What is the switch for coupling transcription and splicing? : Pol II-CTD phosphorylation, phase separation and beyond

Hiroshi Maita* and Shinichi Nakagawa

RNA Biology laboratory, Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060-0812, Japan

*To whom correspondence should be addressed

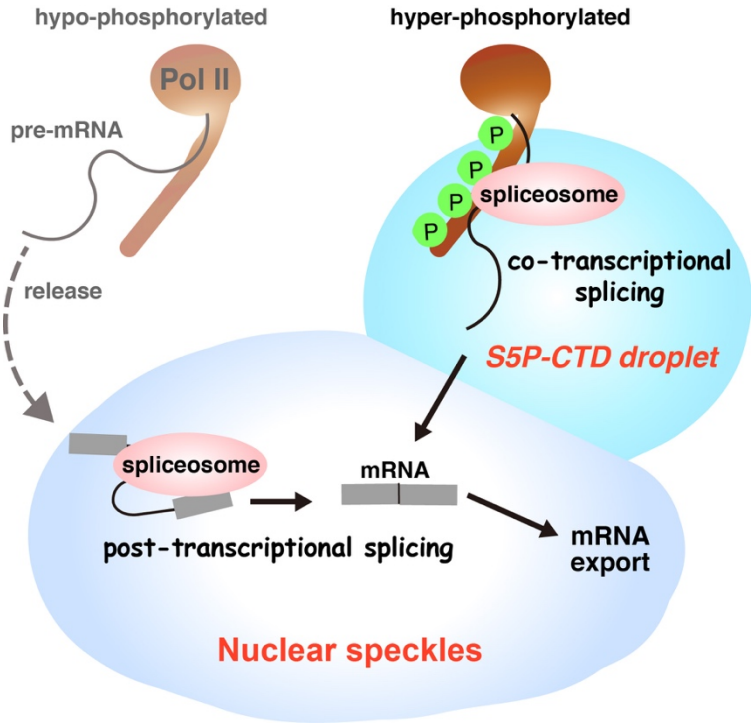
Kita 12 Nishi 6, Kita-ku, Sapporo, Hokkaido 060-0812, Japan
Tel: +81-011-706-3921
Fax: +81-011-706-4988
E-mail: maimai@pharm.hokudai.ac.jp (Hiroshi Maita)
ORCID: https://orcid.org/0000-0002-9620-4651 (Hiroshi Maita)

The authors declare no competing financial interest.

Abstract

Phosphorylation of the RNA Polymerase II C-terminal domain (Pol II CTD) has important roles in the kinetic coupling of splicing with transcription, which is essential for many genes to maintain correct splicing patterns. However, because of the extensively repeated low complexity sequences of Pol II CTD, it was unclear how phosphorylation-dependent molecular interactions were able to provide sufficient specificity to spatiotemporally partition various co-transcriptional events. Here we try to view the molecular mechanisms governing co-transcriptional splicing from the role of phase separation based on recent studies showing the ability of Pol II CTD to form droplets.

Visual summary



Introduction

Pre-mRNA splicing is an essential step for eukaryotic gene expression to remove introns. It is catalyzed by the spliceosome, a macromolecular RNA–protein complex that is highly dynamic, and must be assembled on an intron in every splicing reaction[1]. *In vitro* splicing using an artificially transcribed substrate showed that spliceosomes recognize introns independently from transcription; however, observations that many splicing events occur co-transcriptionally[2–5] indicate that spliceosome assembly is also necessarily a co-transcriptional process.

This process is thought to proceed in a stepwise assembly of pre-built subcomplexes, small nuclear ribonucleoproteins (snRNPs) onto pre-mRNA, but transcription machineries, such as the C-terminal domain of RNA polymerase II (CTD), are also required for the efficient recruitment and assembly of snRNPs[6–10]. Although co-transcriptional splicing is essential for many genes to ensure a correct splicing pattern is achieved[11,12], post-transcriptional splicing is occasionally used for rapid and accurate responses to various stimuli [13–18]. However, the molecular mechanism by which post-transcriptional splicing occurs, particularly in its spatial context, is not fully understood[19].

