

Competition for the conserved branch point sequence influences physiological outcomes in pre-mRNA splicing

Karen Larissa Pereira de Castro¹, Jose M Abril¹, Kuo-Chieh Liao², Haiping Hao³, John Paul Donohue⁴, William K Russell³, William S Fagg^{1,3*}

¹Transplant Division, Department of Surgery, University of Texas Medical Branch, Galveston, United States; ²RNA Genomics and Structure, Genome Institute of Singapore, Agency for Science, Technology, and Research (A*STAR) Singapore, Singapore, Singapore; ³Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, United States; ⁴Sinsheimer Labs, RNA Center for Molecular Biology, Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, Santa Cruz, United States

eLife Assessment

This **important** manuscript provides insights into the competition between Splicing Factor 1 (SF1) and Quaking (QKI) for binding at the ACUAA branch point sequence in a model intron, regulating exon inclusion. The study employs **convincing**, rigorous transcriptomic, proteomic, and reporter assays, with both mammalian cell culture and yeast models.

*For correspondence:

wsfagg@UTMB.EDU

Competing interest: The authors declare that no competing interests exist.

Funding: See page 24

Sent for Review

24 September 2024

Preprint posted

09 October 2024

Reviewed preprint posted

08 November 2024

Reviewed preprint revised

15 December 2025

Version of Record published

20 March 2026

Reviewing Editor: Pablo A Manavella, Universidad Nacional del Litoral-CONICET, Argentina

© Copyright Pereira de Castro et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract Recognition of the intron branch point during spliceosome assembly is a multistep process that can influence mRNA structure and levels. A branch point sequence motif UACUAA is variably conserved in eukaryotic genomes, but in some organisms, more than one protein can recognize it. Here, we show that SF1 and Quaking (QKI) compete for a subset of intron branch sites with the sequence ACUAA in mammalian cells. SF1 activates exon inclusion through this sequence, but QKI represses the inclusion of alternatively spliced exons with this intron branch point sequence. Using mutant reporters derived from a natural intron with two branch site-like sequences, we find that when either branch point sequence is mutated, the other is utilized; however, when both are present, neither is used due to high-affinity binding and strong splicing repression by QKI. QKI occupancy at the dual branch site directly prevents SF1 binding and the subsequent recruitment of spliceosome-associated factors. Finally, ectopic expression of QKI in budding yeast (which lacks QKI) is lethal, at least in part due to widespread splicing repression. In conclusion, QKI can function as a splicing repressor by directly competing with SF1/BBP for a subset of branch point sequences that closely mirror its high-affinity binding site.

Introduction

Up to 95% of human protein-coding genes can be alternatively spliced (Wang et al., 2008; Pan et al., 2008). In contrast, only a minority of *S. cerevisiae* genes contain introns, and under normal growth conditions, most are efficiently spliced (Spingola et al., 1999; Munding et al., 2010). Many influences converge to promote extensive alternative splicing that is observed in higher eukaryotes, including genome complexity, *cis*-acting regulatory elements, and *trans*-acting factors (Fu and Ares,

2014; Ule and Blencowe, 2019; Marasco and Kornblihtt, 2023). An early post-transcriptional step that is required for splicing is branch point (bp) recognition, which, along with 5' splice site (ss), polypyrimidine (pY) tract, and 3'ss recognition, ultimately defines the spliceosome E-complex (Bennett et al., 1992; Seraphin and Rosbash, 1989). Sequence variability in these elements can influence the efficiency with which the E-complex and subsequent splicing complexes mature (Wahl et al., 2009; Zhuang and Weiner, 1986; Michaud and Reed, 1991; Hoskins et al., 2011; Berglund et al., 1997; Zamore et al., 1992). An interesting bp-sequence-specific observation is that different organisms maintain conservation; for example, in *Saccharomyces cerevisiae* (*S. cerevisiae*), it is nearly invariant (UACUAAC), whereas in mammals it is more degenerate (YUNAY) (Moore et al., 1993; Mercer et al., 2015; Zeng et al., 2022; Gould et al., 2016). The underlying functional factors that explain this variability in bp sequence/branch-site (bs) conservation are unknown. This suggests that bs variability might provide a way to control alternative splicing, but how this could be mediated is unclear.

RNA-binding proteins (RBPs) can influence bs recognition by SF1 (mammals) (Berglund et al., 1997; Liu et al., 2001) or MSL5/BBP (*S. cerevisiae*) (Berglund et al., 1997; Jacewicz et al., 2015). When these and other E-complex-associated proteins and snRNAs associate with the substrate, SF1/BBP is evicted and replaced by the SF3B complex, which recruits the 17S U2 snRNP, signaling maturation into the spliceosome A complex (Kastner et al., 2019; Shi, 2017; Martínez-Lumbreras et al., 2024). Disruption of these events by either mutations or non-physiological concentrations of RBPs can lead to splicing defects and disease (Love et al., 2023). For example, the balance between splicing-activating serine-arginine (SR) proteins and splicing-repressing heterogeneous nuclear ribonucleoproteins (hnRNPs) regulates a large set of pre-mRNA substrates and illustrates how competition between RBPs regulates alternative splicing (Busch and Hertel, 2012; Hertel, 2008). However, it is unclear whether RBP competition influences cell type-specific splicing or has evolutionary implications.

The metazoan RBP Signal Transduction and Activator of RNA metabolism (STAR) family members possess KH-type and QUA2 RNA-binding domains and regulate diverse forms of RNA processing, including splicing (Artzt and Wu, 2010). Interestingly, SF1 is the most divergent member of the STAR family, as it lacks the QUA1 dimerization domain that all other members possess (Liu et al., 2001; Ryder and Massi, 2010; Vernet and Artzt, 1997). The QUA1-containing members like Quaking (QKI), KHDRBS1-3, GLD1, and ASD2 exist as dimers and thus bind to a bipartite sequence motif in their target RNAs (Galarneau and Richard, 2005; Galarneau and Richard, 2009). QKI (Teplova et al., 2013; Beuck et al., 2012) and SF1 (Liu et al., 2001) are structurally similar (Figure 1A), and although SF1 lacks the QUA1 dimerization domain, they bind to a similar sequence motif. Interestingly, the SF1 binding motif (Corioni et al., 2011) is more degenerate than the QKI consensus motif (Galarneau and Richard, 2005; Ryder and Williamson, 2004; Hafner et al., 2010; Fagg et al., 2017; Hayakawa-Yano et al., 2017), which is consistent with SF1's role in the recognition of the degenerate mammalian bs.

Analysis of tissue-specific splicing patterns nearly 20 years ago revealed that the UACUAAY motif is associated with muscle-specific exon skipping when located in the intron upstream of alternatively spliced exons or exon inclusion when located in the downstream intron, but the opposite pattern was observed in the brain (Sugnet et al., 2006). It was unclear how the 'conserved bp sequence' UACUAAY might promote exon skipping or inclusion from the downstream intron. Still, it suggested that this element may have an additional and specialized role in tissue-specific alternative splicing. Clarity was provided for this model by the discovery that QKI directly regulates pre-mRNA splicing in C2C12 myoblasts via binding to an ACUAAY element and exhibiting the prototypical 'splicing code' positionally-based regulatory pattern (Fagg et al., 2017; Hall et al., 2013; Barash et al., 2010). This finding is also consistent with the observation that the nuclear isoform QKI5 is expressed at relatively high levels in muscle but at lower levels in the brain (Fagg et al., 2017; Cox et al., 1999; Justice and Hirschi, 2010; Noveroske et al., 2002; Hardy et al., 1996; Lu et al., 2003; Ebersole et al., 1996; Kondo et al., 1999). Thus, QKI5 enforces cell type-specific alternative splicing patterns by binding to a bp sequence that is also a high-affinity QKI binding site, thereby promoting exon skipping.

Studies in lung/lung cells suggested that QKI could repress exon inclusion by competing with SF1 for a bs that might also function as a QKI binding site in the *NUMB* pre-mRNA (Zong et al., 2014). While this raised an intriguing possibility, it was unclear if these two RBPs were in direct competition for binding to a bona fide bs, and how competition for this specialized (but conserved consensus) bs might affect global splicing patterns, modulate splicing in naturally occurring intron substrates, alter

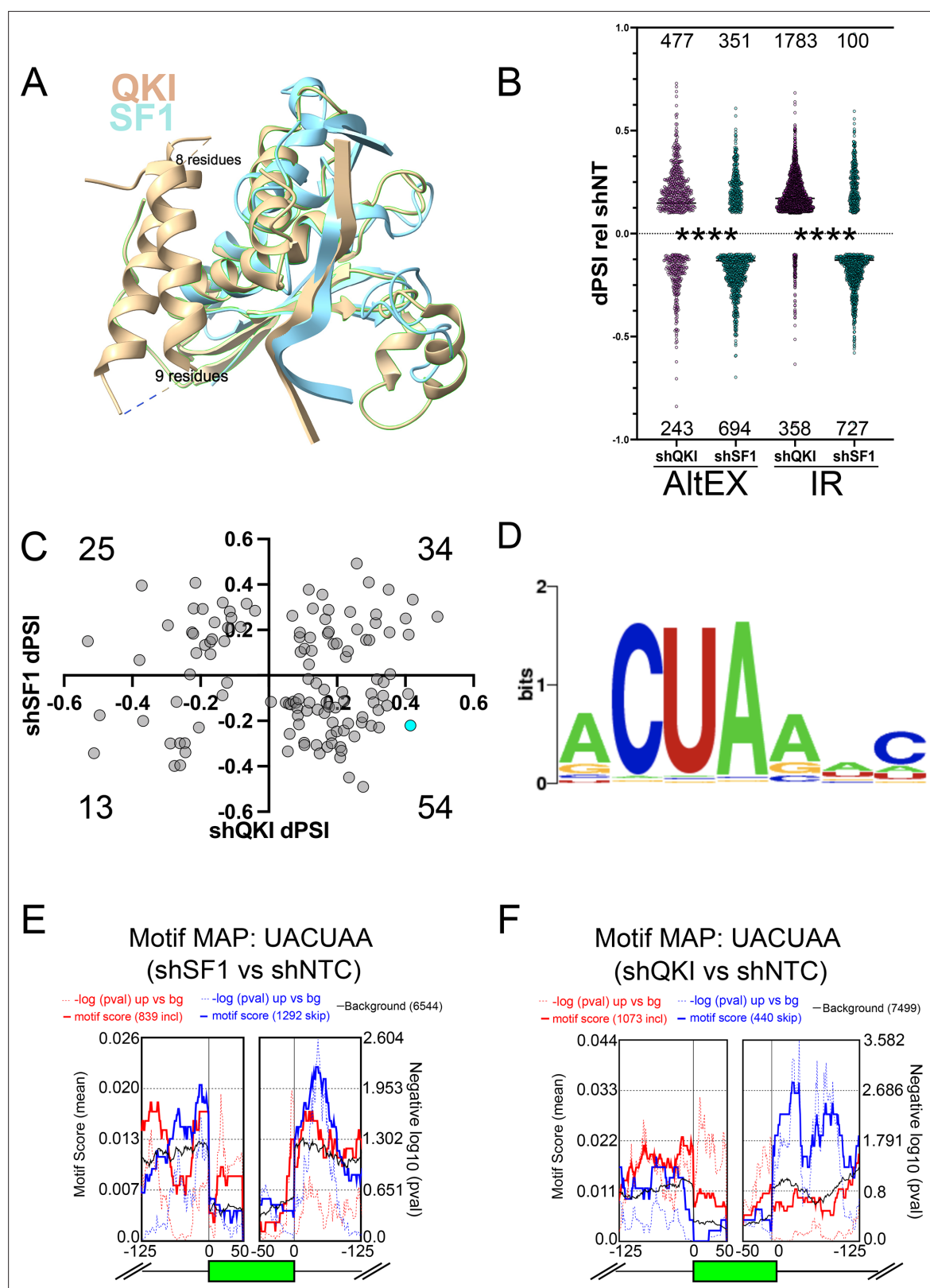


Figure 1. Similarity of Splicing Factor 1 (SF1) and Quaking (QKI) and RNA-seq analysis of SF1 or QKI loss-of-function in HepG2 cells. **(A)** Overlay of SF1 (cyan) and QKI (tan) KH and QUA2 protein domains: green highlighted region indicates KH and QUA2 domains of QKI; the two tan alpha helices on the left side make up the QUA1 domain in QKI (which SF1 lacks; text indicates residues in unstructured regions of QKI). **(B)** Vast-tools analysis of the ENCODE RNA-seq data from SF1 (shSF1) or QKI shRNA (shQKI) knockdown compared to control shRNA (shNT) in HepG2 cells; the y-axis shows dPSI

Figure 1 continued on next page

Figure 1 continued

for shQKI or shSF1 relative to shNT for significantly altered alternatively spliced exons (AltEX; left) or intron retention (IR; right; $dPSI \geq |10|$ and $MVdPSI > 0$), and **** $p < 0.0001$ by Mann-Whitney U when comparing the distribution of changes in shQKI relative to shNT to shSF1 relative to shNT. (C) Scatter plot showing the distribution of AltEX events that changed under depletion of both SF1 ($dPSI$ values relative to shNT on the y-axis) and QKI ($dPSI$ values relative to shNT on the x-axis); the number in each quadrant indicates how many AltEX events were observed; RAI14 exon 11 is shown in cyan. (D) Simple Enrichment Analysis (SEA) of the intron region spanning 60 nt to 20 nt upstream of the 3' splice site in QKI and SF1 regulated AltEX events shown in C ($p < 0.05$). (E) rMAPS motif map for UACUAA generated for AltEX events changing during shSF1 compared to shNT by rMATS; motif scores (solid line) or $-\log_{10}$ p-value (dotted line) is shown in red for exons whose inclusion increases and in blue for exons whose inclusion decreases. (F) rMAPS motif map as described in E. but for shQKI treatment compared to shNT.

spliceosome component recruitment, or influence evolutionary outcomes in splicing. Furthermore, SF1 knockdown in HeLa cells indicates it is required for viability and the splicing of several introns (Tanackovic and Krämer, 2005), but the extent to which its loss affects global splicing patterns is unknown. Here, we test these open questions and find that QKI5 can repress a set of alternatively spliced exons whose bs mirror the high-affinity QKI binding site. We use RAI14 exon 11 as a model and show that ACUAA elements in intron 10 are true bs to which QKI or SF1 can bind to promote exon skipping or inclusion, respectively. Interestingly, QKI binding to RAI14 intron 10 bs leads to the recruitment of paraspeckle-associated proteins, while SF1 binding causes subsequent enrichment of the SF3a/b and U2 17 S complexes. These discoveries expand the scope and clarify the bp competition model. Moreover, the addition of QKI5 into *S. cerevisiae* is lethal and concomitant with widespread splicing repression. Together, these findings suggest that the presence of QKI and degenerate bs may have coevolved to expand the repertoire of cell type-specific alternative splicing in plants and multicellular eukaryotes.

Results

QKI and SF1 co-regulate a set of alternatively spliced exons through a distinct sequence motif

Based on analysis of their individual functions (Berglund et al., 1997; Liu et al., 2001; Corioni et al., 2011; Hall et al., 2013; de Bruin et al., 2016; van der Veer et al., 2013; Darbelli et al., 2016; Selenko et al., 2003; Peled-Zehavi et al., 2001) and common structural features (Figure 1A), we hypothesized that QKI could repress the inclusion, while SF1 could activate the inclusion of a special subset of alternatively spliced exons that have ACUAA for their bs. To test this, we analyzed RNA sequencing (RNA-seq) datasets during QKI or SF1 knockdown (HepG2 cells treated with control non-targeting shRNA or shRNAs targeting QKI (shQKI) or SF1 (shSF1) Van Nostrand et al., 2020a) with VAST-tools to measure changes in alternative pre-mRNA splicing (Irimia et al., 2014; Tapial et al., 2017). Consistent with potentially opposing functions in splicing, more exon inclusion and intron retention (IR) events are observed during QKI knockdown, but more exon skipping and fewer IR events are observed during SF1 knockdown; the distribution of these changes are significantly different when compared to one another (**** $p < 0.0001$ by Mann-Whitney U; Figure 1B, and Supplementary file 1; significance cutoff: change in percent spliced in ($dPSI$) $> |10|$ and minimum value $dPSI$ at 95% confidence interval ($MVdPSI_{95}$) > 0). To measure the distribution of exons regulated by both QKI and SF1 (co-regulated exons), we plotted $dPSI$ values of cassette exons under each condition relative to control for those that change under both knockdown conditions ($dPSI > |10|$ in either shQKI or shSF1 relative to control, and $MVdPSI_{95} > 0$ in both shQKI relative to control and SF1 relative to control). Interestingly, we observed that 54 of the 126 (43%) co-regulated alternatively spliced exons fell into the category in which inclusion increases upon QKI knockdown and decreases upon SF1 knockdown (Figure 1C). Thus, the most highly represented set of these (by quadrant) is QKI-repressed and SF1-activated. We next asked if these co-regulated alternatively spliced exons are associated with any significantly enriched sequences in the upstream intron region in which the bs is found. To do so, we obtained the sequence from 63 to 20 nucleotides upstream of the 3' splice site (these should include putative bs) but exclude pY tracts and 3' splice sites of these co-regulated exons or 1000 control exons from transcripts that are expressed (base mean > 100) but whose splicing is unchanged upon either QKI or SF1 knockdown ($dPSI < |1|$, $MVdPSI = 0$) and used Simple Enrichment Analysis (SEA Bailey and Grant, 2021) to test if any motifs are significantly enriched. This analysis revealed

enrichment of a single ACUAA-like motif in the co-regulated exons ($p=0.033$), but failed to identify any significantly enriched motifs in the control sequences (**Figure 1D**). Interestingly, this motif could serve as a QKI (*Galarneau and Richard, 2005; Ryder and Williamson, 2004; Hafner et al., 2010*) or SF1 (*Corioni et al., 2011; Van Nostrand et al., 2020a*) binding site, or as a bs (*Mercer et al., 2015; Zeng et al., 2022*). Finally, splicing analysis of these RNA-seq datasets using the complementary method rMATS (*Shen et al., 2014*) corroborated these findings by indicating more exon inclusion when QKI is knocked down and more skipping when SF1 is reduced (**Figure 1E and F; Supplementary files 2 and 3**, respectively). Motif enrichment analysis using rMAPS2 (*Park et al., 2016; Hwang et al., 2020*) shows significant enrichment of the UACUAA motif in introns upstream of exons that are skipped more when SF1 is reduced compared to control (**Figure 1E**), and enrichment of this motif in introns upstream of exons is more included when QKI is reduced compared to control (**Figure 1F**). In summary, the most common set of exons co-regulated by SF1 and QKI is SF1-activated and QKI-repressed, and these have enrichment of a bs-like sequence in their proximal upstream intron that also appears to be a QKI binding motif.

