SQANTI3 report

Unique Genes: 25112

Unique Isoforms: 656703

Transcript Classification

Category	Isoforms, count
FSM	151552
ISM	232135
NIC	143517
NNC	119532
Genic Genomic	819
Antisense	2534
Fusion	3334
Intergenic	3212
Genic Intron	0

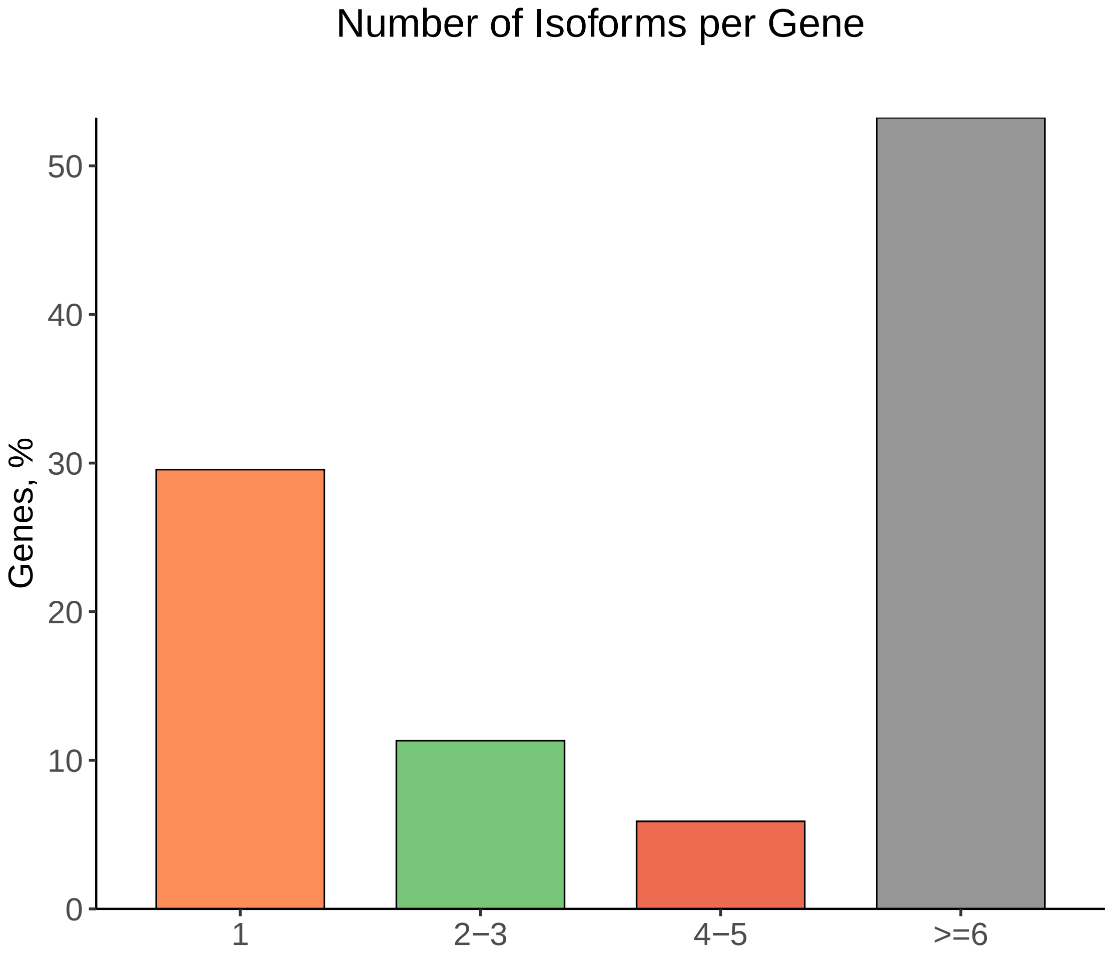
Gene Classification

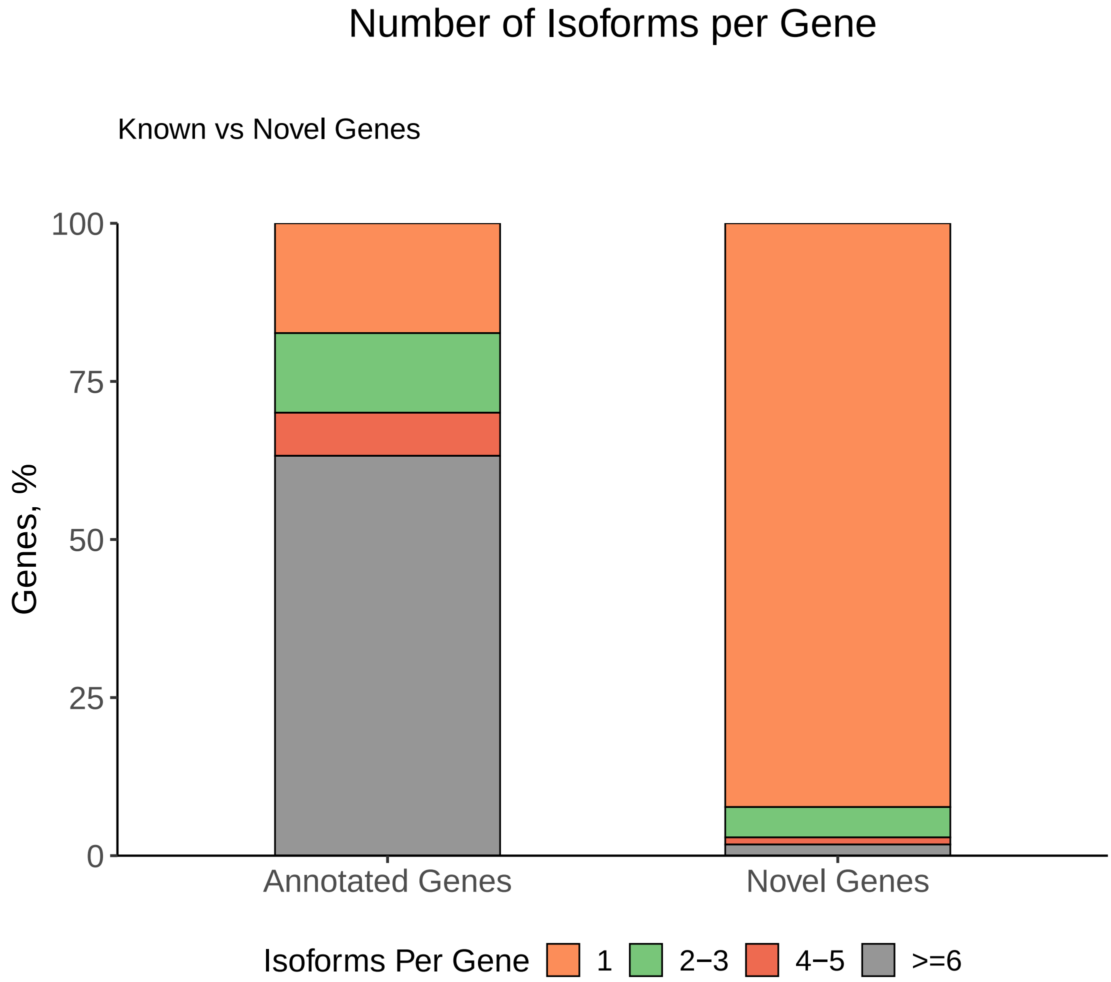
Category	Genes, count
Annotated Genes	21027
Novel Genes	4085

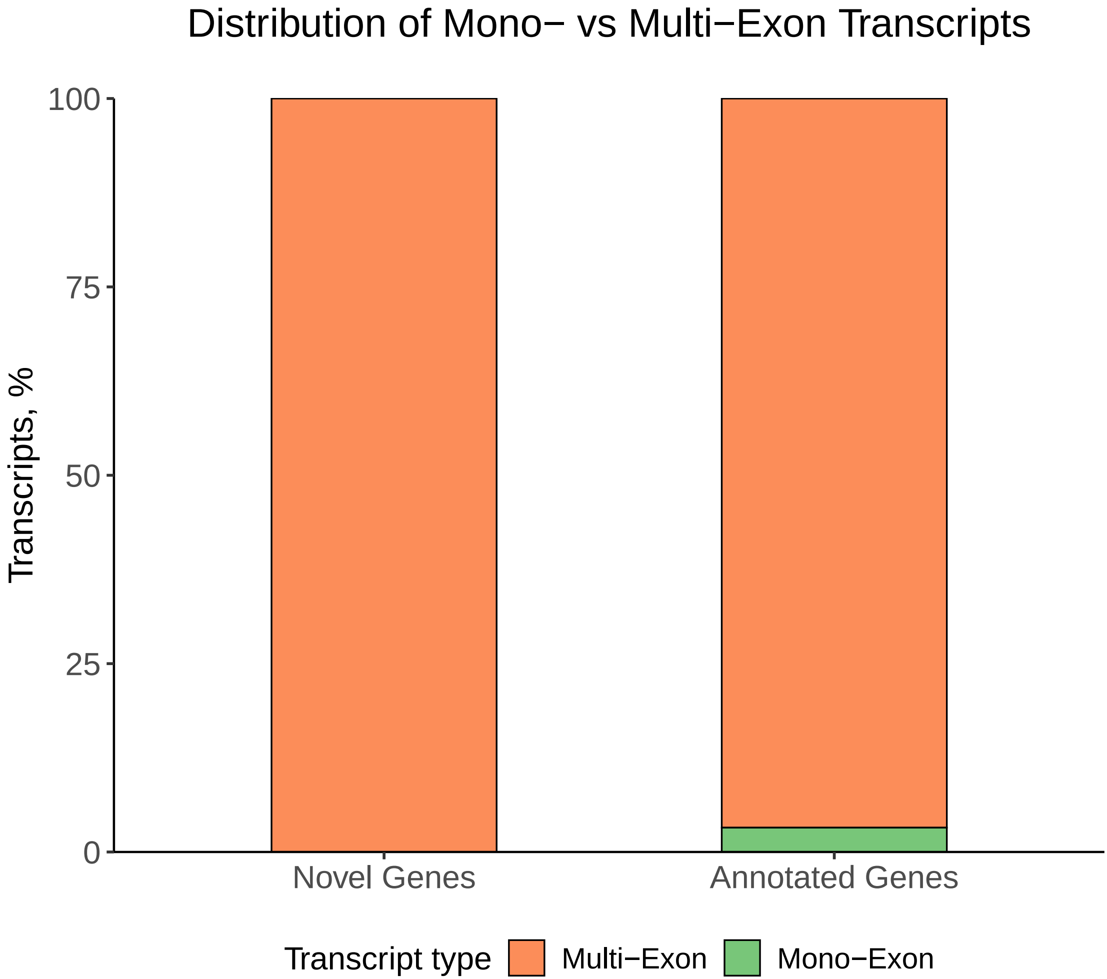
Splice Junction Classification

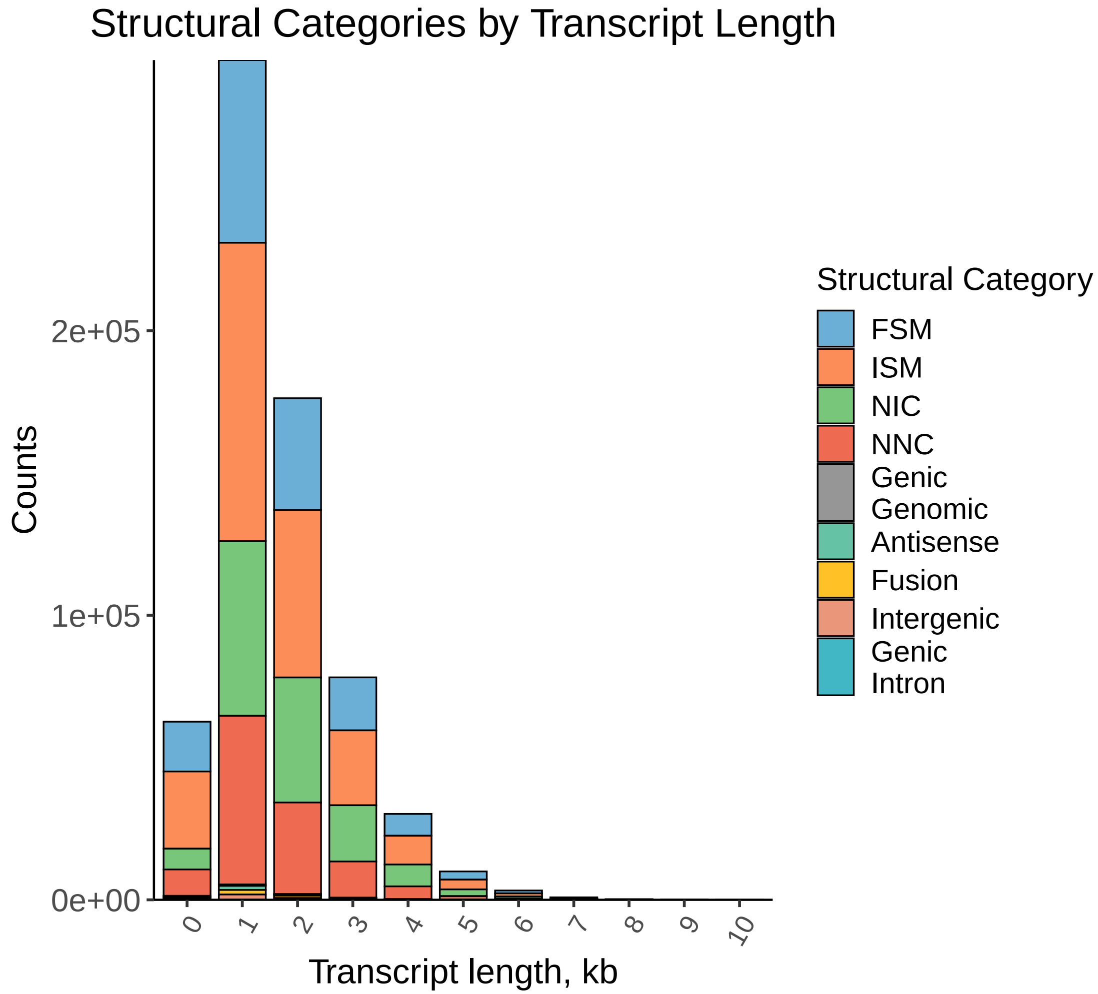
Category	SJs, count	Percent
Known canonical	193264	61.90
Known Non−canonical	88	0.03
Novel canonical	118887	38.08
Novel Non−canonical	0	0.00

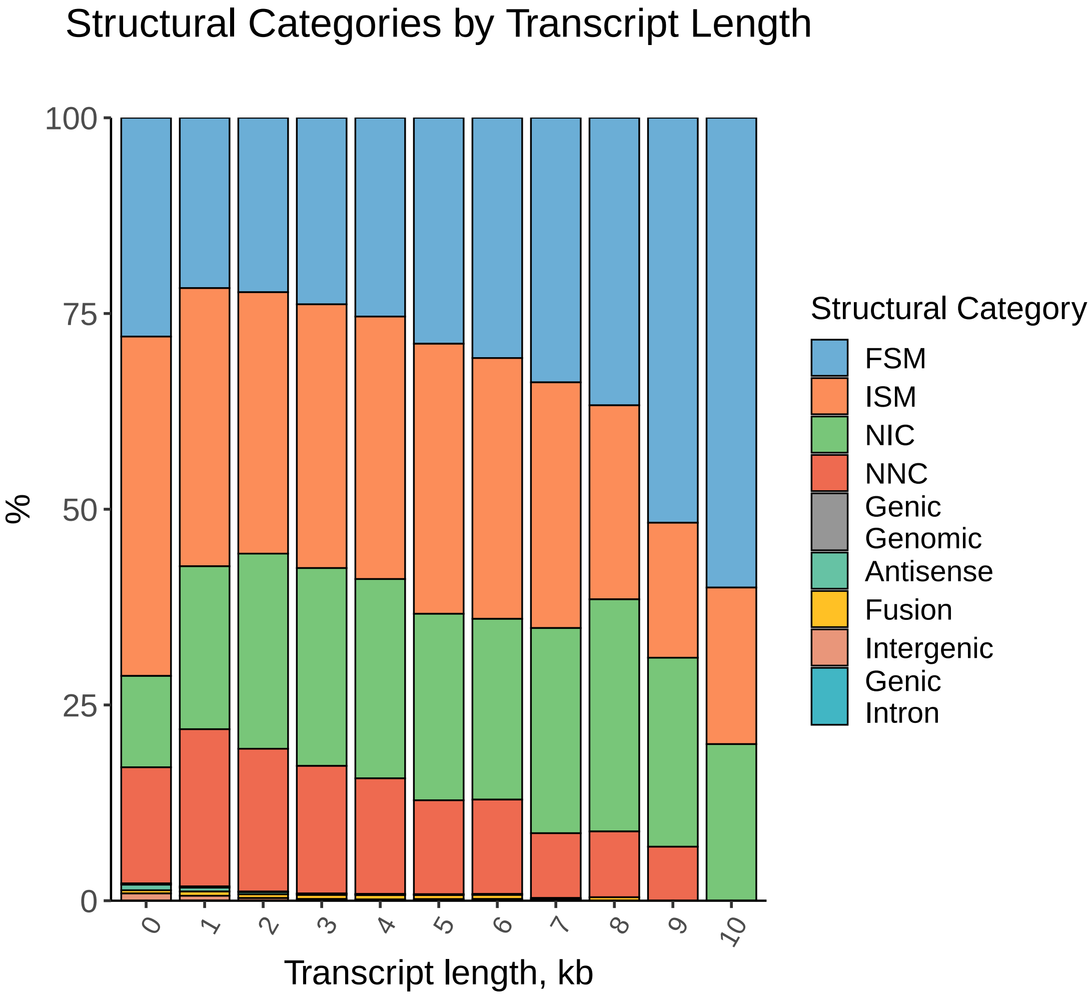
Gene Characterization

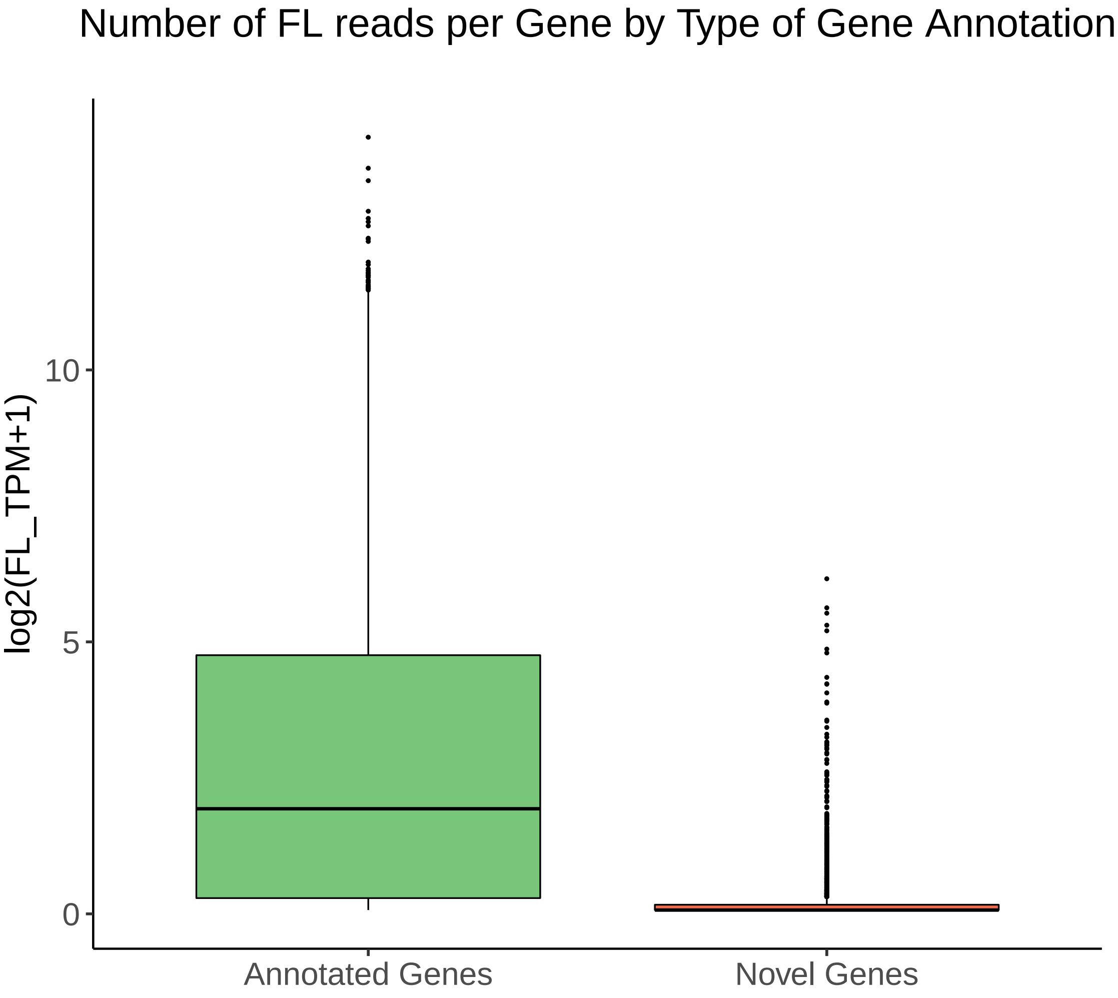

Number of Isoforms per Gene
Genes, %
0
10
20
30
40
50
1
2−3
4−5
>=6
Isoforms per gene

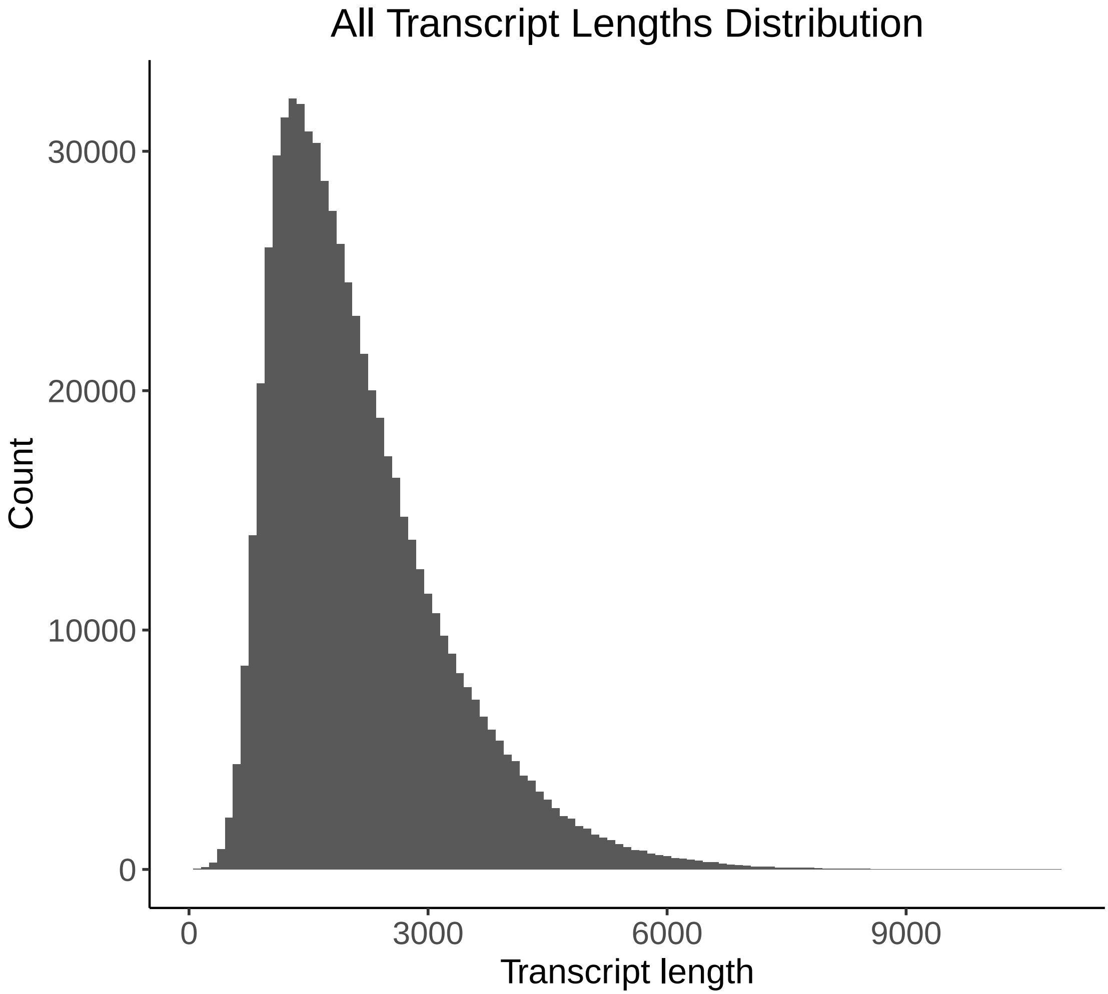

Number of Isoforms per Gene
Known vs Novel Genes
Genes, %
100
75
50
25
0
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3
4−5
>=6

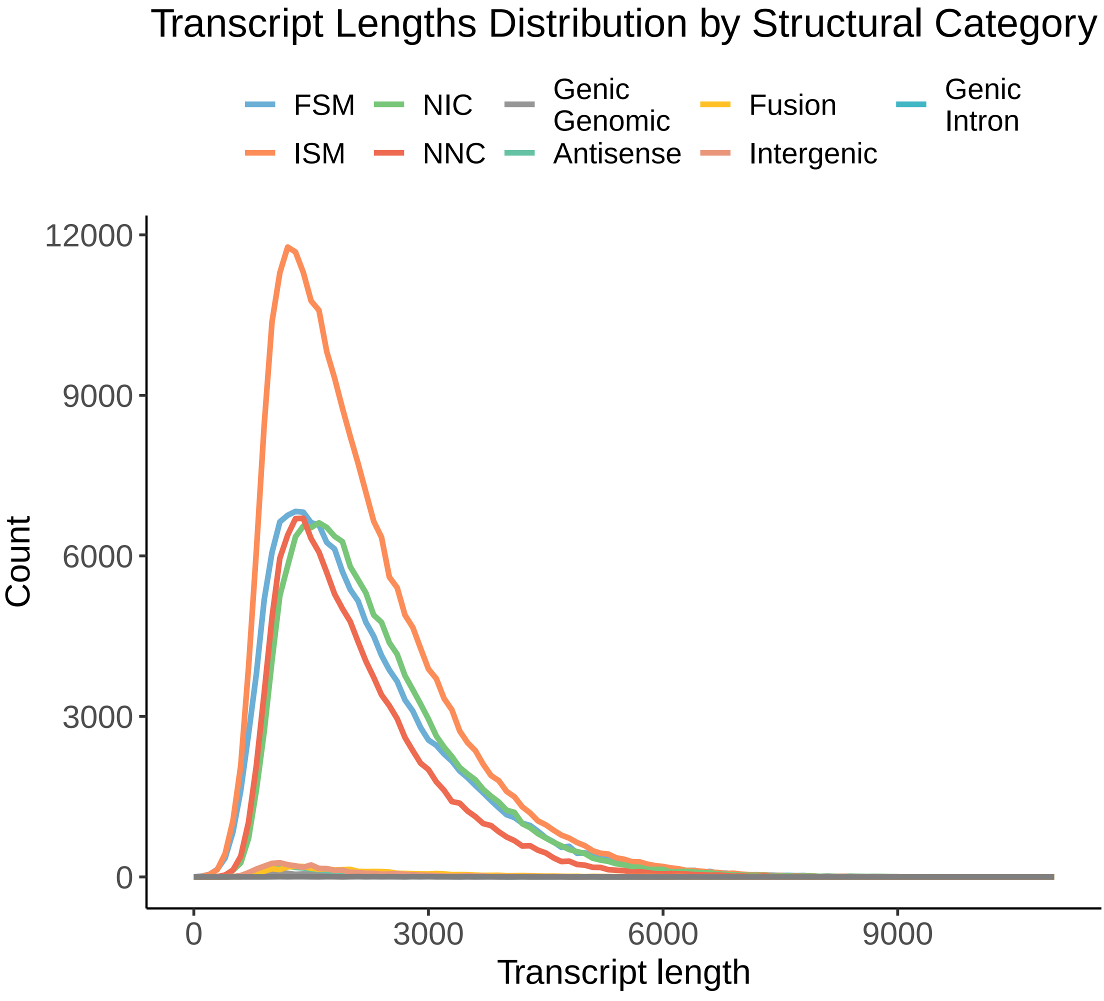

Distribution of Mono− vs Multi−Exon Transcripts
Transcripts, %
100
75
50
25
0
Novel Genes
Annotated Genes
Transcript type
Multi−Exon
Mono−Exon

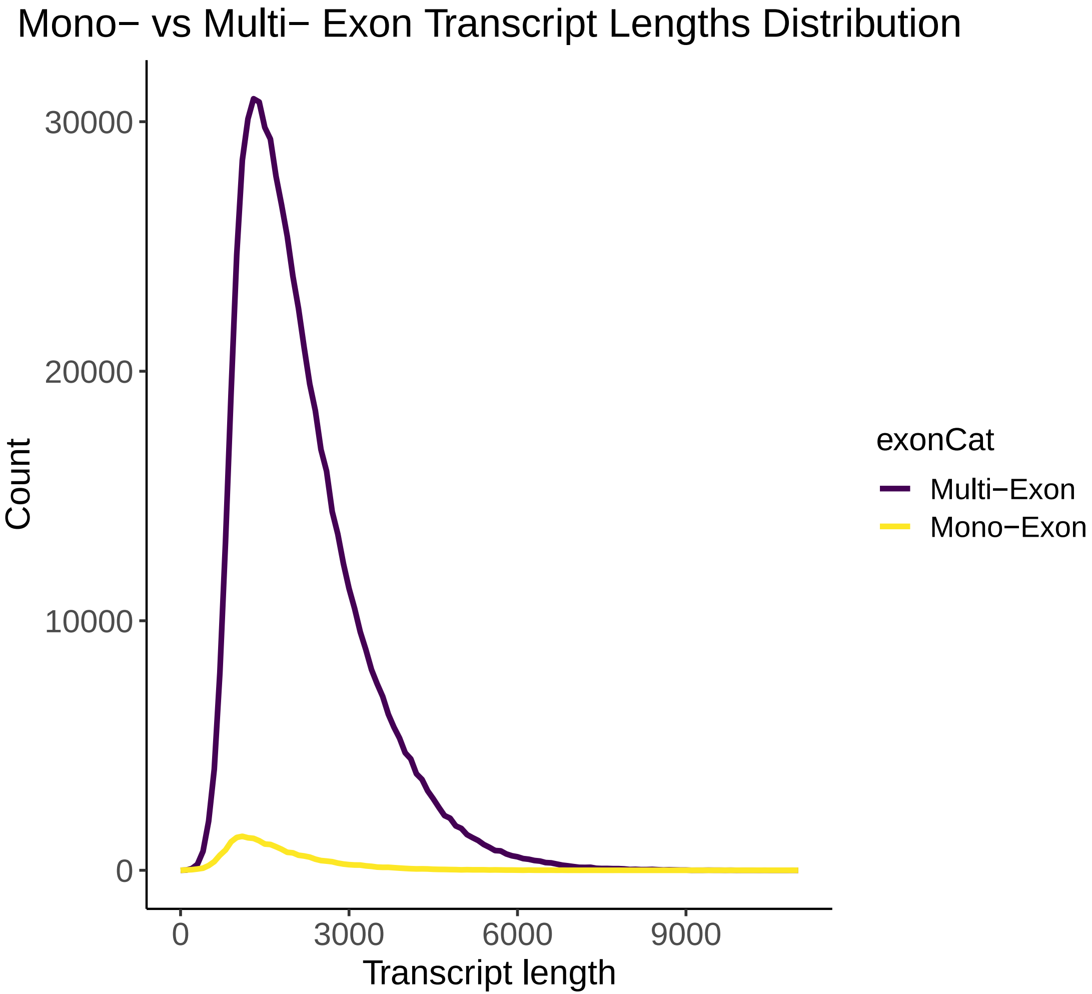

Structural Categories by Transcript Length
Counts
2e+05
1e+05
0e+00
0
1
2
3
4
5
6
7
8
9
10
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