Importantly, splicing is not only the process coupled with transcription; indeed, capping, RNA editing, and 3′ end processing occur in a co-transcriptional manner[20–24], suggesting that cells have mechanisms to select and turn on individual processing machinery at appropriate times. Several mechanisms have been reported to maintain the accuracy of RNA processing. For example, U1 snRNP was shown to compete with the 3′ end processing complex to protect the transcript by binding potential 5′ splice sites. This so-called “tele-scripting” mechanism is thought to avoid pre-mature processing at the cryptic proximal polyA site[25,26]. Moreover, inhibition of Clk kinase often induces conjoined splicing among transcripts derived from two tandem genes. Thus, Clk kinase normally regulates transcription termination and functions to avoid mis-splicing between transcripts from adjacent genes[27]. These results suggest that cells have multi-layered systems to maintain the accuracy of transcription-coupled RNA processing. Pol II CTD is known to coordinate these coupling mechanisms[10,28], but it remains unclear how CTD and its post-translational modifications regulate co-transcriptional processes. In this review, we provide an overview of the molecular mechanisms governing co-transcriptional splicing from the principle of liquid–liquid phase separation, and discuss future directions to uncover the mechanism ensuring post-transcriptional splicing.

Outline of the spliceosome assembly pathway

At the beginning of a splicing reaction, the spliceosome assembles in a step-wise manner on an intron by recognizing three conserved sequences: the 5′ splice site, 3′ splice site, and the branch site. The 5′ and 3′ splice sites are positioned at the 5′ and 3′ ends of introns, respectively, and the branch site is typically located about 20 bases upstream of the 3′ splice site. Spliceosome subunits such as U1 snRNP, the SF1, and U2AF complex first recognize the 5′ splice site, branch site, and 3′ splice site, respectively, then SF3b and U2AF complexes cooperate to form functional U2 snRNP on the 3′ splice site[29]. The resulting A complex consists of U1 and U2 snRNP on an intron, and is further converted to the pre-catalytic B complex upon joining of U4/U6.U5 tri-snRNP. The association of Prp19 complex followed by the release of U1 and U4 snRNP finally triggers catalytic activation of the spliceosome[1,30–32]. This type of assembly pathway is known as the "intron definition" model [33] (Figure 1A), while the "exon definition" model has also been proposed to explain the splicing mechanism of metazoa, which typically have short exons flanked by much longer introns (Figure 1B).

To accurately identify genuine splice sites from the abundant cryptic splice sites that exist in long introns, marking both ends of an exon by the co-transcriptional deposition of U1 and U2 snRNPs is more useful than scanning the entire intron region. This "exon definition" pathway is guaranteed by the assistance of RNA binding proteins and the epigenetic modification of nucleosomes. Furthermore, Pol II CTD is a scaffold for the exonic region of RNA until the spliceosome is formed[28] (Figure 1B). Recent analysis of the purified spliceosome using cryo-electron microscopy and mass spectrometry (MS) has added further details about the structural and compositional changes that occur during spliceosome assembly (reviewed in[1,34,35]). Current remaining questions include how to integrate this knowledge, which was mainly determined *in vitro*, into the co-transcriptional process, and if the assembly pathway differs between co-transcriptional and post-transcriptional splicing.

CTD phosphorylation ensures the productive elongation of transcription

To understand the switching mechanism between co-transcriptional and post-transcriptional splicing, it is necessary to know the molecular mechanism by which co-transcriptional splicing occurs. Pol II CTD phosphorylation has been shown to play important roles that smoothly couple transcription and spliceosome assembly on nascent RNA (reviewed in [36]). CTD consists of a repeated number of heptad amino acid residues, YSPTSPS, which range from 26 in yeast to 42 in flies, and 52 in humans[37]. All residues, including Ser2, Ser5, Thr4, Ser7, and Tyr1 can be phosphorylated at the heptad repeats[38].