RAI14 exon 11 is QKI-repressed and SF1-activated, with dual putative ACUAAC branch sites

To further investigate how competition between QKI and SF1 for the bs is mediated, we sought a prototypical alternatively spliced exon subject to this form of regulation. The criteria by which we narrowed our search were (1) a QKI-repressed and SF1-activated cassette exon with, (2) experimental evidence of a functional ACUAAY bs, and (3) experimental evidence of direct QKI binding. We previously found that QKI5 represses *Rai14* exon 11 and that reads from QKI iCLIP-seq map to the intron immediately upstream of this exon in mouse myoblasts (*Fagg et al., 2017*). Its inclusion is also repressed by QKI and activated by SF1 in HepG2 cells (**Figure 1C** (cyan dot) and **Figure 2A; Van Nostrand et al., 2020a**). Inspection of the upstream intron region proximal to *RAI14* exon 11 revealed intriguing features that fulfilled the above criteria: tandem ACUAAC elements 34 nt or 43 nt upstream of the 3'ss, experimental evidence indicating that either of these could be used as a bs (*Zeng et al., 2022*), and QKI eCLIP peaks from experiments using HepG2 or K562 cells (*Van Nostrand et al., 2020a; Van Nostrand et al., 2020b*) that overlap with these putative ACUAAC bs sequences (**Figure 2A**). We next tested if the co-regulation of *RAI14* exon 11 splicing by QKI and SF1 observed in HepG2 cells could be observed in additional cell types. We measured *RAI14* exon 11 inclusion in HEK293 QKI KO cells and found a significant increase in its inclusion in the KO cells compared to WT ($p<0.0001$ by Student's t-test; **Figure 2B**). To reduce SF1 levels, we used two independent siRNAs targeting it in WT HEK293 cells. The first failed to produce an appreciable knockdown, but the second, or the two combined, reduced the SF1 protein level, leading to greater *RAI14* exon 11 skipping (**Figure 2C**). Overexpression of myc:QKI5 but not a myc:QKI5 construct with reduced RNA binding activity (K120A;R124A *Teplova et al., 2013*) promotes skipping of *RAI14* exon 11 in WT HEK293 cells, and overexpressing SF1 causes more inclusion (**Figure 2—figure supplement 1**). Although the physiological level of *Rai14* exon 11 inclusion is low in mouse myoblasts (perhaps due to the high level of QKI5 in these cells), knocking down QKI5 or SF1 in C2C12 cells leads to a significant increase or decrease, respectively, in its inclusion compared to control ($p<0.0001$ by Student's t-test; **Figure 2D**). Therefore, *RAI14* exon 11 splicing is repressed by QKI and activated by SF1 in various cell types, possibly by binding dual ACUAAC bs elements.

Efficient RAI14 exon 11 skipping requires tandem ACUAAC elements, either of which can be used as a bp

Next, we asked how the ACUAAC elements in *RAI14* intron 10 influence exon 11 alternative splicing. Previous investigations did not stringently discriminate between whether QKI and SF1 competed directly for binding to the bs or if dimeric QKI bound to a bipartite motif that flanked the bs and thus occluded SF1 binding to the true bp (*Zong et al., 2014*). We hypothesized that: (1) efficient exon 11 skipping/splicing repression requires two intact ACUAAC elements (or at least a single ACUAAC element and QKI 'half-site' (which would constitute the previously SELEX-defined high affinity Quaking response element (QRE)) *Galarneau and Richard, 2005*), (2) one ACUAAC element or the other must be present for any exon inclusion (either can be a bs), (3) loss of both ACUAAC elements would lead to no exon inclusion (one or the other is required to have a bs), (4) conversion

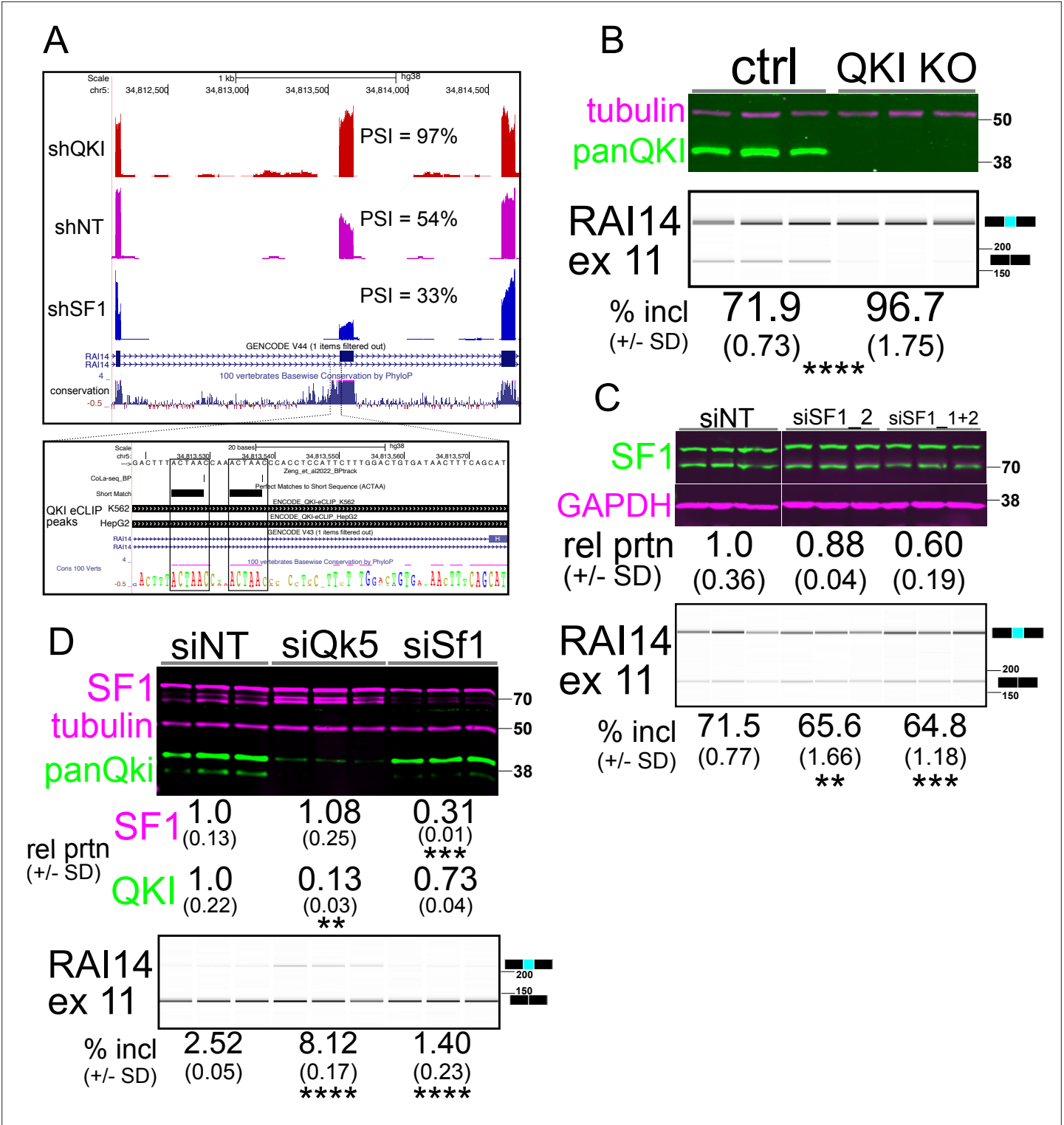


Figure 2. RAI14 exon 11 is repressed by QKI and activated by Splicing Factor 1 (SF1). **(A)** UCSC Genome Browser screenshot (top panel) with RNA-seq reads mapping to RAI14 for shQKI (top), shNT (middle), or shSF1 (bottom); percent spliced in (PSI) values are measured by Vast-tools; the inset shows boxed regions of (from top) two ACUAAC elements with Cola-seq branchpoints mapping to one nucleotide downstream of a branchpoint adenine, ACUAA elements by oligomatch, then Quaking (QKI) eCLIP peaks from K562 (top) or HepG2 (bottom) cells that overlap with these; conservation of 100 vertebrates is shown at the bottom. **(B)** Western blot of protein extracted from HEK293 QKI KO cells (top) probed with anti-tubulin (magenta) or anti-panQKI (green) antibodies (MW in kD is shown on right); RT-PCR of RAI14 exon 11 and BioAnalyzer gel-like image (bottom) of RNA extracted from the cells above showing mean percent included $\pm$ standard deviation below ($n=3$ biological replicates; **** $p<0.001$ by Student's t-test compared to ctrl; Figure 2 continued on next page

Figure 2 continued

size of amplicon in bp is shown on right). (C) Western blot (top) of proteins extracted from HEK293 cells transfected with siNT, siSF1_2, or siSF1_1+2, probed with anti-SF1 (green) or anti-Gapdh (magenta) antibodies; the protein abundance (fold change relative to the siNT control $\pm$ standard deviation) is shown below and MW in kD is shown on right; RT-PCR of *RAI14* exon 11 and BioAnalyzer gel-like image (bottom) of RNA extracted from the cells described above with mean percent included $\pm$ standard deviation below ($n=3$ biological replicates; $**p<0.01$ or $***p<0.001$ by Student's t-test compared to siNT; size of amplicon in bp is shown on right). (D) Western blot (top) of proteins extracted from C2C12 myoblasts transfected with siNT, siQki or siSF1_1+2, probed with anti-SF1 (magenta; top), anti-tubulin (magenta; middle) or anti-panQki (green; bottom); MW in kD is shown on right. The protein abundance (fold change relative to the siNT control $\pm$ standard deviation) is shown below ($n=3$ biological replicates; $**p<0.01$ or $***p<0.001$ by Student's t-test); RT-PCR of *Rai14* exon 11 and BioAnalyzer gel-like image (bottom) of RNA extracted from the C2C12 cells described above with mean percent included $\pm$ standard deviation indicated below ($****p<0.0001$ by Student's t-test; size of amplicon in bp is shown on right).

The online version of this article includes the following source data and figure supplement(s) for figure 2:

Source data 1. Original membranes and BioAnalyzer gel-like images corresponding to **Figure 2A** (left), **Figure 2C** (middle), or **Figure 2D** (right).

Source data 2. Source data containing original uncropped western blots and BioAnalyzer gel-like images/files.

Figure supplement 1. Western blot and RT-PCR of proteins and RNA extracted from wild-type (WT) HEK293 cells transfected with tdTomato, WT myc:Qki5, MT myc:Qki5, and Splicing Factor 1 (SF1).

Figure supplement 1—source data 1. Original uncropped membrane from western blot and uncropped gel-like image from BioAnalyzer shown in **Figure 2—figure supplement 1**.

Figure supplement 1—source data 2. Original uncropped membrane from western blot and uncropped gel-like image from BioAnalyzer shown in **Figure 2—figure supplement 1**.

of ACUAAC elements to another (non-QKI binding motif) bs would lead to more inclusion; regarding RNA stability: (5) removal of either ACUAAC element could destabilize the transcript, but (6) restoring either to a UAAC 'half-site' could rescue this defect and promote exon skipping. To test these, we generated a splicing reporter by cloning 243 nucleotides of *RAI14* intron 10, *RAI14* exon 11, and 95 nucleotides of *RAI14* intron 11 into the beta globin splicing reporter pDUP51 (DUP-*RAI14* exon 11). We first deleted either the first, the second, or both ACUAAC elements (**Figure 3A**), transfected these reporter plasmids into C2C12 cells, and then measured exon inclusion by RT-PCR or total transcript stability by RT-qPCR. Deletion of either the upstream or downstream ACUAAC element causes a ~40% increase in inclusion (**Figure 3B**), indicating that both sites are required for strong splicing repression. Interestingly, we observed the appearance of a mis-spliced product from the upstream deletion mutant compared to WT (**Figure 3B**); the latter may be due to the use of a weak/cryptic bp and 3'ss (see * on **Figure 3B**). Essentially, no inclusion of *RAI14* exon 11 is observed in the absence of both ACUAAC elements, except for the mis-spliced product noted above (**Figure 3B**). Control experiments show that these amplicons are reverse transcriptase-dependent, indicating that the products detected are not due to plasmid DNA or PCR product contamination and corroborate the results described above (**Figure 3—figure supplement 1**). Moreover, loss of the upstream ACUAAC element results in markedly reduced total reporter RNA levels; we observe a similar trend in the other deletion mutants, though to a lesser extent (**Figure 3C**). Therefore, potent splicing repression of *RAI14* exon 11 requires both ACUAAC elements, and one or the other is required for any exon inclusion, and thus they are bona fide bs.

We next asked whether substituting either ACUAAC element, or both, to ACUGAC (which should be a poor substrate for QKI binding but a suitable bp sequence, **Figure 3A**) would also lead to increased inclusion of *RAI14* exon 11 or influence reporter RNA stability. We observe a large increase in exon inclusion and a decrease in skipping for either of the single substitution mutants and for the double substitution mutant (~20–35%; **Figure 3D**). As above, the amplification of these products is RT-dependent (**Figure 3—figure supplement 1**). Transcript abundance is significantly reduced in either single substitution mutant ($****p<0.0001$ by Student's t-test) and is slightly lower in the double substitution mutant (**Figure 3E**). Finally, converting either ACUAAC element to a UAAC 'half-site' results in comparable increases in inclusion/decreases in skipping compared to WT as described in the deletion or substitution mutants (**Figure 3F**), which are also RT-dependent (**Figure 3—figure supplement 1**). Interestingly, these 'half-site' mutants show a higher level of abundance than the WT reporter, especially in the upstream 'half-site' mutant (**Figure 3G**). In summary, removal of either ACUAAC element or conversion to ACUGAC leads to splicing activation of *RAI14* exon 11, indicating a requirement for dual ACUAAC elements for potent splicing repression.

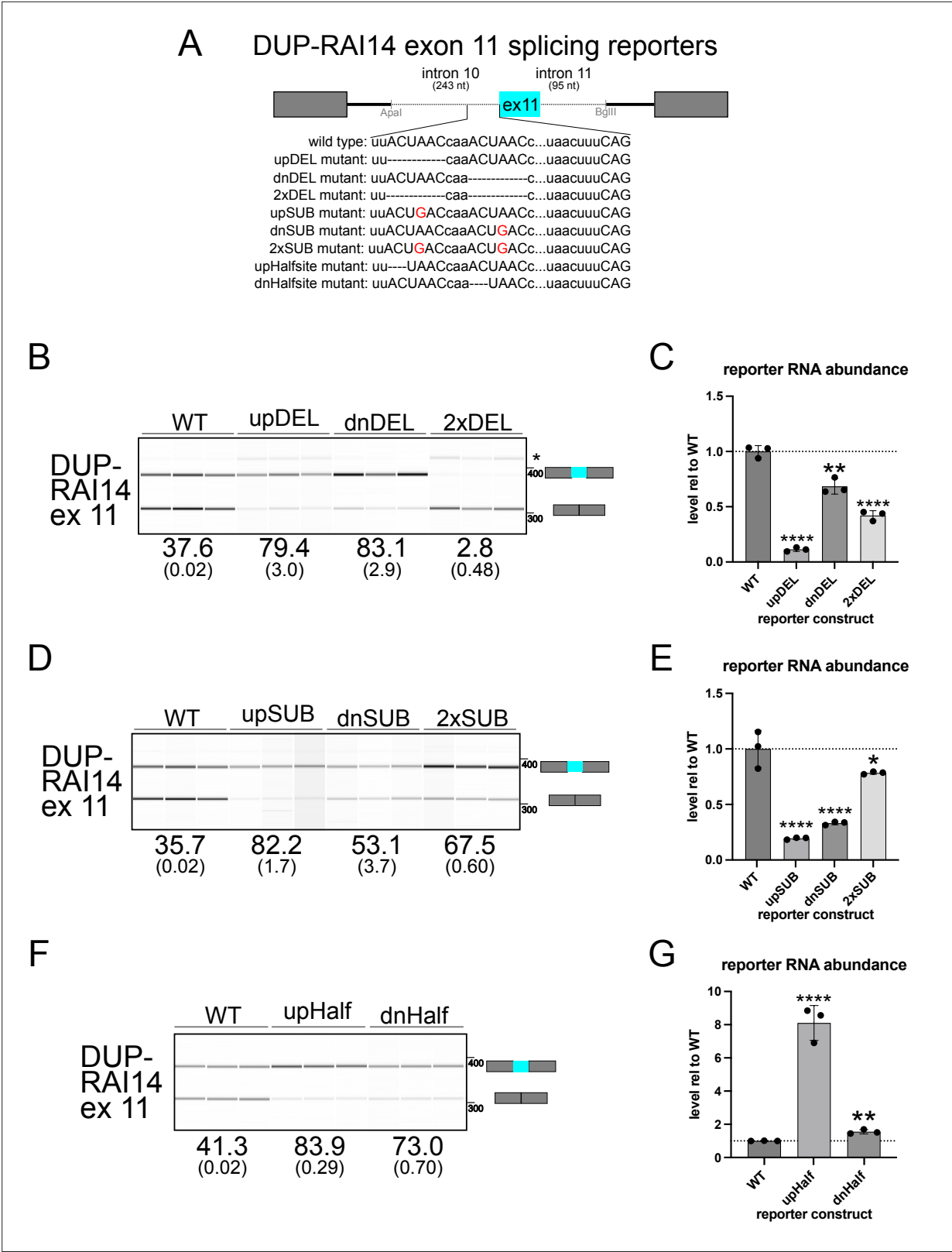


Figure 3. Analysis of the DUP-RAI14 exon 11 (ex 11) splicing reporter. **(A)** Schematic of the beta globin pDUP-RAI14 ex 11 splicing reporters indicating the region of intron 10, exon 11, and intron 11 included; the inset below shows the different constructs (wild-type or mutant) tested. **(B)** RT-PCR and BioAnalyzer gel-like image from RNA extracted from C2C12 cells transfected with *RAI14* ex 11 wild-type or deletion mutant reporters; mean percent included values are shown below (+/- the standard deviation; *indicates an unidentified/spurious product). **(C)** RT-qPCR measuring total reporter RNA

