SQANTI3 report

Unique Genes: 656

Unique Isoforms: 3925

Gene Classification

Category	Genes, count
Annotated Genes	443
Novel Genes	213

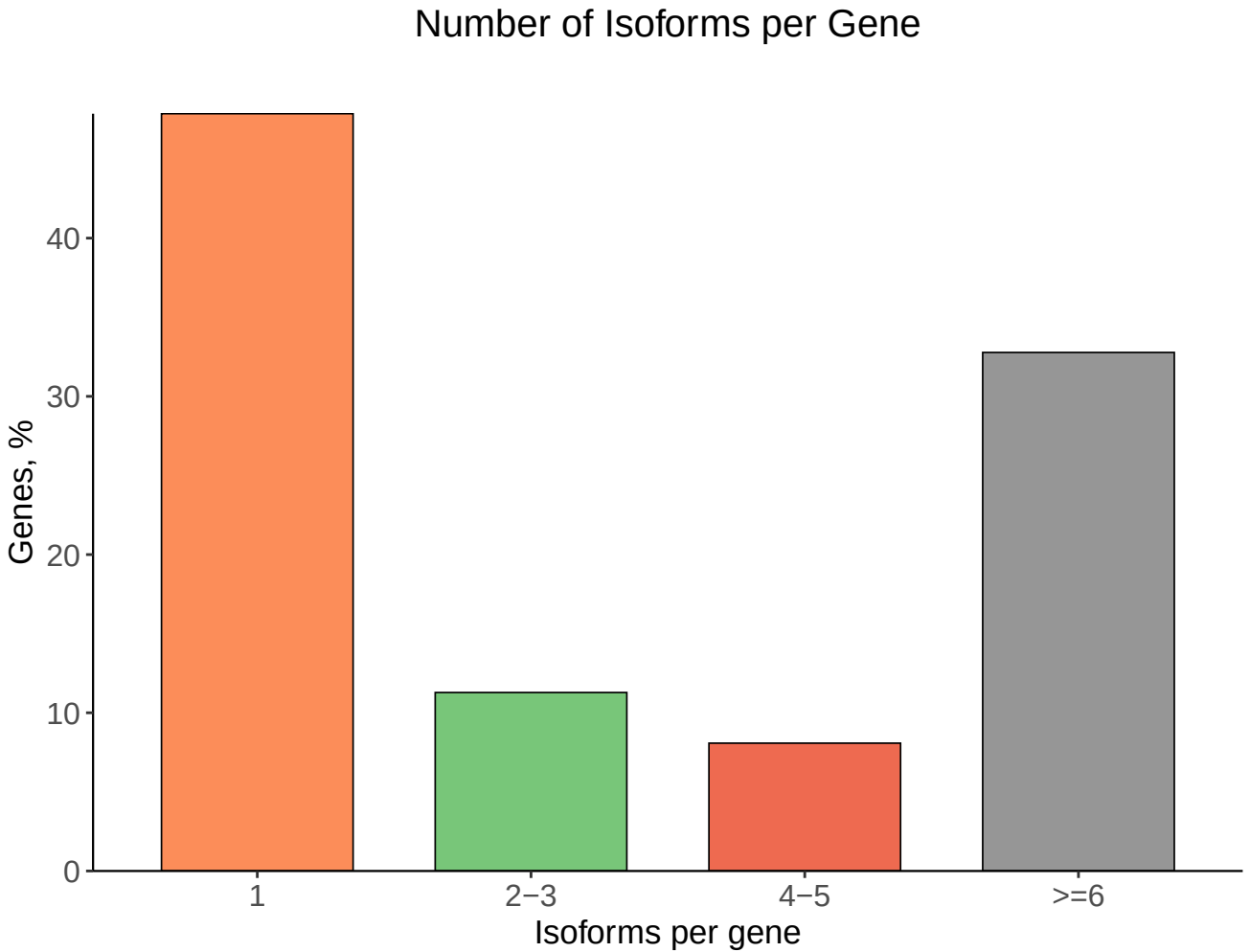
Transcript Classification

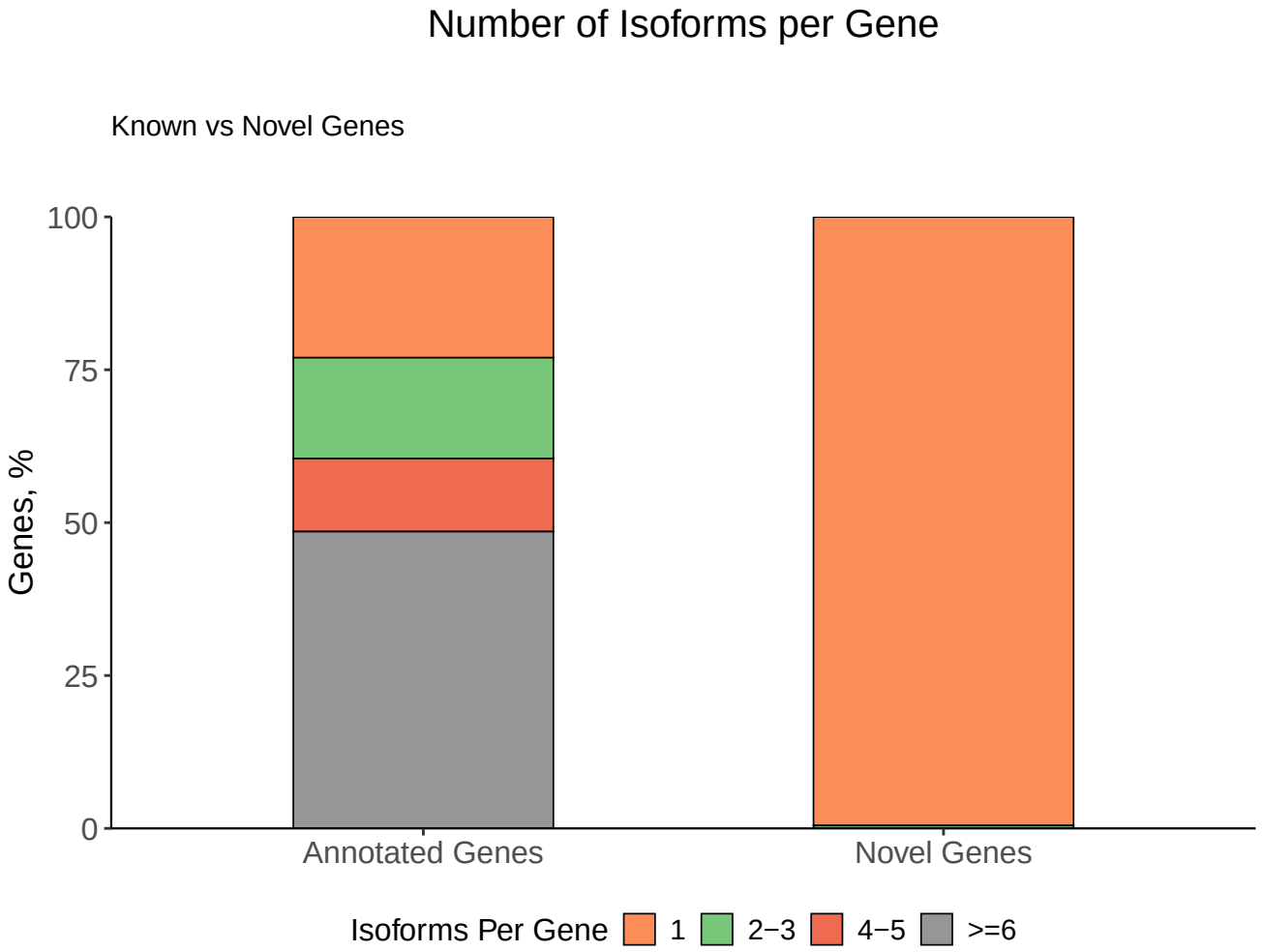
Category	Isoforms, count
FSM	539
ISM	1138
NIC	744
NNC	1139
Genic Genomic	113
Antisense	26
Fusion	38
Intergenic	58
Genic Intron	130

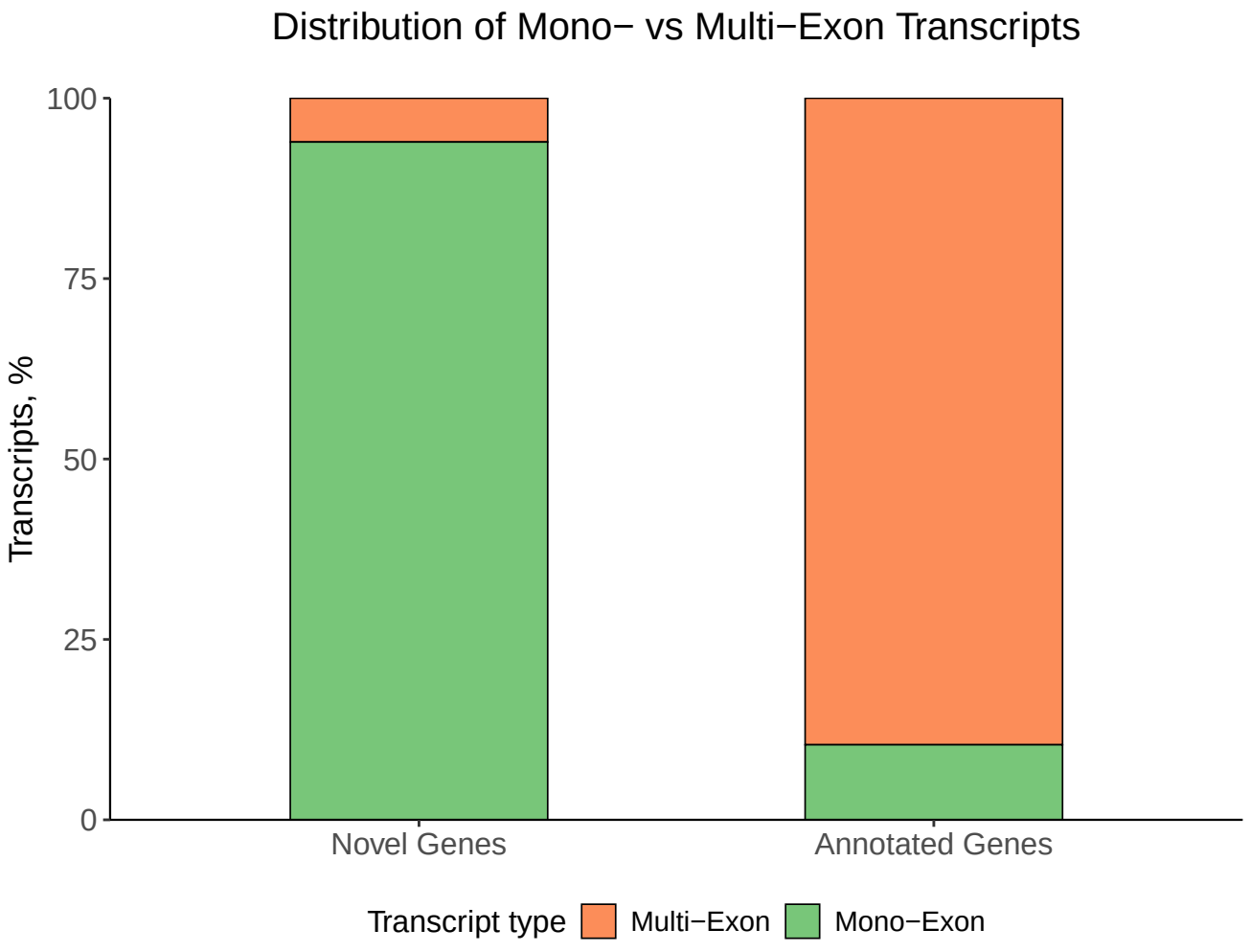
Splice Junction Classification

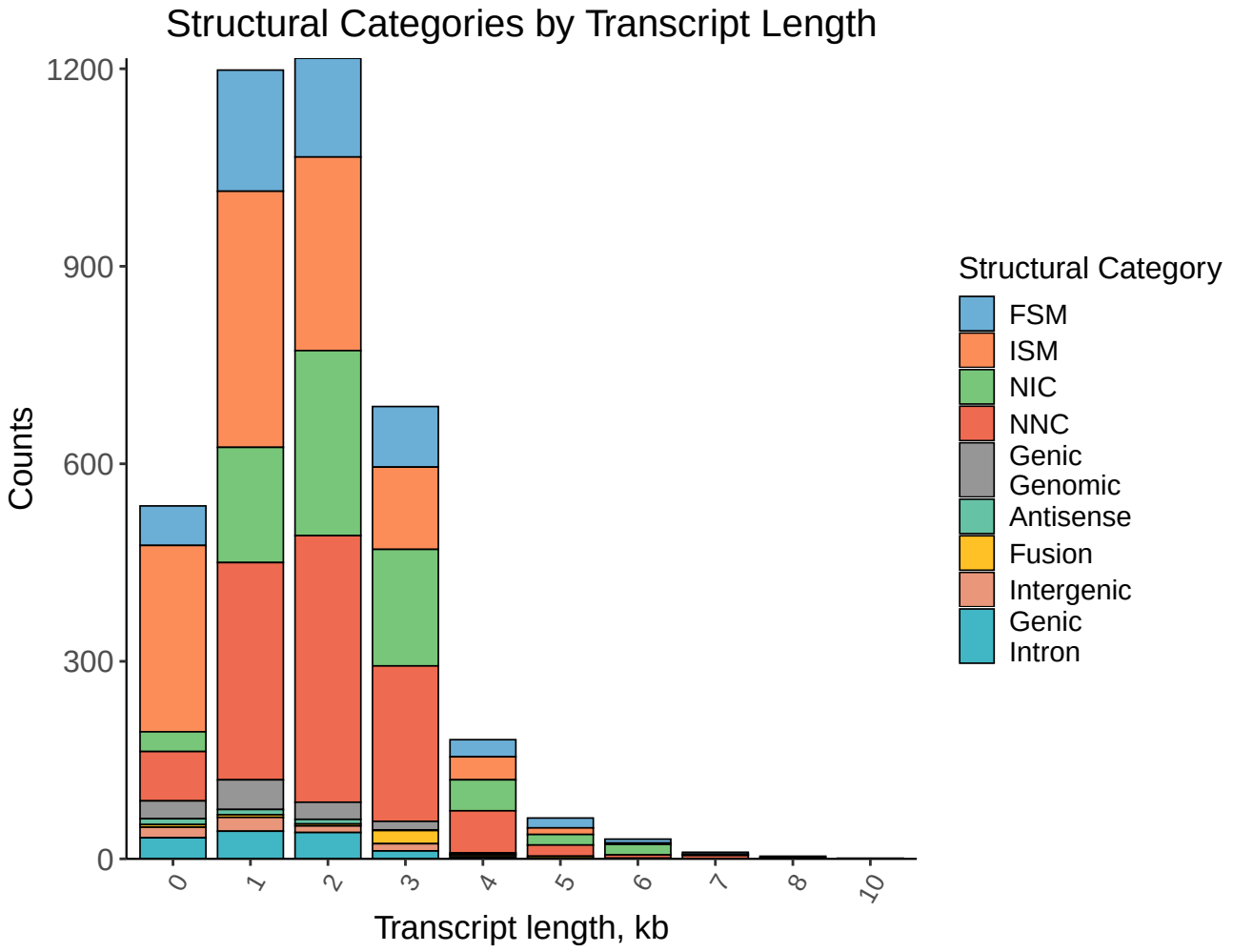
Category	SJs, count	Percent
Known canonical	3374	69.74
Known Non−canonical	4	0.08
Novel canonical	1195	24.70
Novel Non−canonical	265	5.48

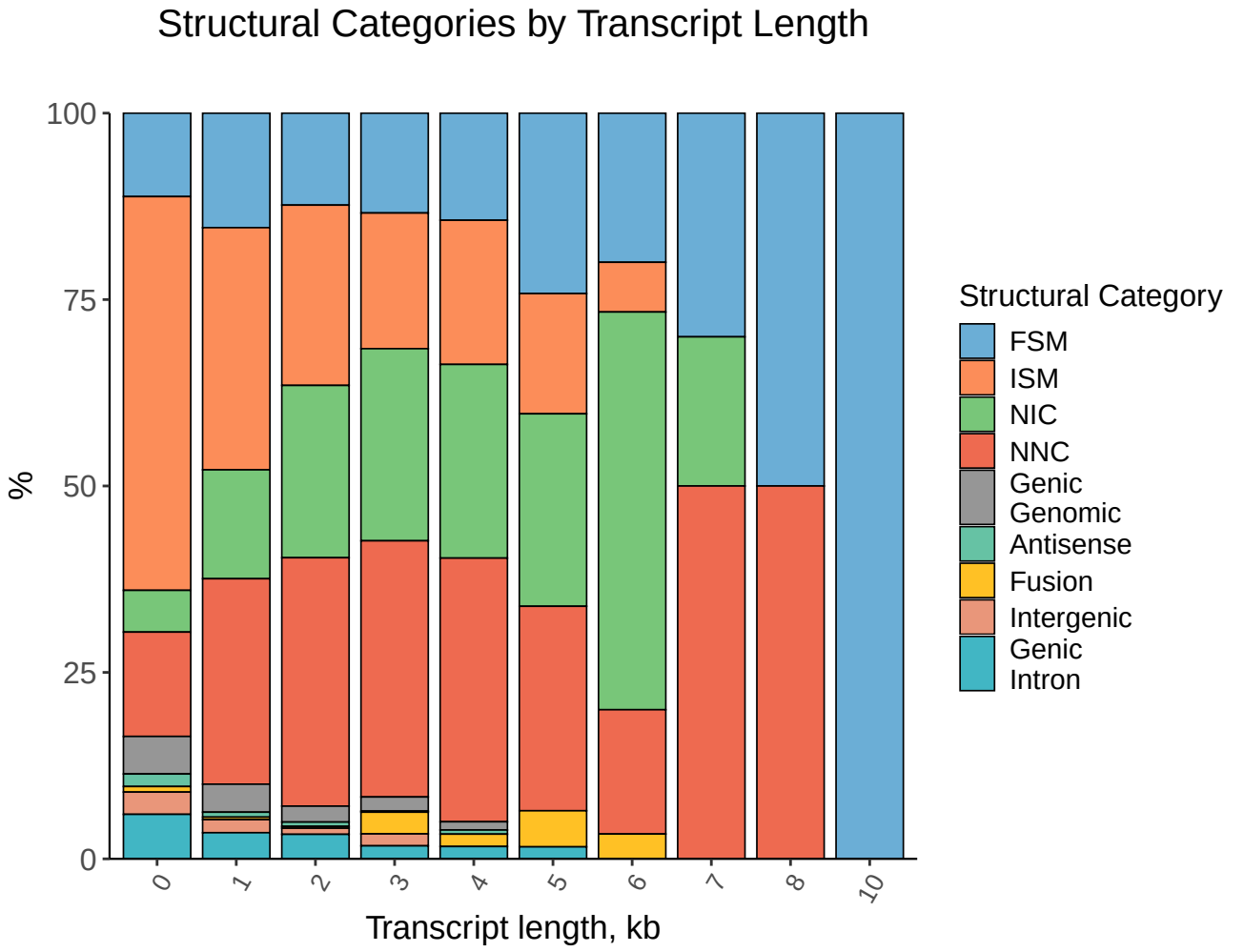
Gene Characterization

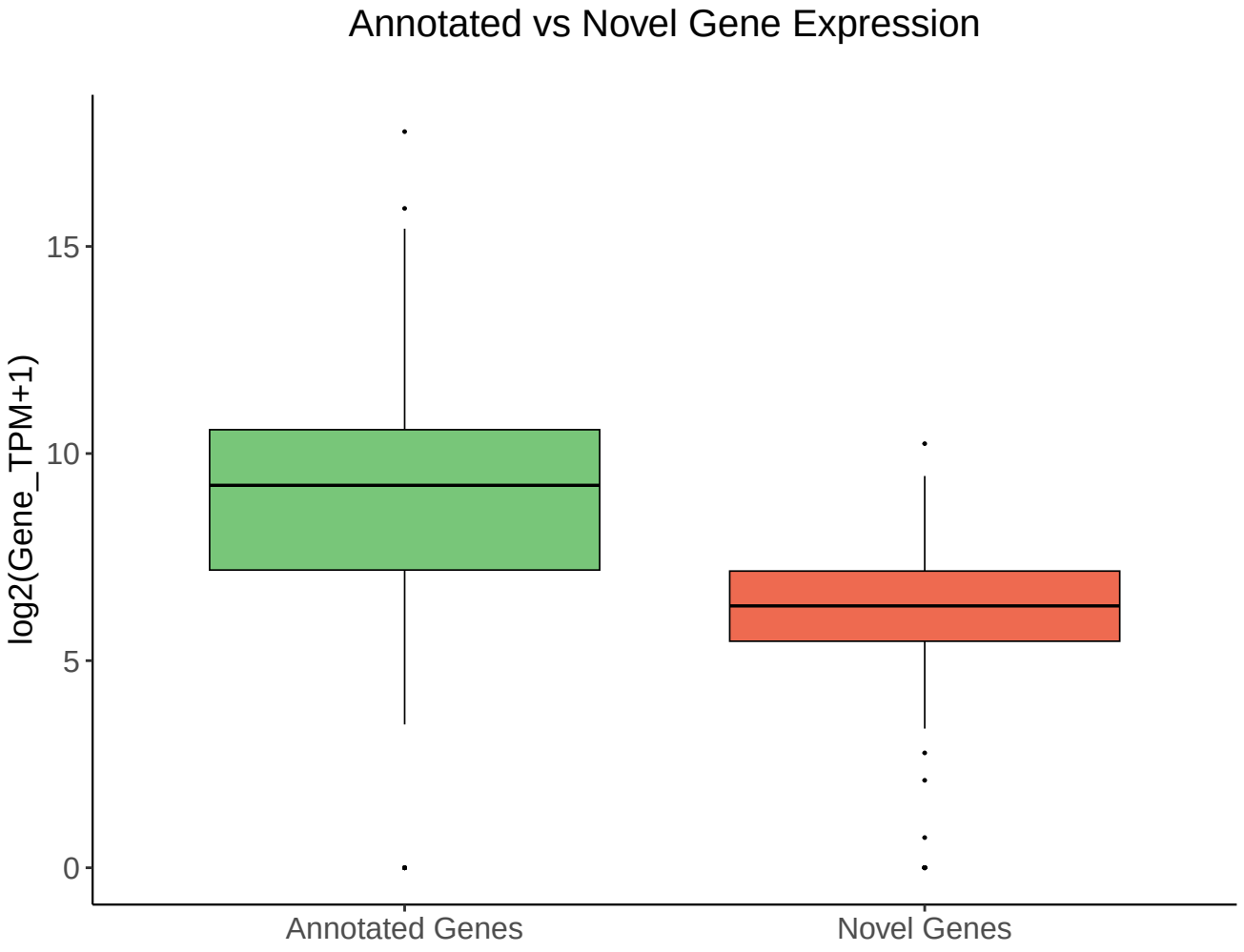

Number of Isoforms per Gene
Genes, %
0
10
20
30
40
1
2−3
4−5
>=6
Isoforms per gene

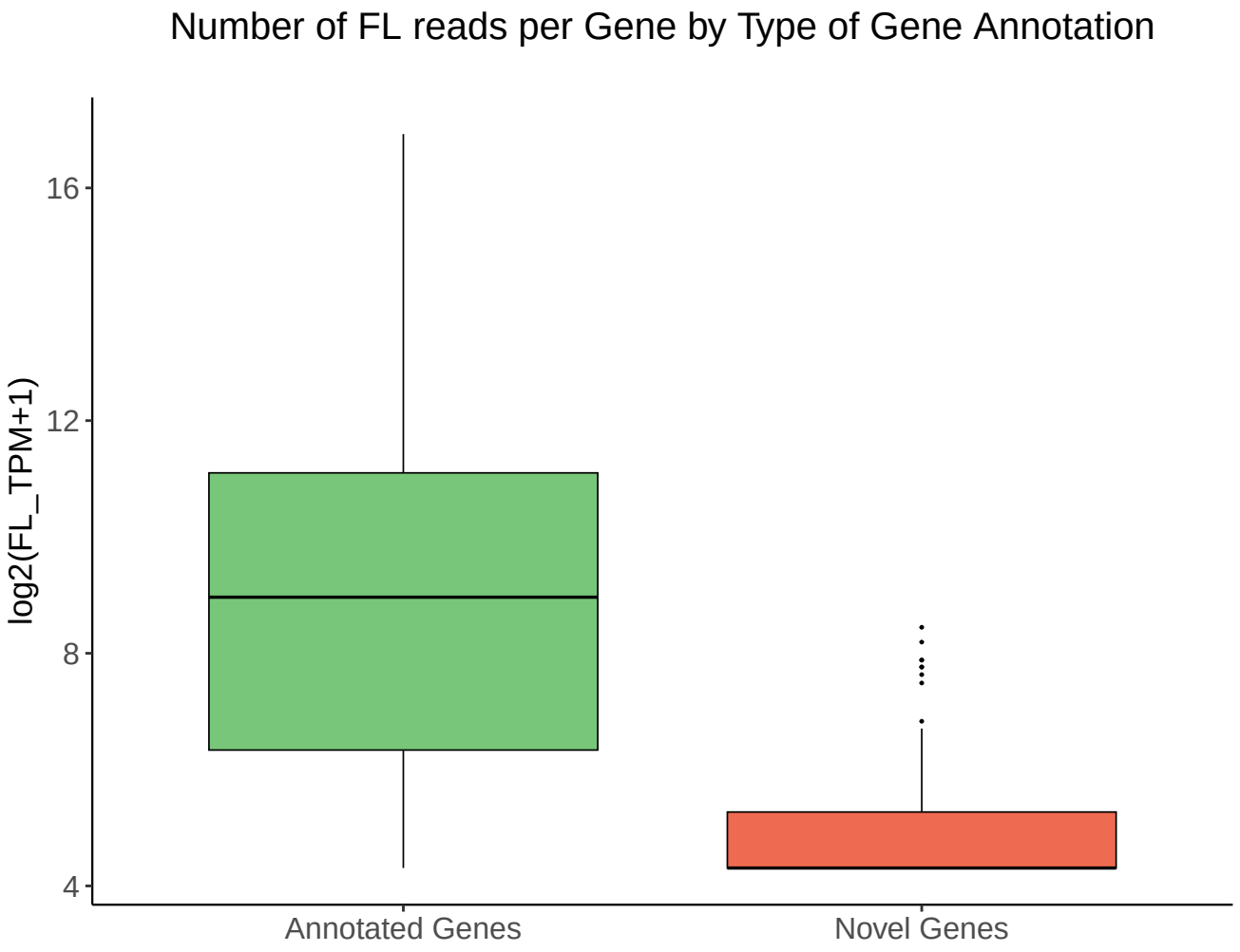

Number of Isoforms per Gene
Known vs Novel Genes
Genes, %
100
75
50
25
0
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3
4−5
>=6

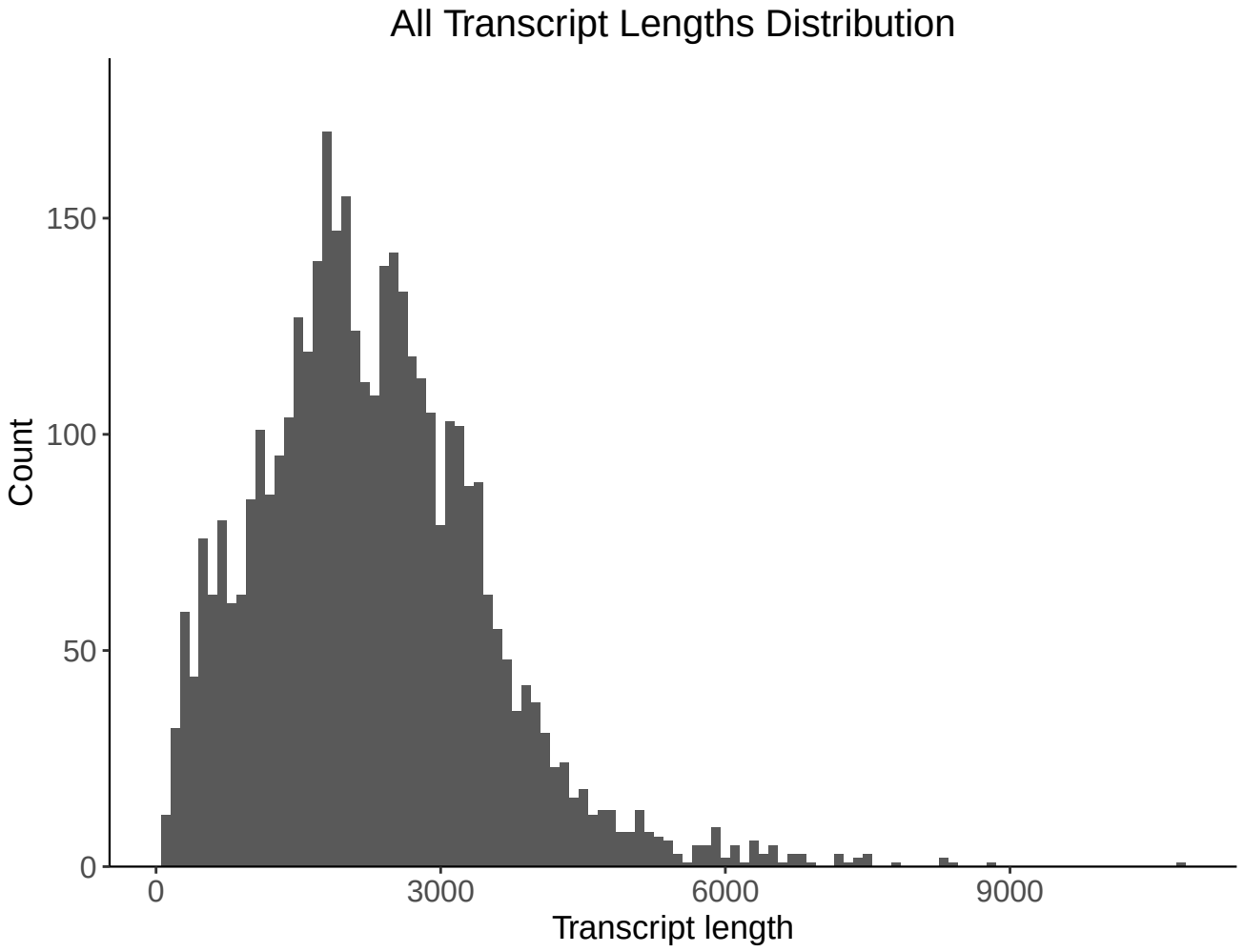

Distribution of Mono− vs Multi−Exon Transcripts
Transcripts, %
100
75
50
25
0
Novel Genes
Annotated Genes
Transcript type
Multi−Exon
Mono−Exon

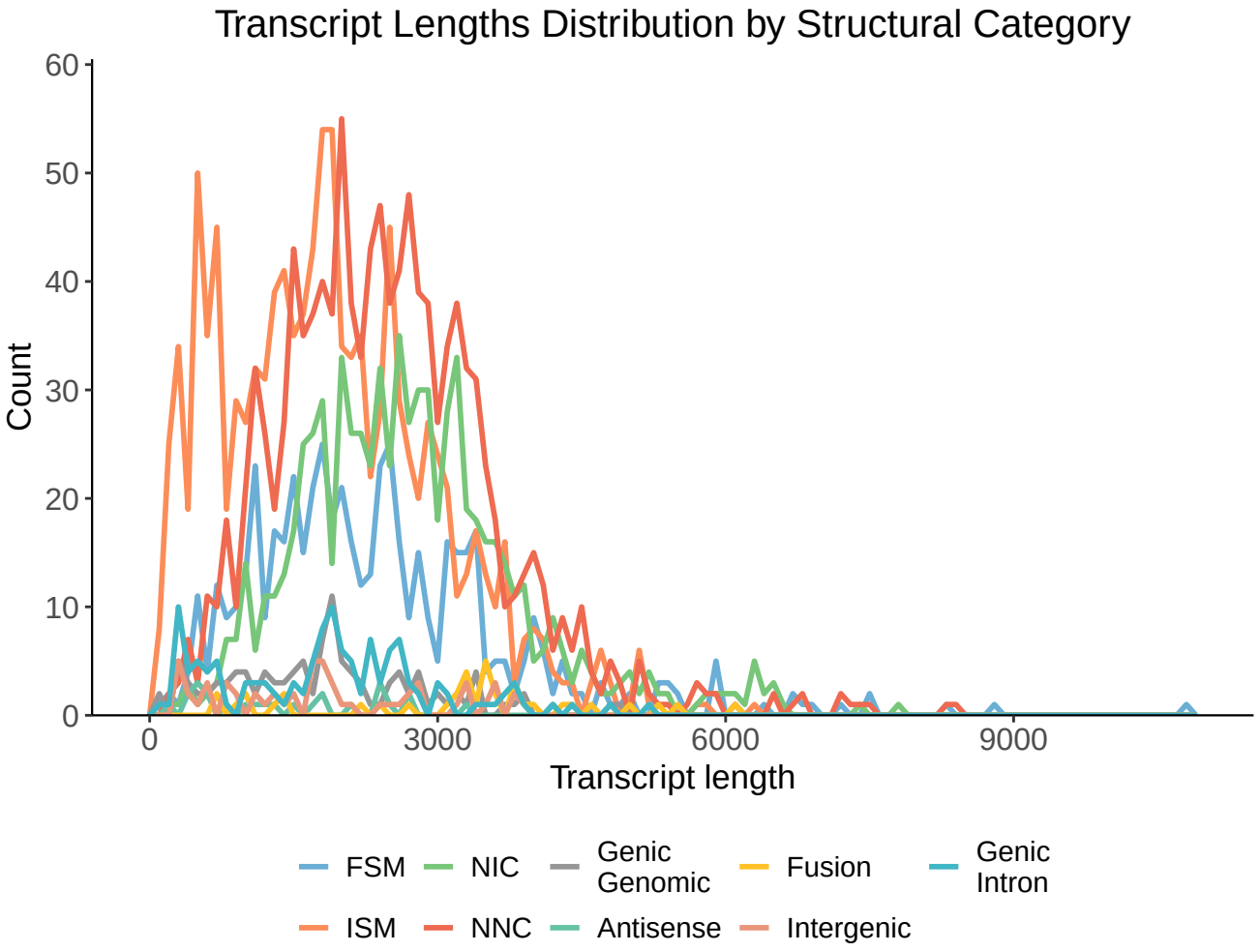

Structural Categories by Transcript Length
Counts
1200
900
600
300
0
0
1
2
3
4
5
6
7
8
10
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

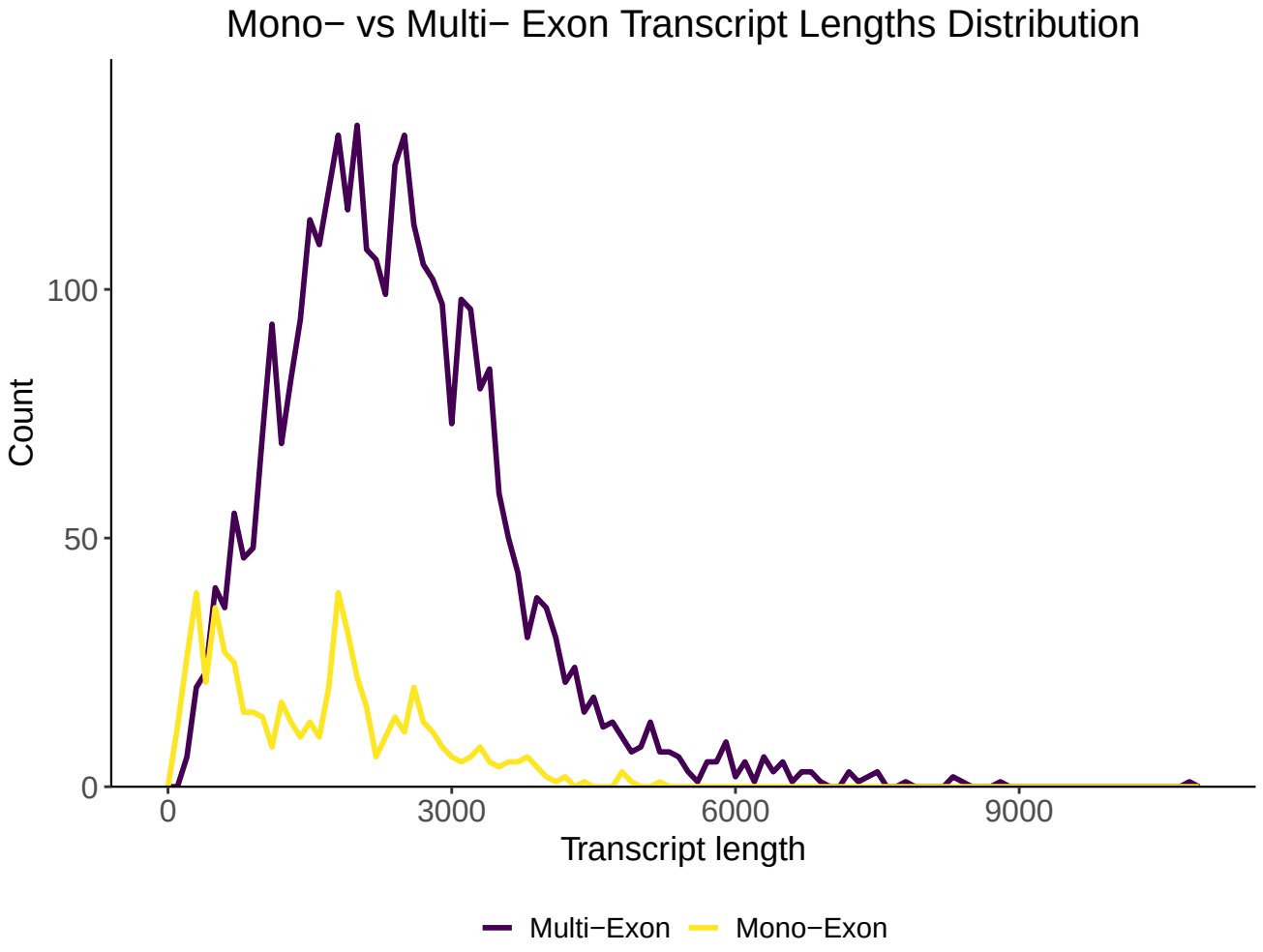

Structural Categories by Transcript Length
100
75
50
25
0
%
0
1
2
3
4
5
6
7
8
10
Transcript length, kb
Structural Category
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron


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

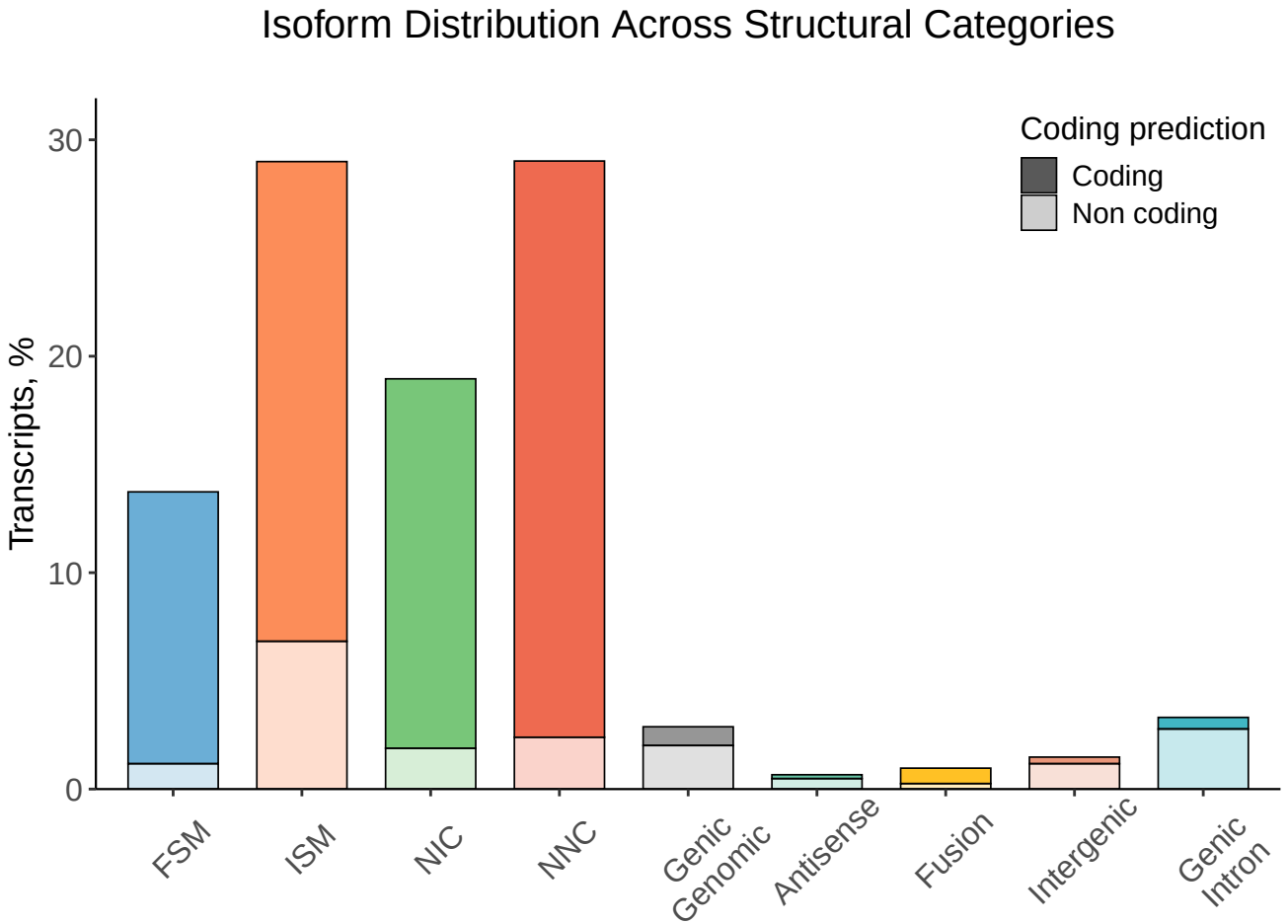

Number of FL reads per Gene by Type of Gene Annotation
log2(FL_TPM+1)
16
12
8
4
Annotated Genes
Novel Genes

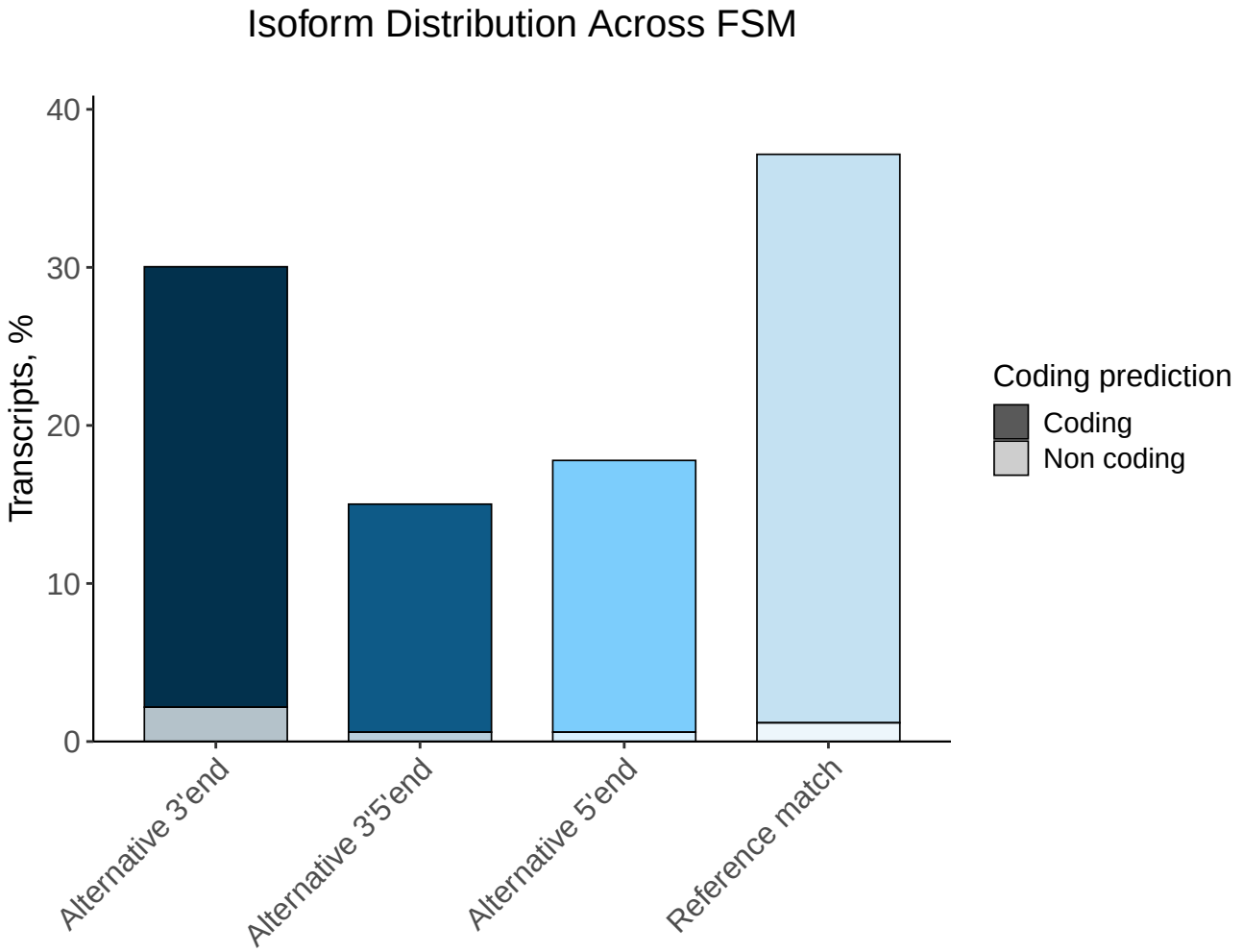

All Transcript Lengths Distribution
Count
150
100
50
0
0
3000
6000
9000
Transcript length