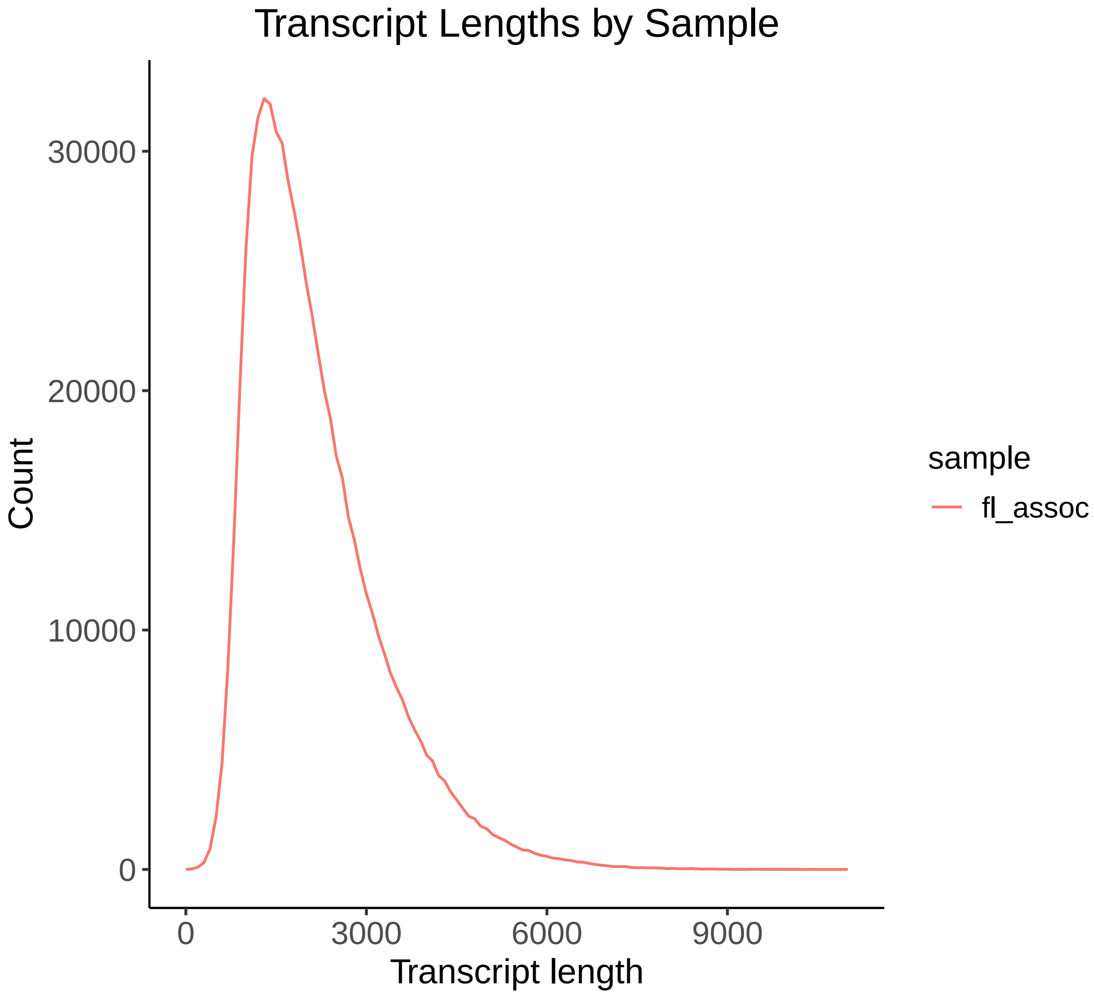

Structural Categories by Transcript Length
100
75
50
25
0
%
0
1
2
3
4
5
6
7
8
9
10
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

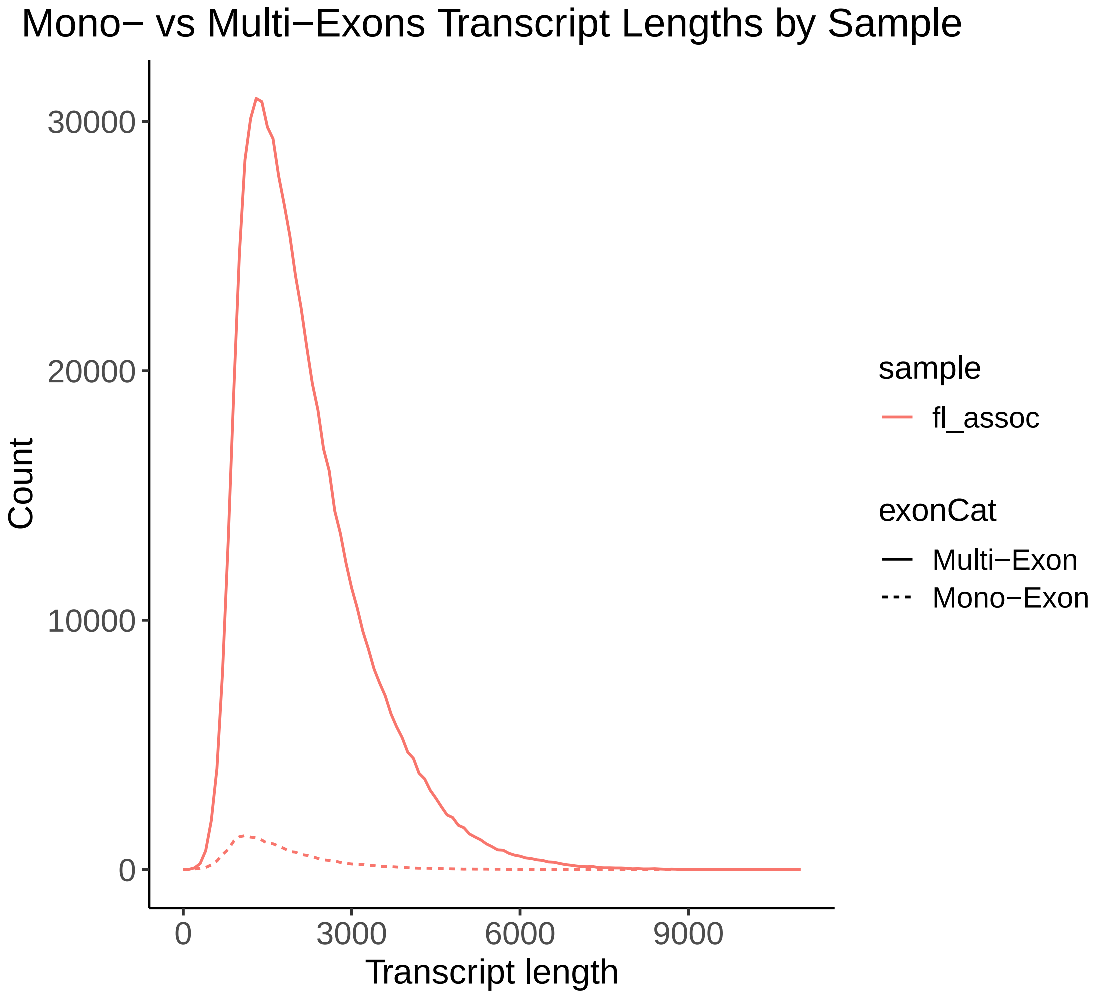

Number of FL reads per Gene by Type of Gene Annotation
log2(FL_TPM+1)
10
5
0
Annotated Genes
Novel Genes


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

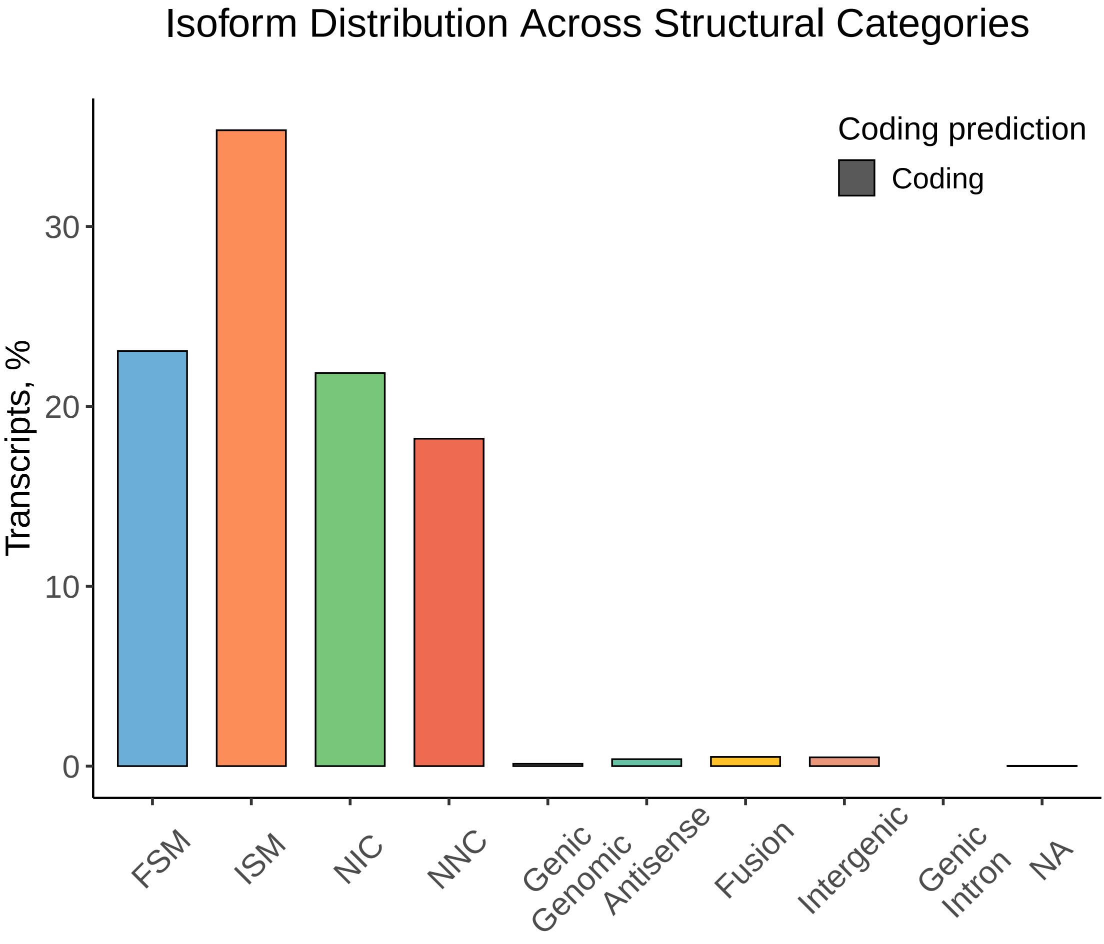

Transcript Lengths Distribution by Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
12000
9000
6000
3000
0
Count
0
3000
6000
9000
Transcript length

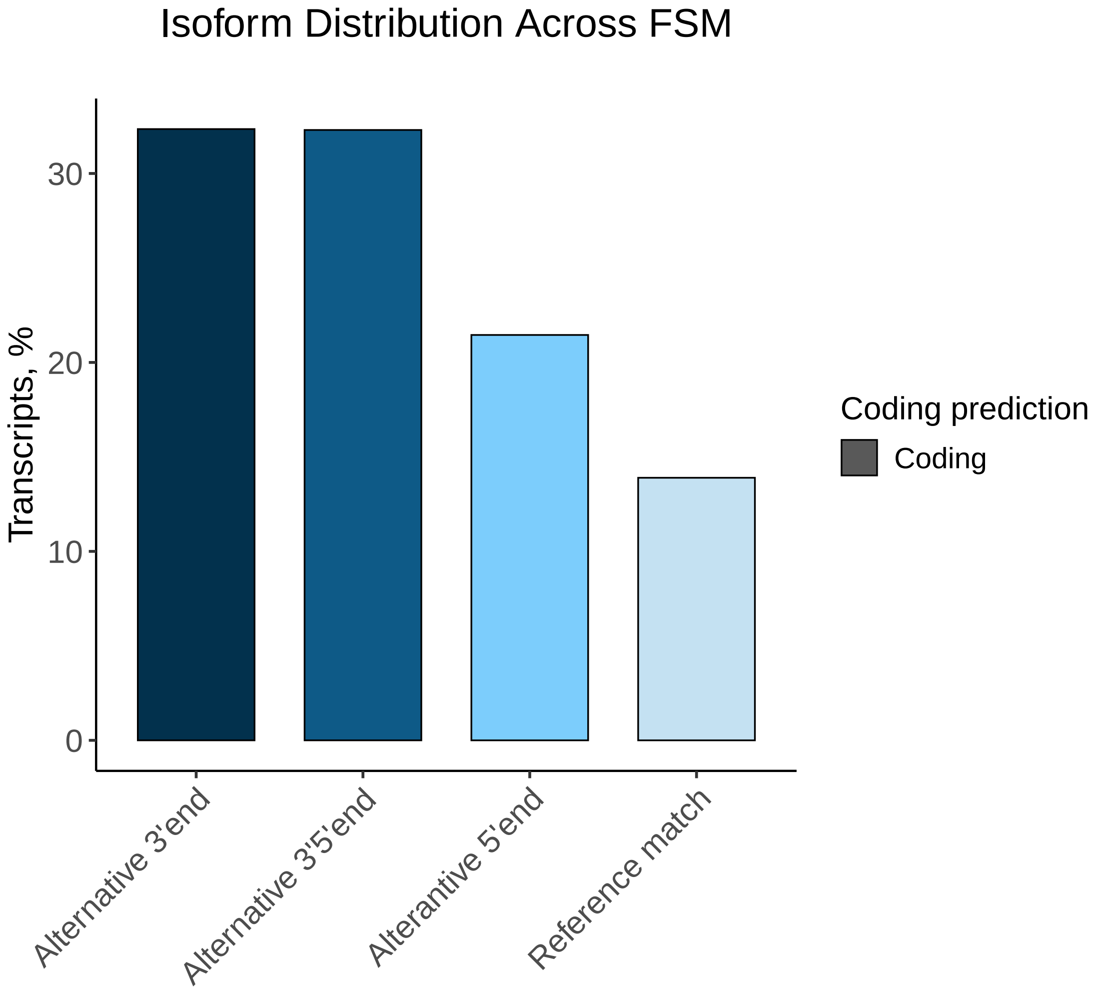

Mono− vs Multi− Exon Transcript Lengths Distribution
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length
exonCat
Multi−Exon
Mono−Exon

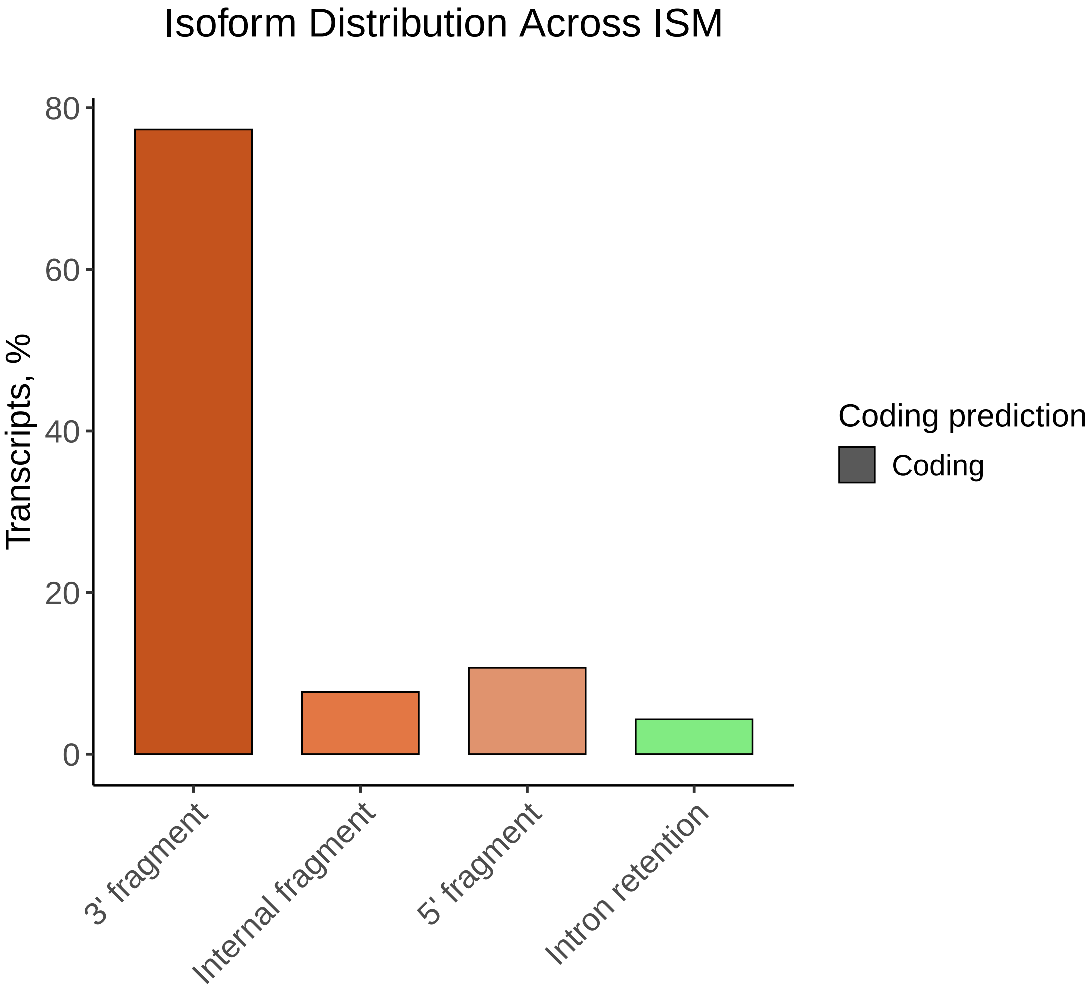

Transcript Lengths by Sample
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc

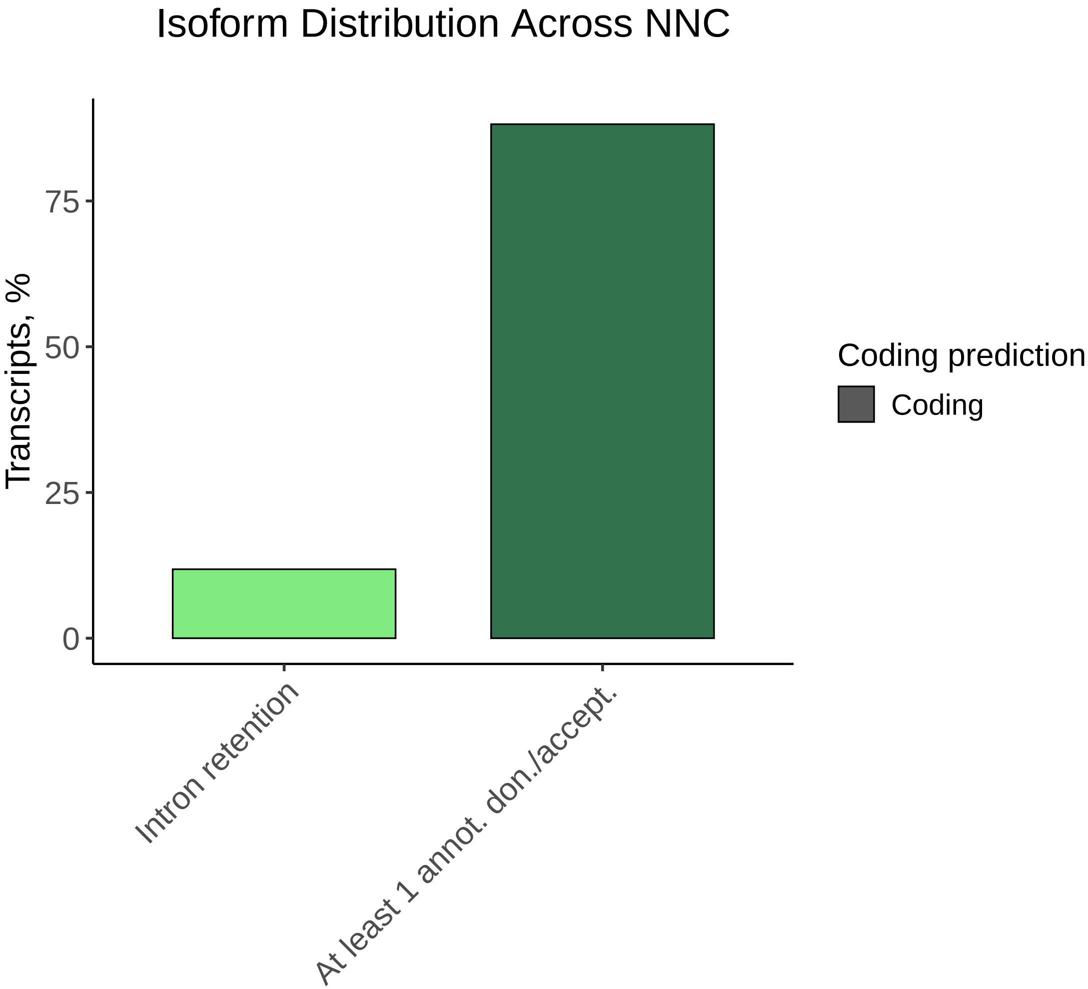

Mono− vs Multi−Exons Transcript Lengths by Sample
Count
30000
20000
10000
0
0
3000
6000
9000
Transcript length
sample
fl_assoc
exonCat
Multi−Exon
Mono−Exon

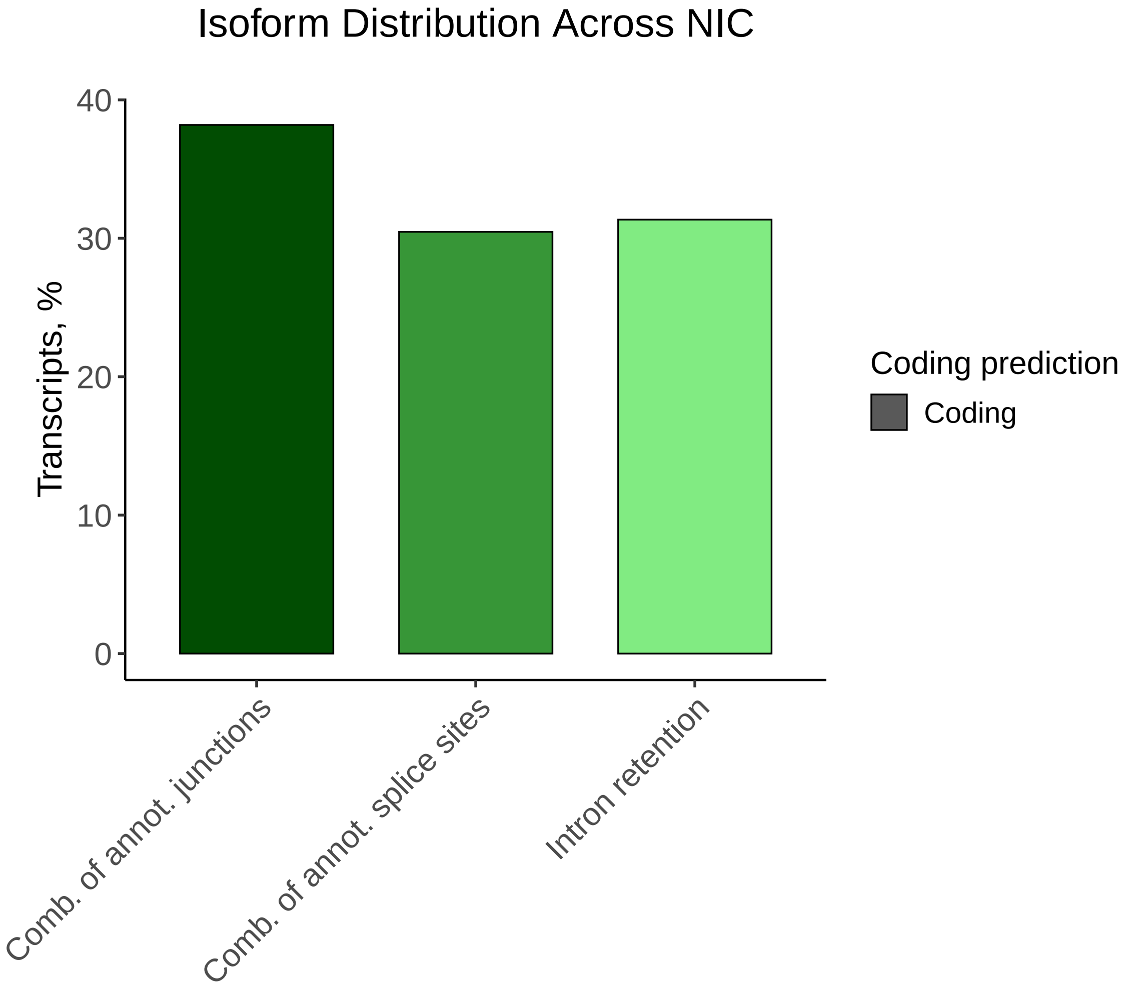
Structural Isoform Characterization

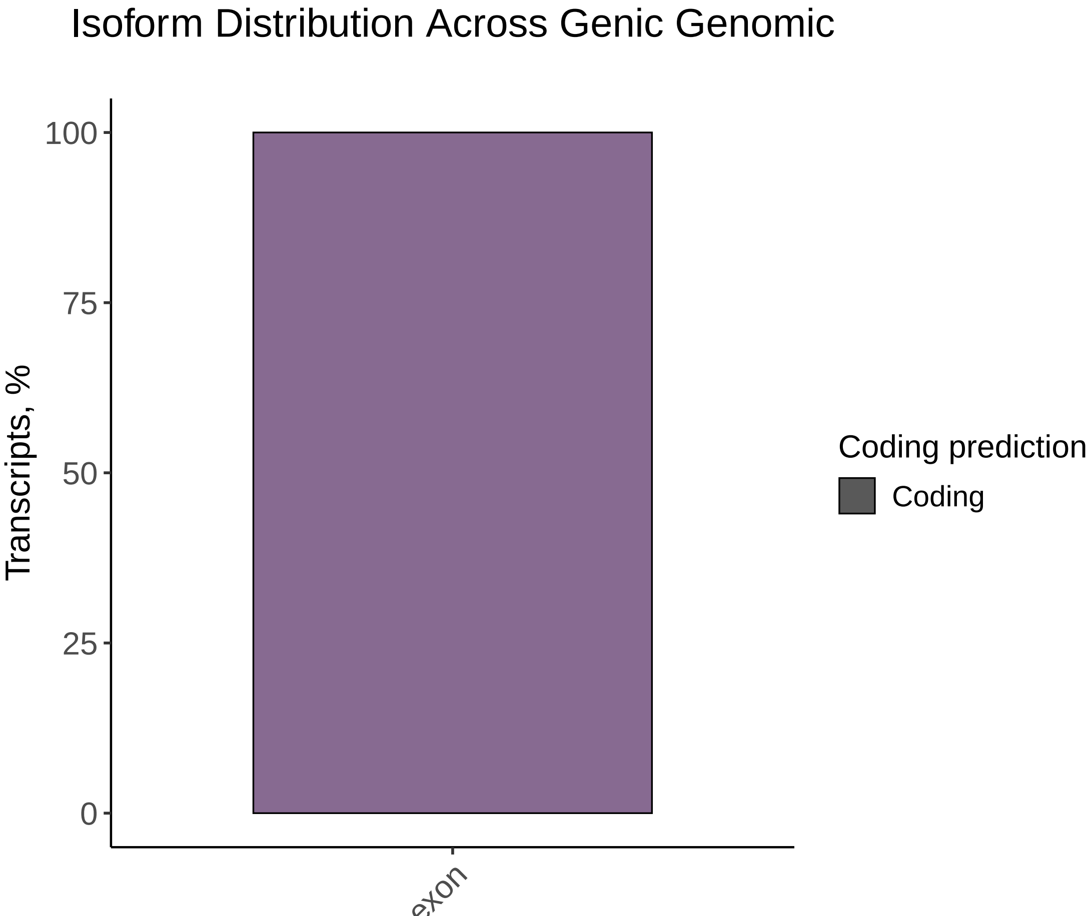

Isoform Distribution Across Structural Categories
Coding prediction
Coding
Transcripts, %
0
10
20
30
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

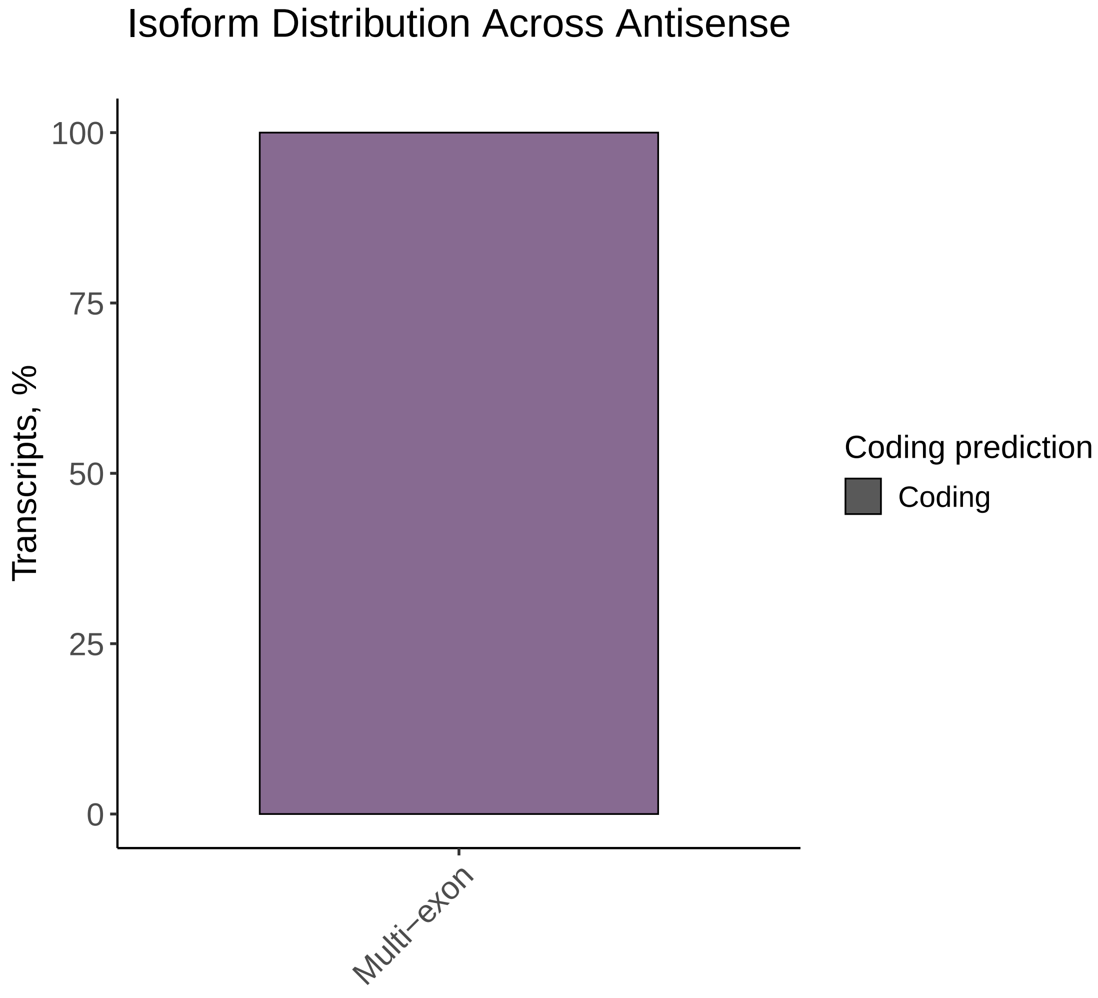

Isoform Distribution Across FSM
Transcripts, %
0
10
20
30
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Coding prediction
Coding

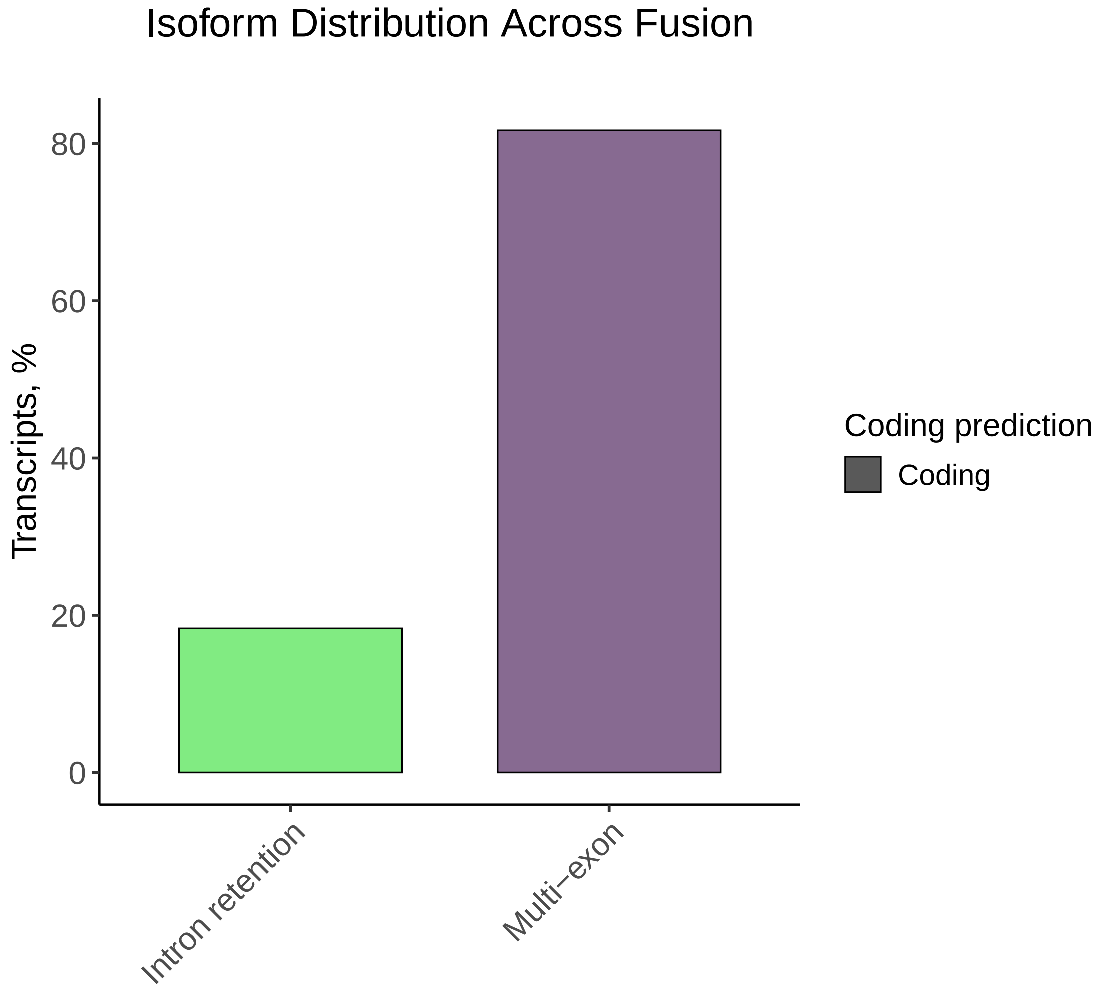

Isoform Distribution Across ISM
Transcripts, %
80
60
40
20
0
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding

