Supplemental Material for

Splicing kinetics and coordination revealed by direct nascent RNA sequencing through nanopores

Heather L. Drexler, Karine Choquet, L. Stirling Churchman

1. Figures S1 to S7
2. Tables S2 to S4

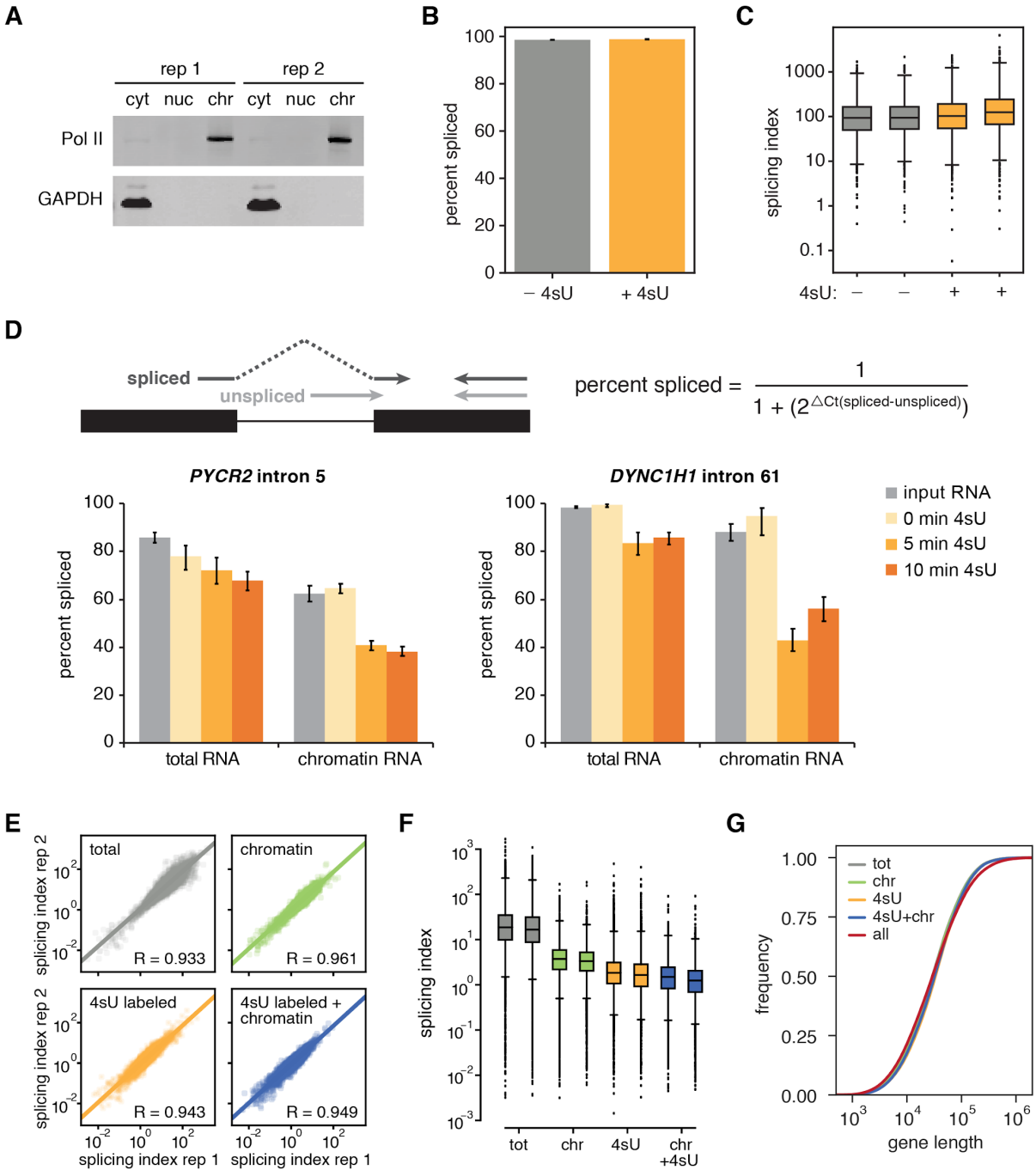
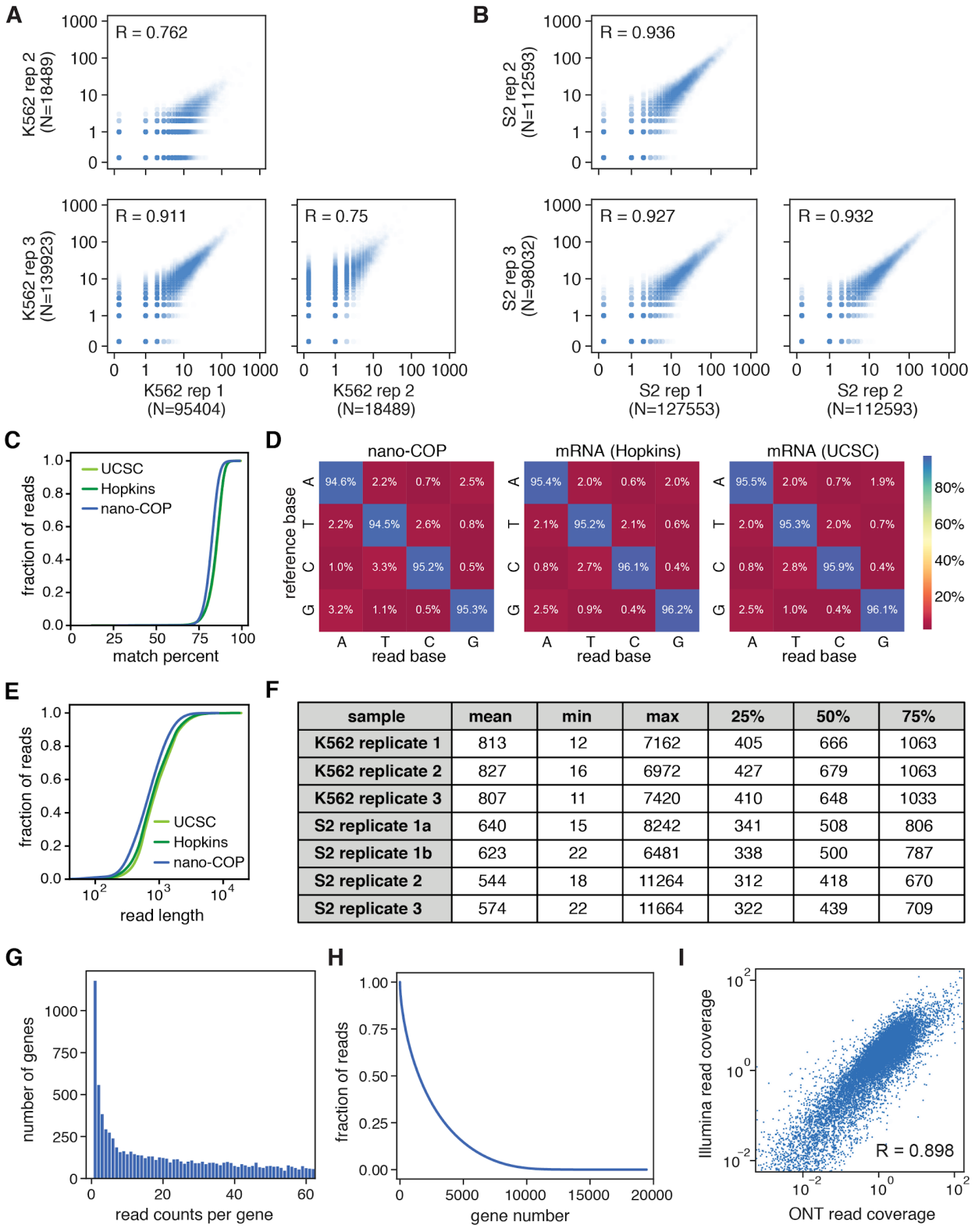