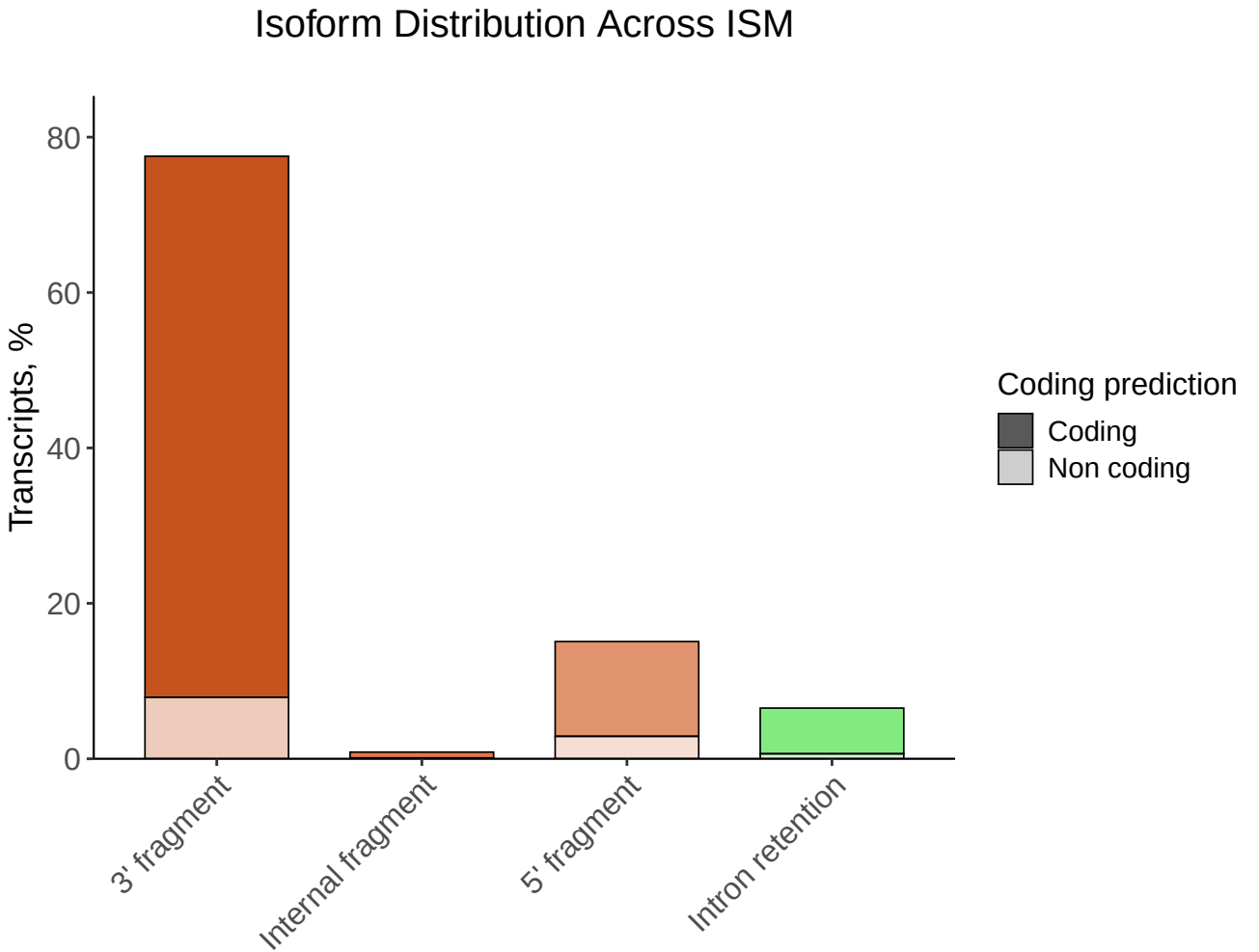

Transcript Lengths Distribution by Structural Category
Count
60
50
40
30
20
10
0
0
3000
6000
9000
Transcript length
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

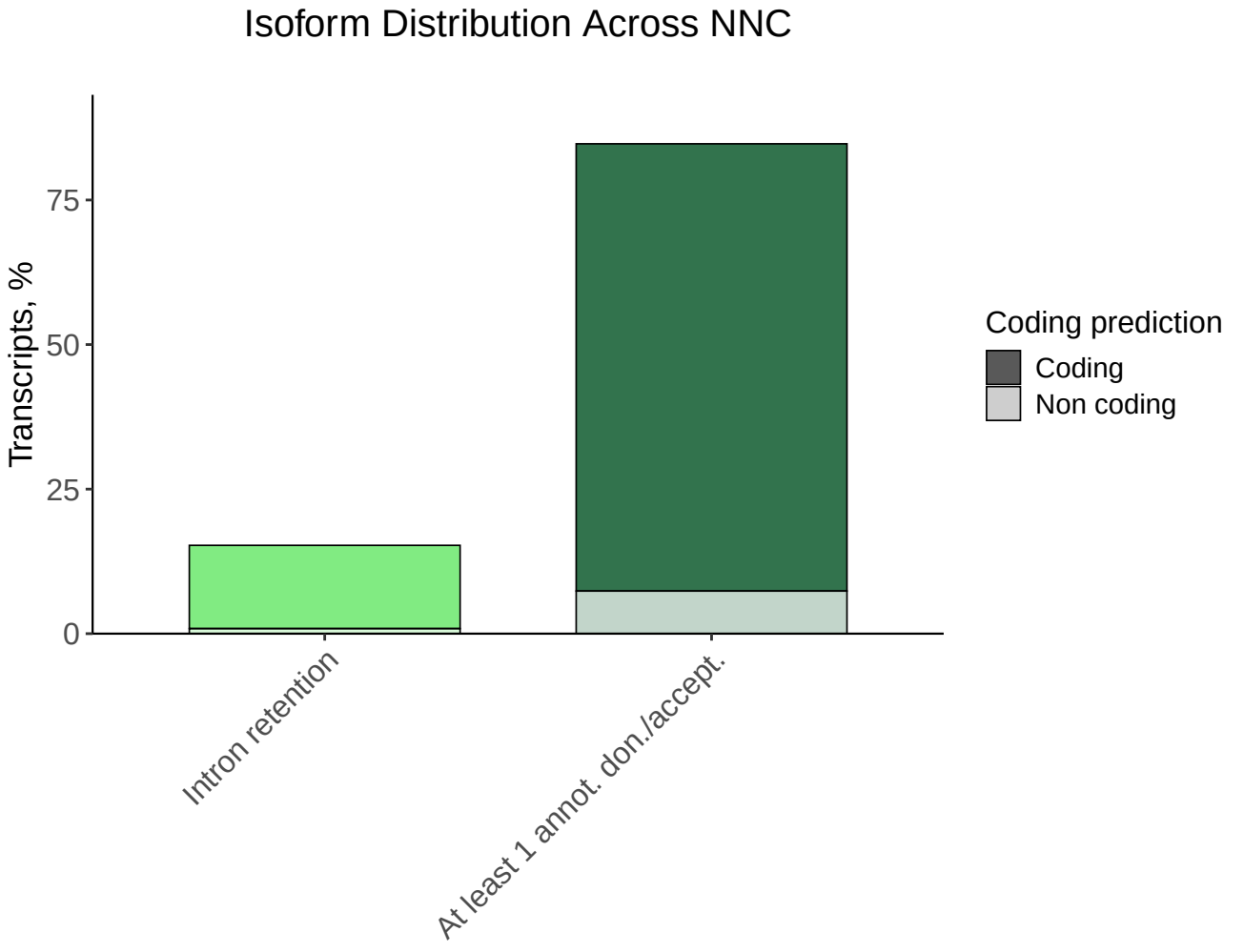

Mono− vs Multi− Exon Transcript Lengths Distribution
Count
100
50
0
0
3000
6000
9000
Transcript length
Multi−Exon
Mono−Exon

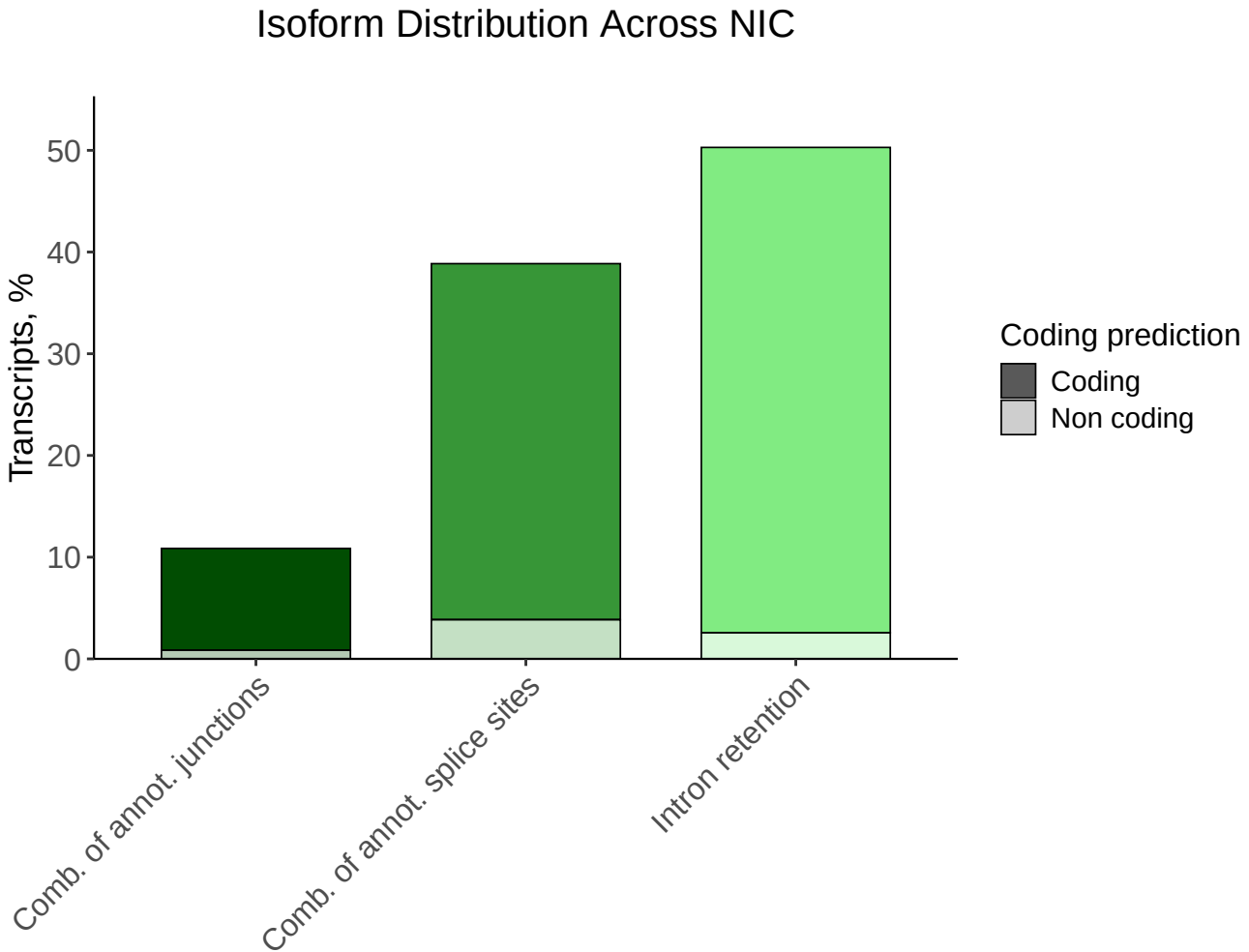
Structural Isoform Characterization

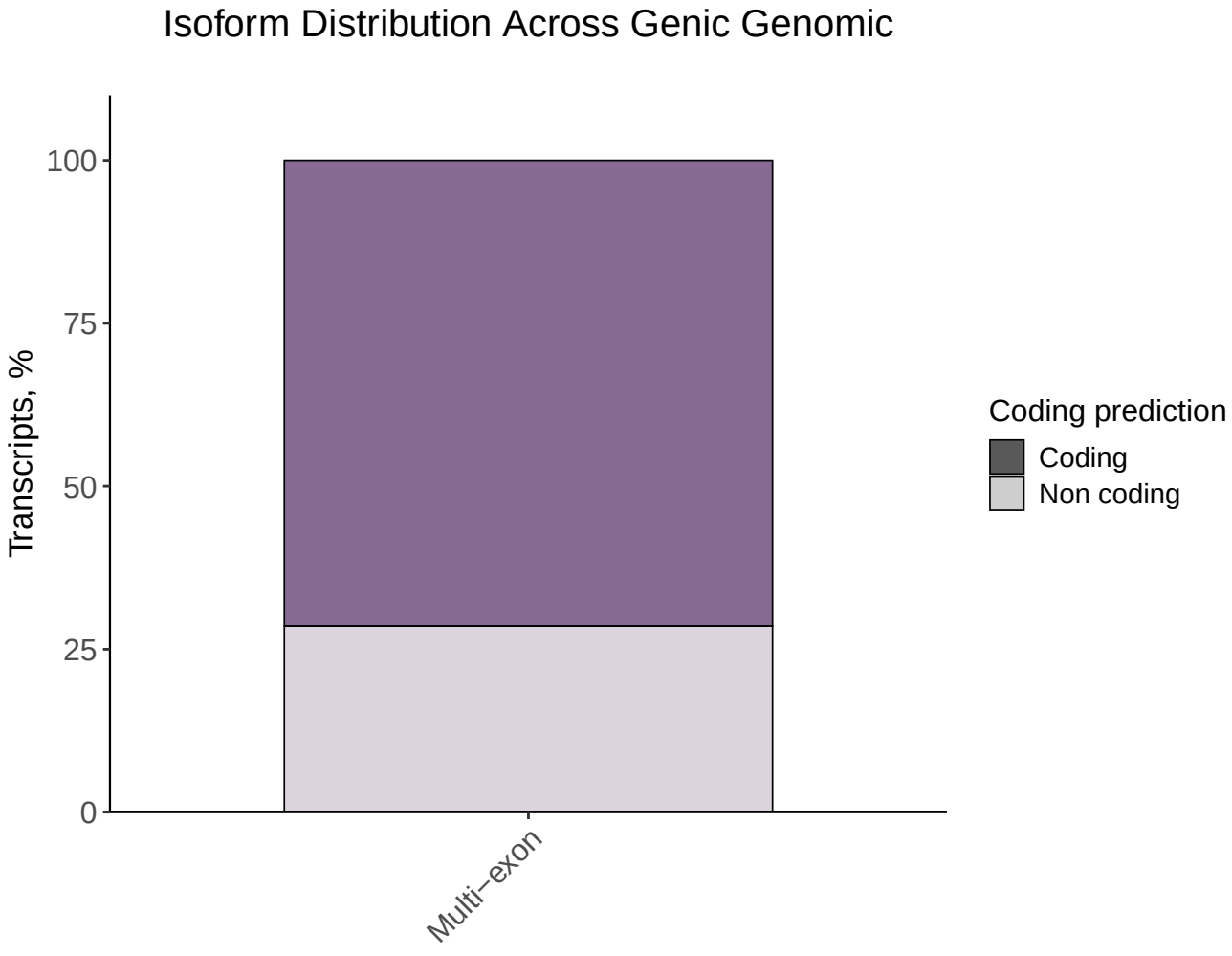

Isoform Distribution Across Structural Categories
Transcripts, %
30
20
10
0
Coding prediction
Coding
Non coding
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

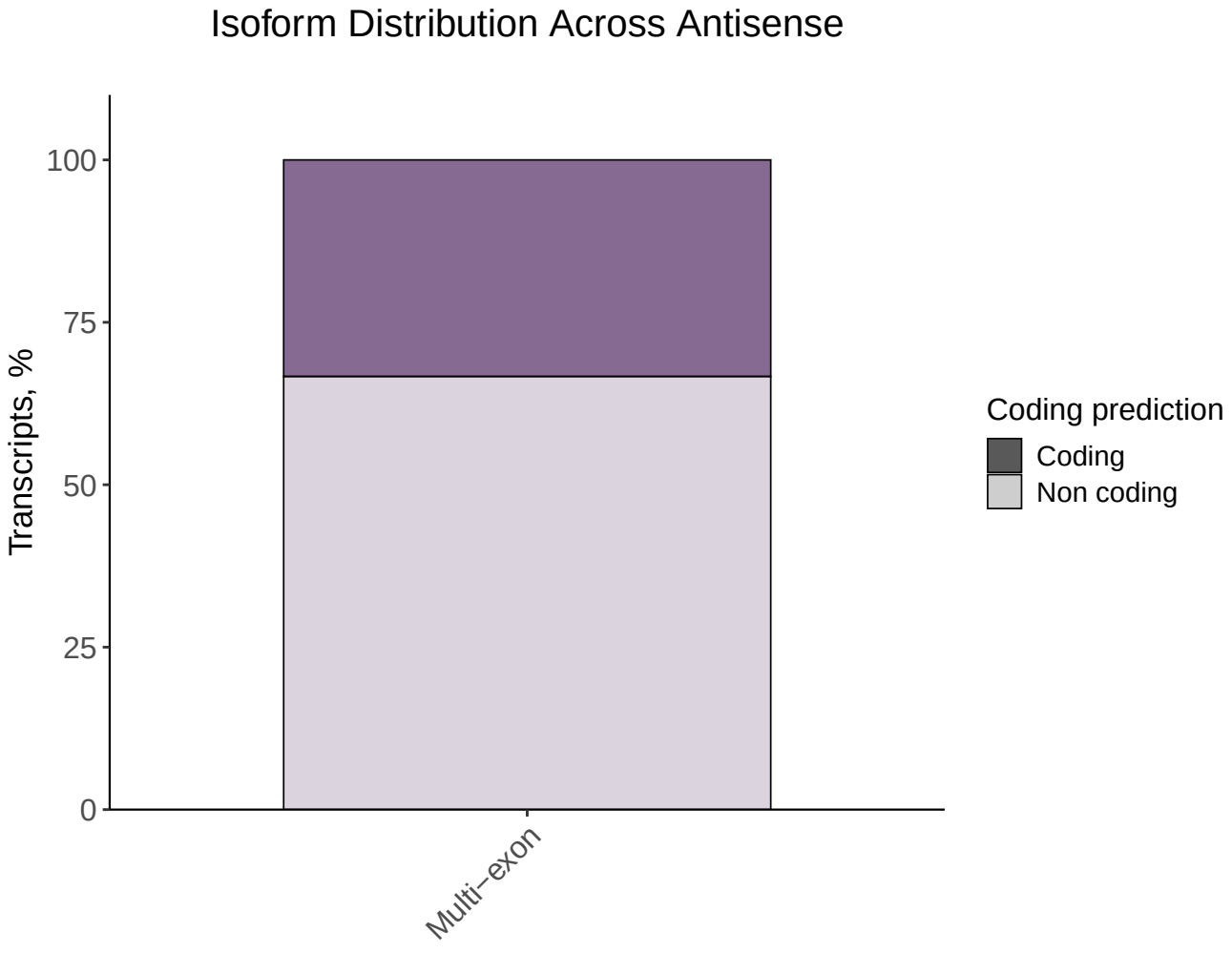

Isoform Distribution Across FSM
Transcripts, %
40
30
20
10
0
Alternative 3'end
Alternative 3'5'end
Alternative 5'end
Reference match
Coding prediction
Coding
Non coding

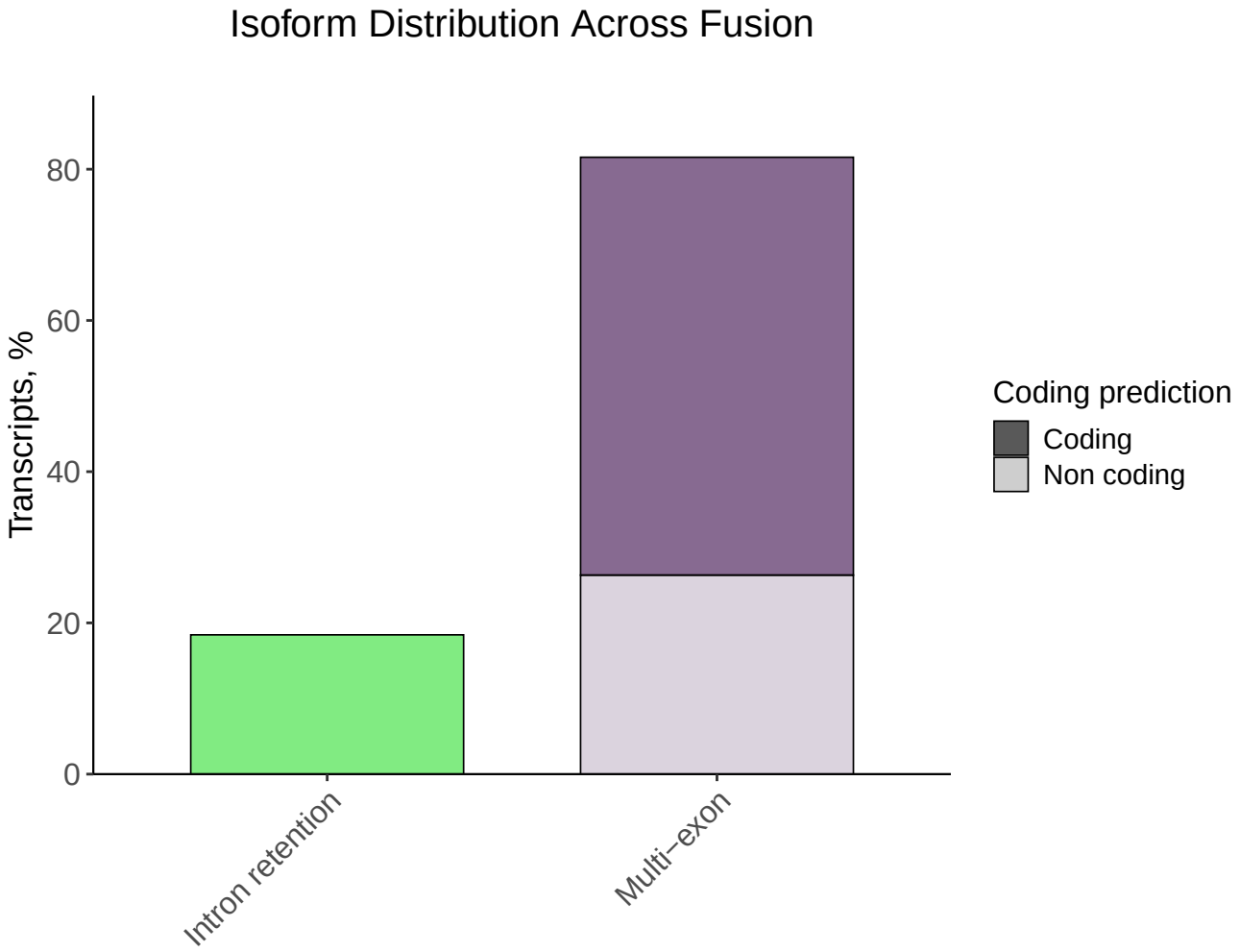

Isoform Distribution Across ISM
Transcripts, %
80
60
40
20
0
3' fragment
Internal fragment
5' fragment
Intron retention
Coding prediction
Coding
Non coding

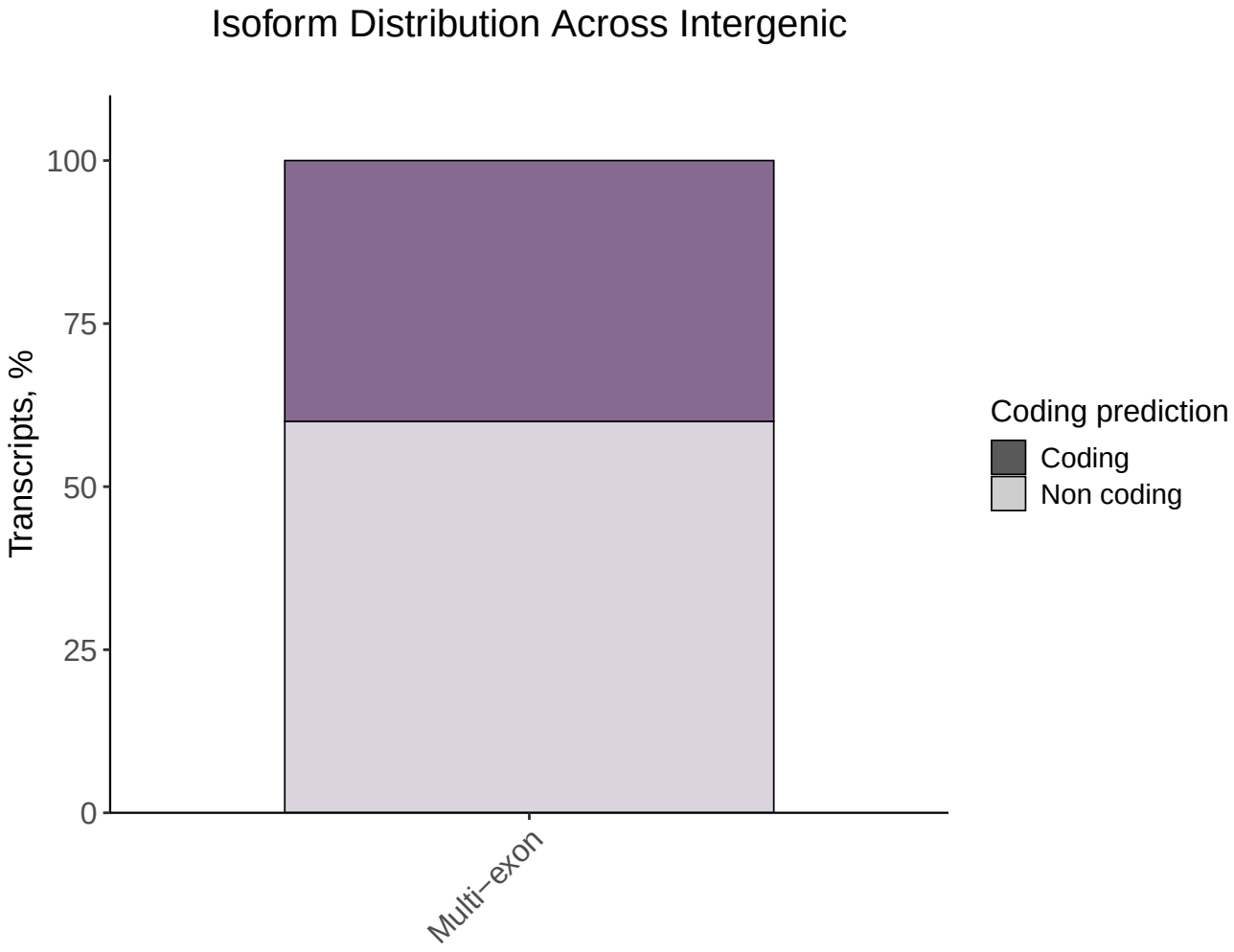

Isoform Distribution Across NNC
Transcripts, %
75
50
25
0
Intron retention
At least 1 annot. don./accept.
Coding prediction
Coding
Non coding

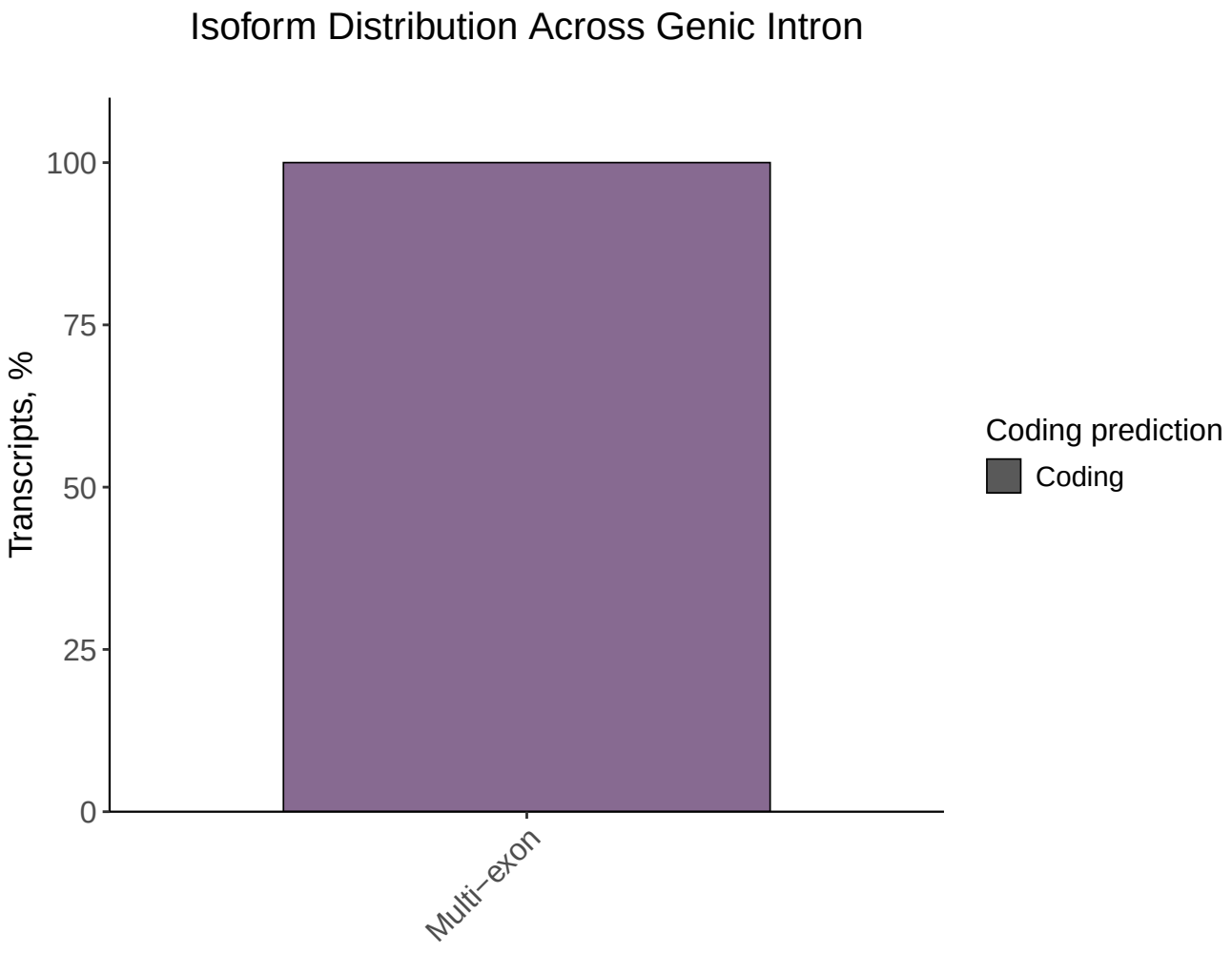

Isoform Distribution Across NIC
Transcripts, %
0
10
20
30
40
50
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Coding prediction
Coding
Non coding

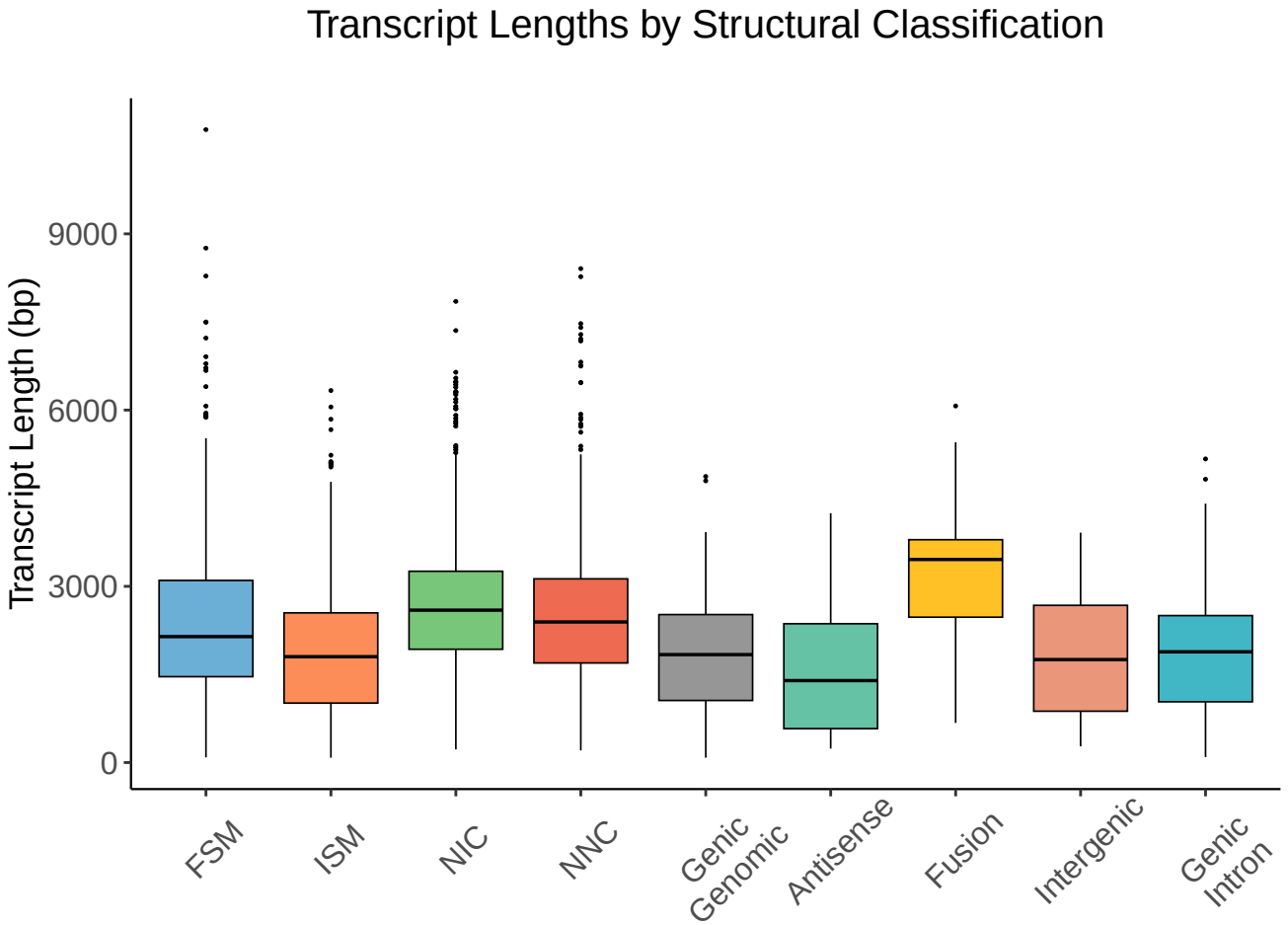

Isoform Distribution Across Genic Genomic
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding
Non coding

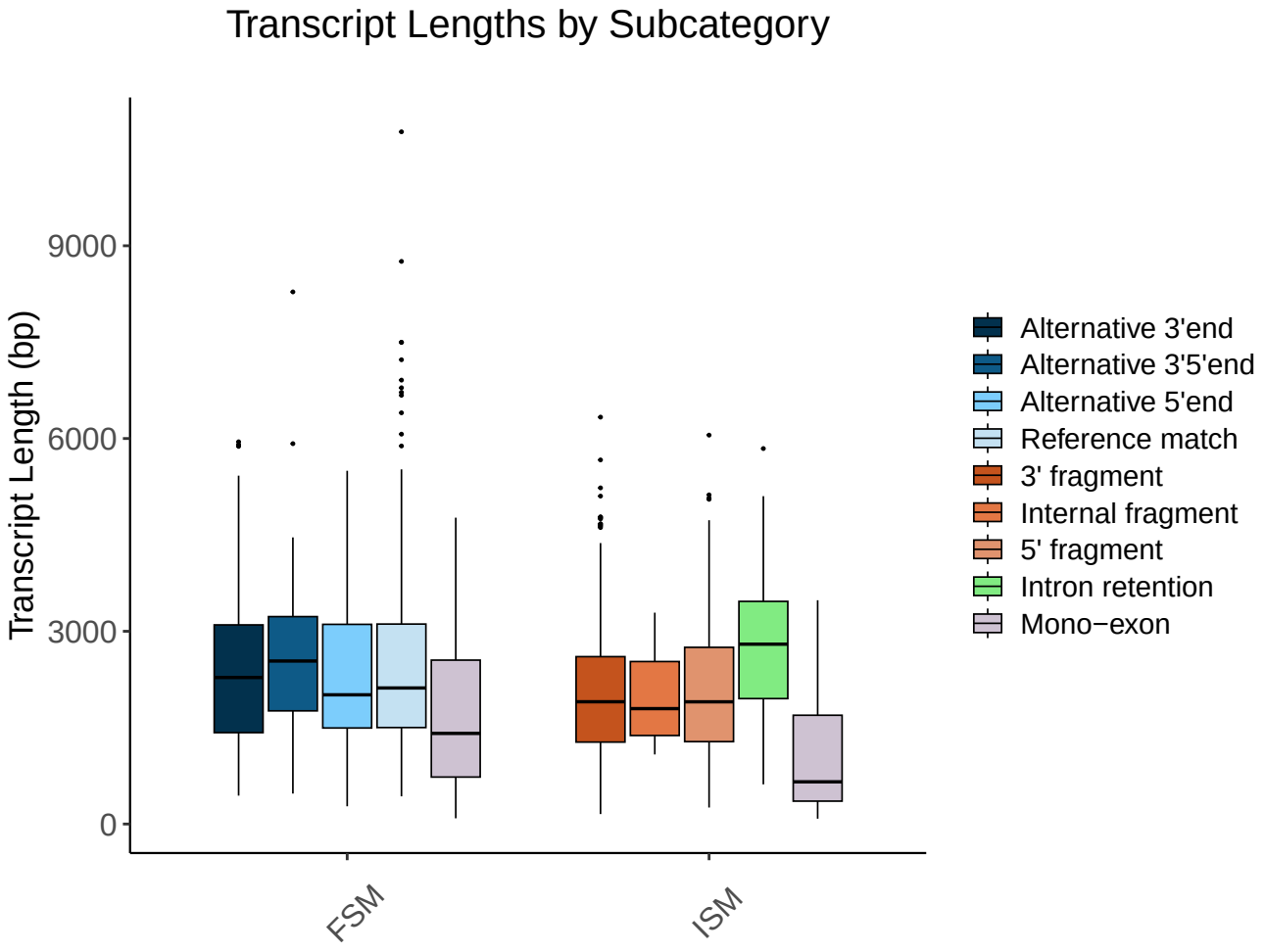

Isoform Distribution Across Antisense
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding
Non coding

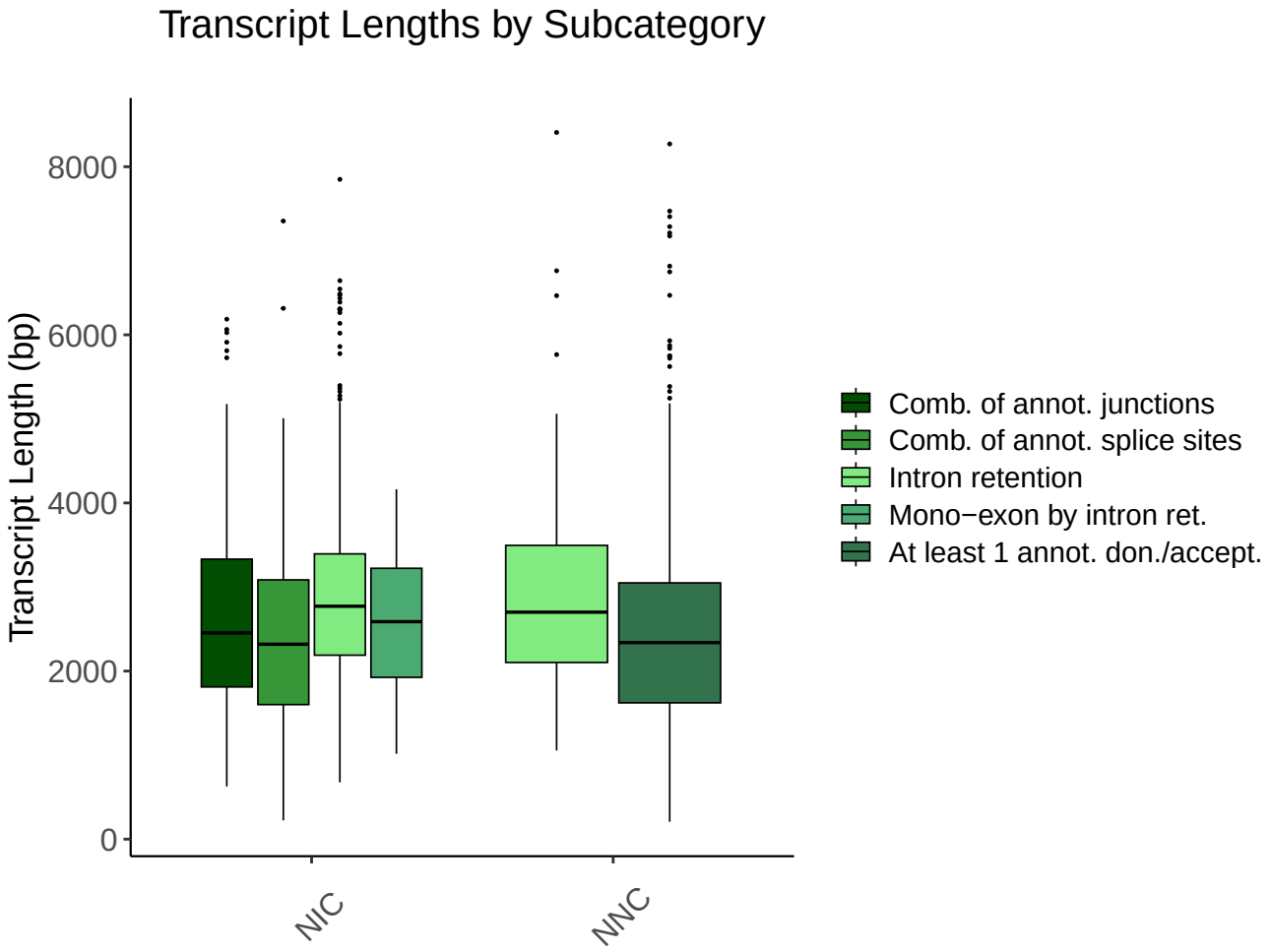

Isoform Distribution Across Fusion
Transcripts, %
80
60
40
20
0
Intron retention
Multi−exon
Coding prediction
Coding
Non coding

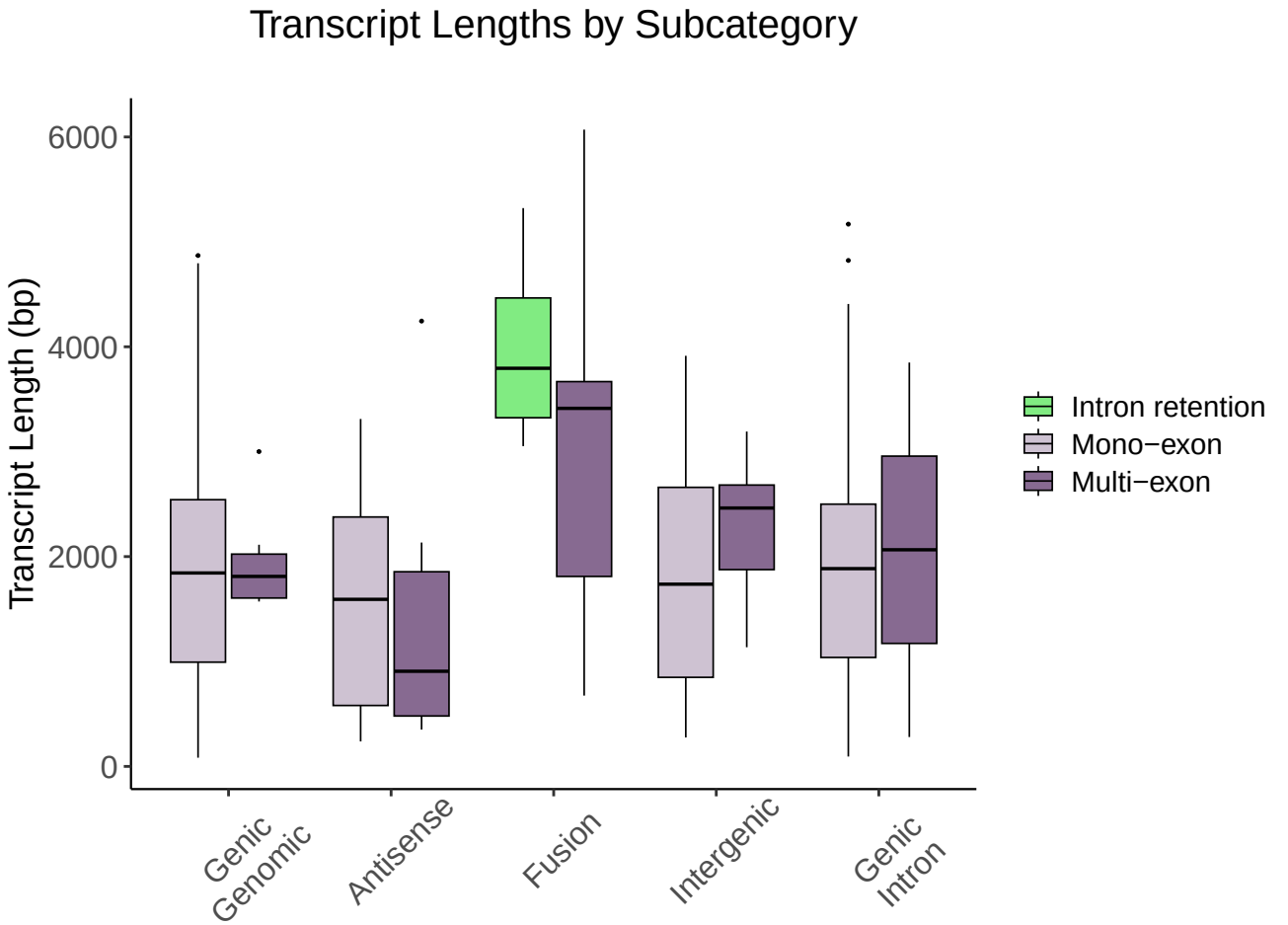

Isoform Distribution Across Intergenic
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding
Non coding

