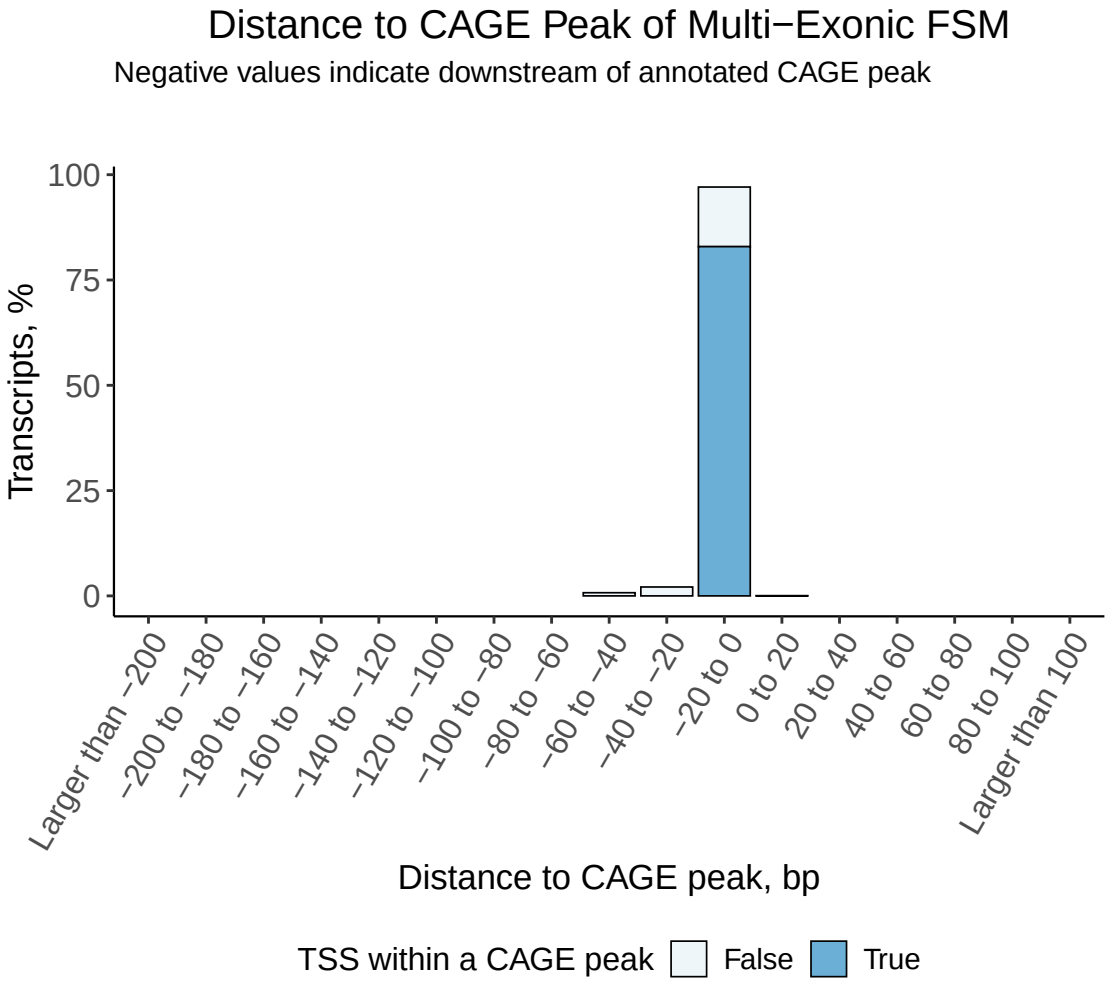
Subcategory	Count	polyA Detected	%
Alternative 3'end	2247	1860	83
Alternative 3'5'end	1020	886	87
Alterantive 5'end	1259	1151	91
Reference match	18624	15162	81
3' fragment	857	756	88
Internal fragment	133	60	45
5' fragment	1589	1014	64
Comb. of annot. junctions	9177	8349	91
Comb. of annot. splice sites	8068	7370	91
Intron retention	8125	7104	87
Mono−exon by intron ret.	41	20	49
At least 1 annot. don./accept.	18119	15875	88
Mono−exon	4897	2354	48
Multi−exon	2861	2123	74

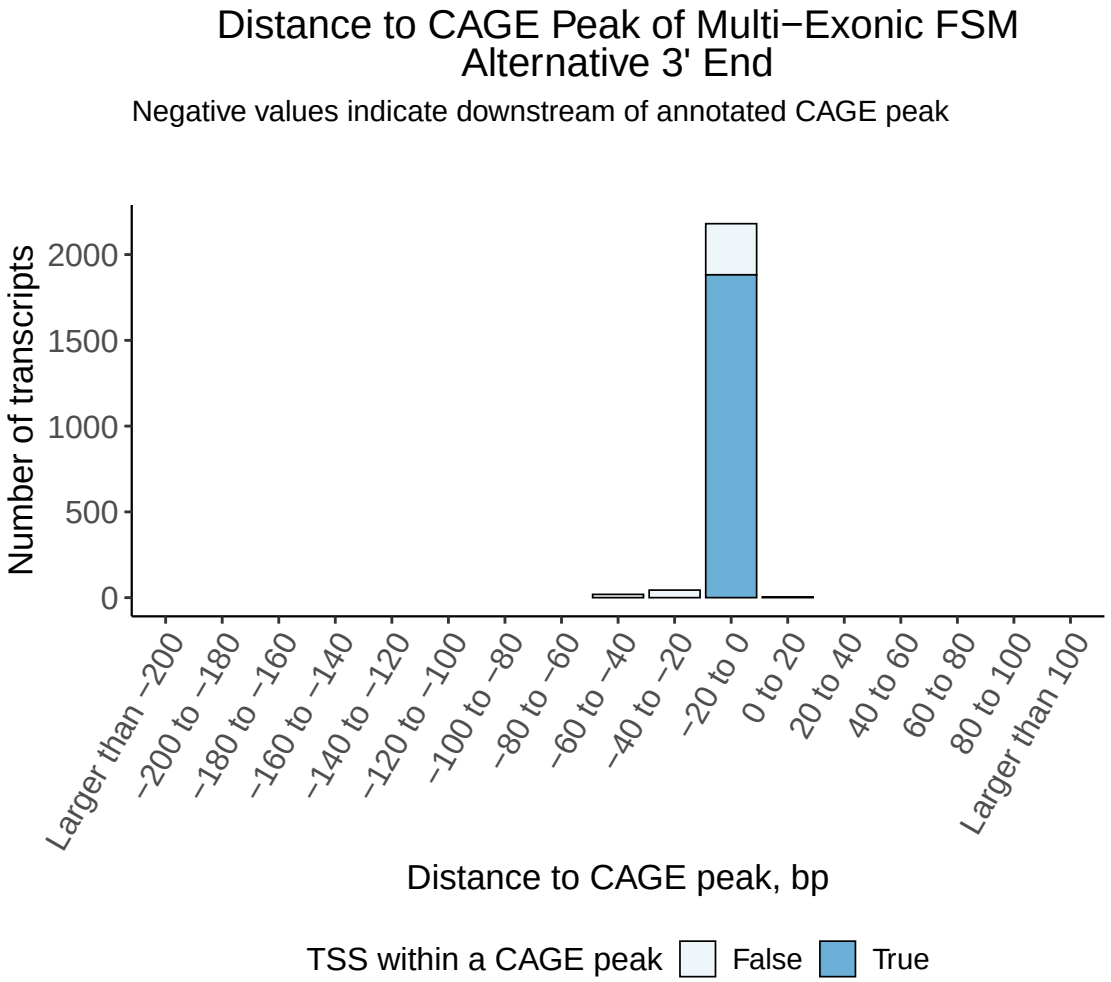
Frequency of PolyA Motifs

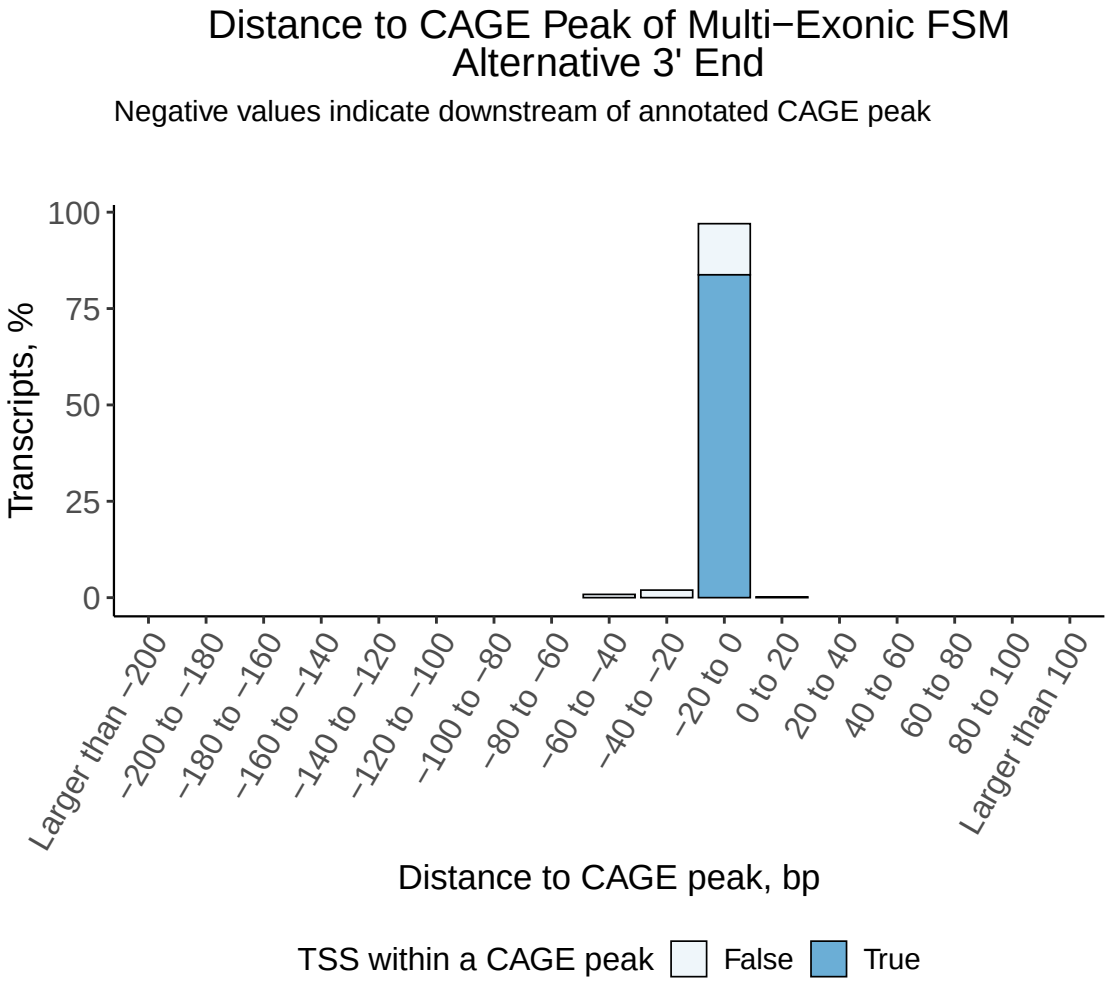
Motif	Count	%
AATAAA	40859	63.8
ATTAAA	9814	15.3
AAAAAG	2077	3.2
TATAAA	1999	3.1
AGTAAA	1846	2.9
AAGAAA	1135	1.8
CATAAA	1035	1.6
TTTAAA	934	1.5
AATACA	901	1.4
AATATA	786	1.2
GATAAA	740	1.2
AAAACA	607	0.9
AATGAA	475	0.7
AATAGA	365	0.6
ACTAAA	357	0.6
GGGGCT	154	0.2

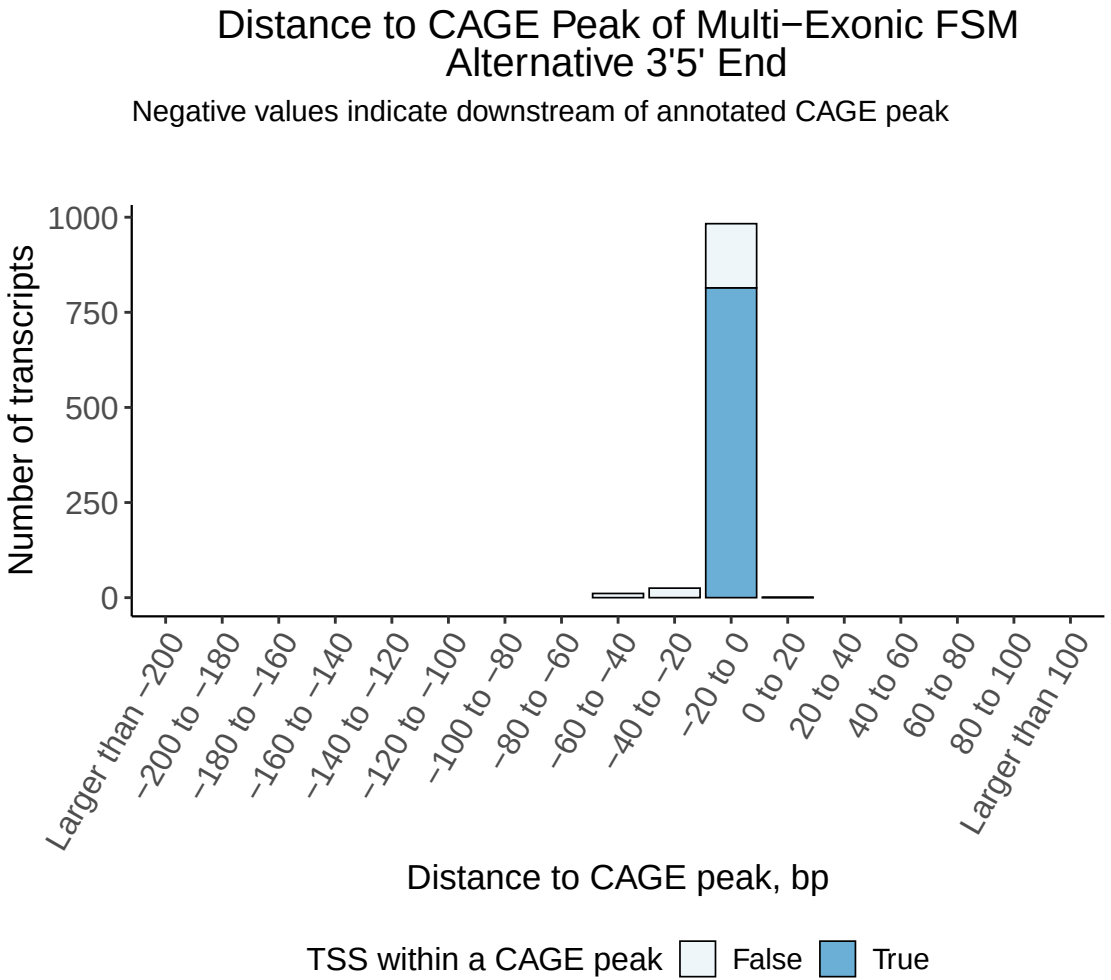
CAGE Distances Analysis

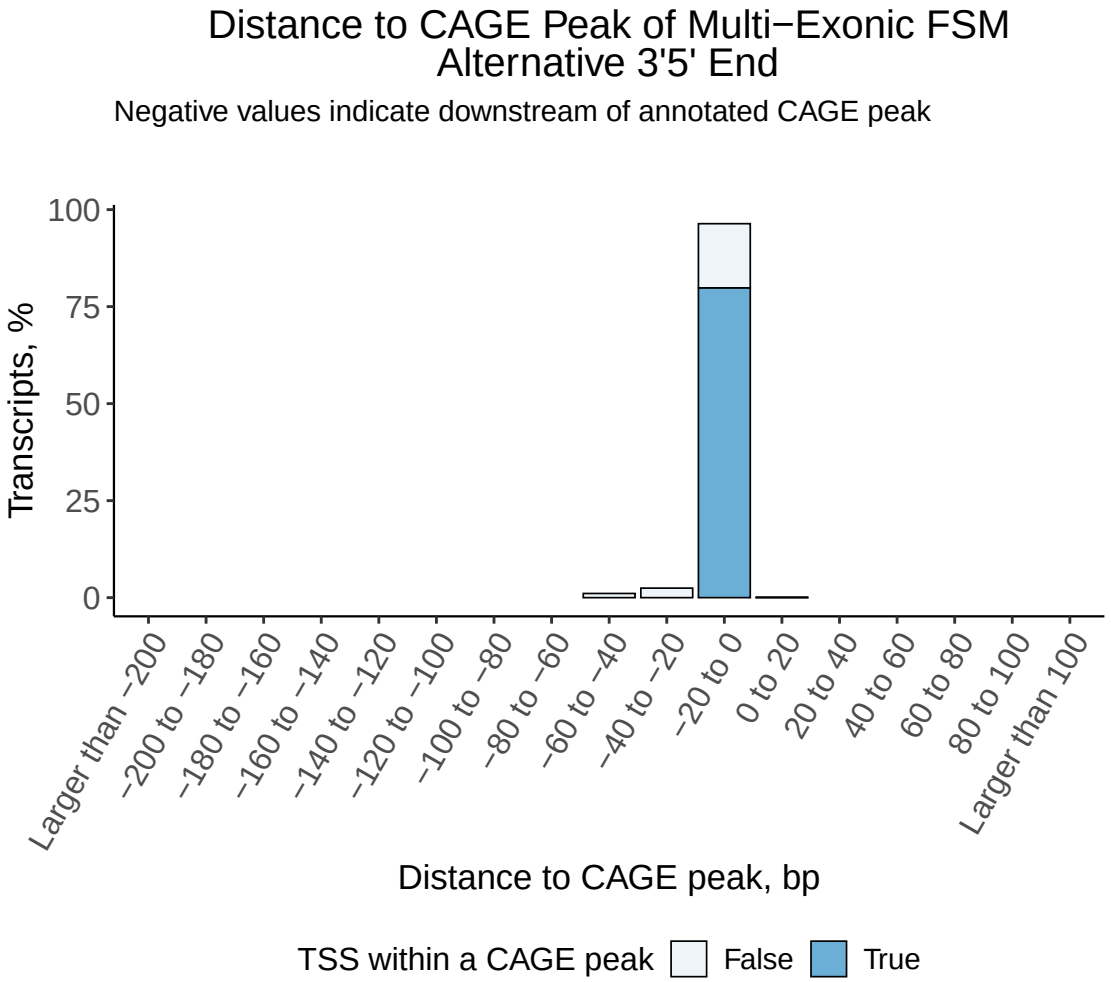

Distance to CAGE Peak of Multi−Exonic FSM
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
20000
15000
10000
5000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

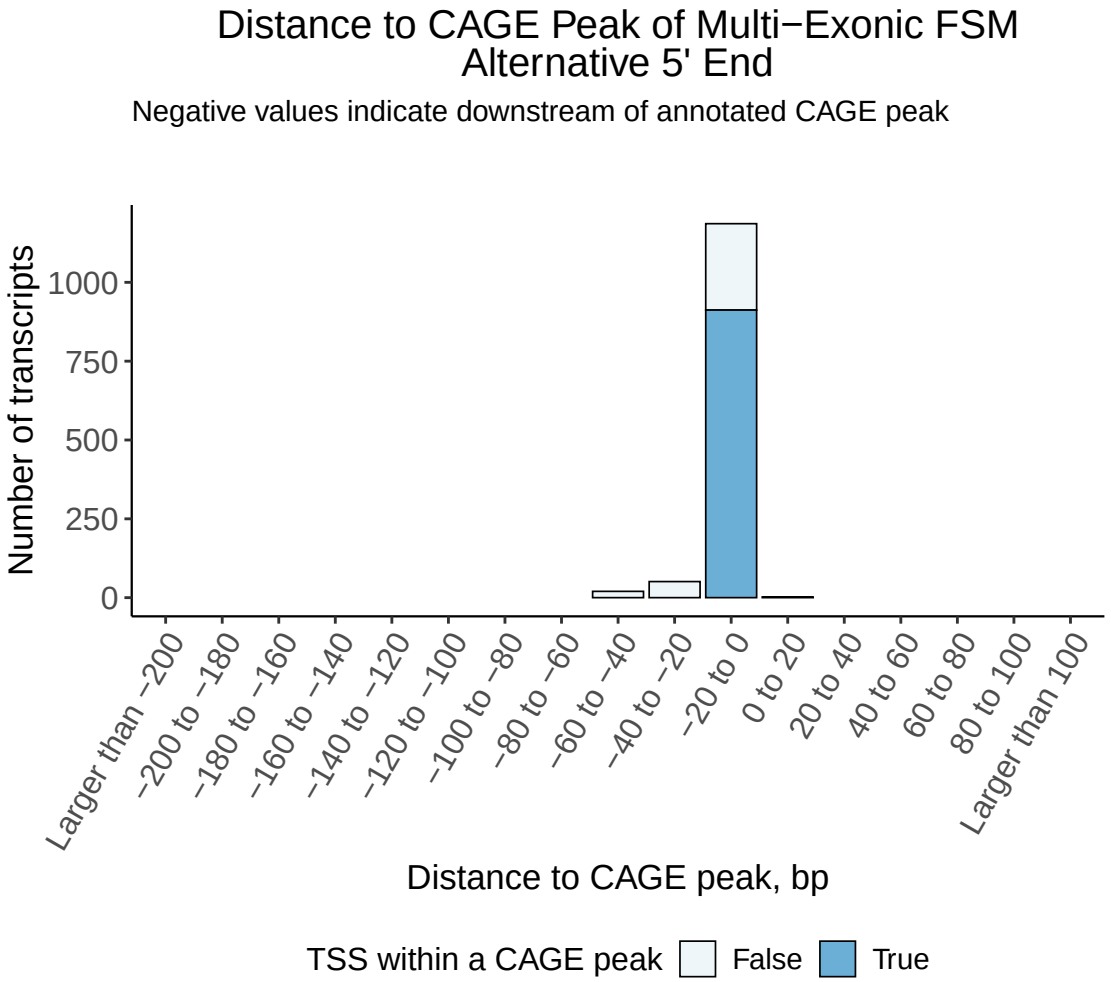

Distance to CAGE Peak of Multi−Exonic FSM
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

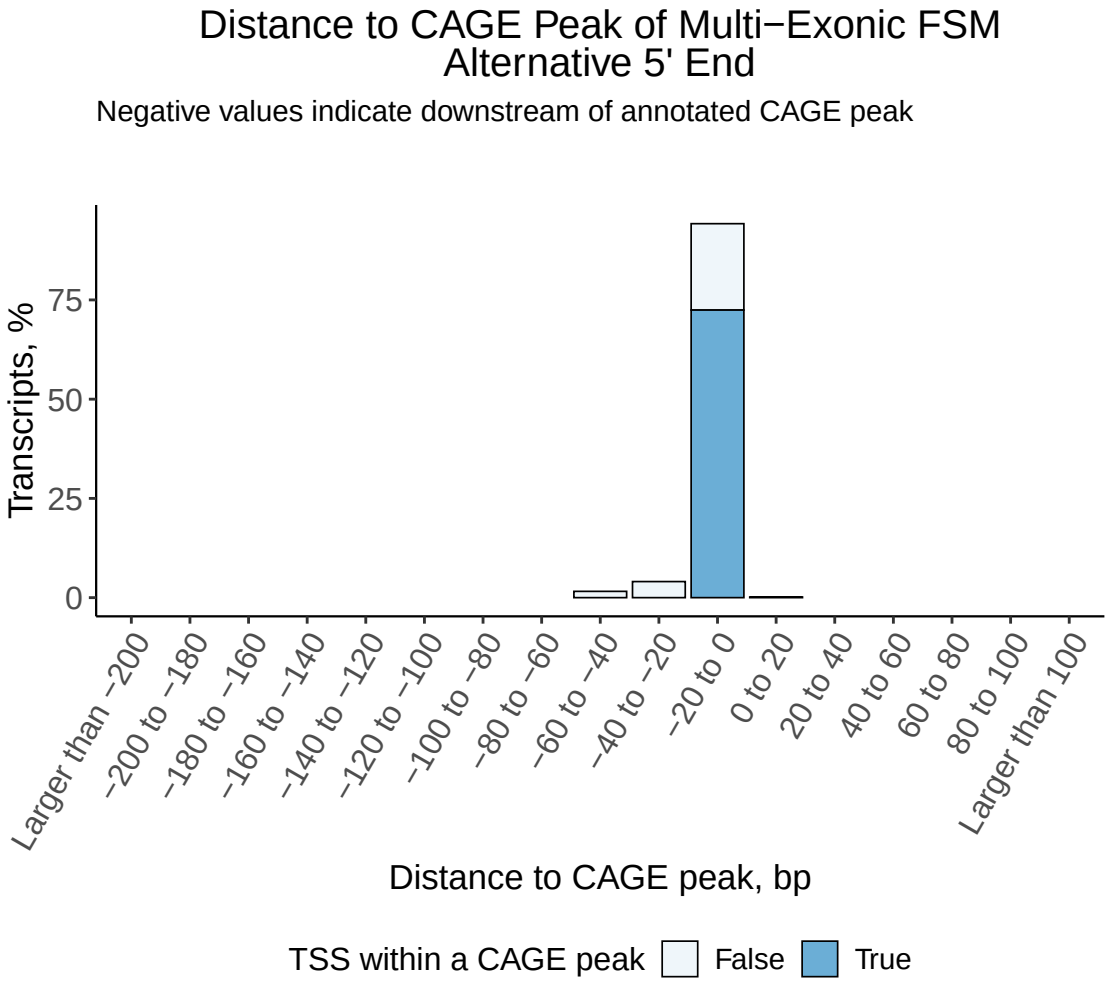

Distance to CAGE Peak of Multi−Exonic FSM Alternative 3' End
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
2000
1500
1000
500
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