Figure S1. Nascent RNA enrichment with cellular fractionation and 4sU metabolic labeling (related to Figure 1). (A) Western blot of subcellular fractions from K562 cells with primary antibodies against Pol II Ser2 phosphoisoform (elongating Pol II) and GAPDH. cyt = cytoplasm; nuc = nucleoplasm; chr = chromatin. (B-C) Total RNA-seq datasets of cells incubated with and without 100 µM 4sU for 4 hours acquired from (Schofield et al., 2018) were compared for (B) global percent of spliced molecules (less than 0.5% increase in the percent of spliced reads in +4sU sample; chi-square p-value < 0.05) and (C) splicing index (increased distribution of splicing index in +4sU samples; t-test p-value < 0.05). (D) Measurement of the percent of spliced

molecules at intron 5 in the *PYCR2* gene and intron 61 of the *DYNC1H1* gene in human K562 cells using a RT-qPCR assay with forward primers in either the upstream exon or tested intron (top). Cells were either left untreated or incubated with 4sU for indicated time points. Total RNA and chromatin-associated RNA were purified from each sample. All 4sU labeled RNA was isolated through biotinylation and separation. (E) Correlation plots of splicing index for each Illumina sequencing library. Pearson's R is labeled at the bottom right of each plot. (F) Distribution of the splicing index for each sample across two biological replicates. tot = total RNA; chr = chromatin-associated RNA; 4sU = 4sU labeled RNA; chr+4sU = 4sU labeled chromatin-associated RNA. The distribution of 4sU+chr splicing index differs significantly (t-test p-value < 0.05) from all other samples in each replicate. (G) Cumulative distribution plot of gene lengths for transcripts with RPKM > 1 in each dataset compared with all multi-exon genes (all). The distribution of gene lengths represented in the 4sU labeled chromatin associated RNA sample is moderately larger than those represented in the chromatin associated RNA sample (t-test p-value < 0.05) due to more opportunity for longer transcripts to have 4sU incorporation (Duffy et al., 2019). Median gene lengths are 29337, 29403, 30040, 33335, and 32121 from all, total, chromatin-associated, 4sU-labeled, and 4sU-labeled chromatin-associated RNA, respectively. Percent spliced = spliced reads / total reads aligning to 3'SS junctions; splicing index = 2 × spliced reads / unspliced reads that align to 5' and 3' splice junctions.



sample	mean	min	max	25%	50%	75%
K562 replicate 1	813	12	7162	405	666	1063
K562 replicate 2	827	16	6972	427	679	1063
K562 replicate 3	807	11	7420	410	648	1033
S2 replicate 1a	640	15	8242	341	508	806
S2 replicate 1b	623	22	6481	338	500	787
S2 replicate 2	544	18	11264	312	418	670
S2 replicate 3	574	22	11664	322	439	709

Figure S2. Key sequencing and alignment statistics of nano-COP (related to Figure 1). (A-B) Pairwise correlation plots of read counts per gene from three nano-COP biological replicates from (A) K562 cells and (B) S2 cells. (C) Cumulative distribution plot of alignment match

percent (the percent of the read that exactly matches the reference sequence) from nano-COP and poly(A)-selected mRNA datasets. nano-COP reads have a slightly lower match percent than poly(A)-selected mRNA (decrease by 3±0.15% match percent; t-test p-value < 0.05). (D) Confusion matrix of read base calls versus reference bases for nano-COP and poly(A)-selected mRNA. (E) Cumulative distribution plot of direct RNA read lengths passing the default basecalling threshold from nano-COP and poly(A)-selected mRNA datasets. nano-COP reads have a slightly shorter read length distribution than poly(A)-selected mRNA (median decrease by 130±35 nt; t-test p-value < 0.05). The direct RNA sequencing datasets of poly(A)-selected mRNA in C-E were generated as part of the Oxford Nanopore RNA consortium and originated from labs at the University of California, Santa Cruz (UCSC) and Johns Hopkins (Hopkins) (Workman et al., 2018). (F) Descriptions of ONT read lengths for all reads that pass the default basecalling threshold from human K562 and *Drosophila* S2 nano-COP datasets. Replicates 1a and 1b from *Drosophila* S2 cells were prepared as technical replicates. (G) Histogram of nano-COP read counts per gene from human K562 cells. (H) Cumulative fraction of total reads that align to individual genes from nano-COP in human K562 cells. (I) Correlation plot of K562 4sU labeled chromatin-associated RNA sequenced on an Illumina instrument and Oxford Nanopore Technologies (ONT) MinION as read coverage per gene.

