SQANTI report

 # Genes: 4

 # Isoforms: 8

Gene classification

category	# genes
Annotated Genes	4
Novel Genes	0

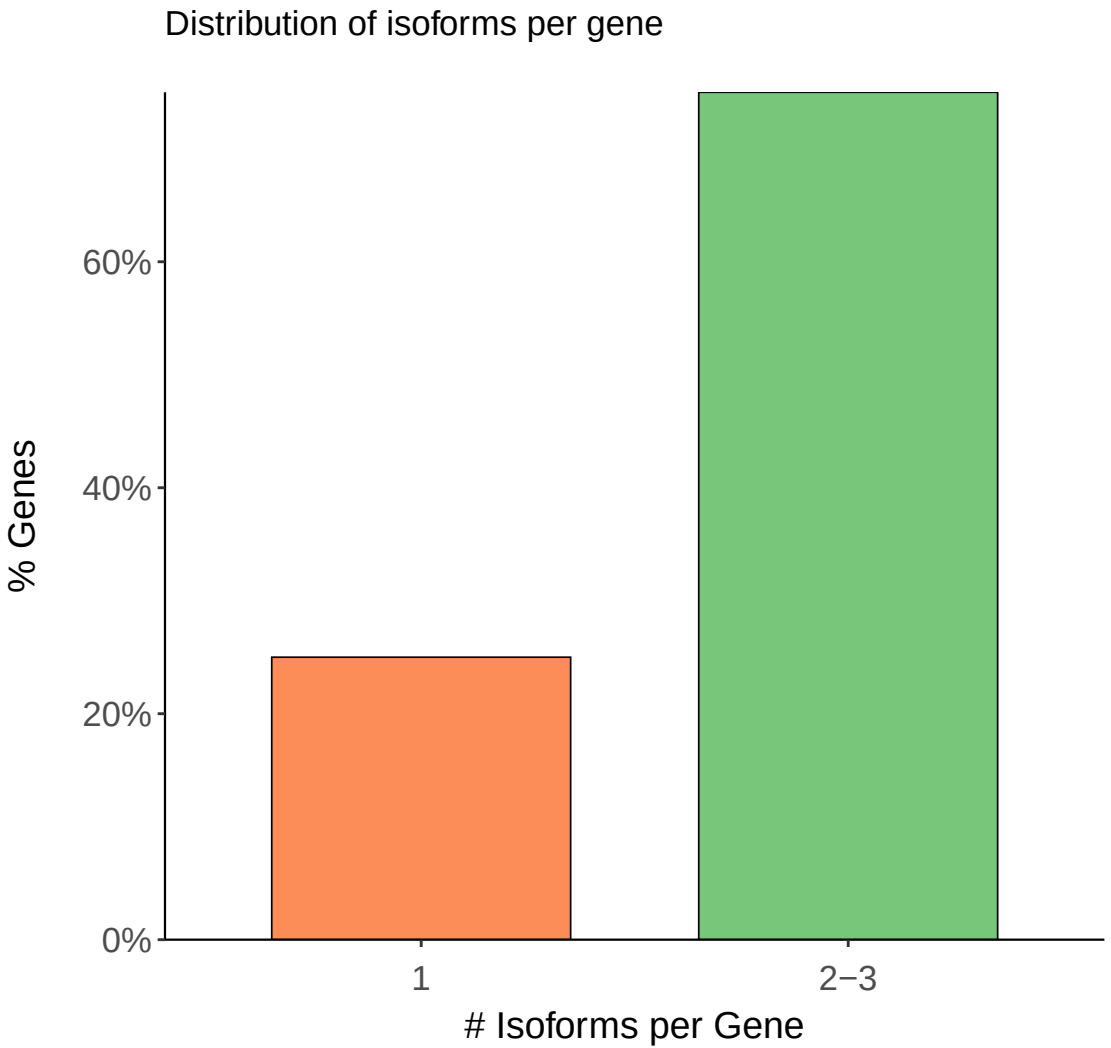
SJ classification

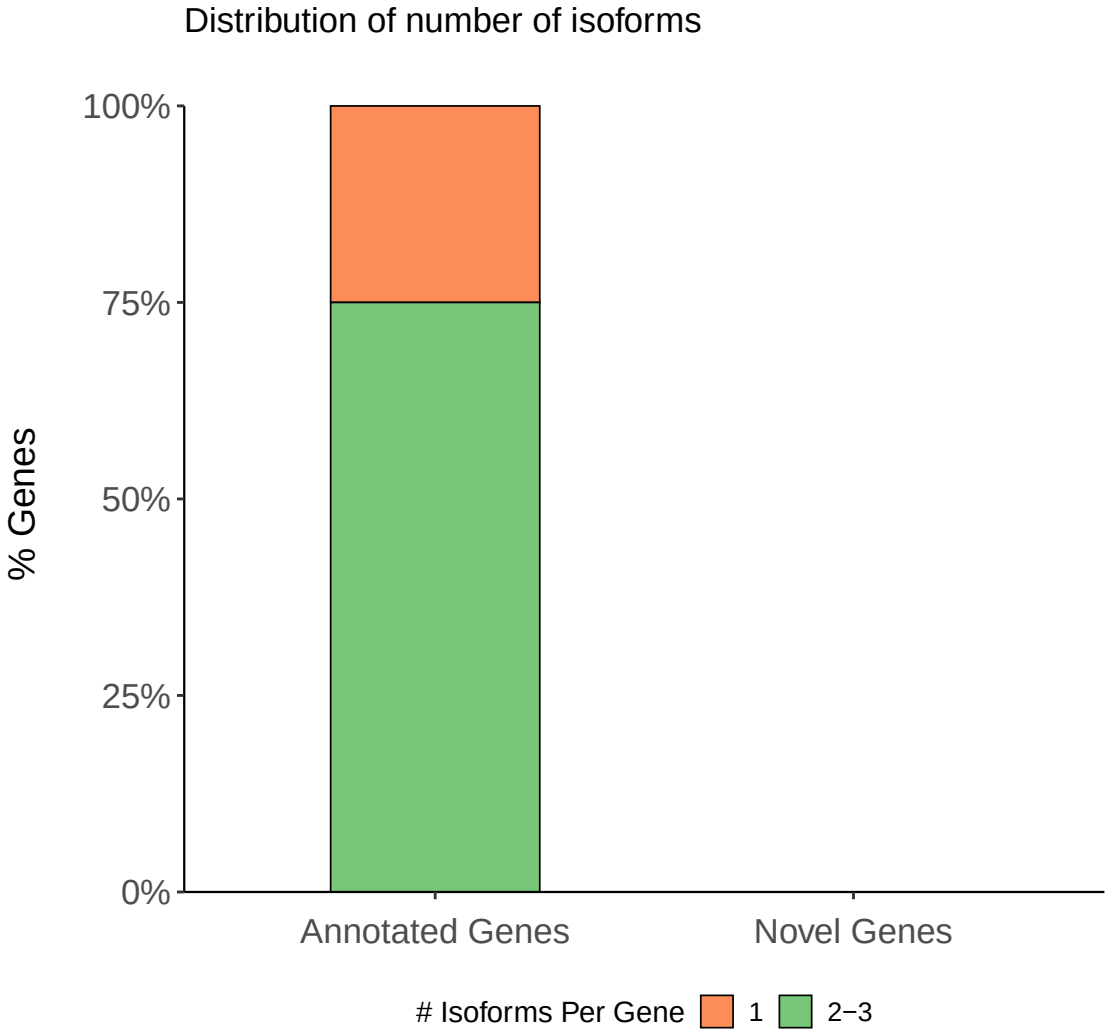
category	# SJ
Known canonical	9
Known Non−canonical	0
Novel canonical	0
Novel Non−canonical	1

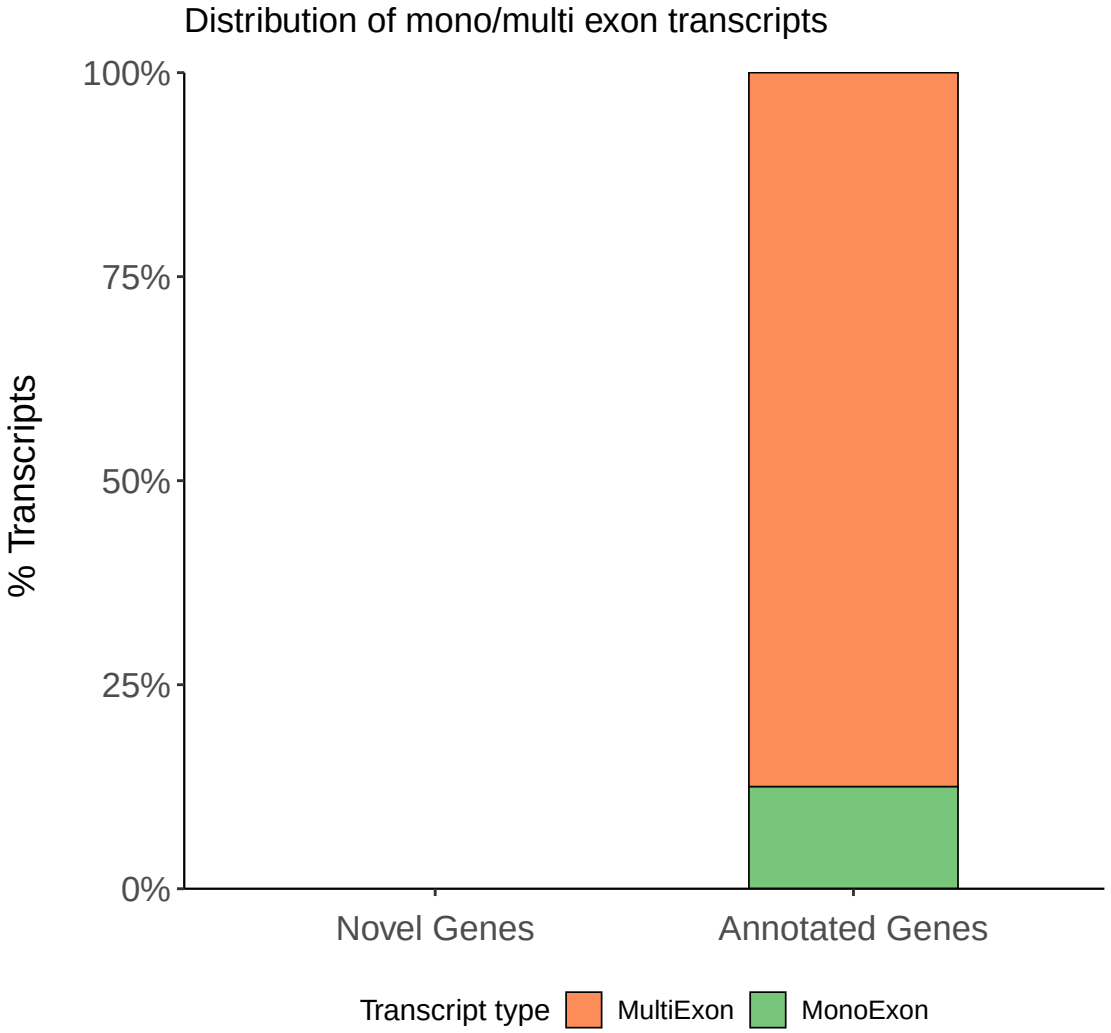
Characterization of transcripts based on splice junctions

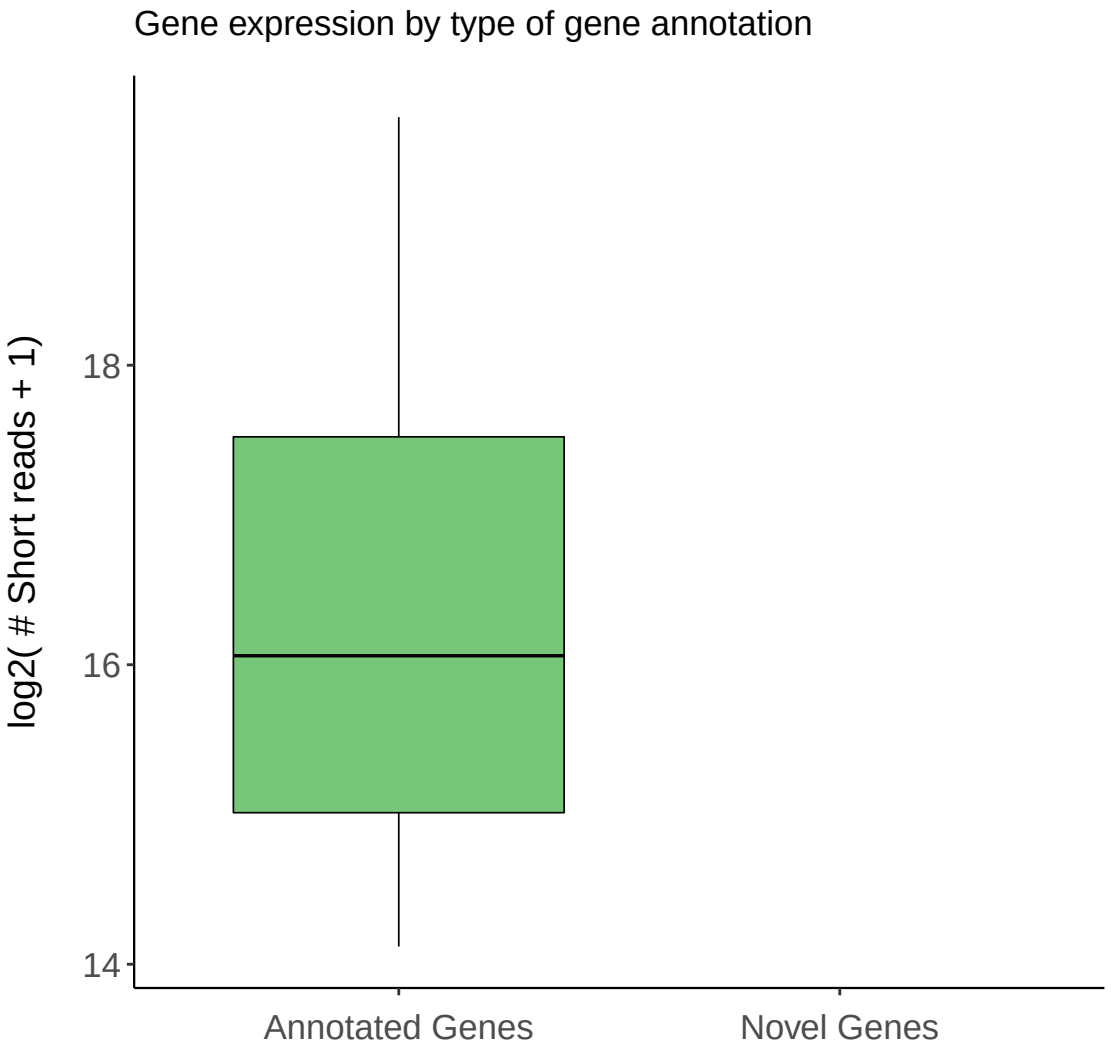
category	# isoforms
FSM	6
NNC	2
ISM	0
NIC	0
Genic Genomic	0
Antisense	0
Fusion	0
Intergenic	0
Genic Intron	0