Figure 3 continued on next page

Figure 3 continued

level, normalized to *Eef1a1*, from RNA described in B., and shown as fold-change relative to WT (** $p < 0.01$ or **** $p < 0.0001$ by Student's t-test). (D) RT-PCR and BioAnalyzer gel-like image from RNA extracted from C2C12 cells transfected with *RAI14* ex 11 wild-type or substitution mutant reporters, analyzed as described in B. (E) RT-qPCR measuring total reporter RNA level, normalized to *Eef1a1*, from RNA described in D., and shown as fold-change relative to WT (* $p < 0.05$ or **** $p < 0.0001$ by Student's t-test). (F) RT-PCR and BioAnalyzer gel-like image from RNA extracted from C2C12 cells transfected with *RAI14* ex 11 wild-type or half-site mutant reporters, analyzed as described in B (*** $p < 0.001$, **** $p < 0.0001$). G. RT-qPCR measuring total reporter RNA level, normalized to *Eef1a1*, from RNA described in F., and shown as fold-change relative to WT (** $p < 0.01$ or **** $p < 0.0001$ by Student's t-test). Each experiment was conducted in biological triplicate, and for RT-PCR with BioAnalyzer measurement, each comparison compared to WT reporter showed $p < 0.0001$ by Student's t-test.

The online version of this article includes the following source data and figure supplement(s) for figure 3:

Source data 1. Original BioAnalyzer gel-like images corresponding to **Figure 3B** (left), **Figure 3D** (middle), or **Figure 3F** (right).

Source data 2. Original BioAnalyzer gel-like images and data files corresponding to **Figure 3B** (left), **Figure 3D** (middle), or **Figure 3F** (right).

Figure supplement 1. Agarose gel showing PCR amplification in the absence (-) or presence (+) of reverse transcriptase.

Figure supplement 1—source data 1. Original uncropped agarose gel images from **Figure 3—figure supplement 1**.

Figure supplement 1—source data 2. Original uncropped agarose gel images from **Figure 3—figure supplement 1**.

QKI binding to *Rai14* intron 10 requires both ACUAAC elements and prevents spliceosome recruitment

We next hypothesized that QKI binding to *Rai14* intron 10 requires both ACUAAC elements, that its binding would prevent spliceosome recruitment by blocking SF1, and that removal of either ACUAAC element would favor SF1 binding and recruitment of spliceosome components. To minimize bias, we initially performed RNA affinity chromatography (RAC) using 64 nt of intron sequence upstream of the 3'ss and including 6 nt of exonic sequence linked to a tobramycin aptamer (WT) (Hartmuth et al., 2002), and the same sequence but with either the upstream ACUAAC (upDEL), downstream ACUAAC (dnDEL), or both ACUAACs (2xDEL), as well as a tobramycin aptamer (APT) only RNA (Figure 4A). Our rationale for this was to identify all proteins bound to these substrates and correlate them with splicing outcomes: WT shows low levels of inclusion, upDEL slightly higher levels of inclusion, dnDEL high levels of inclusion, and 2xDEL lacks a bp and so is completely skipped. We added C2C12 myoblast nuclear extract (NE) to these under conditions that favor splicing (with ATP), then collected the associated proteins and identified them by liquid chromatography with tandem mass spectrometry (LC-MS/MS) (Hartmuth et al., 2002). Quaking is readily identified as associating with the WT RAC substrate and in the input NE but is undetectable in each of the mutant RAC substrates and the APT-only control; SF1 is not detected associating with any RAC substrate but was present in NE (Figure 4B and C, and Supplementary file 4). We used previously published LC-MS/MS datasets to generate a list of early spliceosome (E complex) and 17S U2 snRNP components (E/U2) (Makarov et al., 2012; Sharma et al., 2008; Cvitkovic and Jurica, 2013) in order to focus on these proteins in our RAC-LC-MS/MS datasets to test our hypotheses. Subsequently, we found that the proteomic profiles associated with the WT and 2xDEL substrates are most similar (Figure 4B), suggesting that QKI protein binding or loss of bs leads to similar E/U2 protein binding patterns. Indeed, in either of the single deletion mutants, we observe more enrichment of E/U2 components with the substrate RNA (24 positive and 5 negative values in upDEL and 16 positive values and seven negative values in dnDEL) while the WT and 2xDEL substrates show a more balanced distribution (18 positive and 12 negative or 18 positive and 11 negative values, respectively; Figure 4C). Western blot analysis using the same RAC approach also shows robust QKI association with the WT substrate and undetectable or nearly undetectable SF1 protein, validating our LC-MS/MS findings (Figure 4D).

Previous studies indicate that excluding ATP from RNA splicing/binding assays favors a more stable association of early splicing complexes, including SF1 (Michaud and Reed, 1991; Reed, 1990; Jamison et al., 1992). Therefore, we performed RAC-LC-MS/MS with the same substrates in the absence of ATP (Supplementary file 5). We also employed a data-independent acquisition (DIA) LC-MS/MS method, utilizing NE to construct a peptide search library, thereby enhancing the specificity and sensitivity of detection (see Methods). We observe significant enrichment ($\log_2$ fold change $> |0.2|$ and $p < 0.01$) of E/U2 components (including SF1) binding to the RNA in the absence of the downstream ACUAAC (dnDEL) element compared to WT RNA (Figure 4E). This trend is also observed but to a lesser extent in the absence of the upstream ACUAAC (upDEL) and in the absence of both

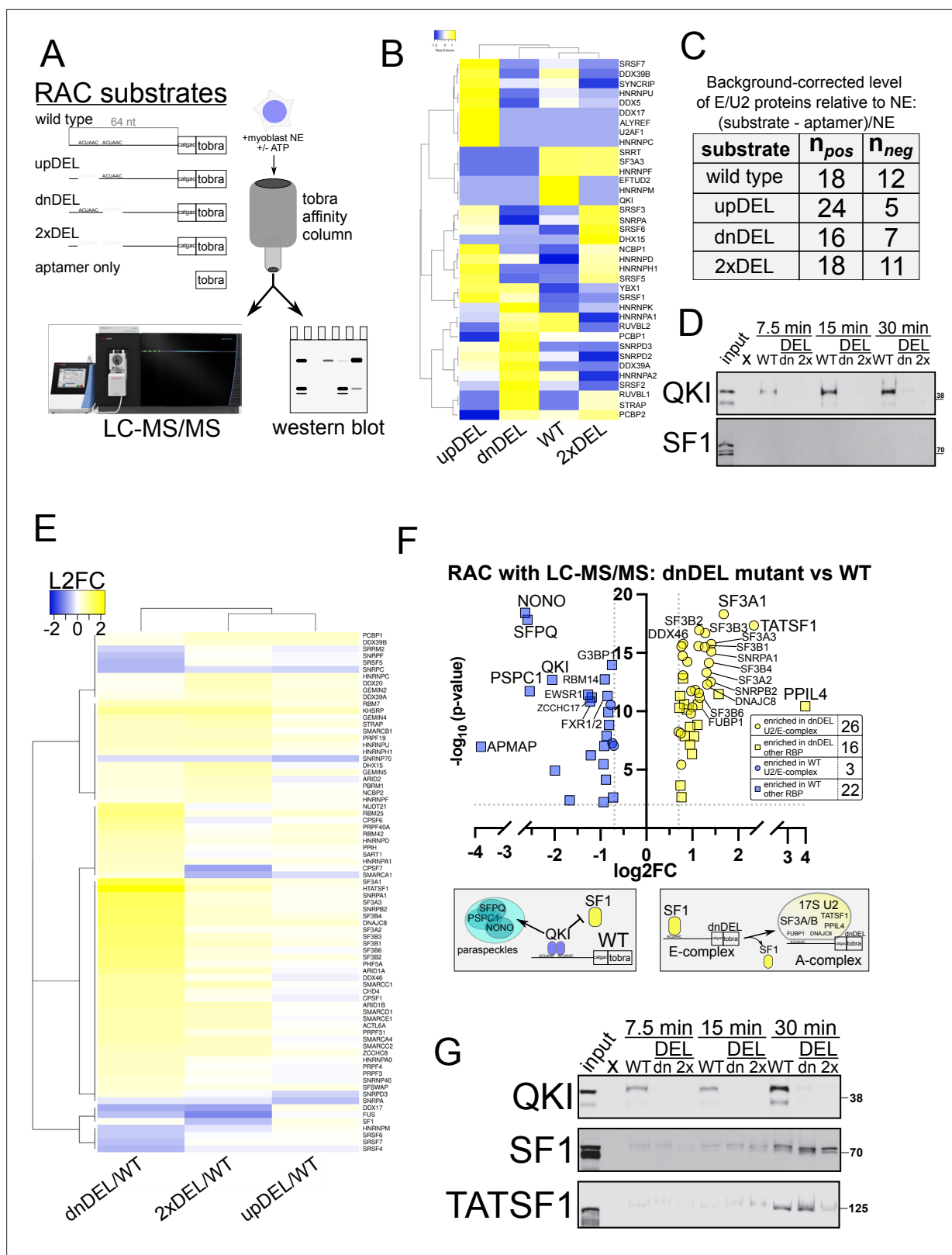


Figure 4. Analysis of the proteins associating with *RAI14* intron 10 RNA. RNA affinity chromatography (RAC) with liquid chromatography and tandem mass spec or western blot analysis. (A) Schematic representation of substrates used for RNA affinity chromatography (RAC) which included 64 nt of *RAI14* intron 10, 6 nt of exon sequence, and the tobramycin aptamer (tobra): wild-type, upDEL, dnDEL, 2xDEL, and aptamer only; C2C12 nuclear extract (NE) was incubated with these, and RAC-associated eluates were analyzed by liquid chromatography with tandem mass spectrometry (LC-MS/MS) or

Figure 4 continued on next page

Figure 4 continued

western blot. **(B)** Heatmap showing hierarchical clustering of early spliceosome and 17S U2 snRNP protein (E/U2) abundance detected in the RAC-LC-MS/MS datasets for each substate shown and in the presence of ATP; these represent background-corrected levels relative to NE (see Methods) and the scale bar shows row Z-score values. **(C)** Numbers observed for relative levels of E/U2 proteins (background corrected relative to NE) observed associating with each RAC substate with either a positive (n_{pos}) or negative (n_{neg}) value. **(D)** Western blot of NE (input) or WT, dnDEL (dn), or 2xDEL (2 x) RAC time-course (+ATP as described in B and C) for 7.5 min (left), 15 min (middle), or 30 min (right) probed with anti-panQki (top) or anti-SF1 (bottom) antibodies; MW shown in kD to the right. **(E)** Heatmap showing hierarchical clustering of early spliceosome and 17S U2 snRNP protein (E/U2) abundance detected in the RAC-LC-MS/MS datasets for each substate shown and in the absence of ATP; these represent data-independent acquisition (DIA; see methods) values normalized to NE and each mutant is shown as log₂ fold change relative to the WT substate and passed cutoff of log₂ fold change > |0.2| and $p < 0.01$. **(F)** Volcano plot comparing LC-MS/MS log₂ protein abundance (log₂ fold-change (log₂FC); x-axis) of E/U2 (circles) and other RNA-binding proteins (RBPs) (squares) observed associating with RAC substates dnDEL compared to WT (y-axis shows $-\log_{10} p$ -value) of enriched proteins (cutoff: L2FC > |0.7| and $p < 0.01$); inset shows the number observed for those enriched in dnDEL (yellow) or WT (blue); schematic below shows model of RAC substates recruitment to distinct protein-associated species. **(G)** Western blot of NE (input) or WT, dnDEL (dn), or 2xDEL (2 x) RAC time-course (-ATP as described in E and F) for 7.5 min (left), 15 min (middle), or 30 min (right) probed with anti-panQki (top), anti-SF1 (middle), or anti-TATSf1 (bottom); MW shown in kD to the right.

The online version of this article includes the following source data for figure 4:

Source data 1. Original membranes corresponding to western blots depicted in **Figure 4D** (left), or **Figure 4G**.

Source data 2. Original membranes corresponding to western blots depicted in **Figure 4D** (left), or **Figure 4G**.

(2xDEL) mutants relative to WT; the proteome profiles associated with these two substrates are also more similar to one another than the dnDEL comparison to WT, consistent with lower splicing efficiency or no splicing, respectively (**Figure 4E**). Closer and more stringent (log₂ fold change $\geq |0.7|$ and $p < 0.01$) inspection of the proteins binding preferentially to the dnDEL mutant RNA compared to the WT RNA reveals enrichment of 26 E/U2 components along with 16 other annotated (non-E/U2) RBPs; we observe depletion (or enrichment in WT) of only three E/U2 proteins but of 22 other annotated RBPs, including QKI (**Figure 4F**). Interestingly, some of the most enriched protein components associating with the dnDEL RNA are TATSf1, Sf3A1, PPIL4, Sf3A3, Sf3B4, DNAJC8, Sf3A2, Sf3B3, Sf3B1, and FUBP1 (**Figure 4F**), many of which are known to interact with the bs and SF1 or are members of the Sf3 complex, which promotes spliceosome maturation by displacing SF1 from the bp and then recruiting the U2 snRNP (*Martínez-Lumbreras et al., 2024; Ebersberger et al., 2023; Schmitzová et al., 2023; Zhan et al., 2024*). The most significantly enriched proteins associating with the WT RAC substate are NONO, SFPO, PSPC1, and QKI; the former three are found in paraspeckles that are formed by liquid-liquid phase separation (*Knott et al., 2016*), and QKI has recently been identified as a paraspeckle component protein (*Dyakov et al., 2024*). Western blot analysis validates the above findings indicating that QKI associates with high affinity to the WT but not dnDEL or 2xDEL RNA, and, in the absence of ATP, SF1 and TATSf1 associate to a greater degree with the dnDEL mutant than either of the other RAC substates (**Figure 4G**). Therefore, both ACUAAC elements are required for QKI binding and repression of Rai14 exon 11 splicing; removing the downstream ACUAAC element causes SF1 binding, subsequent recruitment of the spliceosome A complex machinery, and derepression of Rai14 exon 11 splicing.

Ectopic expression of QKI5 in *S. cerevisiae* is lethal and causes pre-mRNA splicing defects

Our results indicate that QKI and SF1 can directly compete for a subset of ACUAA bp sequences, so we hypothesized that expressing the QKI5 isoform in *S. cerevisiae* (where the bp sequence is nearly invariant UACUAAC and lacks QKI) would be lethal and cause defective splicing. To test this, we inserted a galactose-inducible cDNA encoding either EGFP, mutant K120A;R124A QKI5, or wild-type QKI5 into BY4741 at the *URA3* locus, and each grew normally on glucose-containing media where the transgene is repressed. Galactose induction is tolerated for the EGFP- and mutant QKI5-containing yeast but is lethal for the wild-type QKI5-containing strain (**Figure 5A**). Growth curve analysis indicates that after about 4 hr, the QKI5-expressing BY4741 cells cease proliferation (**Figure 5—figure supplement 1A**). Interestingly, by 24 hr most of these cells have a ‘large-budded’ morphology, which suggests a defect in cell division; EGFP- and mutant QKI5-expressing cells grow normally (**Figure 5—figure supplement 1A and B**).