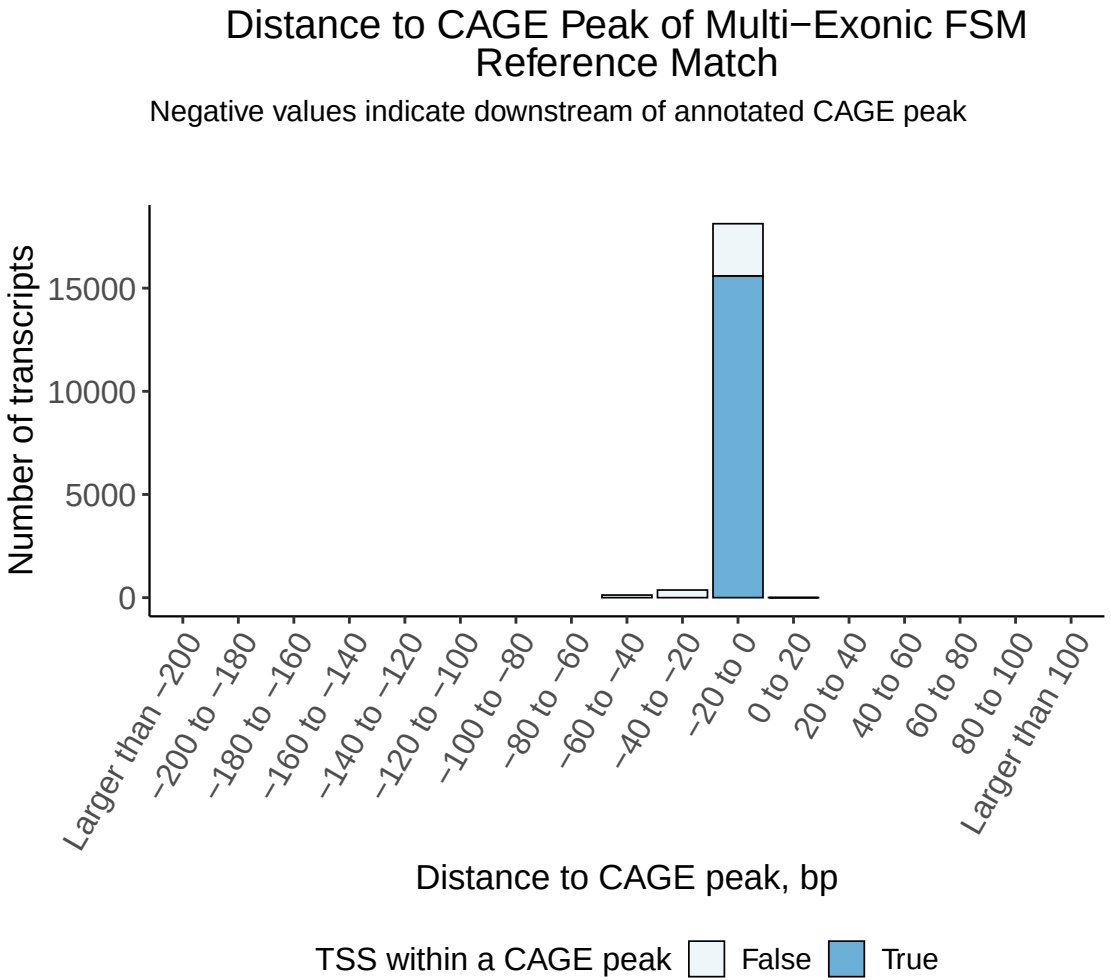

Distance to CAGE Peak of Multi−Exonic FSM Alternative 3' End
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

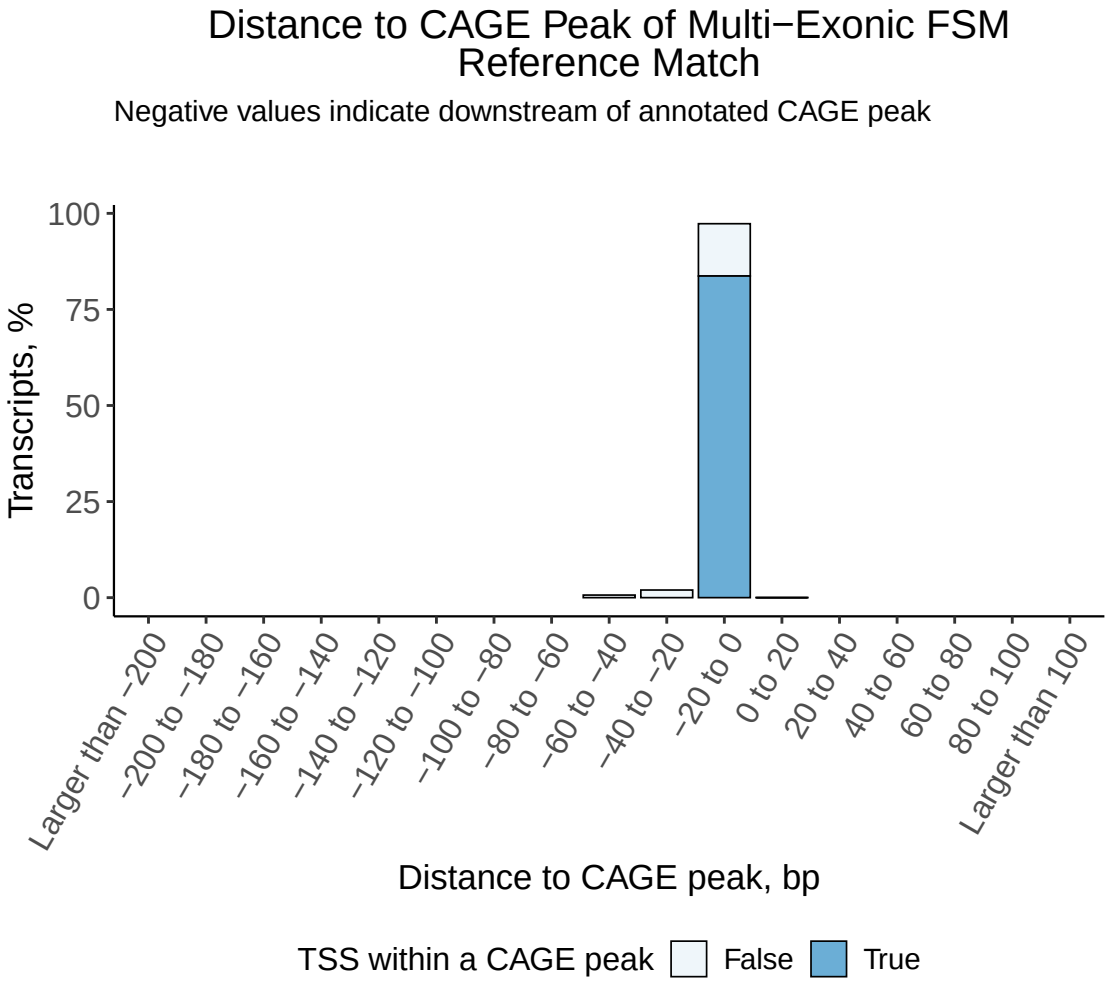

Distance to CAGE Peak of Multi−Exonic FSM Alternative 3'5' End
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
1000
750
500
250
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

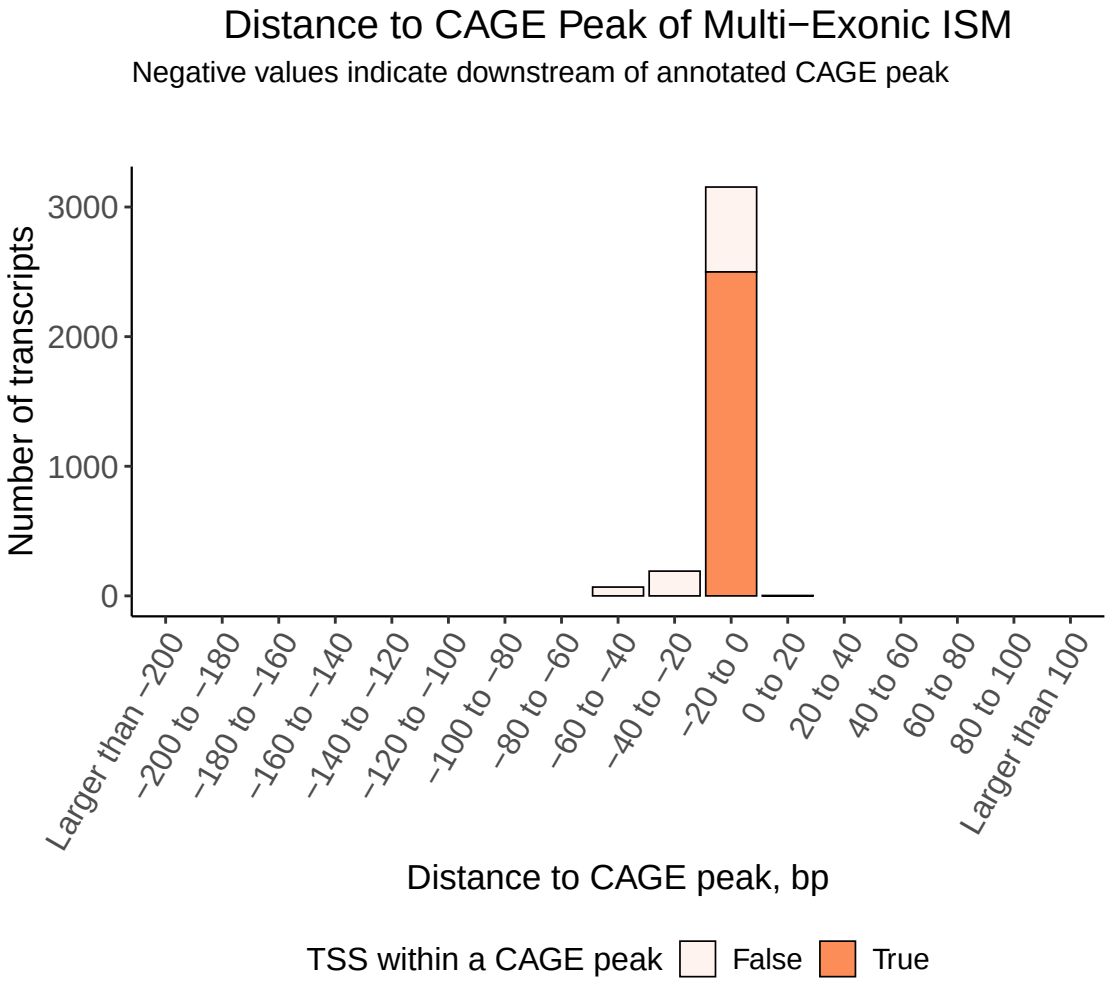

Distance to CAGE Peak of Multi−Exonic FSM Alternative 3'5' End
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

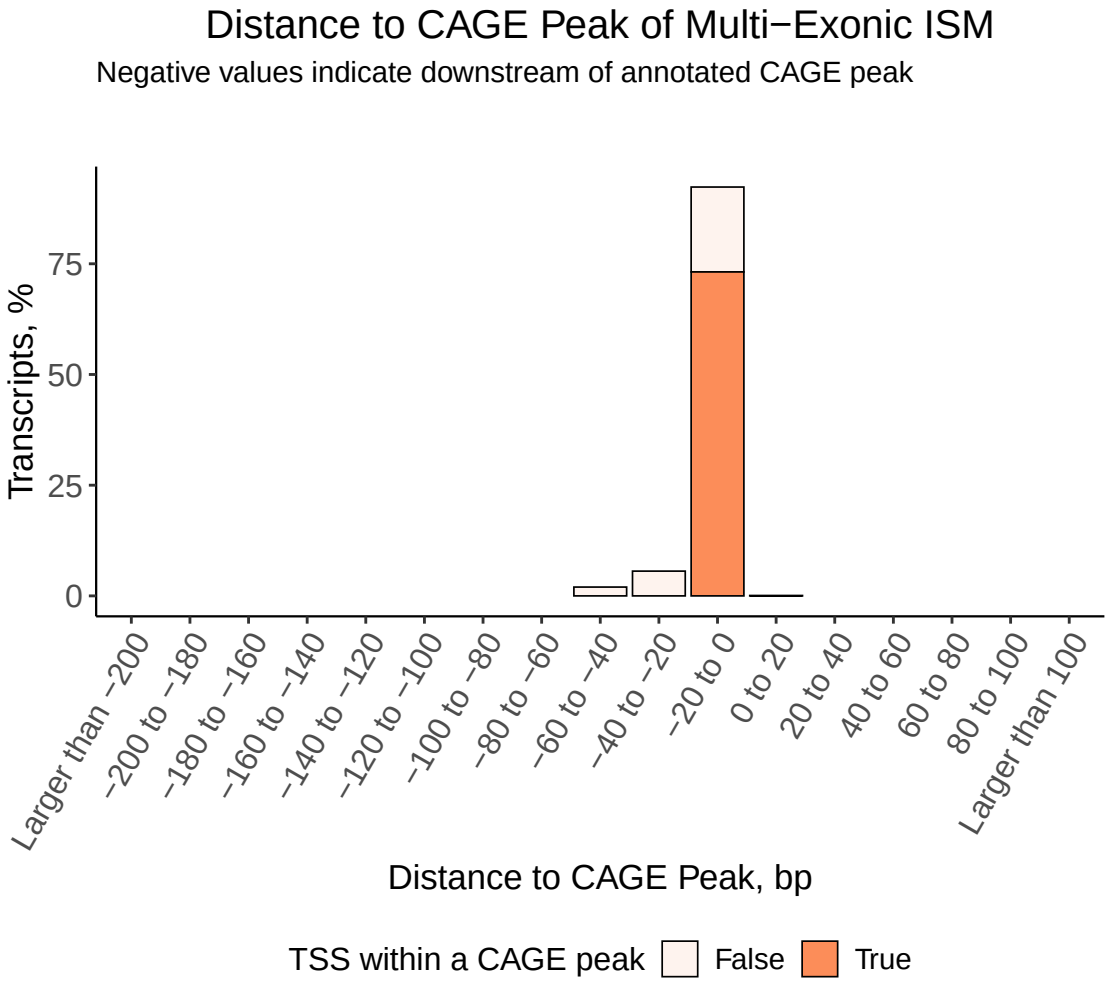

Distance to CAGE Peak of Multi−Exonic FSM Alternative 5' End
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
1000
750
500
250
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

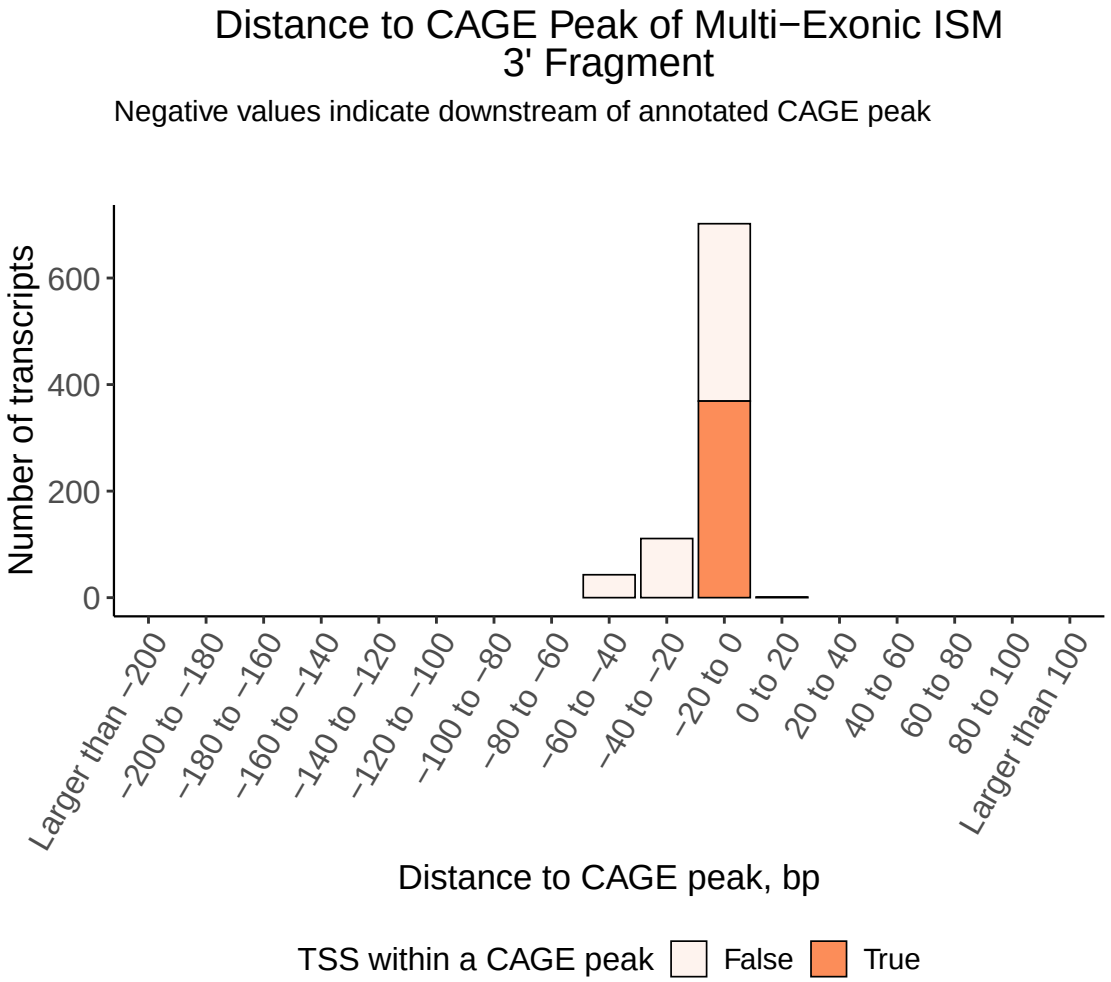

Distance to CAGE Peak of Multi−Exonic FSM Alternative 5' End
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

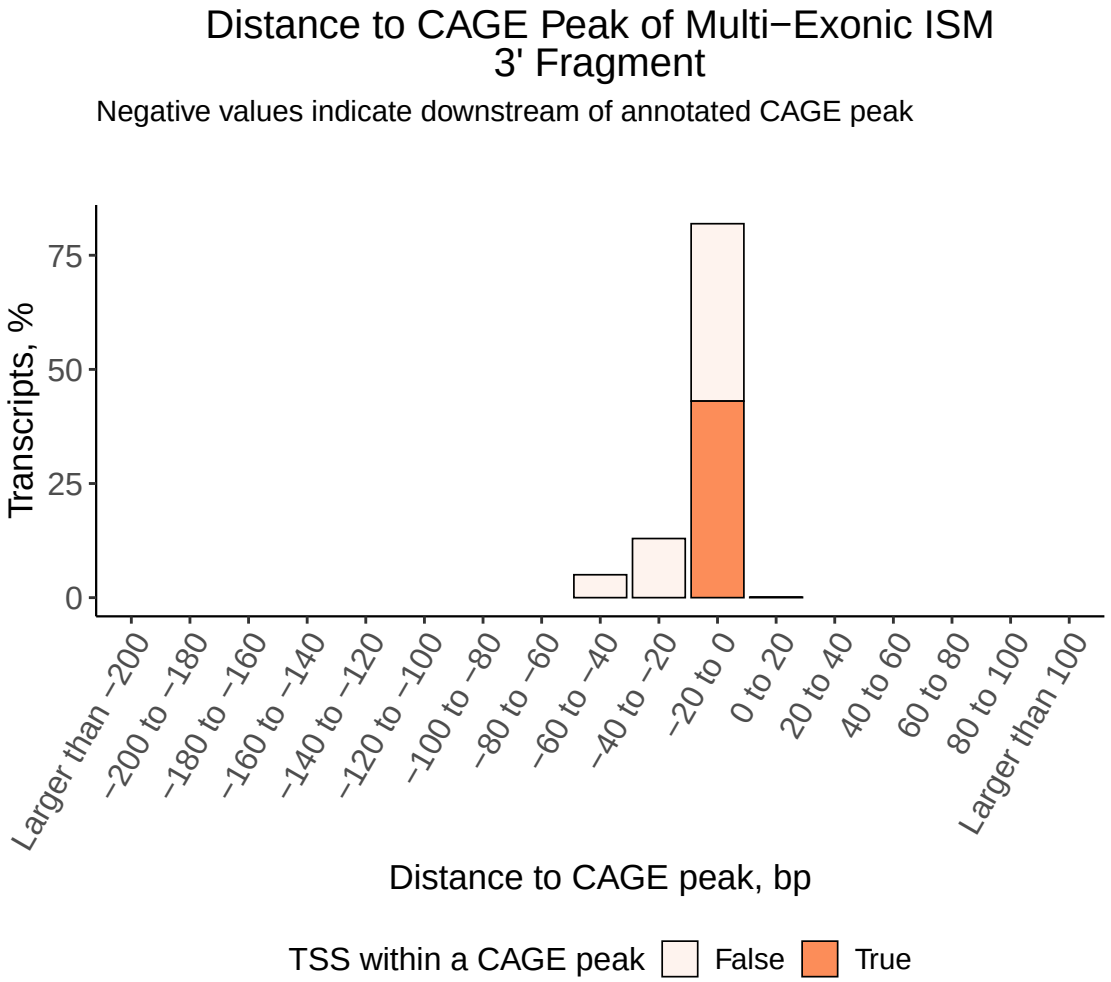

Distance to CAGE Peak of Multi−Exonic FSM Reference Match
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
15000
10000
5000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

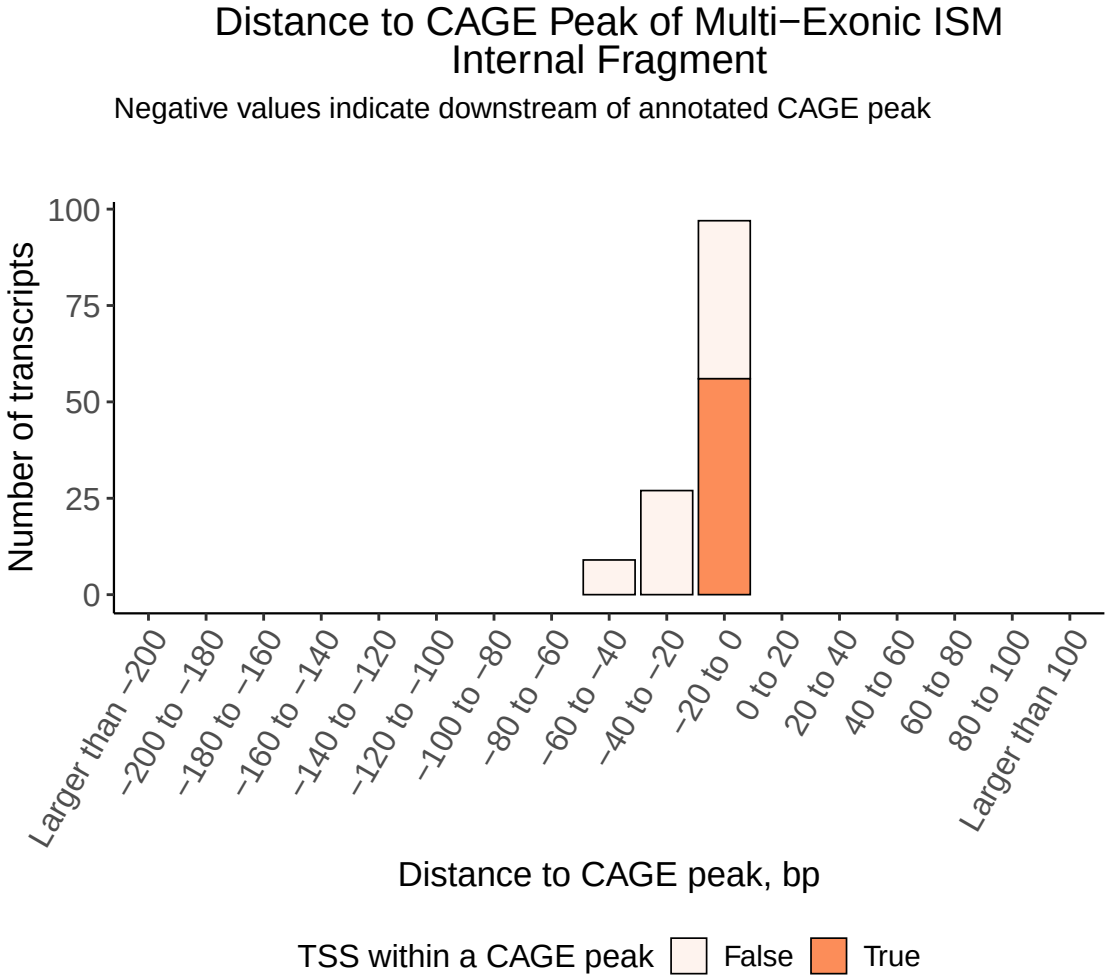

Distance to CAGE Peak of Multi−Exonic FSM Reference Match
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

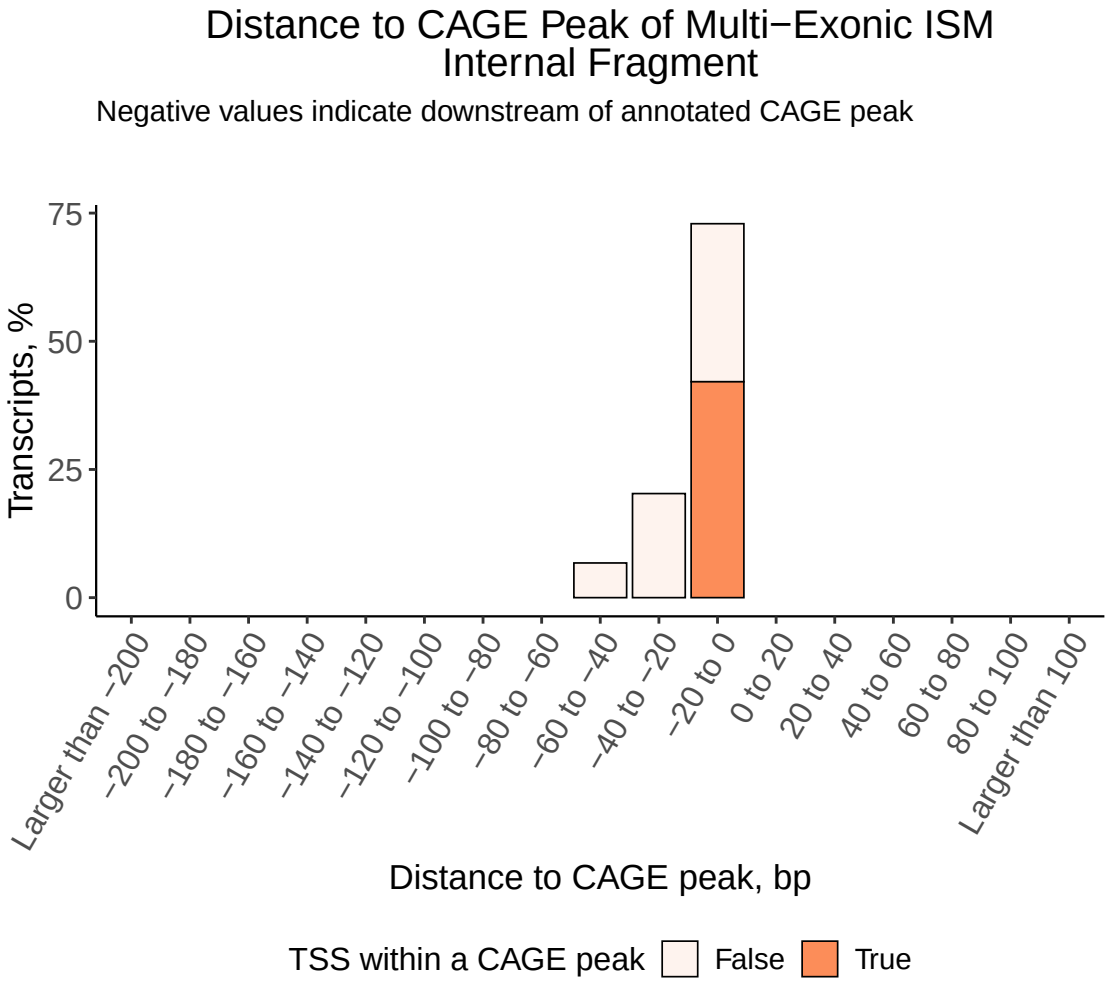

Distance to CAGE Peak of Multi−Exonic ISM
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
3000
2000
1000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

