

SQANTI2 report

Unique Genes: 17568

Unique Isoforms: 289817

Gene classification

Category	# Genes
Annotated Genes	16807
Novel Genes	761

Splice Junction Classification

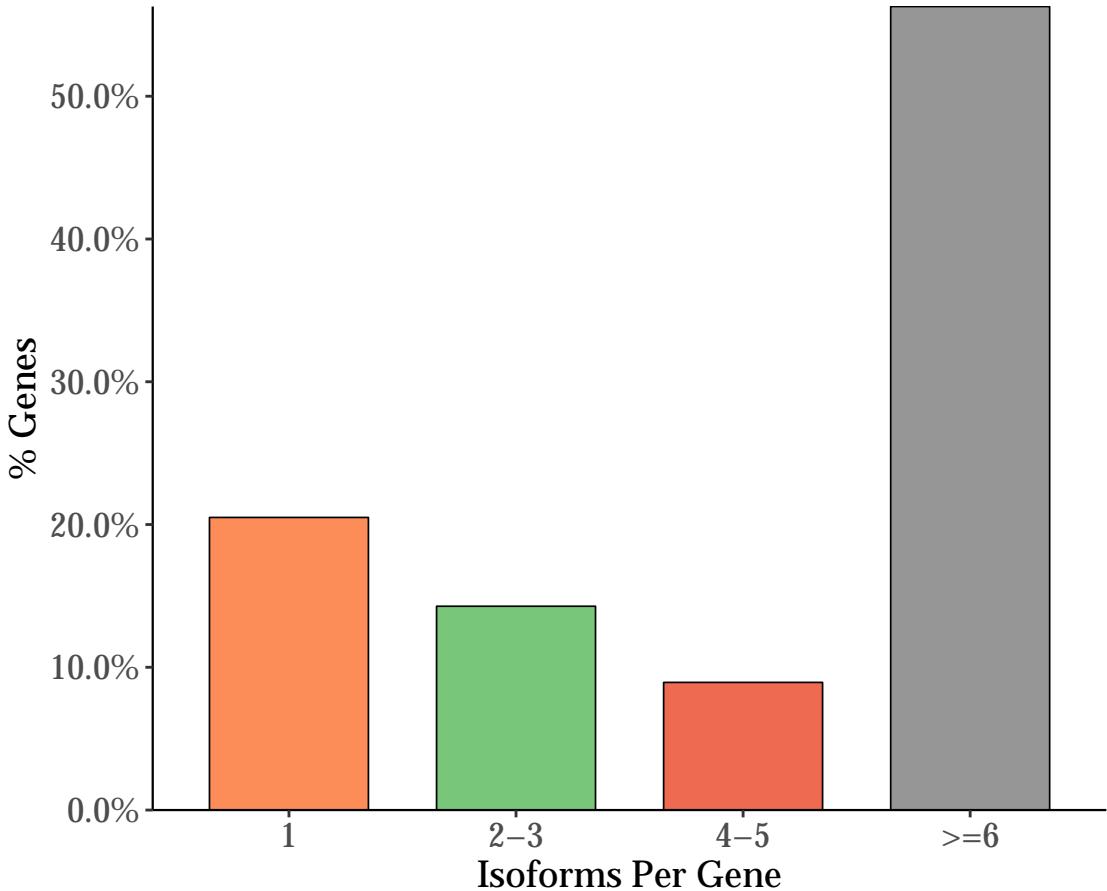
Category	# SJs	Percent
Known canonical	172028	79.85
Known Non-canonical	114	0.05
Novel canonical	43293	20.10
Novel Non-canonical	5	0.00

Characterization of transcripts based on splice junctions

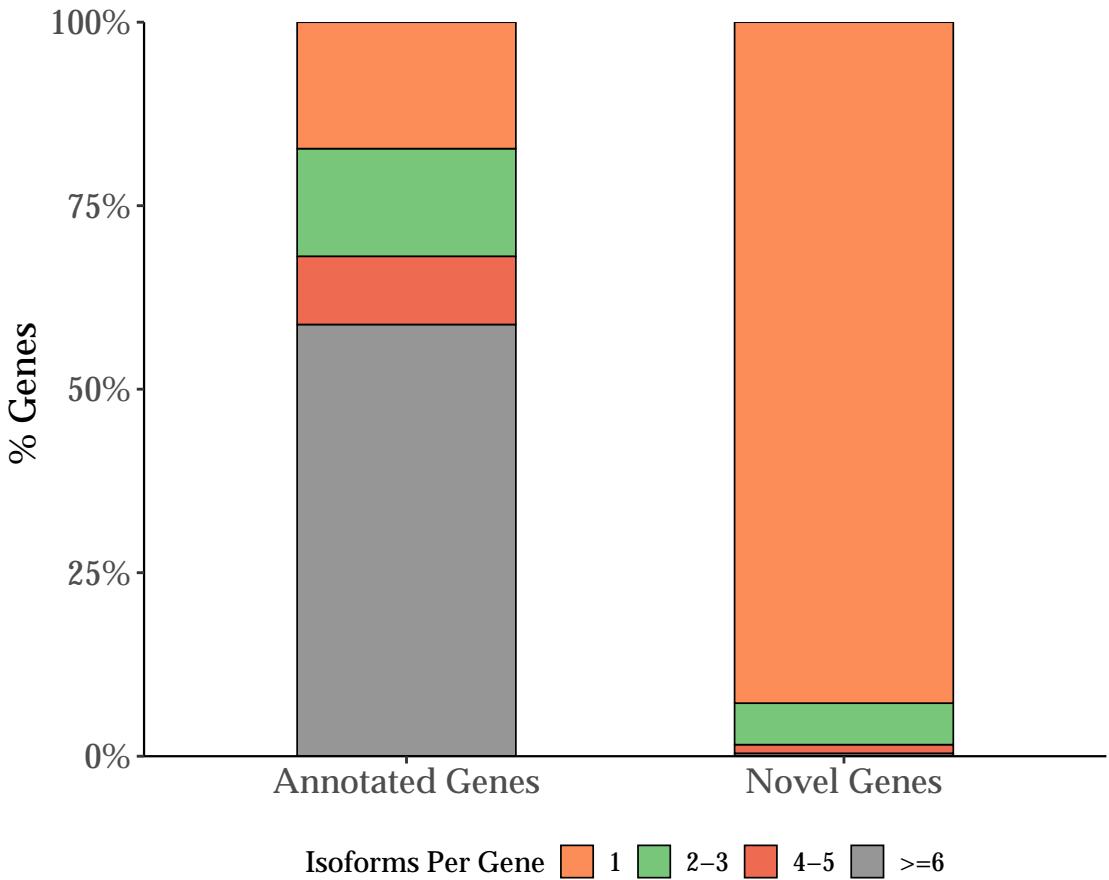
Category	# Isoforms
FSM	90643
ISM	64879
NIC	92008
NNC	41422
Genic Genomic	0
Antisense	317
Fusion	0
Intergenic	548
Genic Intron	0

Gene Characterization

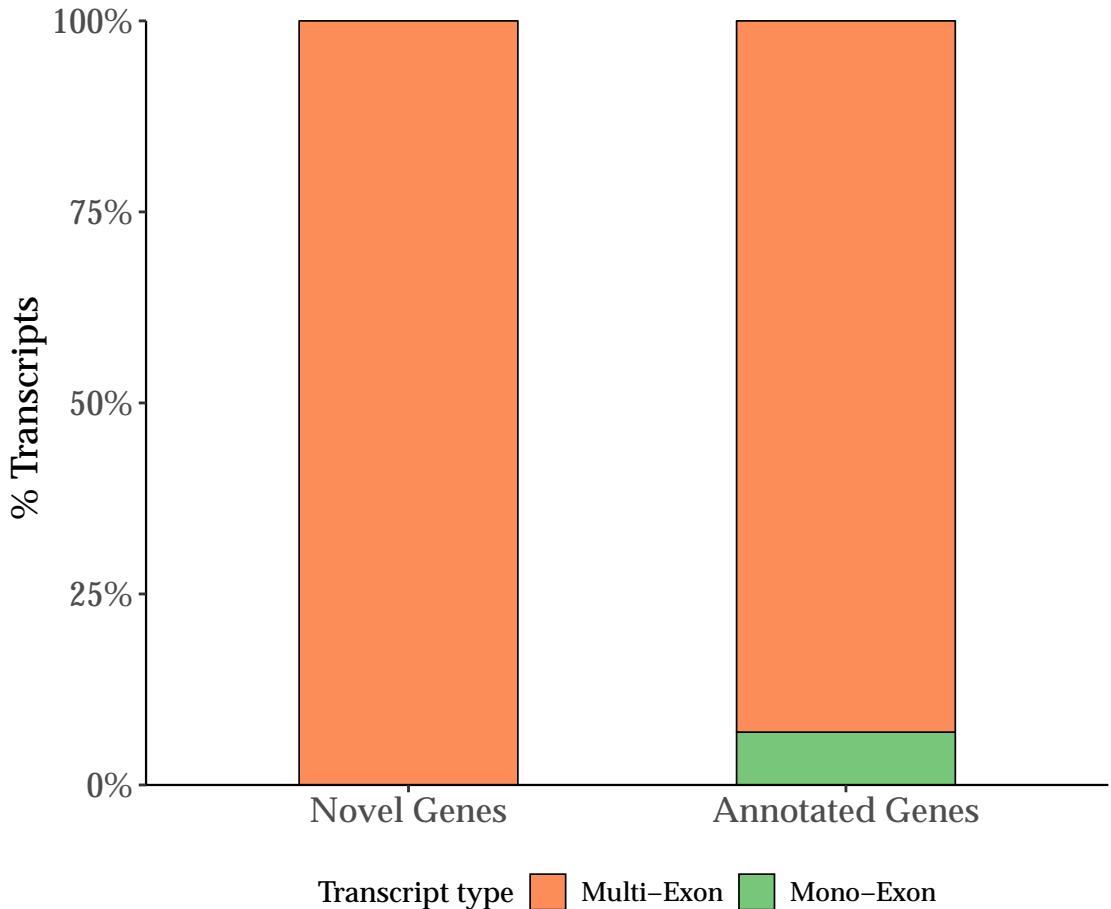
Number of Isoforms per Gene



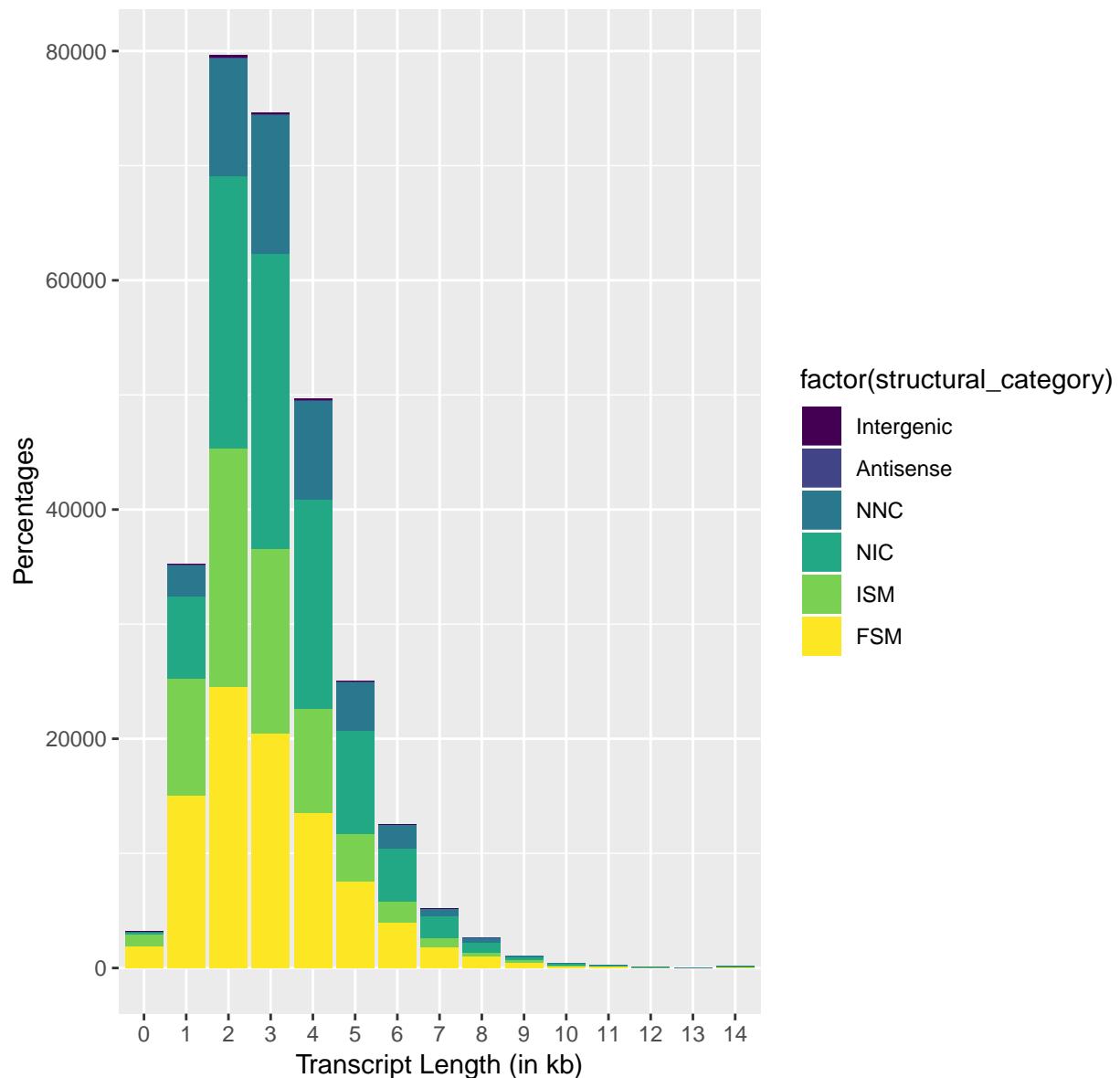
Number of Isoforms per Gene, Novel vs Known Genes



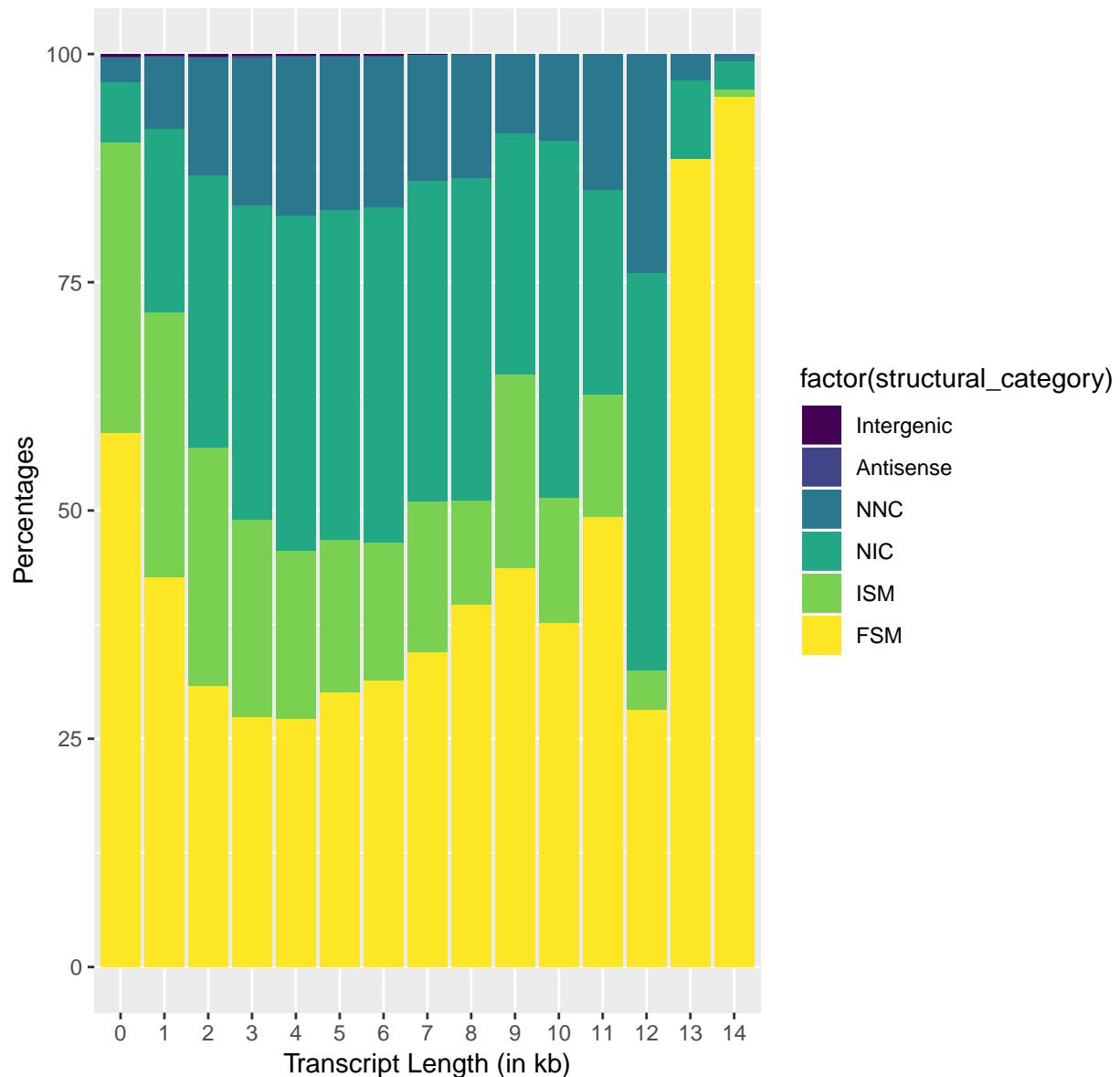
Distribution of Mono– vs Multi–Exon Transcripts



