



Splicing dependency between *EIF4G2* introns is mediated by exon definition and relies on a downstream splicing event

Received for publication, April 18, 2026, and in revised form, June 11, 2026 Published, Papers in Press, June 27, 2026

<https://doi.org/10.1016/j.jbc.2026.113304>

Marie-Lee Pelletier^{1,2}, Joëlle Vincent^{1,2}, Téa Malo^{1,2}, Camille Séguin^{1,2}, Volodymyr Tsybulskyi^{3,4}, Egor Semenchko^{3,4}, Ibrahim Soumana Adamou^{1,2}, Irmtraud M. Meyer^{3,4,5}, and Karine Choquet^{1,2,*}

From the ¹Department of Biochemistry and Functional Genomics, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, Canada; ²Research Center on Aging, CIUSSS de l'Estrie-CHUS, Sherbrooke, Canada; ³Berlin Institute for Medical Systems Biology, Max Delbrück Centre, Berlin, Germany; ⁴Freie Universität Berlin, Department of Biology, Chemistry and Pharmacy, Institute of Chemistry and Biochemistry, Berlin, Germany; ⁵Freie Universität Berlin, Department of Mathematics and Computer Science, Institute of Computer Science, Berlin, Germany

Reviewed by members of the JBC Editorial Board. Edited by Patrick J. O'Brien

In mammalian cells, pre-mRNA splicing requires the excision of multiple introns within the same nascent transcript through the coordinated action of the spliceosome and accessory factors. We and others recently demonstrated that multi-intron splicing follows a predetermined order, and that removal of proximal introns is coordinated. Notably, specific introns that are removed later within a transcript depend on their neighboring introns for efficient excision, but the underlying mechanisms remain unknown. Here, we investigated the determinants of splicing order and coordination using a multi-intron minigene splicing reporter based on the human gene *EIF4G2*. We found that dependency between a later-excised intron and its neighbor is partially mediated through the downstream 5' splice site, consistent with the exon definition model, but is independent from intron length. Importantly, splicing coordination also relies on a prior downstream act of splicing, but not on specific *cis* regulatory elements in neighboring introns. Together, these findings improve our understanding of how distinct splicing reactions are connected and coordinated to produce mature RNA molecules.

Pre-mRNA splicing is an essential step of mammalian gene expression that is catalyzed by the spliceosome, a large complex composed of five distinct small nuclear ribonucleoproteins (snRNPs). Spliceosome assembly is a stepwise process that begins with binding of the U1 snRNP to the 5' splice site, followed by association of U2AF1 and U2AF2 to the 3' splice site and polypyrimidine tract. Next, U2 snRNP binds to the branch point sequence near the 3' splice site. The tri-snRNP, composed of U4, U5 and U6 snRNPs, is then recruited to form the pre-B complex, which transitions to the B complex after removal of U1 snRNP, followed by additional rearrangements that lead to catalytic activation (1). Direct interactions between RNA polymerase (Pol) II and U1 snRNP or U2AF1 suggest that spliceosome assembly initiates co-transcriptionally (2, 3), while intron excision can occur co-

or post-transcriptionally in mammalian cells (reviewed in (4, 5)).

Human genes contain many long introns, yet splicing mechanisms have been mostly studied using synthetic pre-mRNAs (6–9) or minigene constructs (10–16) consisting of one or two introns. However, increasing evidence indicates that splicing reactions within a pre-mRNA do not occur independently from one another. Mutation of a 5' splice site impairs excision of the upstream intron and/or causes skipping of the adjacent exon (17, 18) and *in vitro* binding of U2 snRNP to a 3' splice site is enhanced if U1 snRNPs are bound to the upstream and downstream 5' splice sites (6, 19). Splicing rates and efficiencies tend to be coordinated across introns from the same gene (20, 21), while long-read sequencing studies suggest that excision of proximal introns within single pre-mRNAs is synergistic (22, 23). Furthermore, multiple studies have shown that splicing of mammalian pre-mRNAs does not proceed sequentially from the 5' to the 3' end (10, 18, 22, 24–28). Instead, splicing occurs in a pre-determined order that is the same across transcripts from the same gene and is primarily determined by *cis* regulatory features (10, 18, 22, 27–29). Notably, local “slowpoke” introns were defined as introns that are always excised after their upstream and downstream neighbors (10). In minigene splicing reporter assays, local slowpoke introns had low excision efficiencies when tested alone, while addition of their upstream and/or downstream neighbor to the minigene construct significantly enhanced their excision (10). These findings suggest that interdependency between introns and splicing order are functionally linked. Two proposed models for this functional link are that *i*) efficient excision of local slowpoke introns is enhanced by *cis* elements in their neighboring introns or *ii*) their splicing is directly promoted by a neighboring act of splicing (10), but mechanistic evidence for each model is lacking.

Multiple 5' and 3' splice sites in the same pre-mRNA can compete for recognition by splicing factors and commitment to splicing, resulting in alternative splicing, a regulatory mechanism that plays critical roles in cellular identity and

* For correspondence: Karine Choquet, karine.choquet@usherbrooke.ca.

Splicing dependency between EIF4G2 introns

function (reviewed in (30)). Yet, how splice sites are recognized and paired within long multi-intronic pre-mRNAs is still poorly understood. The intron and exon definition models posit that splice sites are first paired across the shortest unit between them, either the intron or the exon (31, 32). In lower eukaryotes, where genes mostly consist of short introns (<250 nucleotides) separated by longer exons, intron definition is thought to prevail. By contrast, when long introns are separated by shorter exons, as is the case in most vertebrate genes, exon definition is predominant (32, 33). A recent study showed that cross-intron and cross-exon pre-B complexes are structurally and functionally equivalent, and suggested that interaction between the cross-exon pre-B complex and the U1 snRNP bound to an upstream 5' splice site triggers the handover of this splice site from U1 snRNP to U6 snRNP, yielding a cross-intron B complex that is subsequently activated to catalyze excision of the intron (Fig. S1A) (34). Alternatively, the recently proposed "U1 relay model" posits that U1 snRNP and U2AF2 are recruited to the 5' and 3' splice sites of an intron co-transcriptionally and sequentially (35). The 5' splice site of the downstream intron displaces U1 snRNP from the 5' splice site of the upstream intron, which promotes the interaction of the upstream 5' splice site with U6 snRNA and the subsequent transition from the pre-B to the B complex, while U1 snRNP is transferred to the downstream 5' splice site (Fig. S1B) (35). Both models could explain why human genetic variants in splice sites often lead to exon skipping rather than intron retention (32, 35). However, both models require that commitment to splicing proceeds sequentially, such that the downstream 5' splice site is still present when U2 snRNP is recruited to the upstream 3' splice site and/or when U1 snRNP is displaced from the upstream 5' splice site. However, this appears incompatible with the fact that splicing order does not proceed strictly in the direction of transcription, notably in the case of local slowpoke introns, some of which are dependent on prior excision of their downstream intron (10). A unifying model, which has yet to be tested, proposes that splice site commitment and splicing catalysis may be kinetically independent (2, 36). Splice site commitment would occur co-transcriptionally from 5' to 3', with recruitment of U1 and U2 snRNPs to the 5' splice site and the branch point, followed by A complex assembly on the Pol II surface as each intron is transcribed. There may then be a substantial delay until splicing catalysis, which would be functionally independent from transcription, leading to more heterogeneous intron excision timings and delayed removal of certain introns relative to their neighbors (2, 36).

In this study, we developed a multi-intron minigene reporter assay to interrogate the mechanisms underlying the splicing dependency between neighboring introns. We identified *EIF4G2* intron 18 as a local slowpoke intron that is excised after its upstream and downstream neighbors in minigene and endogenous transcripts. Mutagenesis experiments revealed that efficient intron 18 excision is strongly dependent on the initial presence of a downstream intron. This effect is partially mediated through the 5' splice site of its downstream intron, corroborating the role of exon definition

in coordinating adjacent splicing reactions. Furthermore, efficient intron 18 excisions only occurs if its downstream intron is splicing-competent. Our findings support a model by which intron 18 splice site commitment occurs before its neighbors are removed, while its excision necessitates a prior downstream act of splicing, thereby bridging the gap between exon definition, splicing order, and interdependency in intron excision.

Results

EIF4G2 intron 18 is always removed last in endogenous and minigene transcripts

To investigate the mechanisms underlying interdependency between local slowpoke introns and their neighbors, we searched for a model intron that is systematically excised after all other proximal introns. We mined our published dataset of splicing orders computed from direct chromatin-associated RNA nanopore sequencing from human lymphoblastoid cell lines (LCLs), HeLa cells, K562 cells, neuronal progenitor cells and motor neurons derived from human induced pluripotent stem cells (18, 29). We selected *EIF4G2* intron 18 (from RefSeq transcript NM_001418.3) because *i*) it was always removed last within the region spanning introns 15 to 20 across cell types (Figs. 1A and S2A, Table S1); and *ii*) we previously identified a splicing order reversal between introns 17 and 18 in LCL GM19099, which occurred together with a rare single nucleotide polymorphism (SNP, rs200245989) located three base pairs from the junction between intron 17 and exon 18 (Fig. 1B) (29). Overall, these results indicate that *EIF4G2* intron 18 is a *bona fide* local slowpoke intron whose splicing order can be modulated by a specific sequence change, thus representing an ideal model to investigate the underlying mechanisms.

Because slowpoke introns have been found to be dependent on either their upstream or downstream introns (10), we sought to study the mechanisms of *EIF4G2* splicing order in a multi-intronic context. We designed a multi-intron splicing reporter construct consisting of *EIF4G2* introns 15 to 20 ("long minigene") under the control of the cytomegalovirus (CMV) promoter (Fig. 1C). After transfection in HEK293T cells for 24 h, we extracted total RNA and performed RT-PCR with a forward minigene-specific primer and a reverse primer targeting *EIF4G2* exon 21. We obtained a single band of 1201 base pairs, corresponding to the expected spliced product, indicating that all introns were efficiently excised from the minigene (Fig. 1C). Next, we asked whether splicing order was consistent between endogenous and minigene *EIF4G2* transcripts. Since minigene transcripts are not chromatin-associated, making it more challenging to isolate newly synthesized RNA, we used a previously described PCR-based strategy (25, 26) with exon/intron primer pairs to map the removal order of intron 18 and its neighbors (Fig. 1, D and E). For both endogenous and minigene transcripts, PCR with primers in exon 17 and intron 18 yielded two bands corresponding to the unspliced product and a product in which intron 17 had been removed (Fig. 1D, left panel). By contrast,