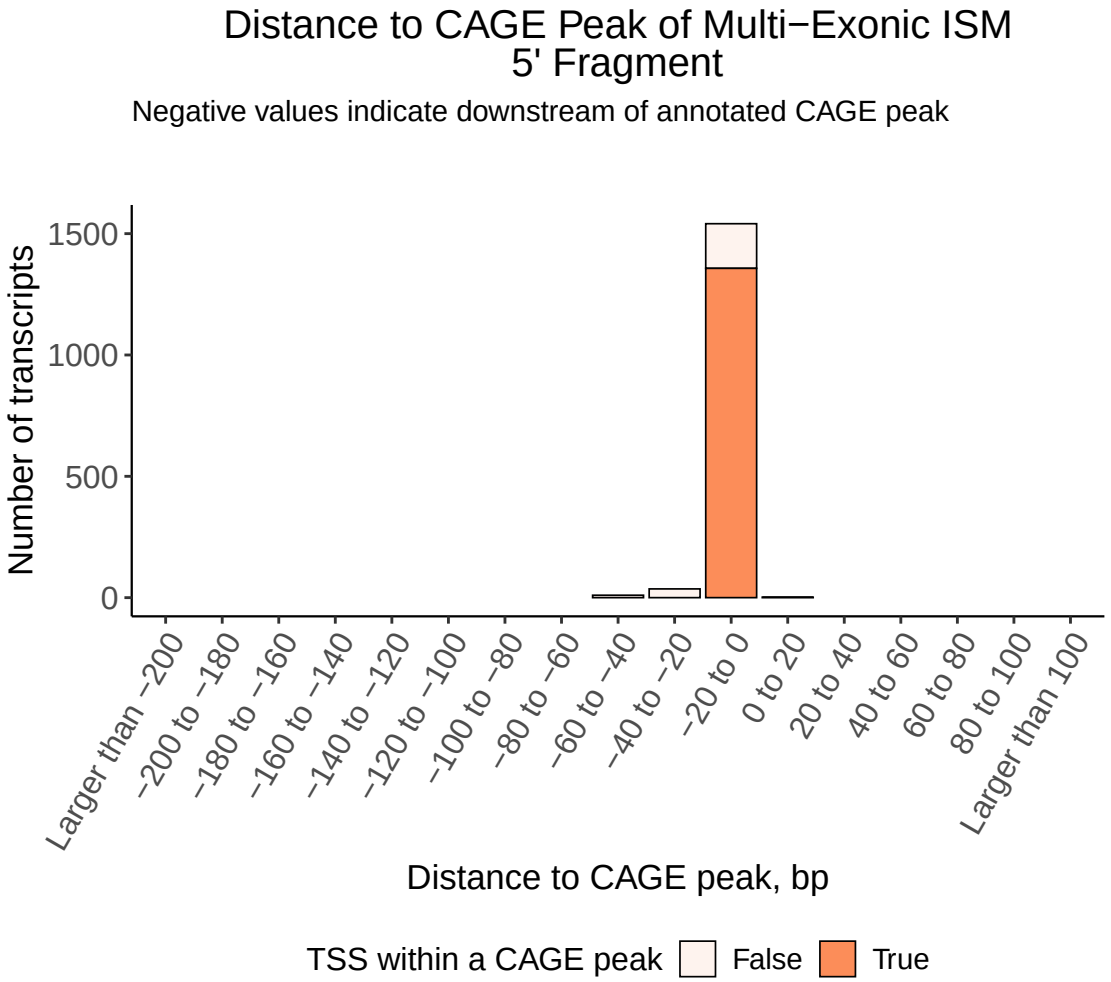

Distance to CAGE Peak of Multi−Exonic ISM
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE Peak, bp
TSS within a CAGE peak
False
True

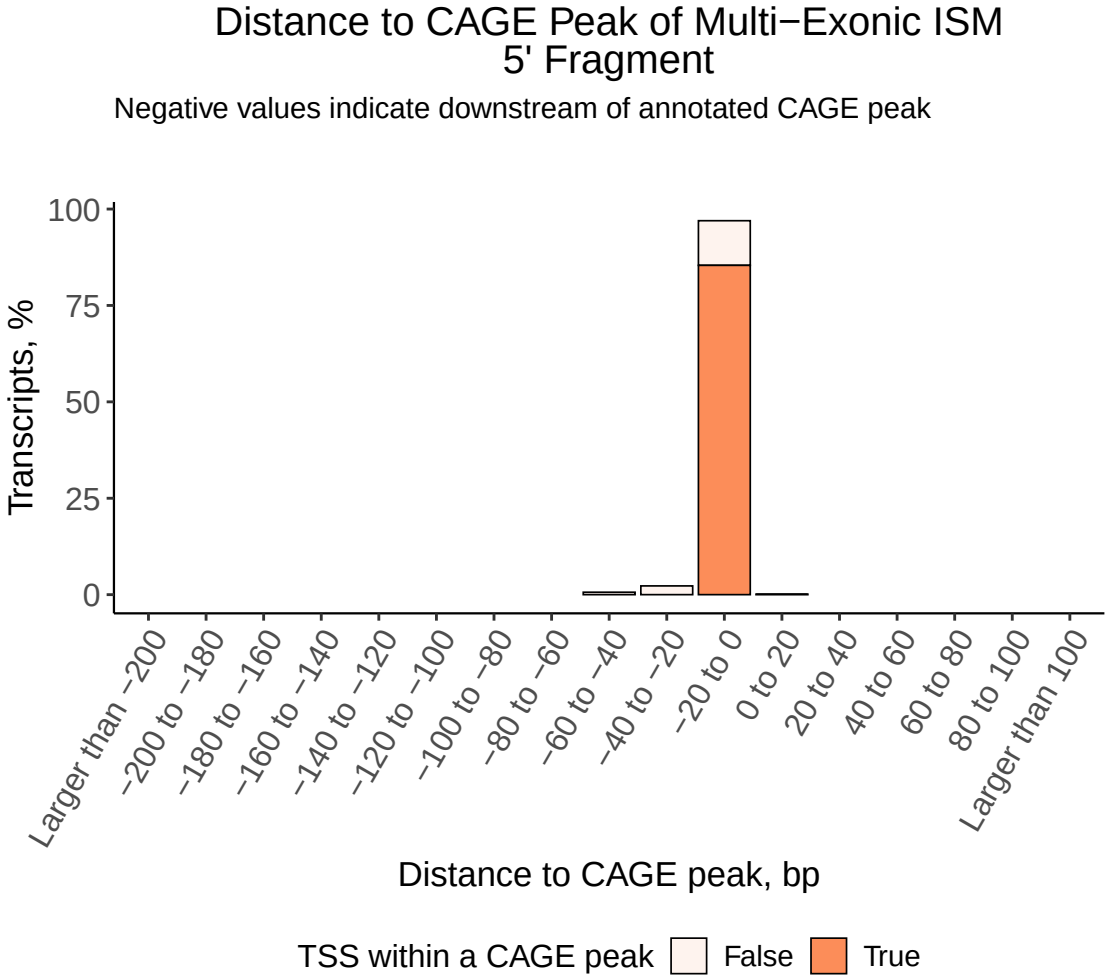

Distance to CAGE Peak of Multi−Exonic ISM 3' Fragment
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
0
200
400
600
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

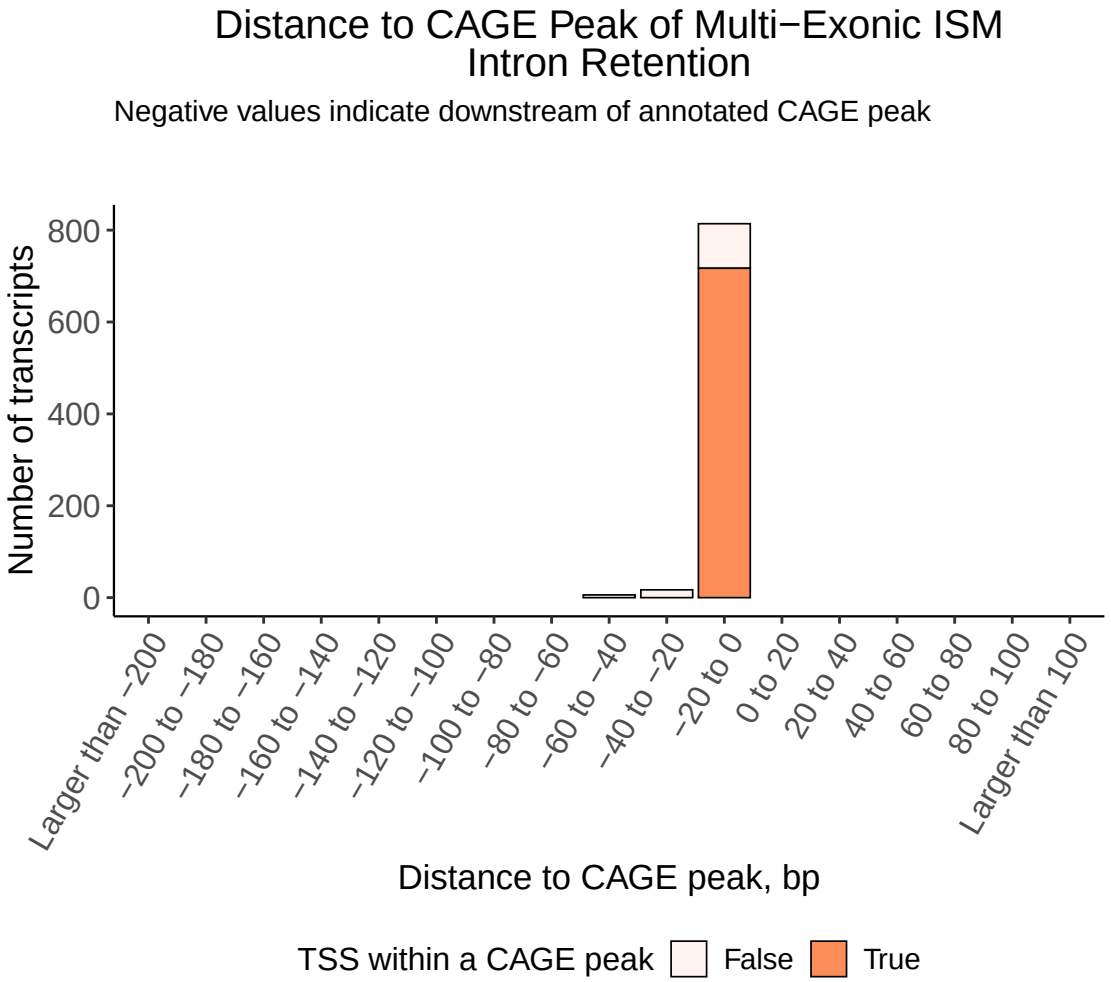

Distance to CAGE Peak of Multi−Exonic ISM 3' Fragment
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

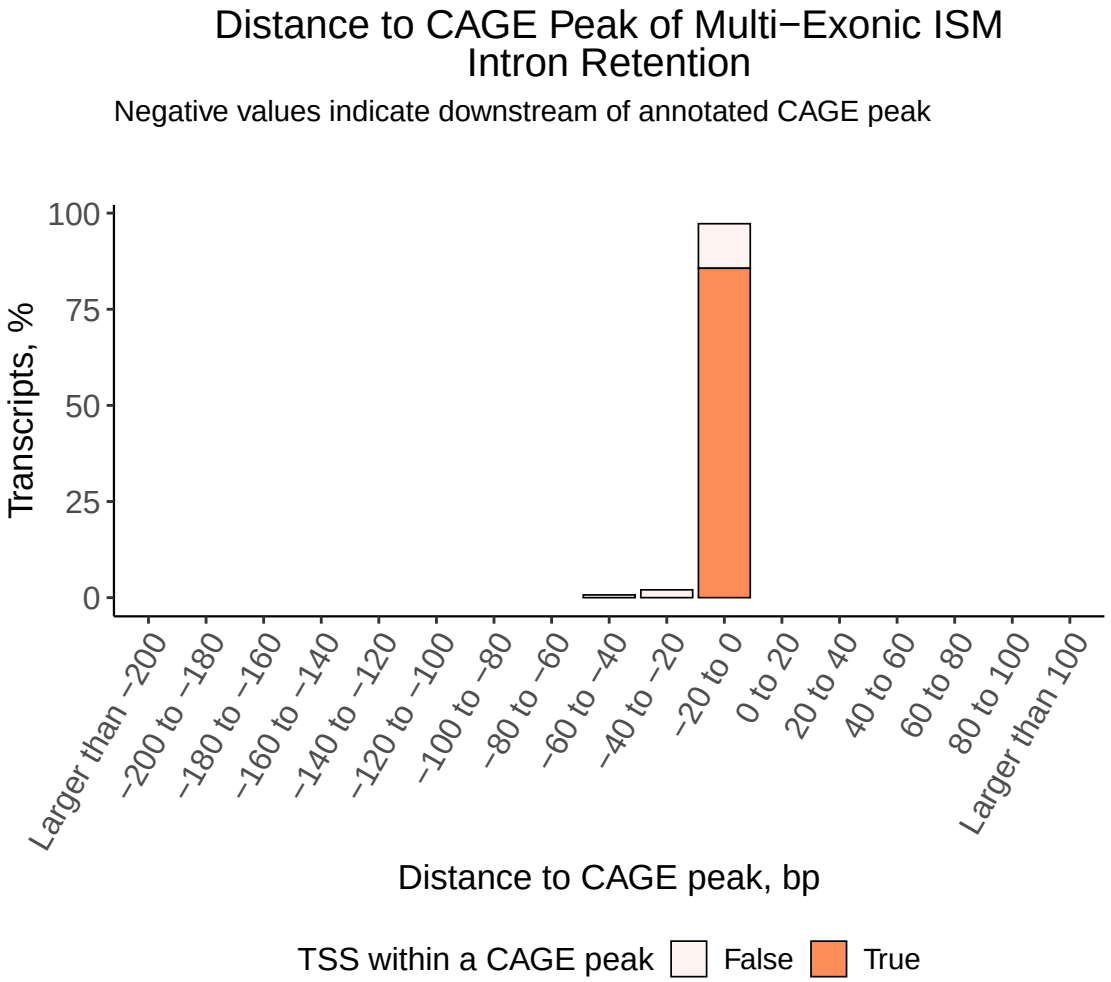

Distance to CAGE Peak of Multi−Exonic ISM Internal Fragment
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

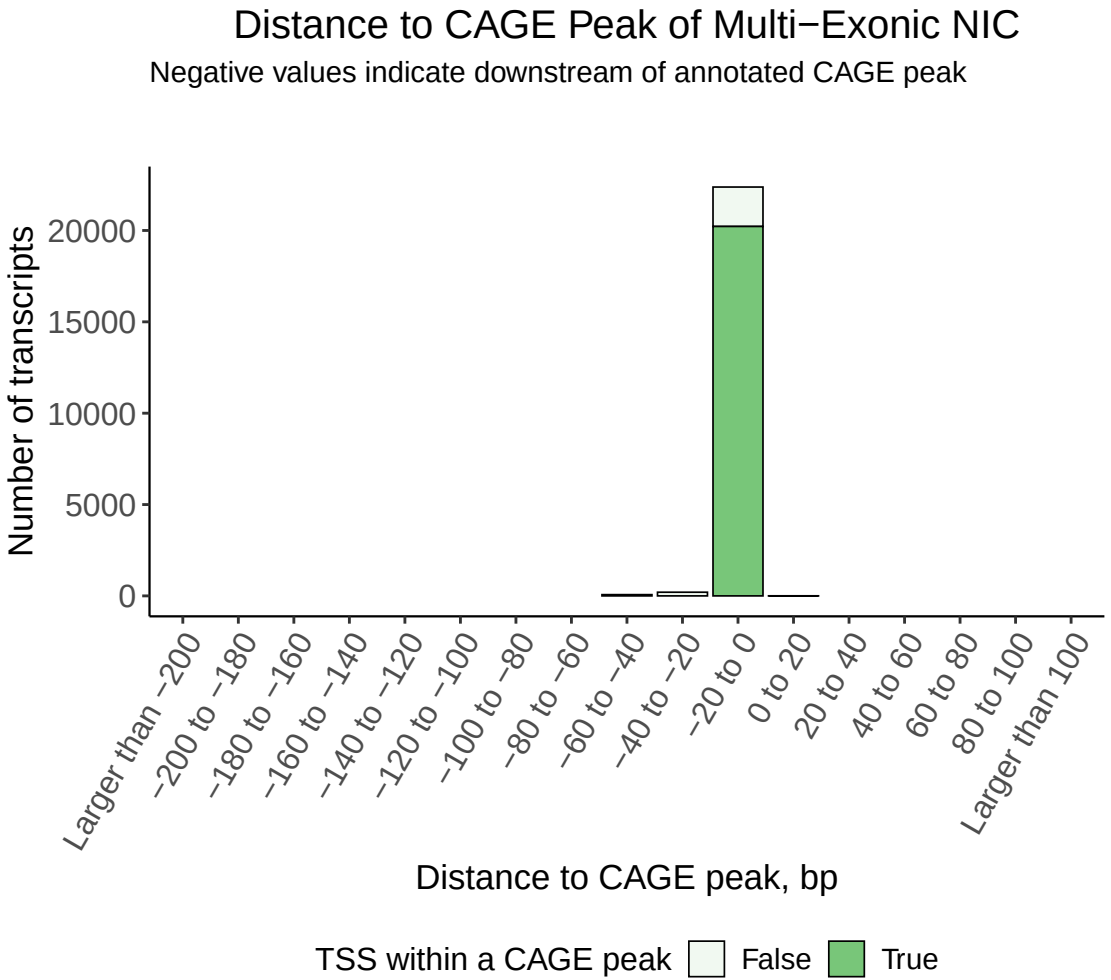

Distance to CAGE Peak of Multi−Exonic ISM Internal Fragment
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

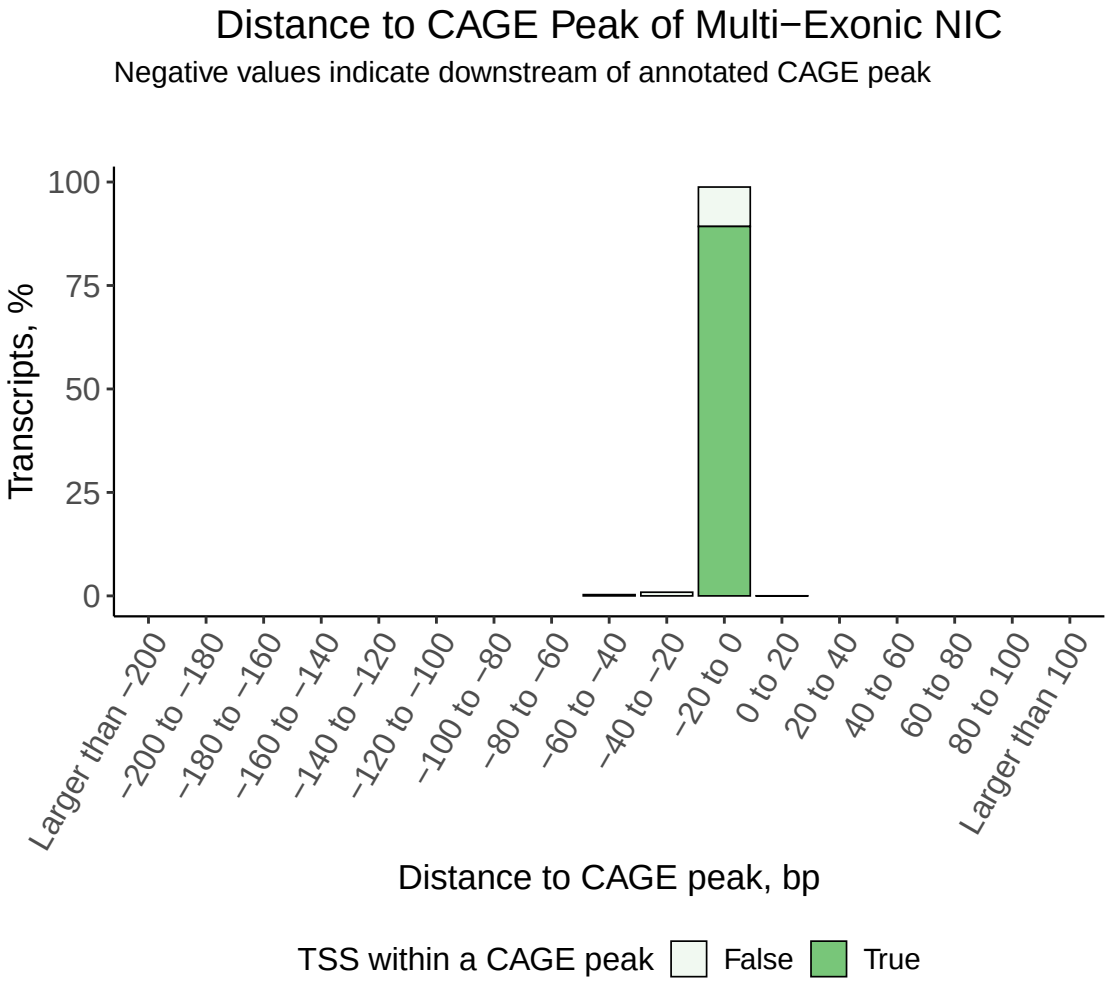

Distance to CAGE Peak of Multi−Exonic ISM 5' Fragment
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
1500
1000
500
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

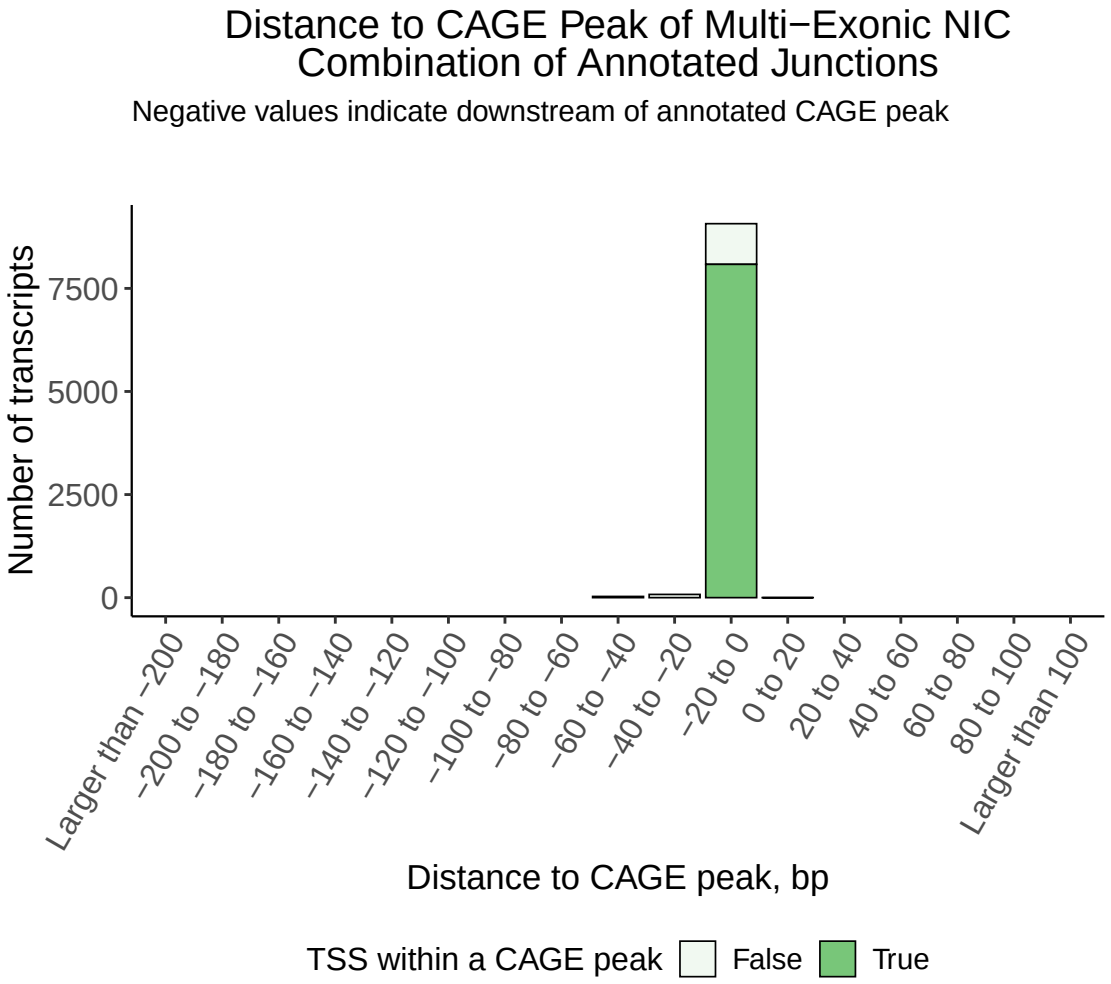

Distance to CAGE Peak of Multi−Exonic ISM 5' Fragment
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