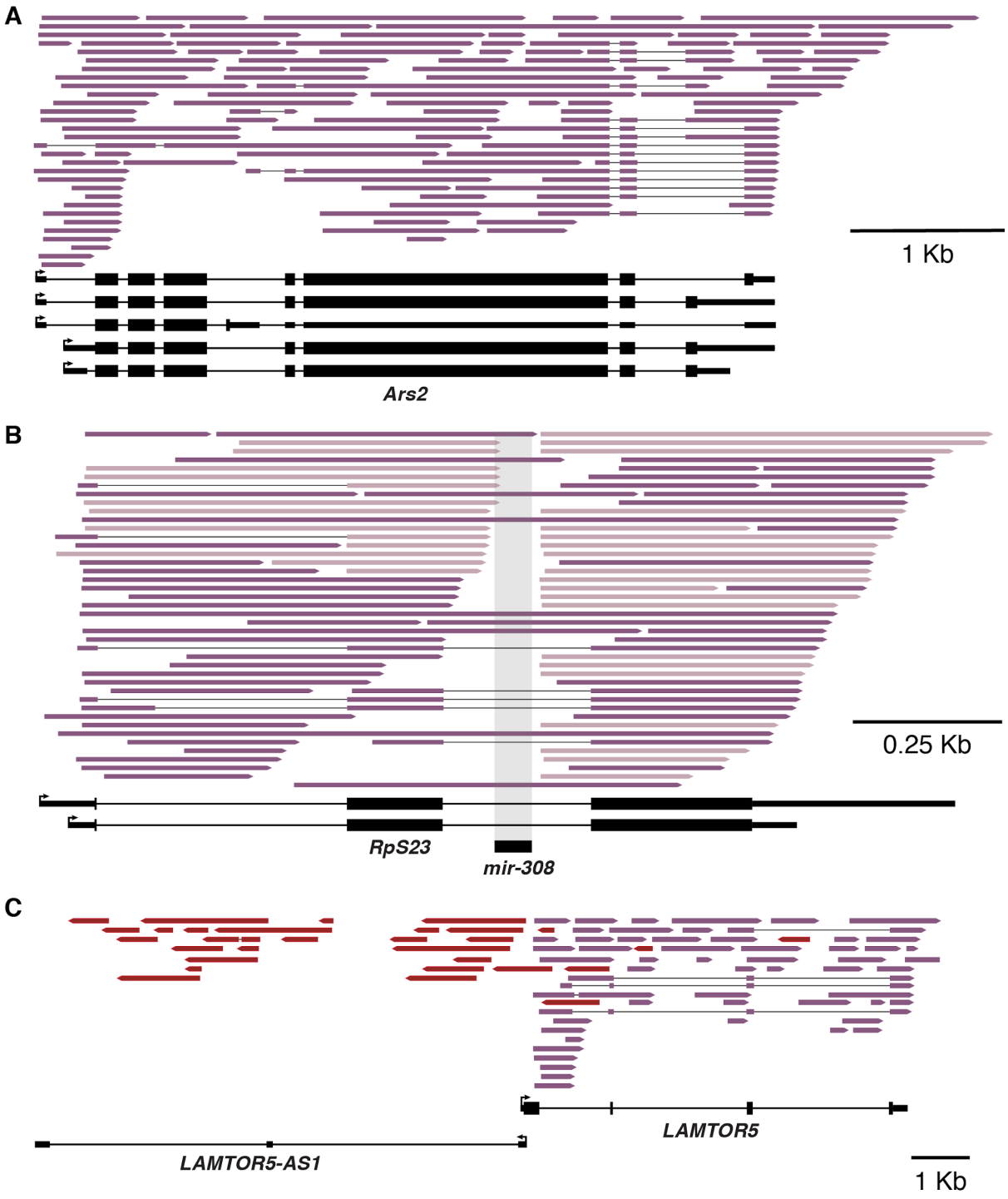


Figure S3. Example nano-COP reads reveal detection of alternative splicing, miRNA processing, and noncoding transcription (related to Figure 1). (A) Reads aligning to the *Ars2* gene in *Drosophila* S2 cells show alternative nascent isoforms in fruit flies. (B) Reads aligning to the *RpS23* gene in *Drosophila* S2 cells display processing around the *mir-308* pre-miRNA. A pileup of reads (designated in light purple) that end and start near the boundary of the *mir-308* pre-miRNA (designated by the light gray box) suggests a mechanism where pre-miRNAs are

excised from intron segments before splicing completes. (C) Reads aligning to the *LAMTOR5* gene and *LAMTOR5-AS1* lncRNA in human K562 cells display nascent noncoding RNAs stemming from antisense transcription. Boxes represent read coverage and black lines represent skipped coverage due to splicing. Arrows on boxes represent the direction of transcription; purple boxes are reads transcribed in the sense direction of the coding gene; red boxes are reads transcribed in the antisense direction of the coding gene.