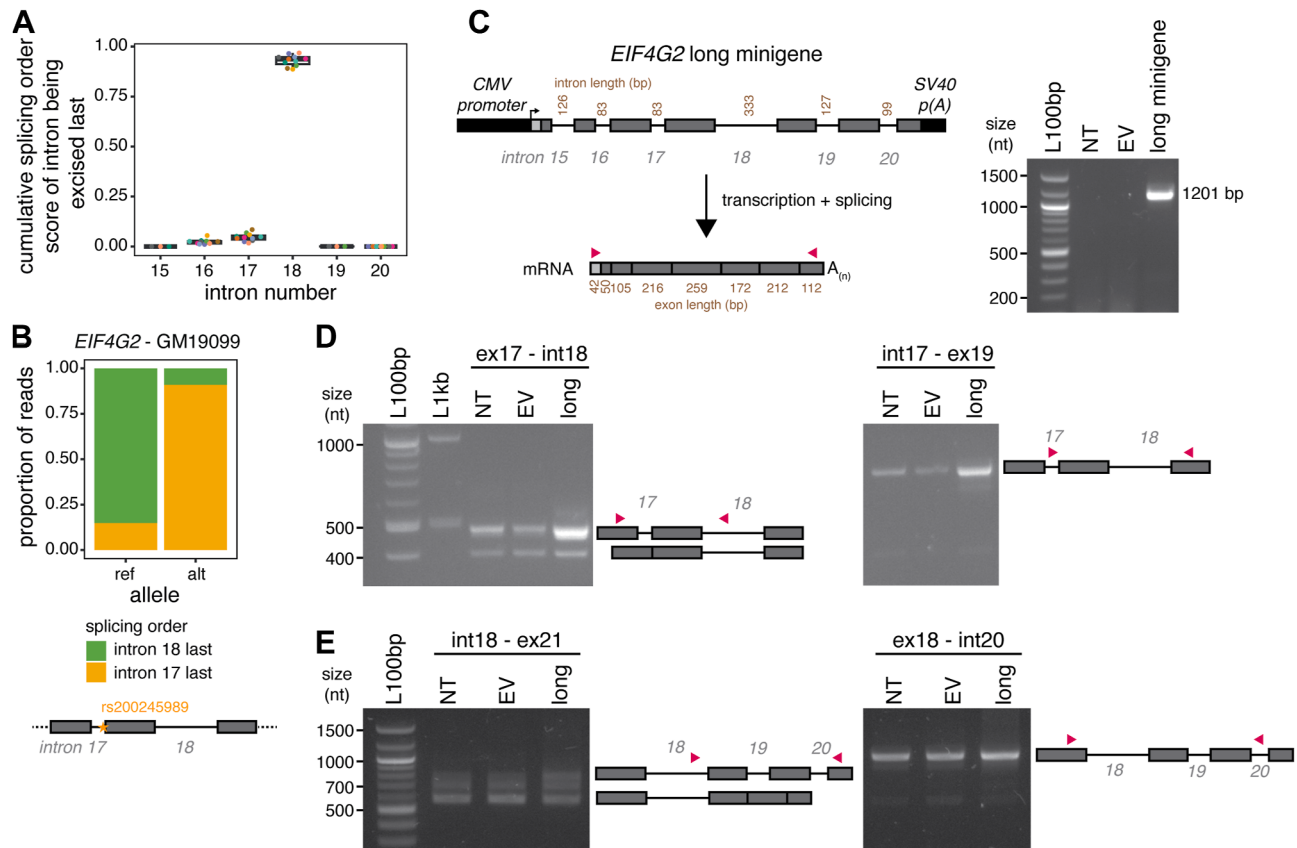


Figure 1. EIF4G2 intron 18 is a slowpoke intron. *A*, frequency of each intron being excised last within EIF4G2 introns 15 to 20, represented using the cumulative splicing order score for all splicing orders in which the given intron is excised last. Individual splicing orders and associated scores are shown in Table S1. Each colored dot represents a distinct human LCL. Intron 18 consistently has the highest score, indicating that it is almost always excised last in each LCL. *B*, allele-specific splicing order for introns 17 and 18 in LCL GM19099. “Ref” and “alt” refer to the genotypes for rs200245989. *C*, left: schematic representation of the EIF4G2 long minigene and the resulting mRNA after expression and splicing in HEK293T cells. Exons are shown as dark gray rectangles and introns as lines. The light gray rectangle represents the plasmid sequence between the cytomegalovirus (CMV) promoter and the start of the EIF4G2 sequence. The dark pink arrowheads indicate the PCR primers used for RT-PCR analysis of minigene transcripts. Right: Agarose gel electrophoresis of the minigene RT-PCR product. The forward primer is specific to the minigene. *D*, agarose gel electrophoresis of RT-PCR products for splicing order analysis of EIF4G2 introns 17 and 18 with the two primer pairs shown as dark pink arrowheads. All products were migrated on the same gel, which was split in two for easier schematic representation of the obtained products. Note that primers are not minigene-specific and amplify the endogenous gene in NT and EV samples. *E*, same as *D*, but for splicing order analysis of introns 18–20. bp, base pair; EV, empty vector; EV, empty vector; L100 bp, 100 bp ladder; L1kb, 1 kb ladder; LCL, lymphoblastoid cell line; NT, no transfection.

PCR with primers in intron 17 and exon 19 yielded only the unspliced product (Fig. 1D, right panel). These results indicate that intron 17 is removed before intron 18. Similar experiments with primers in intron 18/exon 21 and exon 18/intron 20 revealed that introns 19 and 20 are excised before intron 18 (Fig. 1E). Thus, intron 18 is removed after its upstream and downstream neighbors in minigene transcripts, demonstrating that our multi-intron splicing reporter recapitulates the splicing order observed in endogenous transcripts.

EIF4G2 intron 18 excision requires the presence of downstream intron 19

We next asked whether excision of intron 18 is dependent on its upstream and/or downstream introns, as previously reported for other local slowpoke introns. We designed new minigenes consisting of introns 15 to 18 (“short_1”), introns 18 to 20 (“short_2”) and intron 18 alone (Fig. 2A). Transfection of the single intron minigene followed by RT-PCR

resulted in two distinct bands corresponding to the spliced and unspliced transcripts (bands 3 and 4, Fig. 2B), with the latter being more pronounced, suggesting inefficient intron 18 excision. Similarly, absence of introns 19 and 20 in the short_1 minigene yielded an additional band of higher molecular weight compared to the long minigene, indicating an accumulation of incompletely spliced transcripts (band 1, Fig. 2B). Nanopore sequencing of RT-PCR amplicons revealed that most of these partially spliced transcripts showed intron 18 retention only (51–54% of reads), while a smaller proportion also displayed retention of upstream introns (9–19% of reads) (Fig. 2, C and D). To evaluate the impact of multi-intron minigene expression levels on intron 18 retention, we deleted nucleotides 1 to 441 of the CMV enhancer and promoter (37) in the short_1 and long minigenes (Fig. S2B). Despite a >4.5-fold reduction in steady-state minigene transcript levels (Fig. S2C; long minigene: 4.6-fold; short_1 minigene: 4.8-fold), intron 18 retention was consistently observed in the short_1 minigene (Fig. S2B), indicating that this

Splicing dependency between EIF4G2 introns

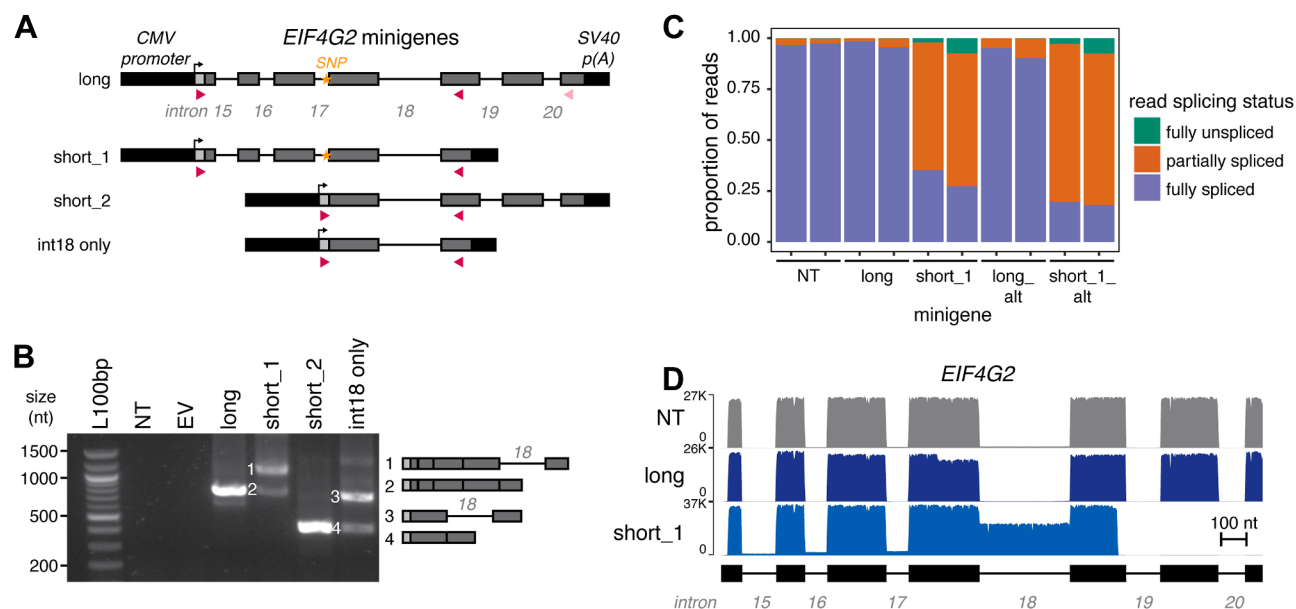


Figure 2. EIF4G2 intron 18 excision depends on its downstream introns. *A*, schematic representation of the minigenes used in the other panels. Exons are shown as dark gray rectangles and introns as lines. The light gray rectangle represents the plasmid-specific 5'-UTR. The orange star represents single nucleotide polymorphism rs200245989 present in long_alt and short_1_alt minigenes. *B*, agarose gel electrophoresis of the RT-PCR products obtained from the expression of each minigene in HEK293T cells, using the primers shown as dark pink arrowheads in *A*. Each major product is numbered and schematized on the right. *C*, nanopore sequencing of RT-PCR amplicons obtained from minigene transcripts, showing the proportion of spliced, partially spliced and unspliced reads. For the long, short_1, long_alt and short_1_alt minigenes, a forward minigene-specific primer was used with a reverse primer in exon 19 (dark pink arrowhead in *A*) or exon 21 (light pink arrowhead in *A*), respectively. RT-PCR in NT was performed with primers in exons 15 and 21. Only reads spanning the full expected amplicons were included. Two biological replicates are shown for each condition. *D*, coverage tracks from nanopore sequencing of RT-PCR amplicons performed in *C*. bp, base pair; int, intron; K, thousands; L100 bp, 100 bp ladder; NT, no transfection.

splicing defect is not due to high minigene overexpression. Together, these results show that efficient intron 18 excision requires the initial presence of its downstream introns.

Turning our attention to the upstream region, we found that absence of introns 15 to 17 did not impair the excision of intron 18 in the short_2 minigene (Fig. 2*B*). We also cloned the alternative allele of GM19099, which carries the SNP rs200245989 (chr11:10800351T>C) in the 3' splice site of intron 17, as long or short_1 minigenes (long_alt and short_1_alt, Fig. 2*A*) to determine whether delayed removal of intron 17 relative to intron 18 affected the excision of either intron. Splicing order analysis of the long_alt minigene showed increased levels of intron 17 being removed after intron 18 compared to the long minigene, as observed in GM19099 (Fig. S3*A*). Intron 18 retention was moderately increased in the short_1_alt minigene (78%) compared to the short_1 minigene (66%, Fig. S3*B*). However, the splicing order reversal did not impact splicing completion in the presence of downstream introns. Despite a slight increase in partially spliced reads due to modest intron 17 retention in the long_alt minigene (Fig. S3*C*), both long and long_alt minigenes were almost fully spliced (>90%, Fig. 2*C*), corroborating the independence between introns 17 and 18. Thus, we focused subsequent experiments on the connection between intron 18 and its downstream introns.

To determine whether intron 18 excision depends on only one or both of its downstream introns, we used site-directed mutagenesis to delete intron 19 or intron 20 from the long minigene (Fig. 3*A*). While deletion of intron 20 had no

detrimental impact on splicing (Fig. 3, *B* and *C*), deletion of intron 19 led to skipping of exons 19 and 20 (Fig. 3*C*). Consequently, although RT-PCR with a reverse primer in exon 19 detected partially spliced minigenes (Fig. 3*B*), these likely correspond to splicing intermediates that still contain exon 19, rather than intron retention. Intron 19 deletion results in a large composite exon (exons 19–20 together) that is > 400 nt and previous reports have shown that such expanded internal exons are often skipped in vertebrates (38). Thus, we cannot distinguish whether intron 18 is alternatively excised because of the absence of intron 19 *per se* or due to the resulting large flanking exon.