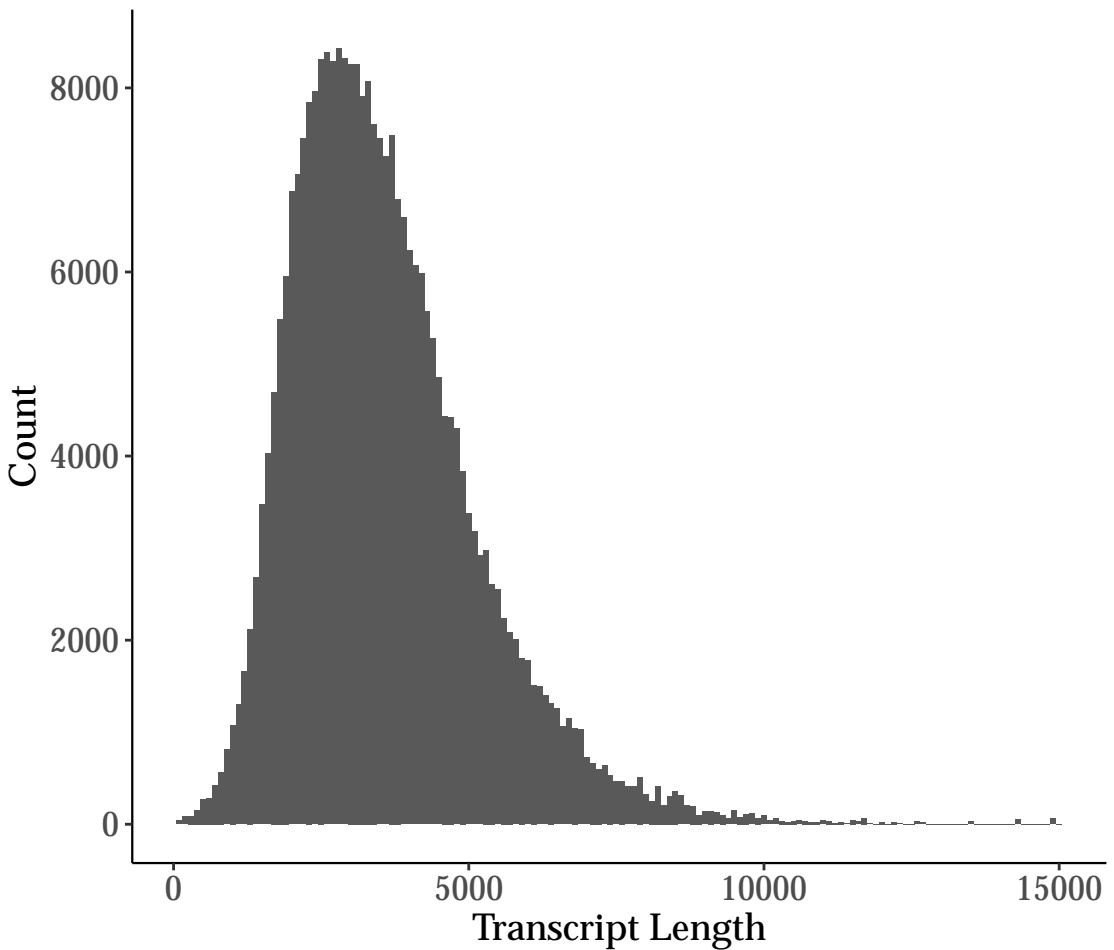
Classifications by Transcript Length



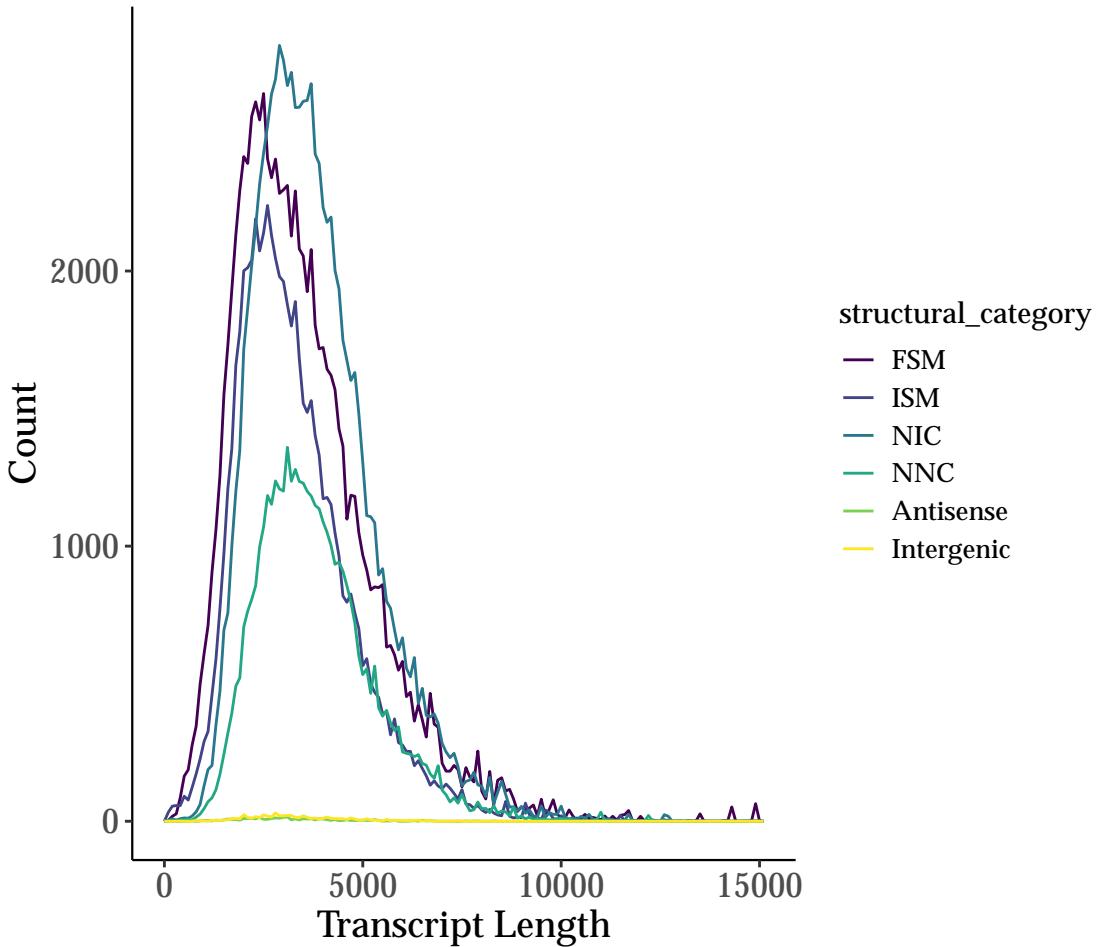
Classifications by Transcript Length, normalized



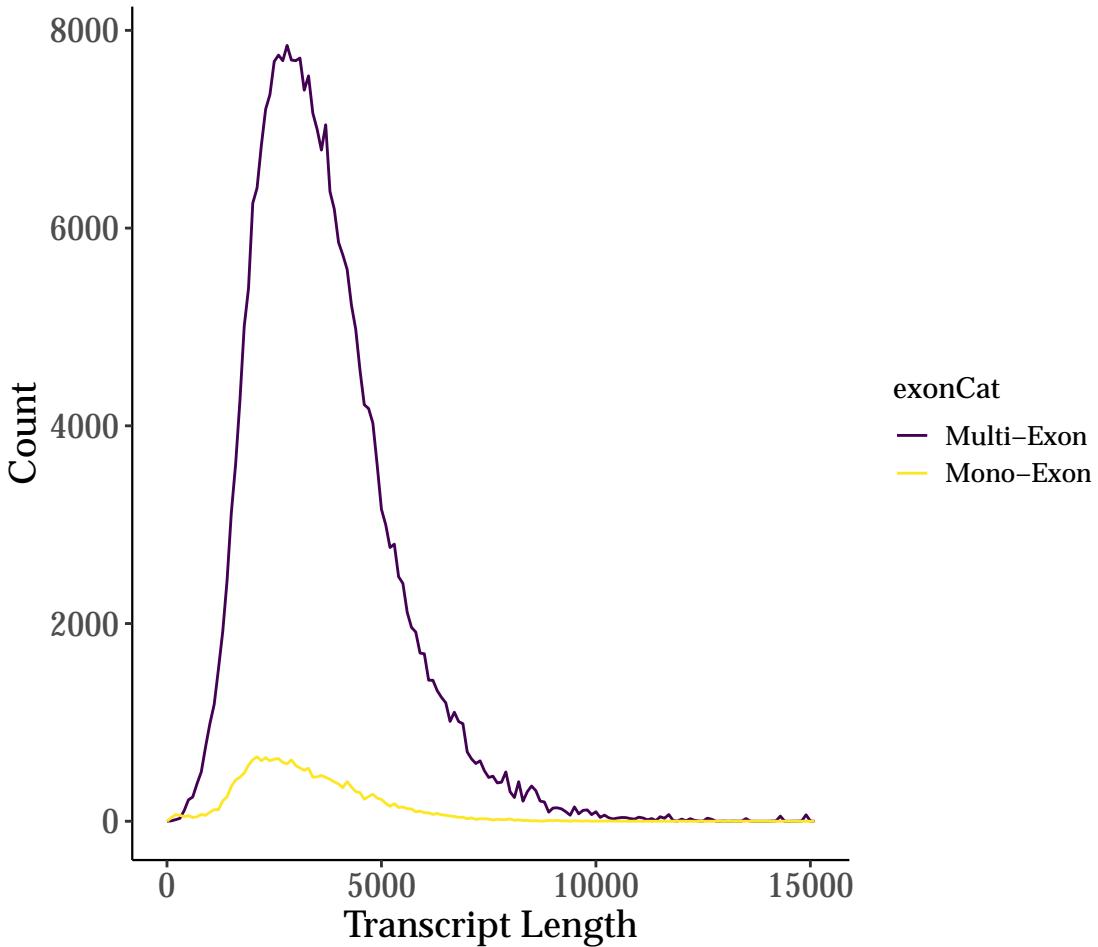
Transcript Lengths, all transcripts



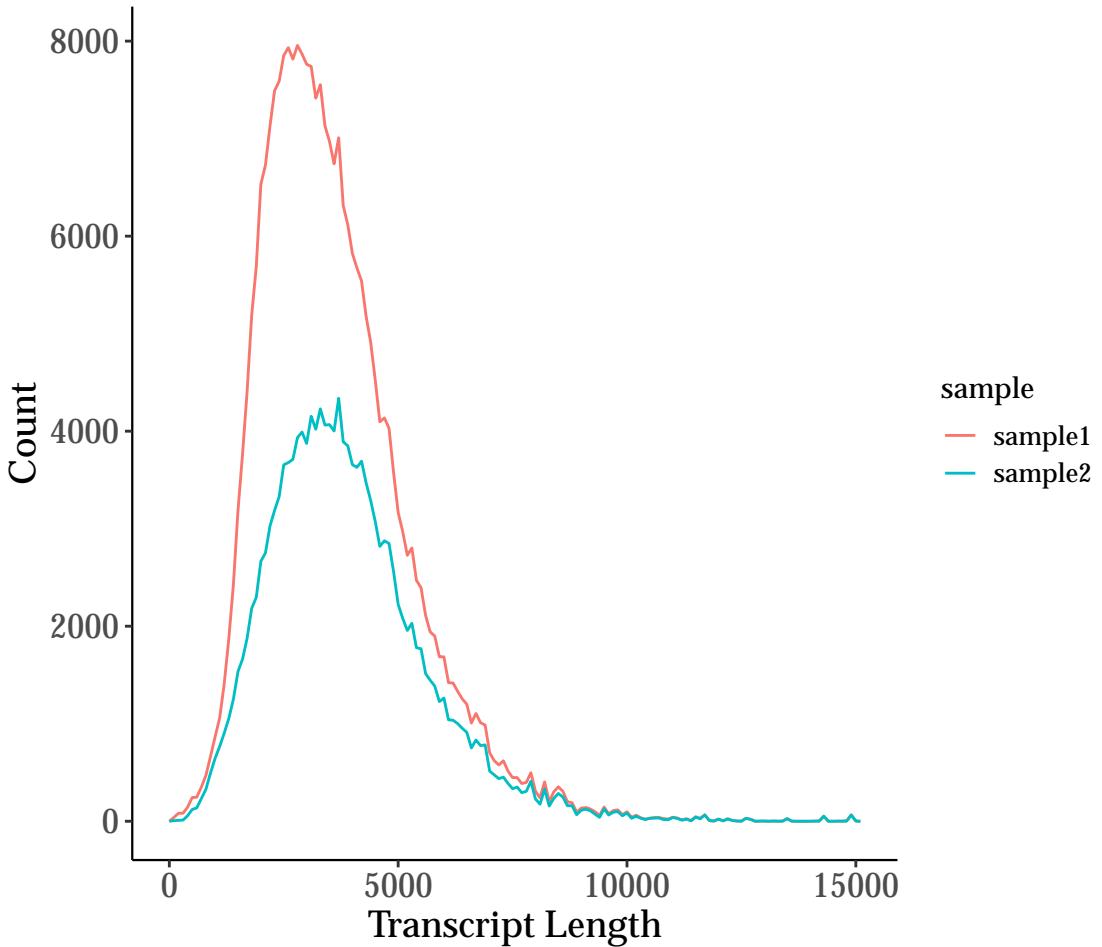
Transcript Lengths, by structural category



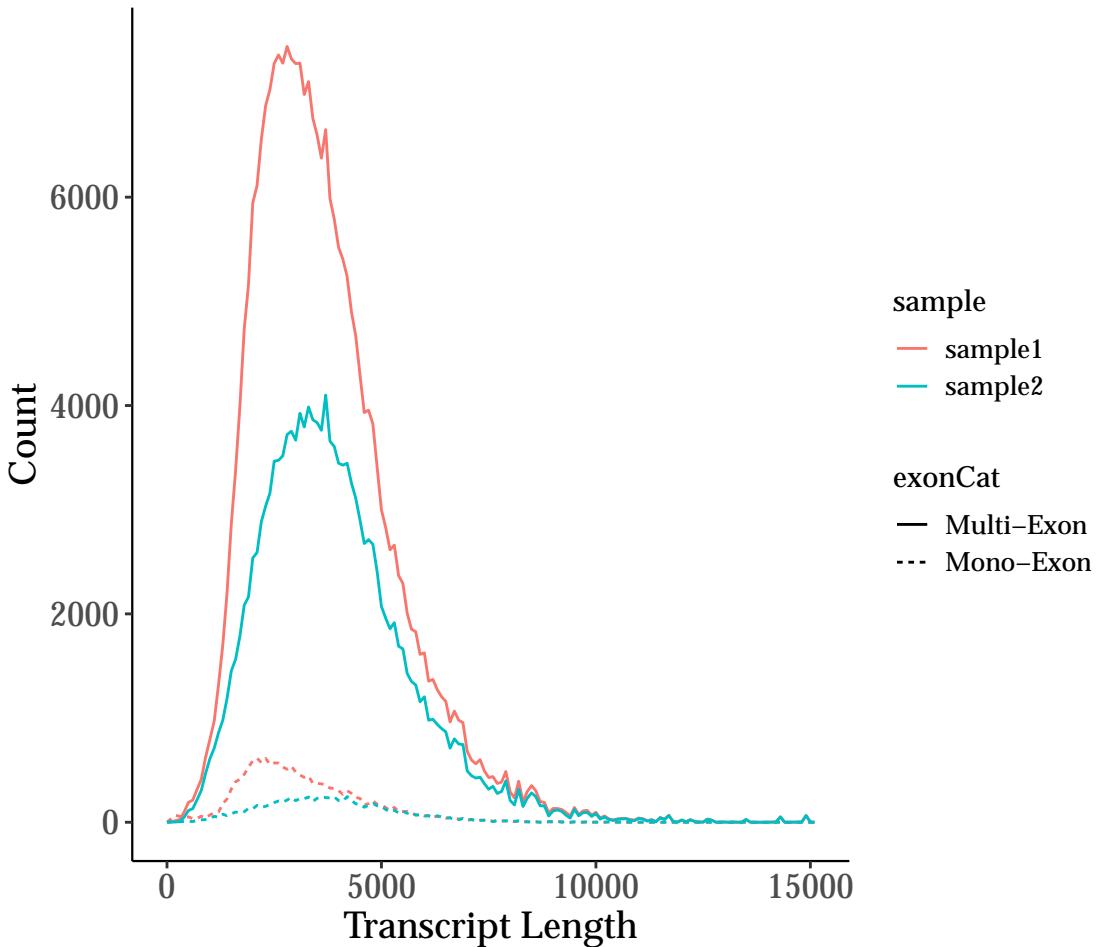
Transcript Lengths, Mono– vs Multi–Exons