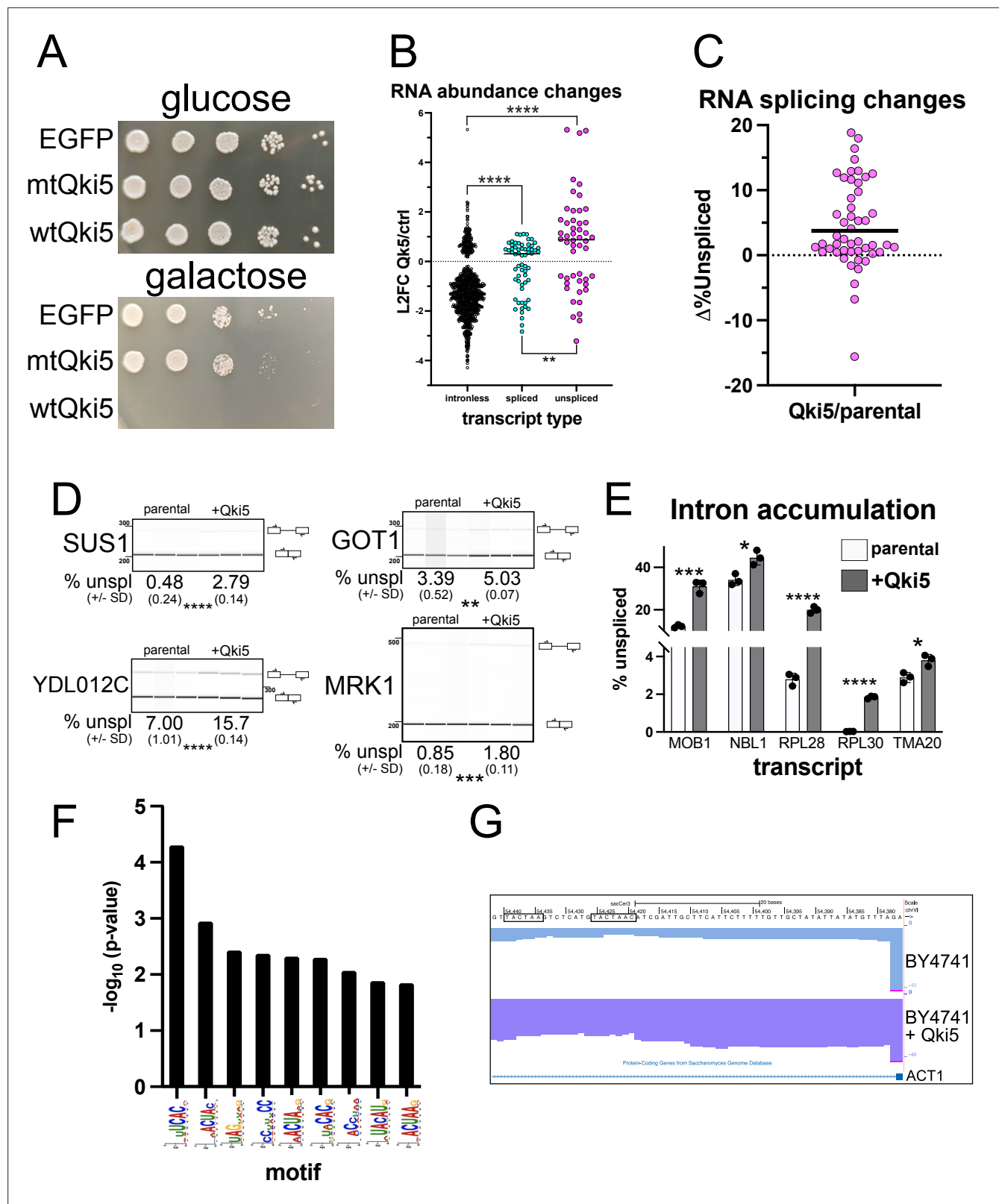


Figure 5. *QKI5* expression is lethal in yeast and represses splicing. **(A)** BY4741 *S. cerevisiae* strain with *EGFP* (top), mutant *QKI5* (mtQKI5; middle), or wild-type *QKI5* (wtQKI5; bottom) GAL-inducible transgene cultures grown on either glucose- (top) or galactose-containing (bottom) YPD plates at decreasing densities (left to right) and incubated at 30 °C for 72 hr. **(B)** Changes in RNA abundance measured by DESeq2 analysis of RNA-seq data for intronless, spliced, or unspliced transcripts were measured from RNA extracted from BY4741 cells with *QKI5* expression (induced by galactose-

Figure 5 continued on next page

Figure 5 continued

containing media for 4 hr), or in the BY4147 parental strain also cultured in galactose-containing media for 4 hr (control), and are shown as L2FC in *QKI5*-induced cells relative to control ($n=3$ biological replicates; cutoff $p<0.1$; abundance cutoff of TPM >0.2 ; $**p<0.01$ or $****p<0.0001$). (C) Splicing changes measuring the change in percent unspliced ($\Delta\%$ Unspliced; y-axis) for BY4741 as described in B.; cutoff $p<0.1$ by Student's t-test and base mean >100 . (D) RT-PCR with primers that span exon-intron-exon junctions and BioAnalyzer gel-like image showing mean percent unspliced ($\pm$ SD; $n=3$) below for introns whose inclusion increased upon *QKI5* ectopic expression and as measured by RNA-seq analysis in C. for the parental or *QKI5*-expressing cells ($**p<0.01$, $***p<0.001$, $****p<0.0001$ by Student's t-test); amplicon size in bp is shown to the left for each except YDL012C, which is shown to the right. (E) RT-qPCR analysis measuring mean percent unspliced transcript from RNA extracted from biological triplicate cultures of either the parental control or BY4741 expressing *QKI5* for each transcript shown ($\pm$ SD; $*p<0.05$, $***p<0.001$, $****p<0.0001$). (F) Bar graph showing $-\log_{10}$ p-values (y-axis) of significantly enriched (SEA; $p<0.01$) motifs observed in 3' proximal ends of introns whose inclusion increases upon ectopic *QKI5* expression in yeast (x-axis). (G) UCSC Genome Browser screenshot showing RNA-seq reads mapping to the *ACT1* transcript intron/exon junction near the 3'ss from RNA extracted from parental control and *QKI5*-induced cells; boxed sequences show two TACTAA elements.

The online version of this article includes the following source data and figure supplement(s) for figure 5:

Source data 1. Original gel-like images from BioAnalyzer used in **Figure 5D**.

Source data 2. Original gel-like images and data file from BioAnalyzer used in **Figure 5D**.

Figure supplement 1. *QKI5* expression is lethal in yeast and represses splicing.

Figure supplement 1—source data 1. Original uncropped agarose gel images of data shown in **Figure 5—figure supplement 1C**.

Figure supplement 1—source data 2. Original uncropped agarose gel images of data shown in **Figure 5—figure supplement 1C**.

Next, we tested how ectopic *QKI5* expression impacts the yeast transcriptome. The parental BY4741 strain or BY4741 with WT *QKI5* was grown in galactose-containing media for 4 hr, and then RNA was collected and sequenced. We used a *S. cerevisiae* genome annotation that specifies intronic regions, pre-mRNAs, spliced mRNAs, and intronless transcripts (Talkish et al., 2019) to map RNA-seq reads to the transcriptome and then measured these transcript types. We observe higher levels of unspliced pre-mRNA and spliced mRNA when comparing the changes in *QKI5* induction to the control, while intronless transcript abundance decreases (cutoff $p<0.1$; **Figure 5B** and **Supplementary file 6**). Unspliced pre-mRNAs accumulate to a greater degree than spliced mRNAs, supporting the notion that ectopic *QKI5* expression perturbs splicing in yeast ($**p<0.01$ by Mann-Whitney U; **Figure 5B**). We measured changes in splicing by calculating the percentage of unspliced pre-mRNA for each intron-containing transcript ($n=304$ (out of 325 total) that passed expression level cutoff of base mean >100) and discovered that 50 changed significantly (cutoff $p<0.1$ by Student's t-test; **Supplementary file 7**) upon *QKI5* expression. Of these, 41 increase in percent unspliced ($\sim 12\%$ of all expressed intron-containing transcripts) while nine decrease compared to control, which indicates a strong bias toward intron accumulation upon ectopic *QKI5* expression (**Figure 5C**). We validated several of these using RT-PCR (**Figure 5D** $n=4$) and (**Figure 5—figure supplement 1C** $n=15$) or RT-qPCR (**Figure 5E** $n=5$) and found that each of these instances of repressed splicing upon *QKI5* expression that is measured by RNA-seq is also observed using these complementary methods. Notably, the *HAC1* intron, which is not removed by the spliceosome (Sidrauski et al., 1996), shows no increase in intron accumulation upon ectopic *QKI5* expression by either RNA-seq or RT-PCR (**Figure 5—figure supplement 1C**). Next, we asked if any intron sequence motifs are enriched that correlate with increased intron retention upon *QKI5* expression by obtaining 80 nt of intron sequence upstream of the 3'ss and compared these to 80 nt of intron sequence in 108 control introns that are expressed but unchanged compared to control using SEA (Bailey and Grant, 2021 see Methods). Nine motifs are significantly enriched ($p<0.01$), four of which have varying degrees of resemblance to the *QKI* binding motif (ACUAA) or 'half-site' (UAAY) (Galarneau and Richard, 2005; Ryder and Williamson, 2004; Hafner et al., 2010). We note that several of these could potentially serve as tandem embedded motifs (for example, UACUAACUAAC, where the first is italicized and the second underlined). A search for these finds that 18% of the introns that accumulated upon *QKI5* induction compared to control have either two independent or tandem embedded motifs. In contrast, examination of the introns that are unchanged upon *QKI5* induction reveals that only 6% had two intron ACUAA motifs within 80 nt of the 3'ss. The occurrence of two ACUAA elements per intron 3'ss-proximal region is significantly overrepresented in the set of introns that accumulate upon *QKI5* ectopic expression ($p=0.035$ by Chi-squared test), suggesting that these are more sensitive to *QKI5* splicing repression. Intriguingly, the essential *ACT1* pre-mRNA has dual ACUAA elements in its intron proximal to the 3'ss, like those observed in *RAI14*

intron 10 (**Figure 2A**), and its splicing is blocked by QKI5 induction (**Figure 5F**). Thus, QKI5 expression in yeast is lethal and is, at least in part, due to splicing inhibition.

Discussion

Our study reveals that QKI can repress the inclusion of alternatively spliced exons by directly competing with SF1/BBP for ACUAA bs. We find that 43% of the alternatively spliced exons co-regulated by both SF1 and QKI fall within the 'QKI repressed and SF1 activated' category and are associated with an ACUAA sequence motif/bs (**Figures 1 and 2**). We show unambiguously that either of the two ACUAAC elements in *RAI14* intron 10 can be used as a bs, but that one or the other is required for any splicing and so must be a bs (**Figure 3**). Interestingly, deleting either ACUAAC element reduces exon skipping (**Figure 3**) and QKI5 binding (**Figure 4**), but increases inclusion (**Figure 3**) and SF1 binding (**Figure 4**). Therefore, QKI binds with high affinity to a bona fide dual bs in *RAI14* intron 10 to prevent SF1 binding and splicing activation. In budding yeast, where the bs is nearly invariant ACUAAC, ectopic expression of QKI5 blocks pre-mRNA splicing and is lethal (**Figure 5**). Together, these findings demonstrate the validity of the bp competition model by showing these two RBPs directly compete for ACUAA bs substates.

Consequences of bp competition for pre-mRNA substates and other RNAs

One of the best-studied examples of competition between RBPs for pre-mRNA substrates is the competition between SR proteins, which promote exon activation, and hnRNPs, which repress exon inclusion (**Mayeda and Krainer, 1992; Lin and Fu, 2007**). This is the predominant manner through which most cell/tissue types carry out alternative pre-mRNA splicing. The model for bp competition between SF1 and QKI that we define here describes a novel but relatively simple pathway through which a specific subset of alternatively spliced exons with ACUAA bs can be regulated. Our findings are consistent with the established role of SF1 in bp recognition and splicing activation and suggest that QKI is more often a splicing repressor. A nuanced implication of this model is that the loss of function of one of these proteins leads to a gain of function in the other, by relieving competitive inhibition of a specific number of ACUAA bs that control the inclusion of certain alternatively spliced exons.

How does competition between SF1 and QKI influence transcript fate? In the case of *RAI14* intron 10, the dual ACUAA bs constitute a high-affinity bipartite binding site for dimeric QKI5 protein. In C2C12 myoblasts where QKI5 levels are high, *Rai14* exon 11 is mostly skipped (**Figure 2**), and this correlates with QKI5 binding (**Figures 2 and 4**). This binding event appears to promote paraspeckle association, as the paraspeckle proteins NONO, SFPO, and PSPC1 (**Knott et al., 2016**) co-associate with QKI5-bound *RAI14* intron 10 (**Figure 4F**). This finding is consistent with a recent report that identifies QKI as a paraspeckle-associated protein (**Dyakov et al., 2024**). It is unclear how paraspeckle localization might influence splicing outcomes, but pre-mRNA localization within the nuclear speckle is associated with more efficient splicing (**Bhat et al., 2024**). It is possible that QKI-directed paraspeckle localization is repressive to pre-mRNA splicing or promotes the ligation of the exons flanking an alternatively spliced one, leading to exon skipping, as we observe in *RAI14*. Further study is required to elucidate how competition between SF1 and QKI for pre-mRNA substates influences their subnuclear localization, splicing, and stability. A recent study suggests that this may be complicated by the fact that QKI itself attenuates paraspeckle biogenesis by promoting the expression of the short *NEAT1* lncRNA isoform, which has lower paraspeckle-promoting activity than the longer isoform (**Zakutansky et al., 2024**). It is unclear if SF1 might have a similar role, which is possible, given that QKI promotes expression of the short isoform of *NEAT1* via ACUAA elements (to which SF1 might also bind). Therefore, this may constitute a complex feedback loop in which QKI (and/or SF1) can regulate splicing in a paraspeckle-dependent manner but also regulates the degree to which paraspeckles accumulate.

Branch site competition defines a subset of cell-type-specific alternative splicing

Competition for alternatively spliced exons between SR proteins and hnRNPs largely explains the splicing patterns observed in most cell types, but not in the most divergent patterns observed in muscle or brain cells (**Merkin et al., 2012; Barbosa-Morais et al., 2012**). This is due in part to

muscle- or brain-specific RBP expression patterns, and our study uncovers the novel finding that bp competition between SF1 and QKI helps explain how these are enforced. SF1 levels are relatively high in most tissues but can vary by up to eightfold (Tapial et al., 2017), while QKI levels are much more dynamic. QKI levels are elevated in muscle, but undetectable in most neurons or predominated by the cytoplasmic isoforms QKI6 and QKI7 in glia (Noveroske et al., 2002; Hardy et al., 1996; Ebersole et al., 1996; Kondo et al., 1999; Feng et al., 2021). Therefore, in muscle, the nuclear QKI:SF1 ratio is high, whereas in many brain cell types, it is low (or zero), resulting in opposite patterns of isoform-specific transcriptomes and proteomes within the context of the bs competition model. This is also relevant in developmentally regulated systems, where QKI5 levels increase during cardiac cell differentiation (Mazin et al., 2021) but decrease as neural stem cells commit to mature cell types (Hayakawa-Yano et al., 2017). Similarly, when embryonic stem cells exit pluripotency and commit to the endoderm lineage, QKI levels decrease; however, upon commitment to the mesodermal lineage, QKI levels increase, and a QKI-associated splicing gain-of-function is observed (Fagg et al., 2022). On the other hand, SF1 reduction, leading to increased unspliced RNA and skipped exons, has been observed in a model of aging and correlates with age-related fitness decline (Heintz et al., 2017). In summary, the competition between SF1 and QKI appears to be a novel molecular mechanism through which different cell-type-, tissue-, lineage-, and developmental-specific, as well as homeostatic, splicing programs can be achieved and regulated.

Evolutionary implications of bp competition

Yeast and other single-celled eukaryotes lack QKI, whereas multicellular eukaryotes and plants have the gene or an ortholog. Strikingly, we found that 12% of the introns expressed in yeast strain BY4741's splicing was inhibited by ectopic QKI5 expression, and that these cells developed a large-budded phenotype due to a failure to divide (Figure 5—figure supplement 1A and B). This phenotype has been observed due to a lack of *TUB1* pre-mRNA splicing (Burns et al., 2002), although we did not observe mis-splicing of it (or *TUB3*) in our study. The introns most sensitive to QKI5 have evidence of the 'bipartite' QKI motif, which is consistent with high-affinity binding of dimeric QKI (Galarneau and Richard, 2005) and similar to that observed in RAI14 intron 10. It is interesting that yeast lack QKI, do not exhibit 'alternative splicing' per se, and have a more invariant bs. This is UACUAAC in *S. cerevisiae* (Berglund et al., 1997; Moore et al., 1993; Gould et al., 2016), but *Schizosaccharomyces pombe* has a slightly more degenerate bs (Kupfer et al., 2004), and its introns share some features with both budding yeast and mammals (Prabhala et al., 1992; Käufer and Potashkin, 2000). Nevertheless, the consensus bs in the latter still represents a strong QKI-binding motif (Galarneau and Richard, 2005; Ryder and Williamson, 2004), suggesting that QKI expression might also be lethal. Thus, it would be evolutionarily unfavorable to select for QKI or an ortholog/close relative.