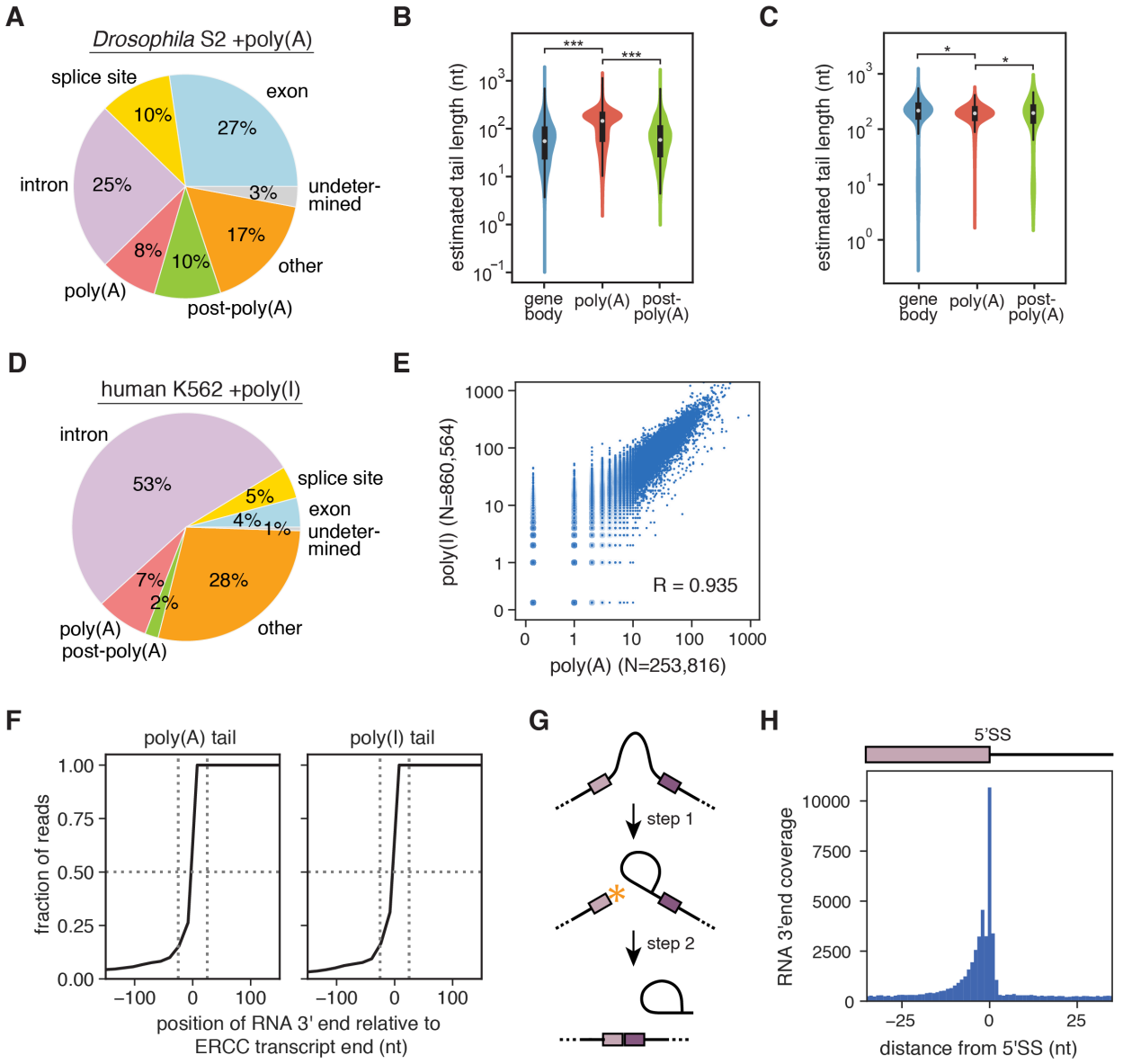


Figure S4. Alignment of nascent transcript 3' ends in nano-COP data (related to Figure 2). (A) Distribution of nano-COP 3' ends by nanopore sequencing in *Drosophila* S2 cells with enzymatic poly(A) tail addition. See Methods for descriptions of 3' end alignment categories. (B-C) The length of poly(A) tails for sequenced RNAs from (B) *Drosophila* S2 cells with enzymatic poly(A) tail addition and (C) human K562 cells with no enzymatic poly(A) tail addition. Tail lengths were estimated using nanopolish-polyA (Loman et al., 2015; Workman et al., 2018) and plotted for RNAs in each sample that have 3' ends aligning within gene bodies (exon, intron, or splice site), at poly(A) sites, or just downstream of poly(A) sites (*** signifies t-test p-value < 1×10^{-30}; * signifies t-test p-value < 1×10^{-5}). (D) Distribution of nascent RNA 3' ends from nano-COP data from human K562 cells with enzymatic poly(I) tail addition. (E) Pairwise correlation plot of nano-COP data generated with poly(A) versus poly(I) tailing strategies (Pearson R = 0.935) (F) Cumulative distribution plot of transcript 3' end positions from direct RNA nanopore sequencing of the *in vitro* transcribed ERCC-00048 transcript with

enzymatic poly(A) or poly(I) tail addition. Vertical dotted lines are at -25 and 25 nt from annotated poly(A) site; horizontal dotted line is at 0.50. (G) Cartoon demonstrating the catalytic steps of RNA splicing. The asterisk (*) represents the available RNA 3' end at the intron 5'SS between the two catalytic steps of the splicing reaction. (H) The frequency of RNA ends (the 3' end of the RNA when aligned) around intronic 5'SS's represented as a histogram in human K562 cells. High coverage of nascent RNA 3' ends at 5'SS's likely results from free exon ends between the first and second catalytic steps of the splicing reaction.

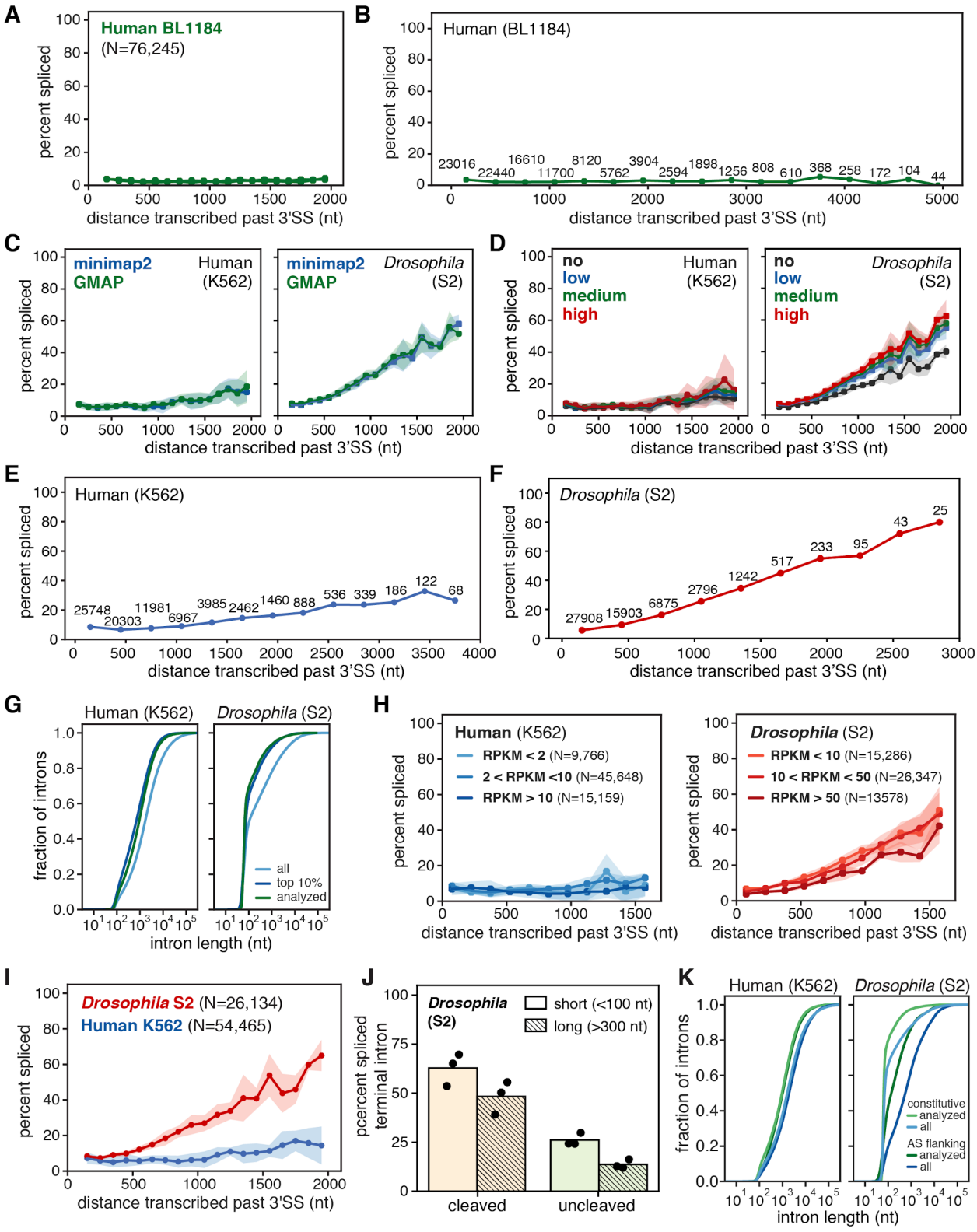


Figure S5. Extensions of plots that reveal the distance transcribed before splicing (related to Figure 2). (A-B) Global analysis of the distance transcribed before splicing in human B lymphoblast cells (BL1184) across (A) two biological replicates up to two kb and (B) combining