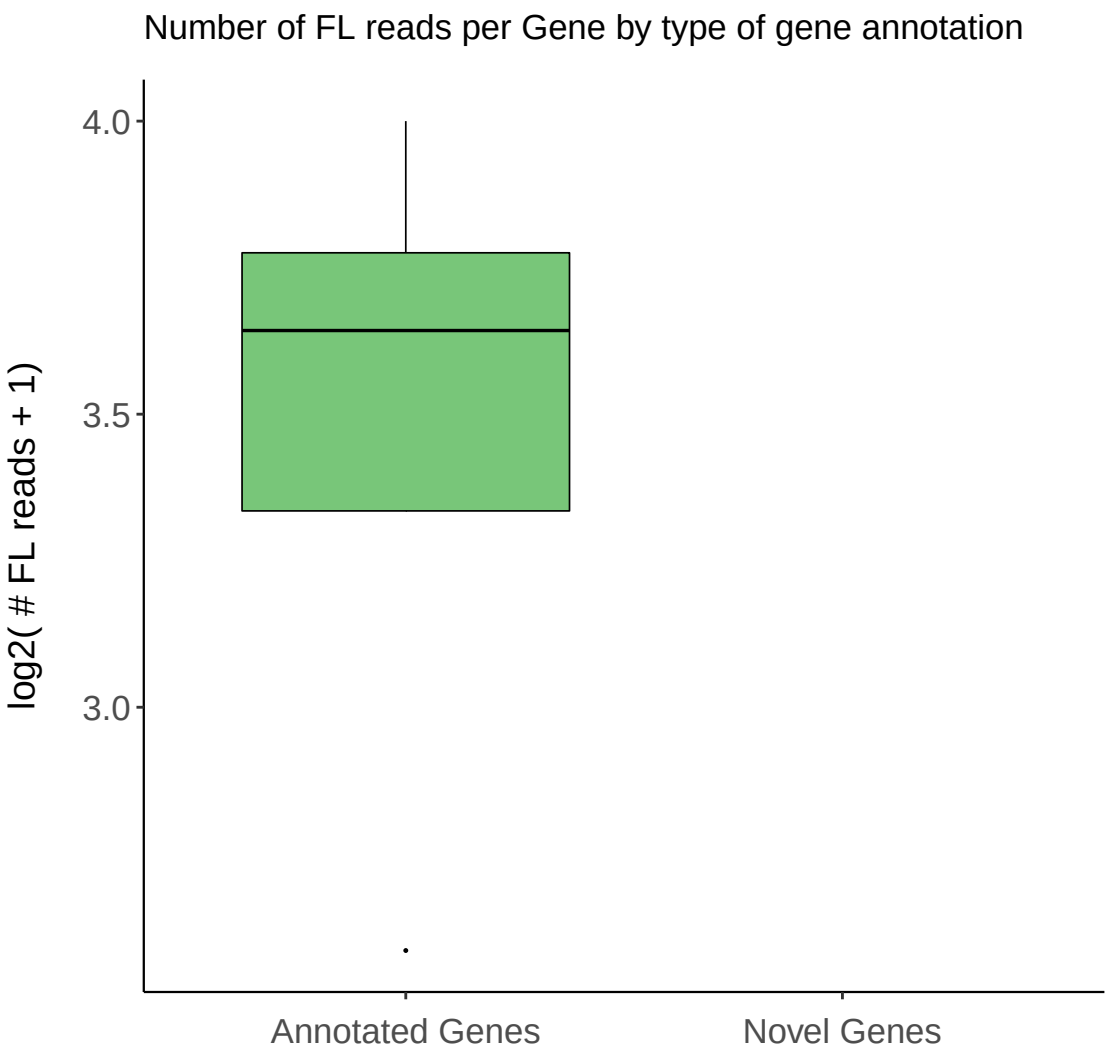
Gene characterization

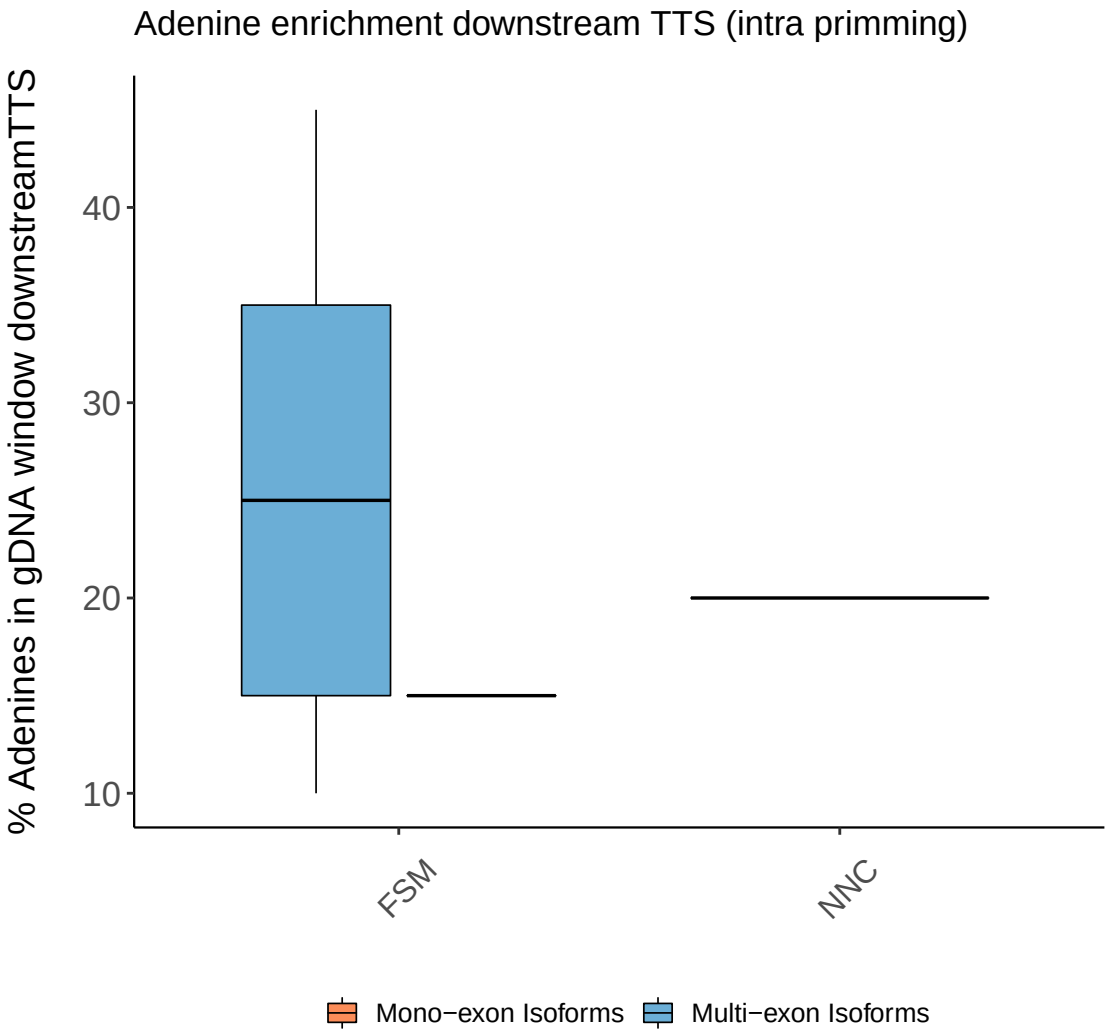

Distribution of isoforms per gene
% Genes
60%
40%
20%
0%
1
2−3
Isoforms per Gene

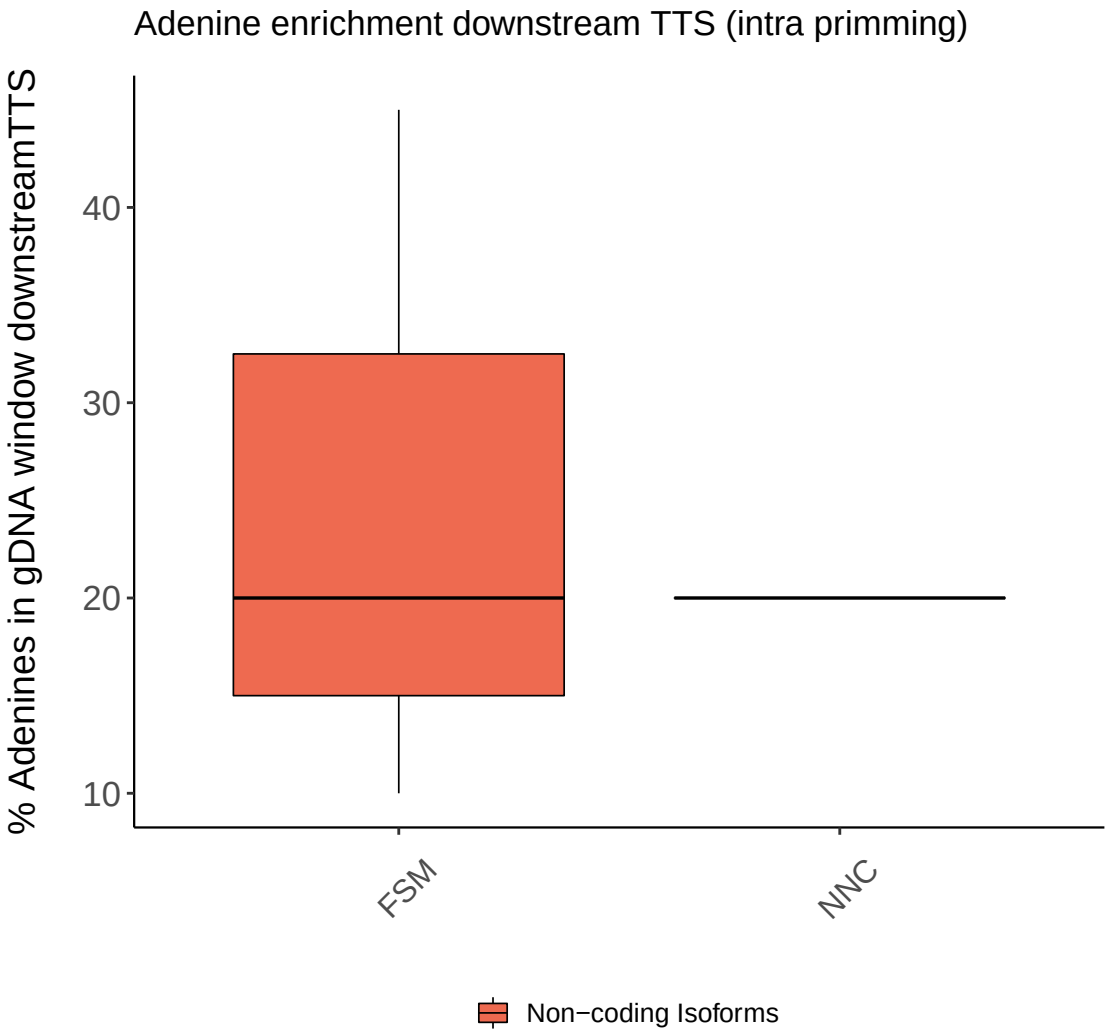

Distribution of number of isoforms
100%
75%
50%
25%
0%
% Genes
Annotated Genes
Novel Genes
Isoforms Per Gene
1
2−3


Distribution of mono/multi exon transcripts
100%
75%
50%
25%
0%
% Transcripts
Novel Genes
Annotated Genes
Transcript type
MultiExon
MonoExon