To circumvent this problem, we deleted both introns 19 and 20 (Δ int19+20), resulting in the absence of internal exons downstream of intron 18. Minigene Δ int19+20 behaved similarly to Δ int19 alone, with retention of intron 18 when using a reverse primer in exon 19 (Fig. 3*B*). Curiously, when using a reverse primer in exon 21, the Δ int19+20 minigene yielded a band of lower molecular weight compared to the long minigene (band #3, Fig. 3*C*), despite the absence of a canonical splice site downstream of intron 18 in the Δ int19+20 minigene. To identify all minigene transcripts without potential biases from the selection of PCR primers, we performed nanopore direct RNA sequencing (DRS) on the long and Δ int19+20 minigenes. In the Δ int19+20 minigene, 57% of reads displayed intron 18 retention (Fig. 3*D*), indicating that PCR with the reverse primer in exon 21 significantly underestimated intron 18 retention, likely due to preferential amplification of a shorter cryptic splicing

Splicing dependency between EIF4G2 introns

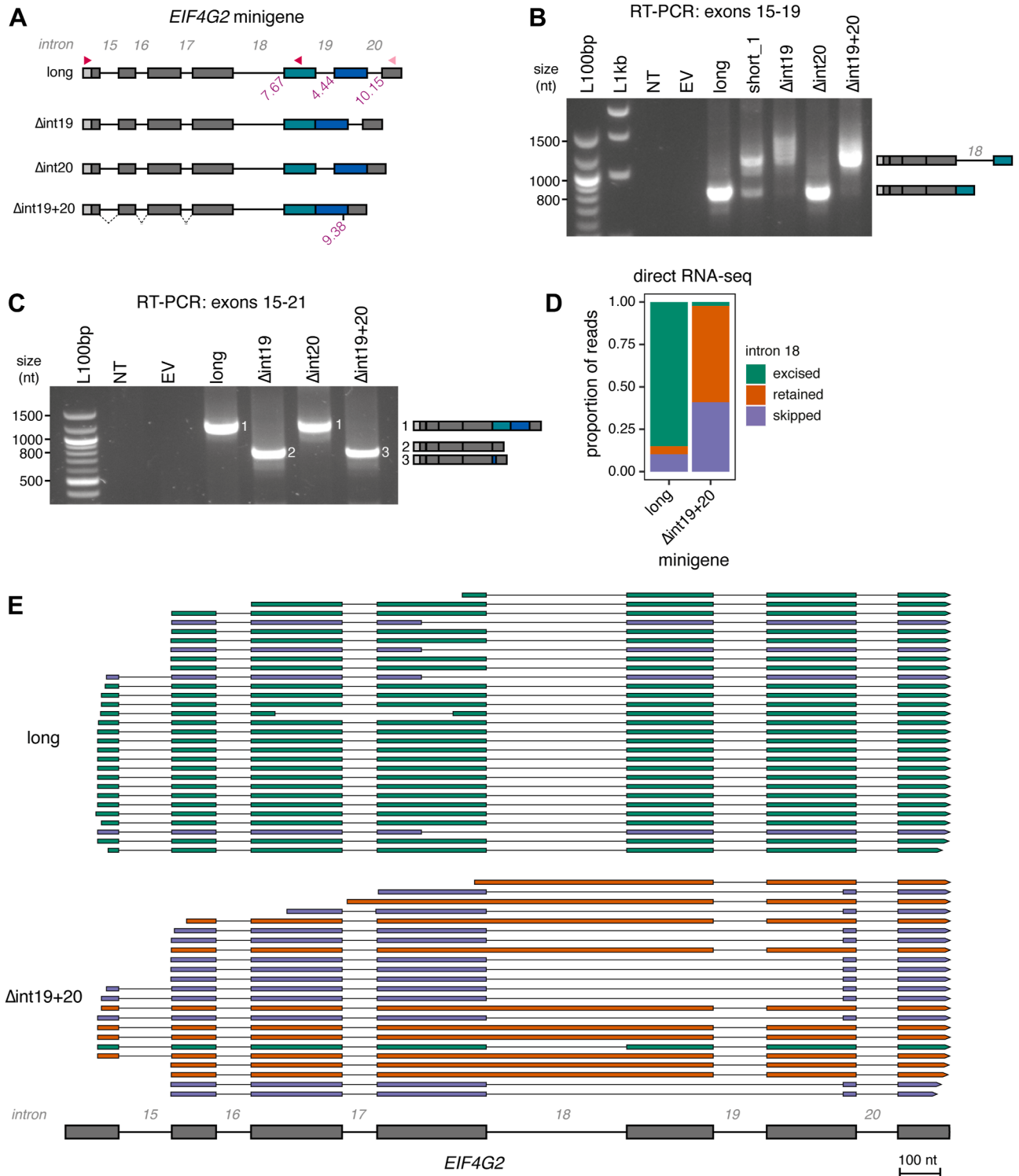


Figure 3. *EIF4G2* intron 18 excision depends on intron 19. *A*, schematic representation of the minigenes used in the other panels, with complete deletion of introns 19 and/or 20. Exons are shown as dark gray rectangles and introns as lines. The light gray rectangle represents the plasmid-specific 5'-UTR. The MaxEnt (68) 3' splice site scores for the canonical 3' splice sites of introns 18 to 20 (long minigene) and the cryptic splice site in exon 20 (Δint19+20 minigene) are shown in purple font. *B*, agarose gel electrophoresis of RT-PCR products obtained from the expression of each minigene in HEK293T cells, using the primers shown as dark pink arrowheads in *A*. Products are schematized on the right. *C*, same as *B*, but using the same minigene-specific forward primer and the reverse primer in exon 21 shown as light pink arrowhead in *A*. The short_1 minigene was not included because it lacks exon 21. *D*, nanopore direct RNA sequencing (DRS) of minigene transcripts, showing the proportion of reads in which intron 18 is excised, retained or its 3' splice site is skipped. Only reads corresponding to the complete minigene transcripts were included. *E*, representative reads from DRS of long and Δint19+20 minigenes. Each arrow represents one read, with lines indicating regions that were excised from the RNA. Displayed reads were randomly subsampled from the complete datasets and are colored according to the intron 18 splicing status as in *D*. bp, base pair; int, intron; L100 bp, 100 bp ladder; L1kb, 1 kb ladder.

Splicing dependency between EIF4G2 introns

product. Indeed, 41% of DRS reads showed exclusion of exon 19 and most of exon 20, with the use of a cryptic 3' splice site at position 181 of exon 20 (Fig. 3E). These observations indicate that the 3' splice site of intron 18 cannot be used if intron 19 is absent from the initial pre-mRNA and that the presence of its downstream exons is not sufficient to promote correct intron 18 excision.

The 5' splice site of intron 19 contributes to efficient and accurate intron 18 excision

Several mechanisms could explain the dependency of intron 18 excision on its downstream intron: 1) The 5' splice site of intron 19 may be required for exon 19 definition and appropriate recruitment of the U2 snRNP to the intron 18 branch point (Fig. 4A); 2) There may be a *cis* element in intron 19 that contributes to efficient intron 18 excision (Fig. 4B); 3) A base pairing interaction between intron 18 and a downstream region may obstruct *cis* elements in intron 18 until intron 19 is removed, leading to a rearrangement of RNA secondary structure that favors binding of splicing factors to intron 18 (Fig. 4C). Indeed, structural elements involving splice sites or the branch point have been shown to impact splicing timing and/or efficiency (39–41); and 4) As proposed by Kim *et al.* (10), the act of removing intron 19 may promote intron 18 excision (Fig. 4D).

To test the contribution of exon definition to the dependency between introns 18 and 19 (model 1, Fig. 4A), we mutated the first four nucleotides (nt) of intron 19 in the long minigene (Fig. 5, A and B). We also mutated the last 26 nucleotides of intron 19, which was not expected to affect intron 18 excision if exon definition underlies its dependency on intron 19. Mutation of the 5' splice site (mut 5'SS) yielded similar results to intron 19 deletion, with skipping of exon 19 (Fig. 5, C and D). Mutation of the intron 19 3' splice site (mut 3'SS) led to exon 20 skipping but did not affect intron 18 excision, while mutation of both splice sites simultaneously

caused skipping of exons 19 and 20 (Fig. 5, C and D). These observations are consistent with the exon definition model, whereby the 5' splice site mutation prevents proper exon 19 definition. To validate the requirement of the downstream 5' splice site for intron 18 excision, we added 72 nt of the 5' portion of intron 19 to the short_1 minigene (short + 72 nt, Fig. 5A). Presence of this segment partially rescued intron 18 retention, but the effect was lost if the added 5' splice site was mutated (short + 72 nt + mut 5'SS), as assessed by both RT-PCR and DRS (Fig. 5, E and F), supporting the model by which exon 19 definition is necessary for intron 18 excision.

Next, we investigated whether shortening intron 18 (333 nt) to <250 nt would switch its excision to a purely intron definition mode and alleviate its dependency on intron 19. We introduced two deletions in intron 18 (nucleotides 50–283 and 167–283) in the long and short_1 minigenes (Fig. 6A), resulting in lengths of 100 and 217 nt. The two flanking exons measured 259 and 172 nt and were either larger or similar in size to the deletion-containing intron 18, yielding an architecture compatible with intron definition. The long minigene transcripts with intron 18 deletion showed complete splicing, indicating that the deletions themselves did not impair intron 18 excision (Fig. 6B). However, neither deletion alleviated intron 18 retention in the short minigenes (Fig. 6B). These results suggest that the coupling between intron 18 and 19 excision is independent from the length of intron 18, and that intron 18 may be removed through exon definition even when it is shorter than its flanking exons.

Because intron 18 excision was not fully rescued by the addition of a downstream 5' splice site in the short+72 nt minigene, we assessed the role of other potential intron 19 *cis* elements in promoting intron 18 removal (model 2, Fig. 4B). We randomly shuffled the internal sequence of intron 19 (positions 7–101) in the long minigene, leaving the 5' splice site, 3' splice site, and branch point intact. However, three distinct shuffling iterations had no impact on intron 18

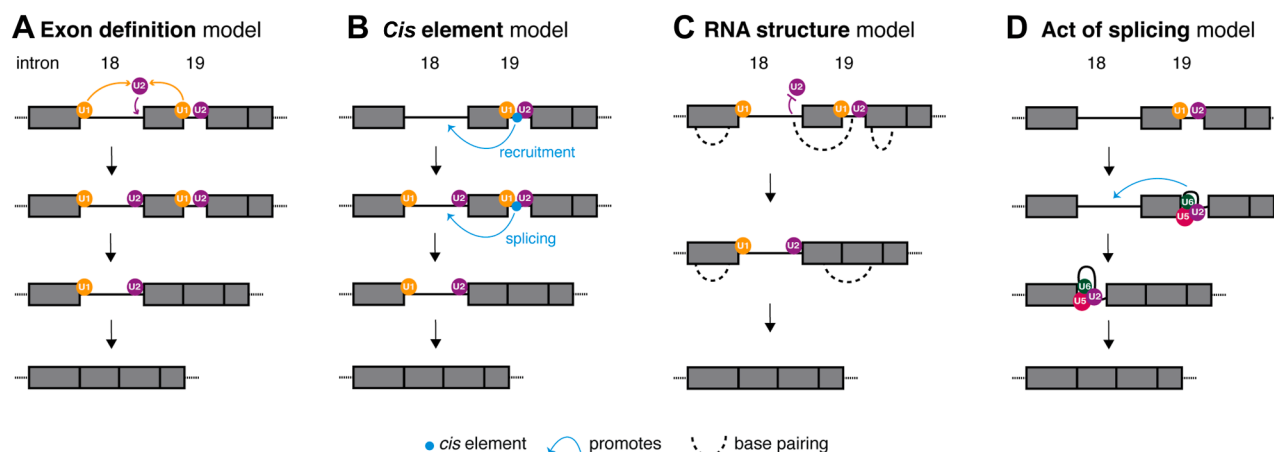


Figure 4. Models for the interdependency between introns 18 and 19. A, in the exon definition model, binding of U1 snRNP to the downstream 5' splice site enhances recruitment of U2 snRNP to the 3' splice site of intron 18. B, in the *cis* element model, a regulatory sequence in intron 19 (distinct from the 5' splice site) enables either recruitment of splicing factors to and/or excision of intron 18. C, in the RNA structure model, the presence of intron 19 results in a secondary structure that blocks binding of splicing factors to intron 18 splice sites or its branch point. Removal of intron 19 results in an alternative structure in which intron 18 is more accessible. D, in the act of splicing model, prior removal of intron 19 through splicing enables excision of intron 18.