Transcript Lengths, By Sample

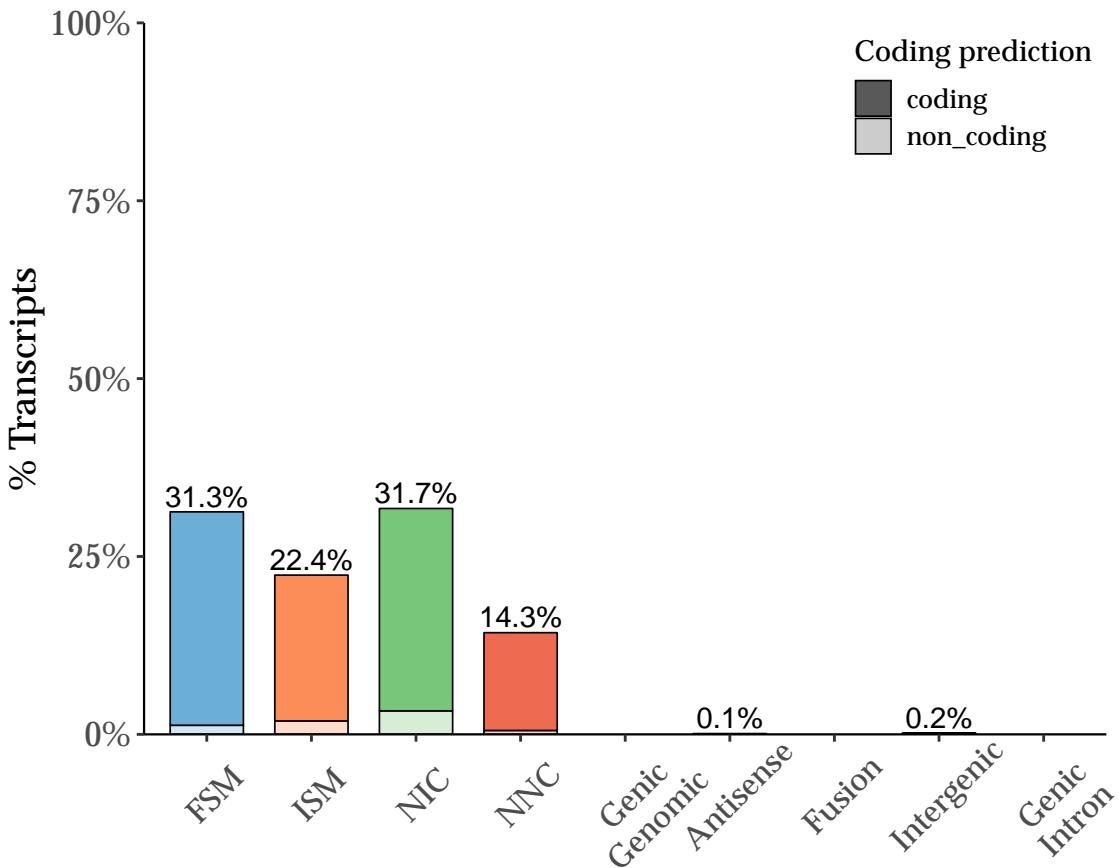


Transcript Lengths, Mono– vs Multi–Exons, By Sample

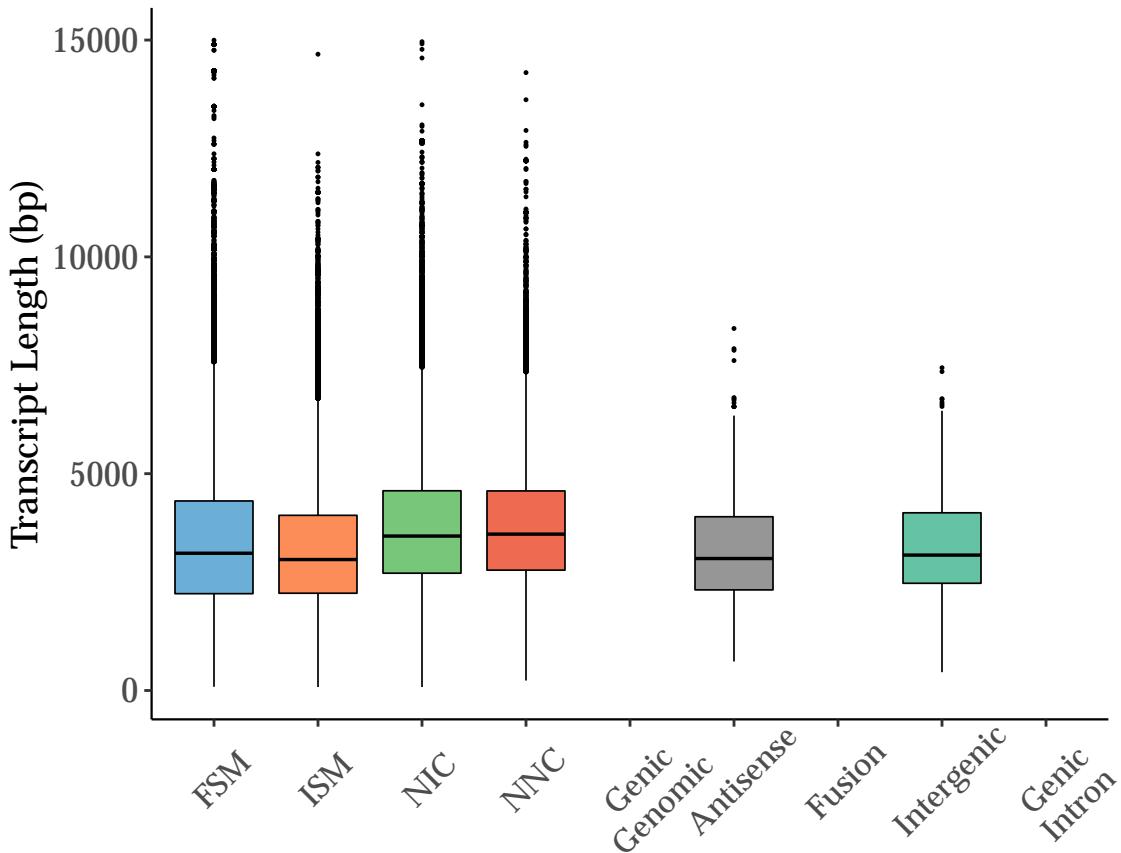


*Structural Isoform Characterization
by Splice Junctions*

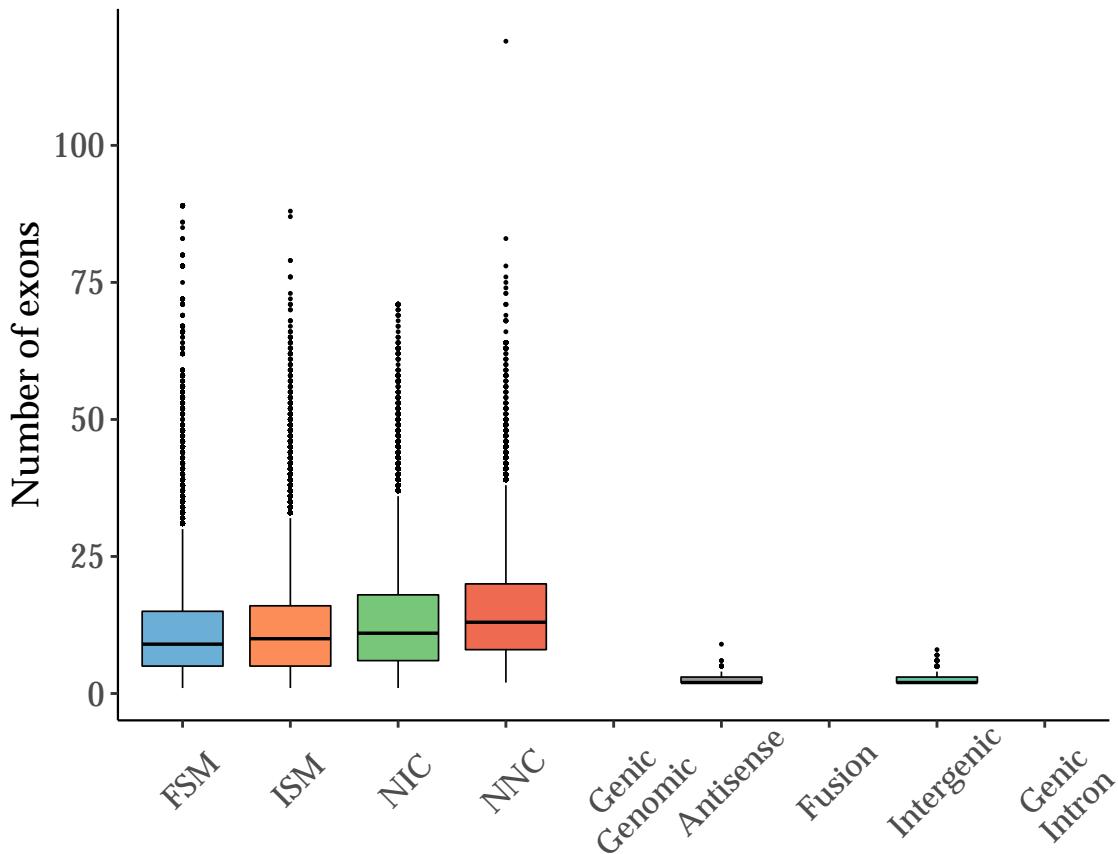
Isoform distribution across structural categories



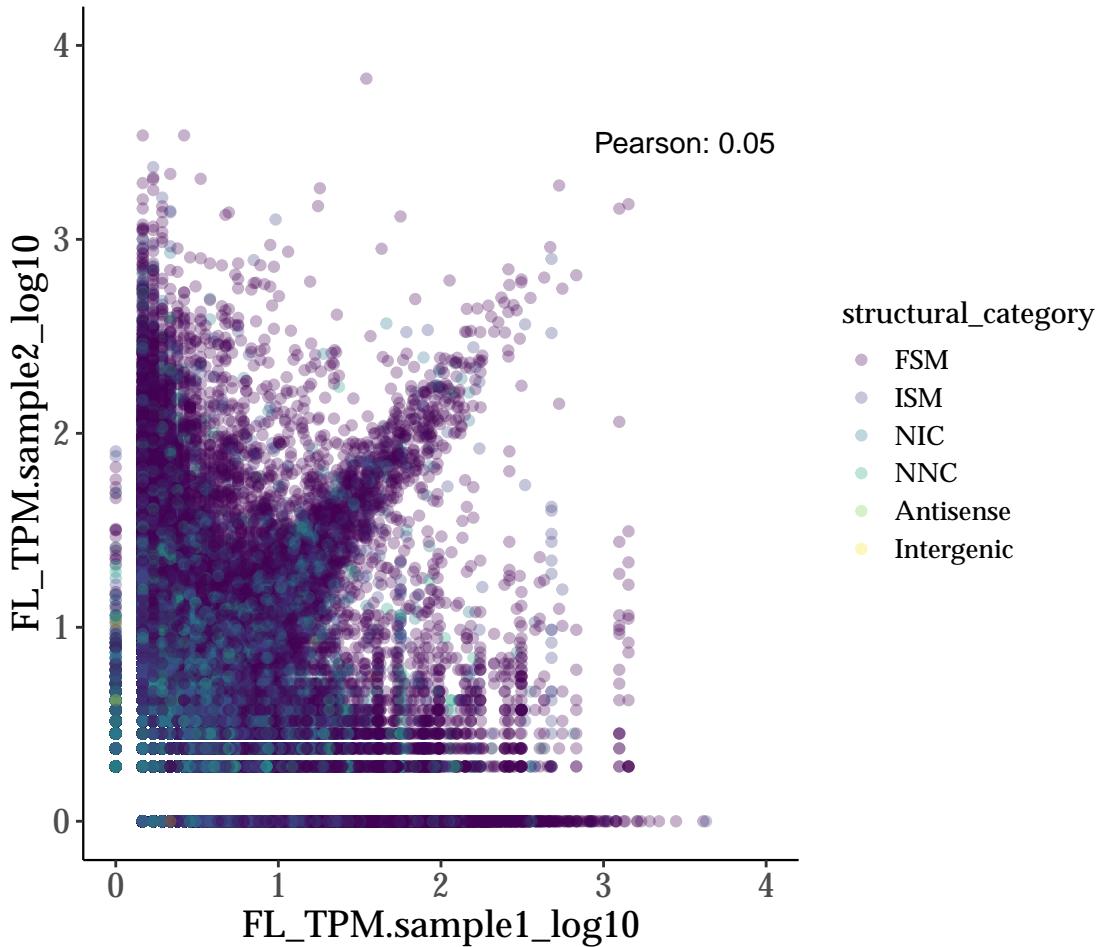
Transcript Lengths by Structural Classification



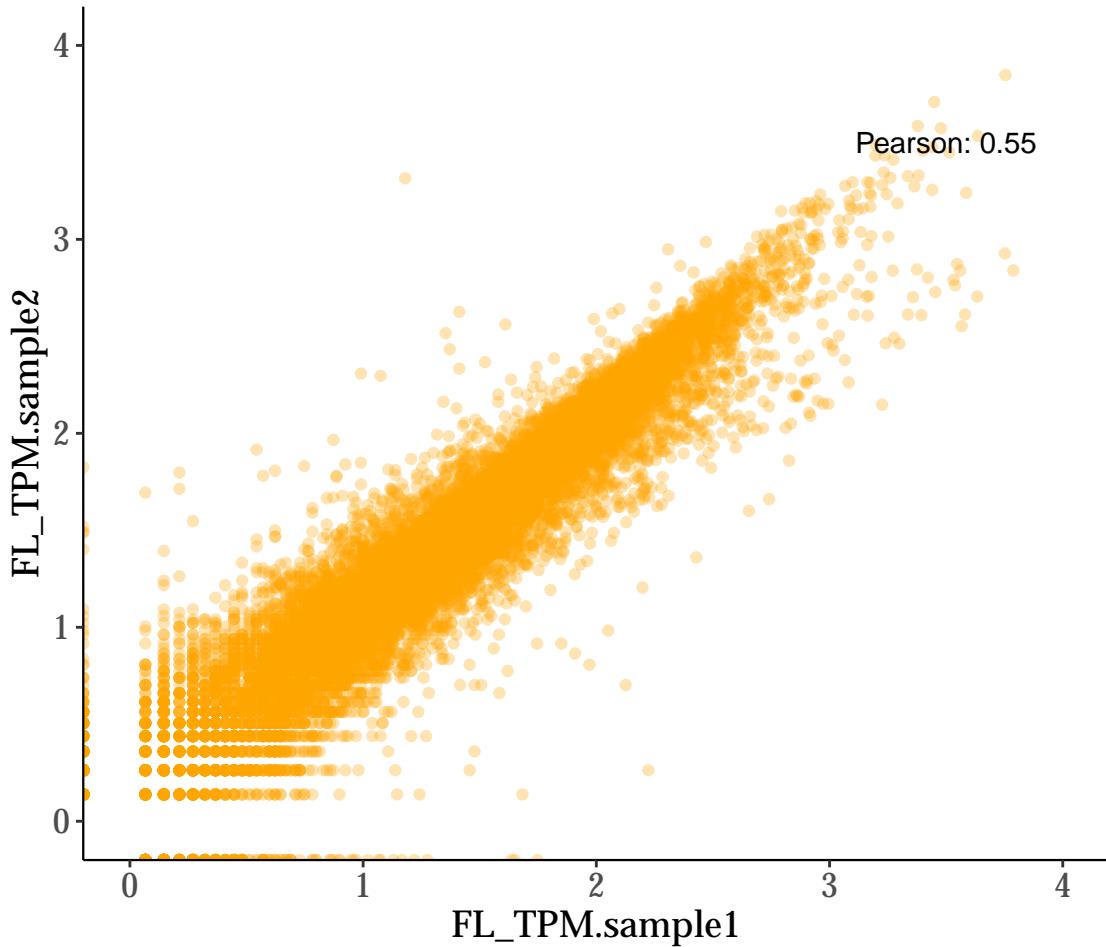
Exon Counts by Structural Classification