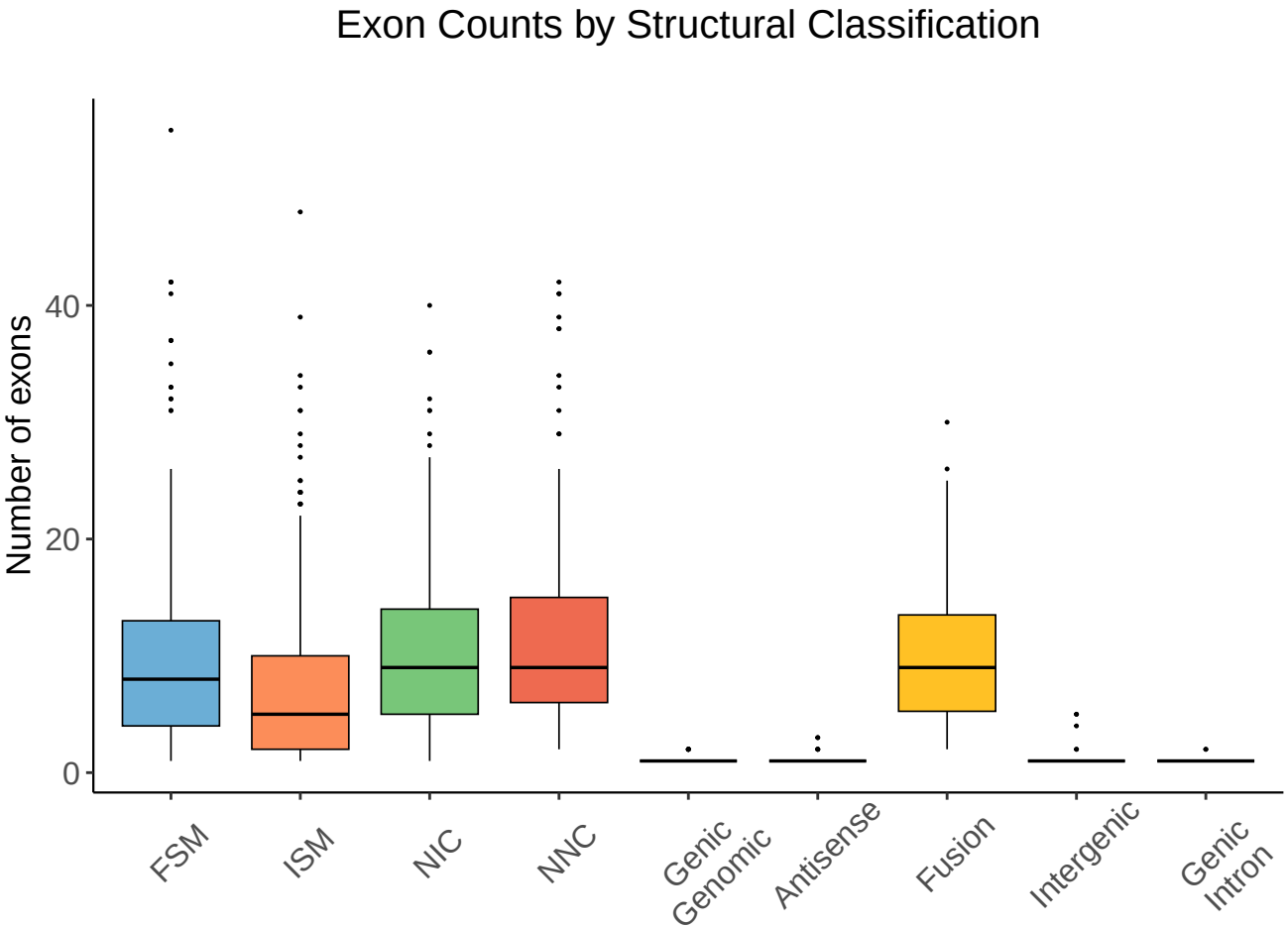

Isoform Distribution Across Genic Intron
Transcripts, %
100
75
50
25
0
Multi−exon
Coding prediction
Coding

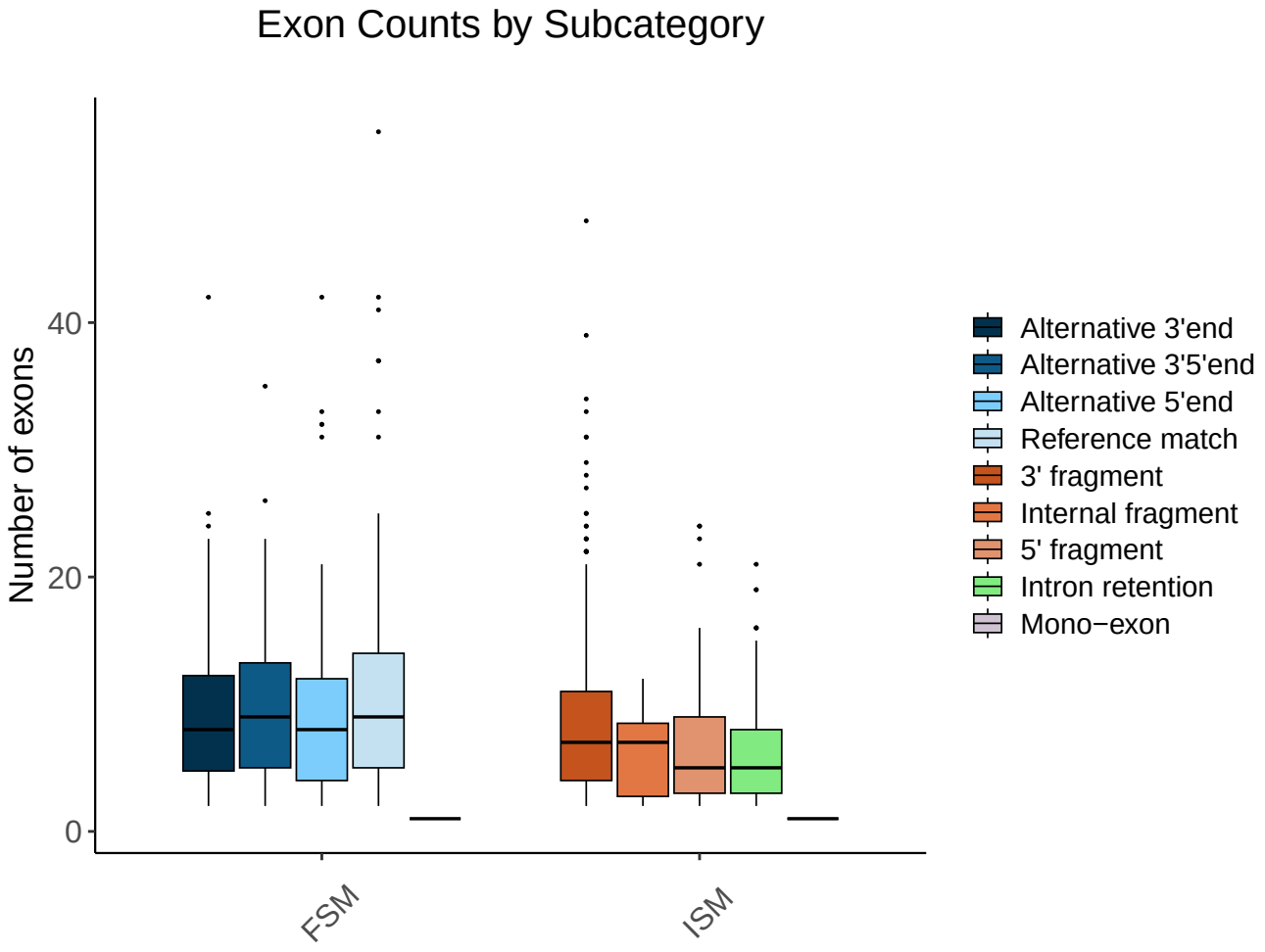

Transcript Lengths by Structural Classification
Transcript Length (bp)
9000
6000
3000
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

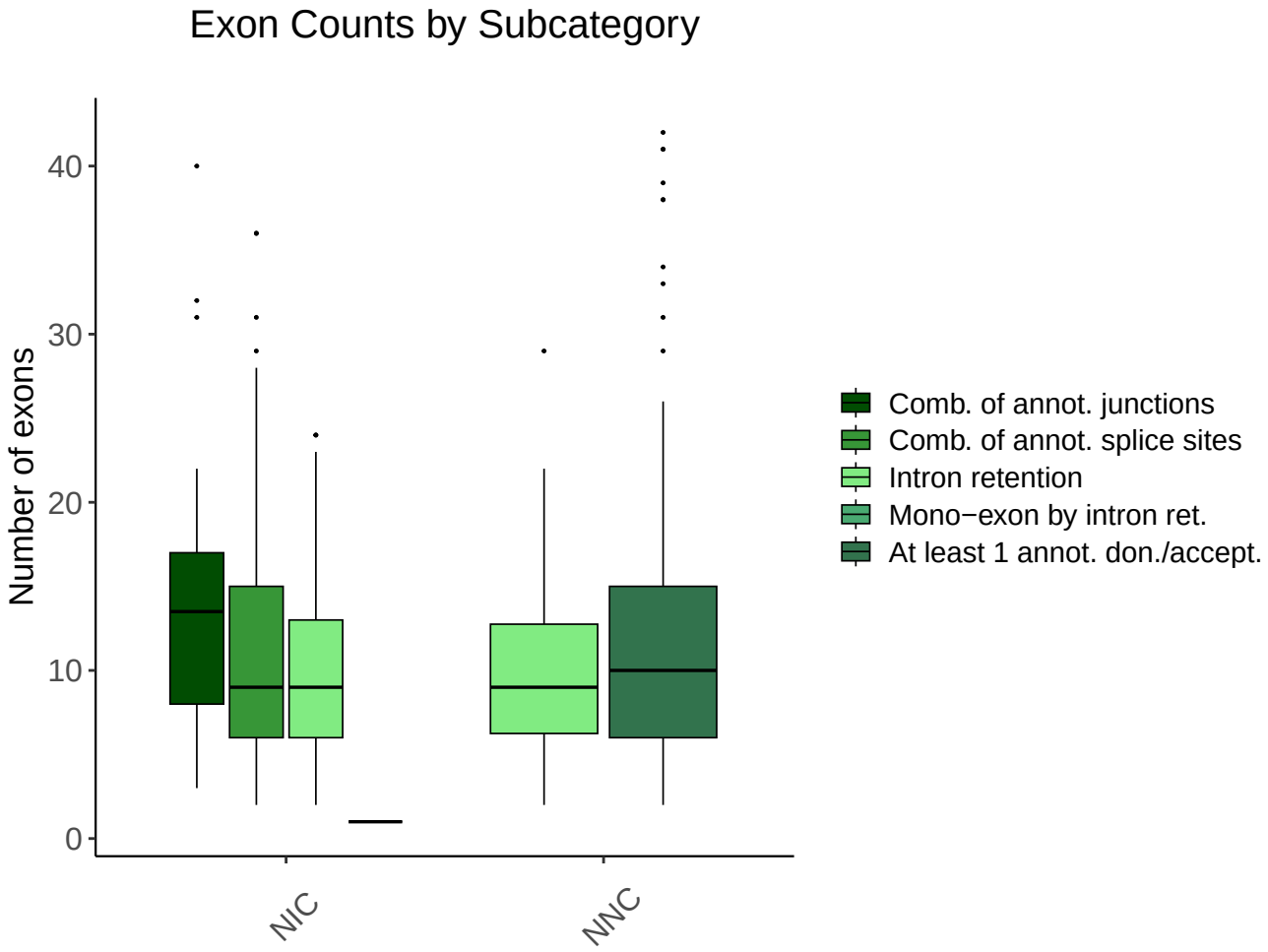

Transcript Lengths by Subcategory
Transcript Length (bp)
9000
6000
3000
0
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alternative 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon

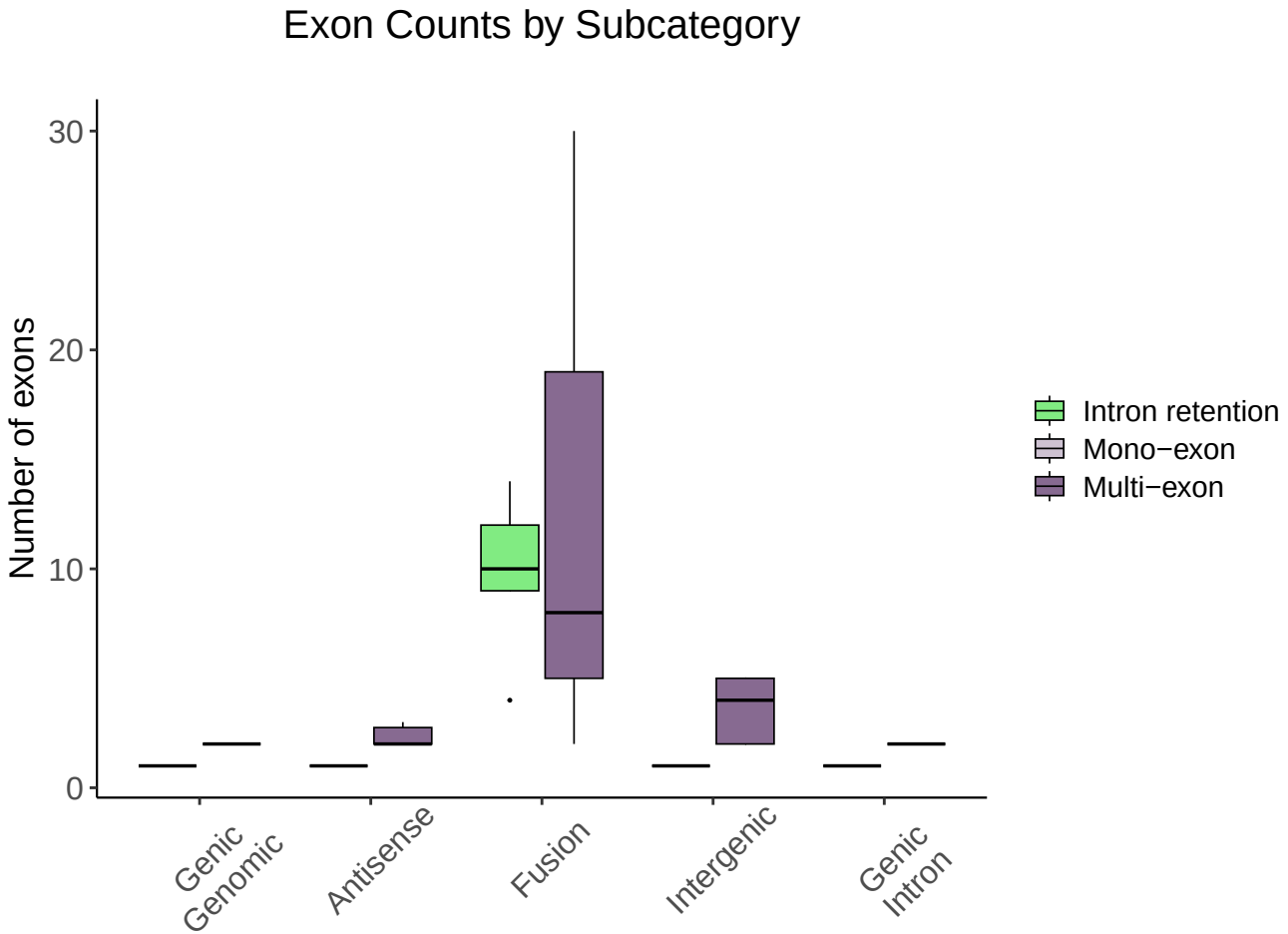

Transcript Lengths by Subcategory
Transcript Length (bp)
8000
6000
4000
2000
0
NIC
NNC
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
At least 1 annot. don./accept.

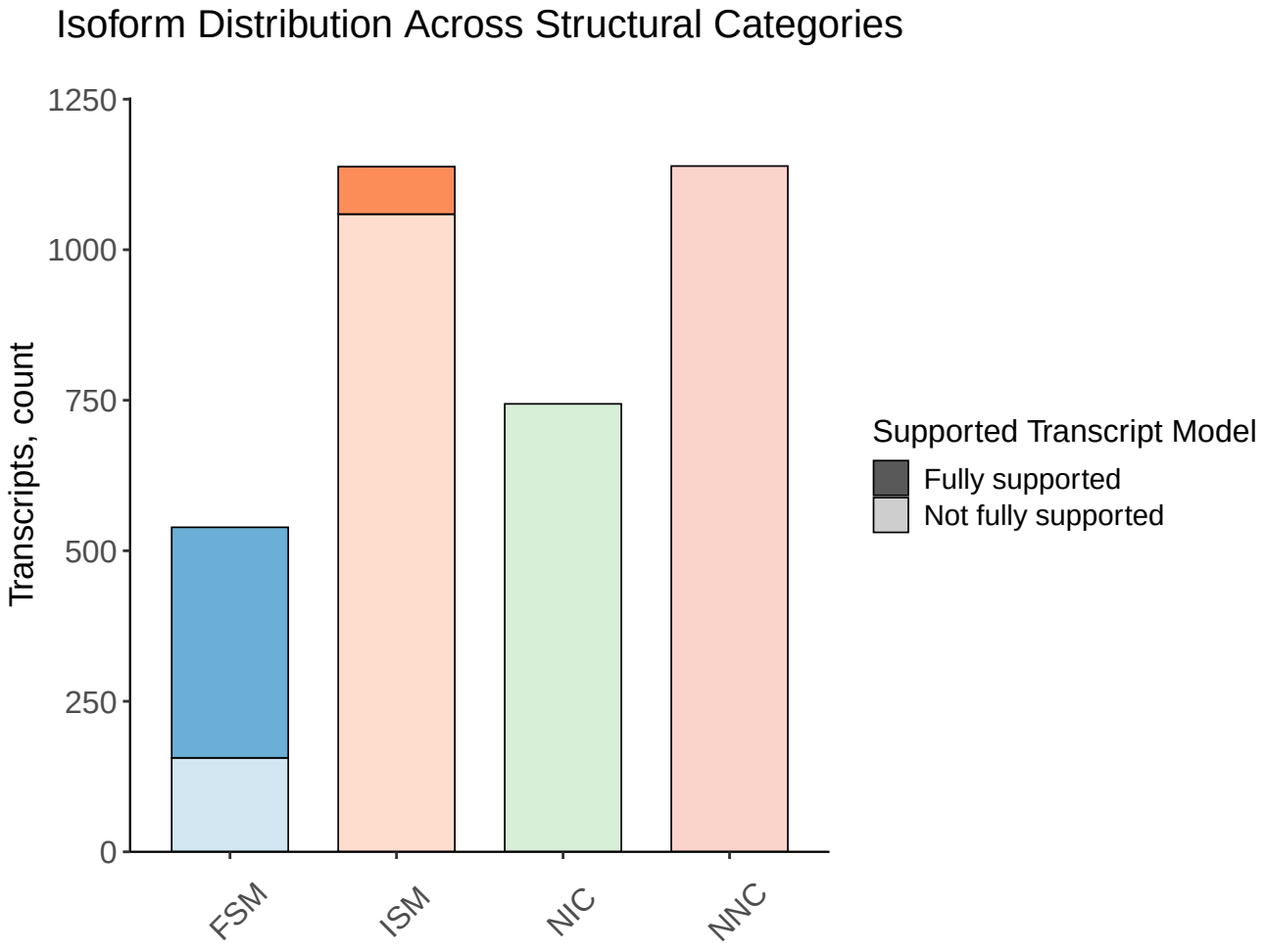

Transcript Lengths by Subcategory
Transcript Length (bp)
6000
4000
2000
0
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Intron retention
Mono−exon
Multi−exon

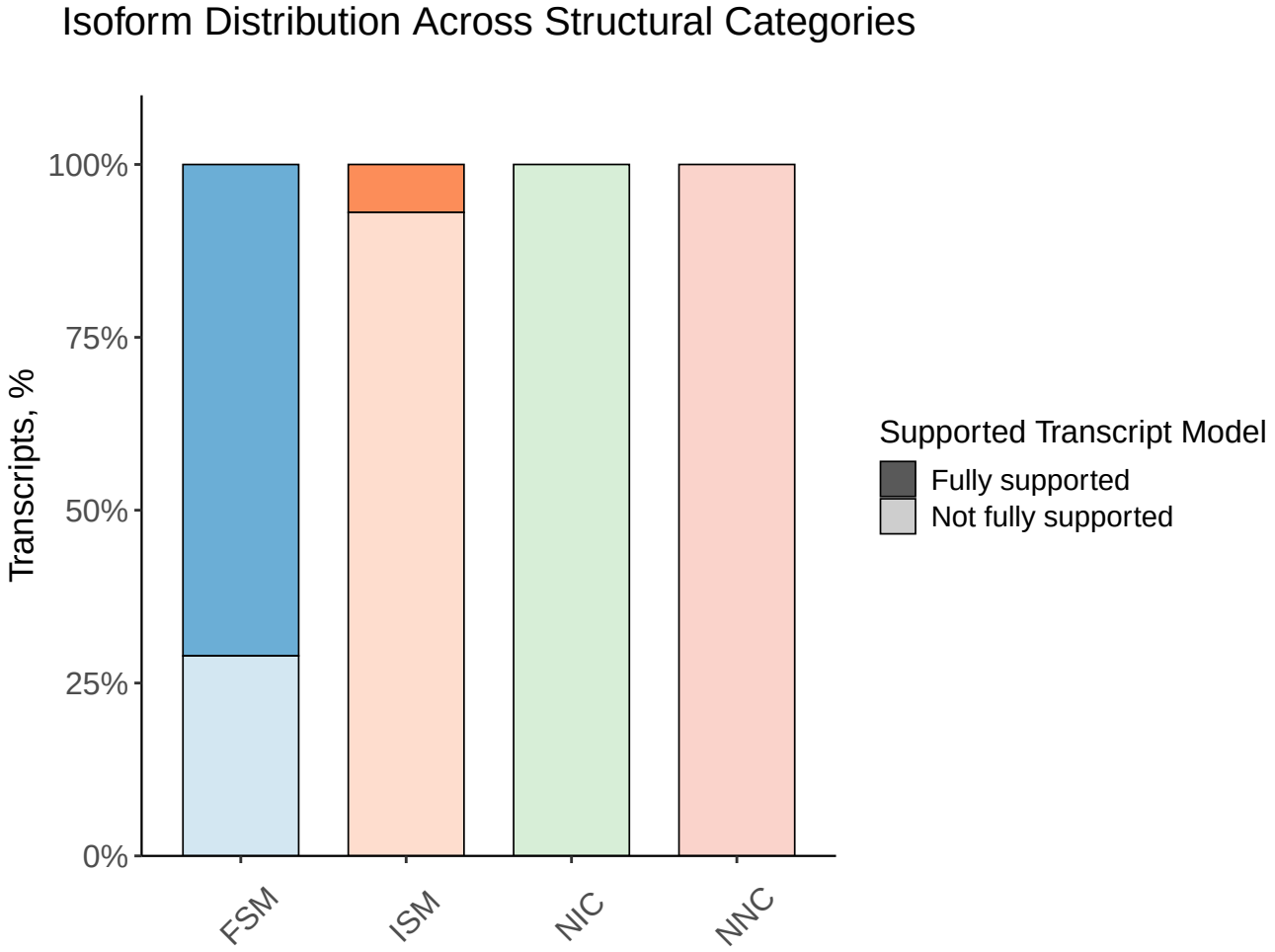

Exon Counts by Structural Classification
Number of exons
0
20
40
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron


Exon Counts by Subcategory
Number of exons
0
20
40
FSM
ISM
Alternative 3'end
Alternative 3'5'end
Alternative 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon

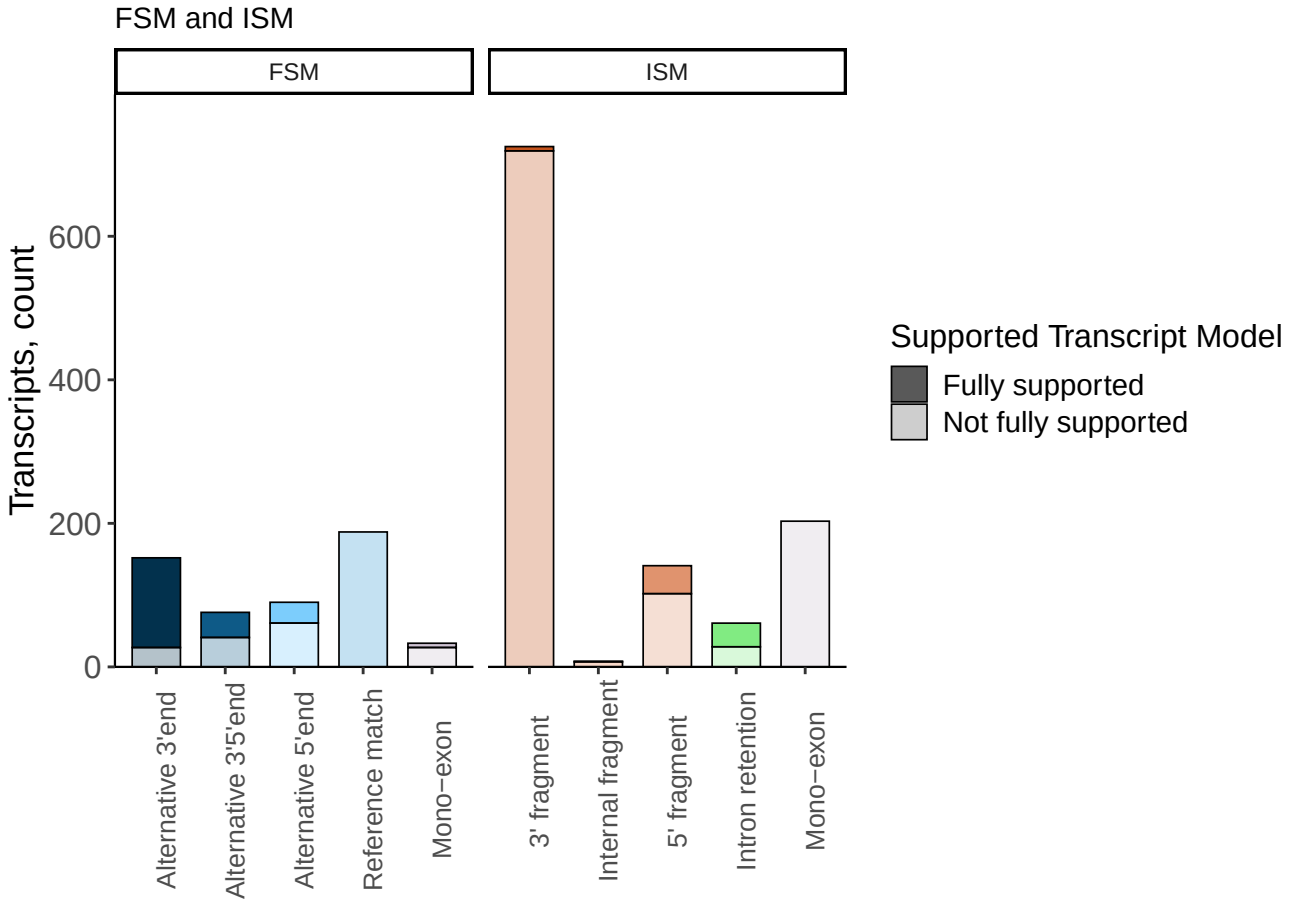

Exon Counts by Subcategory
Number of exons
0
10
20
30
40
NIC
NNC
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
At least 1 annot. don./accept.


Exon Counts by Subcategory
Number of exons
30
20
10
0
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Intron retention
Mono−exon
Multi−exon

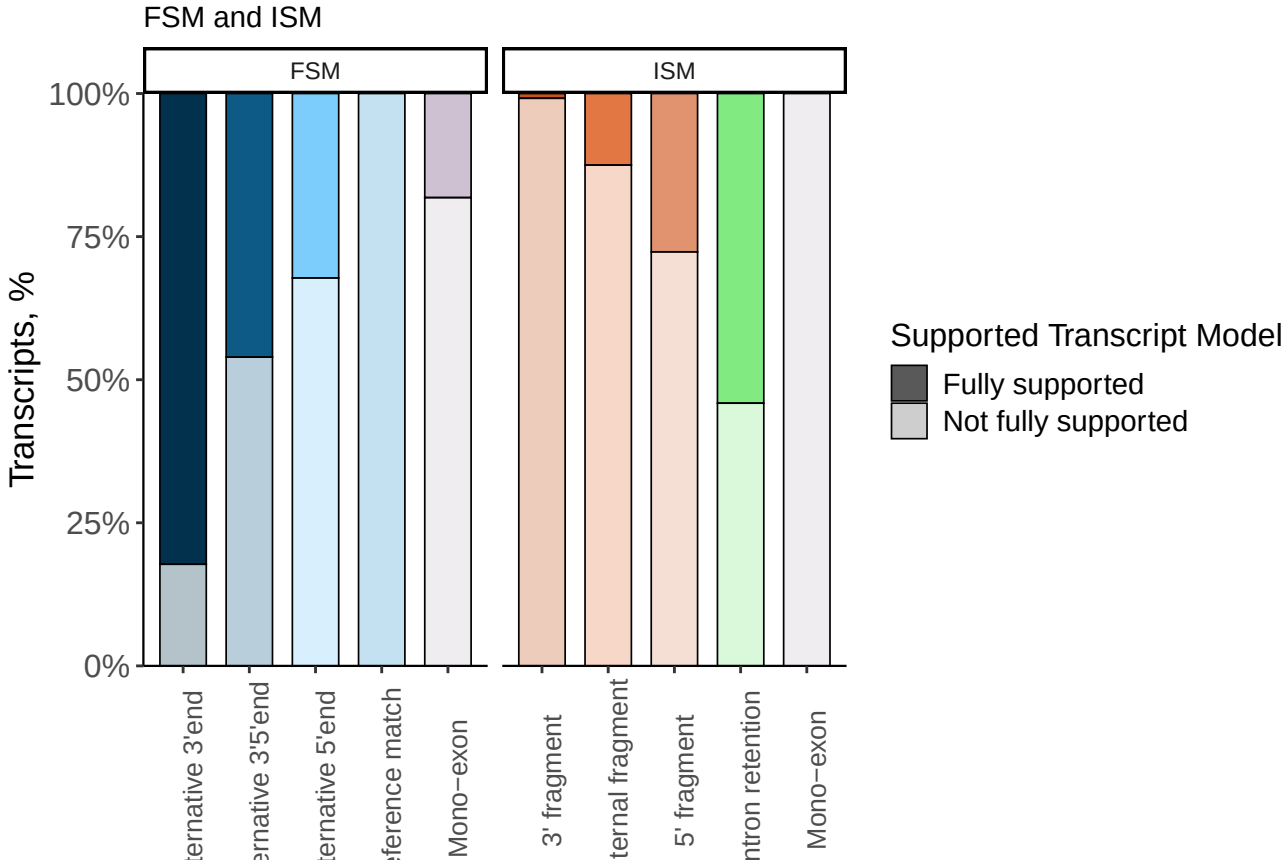

Isoform Distribution Across Structural Categories
Transcripts, count
1250
1000
750
500
250
0
FSM
ISM
NIC
NNC
Supported Transcript Model
Fully supported
Not fully supported


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

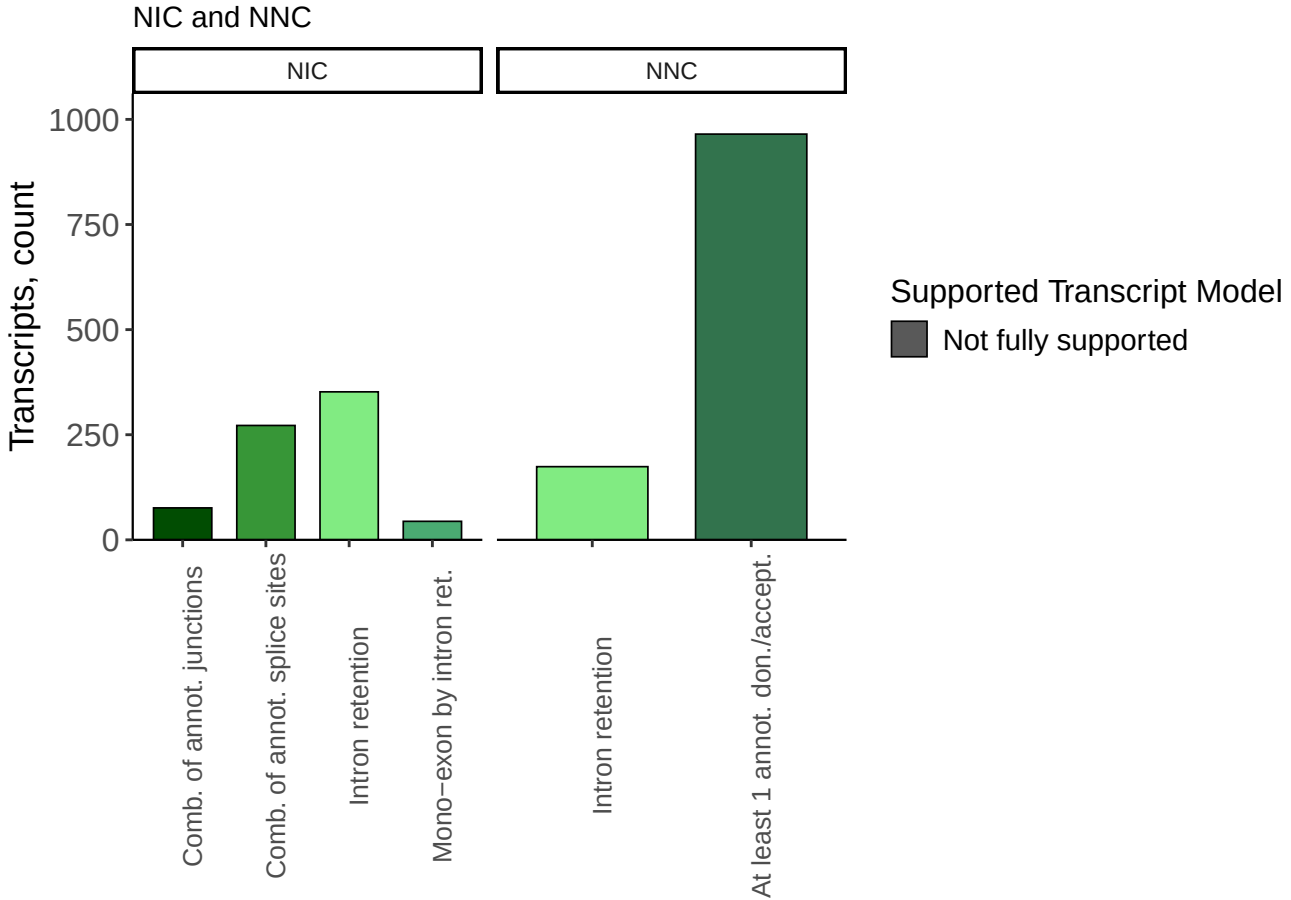
Isoform Distribution Across Structural Subcategories



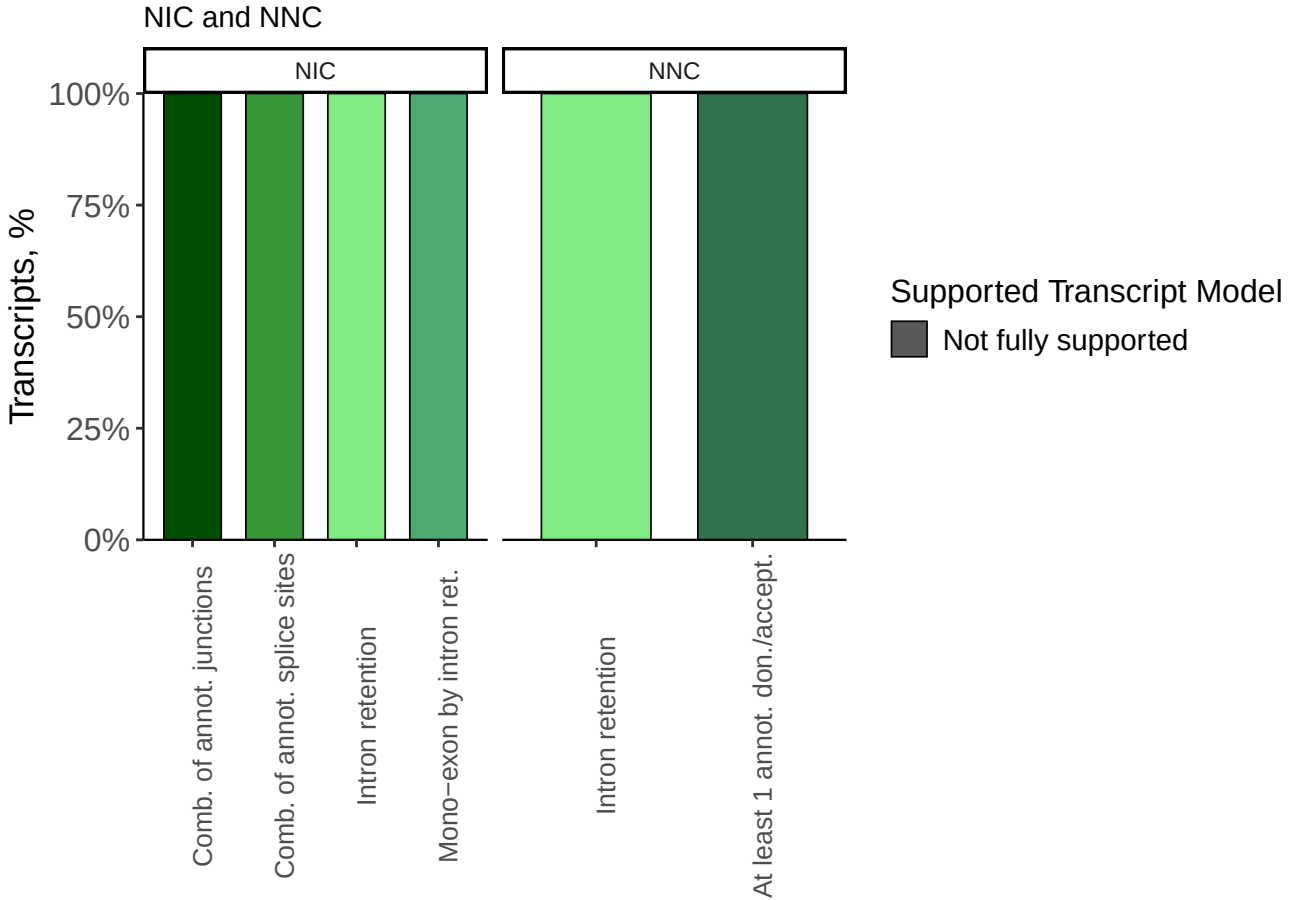
Isoform Distribution Across Structural Subcategories

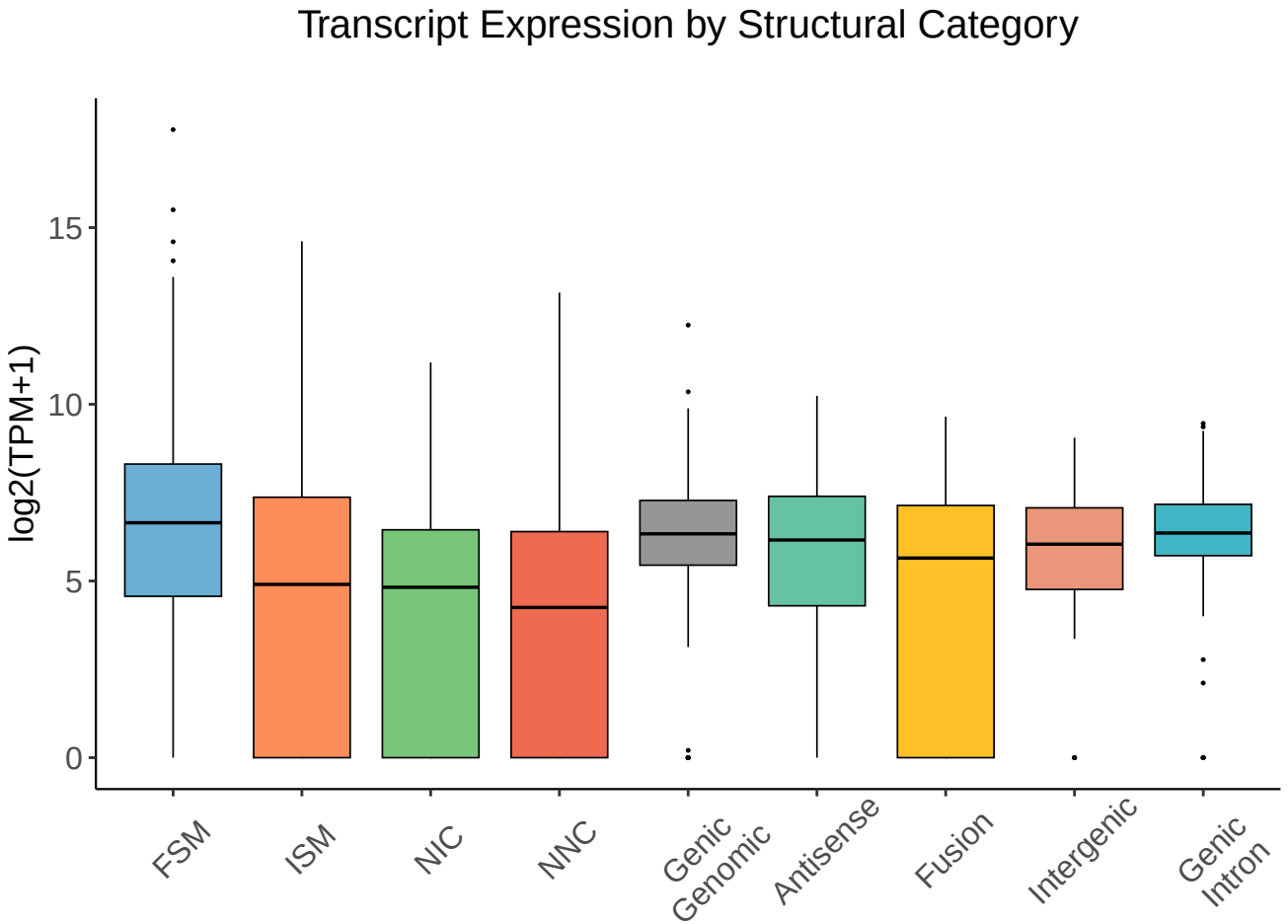


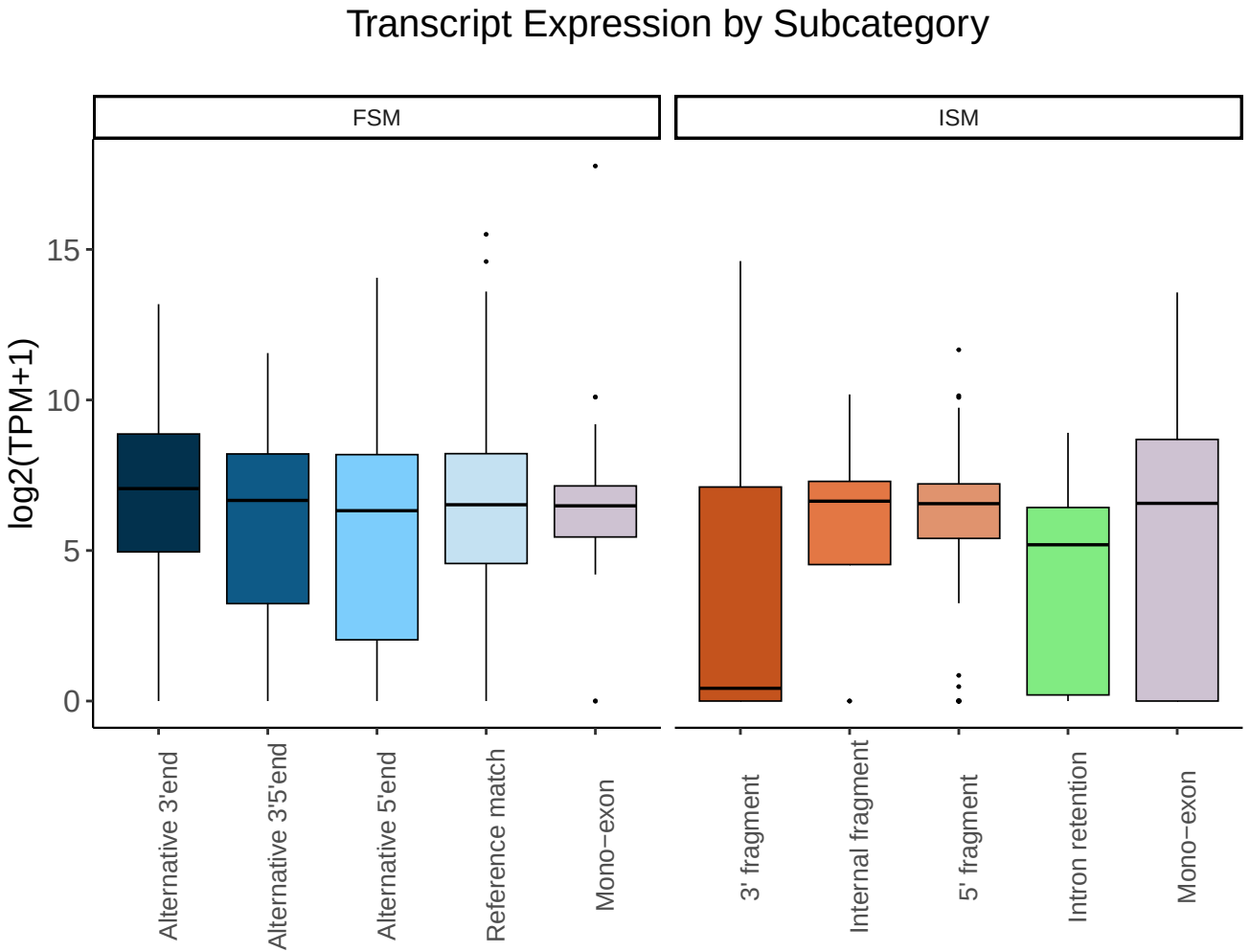
Isoform Distribution Across Structural Subcategories

