SQANTI3 report

Unique Genes: 63670

Unique Isoforms: 336445

Transcript Classification

Category	Isoforms, count
FSM	193864
ISM	52552
NIC	19537
NNC	55277
Genic Genomic	2321
Antisense	1562
Fusion	822
Intergenic	10422
Genic Intron	88

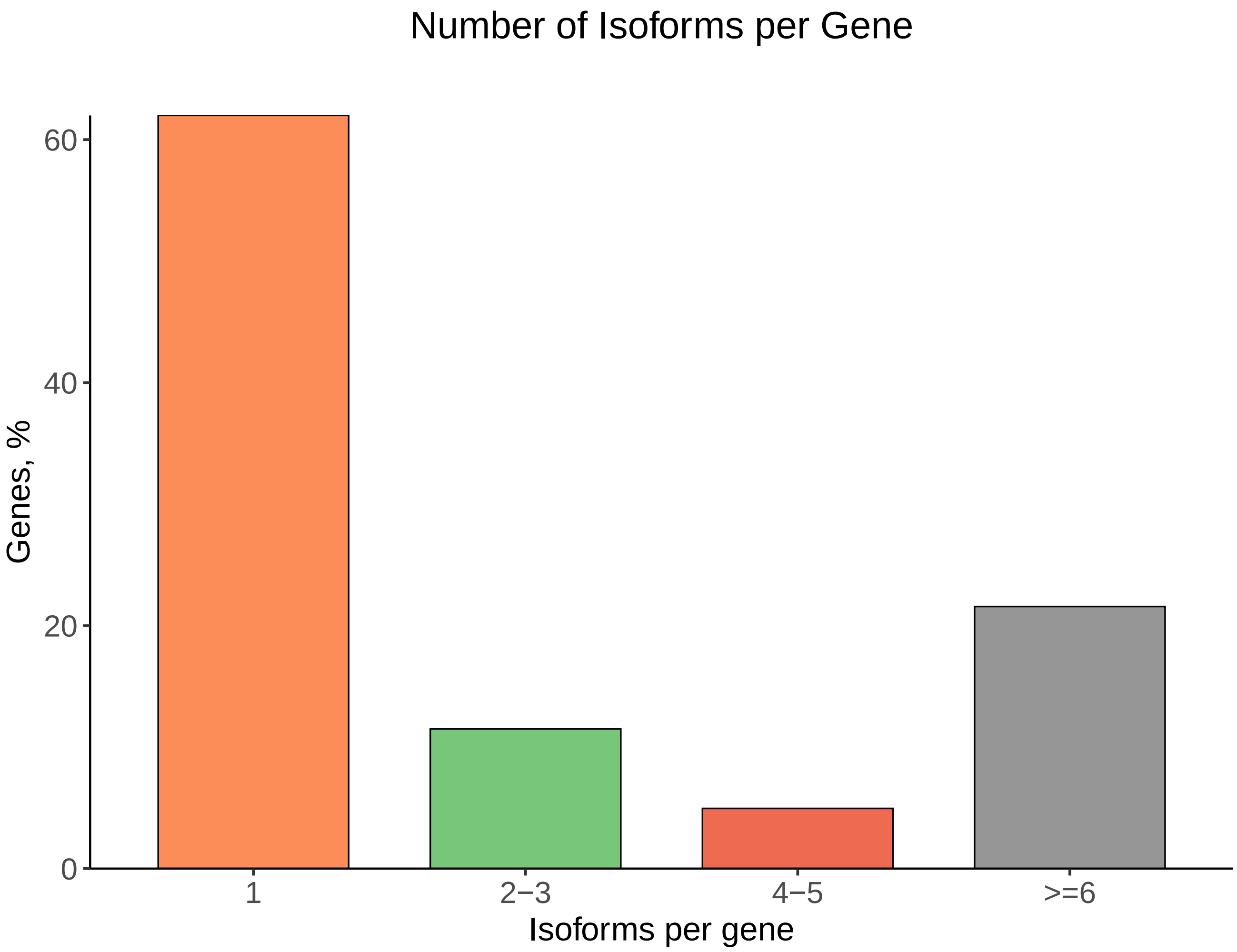
Gene Classification

Category	Genes, count
Annotated Genes	51843
Novel Genes	11827

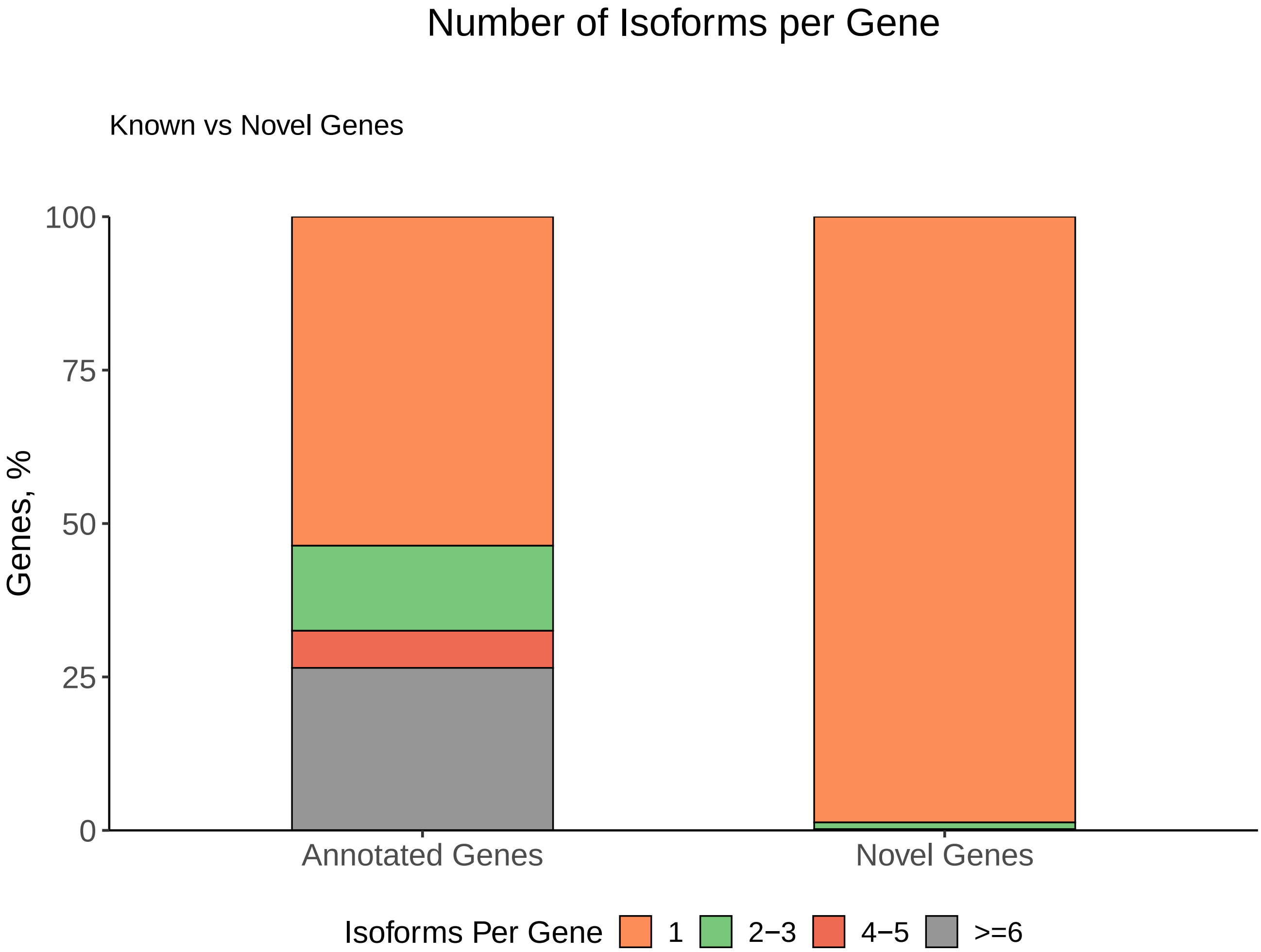
Splice Junction Classification

Category	**SJs, count**	**Percent**
Known canonical	354608	85.34
Known Non−canonical	4950	1.19
Novel canonical	53708	12.93
Novel Non−canonical	2239	0.54

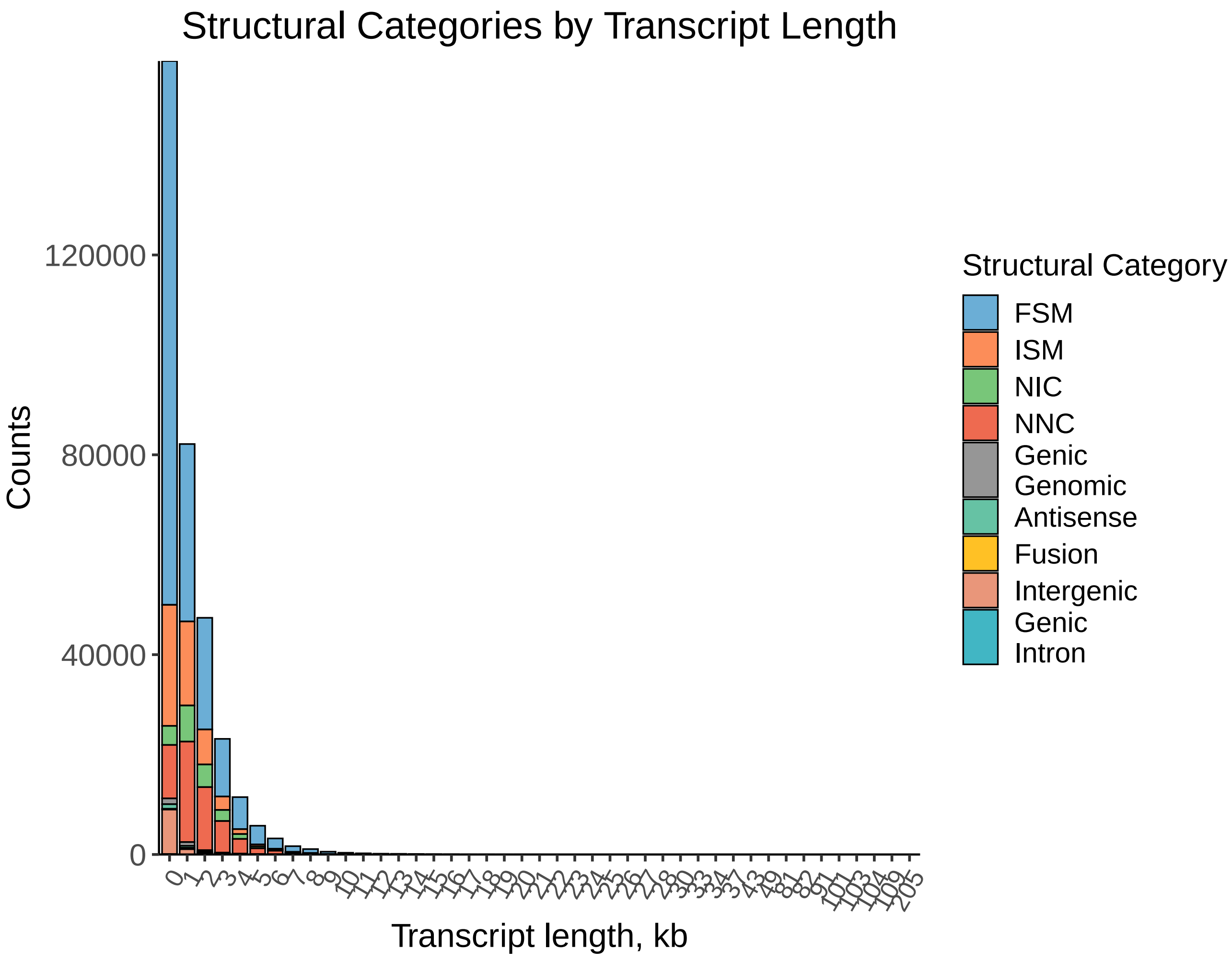
Gene Characterization

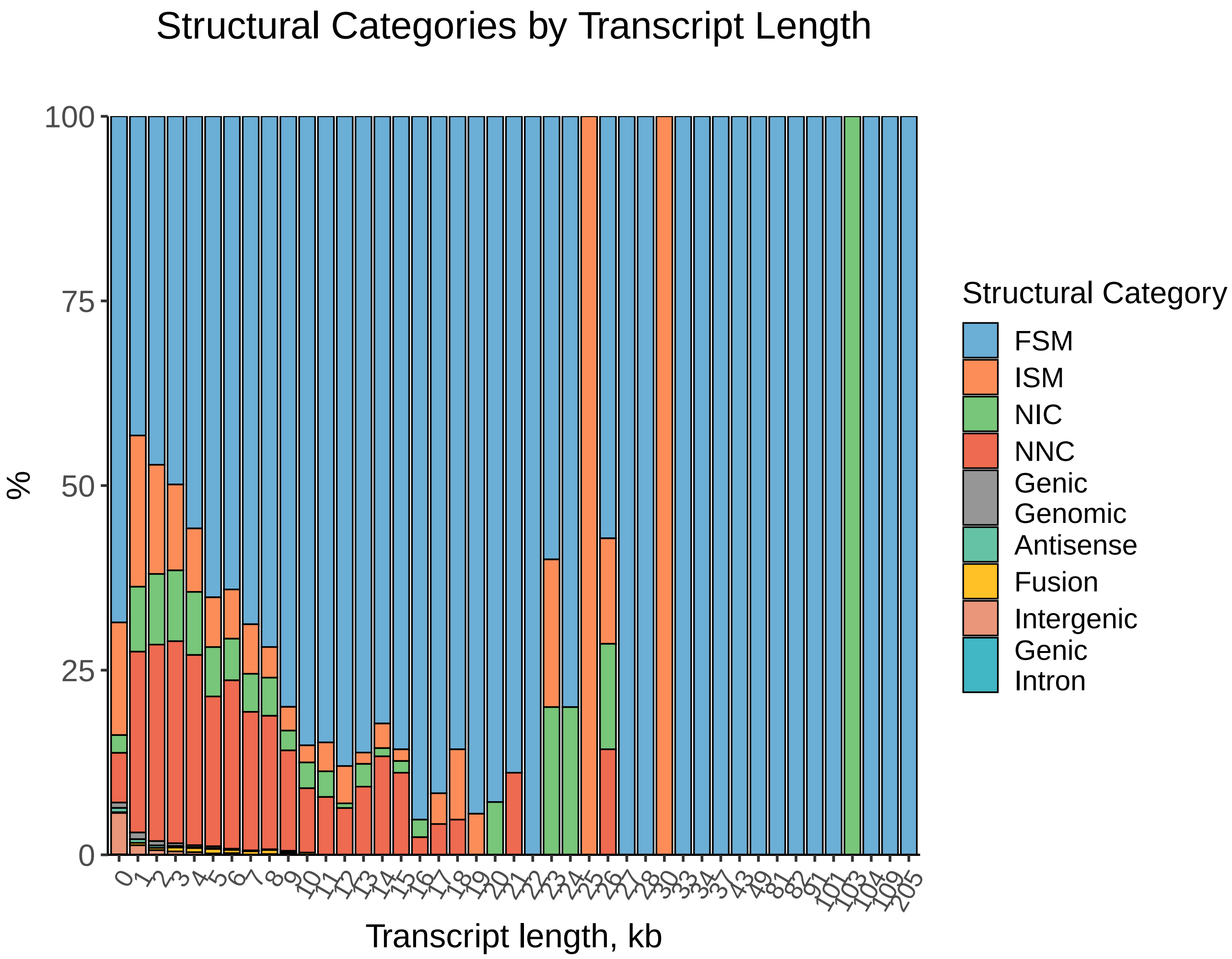


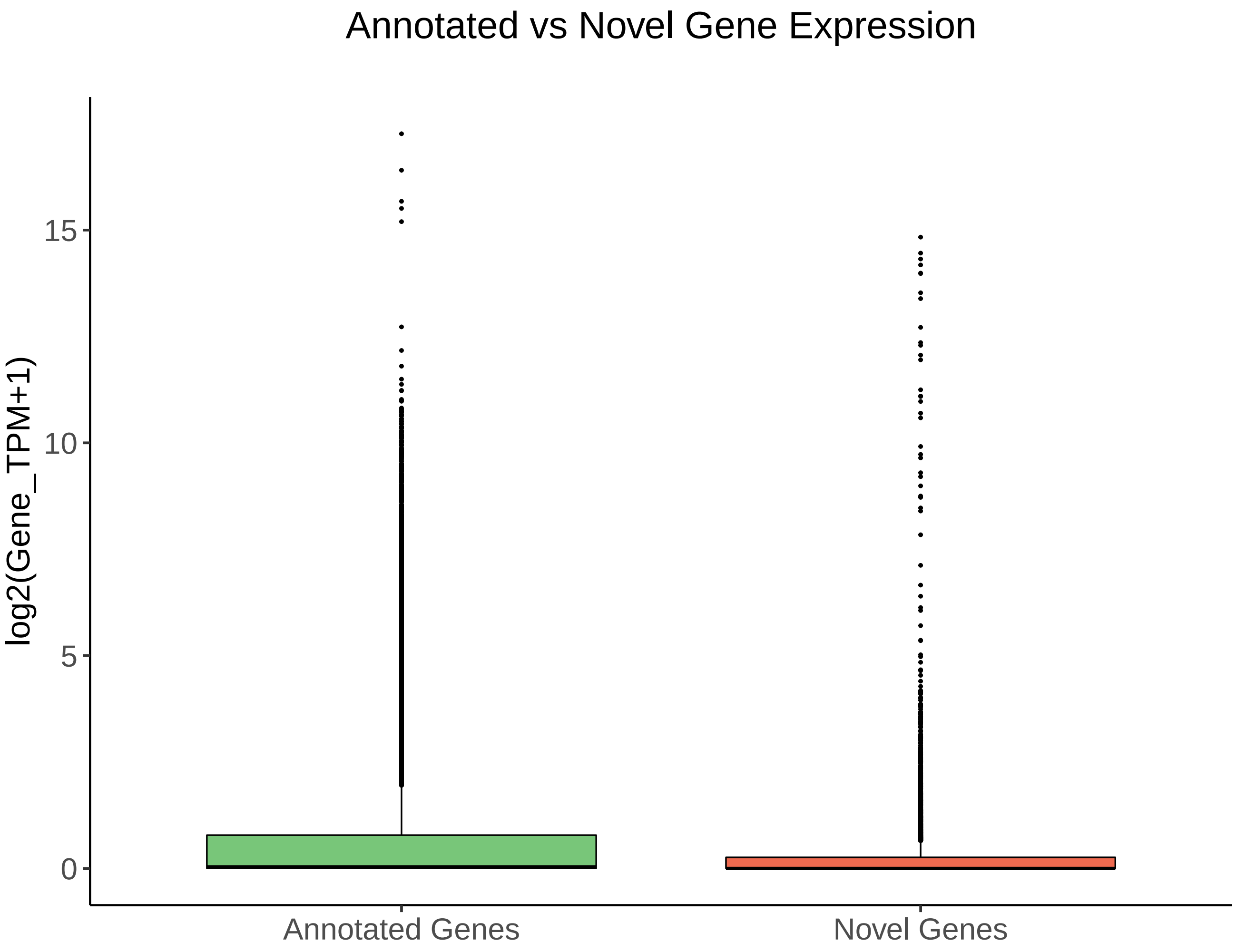
Number of Isoforms per Gene
Genes, %
60
40
20
0
1
2−3
4−5
>=6
Isoforms per gene

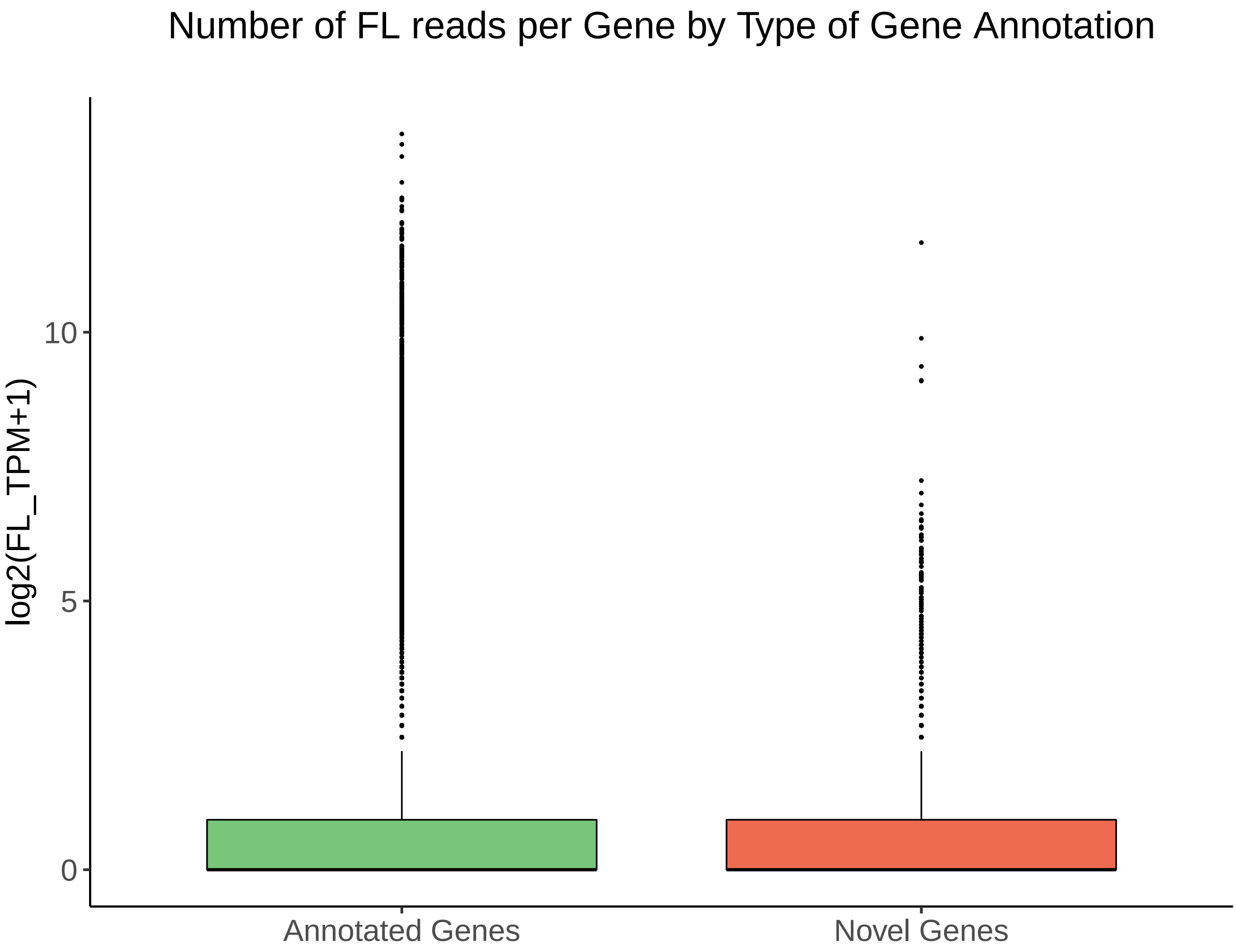

Number of Isoforms per Gene
Known vs Novel Genes
Genes, %
100
75
50
25
0
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3
4−5
>=6


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

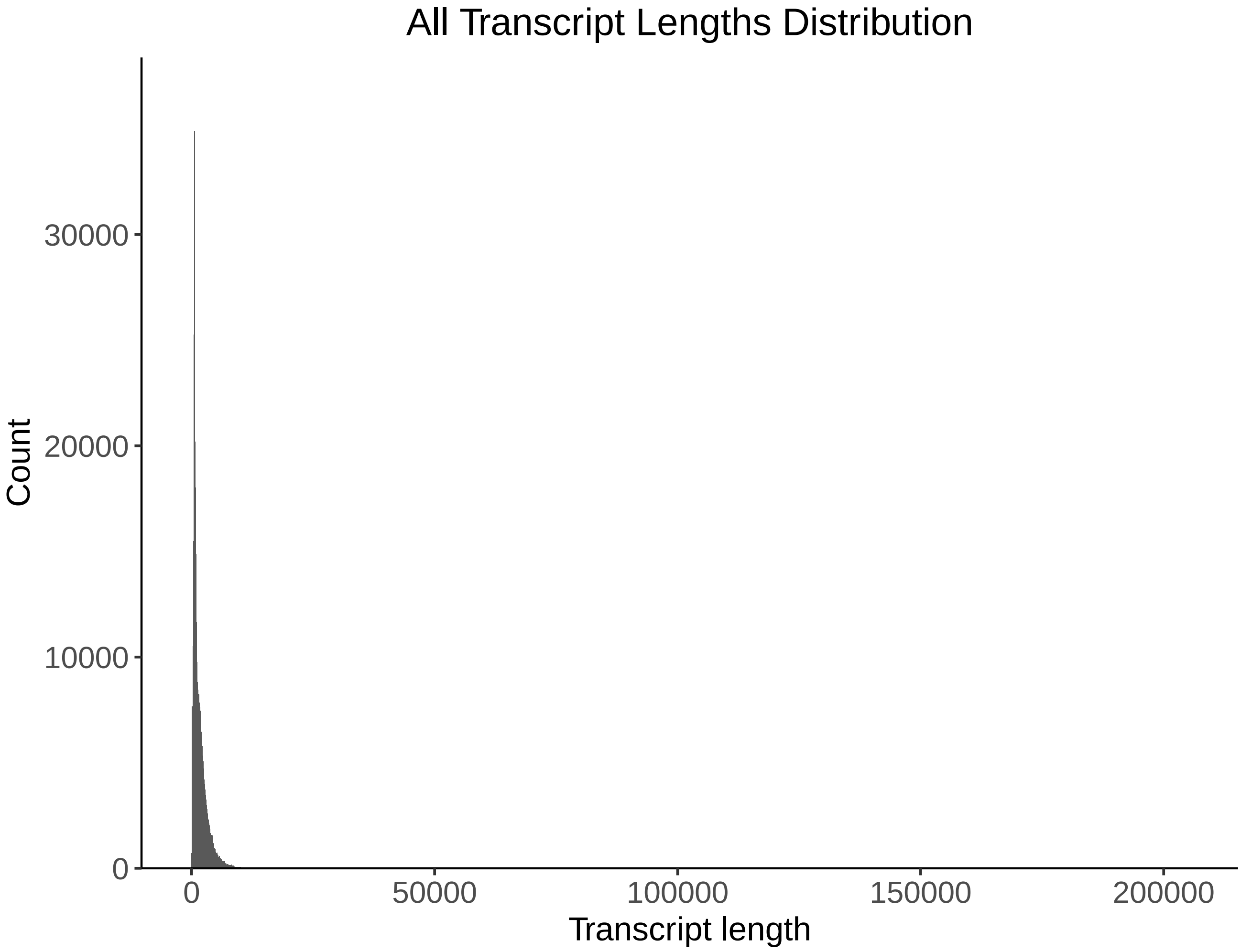

Structural Categories by Transcript Length
Counts
0
40000
80000
120000
Transcript length, kb
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
30
33
34
37
43
49
81
82
91
101
103
104
109
205
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic 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
22
23
24
25
26
27
28
30
33
34
37
43
49
81
82
91
101
103
104
109
205
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

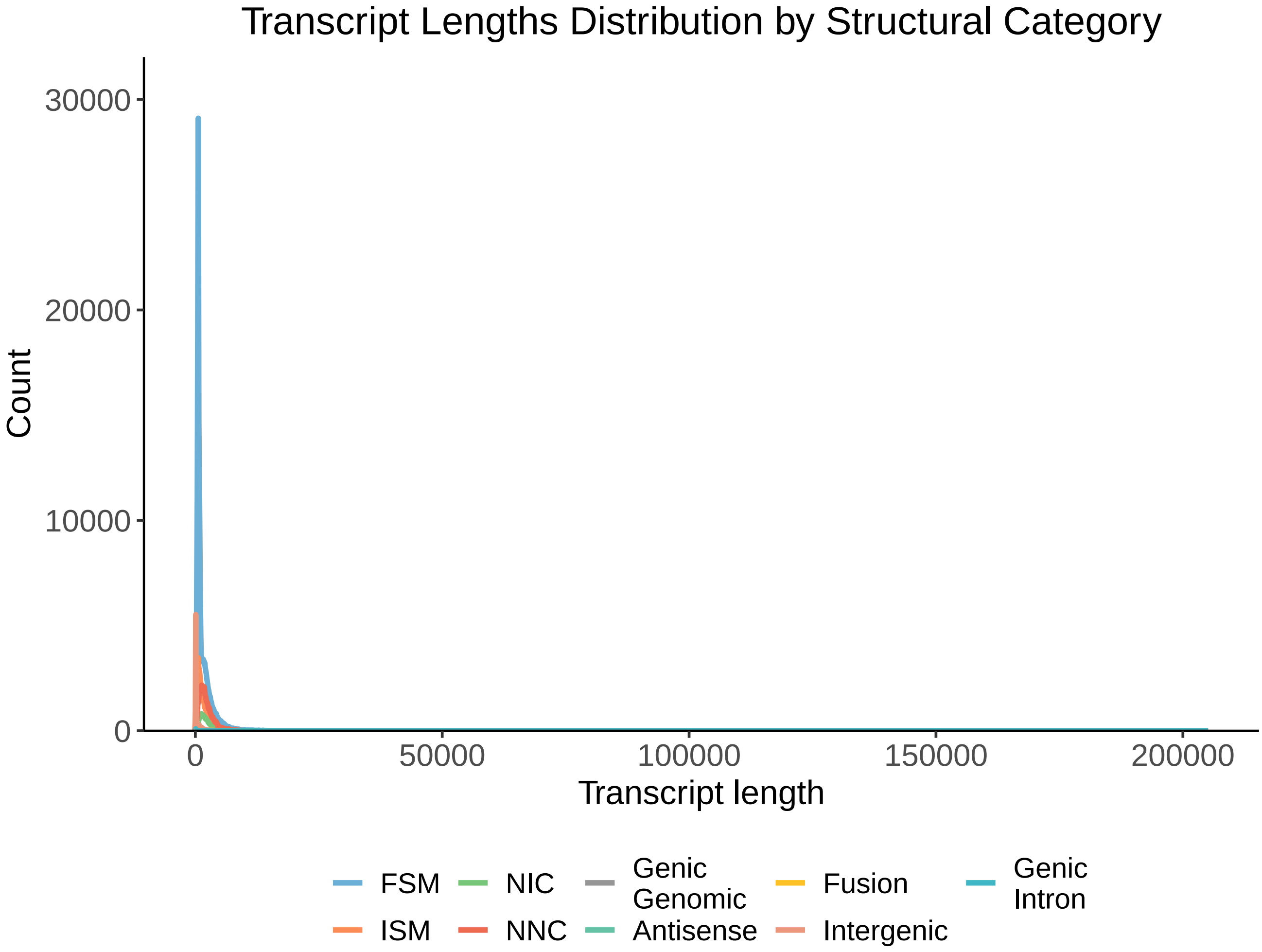

Annotated vs Novel Gene Expression
log2(Gene_TPM+1)
15
10
5
0
Annotated Genes
Novel Genes

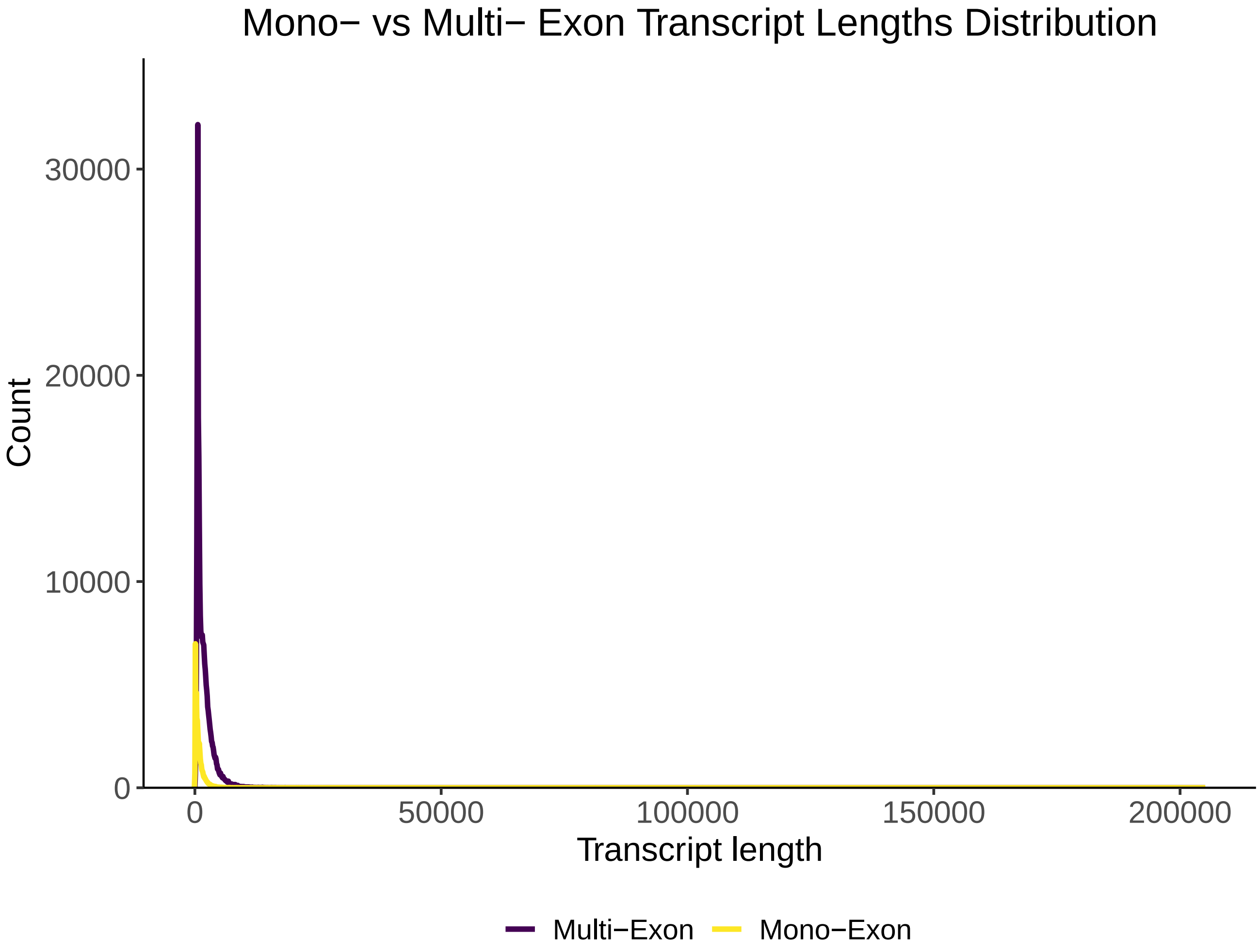


Number of FL reads per Gene by Type of Gene Annotation

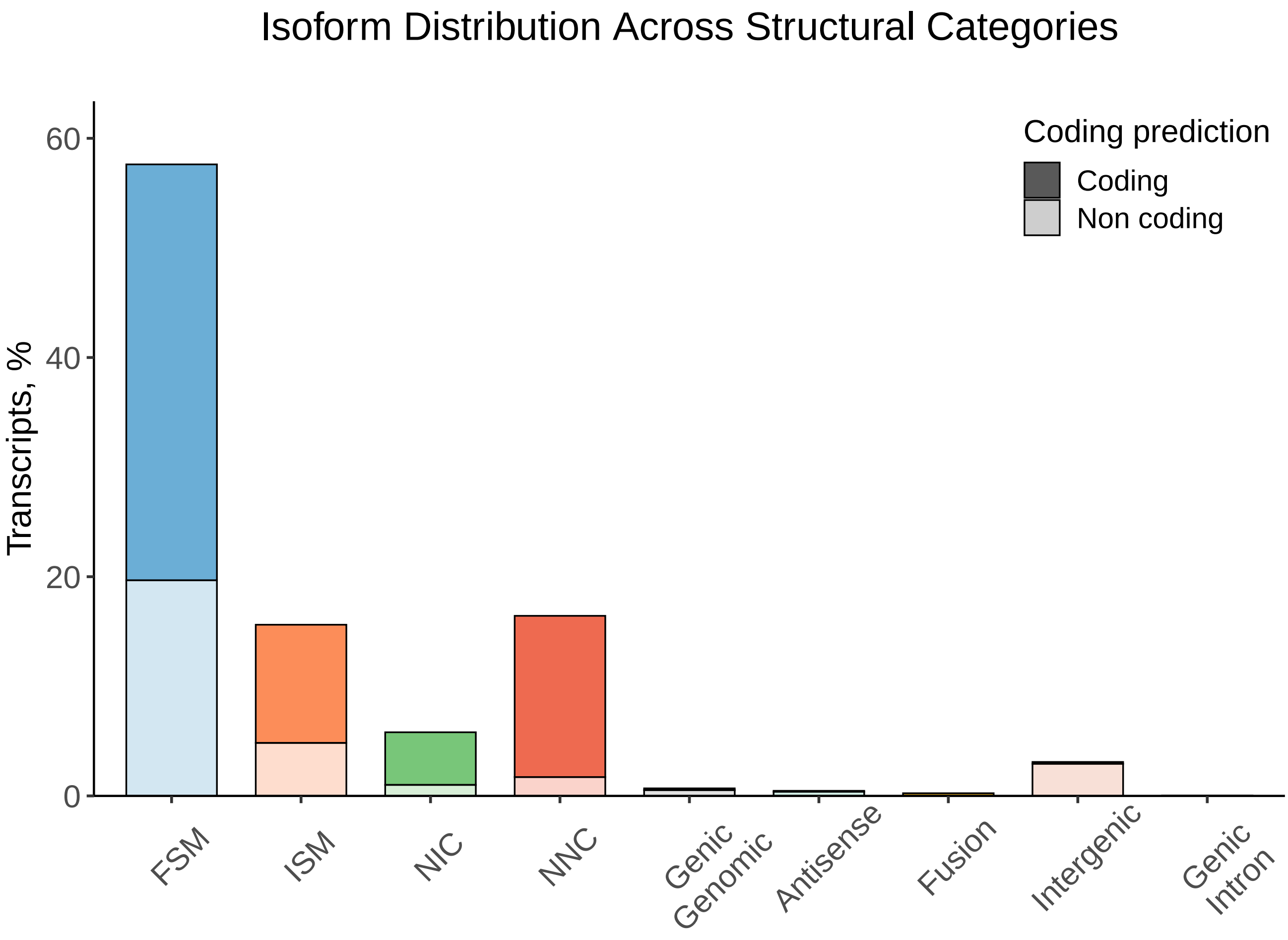


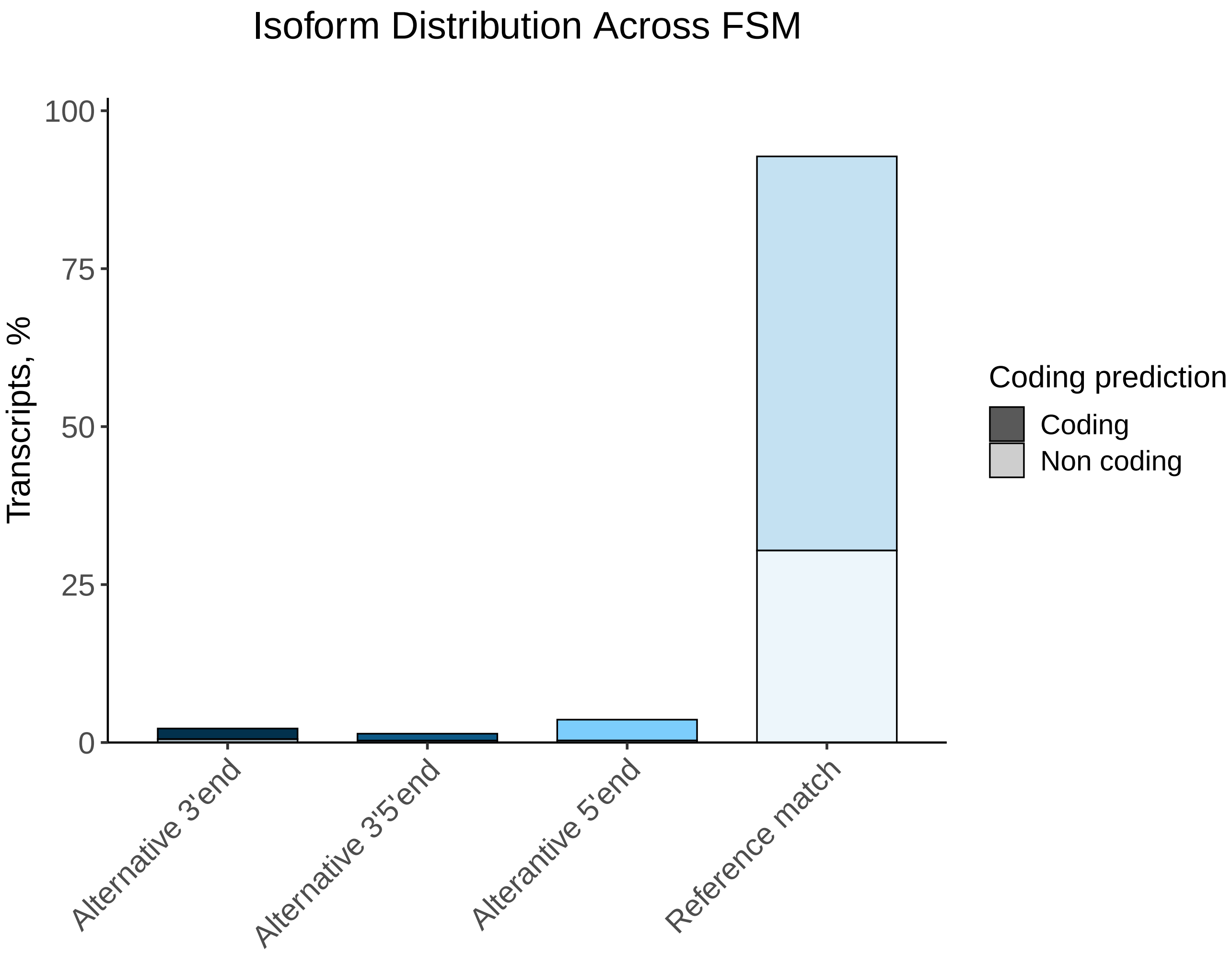
All Transcript Lengths Distribution

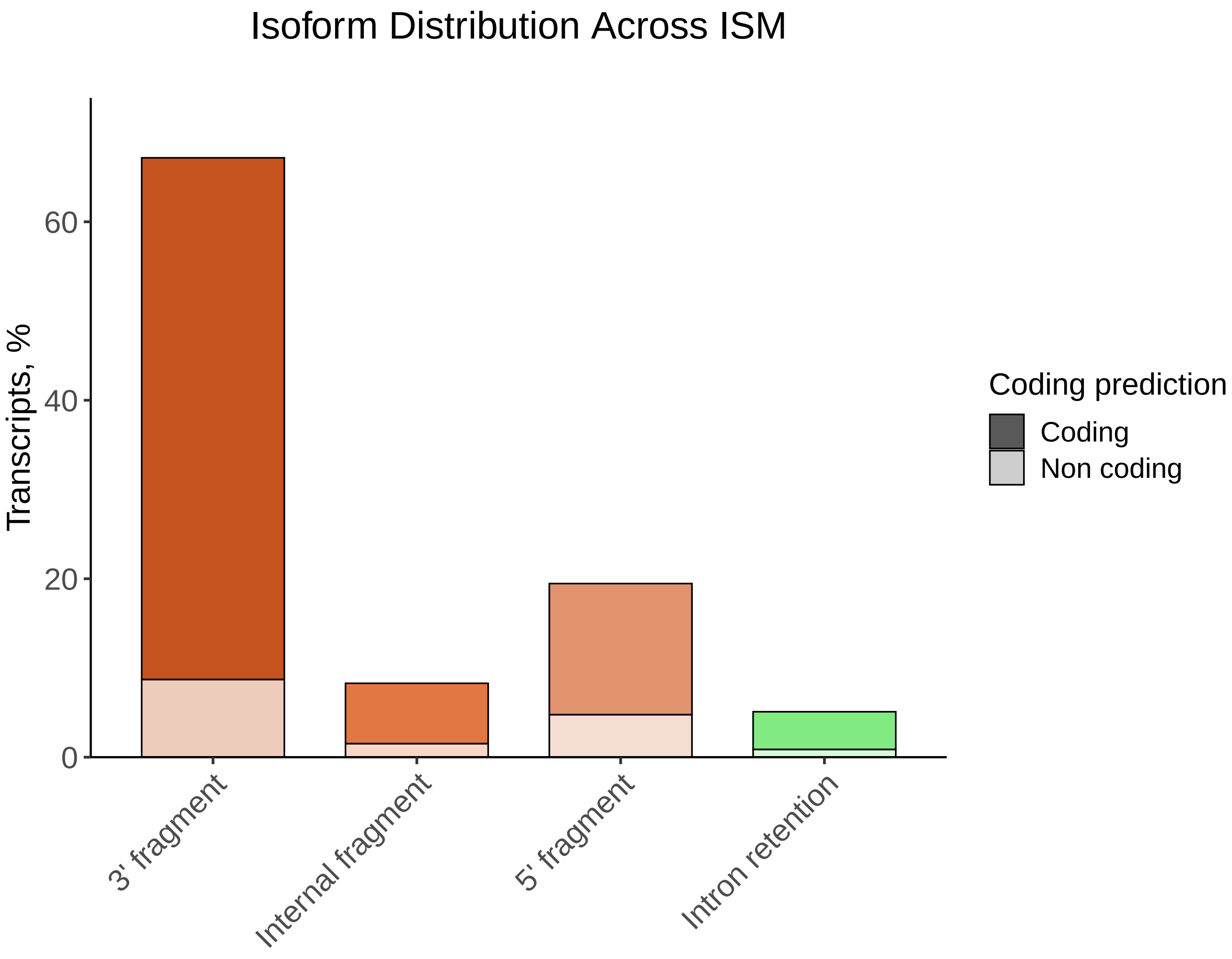

Transcript Lengths Distribution by Structural Category
Count
30000
20000
10000
0
0
50000
100000
150000
200000
Transcript length
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