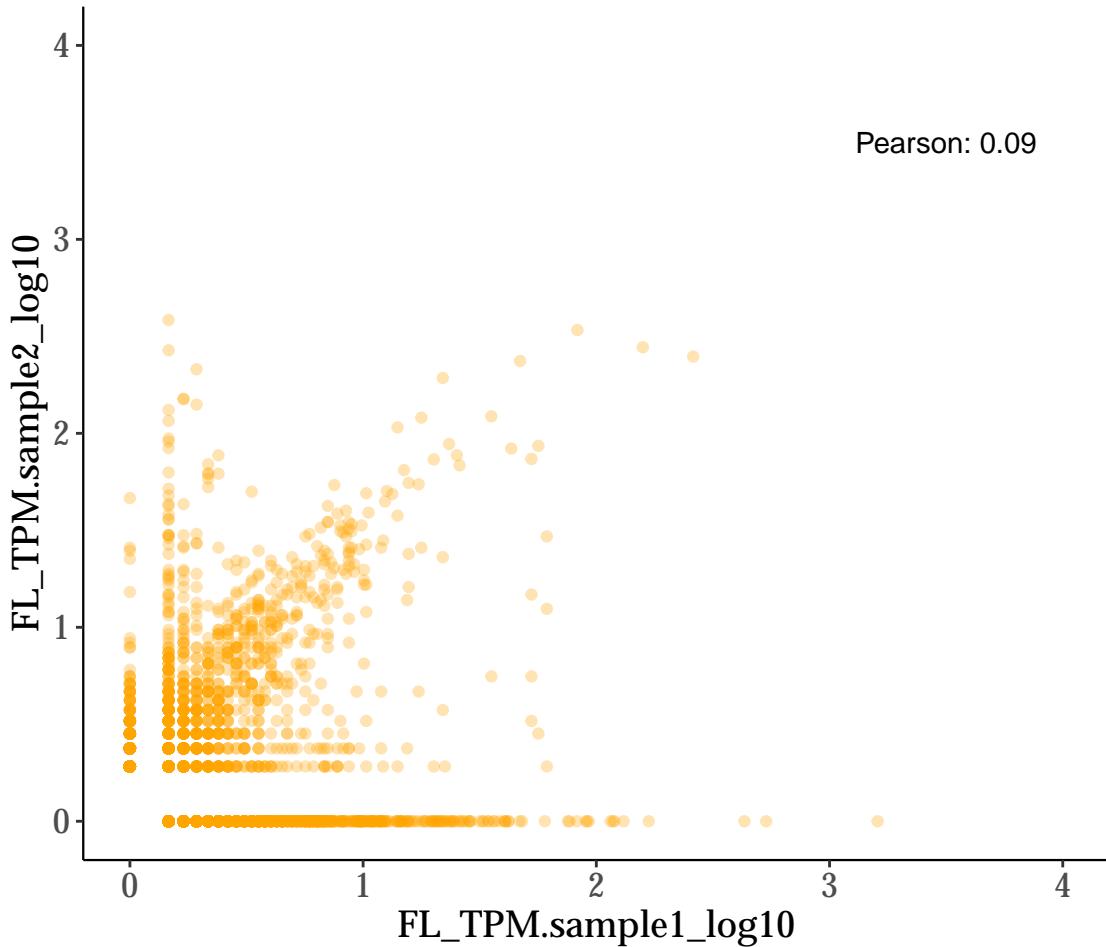
FL TPM (log10 scale) sample1 vs sample2



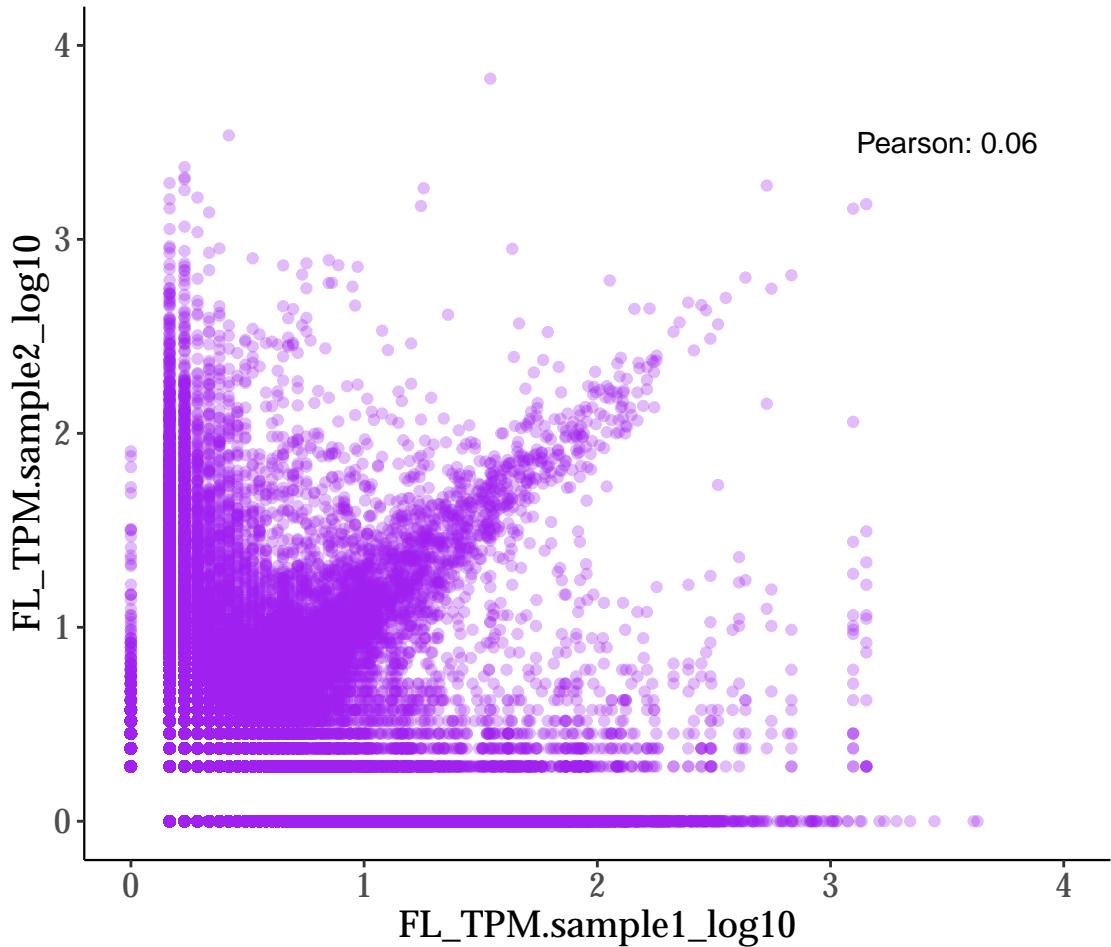
FL TPM (log10 scale) sample1 vs sample2 , grouped by gene



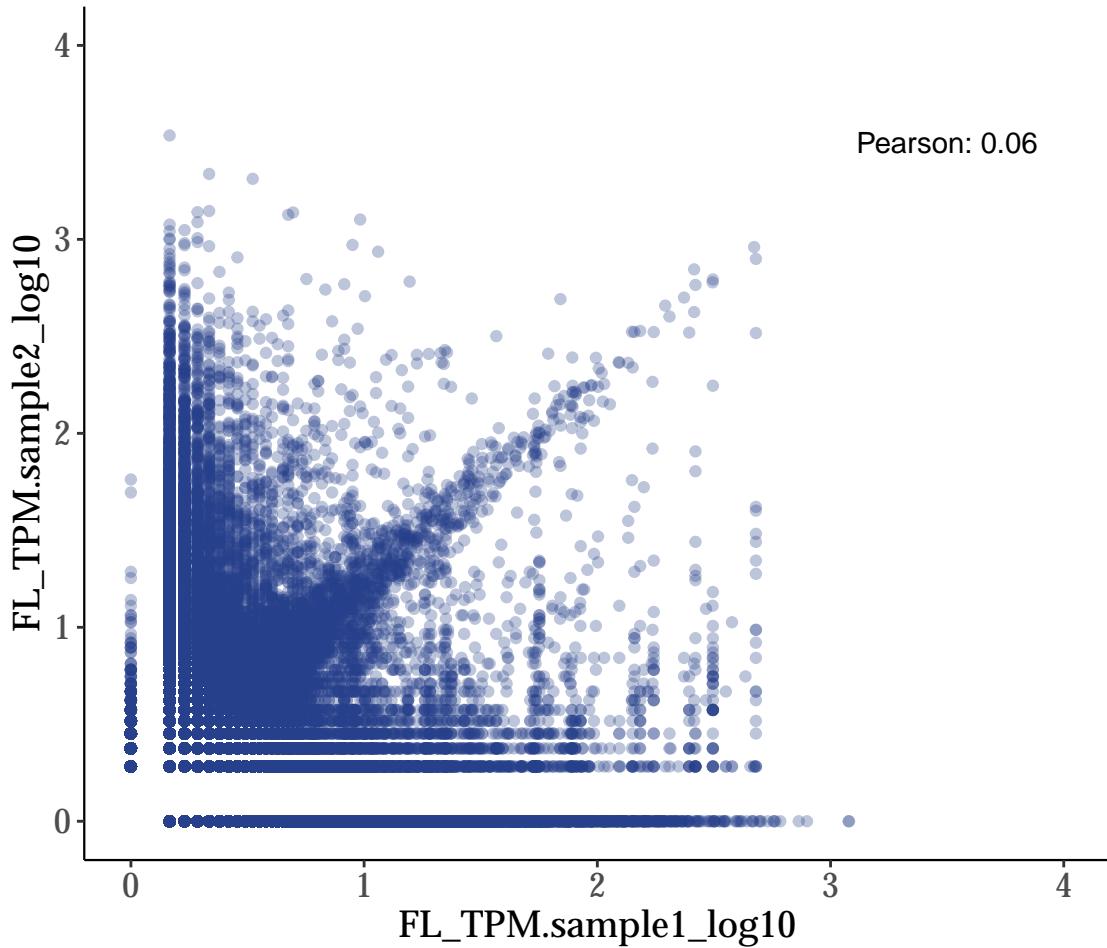
FL TPM (log10 scale) sample1 vs sample2 < 1kb only



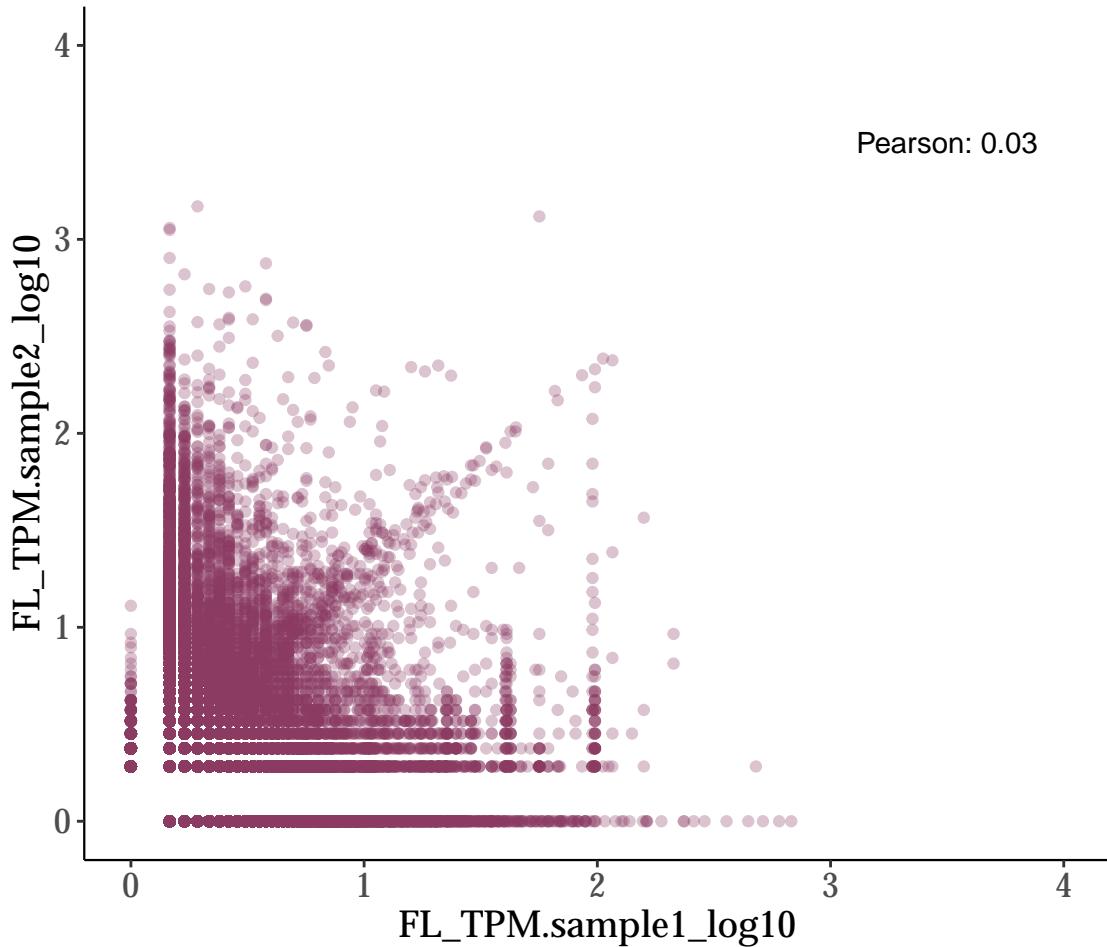
FL TPM (log10 scale) sample1 vs sample2 1–3kb only



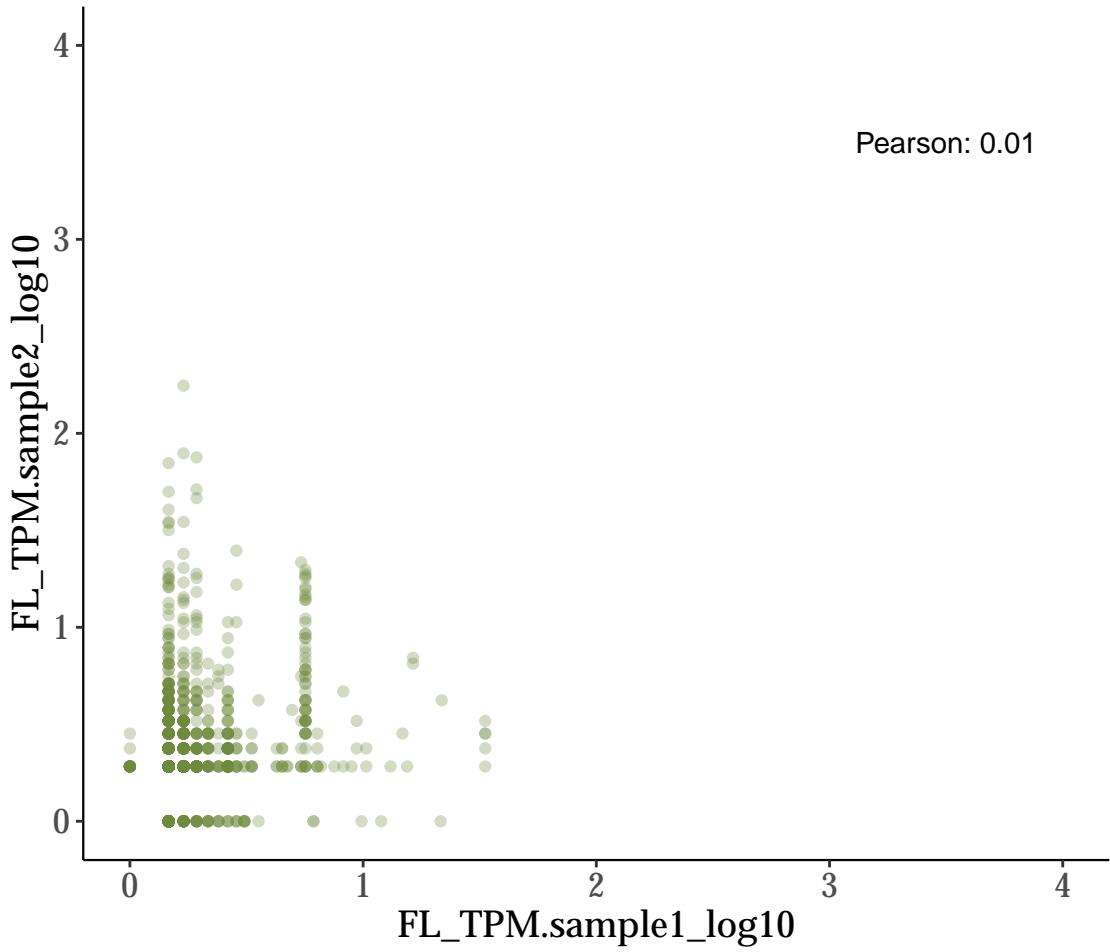
FL TPM (log10 scale) sample1 vs sample2 3–5kb only