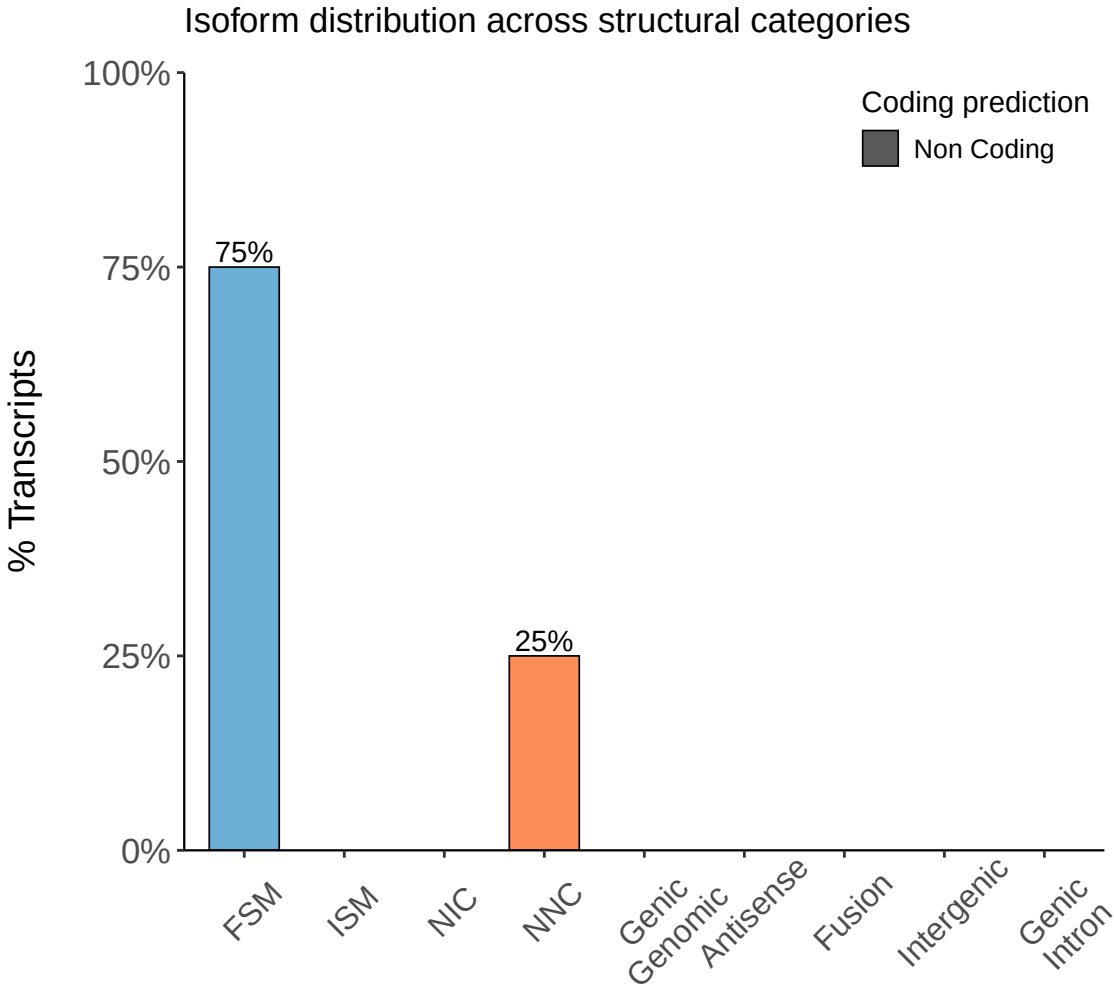

Gene expression by type of gene annotation
log2(# Short reads + 1)
18
16
14
Annotated Genes
Novel Genes

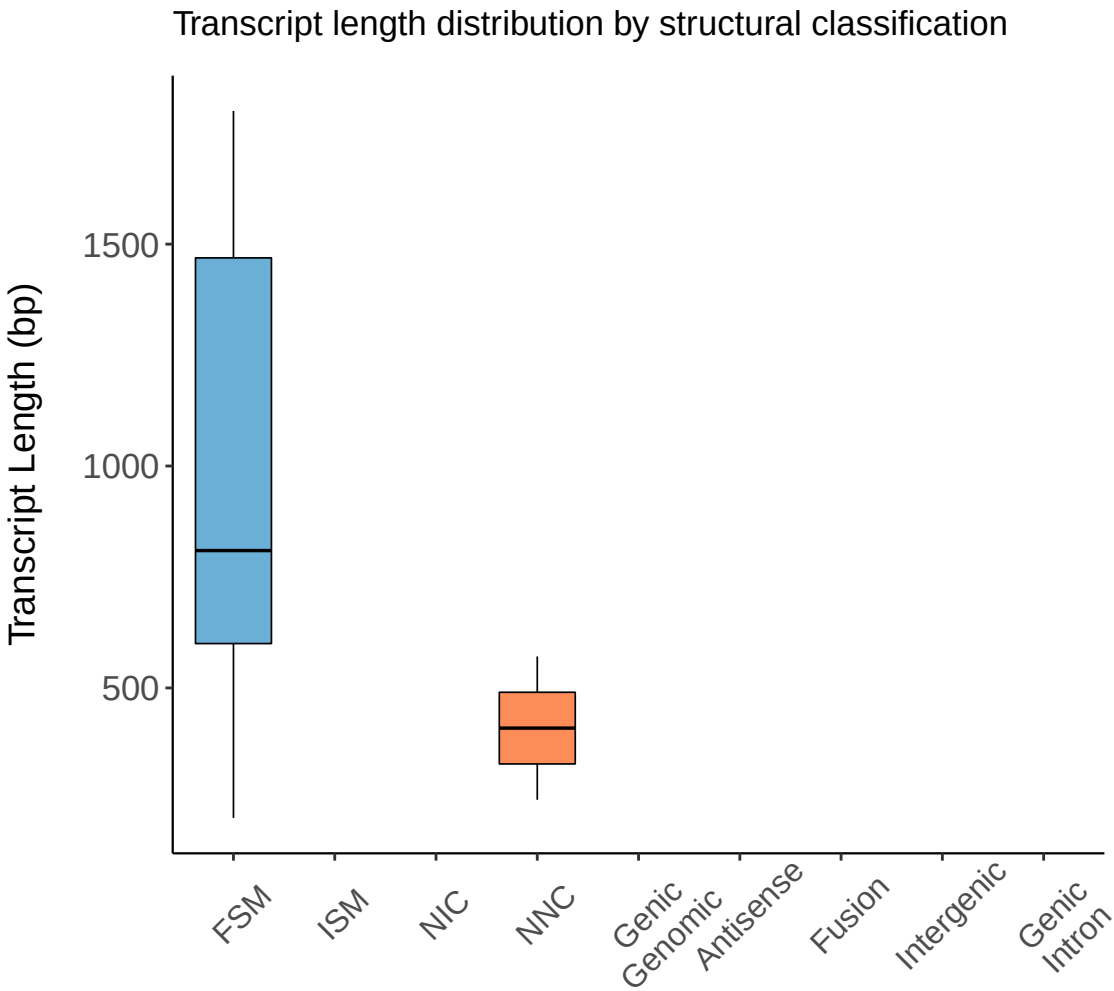

Number of FL reads per Gene by type of gene annotation
log2(# FL reads + 1)
4.0
3.5
3.0
Annotated Genes
Novel Genes

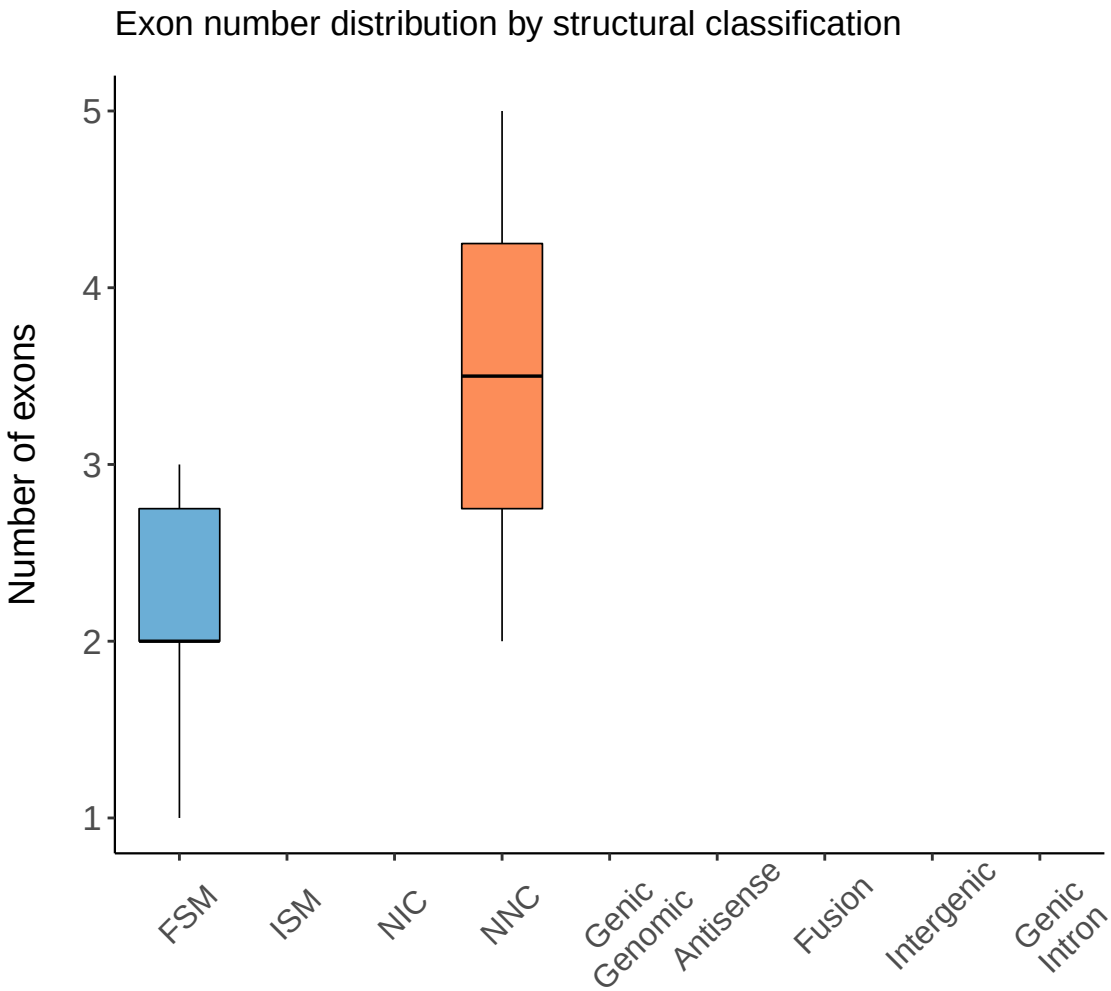

Adenine enrichment downstream TTS (intra primming)
% Adenines in gDNA window downstreamTTS
40
30
20
10
FSM
NNC
Mono−exon Isoforms
Multi−exon Isoforms

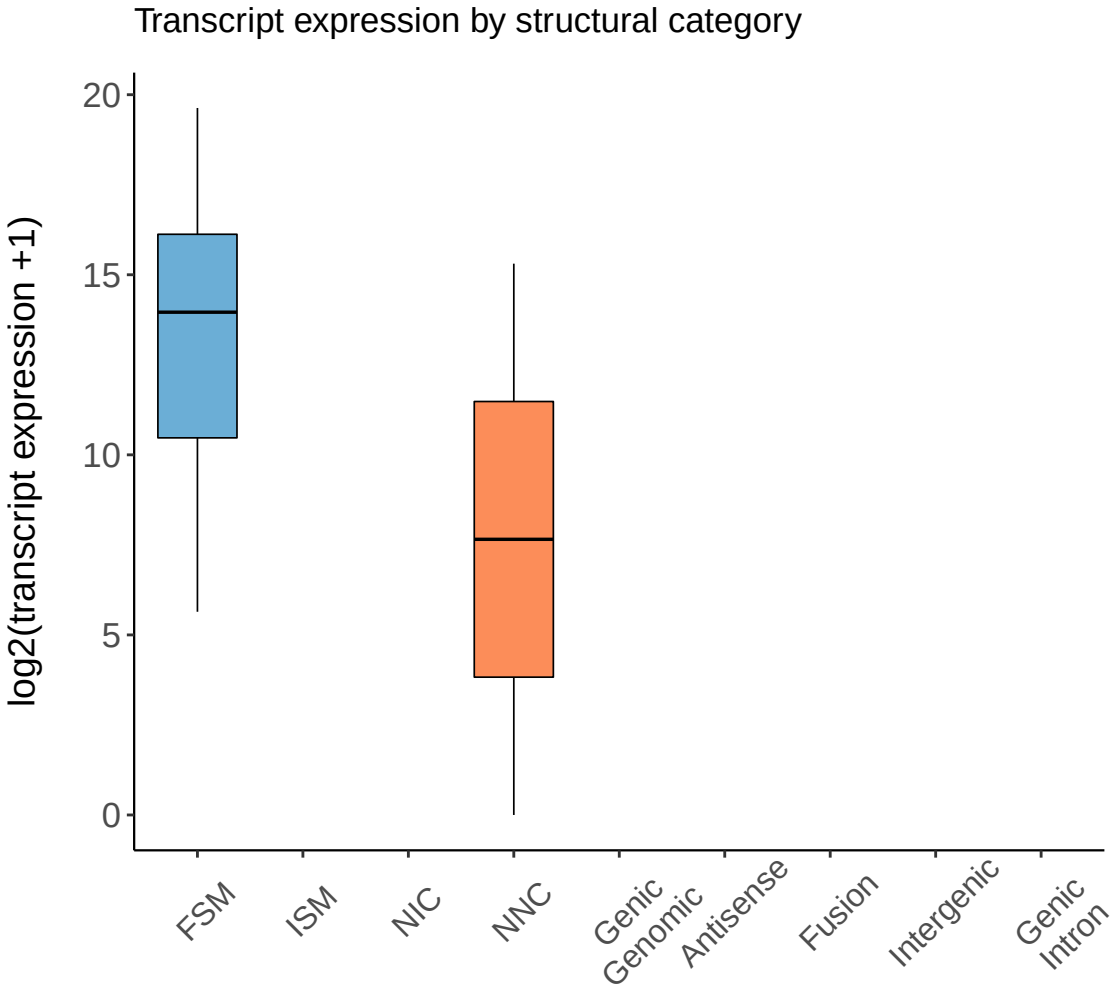

Adenine enrichment downstream TTS (intra primming)
% Adenines in gDNA window downstreamTTS
40
30
20
10
FSM
NNC
Non−coding Isoforms