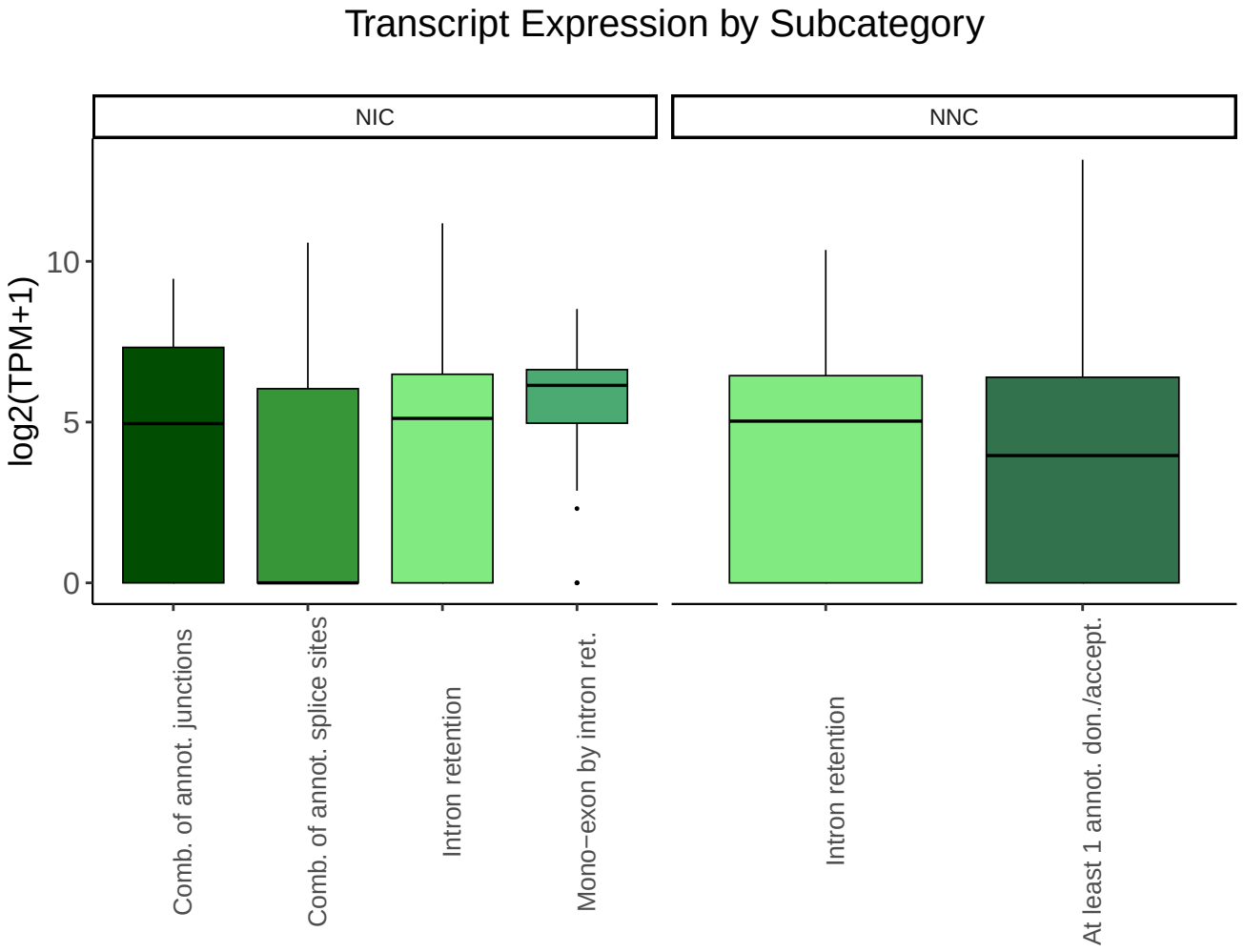


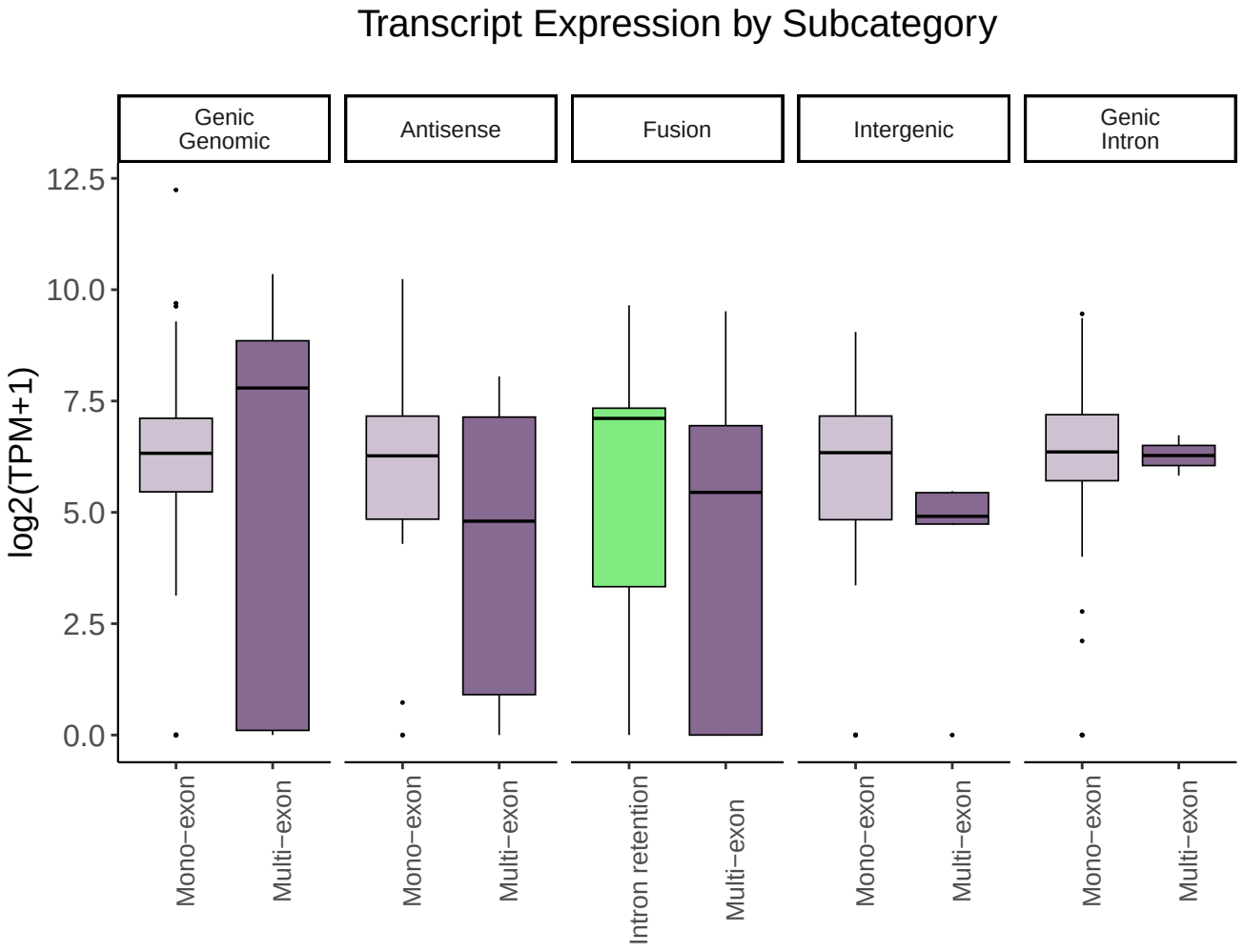
Isoform Distribution Across Structural Subcategories

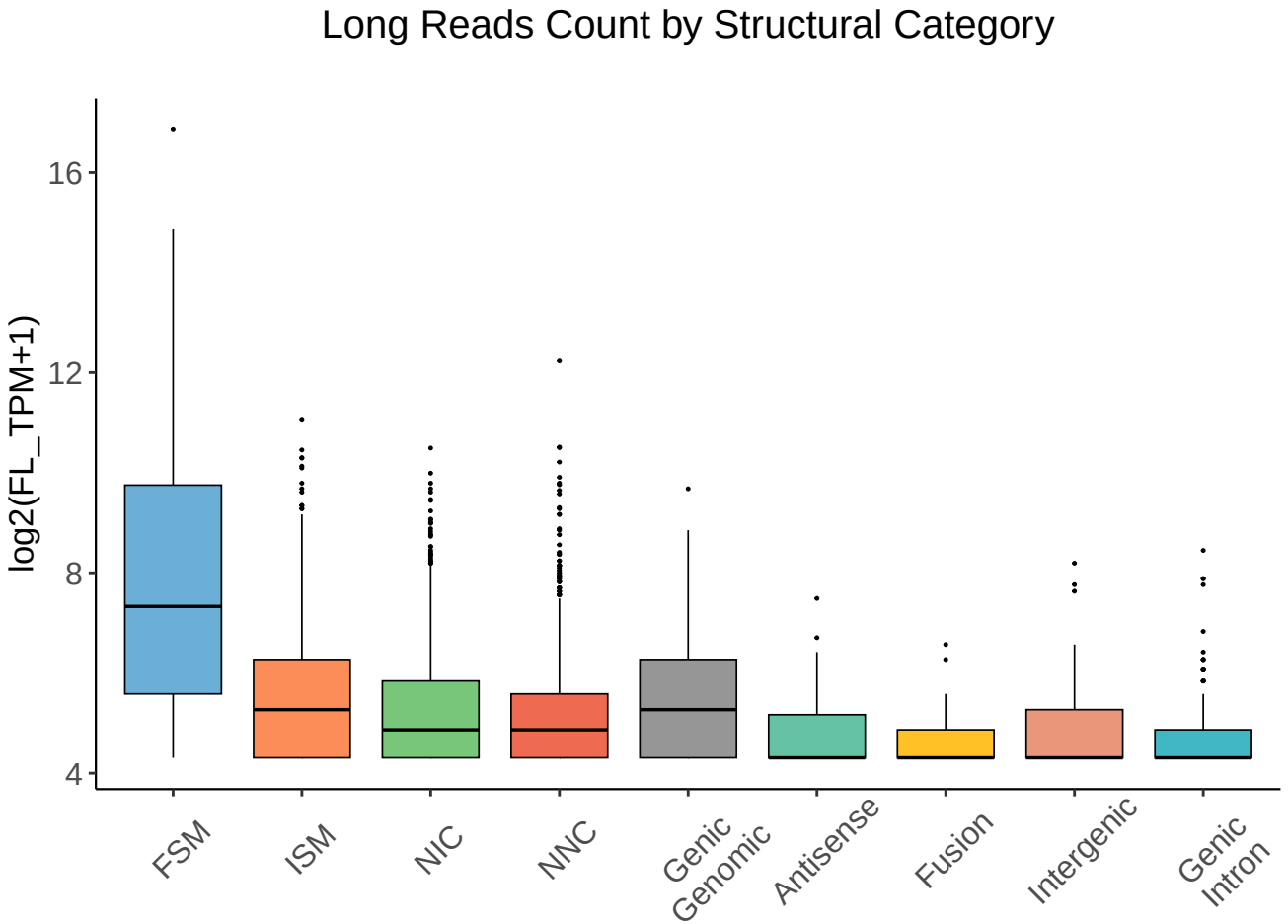


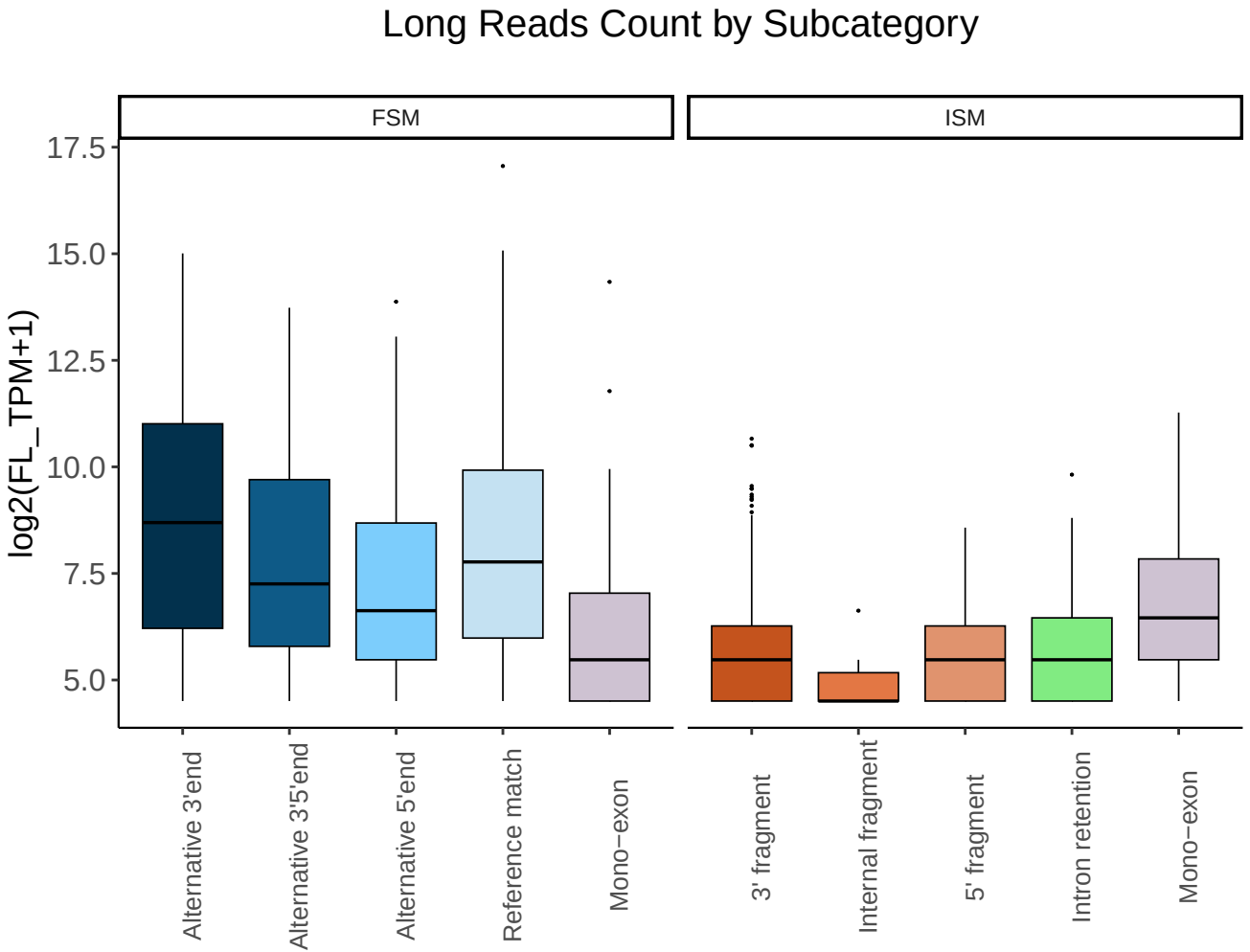

Transcript Expression by Structural Category
log2(TPM+1)
15
10
5
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

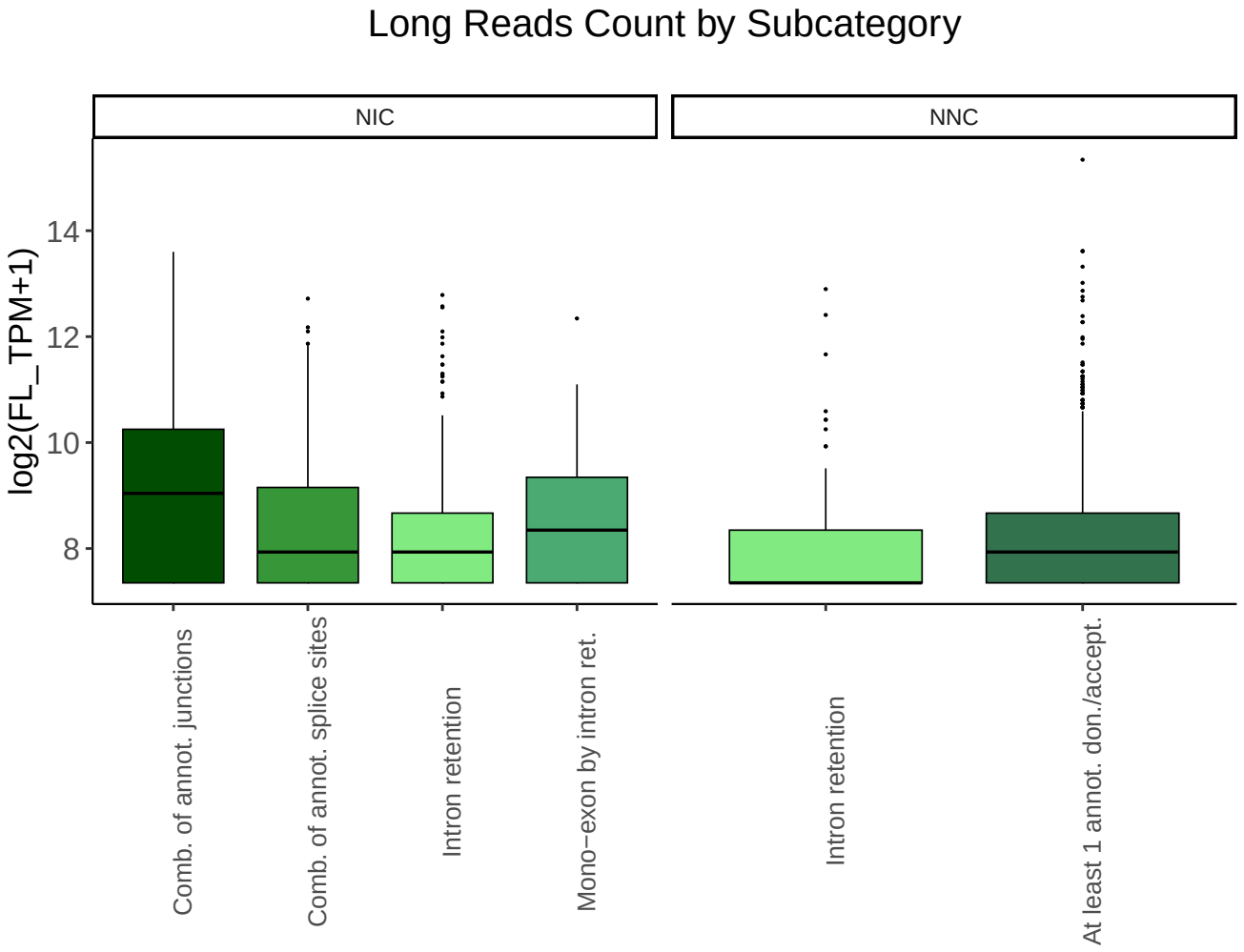

Transcript Expression by Subcategory
FSM
ISM
log2(TPM+1)
15
10
5
0
Alternative 3'end
Alternative 3'5'end
Alternative 5'end
Reference match
Mono−exon
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon

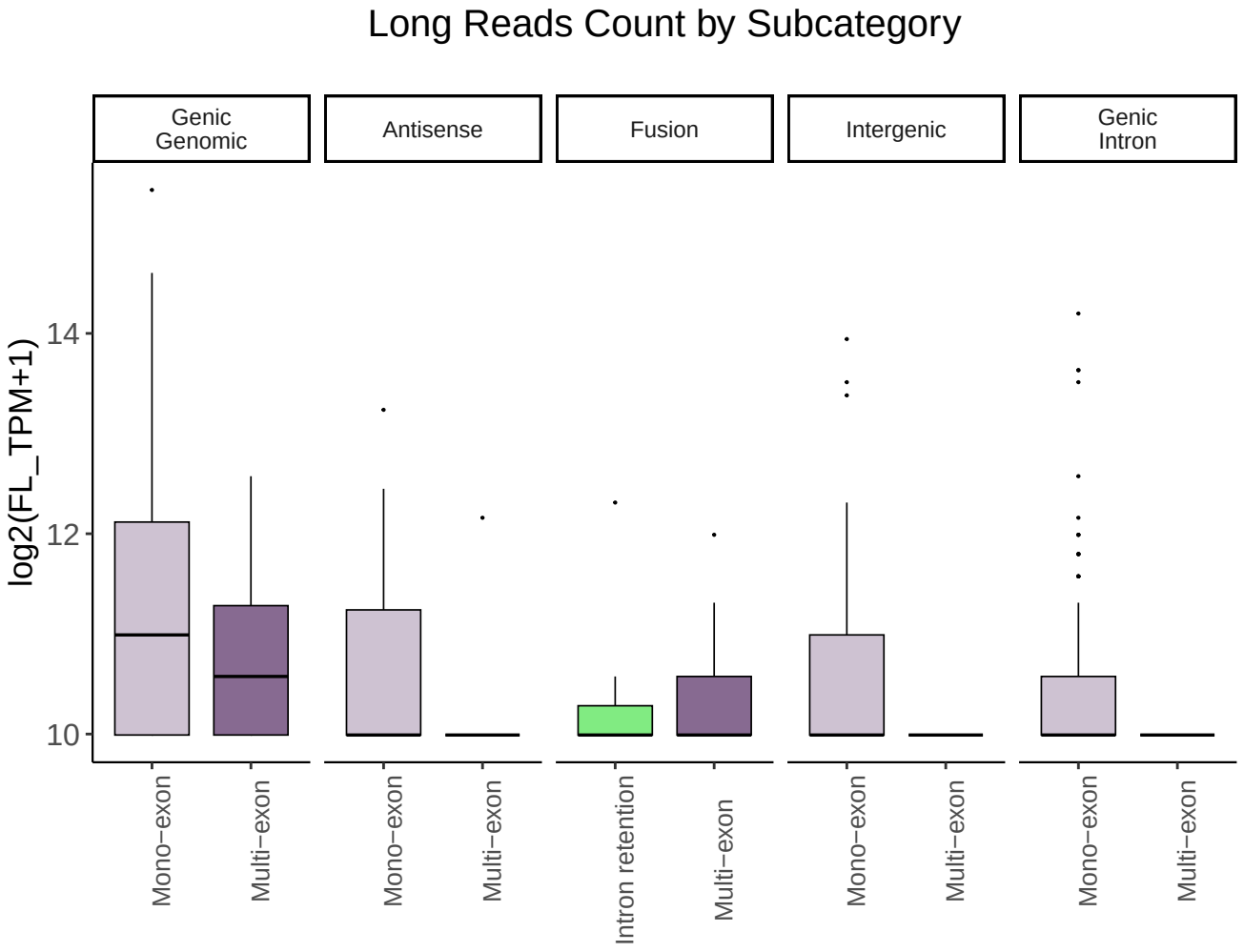

Transcript Expression by Subcategory
NIC
NNC
log2(TPM+1)
10
5
0
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
Intron retention
At least 1 annot. don./accept.

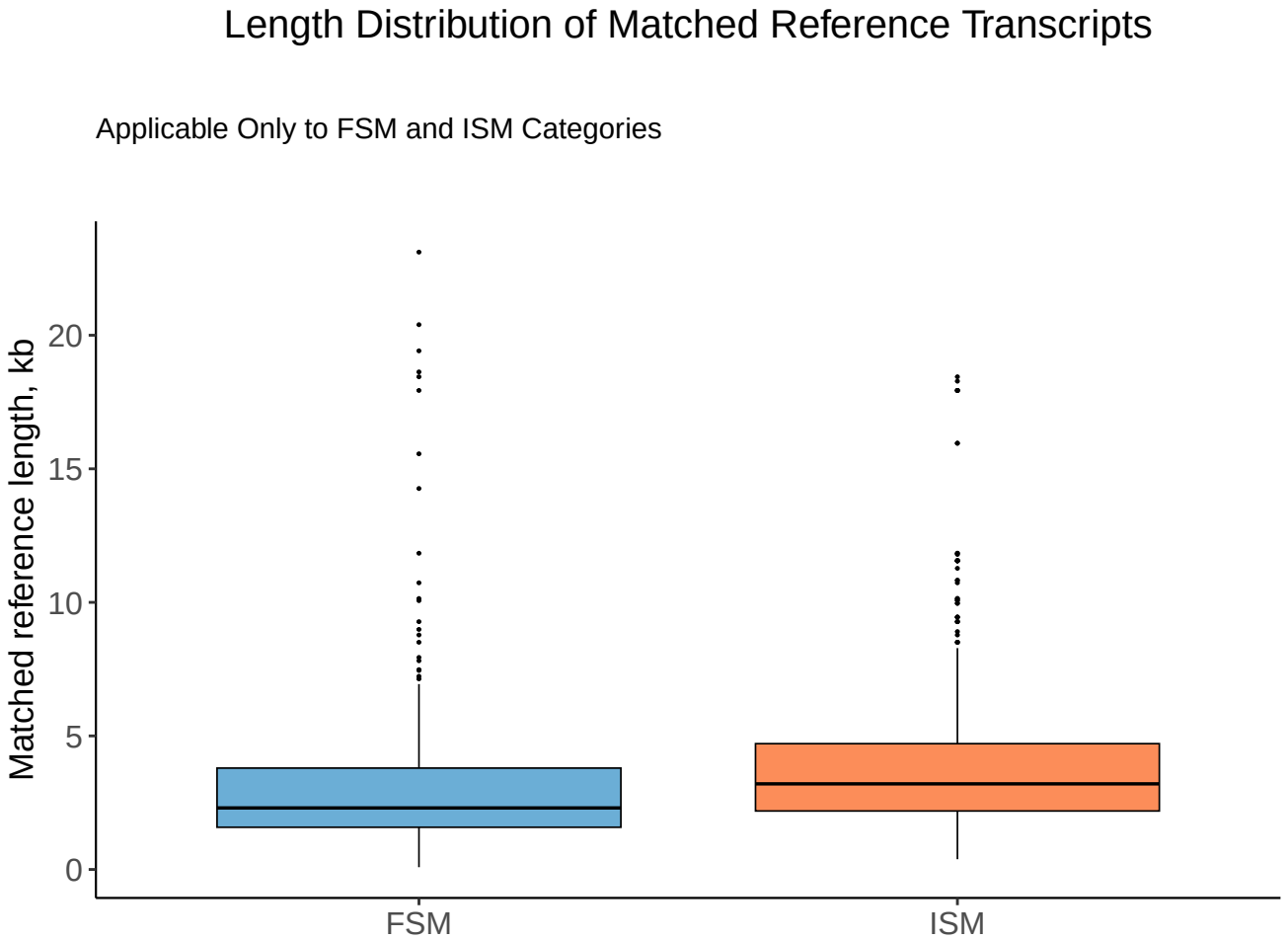

Transcript Expression by Subcategory
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
log2(TPM+1)
12.5
10.0
7.5
5.0
2.5
0.0
Mono−exon
Multi−exon
Mono−exon
Multi−exon
Intron retention
Multi−exon
Mono−exon
Multi−exon
Mono−exon
Multi−exon

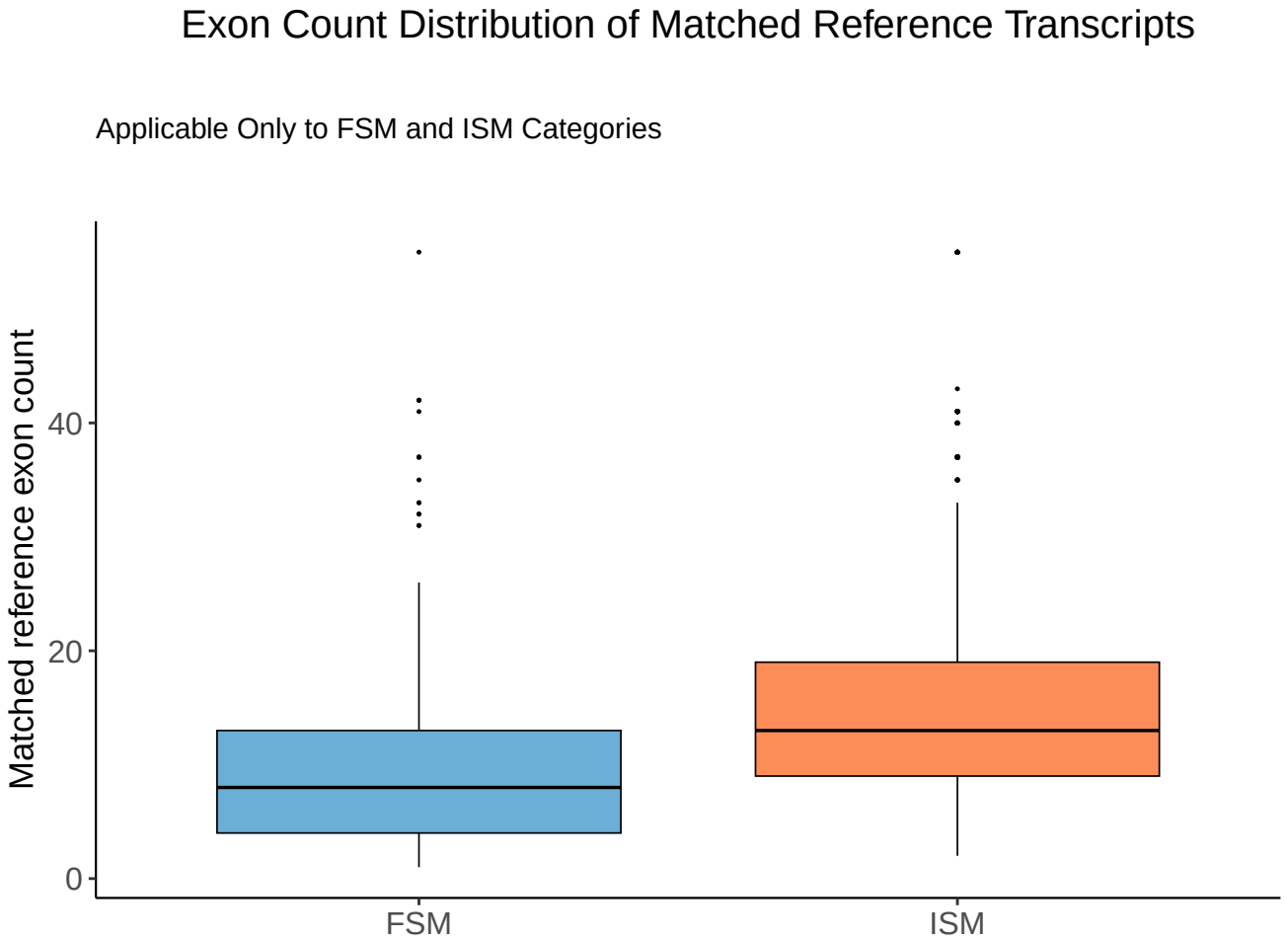

Long Reads Count by Structural Category
log2(FL_TPM+1)
16
12
8
4
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

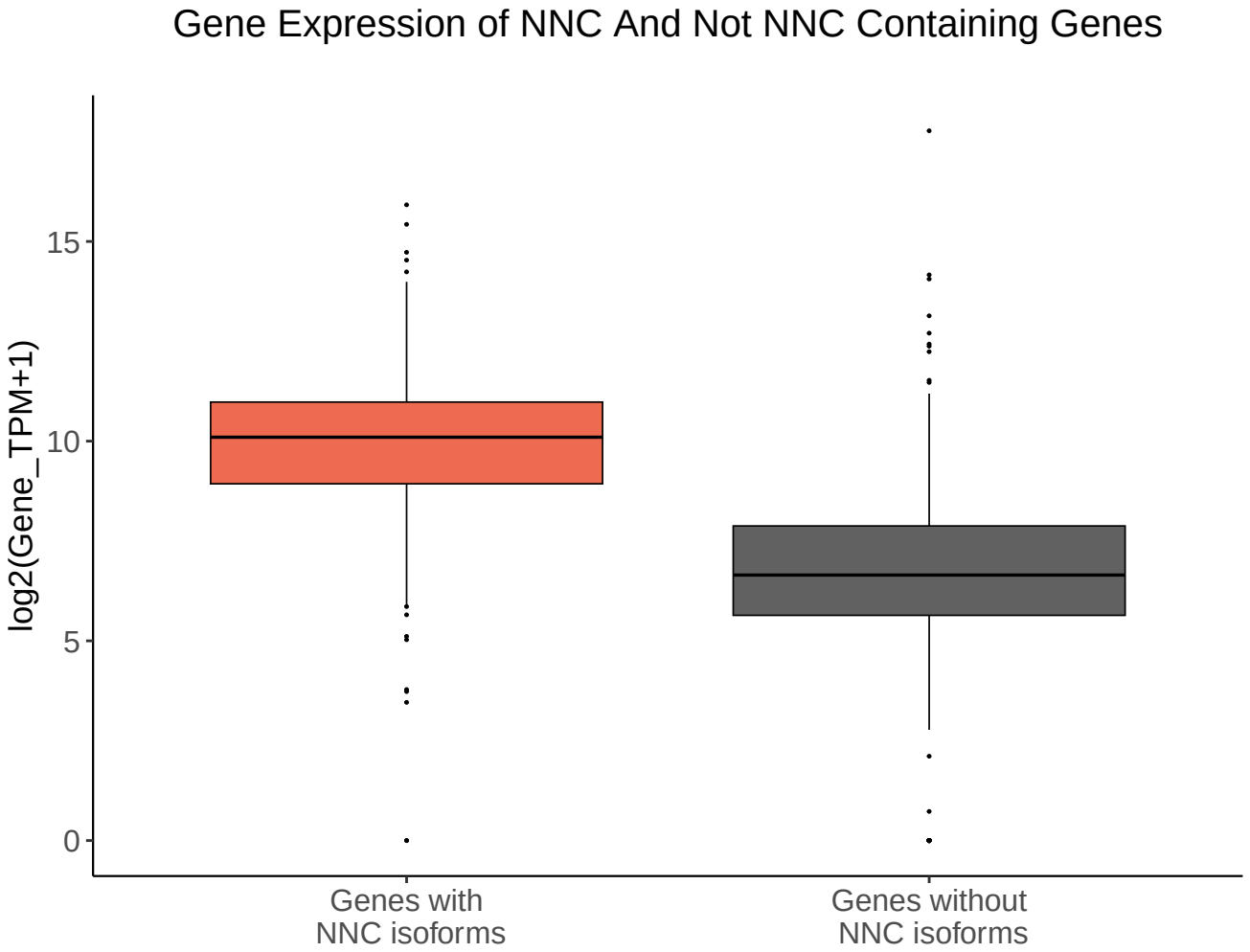

Long Reads Count by Subcategory
FSM
ISM
log2(FL_TPM+1)
17.5
15.0
12.5
10.0
7.5
5.0
Alternative 3'end
Alternative 3'5'end
Alternative 5'end
Reference match
Mono−exon
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon


Long Reads Count by Subcategory
NIC
NNC
log2(FL_TPM+1)
14
12
10
8
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
Intron retention
At least 1 annot. don./accept.

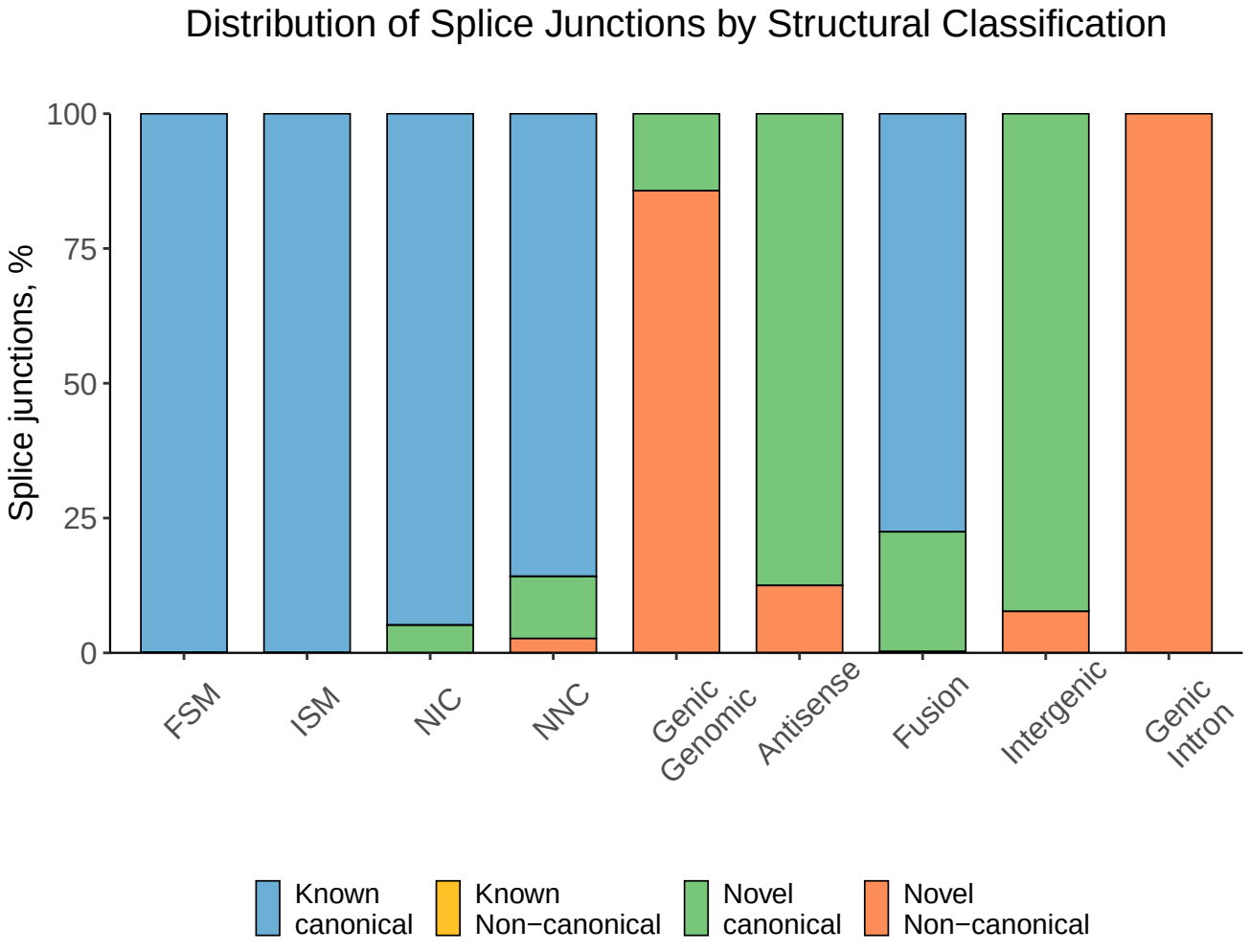

Long Reads Count by Subcategory
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
log2(FL_TPM+1)
14
12
10
Mono−exon
Multi−exon
Mono−exon
Multi−exon
Intron retention
Multi−exon
Mono−exon
Multi−exon
Mono−exon
Multi−exon

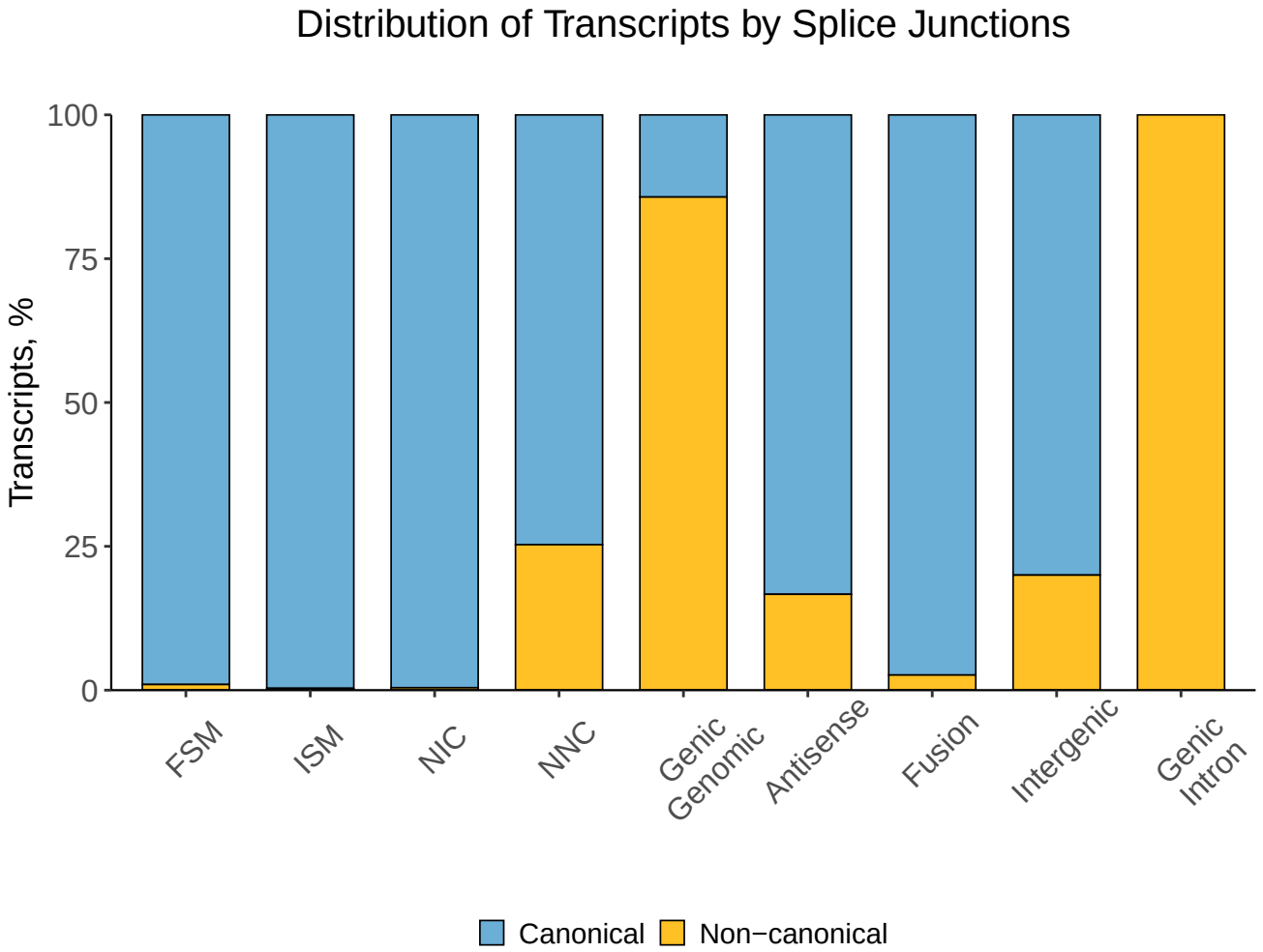

Length Distribution of Matched Reference Transcripts
Applicable Only to FSM and ISM Categories
Matched reference length, kb
0
5
10
15
20
FSM
ISM

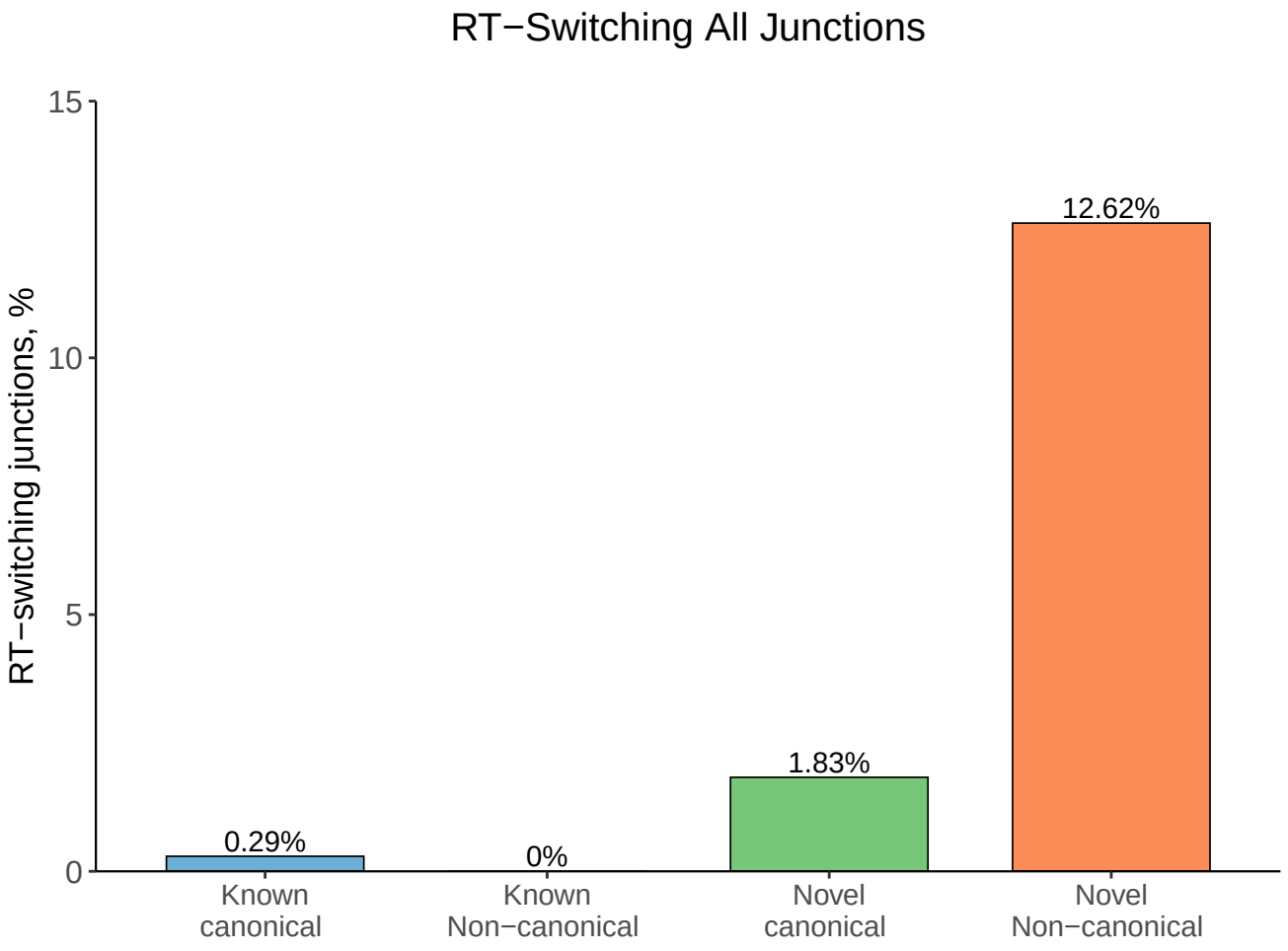

Exon Count Distribution of Matched Reference Transcripts
Applicable Only to FSM and ISM Categories
Matched reference exon count
0
20
40
FSM
ISM

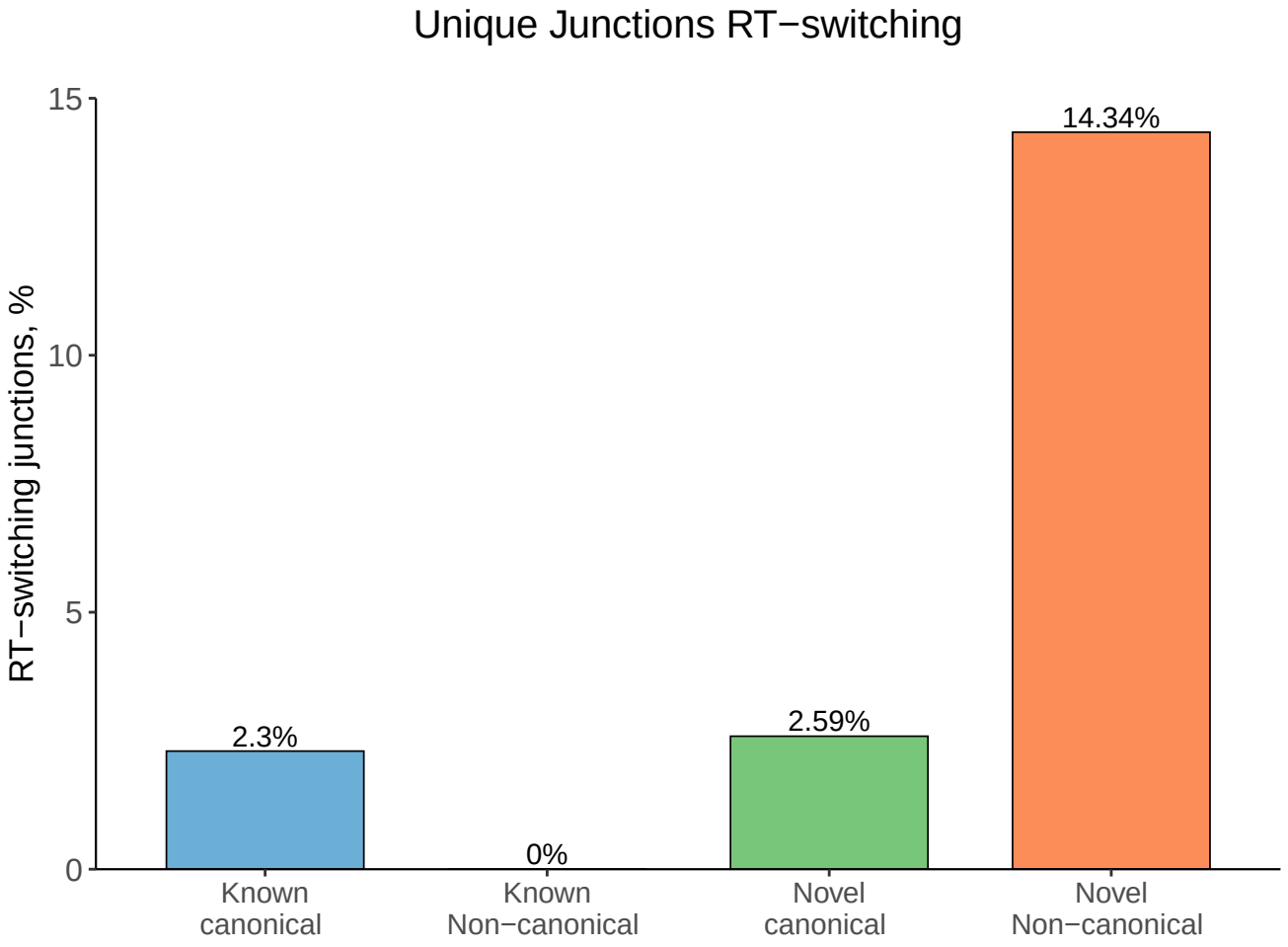

Gene Expression of NNC And Not NNC Containing Genes
log2(Gene_TPM+1)
15
10
5
0
Genes with
NNC isoforms
Genes without
NNC isoforms