FL TPM (log10 scale) sample1 vs sample2 5–10 kb only

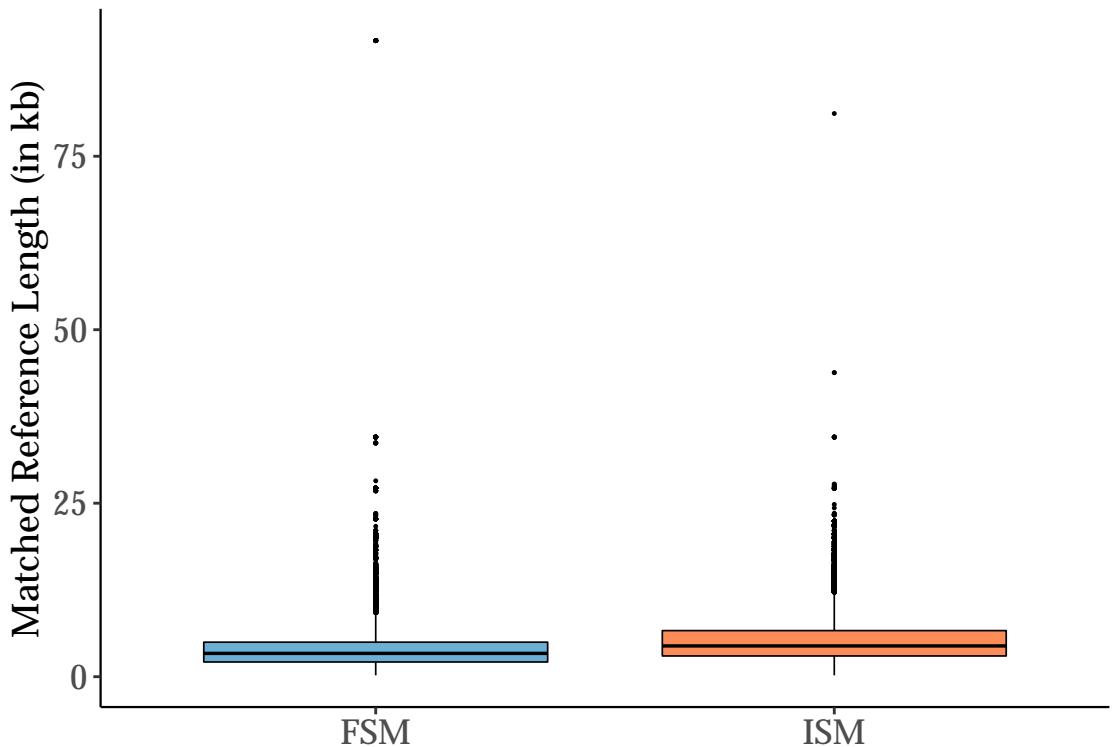


FL TPM (log10 scale) sample1 vs sample2 >10 kb only



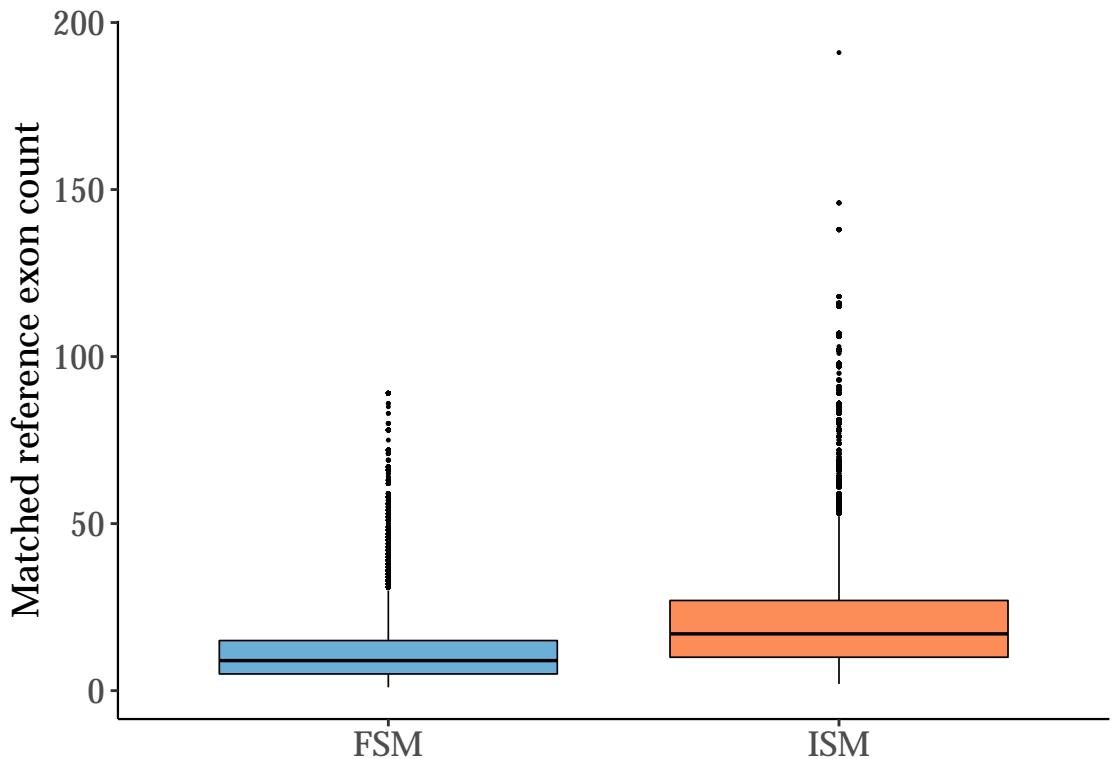
Length Distribution of Matched Reference Transcripts

Applicable only to FSM and ISM categories



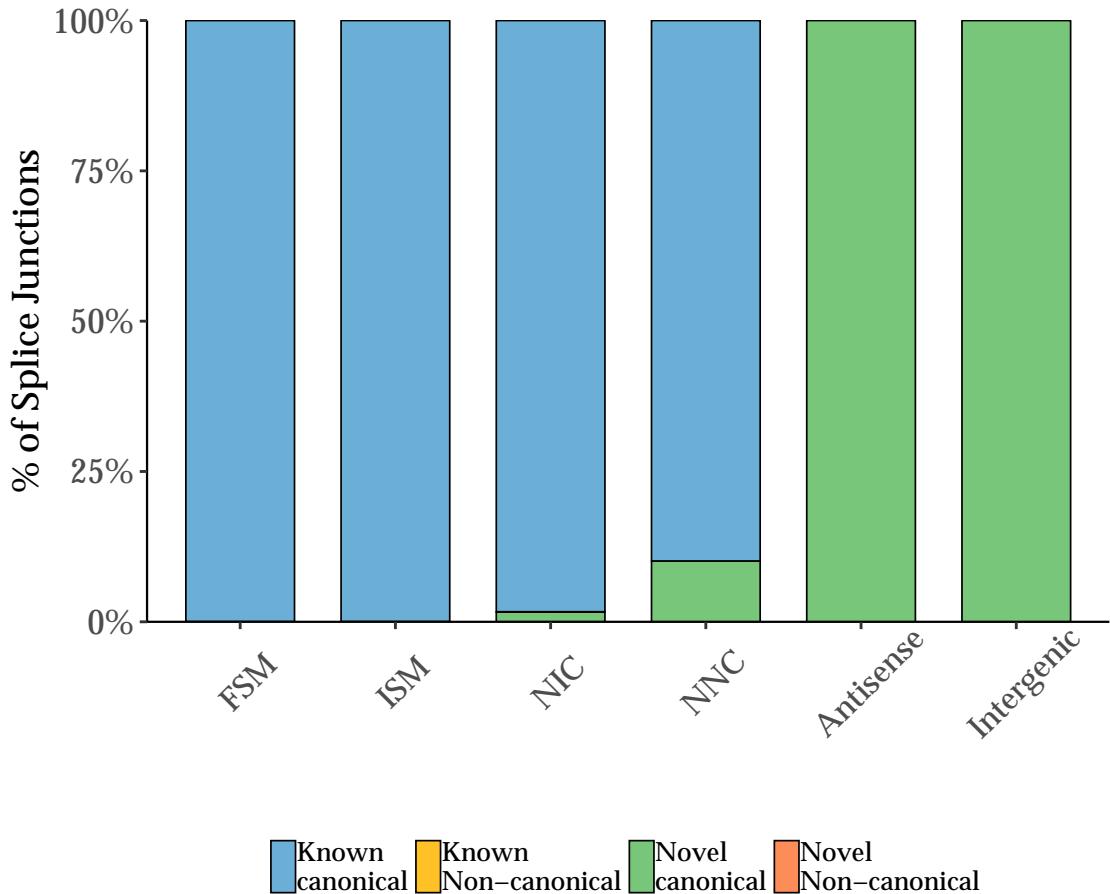
Exon Count Distribution of Matched Reference Transcripts

Applicable only to FSM and ISM categories

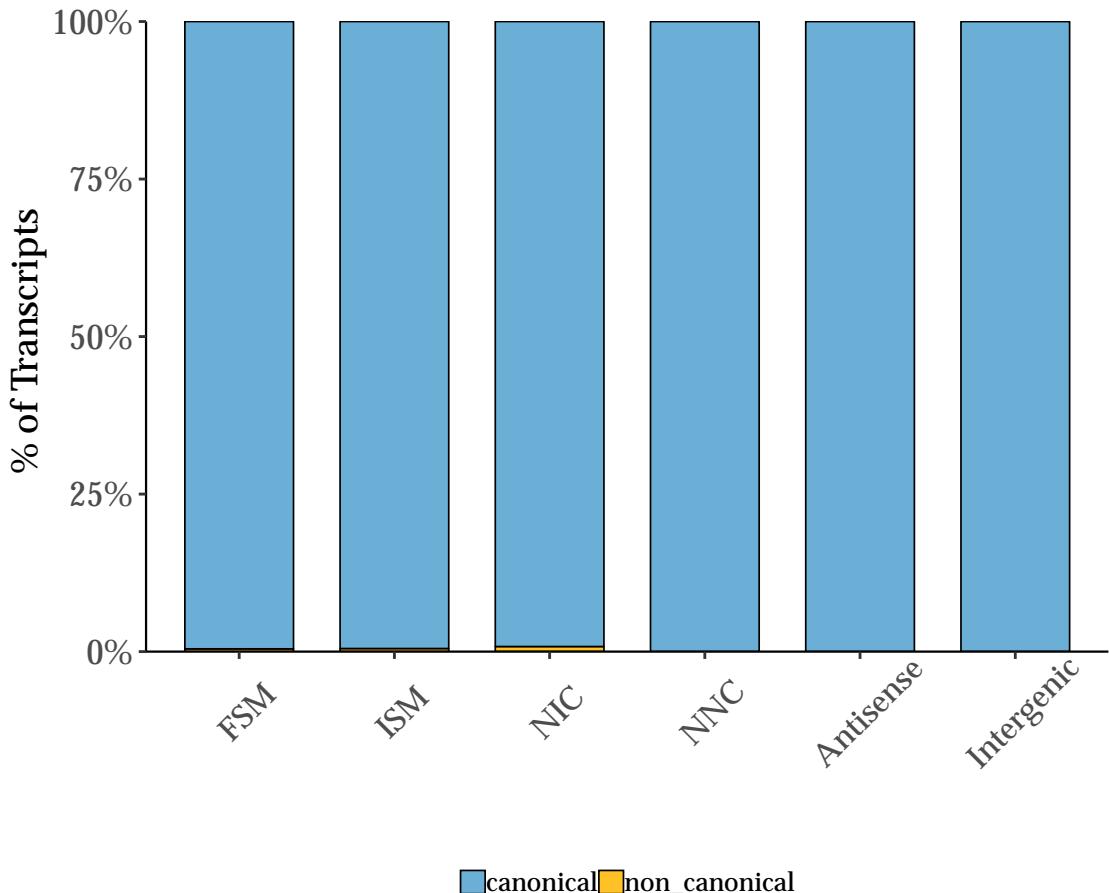


Splice Junction Characterization

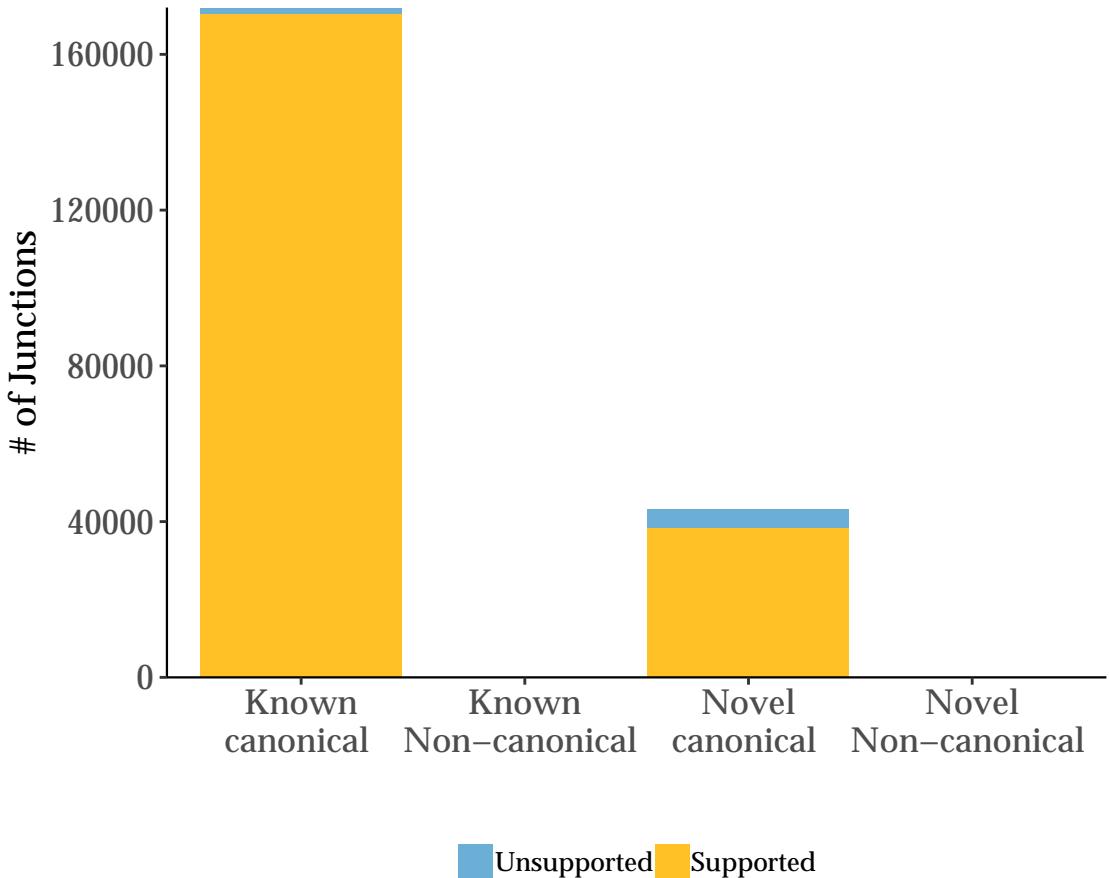
Distribution of Splice Junctions by Structural Classification