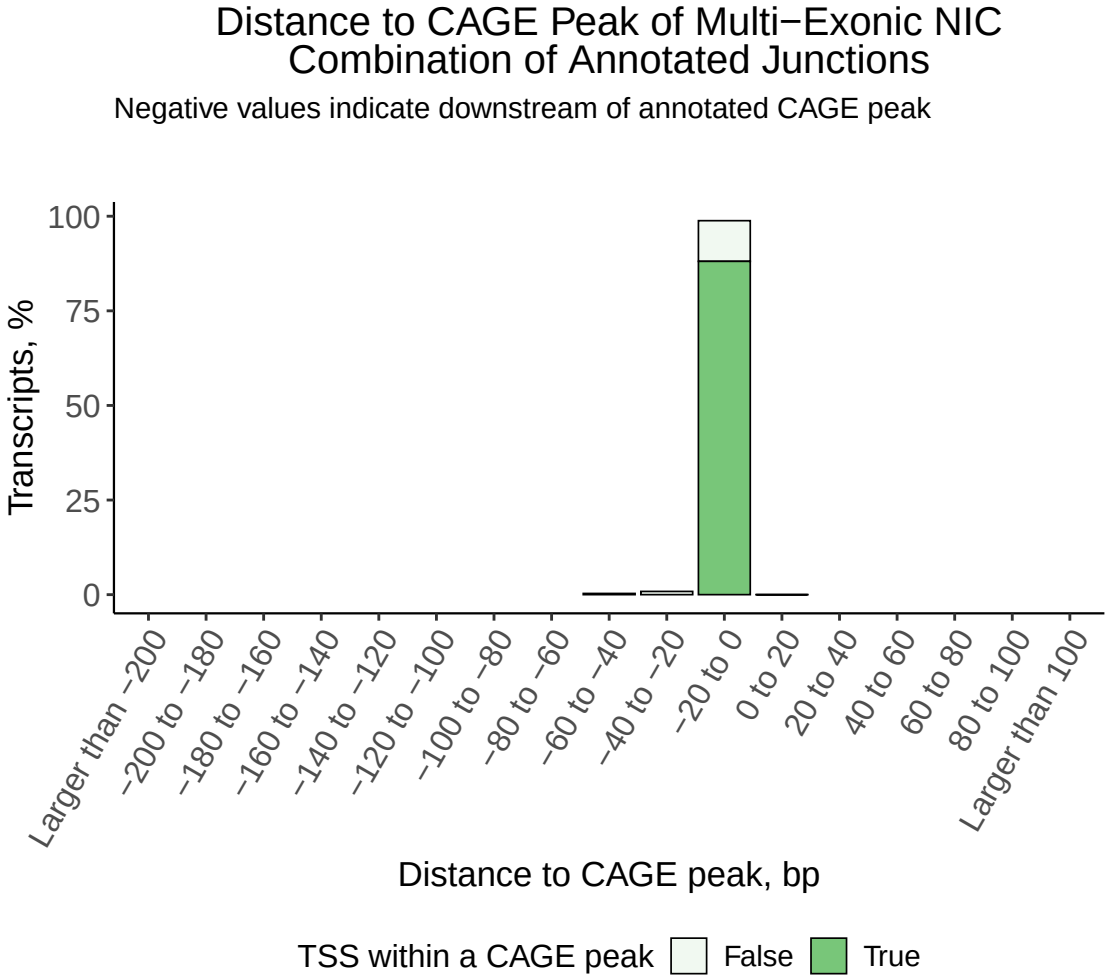

Distance to CAGE Peak of Multi−Exonic ISM Intron Retention
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
800
600
400
200
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

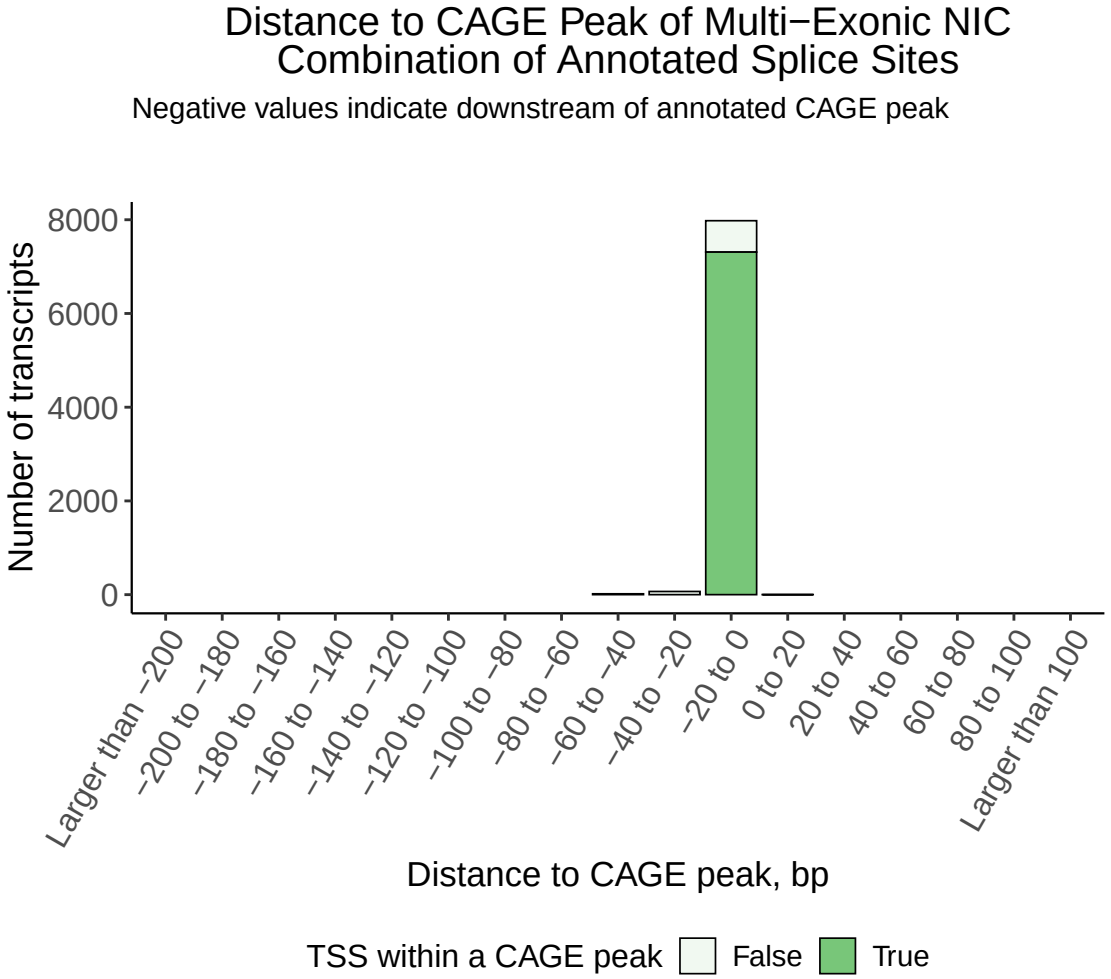

Distance to CAGE Peak of Multi−Exonic ISM Intron Retention
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

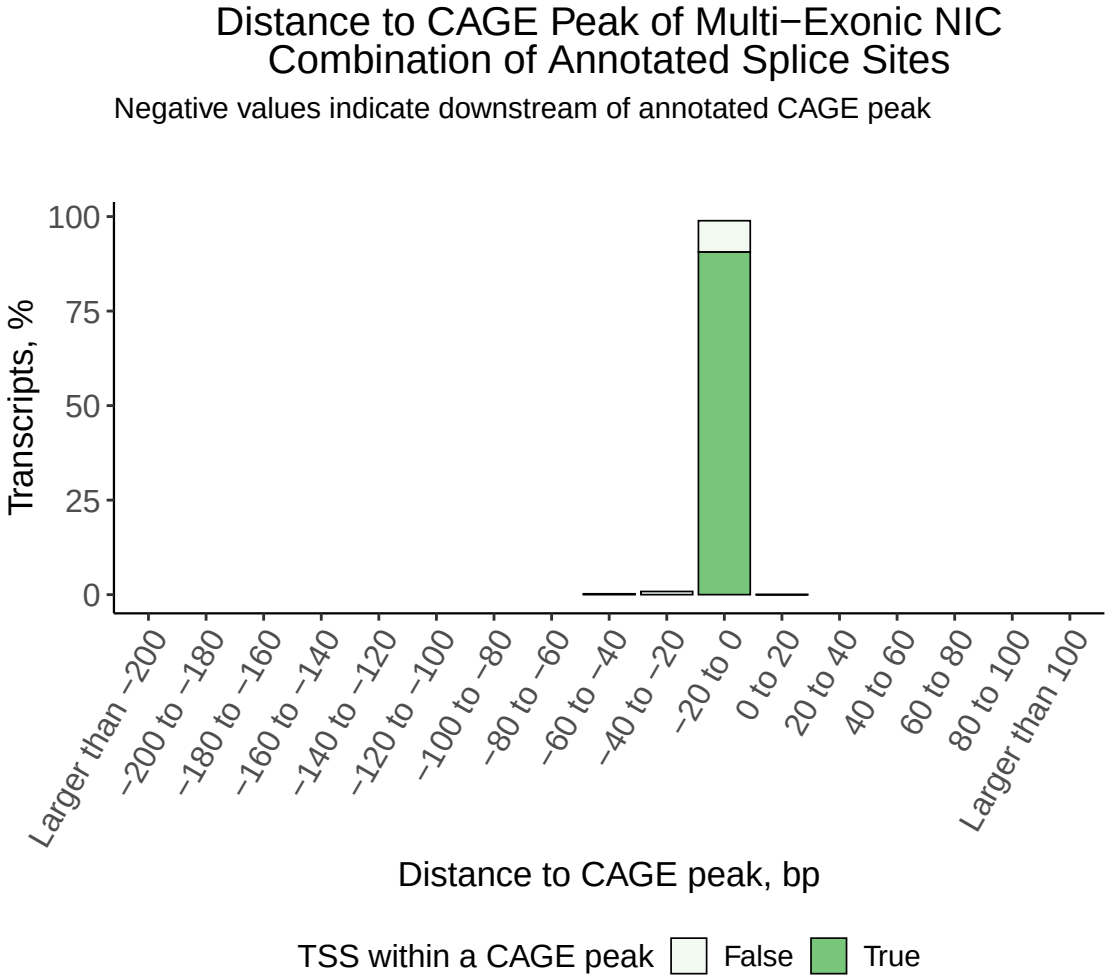

Distance to CAGE Peak of Multi−Exonic NIC
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
20000
15000
10000
5000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

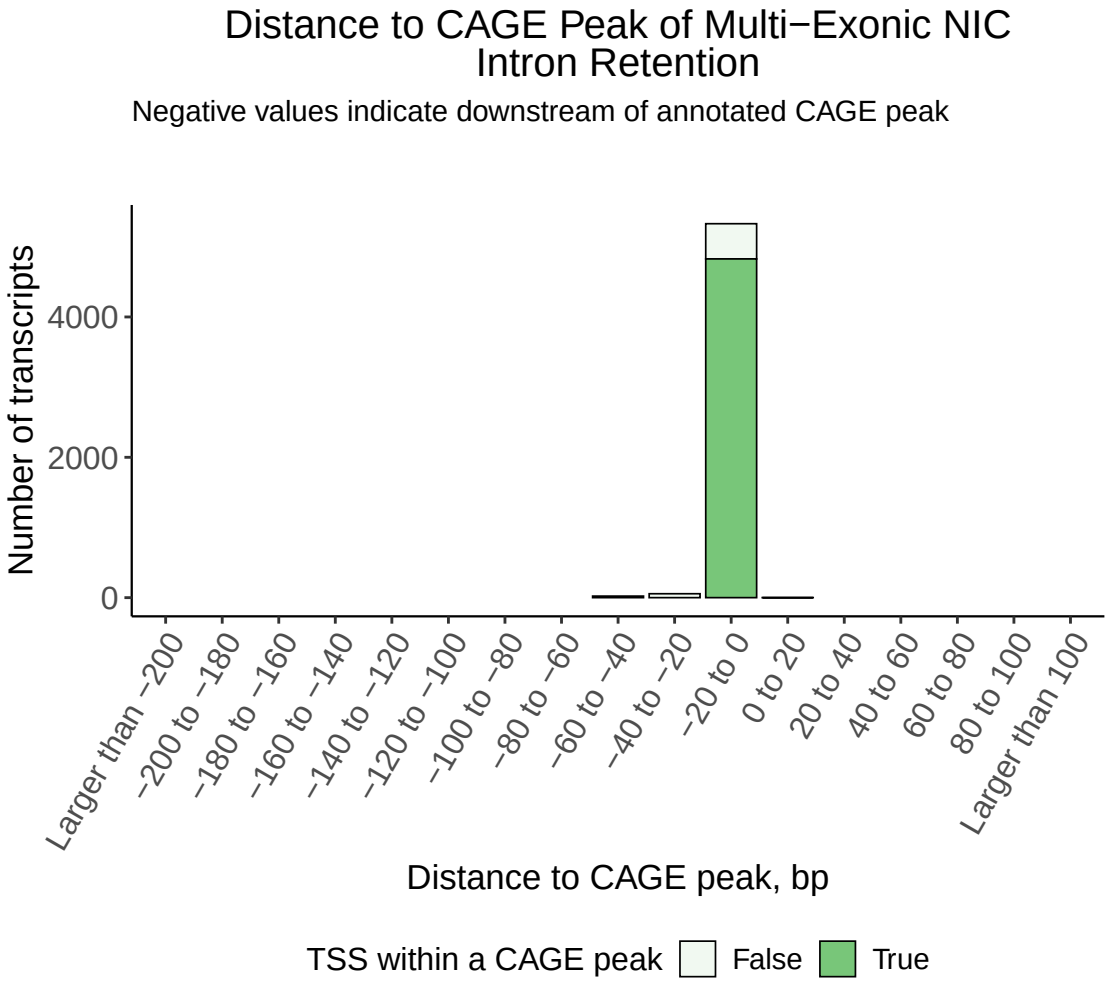

Distance to CAGE Peak of Multi−Exonic NIC
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

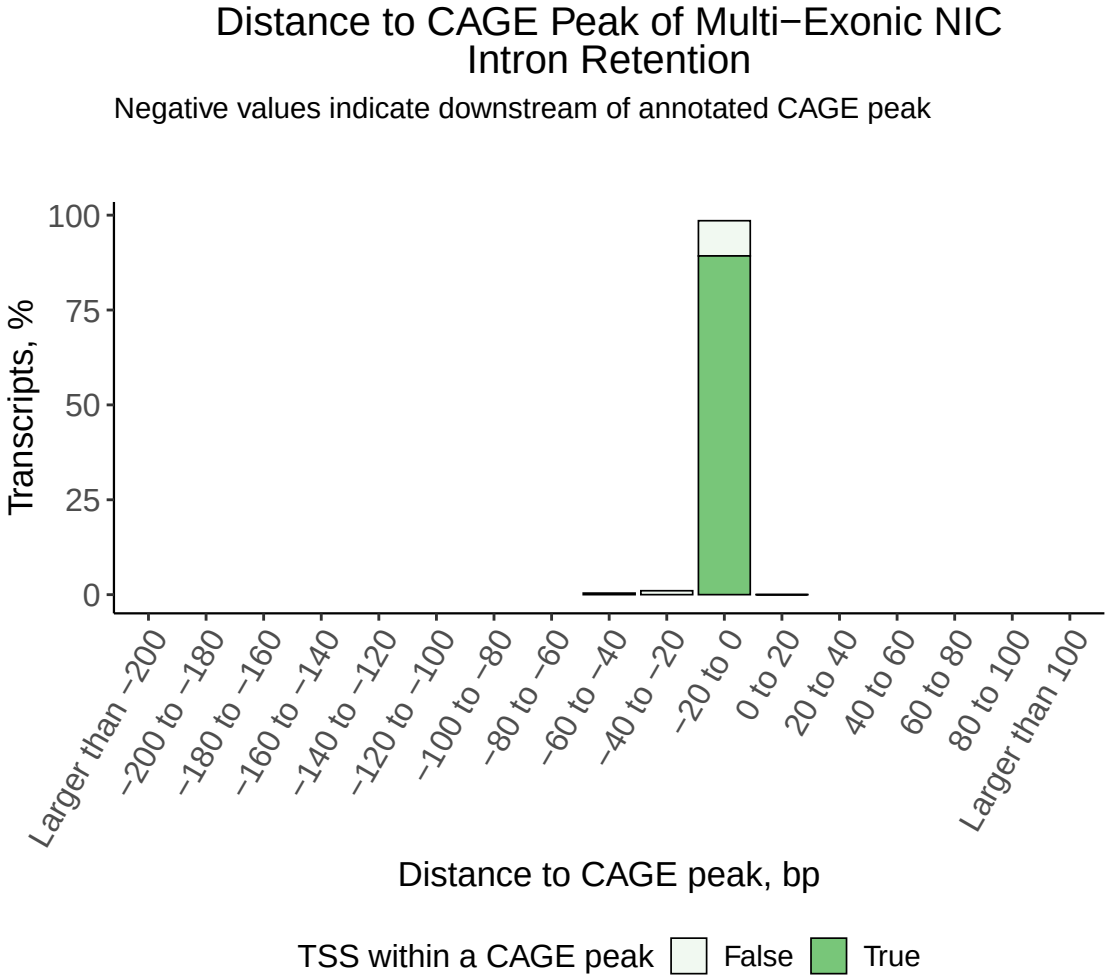

Distance to CAGE Peak of Multi−Exonic NIC Combination of Annotated Junctions
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
0
2500
5000
7500
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

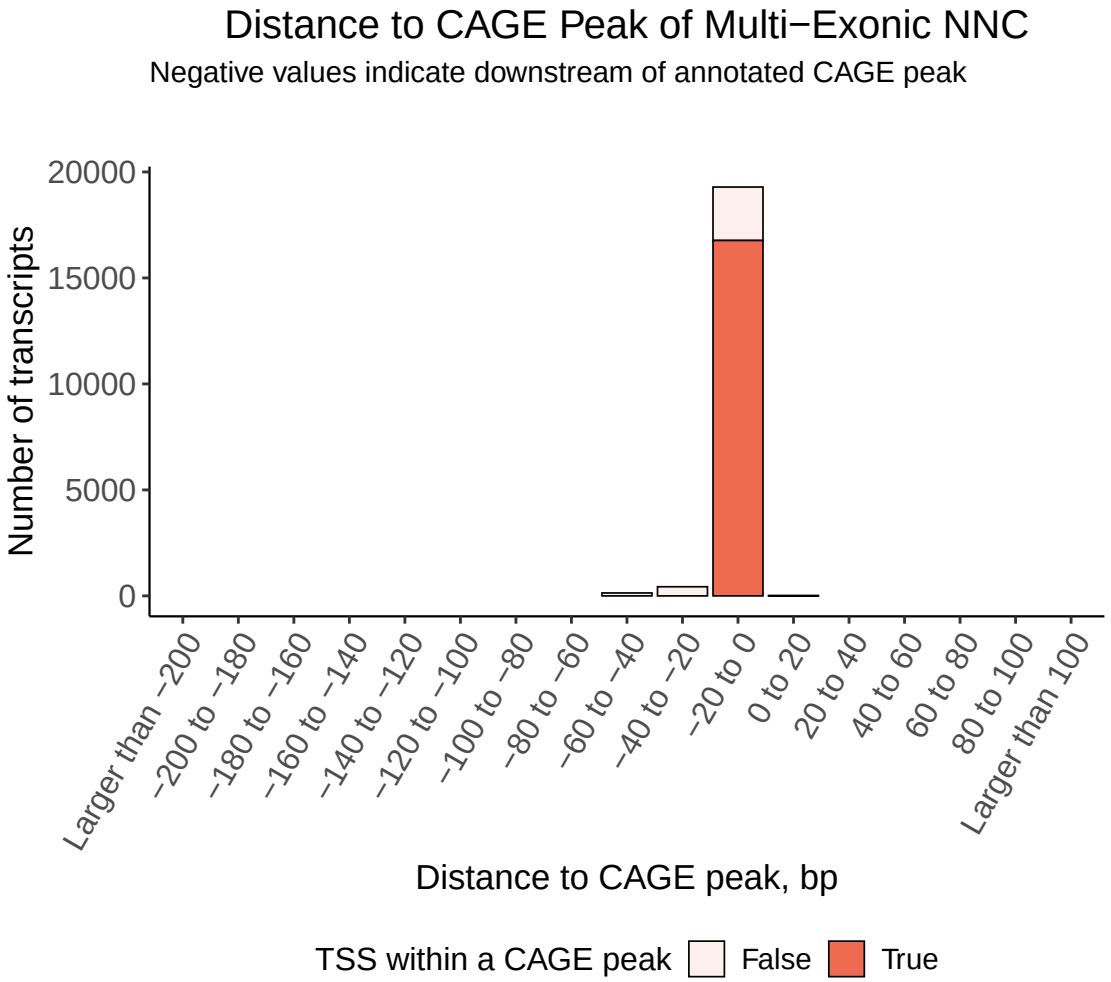

Distance to CAGE Peak of Multi−Exonic NIC Combination of Annotated Junctions
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

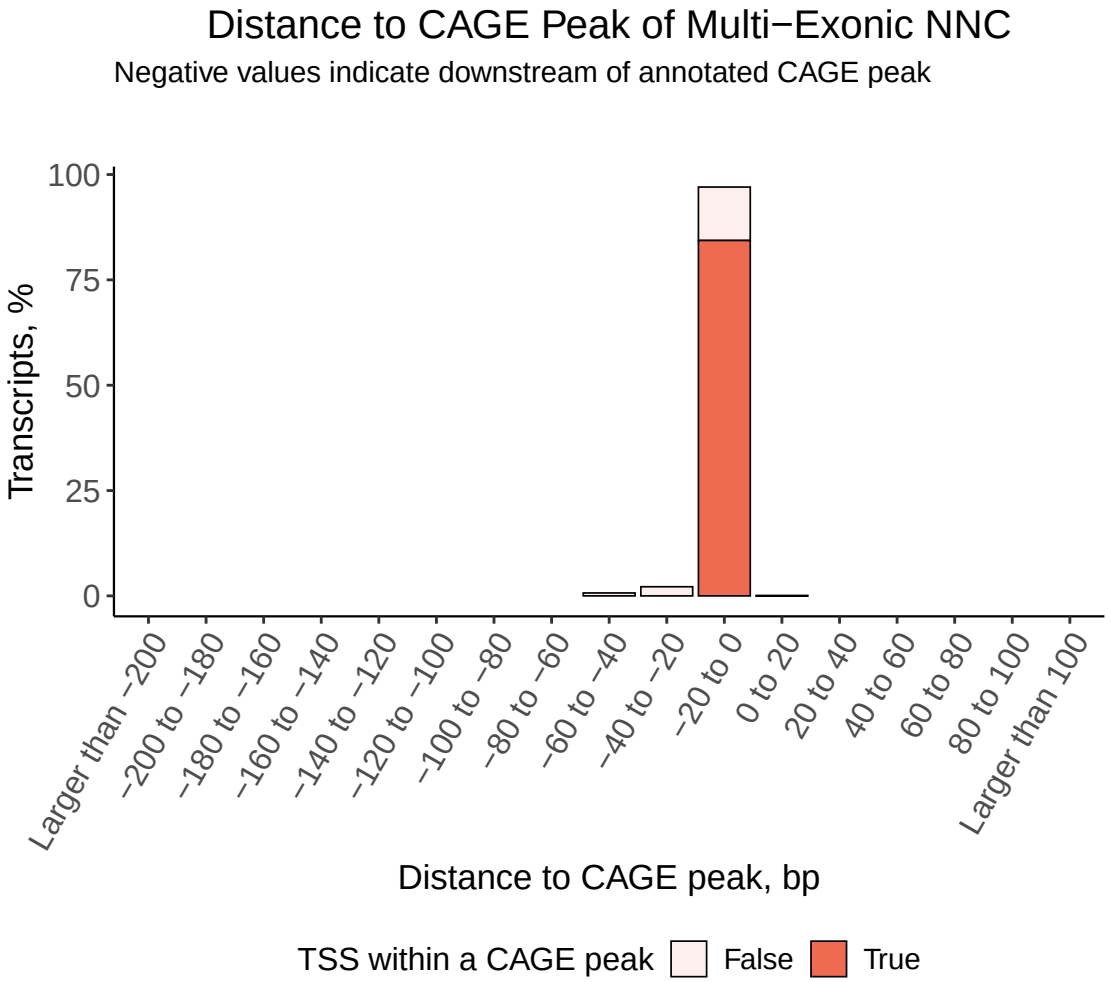

Distance to CAGE Peak of Multi−Exonic NIC Combination of Annotated Splice Sites
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
8000
6000
4000
2000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

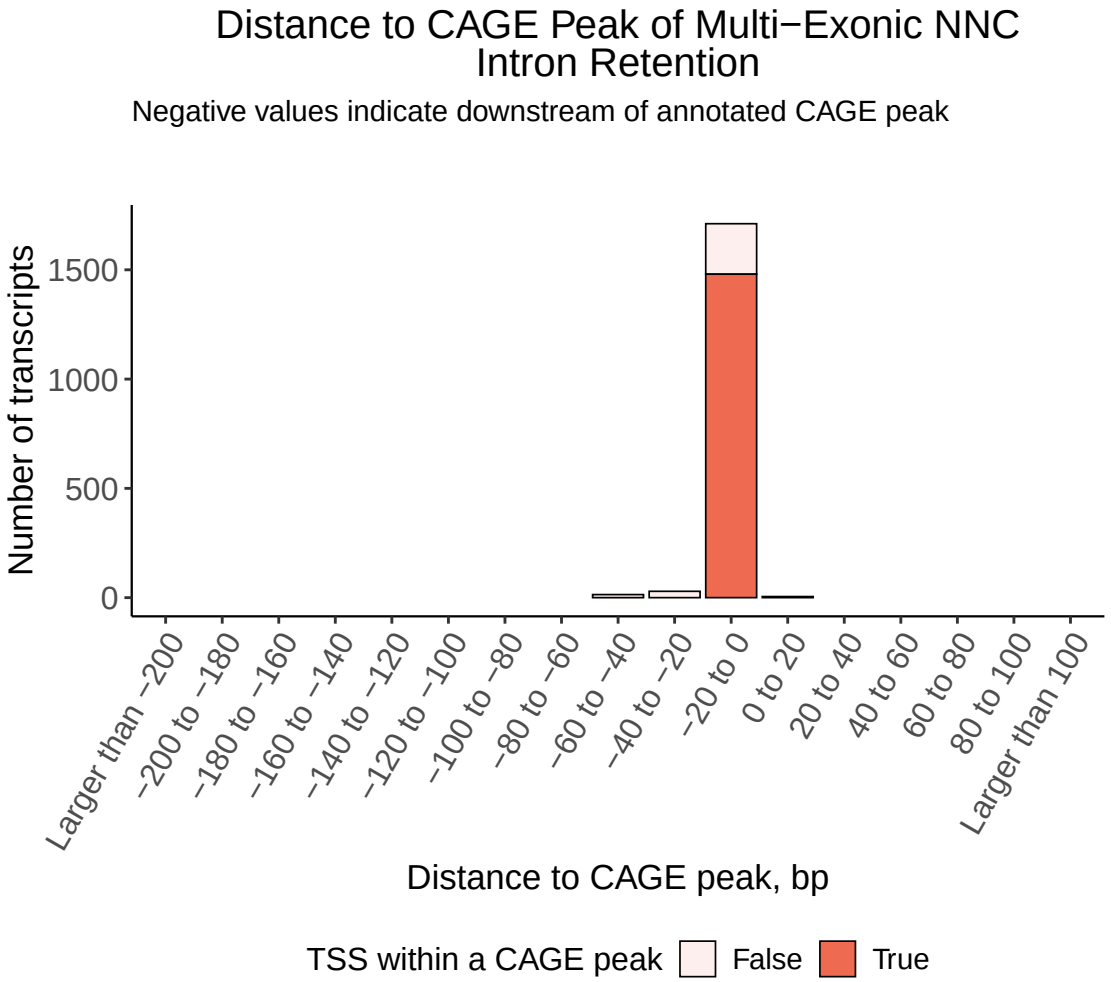

Distance to CAGE Peak of Multi−Exonic NIC Combination of Annotated Splice Sites
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