replicates out to 5 kb. (C) Global analysis of transcribed distance from 3'SS and the percent of spliced molecules with two different alignment programs, GMAP (Wu and Watanabe, 2005) and minimap2 (Li, 2018) (human K562 two-way ANOVA p-value = 0.572 and *Drosophila* S2 two-way ANOVA = 0.715). (D) The distance transcribed before splicing with varying stringency levels of included introns (human K562 two-way ANOVA p-value = 0.005, *Drosophila* S2 two-way ANOVA p-value < 10^{-15}). "No" stringency signifies that all introns were included. "Low", "medium", and "high" stringency levels include varying numbers of introns in the analysis based on intron retention levels from total RNA-seq data (see the Methods section for stringency descriptions). (E, F) Combining all biological and technical replicates to plot the global distance transcribed by percent spliced up to (E) 4 kb in human K562 cells and (F) 3 kb in *Drosophila* S2 cells. Numbers above points represent the total counts used to calculate the percent spliced measurement at each position. (G) Distribution of intron lengths that were utilized in the distance transcribed before splicing plot in Figure 2B. Analyzed introns are compared with all introns (all) and introns in top 10% expressed genes by read count per gene length (top 10%, which are covered by ~30% of the nano-COP dataset) in each respective species. (H) Comparing the distance transcribed before splicing for introns from transcripts with varying gene expression levels as measured by total RNA RPKM from Illumina short read sequencing in human K562 (left) and *Drosophila* S2 (right) cells (Pai et al., 2017) (human K562 two-way ANOVA p-value < 10^{-4} and *Drosophila* S2 two-way ANOVA p-value < 10^{-14}). (I) The distance transcribed before splicing middle introns within genes after computationally removing first and last introns from the analysis (two-way ANOVA p-value < 10^{-15}). (J) The proportion of reads from *Drosophila* libraries with spliced terminal introns and 3' ends in the "cleaved" bin (between 50 nt upstream and 50 nt downstream of poly(A) sites) and "uncleaved" bin (between 50-550 nt downstream of poly(A) sites). Bars are separated by intron length such that short introns (<100 nt) are represented as solid boxes while long introns (>300 nt) are represented as hashed boxes. Black points distinguish the results from three biological replicates. (K) Distribution of intron lengths that were utilized in the distance transcribed before splicing plots in Figure 2E. Analyzed introns neighboring alternative ("AS flanking") versus constitutive exons are compared with all introns in the two categories.

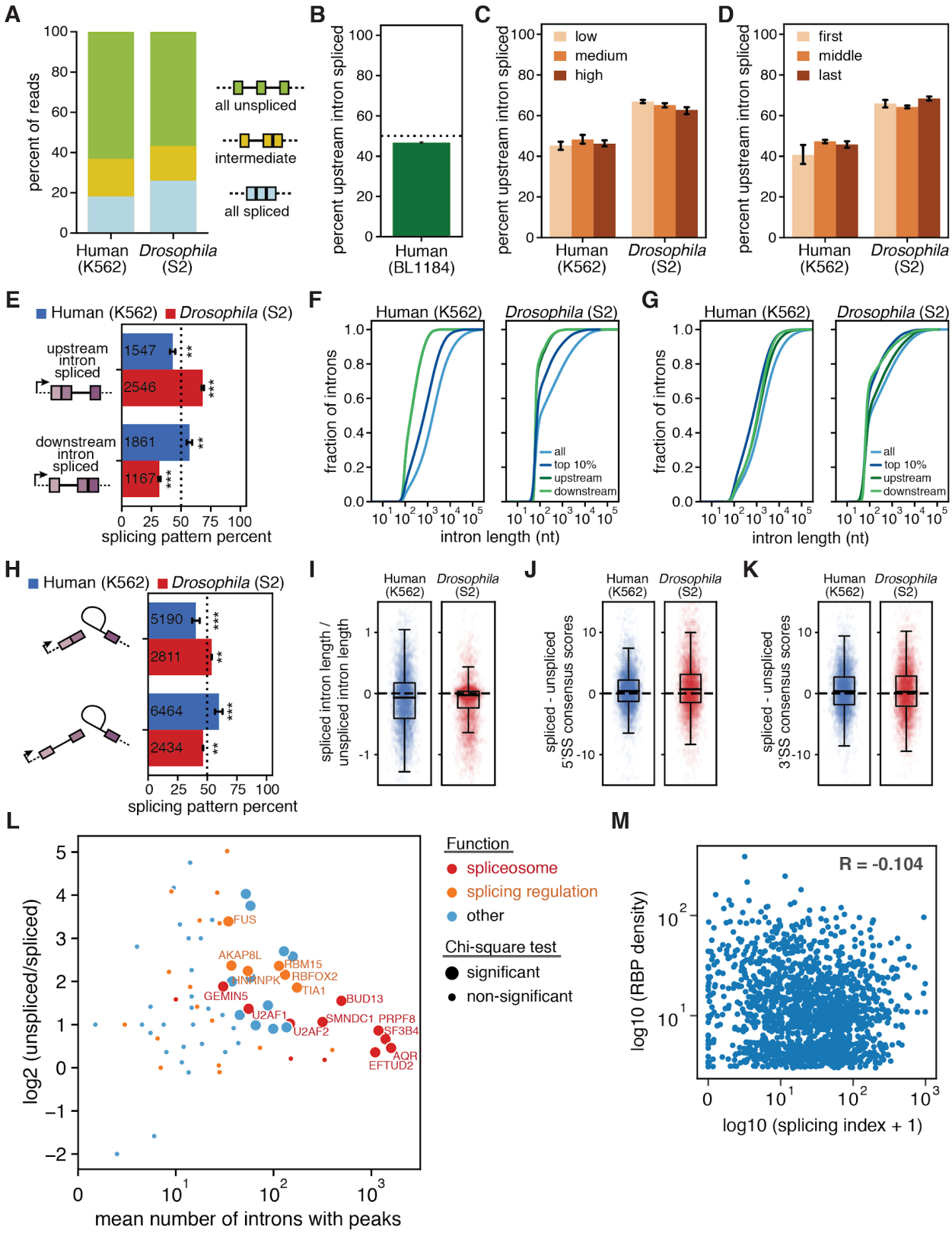
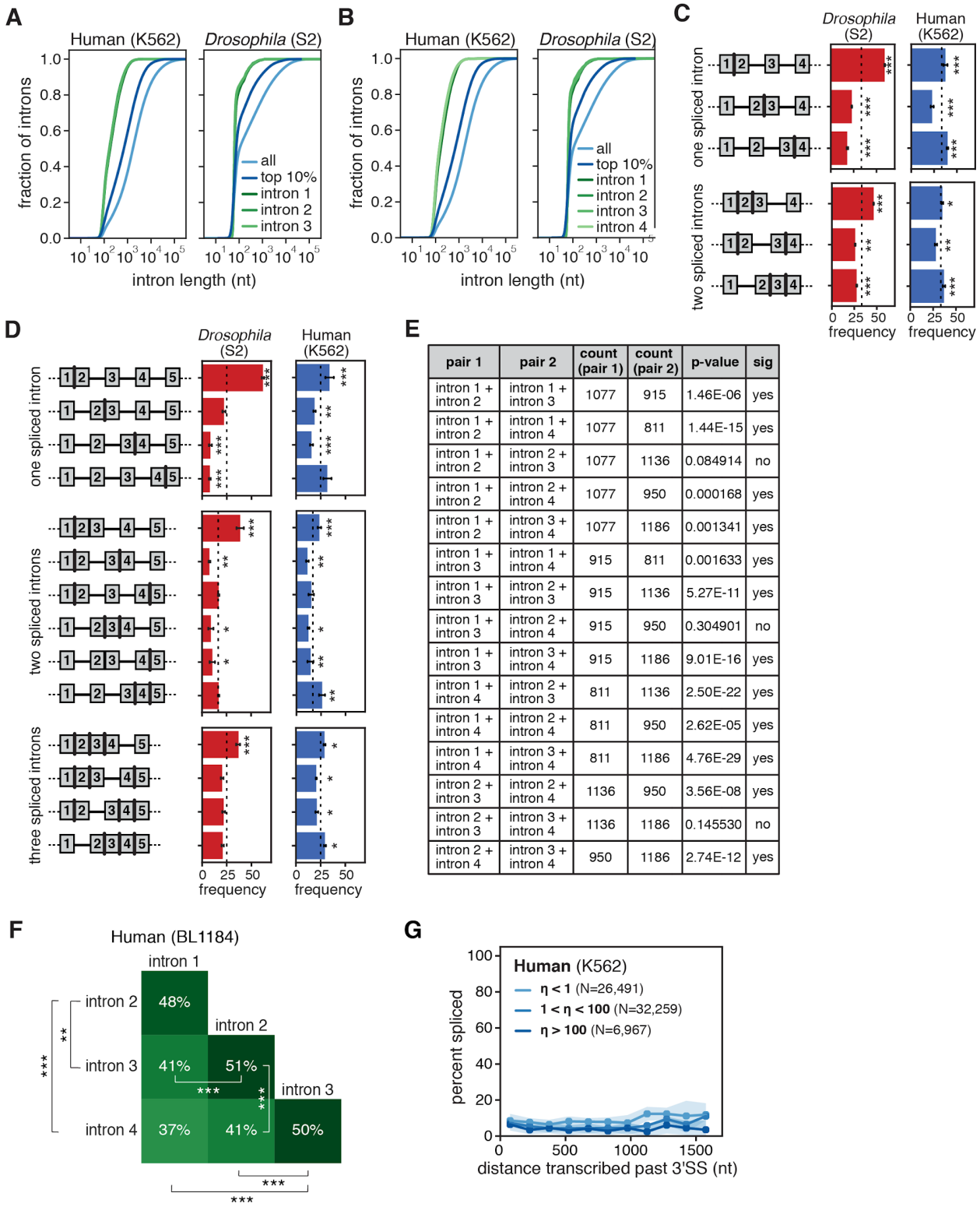