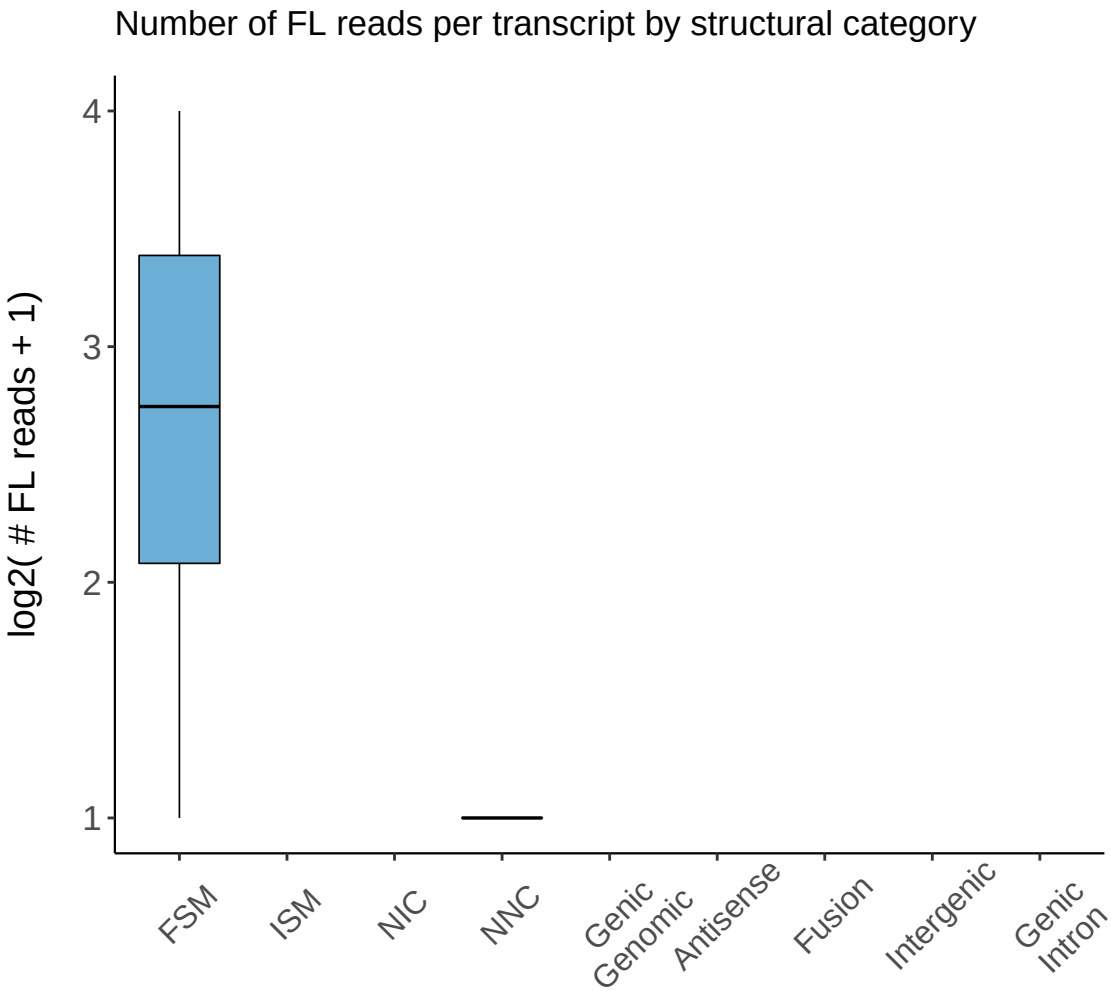
Structrual Isoform characterization based on splice junctions

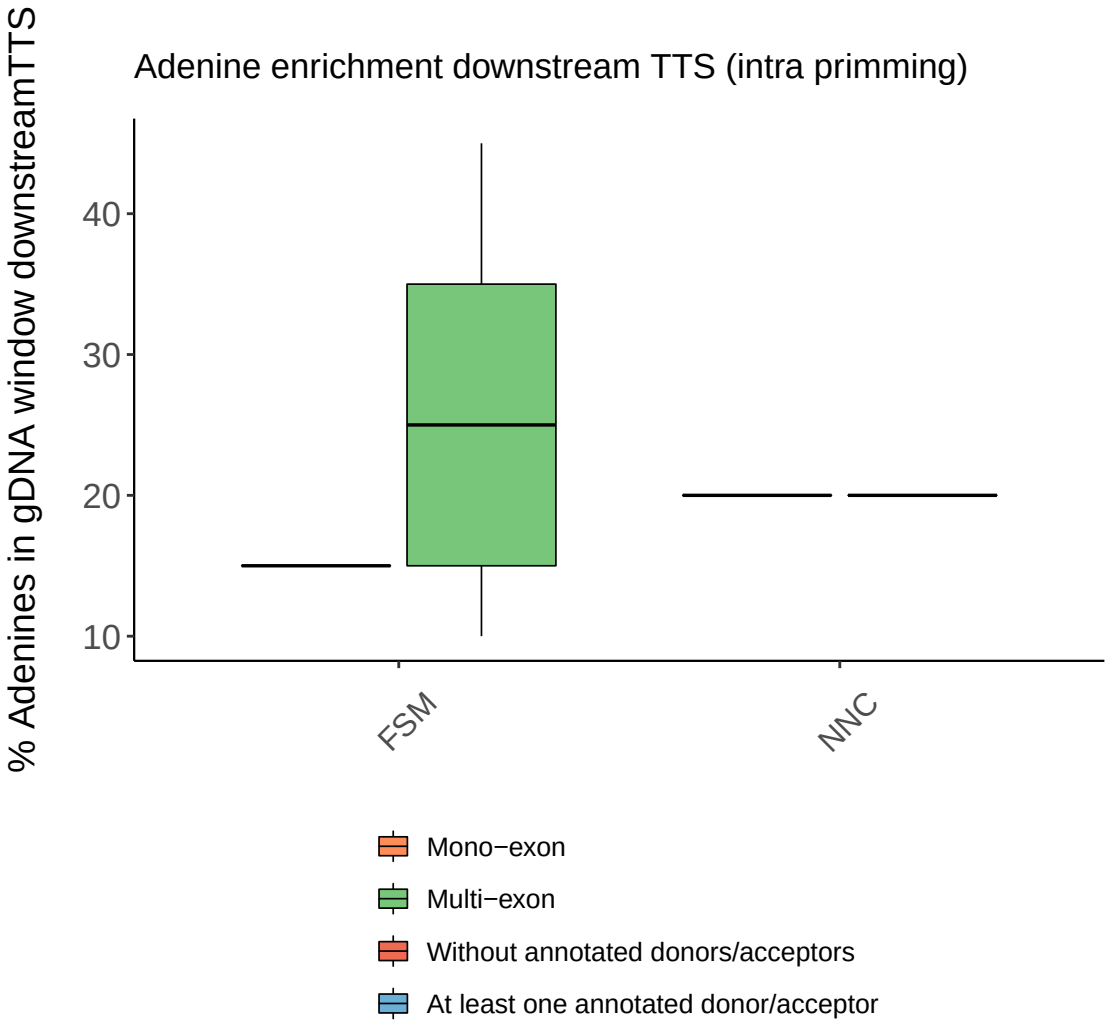

Isoform distribution across structural categories
% Transcripts
100%
75%
50%
25%
0%
Coding prediction
Non Coding
75%
25%
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

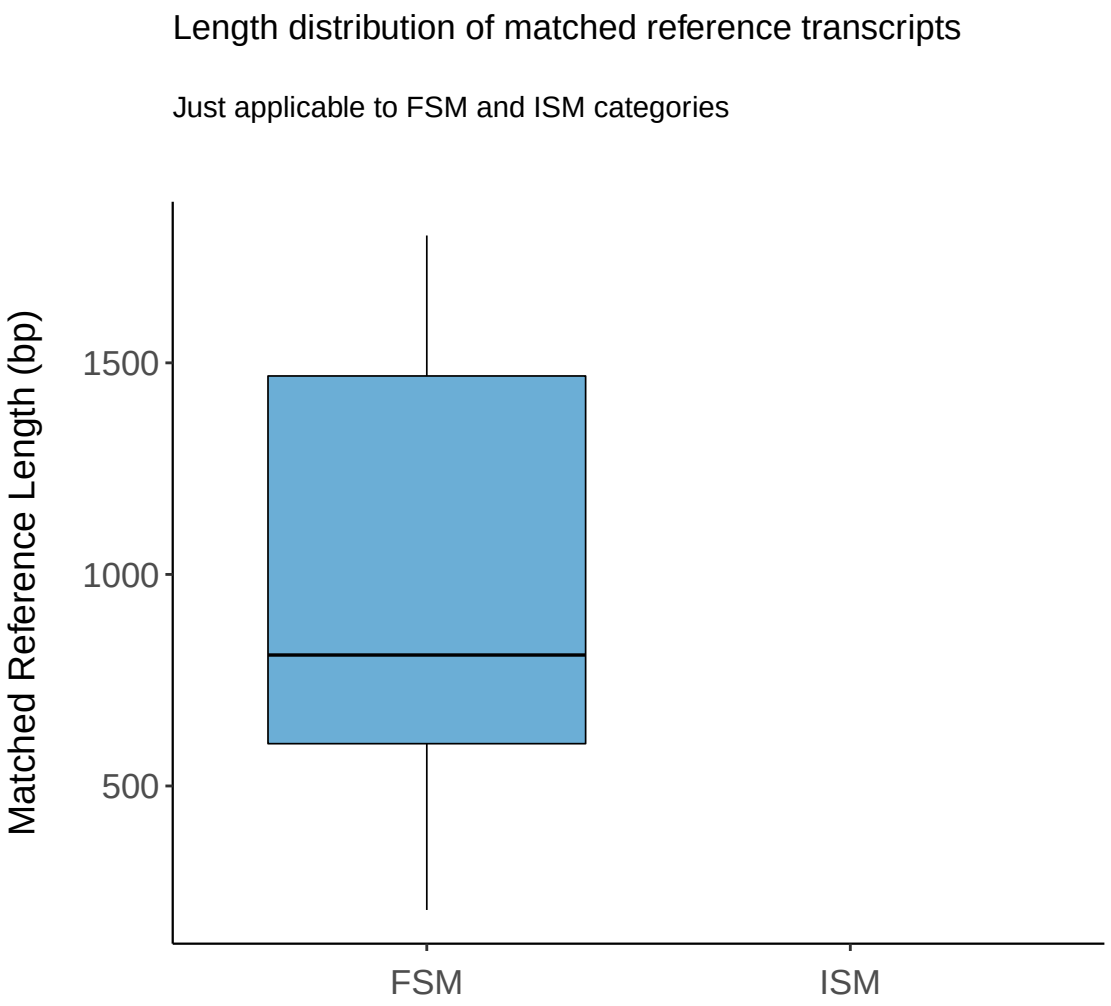

Transcript length distribution by structural classification
Transcript Length (bp)
1500
1000
500
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

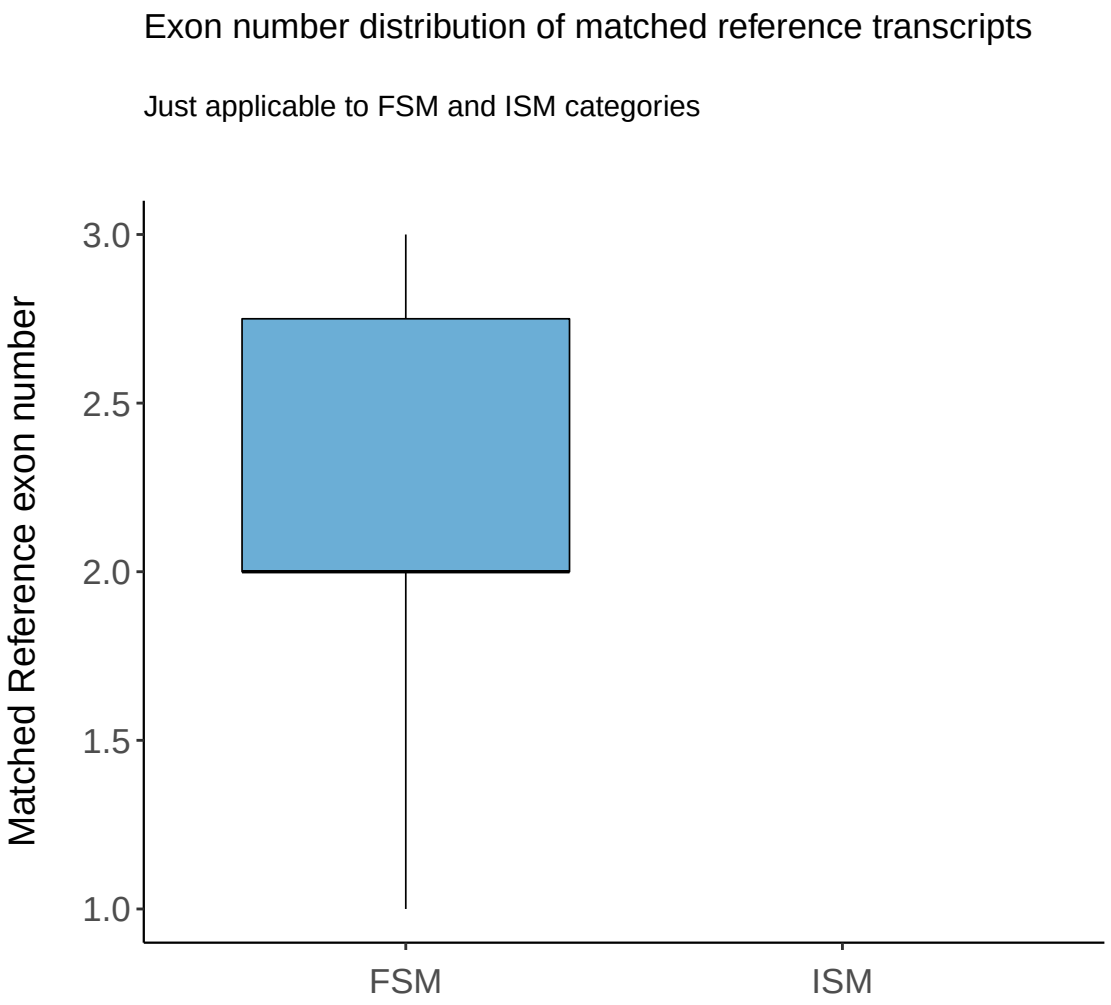

Exon number distribution by structural classification
Number of exons
5
4
3
2
1
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