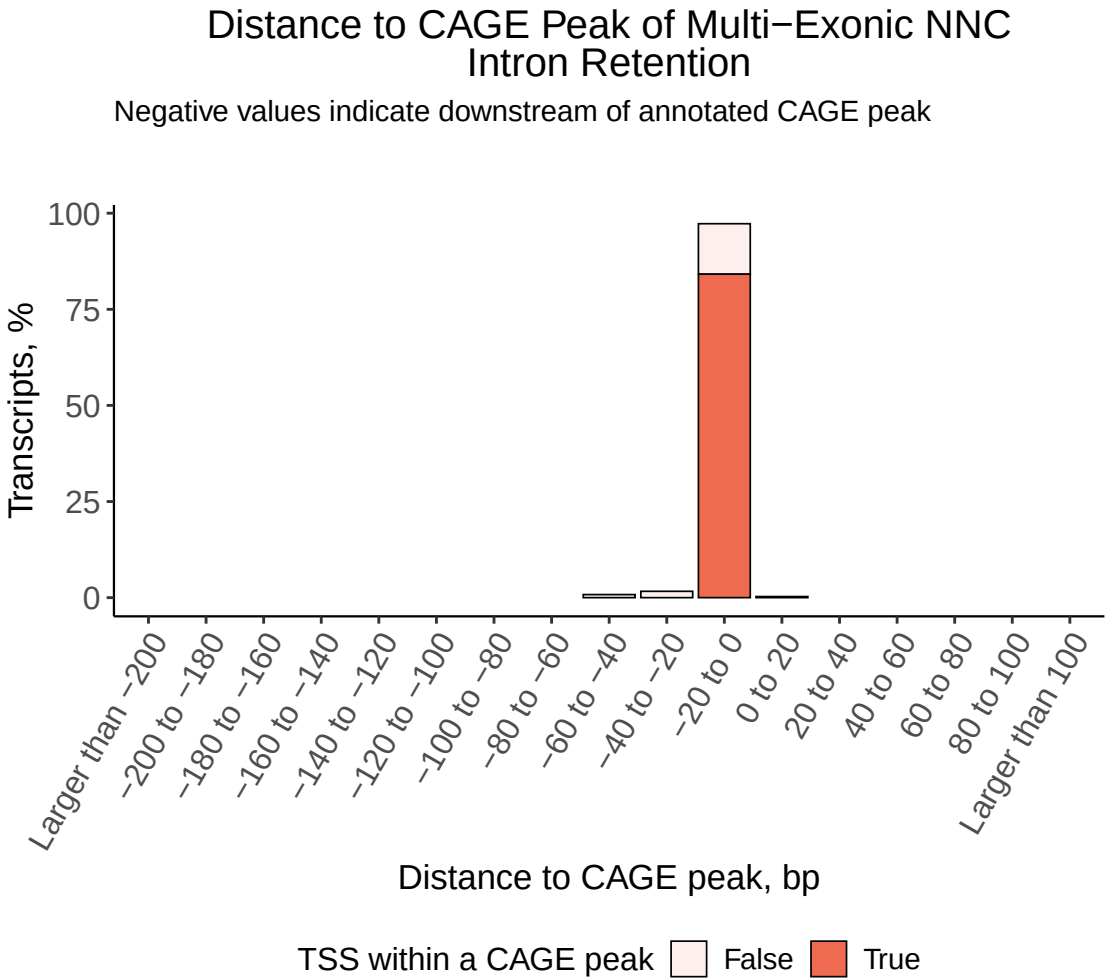

Distance to CAGE Peak of Multi−Exonic NIC Intron Retention
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
4000
2000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

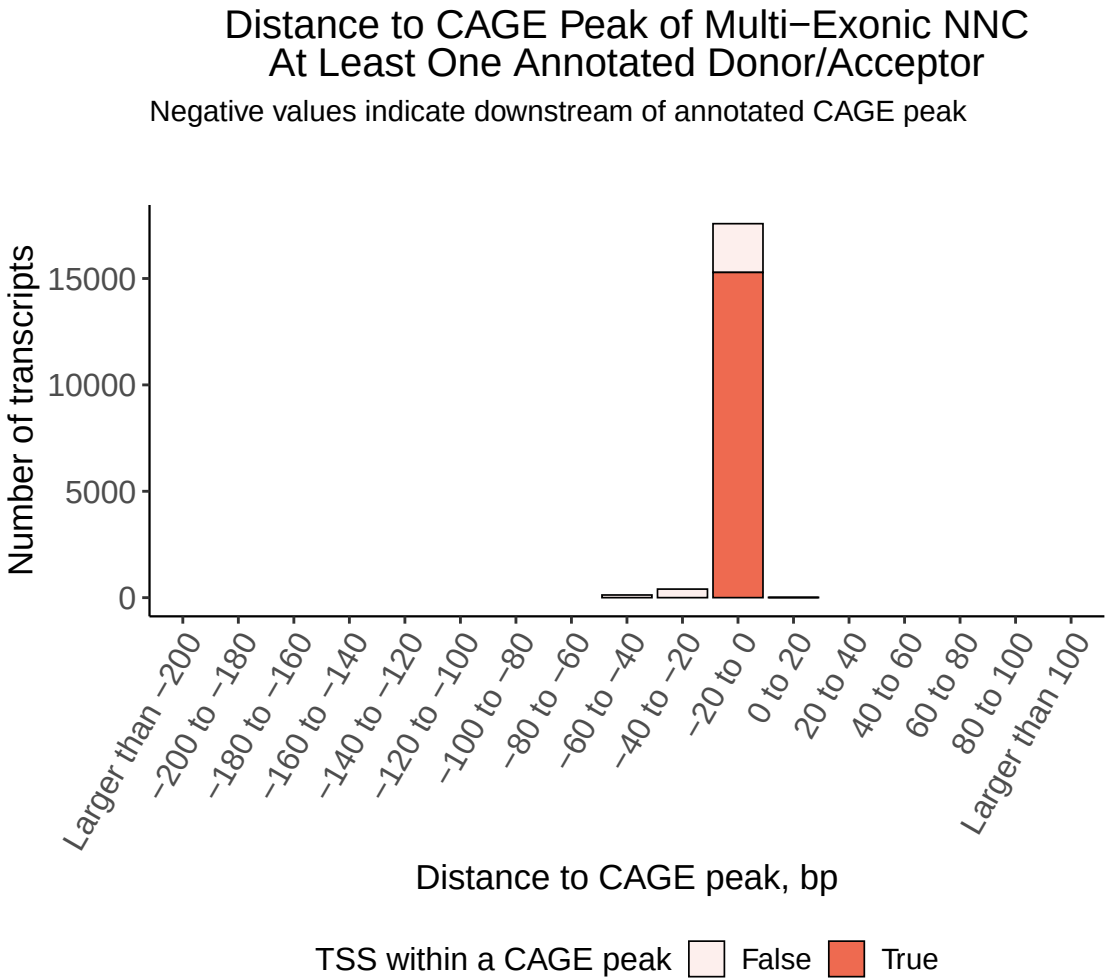

Distance to CAGE Peak of Multi−Exonic NIC Intron Retention
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

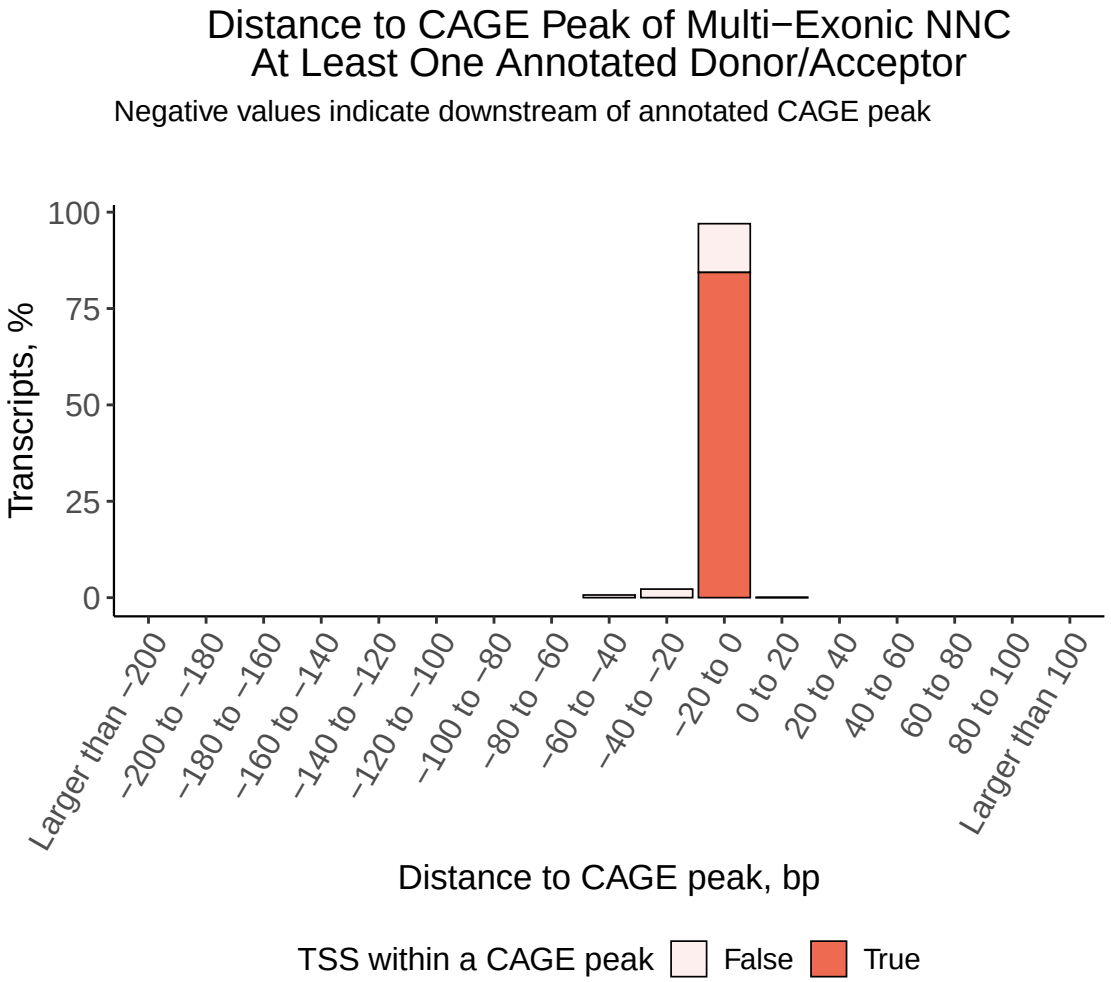

Distance to CAGE Peak of Multi−Exonic NNC
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
20000
15000
10000
5000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True


Distance to CAGE Peak of Multi−Exonic NNC
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True


Distance to CAGE Peak of Multi−Exonic NNC Intron Retention
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
1500
1000
500
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True


Distance to CAGE Peak of Multi−Exonic NNC Intron Retention
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True


Distance to CAGE Peak of Multi−Exonic NNC
At Least One Annotated Donor/Acceptor
Negative values indicate downstream of annotated CAGE peak
Number of transcripts
15000
10000
5000
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True


Distance to CAGE Peak of Multi−Exonic NNC
At Least One Annotated Donor/Acceptor
Negative values indicate downstream of annotated CAGE peak
Transcripts, %
100
75
50
25
0
Larger than −200
−200 to −180
−180 to −160
−160 to −140
−140 to −120
−120 to −100
−100 to −80
−80 to −60
−60 to −40
−40 to −20
−20 to 0
0 to 20
20 to 40
40 to 60
60 to 80
80 to 100
Larger than 100
Distance to CAGE peak, bp
TSS within a CAGE peak
False
True

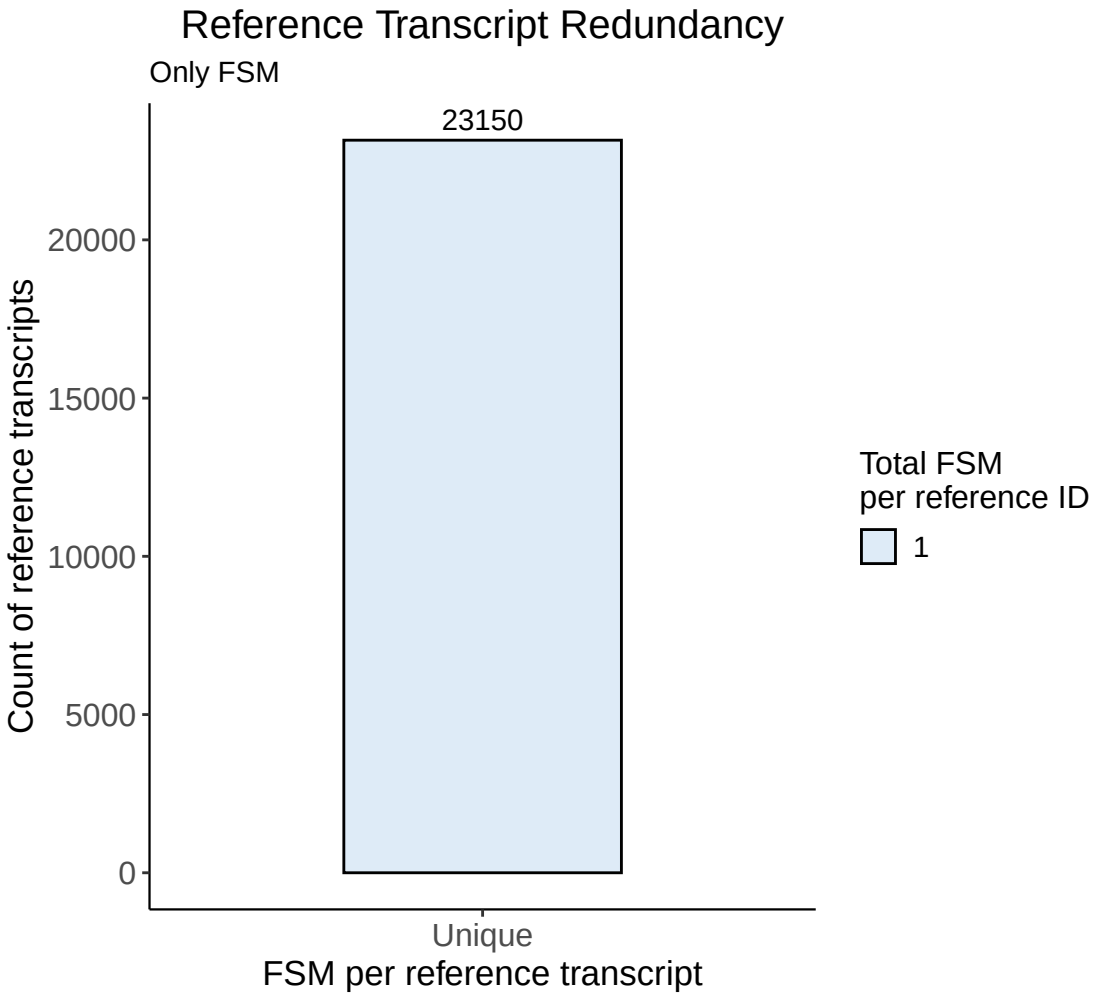
Number of CAGE Detected

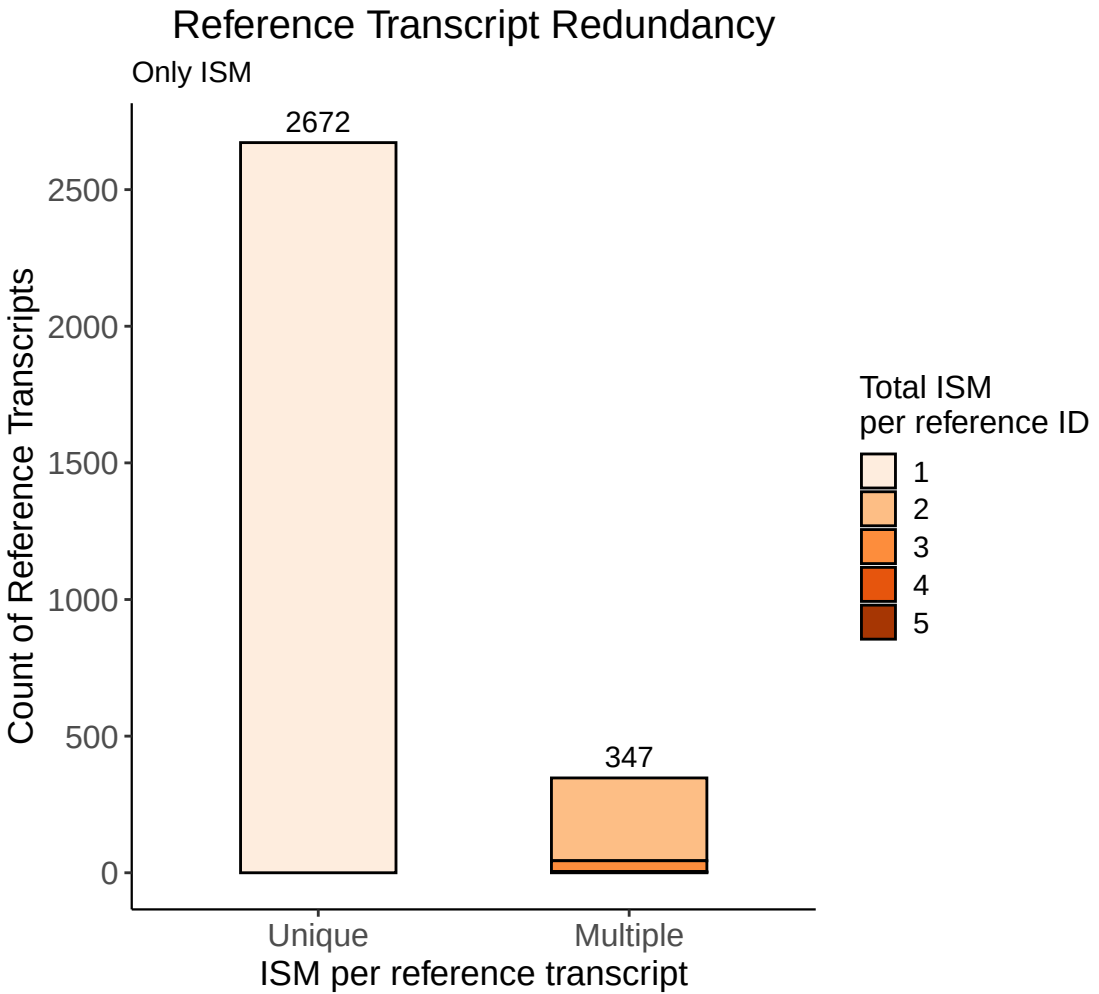
Category	Count	CAGE Detected	%
FSM	25235	20170	80
ISM	3781	2612	69
NIC	22995	20416	89
NNC	19878	16766	84
Genic Genomic	1041	533	51
Antisense	856	392	46
Fusion	549	441	80
Intergenic	2131	878	41
Genic Intron	551	85	15

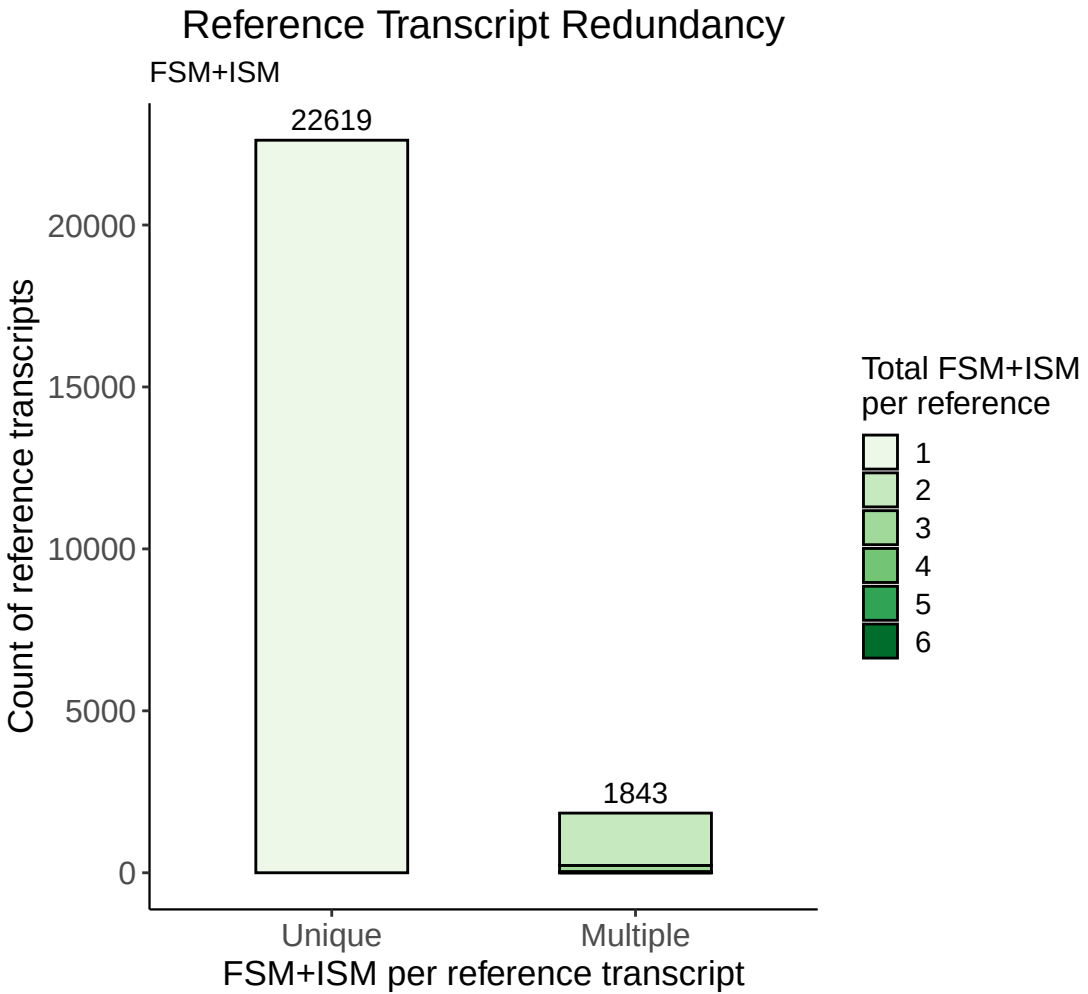
Number of CAGE Detected

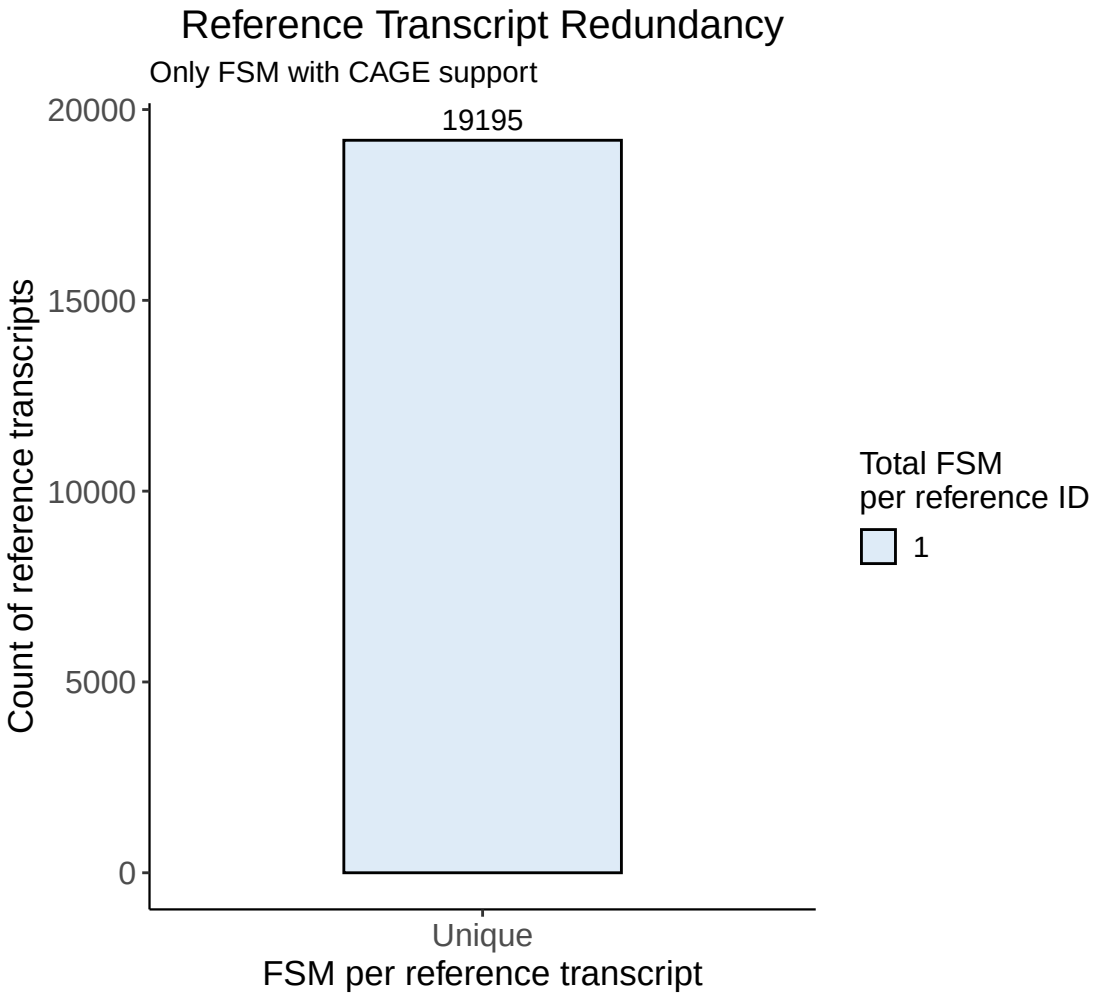
Subcategory	Count	CAGE Detected	%
Alternative 3'end	2247	1882	84
Alternative 3'5'end	1020	814	80
Alterantive 5'end	1259	912	72
Reference match	18624	15587	84
3' fragment	857	369	43
Internal fragment	133	56	42
5' fragment	1589	1357	85
Comb. of annot. junctions	9177	8083	88
Comb. of annot. splice sites	8068	7310	91
Intron retention	8125	7129	88
Mono−exon by intron ret.	41	29	71
At least 1 annot. don./accept.	18119	15286	84
Mono−exon	4897	1847	38
Multi−exon	2861	1632	57