Splice Junction Characterization

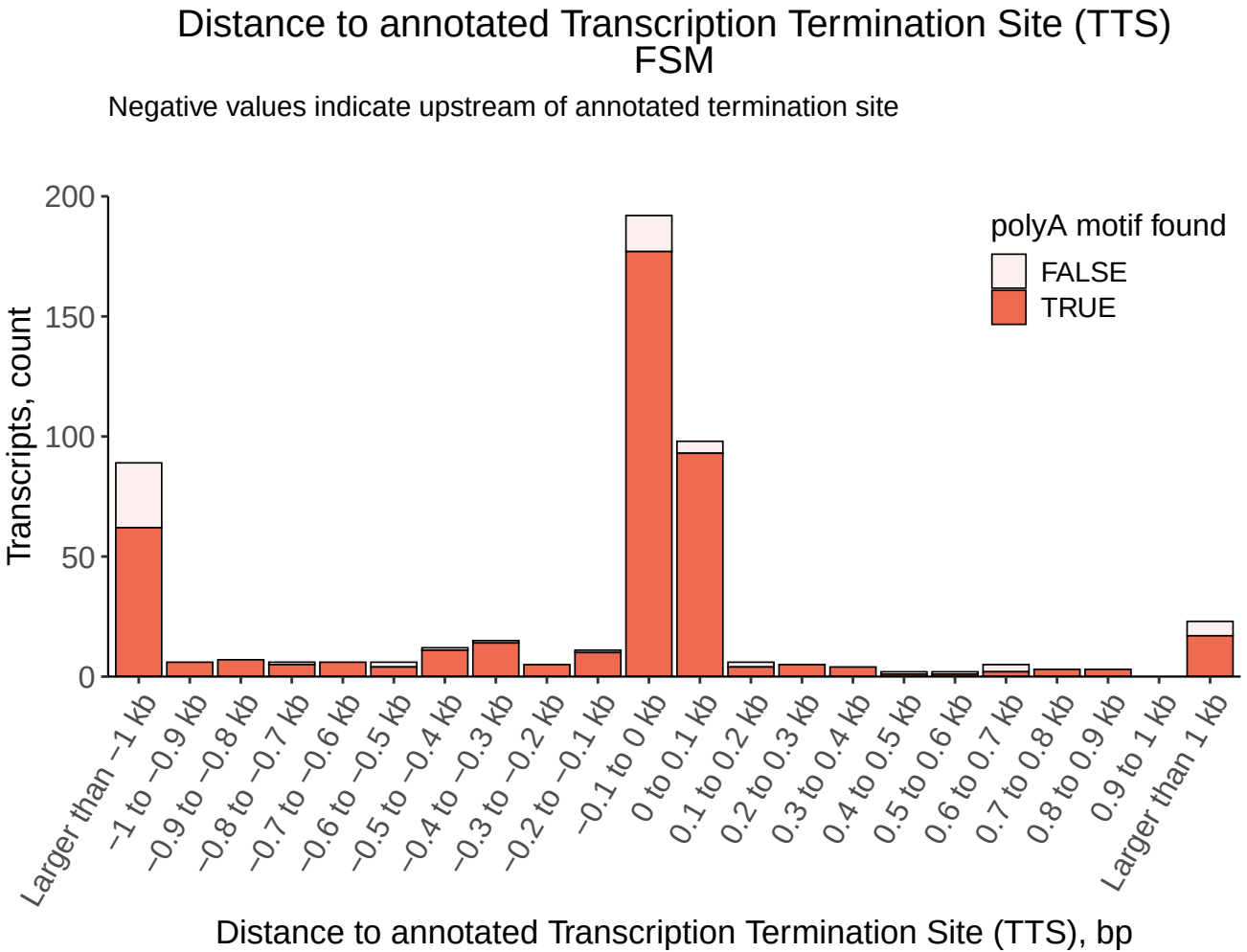

Distribution of Splice Junctions by Structural Classification
Splice junctions, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

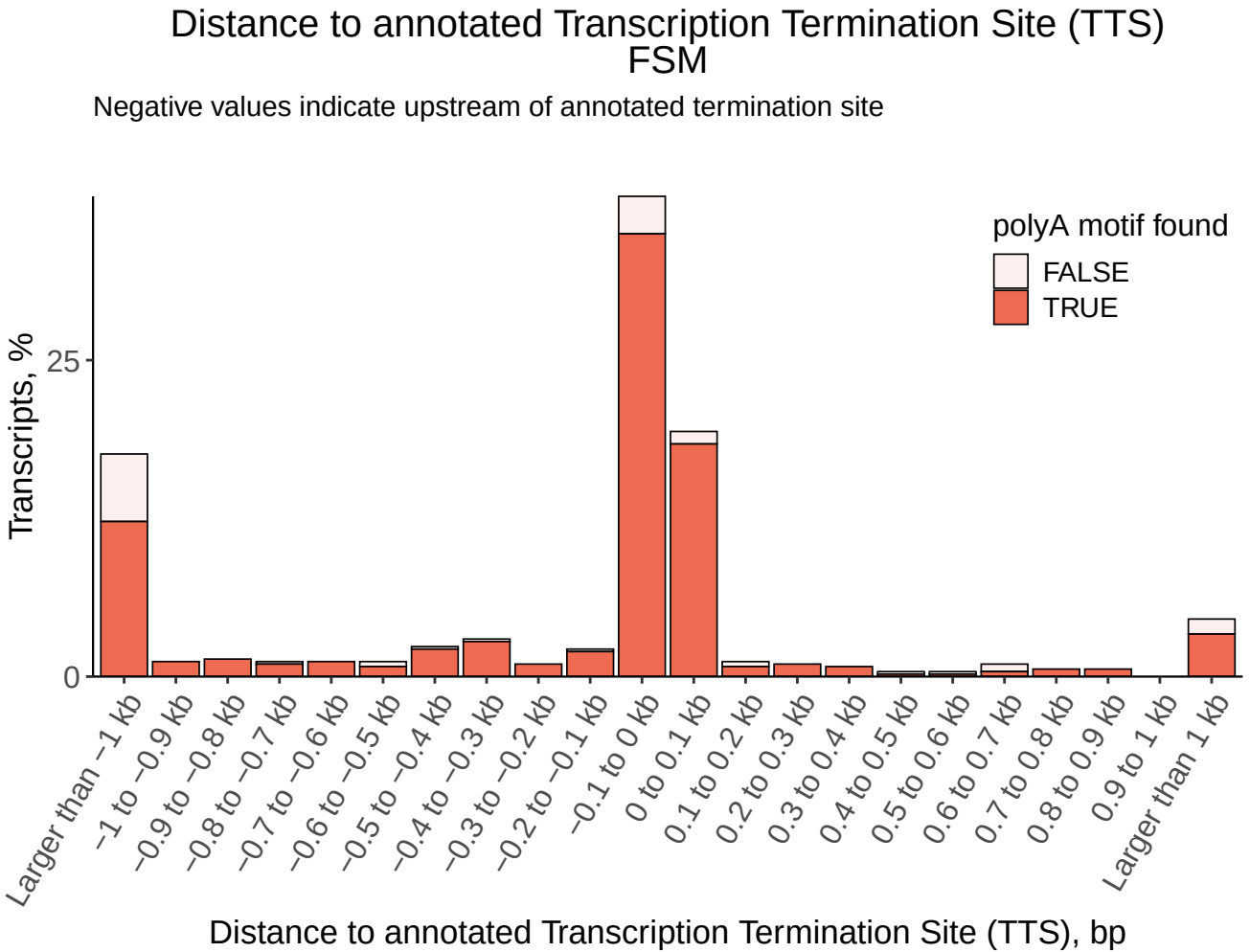

Distribution of Transcripts by Splice Junctions
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Canonical
Non−canonical

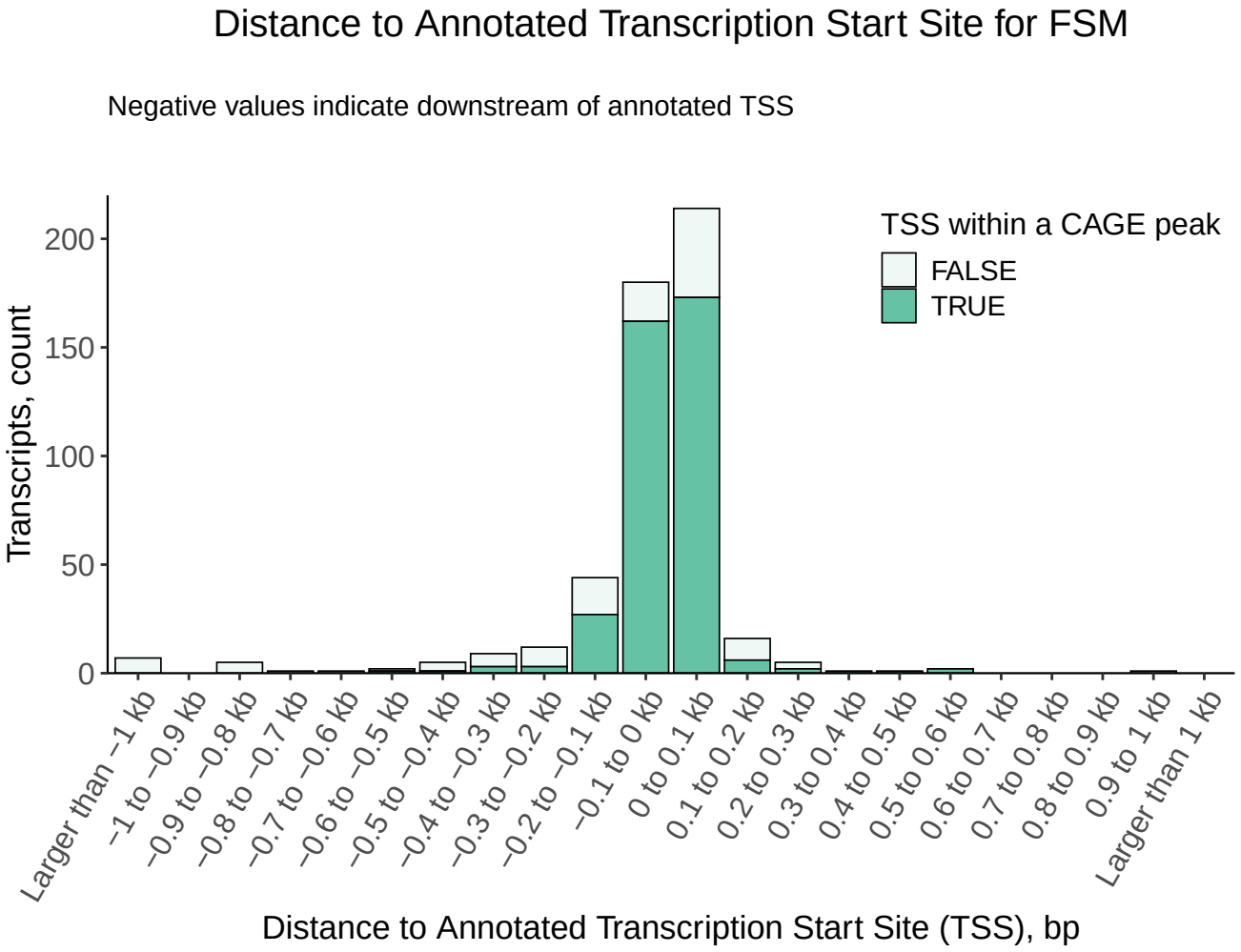

RT−Switching All Junctions
RT−switching junctions, %
15
10
5
0
0.29%
0%
1.83%
12.62%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

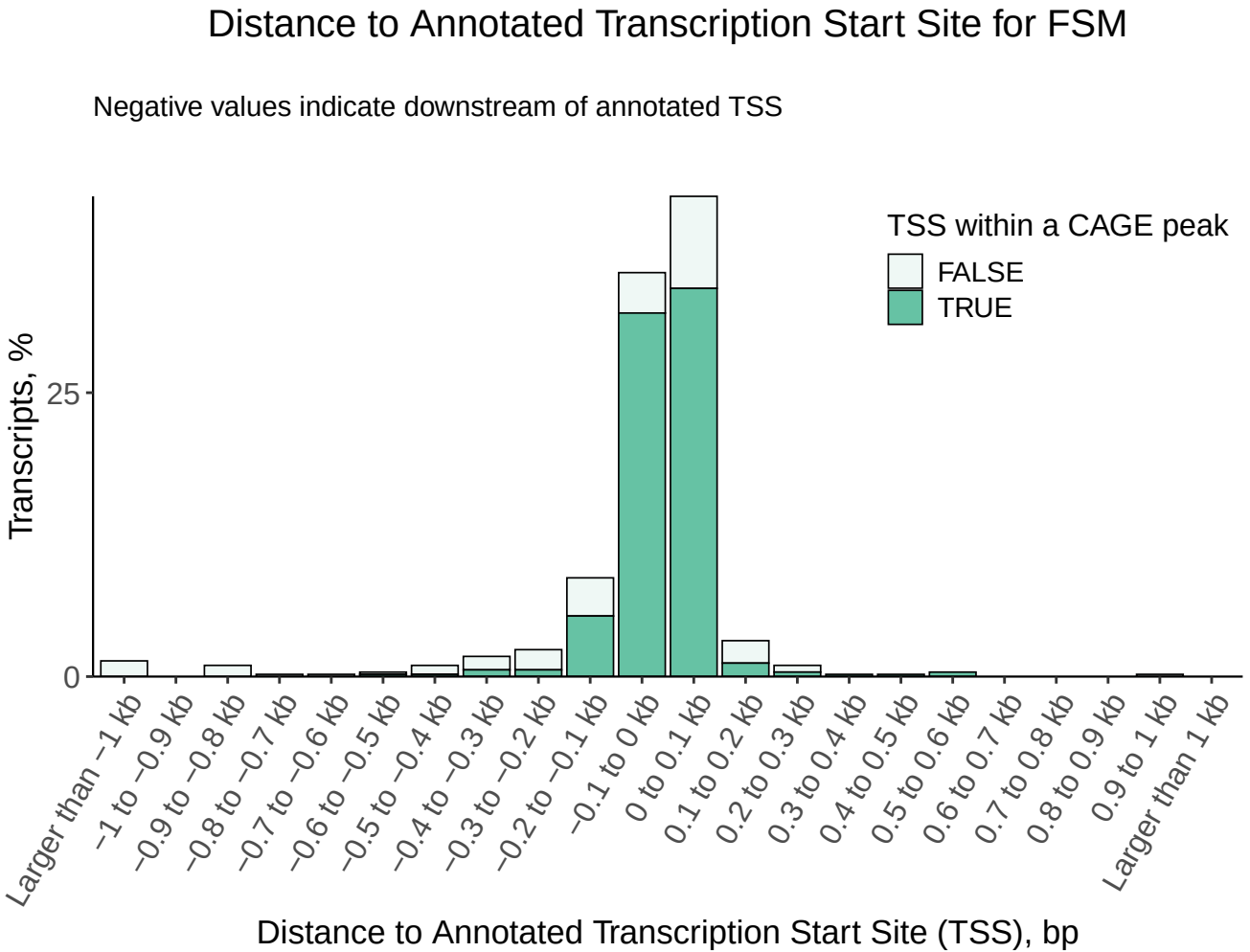

Unique Junctions RT−switching
RT−switching junctions, %
15
10
5
0
2.3%
0%
2.59%
14.34%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

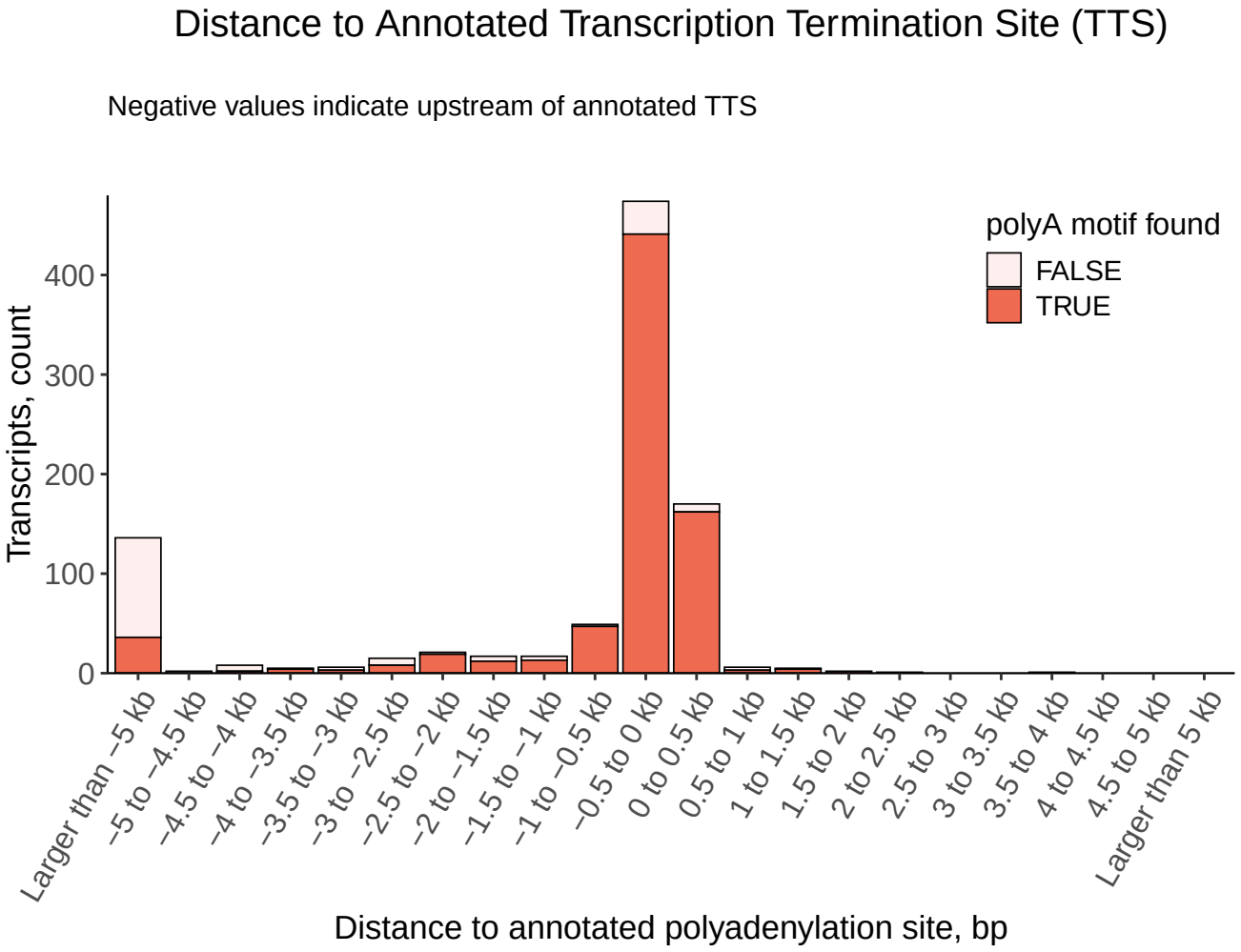
Comparison With Annotated TSS and TTS

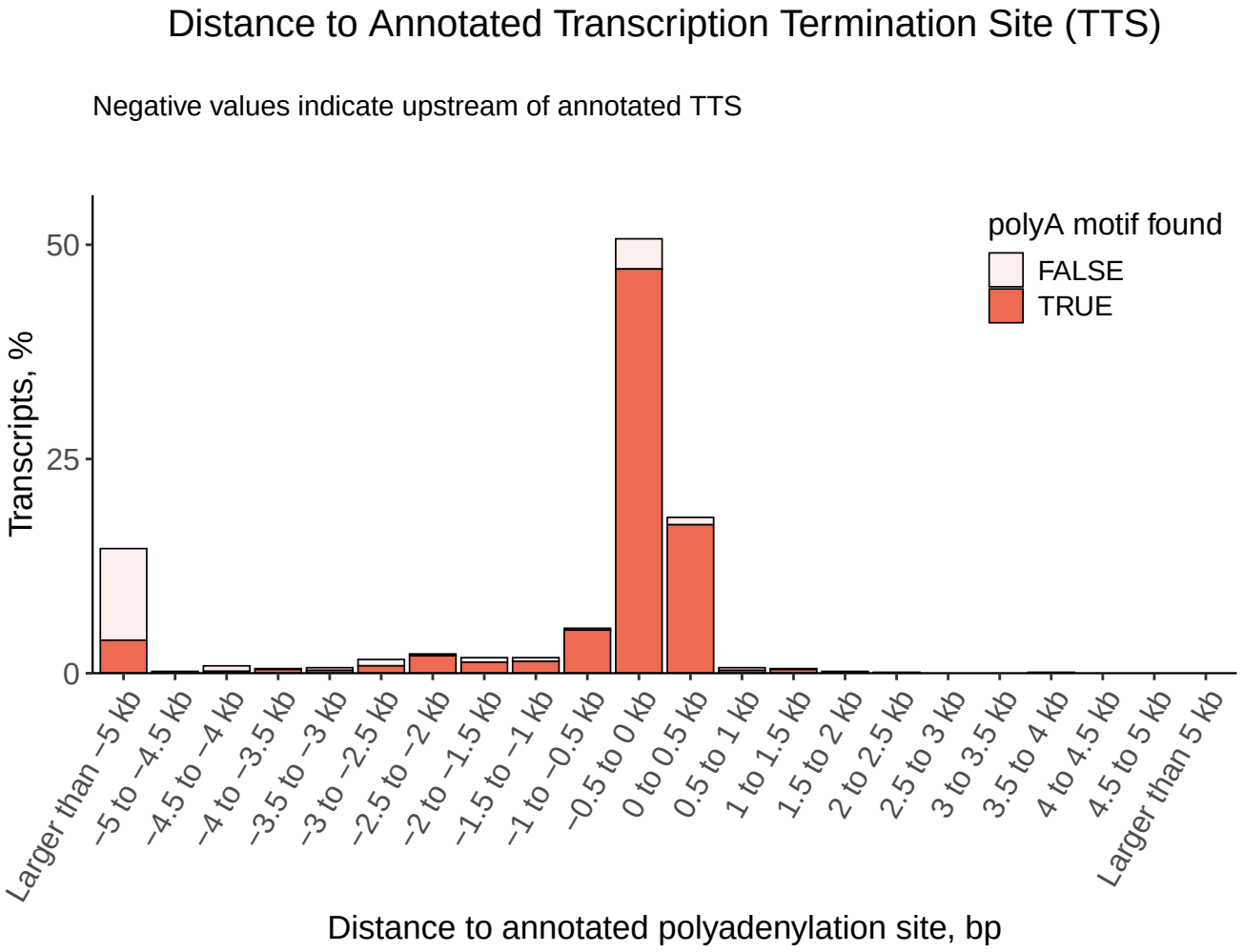

Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
Transcripts, count
200
150
100
50
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 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

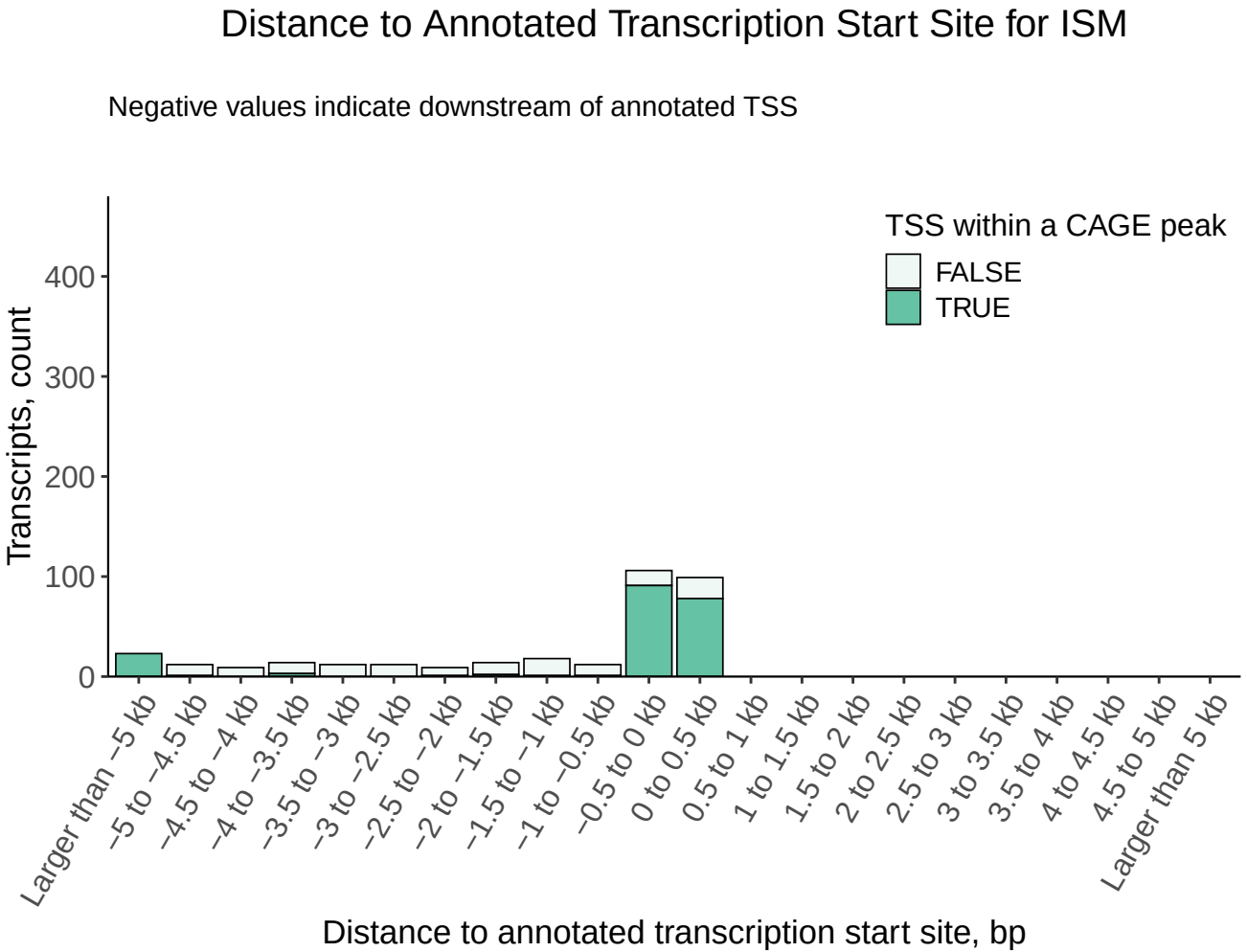

Distance to annotated Transcription Termination Site (TTS)
FSM
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3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

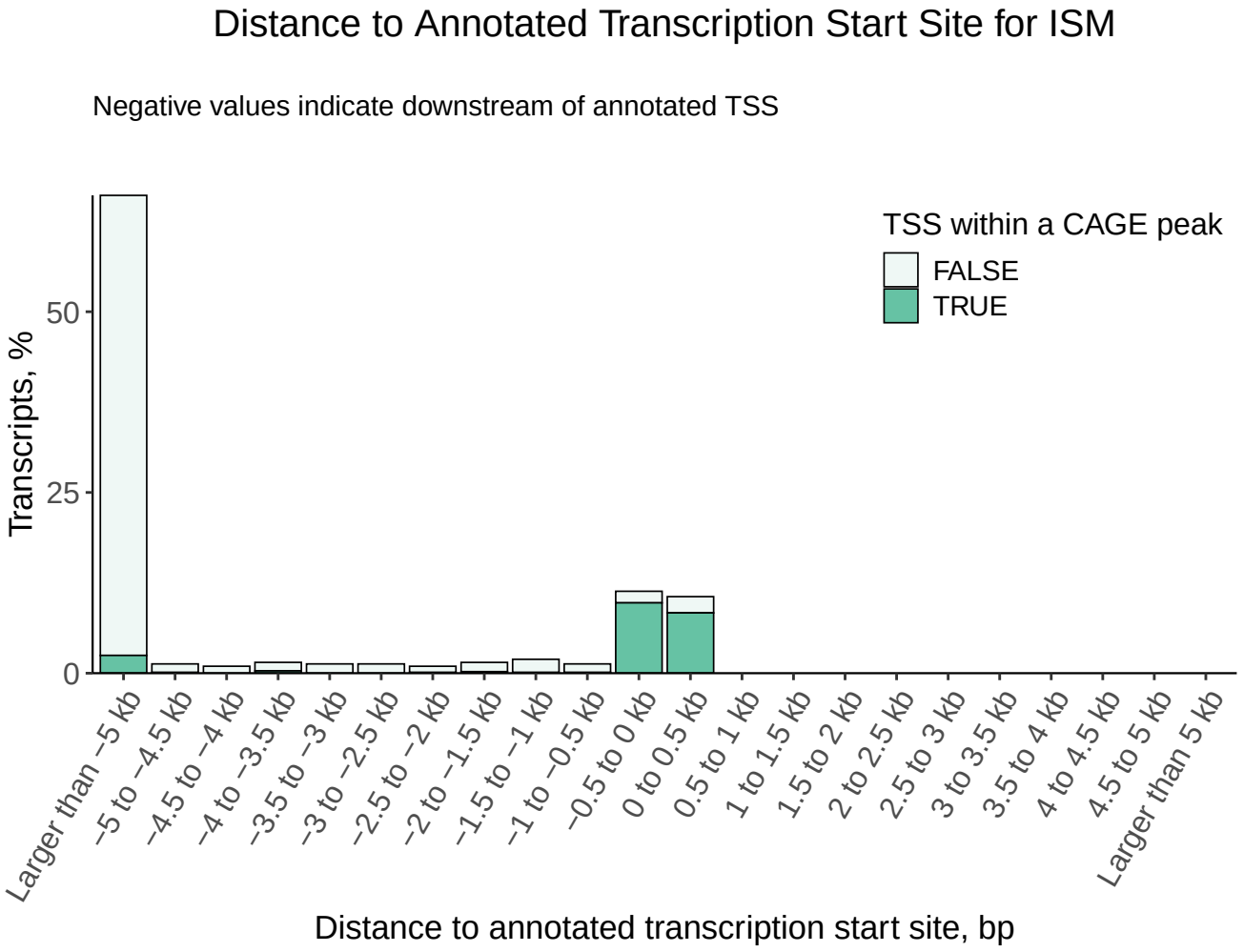

Distance to Annotated Transcription Start Site for FSM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
FALSE
TRUE
Transcripts, count
0
50
100
150
200
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 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 for FSM
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

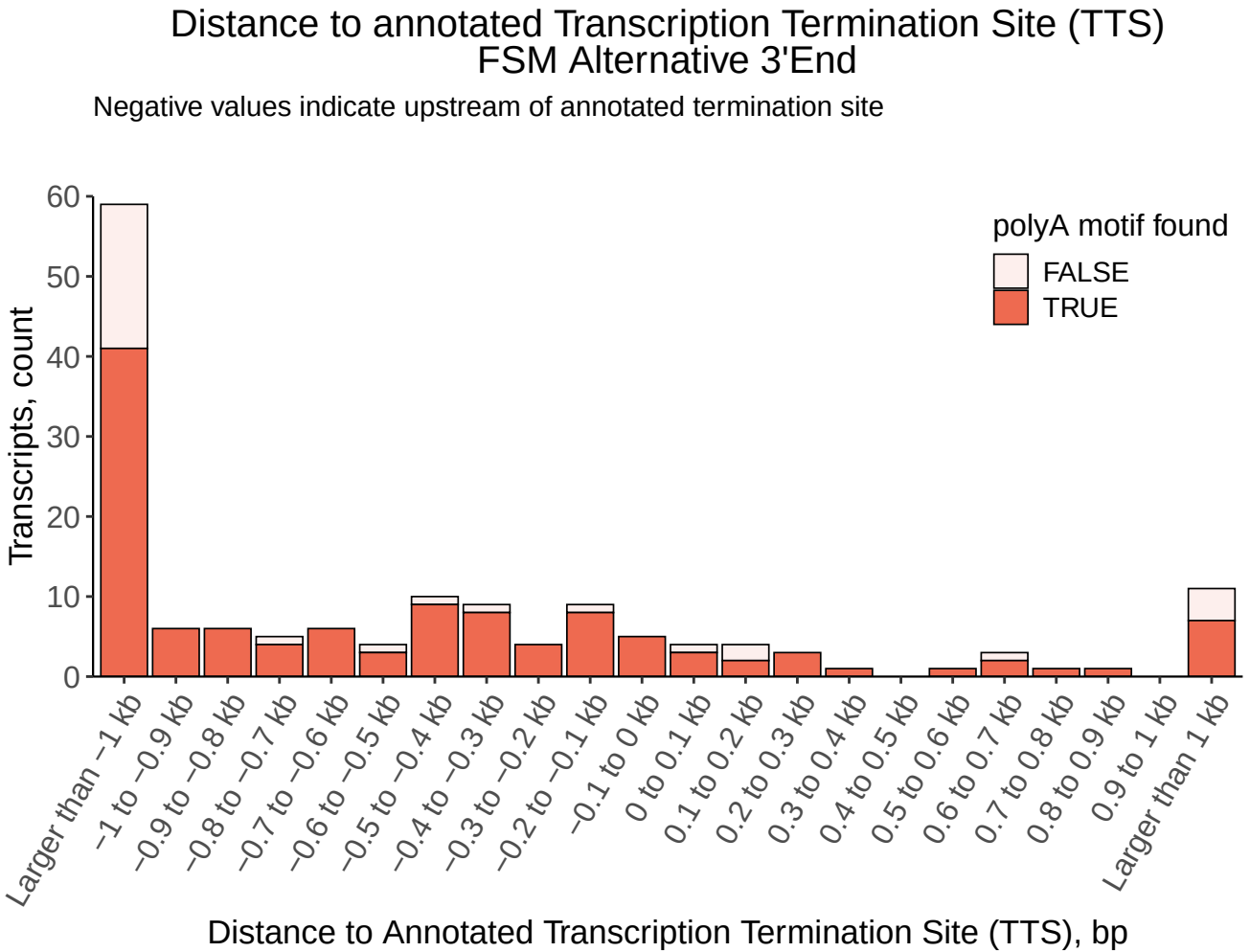

Distance to Annotated Transcription Termination Site (TTS)
Negative values indicate upstream of annotated TTS
polyA motif found
FALSE
TRUE
Transcripts, count
0
100
200
300
400
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 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 polyadenylation site, bp

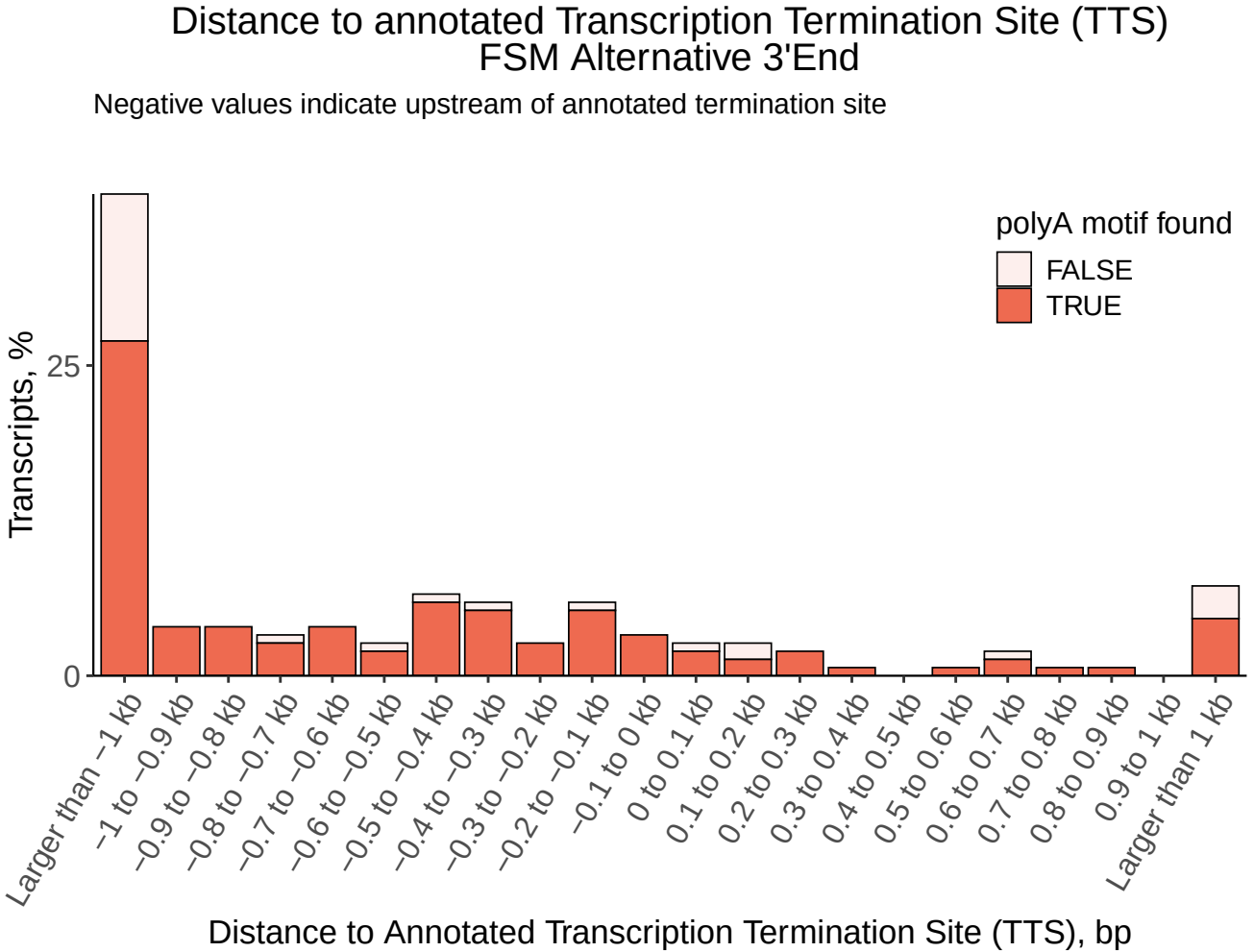

Distance to Annotated Transcription Termination Site (TTS)
Negative values indicate upstream of annotated TTS
polyA motif found
FALSE
TRUE
Transcripts, %
50
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated polyadenylation site, bp

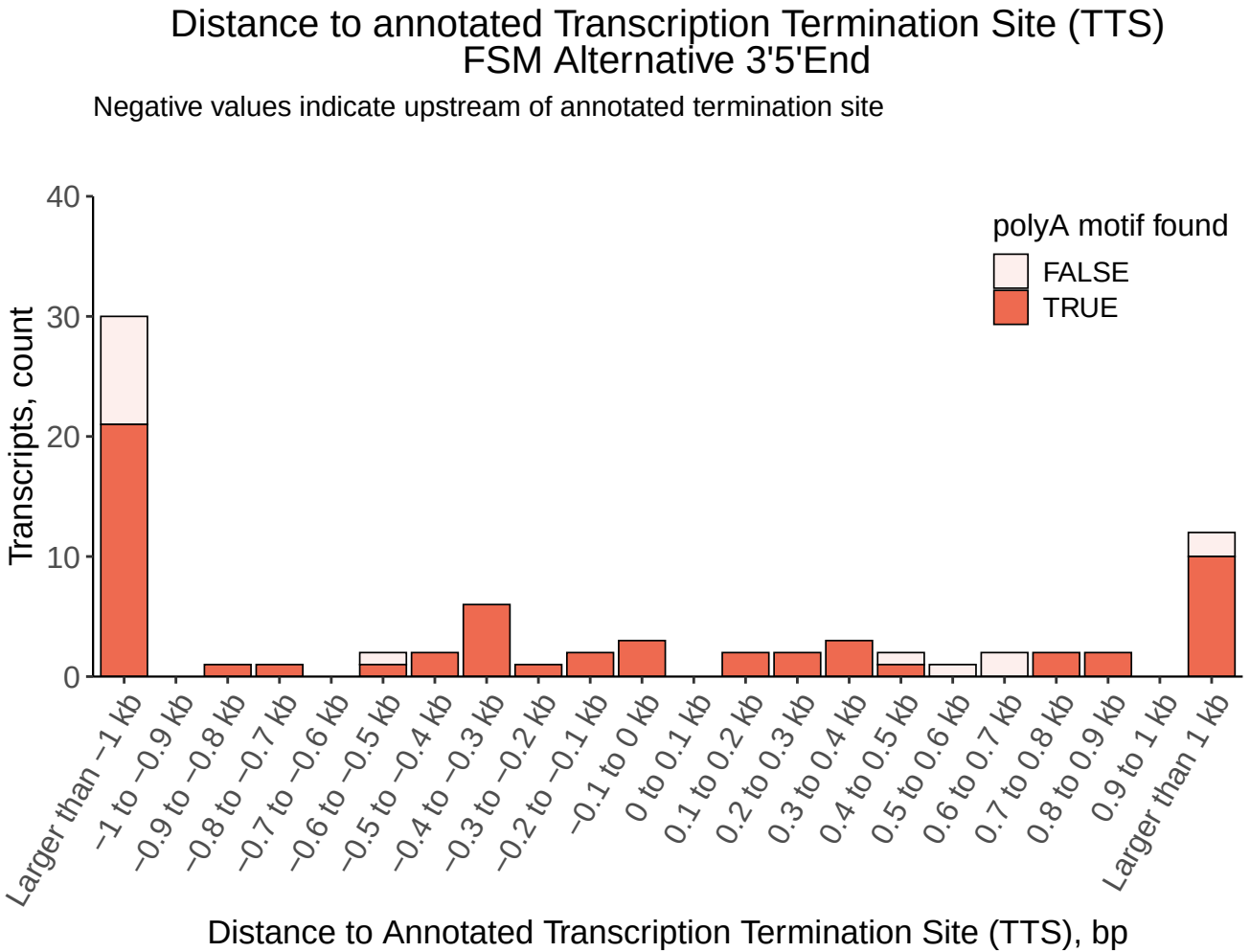

Distance to Annotated Transcription Start Site for ISM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
FALSE
TRUE
Transcripts, count
0
100
200
300
400
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −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

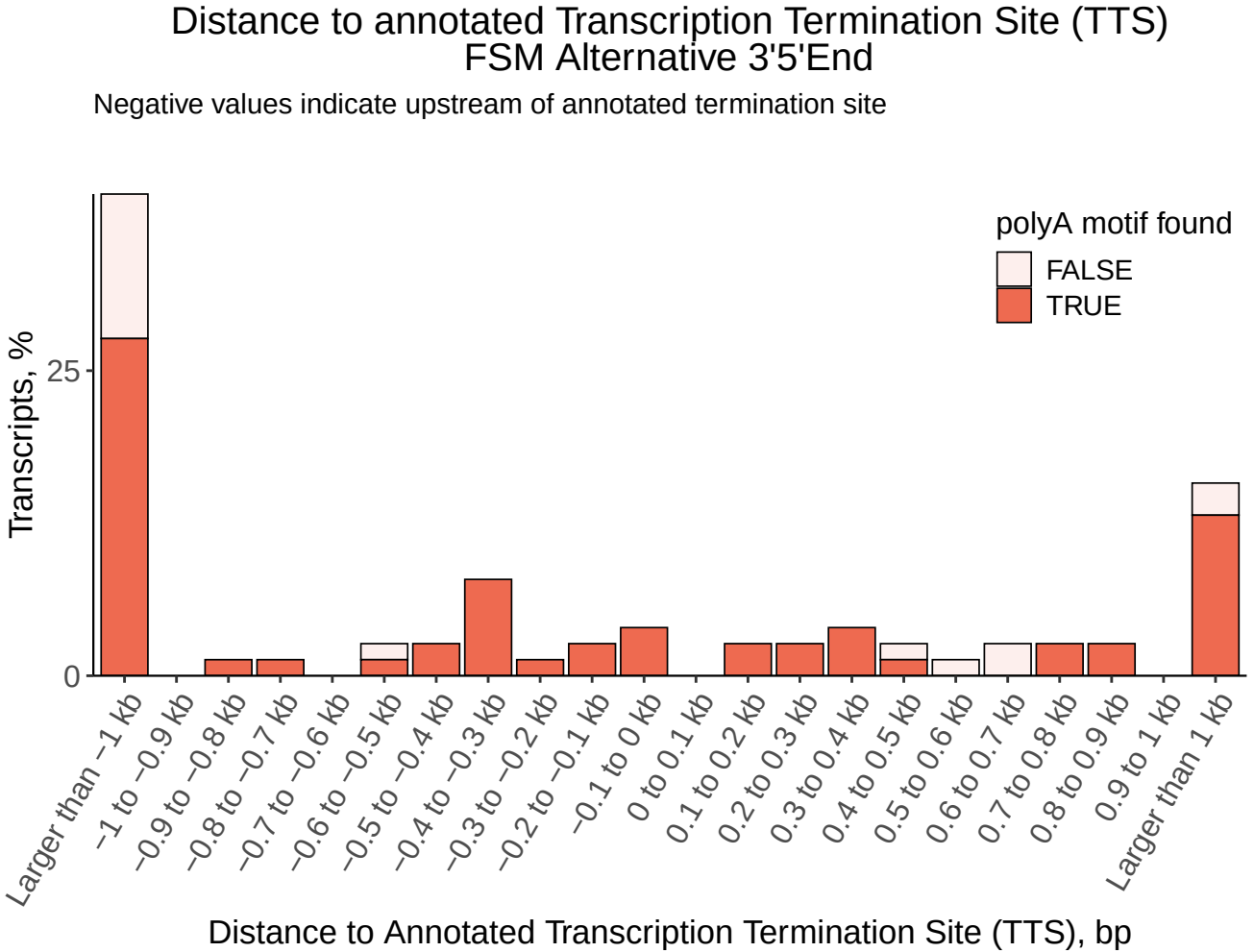

Distance to Annotated Transcription Start Site for ISM
Negative values indicate downstream of annotated TSS
TSS within a CAGE peak
FALSE
TRUE
Transcripts, %
50
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to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