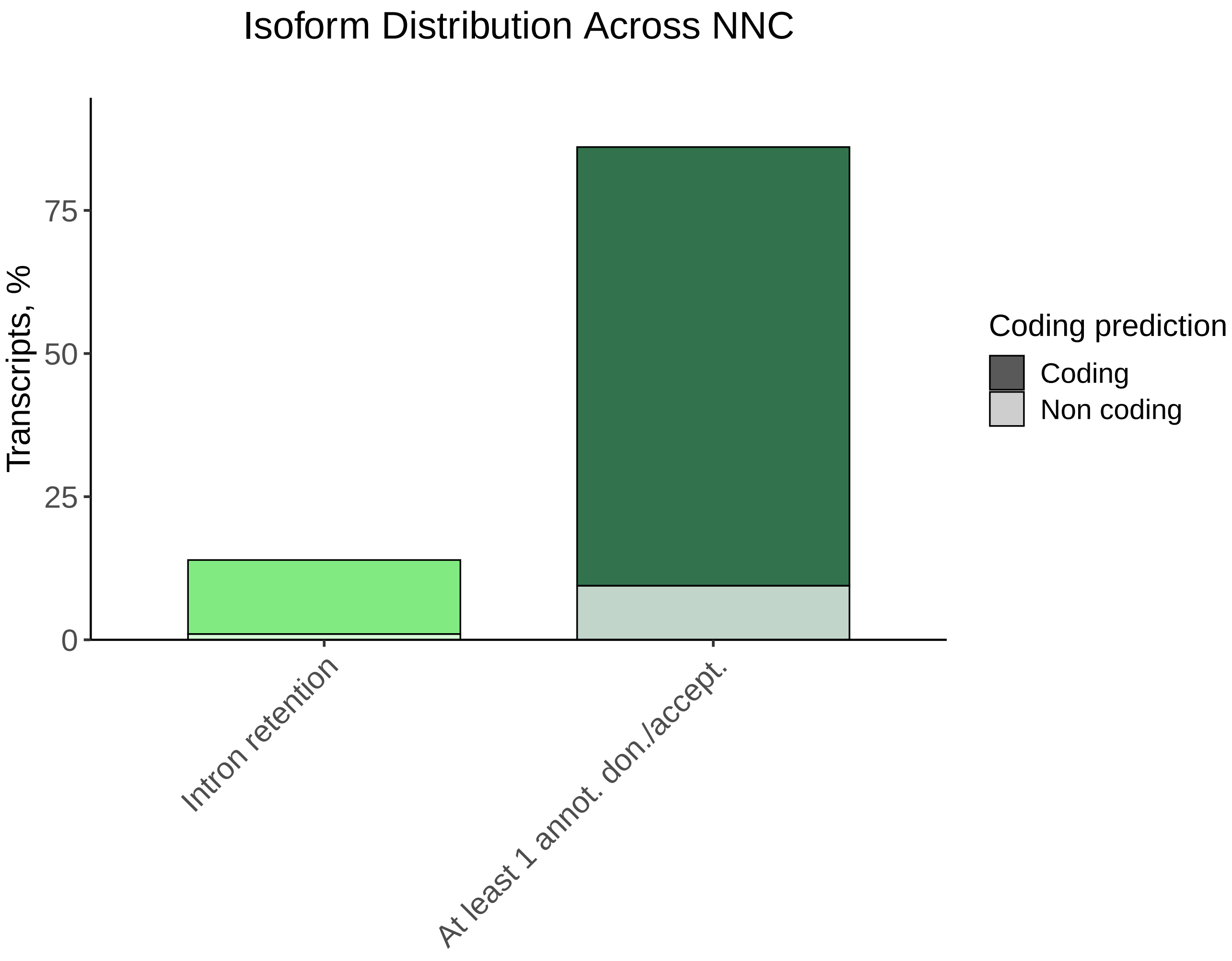

Mono− vs Multi− Exon Transcript Lengths Distribution
Count
30000
20000
10000
0
0
50000
100000
150000
200000
Transcript length
Multi−Exon
Mono−Exon

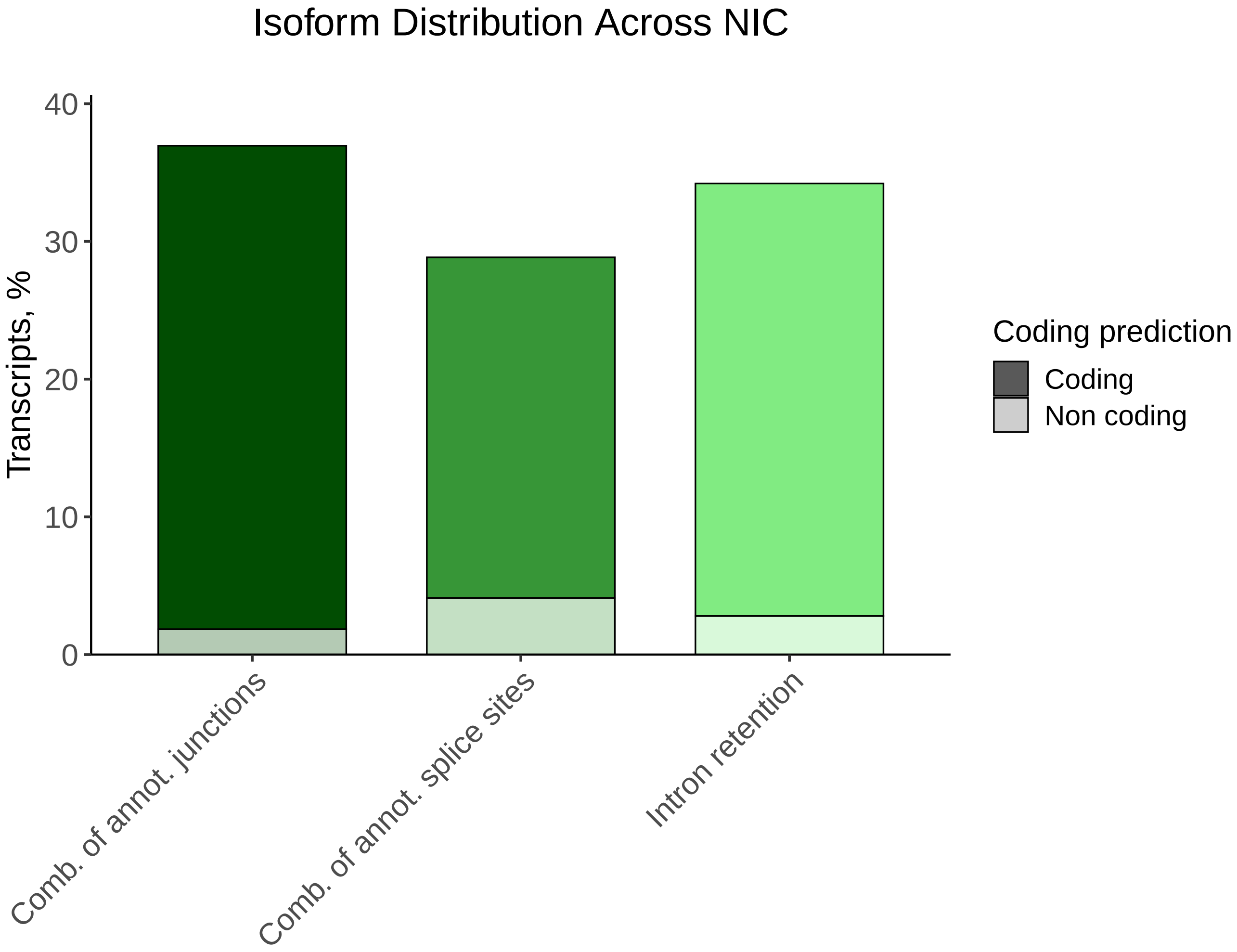
Structural Isoform Characterization

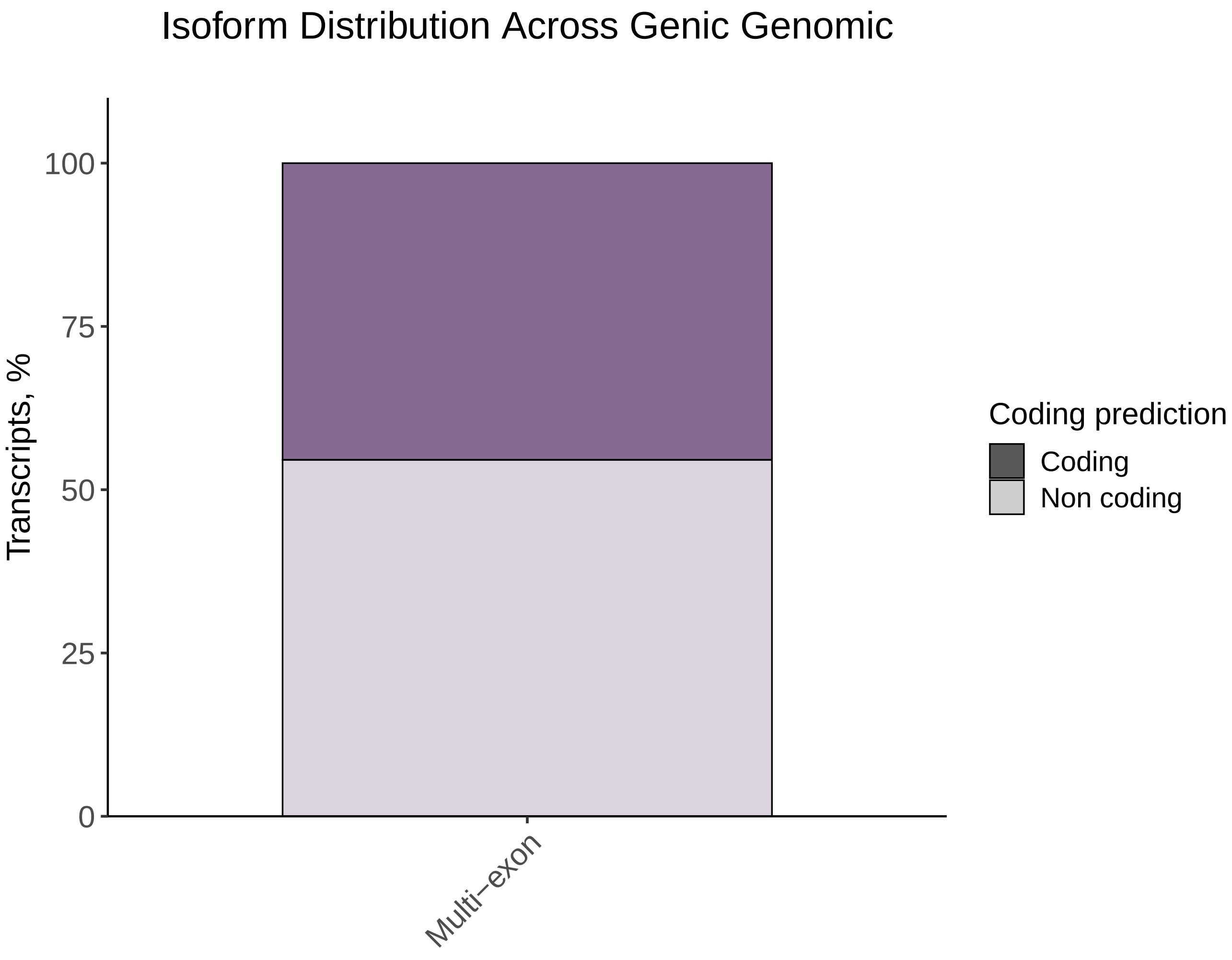
Isoform Distribution Across Structural Categories
Transcripts, %
0
20
40
60
Coding prediction
Coding
Non coding
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

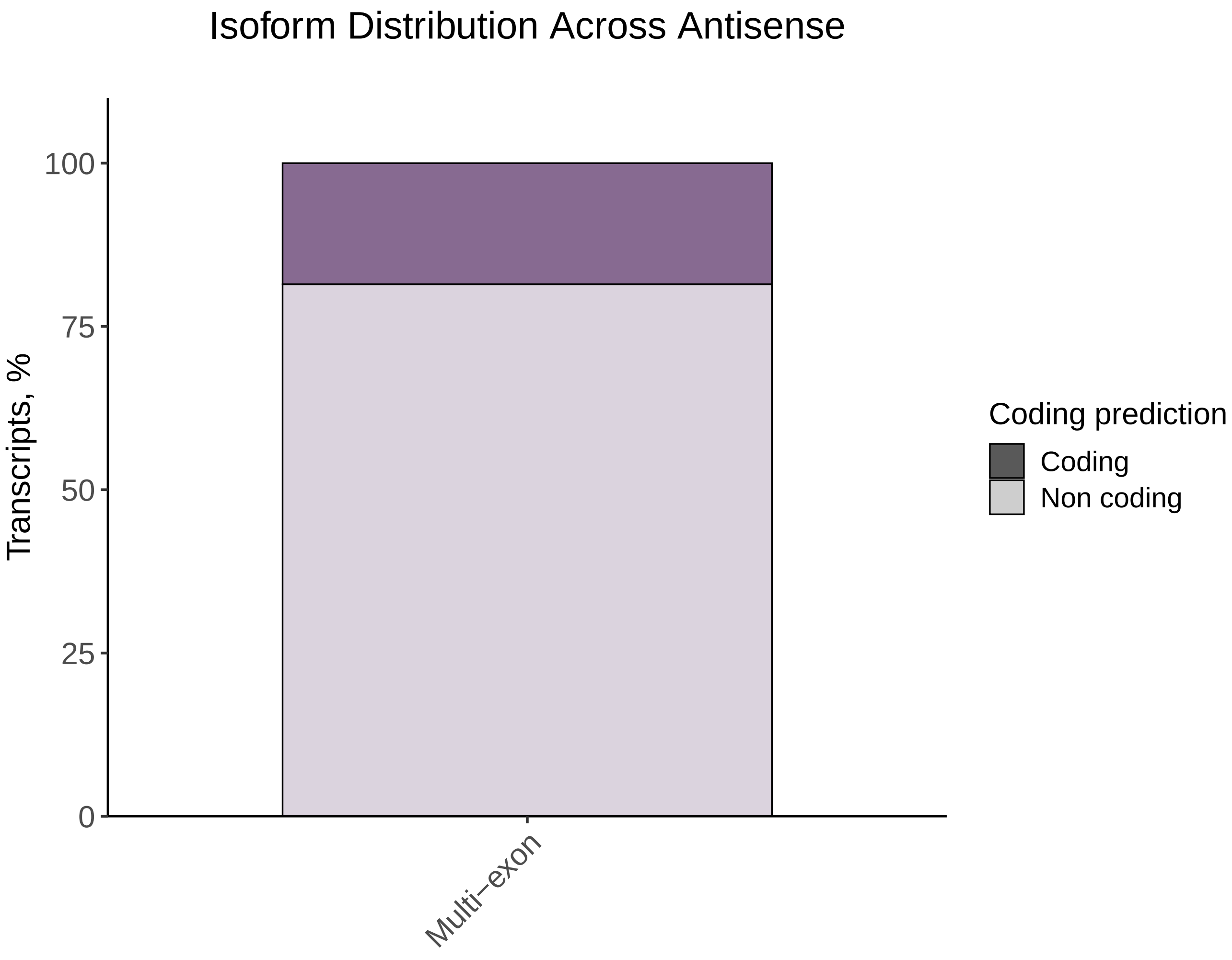

Isoform Distribution Across FSM
Transcripts, %
100
75
50
25
0
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Coding prediction
Coding
Non coding

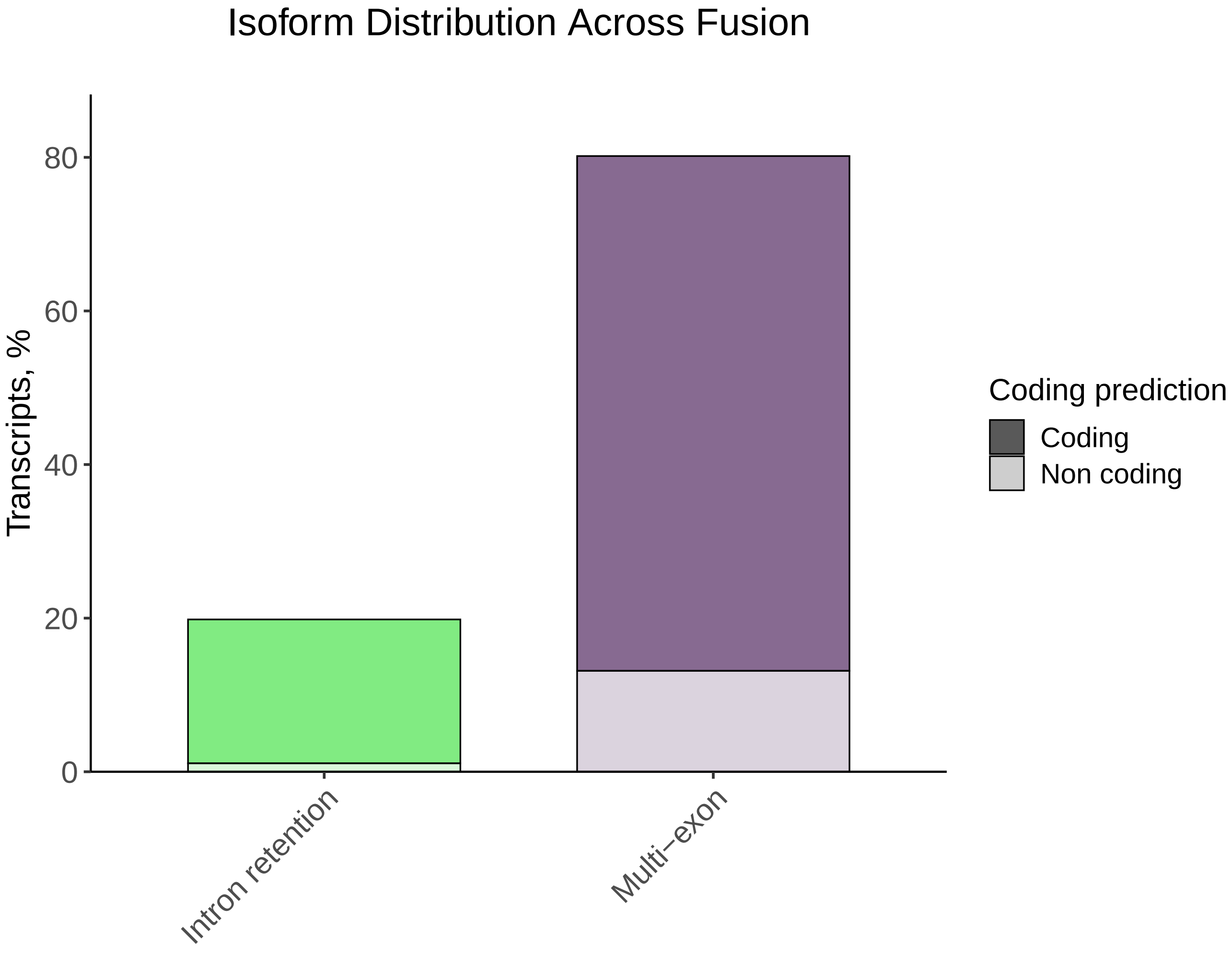

Isoform Distribution Across ISM
Transcripts, %
0
20
40
60
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding
Non coding

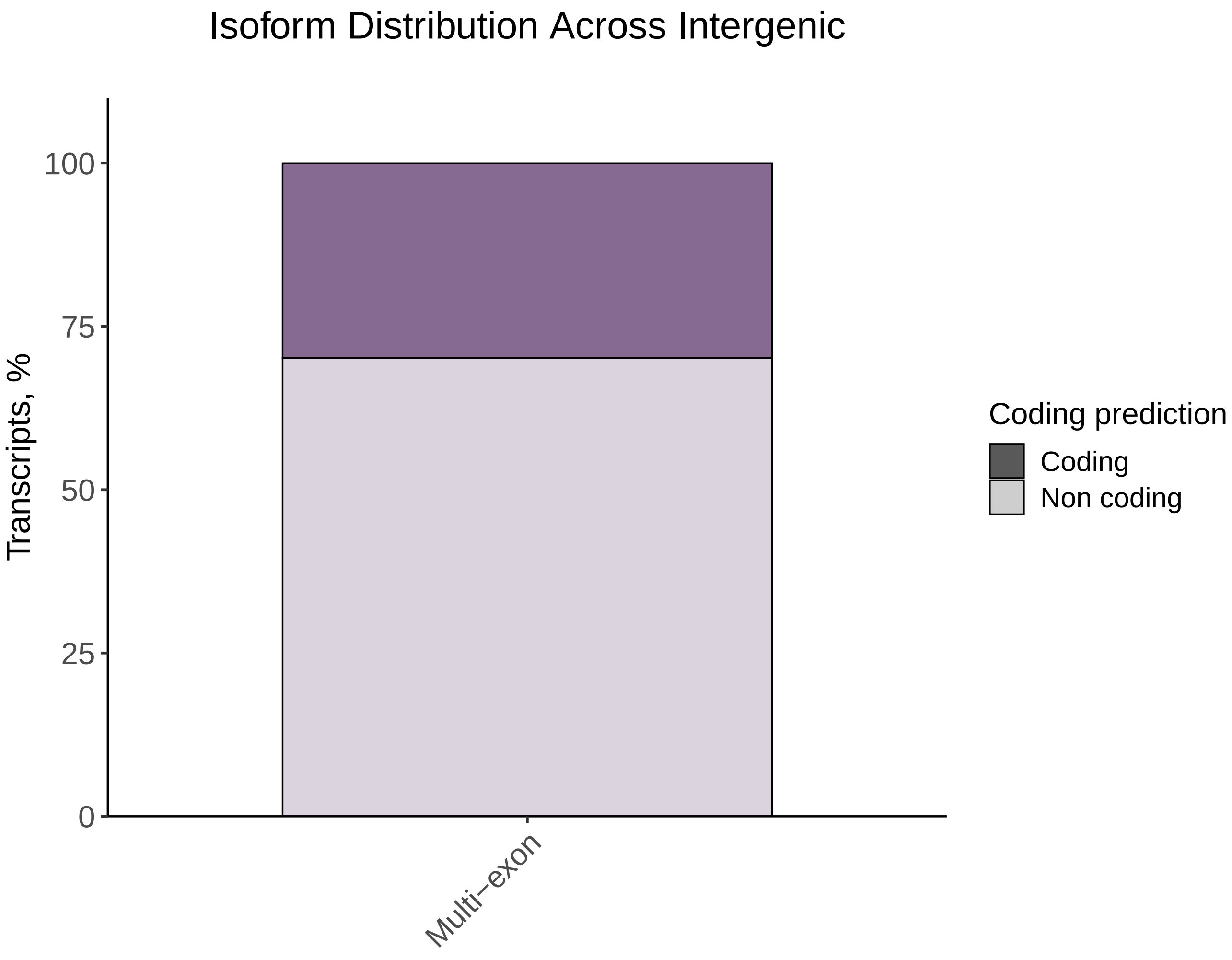

Isoform Distribution Across NNC
Transcripts, %
0
25
50
75
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding
Non coding

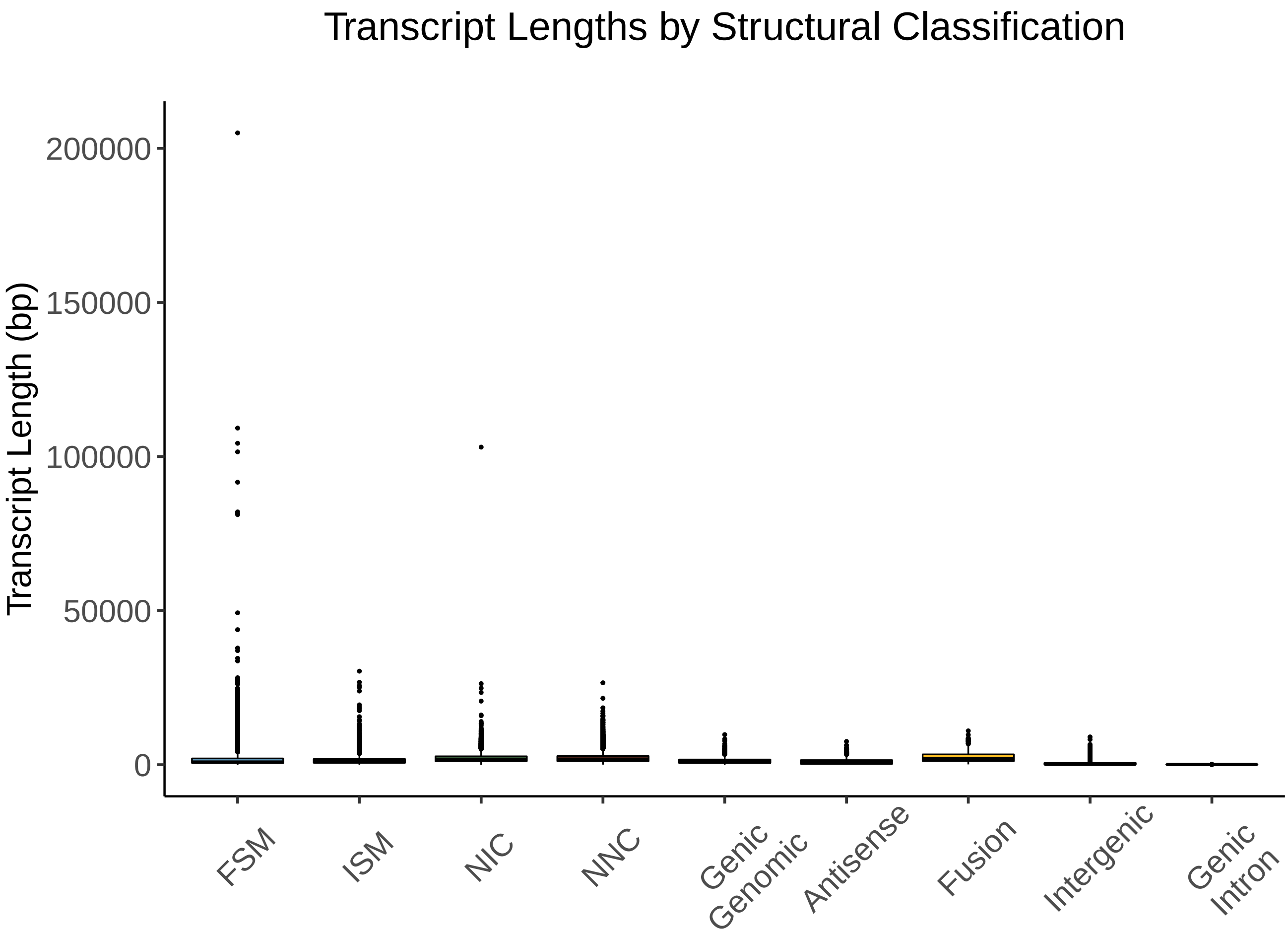

Isoform Distribution Across NIC
Transcripts, %
40
30
20
10
0
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Coding prediction
Coding
Non coding

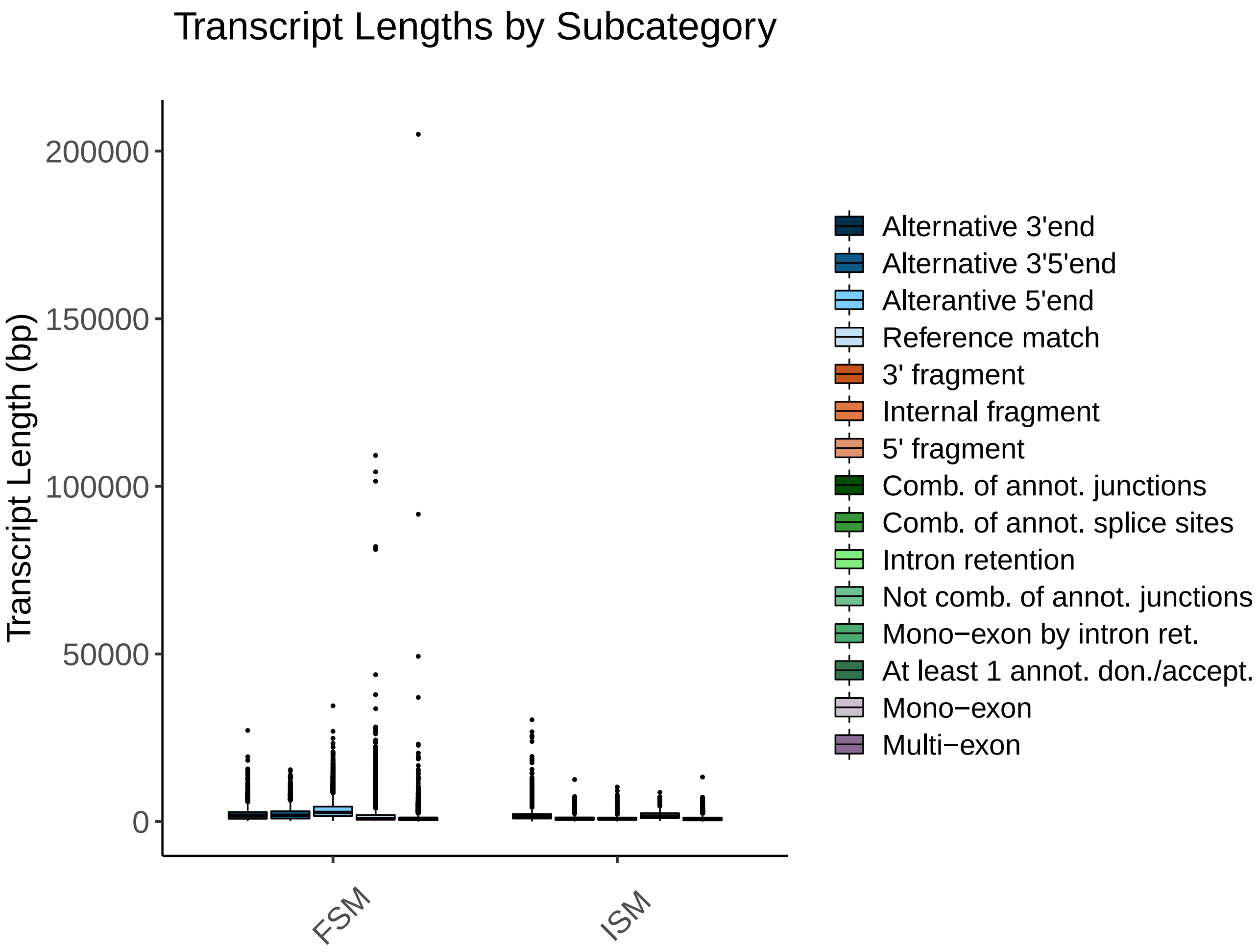

Isoform Distribution Across Genic Genomic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding
Non coding

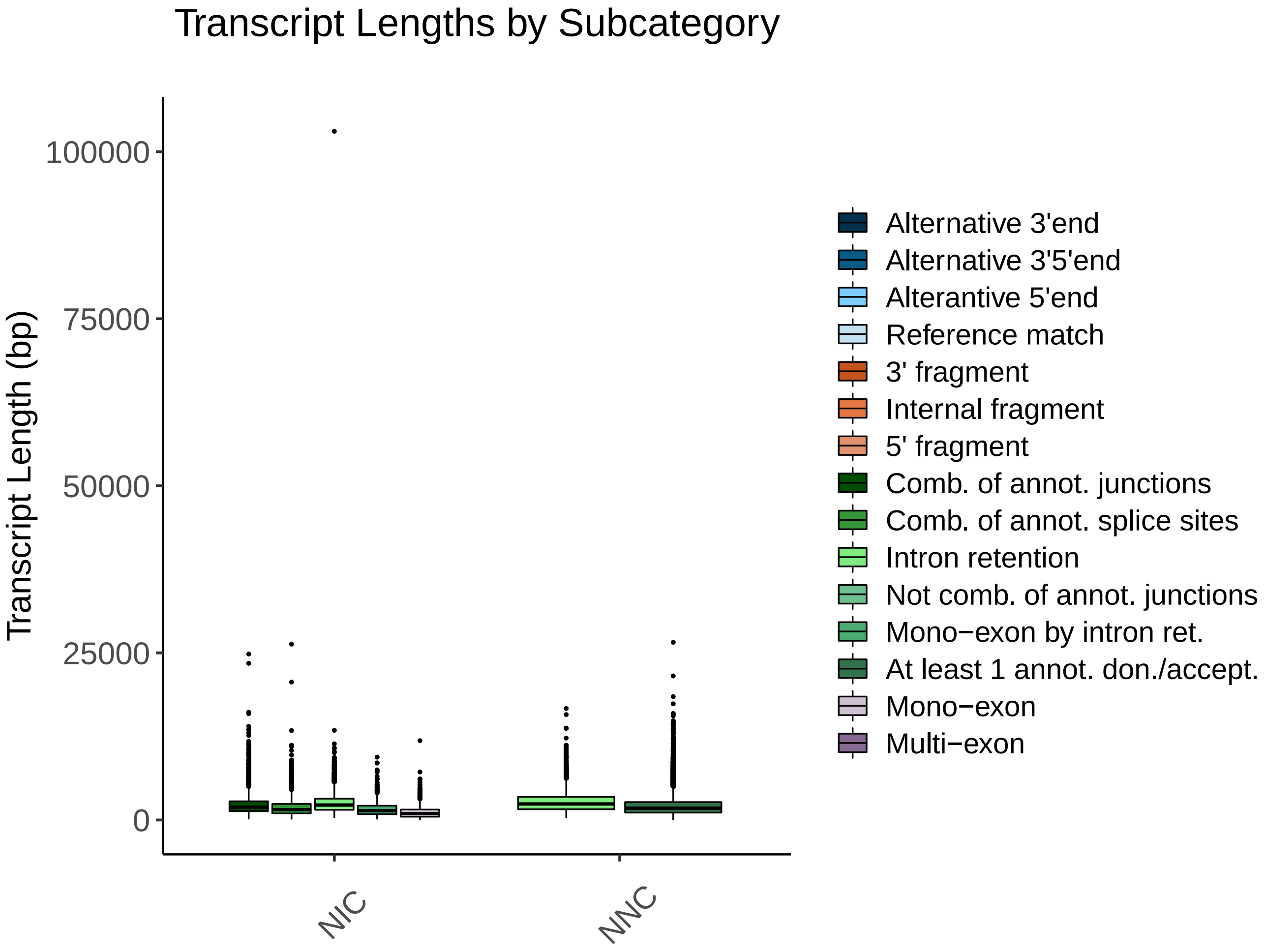
Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding
Non coding

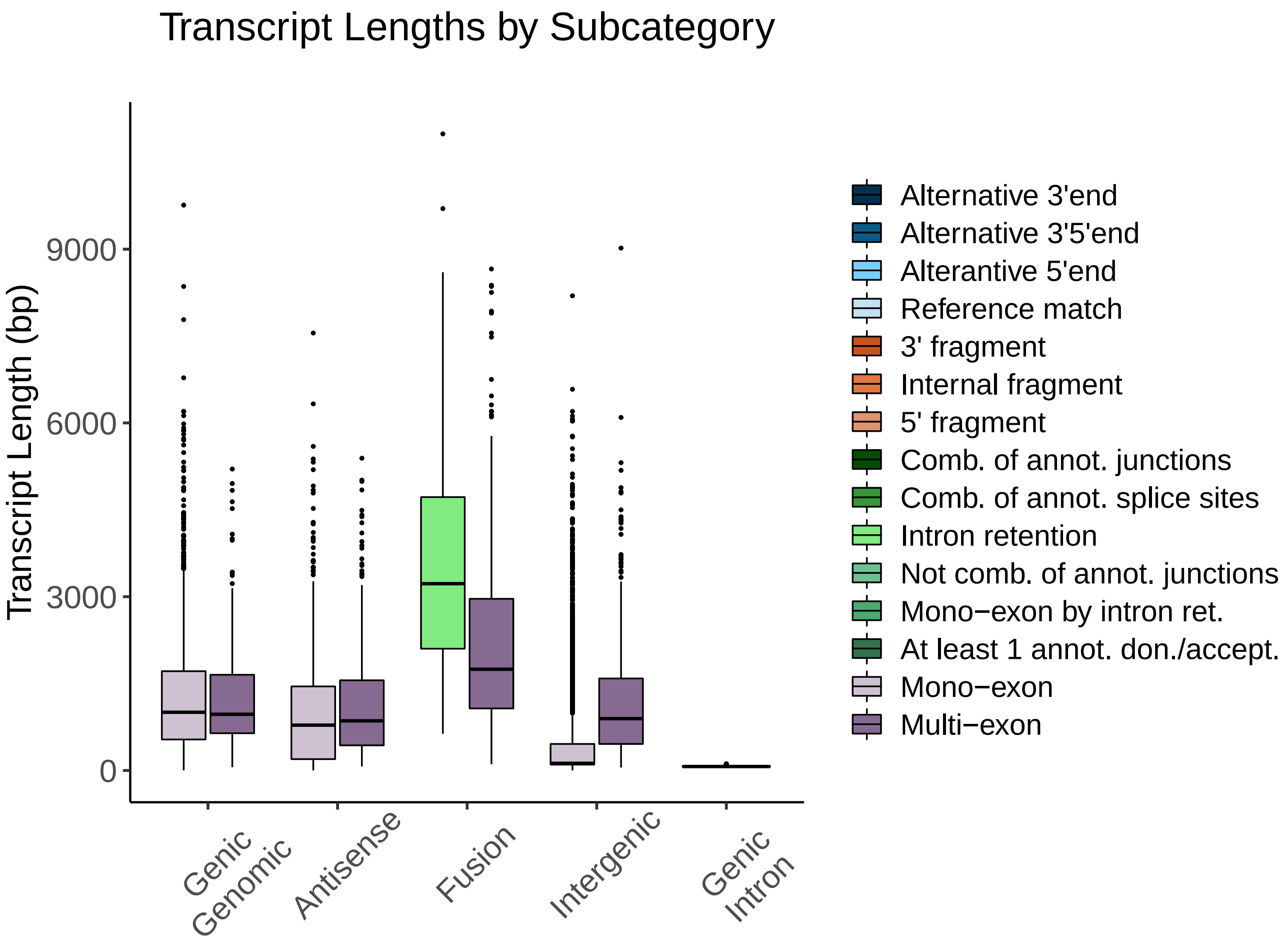

Isoform Distribution Across Fusion
Transcripts, %
0
20
40
60
80
Intron retention
Multi−exon
Coding prediction
Coding
Non coding

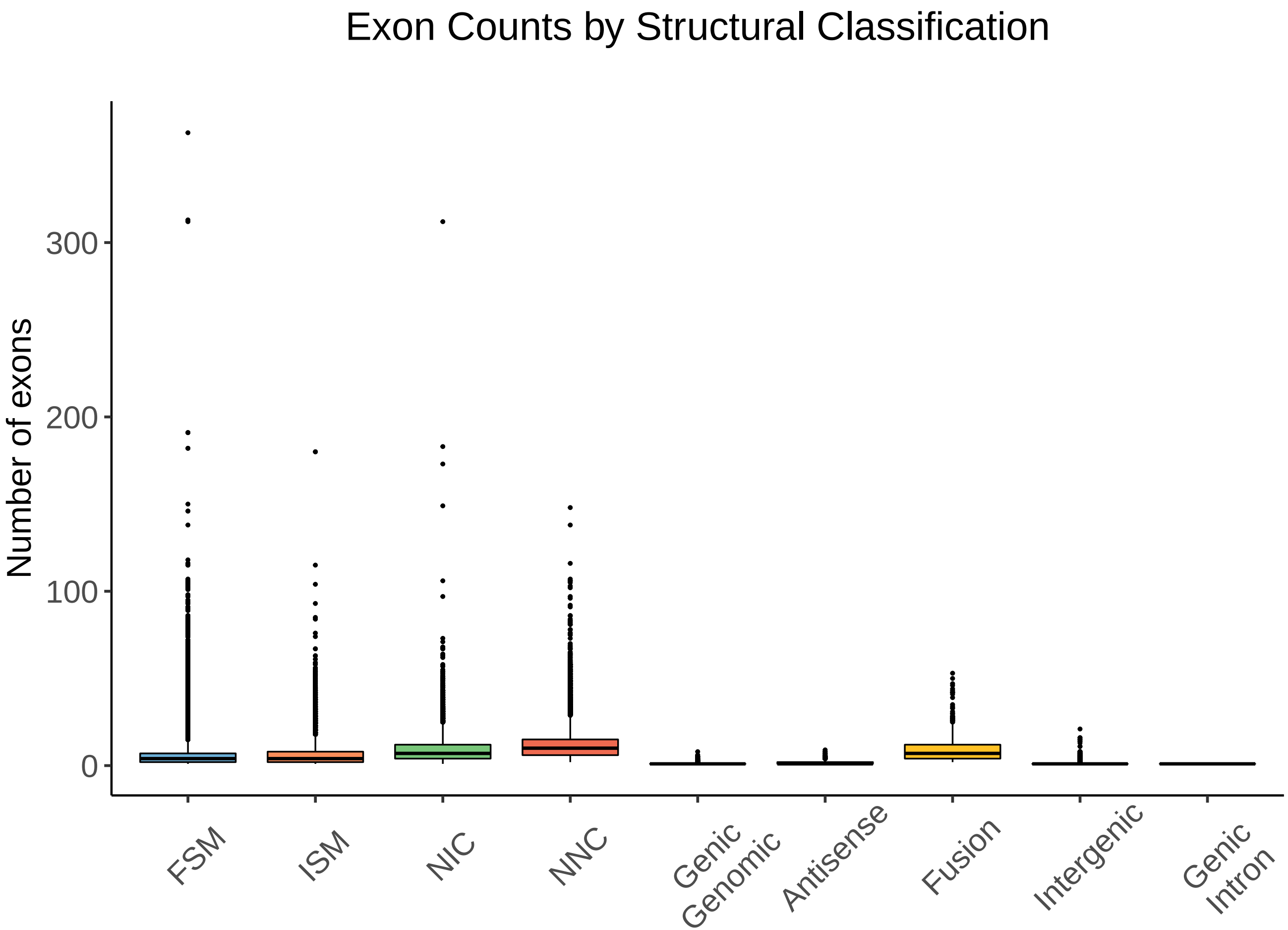

Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi–exon
Coding prediction
Coding
Non coding