The metazoan STAR family of RBPs, with the exception of SF1/BBP, which is the most divergent member, all contain the QUA1 dimerization domain (Vernet and Artzt, 1997). In the case of QKI (or ASD2 in *C. elegans* Ohno et al., 2008) or HOW in *Drosophila* (Zaffran et al., 1997), the binding motif ACUAA, along with the presence of a YAAY 'half site' within about 20 nt, requires more specificity and generates additional points of contact on an RNA, which may explain in part its higher binding affinity than SF1. Perhaps it would not have been evolutionarily advantageous for the majority of bs in multicellular organisms to be UACUAAC/high-affinity QKI motifs (fewer than 20% of human introns have this bs Irimia and Roy, 2008). So more degenerate bs were selected for in metazoans. However, the presence of some ACUAA bs could help diversify the functions of cell-type-specific protein isoforms. For example, many cytoskeletal/contraction-related protein isoforms are specific to muscle (Nakka et al., 2018; Giudice et al., 2016), and many of these are alternatively spliced exons that are repressed by QKI through binding UACUAAY in introns upstream of them (Fagg et al., 2017; Sugnet et al., 2006; Hall et al., 2013; Mazin et al., 2021). Similar observations in worms and flies suggest the intriguing possibility that this may have coevolved with the QKI gene and the more extensive bs degeneracy observed in metazoans and plants. One way this would have been executed, but also constrained, is through bs competition between QKI and SF1/BBP. Consequently, this may have enabled expansion of the transcriptome and proteome through diversification of alternative splicing.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Cell line (yeast)	BY4741	Ares Lab/Prakash Lab	NCBI Taxonomy ID: 1247190	
Cell line (mouse)	C2C12	ATCC	CRL-1772	
Cell line (human)	Flp-In T-Rex 293; WT or QKI KO	Thermo Fisher	RRID:CVCL_U427	
Transfected construct (human)	pDUP-RAI14 (WT and various mutants)	Parental plasmid Dominski and Kole, 1991		
Transfected construct (mouse)	pMyc:Qk5	Sean Ryder		'Wild-type' but contains a translationally silent mutation in the region corresponding to exon 6
Transfected construct (mouse)	pMyc:Qk5 (K120A;R124A)	Fagg et al., 2022	QKI5 defective in RNA binding	
Transfected construct (human)	pcDNA3.1-SF1	This paper		See 'Plasmids and Transfections' section in Methods below
Other	pcDNA3.1-tdTomato	Fagg et al., 2017		Fagg et al., 2017
Other	pJW1666	Addgene plasmid # 112040		Jonathan Weissman (Addgene plasmid # 112040)
Other	pJW1666-Qki5	Addgene plasmid # 112040 with Qki5 inserted		Jonathan Weissman (Addgene plasmid # 112040) with Qki5 inserted
Other	pJW1666-Qki5 (K120A;R124A)	Addgene plasmid # 112040 with Qki5 (K120A;R124A) inserted		Jonathan Weissman (Addgene plasmid # 112040) with Qki5 (K120A;R124A) inserted
Recombinant DNA reagent	pcDNA5-aptamer plasmid	Garcia-Blanco lab		
Antibody	GAPDH; Mouse monoclonal IgM	Sigma	G8795	1:40,000
Antibody	α/β -Tubulin; Rabbit polyclonal	Cell Signalling	2148	1:5000
Antibody	panQKI; Mouse monoclonal IgG2b (clone N147/6)	Sigma	MABN624	1:2000
Antibody	SF1; Rabbit polyclonal	Bethyl	A303-213a	1:5000
Antibody	TATSF1; Rabbit polyclonal	ThermoFisher	20805-1-AP	1:2000
Antibody	IRDye 800CW Goat anti-Mouse IgG2b; Secondary anti-mouse	Licor	P/N 926-32352	1:15,000
Antibody	IRDye 800CW Goat anti-Rabbit IgG Secondary anti-rabbit	Licor	P/N: 926-32211	1:15,000
Antibody	IRDye 680RD Goat anti-Mouse IgM Secondary anti-mouse	Licor	P/N: 926-68180	1:20,000
Antibody	IRDye 680RD Goat anti-Rabbit IgG Secondary anti-rabbit	Licor	P/N: 926-68071	1:20,000
Commercial assay or kit	SYBR Green Universal Master Mix	Applied Biosystems		
Commercial assay or kit	Taq2x Master Mix	New England Biolabs		

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay or kit	ReliaPrep RNA mini prep kit	Promega		
Commercial assay or kit	iScript Reverse Transcription Supermix	BioRad		
Commercial assay or kit	NEBNext poly(A) mRNA Magnetic Isolation module	New England Biolabs	NEB, E7490	
Commercial assay or kit	NEBNext Ultra II Directional RNA Library Prep kit for Illumina	New England Biolabs	NEB, E7760	
Commercial assay or kit	HiScribe T7 High Yield RNA Kit	New England Biolabs	NEB, E2040	
Software, algorithm	UCSF ChimeraX	https://www.rbvi.ucsf.edu/ chimerax	RRID:SCR_015872	Version 1.10.1
Software, algorithm	Vast-tools	<i>Irimia et al., 2014; Tapial et al., 2017</i>		2.0.2, database has16.02.18
Software, algorithm	rMATS	<i>Shen et al., 2014; Park et al., 2016</i>		Version 4.0.2, GENCODE v30 annotation
Software, algorithm	rMAPS2	rMAPS2; https://rmaps. cecsresearch.org/		
Software, algorithm	Trimmomatic	Trimmomatic	RRID:SCR_011848	v.0.39
Software, algorithm	Kallisto	Kallisto	RRID:SCR_016582	v.0.50
Software, algorithm	Deseq2	<i>Love et al., 2023</i>		v1.42.1
Software, algorithm	Simple Enrichment Analysis (SEA)	<i>Bailey and Grant, 2021;</i> https://meme-suite.org/ meme/doc/sea.html		
Software, algorithm	XCalibur	Thermo Scientific	RRID:SCR_014593	V2.5
Software, algorithm	Proteome Discoverer	Thermo Fisher	RRID:SCR_014477	Version 2.2.0388
Software, algorithm	Scaffold	Proteome Software Inc.	RRID:SCR_014321	Version Scaffold_4.11.1
Software, algorithm	Protein Prophet	<i>Nesvizhskii et al., 2003</i>		
Software, algorithm	MSConvert	<i>Chambers et al., 2012</i>		
Software, algorithm	MSFragger	<i>Kong et al., 2017</i>		
Software, algorithm	DIA-NN	https://github.com/ vdemichev/DiaNN ; <i>Demichev et al., 2020</i>		
Software, algorithm	Fragpipe-Analyst	<i>Hsiao et al., 2024</i>		
Sequence-based reagent	Oligonucleotides and RNAi	See <i>Supplementary file 8</i>		

Cell culture

C2C12 mouse myoblasts (obtained from ATCC) and HEK293 cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with high glucose (Invitrogen) and 10% (v/v) heat-inactivated fetal bovine serum (Thermo Fisher). Cells were maintained at 37 °C in a humidified atmosphere with 5% CO₂. The HEK293 QKI KO cells were shared by Kuo-Chieh Liao and generated as previously described (*Liao et al., 2021; Liao et al., 2020*) and used the Flp-In T-Rex 293 cell line. The cells were routinely tested for mycoplasma and tested negative. STR phenotyping was conducted at the UTMB Molecular Genomics Core Facility on QKI KO and control HEK cells, confirming their identity as HEK293T cells.

Protein alignment

The Quaking protein used for alignment is as referenced (*Teplova et al., 2013*) and used PDB accession 4JVH and for SF1 *Liu et al., 2001* used PDB 1K1G. Alignment and visualization were performed using UCSF ChimeraX (<https://www.rbvi.ucsf.edu/chimerax>) version 1.10.1 (*Meng et al., 2023*).

Plasmids and transfections

The pDUP51 splicing reporter plasmid (*Dominski and Kole, 1991*) was used as a backbone to generate the RAI14 exon 11 plasmid. A 243 bp fragment upstream of exon 11, exon 11, and a 95 bp fragment downstream of the exon were PCR-amplified from H9 human embryonic stem cell genomic DNA; the forward primer contained an Apal site, and the reverse primer contained a BglII site. These and the pDUP51 plasmid were digested with Apal and BglII (NEB), then ligated into pDUP51 at the Apal and BglII restriction sites. The resulting pDUP-RAI14 exon 11 plasmid was confirmed by Sanger sequencing. Mutant constructs were generated using the Q5 Site-Directed Mutagenesis Kit and were also verified by Sanger sequencing. The myc:QKI5 plasmids used were described previously and were a kind gift from Sean Ryder (*Fagg et al., 2017; Fagg et al., 2022*). Generation of the pcDNA3.1-tdTomato plasmid was also previously described (*Fagg et al., 2017*). The pcDNA3.1-SF1 plasmid was generated by Gibson Assembly using cDNA from H9 human embryonic stem cells as a template for PCR amplification and confirmed by Sanger sequencing.

Transfections were carried out with Lipofectamine 2000 (Invitrogen) using an 'in tube transfection' protocol, where 2.5×10^5 cells were added to a tube containing Lipofectamine and the appropriate volume of reagent, following the manufacturer's instructions, with 100 ng DNA mix in Gibco Opti-MEM (Thermo) or 30 pmol of siRNA, incubated for 20 min at RT, plated in 12 well plates containing DMEM 10% FBS and incubated overnight. Cells were harvested 24 hr after transfection.

For the RNA affinity chromatography, the j6f1 aptamer under the control of the T7 RNA promoter was obtained by the digestion of the vector pcDNA5-aptamer plasmid and was a gift from Chloe Nagasawa (University of Texas Medical Branch, Galveston, Texas, US) from the lab of Mariano Garcia-Blanco (University of Virginia). RAI14 sequence consisting of 60 bp fragment of RAI14 intron 10 upstream RAI14 exon 11 was obtained from a gBlock (IDT), digested with HindIII and NotI restriction enzymes and cloned into the digested PCNA5 aptamer plasmid.

For yeast transformation, the plasmid pJW1666 (gift from Jonathan Weissman Addgene plasmid # 112040; <http://n2t.net/addgene:112040>; RRID:Addgene_112040; *Costa et al., 2018*) was used to clone the WT or mutant QKI sequences, or the GFP present on WT the plasmid was used as control.

RNA extraction, RT-PCR, and RT-qPCR

RNA was extracted from cells using Trizol (Thermo Fisher). For direct extraction from cell plates, Trizol was added to the wells and samples were vortexed. For extraction from cell lysates, cells were extracted with RSB100 (100 mM Tris-HCl pH 7.4, 0.5% (v/v) NP-40, 0.5% (v/v) Triton X-100, 0.1% (w/v) SDS, 100 mM NaCl, and 2.5 mM MgCl₂). Samples were vortexed, chloroform was added, then the samples were incubated at RT for 2 min. After centrifugation at 13,000 RPM for 15 min at 4 °C, the aqueous phase was transferred to a new tube and samples were processed with the ReliaPrep RNA MiniPrep system (Promega), according to the manufacturer's instructions. Reverse transcription was performed with iScript Reverse Transcription Supermix (BioRad) according to manufacturer instructions. RT-PCR was performed using Taq2x Master Mix (NEB), and cycle numbers were determined empirically to prevent over-amplification. PCR products were analyzed by agarose gel and via Agilent 2100 BioAnalyzer.

For RT-qPCR, cDNA was diluted 1:8 and used with Applied Biosystems SYBR Green Universal Master Mix on a Step One Plus Real-Time PCR machine (Applied Biosystems). RT-qPCR cycling conditions were: 95 °C for 10 min, [95 °C 15 s, 60 °C 60 s] (40 cycles). DUP-RAI14 exon 11 reporter RNA abundances were analyzed by calculating the ΔC_t values normalized to EEF1A1. The $\Delta\Delta C_t$ values were then calculated as $2^{-\Delta\Delta C_t}$ for each sample. The average $\Delta\Delta C_t$ of the wild-type reporter was used as a normalization factor to determine the relative RNA abundance in the mutant reporters.

Western blotting

Proteins were extracted using RSB100 (100 mM Tris-HCl pH 7.4, 0.5% (v/v) NP-40, 0.5% (v/v) Triton X-100, 0.1% (w/v) SDS, 100 mM NaCl, and 2.5 mM MgCl₂) with EDTA-free protease inhibitor cocktail

(SIGMAFAST, Sigma) and 1 mM PMSF. The buffer was added directly to the wells of the plates, scraped, collected, and kept on ice, while vortexing every 5 min for 30 min, followed by centrifugation at 14,000 RPM for 15 min at 4 °C. The supernatant was transferred to a new tube. Samples were stored at –80 °C until use.

Protein concentration was measured using the Bradford assay (BioRad **Bradford, 1976**). The same amount of protein from each sample was loaded onto 10% SDS-PAGE (typically 15–30 µg) and then transferred to a 0.45 µm nitrocellulose membrane (Thermo). The membrane was blocked (with 5% non-fat dry milk dissolved in tris-buffered saline TBS) for 1 hr at room temperature with constant agitation. Primary antibody incubations were performed overnight at 4 °C with constant agitation in TBS 0.01% (v/v) tween (TBS-T) with 5% non-fat dry milk using the following antibodies: Anti-Pan-QKI Antibody, clone N147/6 (MABN624, Sigma-Aldrich) (1:2000), IgM-anti-GAPDH (G8795, clone GAPDH-71.1, Sigma-Aldrich) (1:40000), α/β-Tubulin (#2148, Cell Signaling) (1:5000), anti-SF1 (#A303-213a, Bethyl Laboratories) (1:5000), and anti-HTATSF1 (# 20805–1-AP, Thermo Fisher) (1:2000), all diluted in TBS containing 5% (w/v) milk, and 0.01% (v/v) Tween-20 (TBST). The following day, the membranes were washed three times with TBST with 5% (w/v) milk at RT. Secondary infrared conjugated antibodies: IRDye 800CW Goat anti-Mouse IgG2b (P/N 926–32352), IRDye 680RD Goat anti-Mouse IgM (P/N: 926–68180), IRDye 800CW Goat anti-Rabbit IgG (P/N: 926–32211), and IRDye 680RD Goat anti-Rabbit IgG (P/N: 926–68071) from Li-Cor were diluted according to manufacturer instructions (1:15,000 for 800CW or 1:20,000 for 680RD/LT conjugates) in TBST with 5% (w/v) non-fat dry milk for 1 hr with constant agitation. Membranes were washed three times every 5 min before visualization on the Odyssey CLx imager (Li-Cor).

Yeast transformation

Yeast transformation was performed as described by Ito et al. with modifications (**Ito et al., 1983**) using the BY4741 strain. A 50 mL culture of BY4741 was grown to a density of 5×10^6 overnight. Following centrifugation, the pellet was resuspended in 1 ml of deionized water, centrifuged at 12,000 RPM and resuspended in 0.5 ml of 1 X TE-LiAc (100 mM tris-HCl, 10 mM EDTA pH 7.5; LiAc pH 7.5). Subsequently, cells were resuspended in 3 x volume of 1 X TE-LiAc and incubated at 30 °C with agitation. 2 µg of purified PCR of WT QKI, mutant QKI and GFP plasmids, and 7 µL of salmon sperm carrier DNA were added to 200 µL of the yeast cells and the mixture was incubated for 45 min at 30 °C with agitation. The following steps included a 3 hr incubation at 30 °C with agitation, a 20 min heat shock at 42 °C, centrifugation at 12,000 RPM, and resuspension in YEPD followed by a 45 min incubation. After a 12,000 RPM centrifugation, cells were resuspended in water and spread in URA- selection plates, then incubated at 30 °C until transformants appeared.

Yeast genomic DNA extraction

Genomic DNA extraction was performed using the method developed by Klassen and collaborators with a few modifications (**Klassen et al., 2008**). Cells were patched on URA- selection plates, and 50 µL of cell volume was scraped off the plate and resuspended in zymolase 20T 8 mg/ml. After vortexing the samples, they were incubated at 37 °C for 1 hr. After incubation, the cells were spun down, the supernatants were aspirated, and the cells were resuspended in 250 µL of 10% SDS. Then, they were vortexed and incubated for 30 min at 65 °C. After incubating the samples on ice for 5 min, 100 µL of 5 M potassium acetate, pH 7.5, was added, and the samples were incubated on ice for 1 hr. Tubes were spun at 14,000 RPM for 15 min, and the supernatant was transferred to a new tube. Then 400 µL of ice-cold 95% EtOH was added, vortexed briefly, and spun for 15 min at 14,000 RPM. The supernatants were decanted, and the DNA pellet was dried at 30 °C for 15 min. It was then dissolved in 100 µL of ultra-pure water and placed in a mixer for 20 min at 2000 RPM. The extracted genomic DNA was used to perform PCR with high-fidelity polymerase (Takara Bio Inc) to confirm the transgene expression by Sanger sequencing.

Yeast RNA extraction

RNA was extracted from yeast cells as previously described (**Rio, 2011**). Briefly, 1 mL of yeast culture was pelleted and then resuspended in 400 µL of AE buffer (50 mM, pH 5.2 of NaOAc, 10 Mm EDTA). After the addition of 40 µL of 10% SDS and 400 µL of PCA, the samples were incubated at 65 °C for 10 min. After a 5 min incubation on ice, samples were placed in phase lock gel tubes, centrifuged

at 14,000 RPM for 5 min, and centrifuged again using the same conditions following another chloroform addition; this chloroform wash was repeated, and the samples were centrifuged one final time. The aqueous phase was transferred to a new 1.5 mL tube, 50 μ L of 3 M sodium acetate, pH 5.2, was added, followed by two volumes of 100% ethanol. Samples were centrifuged at 14,000 RPM for 15 min, the ethanol was removed, and the pellet was washed with 70% ethanol. Then, the samples were centrifuged at 14,000 RPM for 5 min, the ethanol was removed, and the samples were air-dried. The pellet was resuspended in RNase-free water and then vortexed. To remove any potential DNA contamination, 1 μ g of RNA was treated with DNase Turbo following the manufacturer's instructions. RNA was extracted after DNase treatment with phenol/chloroform/isoamyl alcohol and then ethanol precipitation. RNA concentration was measured using the NanoDrop (Thermo Fisher).

Yeast spot assay

One milliliter of freshly harvested cells grown overnight in 10 mL of YEPD were centrifuged, washed in ultra-pure water and sonicated. The cells were then diluted to a concentration of 1×10^7 . Serial 10-fold (up to five times) were prepared on both galactose (Gal) and non-galactose plates. The plates were incubated at 30 °C for 3 days.

Yeast growth curve

Yeast was grown in YEPD overnight. When the OD660 reached approximately 0.5, transgene activation was induced by adding galactose dissolved in water to a final concentration of 2% or water as control. For each time point, samples were collected for cell count and RNA extraction.