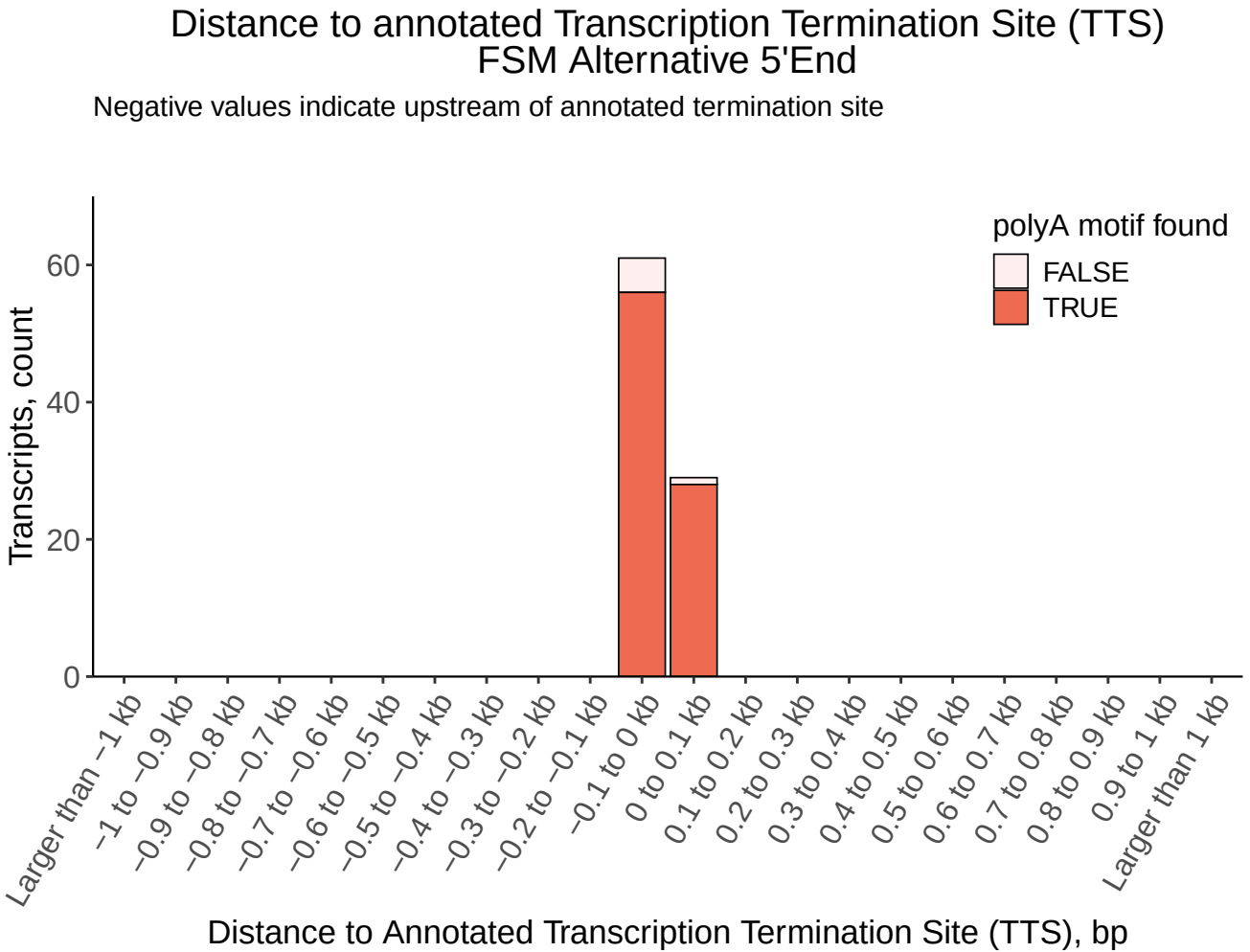
Comparison With Annotated TSS and TTS
by Subcategories

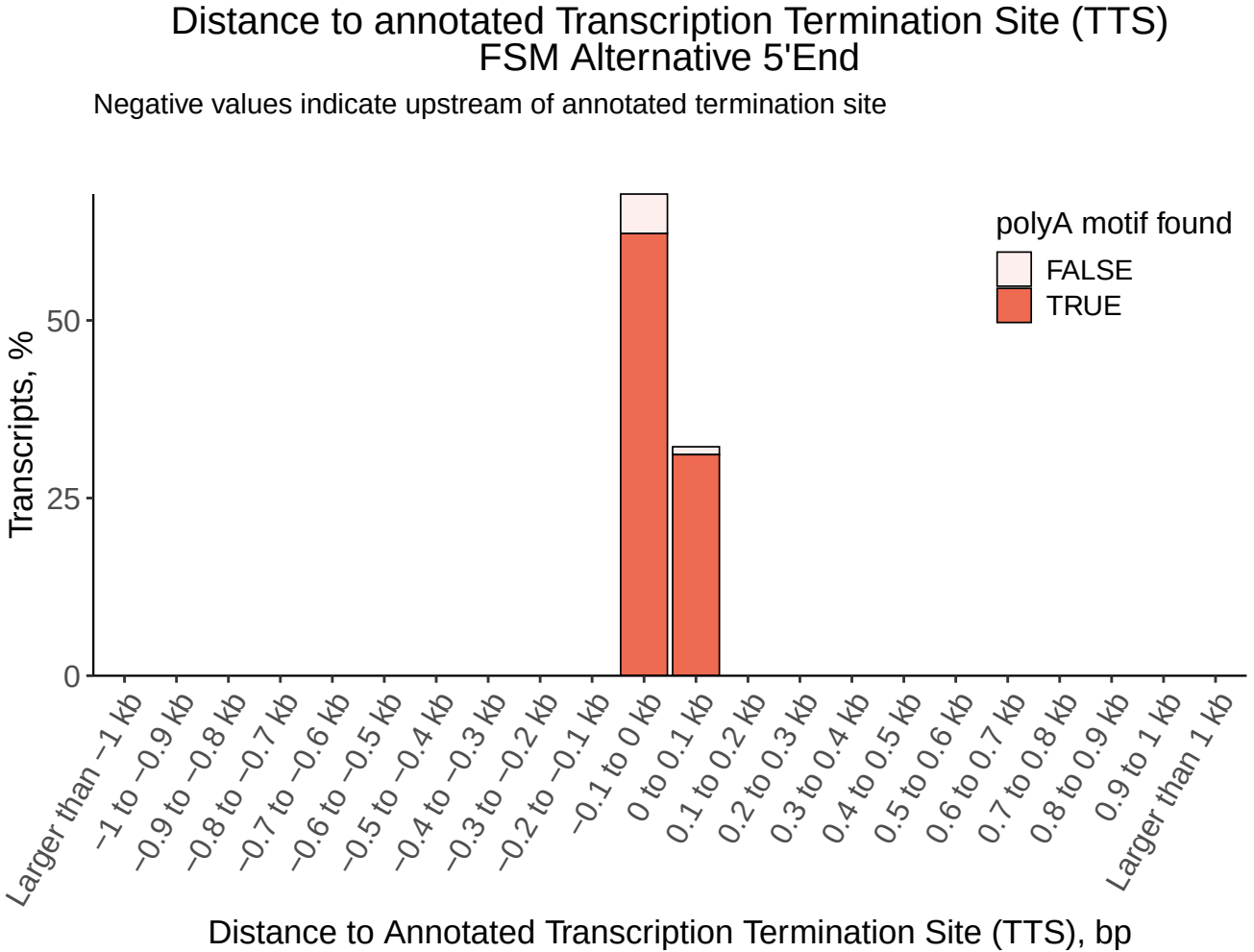

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
Transcripts, count
0
10
20
30
40
50
60
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −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

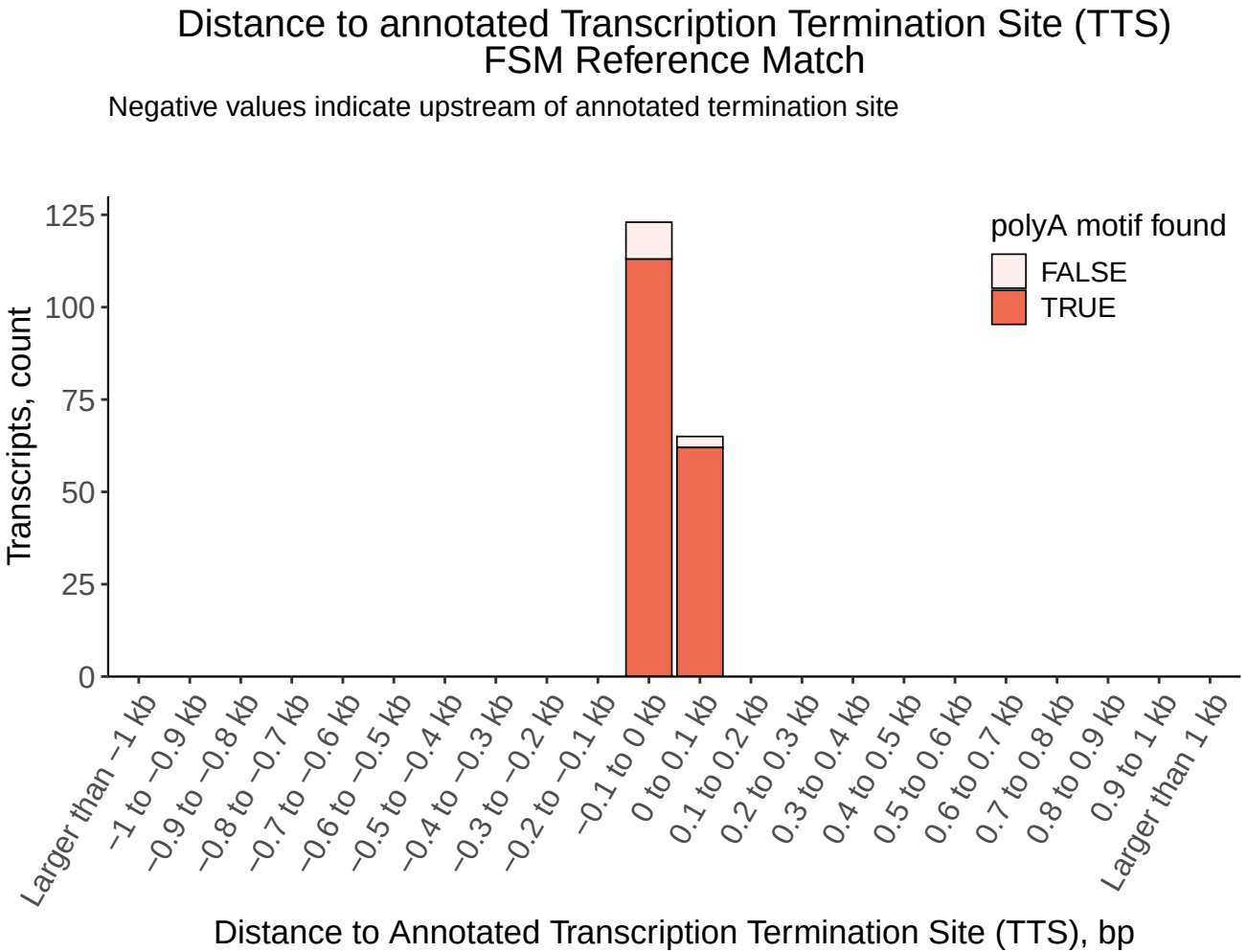

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

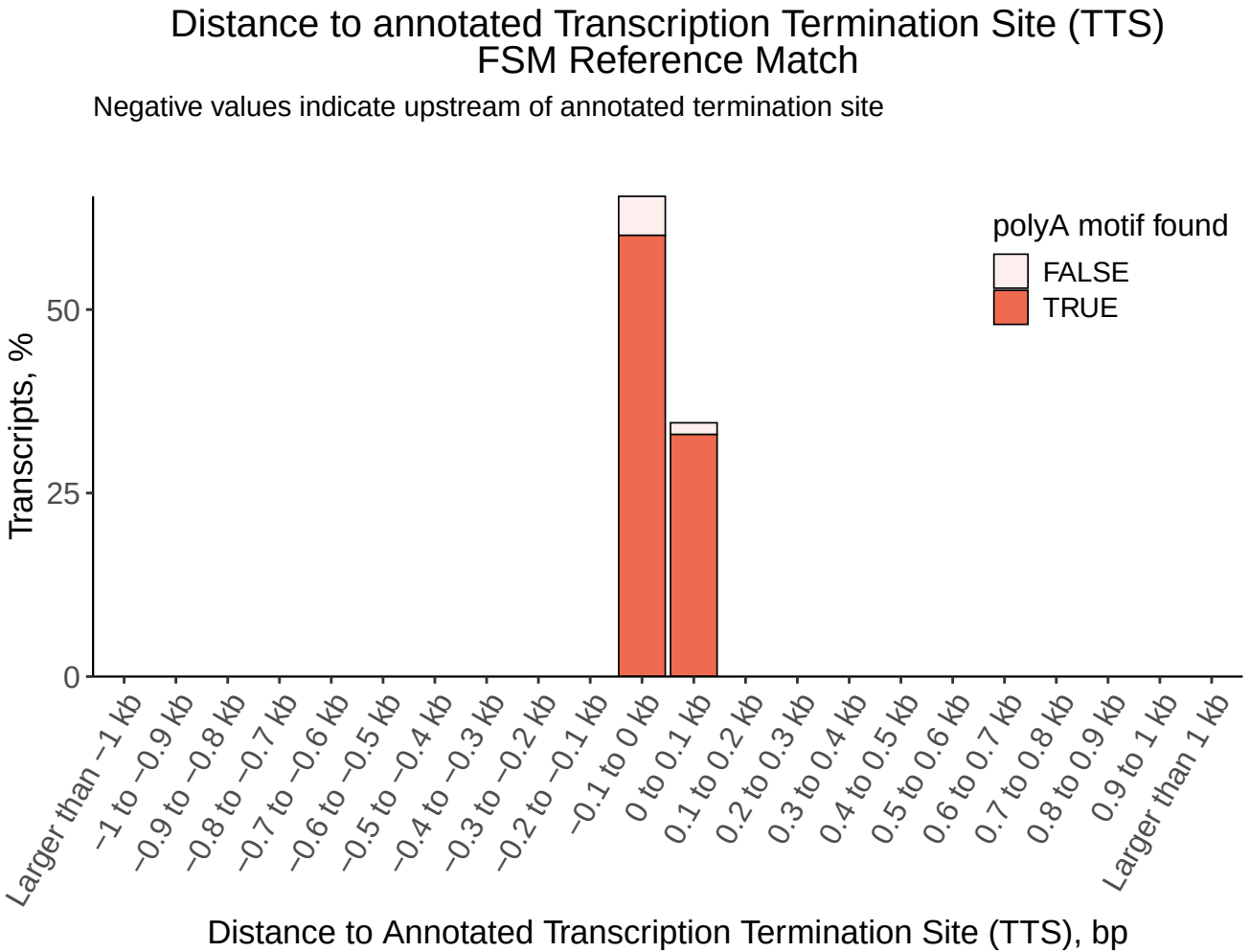

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End
Negative values indicate upstream of annotated termination site
Transcripts, count
0
10
20
30
40
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4 kb
−0.4 to −0.3 kb
−0.3 to −0.2 kb
−0.2 to −0.1 kb
−0.1 to 0 kb
0 to 0.1 kb
0.1 to 0.2 kb
0.2 to 0.3 kb
0.3 to 0.4 kb
0.4 to 0.5 kb
0.5 to 0.6 kb
0.6 to 0.7 kb
0.7 to 0.8 kb
0.8 to 0.9 kb
0.9 to 1 kb
Larger than 1 kb
Distance to Annotated Transcription Termination Site (TTS), bp

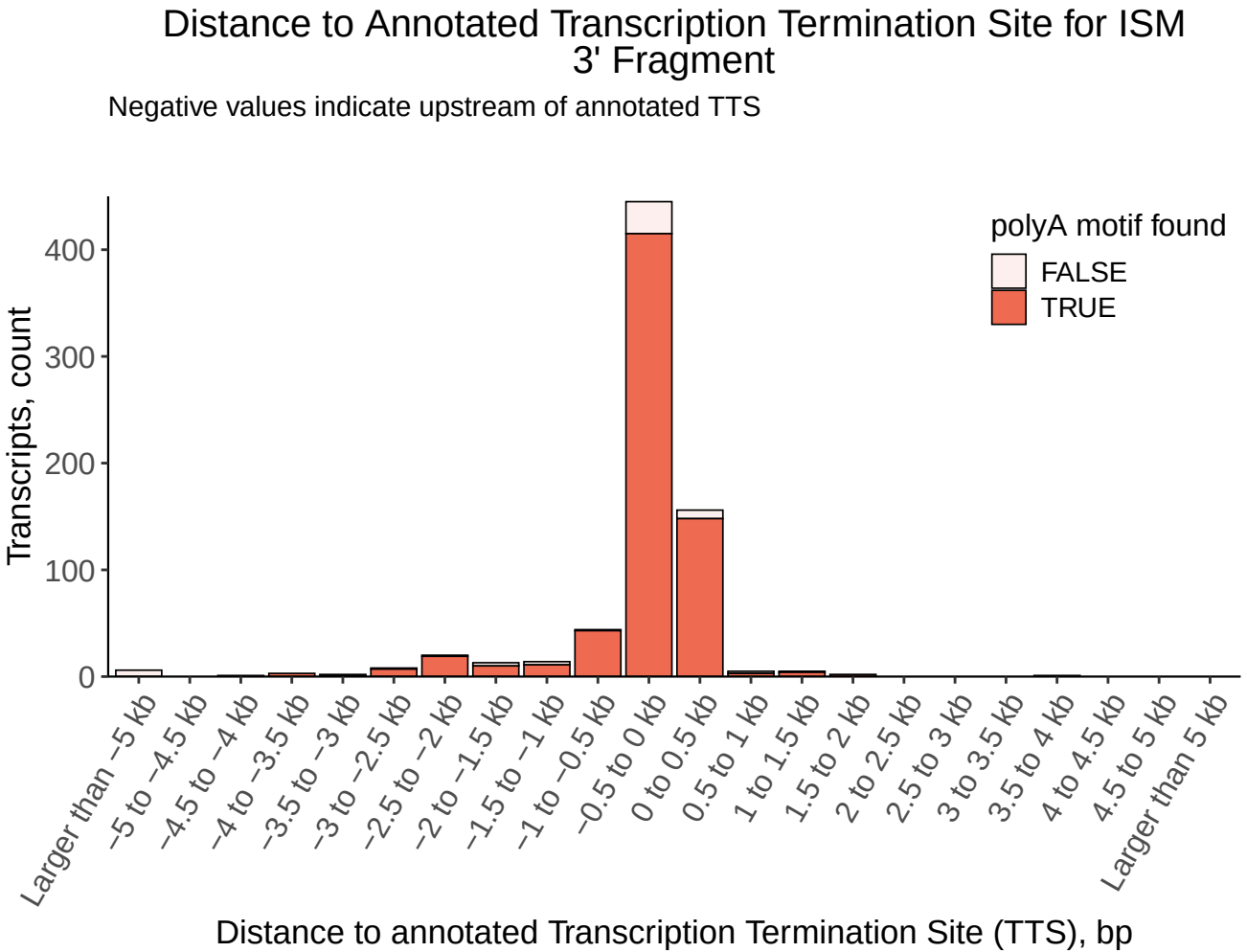

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 3'5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
25
0
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

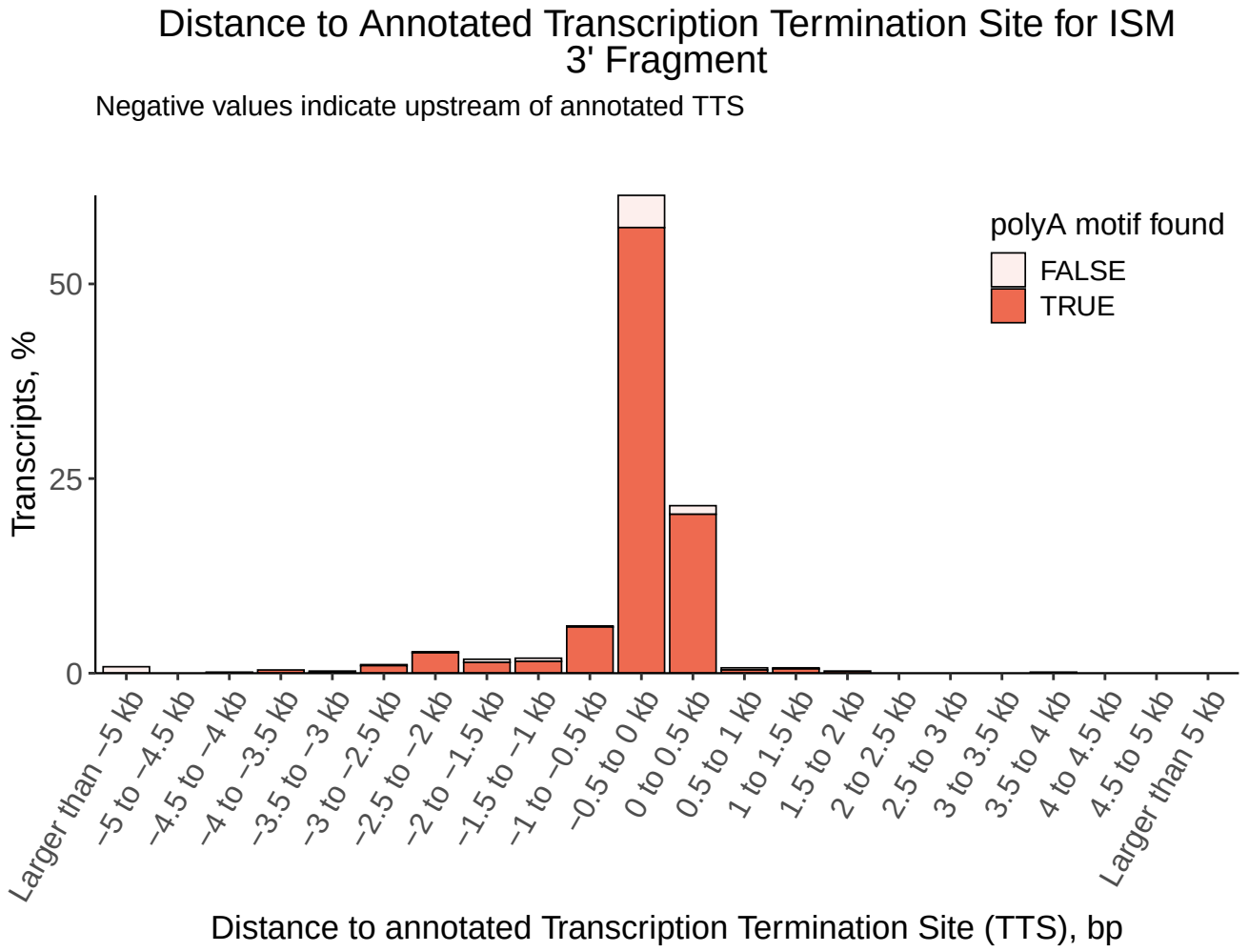

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, count
0
20
40
60
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 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

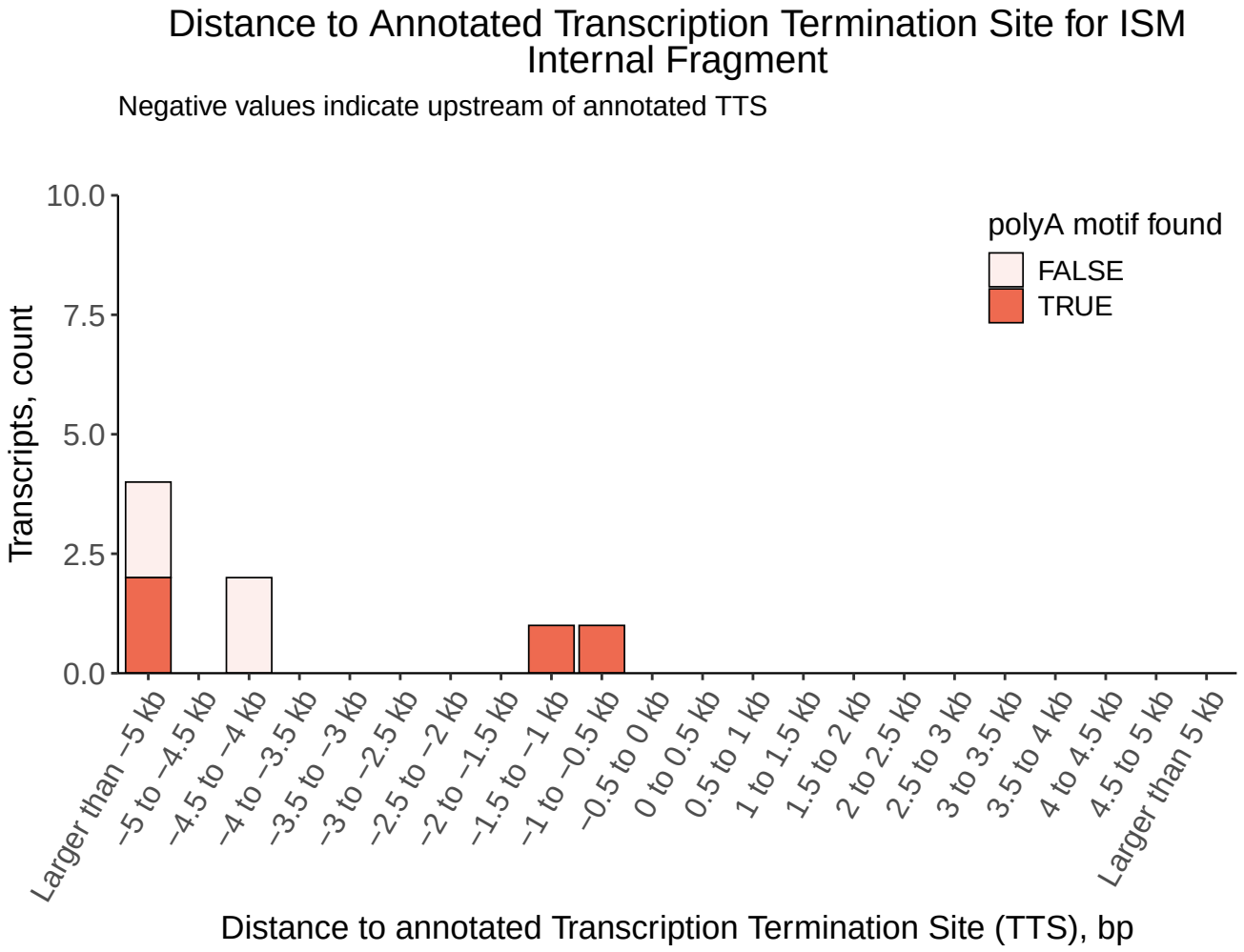

Distance to annotated Transcription Termination Site (TTS)
FSM Alternative 5'End
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to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

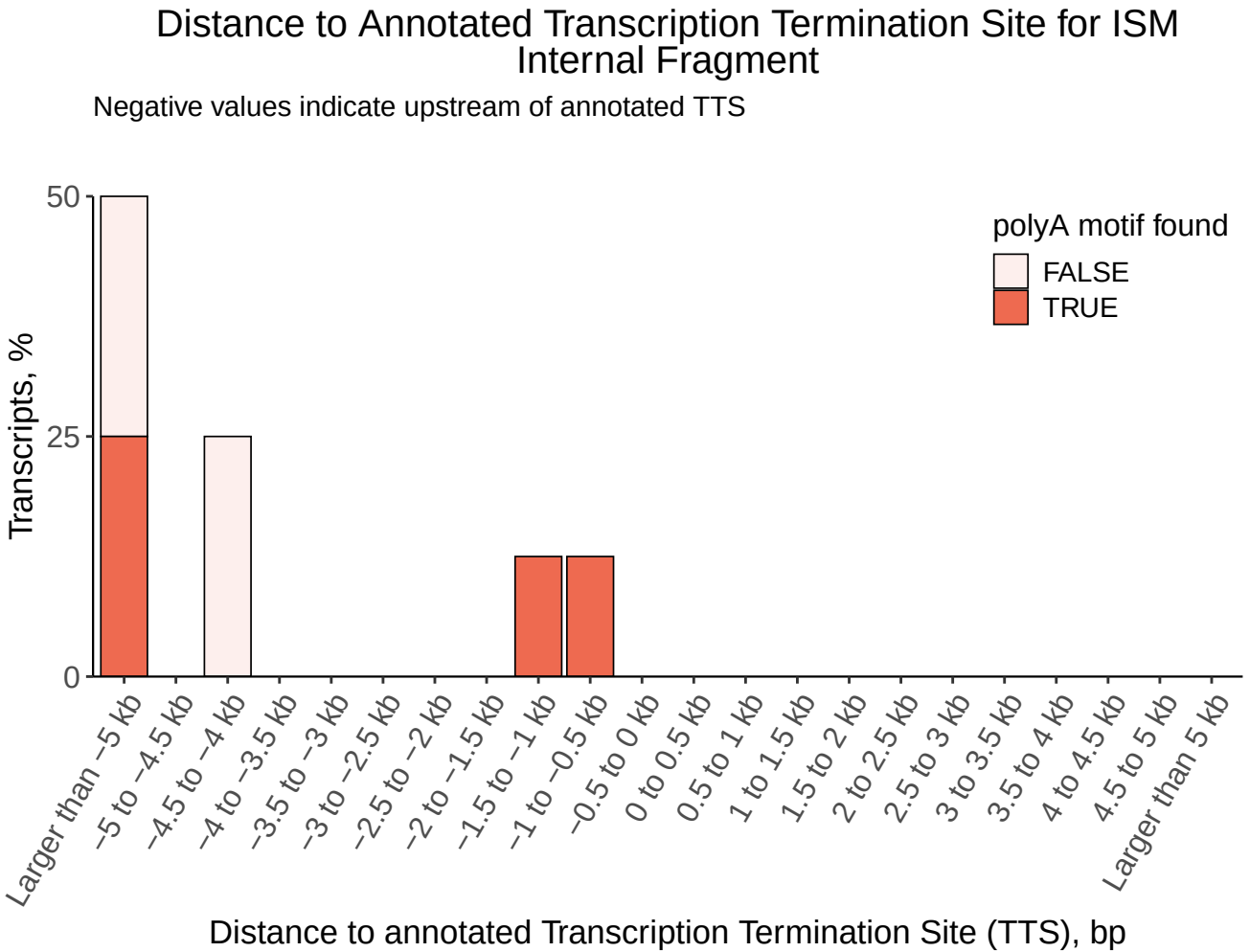

Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
Transcripts, count
125
100
75
50
25
0
polyA motif found
FALSE
TRUE
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 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

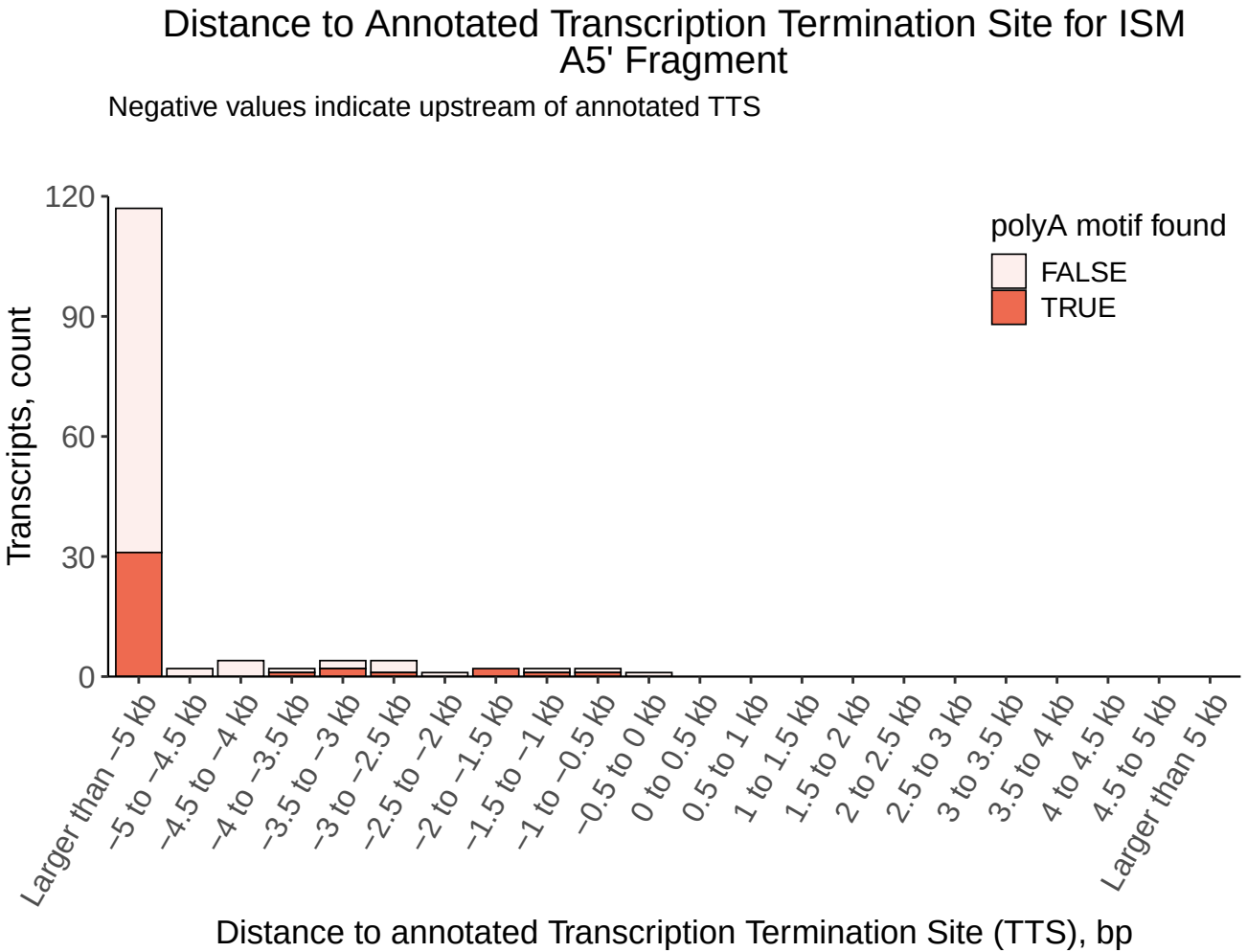

Distance to annotated Transcription Termination Site (TTS)
FSM Reference Match
Negative values indicate upstream of annotated termination site
polyA motif found
FALSE
TRUE
Transcripts, %
0
25
50
Larger than −1 kb
−1 to −0.9 kb
−0.9 to −0.8 kb
−0.8 to −0.7 kb
−0.7 to −0.6 kb
−0.6 to −0.5 kb
−0.5 to −0.4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

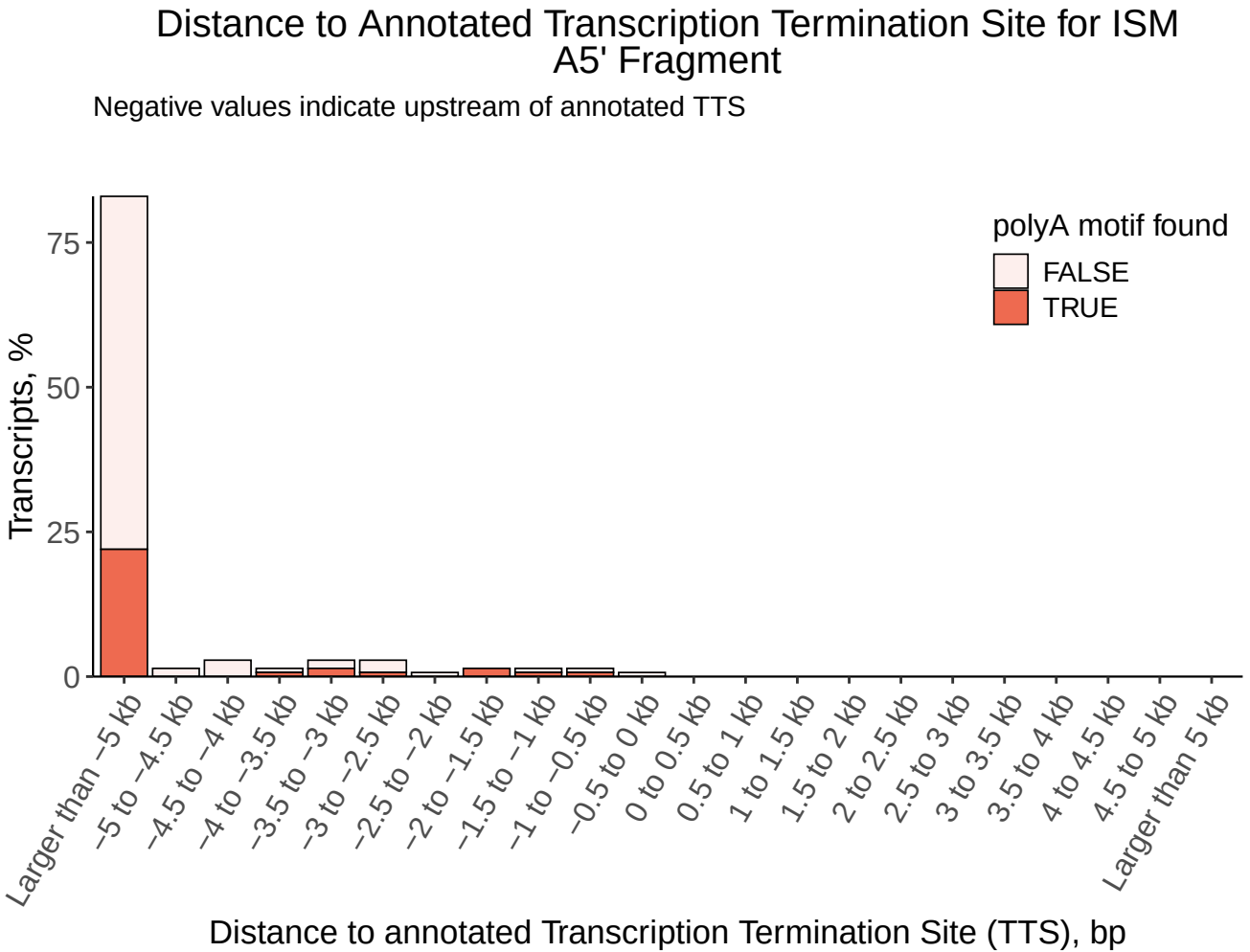

Distance to Annotated Transcription Termination Site for ISM 3' Fragment
Negative values indicate upstream of annotated TTS
Transcripts, count
0
100
200
300
400
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 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 Transcription Termination Site (TTS), bp

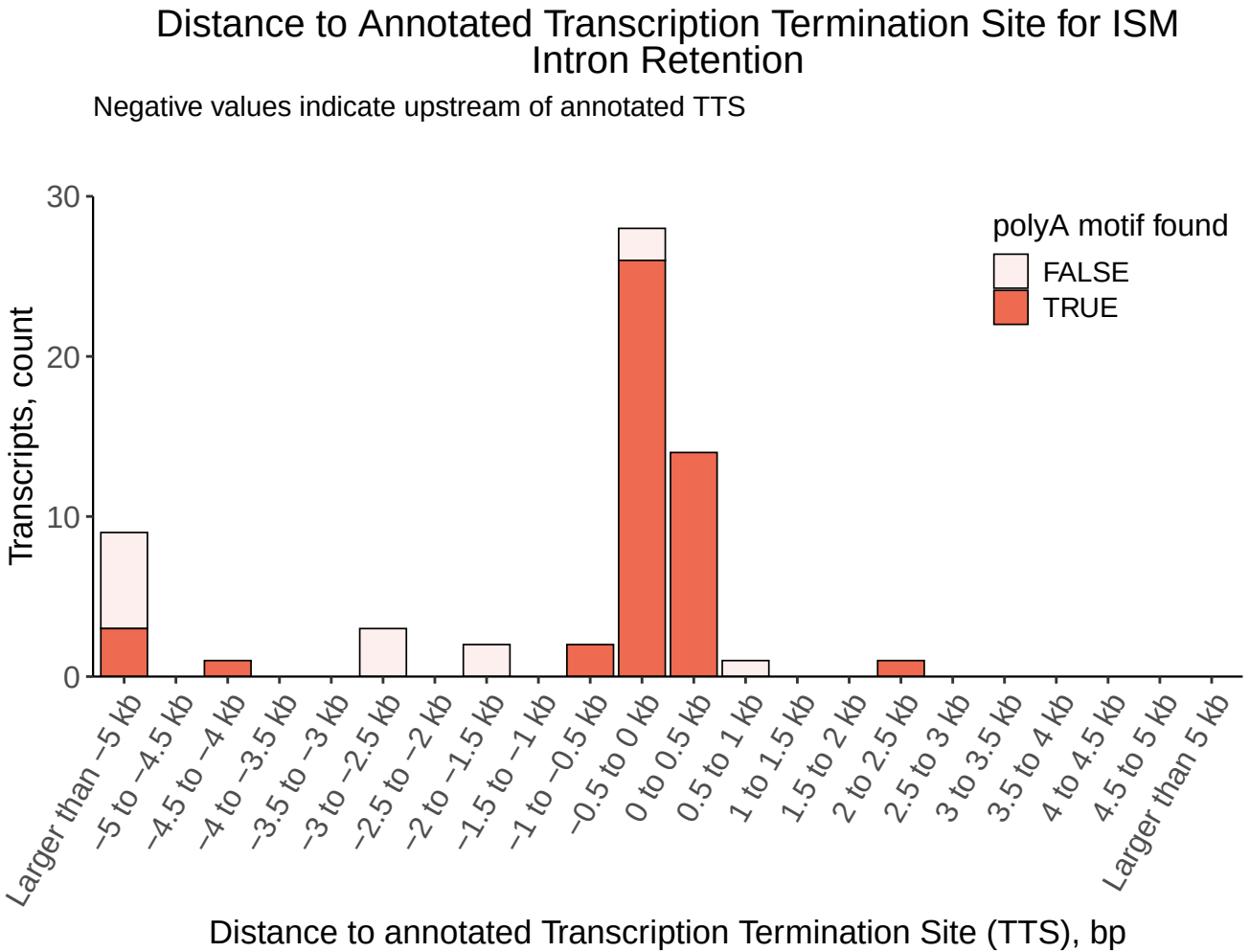

Distance to Annotated Transcription Termination Site for ISM 3' Fragment
Negative values indicate upstream of annotated TTS
polyA motif found
FALSE
TRUE
Transcripts, %
50
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated Transcription Termination Site (TTS), bp

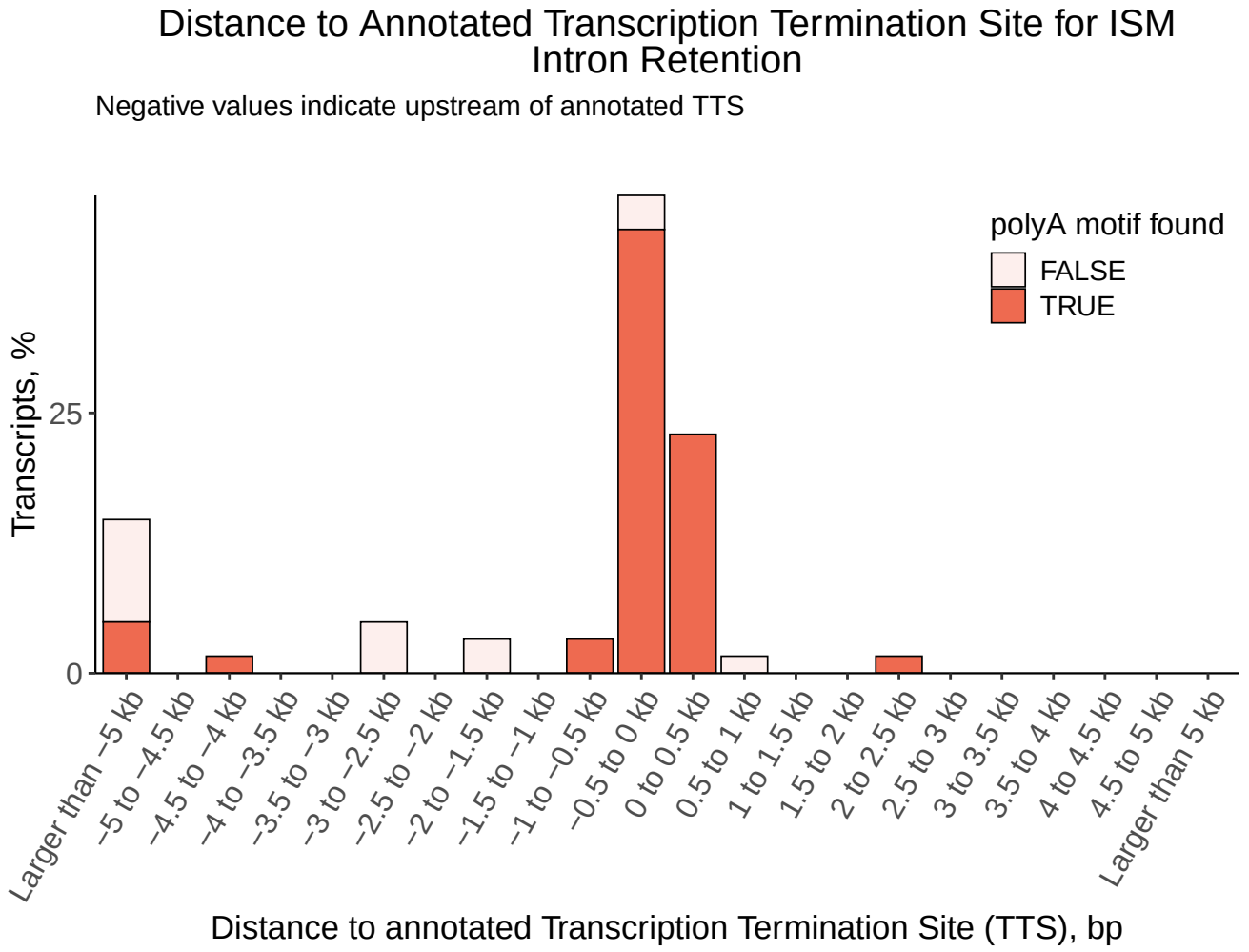

Distance to Annotated Transcription Termination Site for ISM Internal Fragment
Negative values indicate upstream of annotated TTS
Transcripts, count
10.0
7.5
5.0
2.5
0.0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−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 Transcription Termination Site (TTS), bp