A
percent of reads
all unspliced
intermediate
all spliced
Human (K562)
Drosophila (S2)
B
percent upstream intron spliced
Human (BL1184)
C
low
medium
high
D
first
middle
last
E
Human (K562)
Drosophila (S2)
upstream intron spliced
downstream intron spliced
1547
2546
1861
1167
splicing pattern percent
F
fraction of introns
intron length (nt)
all
top 10%
upstream
downstream
G
H
5190
2811
6464
2434
I
spliced intron length / unspliced intron length
J
spliced - unspliced 5'SS consensus scores
K
spliced - unspliced 3'SS consensus scores
L
log2 (unspliced/spliced)
mean number of introns with peaks
FUS
AKAP8L
RBM15
HNRNPK
RBFOX2
TIA1
GEMIN5
BUD13
U2AF1
SMNDC1
PRPF8
U2AF2
SF3B4
AQR
EFTUD2
Function
spliceosome
splicing regulation
other
Chi-square test
significant
non-significant
M
R = -0.104
log10 (RBP density)
log10 (splicing index + 1)

Figure S6. The features and regulators of splicing order (related to Figures 4 and 5). (A) Distribution of splicing patterns in nano-COP reads spanning at least two introns. “all unspliced” represents reads where every intron is present and therefore not spliced; “intermediate” represents reads where at least one intron is spliced and one intron is not spliced; “all spliced” represents reads where every intron within the read is spliced. (B) The proportion of reads that span two or more introns where the upstream intron within the pair is spliced first in human BL1184 cells (N=2973). Black bar represents the range across two biological replicates. (C) Bar plots show changes in the order of splicing across intron pairs from genes with different gene expression patterns in total RNA. “low” includes intron pairs from genes with RPKM < 2 in human K562 cells and < 10 in *Drosophila* S2 cells; “medium” includes intron pairs from genes with RPKM between 2 to 10 in human K562 cells and between 10 to 50 in *Drosophila* S2 cells; “high” includes intron pairs from genes with RPKM > 10 in human K562 cells and > 50 in *Drosophila* S2 cells. (D) Bar plots to show changes in the order of splicing across intron pairs by position within the gene. “First” includes intron pairs where the upstream intron is the first intron within the gene; “middle” includes all intron pairs that do not contain a first or last intron within the gene; “last” includes all intron pairs where the downstream intron is the last intron within the gene. Error bars represent standard error of the mean across three biological replicates. (E) The frequency that the upstream intron is spliced first in reads that spain pairs of introns where both introns are constitutively spliced in total RNA (as defined by meeting the medium stringency splicing criteria described in the Materials & Methods section). A binomial test was used to test whether splicing order percentages differ from random expectations (50%). ** indicates binomial test p-value < 8×10^{-8}; *** indicates binomial test p-value < 2×10^{-115}. (F) Cumulative distribution of intron lengths between all introns (blue), introns from top 10% expressed genes (dark blue), and only the upstream and downstream introns (green) analyzed within pairs. (G) Cumulative distribution of intron lengths between all introns (blue) and the upstream and downstream introns from RNAs that end at an intron 5’SS and are likely derived from splicing intermediates. The downstream intron is assumed to be in the process of splicing (between the first and second catalytic steps) while the upstream intron is either spliced or not spliced. (H) The frequency at which the upstream intron is spliced within reads deriving from splicing intermediates. A binomial test was used to assess whether splicing order percentages differ from random expectations (50%). ** indicates binomial test p-value < 3×10^{-7}; *** indicates binomial test p-value < 4×10^{-32}. (I) The percent change in intron length between first and second spliced introns within pairs. Human first spliced introns are shorter 56.8% of the time; *Drosophila* first spliced introns are shorter 61.1% of the time. (J, K) The difference in (J) 5’SS consensus scores and (K) 3’SS consensus scores between spliced and unspliced introns within pairs. In human cells, the spliced introns within pairs have higher 5’SS scores in 56.0% of cases and 3’SS scores in 54.5% of cases. In *Drosophila* cells, the spliced introns within pairs have higher 5’SS scores in 58.2% of cases and 3’SS scores in 53.0% of cases. (L) MA plot comparing individual RBP binding in spliced and unspliced introns. Each circle represents an RBP. The x-axis represents the mean number of spliced and unspliced introns with peak(s) while the y-axis represents the log2 ratio of the number of unspliced introns with peak(s) and spliced introns with peak(s). Larger circles show RBPs that bind more than five spliced and unspliced introns and that display statistically significant differences in a chi-square test comparing number of spliced and unspliced introns (Bonferroni-corrected p-value threshold = 0.000658) and in a Wilcoxon rank-sum test comparing RBP density in spliced and unspliced introns (Bonferroni-corrected p-value threshold = 0.000495). RBPs are colored based on their annotated functions in Table S5 from