RNA sequencing and analysis

RNA-seq data from the ENCODE project for shQKI and shNT (GSE80878) or shSF1 and shNT (GSE88430) (Van Nostrand et al., 2020a; The ENCODE Project Consortium, 2012) were downloaded from GEO and analyzed as previously described (Fagg et al., 2022) using Vast-tools (Irimia et al., 2014; Tapial et al., 2017) or rMATS (Shen et al., 2014) to compare alternative splicing observed in shQKI relative to shNT, or shSF1 relative to shNT.

The subsequent Vast-tools data files were filtered using a cutoff of dPSI > |10| and MVdPSI95>0 to identify alternatively spliced events that changed significantly upon either QKI or SF1 knockdown relative to the control. For 'co-regulated' exons, the above cutoff was required for either shQKI compared to control or shSF1 compared to control, but in both datasets, the MVdPSI95 had to be >0. rMATS analysis and rMAPS2 (Park et al., 2016; Hwang et al., 2020) motif enrichments were done using default conditions and with custom motifs added to measure various bp-like and potential SF1 binding motifs (ACT[ACTG]AG, [ACTG]CT[AG][CT], TAA[CT], TAA[CT]T[ACTG]A[CT], TACTAAC, TACTAA, ACTAA[CT], TACTAA[CT], CTAAC[ACG]).

Strand-specific RNAseq libraries were prepared from 1 μ g of yeast total RNA using NEBNext poly(A) mRNA Magnetic Isolation module (NEB, E7490) and NEBNext Ultra II Directional RNA Library Prep kit for Illumina (NEB, E7760) following the manufacturer's recommended procedure. The six libraries were pooled in equal molar concentration and sequenced on Illumina NextSeq 550 for PE 150 base pair sequencing, yielding about 30 M paired-end reads each. The associated data files have been uploaded to GEO under accession GSE273838. The RNAseq reads were filtered for low-quality bases and trimmed of adapter sequences using Trimmomatic (v.0.39). The trimmed reads were aligned to the yeast reference genome SacCer3 and a custom yeast annotation file, which accounts for all yeast introns (Hardy et al., 1996) (and allows measurement of unspliced, spliced, and intronless transcripts) using kallisto (v.0.50) to generate the abundance.tsv file for either normalized (TPM) or total (counts) reads for each sample. The TPM file was used as input for DESeq2 analysis (v1.42.1) to identify changes in transcript abundance (abundance cutoff TPM > 0.2 and significance cutoff $p > 0.1$; **Supplementary file 6**), and the read counts file was used to calculate the percentage of unspliced RNA (percent unspliced = (unspliced/(unspliced+spliced))*100) for each transcript that was expressed (base mean > 100) and statistical significance was measured using the Student's t-test with a change in percent unspliced $p < 0.1$ considered significant (**Supplementary file 7**).

Motif analysis of both the yeast and HepG2 RNA-seq datasets were performed using SEA (Bailey and Grant, 2021), and the former used 80 nt of intron sequence upstream of the 3'ss of yeast introns, 38 of which increased ($p < 0.1$) upon ectopic QKI5 expression compared to the parental control (this

represents 38 of the 41 we observed that increased in inclusion upon QKI5 expression, as the other two were located on ChrM and did not overlap with 'Talkish Standard Introns' *Talkish et al., 2019*, which was a requirement for our motif analysis), compared to a background set of 106 introns that were detectable (base mean > 100) but unchanged ($p > 0.2$, change in percent unspliced < |1|) upon ectopic QKI5 expression compared to the parental control. For the HepG2 intron set, we extracted the intron sequence that began 20 nt upstream of the 3'ss (to exclude analyzing potential differences in 3'ss or pY tracts) and spanned 60 nt upstream of the 3'ss of introns that were co-regulated by QKI and SF1 (dPSI > |10| and MVdPSI at 95% confidence interval > 0 in either dataset or the other and at least MVdPSI at 95% confidence interval > 0 in both datasets, as determined by Vast-tools) and performed SEA, and then also performed an identical analysis but using 1000 introns that were detectable (base mean > 100) but unchanged in either shQKI relative to shNT or shSF1 relative to shNT (dPSI < |1|, MVdPSI = 0). In both cases, the motif set that was interrogated was from *Ray et al., 2013*.

RNA affinity chromatography

Tobramycin RNA affinity chromatography was performed as previously described (*Hartmuth et al., 2002*) with several minor modifications. Briefly, the aptamer and RAI14 WT and mutant RNAs were produced after the vectors were digested with NotI (NEB) using the HiScribe T7 High Yield RNA Kit (NEB). Then 120 picomoles of RNA of the aptamer and RAI14 mutants, and 200 picomoles of wt RAI14 were heated in RNA binding buffer (20 mM Tris-HCl, 1 mM CaCl₂, 1 mM MgCl₂, 300 mM KCl, 0.1 mg/ml tRNA, 0.5 mg/ml BSA, 0.01% (v/v) NP-40, 0.2 mM DTT) at 95 °C for 5 min and transferred to RT for 30 min. Samples were incubated with 60 µL of tobramycin-coupled Sepharose matrix at 4 °C for 2 hr with head-over-tail rotation. The beads were washed three times with washing buffer (20 mM Tris-HCl, 1 mM CaCl₂, 1 mM MgCl₂, 145 mM KCl, 0.1% (v/v) NP-40, 0.2 mM DTT) and incubated with a 285 µL solution of 32% of C2C12 NE (as previously described *Dignam et al., 1983*, 32 mM KCl, 2 mM MgCl₂, 2 mM ATP, and 20 mM creatine phosphate (CP) (the reactions were also performed without MgCl₂, ATP and CP in the NE solution) for 7.5, 15, and 30 min at 30 °C with head over tail rotation. After NE incubation, the matrix was washed three times with a higher salt concentration of washing buffer (150 mM KCl). The beads were eluted with 5 mM of tobramycin in 20 mM Tris-HCl, 1 mM CaCl₂, 1 mM MgCl₂, 145 mM KCl, 2 mM MgCl₂, 0.2 mM DTT in a 125 µL solution. The protein eluate was precipitated with acetone and suspended in LDS sample buffer for WB and in 5% (w/v) SDS, 50 mM TEAB pH 7.1 for MS analysis.

Protein precipitation

Acetone was used to precipitate the proteins from the RNA affinity chromatography-eluted samples. Four times the sample volume of cold acetone was added to each sample and vortexed. Samples were incubated for 60 min at -20 °C for 1 hr, following a 10 min centrifugation at 14,000 RPM at 4 °C. The acetone was decanted, and the samples were air-dried for 15 min. Pellets were resuspended with 1 x Laemmli buffer for western blot or with 5% (w/v) SDS, 50 mM TEAB, pH 7.1, for liquid chromatography with tandem mass spectrometry.

Protein digestion

The samples were prepared as previously described (*Herrmann et al., 2021*). Briefly, 25 µg of protein from the above were reduced with 10 mM Tris(2-carboxyethyl) phosphine (TCEP) (77720, Thermo) and incubated at 65 °C for 10 min. The sample was then cooled to room temperature, and 1 µL of 500 mM iodoacetamide acid was added and allowed to react for 30 min in the dark. Then, 3.3 µL of 12% phosphoric acid was added to the protein solution followed by 200 µL of binding buffer (90% Methanol, 100 mM TEAB pH 8.5). The resulting solution was added to the S-Trap spin column (<https://protifi.com/>) and passed through the column using a bench-top centrifuge (60 s spin at 1000×g). The spin column is washed with 150 µL of binding buffer and centrifuged. This is repeated twice. 30 µL of 20 ng/µL trypsin is added to the protein mixture in 50 mM TEAB pH 8.5 and incubated at 37°C overnight. Peptides were eluted twice with 75 µL of 50% acetonitrile, 0.1% (v/v) formic acid. Aliquots of 20 µL of eluted peptides were quantified using the Quantitative Fluorometric Peptide Assay (Pierce, Thermo Fisher Scientific). Eluted volume of peptides corresponding to 5.5 µg of peptides are dried in a speed vac and resuspended in 27.5 µL 1.67% (v/v) acetonitrile, 0.08% (v/v) formic acid, 0.83% (v/v) acetic acid, 97.42% water, and placed in an autosampler vial.

Data-dependent acquisition NanoLC MS/MS analysis

Peptide mixtures were analyzed by nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) using a nano-LC chromatography system (UltiMate 3000 RSLCnano, Dionex), coupled on-line to a Thermo Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA) through a nanospray ion source (Thermo Scientific) similar to as we have described previously (Liu *et al.*, 2022). A trap and elute method was used. The trap column was a C18 PepMap100 (300 $\mu\text{m} \times 5 \text{ mm}$, 5 μm particle size) from Thermo Scientific. The analytical column was an Acclaim PepMap 100 (75 $\mu\text{m} \times 25 \text{ cm}$) from (Thermo Scientific). After equilibrating the column in 98% solvent A (0.1% formic acid in water) and 2% solvent B (0.1% (v/v) formic acid in acetonitrile (ACN)), the samples (2 μL in solvent A) were injected onto the trap column and subsequently eluted (400 nL/min) by gradient elution onto the C18 column as follows: isocratic at 2% B, 0–5 min; 2 to 24% B, 5–86 min; 24 to 44% B, 86–93 min; 44 to 90% B, 93–95 min; 90% B for 1 min, 90 to 10% B, 96–98 min; 10% B for 1 min 10 to 90% B, 99–102 min 90 to 4% B; 90% B for 3 min; 90 to 2%, 105–107 min; and isocratic at 2% B, till 120 min.

All LC-MS/MS data were acquired using XCalibur, version 2.5 (Thermo Fisher Scientific) in positive ion mode using a top speed data-dependent acquisition (DDA) method with a 3 s cycle time. The survey scans (m/z 350–1500) were acquired in the Orbitrap at 120,000 resolution (at $m/z=400$) in profile mode, with a maximum injection time of 100 ms and an AGC target of 400,000 ions. The S-lens RF level is set to 60. Isolation is performed in the quadrupole with a 1.6 Da isolation window, and CID MS/MS acquisition is performed in profile mode using rapid scan rate with detection in the ion-trap, with the following settings: parent threshold = 5000; collision energy = 32%; maximum injection time 56 ms; AGC target 500,000 ions. Monoisotopic precursor selection (MIPS) and charge state filtering were on, with charge states 2–6 included. Dynamic exclusion is used to remove selected precursor ions, with a ± 10 ppm mass tolerance, for 15 s after acquisition of one MS/MS spectrum.

DDA database searching

Tandem mass spectra were extracted and charge state deconvoluted by Proteome Discoverer (Thermo Fisher, version 2.2.0388). Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Sequest (Thermo Fisher Scientific, San Jose, CA, USA; in Proteome Discoverer 2.5.0.402). Sequest was set up to search the canonical mouse proteome and a contaminant database cRAP, assuming the digestion enzyme trypsin. Sequest was searched with a fragment ion mass tolerance of 0.60 Da and a parent ion tolerance of 10.0 PPM. Carbamidomethyl of cysteine was specified in Sequest as a fixed modification. Deamidated of asparagine and glutamine, oxidation of methionine and acetyl of the N-terminus were specified in Sequest as variable modifications.

Criteria for protein identification

Scaffold (version Scaffold_4.11.1, Proteome Software Inc, Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii *et al.*, 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. The resulting normalized spectral counts from the input NE, each RAC substrate (WT, upDEL, dnDEL, 2xDEL, and APT Only control) were obtained, and those values were used to calculate background-corrected levels of enrichment relative to NE ($(\text{NSC}_{\text{RAC}} - \text{NSC}_{\text{APT}})/\text{NSC}_{\text{NE}}$). We used those calculations to construct the heatmap shown in **Figure 4B** and generate the counts shown in **Figure 4C**, by including proteins with experimental LC-MS/MS evidence of being members of early spliceosomes (E-complex or 17S U2 snRNP Makarov *et al.*, 2012; Sharma *et al.*, 2008; Cvitkovic and Jurica, 2013), and that had a positive value of background-corrected levels of enrichment relative to NE in at least 1 of the 4 RAC substrates tested. See **Supplementary file 4**.

Data-independent acquisition NanoLC MS/MS analysis

Peptide mixtures were analyzed by nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) using a nano-LC chromatography system (UltiMate 3000 RSLCnano, Dionex), coupled on-line to a Thermo Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific, San Jose, CA) through a nanospray ion source. A direct injection method is used onto an analytical column; Aurora (75 $\mu\text{m} \times 25 \text{ cm}$, 1.6 μm) from (ionopticks). After equilibrating the column in 98% solvent A (0.1% (v/v) formic acid in water) and 2% solvent B (0.1% (v/v) formic acid in acetonitrile (ACN)), the samples (2 μL in solvent A) were injected (300 nL/min) by gradient elution onto the C18 column as follows: isocratic at 2% B, 0–10 min; 2 to 27% 10–98 min, 27 to 45% B, 98–102 min; 45 to 90% B, 102–103 min; isocratic at 90% B, 103–104 min; 90 to 15%, 104–106 min; 15 to 90% 106–108 min; isocratic for 2 min; 90 to 2%, 110–112 min; and isocratic at 2% B, till 120 min.

All LC-MS/MS data were acquired using an Orbitrap Eclipse in positive ion mode using a data-independent acquisition (DIA) method with a 16 Da window from 400 to 1000 and a loop time of 3 s. The survey scans (m/z 350–1500) were acquired in the Orbitrap at 60,000 resolution (at $m/z=400$) in centroid mode, with a maximum injection time of 118 ms and an AGC target of 100,000 ions. The S-lens RF level was set to 60. Isolation was performed in the quadrupole, and HCD MS/MS acquisition was performed in profile mode using the Orbitrap at a resolution of 30,000 using the following settings: collision energy = 33%, IT 54 ms, AGC target = 50,000. These conditions were duplicated to create six gas-phase fractions of the NE sample using 4 Da fully staggered windows in 100 m/z increments from 400 to 1000 m/z , as described (Searle et al., 2020).

DIA database searching

The raw data was demultiplexed to mzML with 10 ppm accuracy after peak picking in MSConvert (Chambers et al., 2012). The resulting mzML files were searched in MSFragger (Kong et al., 2017) and quantified via DIA-NN (<https://github.com/vdemichev/DiaNN>; Demichev, 2025) using the following settings: peptide length range 7–50, protease set to Trypsin, two missed cleavages, three variable modifications, clip N-term M on, fixed C carbamidomethylation, variable modifications of methionine oxidation and n-terminal acetylation, MS1 and MS2 accuracy set to 20 ppm, 1% FDR, and DIANN quantification strategy set to Robust LC (high accuracy). The files were searched against a database of human proteins acquired from Uniprot (18 December, 2023). The gas-phase fractions were used only to generate the spectral library, which was used for analysis of the individual samples.

Statistical analysis was performed using Fragpipe-Analyst (Hsiao et al., 2024) using an R script based on the ProteinGroup.txt file produced by DIA-NN. First, contaminant proteins, reverse sequences, and proteins identified ‘only by site’ were filtered out. In addition, proteins identified by only a single peptide and those that are not consistently identified or quantified across the same condition are also removed. The DIA data were converted to a $\log_2$ scale, samples were grouped by conditions, and missing values were not imputed. Protein-wise linear models combined with empirical Bayes statistics were used for the differential expression analyses. The limma package from R Bioconductor was used to generate a list of differentially expressed proteins for each pairwise comparison. A cutoff of the adjusted p -value of 0.05 (Benjamini-Hochberg method) and an absolute $\log_2$ fold change of 1 have been applied to determine significantly regulated proteins in each pairwise comparison. We also used the E/U2 list to focus our analysis on early spliceosome proteins and expanded this to include an RBP list (Sasse et al., 2025) to assess non-spliceosome-annotated RBP interactions. The E/U2 or RBPs shown in Figure 4E were included if (1) spectral counts were detected in both the numerator and denominator RAC substrates, (2) $\log_2$ fold change $\geq |0.2|$ and p -value was <0.01 in at least one of the three comparison groups, and (3) data were present for each comparison (no zero or undetectable values). See Supplementary file 5.

Materials availability statement

Newly created materials, such as plasmids, are available upon request to the corresponding author.

Acknowledgements

We thank Manuel Ares Jr. (UCSC) for helpful discussion regarding the experiments in *S. cerevisiae* and for piloting an early version of the experiment. We thank Louise Prakash (UTMB) for allowing us

to use equipment that enabled us to perform the experiments in *S. cerevisiae*, and Robert E Johnson of the Prakash lab for technical assistance. We thank Steve Widen (UTMB, retired) for technical assistance with the QKI and SF1 knockdown RNA-seq datasets. We thank Mariano Garcia-Blanco and Chloe Nagasawa from Mariano’s lab for discussion and technical suggestions on RNA affinity chromatography. We also thank Eric Van Nostrand (Baylor College of Medicine) and Manuel Ares Jr. for pre-submission review of the manuscript. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Additional information

Funding

Funder	Grant reference number	Author
National Institute of General Medical Sciences	R35GM151324	William S Fagg
Cancer Prevention and Research Institute of Texas	RP190682	William K Russell

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Karen Larissa Pereira de Castro, William K Russell, Formal analysis, Investigation, Methodology; Jose M Abril, Investigation, Methodology; Kuo-Chieh Liao, Resources; Haiping Hao, John Paul Donohue, Investigation; William S Fagg, Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

Haiping Hao  <https://orcid.org/0000-0002-8826-1568>
William S Fagg  <https://orcid.org/0000-0003-2909-1666>

Peer review material

Reviewer #2 (Public review): <https://doi.org/10.7554/eLife.103167.3.sa1>
Reviewer #3 (Public review): <https://doi.org/10.7554/eLife.103167.3.sa2>
Author response <https://doi.org/10.7554/eLife.103167.3.sa3>

Additional files

Supplementary files

- Supplementary file 1. Vast-tools splicing analysis datafile from ENCODE RNA-seq data.
- Supplementary file 2. rMATS splicing analysis datafile from ENCODE RNA-seq data comparing shQKI to shControl.
- Supplementary file 3. rMATS splicing analysis datafile from ENCODE RNA-seq data comparing shSF1 to shControl.
- Supplementary file 4. LC-MS/MS datafile from RNA affinity chromatography in the presence of ATP.
- Supplementary file 5. LC-MS/MS datafile from RNA affinity chromatography in the absence of ATP.
- Supplementary file 6. Transcript abundance analysis of RNA-seq data from BY4741 with QKI5 expression compared to parental control.
- Supplementary file 7. Splicing analysis of RNA-seq data from BY4741 with QKI5 expression compared to parental control.
- Supplementary file 8. Oligonucleotide sequence info file.