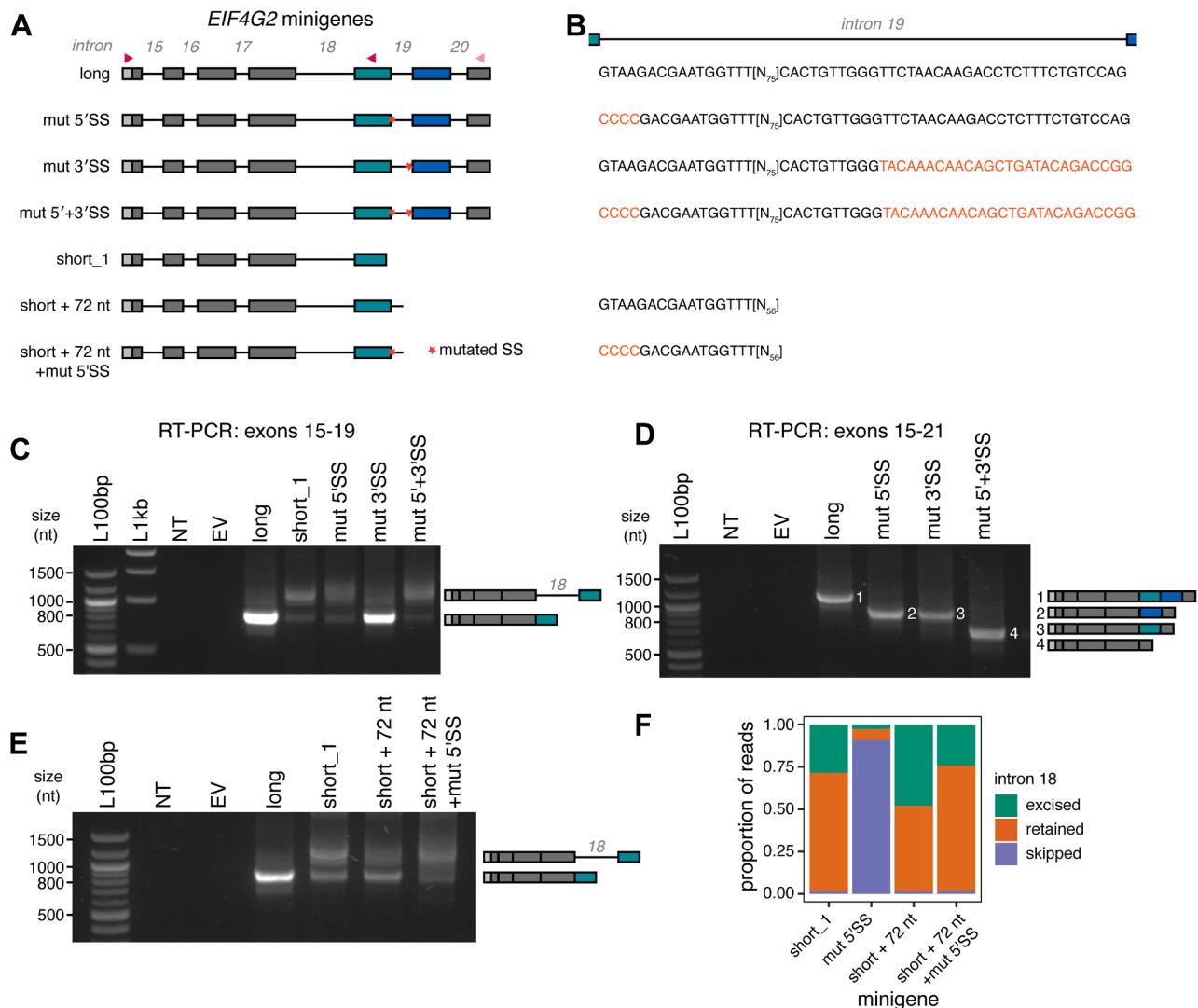


Figure 5. *EIF4G2* intron 18 excision partially depends on a functional downstream 5' splice site. A, schematic representation of the minigenes used in the other panels, with 5' and/or 3' splice site mutations (orange stars). Exons are shown as dark gray rectangles and introns as lines. The light gray rectangle represents the plasmid-specific 5'-UTR. B, intron 18 sequence in each minigene from A, showing the mutated regions in orange. C, agarose gel electrophoresis of RT-PCR products obtained from the expression of the indicated minigenes in HEK293T cells, using the primers shown as dark pink arrowheads in A. Products are schematized on the right. D, Same as C, but using the same minigene-specific forward primer and the reverse primer in exon 21 shown as light pink arrowhead in A. The short_1 minigene was not included because it lacks exon 21. Each product is numbered and schematized on the right. E, Same as C, but for the indicated minigenes. F, Nanopore direct RNA sequencing of the indicated minigene transcripts, showing the proportion of reads in which intron 18 is excised, retained or its 3' splice site is skipped. Only reads corresponding to the complete minigene transcripts were included. SS: splice site; int: intron; mut: mutated; L100 bp: 100 bp ladder.

excision (Fig. S3D), suggesting that the dependency between introns 18 and 19 is not mediated by internal intron 19 sequences.

We also investigated whether RNA-RNA interactions mediate intron 18 excision (model 3, Fig. 4C). If that were the case, we reasoned that these interactions likely involved base pairing between intron 18 and a region in *EIF4G2* that is only present in the long minigene (introns 19–20, exons 20–21), which would get disrupted in the short minigene (Fig. 7, middle panel). To explore this hypothesis, we analyzed published experimental high-throughput RNA-RNA interaction datasets obtained through crosslinking of RNA-RNA duplexes in living cells, followed by intramolecular proximity ligation and next-generation sequencing (42–49). These datasets were

collected from diverse human cell lines, including HEK293T cells, and were uniformly processed using our recently published tool DuplexDiscover (50). We identified 107 high-confidence RNA-RNA duplexes with both arms in *EIF4G2*, indicative of *cis* interactions, of which 35 were in the long minigene region (exons 15–21) (Table S2). To examine interactions that occurred in pre-mRNA, we filtered for duplexes with at least one arm overlapping an intron by at least one nucleotide. We found seven duplexes from five distinct datasets that showed a *cis* interaction between exons 17 and 18 and partially overlapped with intron 17 (Fig. 7), demonstrating our ability to uncover *cis* interactions involving introns in *EIF4G2* using this approach. Nevertheless, a single duplex with an arm in intron 18 was detected, with the second

Splicing dependency between EIF4G2 introns

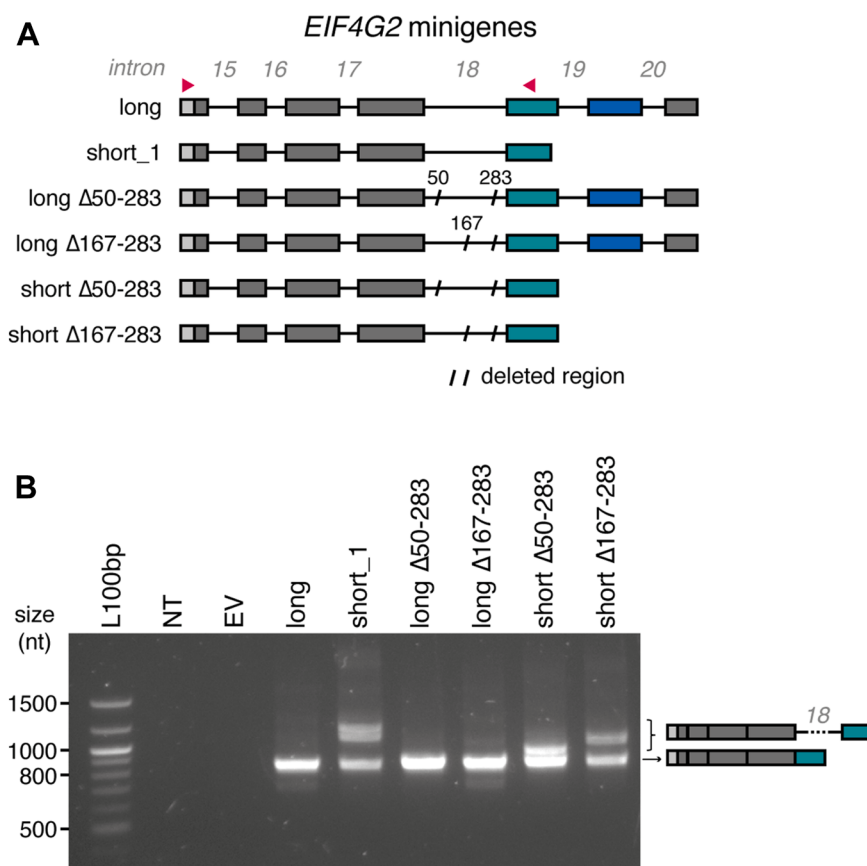


Figure 6. *EIF4G2* intron 18 dependency on its downstream intron is independent from its length. *A*, schematic representation of the minigenes used in *B*. Exons are shown as dark gray rectangles and introns as lines. The light gray rectangle represents the plasmid-specific 5'-UTR. Deleted regions in intron 18 are shown between slanted vertical lines, which are annotated with their position relative to the start of intron 18. *B*, agarose gel electrophoresis of RT-PCR products obtained from the expression of each minigene in *A* in HEK293T cells, using the primers shown as dark pink arrowheads. Products are schematized on the right. Amplicons with intron 18 retention (indicated by the curled bracket) vary in size due to the introduced deletions. bp, base pair; int, intron; L100 bp, 100 bp ladder.

arm mapping to a region in exon 19 that is present in all minigenes (Fig. 7). Moreover, no interactions between intron 18 and regions specific to the long minigene were uncovered. Overall, our findings underscore the importance of proper exon 19 definition for accurate intron 18 excision, while internal intron 19 *cis* elements and secondary structure do not appear to play a major role.

A downstream splicing event is required for efficient intron 18 excision

The partial rescue of intron 18 excision in the short + 72 nt minigene implies that exon definition is not the only determinant of the interdependency between introns 18 and 19. Thus, we tested whether a downstream act of splicing is required for efficient intron 18 removal (model 4, Fig. 4D). First, we added the full intron 19 sequence to the short_1 minigene, but without a downstream exon (short + int19), reasoning that an incomplete 3' splice junction would render intron 19 splicing-incompetent (Fig. 8A). We also generated the equivalent construct with a mutated 3' splice site (Fig. 8A, short + int19 + mut 3'SS). RT-PCR with a reverse primer in the minigene-specific 3' region confirmed that intron 19 was not excised in either construct (Fig. S3E). Consistent with the

presence of an intact downstream 5' splice site, these minigenes displayed intermediate levels of intron 18 retention compared to the short_1 minigene (Fig. 8, B and C). These results corroborate that the presence of intron 19 is not sufficient to promote intron 18 excision if it is not also splicing-competent. Moreover, mutating the 5' splice site of intron 19 in the short + int19 + mut 3'SS construct (Fig. 8A) led to intron 18 retention levels comparable to the short_1 minigene (Figs. 8, B, C and S3E). This indicates that exon 19 definition (model 1) and the act of splicing intron 19 (model 4) are additive and are both required for complete removal of intron 18.