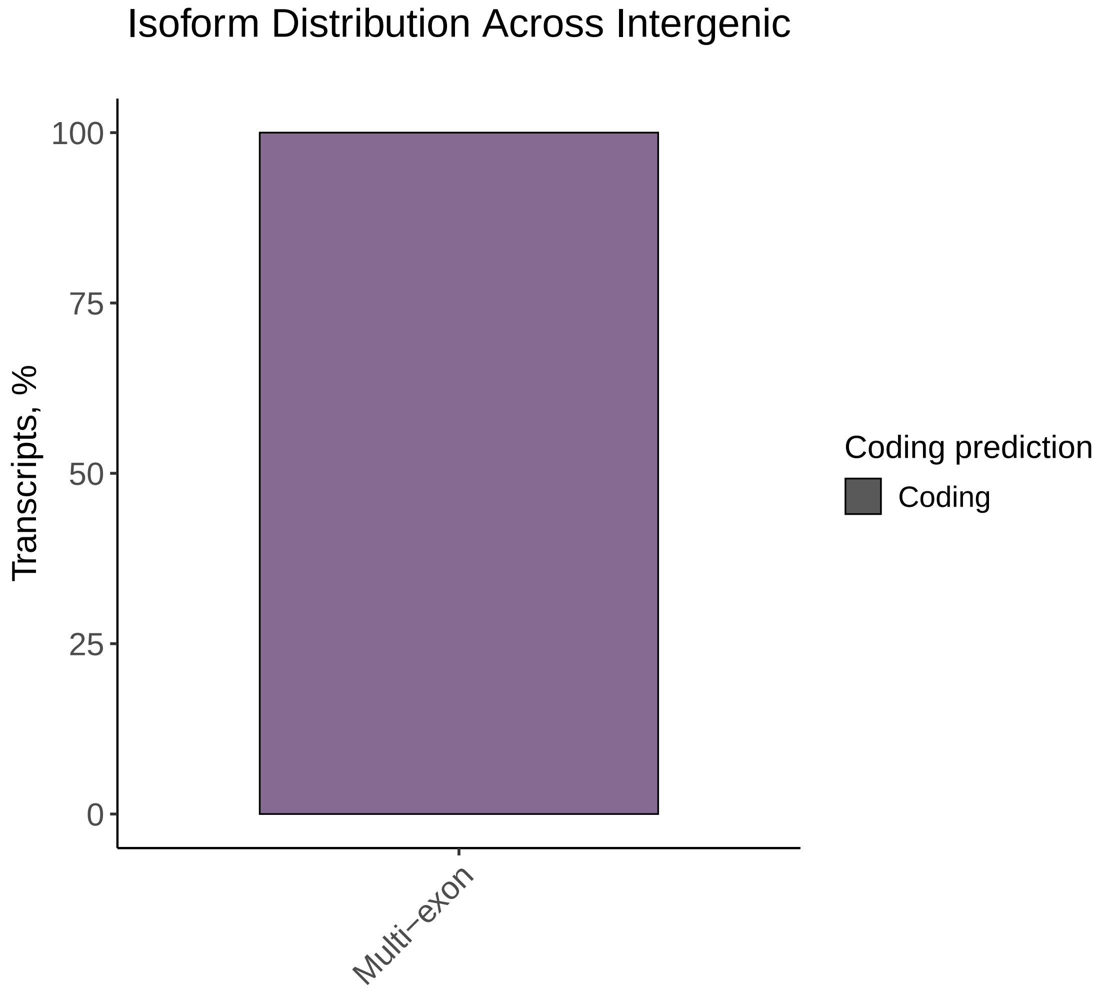

Isoform Distribution Across NNC
Transcripts, %
0
25
50
75
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding


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

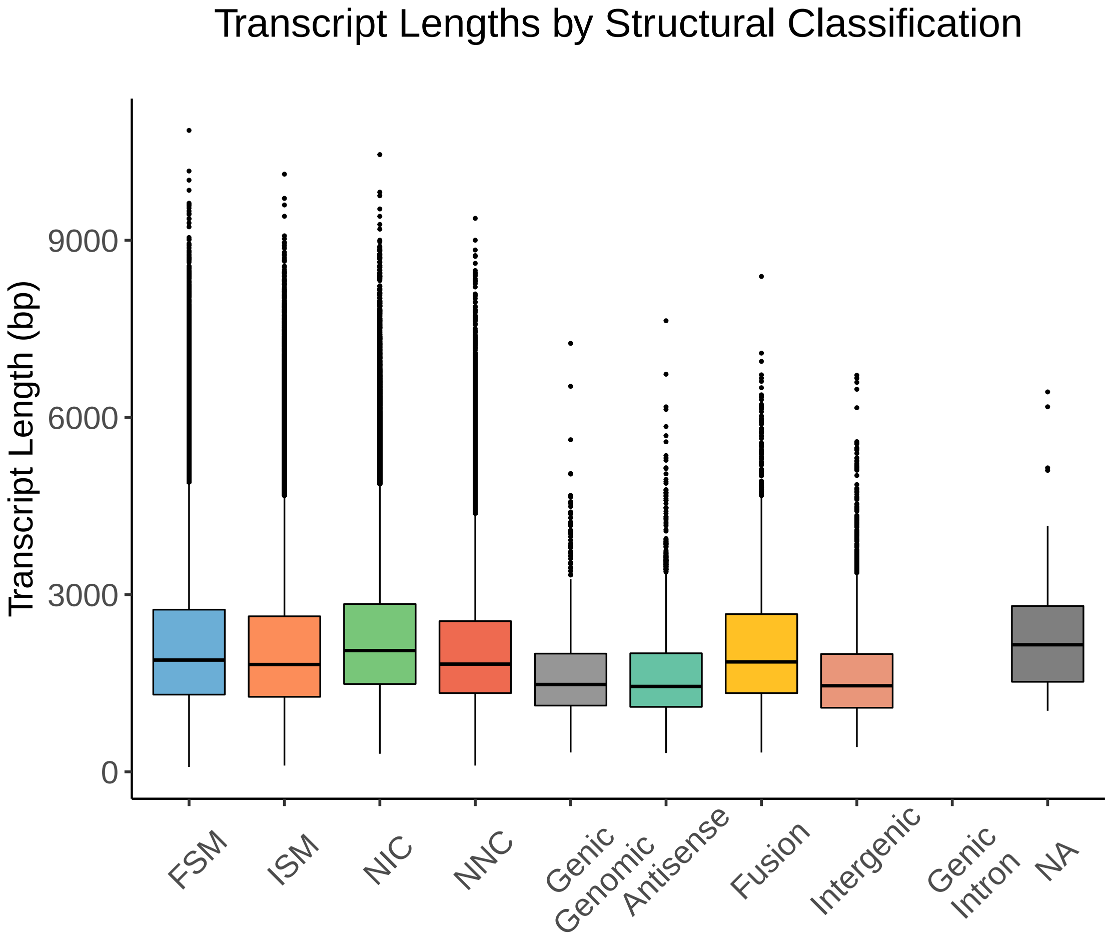

Isoform Distribution Across Genic Genomic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding

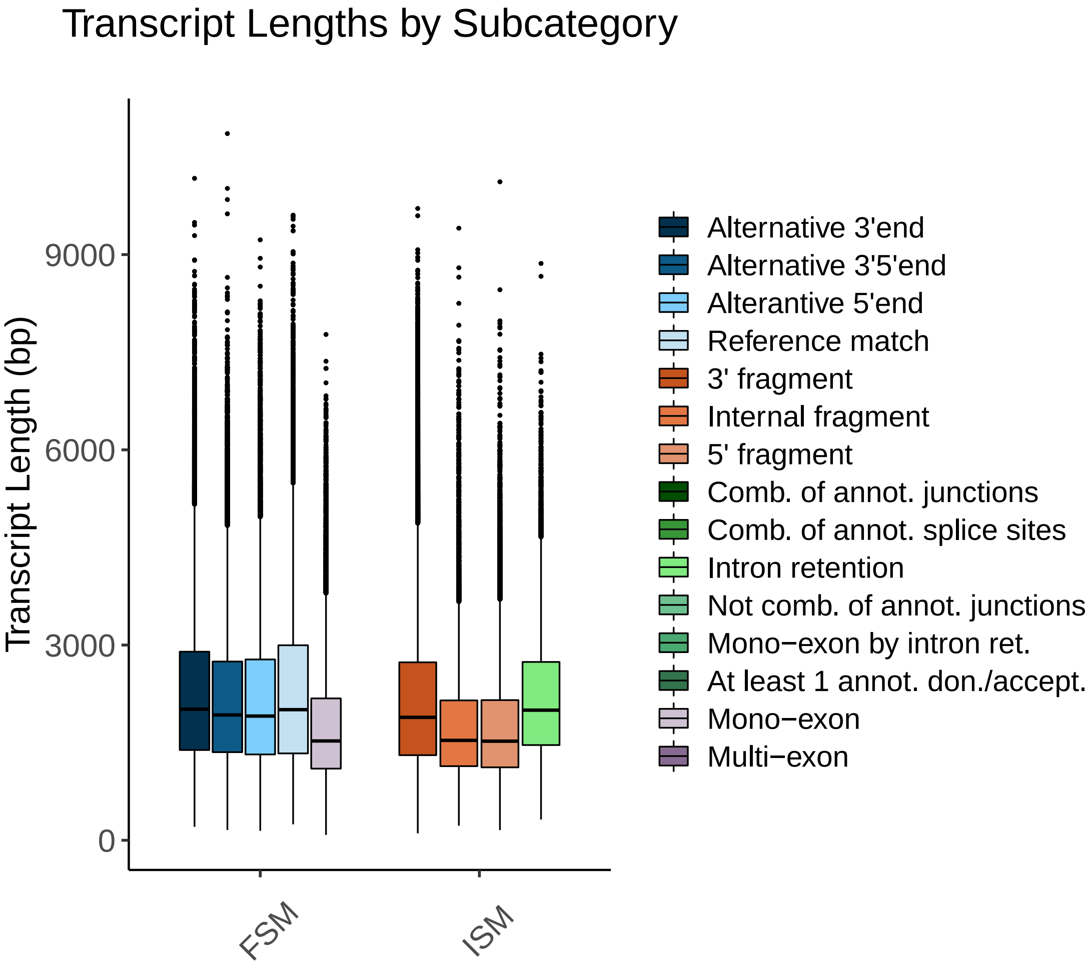

Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding

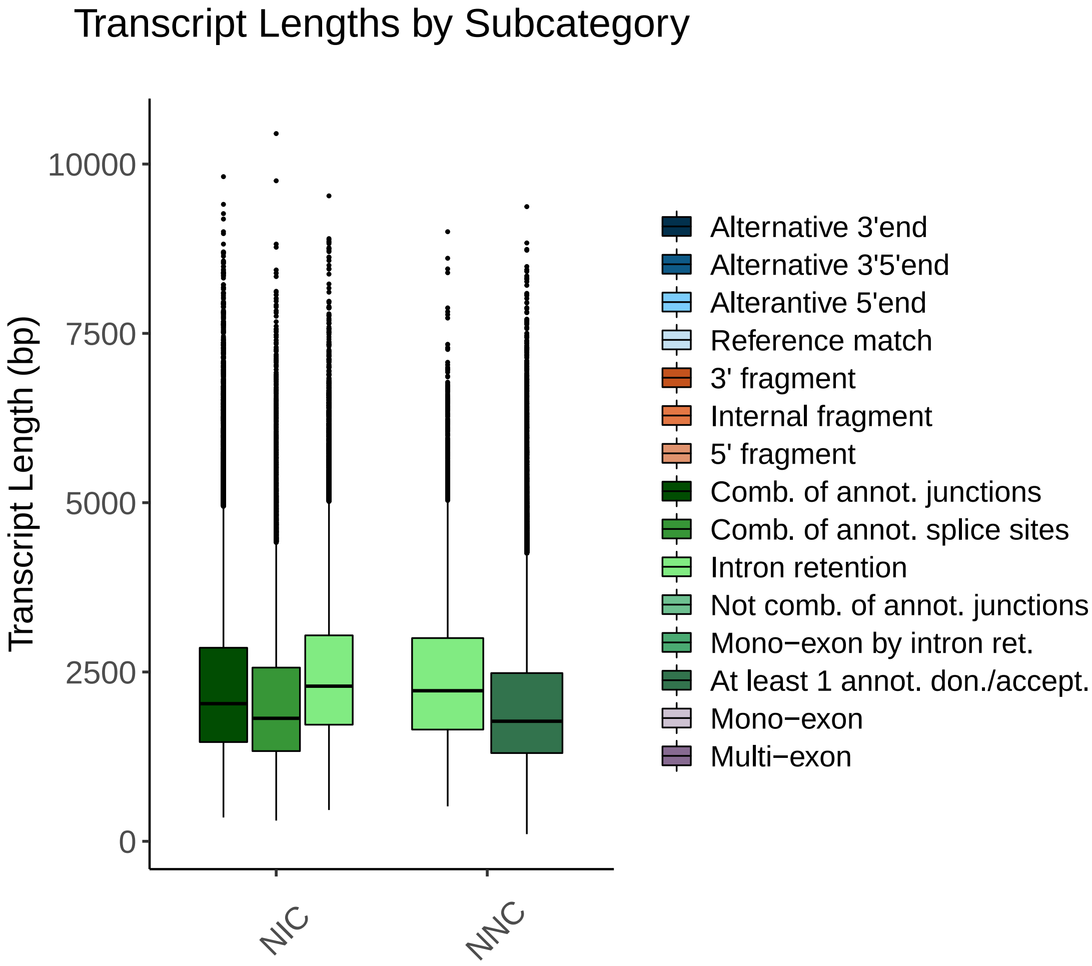

Isoform Distribution Across Fusion
Transcripts, %
0
20
40
60
80
Intron retention
Multi−exon
Coding prediction
Coding



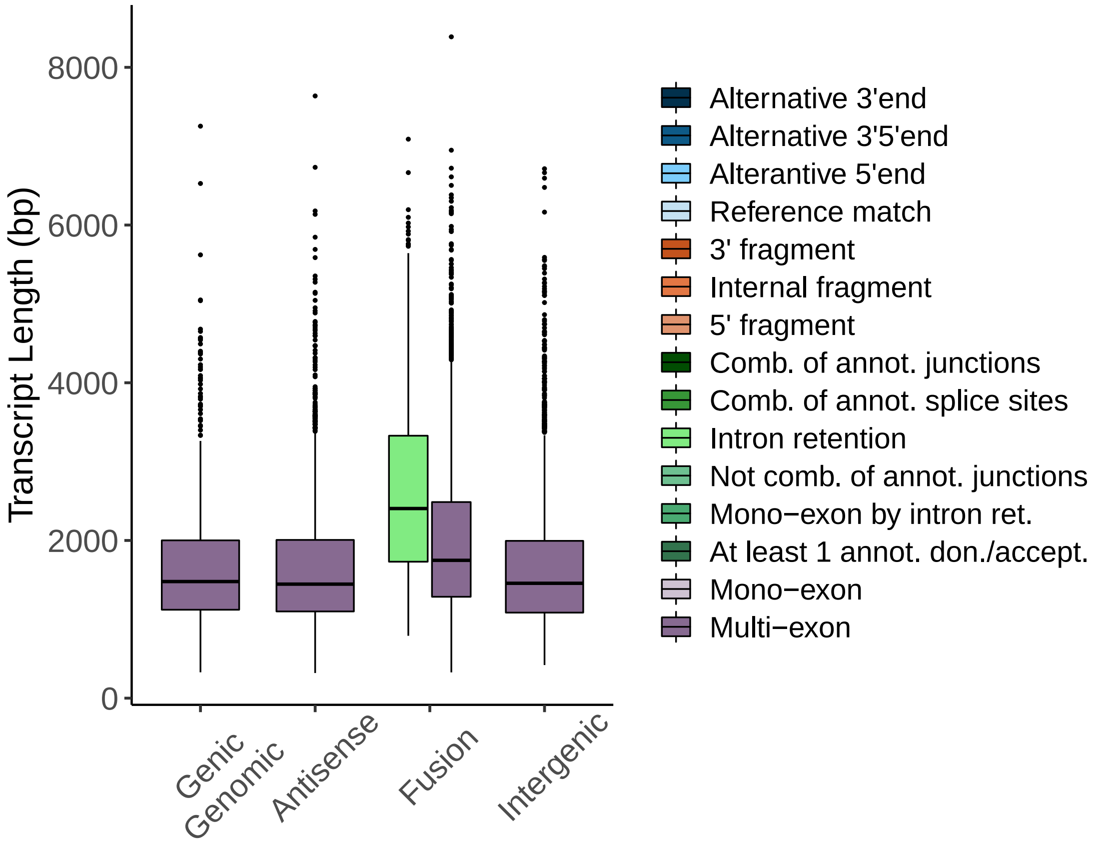
Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding

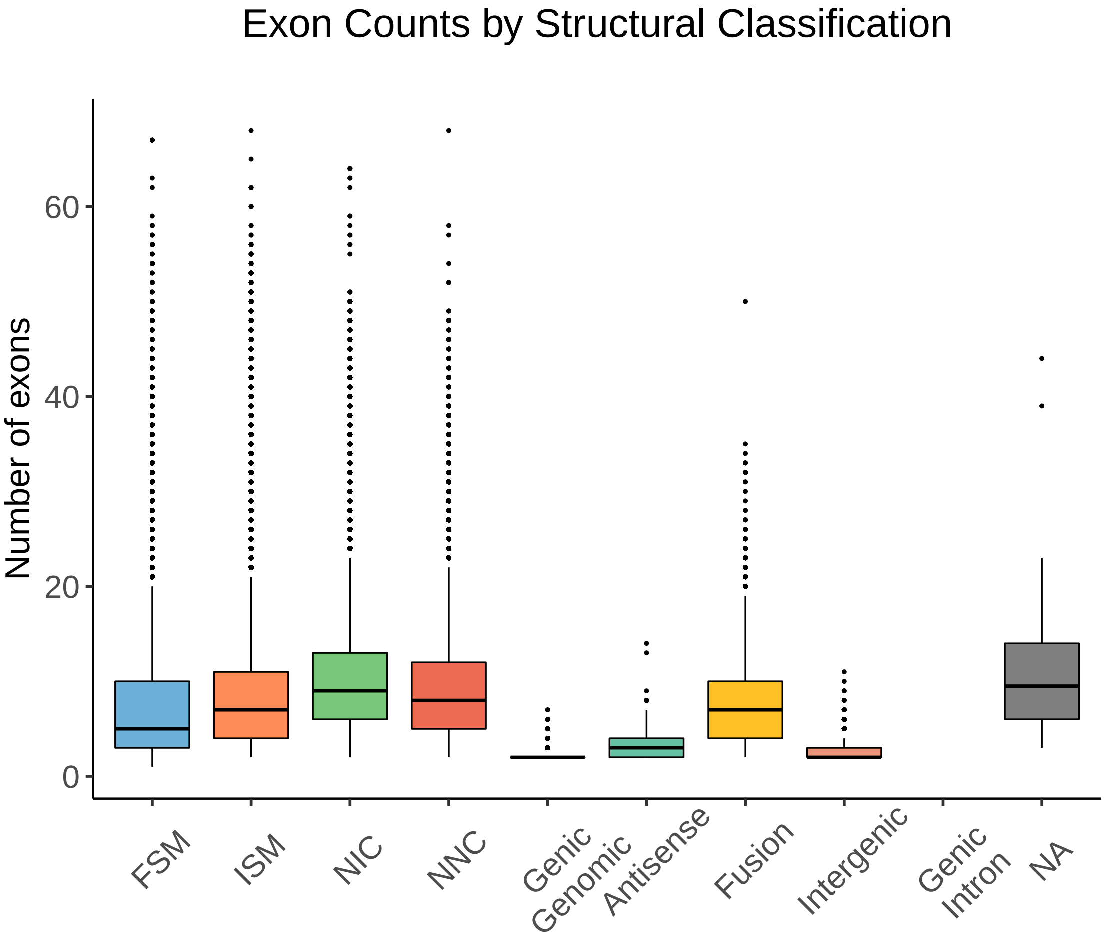

Transcript Lengths by Structural Classification
Transcript Length (bp)
0
3000
6000
9000
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

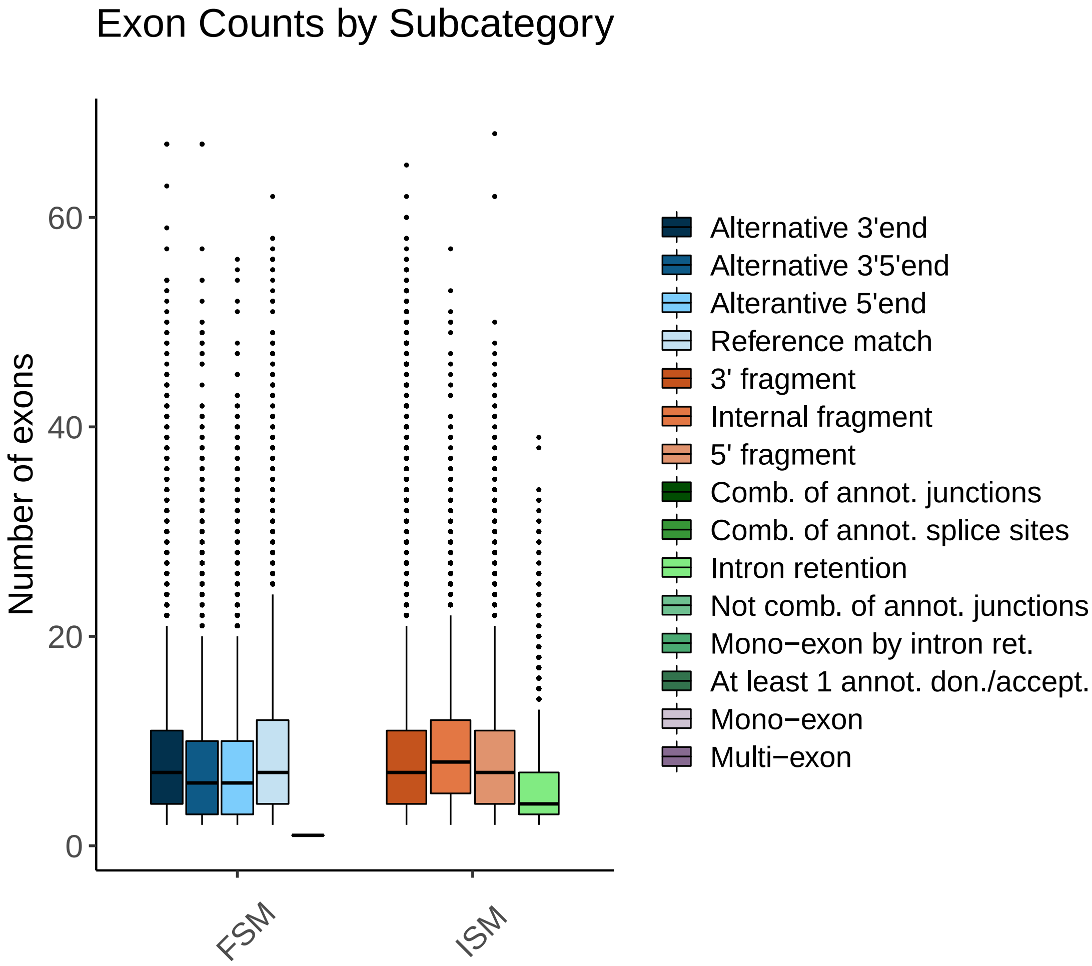

Transcript Lengths by Subcategory
Transcript Length (bp)
9000
6000
3000
0
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon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


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

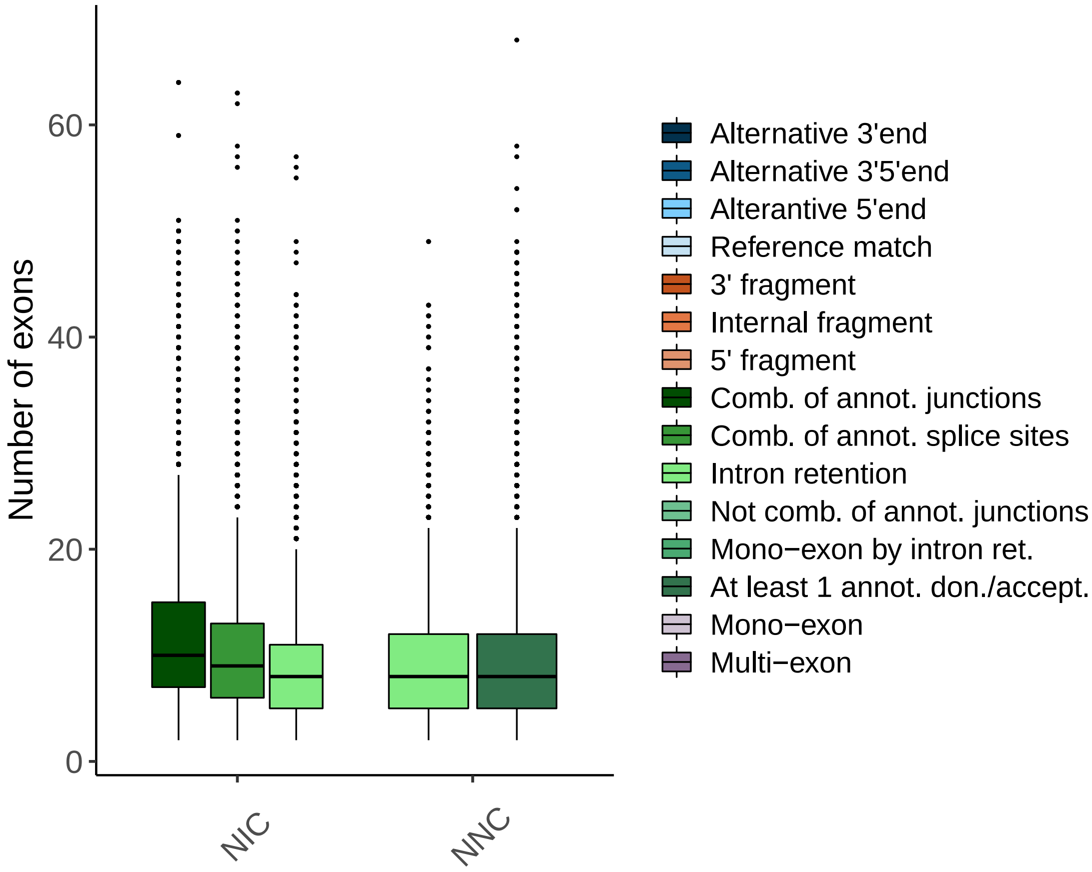
Transcript Lengths by Subcategory

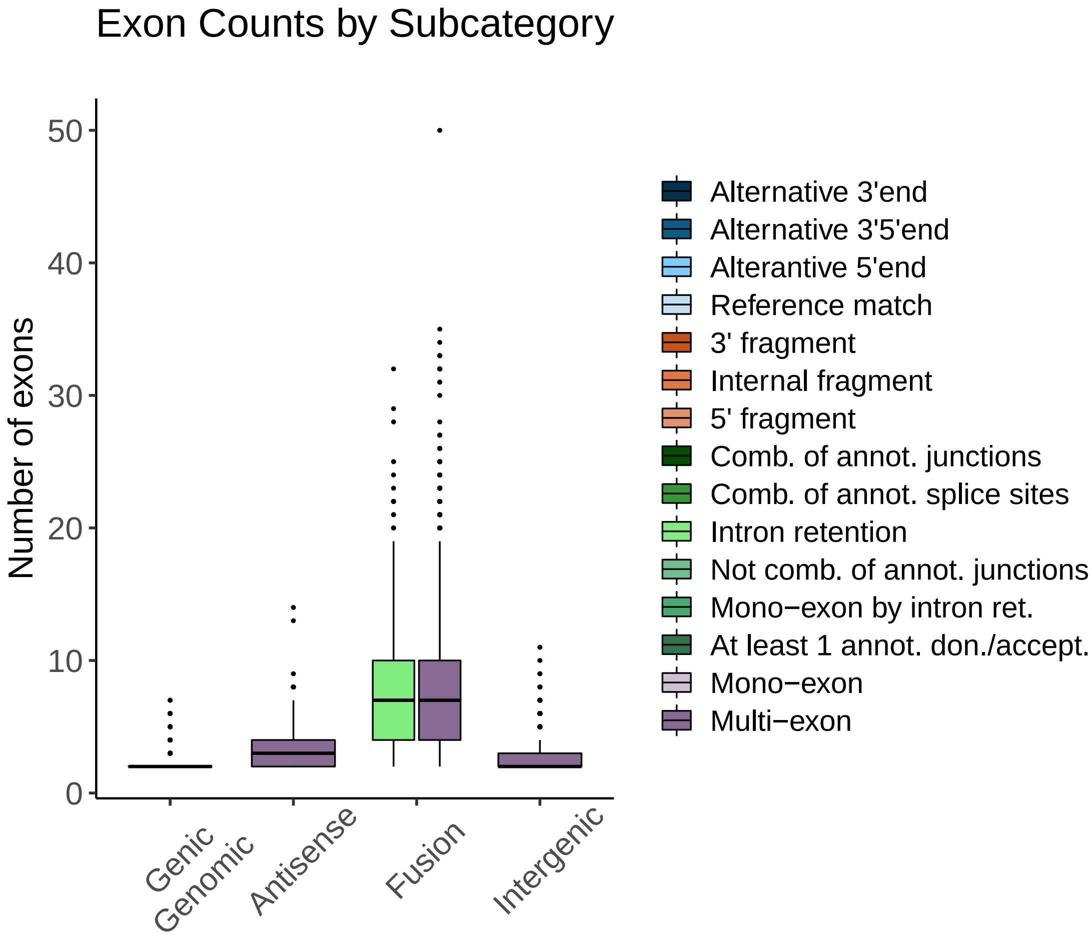


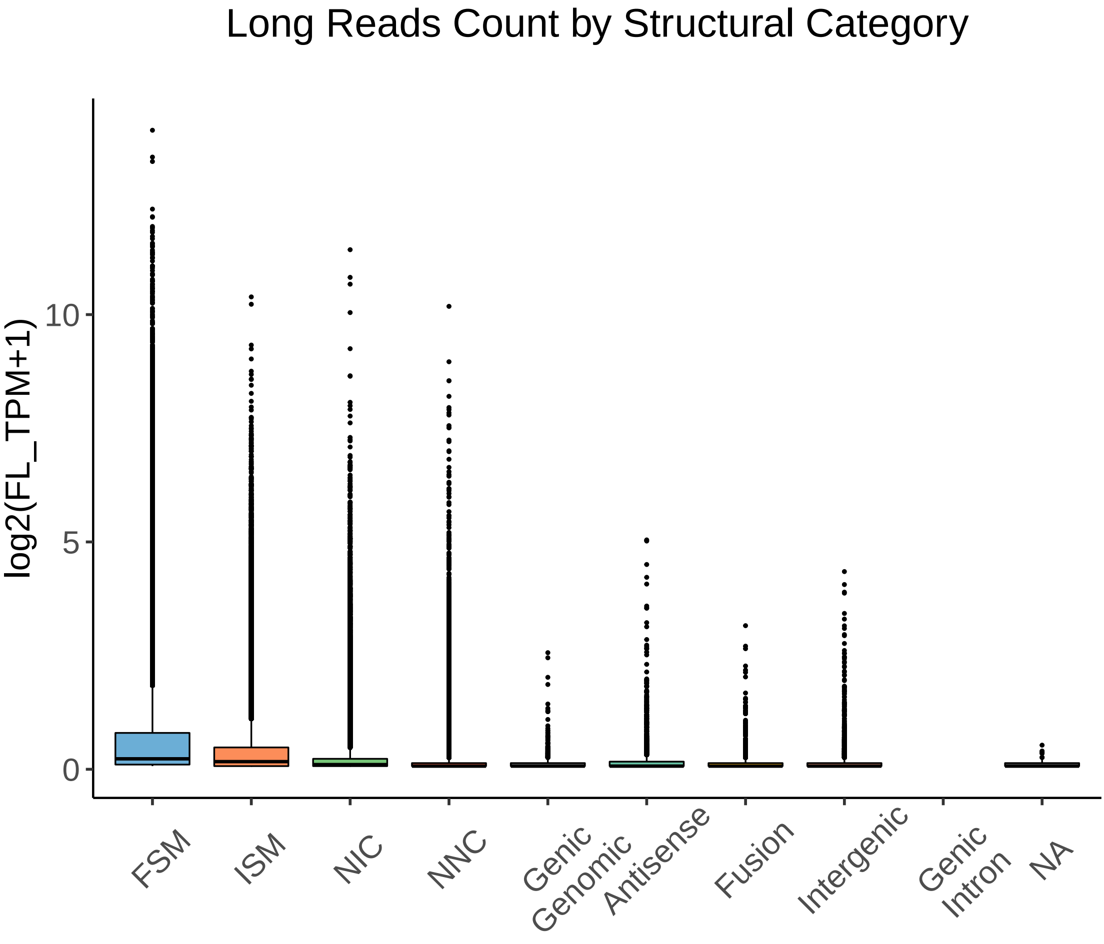

Exon Counts by Structural Classification
Number of exons
0
20
40
60
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA