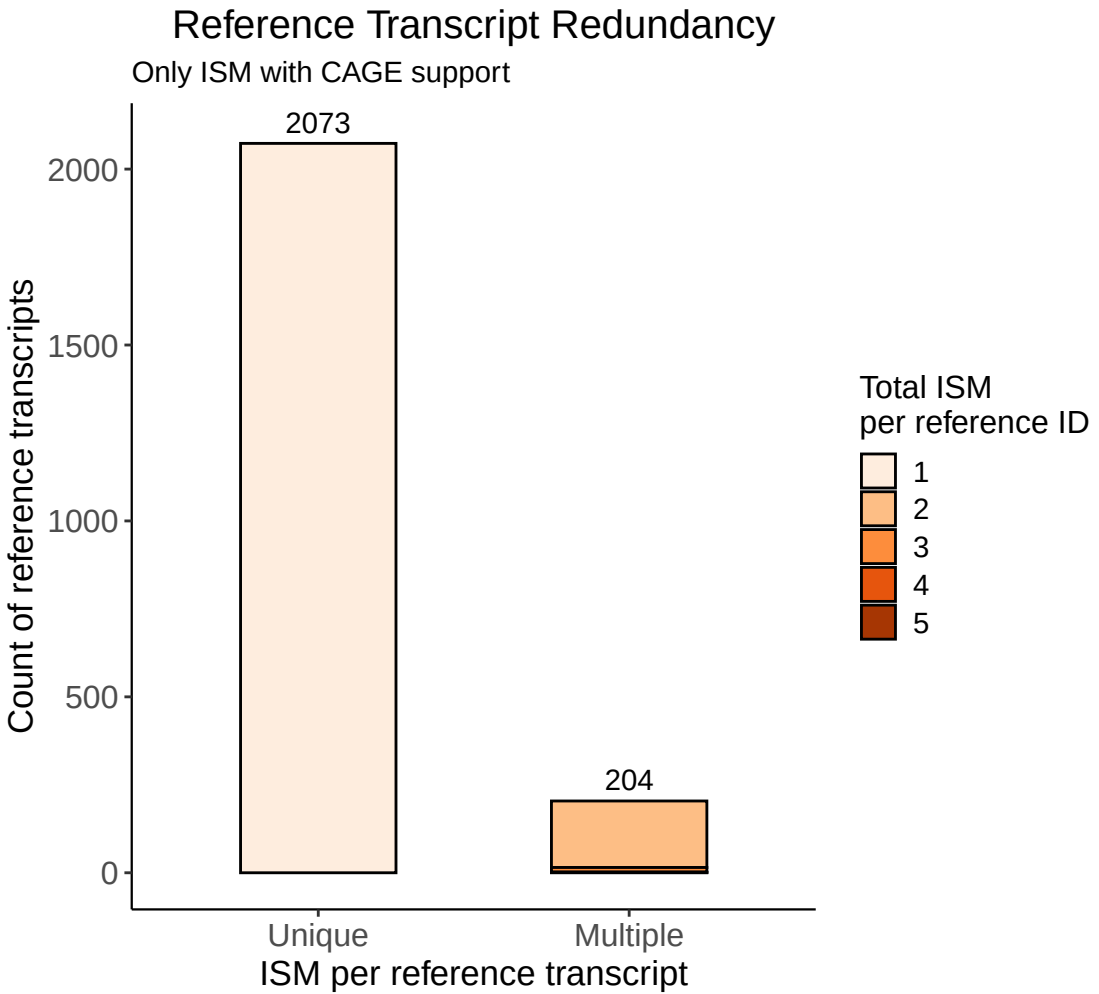
Redundancy Analysis

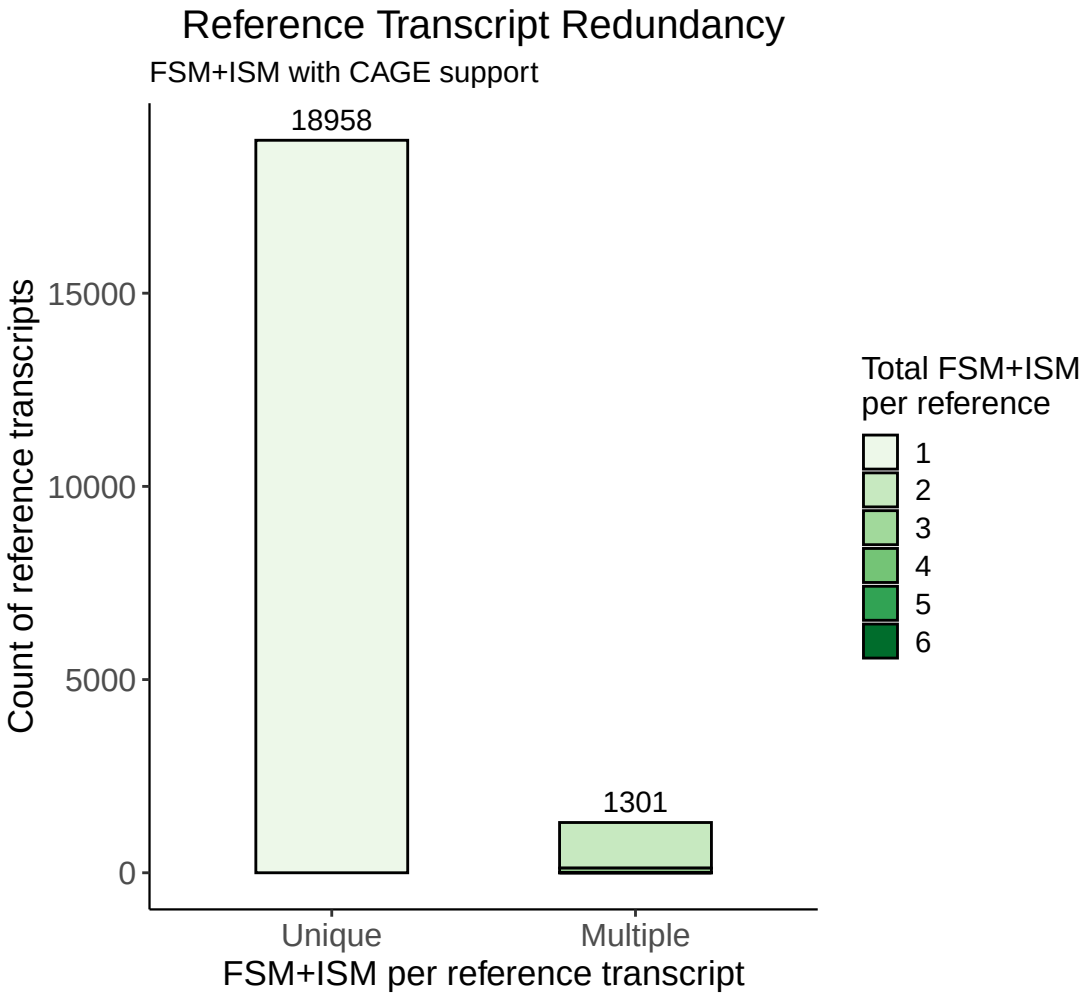

Reference Transcript Redundancy
Only FSM
23150
20000
15000
10000
5000
0
Count of reference transcripts
Unique
FSM per reference transcript
Total FSM
per reference ID
1

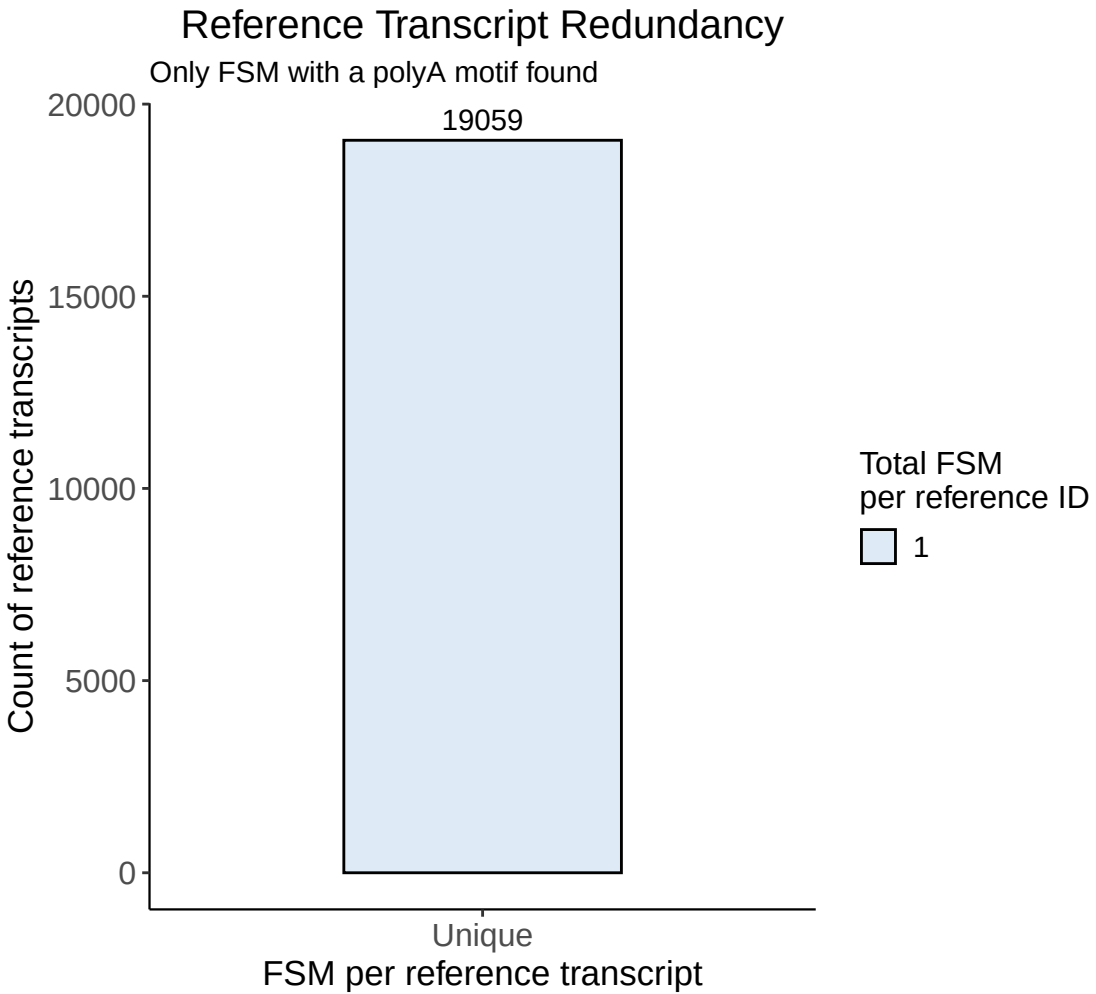

Reference Transcript Redundancy
Only ISM
2672
347
2500
2000
1500
1000
500
0
Count of Reference Transcripts
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5

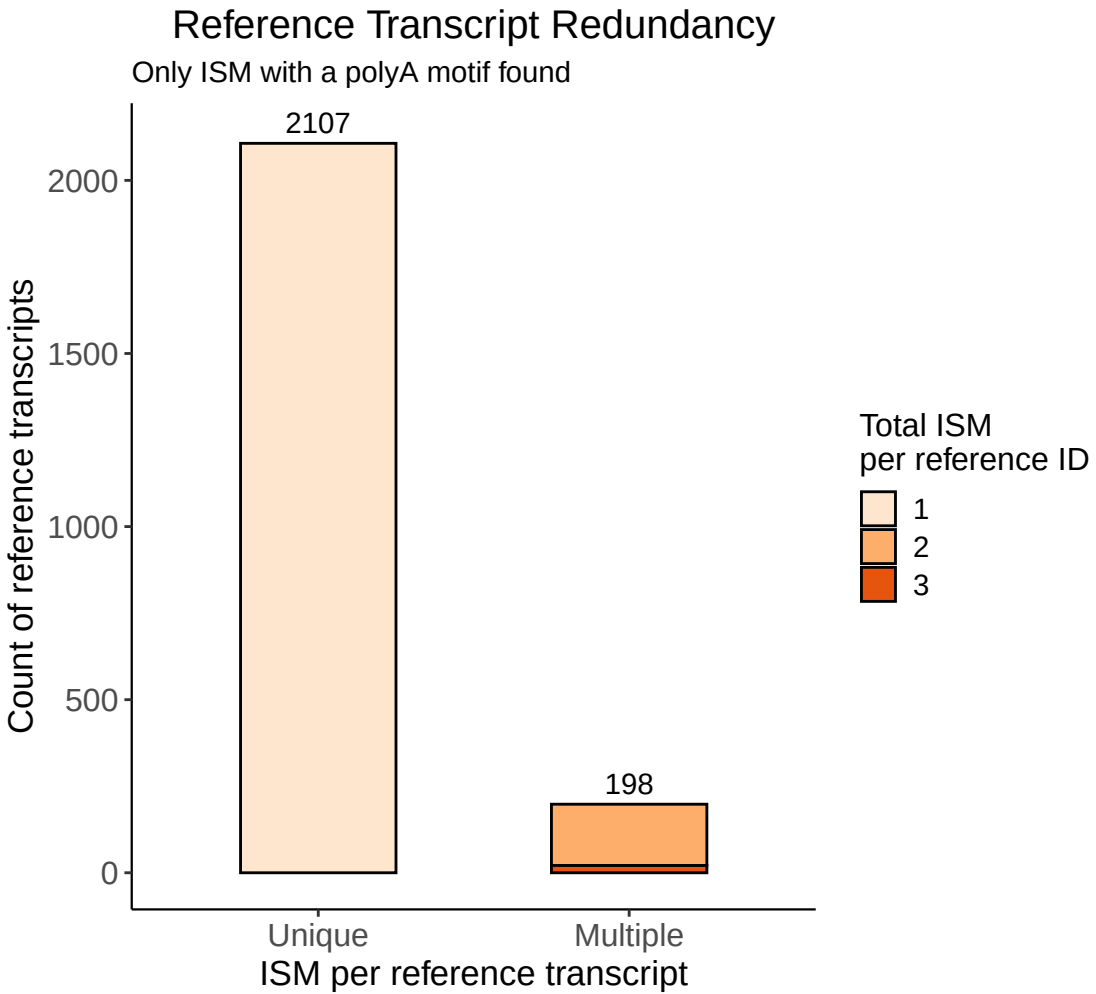

Reference Transcript Redundancy
FSM+ISM
22619
1843
20000
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6

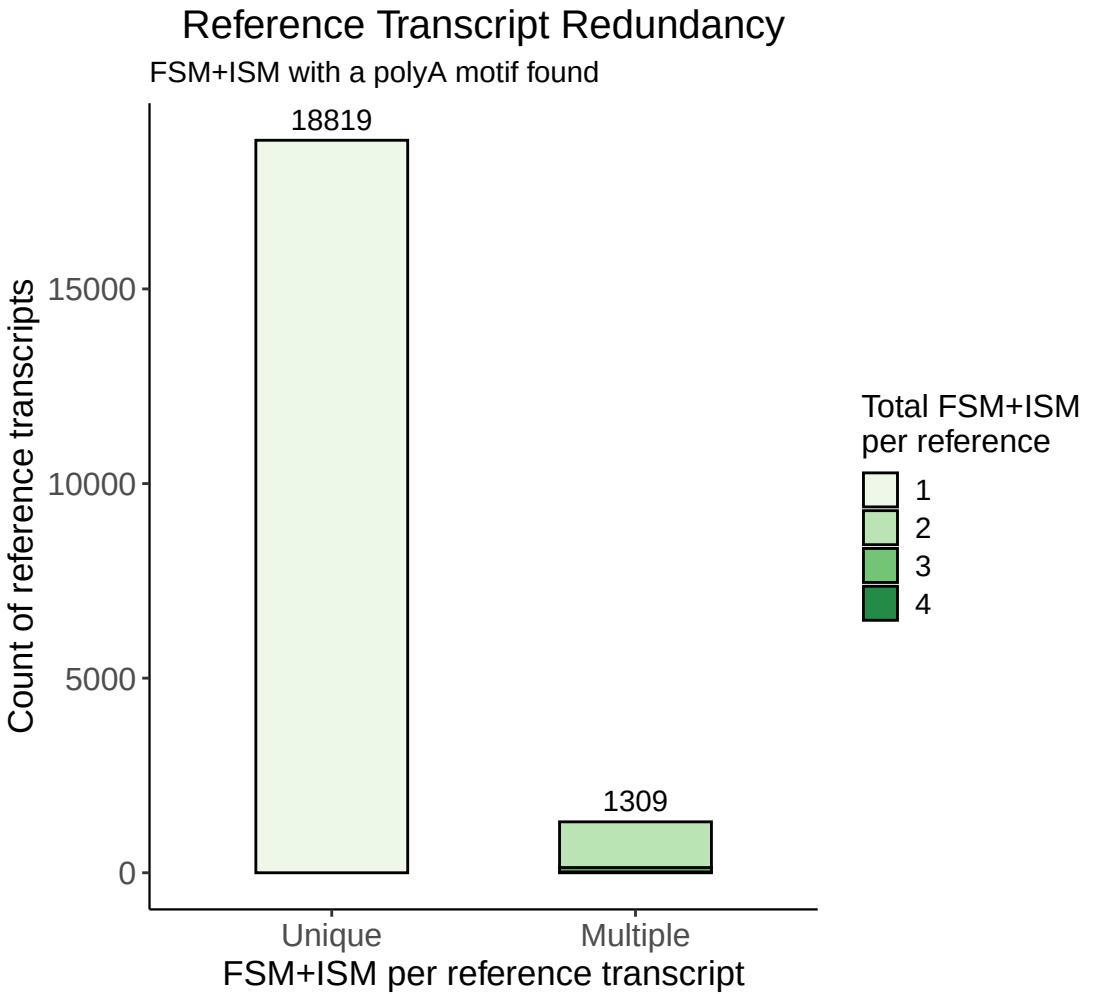

Reference Transcript Redundancy
Only FSM with CAGE support
19195
20000
15000
10000
5000
0
Count of reference transcripts
Unique
FSM per reference transcript
Total FSM
per reference ID
1

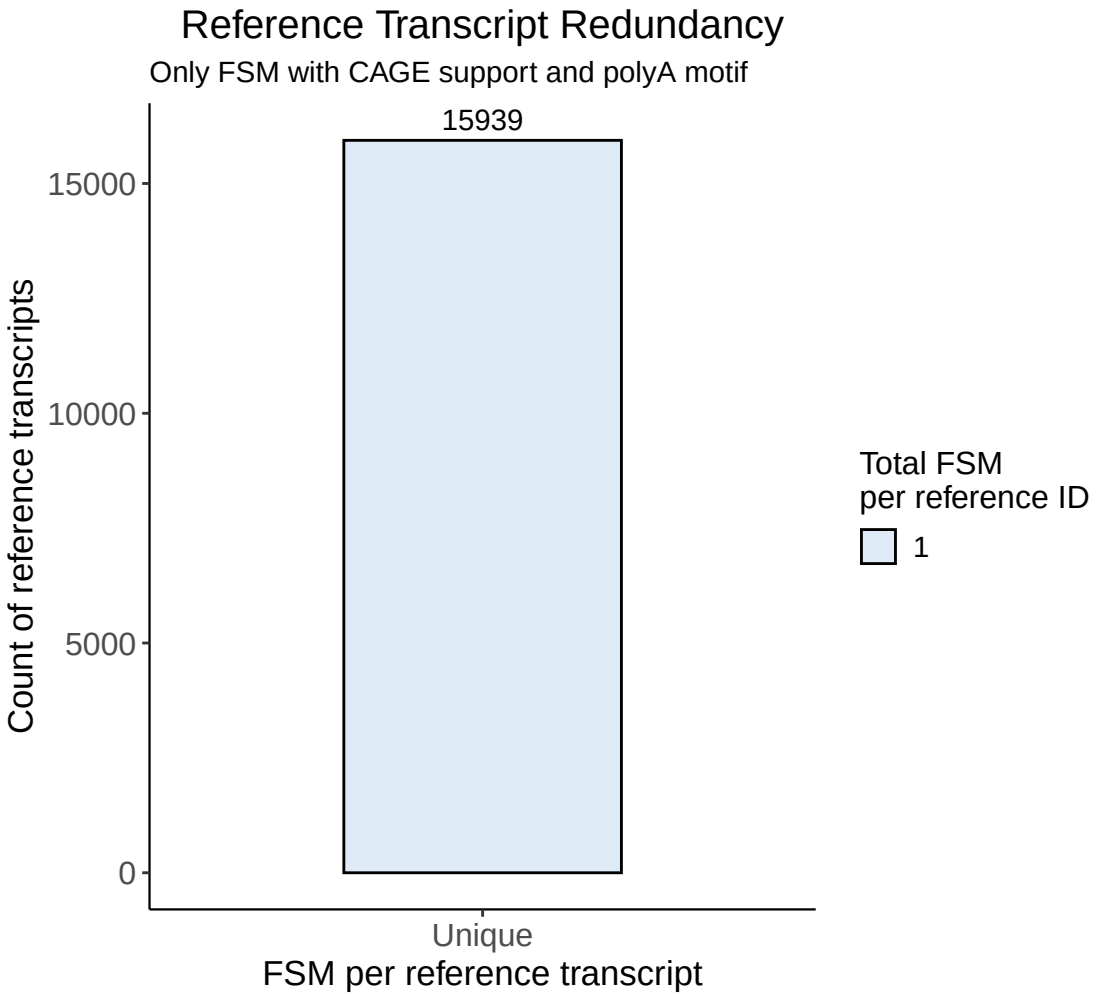

Reference Transcript Redundancy
Only ISM with CAGE support
2073
204
2000
1500
1000
500
0
Count of reference transcripts
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5

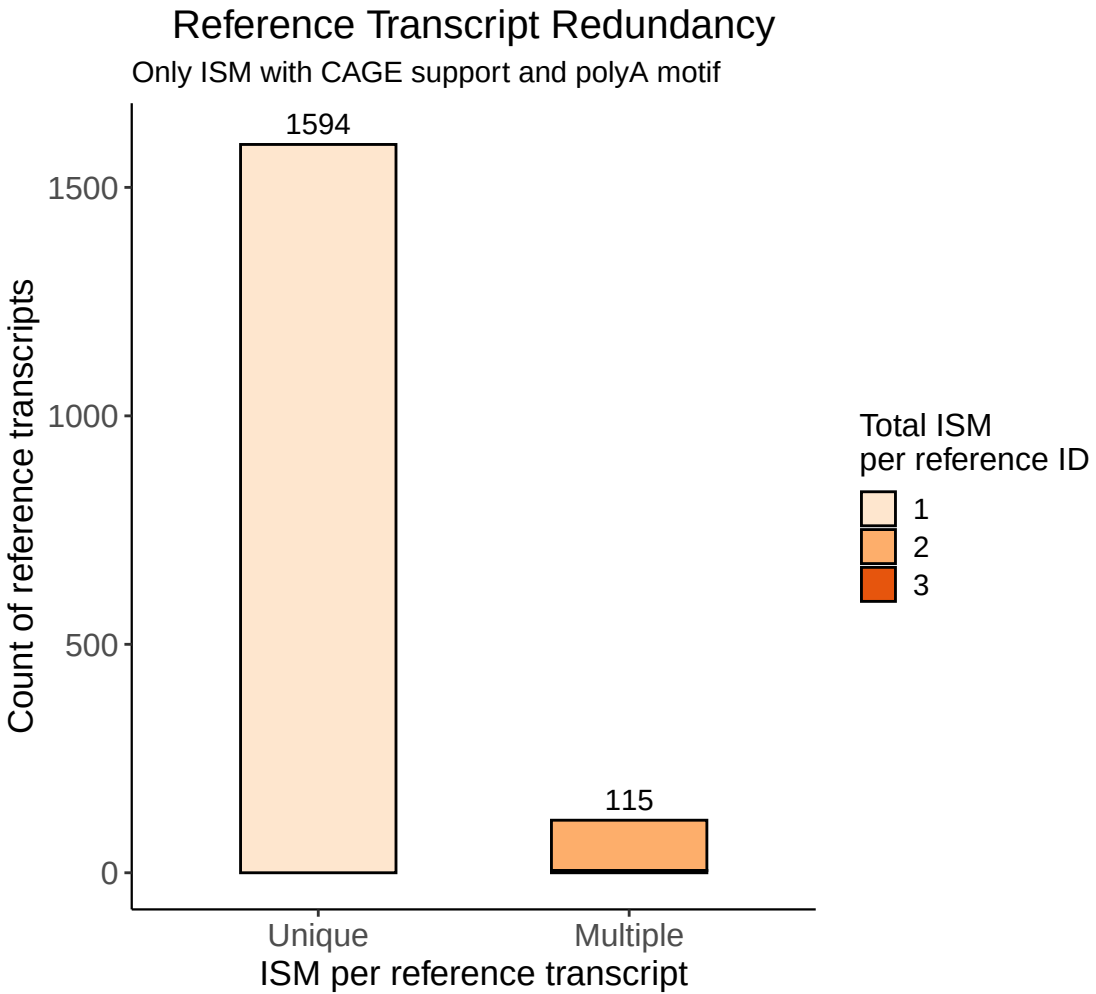

Reference Transcript Redundancy
FSM+ISM with CAGE support
18958
1301
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6