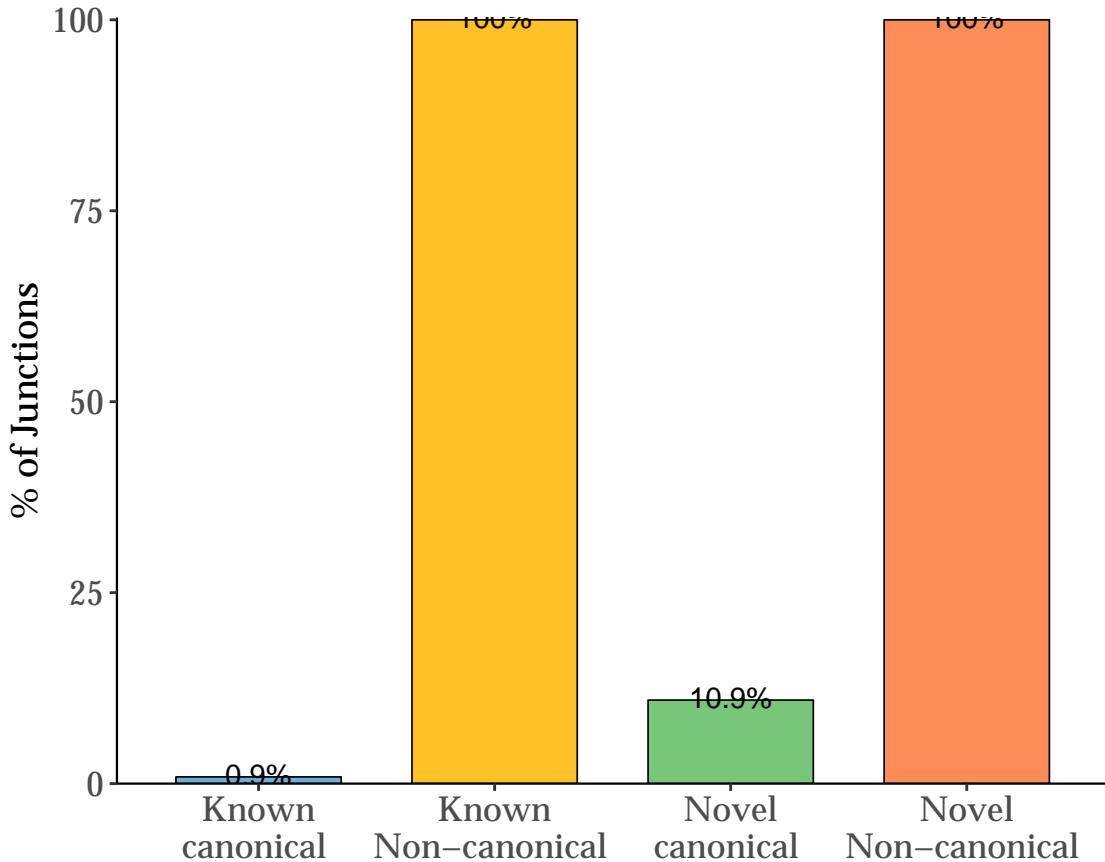
Distribution of Transcripts by Splice Junctions



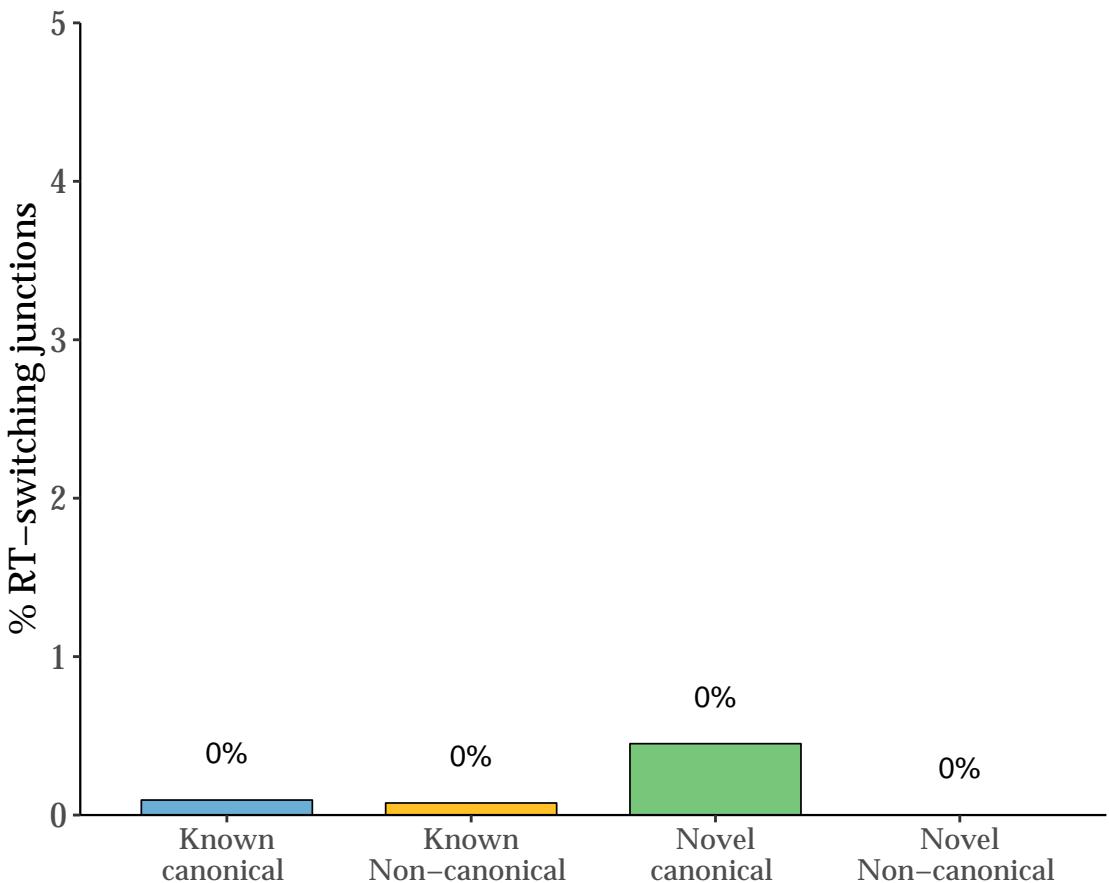
Unique junctions w/ or w/out short read coverage



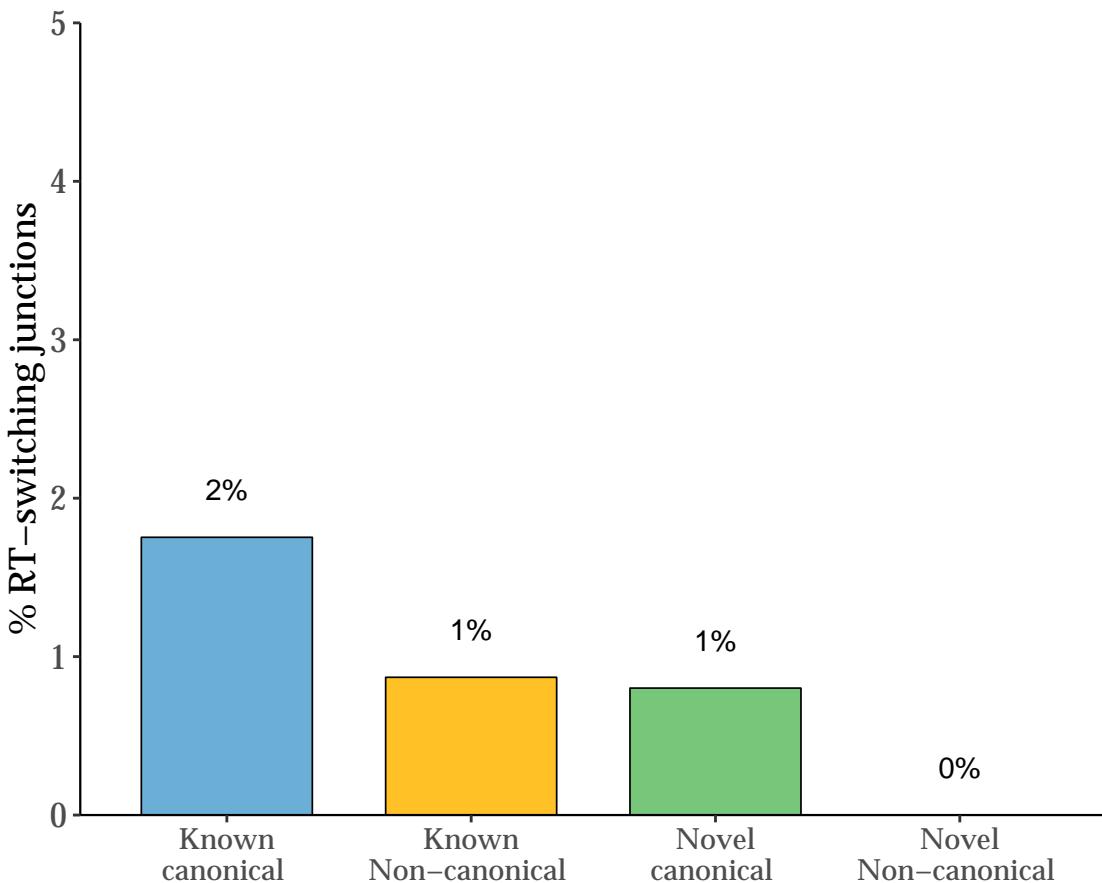
Unique junctions w/out short read coverage (percentage)



RT-switching, all junctions



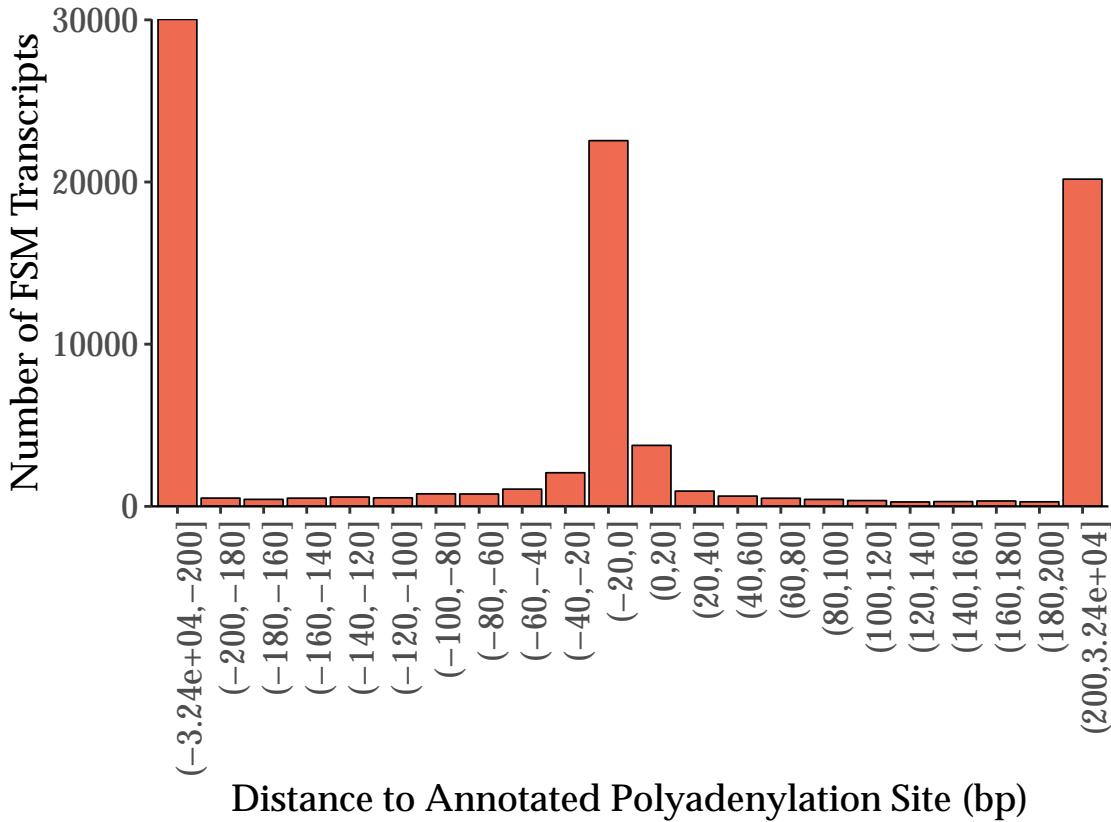
RT-switching, unique junctions



Comparison with Annotated TSS and PolyA Sites

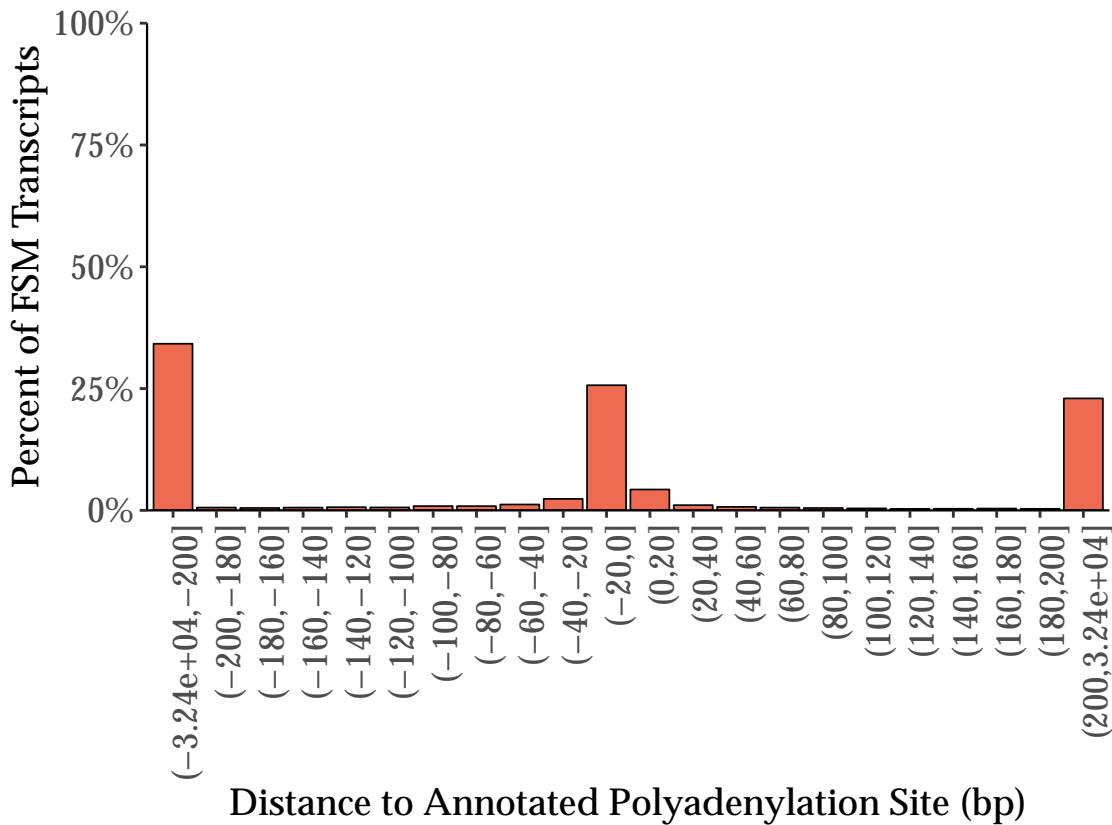
Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