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

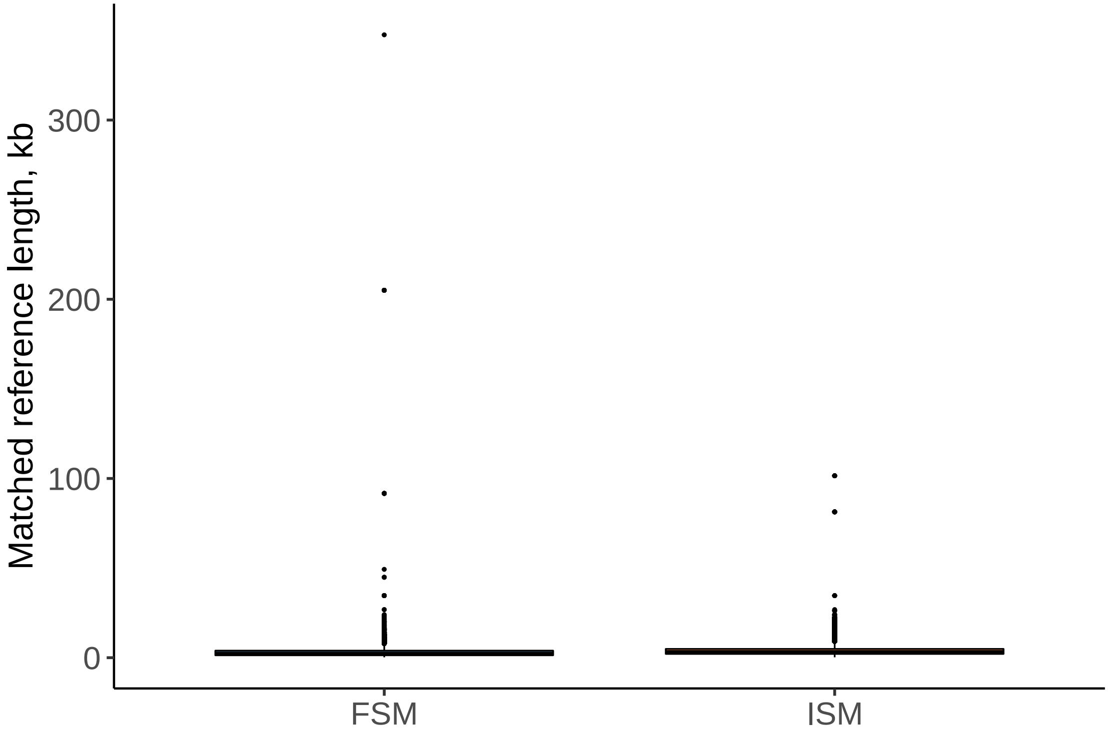


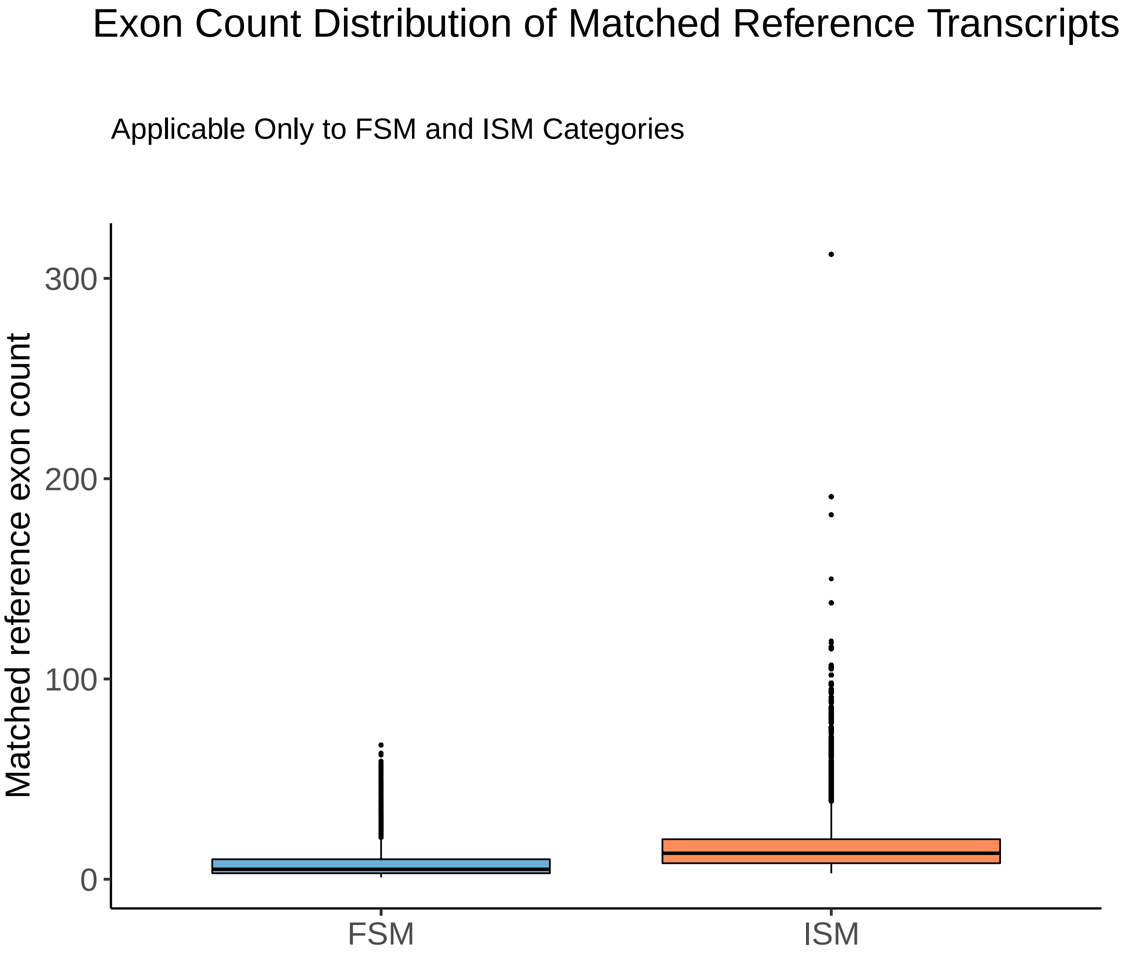

Exon Counts by Subcategory
Number of exons
0
10
20
30
40
50
Genic Genomic
Antisense
Fusion
Intergenic
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of 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)
0
5
10
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
NA

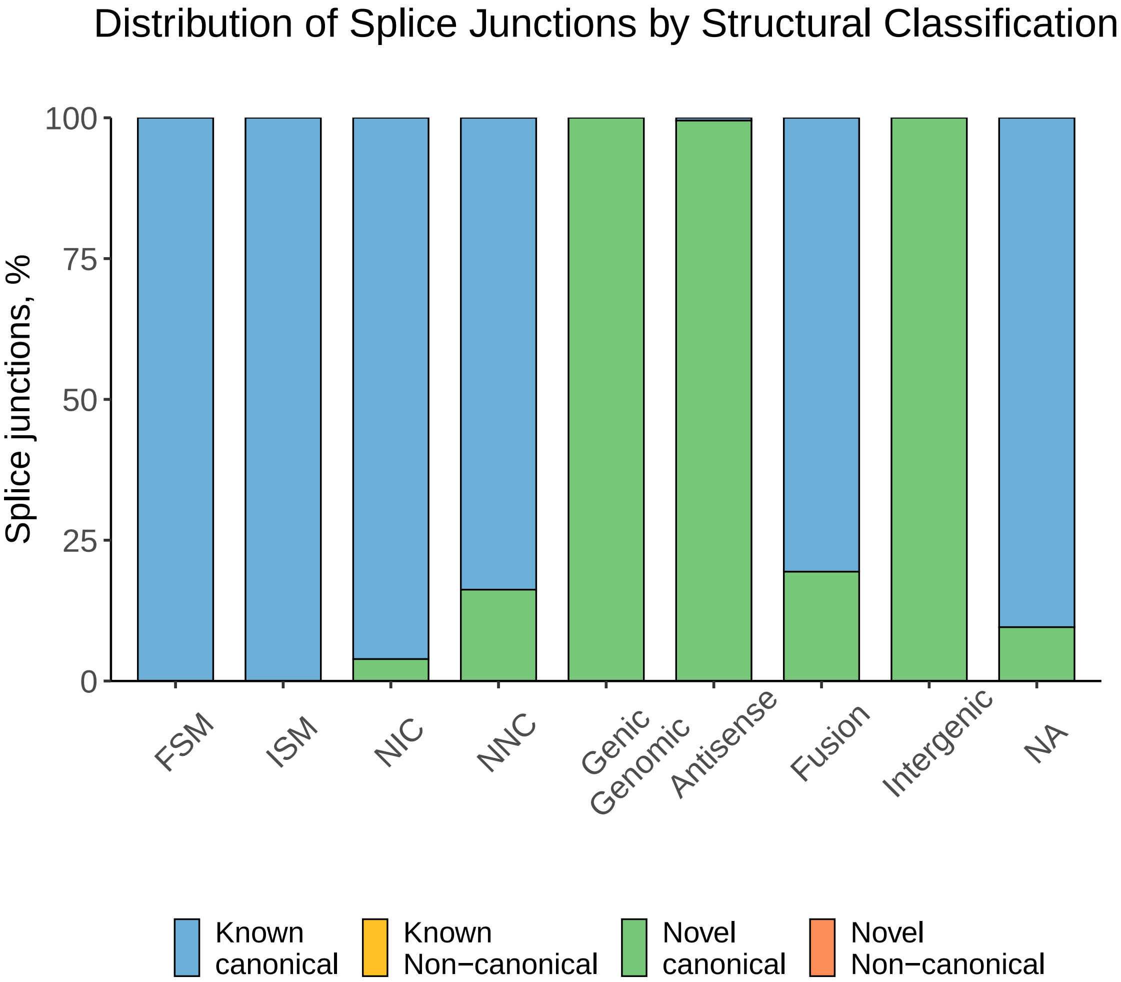
Length Distribution of Matched Reference Transcripts

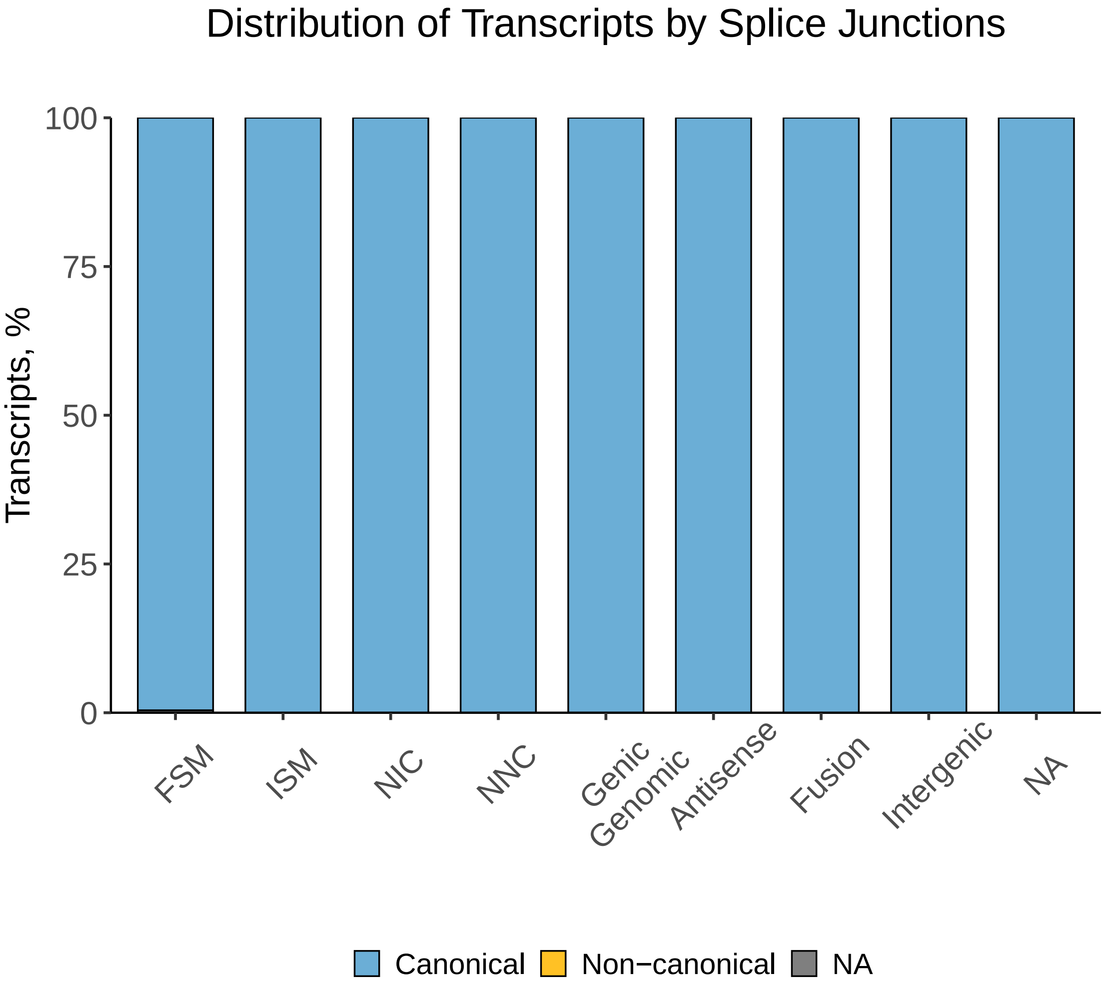
Applicable Only to FSM and ISM Categories

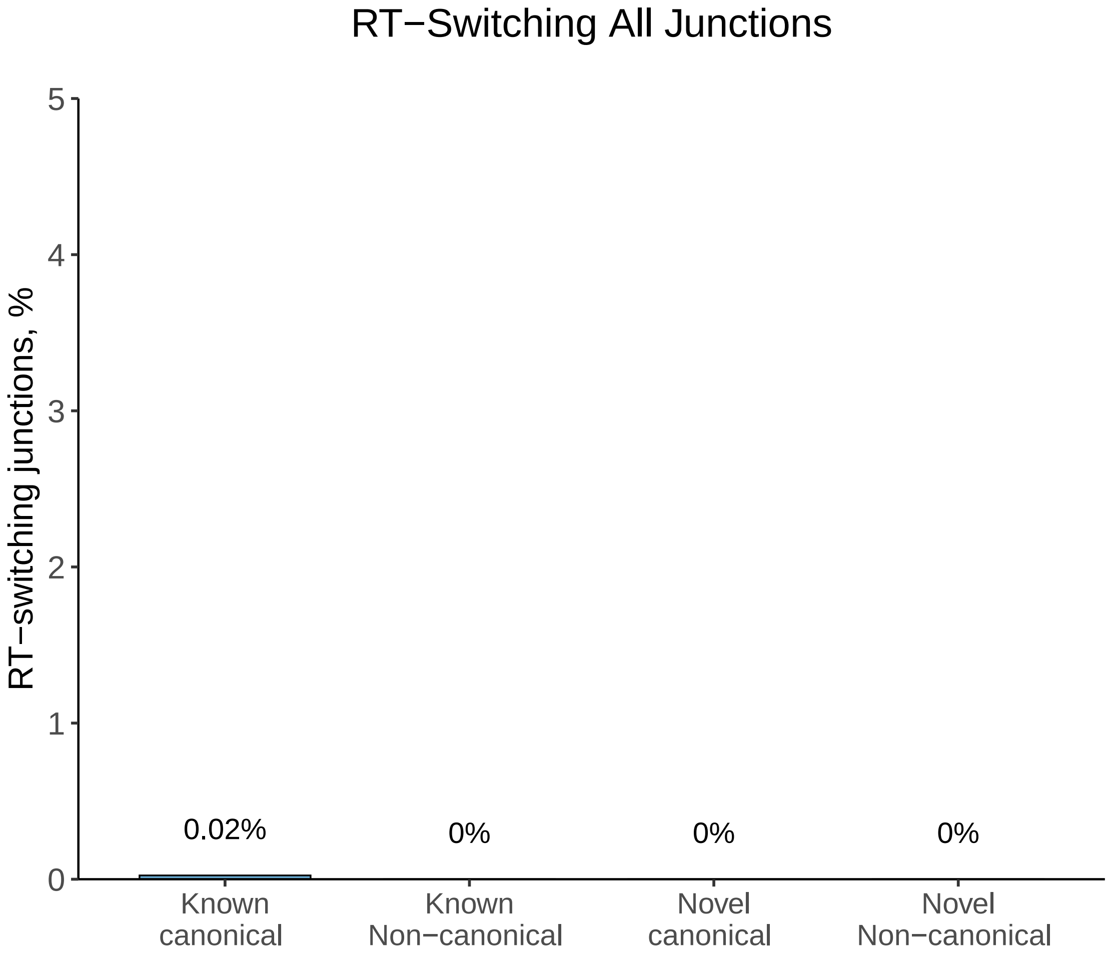


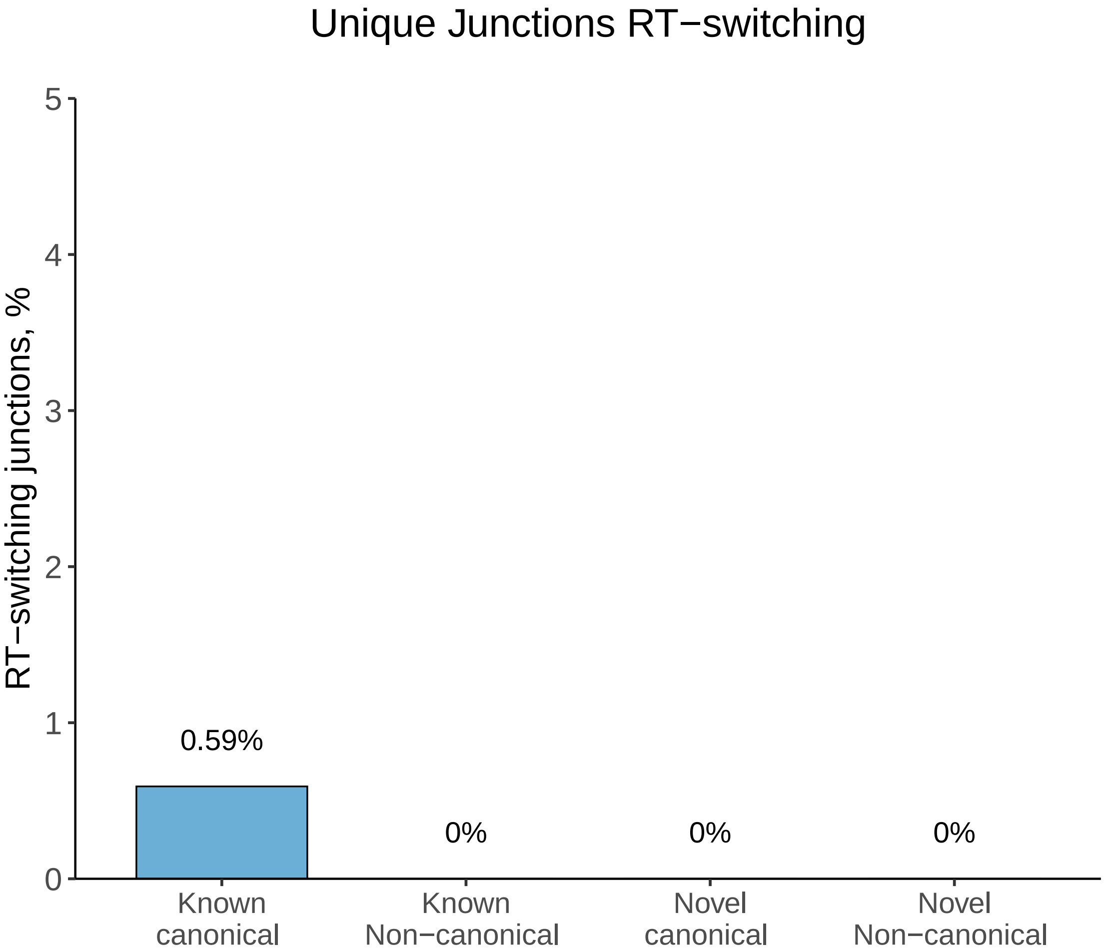

Exon Count Distribution of Matched Reference Transcripts
Applicable Only to FSM and ISM Categories
Matched reference exon count
300
200
100
0
FSM
ISM

Splice Junction Characterization

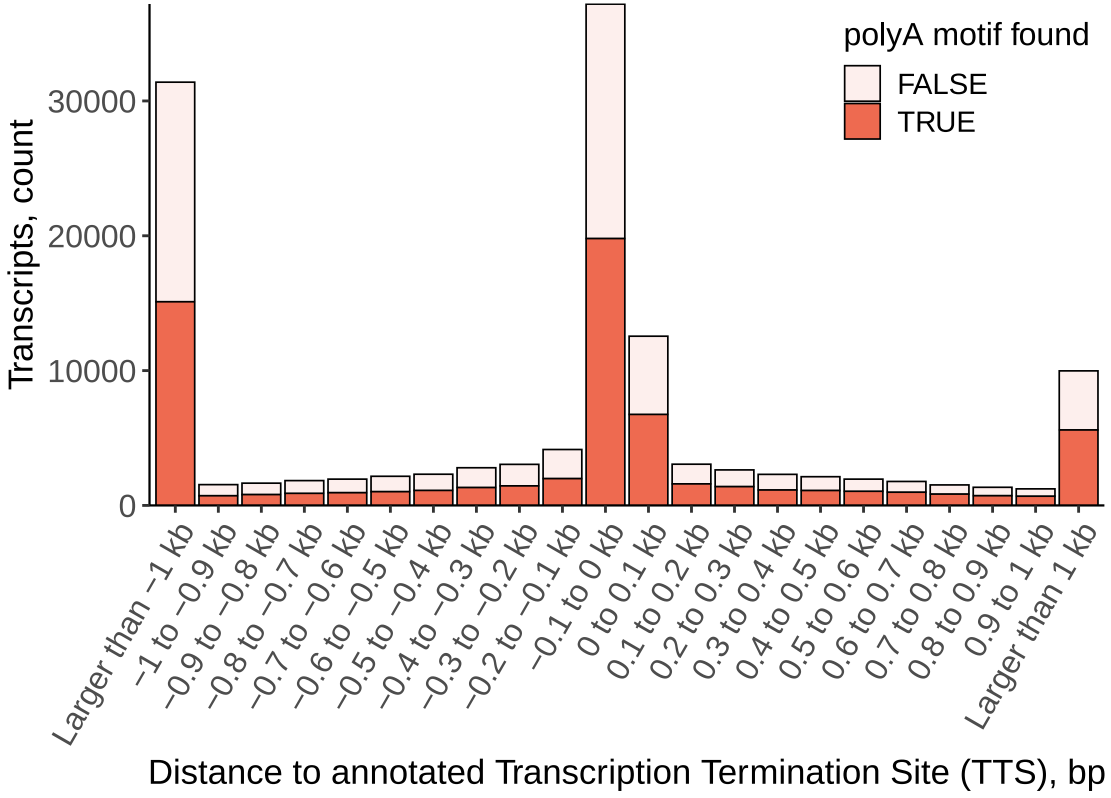

Distribution of Splice Junctions by Structural Classification
Splice junctions, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
NA
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical


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


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

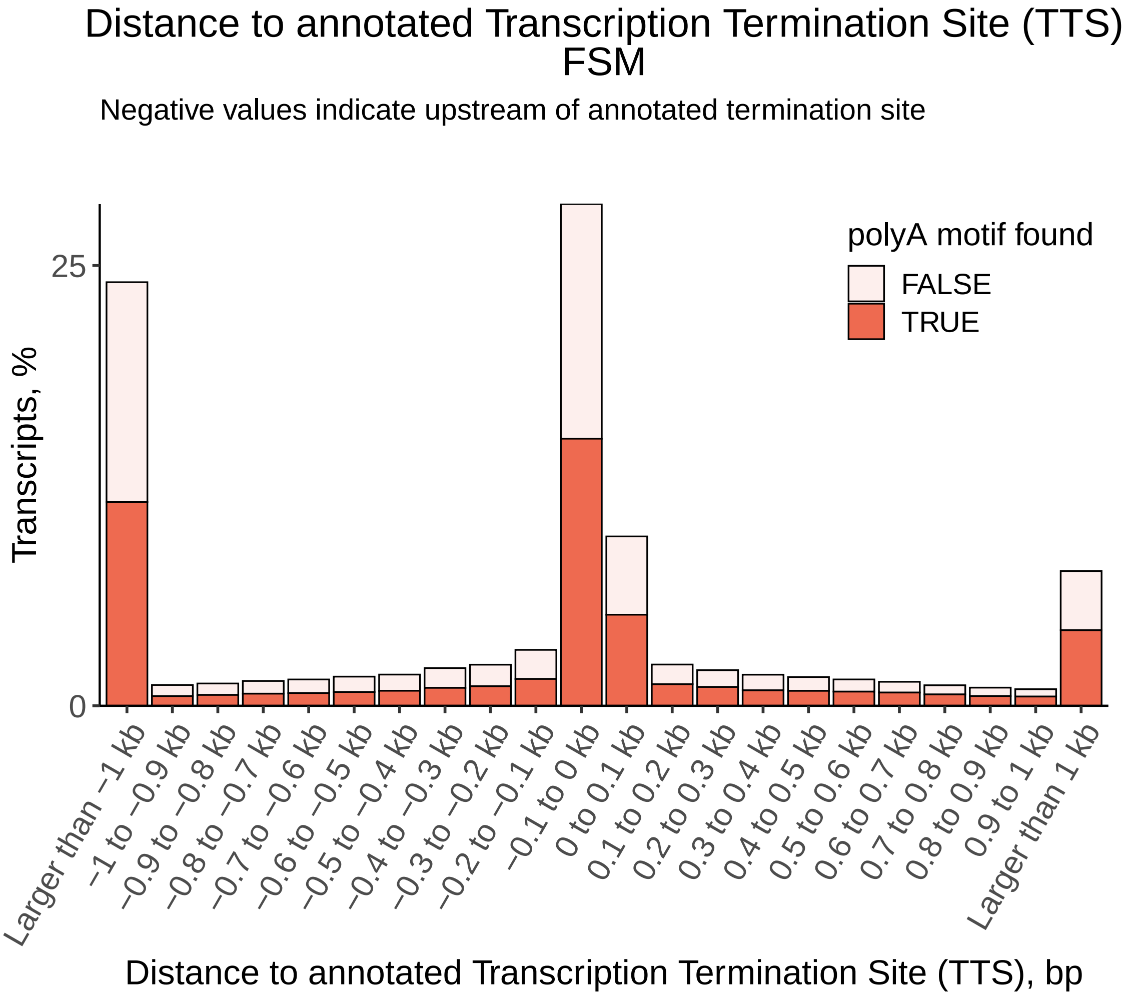

Unique Junctions RT−switching
RT−switching junctions, %
5
4
3
2
1
0
0.59%
0%
0%
0%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

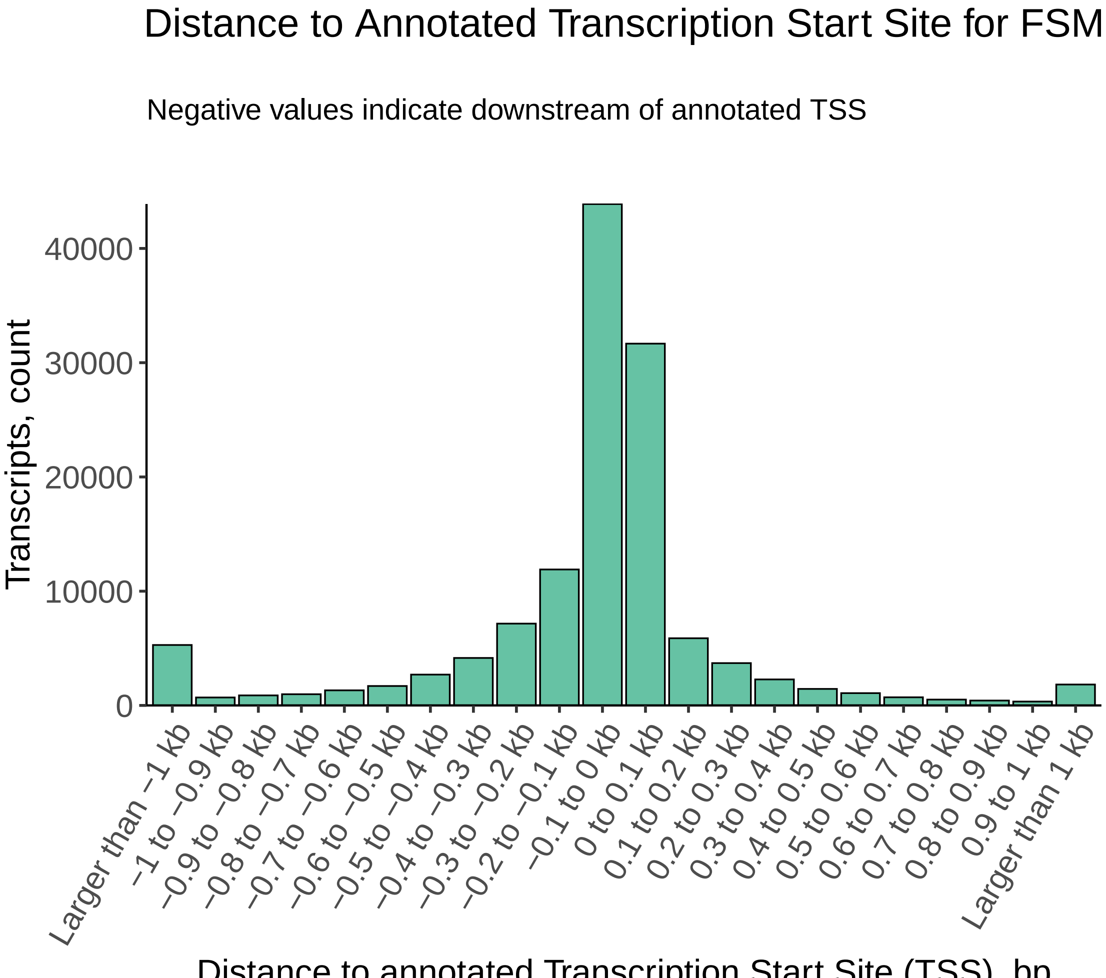
Comparison With Annotated TSS and TTS

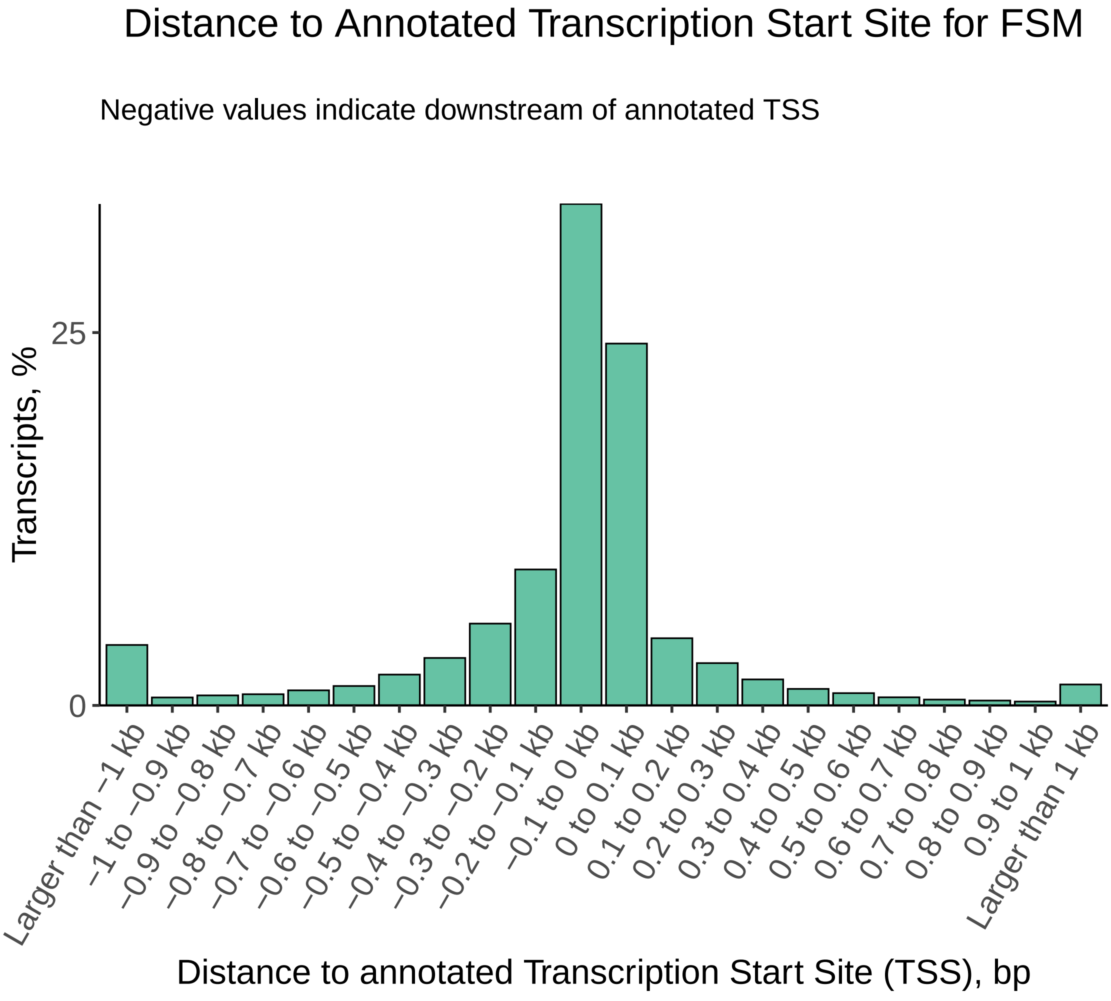


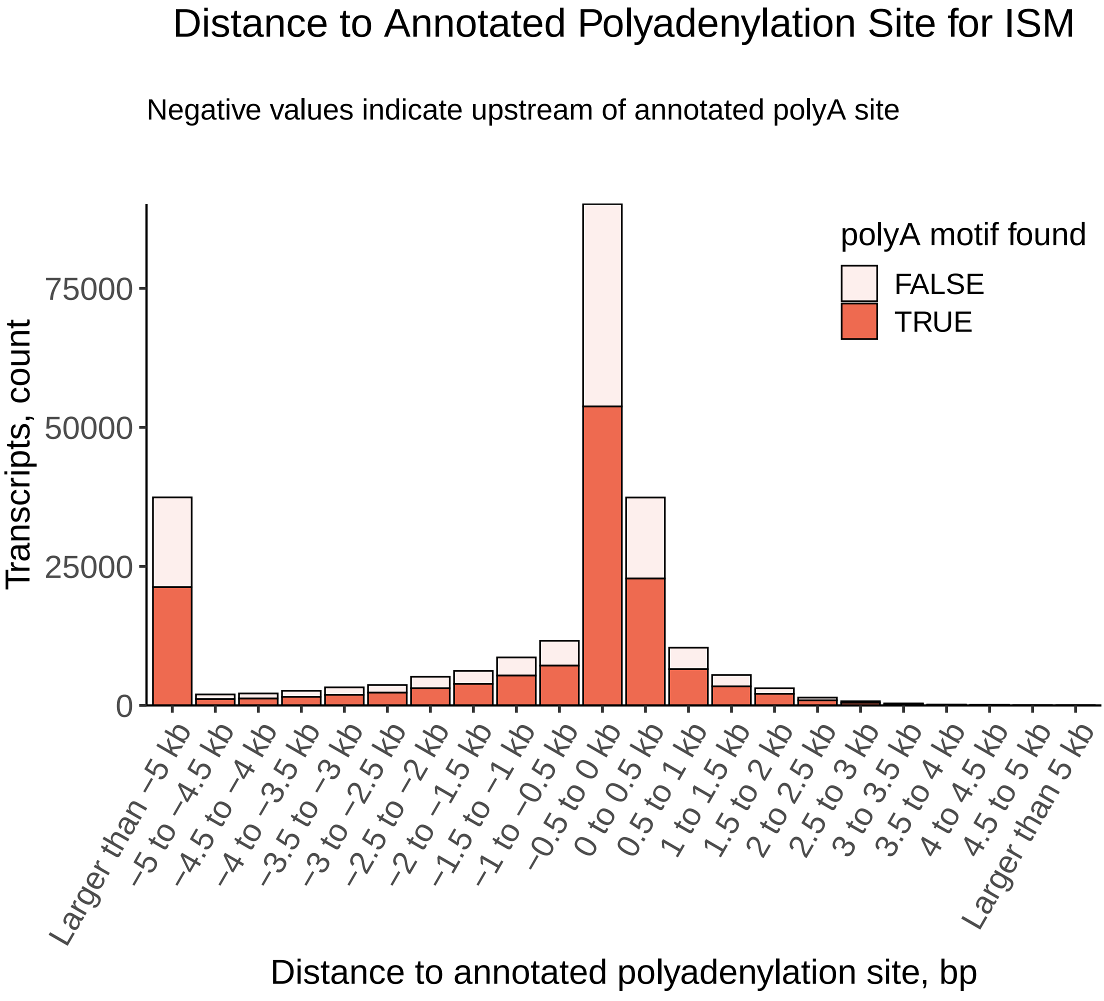
Distance to annotated Transcription Termination Site (TTS)
FSM