Distance to Annotated Transcription Termination Site for ISM Internal Fragment
Negative values indicate upstream of annotated TTS
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 Transcription Termination Site (TTS), bp

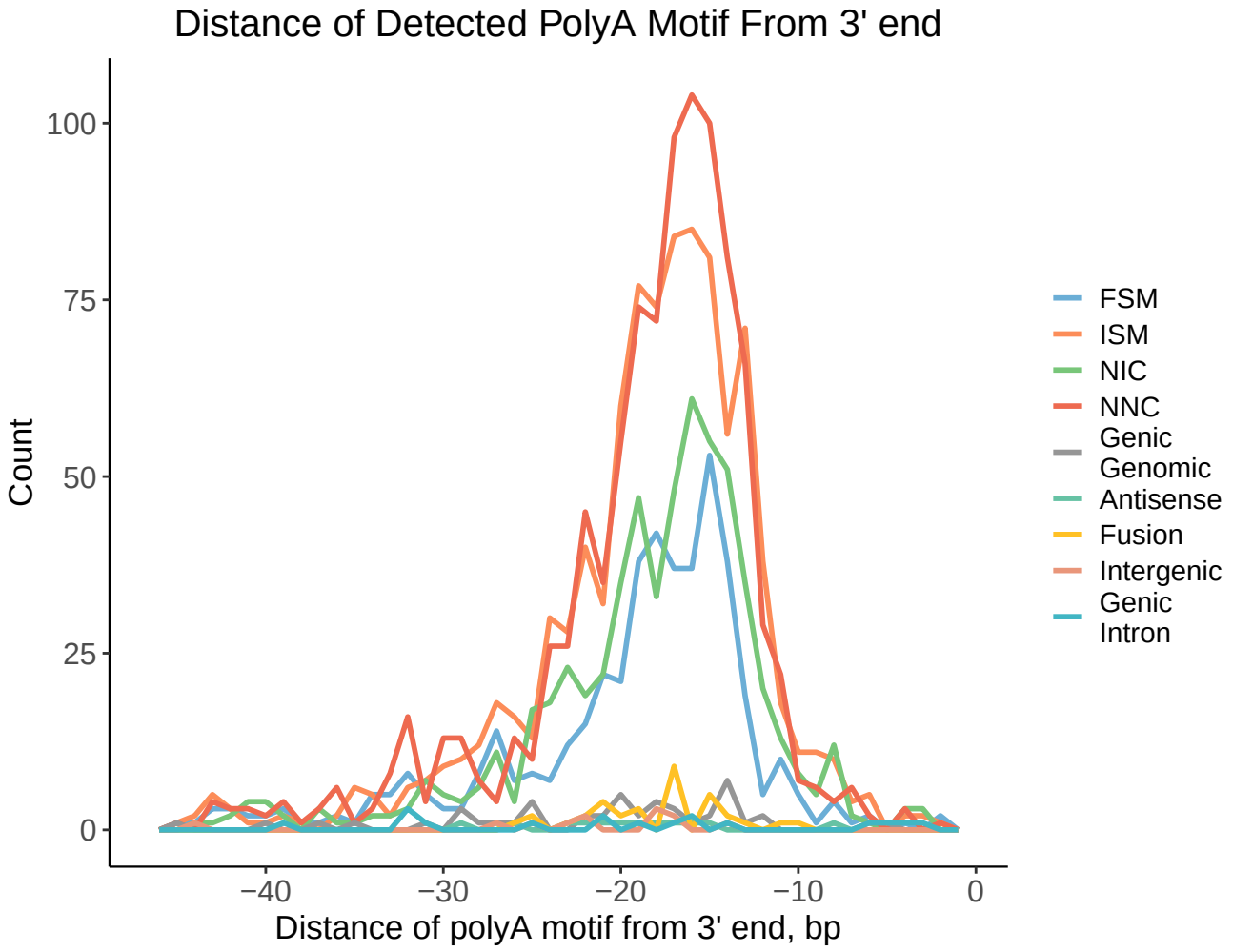

Distance to Annotated Transcription Termination Site for ISM A5' Fragment
Negative values indicate upstream of annotated TTS
Transcripts, count
120
90
60
30
0
polyA motif found
FALSE
TRUE
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 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 Transcription Termination Site (TTS), bp


Distance to Annotated Transcription Termination Site for ISM A5' Fragment
Negative values indicate upstream of annotated TTS
polyA motif found
FALSE
TRUE
Transcripts, %
75
50
25
0
Larger than −5 kb
−5 to −4.5 kb
−4.5 to −4 kb
−4 to −3.5 kb
−3.5 to −3 kb
−3 to −2.5 kb
−2.5 to −2 kb
−2 to −1.5 kb
−1.5 to −1 kb
−1 to −0.5 kb
−0.5 to 0 kb
0 to 0.5 kb
0.5 to 1 kb
1 to 1.5 kb
1.5 to 2 kb
2 to 2.5 kb
2.5 to 3 kb
3 to 3.5 kb
3.5 to 4 kb
4 to 4.5 kb
4.5 to 5 kb
Larger than 5 kb
Distance to annotated Transcription Termination Site (TTS), bp


Distance to Annotated Transcription Termination Site for ISM Intron Retention
Negative values indicate upstream of annotated TTS
Transcripts, count
30
20
10
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 Transcription Termination Site (TTS), bp


Distance to Annotated Transcription Termination Site for ISM Intron Retention
Negative values indicate upstream of annotated TTS
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 Transcription Termination Site (TTS), bp

PolyA Distance Analysis

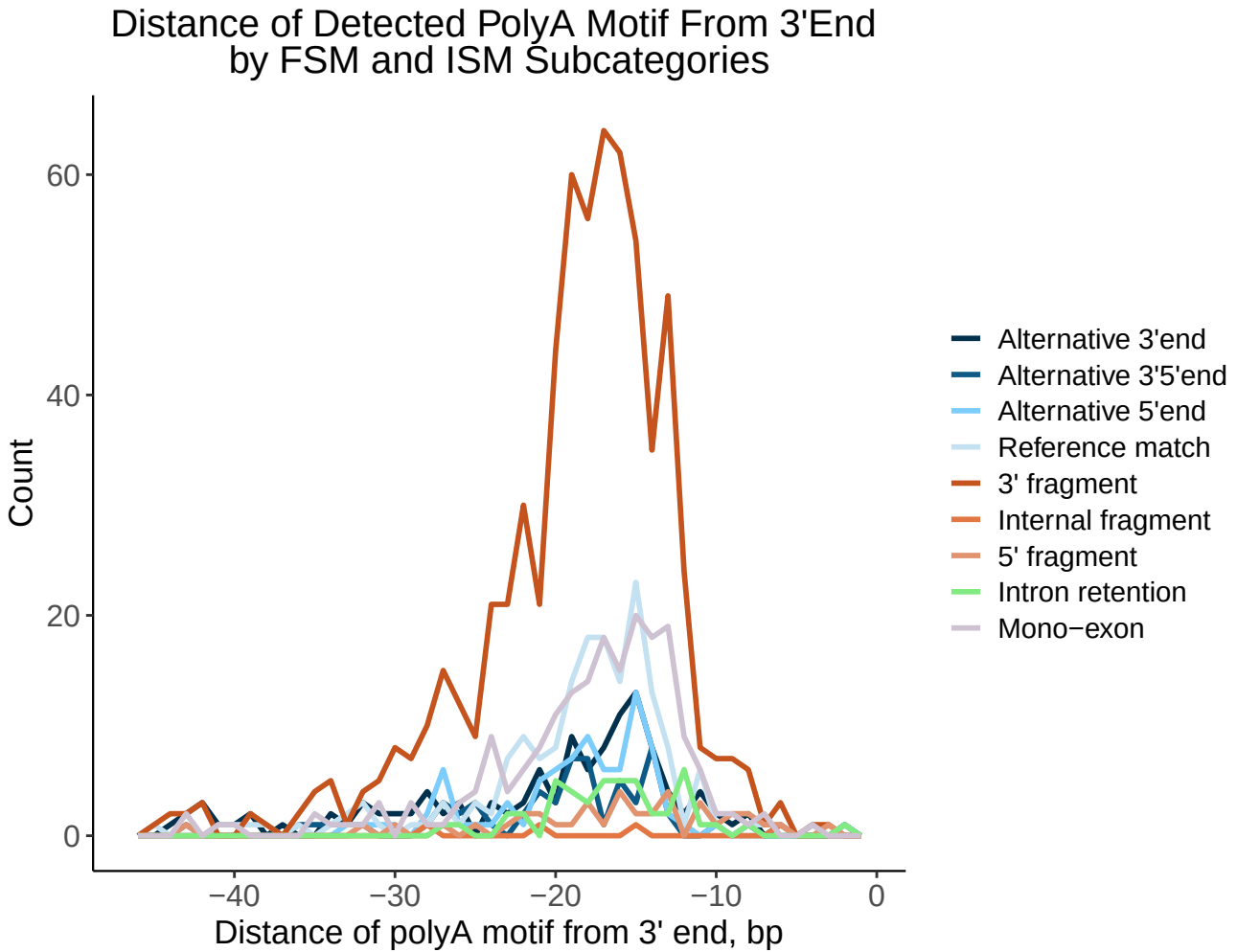

Distance of Detected PolyA Motif From 3' end
Count
100
75
50
25
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

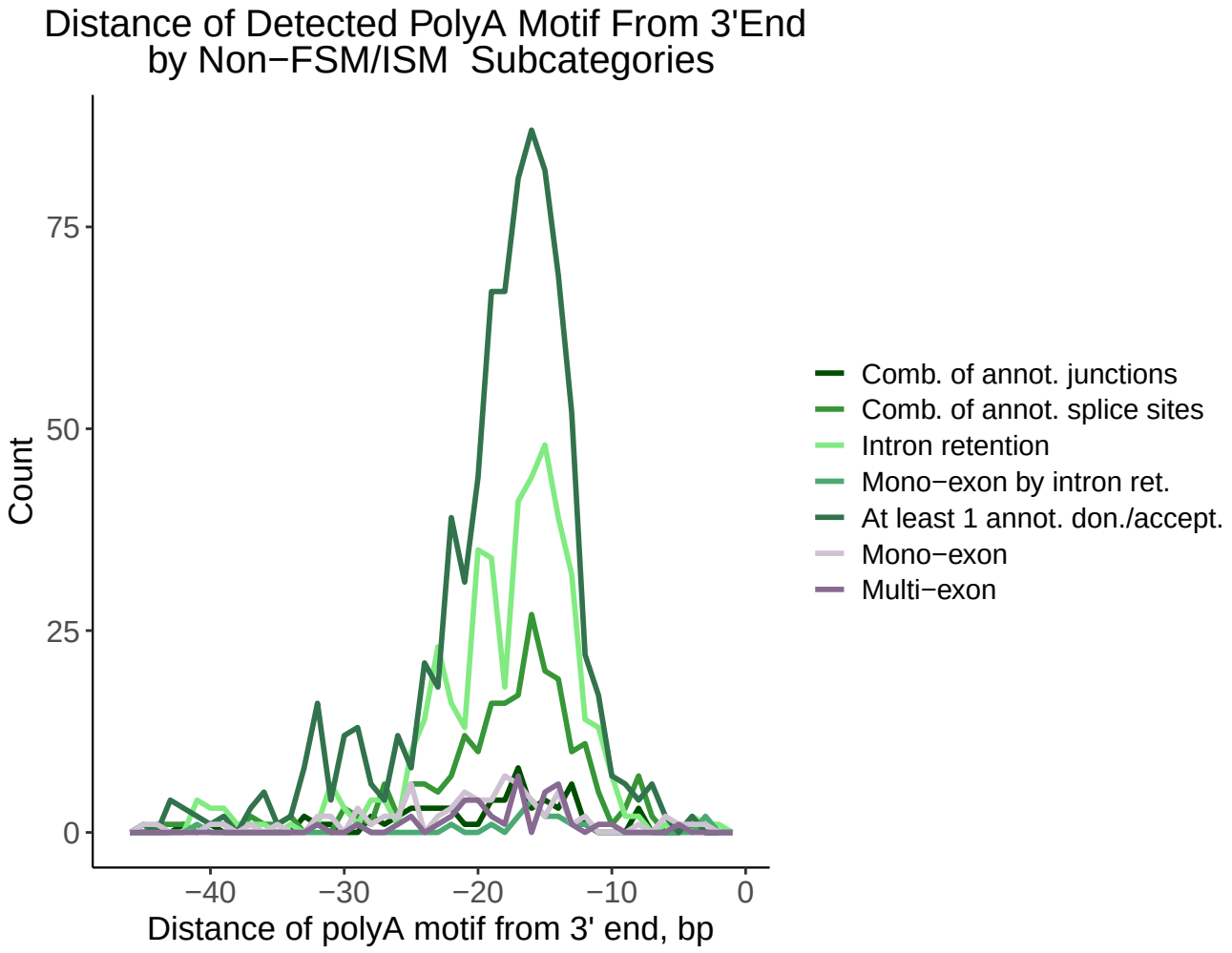
Number of polyA Motifs Detected

Category	Count	polyA Detected	%
FSM	539	458	85
ISM	1138	942	83
NIC	744	594	80
NNC	1139	981	86
Genic Genomic	113	48	42
Antisense	26	12	46
Fusion	38	34	89
Intergenic	58	12	21
Genic Intron	130	17	13

Frequency of PolyA Motifs

Motif	Count	%
AATAAA	2014	65.0
ATTAAA	529	17.1
AGTAAA	161	5.2
TATAAA	90	2.9
GATAAA	54	1.7
AAAACA	41	1.3
AATACA	38	1.2
AAAAAG	32	1.0
AATAGA	28	0.9
CATAAA	26	0.8
AAGAAA	24	0.8
AATATA	20	0.6
TTTAAA	16	0.5
AATGAA	13	0.4
ACTAAA	10	0.3
GGGGCT	2	0.1


Distance of Detected PolyA Motif From 3'End
by FSM and ISM Subcategories
Count
60
40
20
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
Alternative 3'end
Alternative 3'5'end
Alternative 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon


Distance of Detected PolyA Motif From 3'End
by Non−FSM/ISM Subcategories
Count
75
50
25
0
−40
−30
−20
−10
0
Distance of polyA motif from 3' end, bp
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
At least 1 annot. don./accept.
Mono−exon
Multi−exon

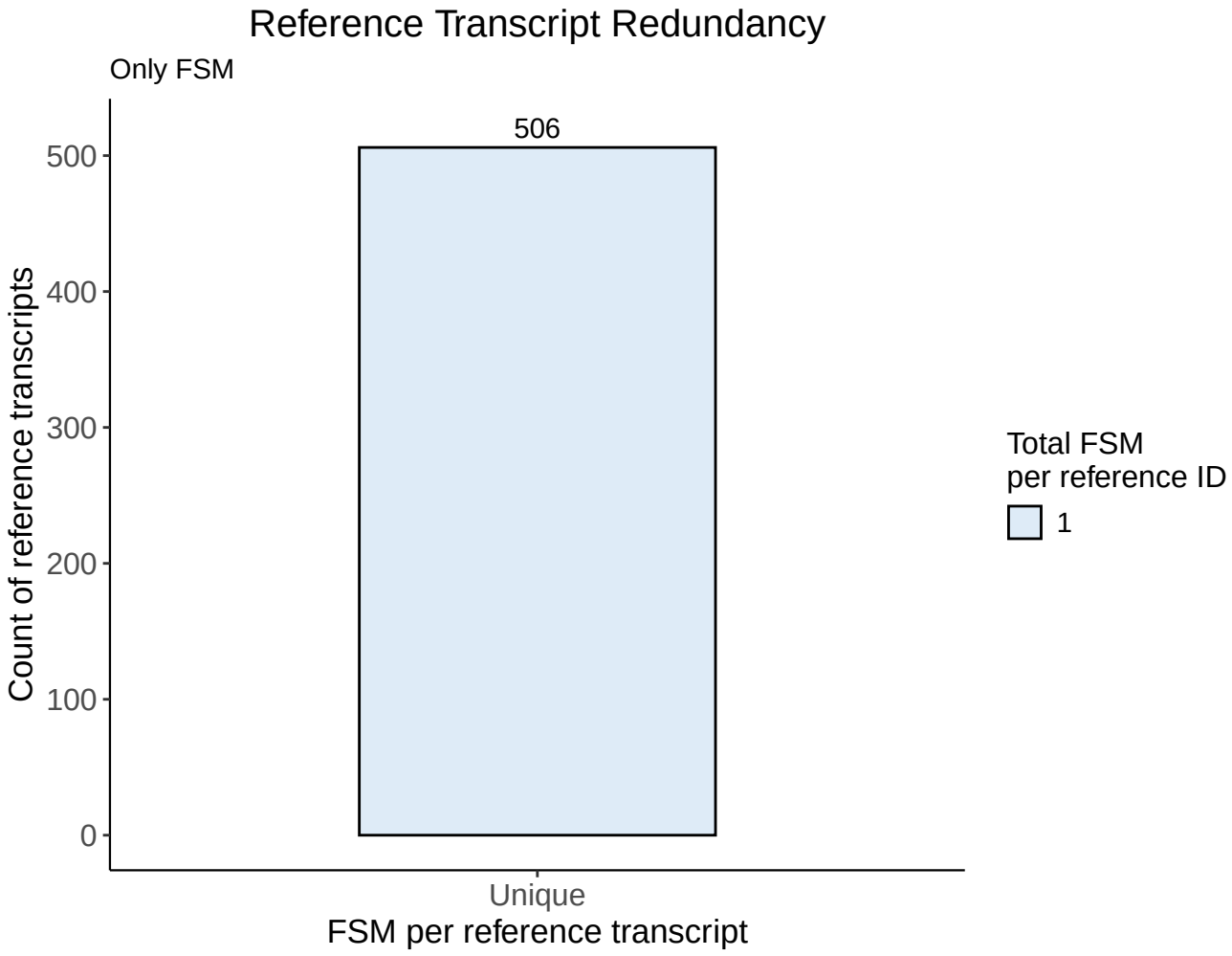
Number of polyA Motifs Detected

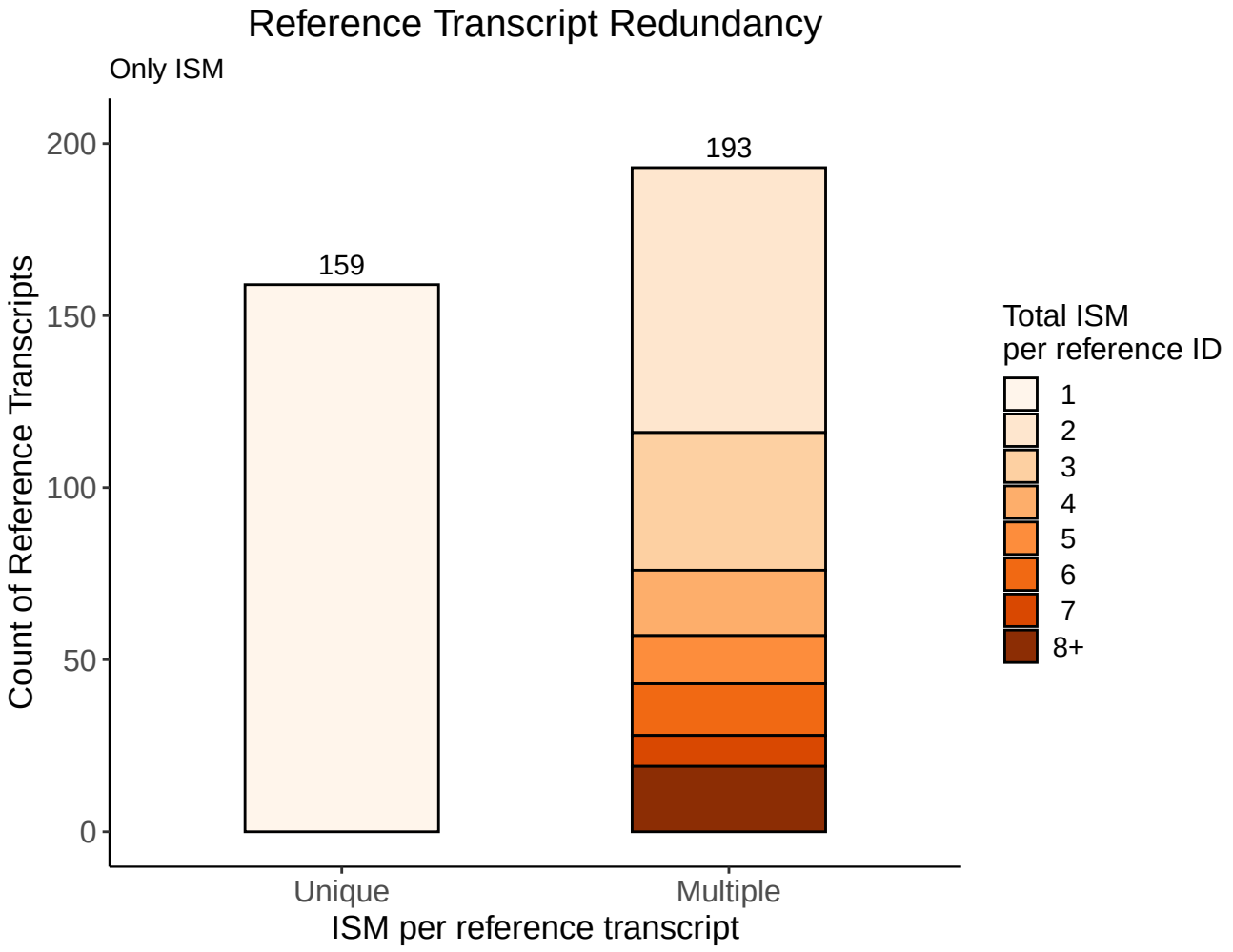
Subcategory	Count	polyA Detected	%
Alternative 3'end	152	121	80
Alternative 3'5'end	76	60	79
Alternative 5'end	90	84	93
Reference match	188	175	93
3' fragment	725	668	92
Internal fragment	8	4	50
5' fragment	141	39	28
Comb. of annot. junctions	76	62	82
Comb. of annot. splice sites	272	227	83
Intron retention	594	491	83
Mono−exon by intron ret.	44	18	41
At least 1 annot. don./accept.	965	831	86
Mono−exon	543	277	51
Multi−exon	51	41	80

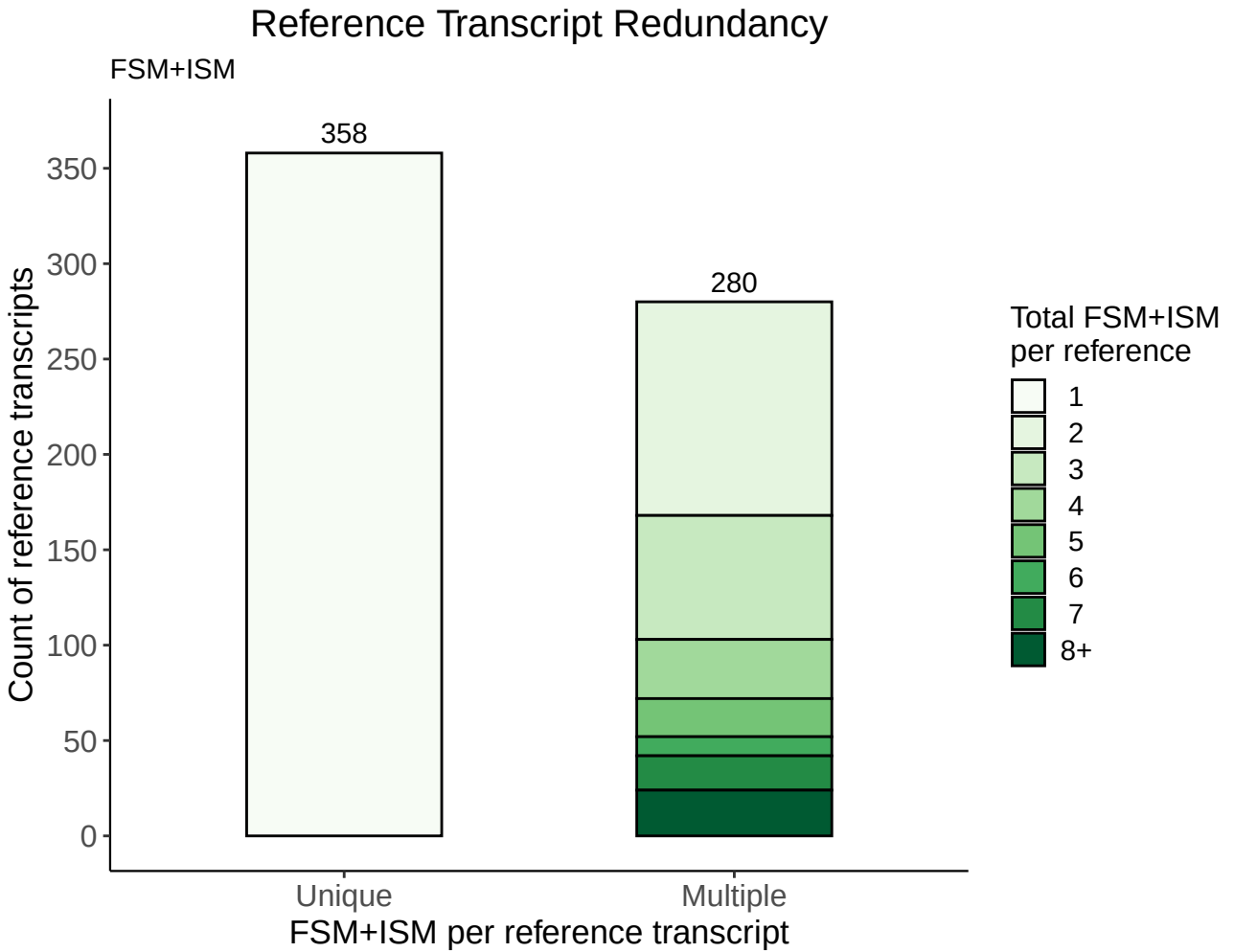
Frequency of PolyA Motifs

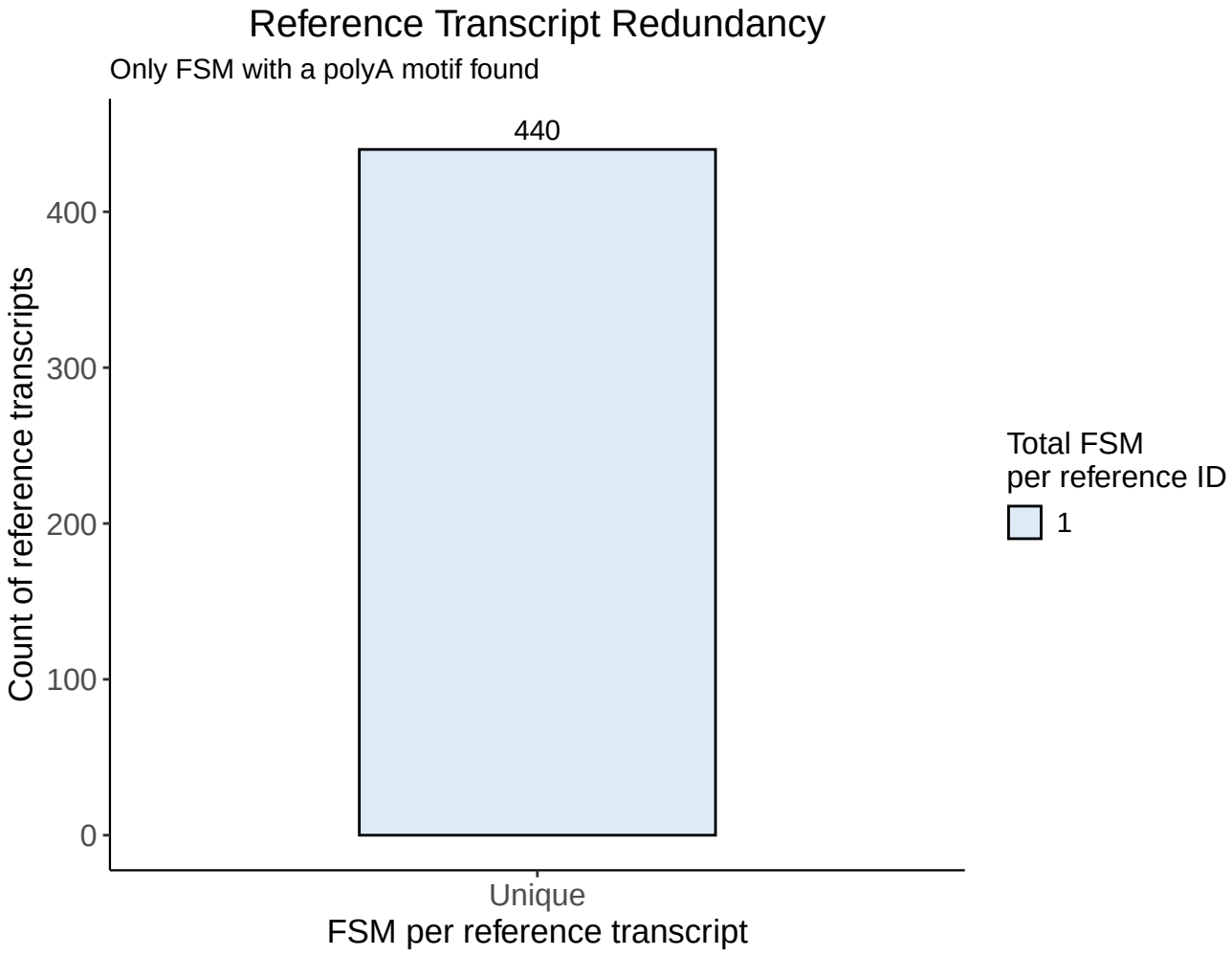
Motif	Count	%
AATAAA	2014	65.0
ATTAAA	529	17.1
AGTAAA	161	5.2
TATAAA	90	2.9
GATAAA	54	1.7
AAAACA	41	1.3
AATACA	38	1.2
AAAAAG	32	1.0
AATAGA	28	0.9
CATAAA	26	0.8
AAGAAA	24	0.8
AATATA	20	0.6
TTTAAA	16	0.5
AATGAA	13	0.4
ACTAAA	10	0.3
GGGGCT	2	0.1

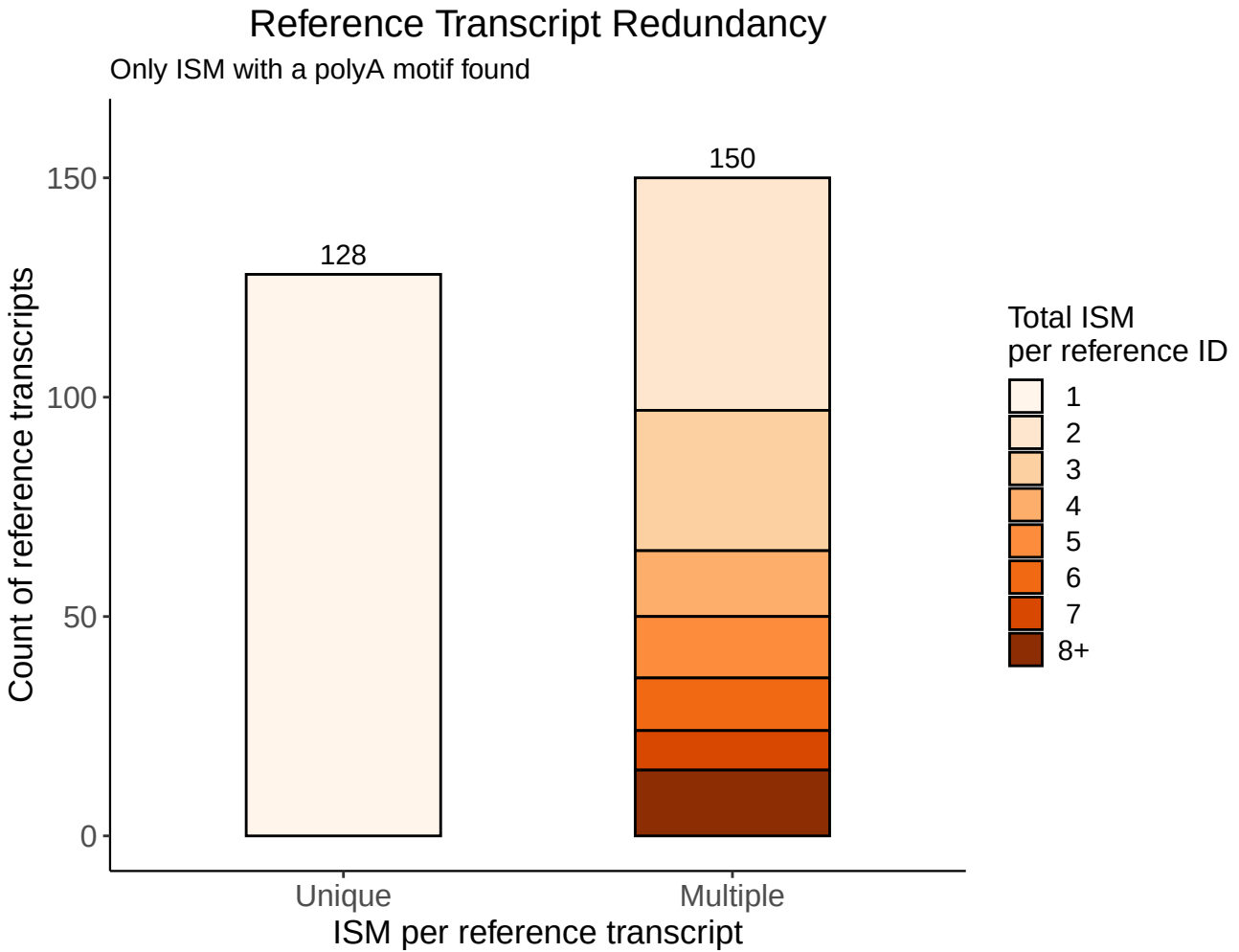
Redundancy Analysis

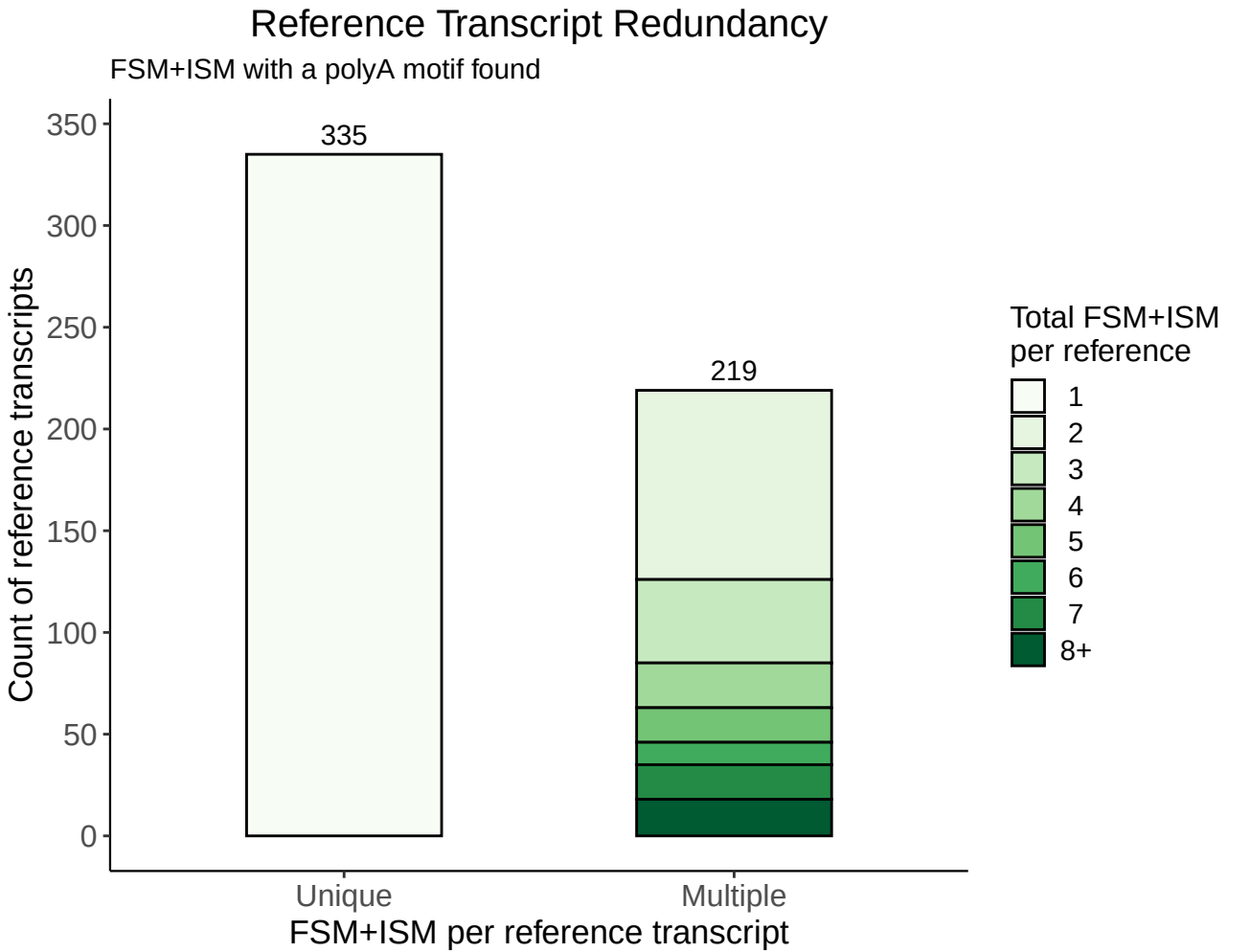

Reference Transcript Redundancy
Only FSM
506
500
400
300
200
100
0
Count of reference transcripts
Unique
FSM per reference transcript
Total FSM
per reference ID
1


Reference Transcript Redundancy
Only ISM
Count of Reference Transcripts
200
150
100
50
0
159
193
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

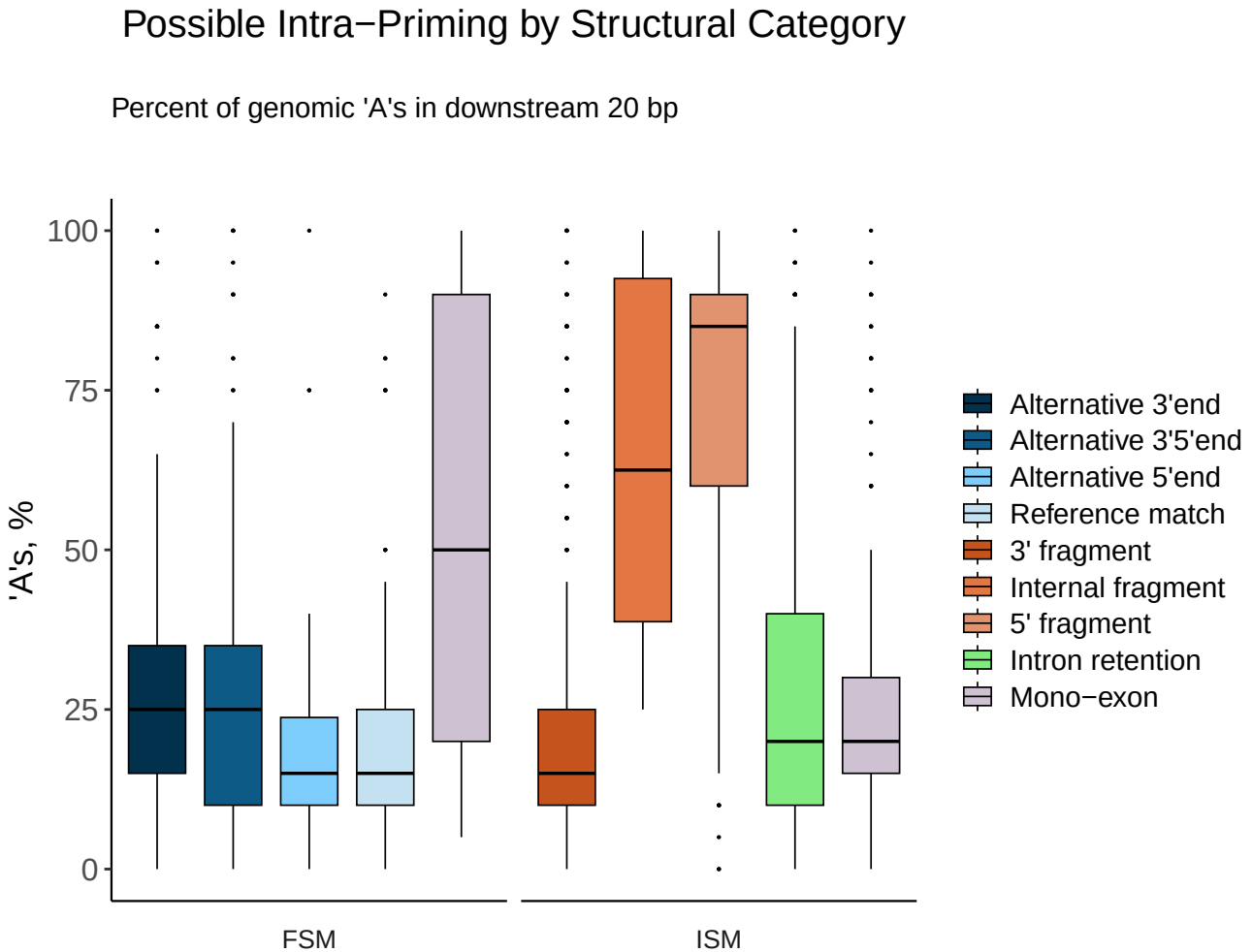

Reference Transcript Redundancy
FSM+ISM
358
280
350
300
250
200
150
100
50
0
Count of reference transcripts
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+

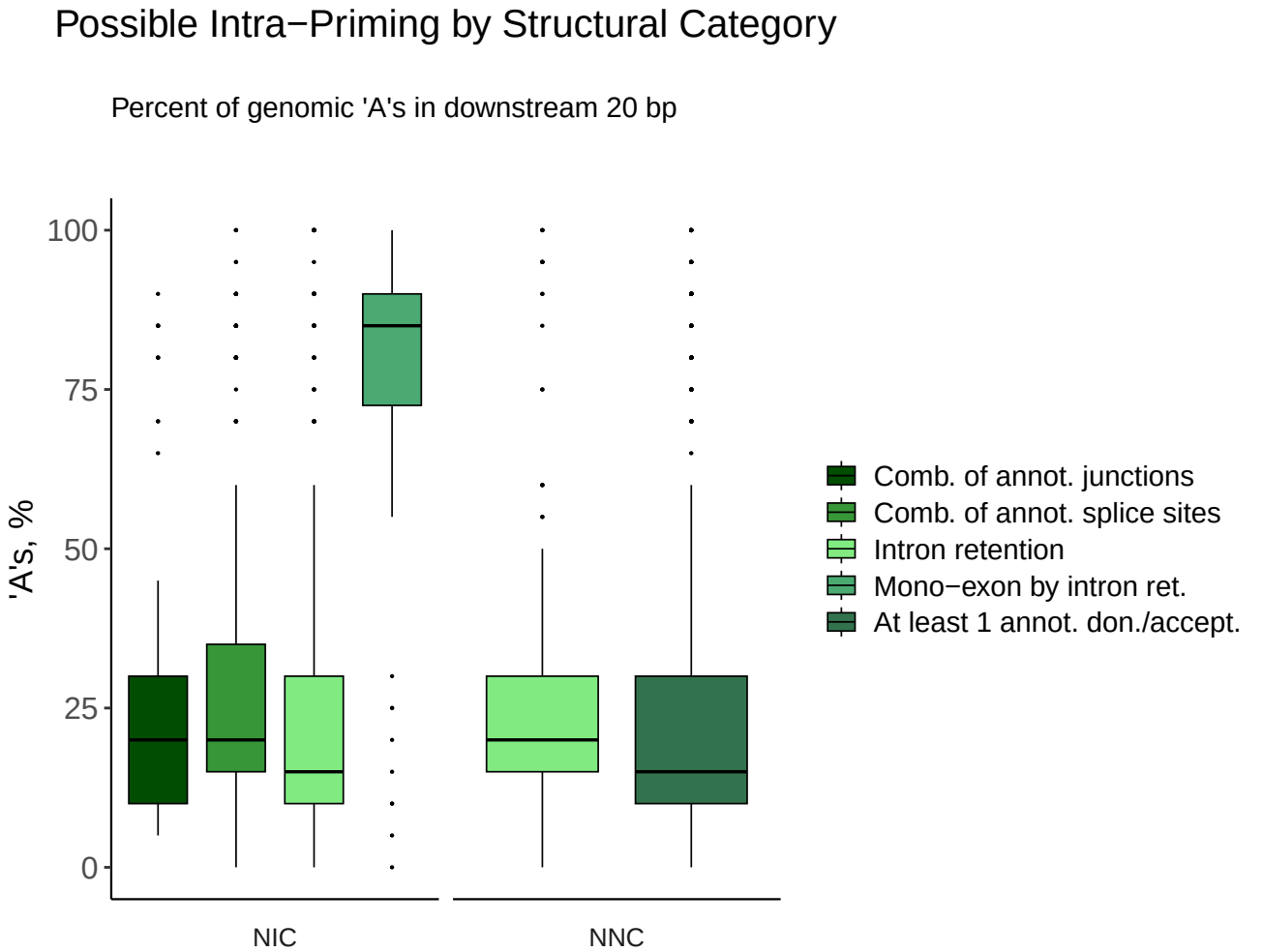

Reference Transcript Redundancy
Only FSM with a polyA motif found
440
400
300
200
100
0
Count of reference transcripts
Unique
FSM per reference transcript
Total FSM
per reference ID
1