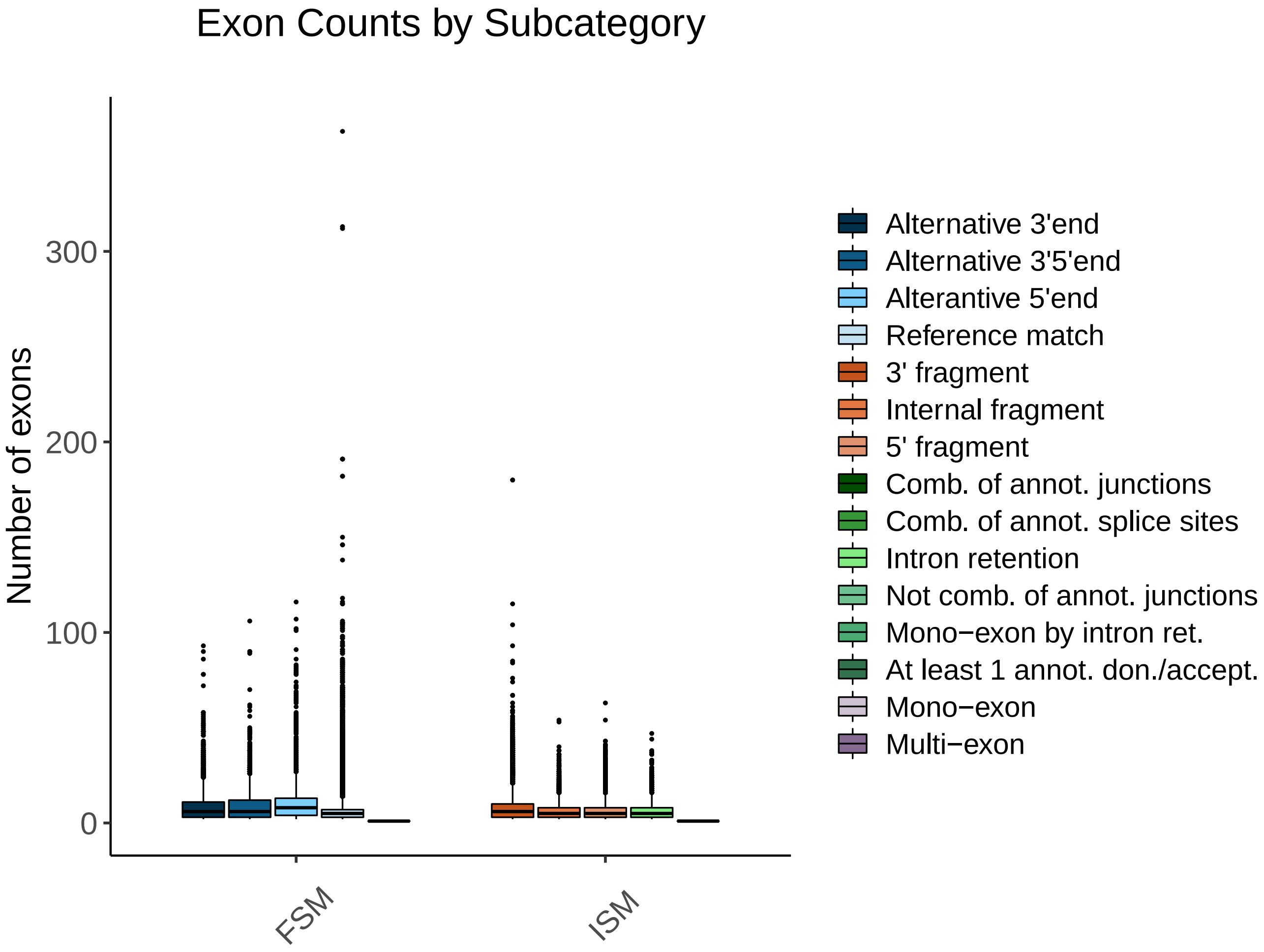

Transcript Lengths by Structural Classification
Transcript Length (bp)
200000
150000
100000
50000
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

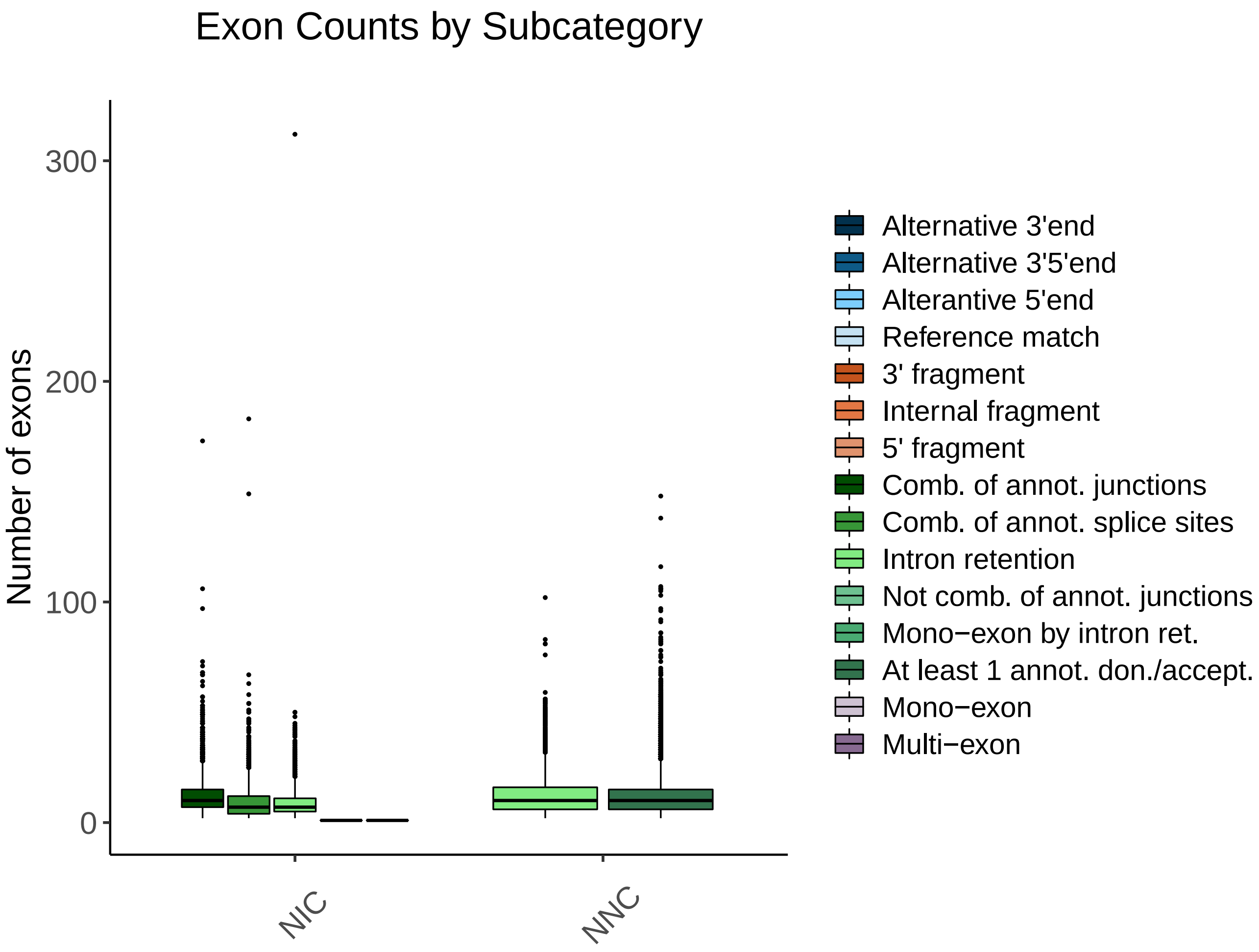

Transcript Lengths by Subcategory
Transcript Length (bp)
200000
150000
100000
50000
0
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

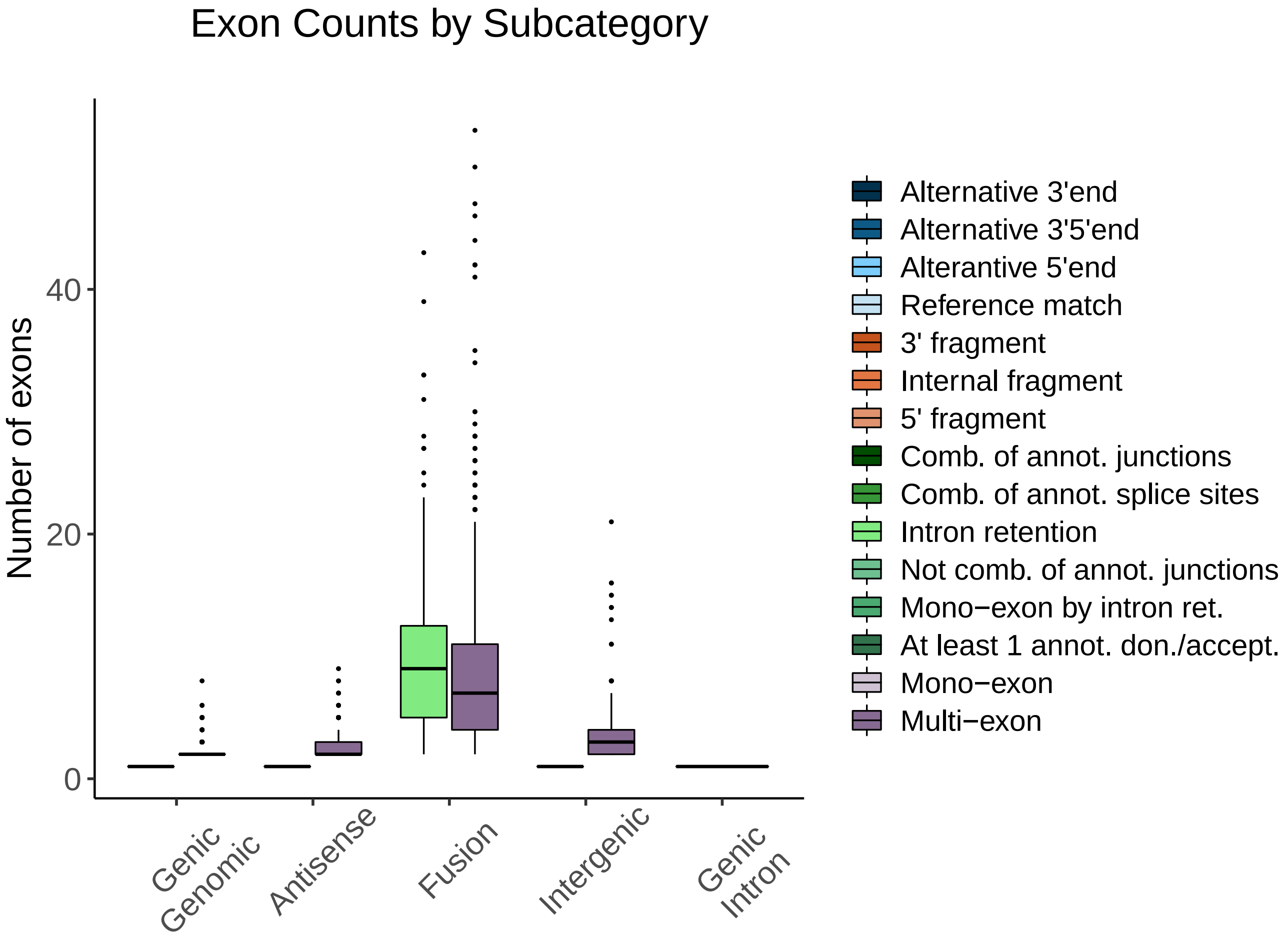

Transcript Lengths by Subcategory
Transcript Length (bp)
100000
75000
50000
25000
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

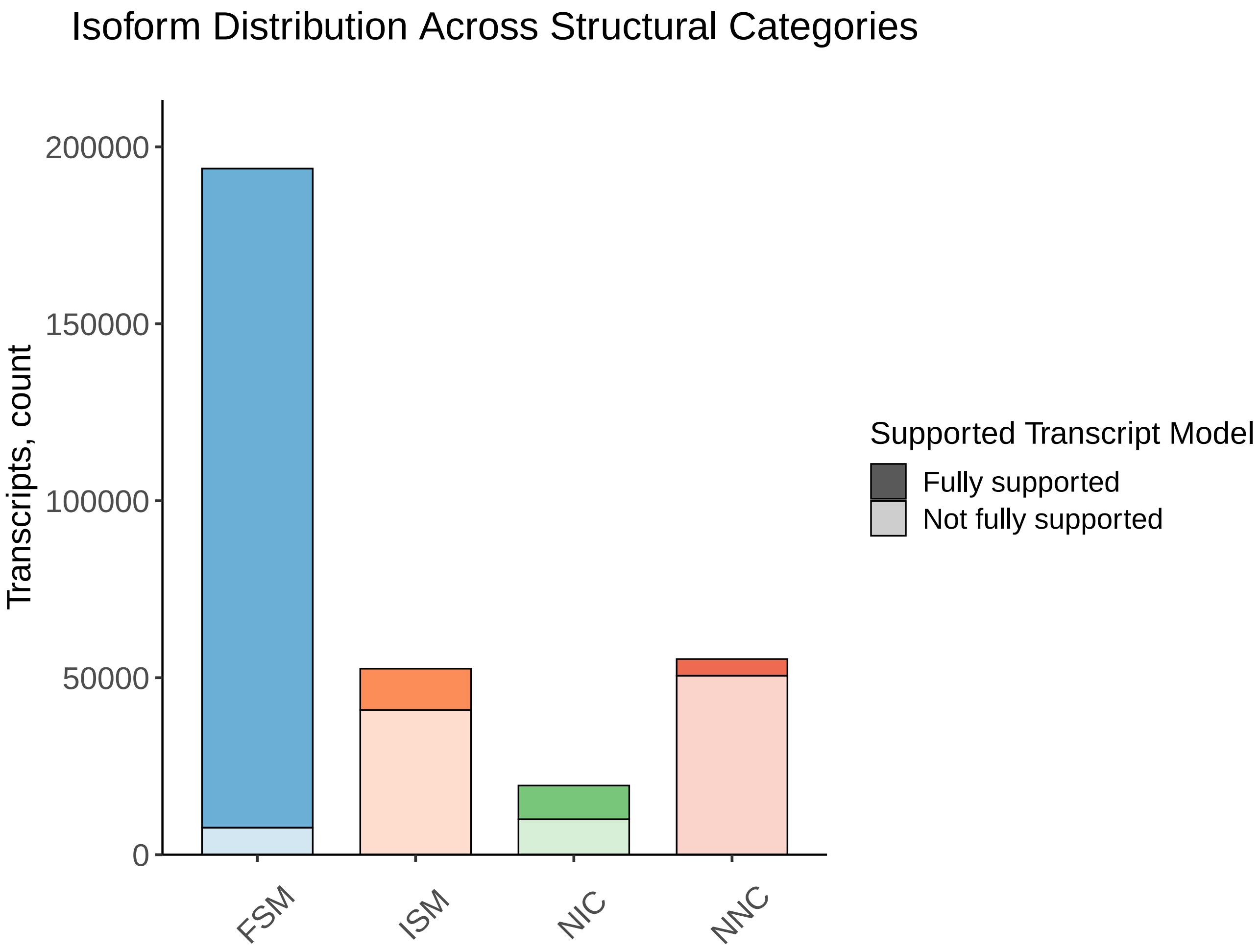
Transcript Lengths by Subcategory
Transcript Length (bp)
0
3000
6000
9000
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron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

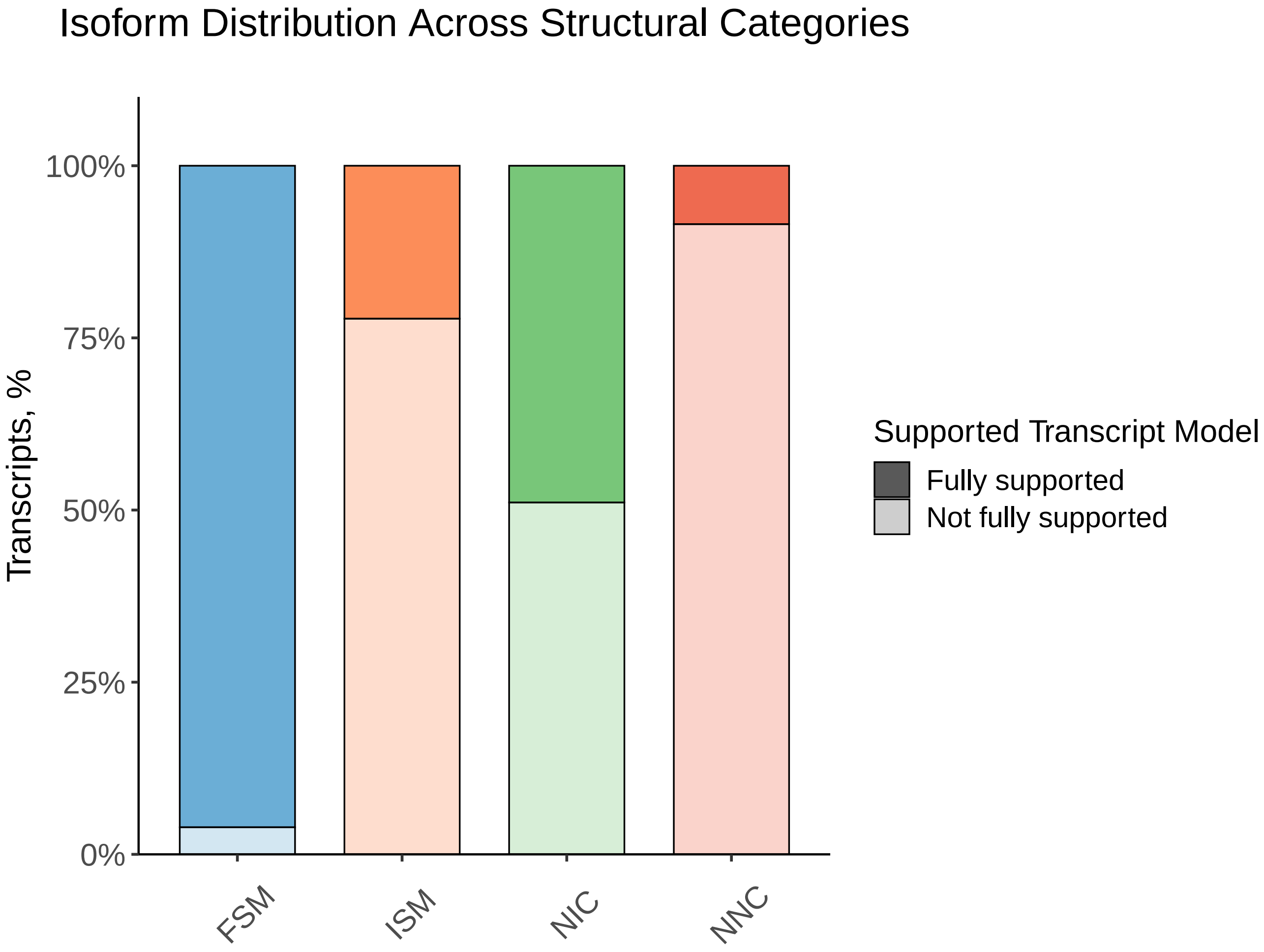

Exon Counts by Structural Classification
Number of exons
300
200
100
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

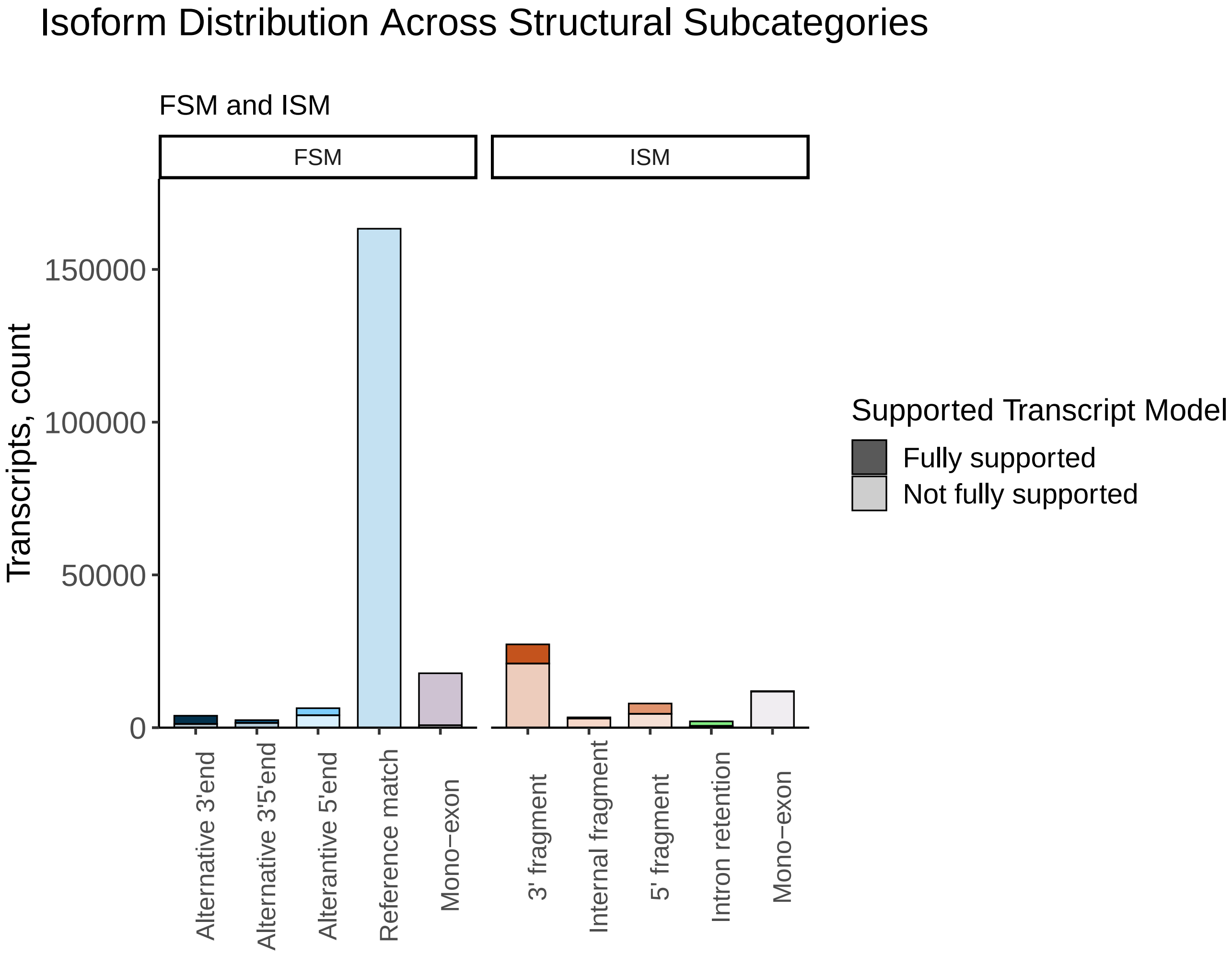

Exon Counts by Subcategory
Number of exons
300
200
100
0
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

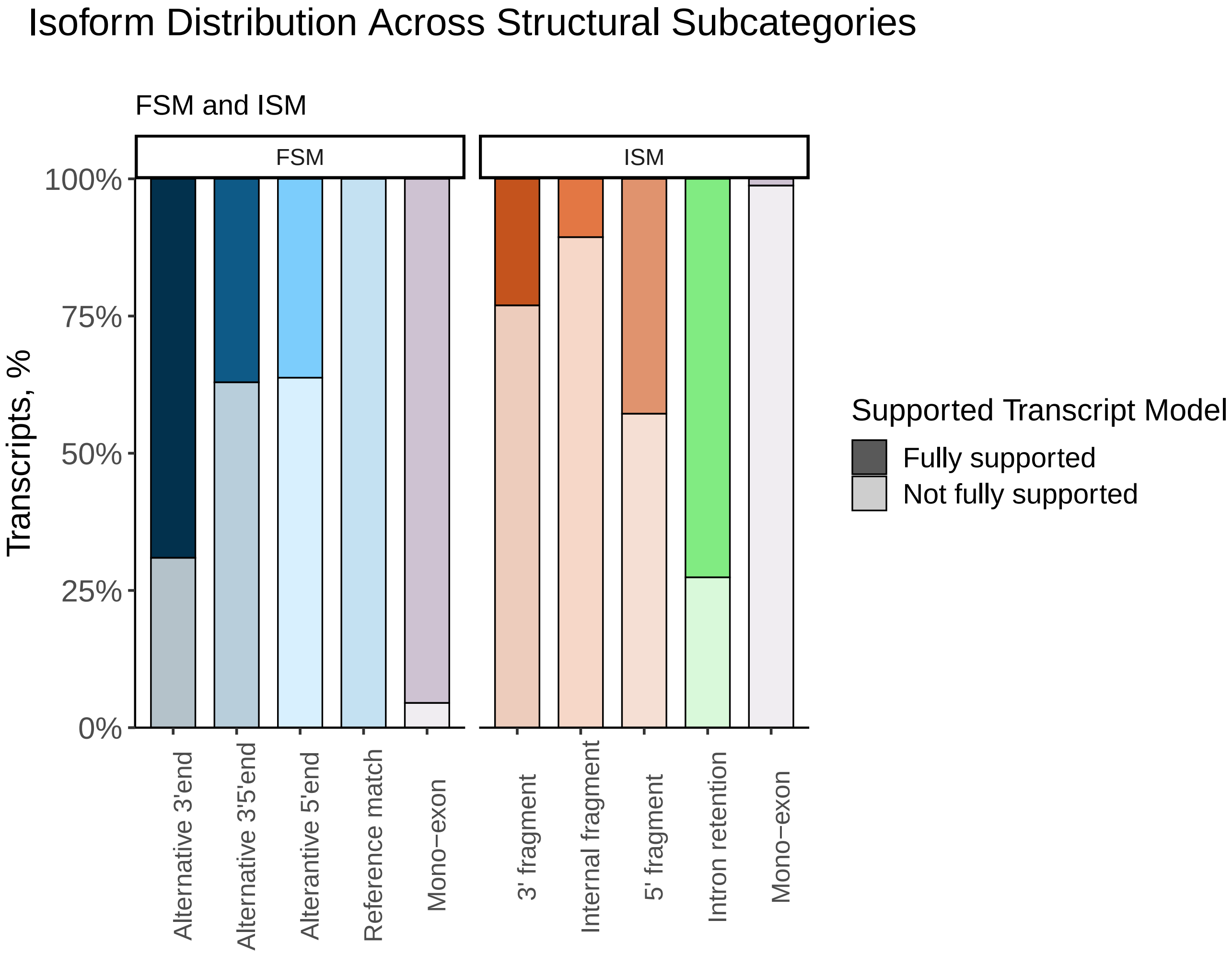

Exon Counts by Subcategory
Number of exons
300
200
100
0
NIC
NNC
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

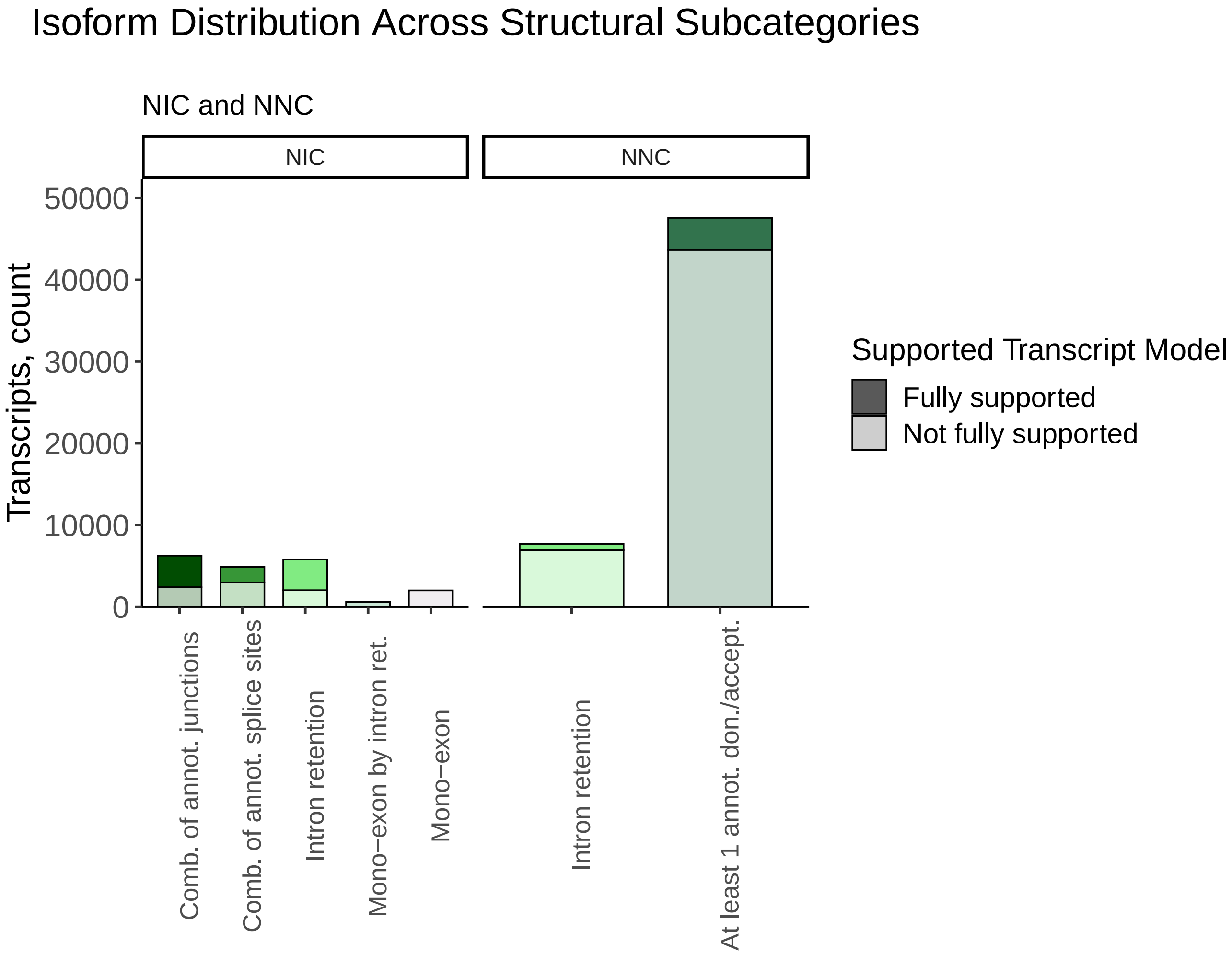

Exon Counts by Subcategory
Number of exons
0
20
40
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb.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

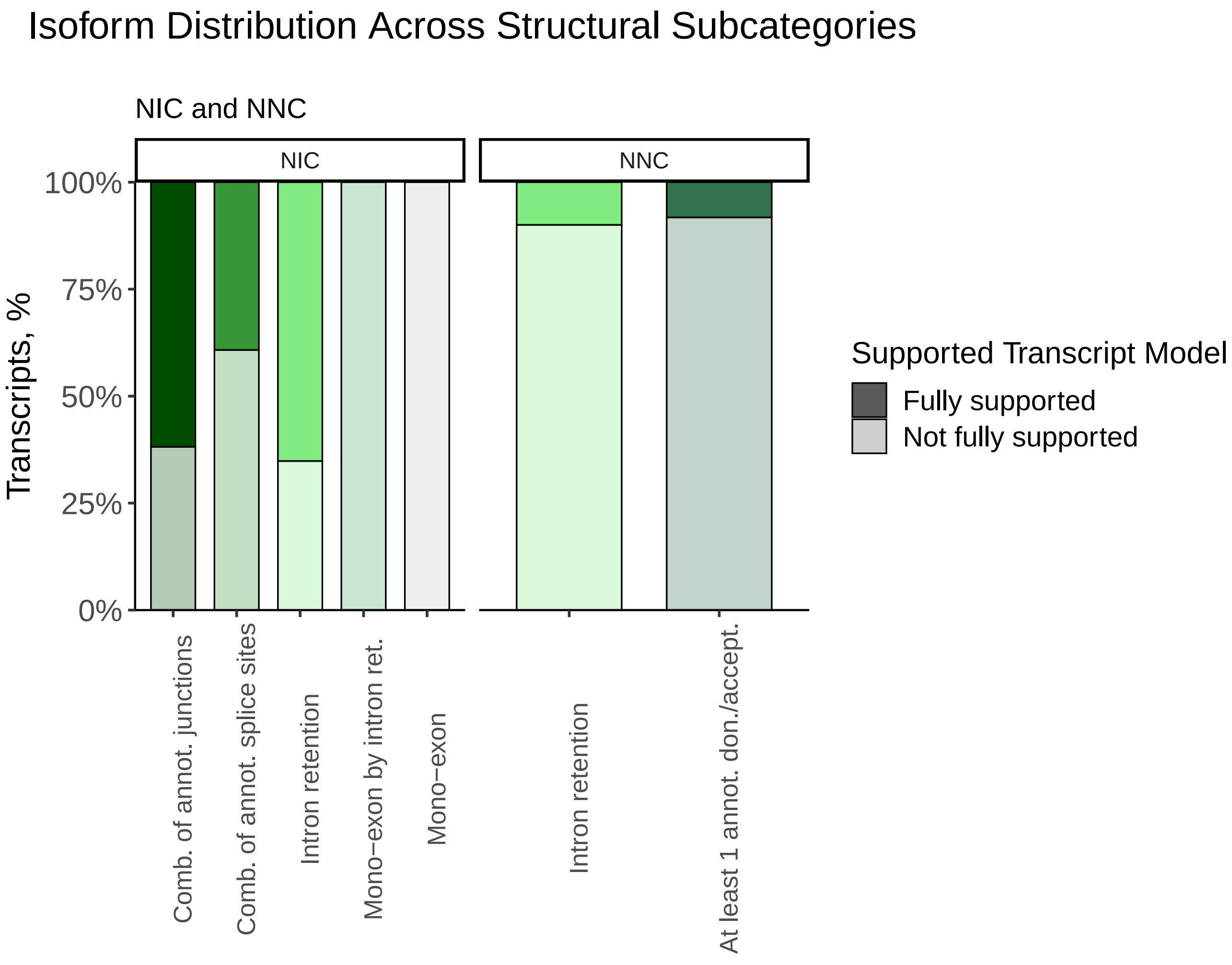

Isoform Distribution Across Structural Categories
Transcripts, count
200000
150000
100000
50000
0
FSM
ISM
NIC
NNC
Supported Transcript Model
Fully supported
Not fully supported


Isoform Distribution Across Structural Categories
Transcripts, %
100%
75%
50%
25%
0%
FSM
ISM
NIC
NNC
Supported Transcript Model
Fully supported
Not fully supported


Isoform Distribution Across Structural Subcategories
FSM and ISM
FSM
ISM
Transcripts, count
150000
100000
50000
0
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Mono−exon
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon
Supported Transcript Model
Fully supported
Not fully supported

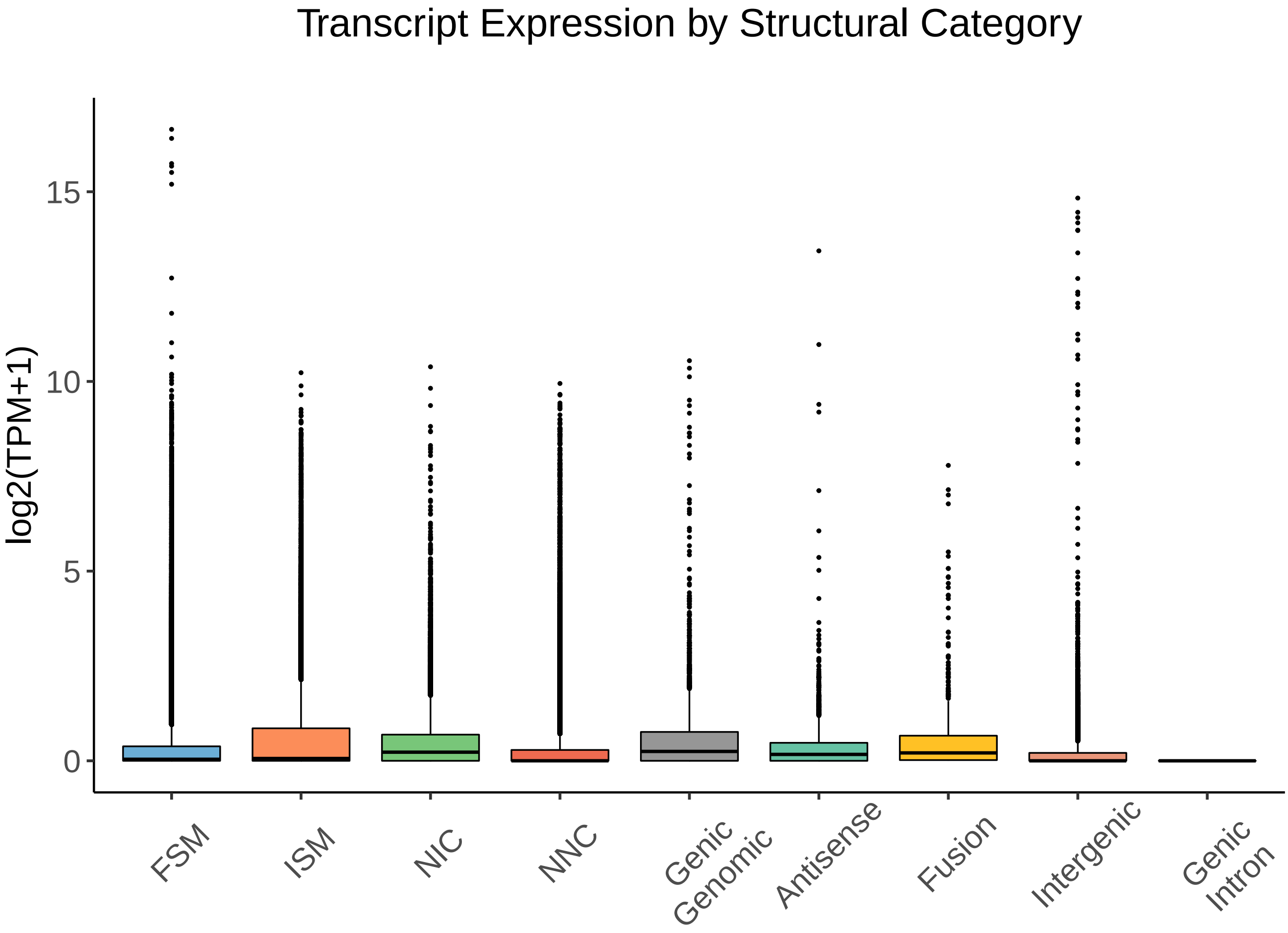

Isoform Distribution Across Structural Subcategories
FSM and ISM
FSM
ISM
Transcripts, %
100%
75%
50%
25%
0%
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Mono−exon
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon
Supported Transcript Model
Fully supported
Not fully supported


Isoform Distribution Across Structural Subcategories
NIC and NNC
NIC
NNC
Transcripts, count
50000
40000
30000
20000
10000
0
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
Mono−exon
Intron retention
At least 1 annot. don./accept.
Supported Transcript Model
Fully supported
Not fully supported

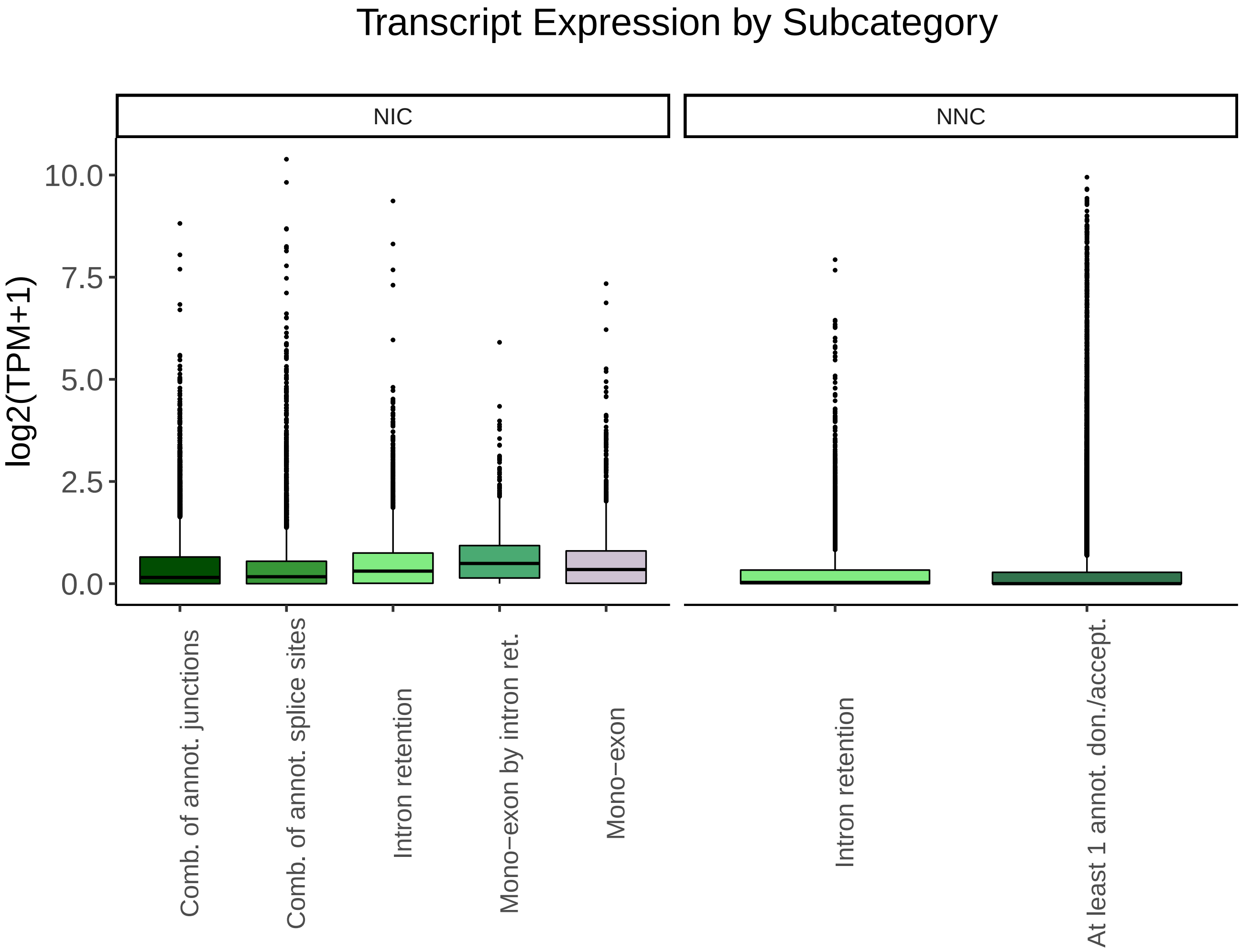


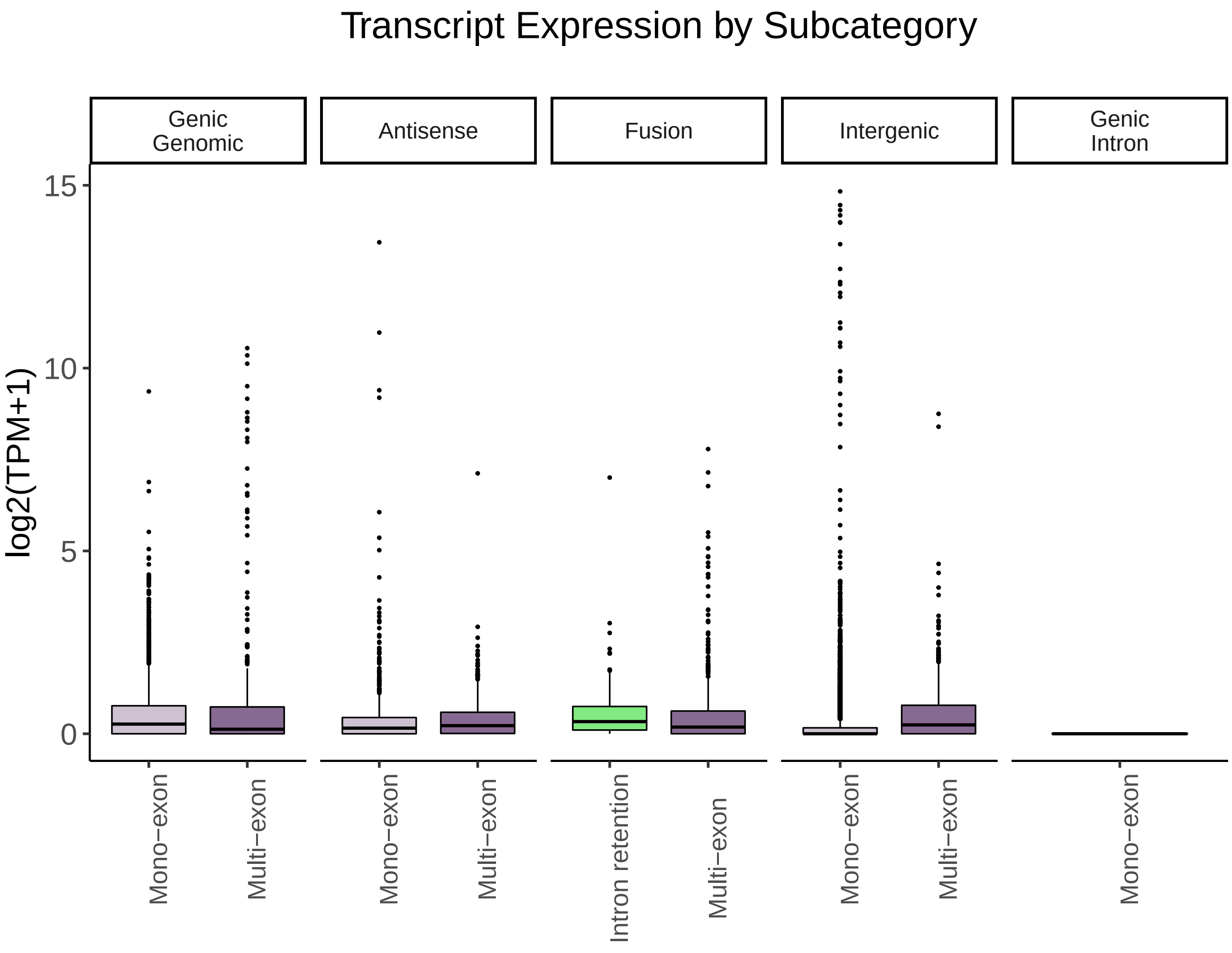
Isoform Distribution Across Structural Subcategories