Lastly, we asked if any downstream intron could promote intron 18 excision or if intron 19 was specifically required. We deleted exon 20 from the Δ int19 minigene (Δ int19-ex20), such that intron 20 was immediately downstream of exon 19. Interestingly, intron 18 was efficiently removed in this minigene but this effect was abolished if the 3' splice site of intron 20 was mutated (Fig. 8, B and C). Of note, some exon 19 skipping was observed in the Δ int19-ex20 minigene (band #3, Fig. S3F), akin to the Δ int19 minigene (band #2, Fig. 3C). However, intron 18 was efficiently excised when exon 19 was not skipped in the Δ int19-ex20 minigene (band #2, Fig. S3F), contrary to the Δ int19 minigene (Fig. 3B), potentially because

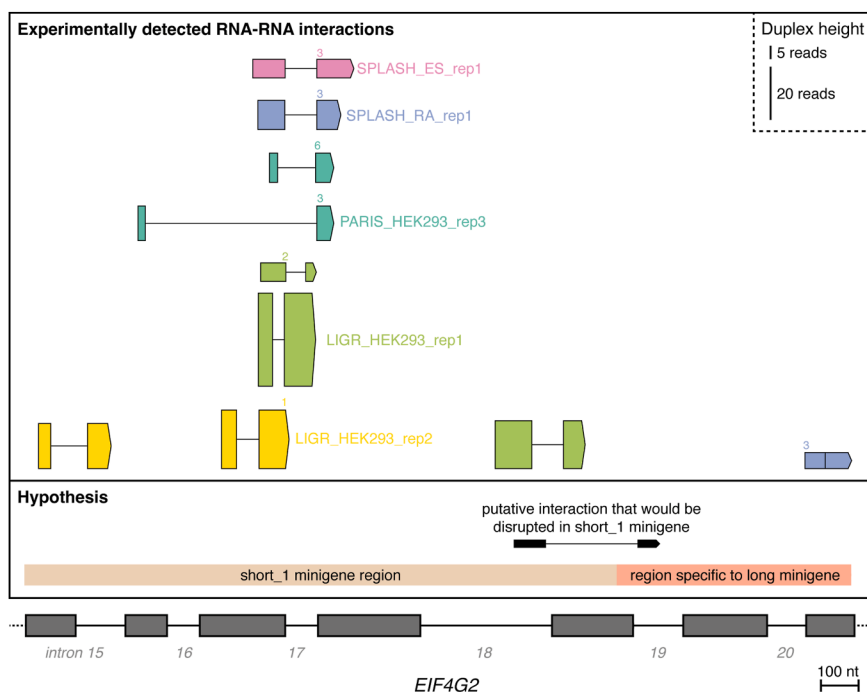


Figure 7. RNA-RNA interactions in EIF4G2. *Top panel:* Duplexes identified in experimental high-throughput RNA-RNA interaction datasets (42–44). Each arrow represents one RNA-RNA duplex, with the two arms (colored rectangles) indicating the two RNA sequences interacting with one another. Only duplexes with at least one arm overlapping an intron by ≥ 1 nucleotide in the region corresponding to the EIF4G2 long minigene are shown. The number of intronic nucleotides is indicated for breakpoints that are closer to splice junctions. The height of the duplex is proportional to the number of supporting reads, as shown in the top right legend. SPLASH, PARIS and LIGR-seq refer to the names of the methods used to map RNA-RNA interactions (42–44). ES: human embryonic stem cells; RA: human embryonic stem cells differentiated with retinoic acid for 5 days; rep: replicate. *Middle panel:* Schematic of the hypothesis in which an RNA-RNA interaction spanning intron 18 and the region specific to the long minigene would get disrupted in the short_1 minigene and contribute to intron 18 retention. *Bottom panel:* EIF4G2 gene structure, with introns shown as lines and exons as rectangles.

of the large composite exon in Δ int19 (exons 19–20 together). These results show that a functional downstream intron is the main requirement for efficient intron 18 excision, while the sequence identity of the downstream intron itself is not critical.

Discussion

Several recent studies have highlighted the coordinated nature of splicing within mammalian pre-mRNAs (6, 10, 17–23), yet the underlying mechanisms remain mostly unknown. Here, we demonstrate that excision of a local slow-poke intron, defined as being removed after its neighbors, is dependent on its downstream intron and we provide evidence that this interdependency is mediated through the combinatorial effect of exon definition and a downstream act of splicing.

Although exon definition is thought to be the predominant mode of splicing in mammalian cells, transcriptome-wide studies have shown that numerous introns are removed after their downstream neighbors (10, 18, 22, 27, 28), raising the question of how exon definition occurs if the downstream 5' splice site is no longer present. In two published examples of a mutation in a 5' splice site impairing excision of the upstream intron, the upstream intron was removed prior to the downstream intron, meaning that the downstream 5' splice site was still present and available for exon definition when the upstream intron was excised (17, 18). By contrast, the present study brings forth an example in which mutation of

the downstream 5' splice site impairs excision of the upstream intron even though the downstream intron is removed first. Given that the late intron 18 excision is observed across multiple cell types, these findings support the hypothesis that splice site commitment and intron excision are distinct steps that can be temporally separated. We propose that splice site recognition occurs in a coordinated manner on several neighboring introns when they are all still present (e.g., EIF4G2 introns 18–20). In the exon definition model (33), after binding of U1 snRNP to each 5' splice site, U2 snRNP is recruited to each 3' splice site through the synergistic action of upstream and downstream U1 snRNPs (6) to form consecutive cross-exon A and then pre-B complexes (Fig. 4A), followed by conversion to cross-intron B complexes (34). Alternatively, in the U1 relay model (35), in which a single U1 snRNP is recycled for adjacent splicing reactions, spliceosome assembly proceeds from 5' to 3' until the cross-intron B complex is formed and U1 snRNP has been handed off to the next intron. In both models, after B complex formation, each intron is removed in a defined order (Fig. 4A), likely based on additional intron-specific regulatory signals. Interestingly, several proteins specific to activated B complexes are enriched on introns that are removed after their neighbors (22) and on detained introns (51), supporting that the pause between splice site commitment and intron excision occurs between the activated B complex and catalytically activated B complexes. This pause may enable fine-tuning of splicing timing and mRNA production in response to cellular needs.

Splicing dependency between EIF4G2 introns

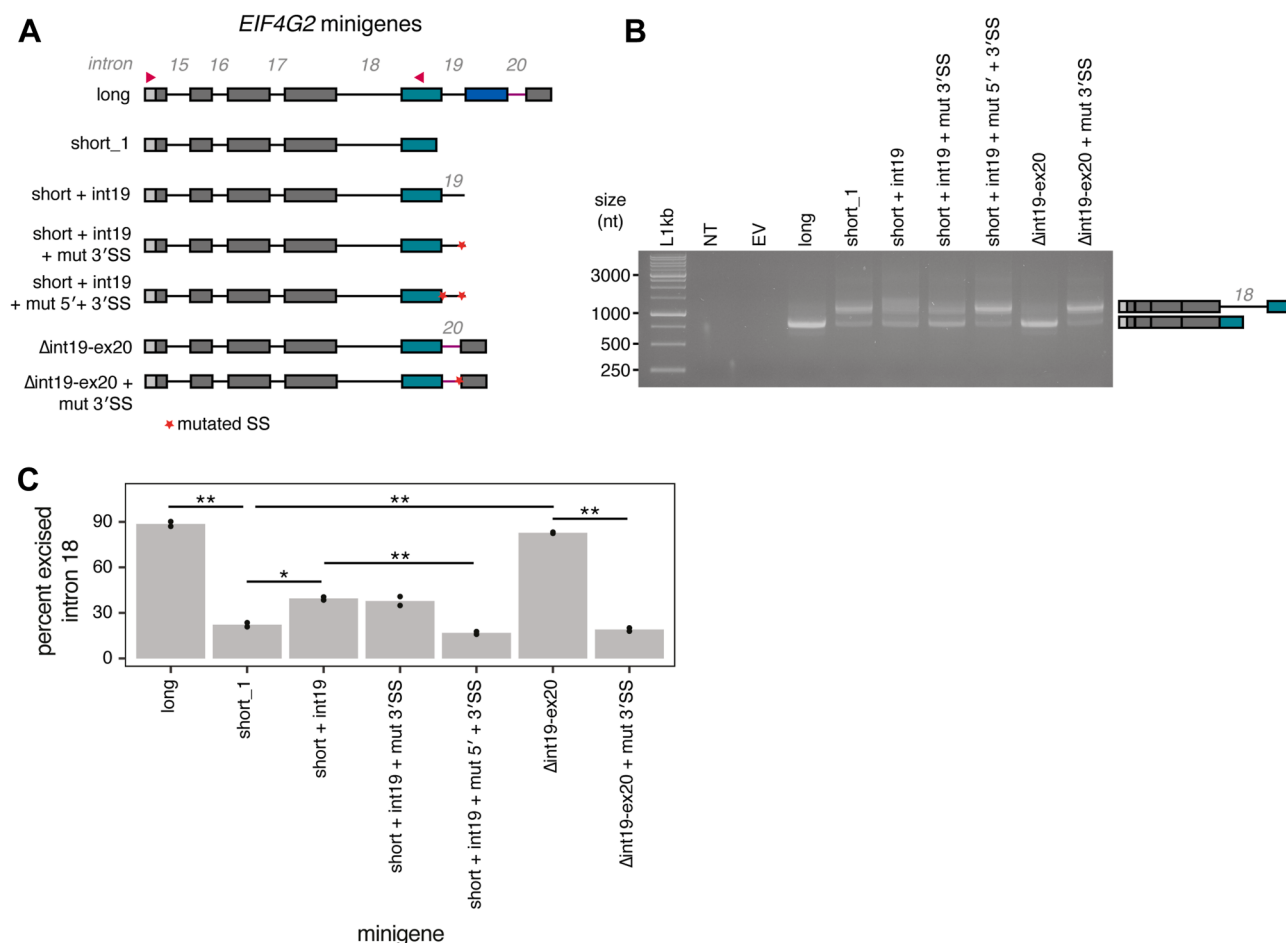


Figure 8. Intron 18 excision requires a downstream act of splicing. A, schematic representation of the minigenes used in B. Exons are shown as dark gray rectangles and introns as lines. The light gray rectangle represents the plasmid-specific 5'-UTR. The orange stars indicate 5' or 3' mutated splice sites, as in Figure 5B. B, agarose gel electrophoresis of RT-PCR products obtained from the expression of the minigenes shown in A in HEK293T cells, using the primers shown as dark pink arrowheads. Products are schematized on the right. C, qRT-PCR of *EIF4G2* intron 18 excision levels in the indicated minigenes. Percent excised levels were calculated as $1/(1 + 2^{\Delta C_t(\text{spliced-unsliced})})$. Dots represent two biological replicates from independent transfections. Minigenes were compared with an unpaired two-sided Student's *t* test. **p*-value < 0.05, ***p*-value < 0.01. SS: splice site; int: intron; mut: mutated; L1kb: 1 kb ladder.

Splicing order is largely conserved between cell types and is modulated by genetic variants, suggesting that it is primarily determined by nucleotide sequences (18, 27, 29). While our experiments show that intron 18 does not depend on specific sequences in intron 19, as its sequence can be shuffled or replaced by intron 20 (Figs. 8, B, C and S3D), regulatory motifs such as splicing enhancers or silencers may mediate delayed intron 18 removal. Previous analysis of binding motifs for 176 RNA-binding proteins (RBP) within intron pairs showed enrichment of a single motif, bound by the splicing factor SRSF3, in last-excised introns (10). Interestingly, addition of this motif further delayed excision of a local slowpoke intron, but had no impact when inserted in an early-excised intron (10), suggesting that multiple factors contribute to delayed intron removal and that these may be intron-specific. Furthermore, 16 RBP motifs were found enriched in first-excised introns, with the majority consisting of U-rich sequences (10). Consistently, U-rich sequences and U2AF2 binding were associated with early splicing in a separate study (28). Intron removal timing may also be mediated by RNA modifications. Indeed, a single intronic

pseudouridine was sufficient to enhance splicing *in vitro*, and pseudouridines were enriched in binding sites of RBPs, including U2AF2 (52). N⁶-methyladenosine modifications were also found to regulate splicing in a context-dependent manner, including for detained introns (5, 53, 54), which are excised in some contexts but not others. In the future, DRS of nascent RNA could be used to interrogate the association between RNA modifications (55) and splicing order. Moreover, combining DRS with perturbation of RBPs, including splicing factors and RNA modification writers, readers and erasers, will help illuminate their role in splicing order determination.