(Van Nostrand et al., 2018); functions that do not relate to “spliceosome” or “splicing regulation” are labeled as “other”. (M) Correlation plot between the intron splicing index (2 × spliced reads / unspliced reads) from Illumina total RNA-seq and RBP density across the intron for introns used in the splicing order analysis with RBP density > 0.



pair 1	pair 2	count (pair 1)	count (pair 2)	p-value	sig
intron 1 + intron 2	intron 1 + intron 3	1077	915	1.46E-06	yes
intron 1 + intron 2	intron 1 + intron 4	1077	811	1.44E-15	yes
intron 1 + intron 2	intron 2 + intron 3	1077	1136	0.084914	no
intron 1 + intron 2	intron 2 + intron 4	1077	950	0.000168	yes
intron 1 + intron 2	intron 3 + intron 4	1077	1186	0.001341	yes
intron 1 + intron 3	intron 1 + intron 4	915	811	0.001633	yes
intron 1 + intron 3	intron 2 + intron 3	915	1136	5.27E-11	yes
intron 1 + intron 3	intron 2 + intron 4	915	950	0.304901	no
intron 1 + intron 3	intron 3 + intron 4	915	1186	9.01E-16	yes
intron 1 + intron 4	intron 2 + intron 3	811	1136	2.50E-22	yes
intron 1 + intron 4	intron 2 + intron 4	811	950	2.62E-05	yes
intron 1 + intron 4	intron 3 + intron 4	811	1186	4.76E-29	yes
intron 2 + intron 3	intron 2 + intron 4	1136	950	3.56E-08	yes
intron 2 + intron 3	intron 3 + intron 4	1136	1186	0.145530	no
intron 2 + intron 4	intron 3 + intron 4	950	1186	2.74E-12	yes

Figure S7. Characteristics of coordinated splicing across neighboring introns (related to Figure 6). (A–B) The size distribution of introns contributing to (A) intron triplet and (B) intron quadruplet plots compared with all introns and introns from top 10% of expressed genes in human K562 and *Drosophila* S2 cells. (C–D) The proportion of reads in *Drosophila* S2 cells

(red) and human K562 cells (blue) that exhibit the same splicing pattern as the schematic to the left across (C) intron triplets and (D) intron quadruplets when first and last introns within genes are not included in the analysis. Error bars represent standard error of the mean across biological replicates. A binomial test was used to assess whether percentages differ from random expectations (dotted lines). * signifies binomial test p-value < 0.05; ** signifies binomial test p-value < 0.01; *** signifies binomial test p-value < 10^{-5} (E) Pairwise comparisons between all compartments of the matrix in Figure 6C. Counts represent the number of reads spanning four or more introns with the same splicing status (both spliced or both unspliced) for the respective intron pairs. For each pair of intron pairs, a chi-square test was used to test the statistical significance of one pair having the same splicing status more frequently than the second pair. The column "sig" indicates whether the difference is statistically significant after multiple testing correction (Bonferroni-corrected p-value threshold = 0.003). (F) Heatmap representing the frequency that two introns within a read that spans at least four introns have the same splicing status (both spliced or both not spliced) in human BL1184 cells. Intron number represents the position of each intron within a read that spans at least four introns such that "intron 1" is the first intron and "intron 4" is the last intron within the set of four introns that a read spans (** signifies chi-square test p-value < 0.003, *** signifies chi-square test p-value < 10^{-5}, Bonferroni-corrected p-value threshold = 0.003). (G) Comparing the distance transcribed before splicing introns in human K562 cells with varying "splicing yield" values (η) that represent the proportion of pre-mRNA introns that get converted into spliced mRNA as measured by (Wachutka et al., 2019) (two-way ANOVA p-value < 10^{-8}).

Table S2. Characteristics of random forest models for splicing order (related to Figure 5). Importance percentages for each intron feature represent the weight of each feature within the prediction model. For each feature, "1" represents the upstream intron (first transcribed) and "2" represents the downstream intron (second transcribed). The prediction accuracy for each random forest classifier from *Drosophila* and human intron pairs represents the percent of time the prediction is correct. Baseline represents the accuracy of predicting splicing order without any information from the feature inputs.

	Importance	
Feature	***Drosophila***	**Human**
intron length 1	14.0%	10.9%
intron length 2	11.2%	11.5%
upstream exon length	8.5%	8.2%
middle exon length	10.4%	8.6%
downstream exon length	8.2%	8.2%
5'SS MaxEnt score 1	9.9%	9.7%
5'SS MaxEnt score 2	9.5%	9.6%
3'SS MaxEnt score 1	8.3%	9.2%
3'SS MaxEnt score 2	8.0%	9.0%
position 1	3.8%	5.7%
position 2	4.0%	5.7%
alternative 1	2.4%	1.8%
alternative 2	2.0%	1.9%

	prediction accuracy	
Feature	Drosophila	Human
intron length	69%	56%
exon length	59%	55%
sequence	62%	55%
positional	62%	49%
combined	74%	60%
Baseline	62%	51%

Table S3. Comparison of RBP binding in pairs of spliced and unspliced introns within the same read (related to Figure 5). RBPs that are significantly more likely to be bound to unspliced introns than spliced introns within pairs are shown. RBP binding was compared in two ways: 1) using a chi-square test comparing the number of spliced and unspliced introns bound by an RBP. Only RBPs with at least one bound spliced intron and one bound unspliced intron were considered (n=76); 2) using a Wilcoxon rank-sum test comparing RBP density normalized by intron length in spliced and unspliced introns (normalized by intron length) within the same pair. Only RBPs with binding in at least one intron of the pair in at least two pairs were considered (n=101). Multiple testing correction was performed using the Bonferroni method using alpha=0.05. The Bonferroni-corrected p-value thresholds for significance were 0.000658 (chi-square test) and 0.000495 (Wilcoxon rank-sum test). The annotated function was obtained from Table S5 of (Van Nostrand et al., 2018) and annotations that do not relate to "spliceosome" or "splicing regulation" are labeled as "other".