Negative values indicate upstream of annotated termination site


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

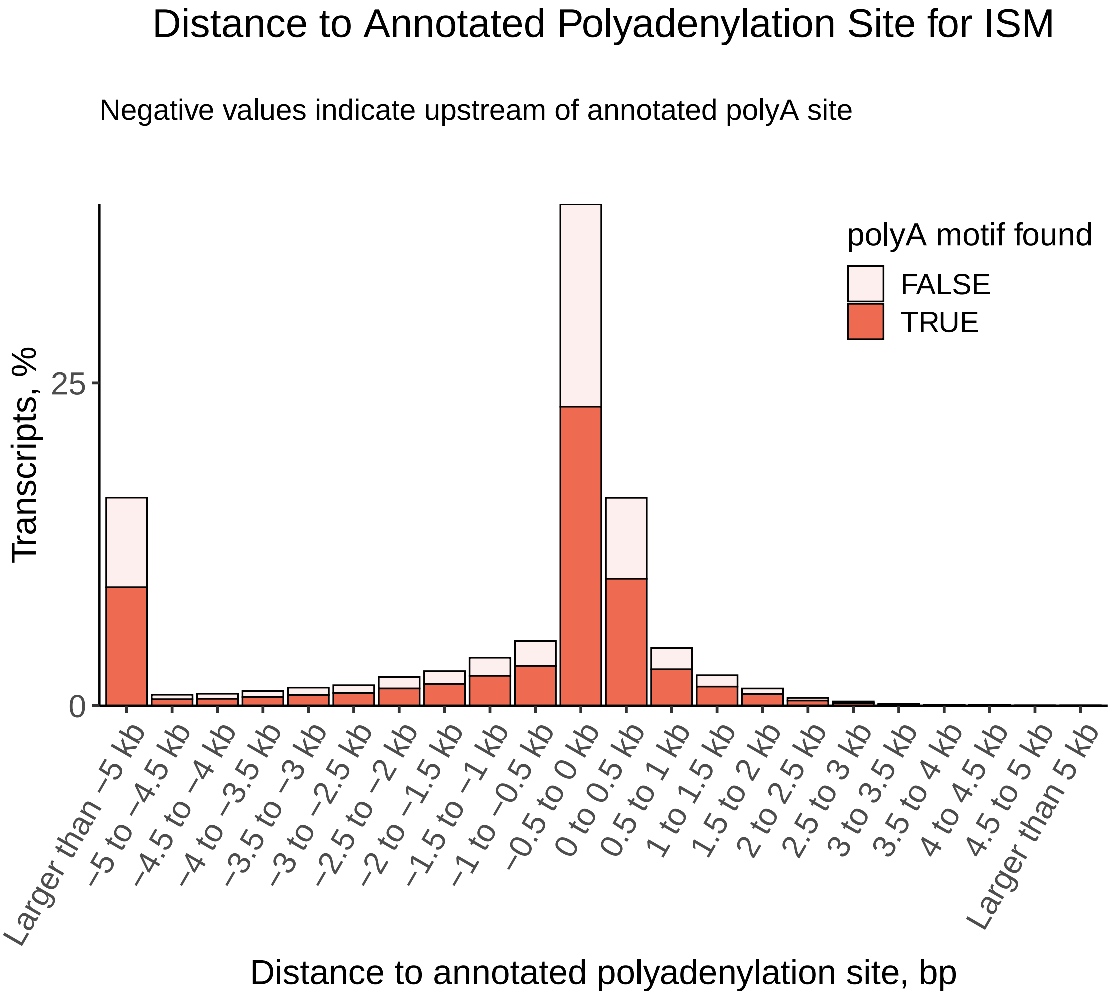

Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
Transcripts, count
40000
30000
20000
10000
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2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


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



Distance to Annotated Polyadenylation Site for ISM

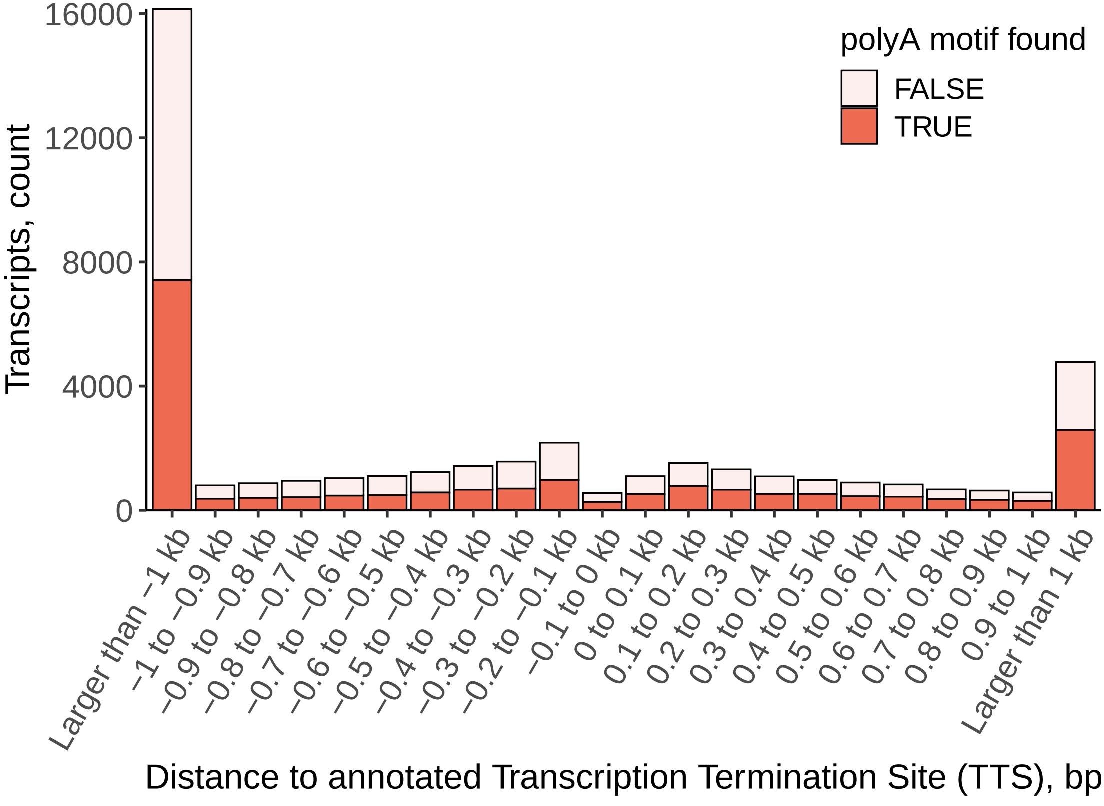
Negative values indicate upstream of annotated polyA site

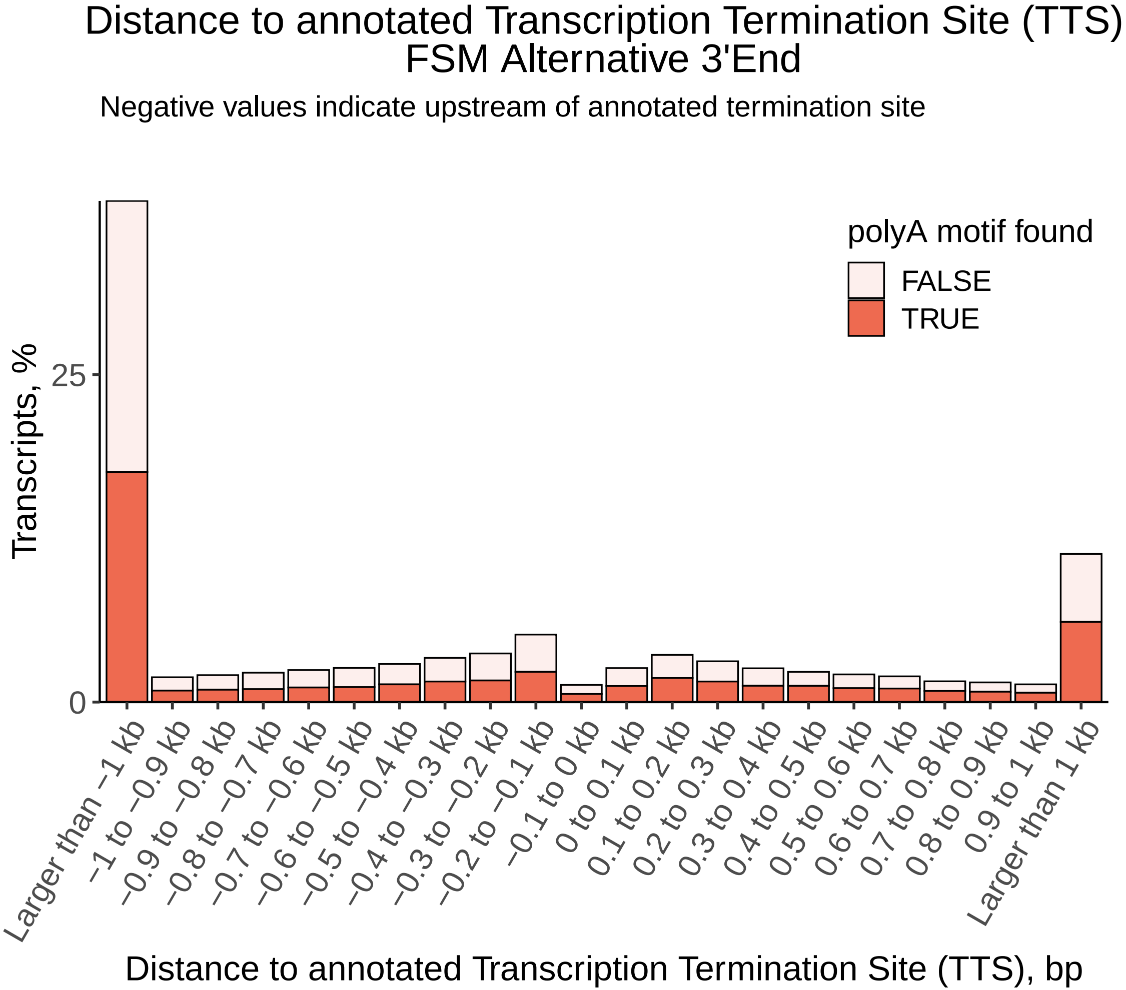

Distance to Annotated Polyadenylation Site for ISM
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb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

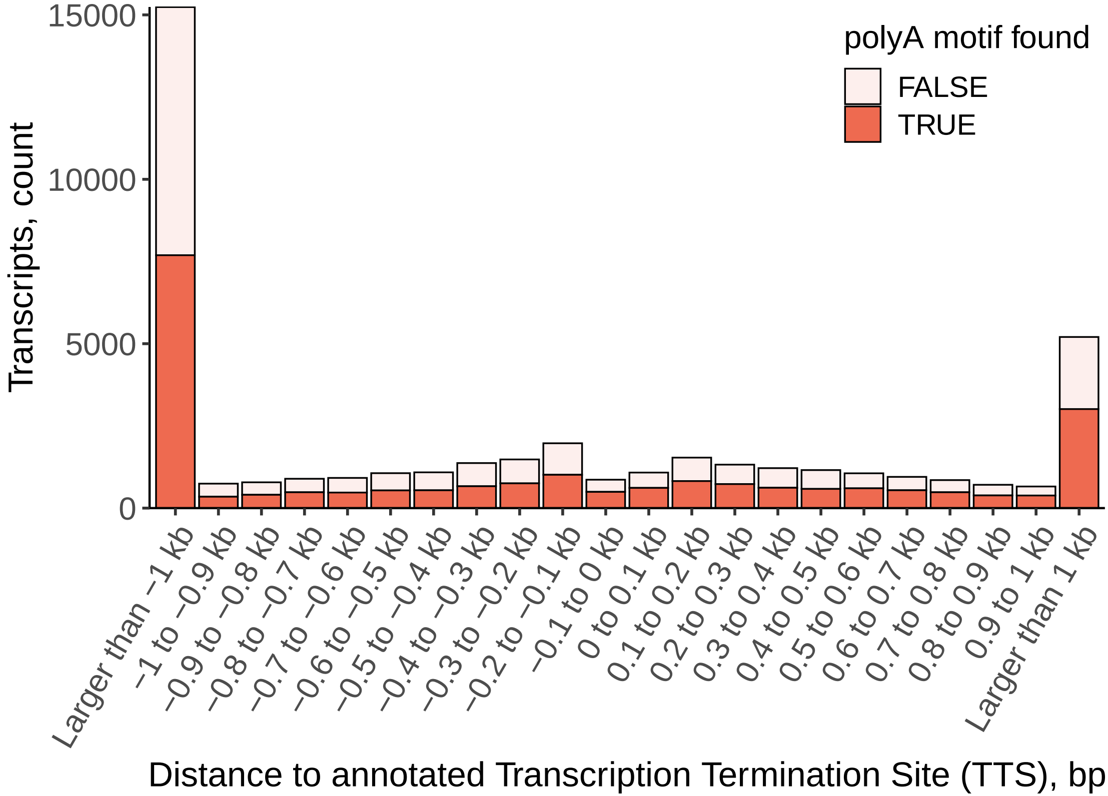
Comparison With Annotated TSS and TTS by Subcategories

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

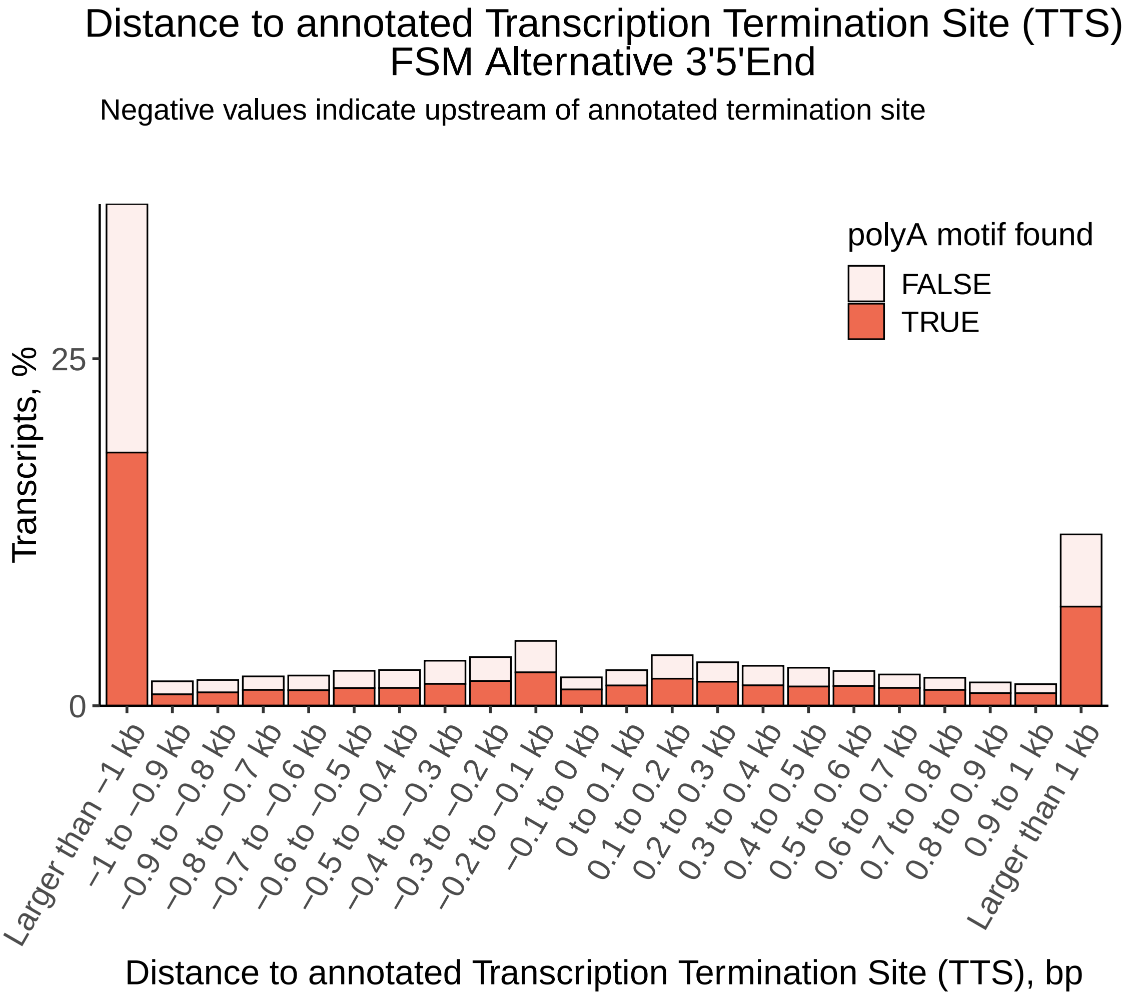
Negative values indicate upstream of annotated termination site




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

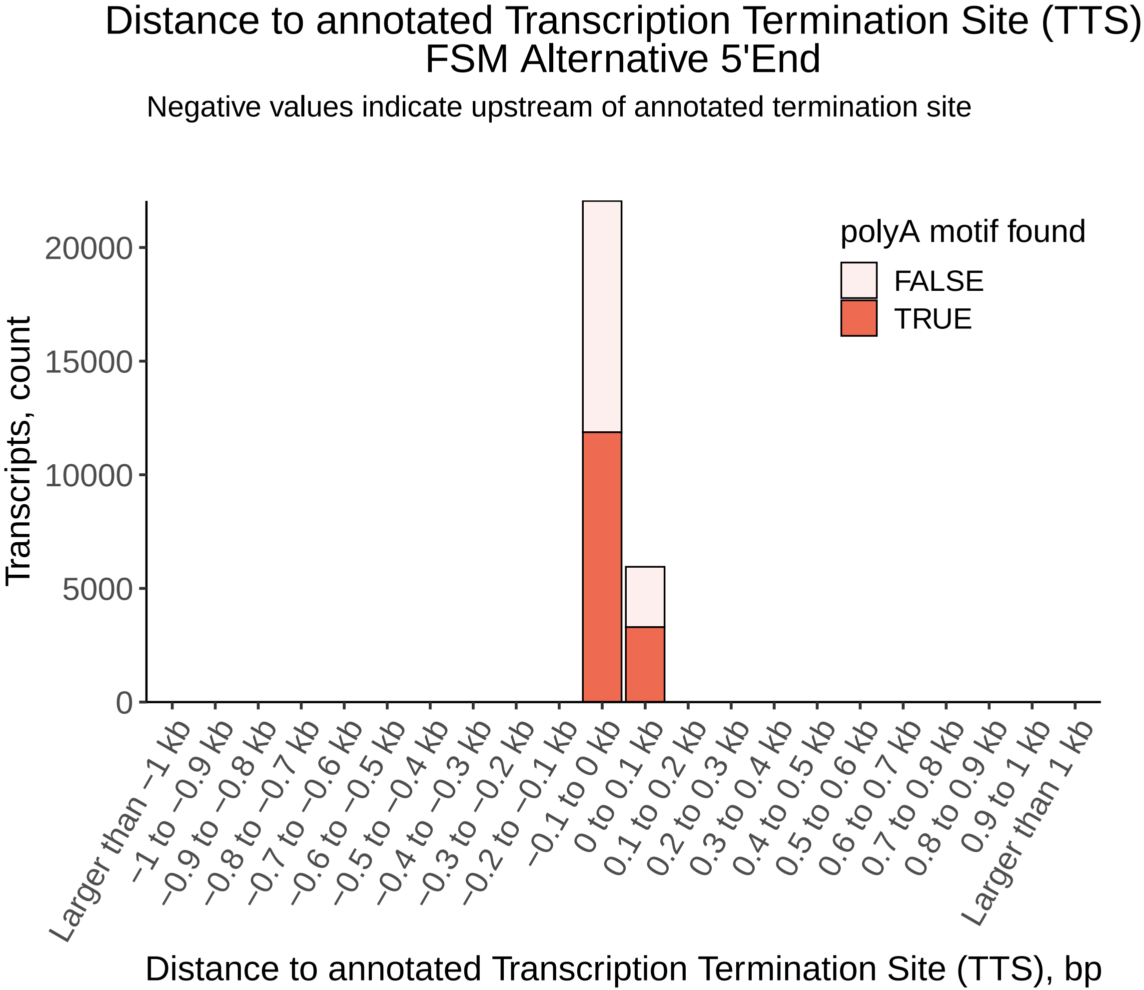

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

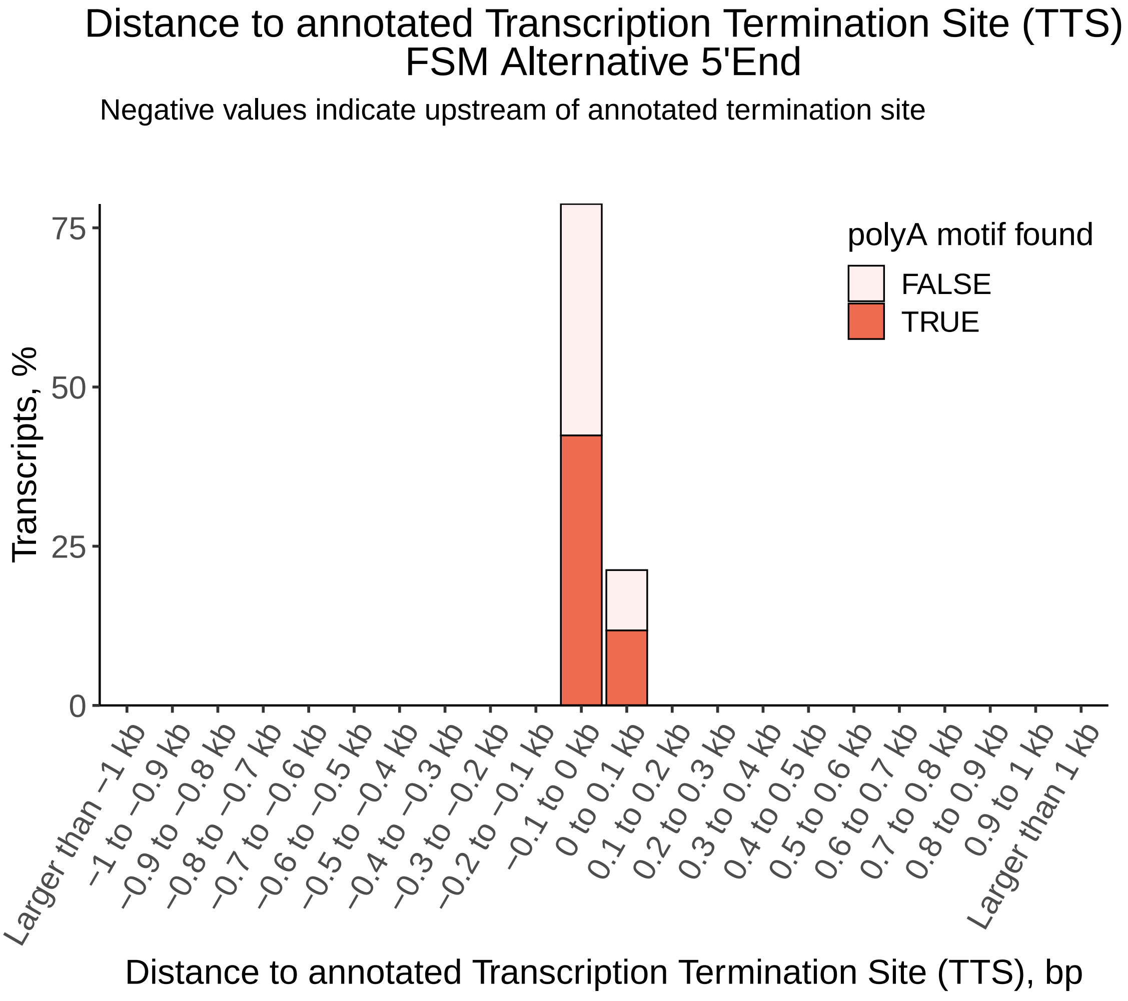
Negative values indicate upstream of annotated termination site

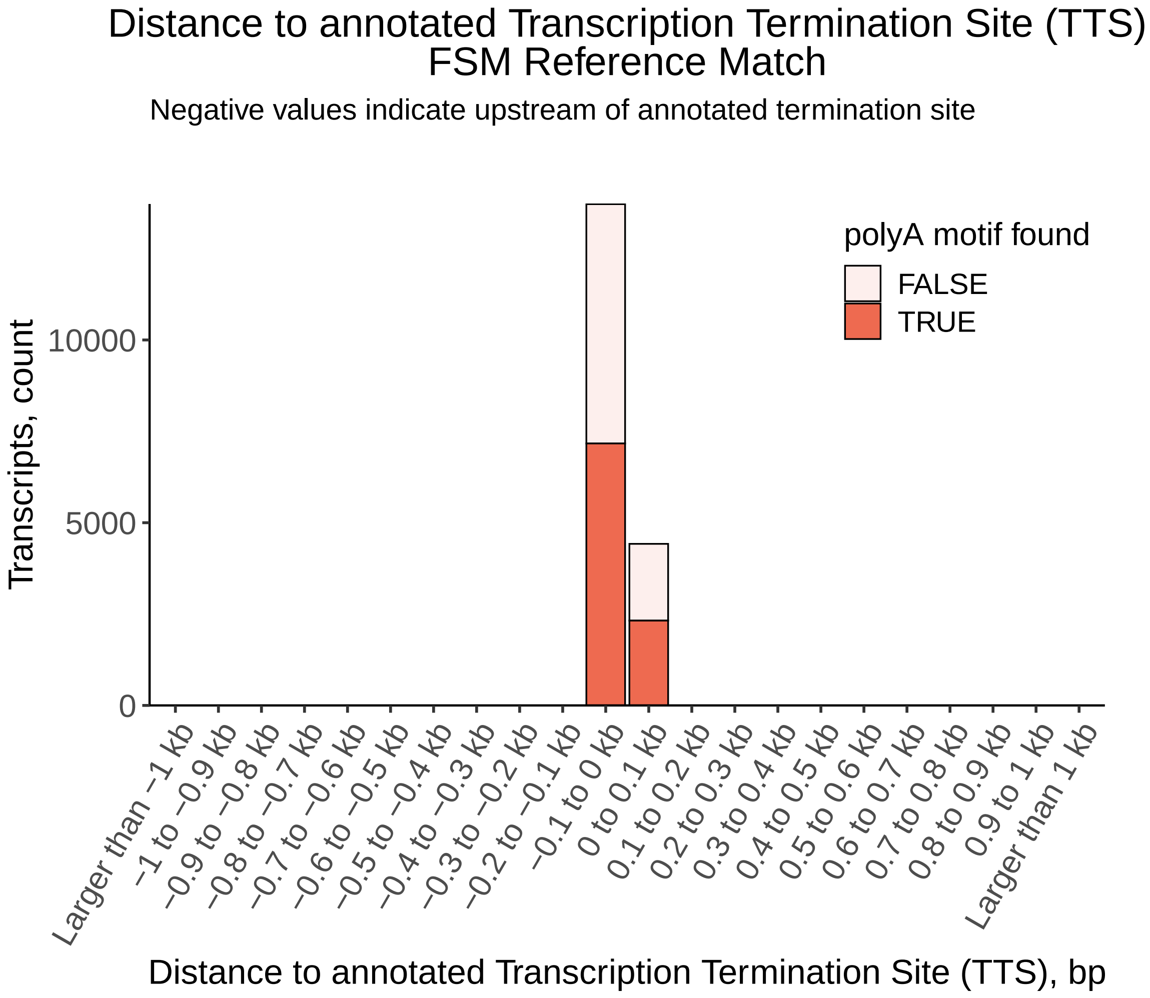


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

Negative values indicate upstream of annotated termination site

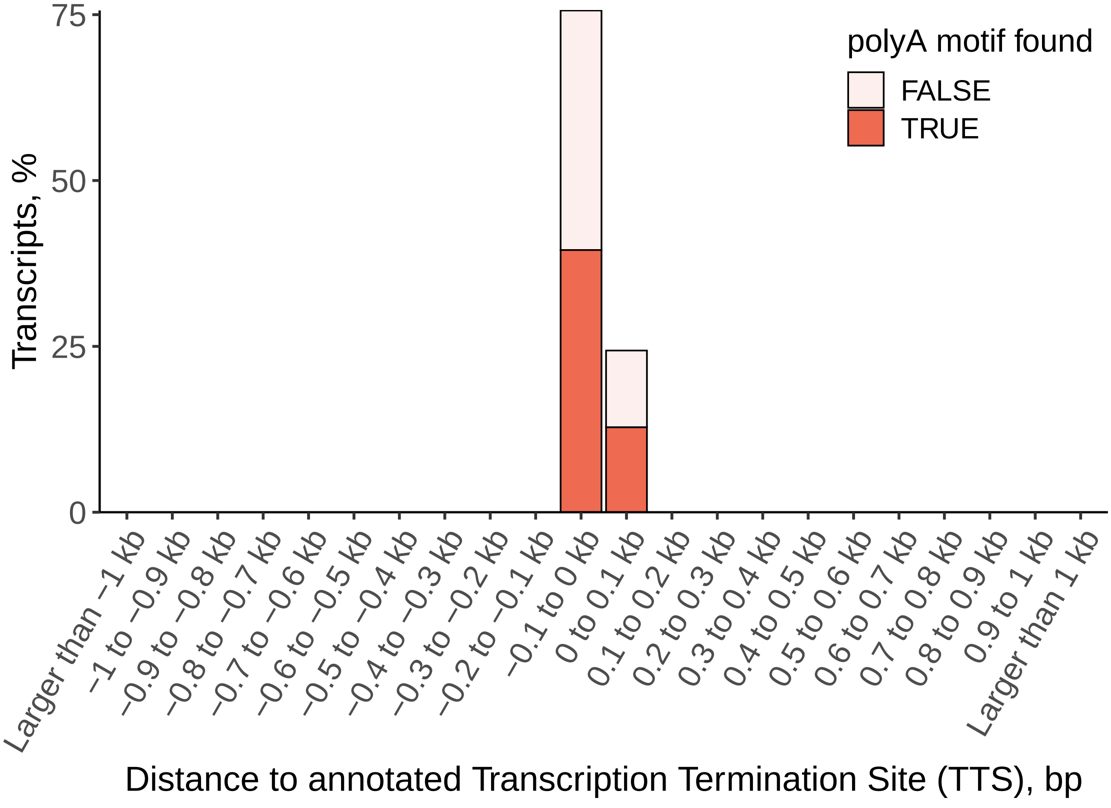

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
Transcripts, count
20000
15000
10000
5000
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6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