Surprisingly, shortening intron 18 did not alleviate its dependency on intron 19, suggesting that intron 18 length – absolute and relative to neighboring exons – is not the only determinant of its excision through exon definition. A recent study demonstrated that the difference in GC content between introns and their flanking exons is a key discriminator between intron and exon definition (11). The differential architecture is characterized by longer introns with low GC content ($\leq 40\%$) and relatively higher exon GC content, while

the leveled architecture is composed of shorter introns with high GC content flanked by exons with similarly high GC content (11). Shortening introns in minigenes with differential architecture did not affect splicing (11), akin to the results obtained here for *EIF4G2*. Indeed, with respective GC contents of 42, 36 and 37%, *EIF4G2* exon 18, intron 18 and exon 19 are more consistent with the differential architecture, which was maintained with the shortened versions of intron 18 (GC contents of 31–33%). Nevertheless, the gene *IFRD2*, in which we previously observed similar dependency of an upstream intron on its downstream 5' splice site for efficient excision (18), suggestive of exon definition, is characterized by a leveled architecture. Thus, additional properties intrinsic to each intron may contribute to removal through intron or exon definition. For instance, co-transcriptional lariat sequencing data indicate that many human introns can be excised independently from the downstream 5' splice site, presumably through intron definition, and that this is favored by a strong polypyrimidine tract enabling optimal U2AF2 binding (28). Future studies will allow determining whether other local slowpoke introns also depend on exon definition.

Importantly, exon definition was not the only requirement for interdependency in intron excision, as our data indicates that adding the 5' splice site of intron 19 or a splicing-incompetent intron 19 to the short_1 minigene is not sufficient to fully alleviate intron 18 retention. In fact, other local slowpoke introns were found to be dependent on their upstream intron for their removal (10), which cannot be explained by exon definition and synergistic U2 snRNP recruitment, as observed for *EIF4G2* intron 18. Our observations that efficient intron 18 excision can occur with a different downstream intron, as long as it is splicing-competent, suggest that local slowpoke introns depend on neighboring acts of splicing rather than specific *cis* elements in their flanking introns. Nevertheless, there appears to be a directional effect that may be specific to each slowpoke intron, as upstream prior splicing events had no impact on *EIF4G2* intron 18 retention. Interestingly, it was recently shown that the exon junction complex (EJC) regulates the coordinated inclusion of blocks of adjacent constitutive exons in a splicing order-dependent manner, with introns separating block exons excised first, while introns flanking the block are removed later (56). The authors postulated that EJC deposition on spliced block exons helped the recognition of flanking weaker splice sites, as the block exons were skipped in the absence of EJC components. These findings argue against a purely 5' to 3' recognition of splice sites, as proposed in the exon definition and U1 relay models. While different models may apply to distinct introns, EJC deposition may also help to promote the excision of later-excised introns that are paused between splice site commitment and splicing catalysis. Such a model could explain why skipping of exons 19 to 20 is favored over intron 18 excision in the Δ int19 *EIF4G2* minigene, since the EJC would not be deposited near the junction between exons 19 and 20. Another possibility is that some splicing factors (e.g. required for catalytic activation or splicing catalysis) are recruited to intron 18 only if they can be

recycled after completing a nearby splicing reaction. However, why such a mechanism would only affect certain introns in a directional manner remains unknown. In the future, experiments that allow tracking splicing factors and their association with RNA over time may provide further insights into the mechanisms underlying interdependency between local slowpoke introns and their neighbors.

Our findings also have important implications for the more than 15% of disease-causing genetic variants that are in splice sites or splicing regulatory elements (57). Two-intron minigene assays are frequently used to determine whether and how such variants disrupt pre-mRNA splicing (58–60). Yet, recent studies have shown that the surrounding sequence context influences how genetic variants or other perturbations impact splicing, including in minigene assays (61–64). Our findings provide further evidence that some introns are not correctly excised in the absence of their neighbors. Thus, human genetics studies should prioritize minigenes with larger genomic context when feasible, including splicing-competent neighboring introns, as this is more likely to faithfully recapitulate splicing of endogenous transcripts and to accurately reveal how variants impact pre-mRNA splicing in genetic disorders.

Experimental procedures

Splicing order analysis of endogenous *EIF4G2* transcripts

EIF4G2 splicing order scores were computed by multiplying the frequency of intermediate isoforms required to proceed through a given splicing order, as previously described (18, 29), using published DRS data from LCLs, K562 cells, HeLa cells, and induced pluripotent stem cells-derived neuron progenitors and motor neurons (GEO accession numbers GSE256190, GSE208225, GSE232455). Splicing order scores are bound by 0 and 1, with the most frequent splicing orders have higher scores. In all LCLs except GM19099, splicing order was analyzed simultaneously for introns 15 to 20 (Fig. 1A, Table S1). Due to the lower sequencing depth in datasets from other cell types, splicing order that had previously been separately computed for introns 16 to 18 and 18 to 20 was extracted from Table S5 of reference (18) (Fig. S2A). The cumulative splicing order score for each intron being excised last (Figs. 1A and S2A) was calculated by summing scores from all splicing orders in which the considered intron was excised last. For allele-specific splicing order analysis in GM19099, reads were mapped to their parental allele of origin and splicing statuses were computed as previously described (29). We extracted splicing statuses from all reads spanning introns 17 and 18 and filtered for reads in which one intron was excised and one intron was still present. The frequency of intron 18 or intron 17 being still present in the read (excised last) was computed for each allele (Fig. 1B).

Cloning of *EIF4G2* minigenes

Genomic DNA was extracted from LCL GM19099 (Coriell) using DNeasy Blood & Tissue kit (Qiagen, cat. no.

Splicing dependency between EIF4G2 introns

69504). *EIF4G2* regions spanning exons 15 to 19 and exons 15 to 21 were amplified from genomic DNA using primers with MfeI and XhoI restriction sites (Table S3) and the Phusion High-Fidelity PCR Master Mix [New England Biolabs (NEB), cat. no. M0531] according to the manufacturer's instructions. PCR products were run on an agarose gel for 1 h at 130V. Bands of the expected size were purified using the MinElute Gel Extraction kit (Qiagen, cat. no. 28604). PCR products or 5 µg of plasmid pCMV-M1 (Addgene, cat. no. 23007) were digested with 1 µl of MfeI and 1 µl of XhoI at 37 °C for 1 h. The linearized plasmid was run on an agarose gel for 1 h at 130V and purified using the MinElute Gel Extraction kit. Digested PCR products were purified using the Monarch PCR & DNA Cleanup kit (NEB, cat. no. T1030). Ligation was performed with 7 µl of purified PCR product, 50 ng of digested plasmid, 1 µl of T4 DNA Ligase (NEB, cat. no. M0202S) and 1 µl of T4 DNA Ligase buffer in a total volume of 10 µl. Ligation reactions were incubated at room temperature for 1 h, transformed into 5- α competent *Escherichia coli* (NEB, cat. no. C2987), plated on LB-agar plates with 50 mg/ml of kanamycin and incubated at 37 °C overnight. Positive colonies were picked and individually grown in LB-agar with 50 mg/ml kanamycin. Plasmid DNA was extracted using the Monarch Plasmid Miniprep kit (NEB, cat. no. T1010) or the Plasmid Plus Midi Kit (Qiagen, cat. no. 12943). Full plasmid sequences were verified by nanopore sequencing through Primordium or the RNomics Platform at Université de Sherbrooke. Plasmids corresponding to the GM19099 reference allele (genotype A at SNP rs200245989) were used for all experiments (long and short_1, unless otherwise specified (long_alt and short_1_alt carrying genotype G at SNP rs200245989, Figs. 2C and S2A). Deletions and mutations were introduced using the Q5 site-directed mutagenesis kit (NEB, cat. no. E0554) according to manufacturer's instructions. Primers used for mutagenesis are listed in Table S3. 5' and 3' splice site mutated sequences were adapted from (6).

Cell culture

HEK293T cells (kind gift from François Bachand, Université de Sherbrooke) were grown in DMEM (Wisent, cat. no. 319-005-CL) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin and tested negative for *mycoplasma*. For minigene transfections, 0.6 million cells were plated in each well of a 6-well plate. Twenty-four hours later, cells were transfected with 2.5 µg of plasmid using 7.5 µl of Lipofectamine 3000 (Life Technologies, cat. no. L0000008). Cells were harvested 24 hours after transfection. All transfection and subsequent experiments were performed in ≥ 2 biological replicates.

RNA extraction

Cells were lysed in 1 ml of QIAzol lysis reagent (Qiagen, cat. no. 79306) and incubated at room temperature for 5 min. After addition of 200 µl of chloroform, samples were vigorously mixed by hand for 15 s, incubated at room temperature

for 2 to 3 min and centrifuged at 12,000 $\times$ g for 15 min at 4 °C. The aqueous phase was transferred to a new tube, mixed with 500 µl of isopropanol and 2 µl of GlycoBlue (Thermo Fisher Scientific, cat. no. AM9515) and incubated for 10 min at room temperature. Samples were centrifuged at 12,000 $\times$ g for 10 min at 4 °C. RNA pellets were washed twice by adding 1 ml of 75% ethanol and centrifuging at 7000 $\times$ g for 5 min at 4 °C. Air-dried pellets were resuspended in 30 to 50 µl of nuclease-free water.

DNase treatment and RT-PCR

DNase treatment was performed on 2 µg of RNA using amplification-grade DNase I (Thermo Fisher Scientific cat. no. 18068015) according to the manufacturer's instructions. For reverse transcription, 700 ng of DNase-treated RNA was mixed with 2 µl of Random Primer Mix (60 µM, NEB cat. no. S1330) and 1 µl of 10 mM dNTPs in a total volume of 10 µl. After incubation at 65 °C for 5 min, 5 µl of 5X Induro RT reaction buffer, 0.2 µl of murine RNase Inhibitor (40 U/µl, NEB cat. no. M0314), 1 µl of Induro Reverse Transcriptase (200 U/µl, NEB, cat. no. M0681) and 4.8 µl of nuclease-free water were added to the samples. Reverse transcription was carried out by incubating at 25 °C for 2 min, 55 °C for 30 min and 95 °C for 1 min. PCR was performed with 1 µl of cDNA and the Q5 Hot Start High-Fidelity DNA Polymerase (NEB, cat. no. M0491) in 25 µl reactions according to the manufacturer's instructions, including use of the High GC enhancer. Primers used for RT-PCR are listed in Table S4. The following cycling conditions were used: 98 °C for 30 s, 35 cycles of 98 °C for 10 s, 60 °C for 30 s and 72 °C for 1 min, 72 °C for 2 min. For agarose gel electrophoresis, 10 µl of the PCR reactions were mixed with 2 µl of 6X gel loading dye (NEB, cat. no. B7024S) and migrated on a 1% agarose gel for 45 min at 120V, along with 100 bp and/or 1 kb DNA Ladders (NEB, cat. no. N3231L and N3232L or FroggaBio, cat. no. DM010). qRT-PCR to measure *EIF4G2* expression levels (Fig. S2C) was performed by combining 3 µl of cDNA diluted in water at a ratio of 1:250, 10 µl of Luna Universal qPCR master mix (NEB, cat. no. M3003) and 0.5 µl of forward and reverse primers (10 µM, Table S4) in a total volume of 20 µl with the following cycling conditions: 95 °C for 30 s, 40 cycles of 95 °C for 5 s, 60 °C for 10 s, and a final melt curve (65 °C to 95 °C with 0.5 °C increment for 5 s). qRT-PCR to measure intron 18 excision levels (Fig. 8C) was carried out as described above, but a minigene-specific reverse transcription was performed with 1.2 µl of a reverse primer targeting the SV40 sequence (100 µM, Table S4) instead of the Random Primer Mix. The resulting cDNA was diluted at a ratio of 1:100 and 1.5 µl was combined with 5 µl of Luna Universal qPCR master mix and 0.25 µl of forward and reverse primers (10 µM, Table S4) in a total volume of 10 µl with the following cycling conditions: 95 °C for 30 s, 40 cycles of 95 °C for 5 s, 58 °C for 10 s, and a final melt curve (65 °C to 95 °C with 0.5 °C increment for 5 s). Two primer pairs were used, which respectively targeted the unspliced intron 18-exon 19 junction and the spliced exon 18-exon 19 junction. Intron 18 percent excised levels were

calculated as $1/(1 + 2^{\Delta Ct(\text{spliced-unspliced})})$, as previously done (65).