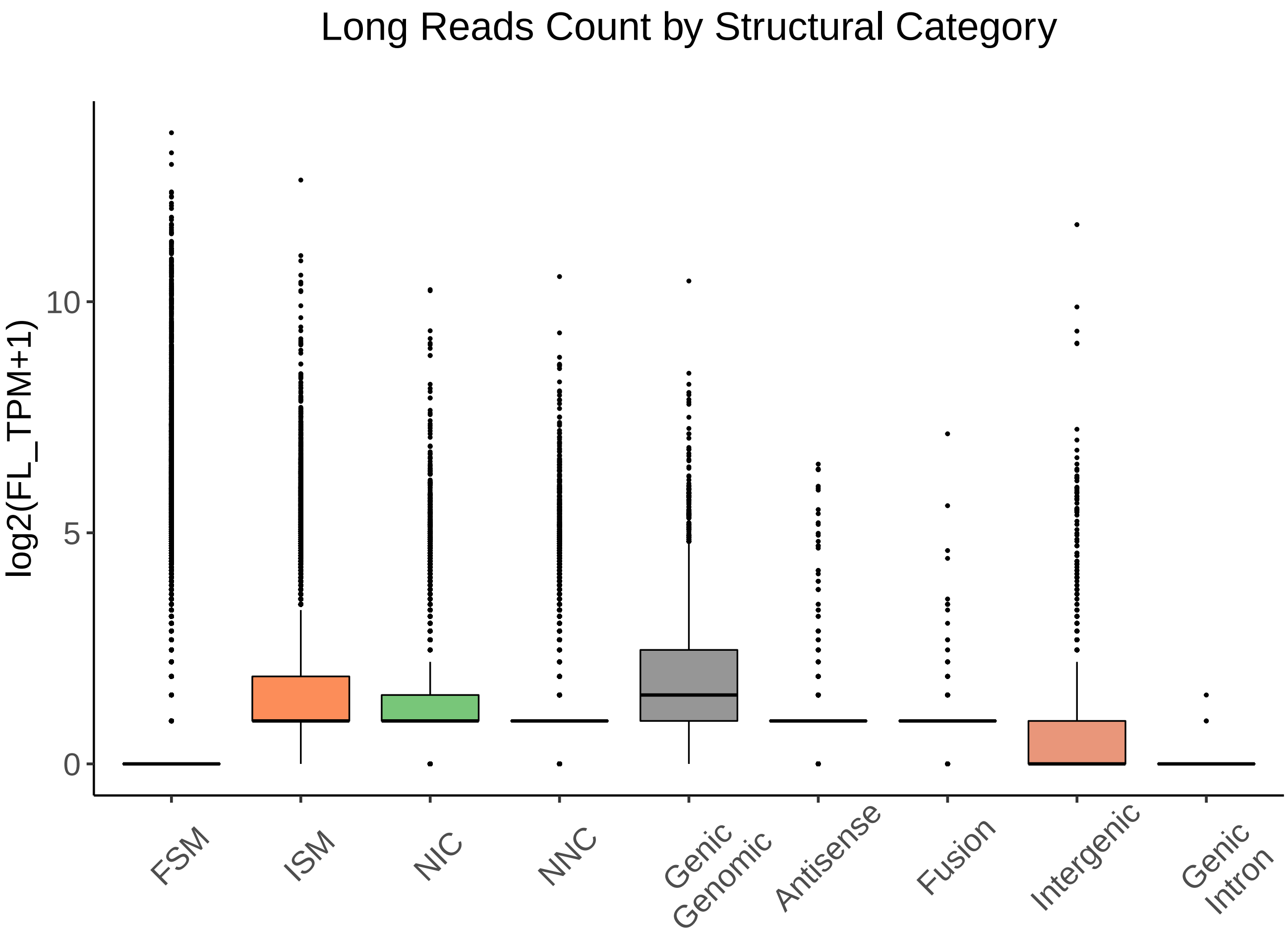
NIC and NNC

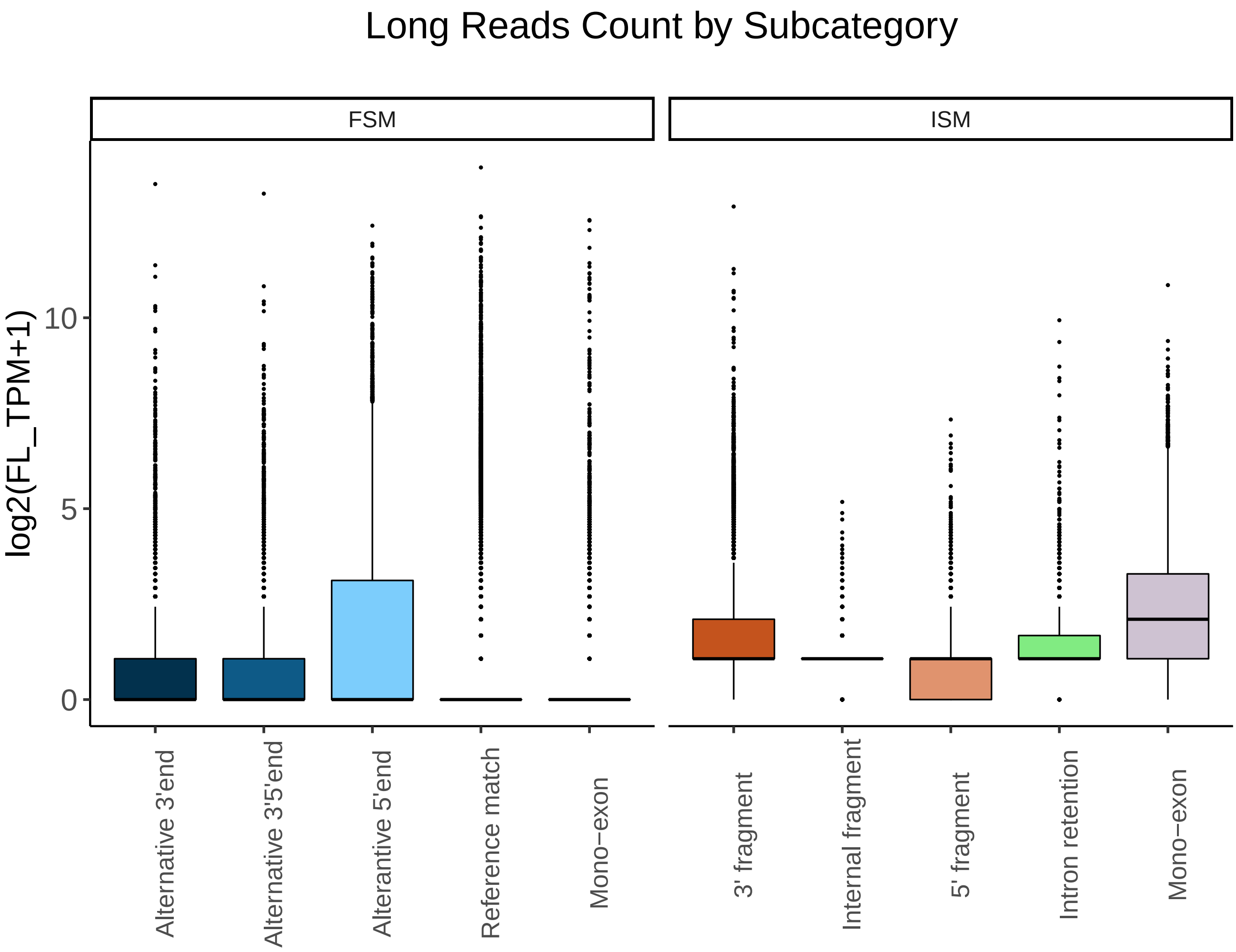

Transcript Expression by Structural Category
log2(TPM+1)
15
10
5
0
FSM
ISM
NIC
NNC
Genic
Genomic
Antisense
Fusion
Intergenic
Genic
Intron

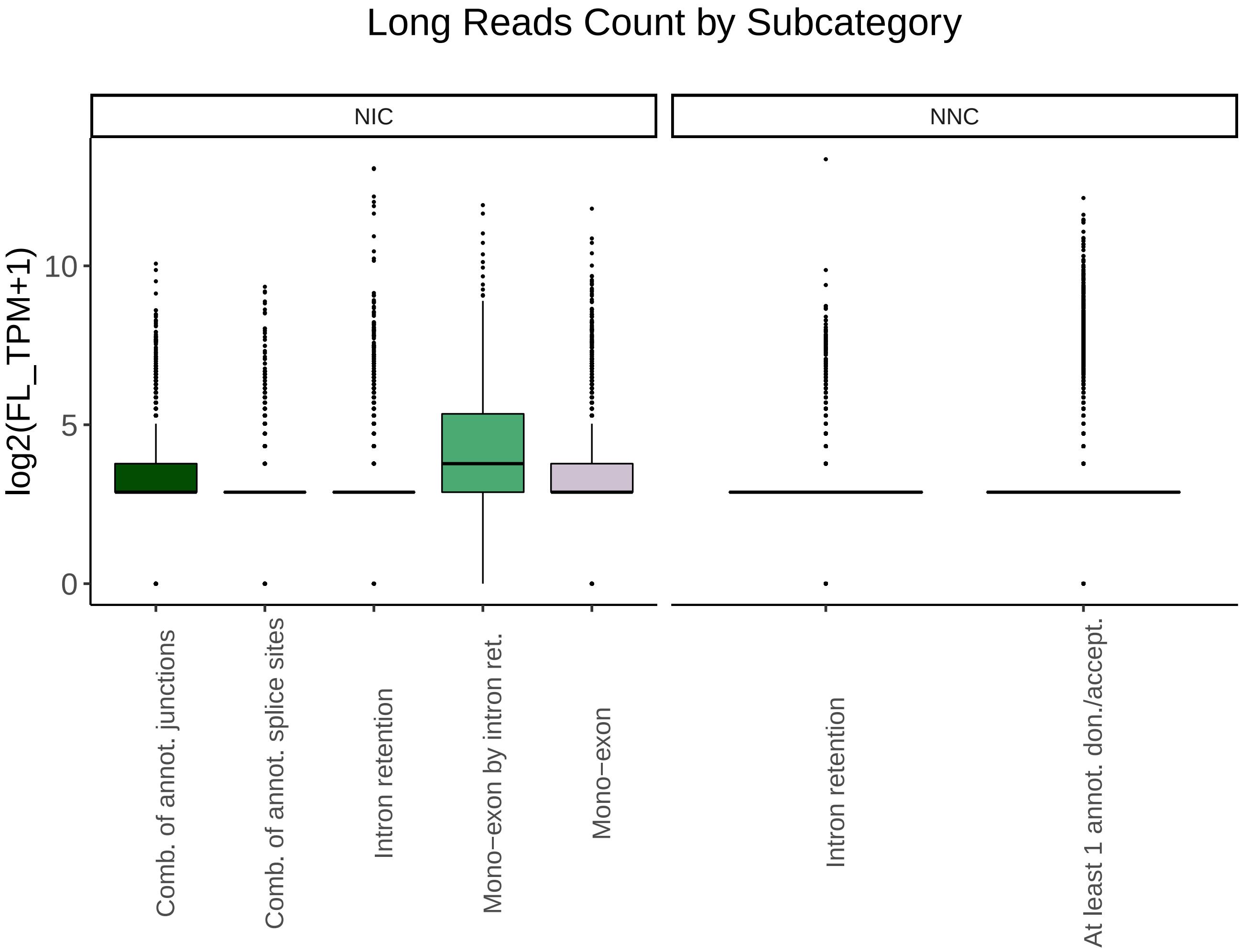

Transcript Expression by Subcategory
FSM
ISM
log2(TPM+1)
15
10
5
0
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Mono−exon
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon

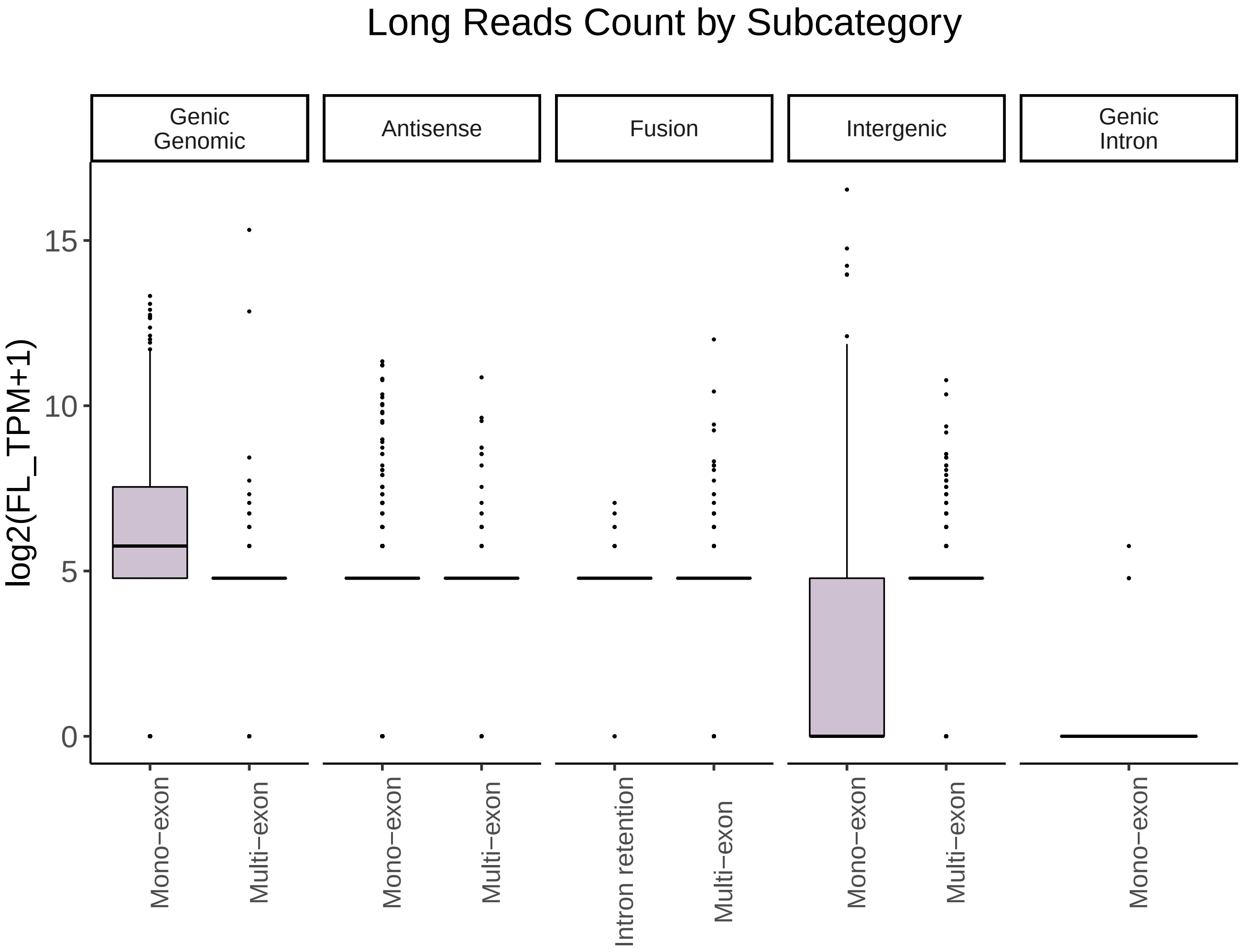

Transcript Expression by Subcategory
NIC
NNC
log2(TPM+1)
10.0
7.5
5.0
2.5
0.0
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
Mono−exon
Intron retention
At least 1 annot. don./accept.

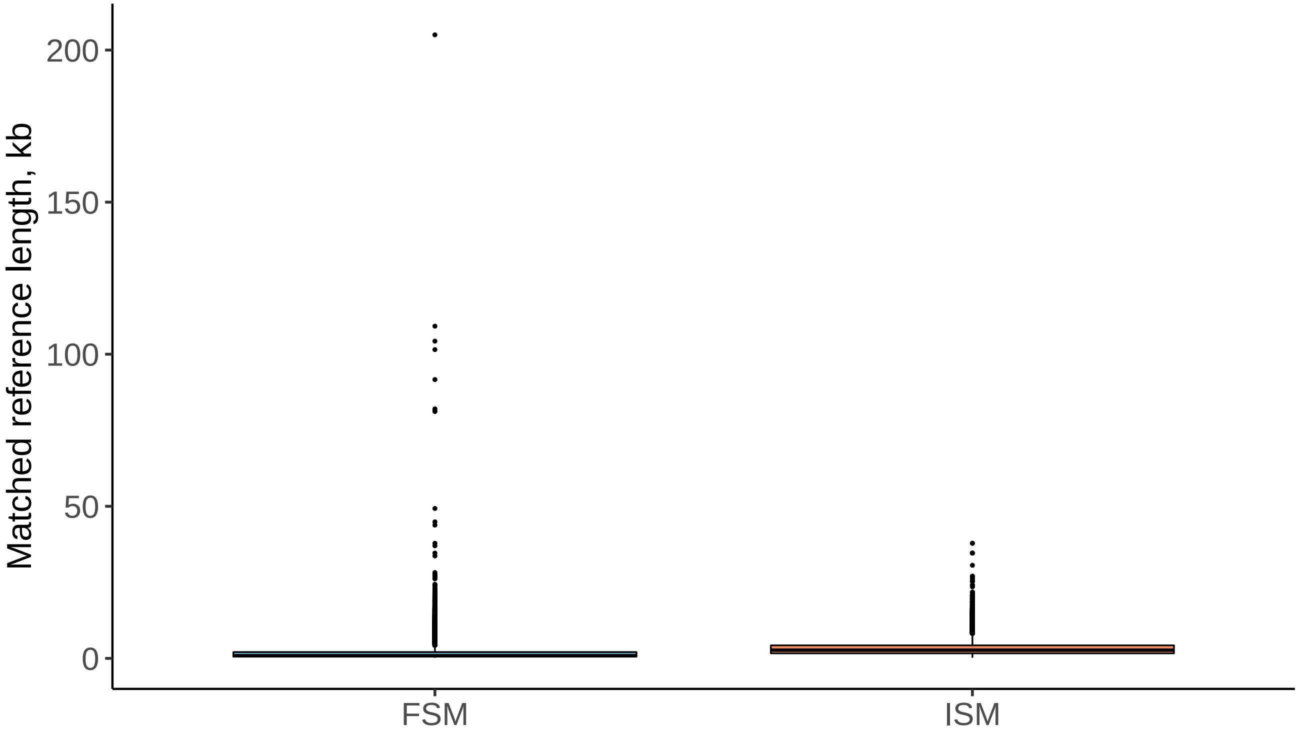

Transcript Expression by Subcategory
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
log2(TPM+1)
15
10
5
0
Mono−exon
Multi−exon
Mono−exon
Multi−exon
Intron retention
Multi−exon
Mono−exon
Multi−exon
Mono−exon


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


Long Reads Count by Subcategory
FSM
ISM
log2(FL_TPM+1)
10
5
0
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
Mono−exon
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon

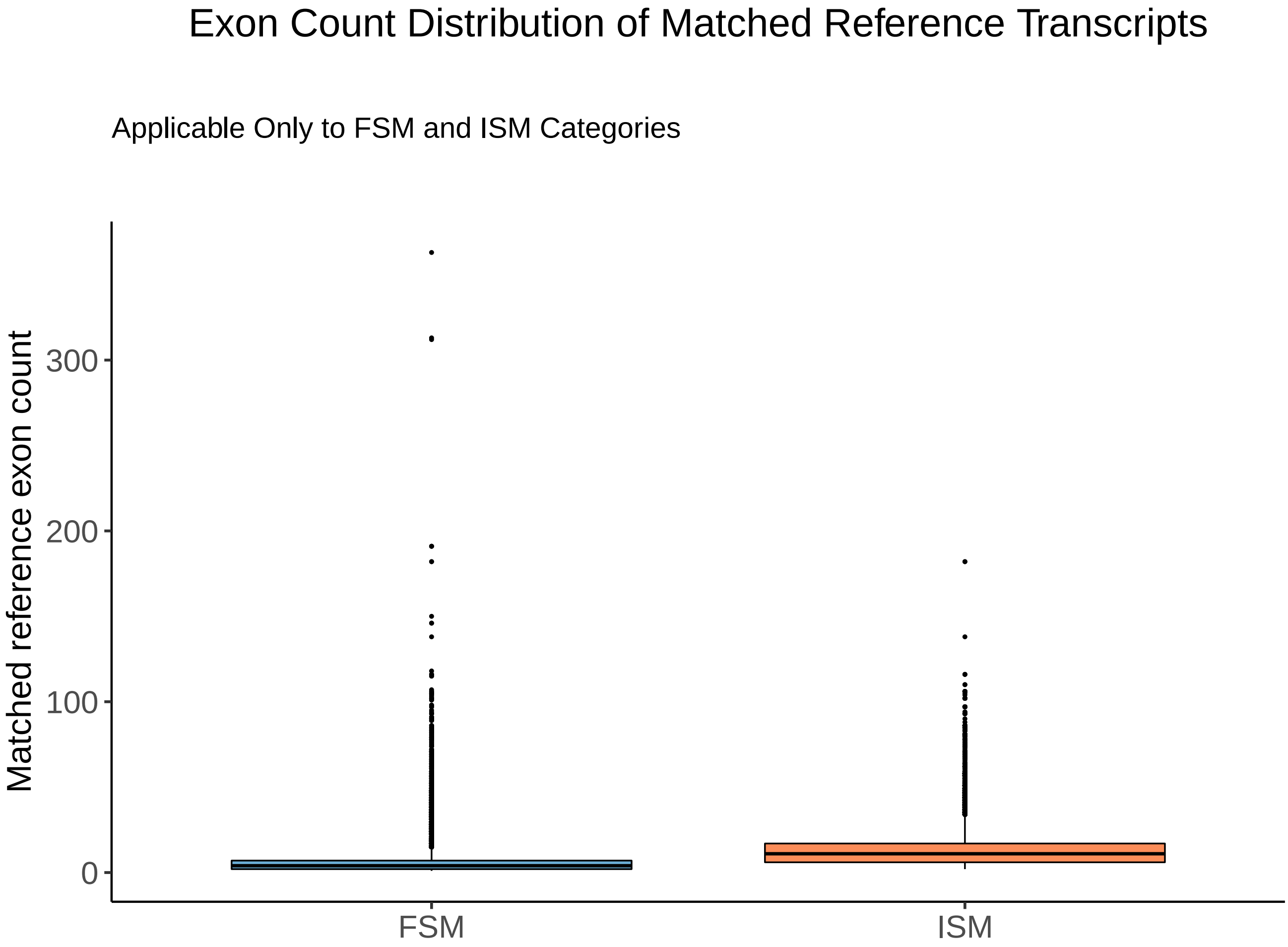

Long Reads Count by Subcategory
NIC
NNC
log2(FL_TPM+1)
0
5
10
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
Mono−exon
Intron retention
At least 1 annot. don./accept.


Long Reads Count by Subcategory
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
log2(FL_TPM+1)
0
5
10
15
Mono−exon
Multi−exon
Mono−exon
Multi−exon
Intron retention
Multi−exon
Mono−exon
Multi−exon
Mono−exon



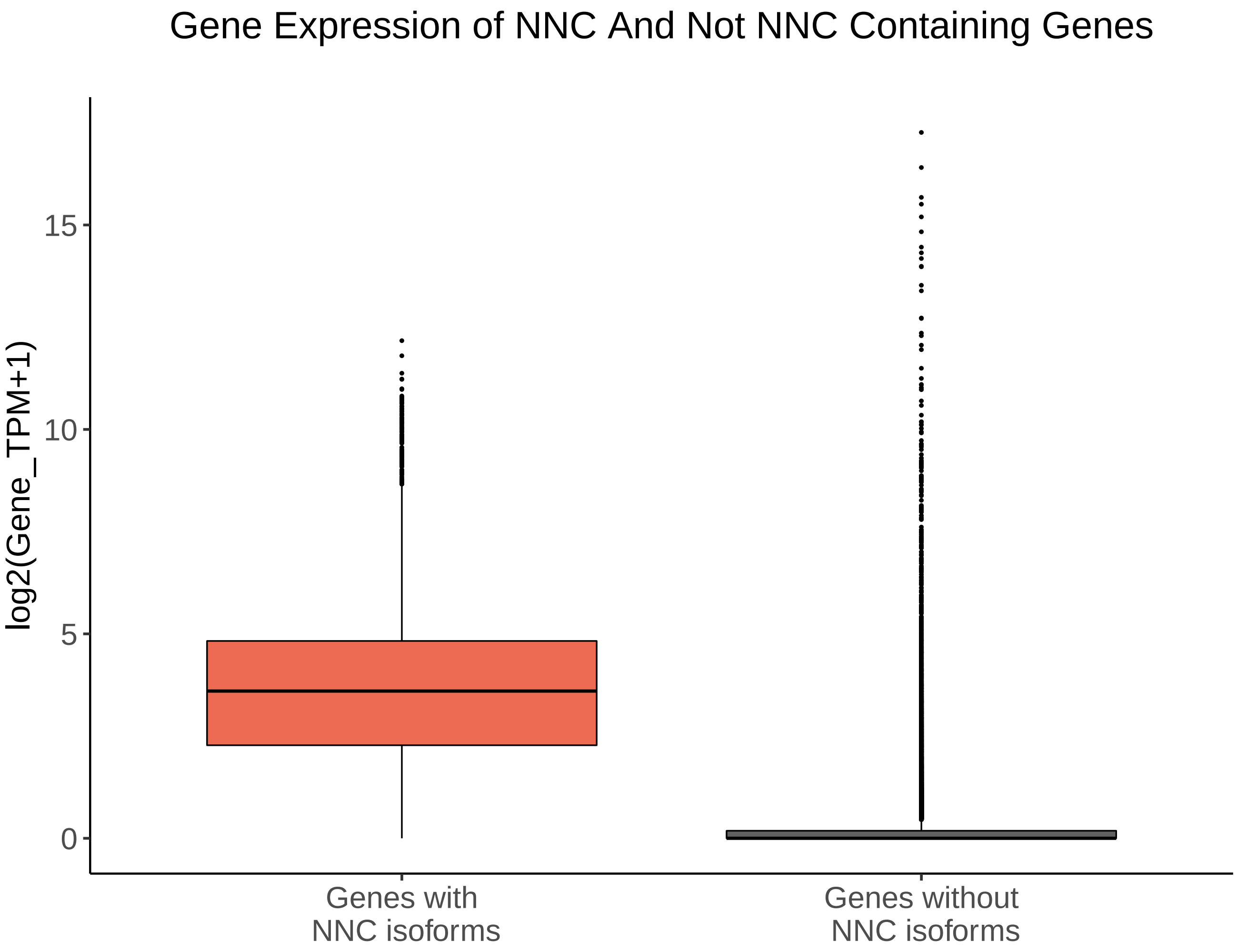
Length Distribution of Matched Reference Transcripts

Applicable Only to FSM and ISM Categories

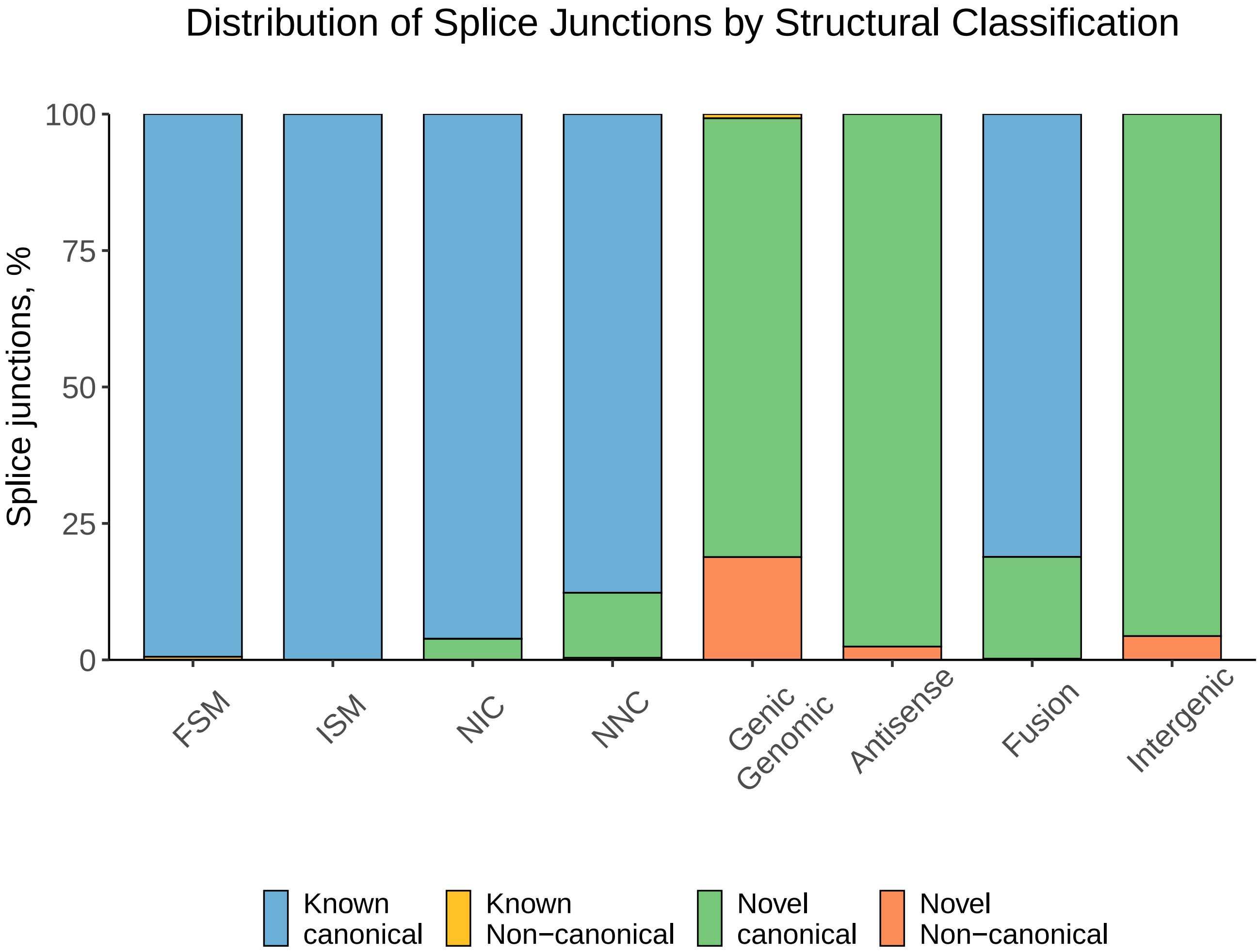


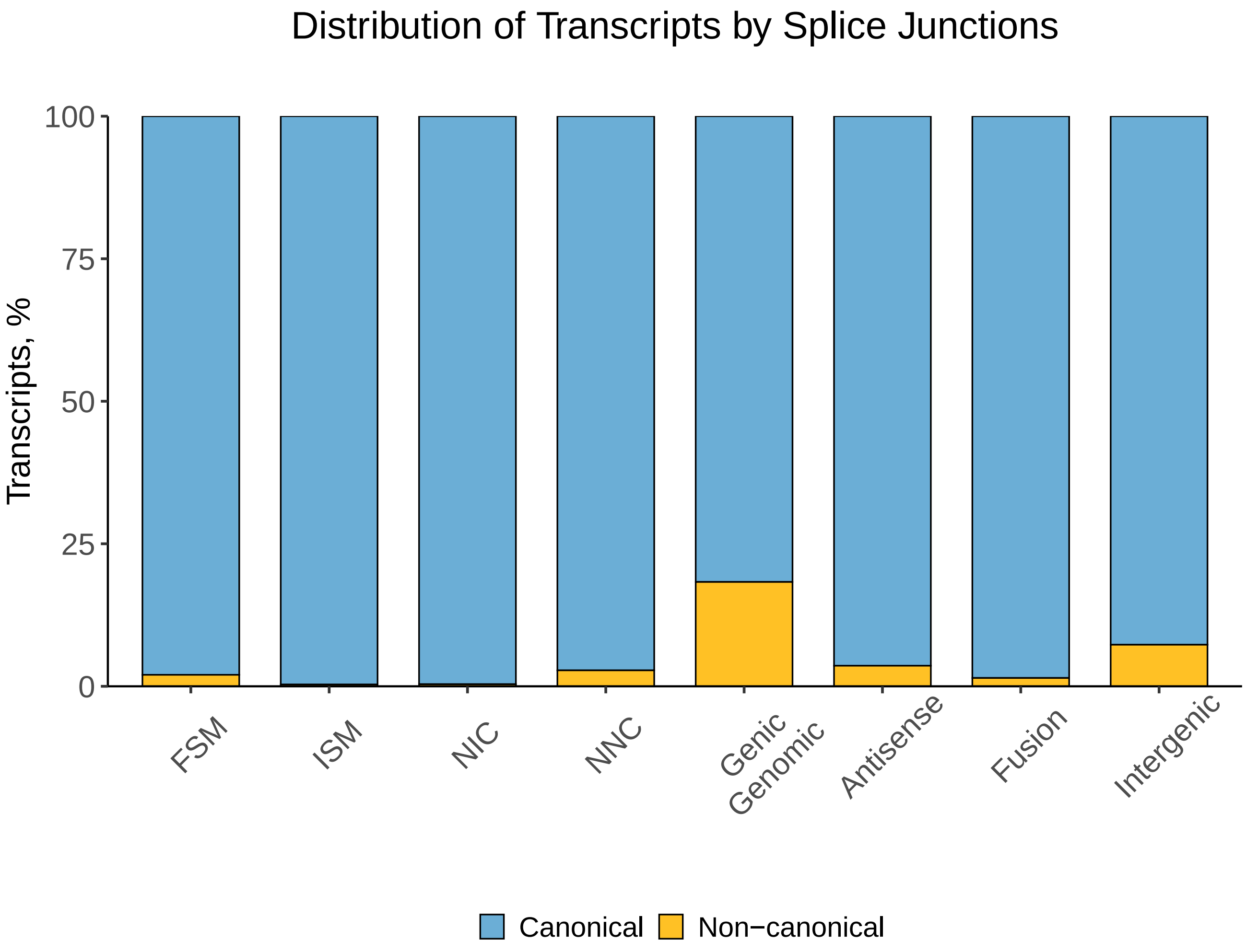
Exon Count Distribution of Matched Reference Transcripts

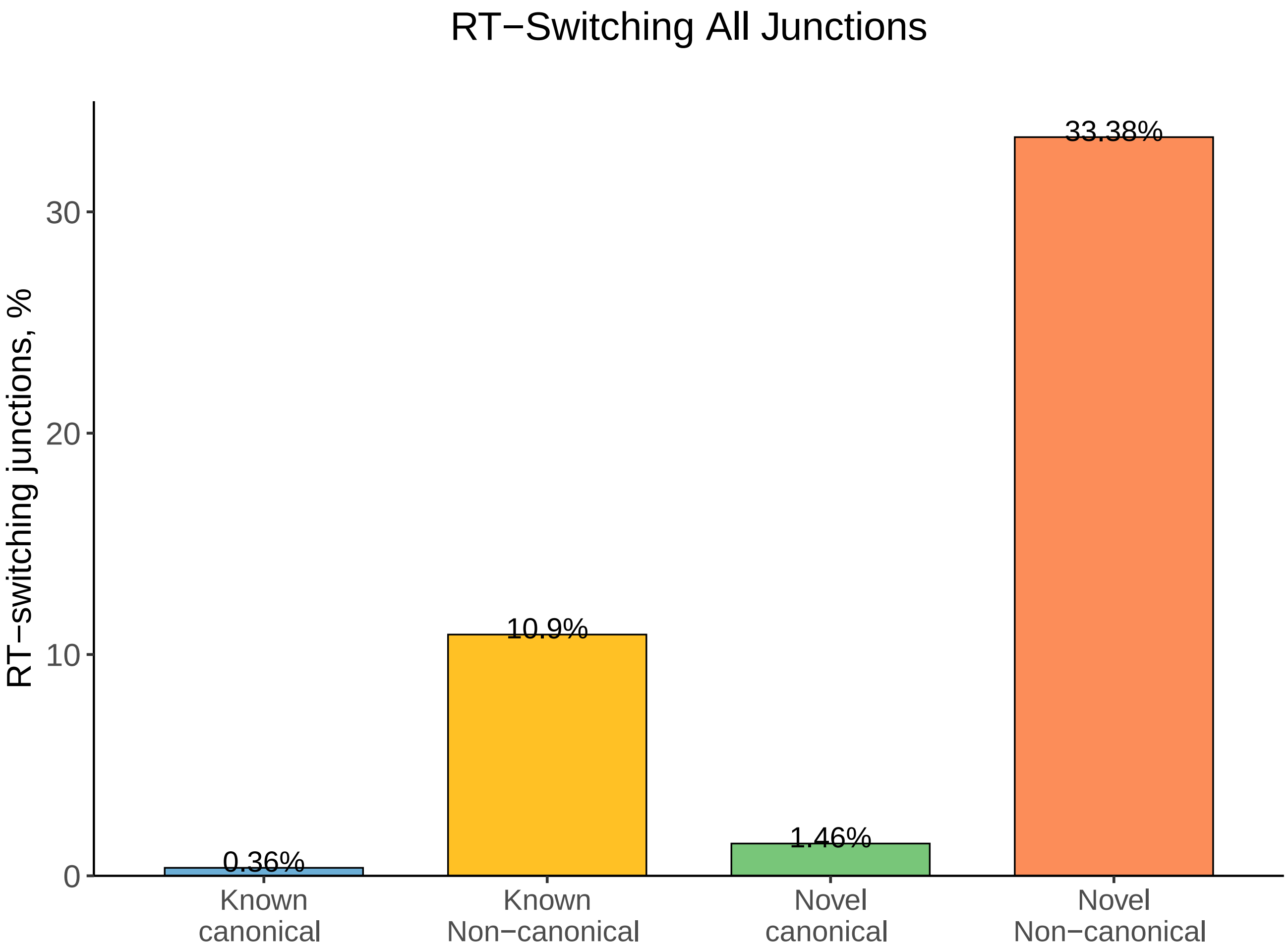
Applicable Only to FSM and ISM Categories

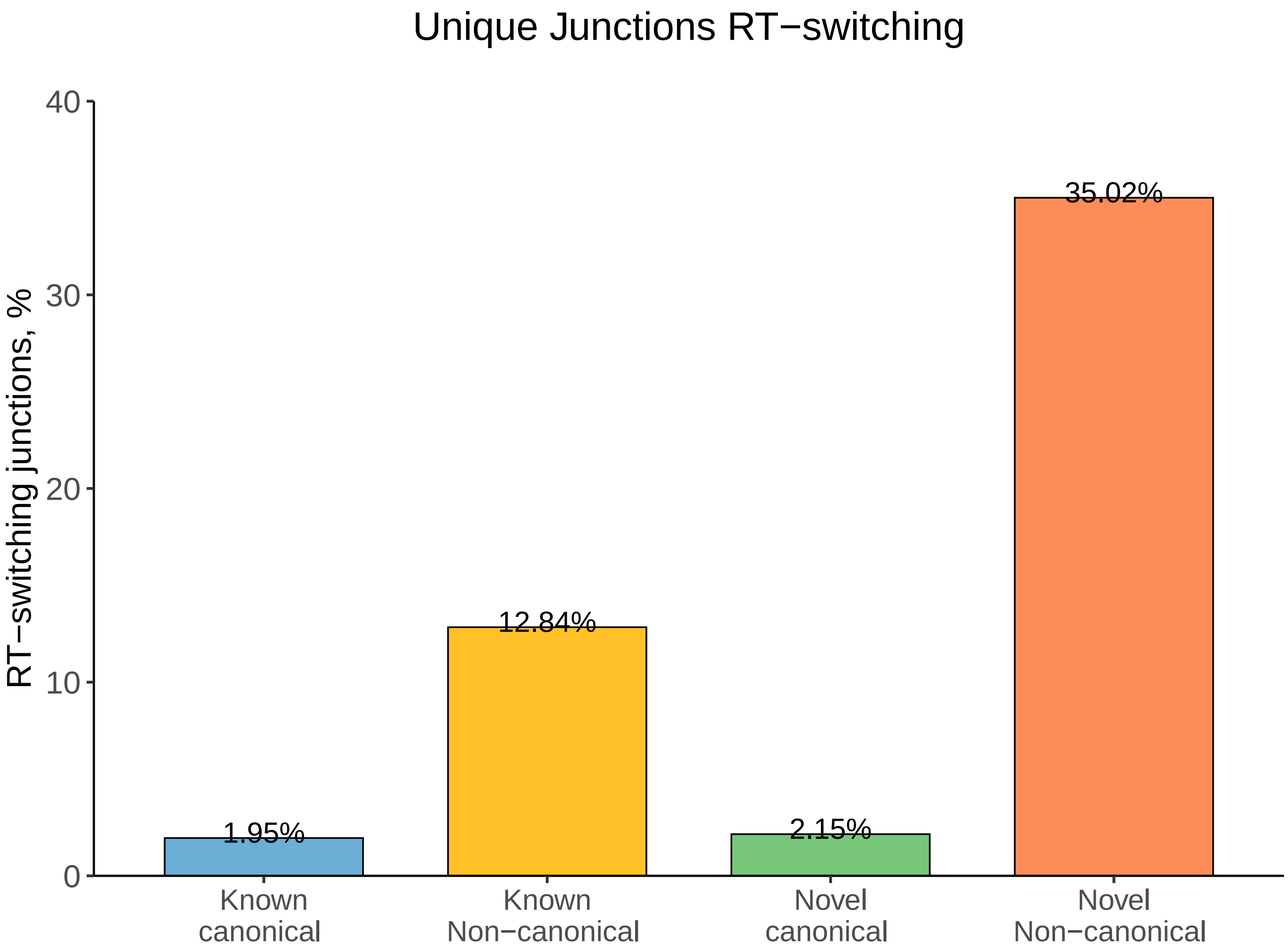

Gene Expression of NNC And Not NNC Containing Genes
log2(Gene_TPM+1)
15
10
5
0
Genes with NNC isoforms
Genes without NNC isoforms