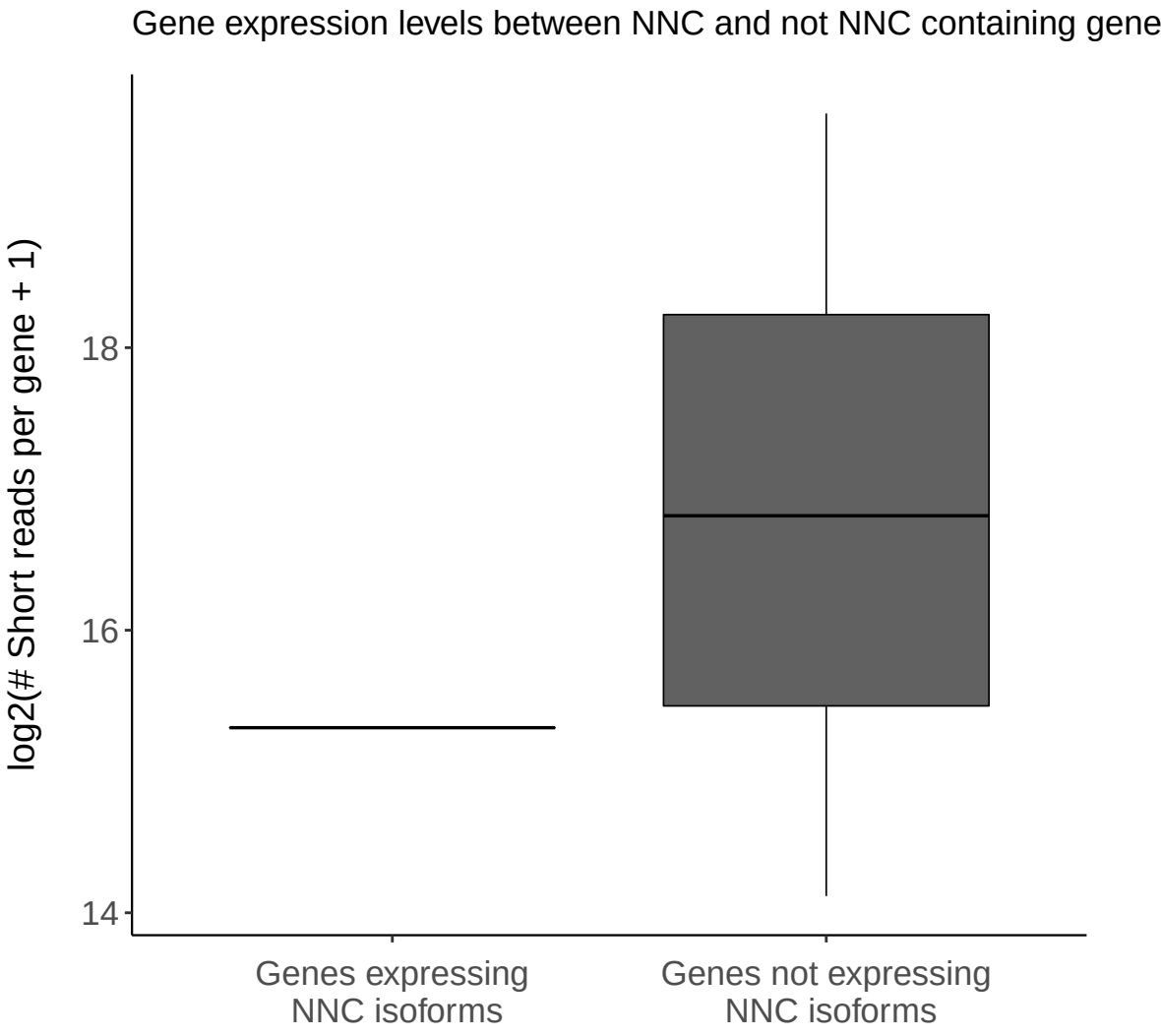

Transcript expression by structural category
log2(transcript expression +1)
20
15
10
5
0
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

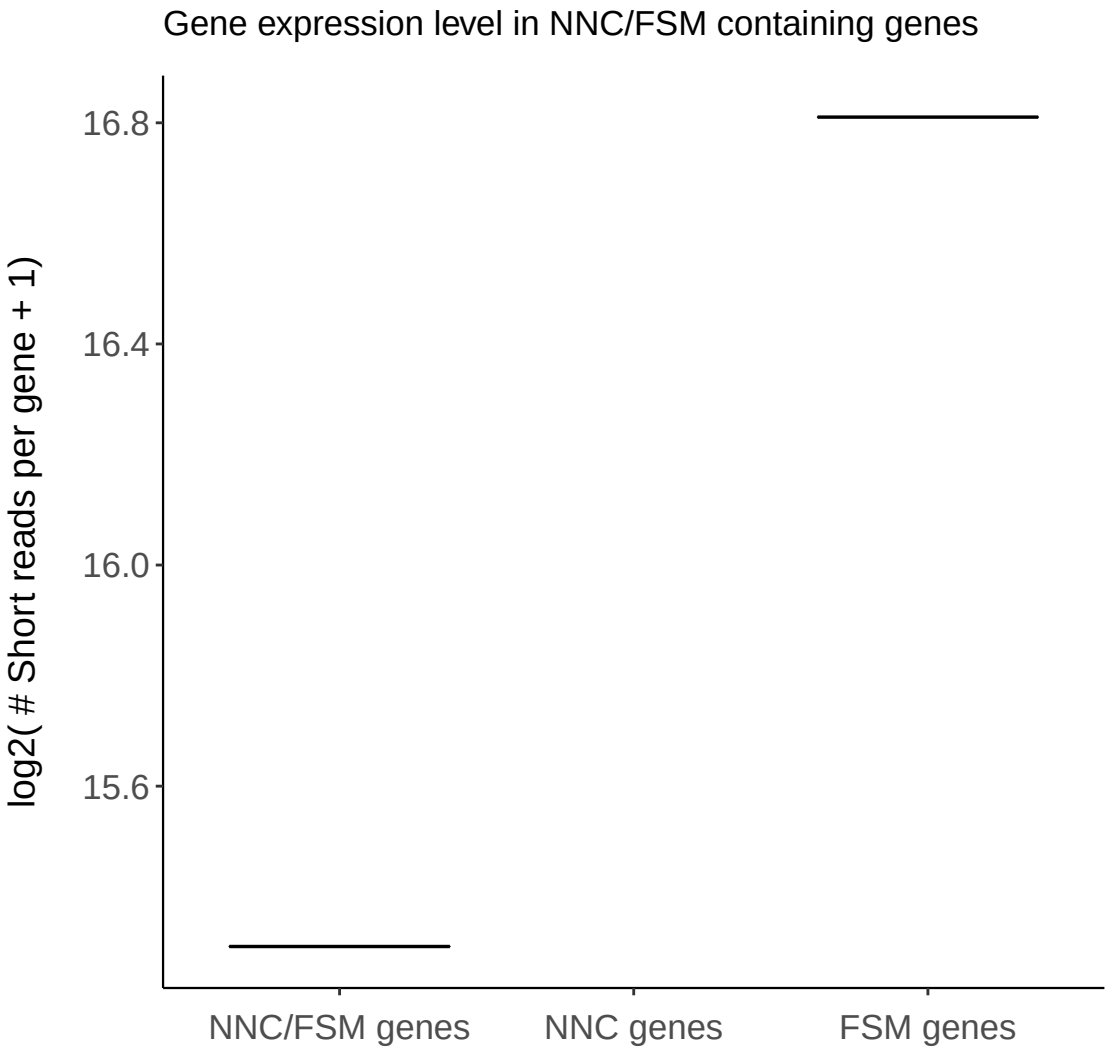

Number of FL reads per transcript by structural category
log2(# FL reads + 1)
4
3
2
1
FSM
ISM
NIC
NNC
Genic Genomic
Antisense
Fusion
Intergenic
Genic Intron

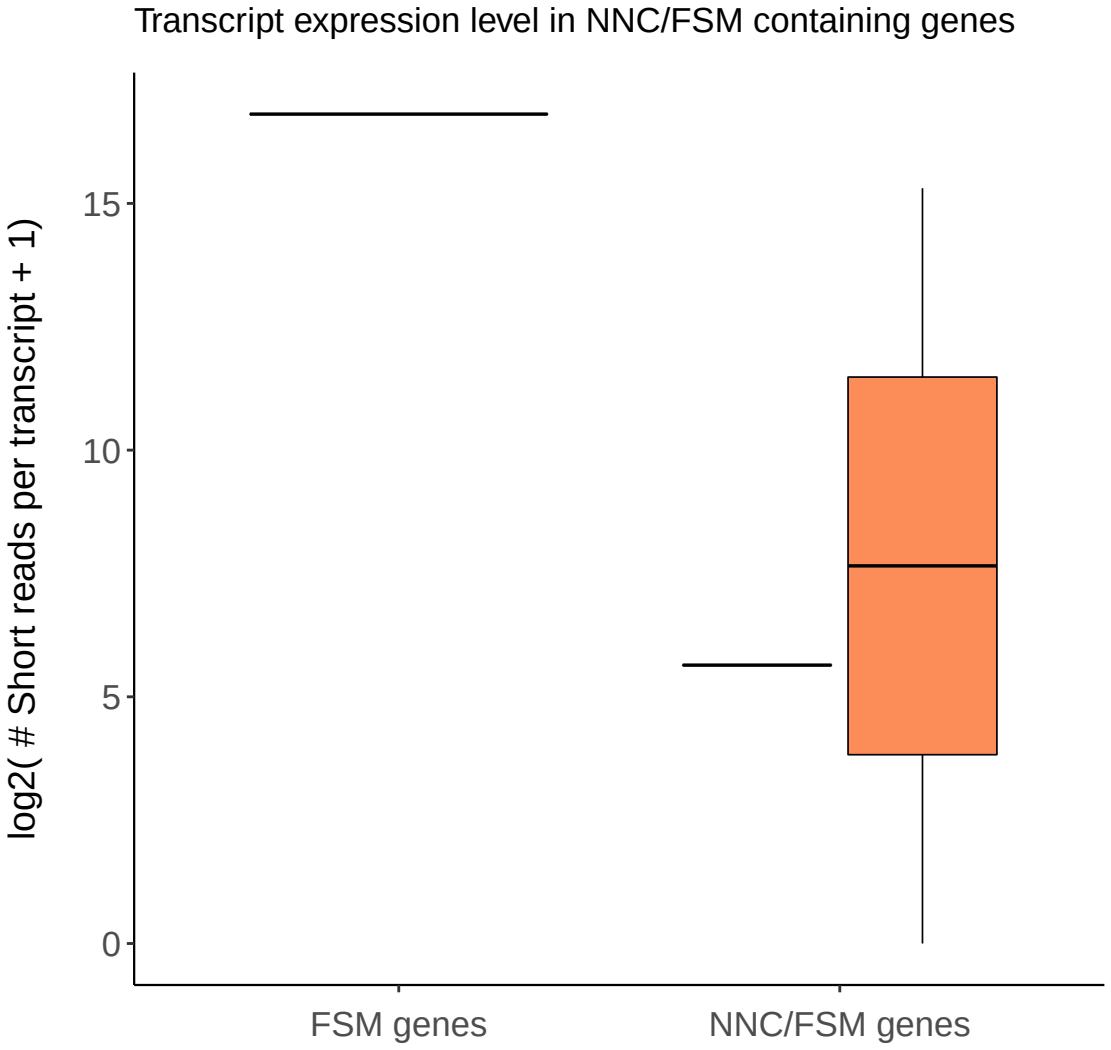

Adenine enrichment downstream TTS (intra primming)
% Adenines in gDNA window downstreamTTS
40
30
20
10
FSM
NNC
Mono−exon
Multi−exon
Without annotated donors/acceptors
At least one annotated donor/acceptor


Length distribution of matched reference transcripts
Just applicable to FSM and ISM categories
Matched Reference Length (bp)
1500
1000
500
FSM
ISM