Nanopore sequencing of RT-PCR amplicons

PCR was performed with 1 μ l of cDNA as described above, but with only 20 cycles. For the endogenous transcripts and the long minigene, a reverse primer in exon 21 was used (Table S4). For the short minigene, a reverse primer in exon 19 was used (Table S4). PCR reactions were cleaned-up using the Monarch PCR and DNA clean-up kit (NEB, cat. no. T1030S). 10 to 130 ng of DNA from each sample was carried forward for end-repair using the NEBNext Ultra II End repair/dA-tailing Module (NEB, cat. no. E7546S). The DNA was mixed with 1X volume of AMPure XP beads (Beckman Coulter, cat. no. A63881), incubated on a rotator for 5 min at room temperature, washed twice with 80% ethanol and eluted in 10 μ l water. For ligation of unique sample barcodes, 10 ng of end-prepped amplicon in a total volume of 7.5 μ l was combined with 2.5 μ l native barcode [Oxford Nanopore Technologies (ONT), cat. no. SQK-NBD114.24] and 10 μ l of Blunt/TA Ligase Master Mix (NEB, cat. no. M0367). The ligation reaction was incubated for 20 min at room temperature and inactivated with EDTA. All samples were pooled together and mixed with 0.4X volume of AMPure XP beads. The pooled sample was incubated on a rotator for 10 min at room temperature, washed twice with 80% ethanol and eluted in 35 μ l water. 30 μ l was combined with 5 μ l of native adapter (ONT, cat. no. SQK-NBD114.24), 10 μ l of NEBNext Quick Ligation Reaction Buffer and 5 μ l of Quick T4 Ligase (NEBNext Quick Ligation Module, NEB, cat. no. E6056). The ligation reaction was incubated for 20 min at room temperature, mixed with 0.4X volume of AMPure XP beads and incubated on a rotator for 10 min at room temperature. Sample clean-up and loading onto a FLO-MIN114 flow cell was performed with the Native Barcoding Kit 24 V14 (ONT, cat. no. SQK-NBD114.24) according to manufacturer's instructions. Sequencing was performed on a MinION device.

Direct RNA nanopore sequencing

DRS library preparation was performed using the kit SQK-RNA004 (ONT) with 1 μ g of total RNA according to manufacturer's instructions. Sequencing was performed on a FLO-PRO004RA flow cell using a MinION device. Sequencing was carried out until >130,000 reads were obtained, which was sufficient to analyze minigene splicing given its high overexpression level (Fig. S2C). Each flow cell was used three times. Between each sample, the flow cell was washed with the kit EXP-WSH004 (ONT) according to manufacturer's instructions.

Nanopore sequencing data processing and analysis

For cDNA-PCR sequencing, basecalling and demultiplexing were performed with guppy v6.5.7 (ONT). All reads that passed basecalling were aligned to the reference human genome GRCh38 using minimap2 v2.28 (66) with parameters

-ax splice. For DRS, fast basecalling was performed during sequencing with MinKNOW or after sequencing with Dorado v0.7.3 (ONT) and model rna004_130bps_fast@v5.0.0. All reads that passed basecalling were aligned to the reference human genome using minimap2 v2.28 (66) with parameters -ax splice -uf -k14. Uniquely mapped reads were extracted as detailed in (65). Coverage tracks in Figures 2D and 3E were generated with pyGenomeTracks (67). Intron excision statuses (excised, retained or skipped) were determined using the alignment CIGAR strings as previously described (18, 65). For cDNA-PCR, only reads spanning all expected amplicon introns (*i.e.* 15–20 for no transfection control and the long minigene, introns 15–18 for the short_1 minigene) were kept in the analysis. For DRS, only reads spanning all expected minigene introns (*i.e.* introns 15–20 for the long, Δ int19+20 and mut 5'SS minigenes, introns 15–18 for the short_1, short + 72 nt and short + 72 nt + mut 5'SS minigenes) were considered in order to exclude endogenous *EIF4G2* transcripts.

Analysis of RNA-RNA interaction datasets

Previously published high-throughput RNA-RNA interactions datasets (42–49) were processed using DuplexDiscoverer (50) to identify duplexes in *EIF4G2*. First, we compiled all publicly available datasets of RNA-RNA interaction profiling experiments, including RIC-seq, SPLASH, PARIS, LIGR-seq, PARIS2, KARR-seq, and CAR-SPLASH (42–49). For each sample, we ran DuplexDiscoverer using parameters in accordance with the processing guidelines described in the original publications. High-confidence duplexes with more than five supporting reads were subsequently filtered for hypothetical RNA-RNA interactions to include only those involving *cis* interactions within the minigene region (exons 15–21 and intervening introns), for which at least one anchor spanned an intron by ≥ 1 nucleotides. Resulting duplexes were converted to BED12 format, with each anchor representing one block, and plotted using pyGenomeTracks (67).

Data availability

The *EIF4G2* nanopore sequencing data generated in this article is available at the Gene Expression Omnibus (accession number GSE324595).

Supporting information—This article contains supporting information (34, 35, 37, 42–49).

Acknowledgments—We thank members of the Choquet and Meyer labs for helpful discussions; Carole-Anne Martineau, Salomé Sabatié, Mark Bayfield and François Bachand for critical reading of the manuscript; L. Stirling Churchman (Harvard University) for sharing the long and short_1 minigenes that were initially cloned in her laboratory by K. C.; and the RNomics Platform at Université de Sherbrooke and the Sanger sequencing platform at the Centre de recherche du CHU de Québec-Université Laval for sequencing services.

Splicing dependency between EIF4G2 introns

Author contributions—M. L. P., J. V., T. M., C. S., V. T., E. S., I. S. A., I. M. M. and K. C. writing—review and editing; M. L. P., V. T., and E. S. visualization; M. L. P., J. B., T. M., C. S., V. T., E. S., I. S. A., I. M. M., and K. C. investigation; M. L. P., T. M., V. T., E. S., I. M. M., and K. C. formal analysis; J. V. validation; J. V. and K. C. methodology; V. T., E. S., I. M. M., and K. C. software; V. T., E. S., I. M. M., and K. C. conceptualization; I. M. M. and K. C. supervision; K. C. writing—original draft; K. C. funding acquisition.

Funding and additional information—This research was undertaken thanks in part to funding from the Canada First Research Excellence Fund awarded to the D2R Initiative at McGill University (start-up funds to K. C. and doctoral award to I. S. A.), to start-up funds from the Faculty of Medicine and Health Sciences of Université de Sherbrooke and the Research Center on Aging, to a Discovery Grant from the National Science and Engineering Research Council of Canada, and to a Project Grant from the Canadian Institutes of Health Research. K. C. is the recipient of a Junior 1 Research Scholar award from the Fonds de recherche du Québec (<https://doi.org/10.69777/365832>). T. M. was supported by an undergraduate research fellowship from the Research Center on Aging. V. T., E. S. and I. M. M. gratefully acknowledge funding by the Helmholtz Association.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CMV, cytomegalovirus; DRS, direct RNA sequencing; EJC, exon junction complex; LCL, lymphoblastoid cell line; RBP, RNA-binding protein; SNP, single nucleotide polymorphism; snRNP, small nuclear ribonucleoprotein.

References

- Kastner, B., Will, C. L., Stark, H., and Lührmann, R. (2019) Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harb. Perspect. Biol.* <https://doi.org/10.1101/cshperspect.a032417>
- Zhang, S., Aibara, S., Vos, S. M., Agafonov, D. E., Lührmann, R., and Cramer, P. (2021) Structure of a transcribing RNA polymerase II-U1 snRNP complex. *Science* **371**, 305–309
- Shao, C., Hao, Y., Jiang, L., Wang, D., Wang, R., Li, Y., *et al.* (2025) Dynamic U2AF cycling defines two phases of cotranscriptional pre-mRNA splicing. *Science* **389**, ead9141
- Shine, M., Gordon, J., Schärffen, L., Zigackova, D., Herzel, L., and Neugebauer, K. M. (2024) Co-transcriptional gene regulation in eukaryotes and prokaryotes. *Nat. Rev. Mol. Cell Biol.* **25**, 534–554
- Choquet, K., Patop, I. L., and Churchman, L. S. (2025) The regulation and function of post-transcriptional RNA splicing. *Nat. Rev. Genet.* **26**, 378–394
- Braun, J. E., Friedman, L. J., Gelles, J., and Moore, M. J. (2018) Synergistic assembly of human pre-spliceosomes across introns and exons. *Elife* **7**, e37751
- Krainer, A. R., Maniatis, T., Ruskin, B., and Green, M. R. (1984) Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell* **36**, 993–1005
- Black, D. L., Chabot, B., and Steitz, J. A. (1985) U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* **42**, 737–750
- Hoskins, A. A., Friedman, L. J., Gallagher, S. S., Crawford, D. J., Anderson, E. G., Wombacher, R., *et al.* (2011) Ordered and dynamic assembly of single spliceosomes. *Science* **331**, 1289–1295
- Kim, S. W., Taggart, A. J., Heintzelman, C., Cygan, K. J., Hull, C. G., Wang, J., *et al.* (2017) Widespread intra-dependencies in the removal of introns from human transcripts. *Nucleic Acids Res.* **45**, 9503–9513