Negative values indicate upstream of annotated termination site

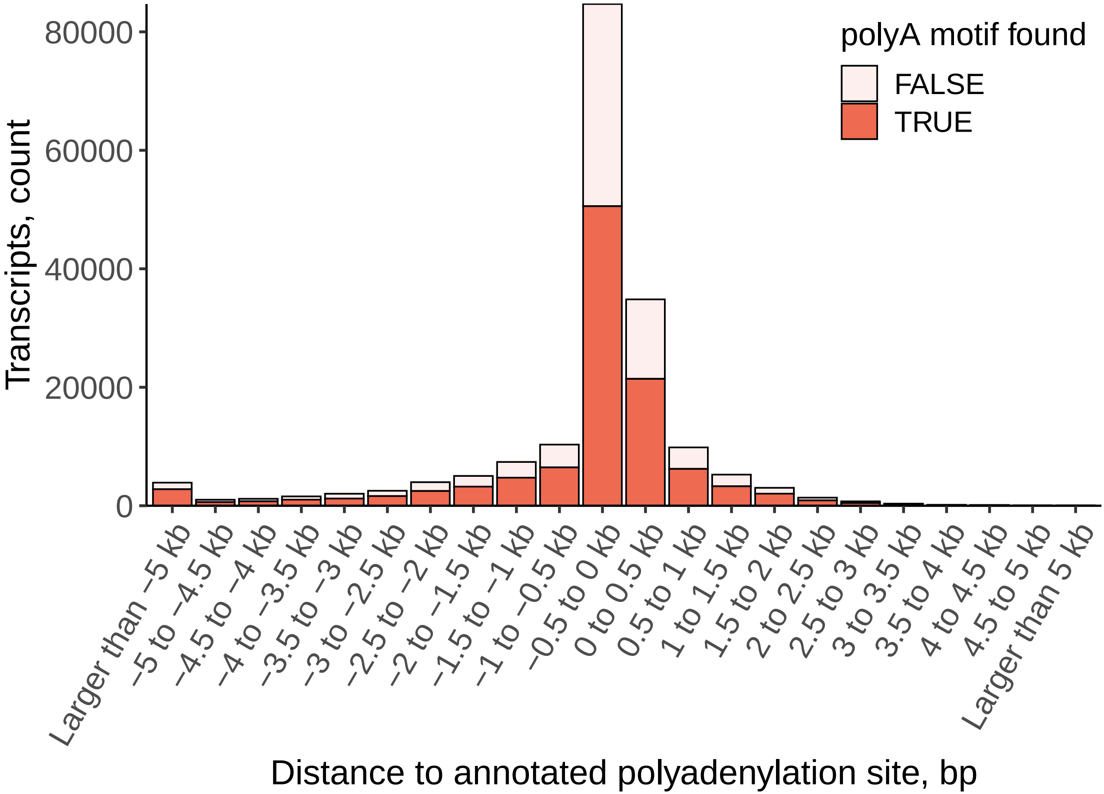


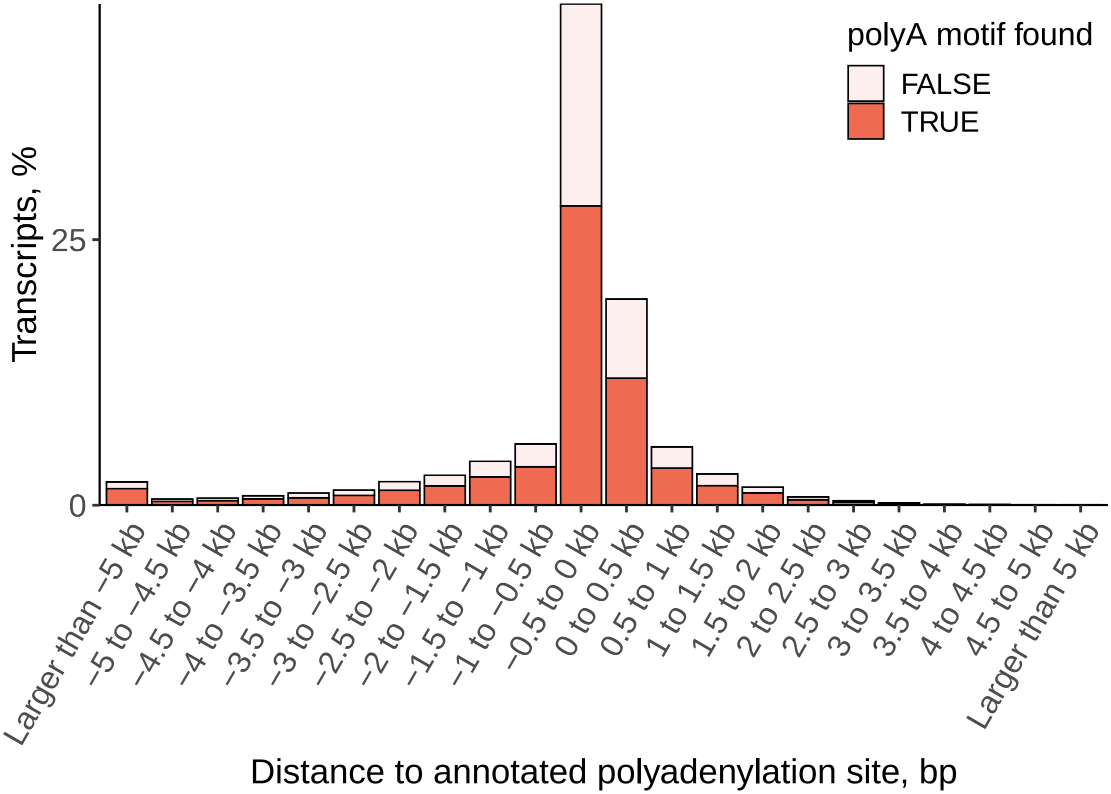
Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match

Negative values indicate upstream of annotated termination site

Distance to Annotated Polyadenylation Site for ISM 3' Fragment

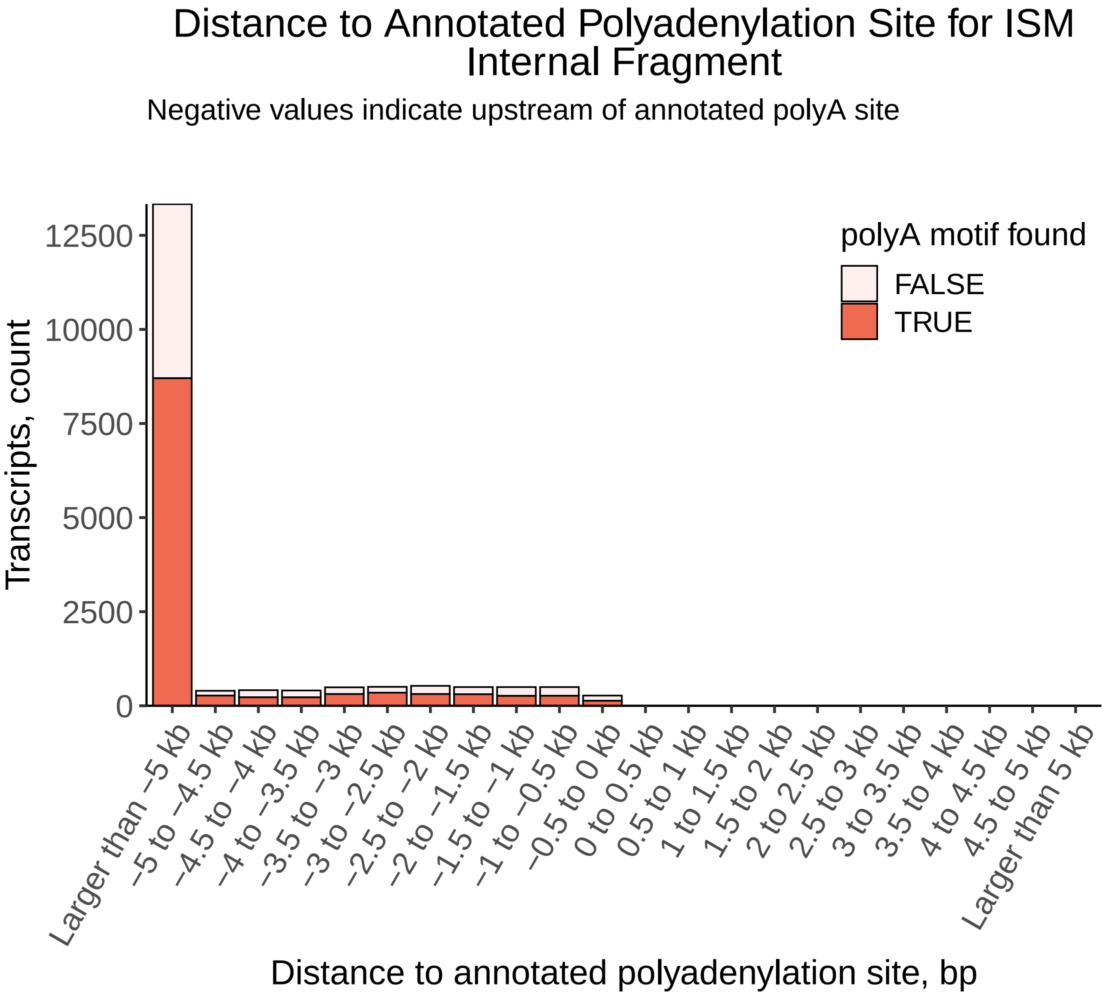
Negative values indicate upstream of annotated polyA site

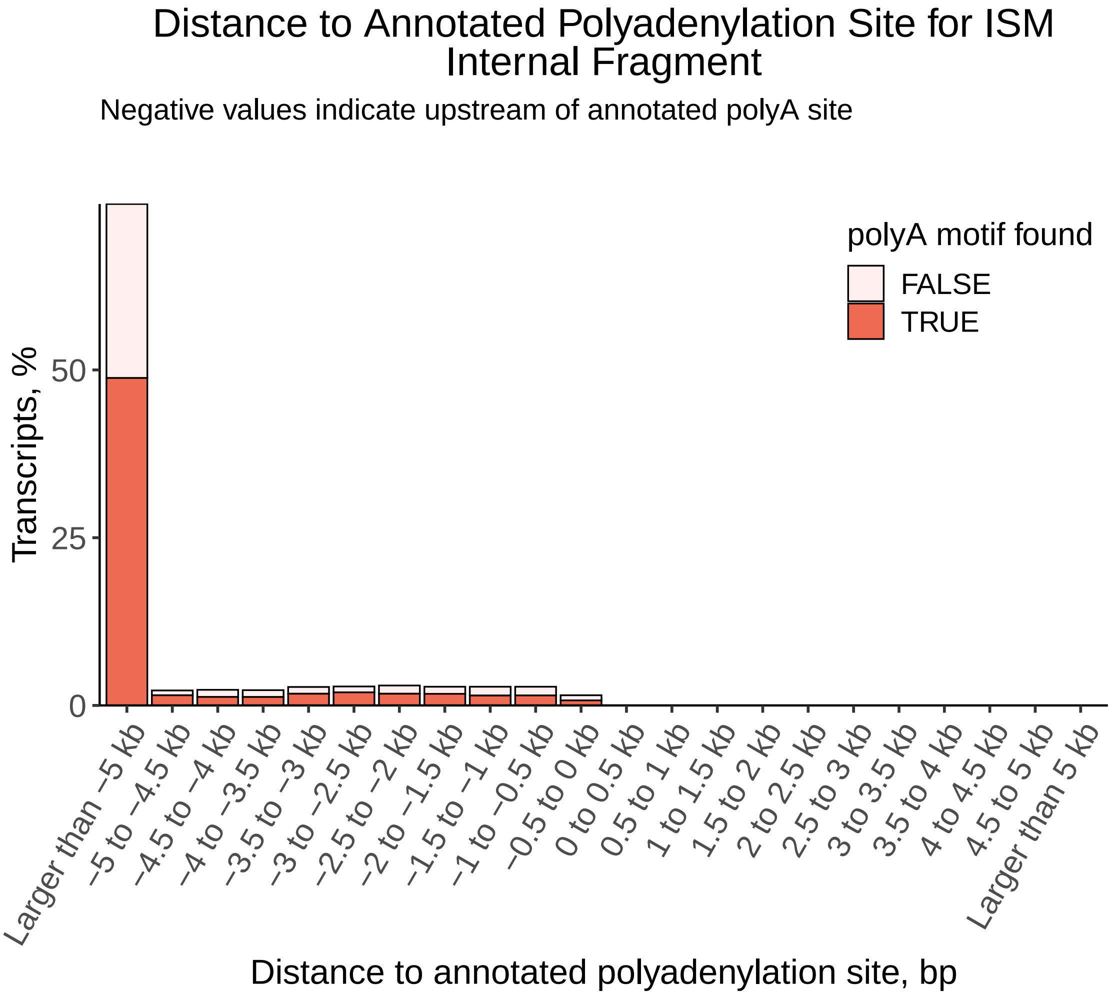


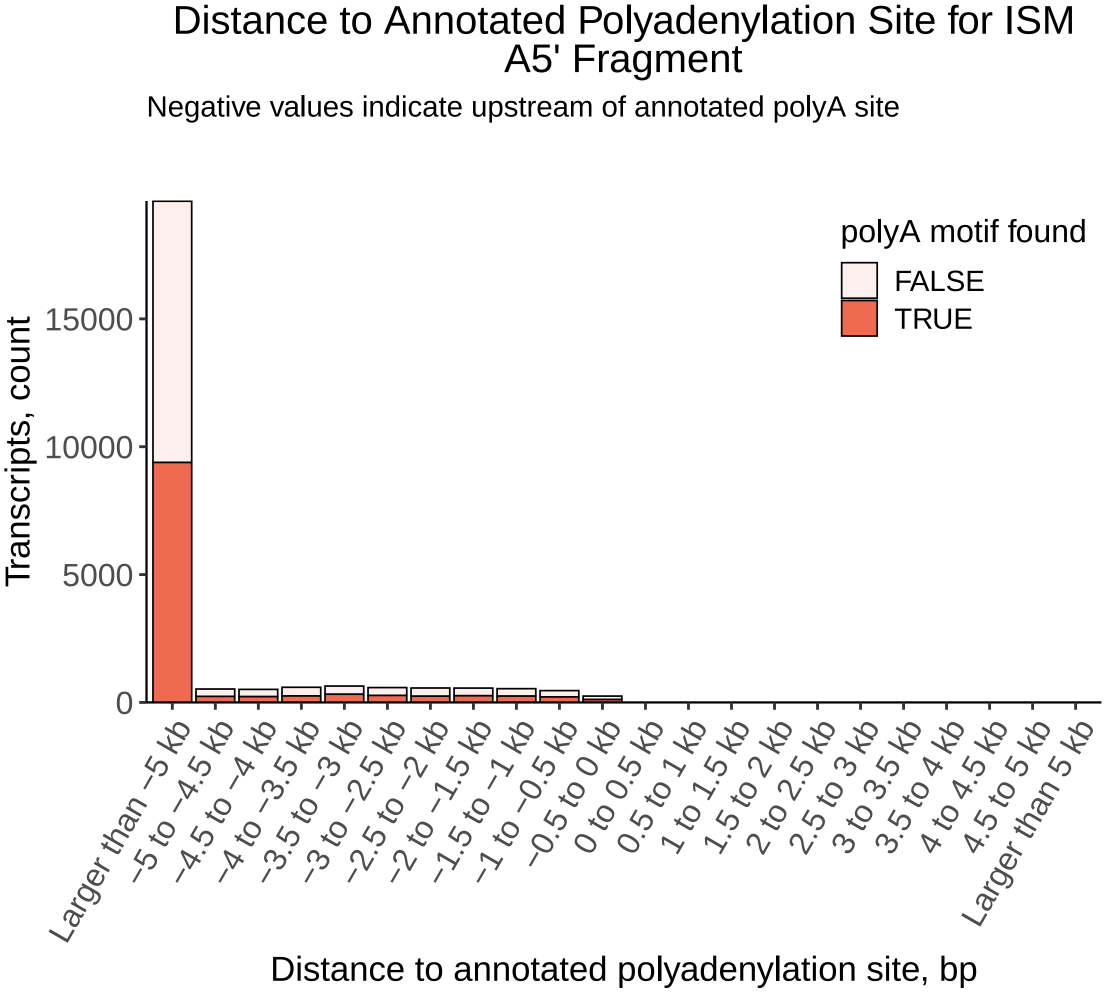


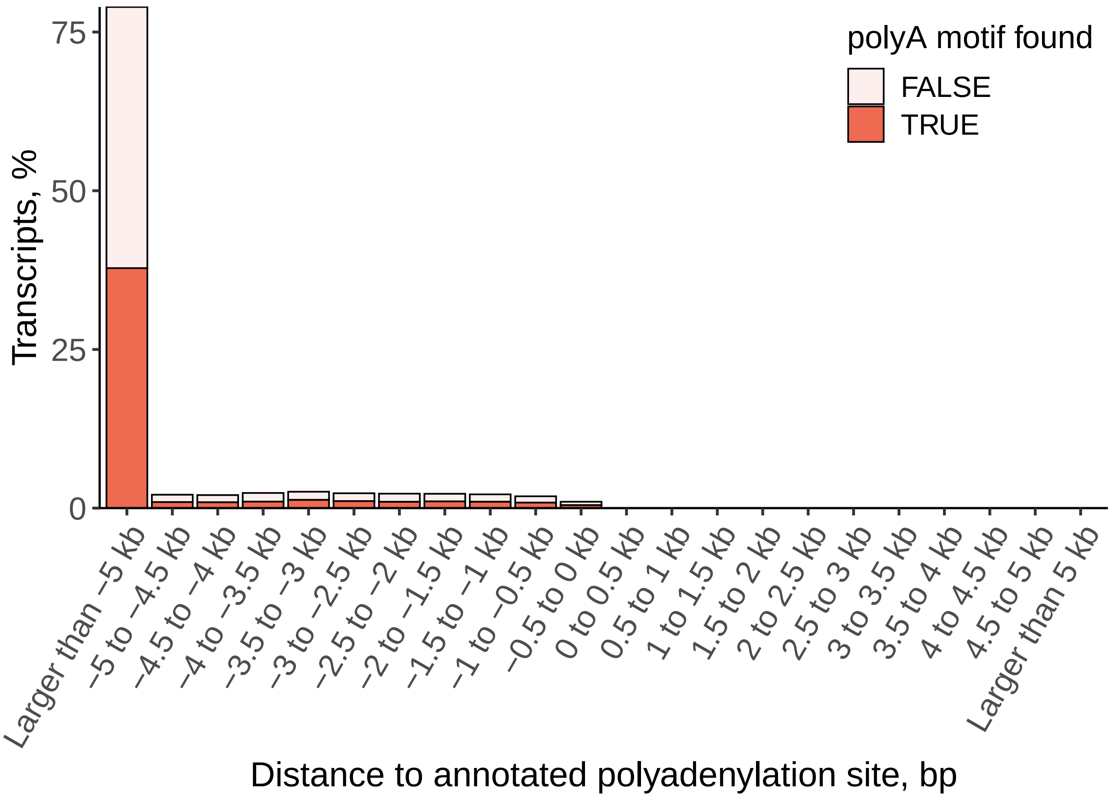
Distance to Annotated Polyadenylation Site for ISM 3' Fragment

Negative values indicate upstream of annotated polyA site


Distance to Annotated Polyadenylation Site for ISM Internal Fragment
Negative values indicate upstream of annotated polyA site
Transcripts, count
12500
10000
7500
5000
2500
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

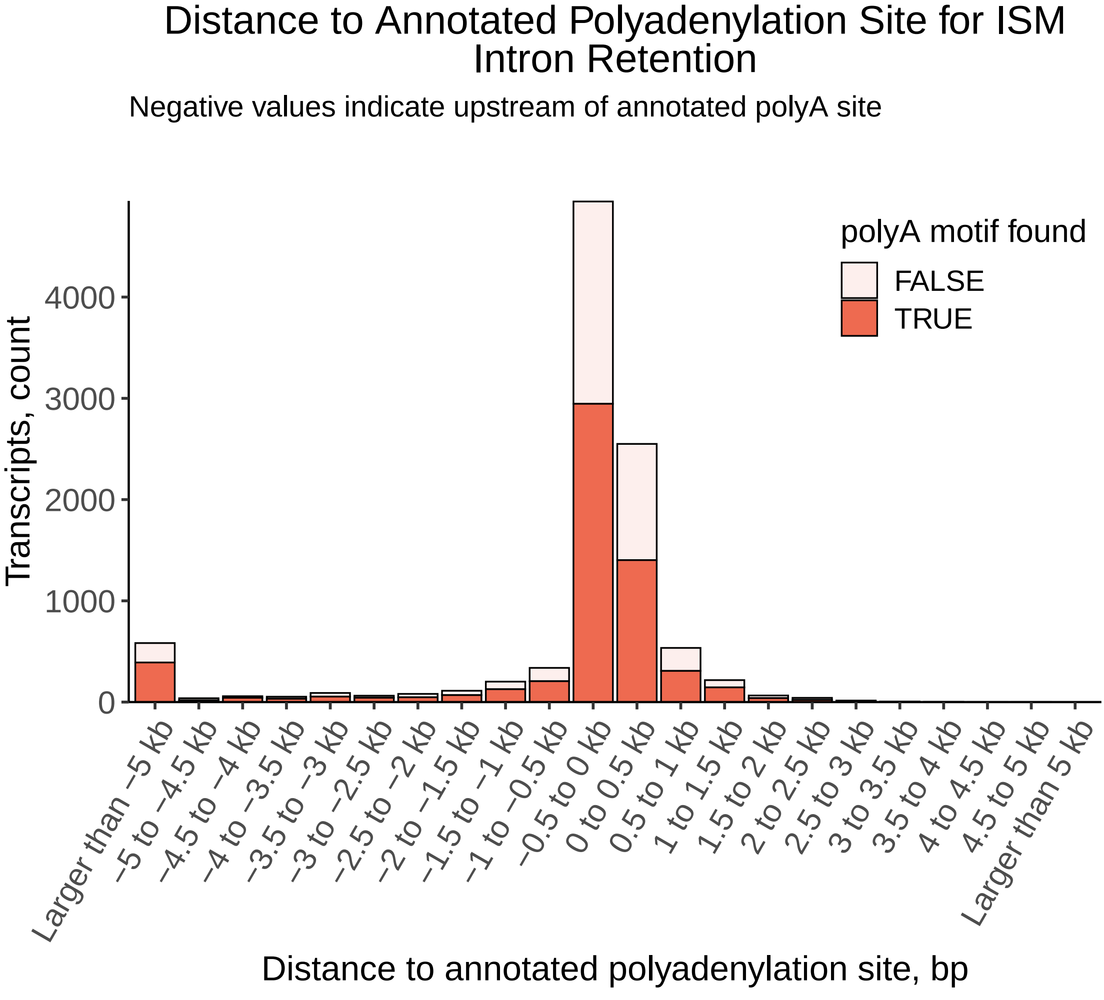

Distance to Annotated Polyadenylation Site for ISM Internal Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to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

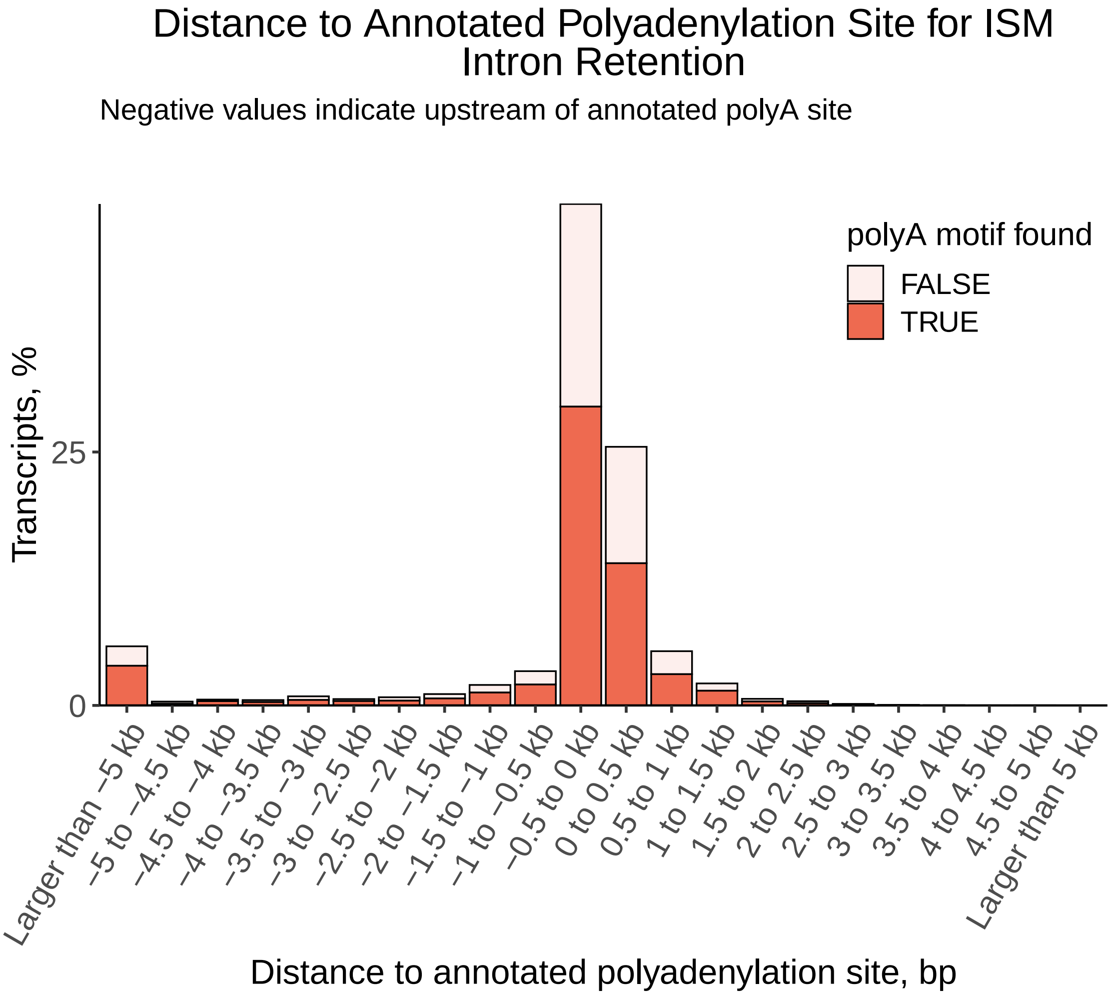

Distance to Annotated Polyadenylation Site for ISM A5' Fragment
Negative values indicate upstream of annotated polyA site
polyA motif found
FALSE
TRUE
Transcripts, count
0
5000
10000
15000
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5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



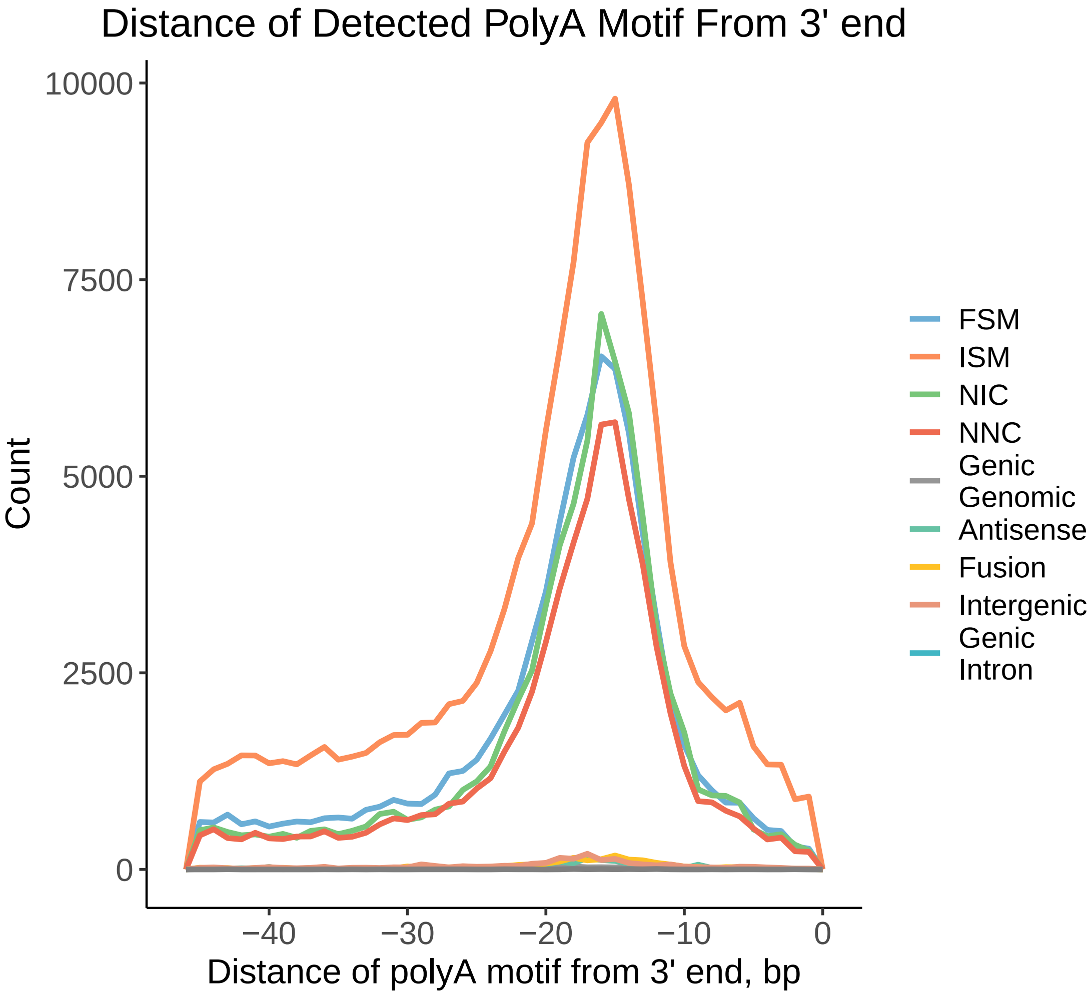
Distance to Annotated Polyadenylation Site for ISM A5' Fragment

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
4000
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, %
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

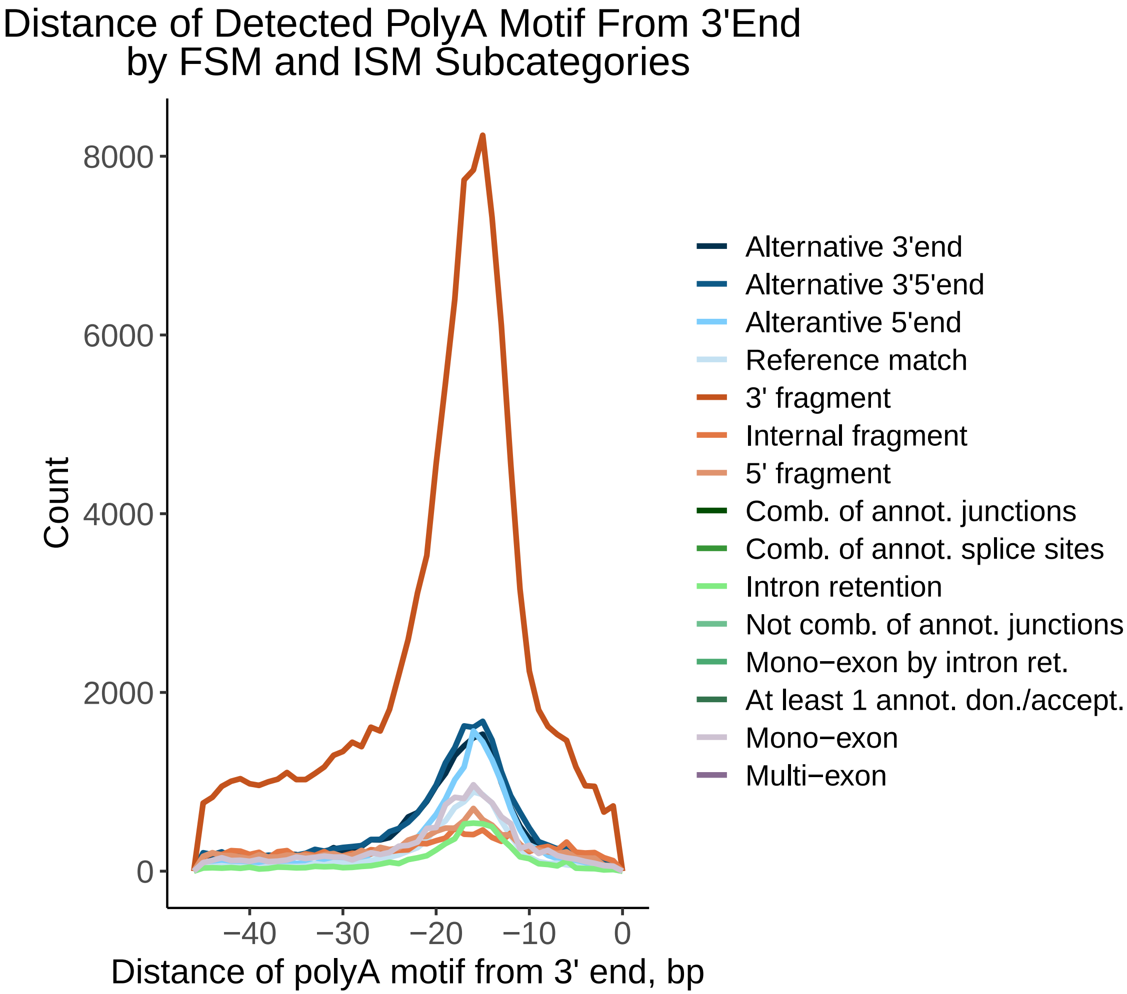

Distance of Detected PolyA Motif From 3' end
Count
10000
7500
5000
2500
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