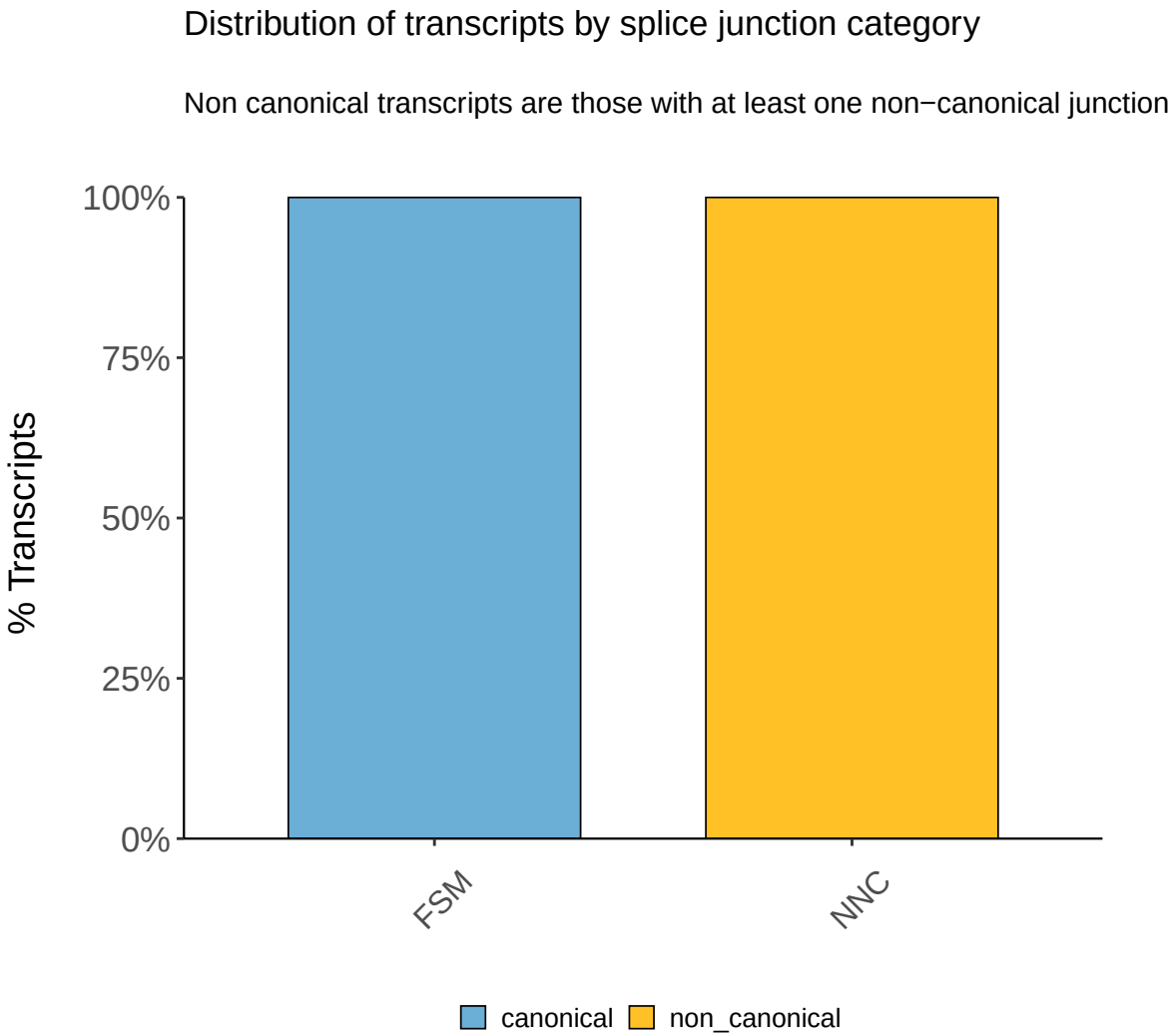

Exon number distribution of matched reference transcripts
Just applicable to FSM and ISM categories
Matched Reference exon number
3.0
2.5
2.0
1.5
1.0
FSM
ISM

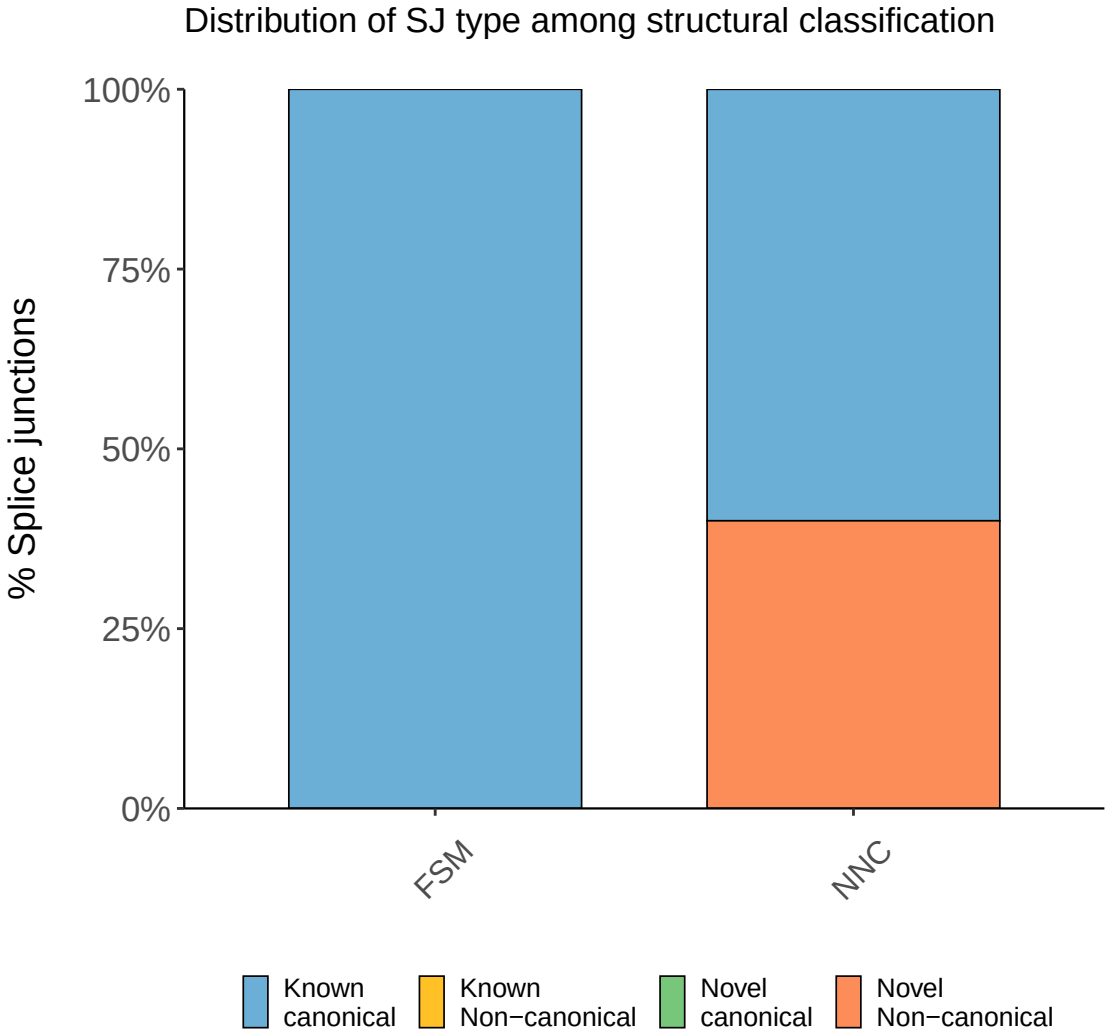

Gene expression levels between NNC and not NNC containing gene
log2(# Short reads per gene + 1)
18
16
14
Genes expressing NNC isoforms
Genes not expressing NNC isoforms

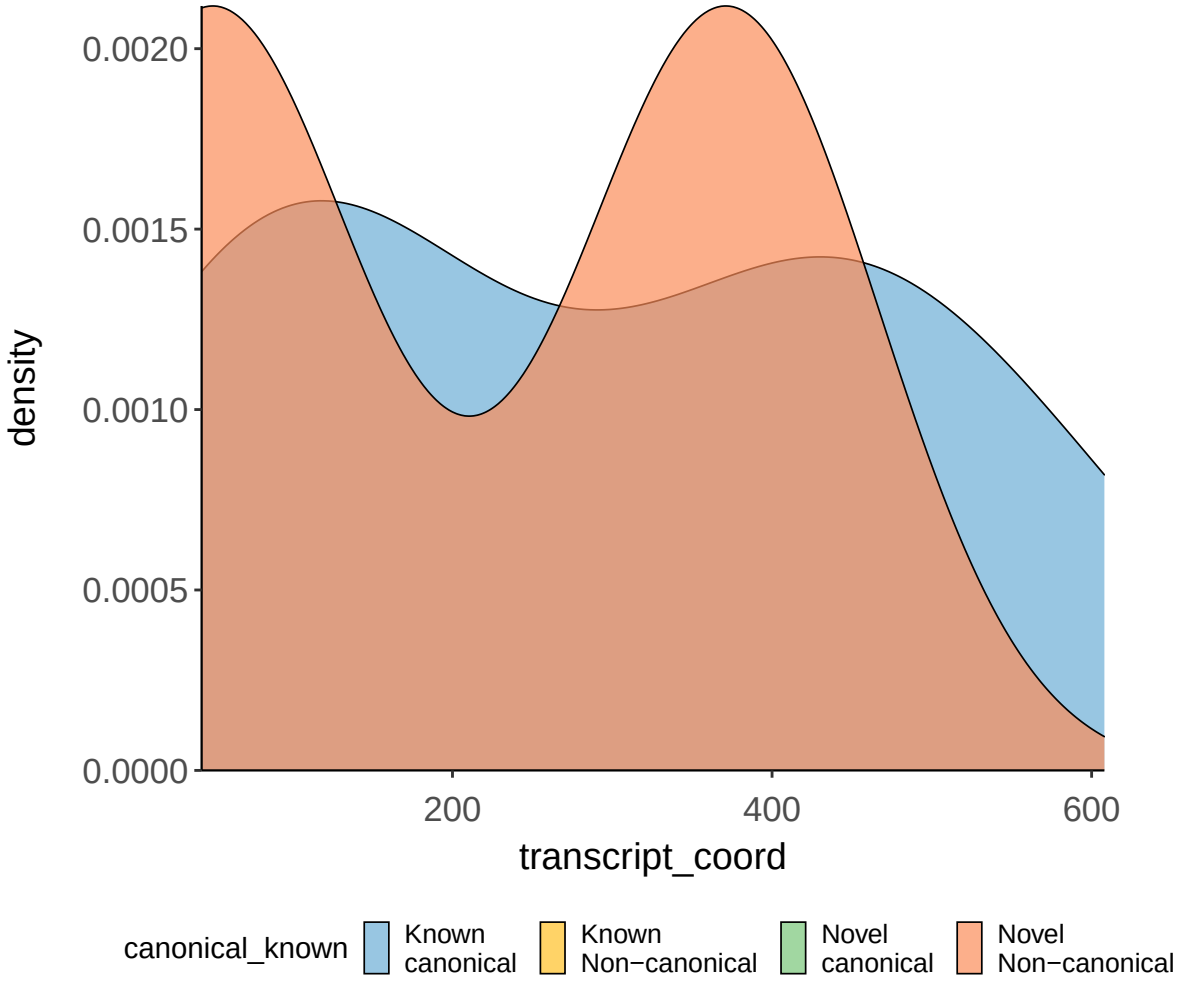

Gene expression level in NNC/FSM containing genes
log2(# Short reads per gene + 1)
16.8
16.4
16.0
15.6
NNC/FSM genes
NNC genes
FSM genes

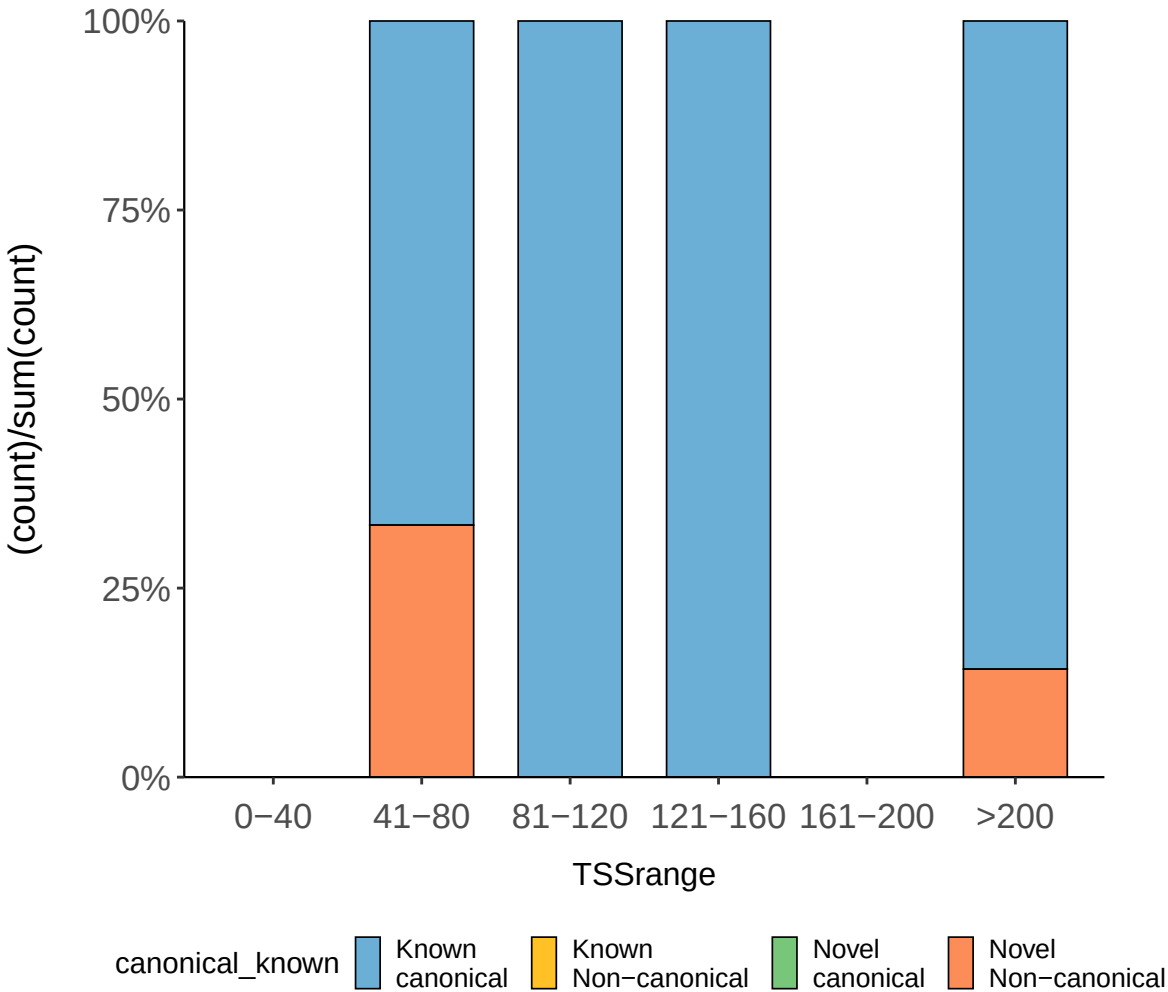

Transcript expression level in NNC/FSM containing genes
log2(# Short reads per transcript + 1)
15
10
5
0
FSM genes
NNC/FSM genes