Splice Junction Characterization

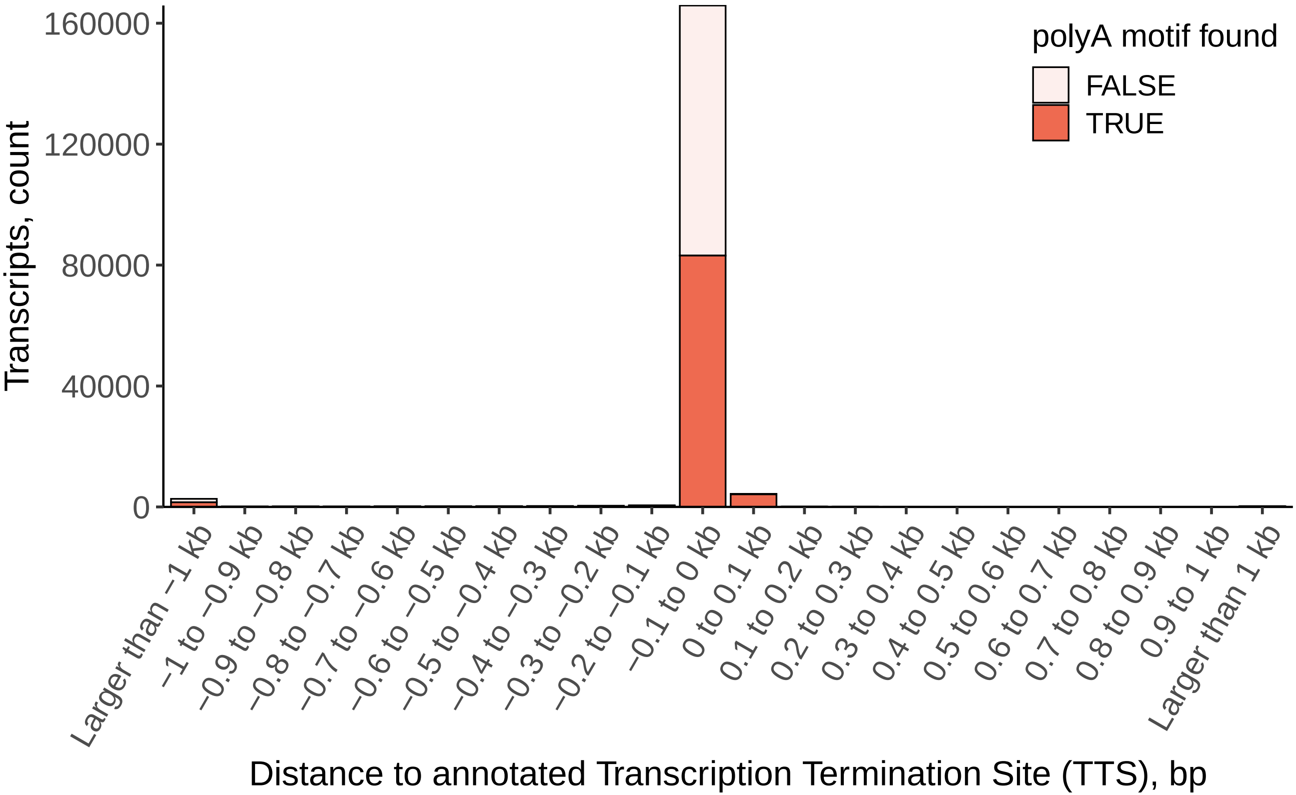

Distribution of Splice Junctions by Structural Classification
Splice junctions, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical


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, %
0
10
20
30
0.36%
10.9%
1.46%
33.38%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

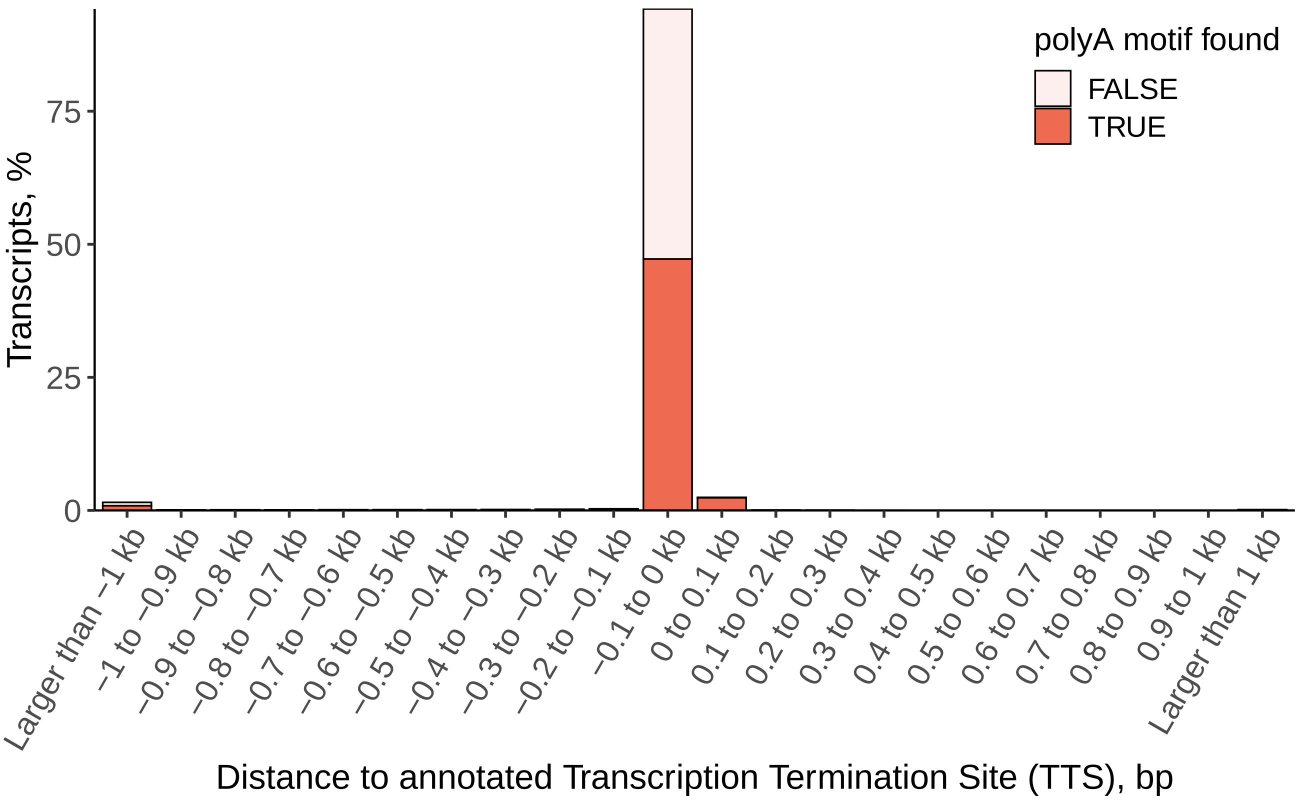

Unique Junctions RT−switching
RT−switching junctions, %
40
30
20
10
0
1.95%
12.84%
2.15%
35.02%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

Comparison With Annotated TSS and TTS

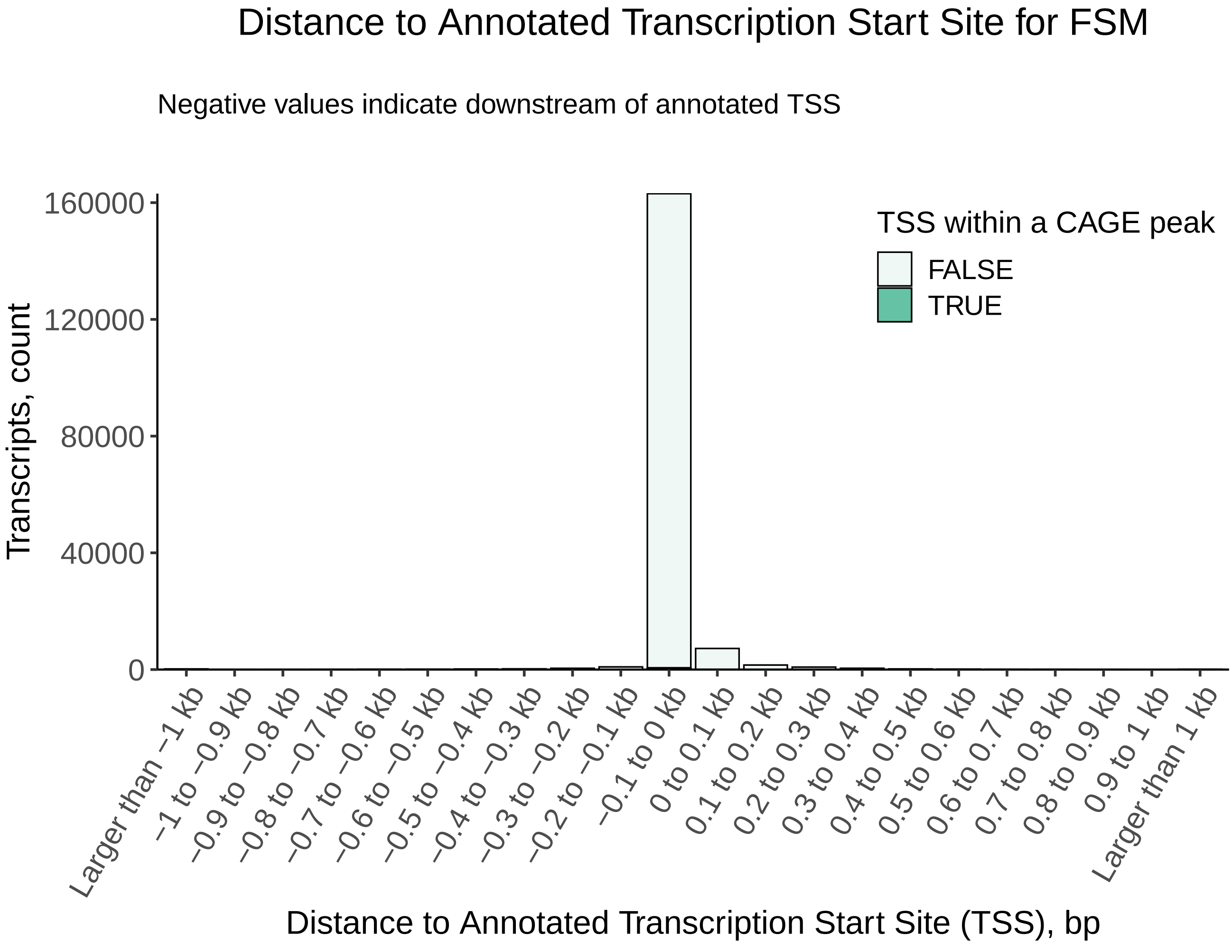


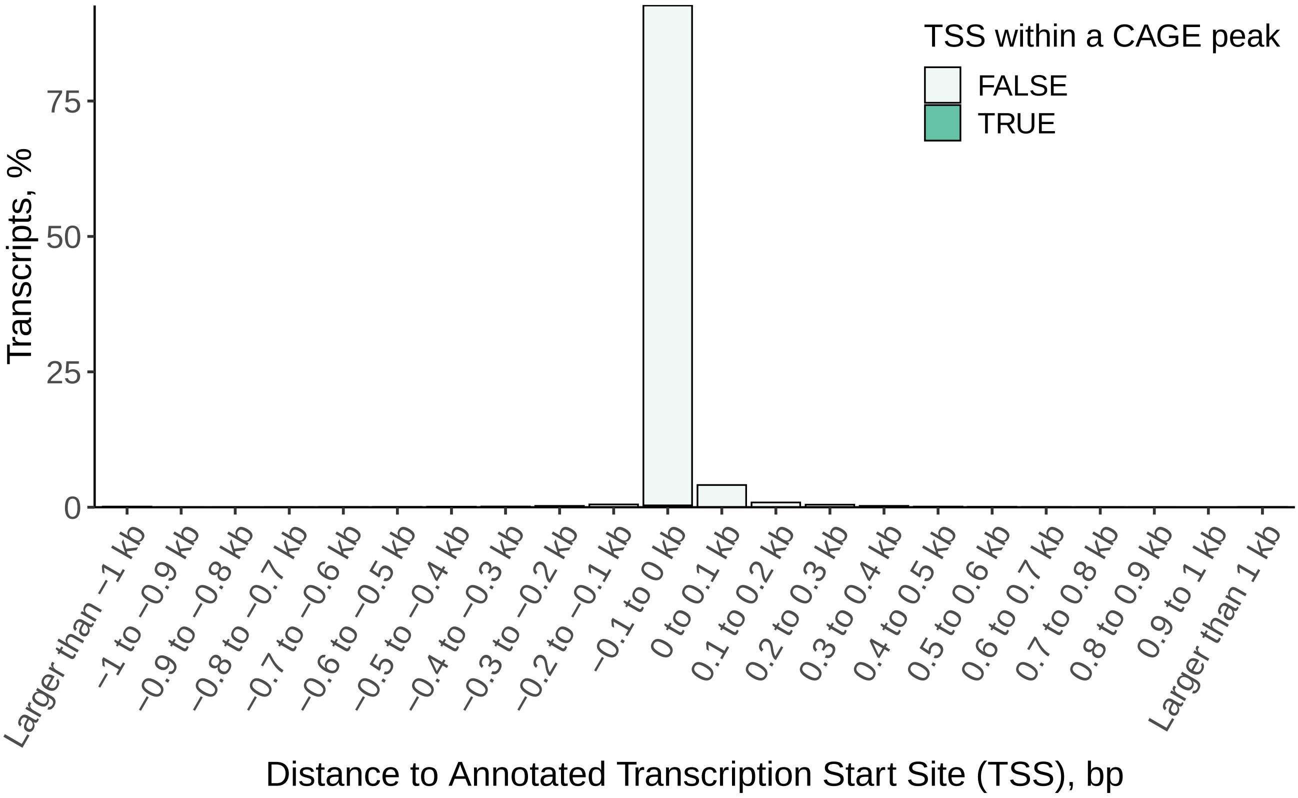
Distance to annotated Transcription Termination Site (TTS)
FSM

Negative values indicate upstream of annotated termination site


Distance to annotated Transcription Termination Site (TTS)
FSM

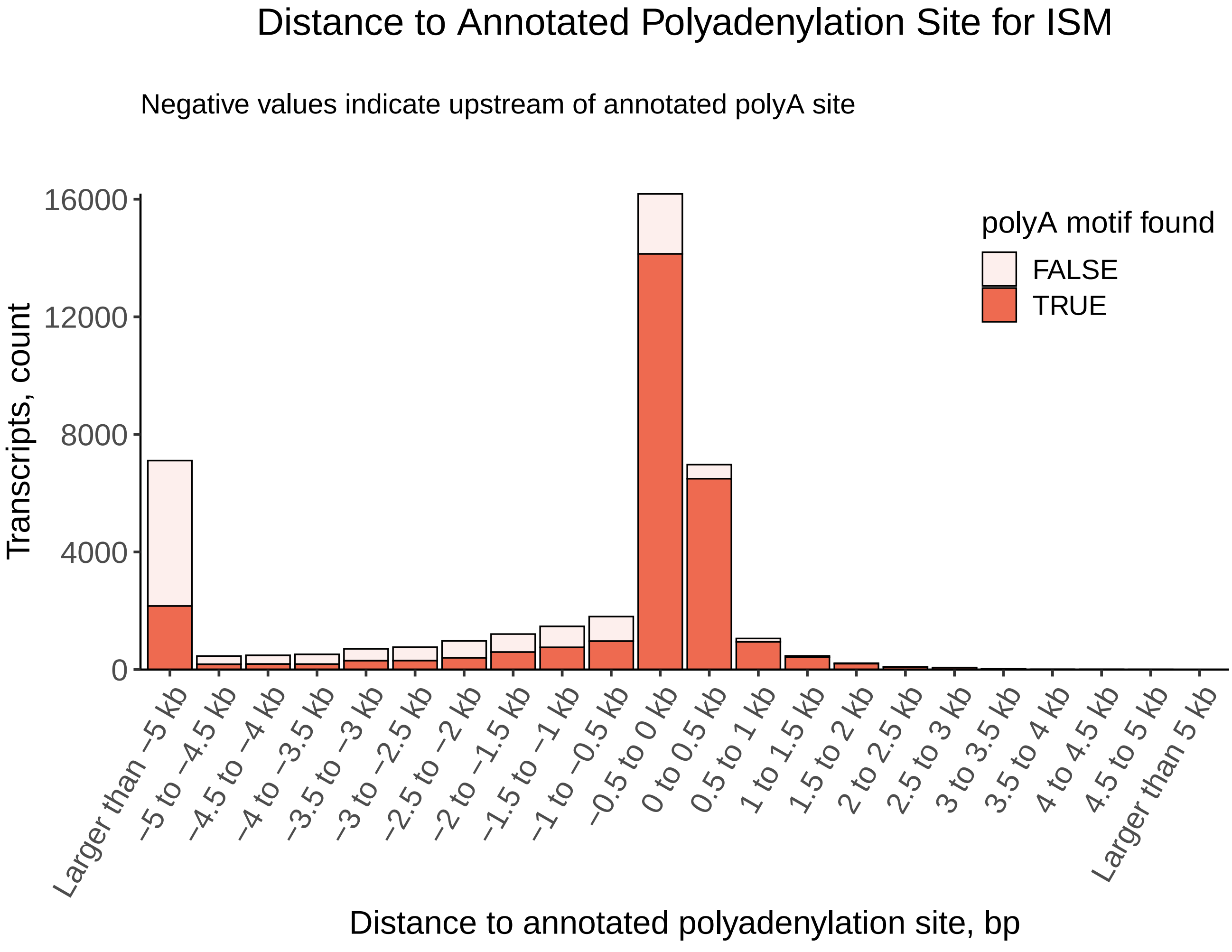
Negative values indicate upstream of annotated termination site

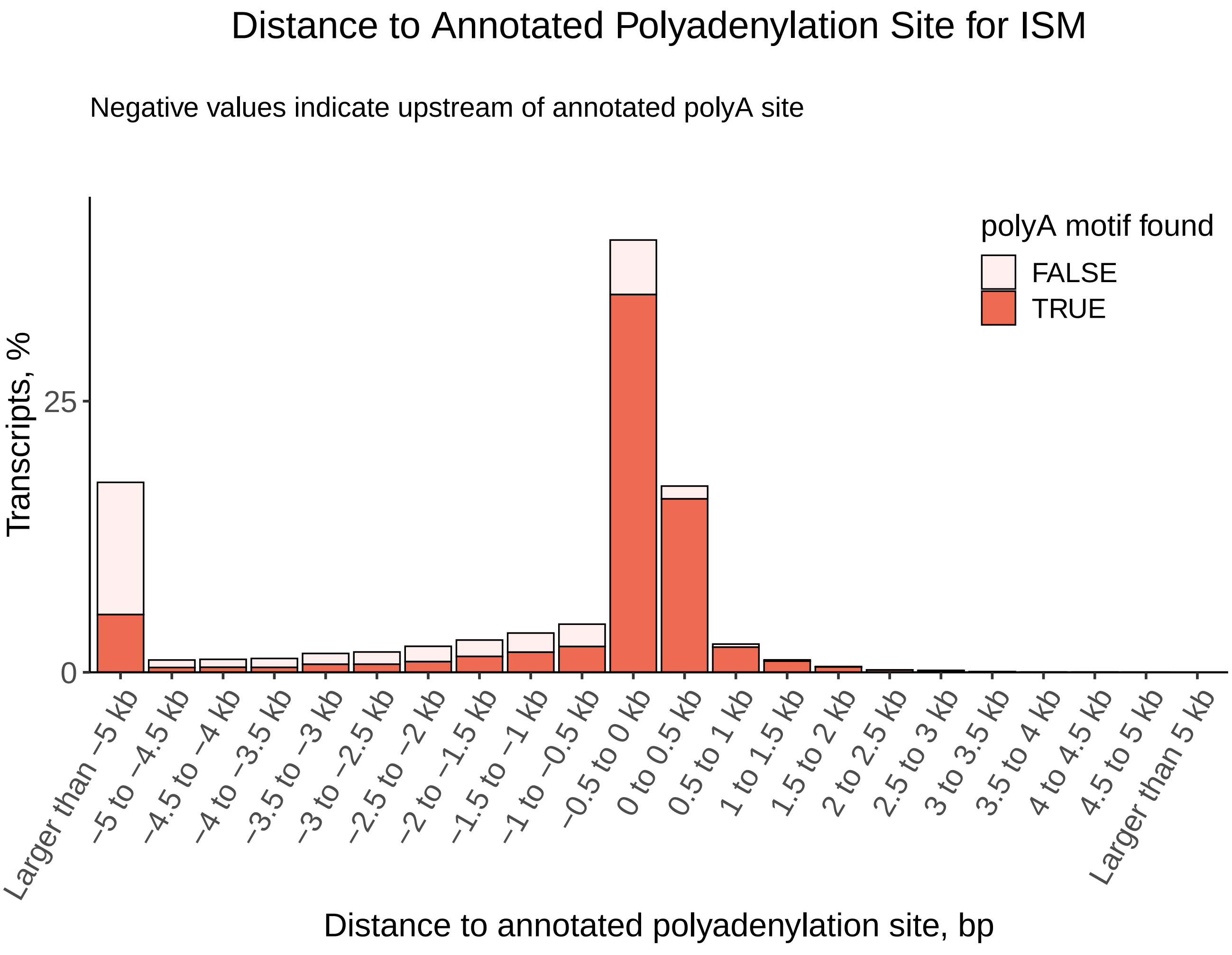

Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
Transcripts, count
160000
120000
80000
40000
0
TSS within a CAGE peak
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−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 for FSM

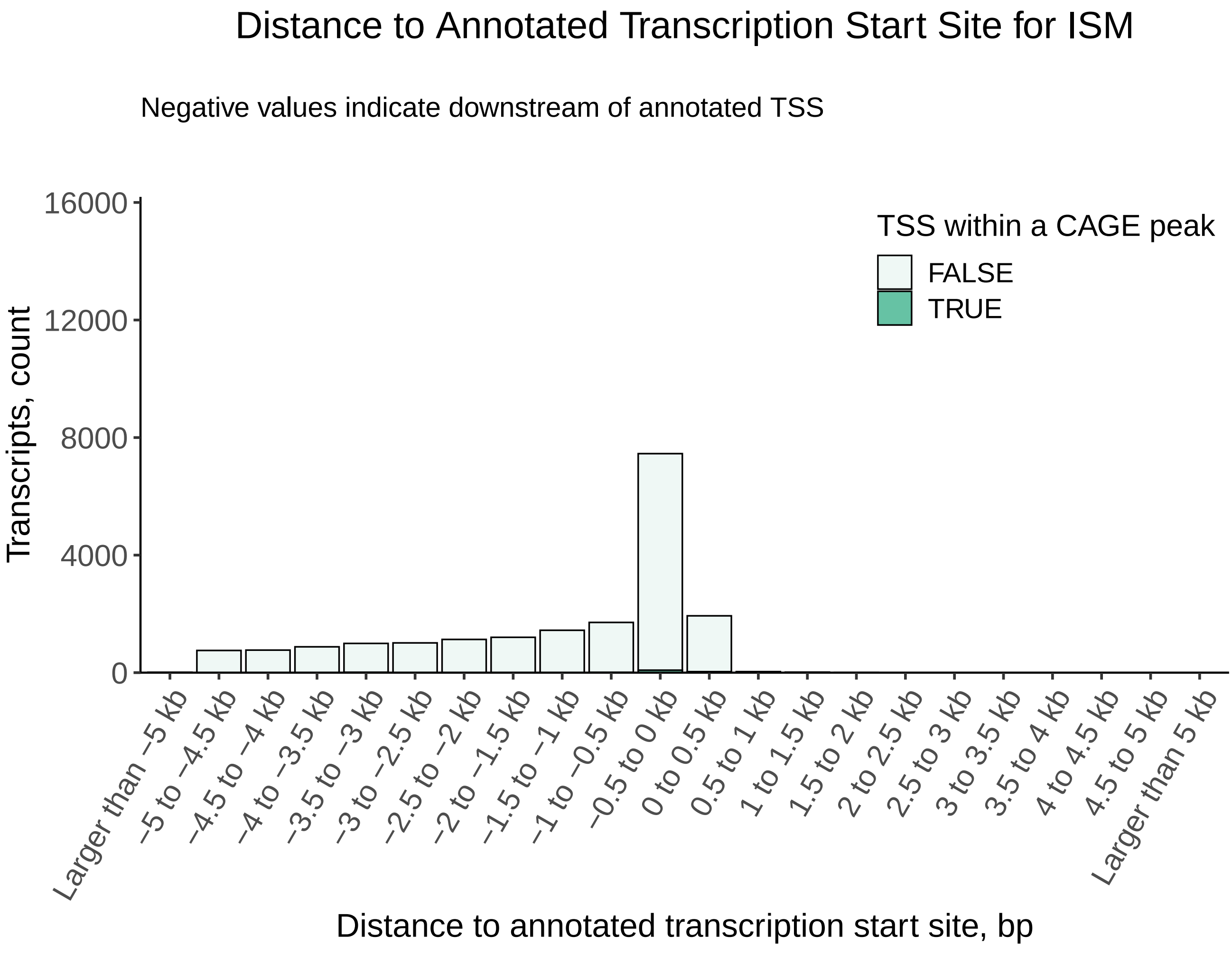
Negative values indicate downstream of annotated TSS


Distance to Annotated Polyadenylation Site for ISM
Negative values indicate upstream of annotated polyA site
Transcripts, count
0
4000
8000
12000
16000
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp



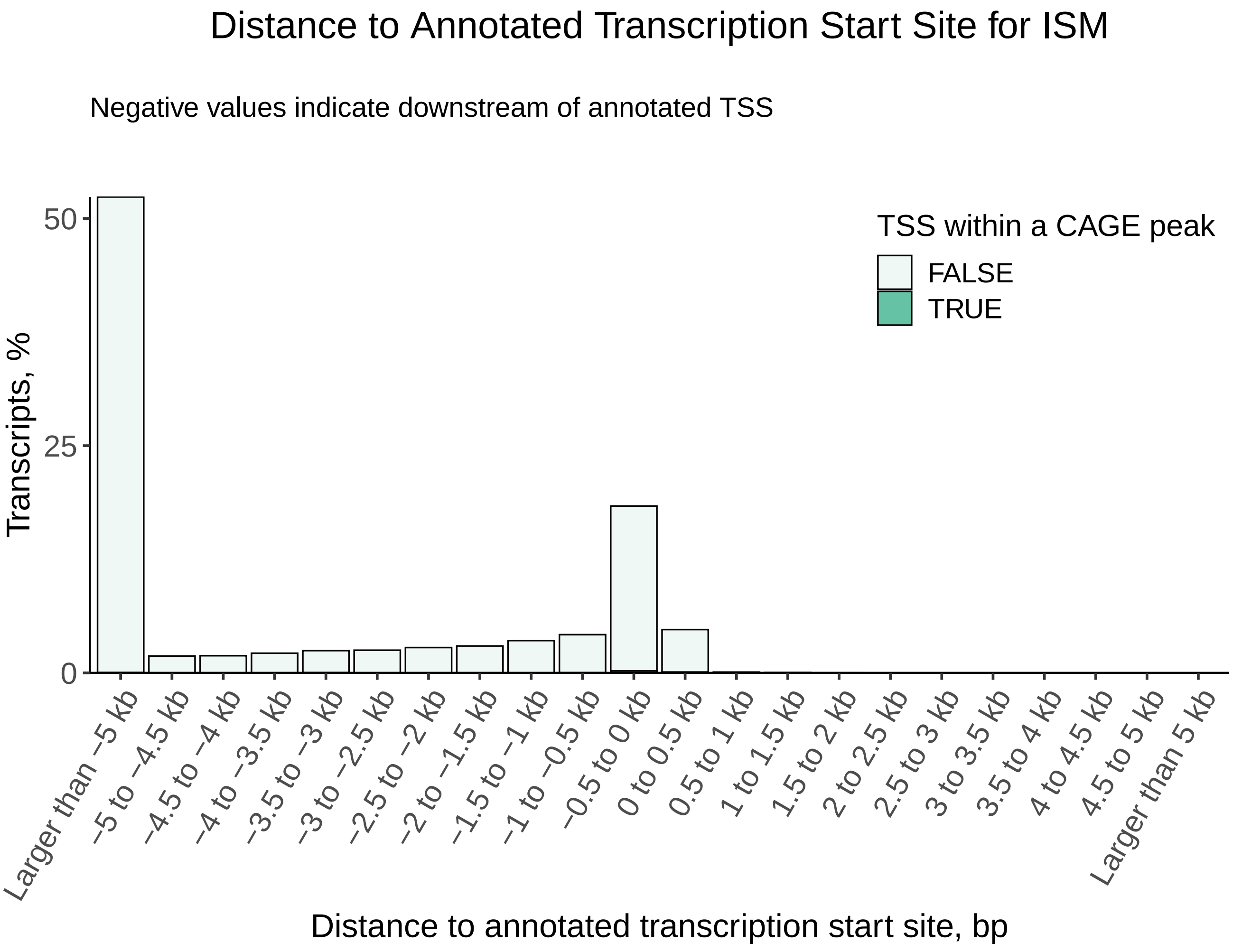
Distance to Annotated Polyadenylation Site for ISM

Negative values indicate upstream of annotated polyA site

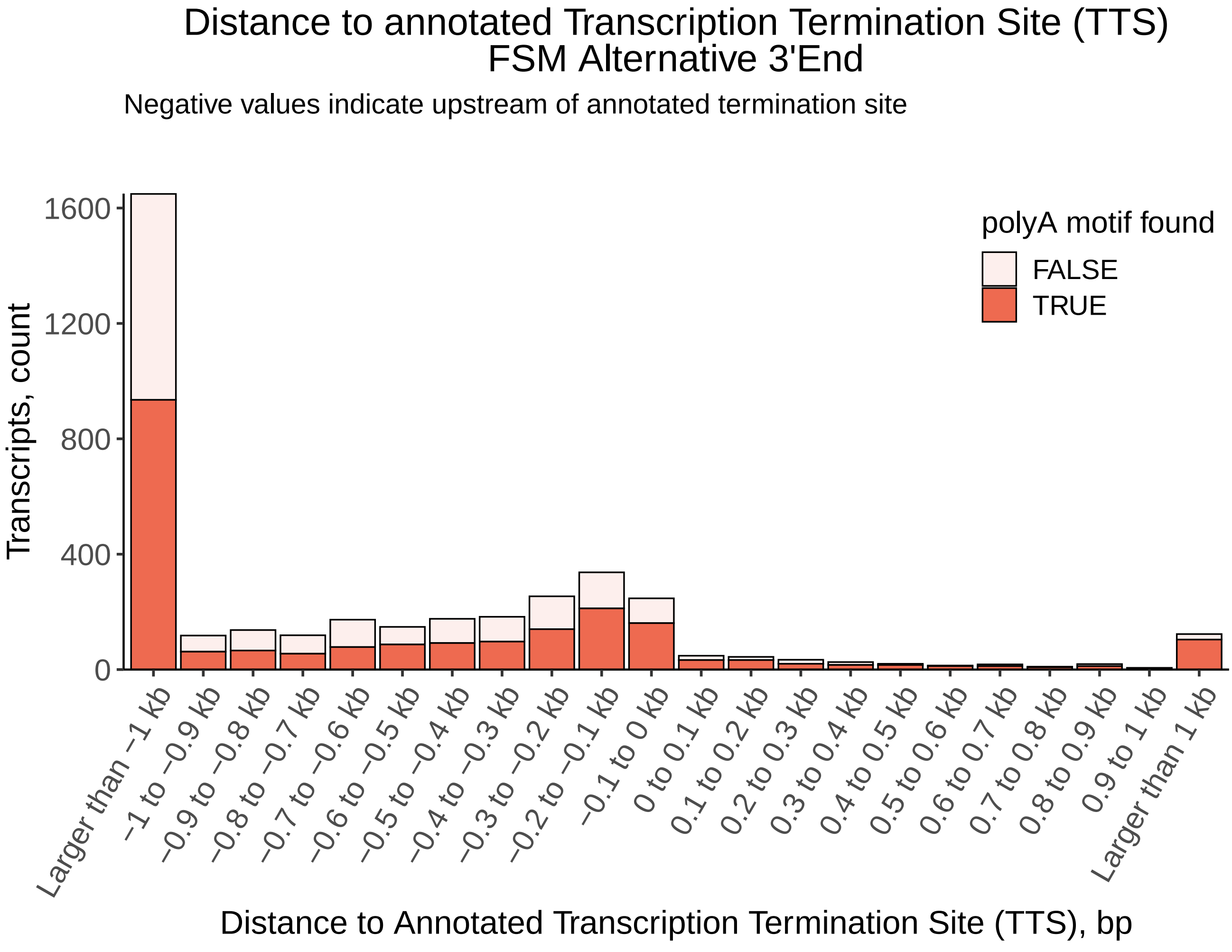


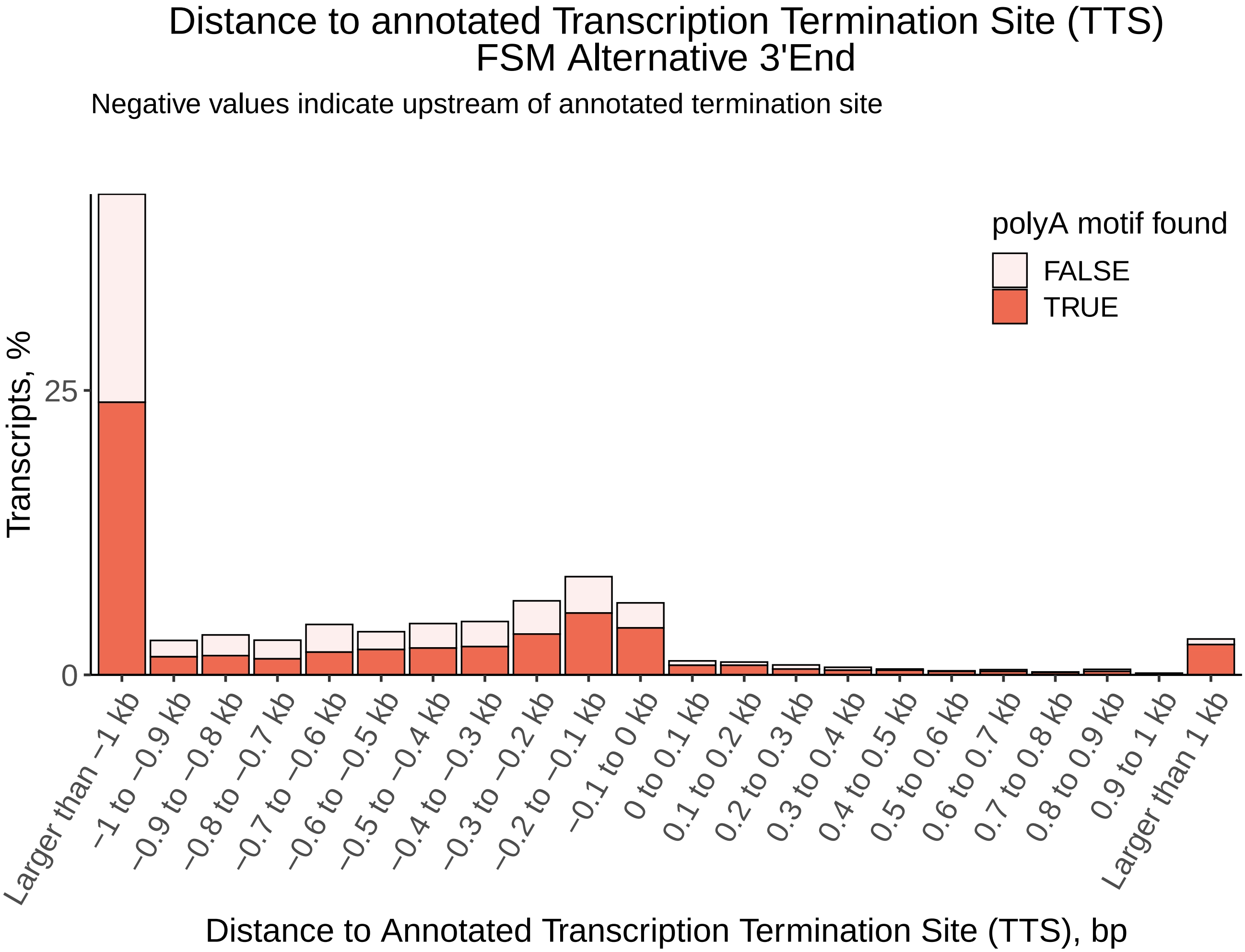
Distance to Annotated Transcription Start Site for ISM

Negative values indicate downstream of annotated TSS

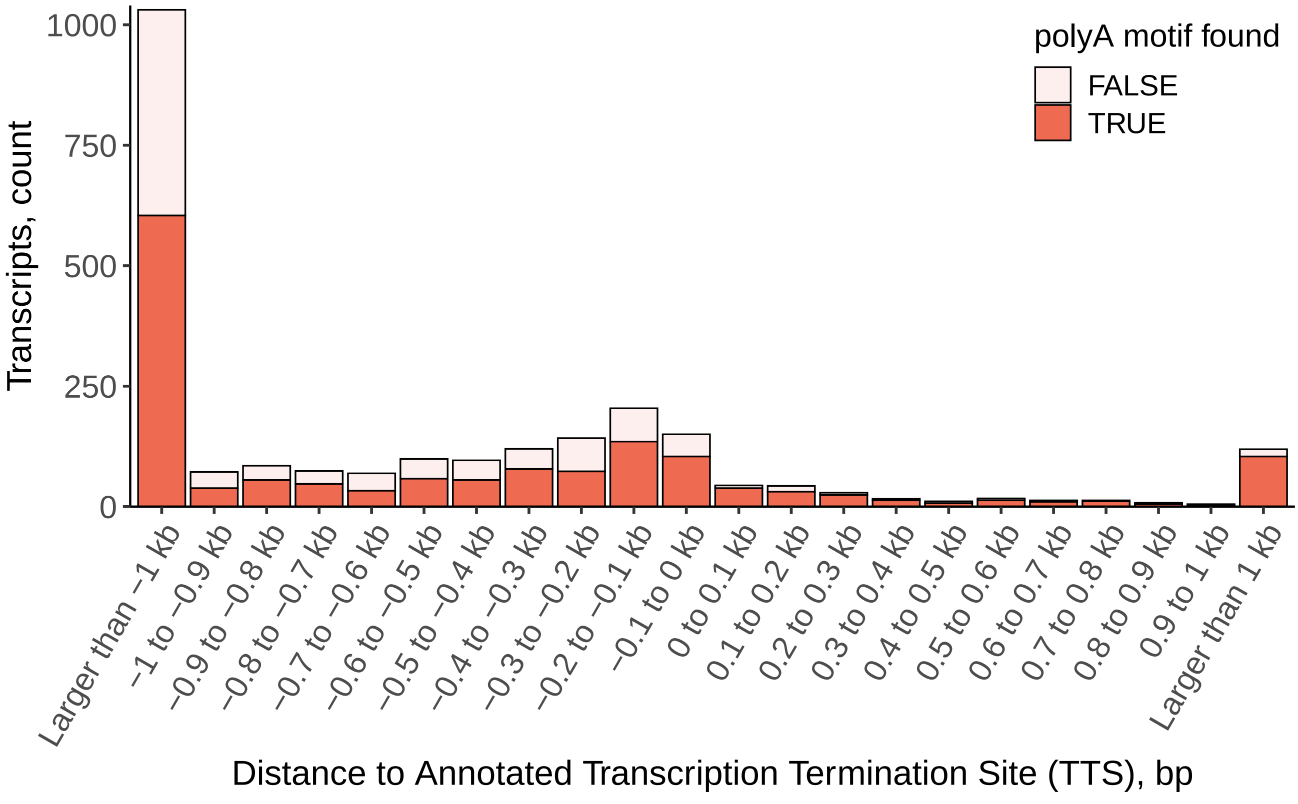

Distance to Annotated Transcription Start Site for ISM
Negative values indicate downstream of annotated TSS
Transcripts, %
50
25
0
TSS within a CAGE peak
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −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

Comparison With Annotated TSS and TTS by Subcategories


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
Transcripts, count
1600
1200
800
400
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to Annotated Transcription Termination Site (TTS), bp

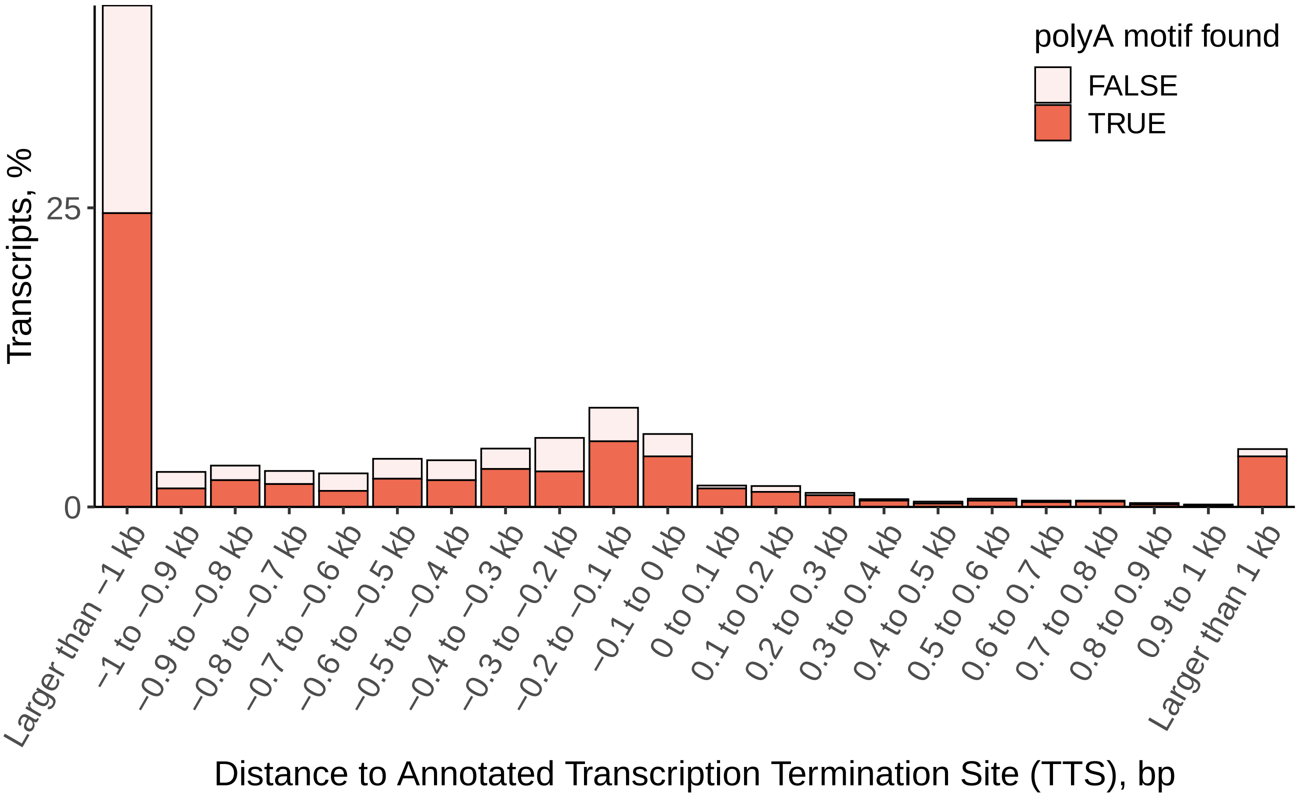


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

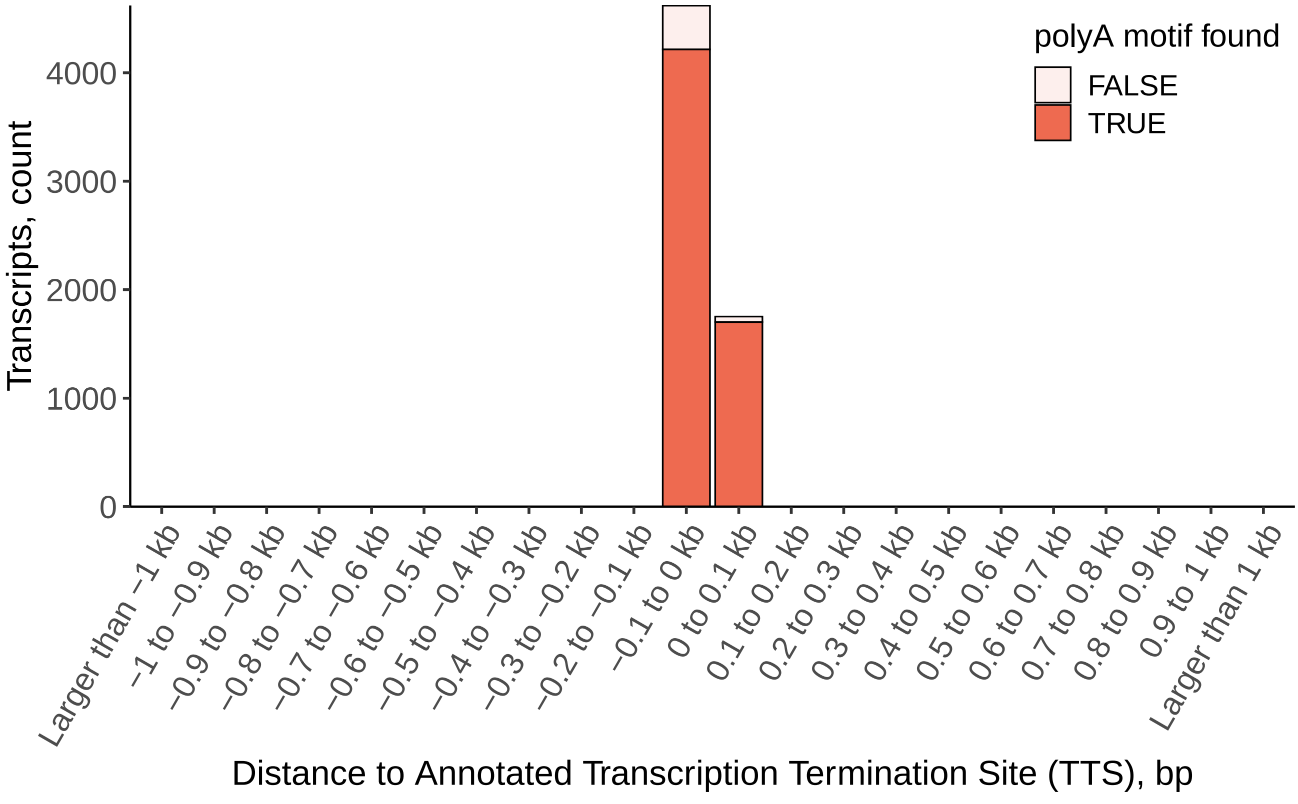


Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End

Negative values indicate upstream of annotated termination site

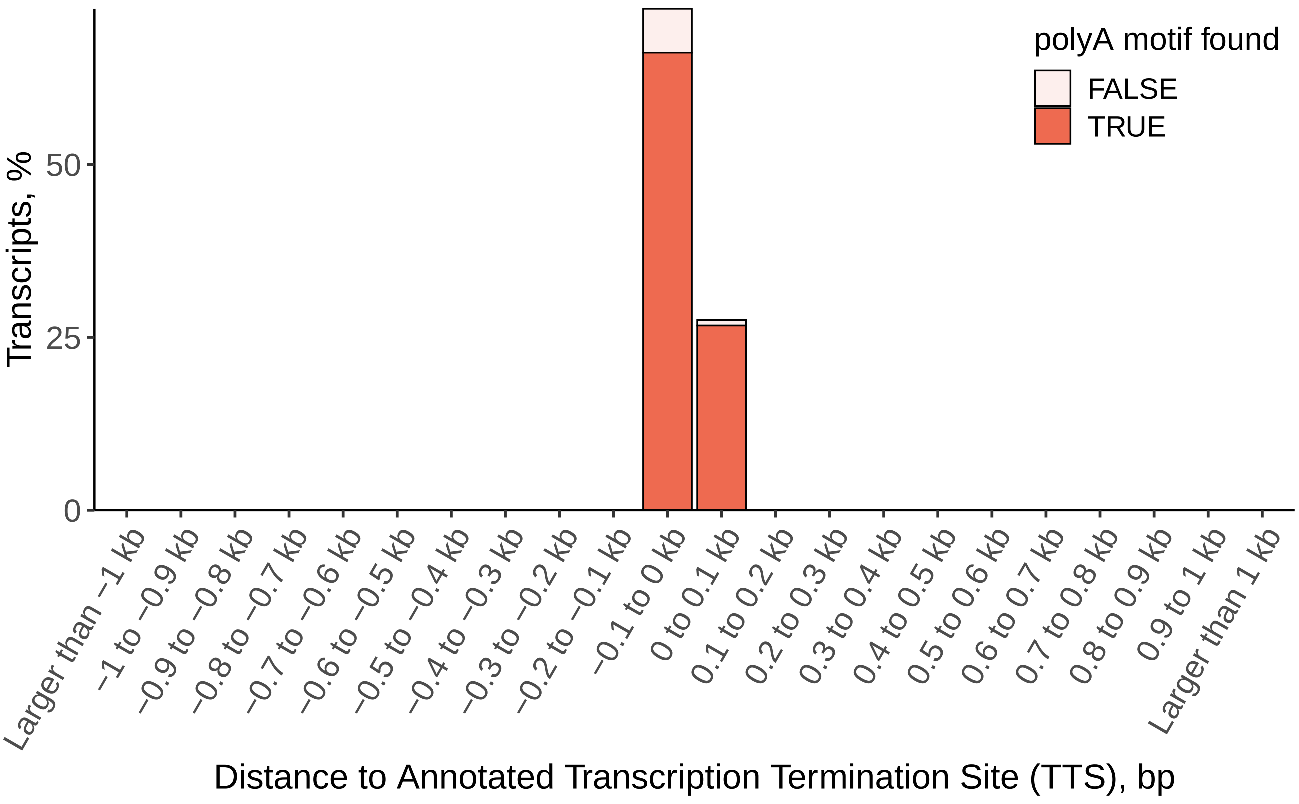
Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End