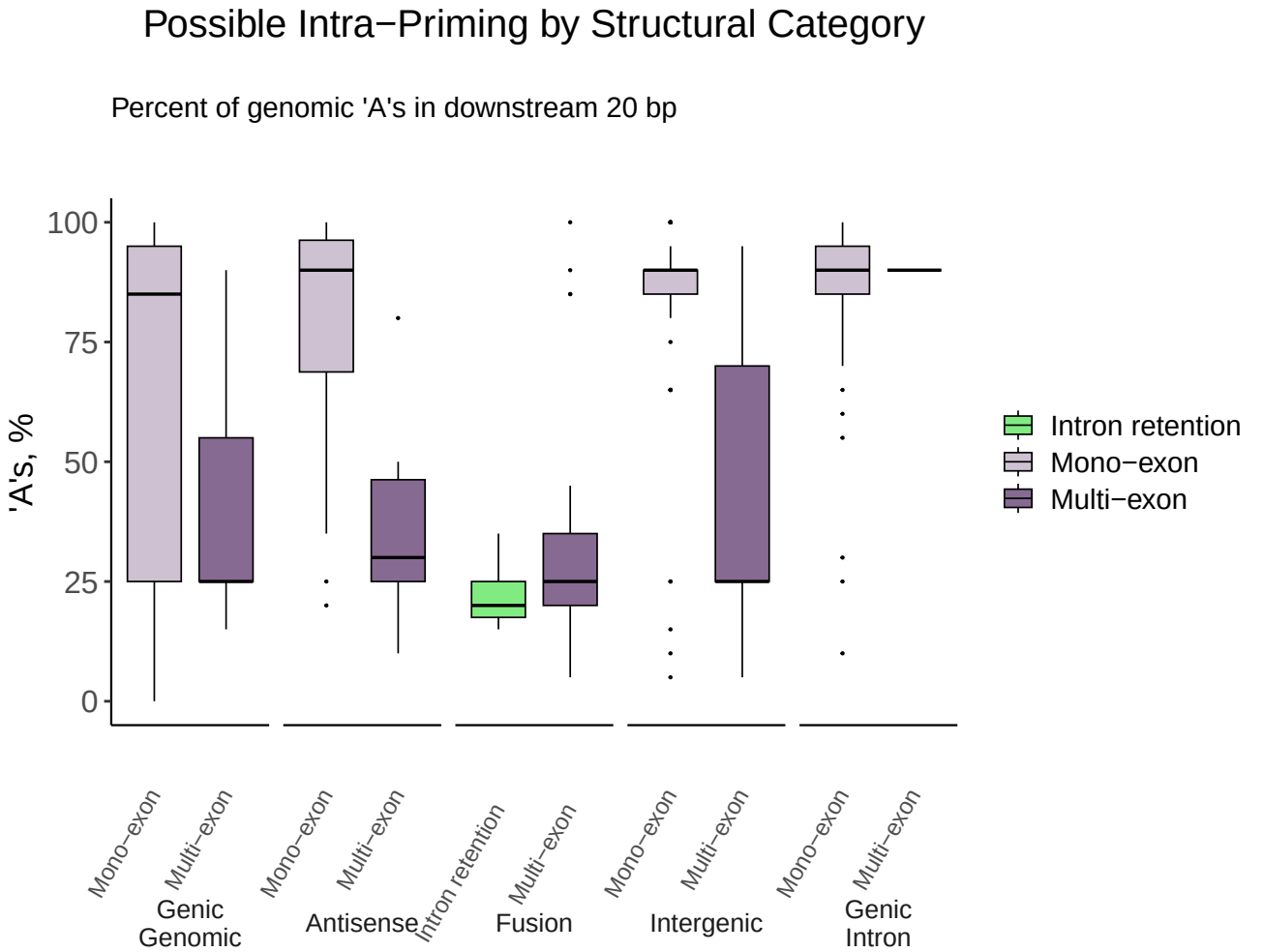

Reference Transcript Redundancy
Only ISM with a polyA motif found
150
128
150
100
50
0
Count of reference transcripts
Unique
Multiple
ISM per reference transcript
Total ISM
per reference ID
1
2
3
4
5
6
7
8+

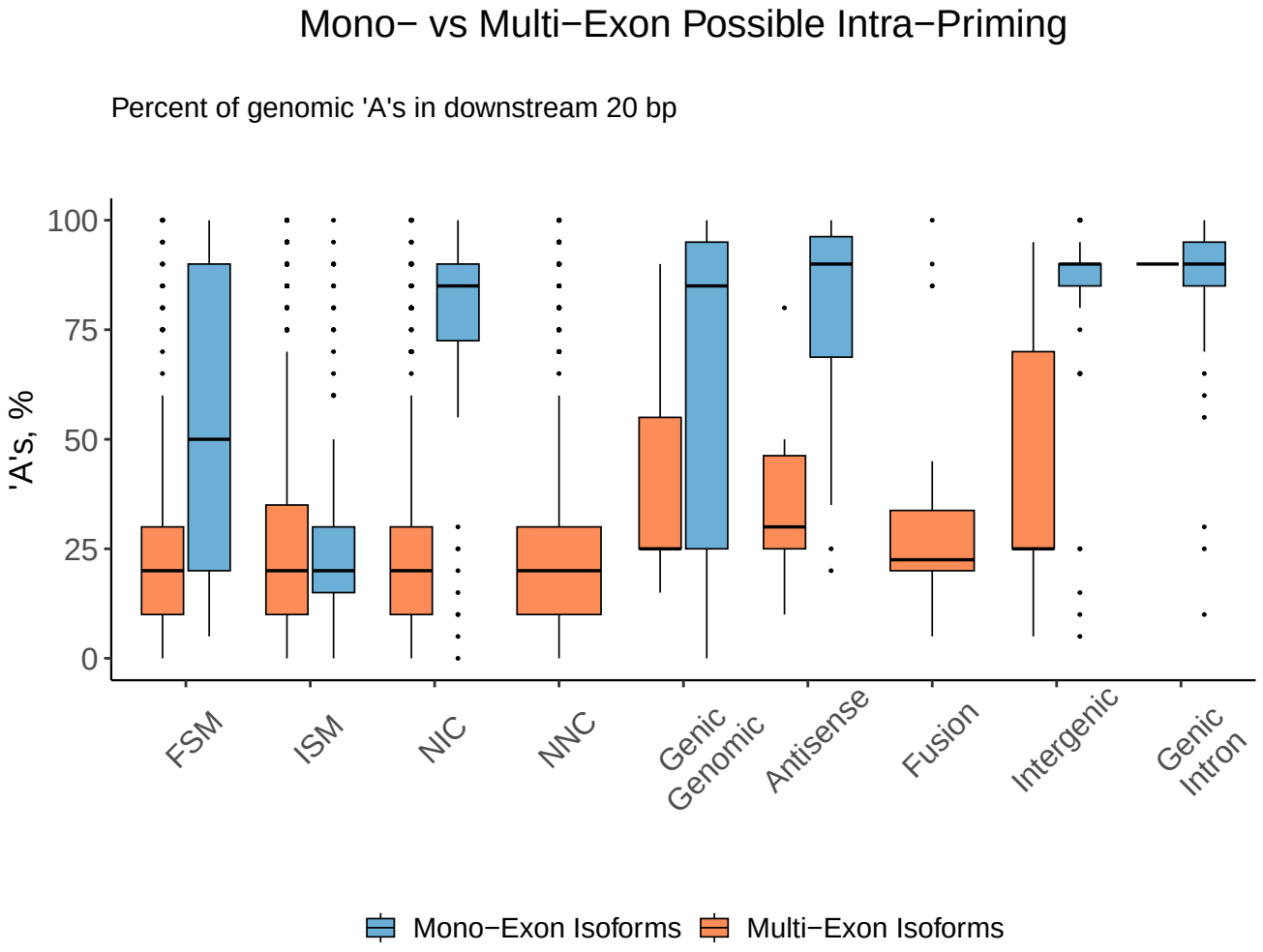

Reference Transcript Redundancy
FSM+ISM with a polyA motif found
Count of reference transcripts
350
300
250
200
150
100
50
0
335
219
Unique
Multiple
FSM+ISM per reference transcript
Total FSM+ISM per reference
1
2
3
4
5
6
7
8+

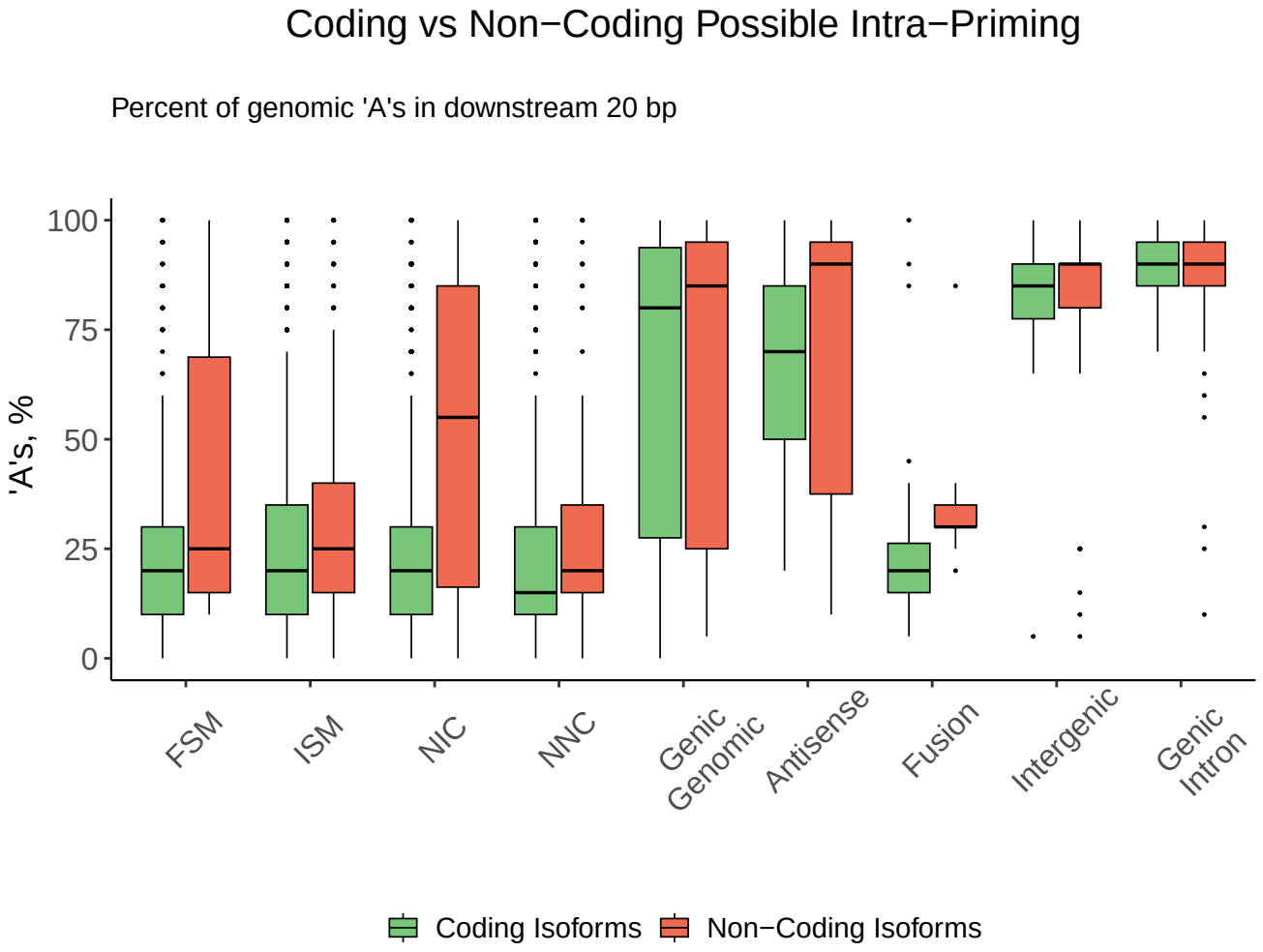
Intra−Priming Quality Check


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
Alternative 5'end
Reference match
3' fragment
Internal fragment
5' fragment
Intron retention
Mono−exon

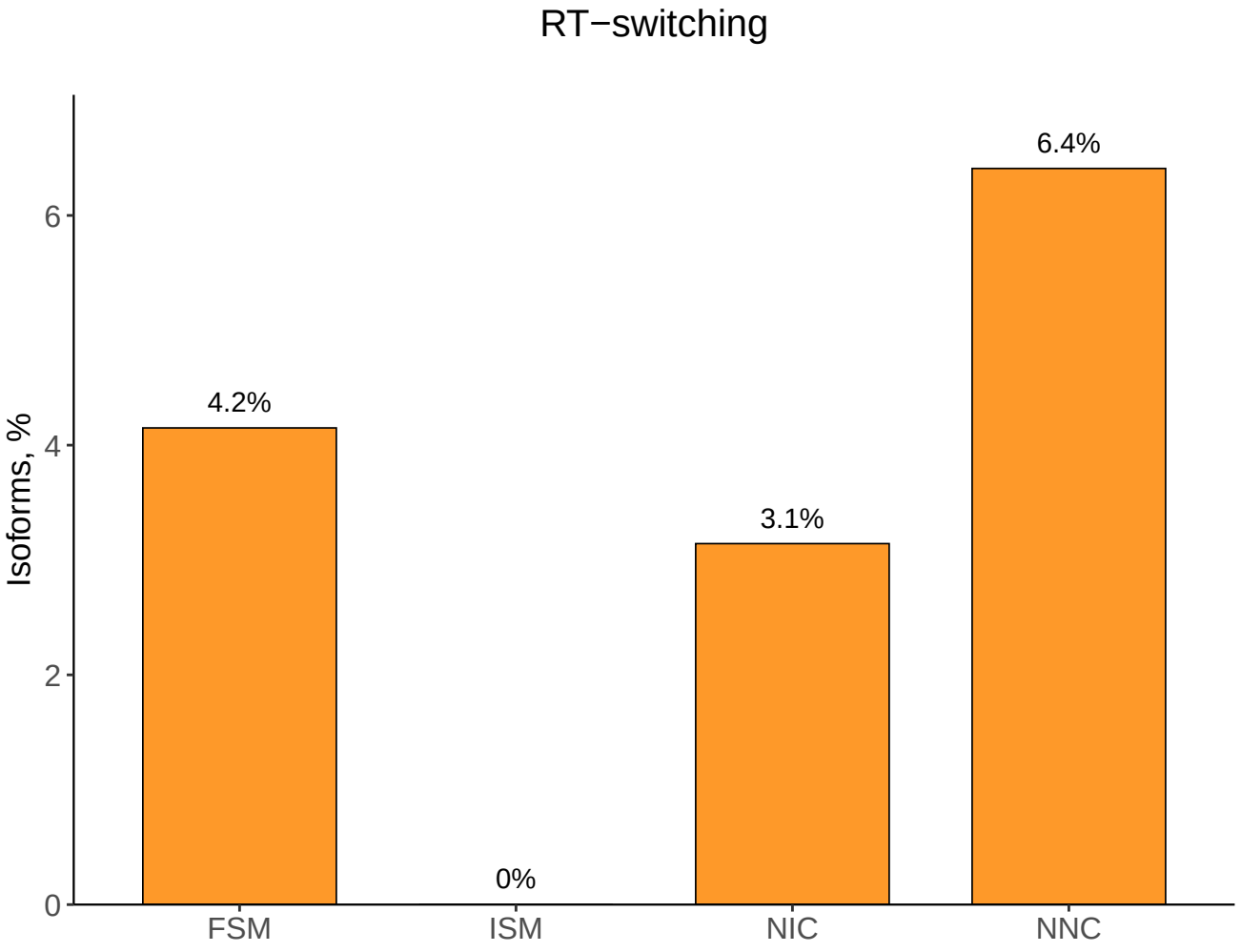

Possible Intra−Priming by Structural Category
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
NIC
NNC
Comb. of annot. junctions
Comb. of annot. splice sites
Intron retention
Mono−exon by intron ret.
At least 1 annot. don./accept.

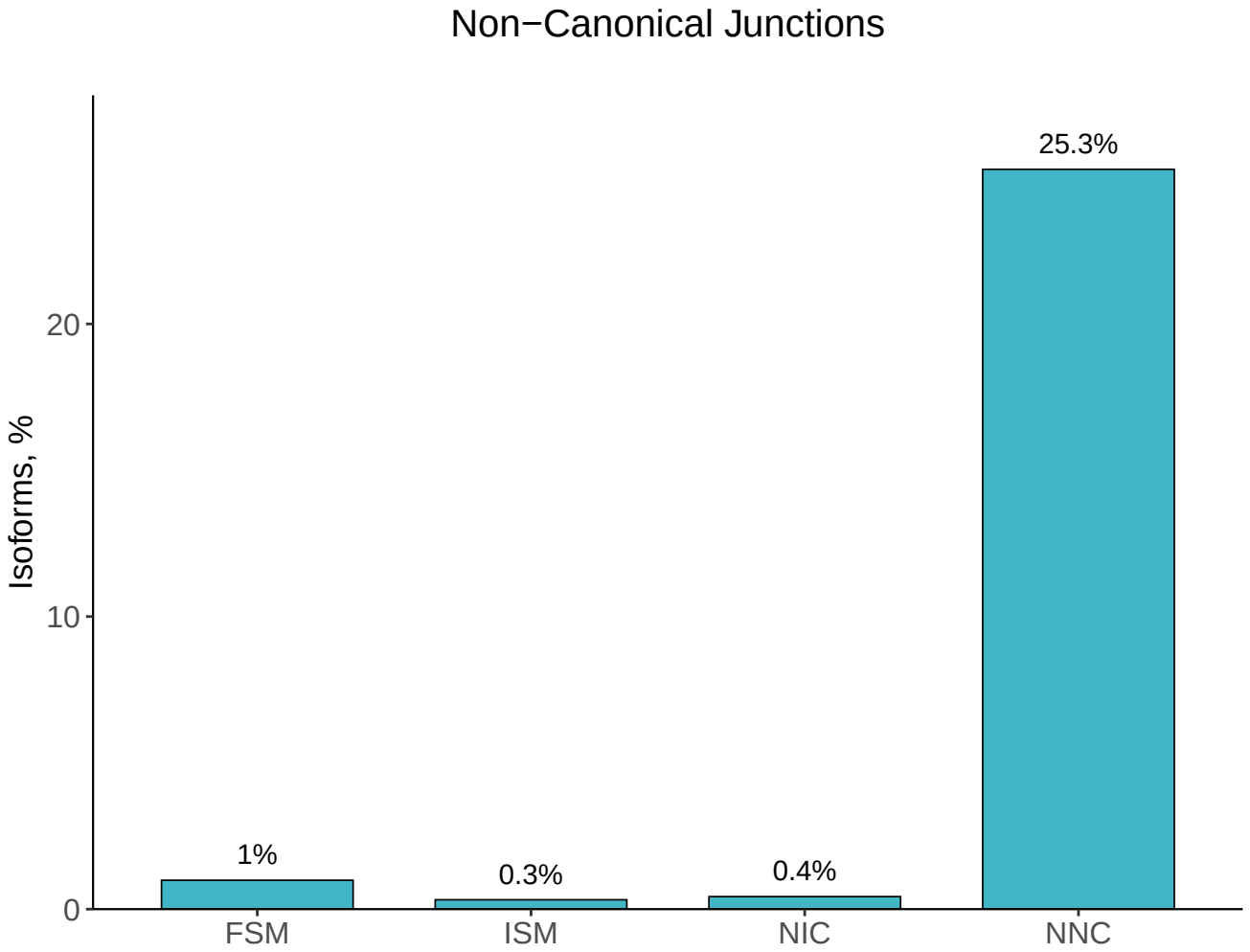

Possible Intra−Priming by Structural Category
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
Mono−exon
Multi−exon
Mono−exon
Multi−exon
Intron retention
Multi−exon
Mono−exon
Multi−exon
Mono−exon
Multi−exon
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Intron retention
Mono−exon
Multi−exon

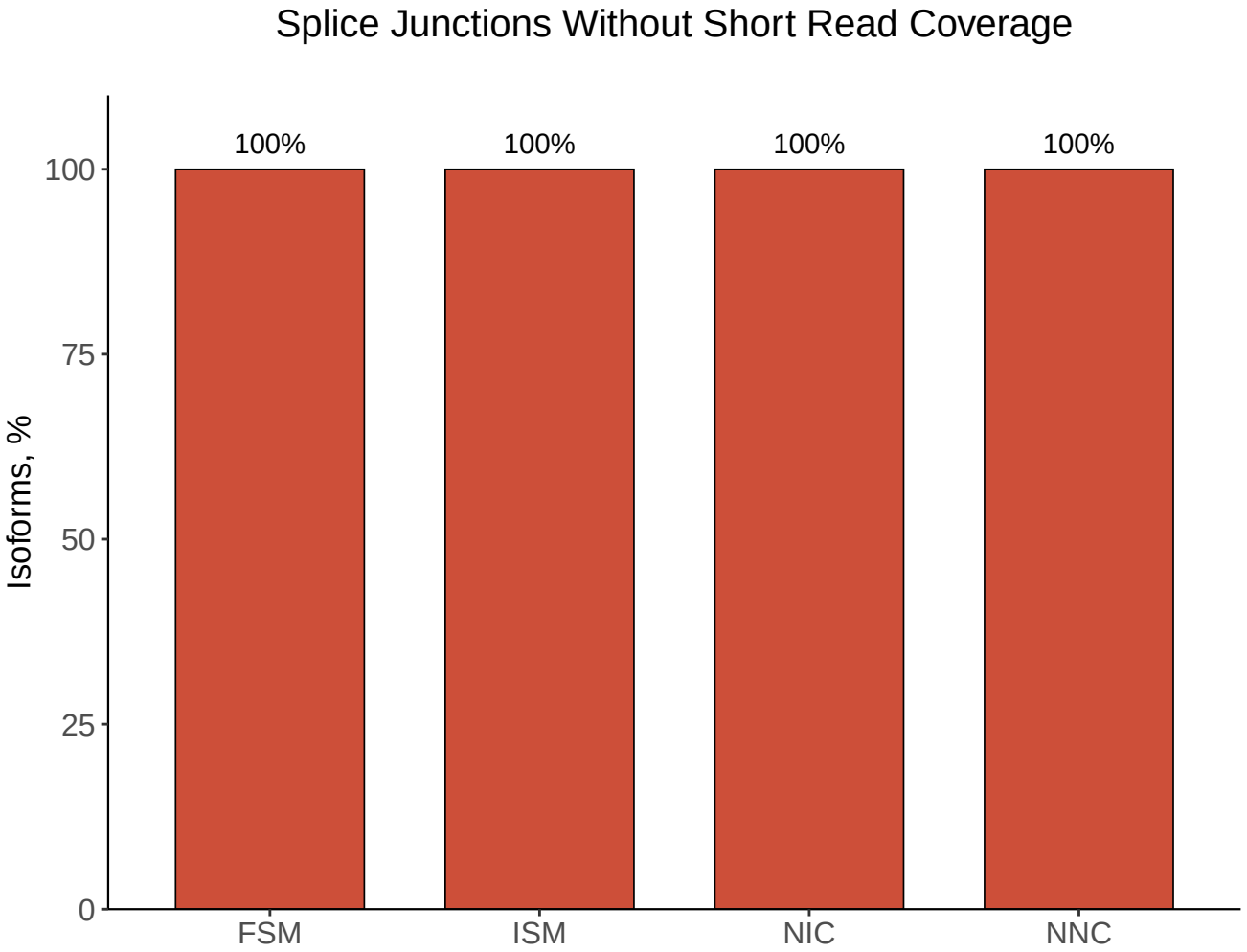

Mono− vs Multi−Exon Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Mono−Exon Isoforms
Multi−Exon Isoforms

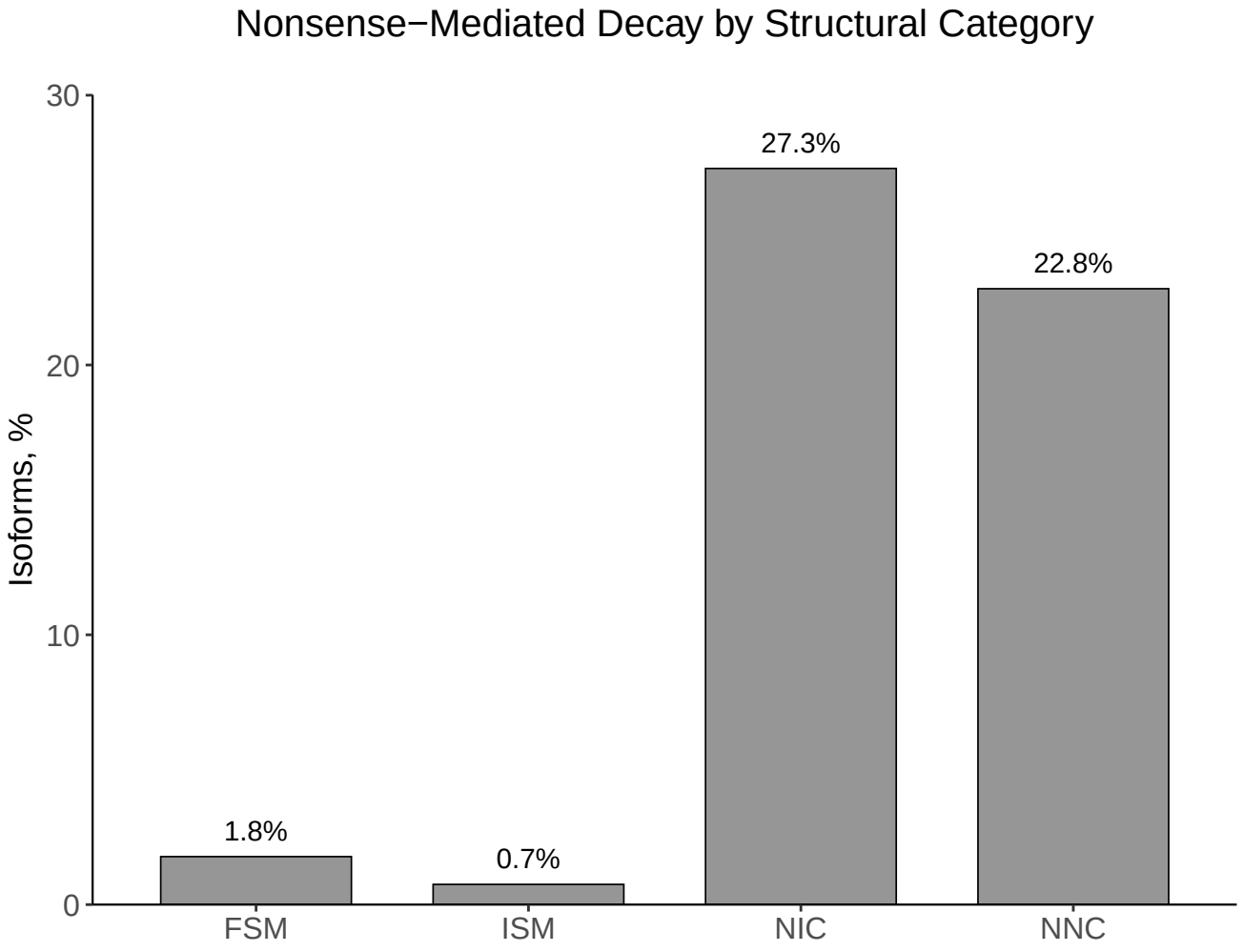

Coding vs Non−Coding Possible Intra−Priming
Percent of genomic 'A's in downstream 20 bp
'A's, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron
Coding Isoforms
Non−Coding Isoforms

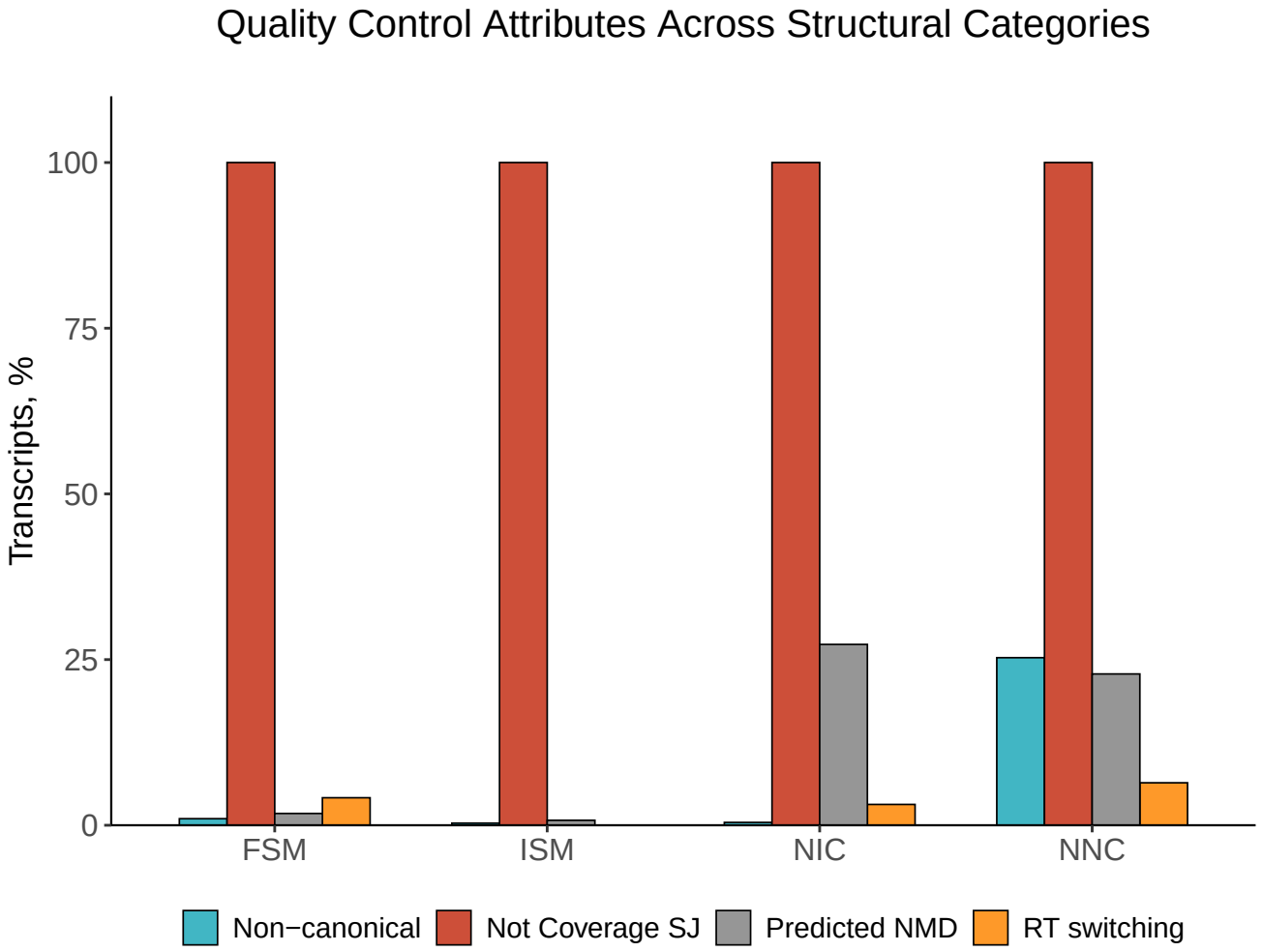
Features of Bad Quality


RT−switching
Isoforms, %
0
2
4
6
4.2%
0%
3.1%
6.4%
FSM
ISM
NIC
NNC

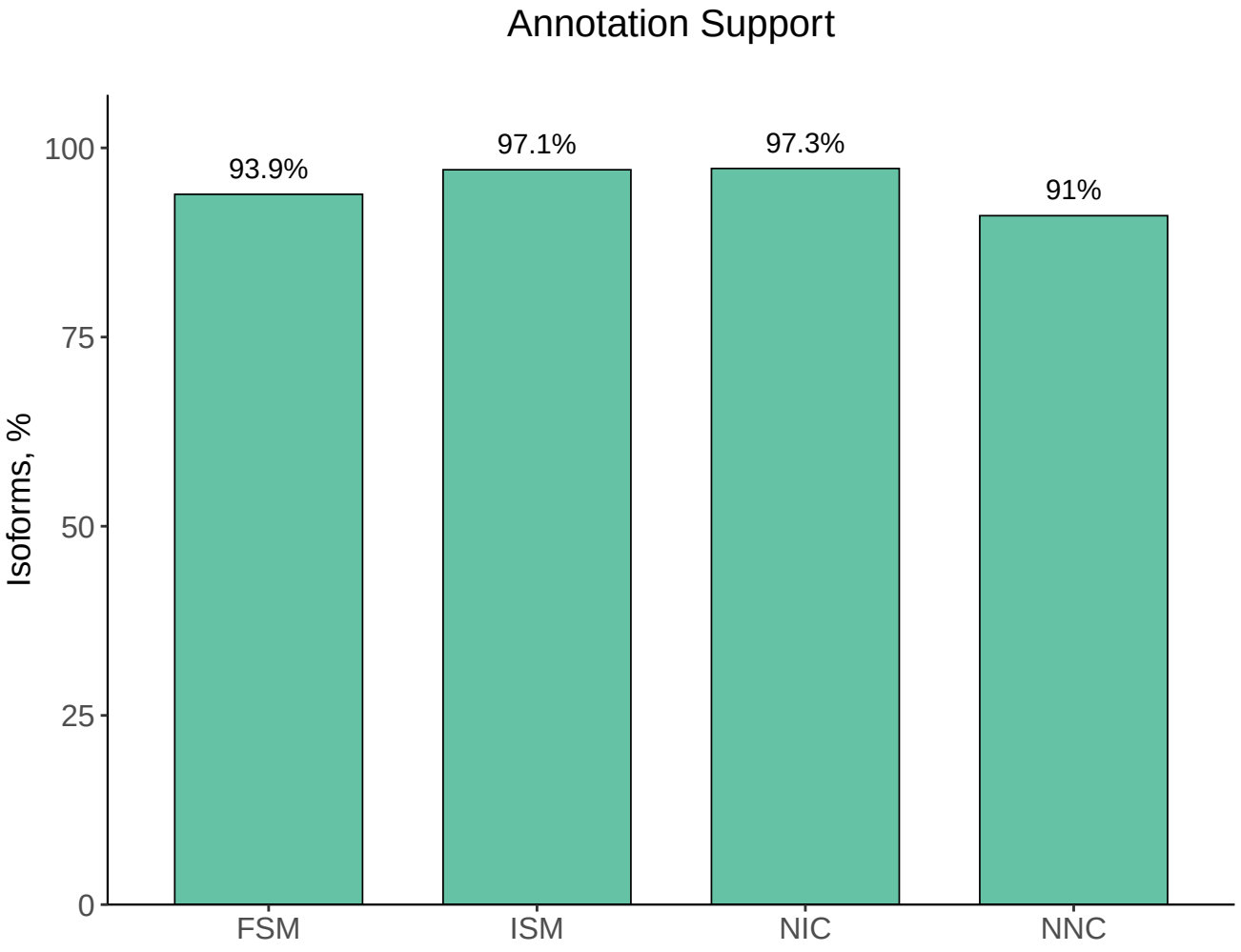

Non−Canonical Junctions
Isoforms, %
0
10
20
1%
0.3%
0.4%
25.3%
FSM
ISM
NIC
NNC

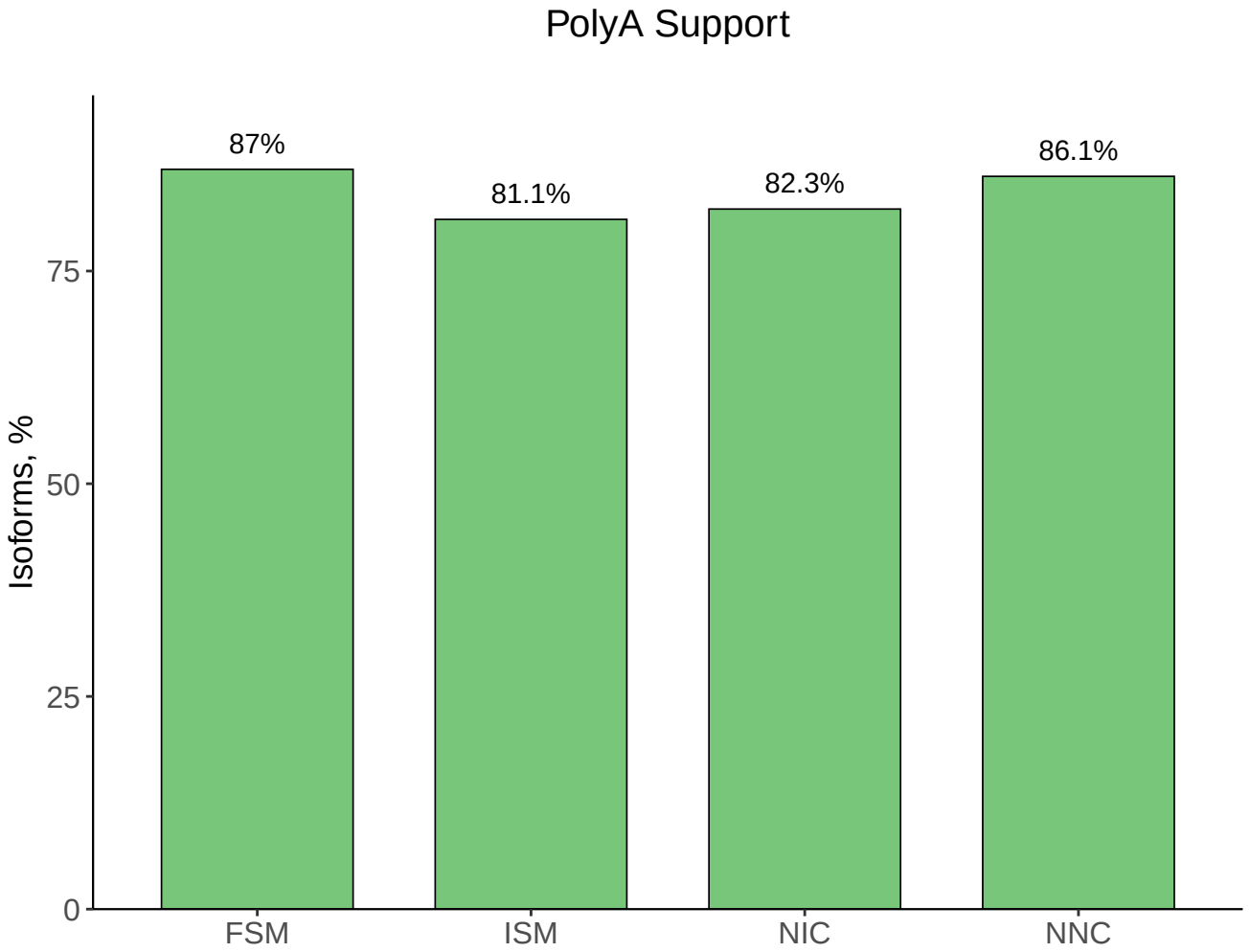

Splice Junctions Without Short Read Coverage
Isoforms, %
100
75
50
25
0
100%
100%
100%
100%
FSM
ISM
NIC
NNC

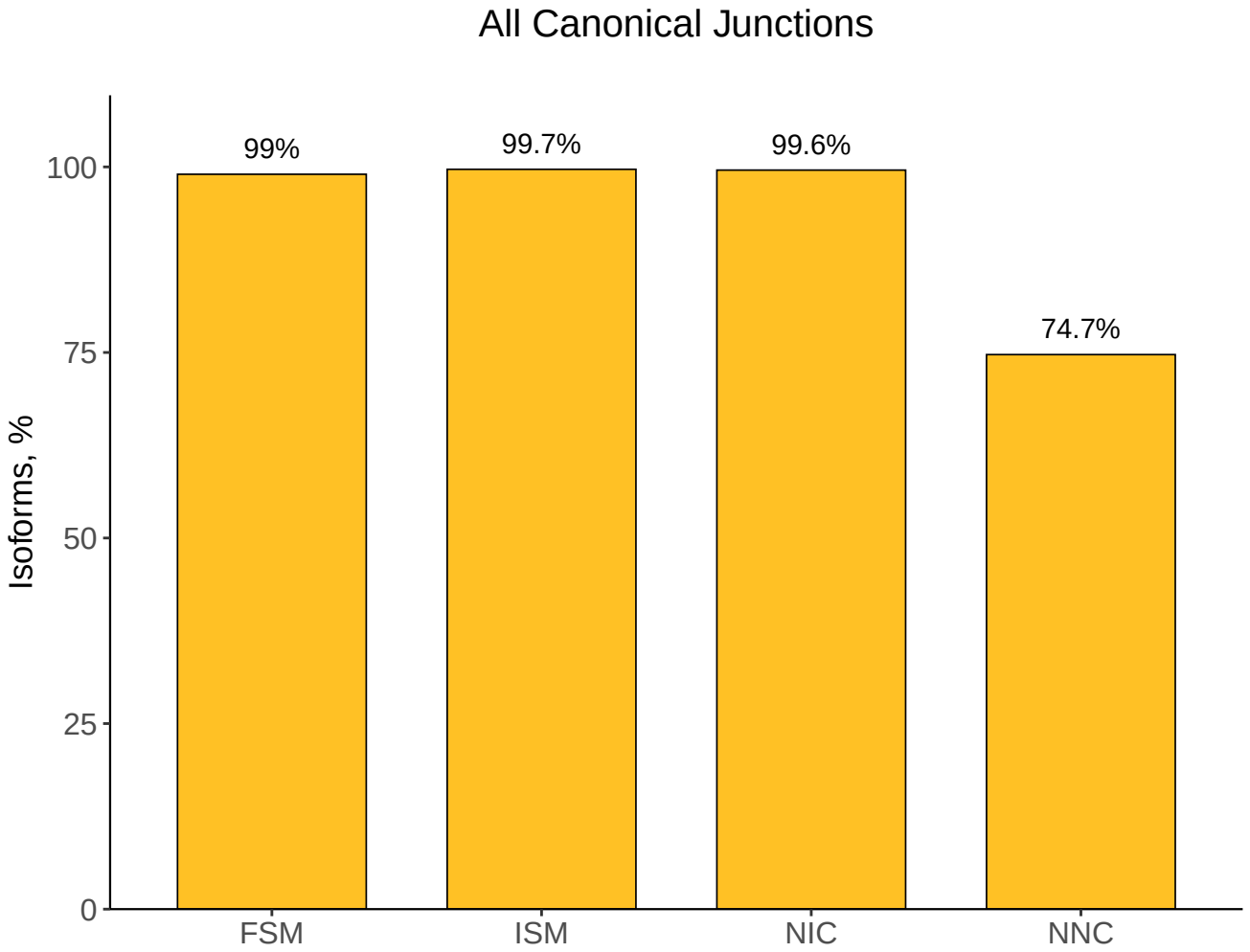

Nonsense−Mediated Decay by Structural Category
Isoforms, %
30
20
10
0
1.8%
0.7%
27.3%
22.8%
FSM
ISM
NIC
NNC


Quality Control Attributes Across Structural Categories
Transcripts, %
100
75
50
25
0
FSM
ISM
NIC
NNC
Non−canonical
Not Coverage SJ
Predicted NMD
RT switching

Features of Good Quality

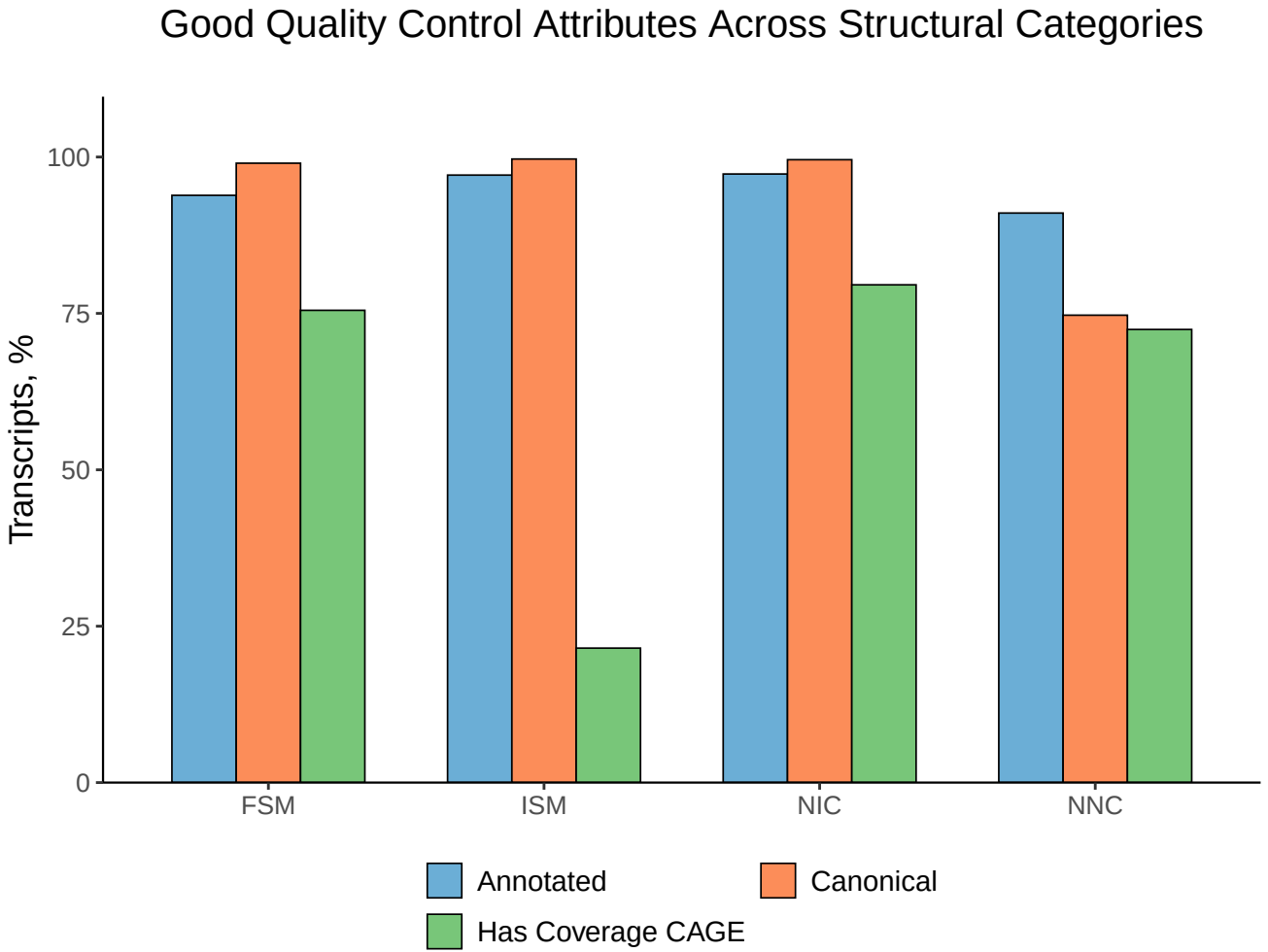

Annotation Support
Isoforms, %
100
75
50
25
0
93.9%
97.1%
97.3%
91%
FSM
ISM
NIC
NNC


PolyA Support
Isoforms, %
0
25
50
75
87%
81.1%
82.3%
86.1%
FSM
ISM
NIC
NNC


All Canonical Junctions
Isoforms, %
100
75
50
25
0
99%
99.7%
99.6%
74.7%
FSM
ISM
NIC
NNC

Splice Junctions With Short Read Coverage

Isoforms, %


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