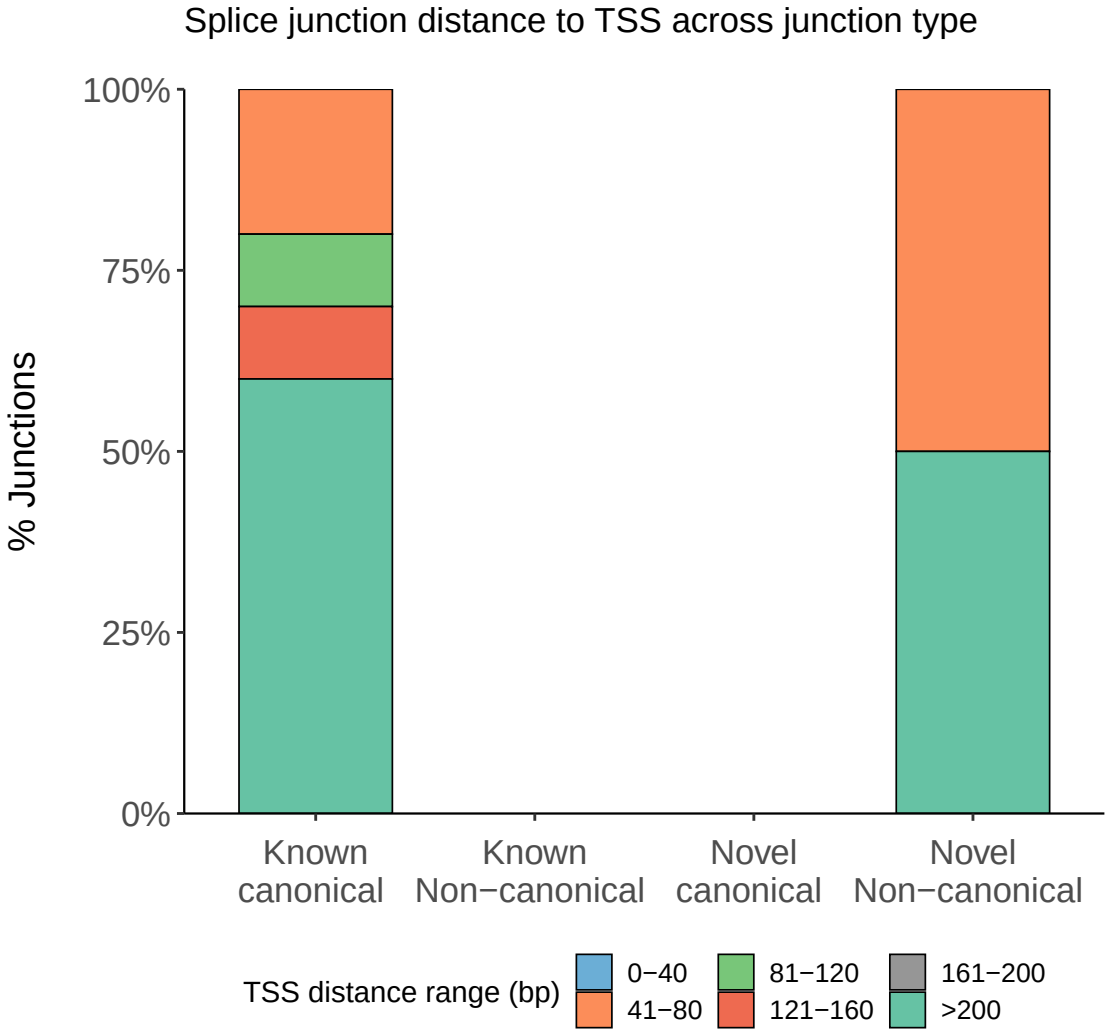
Splice junction characterization


Distribution of transcripts by splice junction category
Non canonical transcripts are those with at least one non−canonical junction
% Transcripts
100%
75%
50%
25%
0%
FSM
NNC
canonical
non_canonical

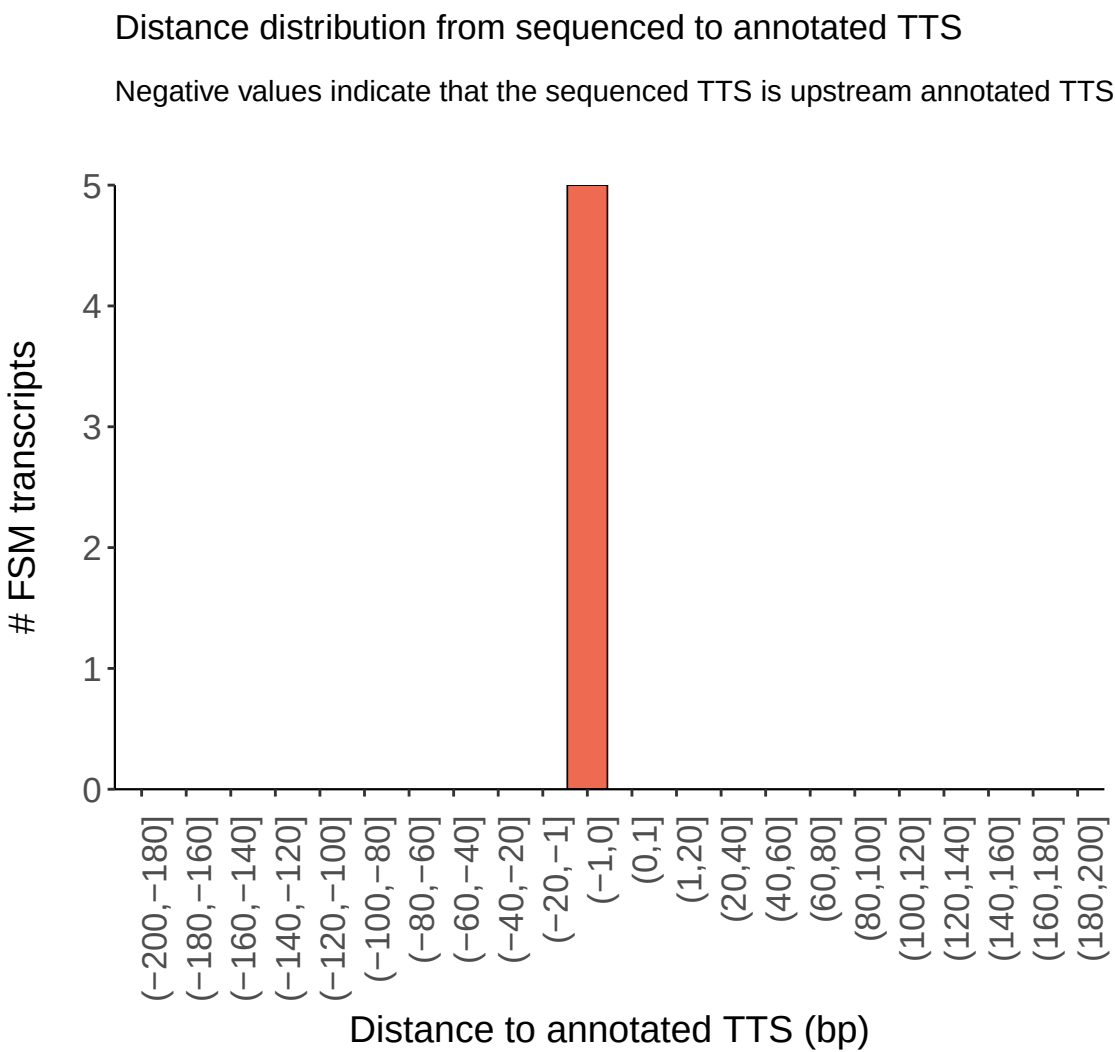

Distribution of SJ type among structural classification
% Splice junctions
100%
75%
50%
25%
0%
FSM
NNC
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

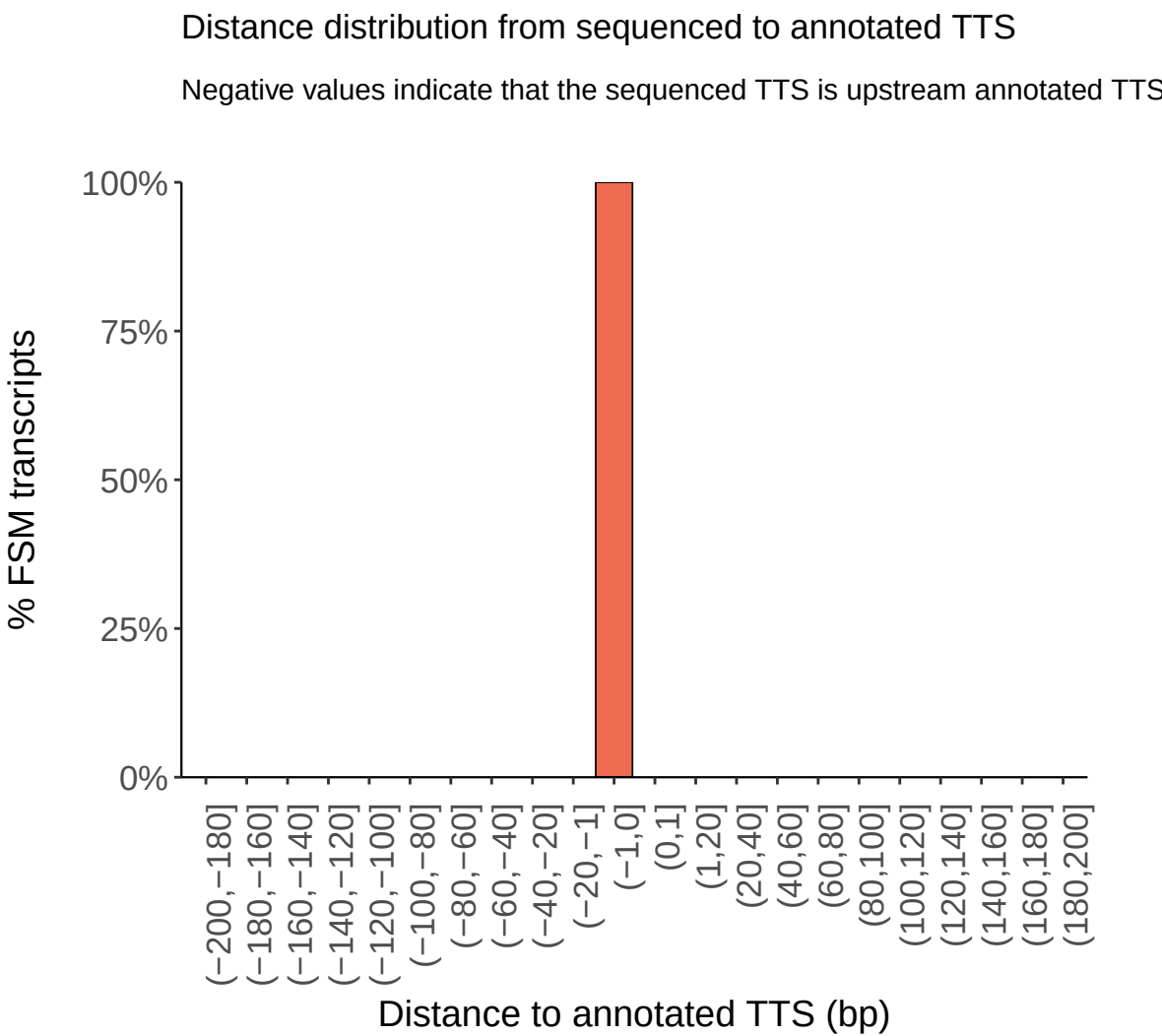

density
0.0020
0.0015
0.0010
0.0005
0.0000
200
400
600
transcript_coord
canonical_known
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