- Tammer, L., Hameiri, O., Keydar, I., Roy, V. R., Ashkenazy-Titelman, A., Custódio, N., *et al.* (2022) Gene architecture directs splicing outcome in separate nuclear spatial regions. *Mol. Cell* **82**, 1021–1034.e8
- Erkelenz, S., Theiss, S., Kaisers, W., Ptok, J., Walotka, L., Müller, L., *et al.* (2018) Ranking noncanonical 5' splice site usage by genome-wide RNA-seq analysis and splicing reporter assays. *Genome Res.* **28**, 1826–1840
- Stoilov, P., Lin, C.-H., Damoiseaux, R., Nikolic, J., and Black, D. L. (2008) A high-throughput screening strategy identifies cardiotoxic steroids as alternative splicing modulators. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 11218–11223
- Zheng, S., Damoiseaux, R., Chen, L., and Black, D. L. (2013) A broadly applicable high-throughput screening strategy identifies new regulators of Dlg4 (Psd-95) alternative splicing. *Genome Res.* **23**, 998–1007
- O'Neill, M. J., Yang, T., Laudeman, J., Calandranis, M. E., Harvey, M. L., Solus, J. F., *et al.* (2024) ParSE-seq: a calibrated multiplexed assay to facilitate the clinical classification of putative splice-altering variants. *Nat. Commun.* **15**, 8320
- Souček, P., Réblová, K., Kramárek, M., Radová, L., Grymová, T., Hujová, P., *et al.* (2019) High-throughput analysis revealed mutations' diverging effects on SMN1 exon 7 splicing. *RNA Biol.* **16**, 1364–1376
- Talerico, M., and Berget, S. M. (1990) Effect of 5' splice site mutations on splicing of the preceding intron. *Mol. Cell. Biol.* **10**, 6299–6305
- Choquet, K., Baxter-Koenigs, A. R., Dülk, S.-L., Smalec, B. M., Rouskin, S., and Churchman, L. S. (2023) Pre-mRNA splicing order is pre-determined and maintains splicing fidelity across multi-intronic transcripts. *Nat. Struct. Mol. Biol.* **30**, 1064–1076
- Grabowski, P. J., Nasim, F. U., Kuo, H. C., and Burch, R. (1991) Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.* **11**, 5919–5928
- Sánchez-Escabias, E., Guerrero-Martínez, J. A., and Reyes, J. C. (2022) Co-transcriptional splicing efficiency is a gene-specific feature that can be regulated by TGFβ. *Commun. Biol.* **5**, 277
- [preprint] Casill, A. D., Haimowitz, A. J., Kosmyna, B., Query, C. C., Ye, K., and Gamble, M. J. (2021) Spatial organization of transcript elongation and splicing kinetics. *bioRxiv*. <https://doi.org/10.1101/2021.01.28.428713>
- Drexler, H. L., Choquet, K., and Churchman, L. S. (2020) Splicing kinetics and coordination revealed by direct nascent RNA sequencing through nanopores. *Mol. Cell* **77**, 985–998.e8
- Reimer, K. A., Mimoso, C. A., Adelman, K., and Neugebauer, K. M. (2021) Co-transcriptional splicing regulates 3' end cleavage during mammalian erythropoiesis. *Mol. Cell* **81**, 998–1012.e7
- Sousa-Luís, R., Dujardin, G., Zukher, I., Kimura, H., Weldon, C., Carmo-Fonseca, M., *et al.* (2021) POINT technology illuminates the processing of polymerase-associated intact nascent transcripts. *Mol. Cell* **81**, 1935–1950.e6
- Kessler, O., Jiang, Y., and Chasin, L. A. (1993) Order of intron removal during splicing of endogenous adenine phosphoribosyltransferase and dihydrofolate reductase pre-mRNA. *Mol. Cell Biol.* **13**, 6211–6222
- Takahara, K., Schwarze, U., Imamura, Y., Hoffman, G. G., Toriello, H., Smith, L. T., *et al.* (2002) Order of Intron Removal influences multiple splice outcomes, including a two-exon skip, in a COL5A1 acceptor-site mutation that results in abnormal Pro-α1(V) N-Propeptides and Ehlers-Danlos Syndrome type I. *Am. J. Hum. Genet.* **71**, 451–465
- Gohr, A., Iñiguez, L. P., Torres-Méndez, A., Bonnal, S., and Irimia, M. (2023) InSplico: effective computational tool for studying splicing order of adjacent introns genome-wide with short and long RNA-seq reads. *Nucleic Acids Res.* **51**, e56
- Zeng, Y., Fair, B. J., Zeng, H., Krishnamohan, A., Hou, Y., Hall, J. M., *et al.* (2022) Profiling lariat intermediates reveals genetic determinants of early and late co-transcriptional splicing. *Mol. Cell* **82**, 4681–4699.e8
- Choquet, K., Chaumont, L.-P., Bache, S., Baxter-Koenigs, A. R., and Churchman, L. S. (2025) Genetic regulation of nascent RNA maturation revealed by direct RNA nanopore sequencing. *Genome Res.* **35**, 712–724
- Baralle, F. E., and Giudice, J. (2017) Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **18**, 437–451
- Talerico, M., and Berget, S. M. (1994) Intron definition in splicing of small Drosophila introns. *Mol. Cell. Biol.* **14**, 3434–3445

32. Berget, S. M. (1995) Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**, 2411–2414
33. De Conti, L., Baralle, M., and Buratti, E. (2013) Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* **4**, 49–60
34. Zhang, Z., Kumar, V., Dybkov, O., Will, C. L., Zhong, J., Ludwig, S. E. J., *et al.* (2024) Structural insights into the cross-exon to cross-intron spliceosome switch. *Nature* **630**, 1012–1019
35. Yoon, Y., Quan, C., Soles, L. V., and Shi, Y. (2026) Coordinating mRNA maturation: the U1 relay model. *Mol. Cell.* <https://doi.org/10.1016/j.molcel.2026.01.006>
36. de la Mata, M., Lafaille, C., and Kornblihtt, A. R. (2010) First come, first served revisited: factors affecting the same alternative splicing event have different effects on the relative rates of intron removal. *RNA* **16**, 904–912
37. Morita, E., Arai, J., Christensen, D., Votteler, J., and Sundquist, W. I. (2012) Attenuated protein expression vectors for use in siRNA rescue experiments. *BioTechniques*. **0**, 1–5
38. Sterner, D. A., Carlo, T., and Berget, S. M. (1996) Architectural limits on split genes. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15081–15085
39. Saldi, T., Riemondy, K., Erickson, B., and Bentley, D. L. (2021) Alternative RNA structures formed during transcription depend on elongation rate and modify RNA processing. *Mol. Cell.* **81**, 1789–1801.e5
40. [preprint] Schärffen, L., Bech, P., Podsiwałow-Bartnicka, P., and Neugebauer, K. M. (2025) Control of gene output by intron RNA structure. *Biorxiv*. <https://doi.org/10.1101/2025.11.08.687378>
41. Rangan, R., Huang, R., Hunter, O., Pham, P., Ares, M., and Das, R. (2025) Comprehensive analysis of *Saccharomyces cerevisiae* intron structures in vivo. *Nat. Struct. Mol. Biol.* **32**, 1488–1502
42. Aw, J. G. A., Shen, Y., Wilm, A., Sun, M., Lim, X. N., Boon, K.-L., *et al.* (2016) In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order Organization and regulation. *Mol. Cell.* **62**, 603–617
43. Lu, Z., Zhang, Q. C., Lee, B., Flynn, R. A., Smith, M. A., Robinson, J. T., *et al.* (2016) RNA Duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279
44. Sharma, E., Sterne-Weiler, T., O'Hanlon, D., and Blencowe, B. J. (2016) Global mapping of human RNA-RNA interactions. *Mol. Cell.* **62**, 618–626
45. Cai, Z., Cao, C., Ji, L., Ye, R., Wang, D., Xia, C., *et al.* (2020) RIC-seq for global in situ profiling of RNA-RNA spatial interactions. *Nature* **582**, 432–437
46. Liang, L., Cao, C., Ji, L., Cai, Z., Wang, D., Ye, R., *et al.* (2023) Complementary Alu sequences mediate enhancer-promoter selectivity. *Nature* **619**, 868–875
47. Wu, T., Cheng, A. Y., Zhang, Y., Xu, J., Wu, J., Wen, L., *et al.* (2024) KARR-seq reveals cellular higher-order RNA structures and RNA-RNA interactions. *Nat. Biotechnol.* **42**, 1909–1920
48. Zhang, M., Li, K., Bai, J., Velema, W. A., Yu, C., van Damme, R., *et al.* (2021) Optimized photochemistry enables efficient analysis of dynamic RNA structures and interactomes in genetic and infectious diseases. *Nat. Commun.* **12**, 2344
49. Shenasa, H., Fong, N., Erickson, B., Iosub, I. A., Ule, J., and Bentley, D. L. (2025) CAR-SPLASH identifies nascent pre-mRNA structures implicated in kinetic coupling and alternative splicing. *Proc. Natl. Acad. Sci. U. S. A.* **122**, e2504034122
50. Semenchenko, E., Tsybulskiy, V., and Meyer, I. M. (2025) DuplexDiscoverer: a computational method for the analysis of experimental duplex RNA-RNA interaction data. *Nucleic Acids Res.* **53**, gkaf266
51. Meng, D., Zheng, Q., Zhang, X., Luo, L., and Jia, Y. (2022) A molecular brake that modulates spliceosome pausing at detained introns contributes to neurodegeneration. *Protein Cell.* <https://doi.org/10.1093/procel/pwac008>
52. Martinez, N. M., Su, A., Burns, M. C., Nussbacher, J. K., Schaening, C., Sathe, S., *et al.* (2022) Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect pre-mRNA processing. *Mol. Cell.* **82**, 645–659.e9
53. Mendel, M., Delaney, K., Pandey, R. R., Chen, K.-M., Wenda, J. M., Vågbo, C. B., *et al.* (2021) Splice site m6A methylation prevents binding of U2AF35 to inhibit RNA splicing. *Cell* **184**, 3125–3142.e25
54. González-Iglesias, A., Arcas, A., Domingo-Muelas, A., Mancini, E., Galcerán, J., Valcárcel, J., *et al.* (2024) Intron detention tightly regulates the stemness/differentiation switch in the adult neurogenic niche. *Nat. Commun.* **15**, 2837
55. Cruciani, S., and Novoa, E. M. (2026) The new era of single-molecule RNA modification detection through nanopore base-calling models. *Nat. Rev. Mol. Cell Biol.* **27**, 10–18
56. Bergfort, A., Gordon, J. M., Gazzara, M. R., Hung, C.-T., Lee, B., Barash, Y., *et al.* (2025) The exon junction complex coordinates the cotranscriptional inclusion of blocks of neighboring exons. *Genes Dev.* <https://doi.org/10.1101/gad.353081.125>
57. Park, E., Pan, Z., Zhang, Z., Lin, L., and Xing, Y. (2018) The expanding landscape of alternative splicing variation in human populations. *Am. J. Hum. Genet.* **102**, 11–26
58. Soens, Z. T., Branch, J., Wu, S., Yuan, Z., Li, Y., Li, H., *et al.* (2017) Leveraging splice-affecting variant predictors and a minigene validation system to identify Mendelian disease-causing variants among exon-captured variants of uncertain significance. *Hum. Mutat.* **38**, 1521–1533
59. Desviat, L. R., Pérez, B., and Ugarte, M. (2012) Minigenes to confirm exon skipping mutations. *Methods Mol. Biol. Clifton NJ.* **867**, 37–47
60. Rosso, L. E., Pianigiani, G., Morgan, A., Rubinato, E., Paccagnella, E., Lenarduzzi, S., *et al.* (2025) Unraveling the functional impact of splicing variants in inherited hearing disorders through minigene splicing assays. *Biomedicines* **13**, 2245
61. Holm, L. L., Doktor, T. K., Hansen, M. B., Petersen, U. S. S., and Andresen, B. S. (2022) Vulnerable exons, like ACADM exon 5, are highly dependent on maintaining a correct balance between splicing enhancers and silencers. *Hum. Mutat.* **43**, 253–265
62. Holm, L. L., Doktor, T. K., Flugt, K. K., Petersen, U. S. S., Petersen, R., and Andresen, B. S. (2024) All exons are not created equal-exon vulnerability determines the effect of exonic mutations on splicing. *Nucleic Acids Res.* **52**, 4588–4603
63. Karjosukarso, D. W., Kieffmann, J. F., Bukkems, F., Duijkers, L., and Collin, R. W. J. (2025) Context matters: the importance of a comprehensive Genomic Region when assessing the therapeutic potential of antisense oligonucleotides in splicing assays. *Nucleic Acid Ther.* **35**, 241–245
64. Lin, J.-H., Wu, H., Zou, W.-B., Masson, E., Fichou, Y., Le Gac, G., *et al.* (2021) Splicing outcomes of 5' splice site GT>GC variants that generate wild-type transcripts differ significantly between full-length and minigene splicing assays. *Front. Genet.* **12**, 701652
65. Drexler, H. L., Choquet, K., Merens, H. E., Tang, P. S., Simpson, J. T., and Churchman, L. S. (2021) Revealing nascent RNA processing dynamics with nano-COP. *Nat. Protoc.* **16**, 1343–1375
66. Li, H. (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100
67. Lopez-Delisle, L., Rabbani, L., Wolff, J., Bhardwaj, V., Backofen, R., Grüning, B., *et al.* (2021) pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422–423
68. Yeo, G., and Burge, C. B. (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **11**, 377–394