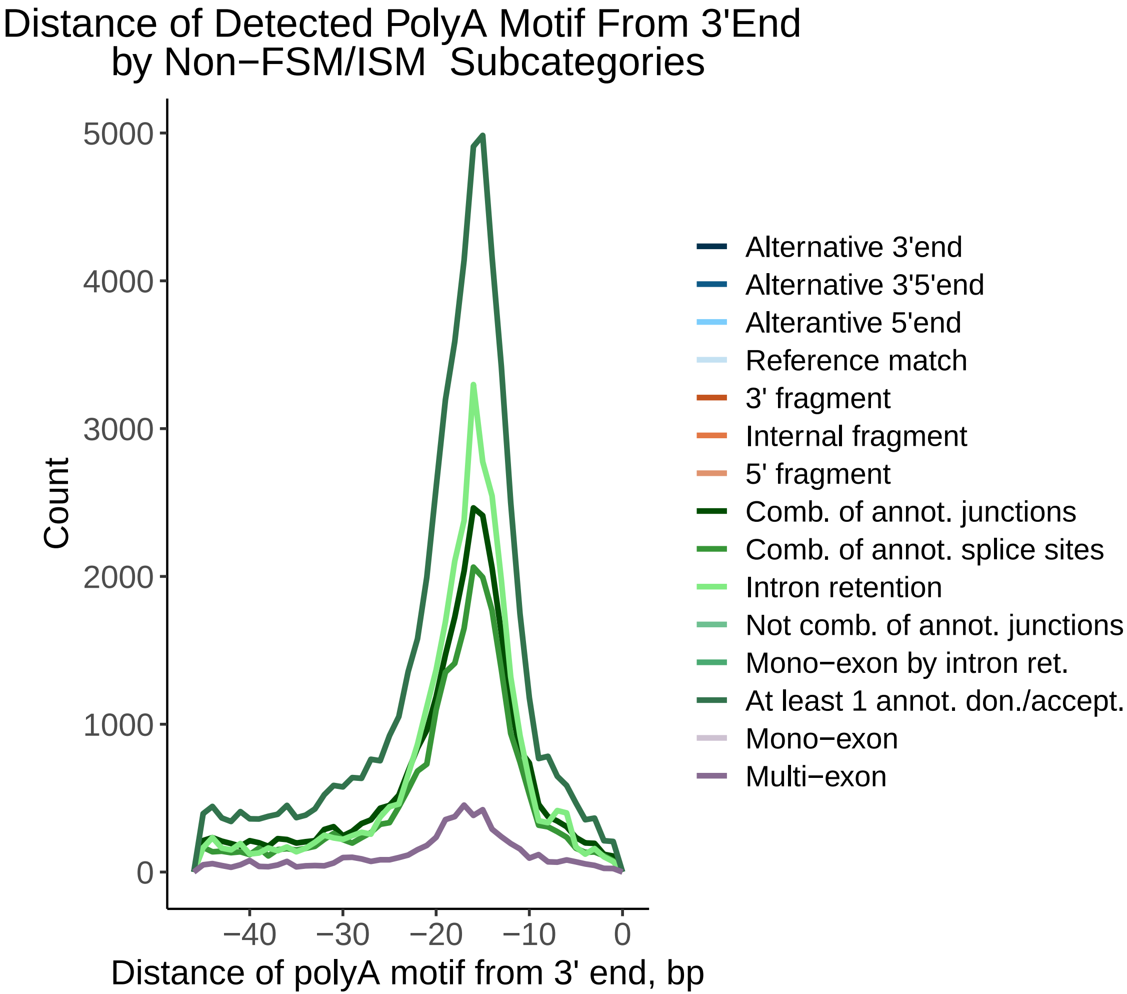
Frequency of PolyA Motifs

Motif	Count	%
AATAAA	180961	49.7
ATTAAA	53380	14.7
TATAAA	14742	4.0
AGTAAA	13638	3.7
AAAAAG	13211	3.6
AAGAAA	12949	3.6
GGGGCT	11038	3.0
TTTAAA	9813	2.7
AATATA	8255	2.3
CATAAA	8091	2.2
GATAAA	7737	2.1
AATACA	7642	2.1
AAAACA	7334	2.0
AATGAA	7000	1.9
AATAGA	4272	1.2
ACTAAA	3979	1.1

Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	151552	79768	53
ISM	232135	139402	60
NIC	143517	74422	52
NNC	119532	64557	54
Genic Genomic	819	437	53
Antisense	2534	1448	57
Fusion	3334	1903	57
Intergenic	3212	2061	64
NA	68	44	65


Distance of Detected PolyA Motif From 3'End
by FSM and ISM Subcategories
Count
8000
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
5000
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

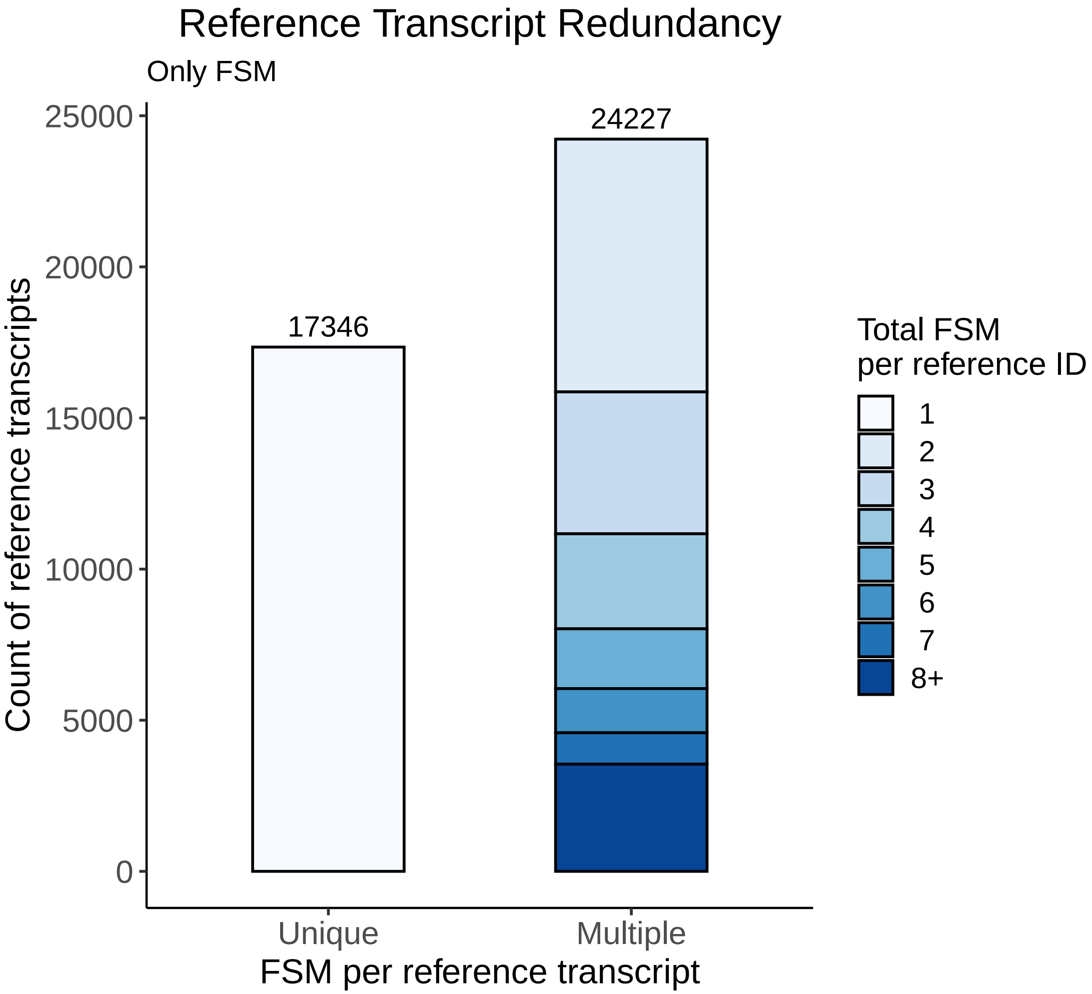
Number of polyA Motifs Detected

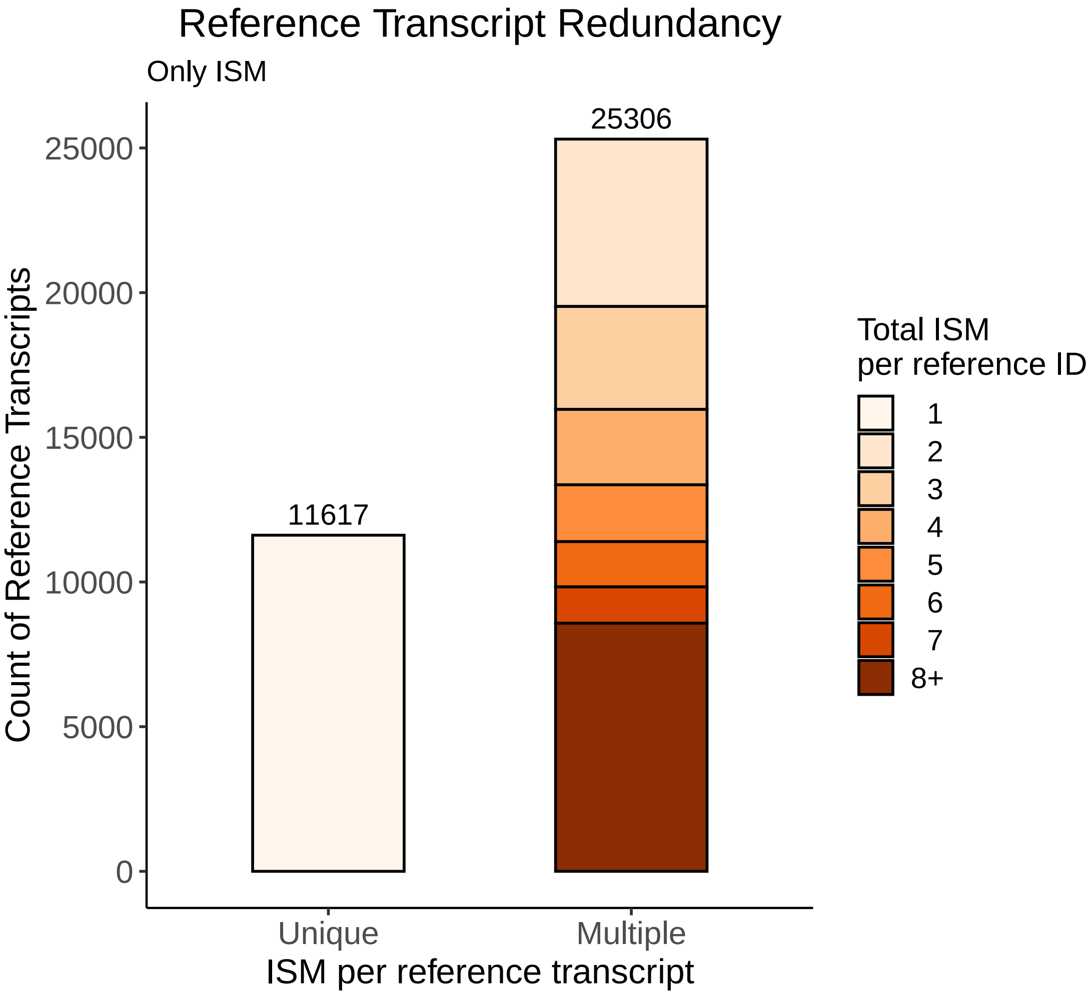
Subcategory	Count	polyA Detected	%
Alternative 3'end	42223	20188	48
Alternative 3'5'end	42160	22184	53
Alterantive 5'end	28000	15168	54
Reference match	18137	9489	52
3' fragment	179475	110349	61
Internal fragment	17836	11365	64
5' fragment	24829	11783	47
Comb. of annot. junctions	54805	28353	52
Comb. of annot. splice sites	43723	23024	53
Intron retention	69745	36547	52
At least 1 annot. don./accept.	105389	57271	54
Mono−exon	21032	12739	61
Multi−exon	9288	5545	60
NA	61	37	61

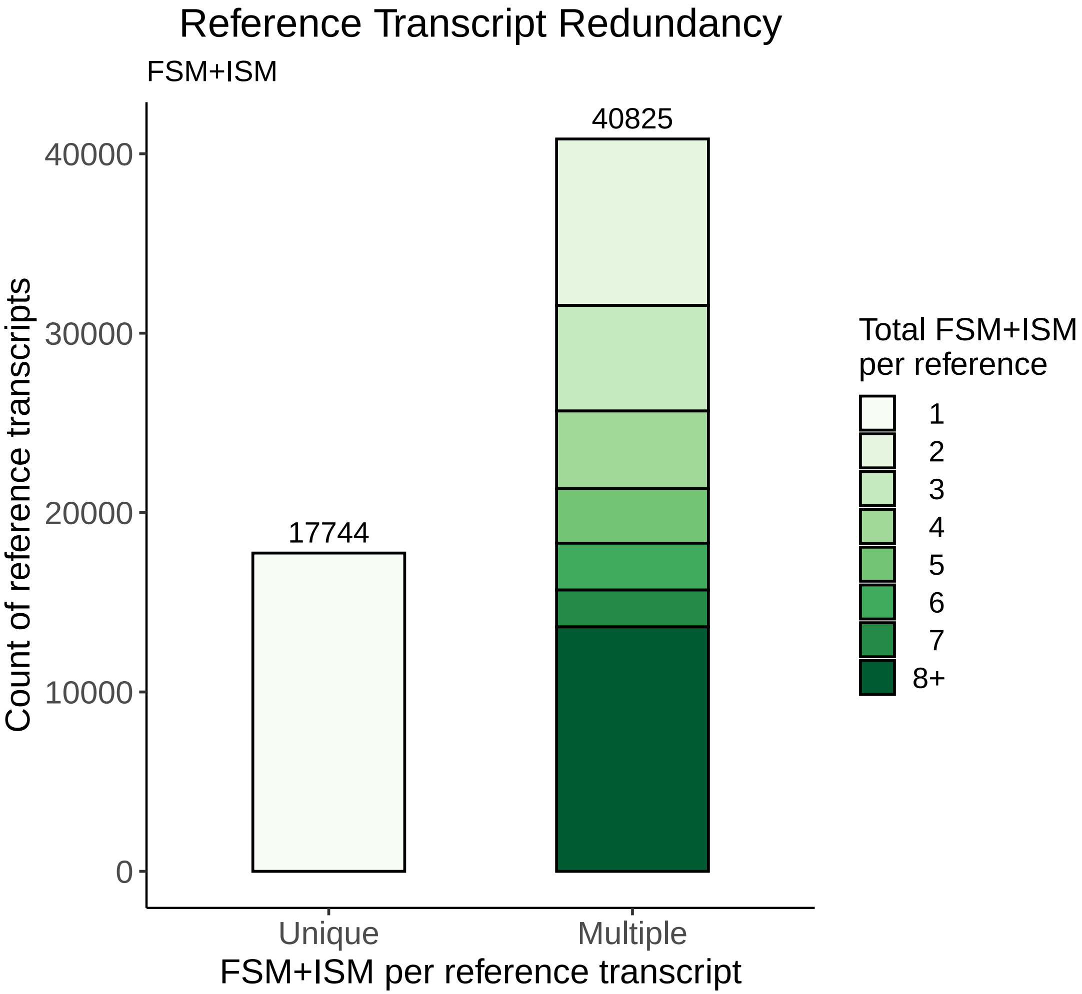
Frequency of PolyA Motifs

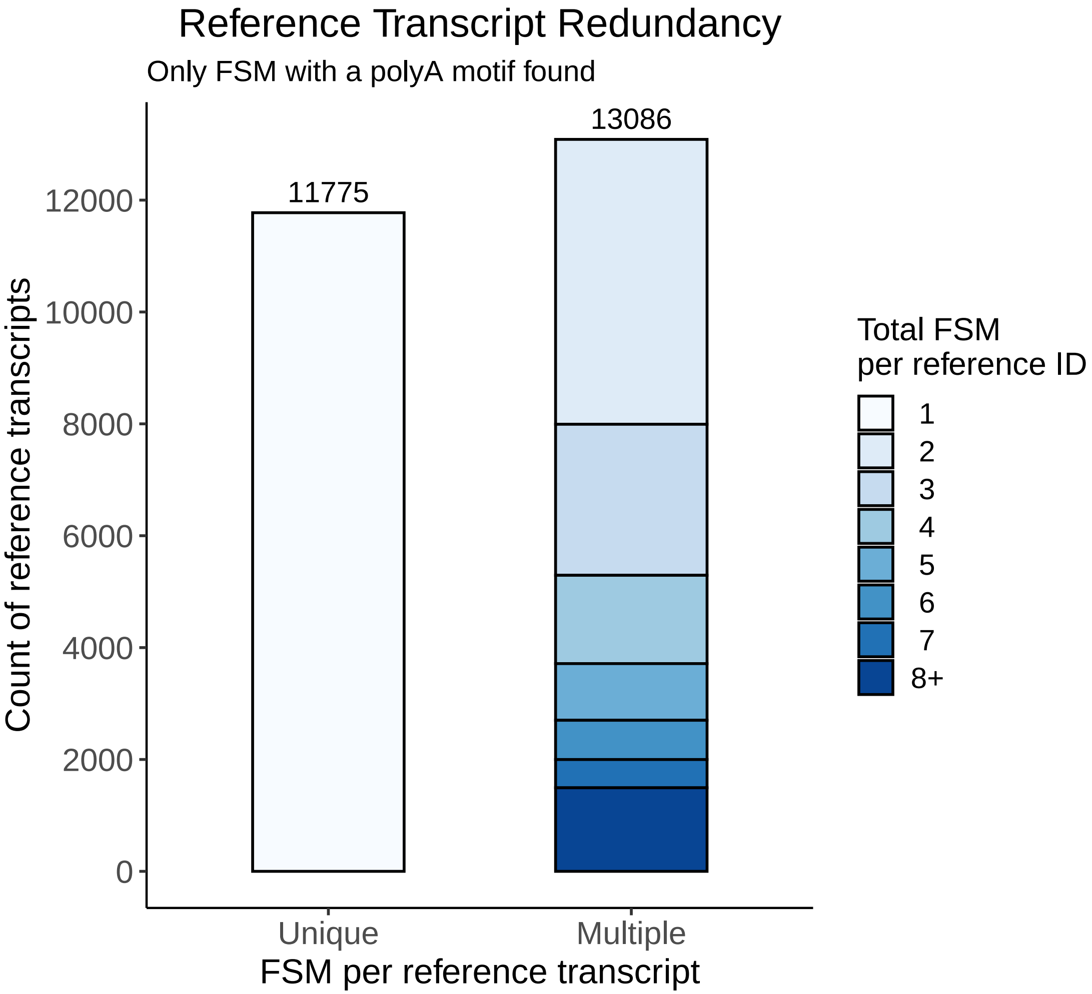
Motif	Count	%
AATAAA	180961	49.7
ATTAAA	53380	14.7
TATAAA	14742	4.0
AGTAAA	13638	3.7
AAAAAG	13211	3.6
AAGAAA	12949	3.6
GGGGCT	11038	3.0
TTTAAA	9813	2.7
AATATA	8255	2.3
CATAAA	8091	2.2
GATAAA	7737	2.1
AATACA	7642	2.1
AAAACA	7334	2.0
AATGAA	7000	1.9
AATAGA	4272	1.2
ACTAAA	3979	1.1

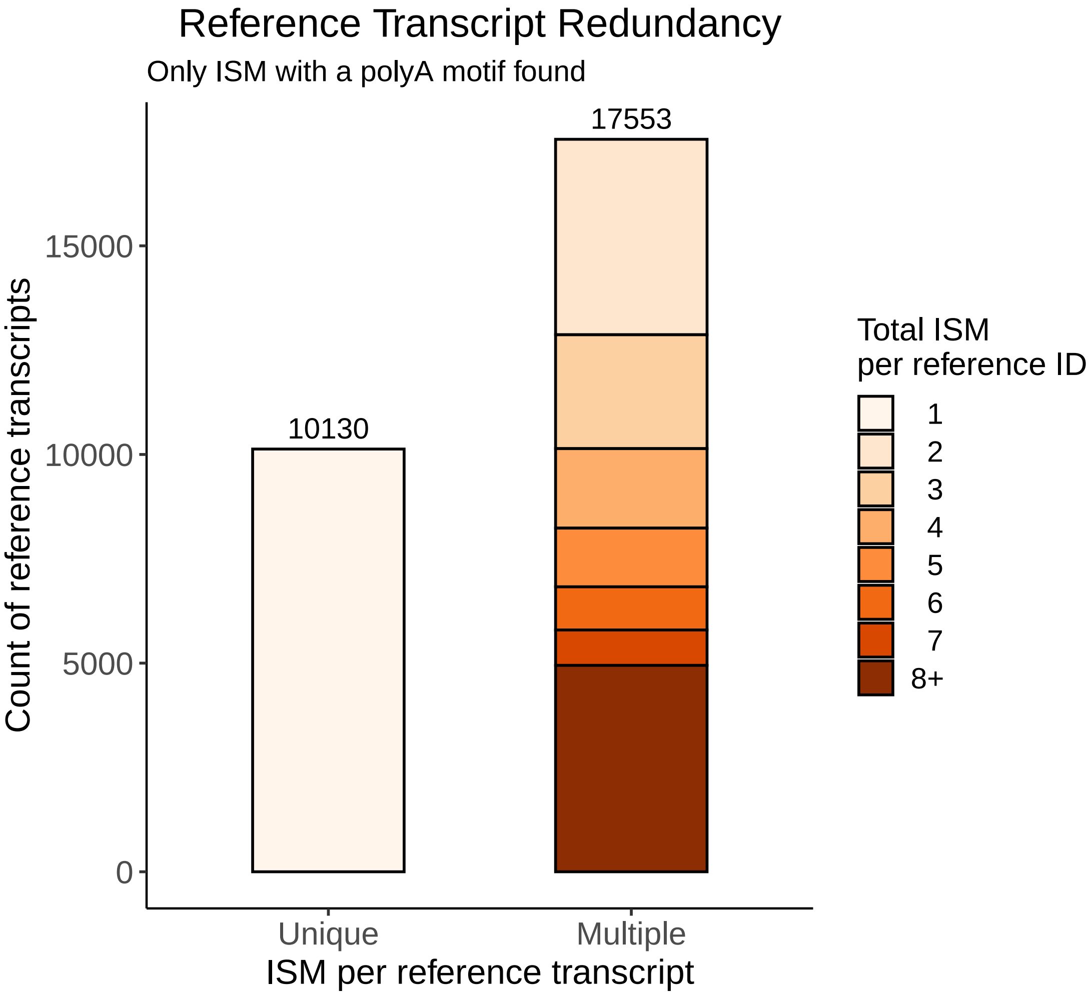
Redundancy Analysis

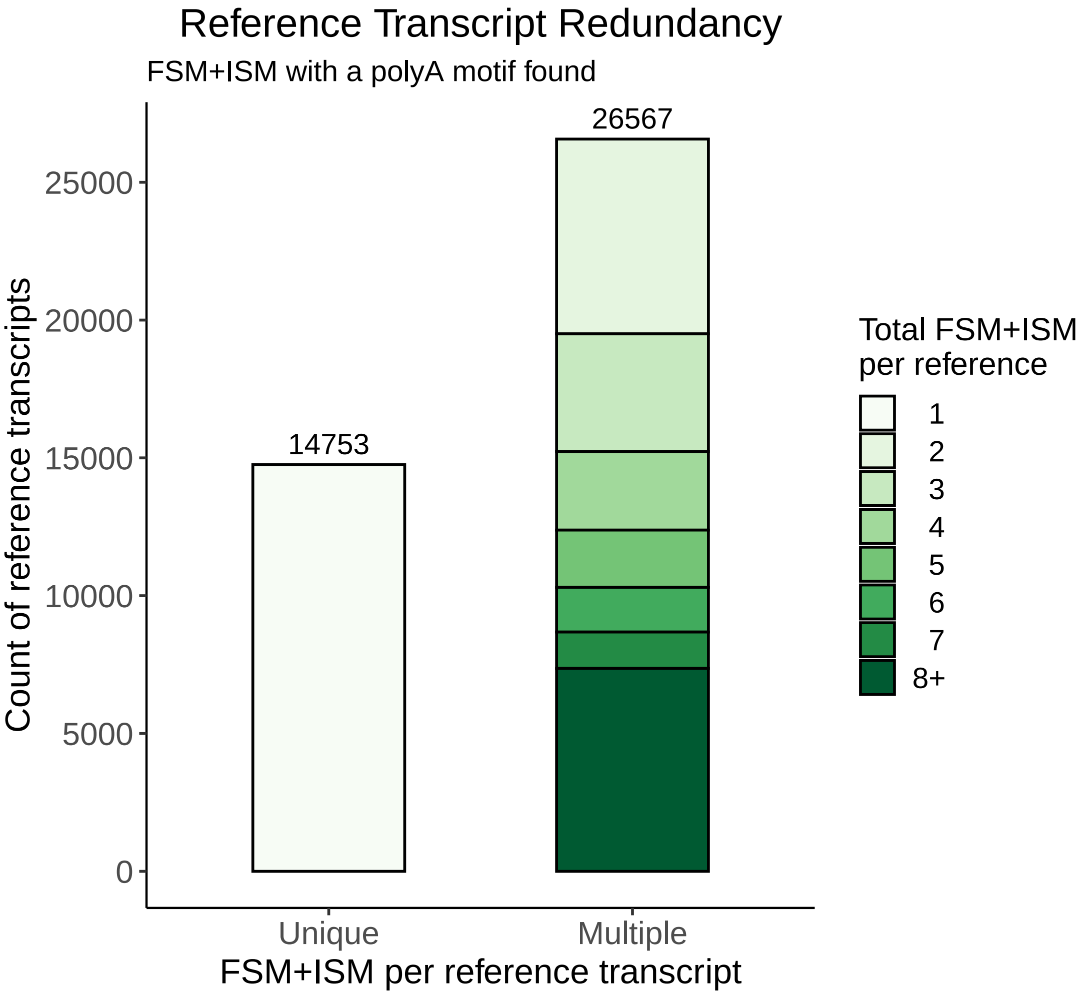

Reference Transcript Redundancy
Only FSM
Count of reference transcripts
25000
20000
15000
10000
5000
0
17346
24227
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only ISM
Count of Reference Transcripts
0
5000
10000
15000
20000
25000
11617
25306
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

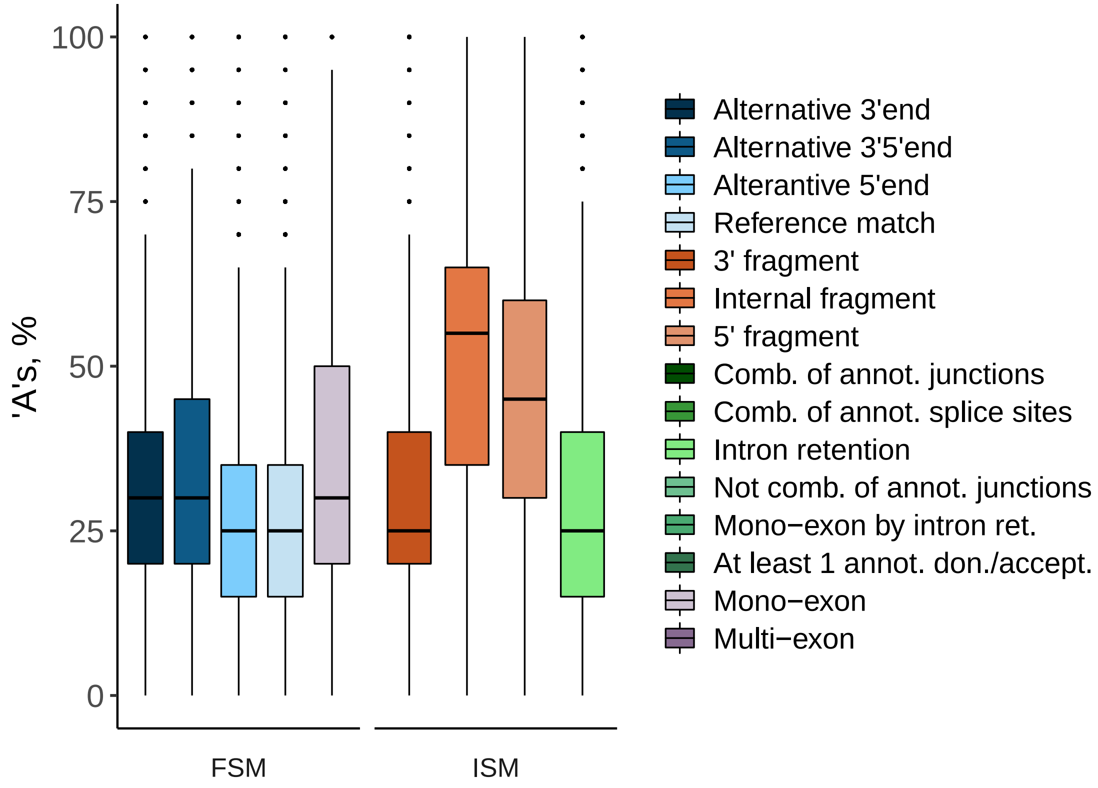

Reference Transcript Redundancy
FSM+ISM
40825
17744
40000
30000
20000
10000
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only FSM with a polyA motif found
11775
13086
12000
10000
8000
6000
4000
2000
0
Count of reference transcripts
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3
4
5
6
7
8+


Reference Transcript Redundancy
Only ISM with a polyA motif found
Count of reference transcripts
0
5000
10000
15000
10130
17553
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
0
5000
10000
15000
20000
25000
14753
26567
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+

Intra−Priming Quality Check

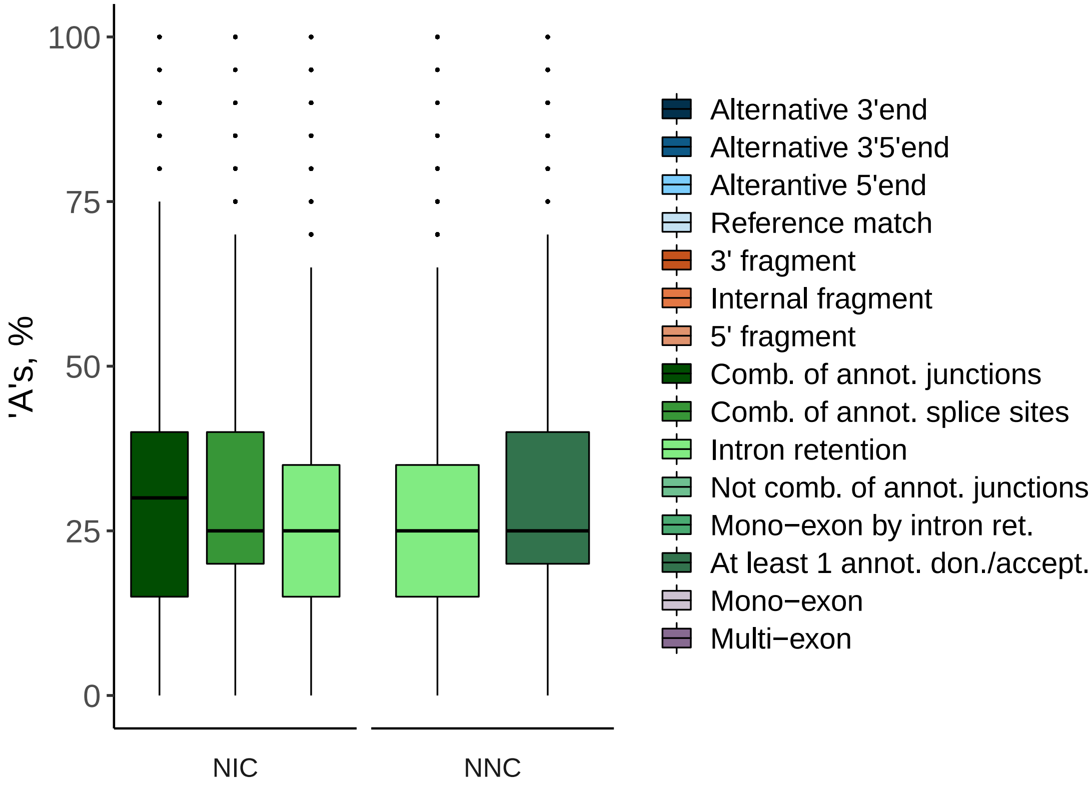


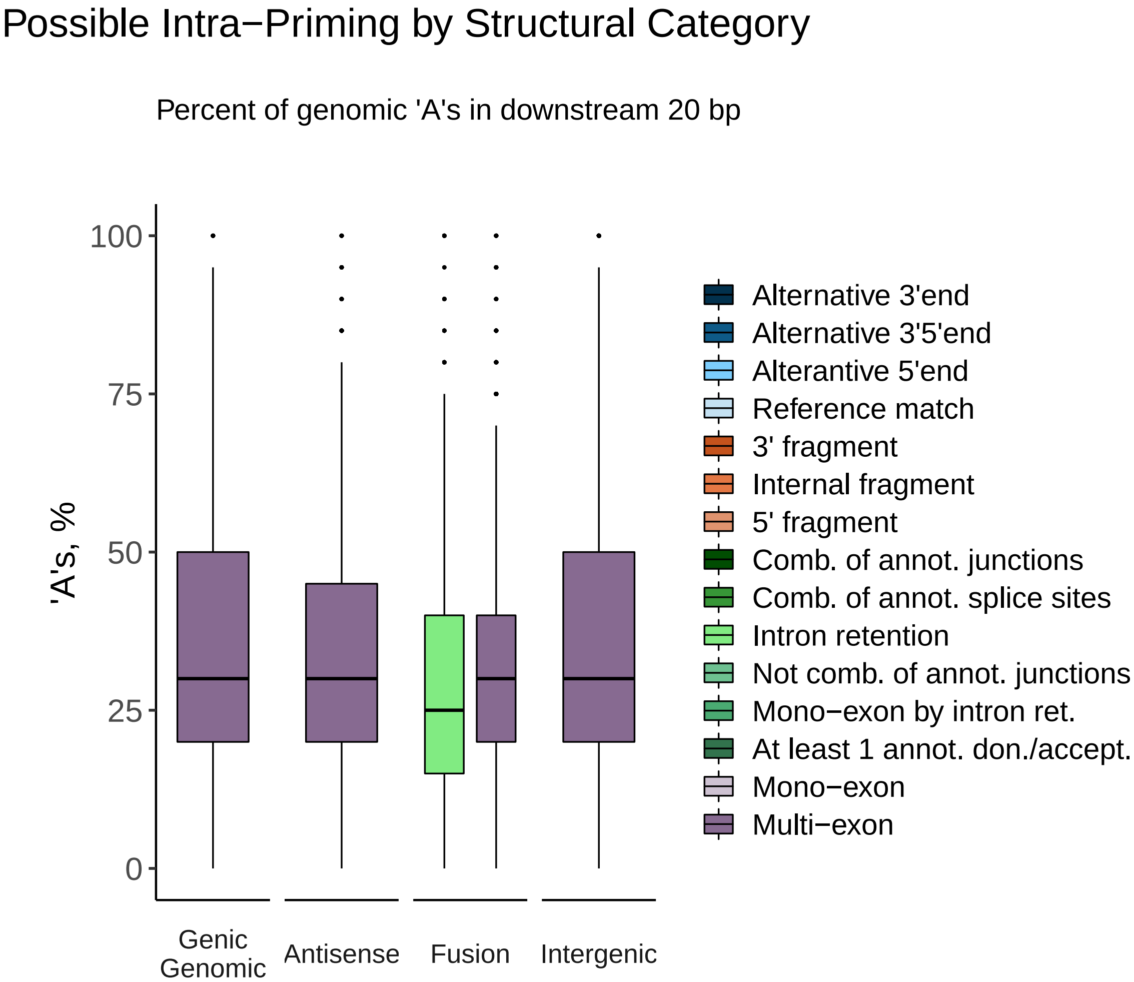
Possible Intra−Priming by Structural Category