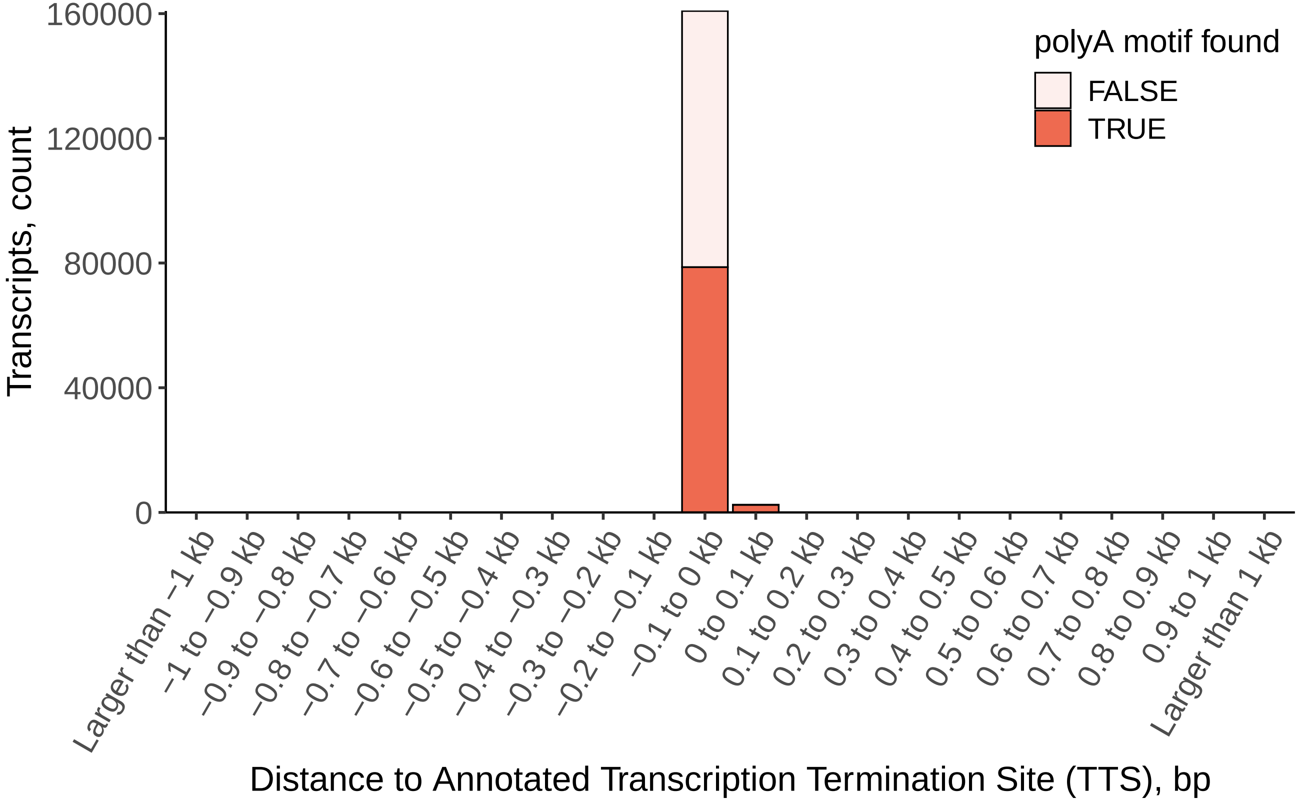
Negative values indicate upstream of annotated termination site



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

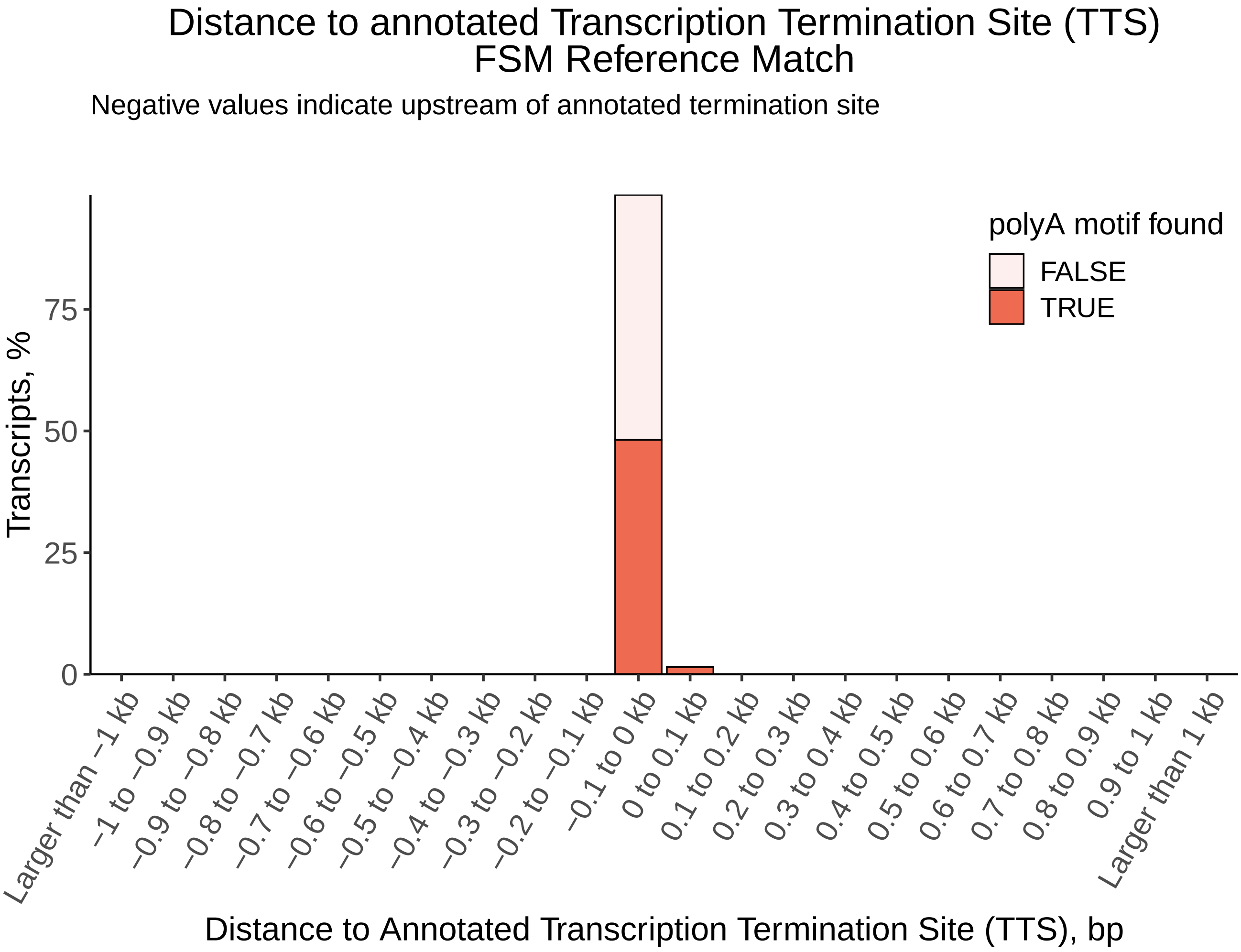
Negative values indicate upstream of annotated termination site

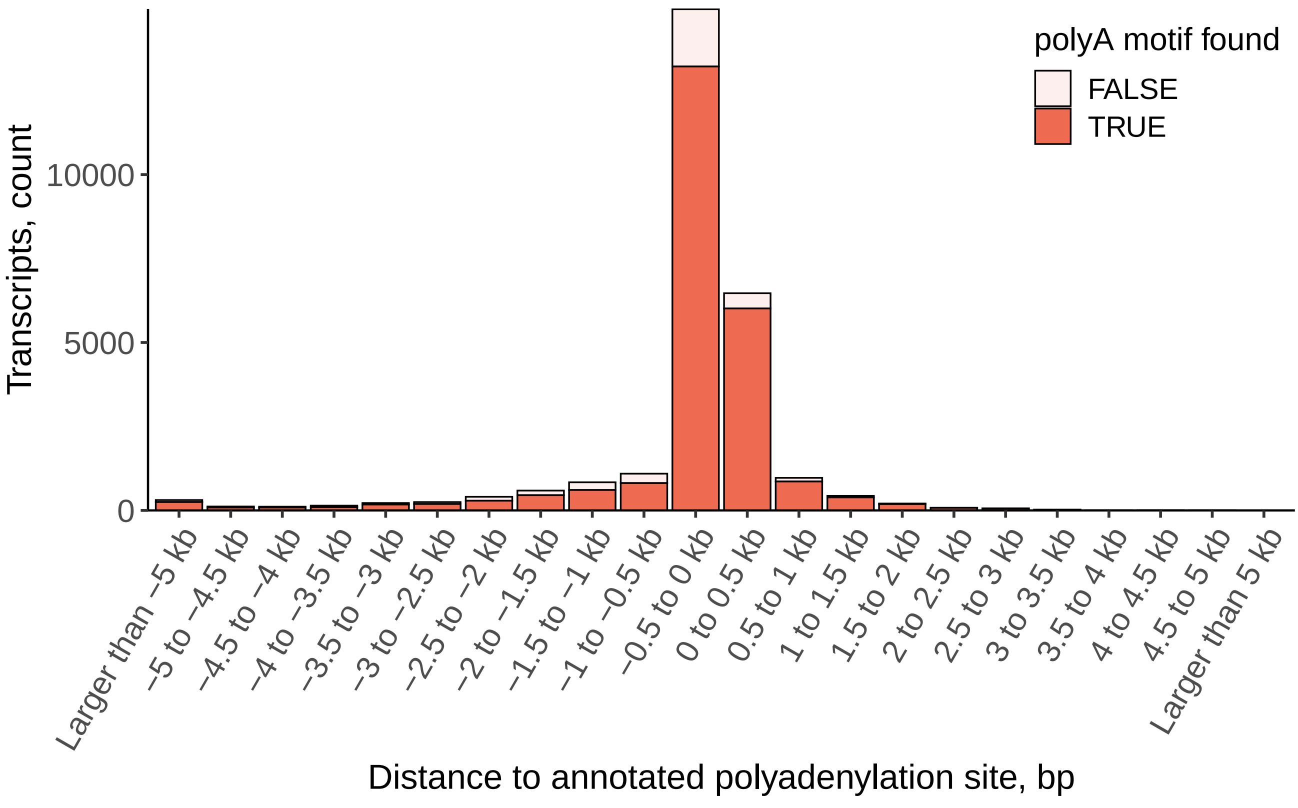




Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match

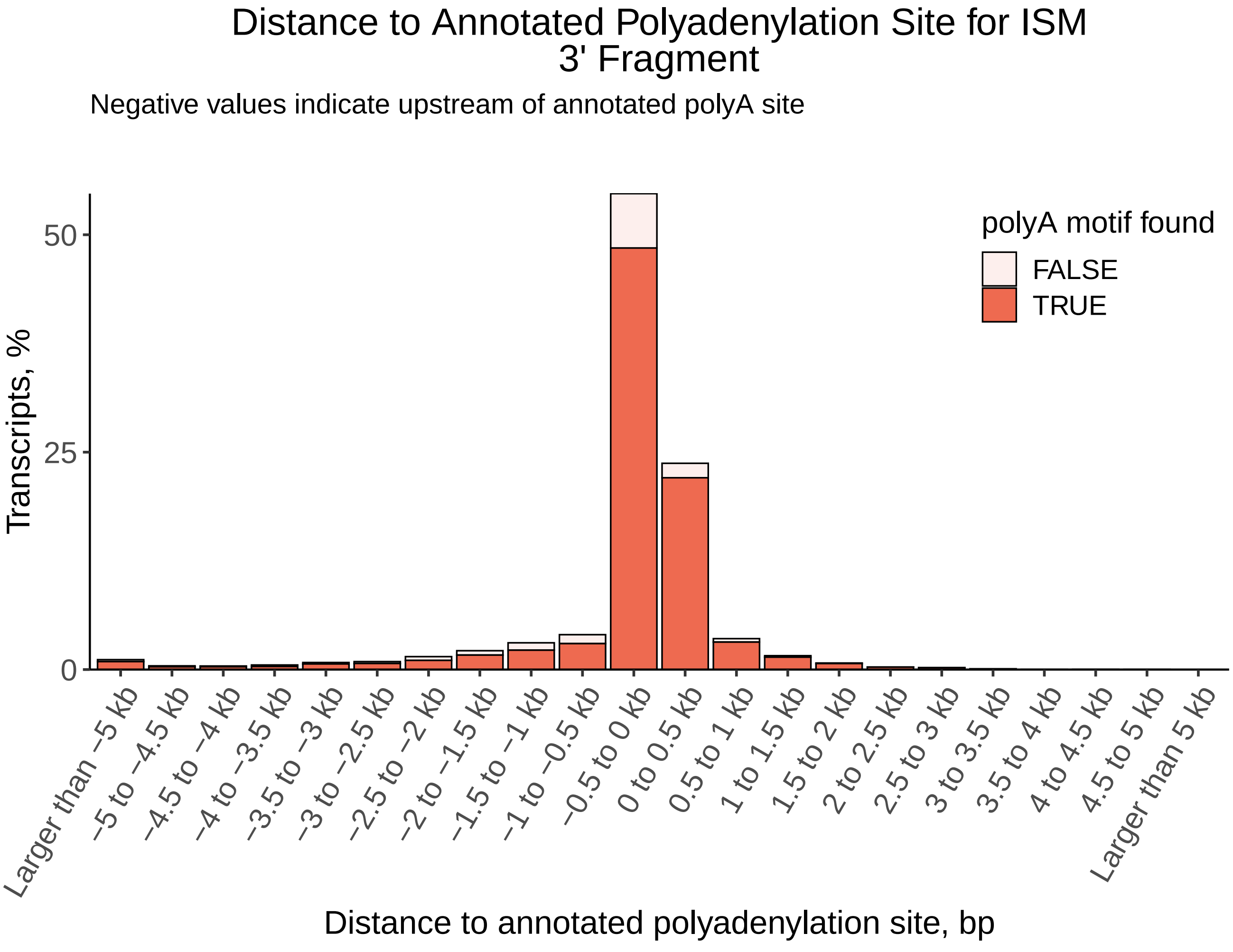
Negative values indicate upstream of annotated termination site


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



Distance to Annotated Polyadenylation Site for ISM
3' Fragment

Negative values indicate upstream of annotated polyA site

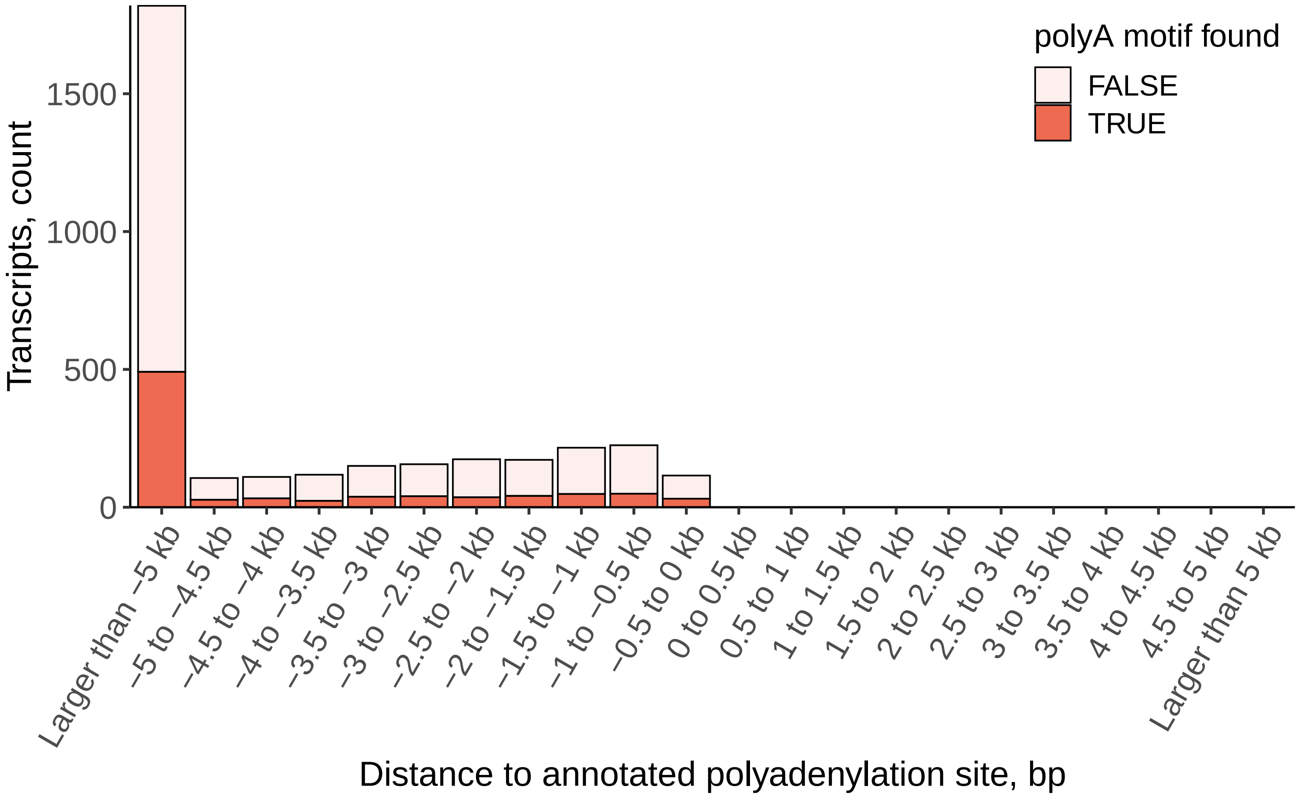


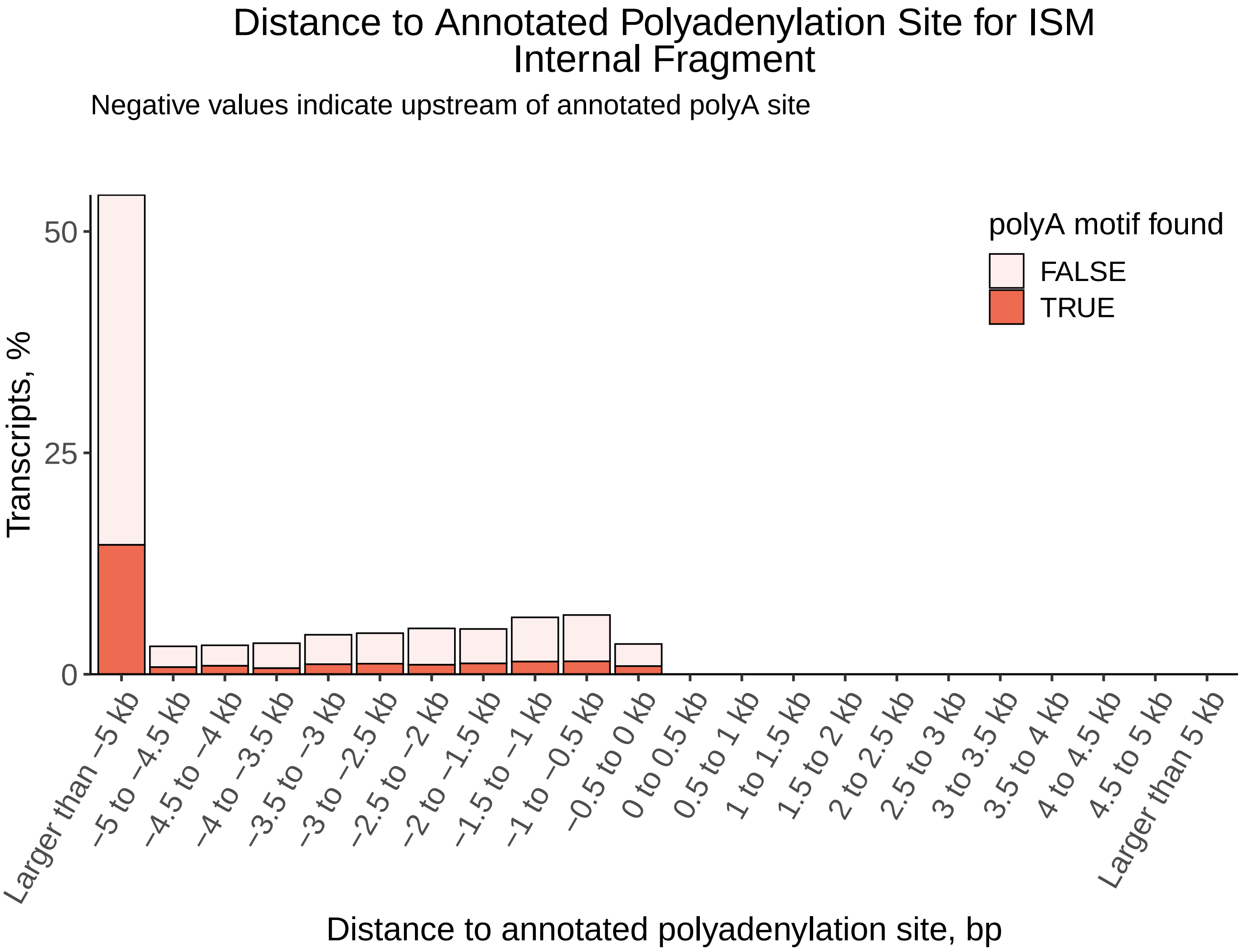
Distance to Annotated Polyadenylation Site for ISM 3' Fragment

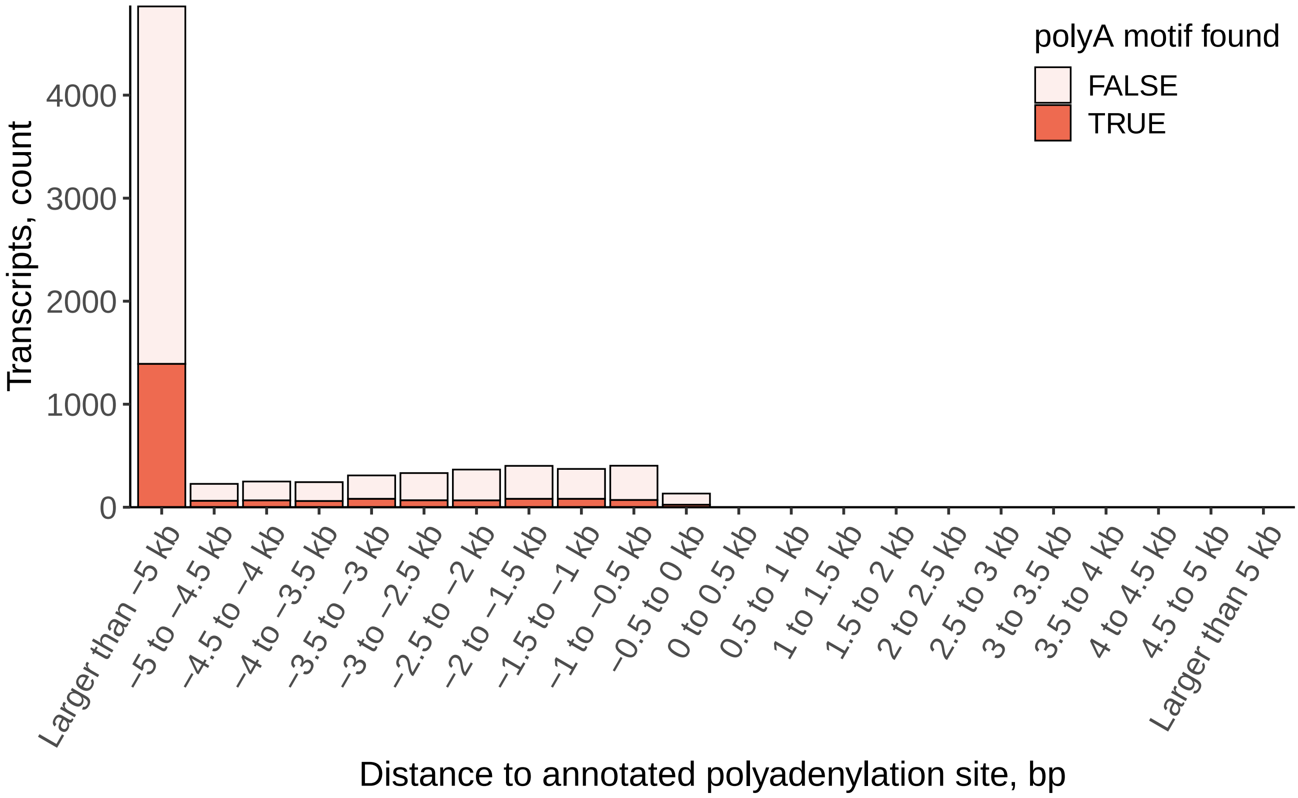
Negative values indicate upstream of annotated polyA site

Distance to Annotated Polyadenylation Site for ISM Internal Fragment

Negative values indicate upstream of annotated polyA site

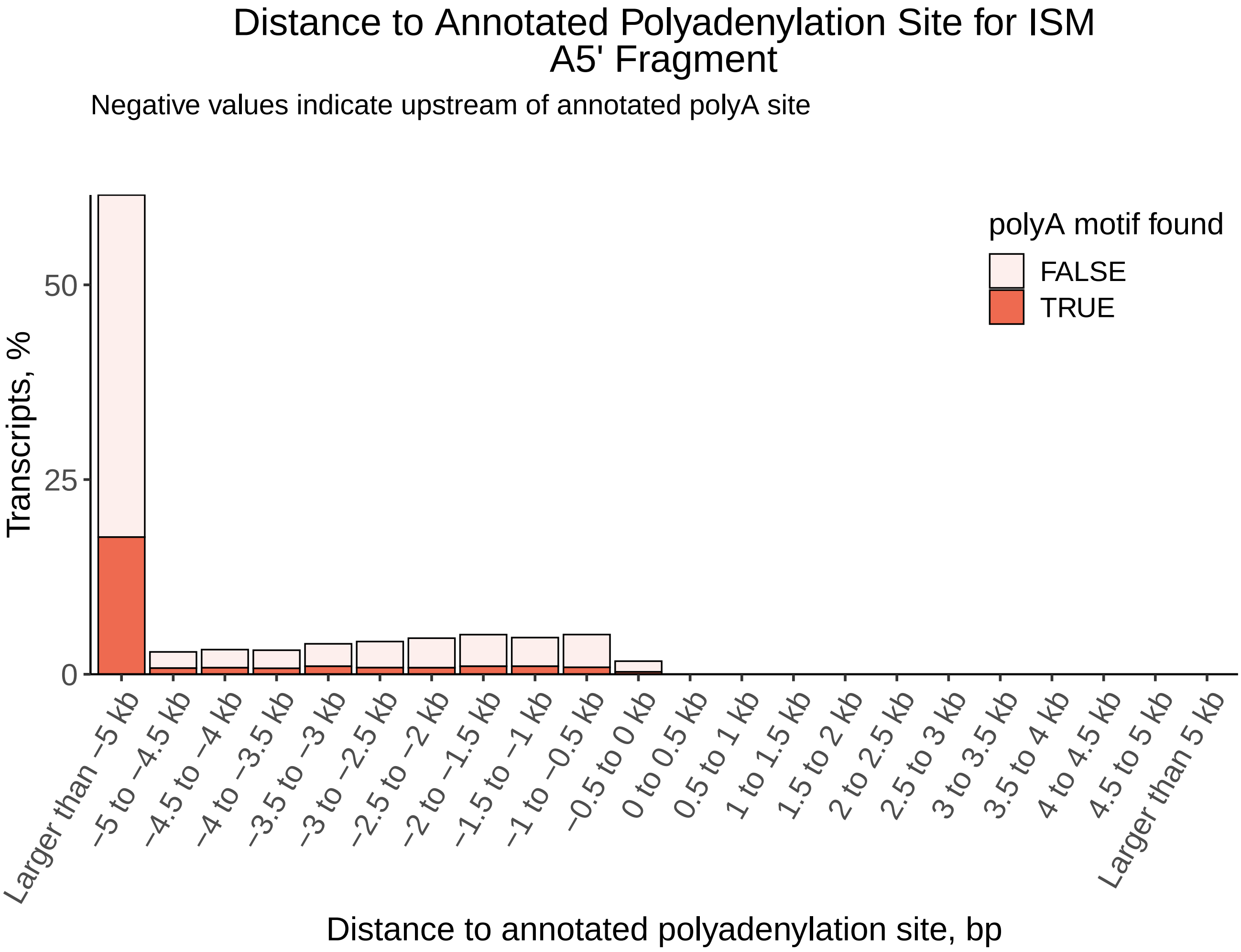



Distance to Annotated Polyadenylation Site for ISM
Internal Fragment
Negative values indicate upstream of annotated polyA site
Transcripts, %
50
25
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to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



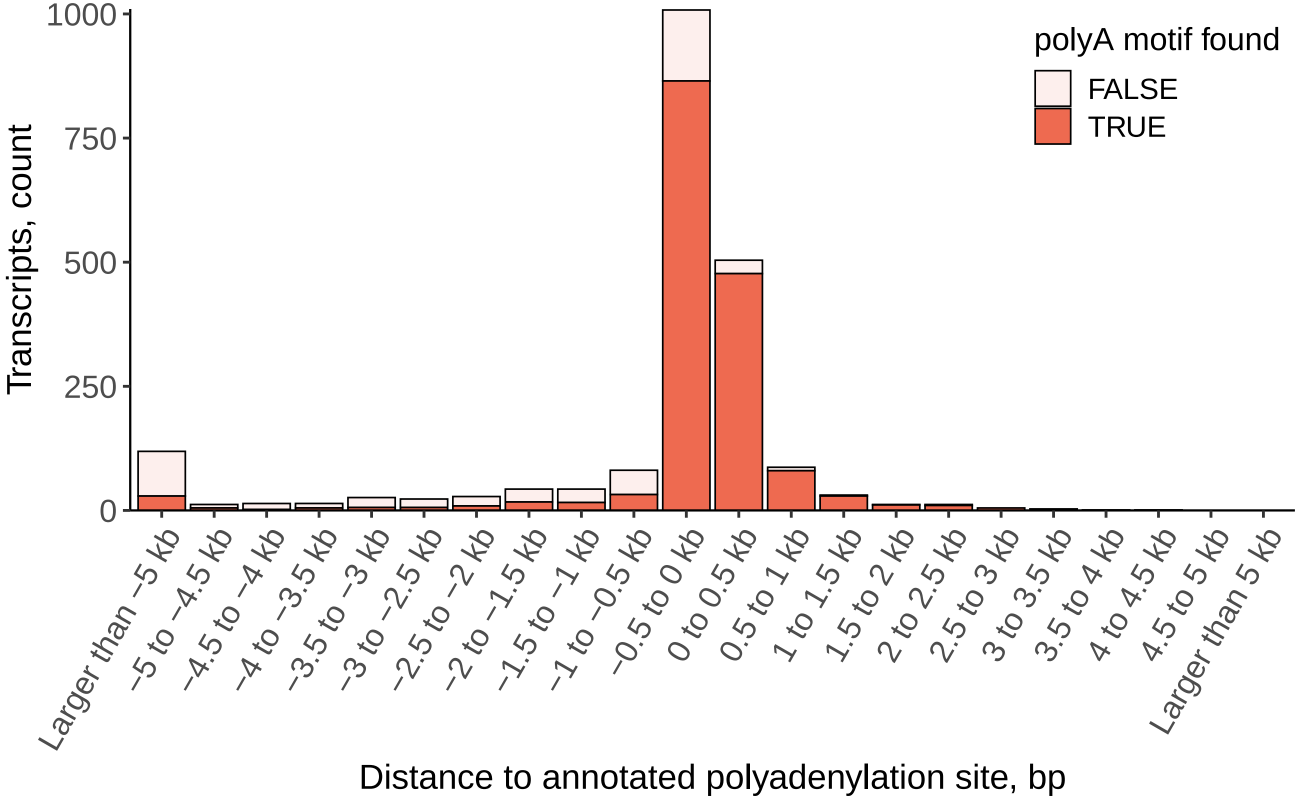
Distance to Annotated Polyadenylation Site for ISM A5' Fragment

Negative values indicate upstream of annotated polyA site



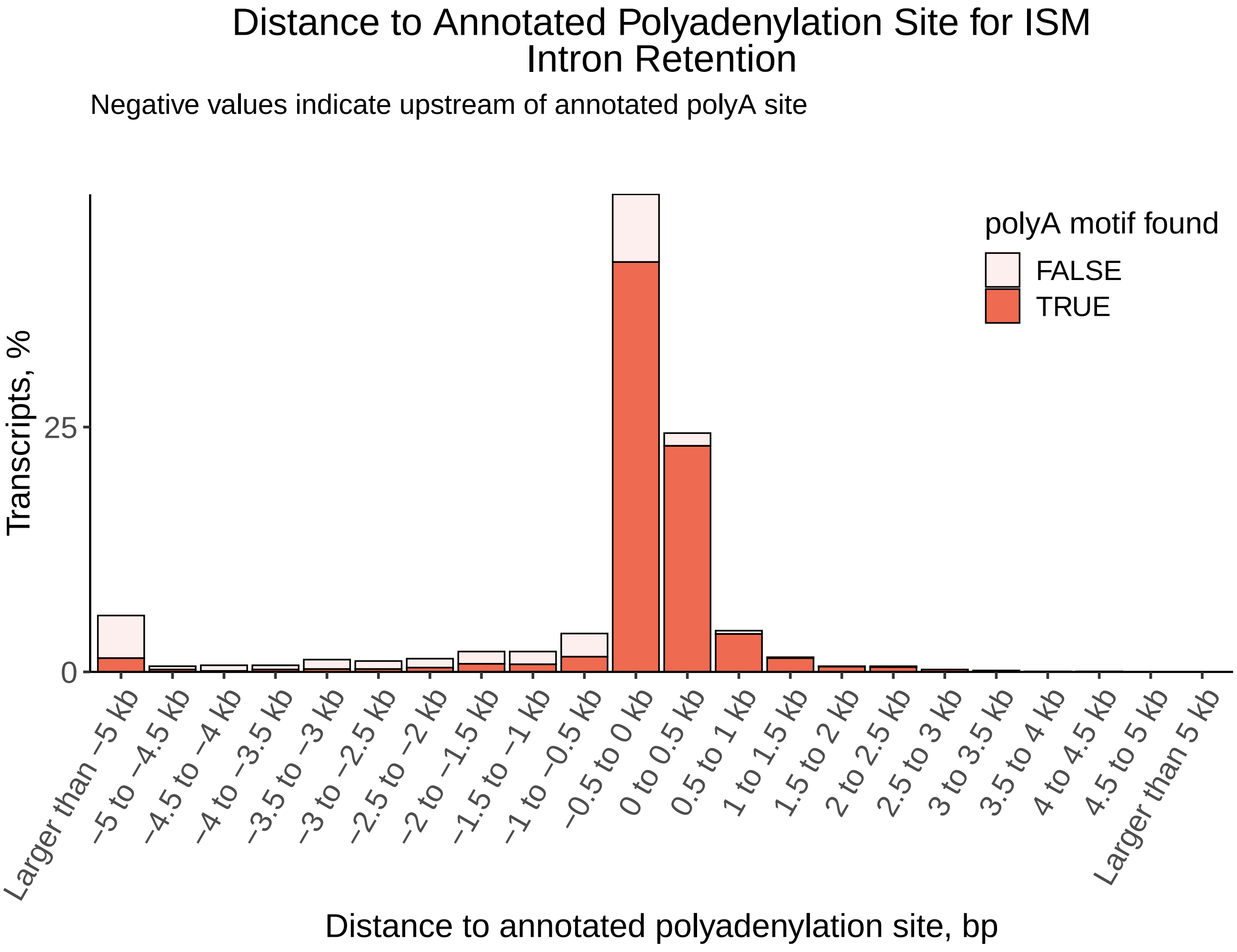
Distance to Annotated Polyadenylation Site for ISM A5' Fragment

Negative values indicate upstream of annotated polyA site



Distance to Annotated Polyadenylation Site for ISM Intron Retention

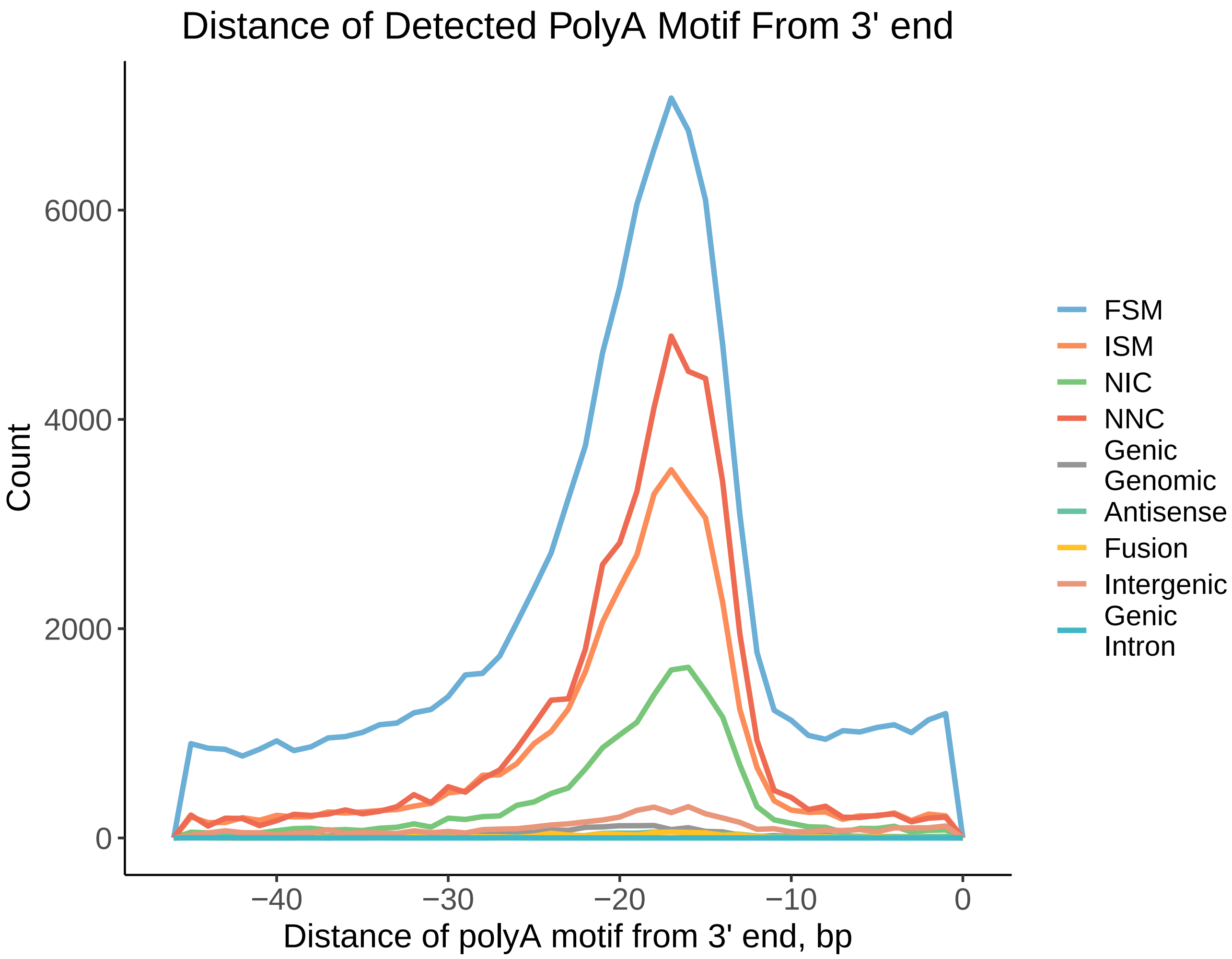
Negative values indicate upstream of annotated polyA site