MDAR checklist

Data availability

New RNA-seq data files have been uploaded to GEO under accession GSE273838. All data generated or analyzed in this study are also included in the manuscript and supporting files; source data files have been provided for all figures.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
de Castro KL, Johnson RE, Abril J, Kuo-Chieh L, Widen SG, Hao H, Donohue JP, Russell WK, Fagg WS	2024	An ancient competition for the conserved branchpoint sequence influences physiological and evolutionary outcomes in splicing	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE273838	NCBI Gene Expression Omnibus, GSE273838

References

Artzt K, Wu JI. 2010. STAR trek: an introduction to STAR family proteins and review of quaking (QKI). *Advances in Experimental Medicine and Biology* **693**:1–24 PMID: 21189682.

Bailey TL, Grant CE. 2021. SEA: simple enrichment analysis of motifs. *bioRxiv*. DOI: <https://doi.org/10.1101/2021.08.23.457422>

Barash Y, Calarco JA, Gao W, Pan Q, Wang X, Shai O, Blencowe BJ, Frey BJ. 2010. Deciphering the splicing code. *Nature* **465**:53–59. DOI: <https://doi.org/10.1038/nature09000>, PMID: 20445623

Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, Slobodeniuc V, Kutter C, Watt S, Colak R, Kim T, Misquitta-Ali CM, Wilson MD, Kim PM, Odom DT, Frey BJ, Blencowe BJ. 2012. The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**:1587–1593. DOI: <https://doi.org/10.1126/science.1230612>, PMID: 23258890

Bennett M, Michaud S, Kingston J, Reed R. 1992. Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes & Development* **6**:1986–2000. DOI: <https://doi.org/10.1101/gad.6.10.1986>, PMID: 1398075

Berglund JA, Chua K, Abovich N, Reed R, Rosbash M. 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89**:781–787. DOI: [https://doi.org/10.1016/s0092-8674\(00\)80261-5](https://doi.org/10.1016/s0092-8674(00)80261-5), PMID: 9182766

Beuck C, Qu S, Fagg WS, Ares M, Williamson JR. 2012. Structural analysis of the quaking homodimerization interface. *Journal of Molecular Biology* **423**:766–781. DOI: <https://doi.org/10.1016/j.jmb.2012.08.027>, PMID: 22982292

Bhat P, Chow A, Emert B, Ettlin O, Quinodoz SA, Strehle M, Takei Y, Burr A, Goronzy IN, Chen AW, Huang W, Ferrer JLM, Soehalim E, Goh S-T, Chari T, Sullivan DK, Blanco MR, Guttman M. 2024. Genome organization around nuclear speckles drives mRNA splicing efficiency. *Nature* **629**:1165–1173. DOI: <https://doi.org/10.1038/s41586-024-07429-6>, PMID: 38720076

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**:248–254. DOI: [https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3), PMID: 942051

Burns CG, Ohi R, Mehta S, O'Toole ET, Winey M, Clark TA, Sugnet CW, Ares M Jr, Gould KL. 2002. Removal of a single alpha-tubulin gene intron suppresses cell cycle arrest phenotypes of splicing factor mutations in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **22**:801–815. DOI: <https://doi.org/10.1128/MCB.22.3.801-815.2002>, PMID: 11784857

Busch A, Hertel KJ. 2012. Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdisciplinary Reviews. RNA* **3**:1–12. DOI: <https://doi.org/10.1002/wrna.100>, PMID: 21898828

Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, et al. 2012. A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology* **30**:918–920. DOI: <https://doi.org/10.1038/nbt.2377>, PMID: 23051804

Corioni M, Antih N, Tanackovic G, Zavolan M, Krämer A. 2011. Analysis of in situ pre-mRNA targets of human splicing factor SF1 reveals a function in alternative splicing. *Nucleic Acids Research* **39**:1868–1879. DOI: <https://doi.org/10.1093/nar/gkq1042>, PMID: 21062807

Costa EA, Subramanian K, Nunnari J, Weissman JS. 2018. Defining the physiological role of SRP in protein-targeting efficiency and specificity. *Science* **359**:689–692. DOI: <https://doi.org/10.1126/science.aar3607>, PMID: 29348368

Cox RD, Hugill A, Shedlovsky A, Noveroske JK, Best S, Justice MJ, Lehrach H, Dove WF. 1999. Contrasting effects of ENU induced embryonic lethal mutations of the quaking gene. *Genomics* **57**:333–341. DOI: <https://doi.org/10.1006/geno.1999.5804>

- Cvitkovic I**, Jurica MS. 2013. Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic Acids Research* **41**:D132–D141. DOI: <https://doi.org/10.1093/nar/gks999>, PMID: 23118483
- Darbelli L**, Vogel G, Almazan G, Richard S. 2016. Quaking regulates neurofascin 155 expression for myelin and axoglial junction maintenance. *The Journal of Neuroscience* **36**:4106–4120. DOI: <https://doi.org/10.1523/JNEUROSCI.3529-15.2016>, PMID: 27053216
- de Bruin RG**, Shiue L, Prins J, de Boer HC, Singh A, Fagg WS, van Gils JM, Duijs JMGJ, Katzman S, Kraaijeveld AO, Böhringer S, Leung WY, Kielbasa SM, Donahue JP, van der Zande PHJ, Sijbom R, van Alem CMA, Bot I, van Kooten C, Jukema JW, et al. 2016. Quaking promotes monocyte differentiation into pro-atherogenic macrophages by controlling pre-mRNA splicing and gene expression. *Nature Communications* **7**:10846. DOI: <https://doi.org/10.1038/ncomms10846>, PMID: 27029405
- Demichev V**, Messner CB, Vernardis SI, Lilley KS, Ralser M. 2020. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nature Methods* **17**:41–44. DOI: <https://doi.org/10.1038/s41592-019-0638-x>, PMID: 31768060
- Demichev V**. 2025. DiaNN. GitHub. <https://github.com/vdemichev/DiaNN>
- Dignam JD**, Lebovitz RM, Roeder RG. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research* **11**:1475–1489. DOI: <https://doi.org/10.1093/nar/11.5.1475>
- Dominski Z**, Kole R. 1991. Selection of splice sites in pre-mRNAs with short internal exons. *Molecular and Cellular Biology* **11**:6075–6083. DOI: <https://doi.org/10.1128/mcb.11.12.6075-6083.1991>, PMID: 1944277
- Dyakov BJA**, Kobelke S, Ahmed BR, Wu M, Roth JF, Kasmaeifar V. 2024. Spatial Proteomic Mapping of Human Nuclear Bodies Reveals New Functional Insights into RNA Regulation. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.07.03.601239>
- Ebersberger S**, Hipp C, Mulorz MM, Buchbender A, Hubrich D, Kang H-S, Martínez-Lumbreras S, Kristofori P, Sutandy FXR, Llacsahuanga Allica L, Schönfeld J, Bakisoglu C, Busch A, Hänel H, Tretow K, Welzel M, Di Liddo A, Möckel MM, Zarnack K, Ebersberger I, et al. 2023. FUBP1 is a general splicing factor facilitating 3' splice site recognition and splicing of long introns. *Molecular Cell* **83**:2653–2672. DOI: <https://doi.org/10.1016/j.molcel.2023.07.002>, PMID: 37506698
- Ebersole TA**, Chen Q, Justice MJ, Artzt K. 1996. The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins. *Nature Genetics* **12**:260–265. DOI: <https://doi.org/10.1038/ng0396-260>, PMID: 8589716
- Fagg WS**, Liu N, Fair JH, Shiue L, Katzman S, Donohue JP, Ares M Jr. 2017. Autogenous cross-regulation of Quaking mRNA processing and translation balances Quaking functions in splicing and translation. *Genes & Development* **31**:1894–1909. DOI: <https://doi.org/10.1101/gad.302059.117>
- Fagg WS**, Liu N, Braunschweig U, Pereira de Castro KL, Chen X, Ditmars FS, Widen SG, Donohue JP, Modis K, Russell WK, Fair JH, Weirauch MT, Blencowe BJ, Garcia-Blanco MA. 2022. Definition of germ layer cell lineage alternative splicing programs reveals a critical role for Quaking in specifying cardiac cell fate. *Nucleic Acids Research* **50**:5313–5334. DOI: <https://doi.org/10.1093/nar/gkac327>, PMID: 35544276
- Feng H**, Moakley DF, Chen S, McKenzie MG, Menon V, Zhang C. 2021. Complexity and graded regulation of neuronal cell-type-specific alternative splicing revealed by single-cell RNA sequencing. *PNAS* **118**:2013056118. DOI: <https://doi.org/10.1073/pnas.2013056118>
- Fu XD**, Ares M Jr. 2014. Context-dependent control of alternative splicing by RNA-binding proteins. *Nature Reviews Genetics* **15**:689–701. DOI: <https://doi.org/10.1038/nrg3778>
- Galarneau A**, Richard S. 2005. Target RNA motif and target mRNAs of the Quaking STAR protein. *Nature Structural & Molecular Biology* **12**:691–698. DOI: <https://doi.org/10.1038/nsmb963>, PMID: 16041388
- Galarneau A**, Richard S. 2009. The STAR RNA binding proteins GLD-1, QKI, SAM68 and SLM-2 bind bipartite RNA motifs. *BMC Molecular Biology* **10**:47. DOI: <https://doi.org/10.1186/1471-2199-10-47>, PMID: 19457263
- Giudice J**, Loehr JA, Rodney GG, Cooper TA. 2016. Alternative splicing of four trafficking genes regulates myofiber structure and skeletal muscle physiology. *Cell Reports* **17**:1923–1933. DOI: <https://doi.org/10.1016/j.celrep.2016.10.072>, PMID: 27851958
- Gould GM**, Paggi JM, Guo Y, Phizicky DV, Zinshteyn B, Wang ET, Gilbert WV, Gifford DK, Burge CB. 2016. Identification of new branch points and unconventional introns in *Saccharomyces cerevisiae*. *RNA* **22**:1522–1534. DOI: <https://doi.org/10.1261/rna.057216.116>, PMID: 27473169
- Hafner M**, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**:129–141. DOI: <https://doi.org/10.1016/j.cell.2010.03.009>, PMID: 20371350
- Hall MP**, Nagel RJ, Fagg WS, Shiue L, Cline MS, Perriman RJ, Donohue JP, Ares M Jr. 2013. Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation. *RNA* **19**:627–638. DOI: <https://doi.org/10.1261/rna.038422.113>, PMID: 23525800
- Hardy RJ**, Loushin CL, Friedrich VL Jr, Chen Q, Ebersole TA, Lazzarini RA, Artzt K. 1996. Neural cell type-specific expression of QKI proteins is altered in quakingviable mutant mice. *The Journal of Neuroscience* **16**:7941–7949. DOI: <https://doi.org/10.1523/JNEUROSCI.16-24-07941.1996>, PMID: 8987822
- Hartmuth K**, Urlaub H, Vornlocher H-P, Will CL, Gentzel M, Wilm M, Lührmann R. 2002. Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *PNAS* **99**:16719–16724. DOI: <https://doi.org/10.1073/pnas.262483899>, PMID: 12477934
- Hayakawa-Yano Y**, Suyama S, Nogami M, Yugami M, Koya I, Furukawa T, Zhou L, Abe M, Sakimura K, Takebayashi H, Nakanishi A, Okano H, Yano M. 2017. An RNA-binding protein, Qki5, regulates embryonic

- neural stem cells through pre-mRNA processing in cell adhesion signaling. *Genes & Development* **31**:1910–1925. DOI: <https://doi.org/10.1101/gad.300822.117>, PMID: 29021239
- Heintz C**, Doktor TK, Lanjuin A, Escoubas C, Zhang Y, Weir HJ, Dutta S, Silva-García CG, Bruun GH, Morantte I, Hoxhaj G, Manning BD, Andresen BS, Mair WB. 2017. Splicing factor 1 modulates dietary restriction and TORC1 pathway longevity in *C. elegans*. *Nature* **541**:102–106. DOI: <https://doi.org/10.1038/nature20789>, PMID: 27919065
- Herrmann GK**, Russell WK, Garg NJ, Yin YW. 2021. Poly(ADP-ribose) polymerase 1 regulates mitochondrial DNA repair in an NAD-dependent manner. *Journal of Biological Chemistry* **296**:100309. DOI: <https://doi.org/10.1016/j.jbc.2021.100309>
- Hertel KJ**. 2008. Combinatorial control of exon recognition. *The Journal of Biological Chemistry* **283**:1211–1215. DOI: <https://doi.org/10.1074/jbc.R700035200>, PMID: 18024426
- Hoskins AA**, Friedman LJ, Gallagher SS, Crawford DJ, Anderson EG, Wombacher R, Ramirez N, Cornish VW, Gelles J, Moore MJ. 2011. Ordered and dynamic assembly of single spliceosomes. *Science* **331**:1289–1295. DOI: <https://doi.org/10.1126/science.1198830>, PMID: 21393538
- Hsiao Y**, Zhang H, Li GX, Deng Y, Yu F, Kahrood HV, Steele JR, Schittenhelm RB, Nesvizhskii AI. 2024. Analysis and Visualization of Quantitative Proteomics Data Using FragPipe-Analyst. *bioRxiv*. DOI: <https://doi.org/10.1101/2024.03.05.583643>
- Hwang JY**, Jung S, Kook TL, Rouchka EC, Bok J, Park JW. 2020. rMAPS2: an update of the RNA map analysis and plotting server for alternative splicing regulation. *Nucleic Acids Research* **48**:W300–W306. DOI: <https://doi.org/10.1093/nar/gkaa237>, PMID: 32286627
- Irimia M**, Roy SW. 2008. Evolutionary convergence on highly-conserved 3' intron structures in intron-poor eukaryotes and insights into the ancestral eukaryotic genome. *PLOS Genetics* **4**:e1000148. DOI: <https://doi.org/10.1371/journal.pgen.1000148>, PMID: 18688272
- Irimia M**, Weatheritt RJ, Ellis JD, Parikhshak NN, Gonatopoulos-Pournatzis T, Babor M, Quesnel-Vallières M, Tapial J, Raj B, O'Hanlon D, Barrios-Rodiles M, Sternberg MJE, Cordes SP, Roth FP, Wrana JL, Geschwind DH, Blencowe BJ. 2014. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**:1511–1523. DOI: <https://doi.org/10.1016/j.cell.2014.11.035>, PMID: 25525873
- Ito H**, Fukuda Y, Murata K, Kimura A. 1983. Transformation of intact yeast cells treated with alkali cations. *Journal of Bacteriology* **153**:163–168. DOI: <https://doi.org/10.1128/jb.153.1.163-168.1983>, PMID: 6336730
- Jacewicz A**, Chico L, Smith P, Schwer B, Shuman S. 2015. Structural basis for recognition of intron branchpoint RNA by yeast Msl5 and selective effects of interfacial mutations on splicing of yeast pre-mRNAs. *RNA* **21**:401–414. DOI: <https://doi.org/10.1261/rna.048942.114>, PMID: 25587180
- Jamison SF**, Crow A, Garcia-Blanco MA. 1992. The spliceosome assembly pathway in mammalian extracts. *Molecular and Cellular Biology* **12**:4279–4287. DOI: <https://doi.org/10.1128/mcb.12.10.4279-4287.1992>, PMID: 1383687
- Justice MJ**, Hirschi KK. 2010. The role of quaking in mammalian embryonic development. *Advances in Experimental Medicine and Biology* **693**:82–92. DOI: https://doi.org/10.1007/978-1-4419-7005-3_6, PMID: 21189687
- Kastner B**, Will CL, Stark H, Lührmann R. 2019. Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harbor Perspectives in Biology* **11**:a032417. DOI: <https://doi.org/10.1101/cshperspect.a032417>, PMID: 30765414
- Käufer NF**, Potashkin J. 2000. Analysis of the splicing machinery in fission yeast: a comparison with budding yeast and mammals. *Nucleic Acids Research* **28**:3003–3010. DOI: <https://doi.org/10.1093/nar/28.16.3003>, PMID: 10931913
- Klassen R**, Paluszynski JP, Wemhoff S, Pfeiffer A, Fricke J, Meinhardt F. 2008. The primary target of the killer toxin from *Pichia acaciae* is tRNA(Gln). *Molecular Microbiology* **69**:681–697. DOI: <https://doi.org/10.1111/j.1365-2958.2008.06319.x>, PMID: 18532979
- Knott GJ**, Bond CS, Fox AH. 2016. The DBHS proteins SFPO, NONO and PSPC1: a multipurpose molecular scaffold. *Nucleic Acids Research* **44**:3989–4004. DOI: <https://doi.org/10.1093/nar/gkw271>, PMID: 27084935
- Kondo T**, Furuta T, Mitsunaga K, Ebersole TA, Shichiri M, Wu J, Artzt K, Yamamura K, Abe K. 1999. Genomic organization and expression analysis of the mouse qkl locus. *Mammalian Genome* **10**:662–669. DOI: <https://doi.org/10.1007/s003359901068>, PMID: 10384037
- Kong AT**, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. 2017. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nature Methods* **14**:513–520. DOI: <https://doi.org/10.1038/nmeth.4256>, PMID: 28394336
- Kupfer DM**, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, Roe BA, Murphy JW. 2004. Introns and splicing elements of five diverse fungi. *Eukaryotic Cell* **3**:1088–1100. DOI: <https://doi.org/10.1128/EC.3.5.1088-1100.2004>, PMID: 15470237
- Liao KC**, Chuo V, Fagg WS, Bradrick SS, Pompon J, Garcia-Blanco MA. 2020. The RNA binding protein Quaking represses host interferon response by downregulating MAVS. *RNA Biology* **17**:366–380. DOI: <https://doi.org/10.1080/15476286.2019.1703069>, PMID: 31829086
- Liao K-C**, Chuo V, Fagg WS, Modahl CM, Widen S, Garcia-Blanco MA. 2021. The RNA binding protein Quaking represses splicing of the Fibronectin EDA exon and downregulates the interferon response. *Nucleic Acids Research* **49**:10034–10045. DOI: <https://doi.org/10.1093/nar/gkab732>, PMID: 34428287
- Lin S**, Fu XD. 2007. SR proteins and related factors in alternative splicing. *Advances in Experimental Medicine and Biology* **623**:107–122. DOI: https://doi.org/10.1007/978-0-387-77374-2_7, PMID: 18380343