RBP	number of spliced introns w/ peak(s)	number of unspliced introns w/ peak(s)	chi-square test p-value	mean density / intron length spliced introns	mean density / intron length unspliced introns	rank-sum test p-value	annotated function
BUD13	251	736	3.61E-58	0.008548508	0.020045879	3.81E-76	spliceosome
PRPF8	839	1524	1.66E-55	0.015095547	0.021645688	1.11E-63	spliceosome
DDX24	45	269	3.04E-37	0.003291424	0.020524135	8.85E-54	other
UCHL5	34	221	5.38E-32	0.00278193	0.016504893	7.70E-47	other
SF3B4	1081	1710	6.79E-42	0.02068125	0.026350117	3.06E-37	spliceosome
RBFOX2	48	214	6.59E-25	0.003842254	0.017749303	4.02E-37	splicing_regulation
RBM15	37	190	2.31E-24	0.005473396	0.021331966	6.54E-33	splicing_regulation
TIA1	75	272	1.31E-26	0.007602917	0.015705973	3.19E-32	splicing_regulation
AQR	1340	1842	4.04E-25	0.017805771	0.022669265	6.77E-30	spliceosome
ZNF622	6	98	3.19E-19	0.001366442	0.019304699	2.05E-28	other
GRWD1	8	108	2.55E-20	0.002530093	0.021980754	5.91E-28	other
SMNDC1	205	428	1.25E-19	0.014404601	0.02223641	4.82E-21	spliceosome
XRCC6	22	94	3.49E-11	0.002099768	0.010355444	2.69E-18	other
FUS	6	63	1.37E-11	0.000748567	0.008690731	7.99E-17	splicing_regulation
HNRNPK	19	90	1.64E-11	0.003536965	0.010888288	1.36E-14	splicing_regulation
AKAP8L	12	62	1.11E-08	0.001775097	0.015196145	3.29E-13	splicing_regulation
NIPBL	47	128	1.12E-09	0.005944474	0.013855176	8.32E-13	other
GTF2F1	15	60	3.46E-07	0.002858943	0.012792316	1.50E-11	other
U2AF2	97	197	5.05E-09	0.010035114	0.018562803	1.36E-09	spliceosome
PHF6	30	78	5.55E-06	0.009182859	0.021527962	6.22E-09	other
GEMIN5	13	48	1.28E-05	0.003424917	0.012838977	2.93E-08	spliceosome
XRN2	93	178	2.46E-07	0.014877535	0.018720319	7.79E-08	other
EFTUD2	960	1230	2.09E-10	0.023178007	0.026148912	1.86E-06	spliceosome

DROSHA	27	63	0.00021291	0.006117862	0.015816022	2.42E-06	other
AATF	69	129	2.36E-05	0.017330203	0.026003891	3.06E-06	other
ZC3H8	44	87	0.00022454	0.011394434	0.016989358	2.49E-05	other
U2AF1	31	80	4.71E-06	0.007087029	0.011201301	7.35E-05	spliceosome

Table S4. Primers and oligonucleotides used in this study (related to STAR Methods). A list of all primer and oligo sequences that were utilized in this study.

Name	Sequence	Species
BRD2_F	CAAAATTATAAAACAGCCTATGGACATG	Human
BRD2_R	TTTTCCAGCGTTTGTGCCATTAGGA	Human
Set_F	AAATTGATGCCTGCCAGAAC	Drosophila
Set_R	GGATTCTCGTCGAAATGGAA	Drosophila
PYCR2_e5e6_F	GGCTTTGCTGGGAGCTGCCAAGATG	Human
PYCR2_i5e6_F	TTCCCCATAGGGAGCTGCCAAGAT	Human
PYCR2_e6_R	TGATGAGCAGAGAGCGGAAG	Human
DYNC1H1_e61e62_F	CAAAGGACCTCTTCCAGGTG	Human
DYNC1H1_i61e62_F	TATGGAACAACATCGTCTCCTG	Human
DYNC1H1_e62_R	TGCCCTTCAGTTTGATTCTTG	Human
ONT_oligoA	/5PHOS/GGCTTCTTCTTGCTCTTAGGTAGTAGGTTC	-
ONT_oligoC_C10	GAGGCGAGCGGTCAATTTTCCTAAGAGCAAGAAGAAGCCCCCCCCCCCC	-
T7_ERCC-00048_Gblock	TAATACGACTCACTATAGGGAGATCTGTAAATCCCGTAAACGAGTAGTACGAATCCGGACTTGAATACACGCGTCAATCCCTTTTTATATCCTAGAATGGACCGTGTGGACGGCAACTCAGAGATAACGCATATCTATGTGCTCGCTTGCCCATCAAAGAAGAGACGGCGACCAAACGGACGACATATAGTGACATGGTCAACCCGTACGCCTGCTTCGTAAGCCGACGGTCCTTTGAAGAGGCTGGCGAATCATGTCGTTTGTGCTTACTATTACATGCTAGCTTGGTTGGGGCATCTCGGGACAACGTCTATGTACAATAAACACAAAGCCGCGTAGTTATCTTCCGCGAGTTCCGCCCAATACATTGGCGGTGACTTGAGACCGCTAAAATGCACATAGAAGCCTCAAACATGGTAAGACTATAGATAAGCGGCGCGAAAACACGGCATTTGGAATGATGTGTACTGGGAATAAGACGACGTCGCTATGGCCTCTCCGGAAGGCGGTGTATGTGCCAAGCGATGTTTCATTAATGTAACGGACAGGTCGCTGAGGTGGCTTTCGTTGGGGGCGCCGTCTTTGGGGGAGATTGCGTCAATTTTGACTGTCAGATCAGCGACTAGATTTTAGGCAGATTAGTGTGCCACCTGAATCAATAGAACAATATCAGTTATGGCGGTGCAGTATACTATACAATGGGTTGGGCGCATCTGCATGTCTCATGCTGTCATGGCAATCGACCTCTAGTCTGGGGTGATCCGAGGCGCTTCTCTTATTAGGAATAGTGCAGGACCCGAAACCGCCATAGGGAAAGGGTGAGCGAGGTAGCAGCGTAATAATTCGCGGTGGGCAGGAAATGCTTAGTGTTCTGTCTCAAGACCTAAGCGACAGCGTGACCTTGTTTCACTTACCTCTGAAGCTCTTCGACGTTATAGATATTGGCATCCCTAAACAACGAGTACCTTGTGCTACGACAGAAAAGTGACCTG	-