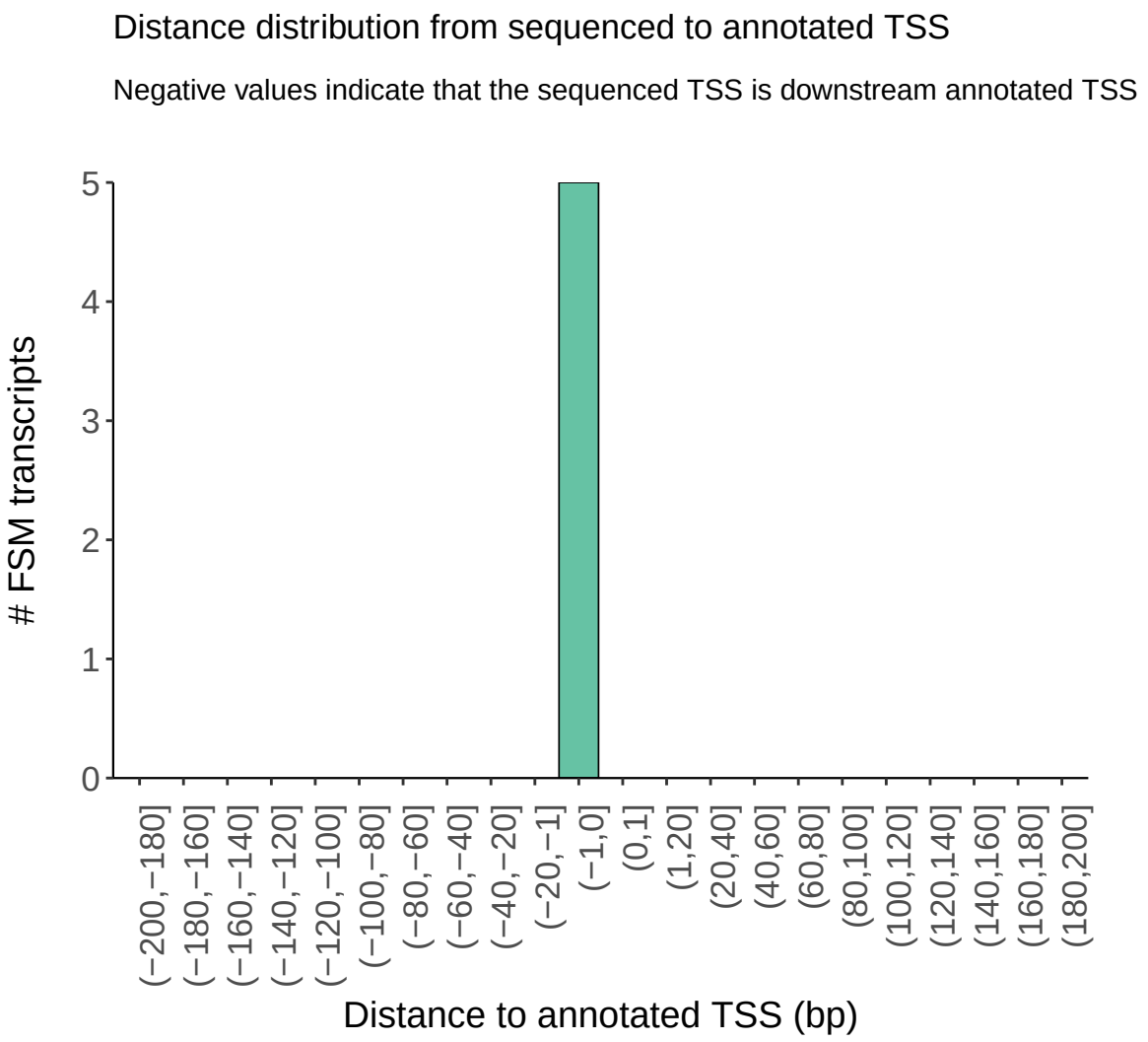

100%
75%
50%
25%
0%
(count)/sum(count)
0−40
41−80
81−120
121−160
161−200
>200
TSSrange
canonical_known
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical

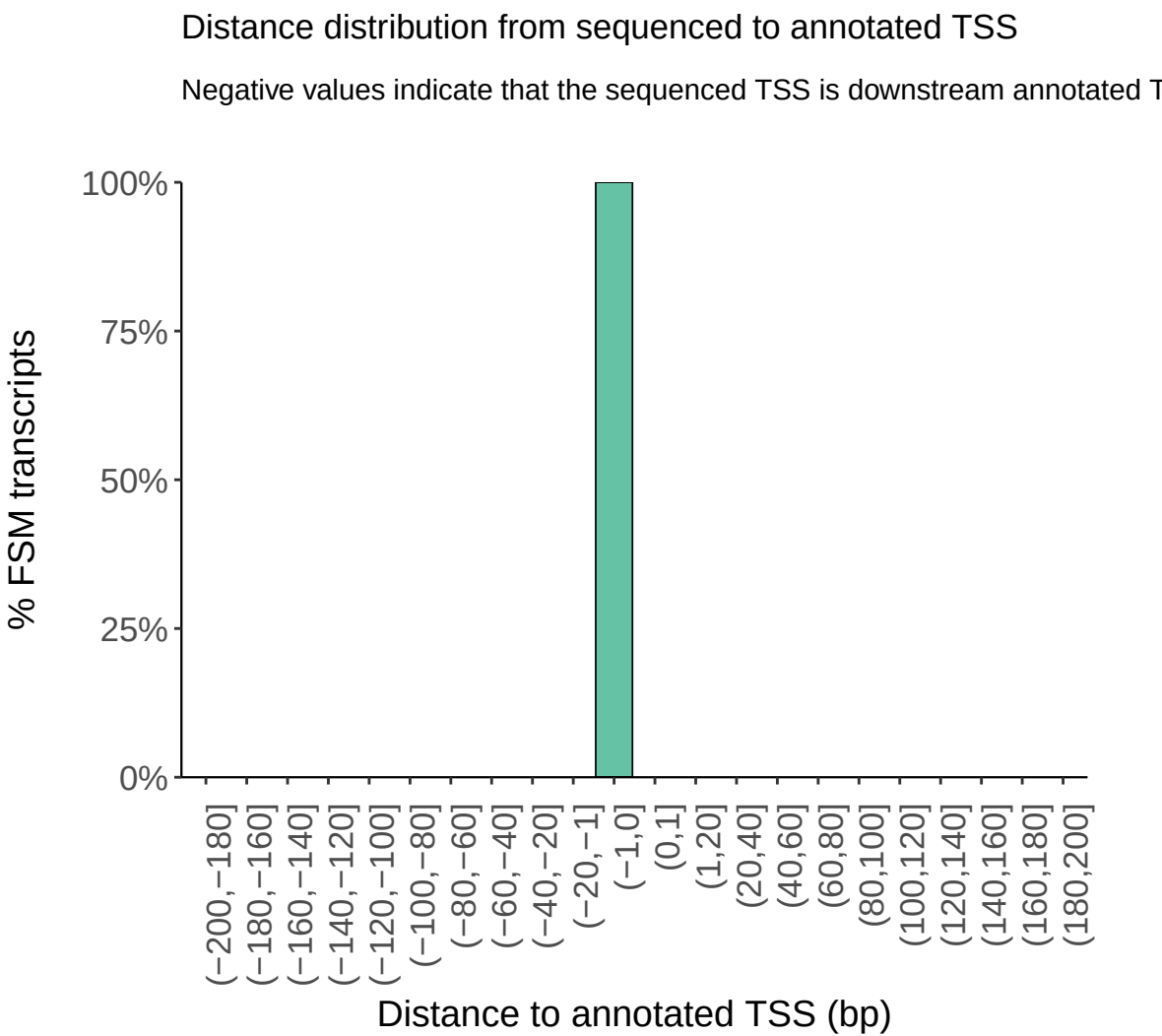

Splice junction distance to TSS across junction type
% Junctions
100%
75%
50%
25%
0%
Known canonical
Known Non−canonical
Novel canonical
Novel Non−canonical
TSS distance range (bp)
0−40
41−80
81−120
121−160
161−200
>200

Full−lengthness characterization of isoforms

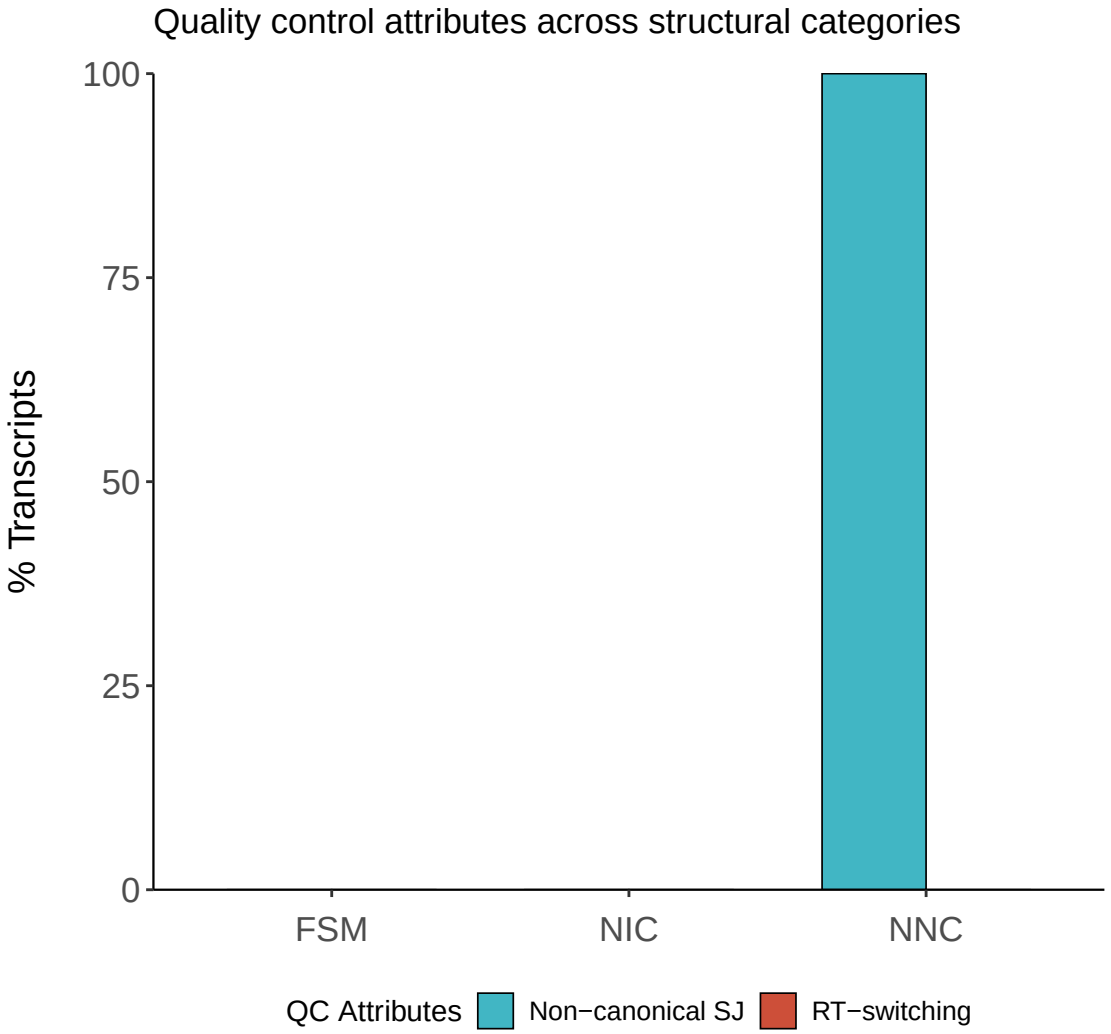

Distance distribution from sequenced to annotated TTS
Negative values indicate that the sequenced TTS is upstream annotated TTS
FSM transcripts
0
1
2
3
4
5
(−200,−180]
(−180,−160]
(−160,−140]
(−140,−120]
(−120,−100]
(−100,−80]
(−80,−60]
(−60,−40]
(−40,−20]
(−20,−1]
(−1,0]
(0,1]
(1,20]
(20,40]
(40,60]
(60,80]
(80,100]
(100,120]
(120,140]
(140,160]
(160,180]
(180,200]
Distance to annotated TTS (bp)


Distance distribution from sequenced to annotated TTS
Negative values indicate that the sequenced TTS is upstream annotated TTS
% FSM transcripts
100%
75%
50%
25%
0%
(−200,−180]
(−180,−160]
(−160,−140]
(−140,−120]
(−120,−100]
(−100,−80]
(−80,−60]
(−60,−40]
(−40,−20]
(−20,−1]
(−1,0]
(0,1]
(1,20]
(20,40]
(40,60]
(60,80]
(80,100]
(100,120]
(120,140]
(140,160]
(160,180]
(180,200]
Distance to annotated TTS (bp)


Distance distribution from sequenced to annotated TSS
Negative values indicate that the sequenced TSS is downstream annotated TSS
FSM transcripts
0
1
2
3
4
5
(−200,−180]
(−180,−160]
(−160,−140]
(−140,−120]
(−120,−100]
(−100,−80]
(−80,−60]
(−60,−40]
(−40,−20]
(−20,−1]
(−1,0]
(0,1]
(1,20]
(20,40]
(40,60]
(60,80]
(80,100]
(100,120]
(120,140]
(140,160]
(160,180]
(180,200]
Distance to annotated TSS (bp)


Distance distribution from sequenced to annotated TSS
Negative values indicate that the sequenced TSS is downstream annotated T
% FSM transcripts
100%
75%
50%
25%
0%
(−200,−180]
(−180,−160]
(−160,−140]
(−140,−120]
(−120,−100]
(−100,−80]
(−80,−60]
(−60,−40]
(−40,−20]
(−20,−1]
(−1,0]
(0,1]
(1,20]
(20,40]
(40,60]
(60,80]
(80,100]
(100,120]
(120,140]
(140,160]
(160,180]
(180,200]
Distance to annotated TSS (bp)

Quality control attributes


Quality control attributes across structural categories
% Transcripts
100
75
50
25
0
FSM
NIC
NNC
QC Attributes
Non−canonical SJ
RT−switching