Distance to Annotated Polyadenylation Site for ISM Intron Retention

Negative values indicate upstream of annotated polyA site

PolyA Distance Analysis

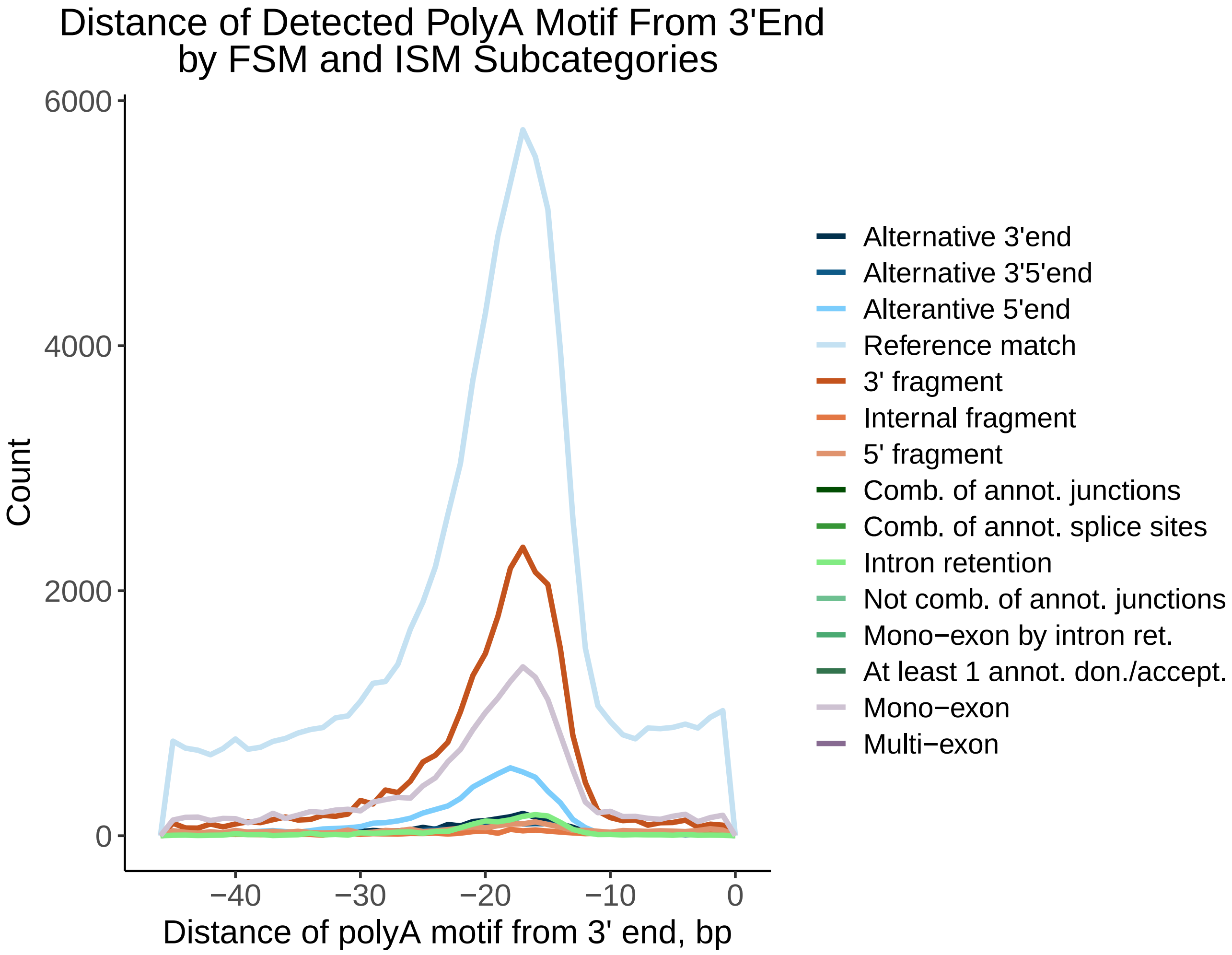

Distance of Detected PolyA Motif From 3' end
Count
6000
4000
2000
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

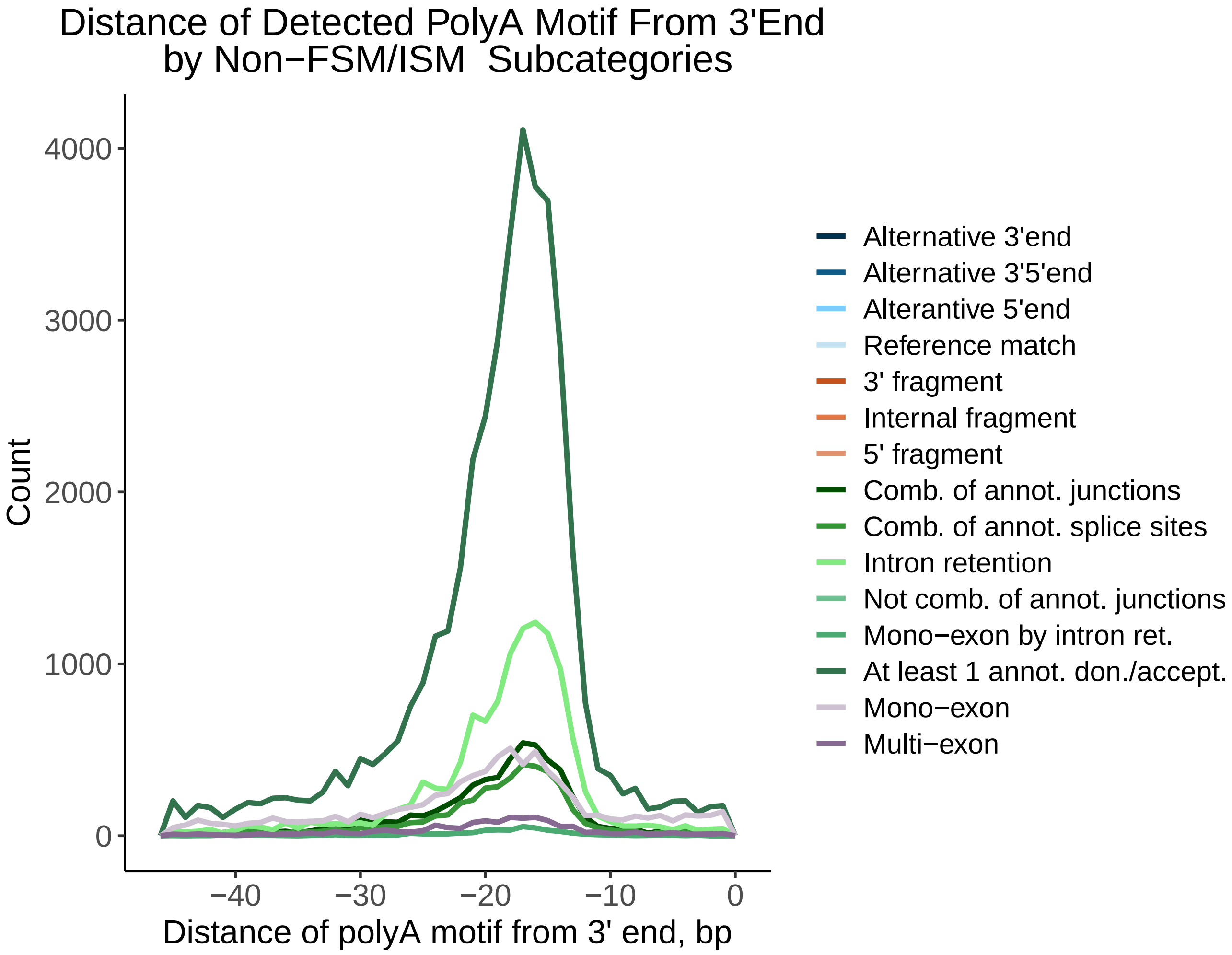
Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	193864	98601	51
ISM	52552	37941	72
NIC	19537	16353	84
NNC	55277	47586	86
Genic Genomic	2321	1698	73
Antisense	1562	869	56
Fusion	822	646	79
Intergenic	10422	4644	45
Genic Intron	88	6	7

Frequency of PolyA Motifs

Motif	Count	%
AATAAA	122788	58.9
ATTAAA	30394	14.6
AGTAAA	6689	3.2
TATAAA	6498	3.1
AAGAAA	5424	2.6
AAAAAG	5035	2.4
TTTAAA	4080	2.0
CATAAA	4055	1.9
AATACA	3848	1.8
AATATA	3394	1.6
GATAAA	3352	1.6
AATGAA	3150	1.5
AAAACA	2910	1.4
GGGGCT	2650	1.3
ACTAAA	2166	1.0
AATAGA	1911	0.9


Distance of Detected PolyA Motif From 3'End
by FSM and ISM Subcategories
Count
6000
4000
2000
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon


Distance of Detected PolyA Motif From 3'End
by Non−FSM/ISM Subcategories
Count
4000
3000
2000
1000
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
Alternative 3'end
Alternative 3'5'end
Alterantive 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Not comb. of annot. junctions
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

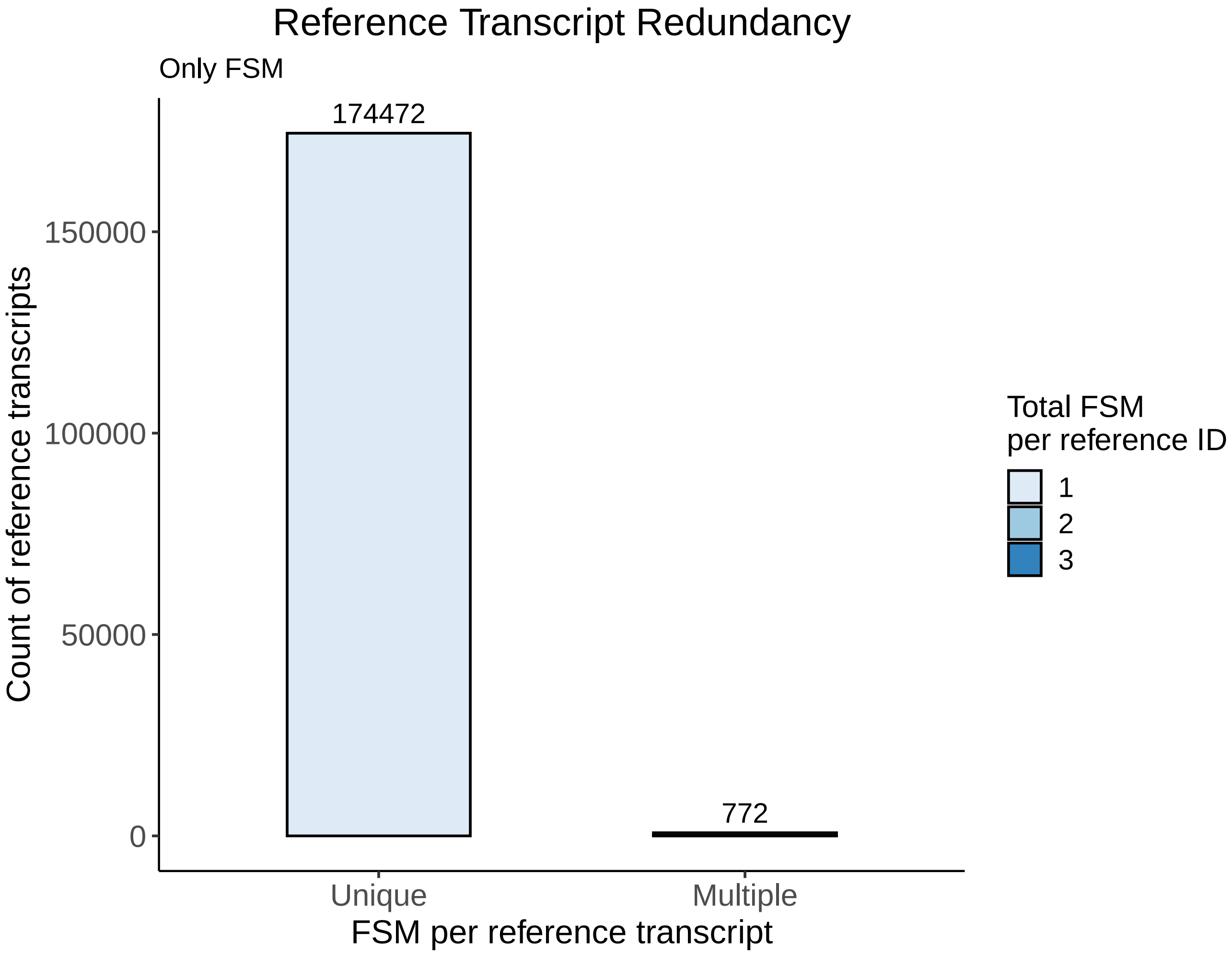
Number of polyA Motifs Detected

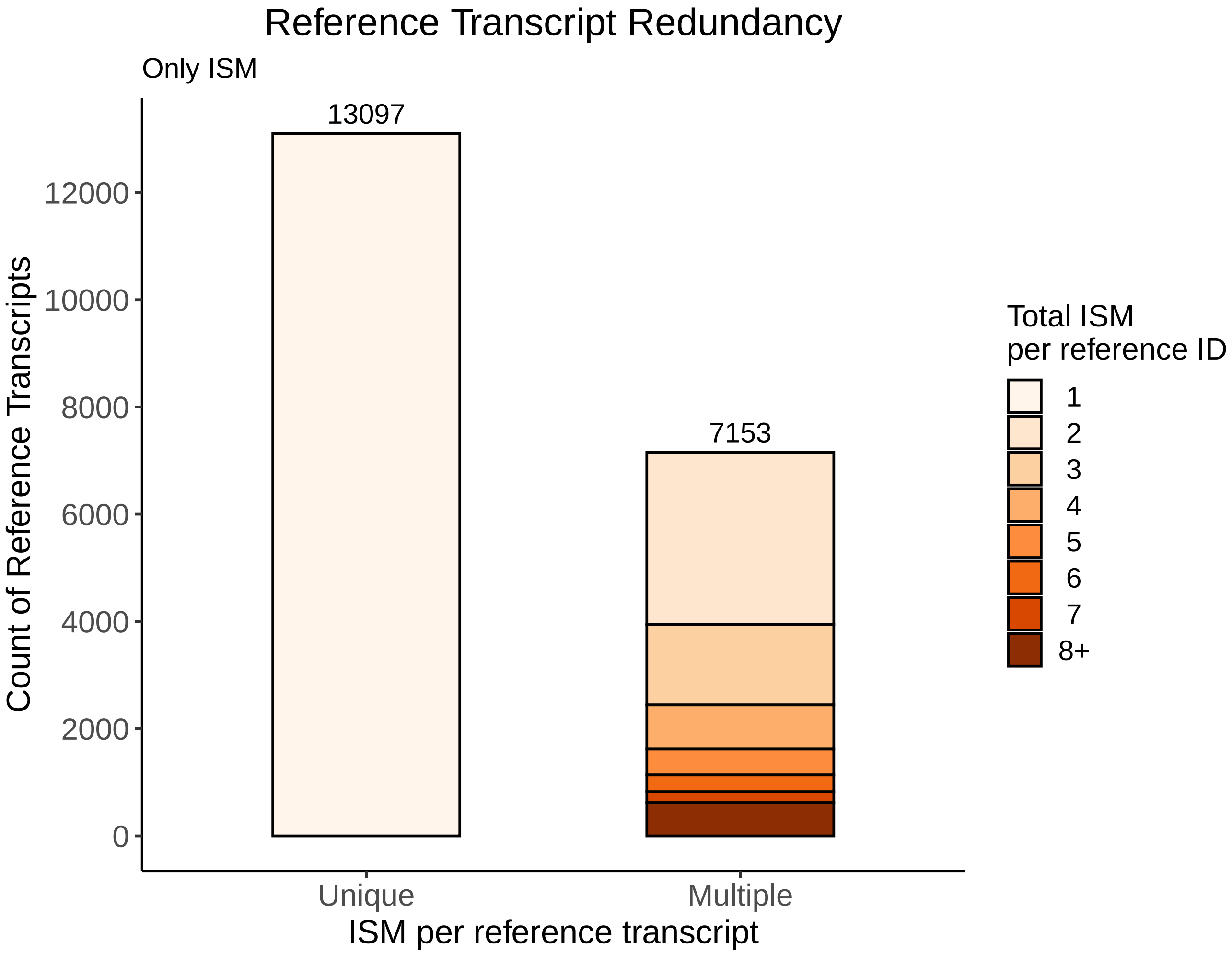
Subcategory	Count	polyA Detected	%
Alternative 3'end	3903	2257	58
Alternative 3'5'end	2460	1539	63
Alterantive 5'end	6371	5916	93
Reference match	163319	81053	50
3' fragment	27269	23895	88
Internal fragment	3361	856	25
5' fragment	7899	2049	26
Comb. of annot. junctions	6253	5488	88
Comb. of annot. splice sites	4882	4174	85
Intron retention	15722	13446	86
Mono−exon by intron ret.	608	431	71
At least 1 annot. don./accept.	47573	40834	86
Mono−exon	44812	25042	56
Multi−exon	2013	1364	68

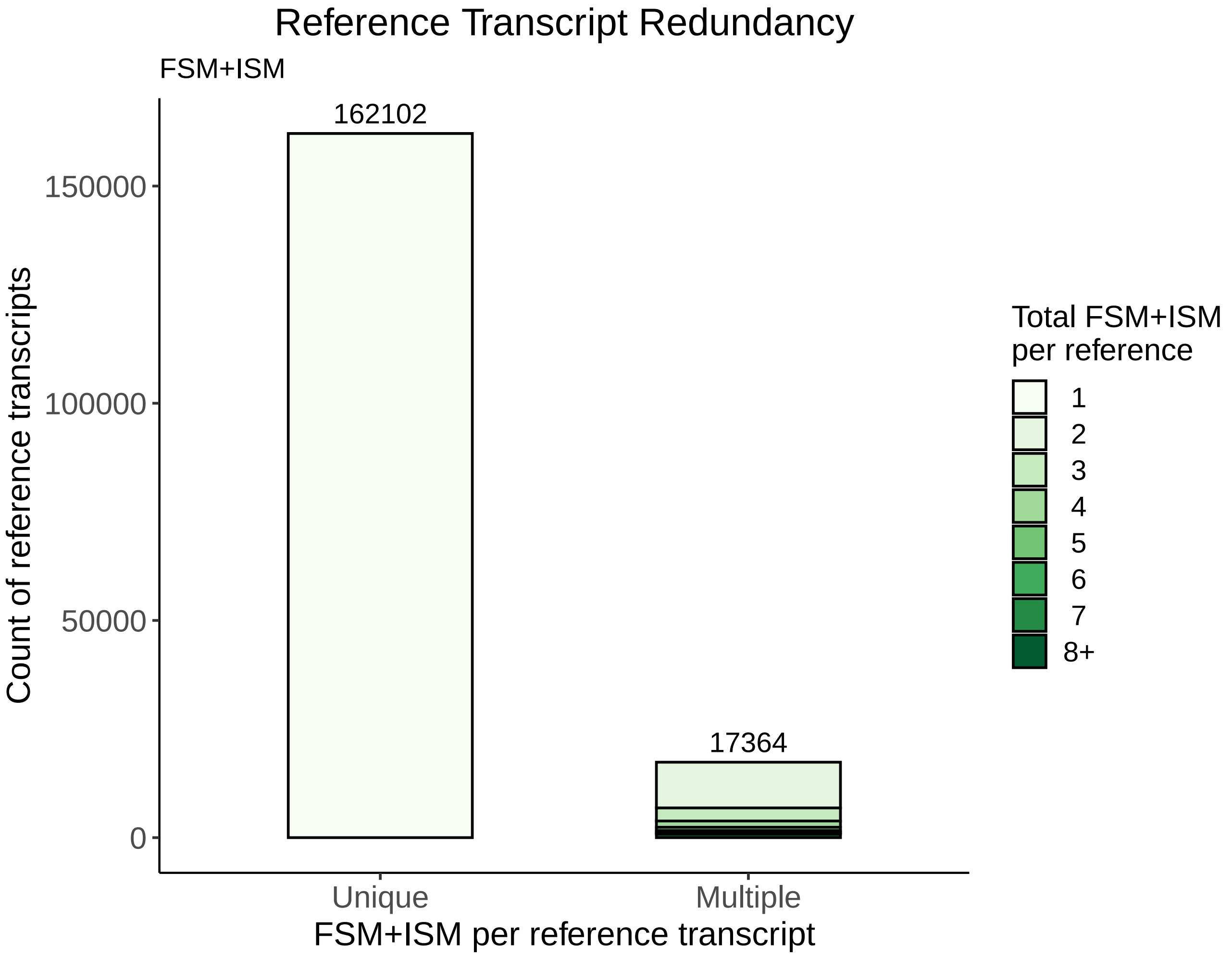
Frequency of PolyA Motifs

Motif	Count	%
AATAAA	122788	58.9
ATTAAA	30394	14.6
AGTAAA	6689	3.2
TATAAA	6498	3.1
AAGAAA	5424	2.6
AAAAAG	5035	2.4
TTTAAA	4080	2.0
CATAAA	4055	1.9
AATACA	3848	1.8
AATATA	3394	1.6
GATAAA	3352	1.6
AATGAA	3150	1.5
AAAACA	2910	1.4
GGGGCT	2650	1.3
ACTAAA	2166	1.0
AATAGA	1911	0.9

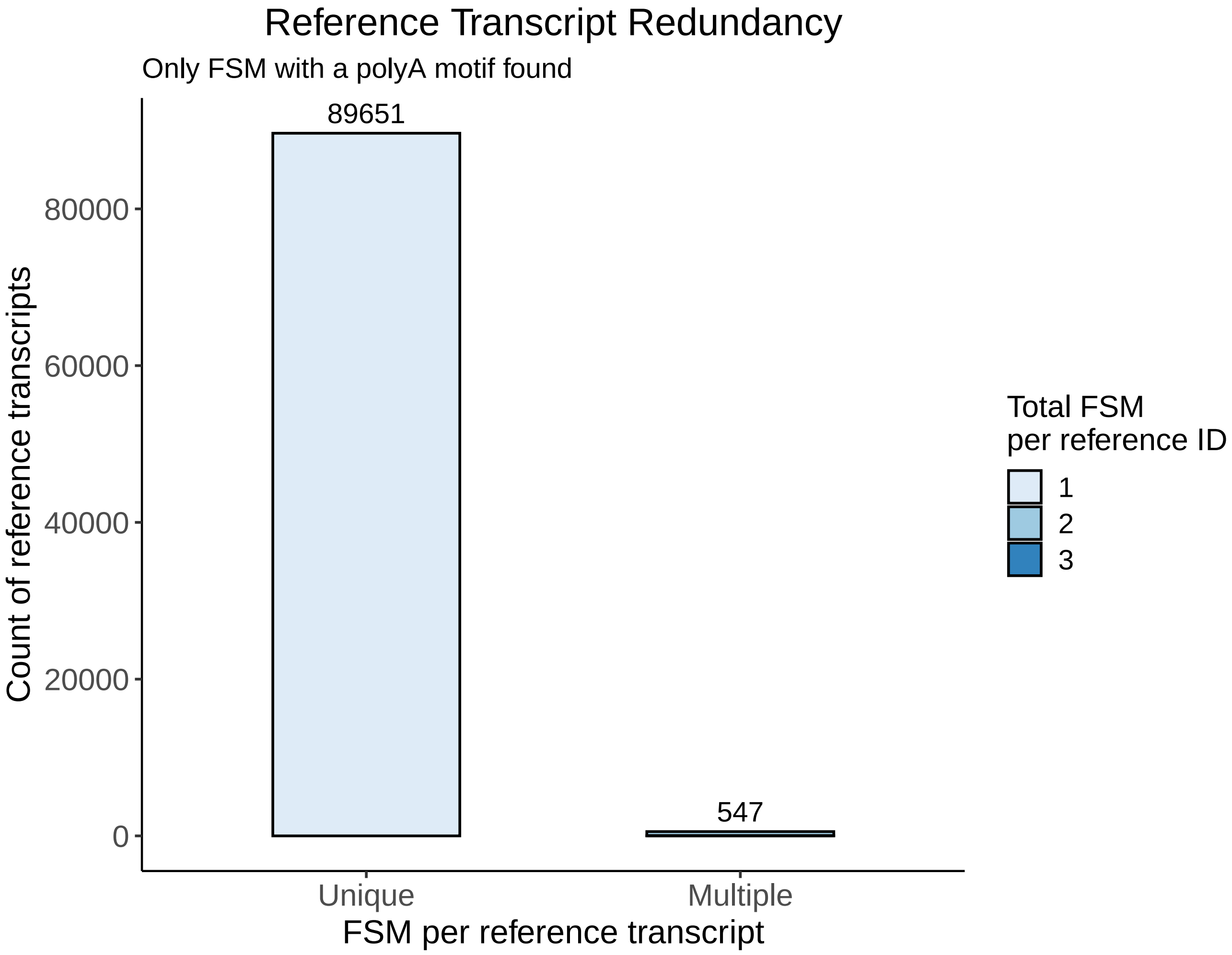
Redundancy Analysis


Reference Transcript Redundancy
Only FSM
174472
772
150000
100000
50000
0
Count of reference transcripts
Unique
Multiple
FSM per reference transcript
Total FSM
per reference ID
1
2
3

Reference Transcript Redundancy
Only ISM
13097
7153
Count of Reference Transcripts
12000
10000
8000
6000
4000
2000
0
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

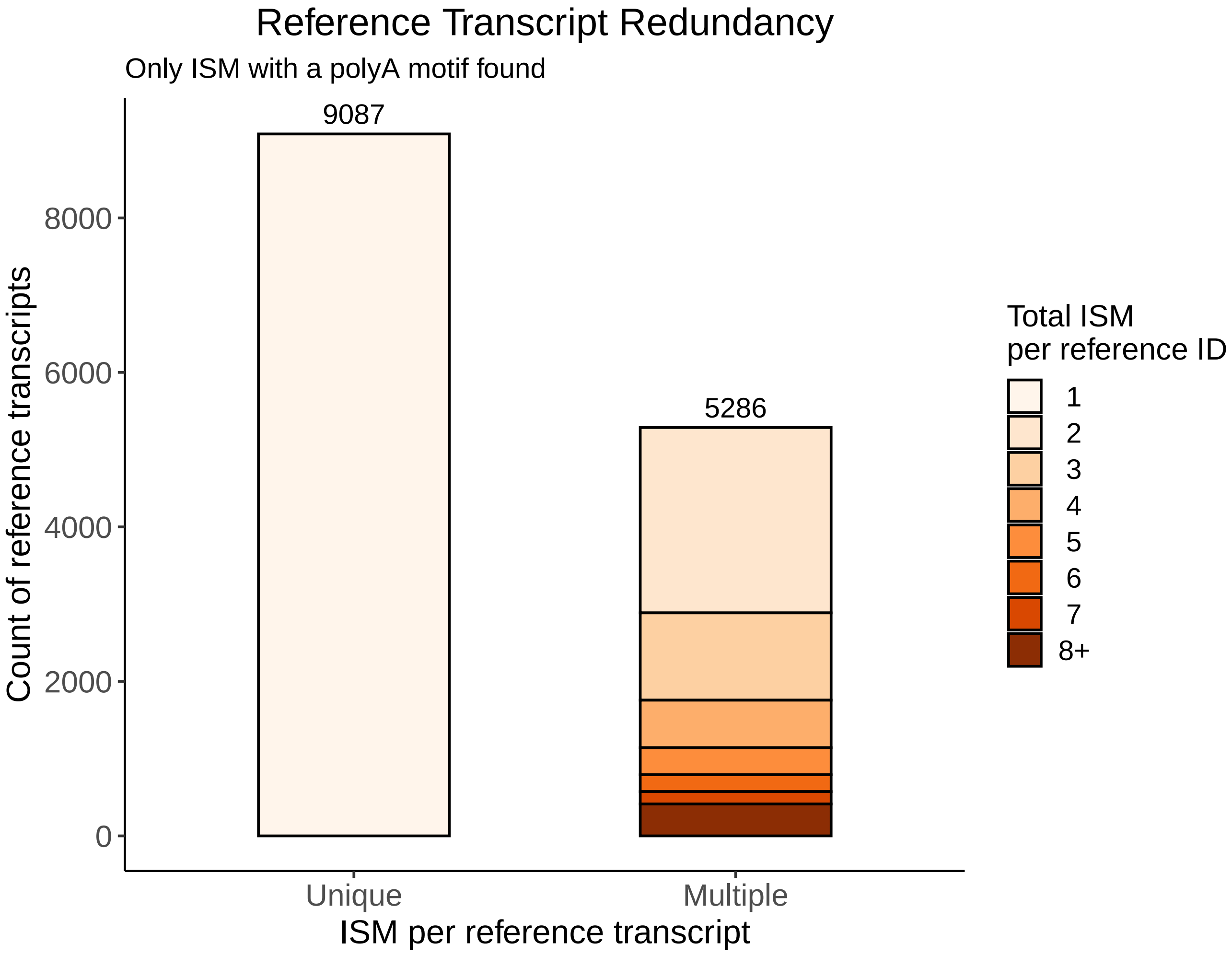


Reference Transcript Redundancy

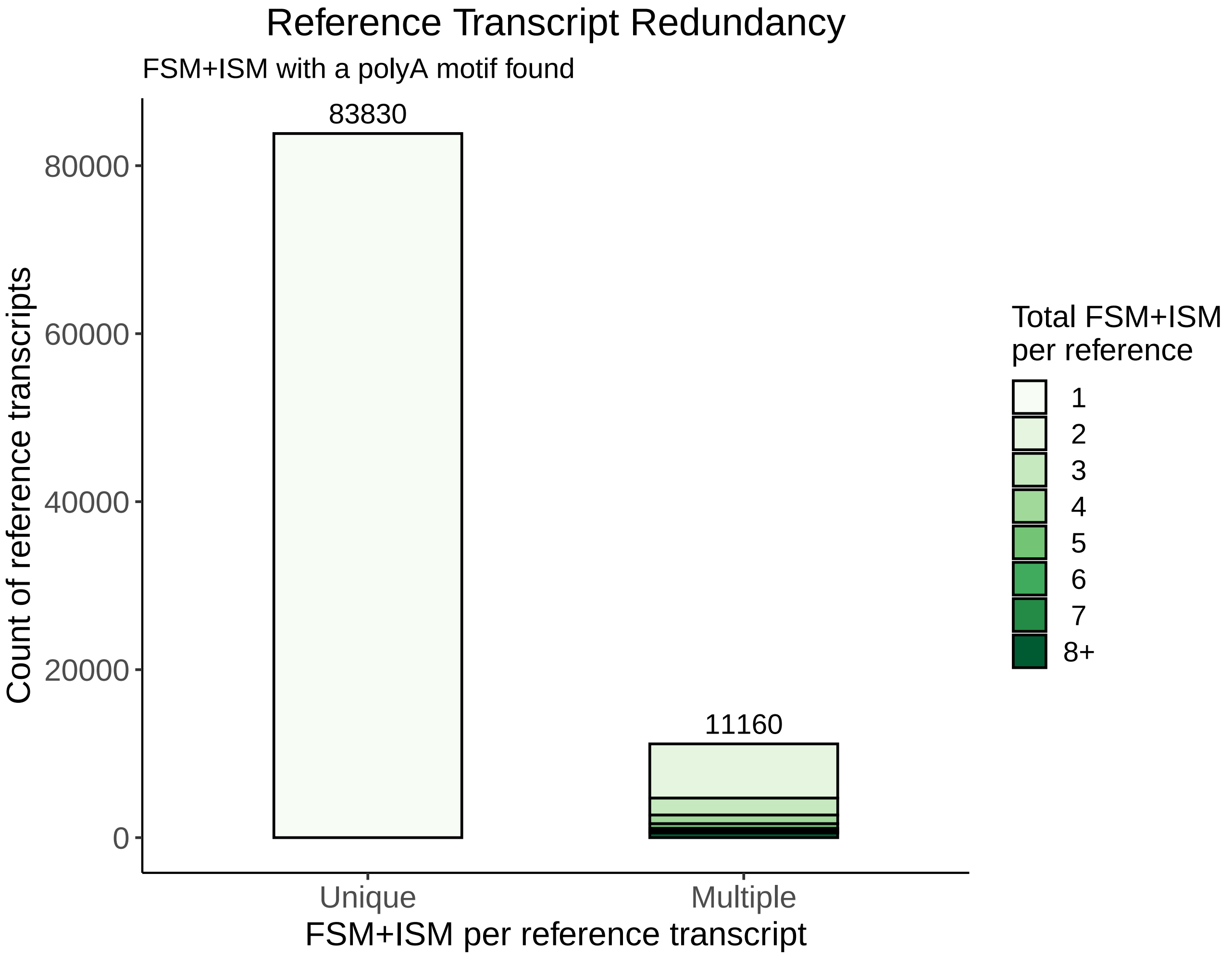


Reference Transcript Redundancy

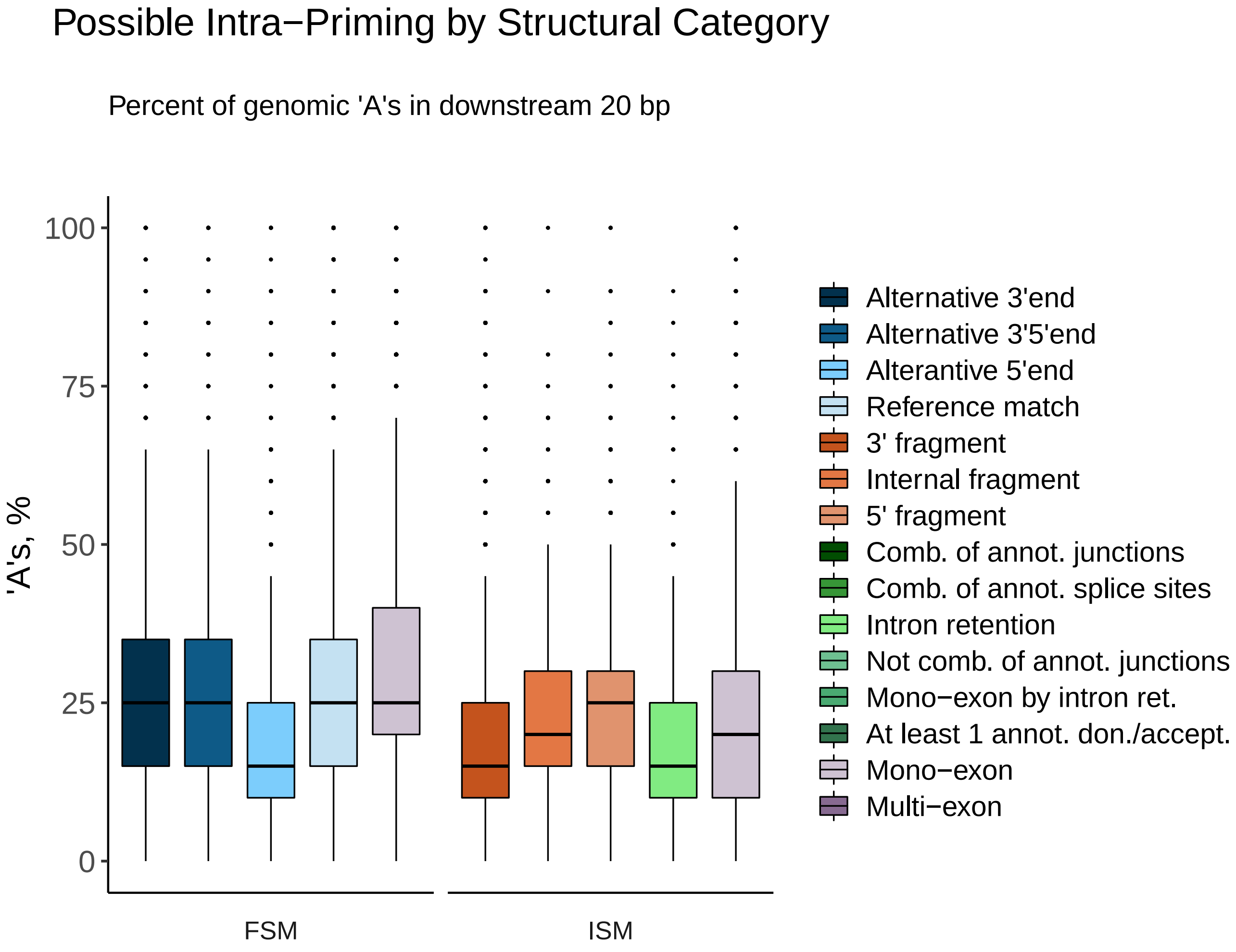
Only FSM with a polyA motif found

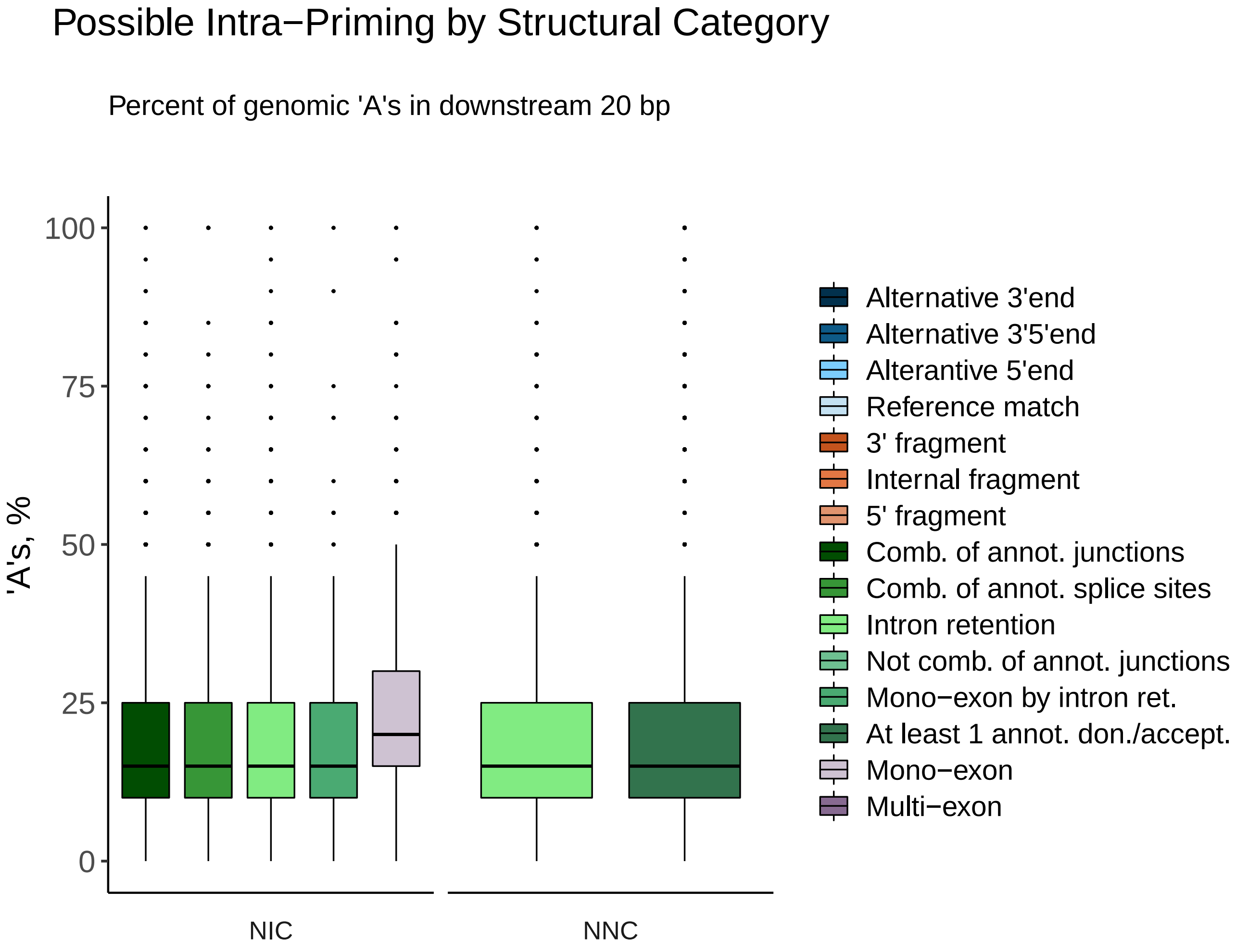

Reference Transcript Redundancy

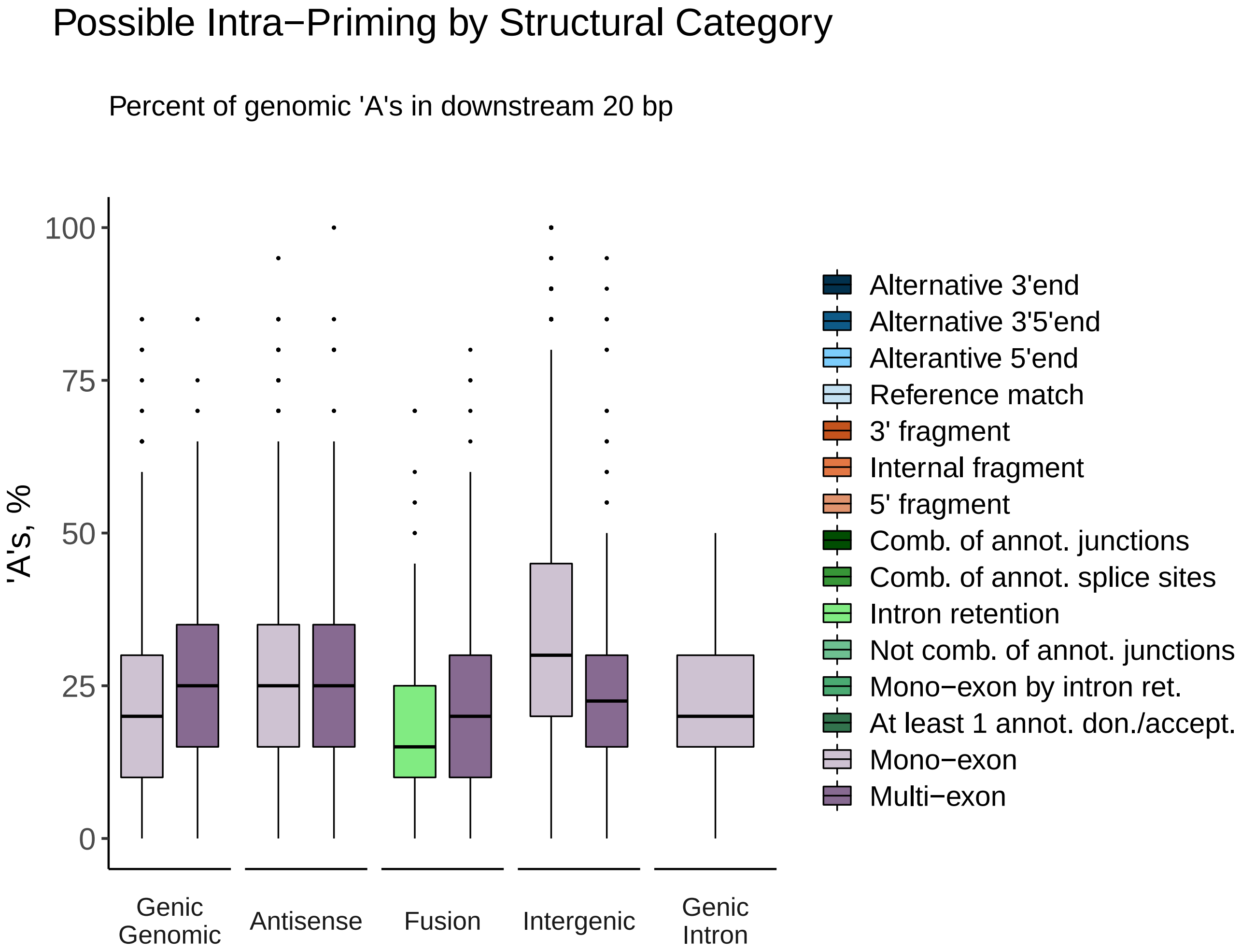
Only ISM with a polyA motif found


Reference Transcript Redundancy
FSM+ISM with a polyA motif found
83830
11160
80000
60000
40000
20000
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+