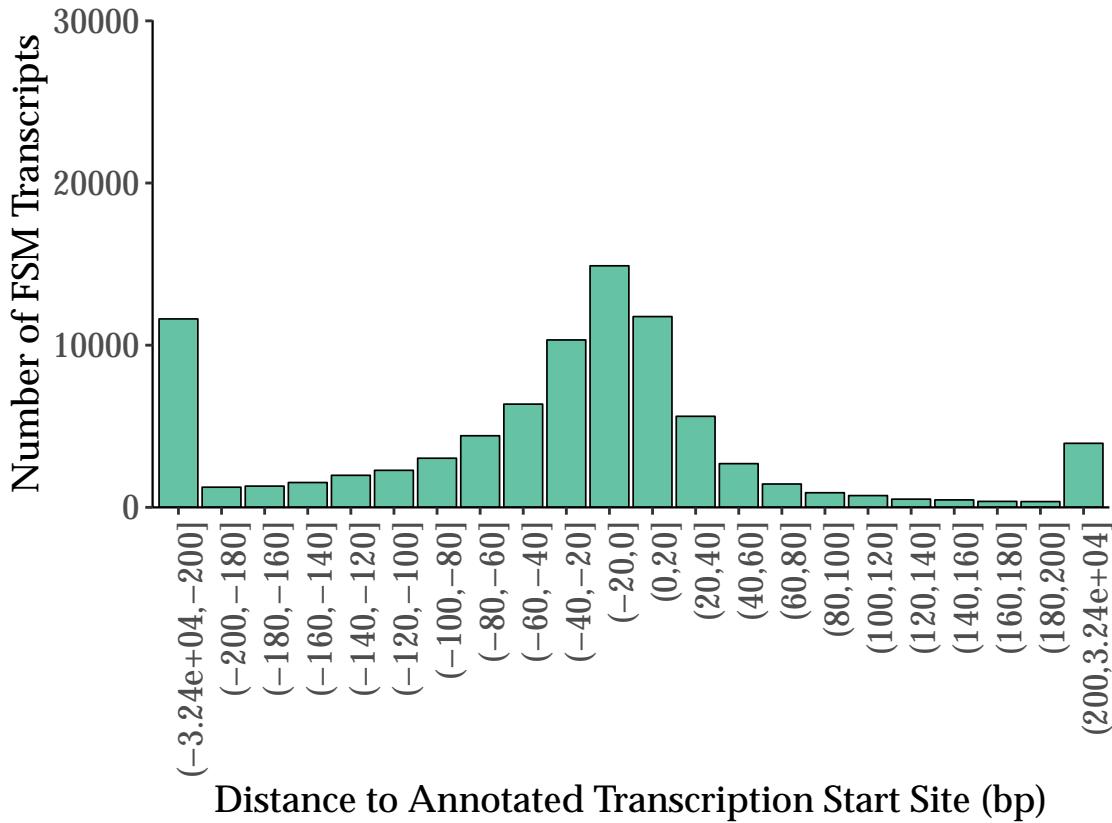
Distance to Annotated Polyadenylation Site, FSM only

Negative values indicate upstream of annotated polyA site



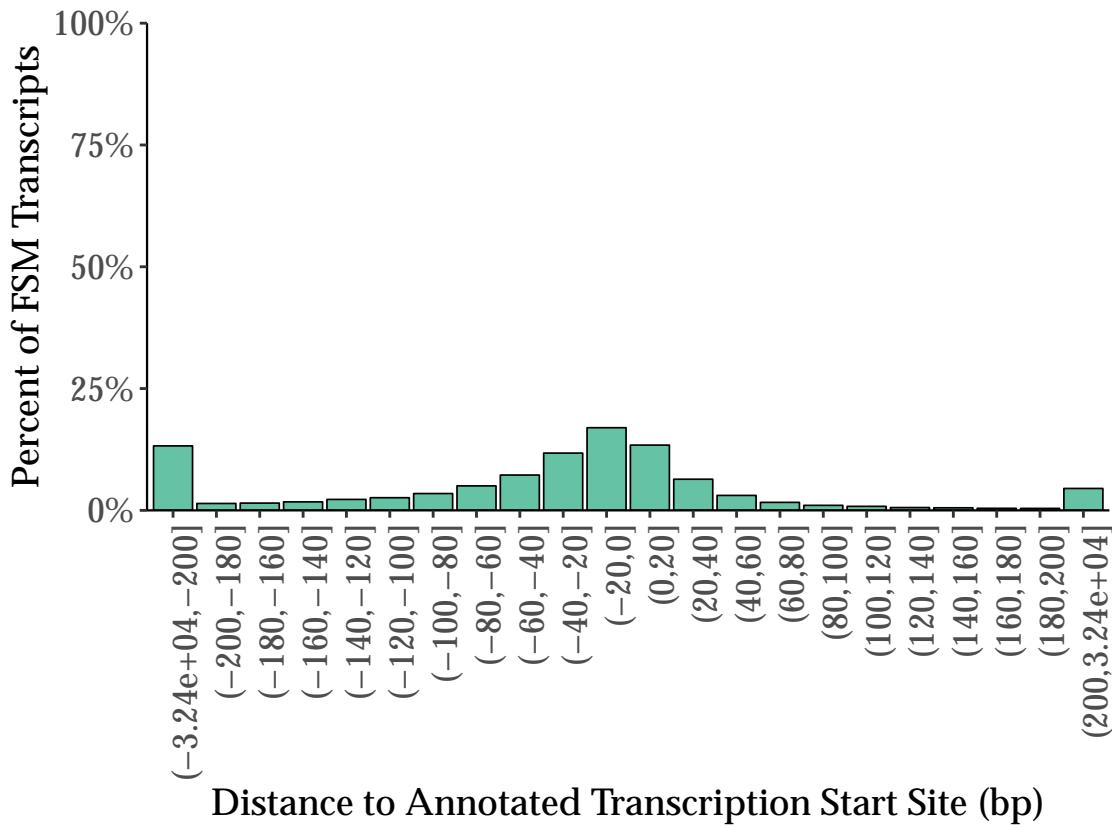
Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



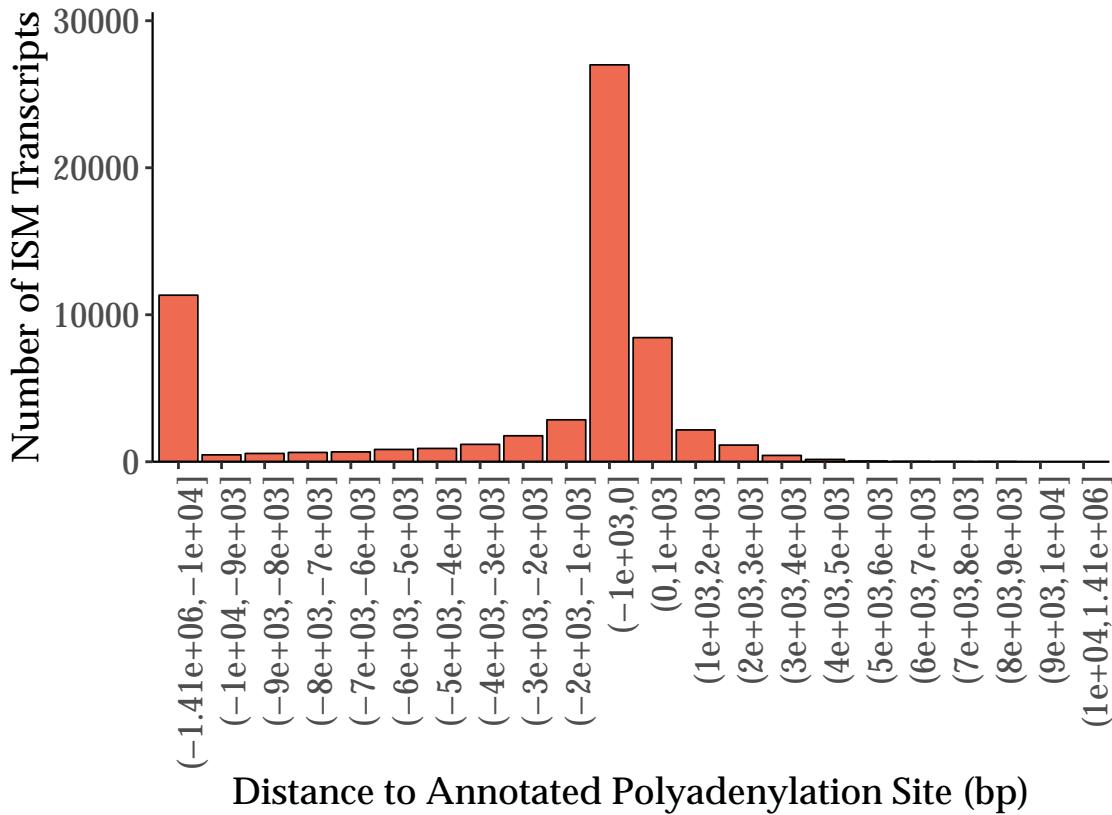
Distance to Annotated Transcription Start Site, FSM only

Negative values indicate downstream of annotated TSS



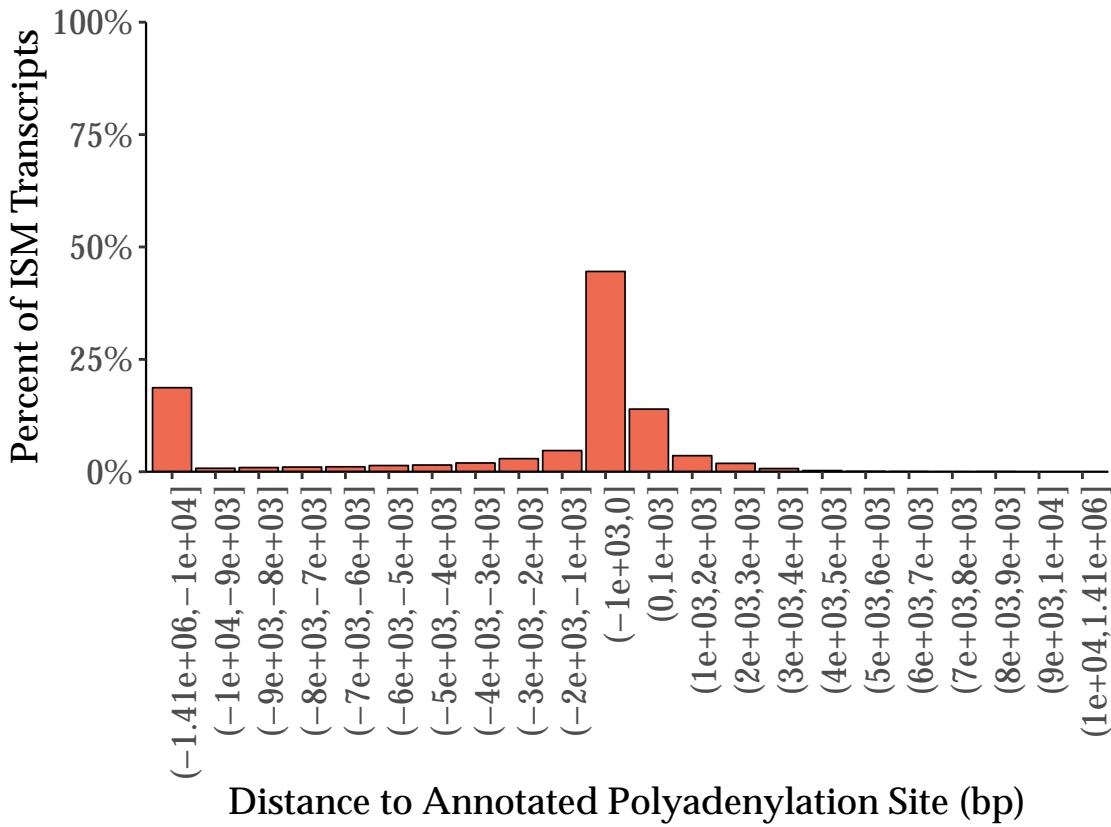
Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



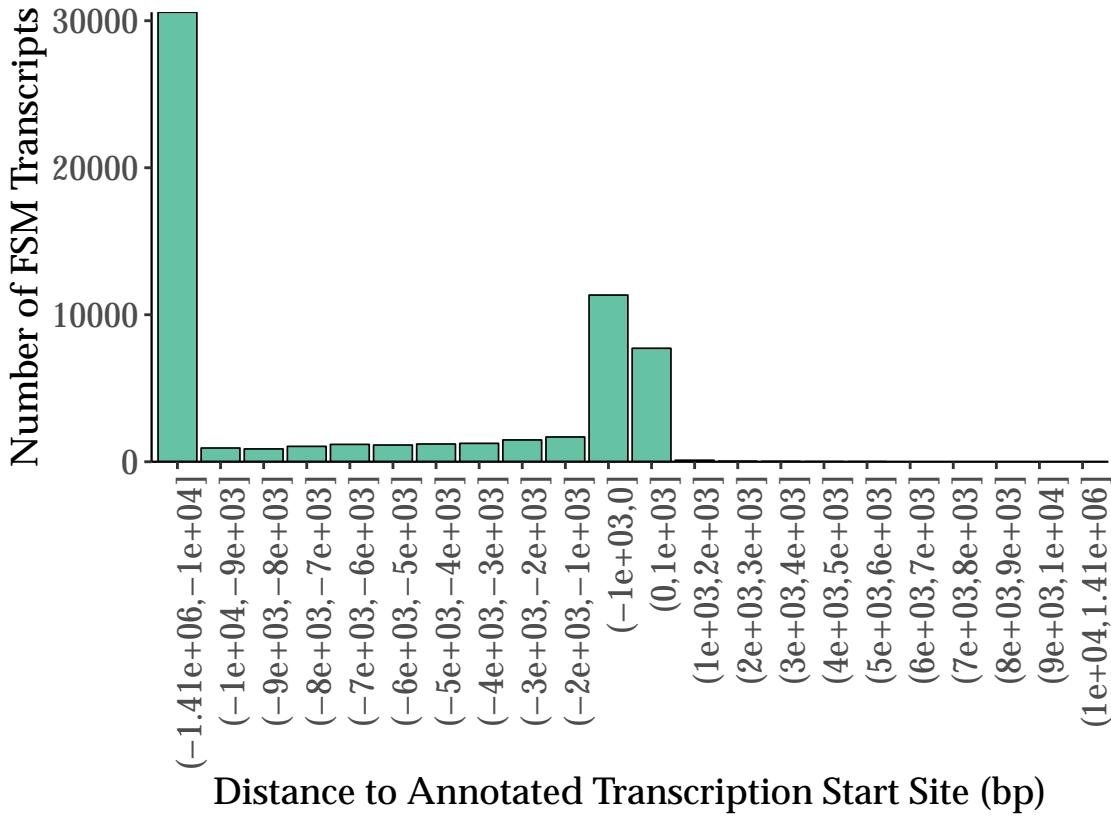
Distance to Annotated Polyadenylation Site, ISM only

Negative values indicate upstream of annotated polyA site



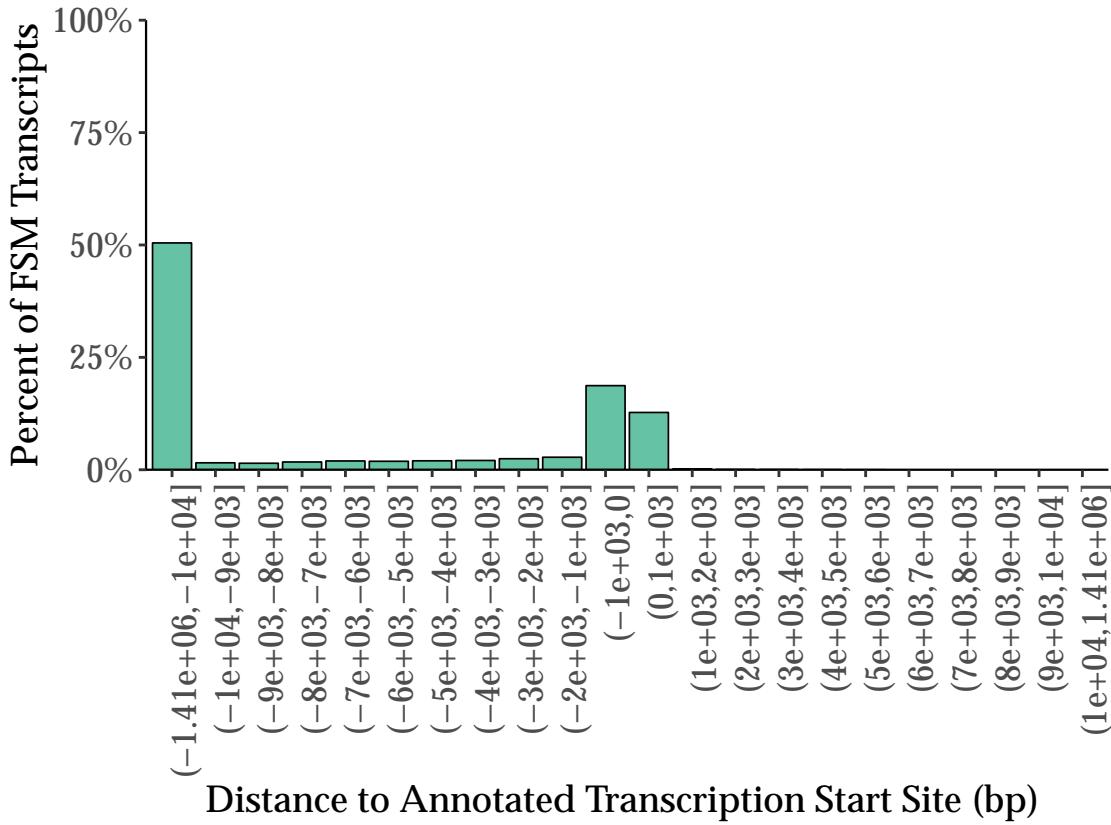
Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS

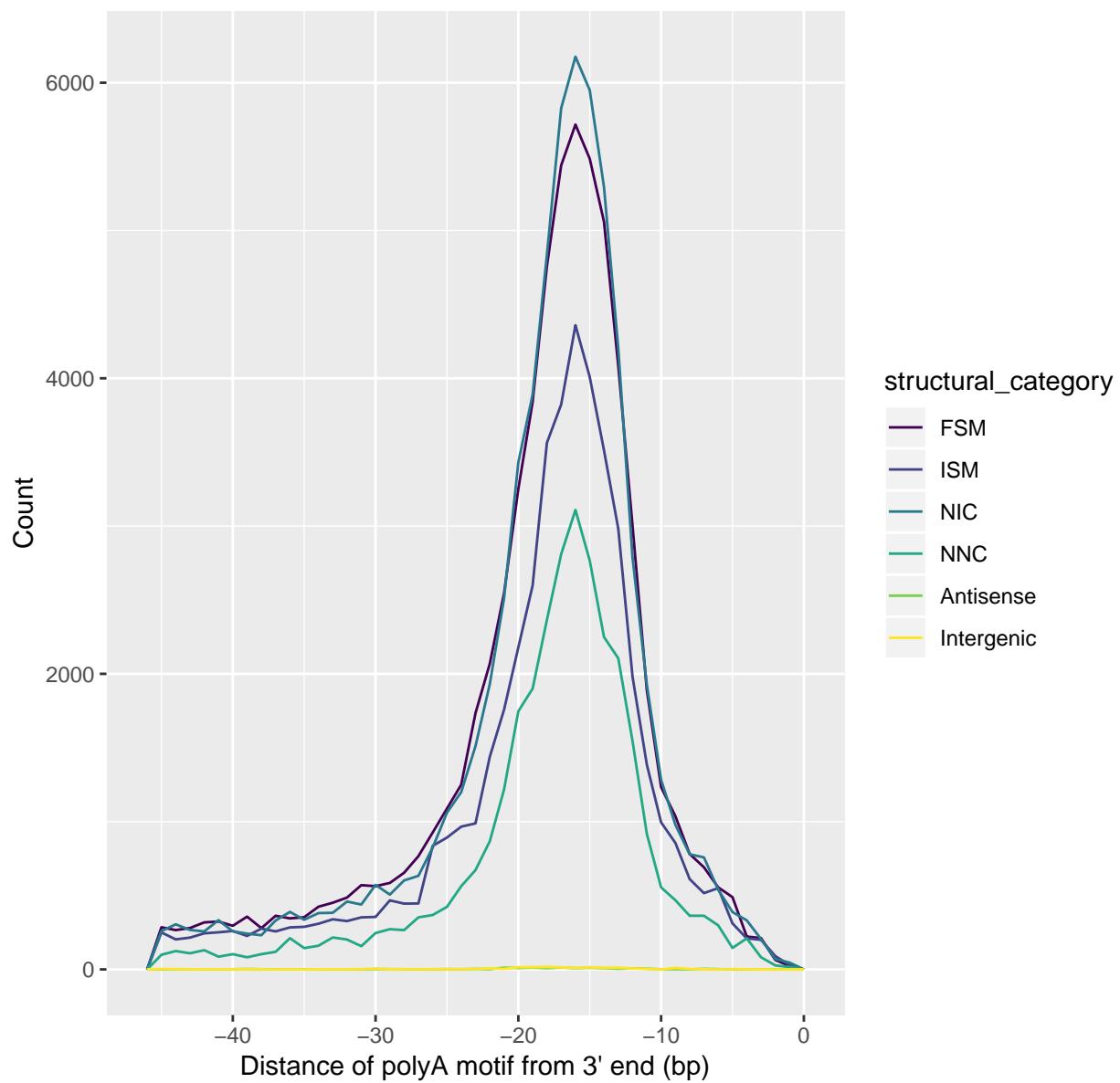


Distance to Annotated Transcription Start Site, ISM only

Negative values indicate downstream of annotated TSS



Distance of detected polyA motif from 3' end



Frequency of polyA motifs

Number of polyA Motifs Detected

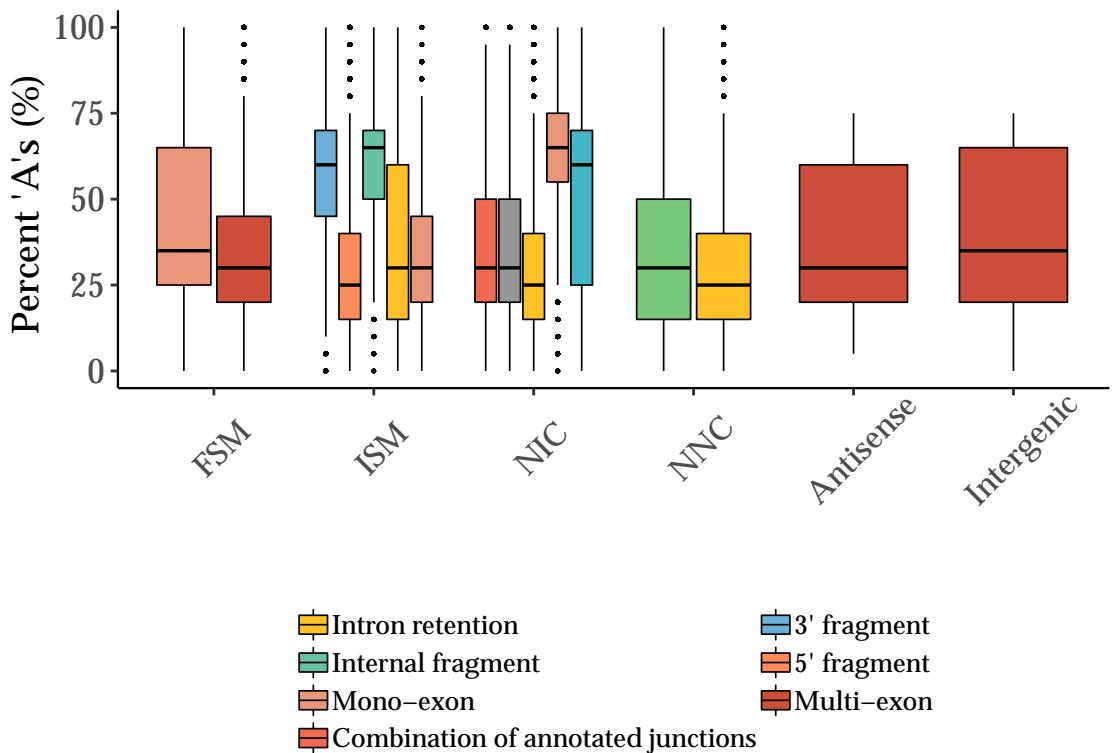
Category	Count	polyA Detected	%
FSM	90643	65410	72
ISM	64879	47431	73
NIC	92008	65917	72
NNC	41422	31332	76
Antisense	317	130	41
Intergenic	548	181	33

Motif	Count	%
AATAAA	132254	62.9
ATTAAA	33713	16.0
AGTAAA	6813	3.2
TATAAA	6297	3.0
AAGAAA	4211	2.0
TTTAAA	3859	1.8
AATACA	3496	1.7
AATATA	3360	1.6
GATAAA	3084	1.5
CATAAA	3032	1.4
AAAAAG	3001	1.4
AATGAA	2132	1.0
AAAACA	1837	0.9
AATAGA	1584	0.8
ACTAAA	1414	0.7
GGGGCT	314	0.1

Intra-Priming Quality Check

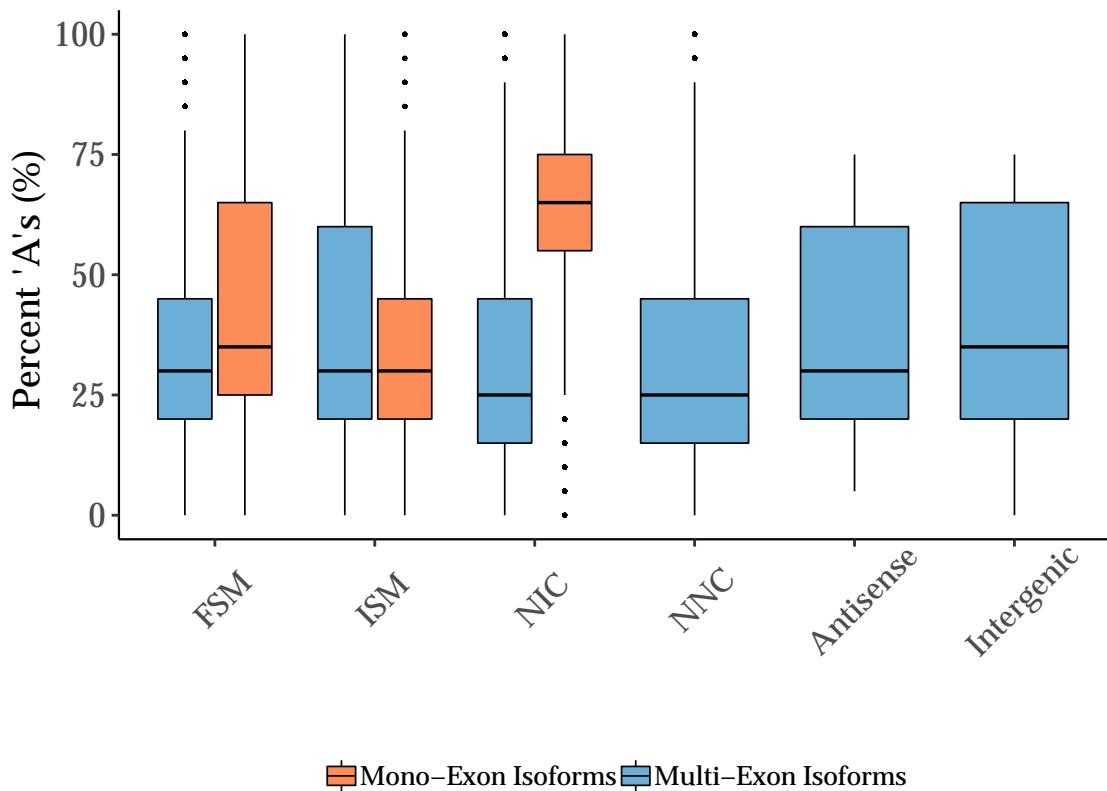
Possible Intra–Priming by Structural Category

Percent of genomic 'A's in downstream 20 bp



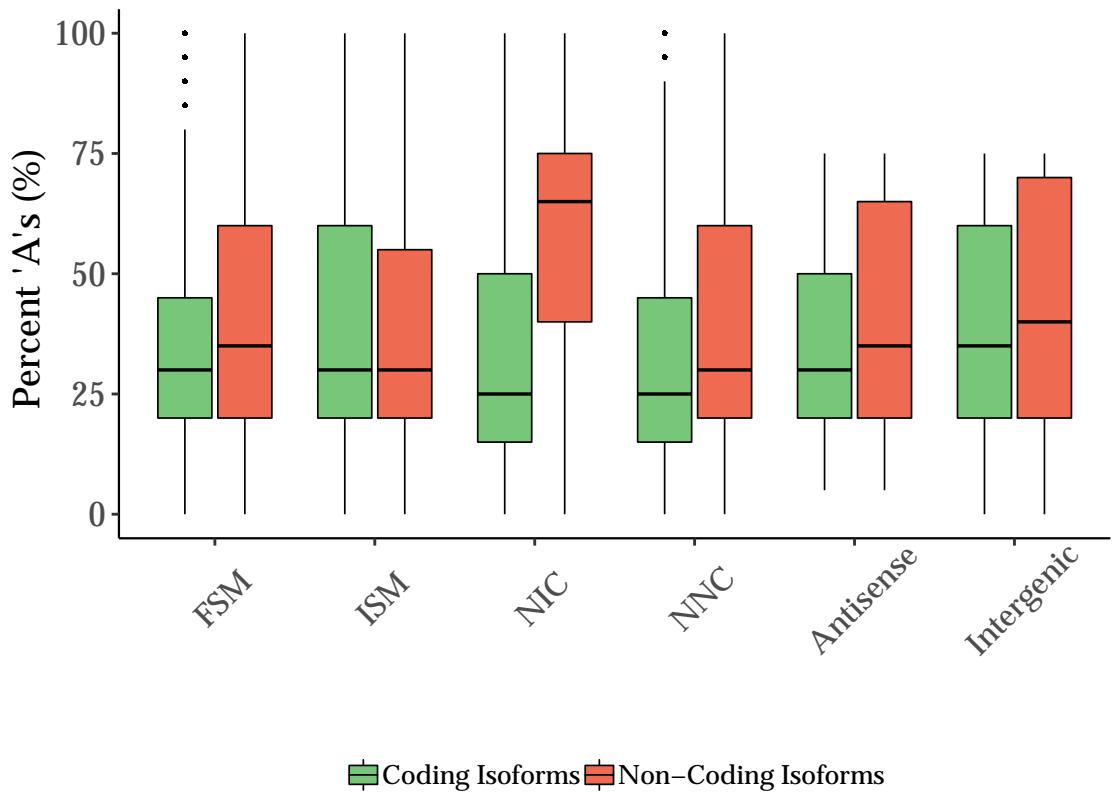
Possible Intra–Priming, Mono– vs Multi–Exon

Percent of genomic 'A's in downstream 20 bp



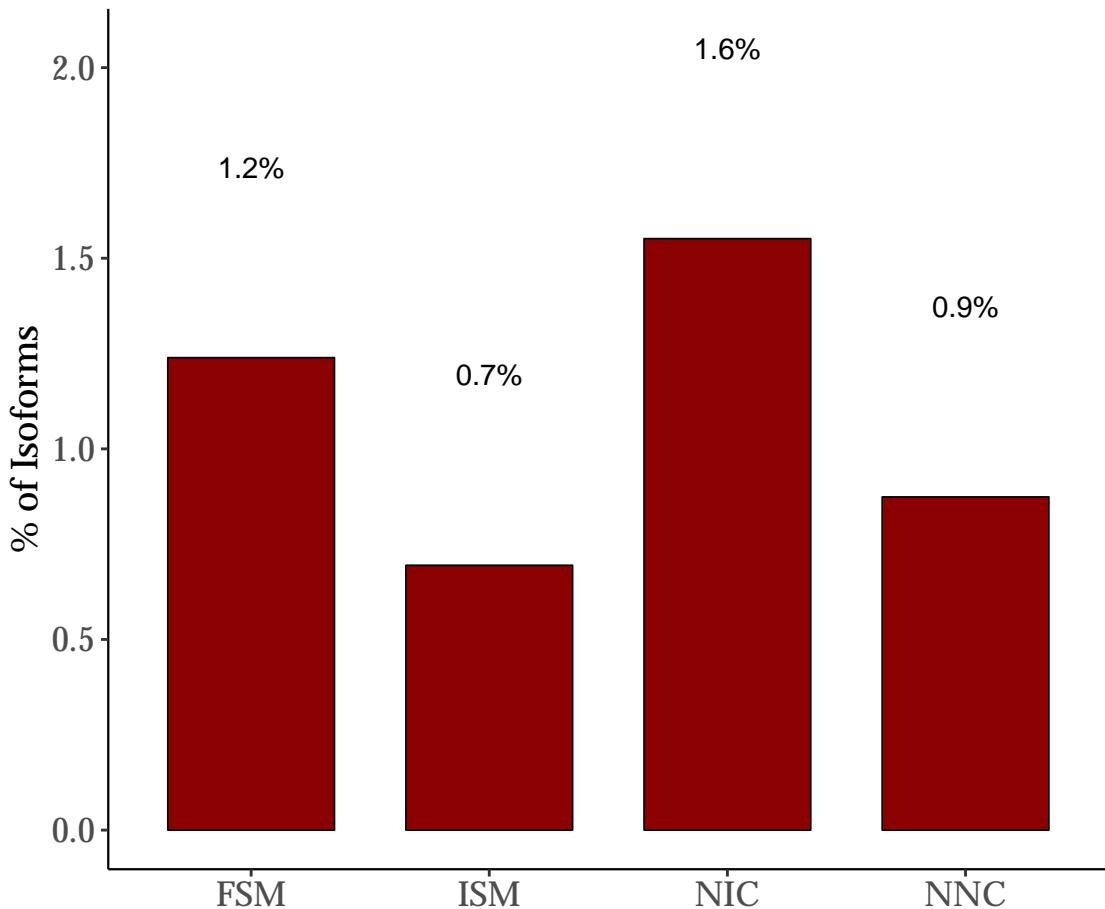
Possible Intra-Priming, Coding vs Non-Coding

Percent of genomic 'A's in downstream 20 bp



Quality Controls

Incidence of RT-switching



Incidence of Non-Canonical Junctions