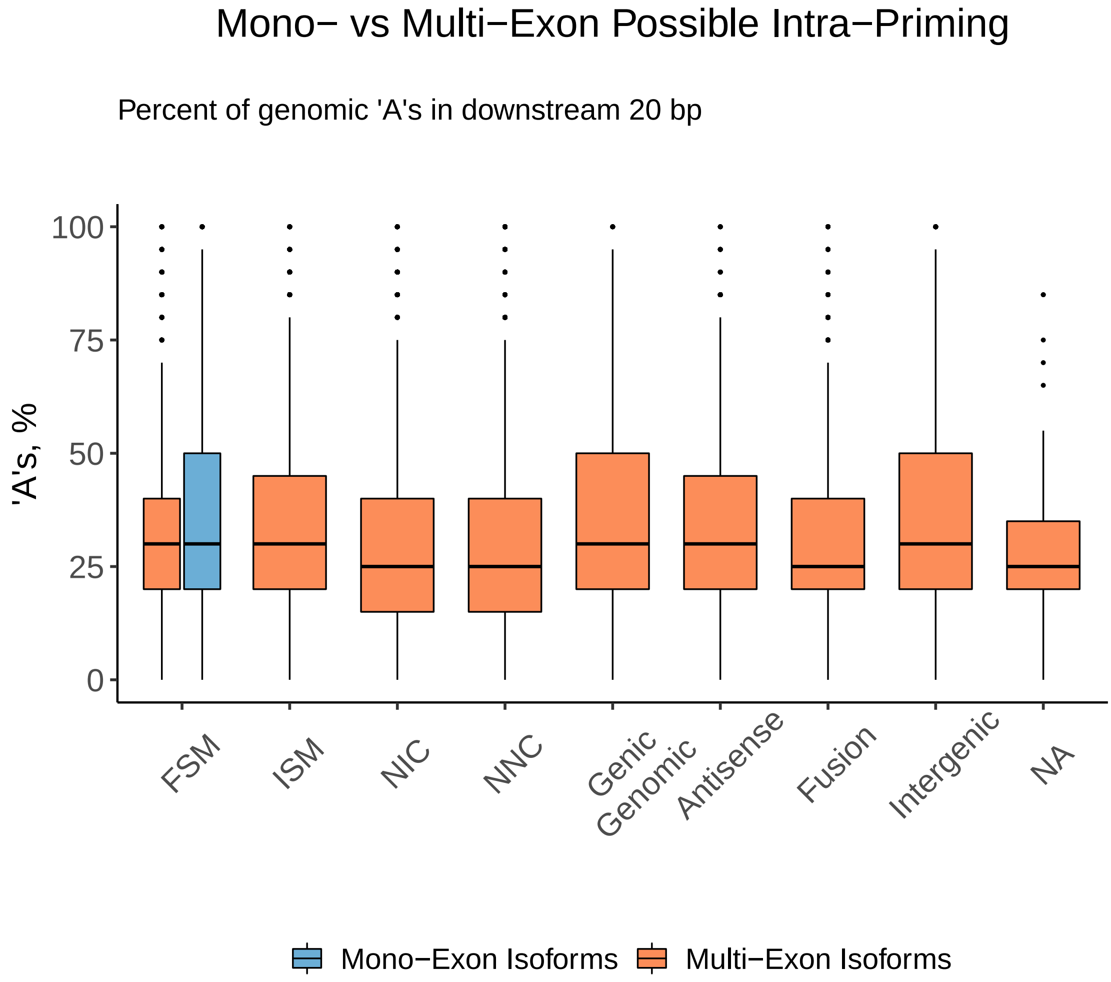
Percent of genomic 'A's in downstream 20 bp

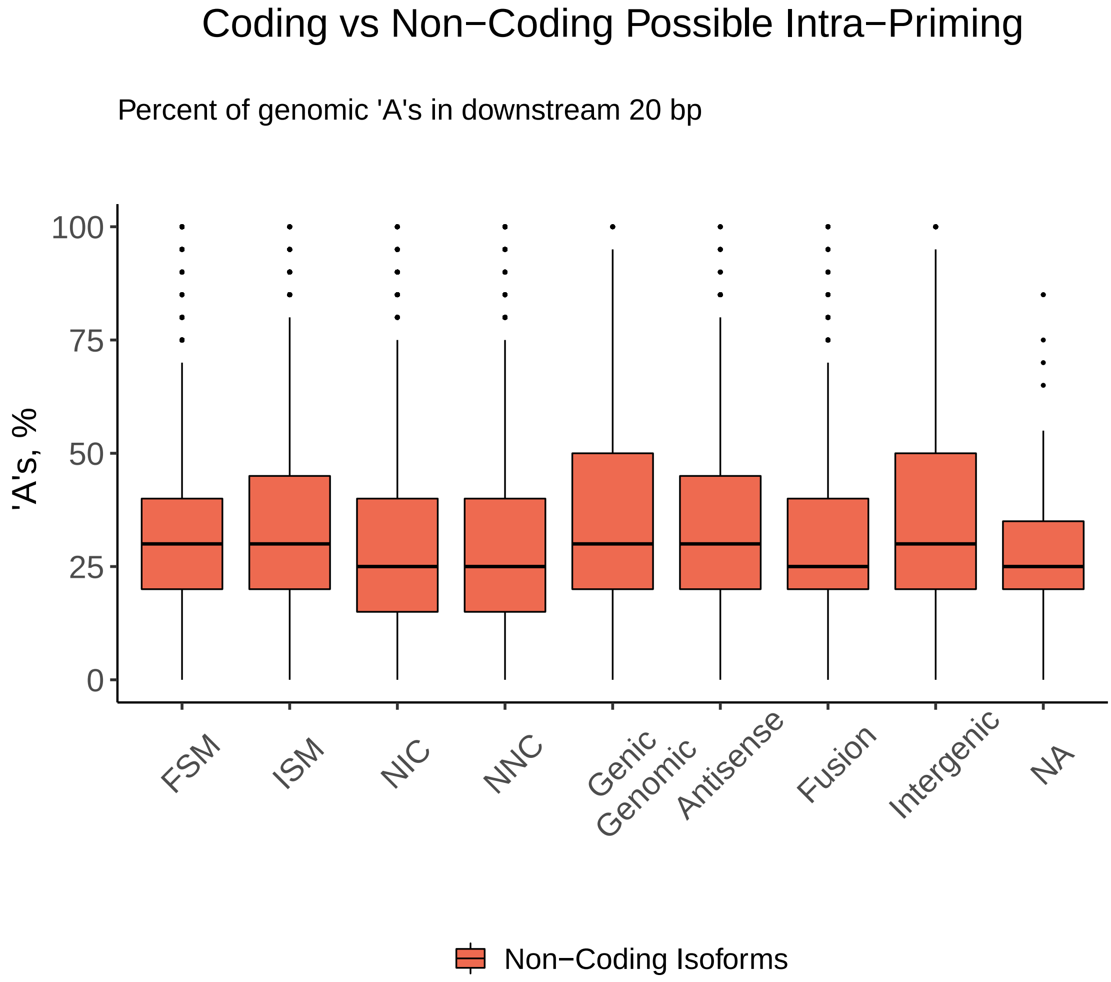
Possible Intra−Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp




Possible Intra−Priming by Structural Category
Percent of genomic 'A's in downstream 20 bp
'A's, %
0
25
50
75
100
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

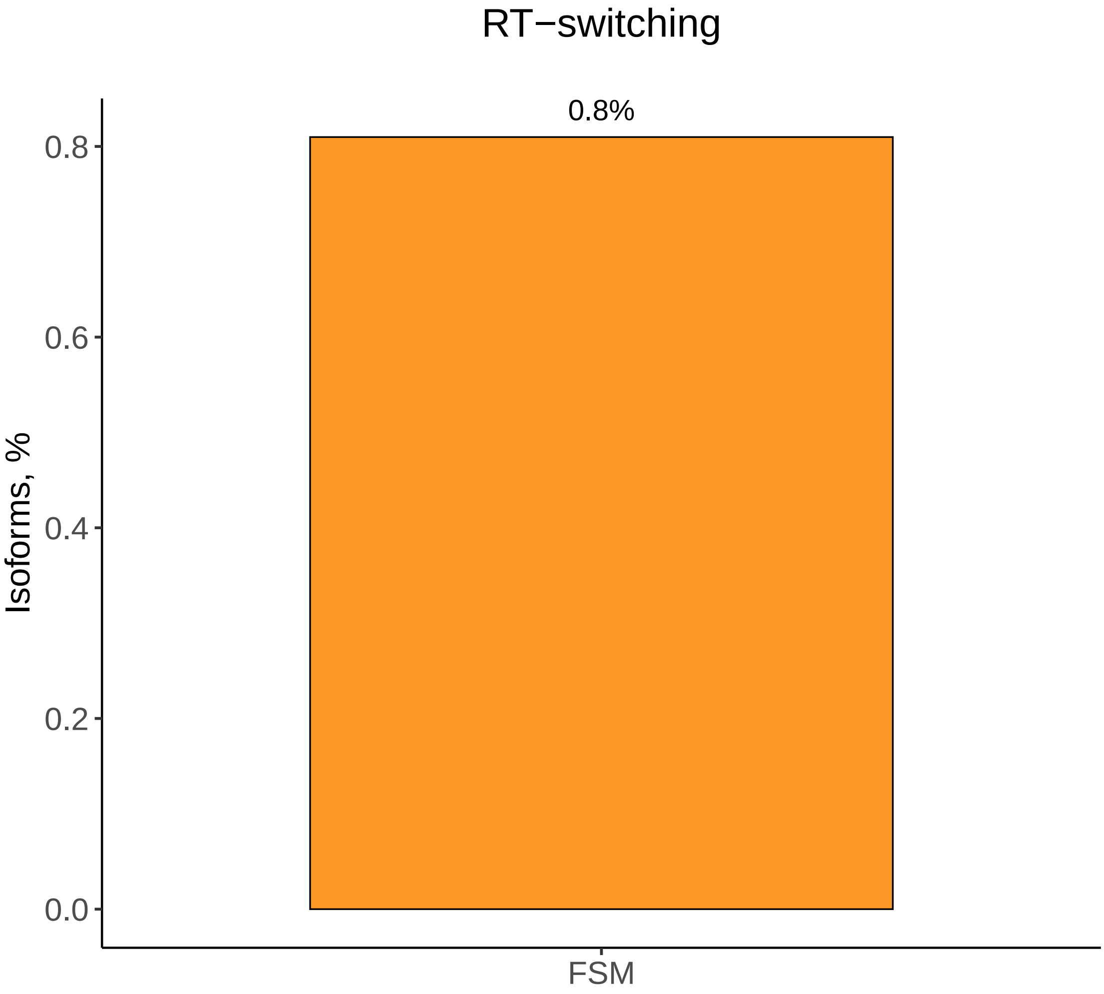

Mono− vs Multi−Exon Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
NA
Mono−Exon Isoforms
Multi−Exon Isoforms

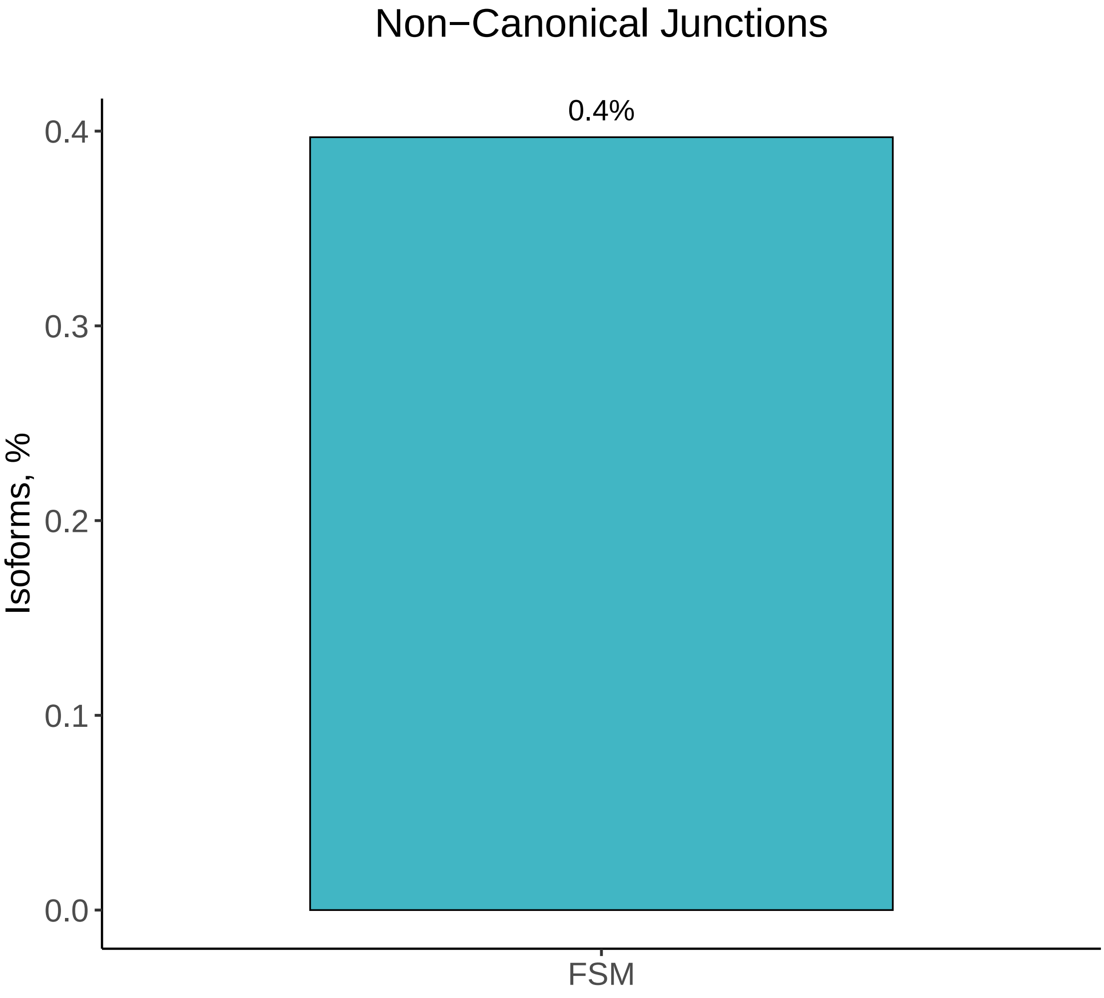


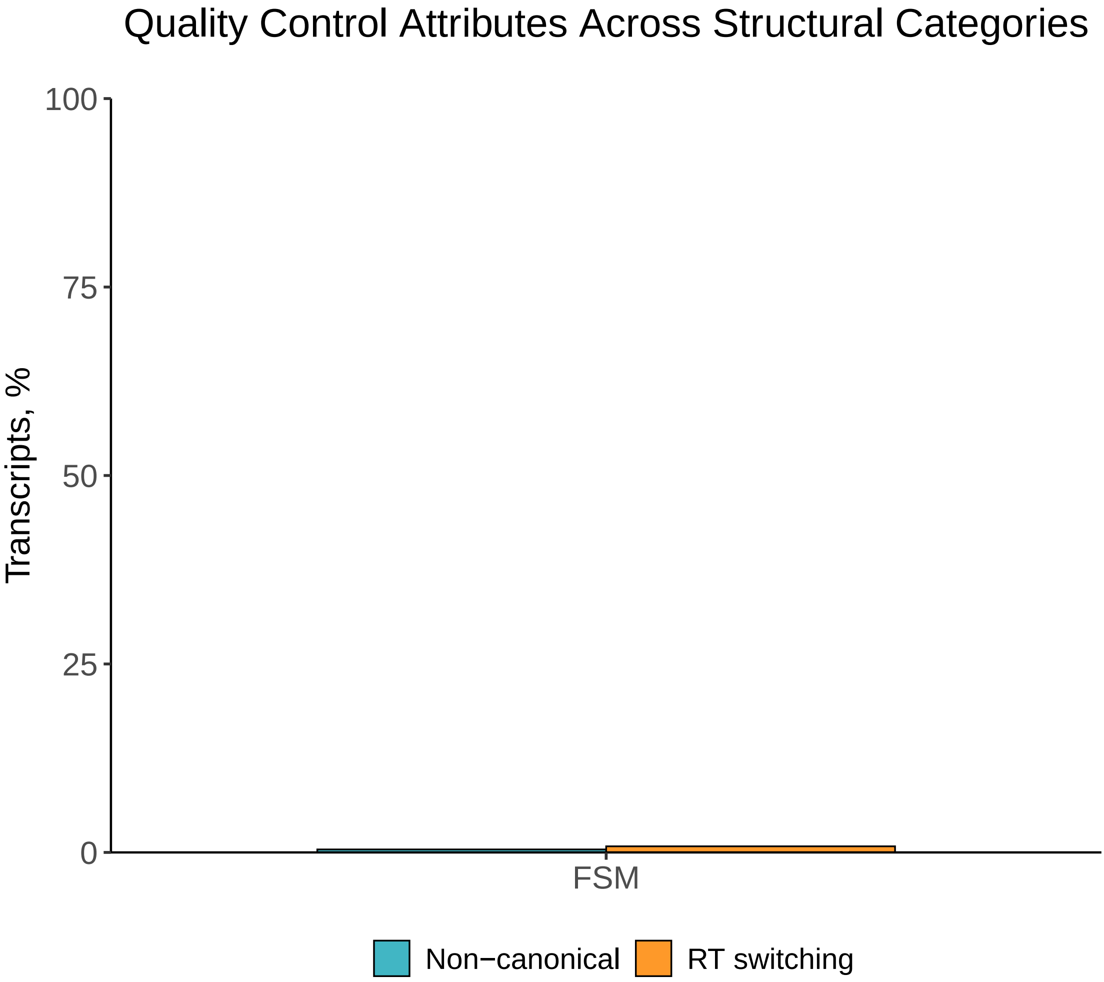
Coding vs Non−Coding Possible Intra−Priming

Percent of genomic 'A's in downstream 20 bp

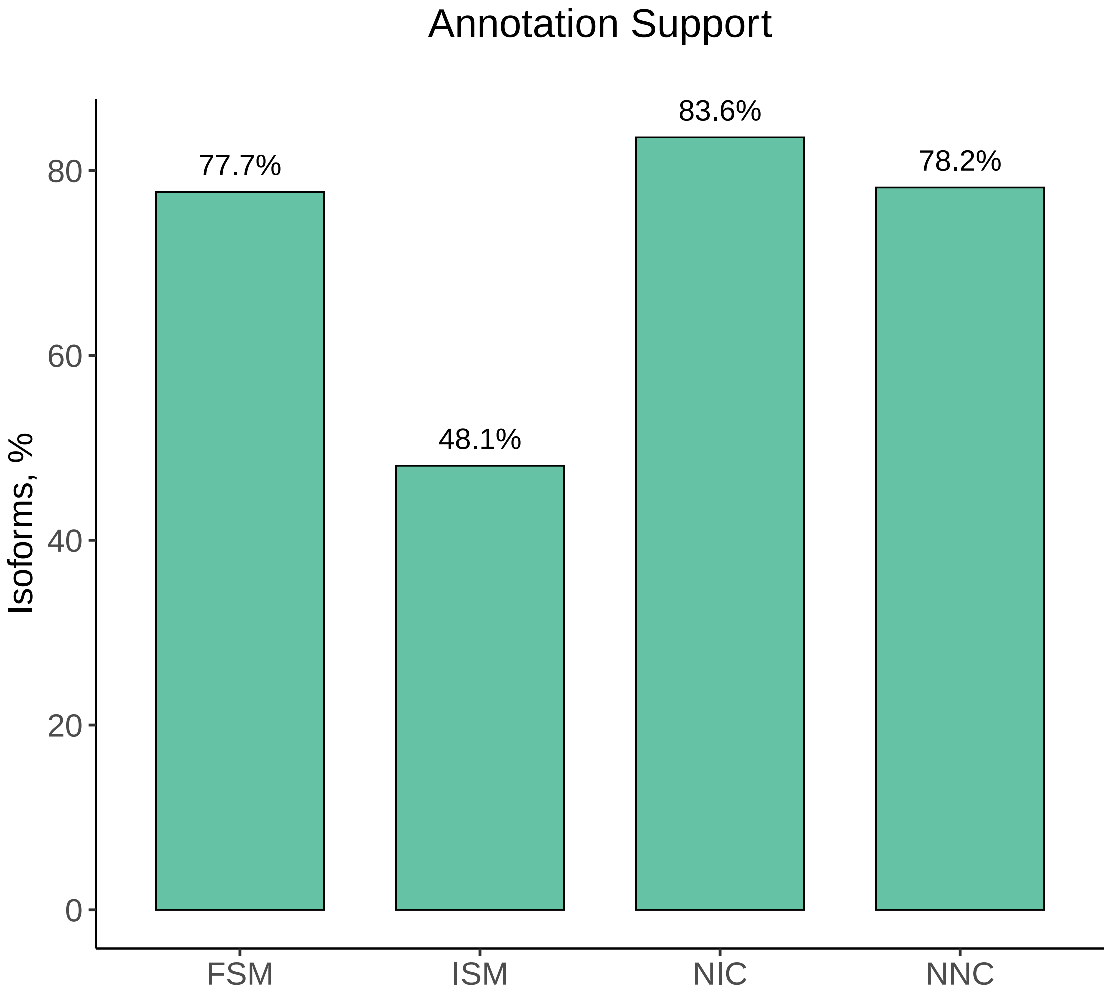
Features of Bad Quality

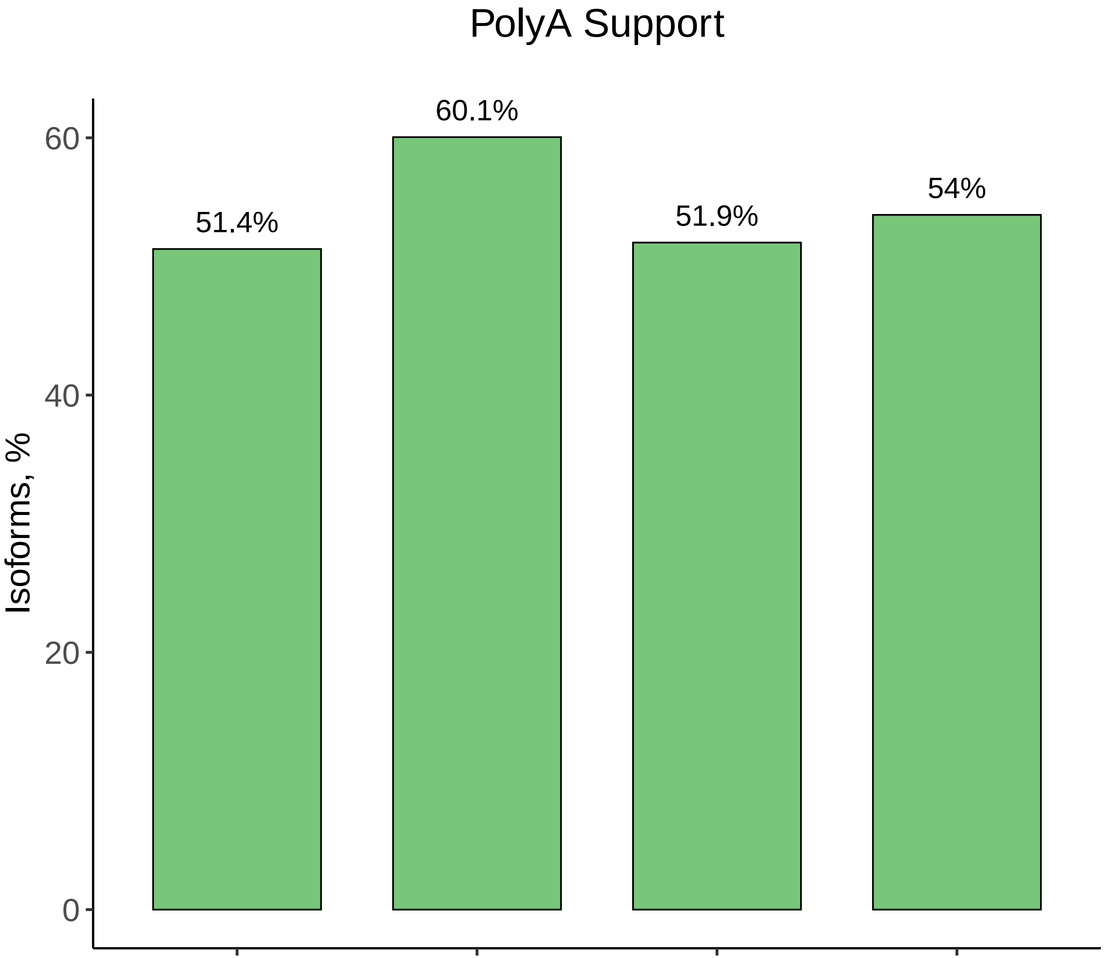

RT−switching
0.8%
Isoforms, %
0.0
0.2
0.4
0.6
0.8
FSM

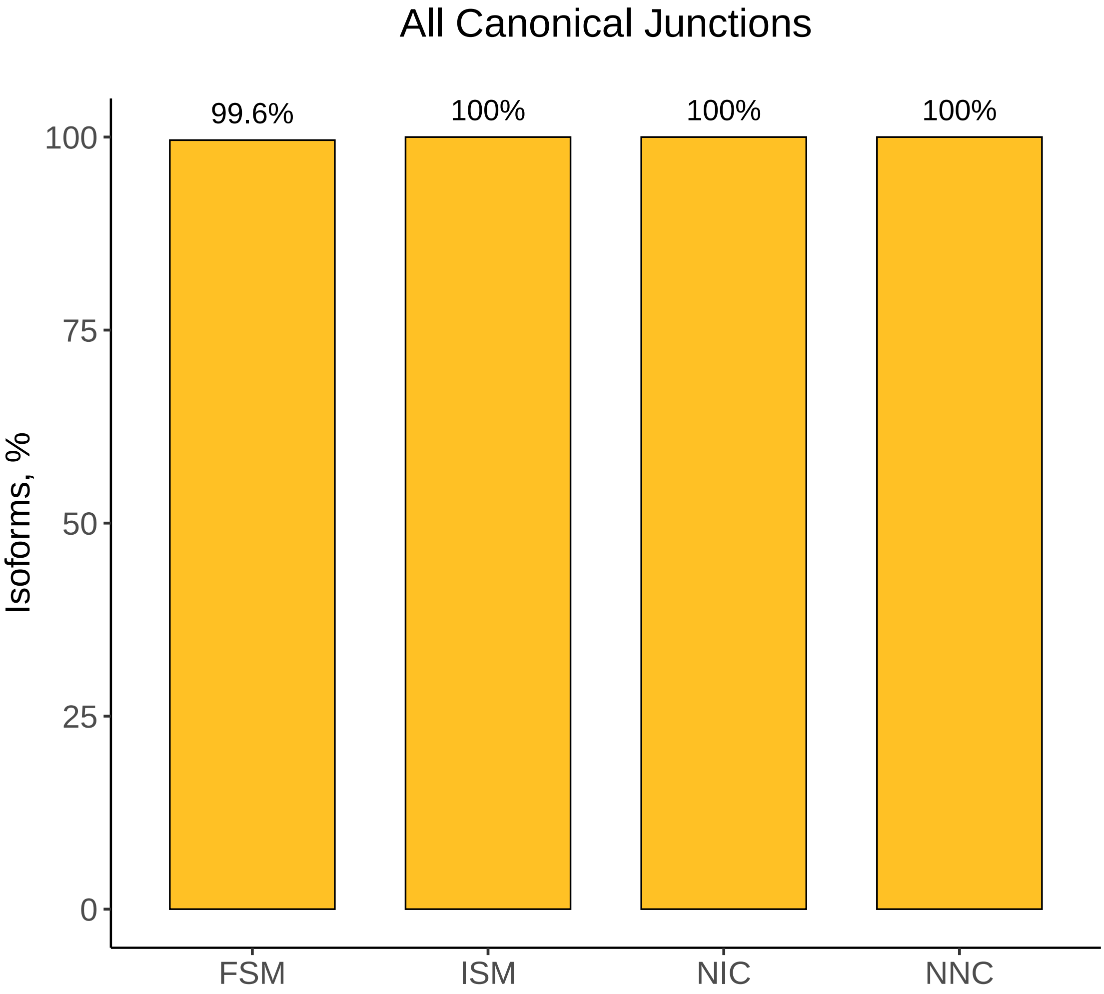

Non−Canonical Junctions
0.4%
Isoforms, %
0.4
0.3
0.2
0.1
0.0
FSM

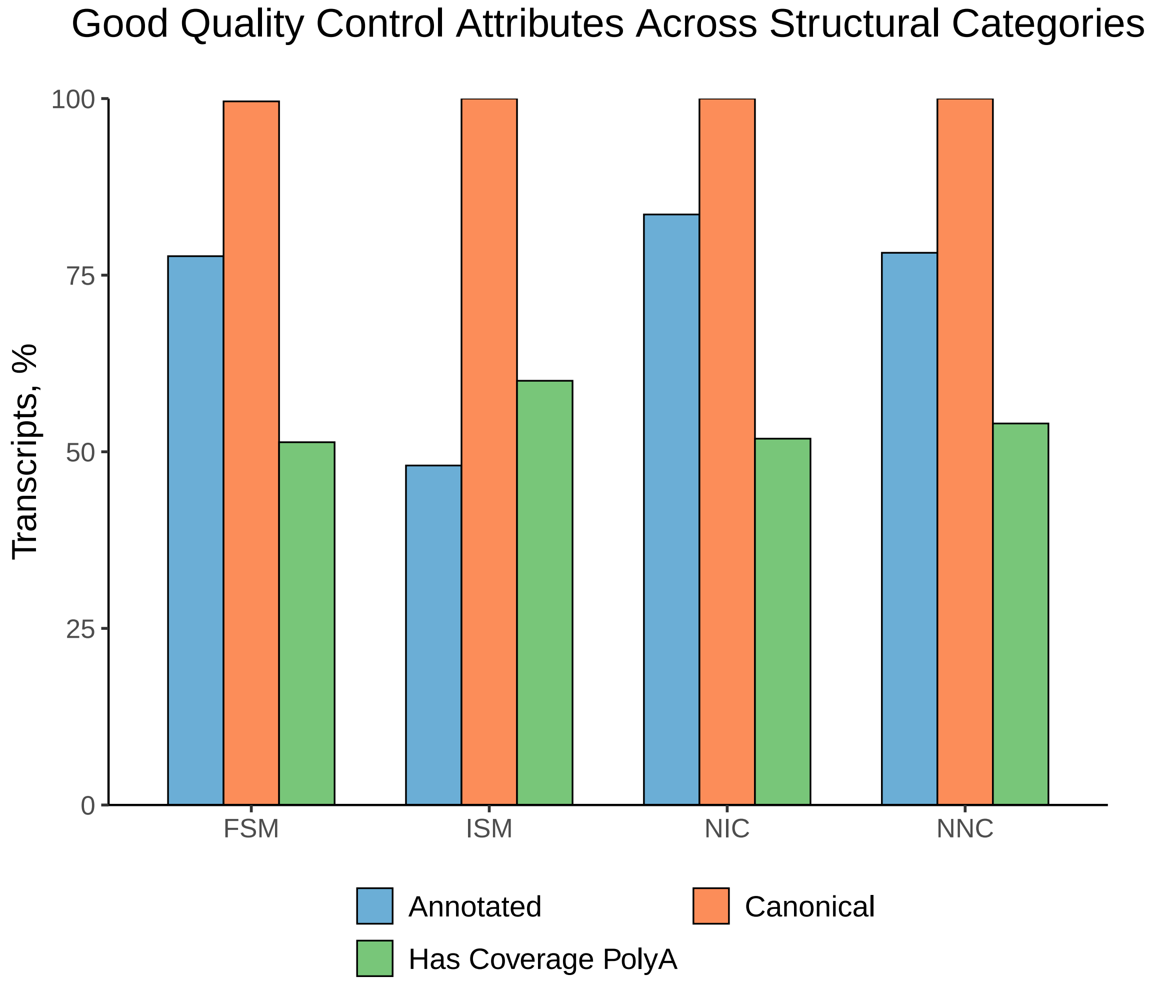

Quality Control Attributes Across Structural Categories
Transcripts, %
100
75
50
25
0
FSM
Non−canonical
RT switching

Features of Good Quality


Annotation Support
Isoforms, %
0
20
40
60
80
77.7%
48.1%
83.6%
78.2%
FSM
ISM
NIC
NNC


PolyA Support
Isoforms, %
60
40
20
0
51.4%
60.1%
51.9%
54%
FSM
ISM
NIC
NNC


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


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