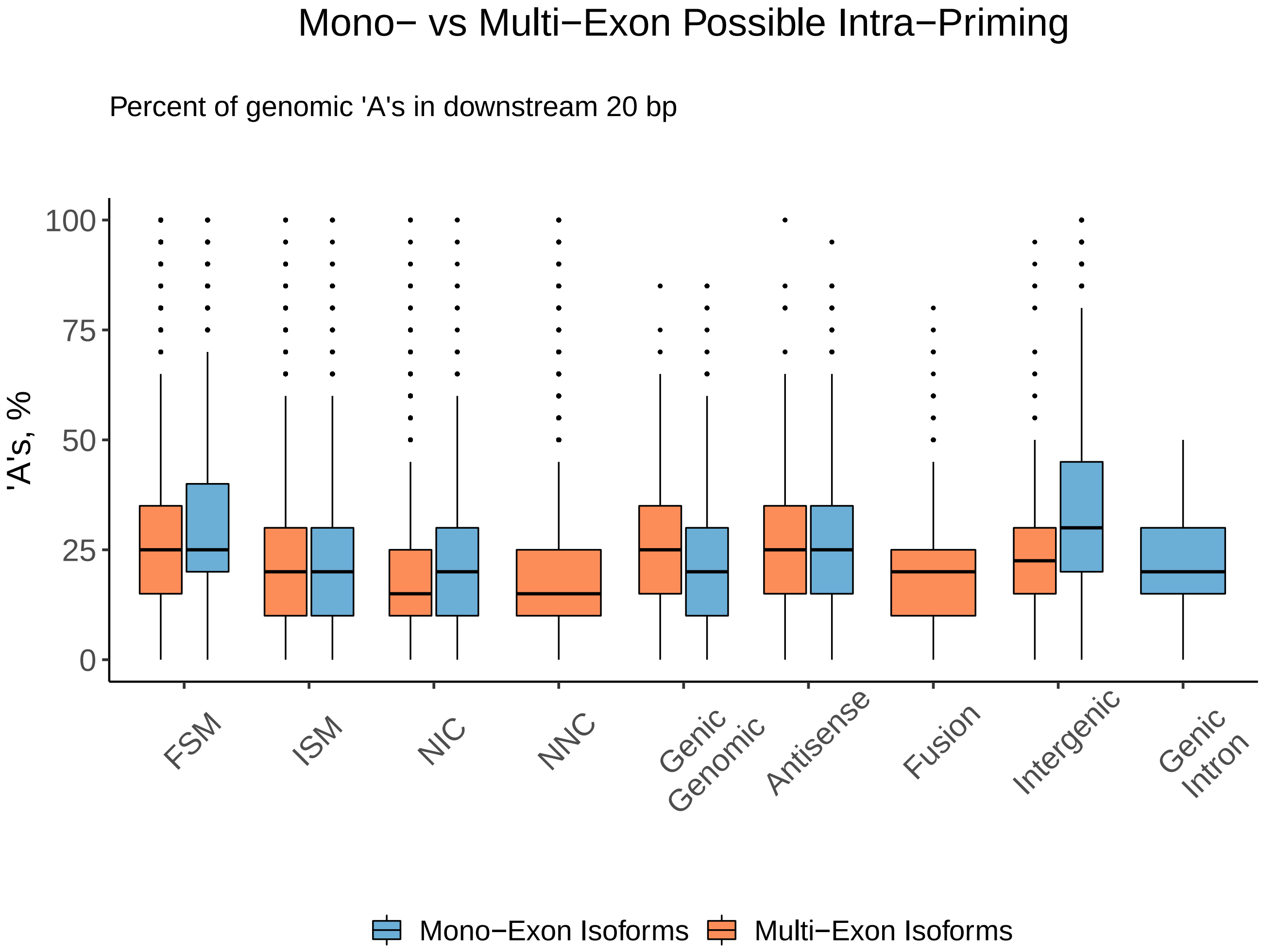
Intra−Priming Quality Check


Possible Intra−Priming by Structural Category
Percent of genomic 'A's in downstream 20 bp
'A's, %
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


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

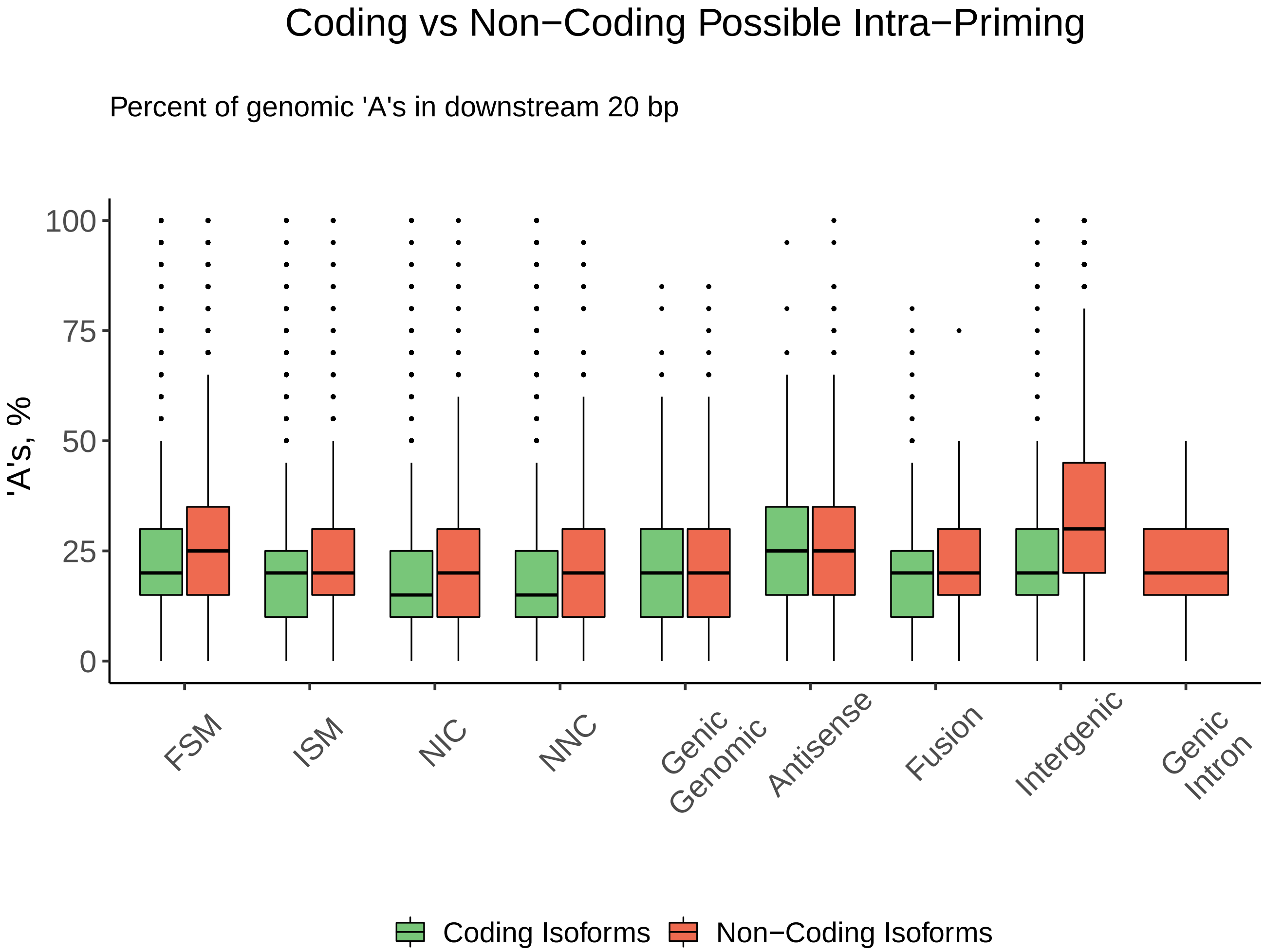

Possible Intra−Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp

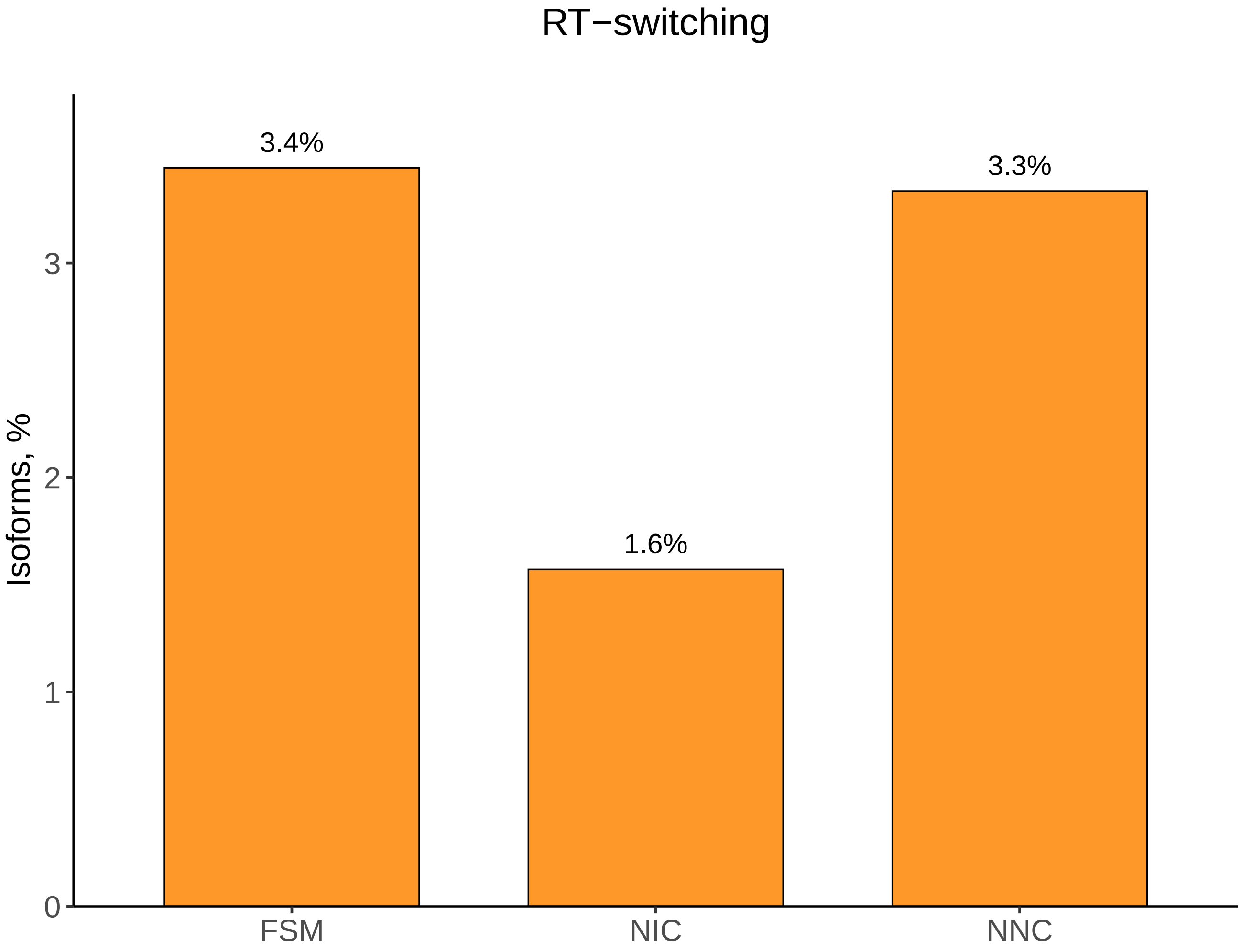


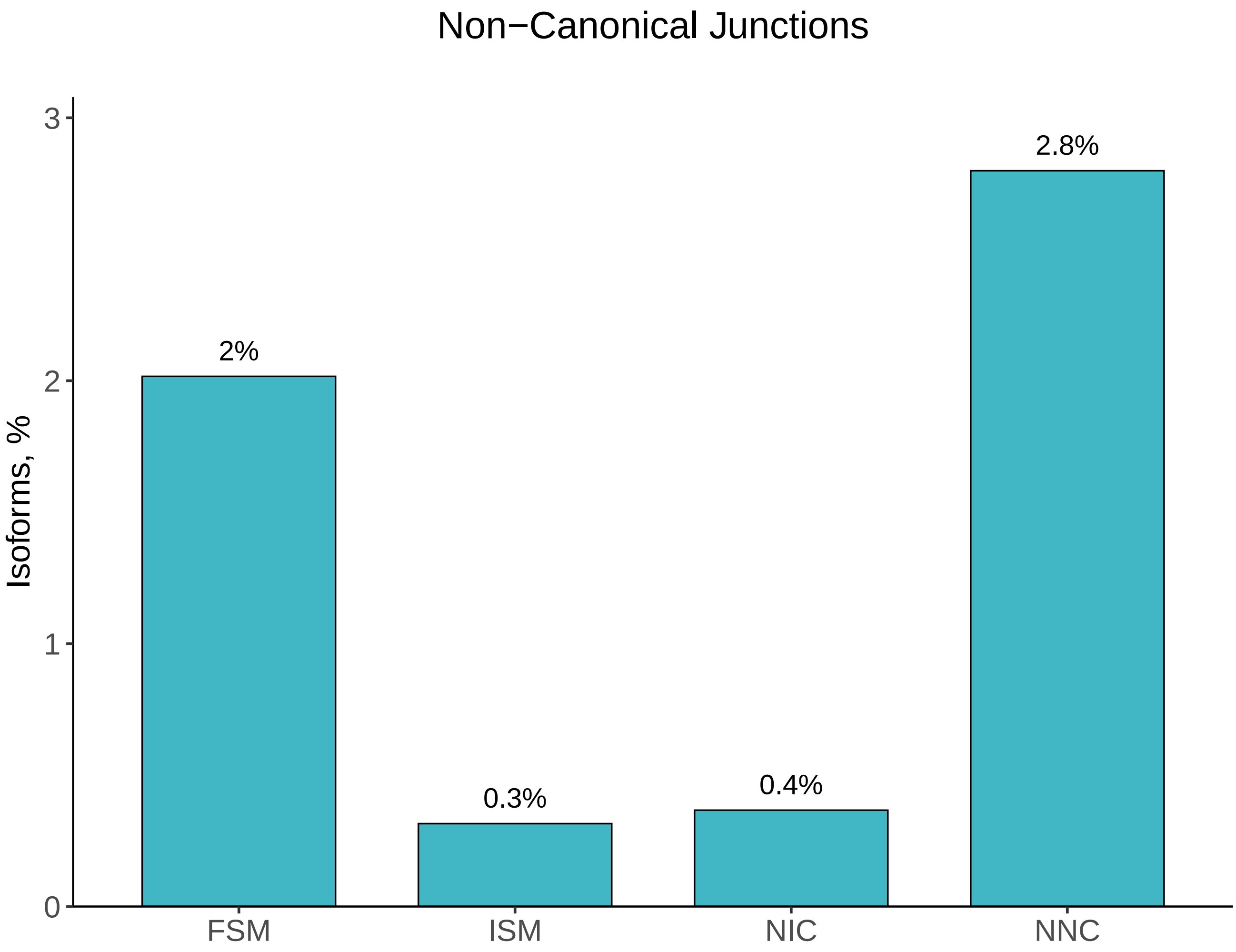
Mono− vs Multi−Exon Possible Intra−Priming

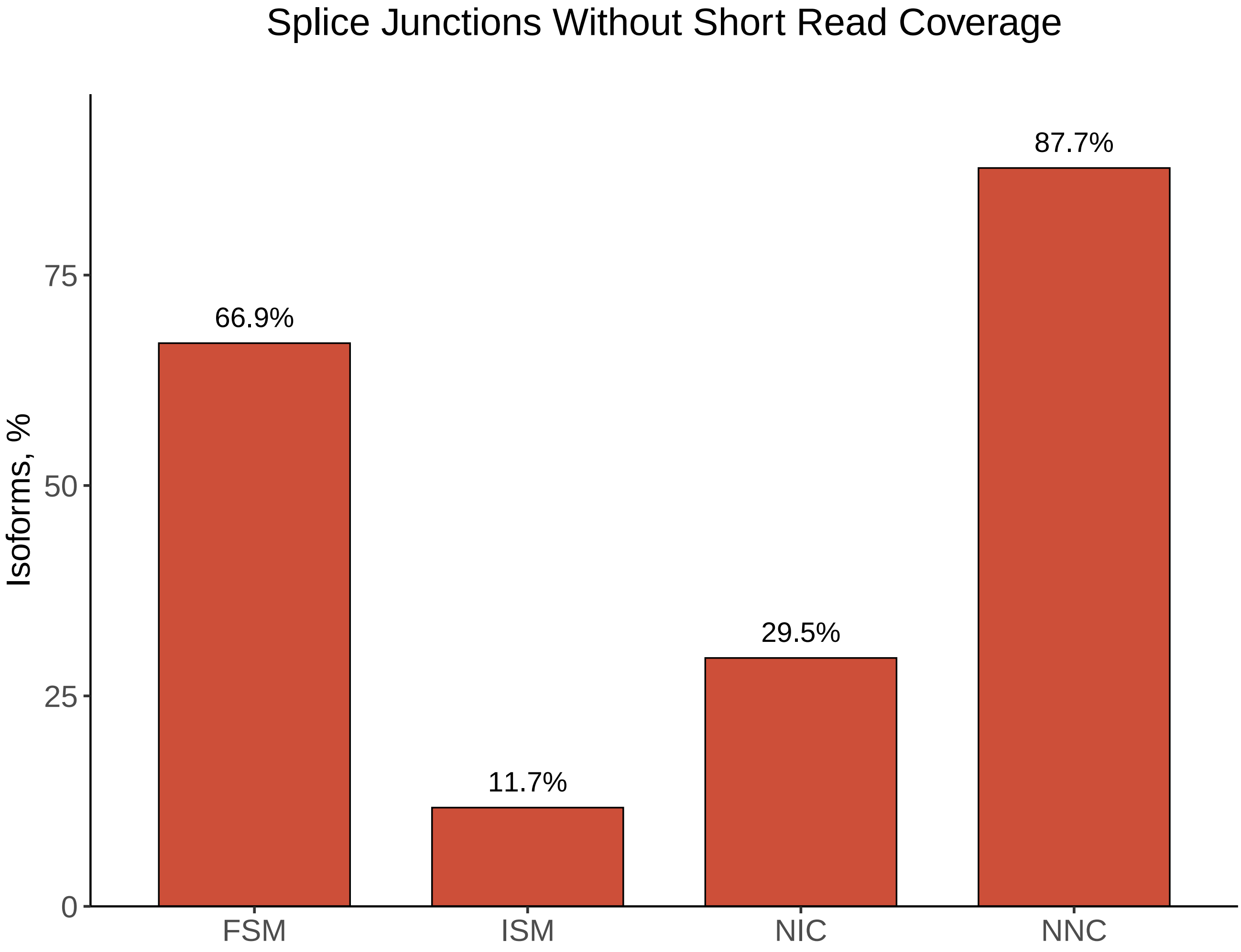
Percent of genomic 'A's in downstream 20 bp

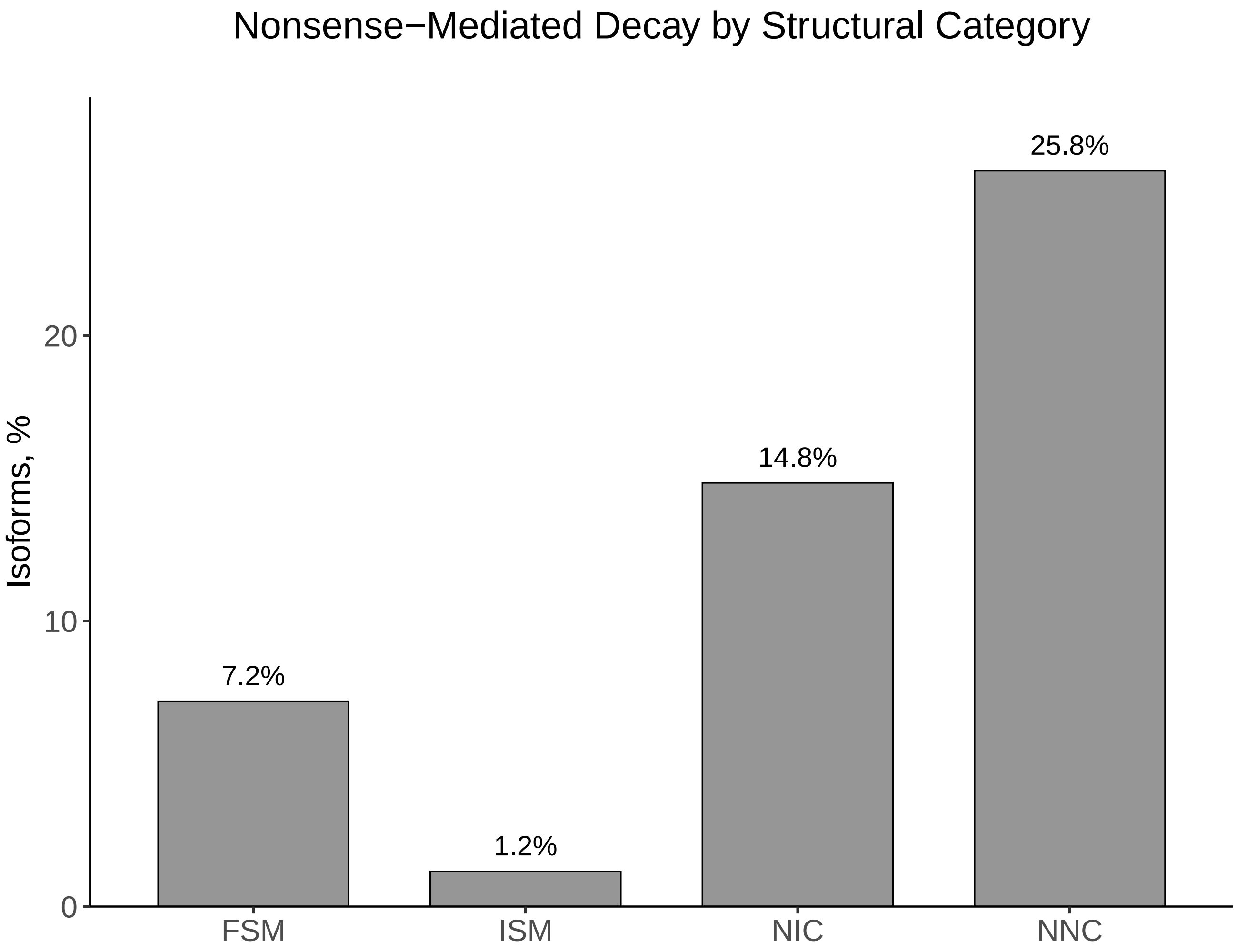

Coding vs Non−Coding Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Coding Isoforms
Non−Coding Isoforms

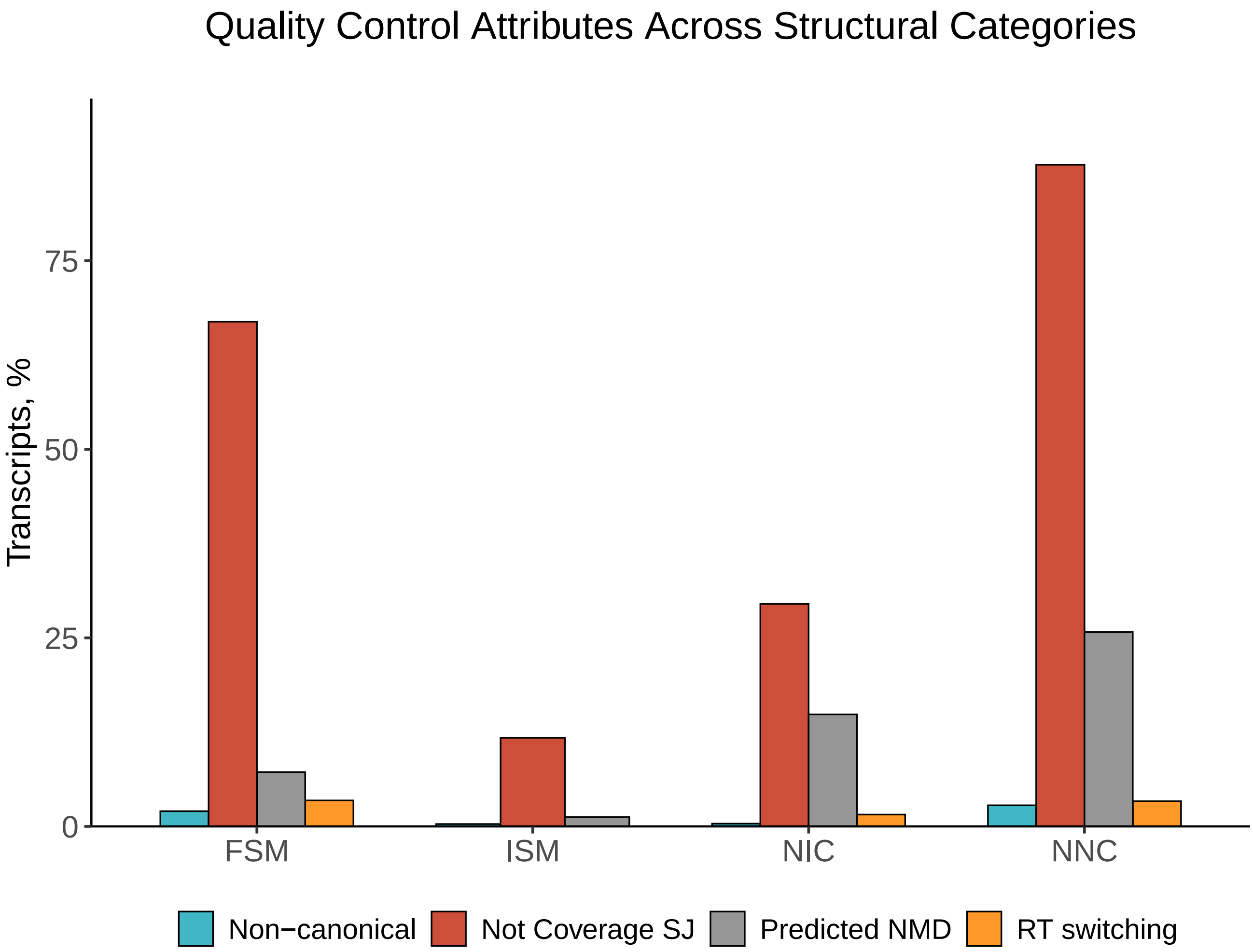
Features of Bad Quality


RT−switching
Isoforms, %
3.4%
1.6%
3.3%
0
1
2
3
FSM
NIC
NNC

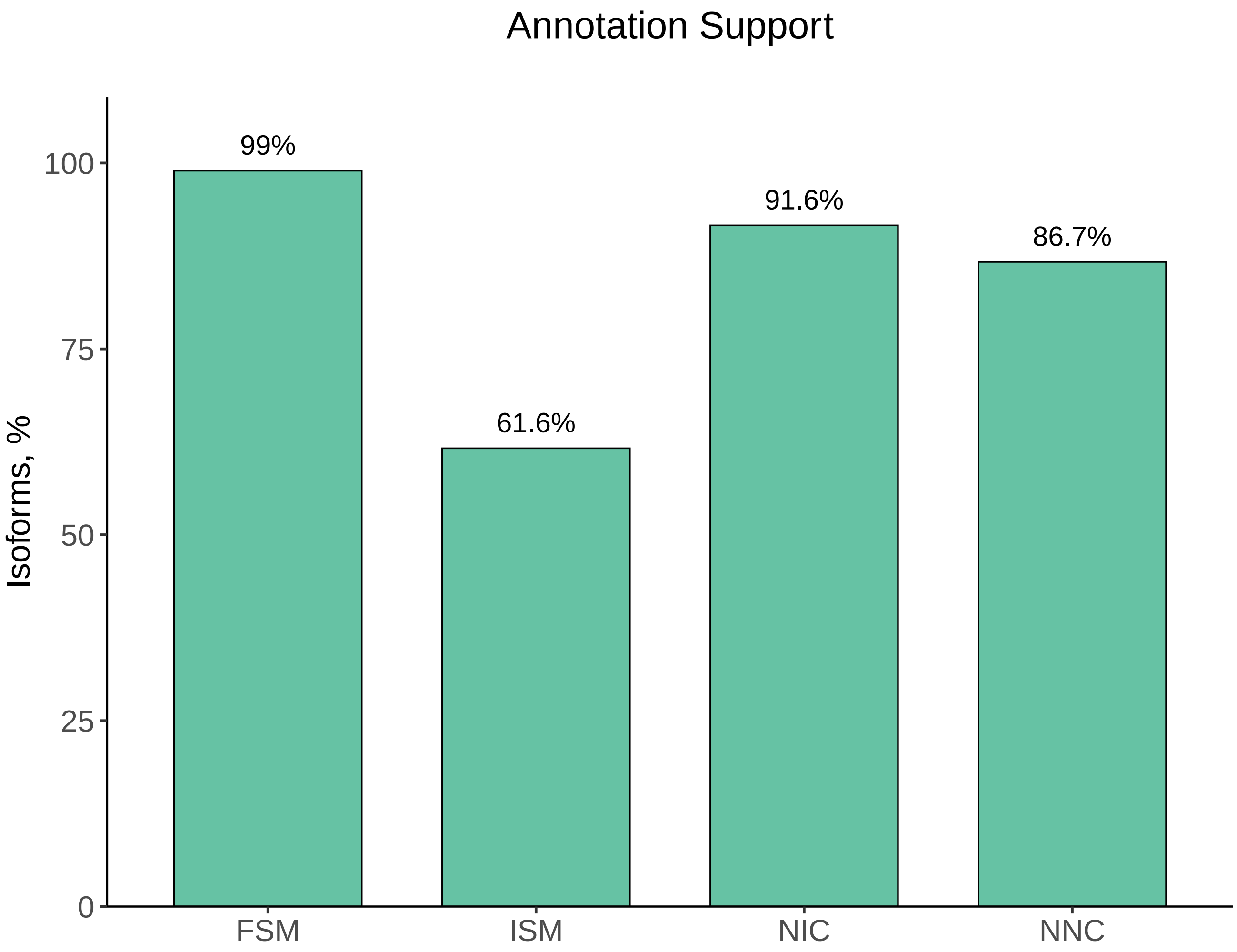

Non−Canonical Junctions
Isoforms, %
3
2
1
0
2%
0.3%
0.4%
2.8%
FSM
ISM
NIC
NNC

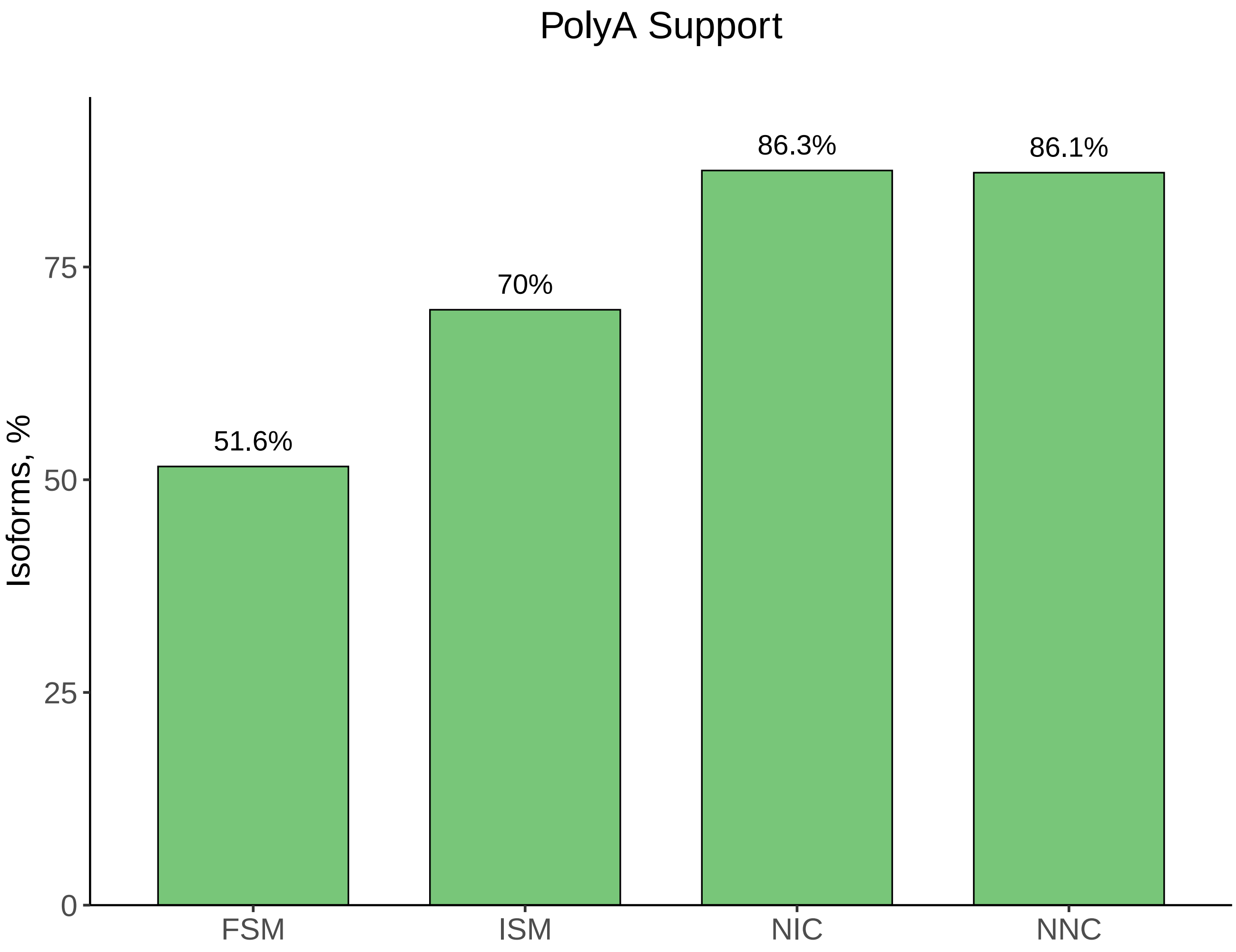

Splice Junctions Without Short Read Coverage
Isoforms, %
75
50
25
0
66.9%
11.7%
29.5%
87.7%
FSM
ISM
NIC
NNC

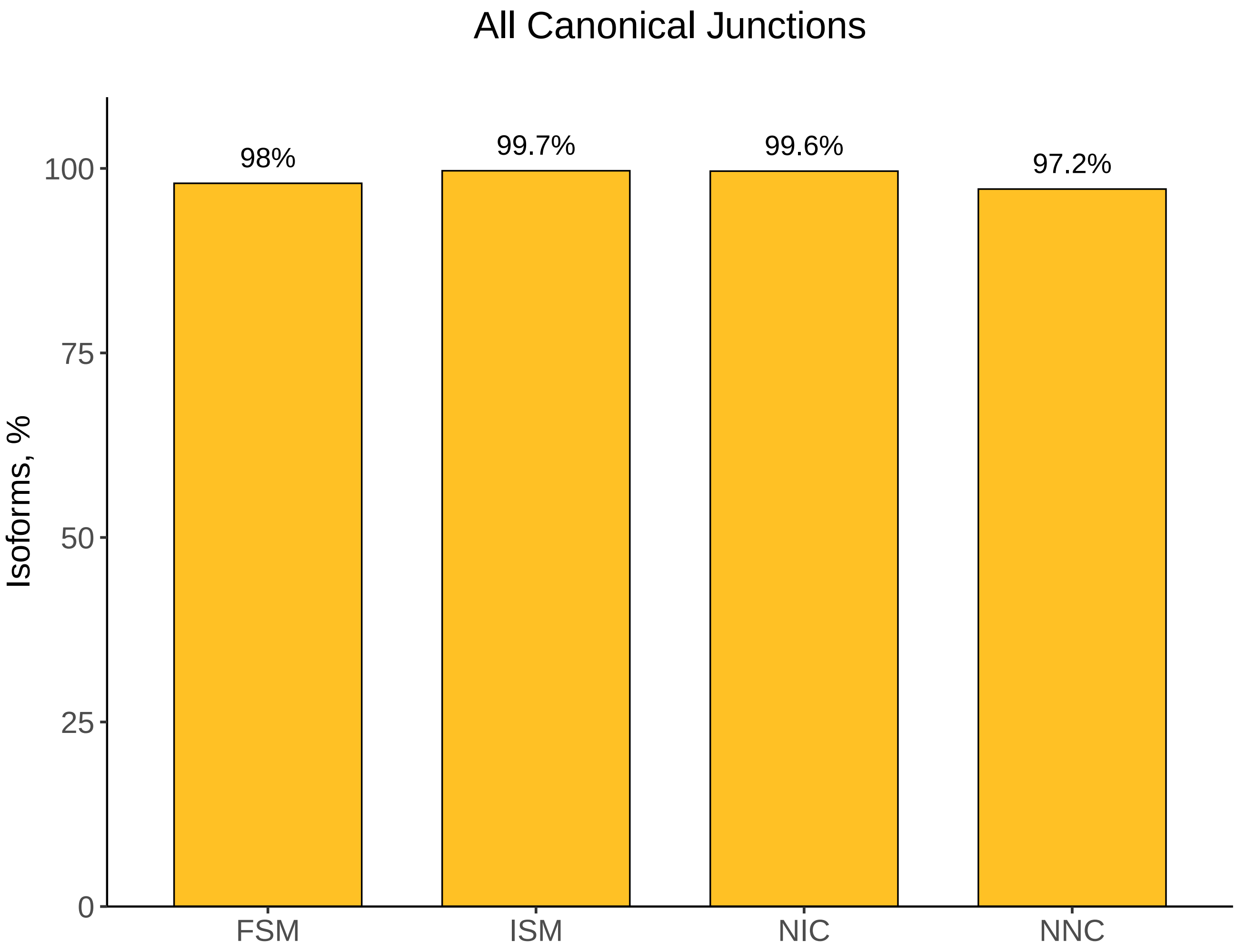

Nonsense−Mediated Decay by Structural Category
Isoforms, %
0
10
20
7.2%
1.2%
14.8%
25.8%
FSM
ISM
NIC
NNC

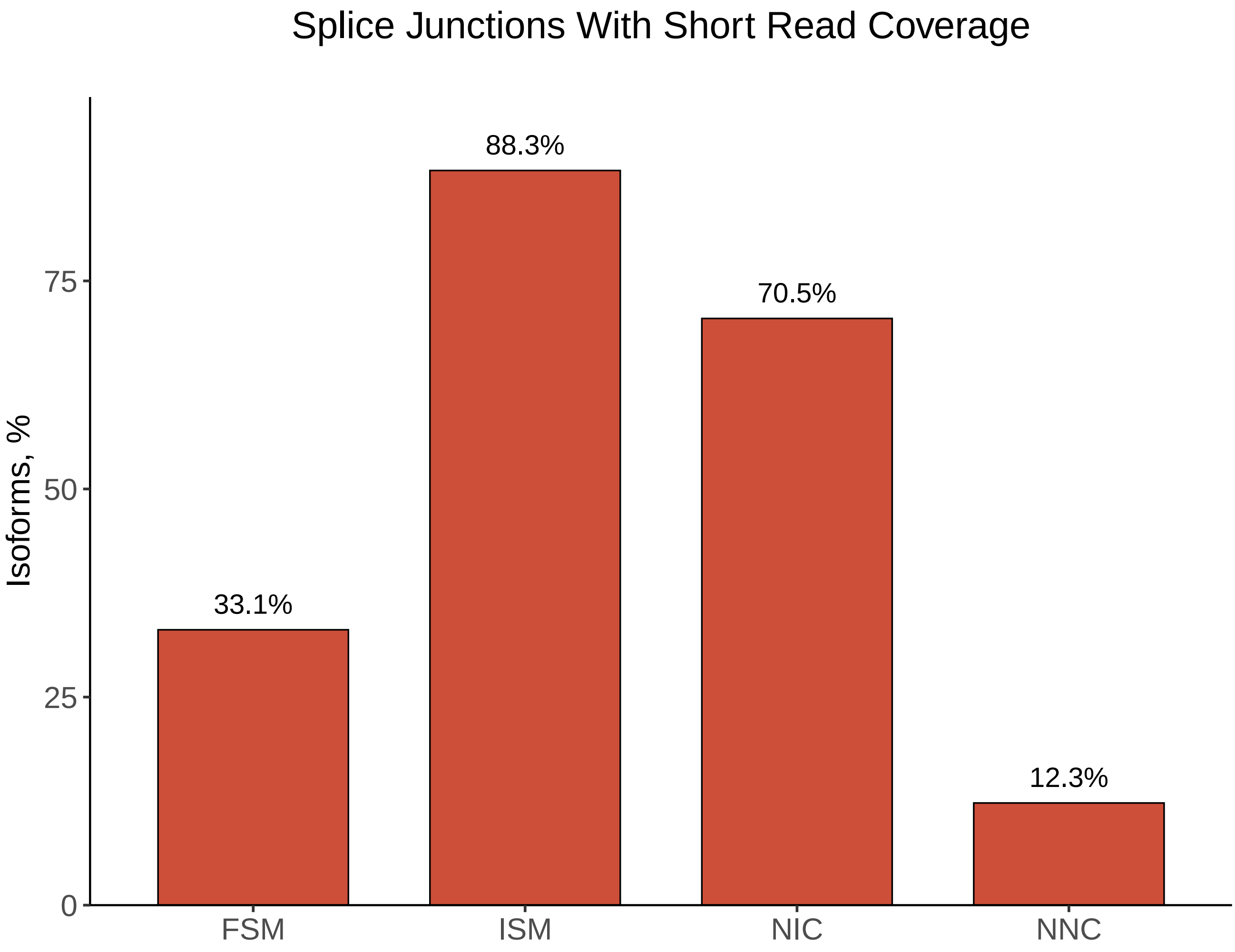

Quality Control Attributes Across Structural Categories
Transcripts, %
0
25
50
75
FSM
ISM
NIC
NNC
Non−canonical
Not Coverage SJ
Predicted NMD
RT switching

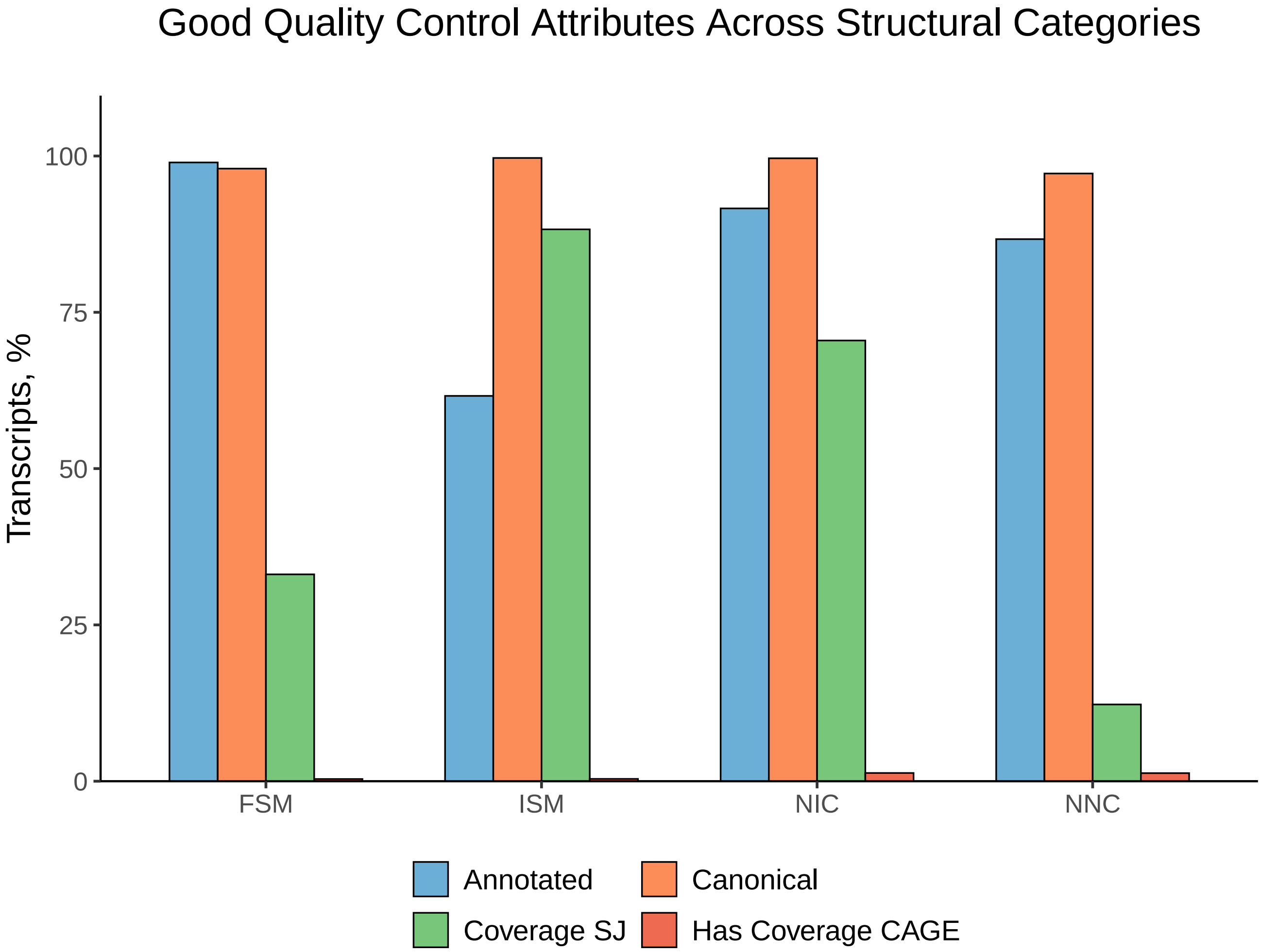
Features of Good Quality


Annotation Support
Isoforms, %
0
25
50
75
100
99%
61.6%
91.6%
86.7%
FSM
ISM
NIC
NNC


PolyA Support
Isoforms, %
0
25
50
75
51.6%
70%
86.3%
86.1%
FSM
ISM
NIC
NNC


All Canonical Junctions
Isoforms, %
0
25
50
75
100
98%
99.7%
99.6%
97.2%
FSM
ISM
NIC
NNC


Splice Junctions With Short Read Coverage
Isoforms, %
0
25
50
75
33.1%
88.3%
70.5%
12.3%
FSM
ISM
NIC
NNC


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