- Liu Z, Luyten I, Bottomley MJ, Messias AC, Houngrinou-Molango S, Sprangers R, Zanier K, Krämer A, Sattler M. 2001. Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science* **294**:1098–1102. DOI: <https://doi.org/10.1126/science.1064719>, PMID: 11691992
- Liu N, Bowen CM, Shoja MM, Castro de Pereira KL, Dongur LP, Saad A, Russell WK, Broderick TC, Fair JH, Fagg WS. 2022. Comparative analysis of co-cultured amniotic cell-conditioned media with cell-free amniotic fluid reveals differential effects on epithelial-mesenchymal transition and myofibroblast activation. *Biomedicine* **10**:2189. DOI: <https://doi.org/10.3390/biomedicine10092189>, PMID: 36140291
- Love SL, Emerson JD, Koide K, Hoskins AA. 2023. Pre-mRNA splicing-associated diseases and therapies. *RNA Biology* **20**:525–538. DOI: <https://doi.org/10.1080/15476286.2023.2239601>, PMID: 37528617
- Lu Z, Zhang Y, Ku L, Wang H, Ahmadian A, Feng Y. 2003. The quakingviable mutation affects qkl mRNA expression specifically in myelin-producing cells of the nervous system. *Nucleic Acids Research* **31**:4616–4624. DOI: <https://doi.org/10.1093/nar/gkg635>, PMID: 12888522
- Makarov EM, Owen N, Bottrill A, Makarova OV. 2012. Functional mammalian spliceosomal complex E contains SMN complex proteins in addition to U1 and U2 snRNPs. *Nucleic Acids Research* **40**:2639–2652. DOI: <https://doi.org/10.1093/nar/gkr1056>, PMID: 22110043
- Marasco LE, Kornblihtt AR. 2023. The physiology of alternative splicing. *Nature Reviews. Molecular Cell Biology* **24**:242–254. DOI: <https://doi.org/10.1038/s41580-022-00545-z>, PMID: 36229538
- Martínez-Lumbreras S, Morguet C, Sattler M. 2024. Dynamic interactions drive early spliceosome assembly. *Current Opinion in Structural Biology* **88**:102907. DOI: <https://doi.org/10.1016/j.sbi.2024.102907>, PMID: 39168044
- Mayeda A, Krainer AR. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68**:365–375. DOI: [https://doi.org/10.1016/0092-8674\(92\)90477-t](https://doi.org/10.1016/0092-8674(92)90477-t), PMID: 1531115
- Mazin PV, Khaitovich P, Cardoso-Moreira M, Kaessmann H. 2021. Alternative splicing during mammalian organ development. *Nature Genetics* **53**:925–934. DOI: <https://doi.org/10.1038/s41588-021-00851-w>, PMID: 33941934
- Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* **32**:e4792. DOI: <https://doi.org/10.1002/pro.4792>, PMID: 37774136
- Mercer TR, Clark MB, Andersen SB, Brunck ME, Haerty W, Crawford J, Taft RJ, Nielsen LK, Dinger ME, Mattick JS. 2015. Genome-wide discovery of human splicing branchpoints. *Genome Research* **25**:290–303. DOI: <https://doi.org/10.1101/gr.182899.114>, PMID: 25561518
- Merkin J, Russell C, Chen P, Burge CB. 2012. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* **338**:1593–1599. DOI: <https://doi.org/10.1126/science.1228186>, PMID: 23258891
- Michaud S, Reed R. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes & Development* **5**:2534–2546. DOI: <https://doi.org/10.1101/gad.5.12b.2534>, PMID: 1836445
- Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNAs by the spliceosome. *Cold Spring Harbor Monograph Series* **24**:303. DOI: <https://doi.org/10.1101/0.525-560>
- Munding EM, Igel AH, Shiue L, Dorigi KM, Treviño LR, Ares M Jr. 2010. Integration of a splicing regulatory network within the meiotic gene expression program of *Saccharomyces cerevisiae*. *Genes & Development* **24**:2693–2704. DOI: <https://doi.org/10.1101/gad.1977410>
- Nakka K, Ghigna C, Gabellini D, Dilworth FJ. 2018. Diversification of the muscle proteome through alternative splicing. *Skeletal Muscle* **8**:8. DOI: <https://doi.org/10.1186/s13395-018-0152-3>, PMID: 29510724
- Nesvizhskii AI, Keller A, Kolker E, Aebersold R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry* **75**:4646–4658. DOI: <https://doi.org/10.1021/ac0341261>, PMID: 14632076
- Noveroske JK, Lai L, Gaussin V, Northrop JL, Nakamura H, Hirschi KK, Justice MJ. 2002. Quaking is essential for blood vessel development. *Genesis* **32**:218–230. DOI: <https://doi.org/10.1002/gene.10060>, PMID: 11892011
- Ohno G, Hagiwara M, Kuroyanagi H. 2008. STAR family RNA-binding protein ASD-2 regulates developmental switching of mutually exclusive alternative splicing in vivo. *Genes & Development* **22**:360–374. DOI: <https://doi.org/10.1101/gad.1620608>, PMID: 18230701
- Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics* **40**:1413–1415. DOI: <https://doi.org/10.1038/ng.259>, PMID: 18978789
- Park JW, Jung S, Rouchka EC, Tseng Y-T, Xing Y. 2016. rMAPS: RNA map analysis and plotting server for alternative exon regulation. *Nucleic Acids Research* **44**:W333–W338. DOI: <https://doi.org/10.1093/nar/gkw410>, PMID: 27174931
- Peled-Zehavi H, Berglund JA, Rosbash M, Frankel AD. 2001. Recognition of RNA branch point sequences by the KH domain of splicing factor 1 (mammalian branch point binding protein) in a splicing factor complex. *Molecular and Cellular Biology* **21**:5232–5241. DOI: <https://doi.org/10.1128/MCB.21.15.5232-5241.2001>, PMID: 11438677
- Prabhala G, Rosenberg GH, Käufer NF. 1992. Architectural features of pre-mRNA introns in the fission yeast *Schizosaccharomyces pombe*. *Yeast* **8**:171–182. DOI: <https://doi.org/10.1002/yea.320080303>, PMID: 1574925
- Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, Na H, Irimia M, Matzat LH, Dale RK, Smith SA, Yarosh CA, Kelly SM, Nabat B, Mecnas D, Li W, et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**:172–177. DOI: <https://doi.org/10.1038/nature12311>, PMID: 23846655

- Reed R. 1990. Protein composition of mammalian spliceosomes assembled in vitro. *PNAS* **87**:8031–8035. DOI: <https://doi.org/10.1073/pnas.87.20.8031>, PMID: 2146679
- Rio DC. 2011. *RNA: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Ryder SP, Williamson JR. 2004. Specificity of the STAR/GSG domain protein Qk1: implications for the regulation of myelination. *RNA* **10**:1449–1458. DOI: <https://doi.org/10.1261/rna.7780504>, PMID: 15273320
- Ryder SP, Massi F. 2010. Insights into the structural basis of RNA recognition by STAR domain proteins. *Advances in Experimental Medicine and Biology* **693**:37–53. DOI: https://doi.org/10.1007/978-1-4419-7005-3_3, PMID: 21189684
- Sasse A, Ray D, Laverty KU, Tam CL, Albu M, Zheng H, Leviansky Y, Lyudovik O, Dalal T, Nie K, Magis C, Notredame C, Valkov E, Weirauch MT, Hughes TR, Morris Q. 2025. A resource of RNA-binding protein motifs across eukaryotes reveals evolutionary dynamics and gene-regulatory function. *Nature Biotechnology*. DOI: <https://doi.org/10.1038/s41587-025-02733-6>, PMID: 40715457
- Schmitzová J, Cretu C, Dienemann C, Urlaub H, Pena V. 2023. Structural basis of catalytic activation in human splicing. *Nature* **617**:842–850. DOI: <https://doi.org/10.1038/s41586-023-06049-w>, PMID: 37165190
- Searle BC, Swearingen KE, Barnes CA, Schmidt T, Gessulat S, Küster B, Wilhelm M. 2020. Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nature Communications* **11**:1548. DOI: <https://doi.org/10.1038/s41467-020-15346-1>, PMID: 32214105
- Selenko P, Gregorovic G, Sprangers R, Stier G, Rhani Z, Krämer A, Sattler M. 2003. Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. *Molecular Cell* **11**:965–976. DOI: [https://doi.org/10.1016/s1097-2765\(03\)00115-1](https://doi.org/10.1016/s1097-2765(03)00115-1), PMID: 12718882
- Seraphin B, Rosbash M. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**:349–358. DOI: [https://doi.org/10.1016/0092-8674\(89\)90296-1](https://doi.org/10.1016/0092-8674(89)90296-1), PMID: 2529976
- Sharma S, Kohlstaedt LA, Damianov A, Rio DC, Black DL. 2008. Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nature Structural & Molecular Biology* **15**:183–191. DOI: <https://doi.org/10.1038/nsmb.1375>, PMID: 18193060
- Shen S, Park JW, Lu Z, Lin L, Henry MD, Wu YN, Zhou Q, Xing Y. 2014. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *PNAS* **111**:E5593–E5601. DOI: <https://doi.org/10.1073/pnas.1419161111>, PMID: 25480548
- Shi Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nature Reviews Molecular Cell Biology* **18**:655–670. DOI: <https://doi.org/10.1038/nrm.2017.86>
- Sidrauski C, Cox JS, Walter P. 1996. tRNA ligase is required for regulated mRNA splicing in the unfolded protein response. *Cell* **87**:405–413. DOI: [https://doi.org/10.1016/s0092-8674\(00\)81361-6](https://doi.org/10.1016/s0092-8674(00)81361-6), PMID: 8898194
- Spingola M, Grate L, Haussler D, Ares M Jr. 1999. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *RNA* **5**:221–234. DOI: <https://doi.org/10.1017/S1355838299981682>
- Sugnet CW, Srinivasan K, Clark TA, O'Brien G, Cline MS, Wang H, Williams A, Kulp D, Blume JE, Haussler D, Ares M Jr. 2006. Unusual intron conservation near tissue-regulated exons found by splicing microarrays. *PLOS Computational Biology* **2**:e4. DOI: <https://doi.org/10.1371/journal.pcbi.0020004>, PMID: 16424921
- Talkish J, Igel H, Perriman RJ, Shiue L, Katzman S, Munding EM, Shelansky R, Donohue JP, Ares M Jr. 2019. Rapidly evolving protointrons in *Saccharomyces* genomes revealed by a hungry spliceosome. *PLOS Genetics* **15**:e1008249. DOI: <https://doi.org/10.1371/journal.pgen.1008249>, PMID: 31437148
- Tanackovic G, Krämer A. 2005. Human splicing factor SF3a, but not SF1, is essential for pre-mRNA splicing in vivo. *Molecular Biology of the Cell* **16**:1366–1377. DOI: <https://doi.org/10.1091/mbc.e04-11-1034>, PMID: 15647371
- Tapial J, Ha KCH, Sterne-Weiler T, Gohr A, Braunschweig U, Hermoso-Pulido A, Quesnel-Vallières M, Permanyer J, Sodaei R, Marquez Y, Cozzuto L, Wang X, Gómez-Velázquez M, Rayon T, Manzanera M, Ponomarenko J, Blencowe BJ, Irimia M. 2017. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Research* **27**:1759–1768. DOI: <https://doi.org/10.1101/gr.220962.117>, PMID: 28855263
- Teplova M, Hafner M, Teplov D, Essig K, Tuschl T, Patel DJ. 2013. Structure-function studies of STAR family Quaking proteins bound to their in vivo RNA target sites. *Genes & Development* **27**:928–940. DOI: <https://doi.org/10.1101/gad.216531.113>, PMID: 23630077
- The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**:57–74. DOI: <https://doi.org/10.1038/nature11247>
- Ule J, Blencowe BJ. 2019. Alternative splicing regulatory networks: functions, mechanisms, and evolution. *Molecular Cell* **76**:329–345. DOI: <https://doi.org/10.1016/j.molcel.2019.09.017>, PMID: 31626751
- van der Veer EP, de Bruin RG, Kraaijeveld AO, de Vries MR, Bot I, Pera T, Segers FM, Trompet S, van Gils JM, Roeten MK, Beckers CM, van Santbrink PJ, Janssen A, van Solingen C, Swildens J, de Boer HC, Peters EA, Bijkerk R, Rousch M, Doop M, et al. 2013. Quaking, an RNA-binding protein, is a critical regulator of vascular smooth muscle cell phenotype. *Circulation Research* **113**:1065–1075. DOI: <https://doi.org/10.1161/CIRCRESAHA.113.301302>, PMID: 23963726
- Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen JY, Cody NAL, Dominguez D, Olson S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Bergalet J, Duff MO, Garcia KE, Gelboin-Burkhart C, Hochman M, et al. 2020a. A large-scale binding and functional map of human RNA-binding proteins. *Nature* **583**:711–719. DOI: <https://doi.org/10.1038/s41586-020-2077-3>
- Van Nostrand EL, Pratt GA, Yee BA, Wheeler EC, Blue SM, Mueller J, Park SS, Garcia KE, Gelboin-Burkhart C, Nguyen TB, Rabano I, Stanton R, Sundararaman B, Wang R, Fu X-D, Graveley BR, Yeo GW. 2020b. Principles of

- RNA processing from analysis of enhanced CLIP maps for 150 RNA binding proteins. *Genome Biology* **21**:90. DOI: <https://doi.org/10.1186/s13059-020-01982-9>, PMID: 32252787
- Vernet C, Artzt K. 1997. STAR, a gene family involved in signal transduction and activation of RNA. *Trends in Genetics* **13**:479–484. DOI: [https://doi.org/10.1016/s0168-9525\(97\)01269-9](https://doi.org/10.1016/s0168-9525(97)01269-9), PMID: 9433137
- Wahl MC, Will CL, Lührmann R. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**:701–718. DOI: <https://doi.org/10.1016/j.cell.2009.02.009>, PMID: 19239890
- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**:470–476. DOI: <https://doi.org/10.1038/nature07509>, PMID: 18978772
- Zaffran S, Astier M, Gratecos D, Sémériva M. 1997. The held out wings (how) *Drosophila* gene encodes a putative RNA-binding protein involved in the control of muscular and cardiac activity. *Development* **124**:2087–2098. DOI: <https://doi.org/10.1242/dev.124.10.2087>, PMID: 9169854
- Zakutansky PM, Ku L, Zhang G, Shi L, Li Y, Yao B, Bassell GJ, Read RD, Feng Y. 2024. Isoform balance of the long noncoding RNA NEAT1 is regulated by the RNA-binding protein QKI, governs the glioma transcriptome, and impacts cell migration. *The Journal of Biological Chemistry* **300**:107595. DOI: <https://doi.org/10.1016/j.jbc.2024.107595>, PMID: 39032650
- Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355**:609–614. DOI: <https://doi.org/10.1038/355609a0>, PMID: 1538748
- Zeng Y, Fair BJ, Zeng H, Krishnamohan A, Hou Y, Hall JM, Ruthenburg AJ, Li YI, Staley JP. 2022. Profiling lariat intermediates reveals genetic determinants of early and late co-transcriptional splicing. *Molecular Cell* **82**:4681–4699. DOI: <https://doi.org/10.1016/j.molcel.2022.11.004>, PMID: 36435176
- Zhan X, Lu Y, Shi Y. 2024. Molecular basis for the activation of human spliceosome. *Nature Communications* **15**:6348. DOI: <https://doi.org/10.1038/s41467-024-50785-0>, PMID: 39068178
- Zhuang Y, Weiner AM. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**:827–835. DOI: [https://doi.org/10.1016/0092-8674\(86\)90064-4](https://doi.org/10.1016/0092-8674(86)90064-4), PMID: 3757028
- Zong F-Y, Fu X, Wei W-J, Luo Y-G, Heiner M, Cao L-J, Fang Z, Fang R, Lu D, Ji H, Hui J. 2014. The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. *PLOS Genetics* **10**:e1004289. DOI: <https://doi.org/10.1371/journal.pgen.1004289>, PMID: 24722255