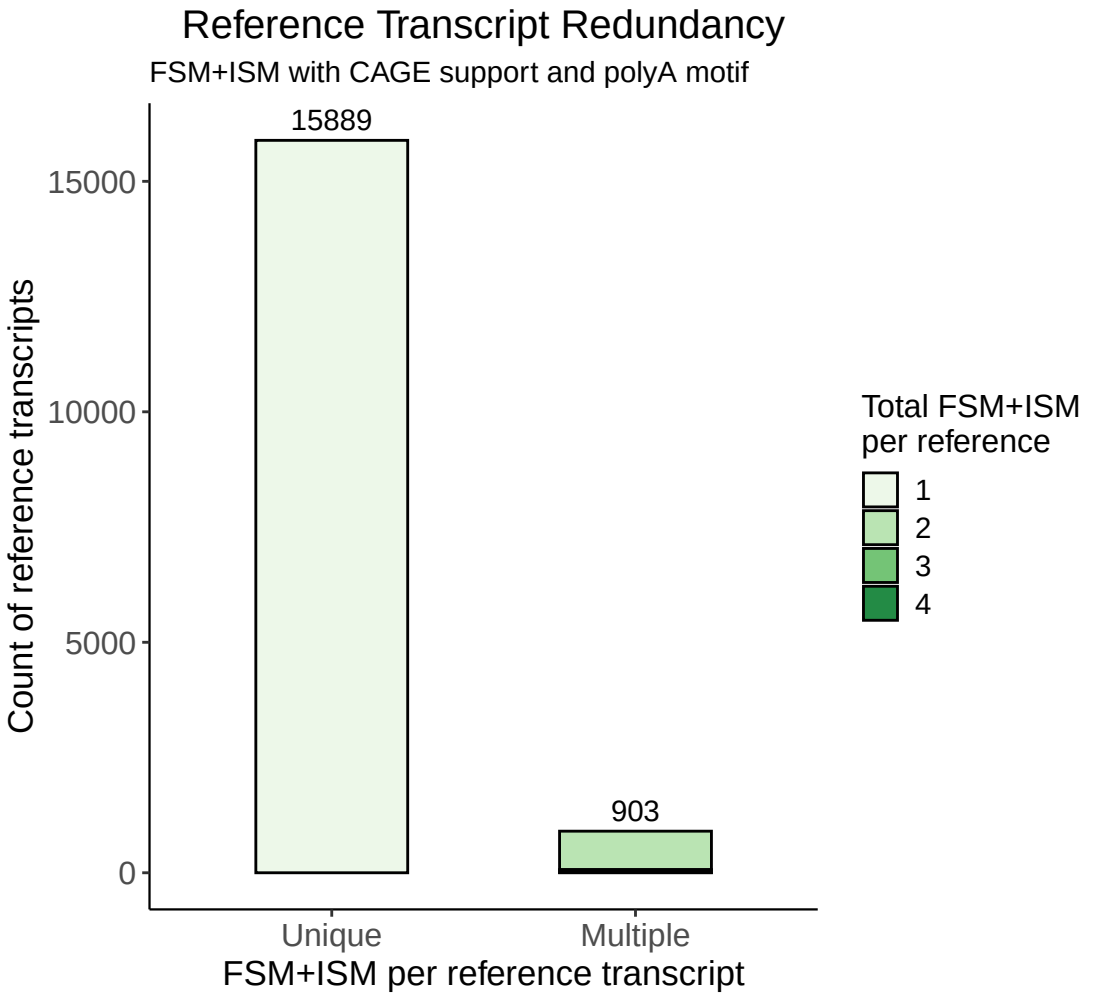

Reference Transcript Redundancy
Only FSM with a polyA motif found
19059
20000
15000
10000
5000
0
Count of reference transcripts
Unique
FSM per reference transcript
Total FSM
per reference ID
1


Reference Transcript Redundancy
Only ISM with a polyA motif found
2107
198
2000
1500
1000
500
0
Count of reference transcripts
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3


Reference Transcript Redundancy
FSM+ISM with a polyA motif found
18819
1309
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4

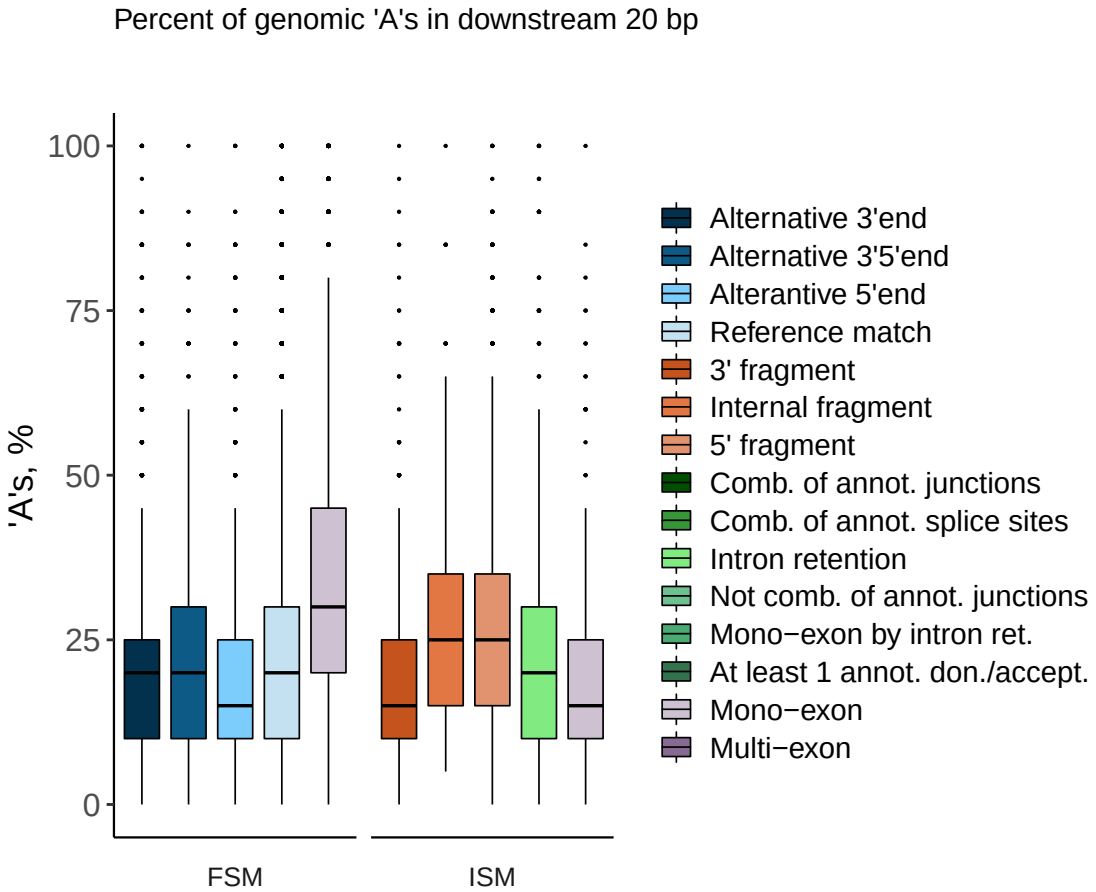

Reference Transcript Redundancy
Only FSM with CAGE support and polyA motif
15939
15000
10000
5000
0
Count of reference transcripts
Unique
FSM per reference transcript
Total FSM
per reference ID
1


Reference Transcript Redundancy
Only ISM with CAGE support and polyA motif
1594
115
1500
1000
500
0
Count of reference transcripts
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3

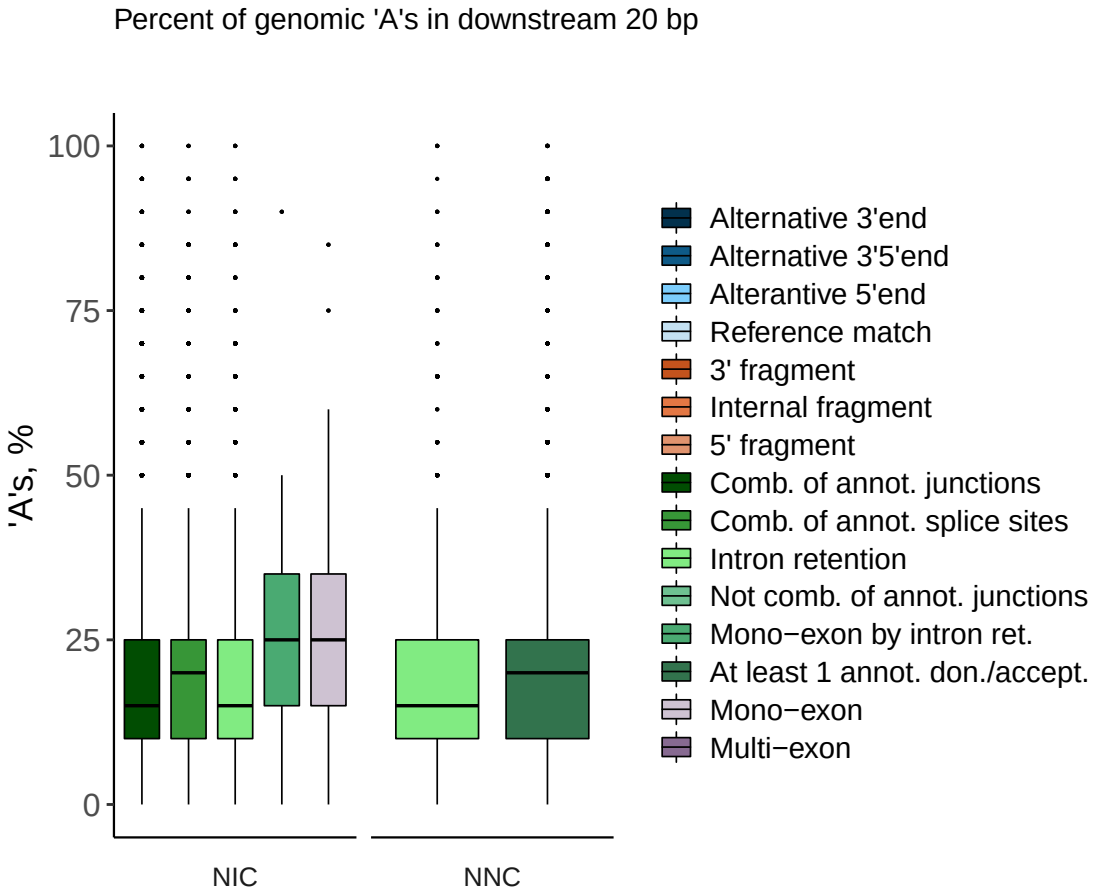

Reference Transcript Redundancy
FSM+ISM with CAGE support and polyA motif
15889
903
15000
10000
5000
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4

Intra−Priming Quality Check

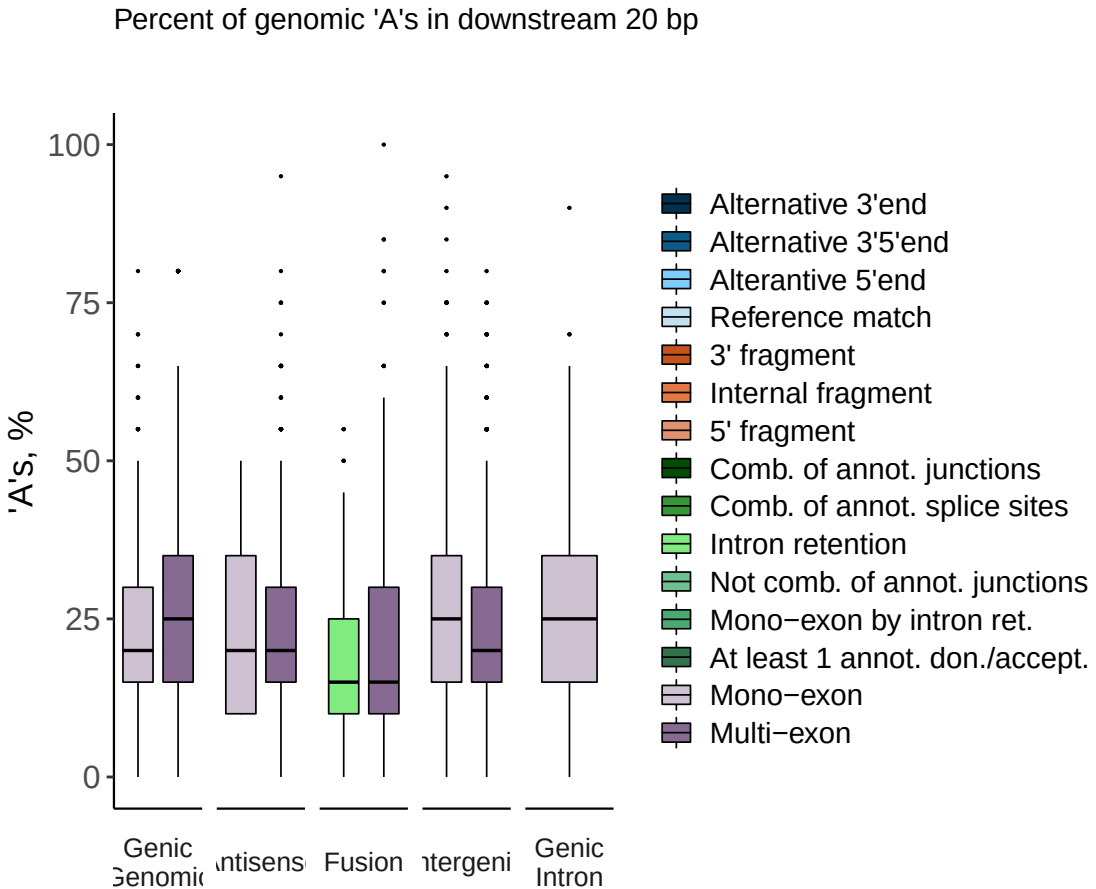
Possible Intra−Priming by Structural Category

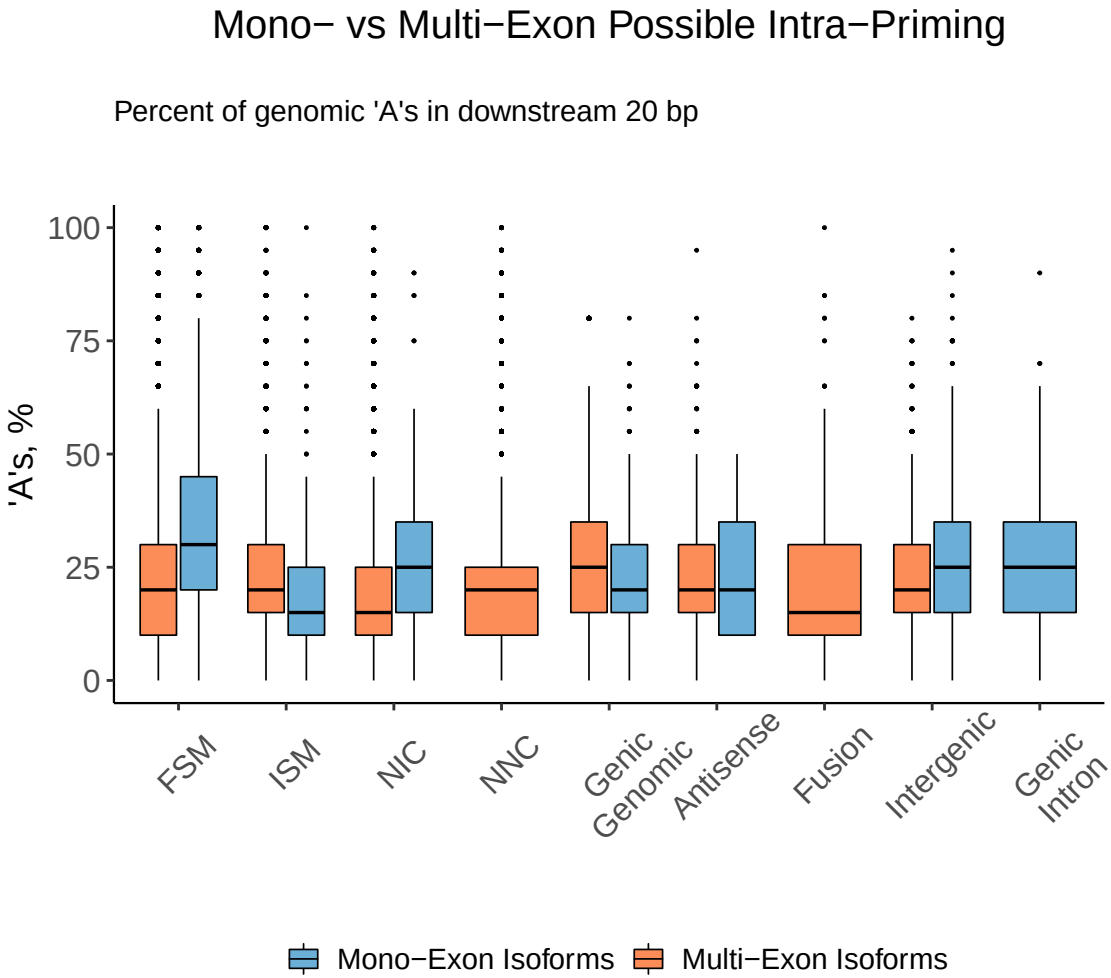


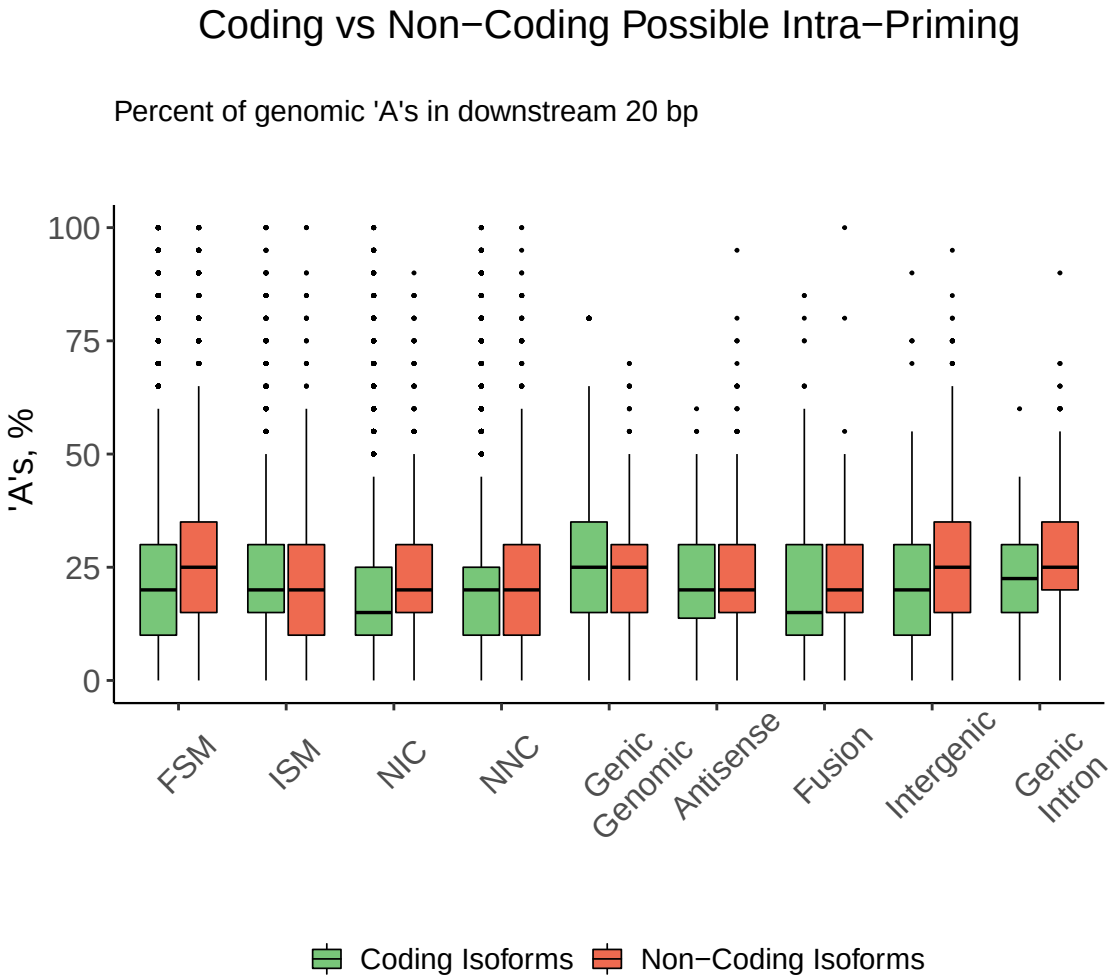
Possible Intra−Priming by Structural Category



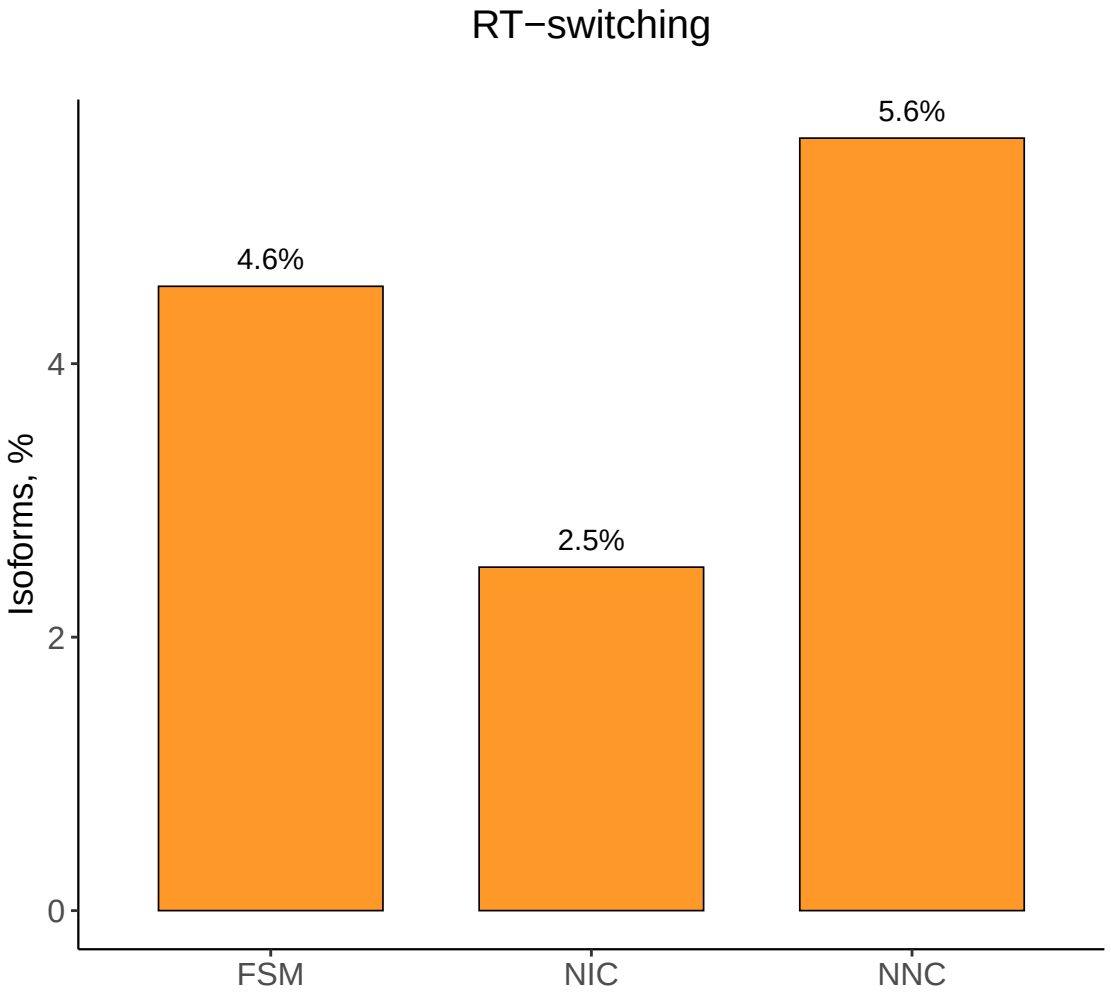
Possible Intra−Priming by Structural Category

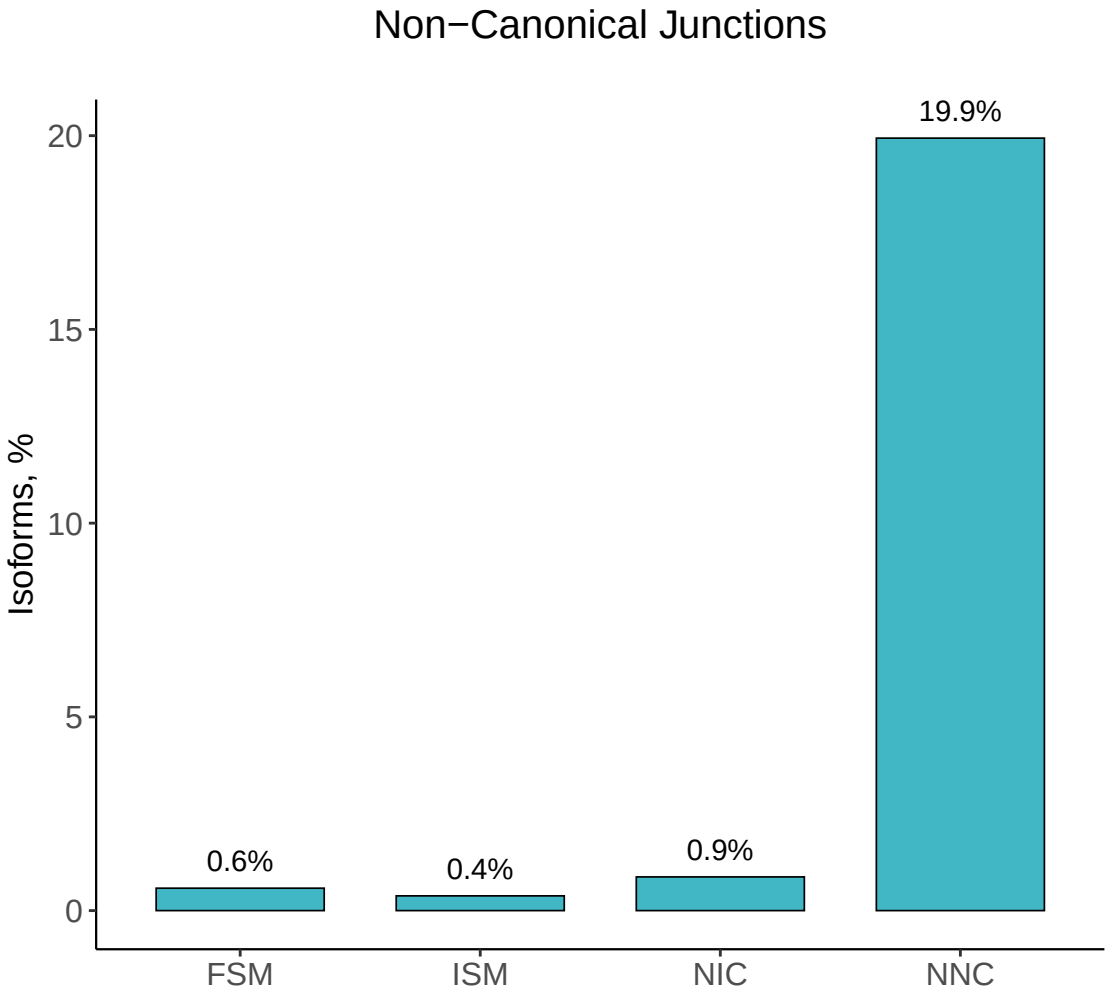


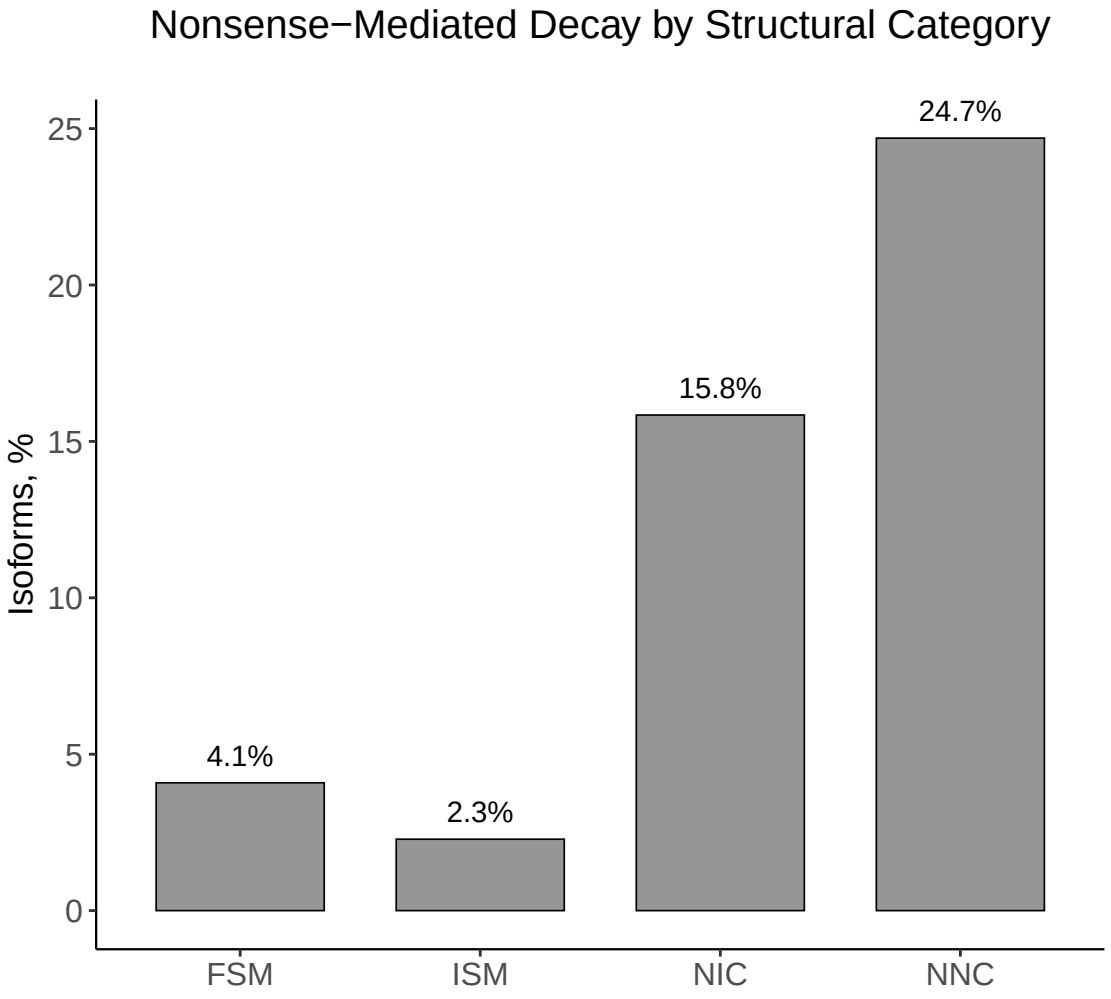

Mono− vs Multi−Exon Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Mono−Exon Isoforms
Multi−Exon Isoforms

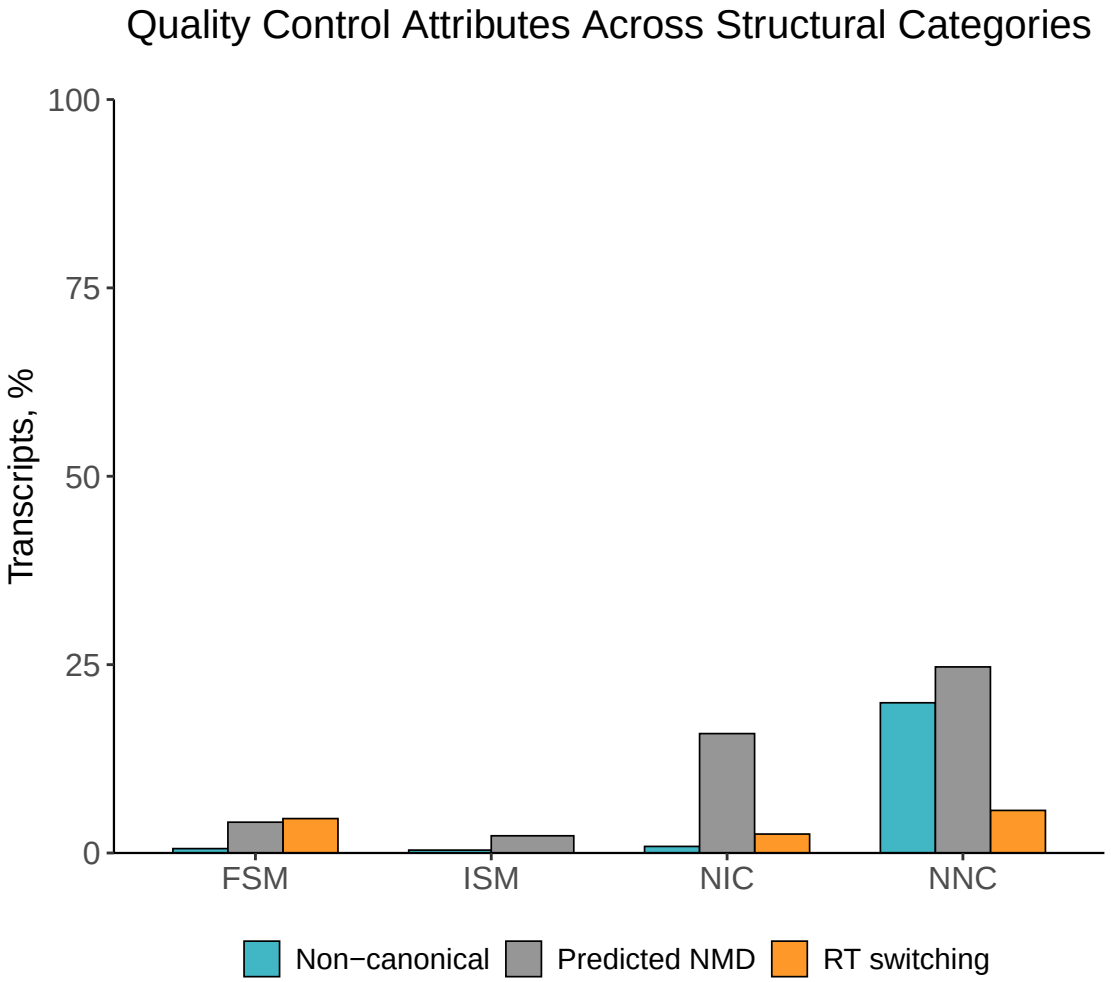

Coding vs Non−Coding Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Coding Isoforms
Non−Coding Isoforms

Features of Bad Quality

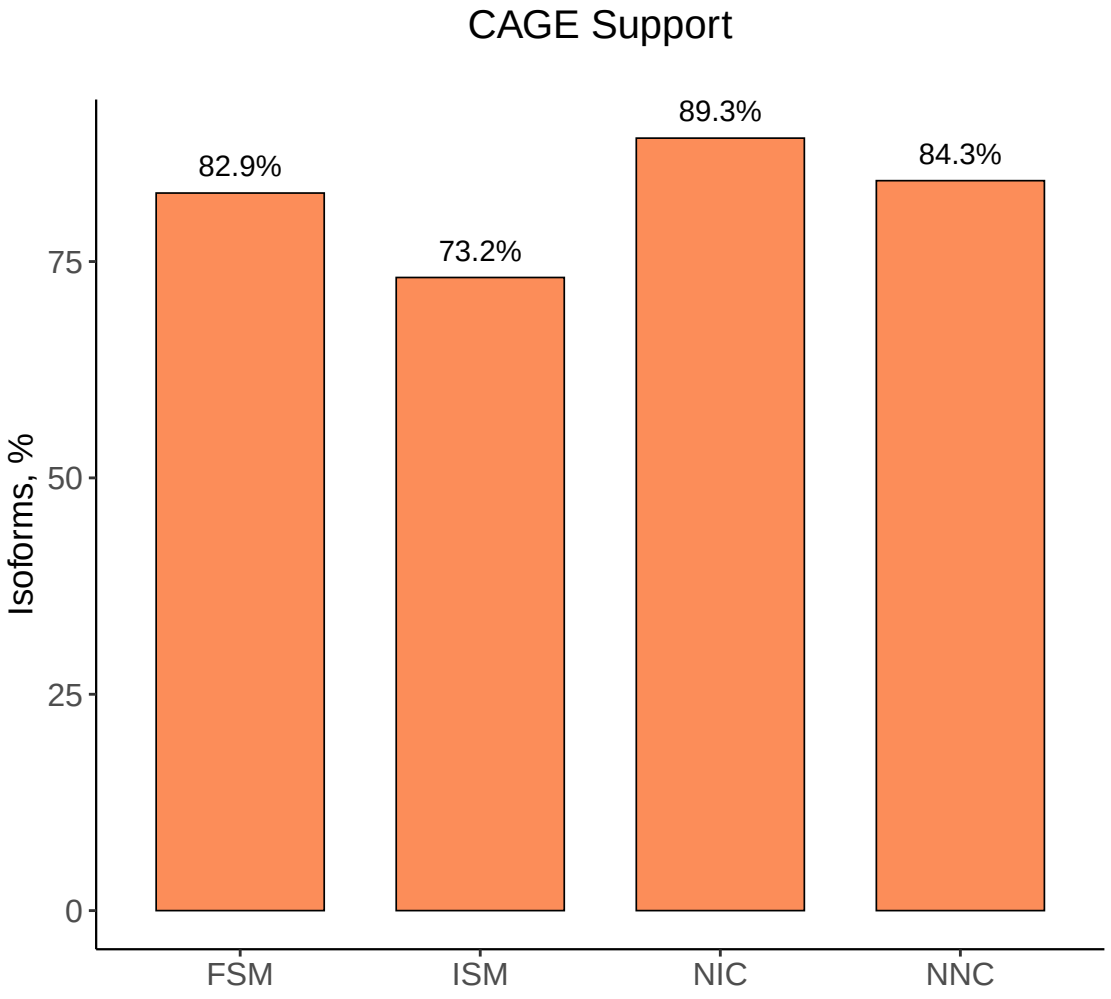

RT−switching
Isoforms, %
0
2
4
4.6%
2.5%
5.6%
FSM
NIC
NNC

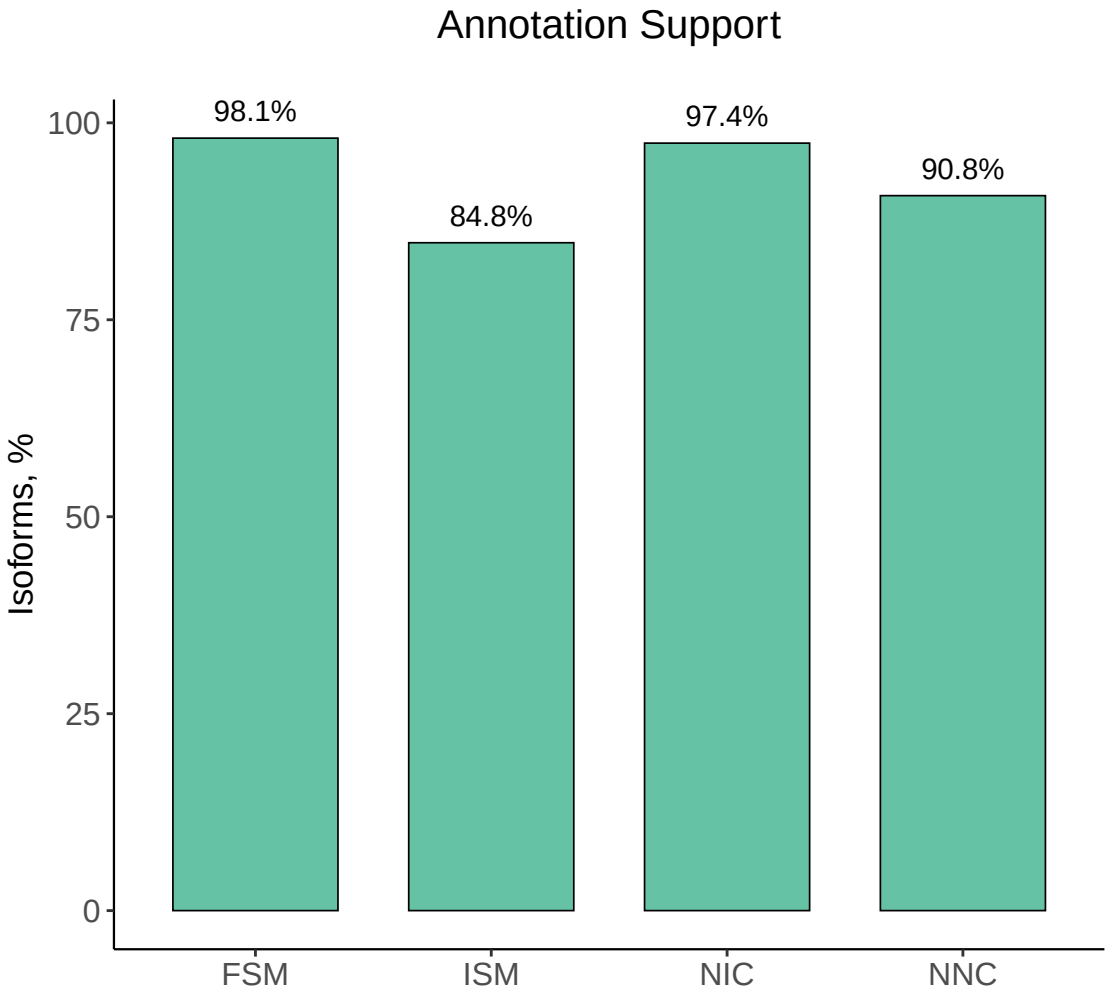

Non−Canonical Junctions
Isoforms, %
20
15
10
5
0
19.9%
0.6%
0.4%
0.9%
FSM
ISM
NIC
NNC

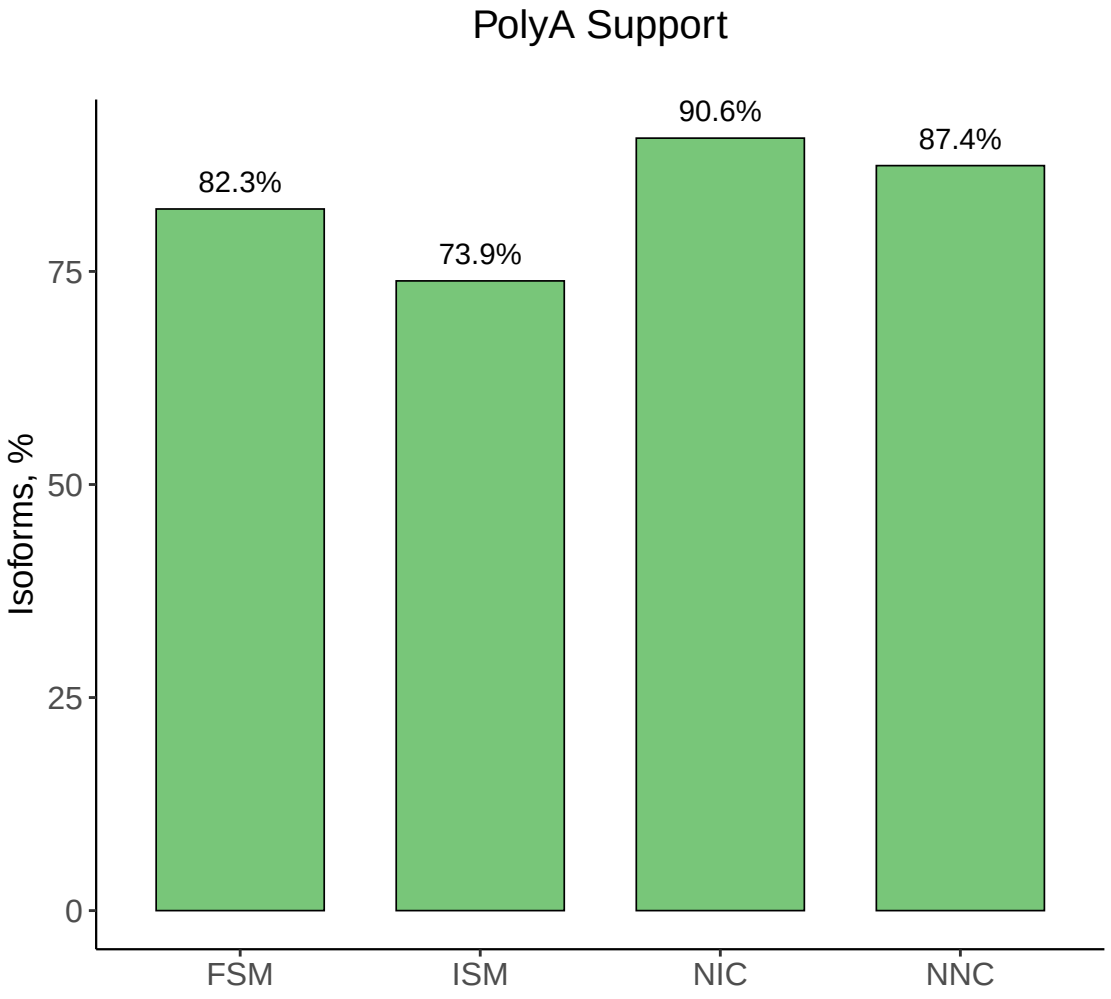

Nonsense−Mediated Decay by Structural Category
Isoforms, %
25
20
15
10
5
0
4.1%
2.3%
15.8%
24.7%
FSM
ISM
NIC
NNC

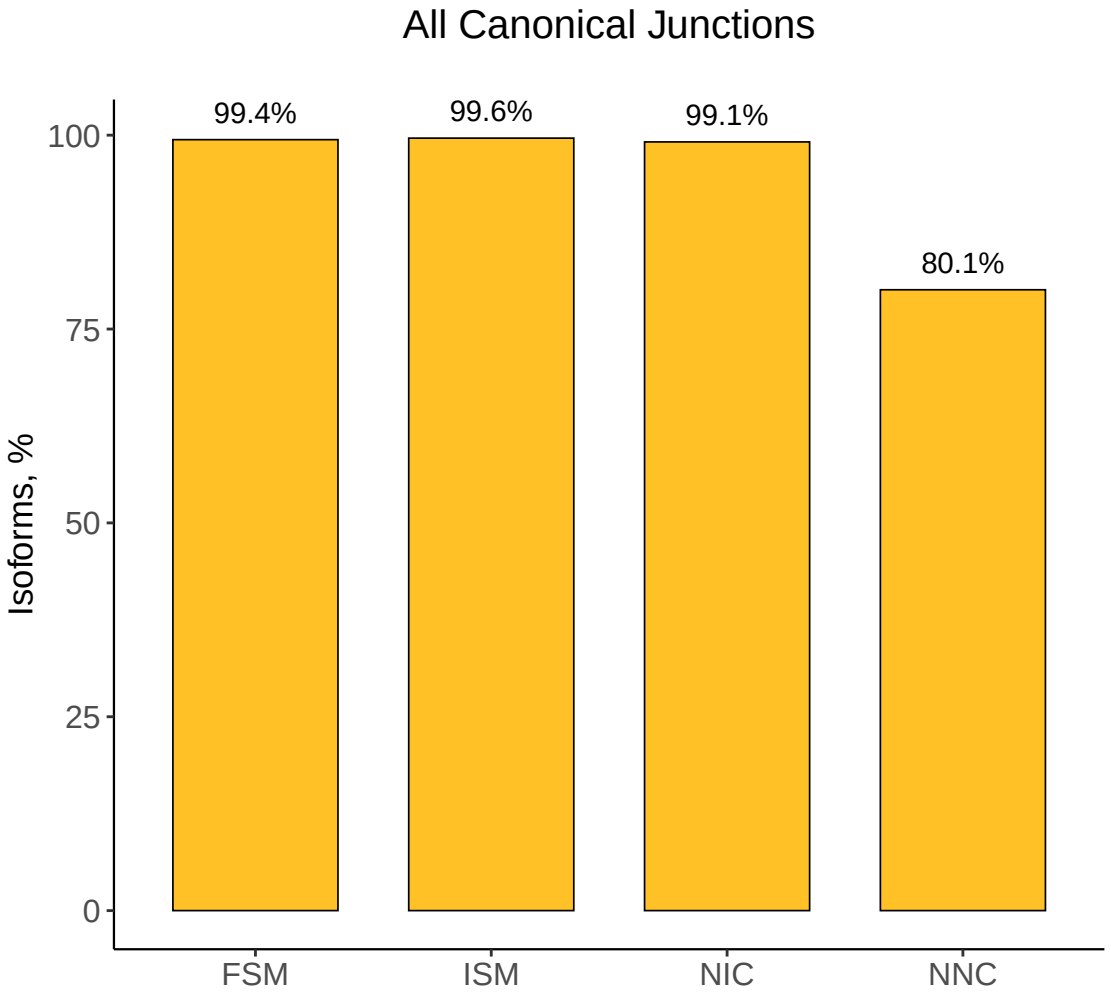

Quality Control Attributes Across Structural Categories
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Non−canonical
Predicted NMD
RT switching

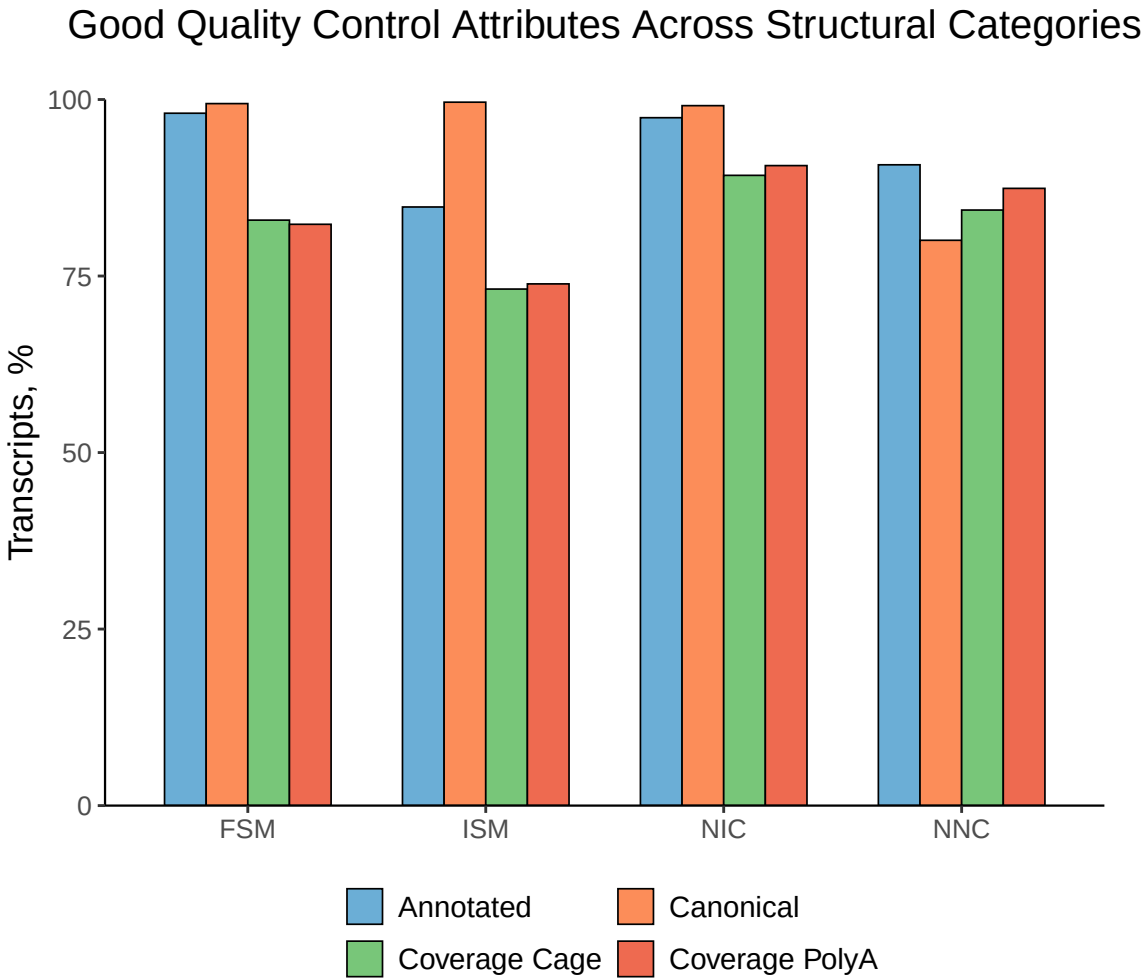
Features of Good Quality


CAGE Support
Isoforms, %
0
25
50
75
82.9%
73.2%
89.3%
84.3%
FSM
ISM
NIC
NNC


Annotation Support
Isoforms, %
100
75
50
25
0
98.1%
84.8%
97.4%
90.8%
FSM
ISM
NIC
NNC


PolyA Support
Isoforms, %
0
25
50
75
82.3%
73.9%
90.6%
87.4%
FSM
ISM
NIC
NNC


All Canonical Junctions
99.4%
99.6%
99.1%
80.1%
100
75
50
25
0
Isoforms, %
FSM
ISM
NIC
NNC


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