Phosphorylation at distinct positions is thought to function differently in co-transcriptional RNA processing, although phosphorylation levels vary according to position. Moreover, recent MS analysis of CTD phosphorylation revealed that the phosphorylation code is not complicated; rather, single repeats often contain a single phosphorylation[39].

Studies about the molecular interaction networks of CTD, particularly involving the comprehensive identification of phospho-CTD interacting proteins [36,40,41], suggest that phosphorylated CTD acts as a scaffold to assemble RNA processing machineries during transcription. Phosphorylation levels dynamically change together with the relative position of RNA pol II on the gene. For example, Ser2 and Ser5 are the most abundantly phosphorylated residues, but their phosphorylation levels peak at different positions in a gene. Ser5 phosphorylation, which is thought to be mainly mediated by CDK7 in transcription factor TFIIH, reaches a maximum level around the transcription start site (TSS), and is an important cue for transcriptional initiation by releasing Pol II from the mediator complex[42–44]. After initiating transcription, Ser5 phosphorylation levels are thought to be reduced rapidly, and maintained at lower levels until transcription termination for protein-coding genes[38,45]. Ser2 phosphorylation is another cue required to progress transcription into the elongation phase. Shortly after initiation, Pol II normally pauses around 30–60 nucleotides downstream of the TSS in mammalian cells. In the paused state, DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) bind Pol II, thereby replacing initiation factors and blocking transcriptional elongation by changing the conformation of the DNA–RNA hybrid formed on the Pol II funnel[46]. The positive transcription elongation factor (P-TEFb), a complex consisting of CDK9 and cyclin T1, releases Pol II from the paused state with the assistance of PAF and SPT6 [47–49]. The phosphorylation target of P-TEFb is not limited to Pol II CTD, but includes many proteins in the paused elongation complex such as DSIF and NELF[50–52]. Interestingly, a recent study revealed that a linker region of Pol II between the core domain and CTD is also phosphorylated by P-TEFb, and that SPT6 binds the phosphorylated linker region[48]. This interaction places CTD closer to the Pol II core domain and nascent exiting RNA, which is advantageous for the role of CTD as a scaffold to recruit RNA processing factors on nascent RNA[48].

In comparison with mechanisms by which CTD phosphorylation functions at the initial stage of transcription, little is known about the in vivo function of CTD phosphorylation in coupling splicing to transcription. This is possibly because the inhibition of CTD kinases necessarily reduces productive elongation of the Pol II complex, making analysis difficult. Additionally, reduced CTD phosphorylation indirectly affects

the kinetic coupling of transcription and splicing because it decreases elongation rates. The importance of a proper elongation rate to maintain accurate splicing patterns is explained by the "window of opportunity" model[53] and "Goldilocks" model[11,12,53–56]. A study combining DRB and UV treatment in alternative splicing provided interesting observations[57]. UV enhances CTD phosphorylation but also triggers the reduction of elongation rates, while DRB reduces both CTD phosphorylation and elongation rates. Thus, combined DRB and UV treatment would reduce the elongation rate but produce opposing effects on CTD phosphorylation. In some genes, DRB blocks UV-induced alternative splicing change[57], suggesting that CTD phosphorylation has a regulatory role in alternative splicing apart from kinetic coupling. This is consistent with the finding that P-TEFb reduction affects alternative splicing[58], although the detailed molecular roles of CTD phosphorylation in co-transcriptional splicing are unclear.

S5P CTD complex involves the spliceosome

Recently, the novel mammalian native elongating transcript sequencing (mNET-seq) approach was developed to read nascent RNA attached to the elongating Pol II complex. It revealed that the 5′ exon, an intermediate product of the first step reaction of splicing, is associated with Ser5-phosphorylated (S5P) CTD[40,59]. This directly indicates that the existence of S5P in the gene body and the enrichment of splicing intermediates specific to the S5P complex function in co-transcriptional splicing in mammalian cells. In yeast, expression of the Pol II CTD mutant in which the first eight Ser5 residues are replaced with alanine increased intron retention[41], suggesting that the role of S5P in co-transcriptional splicing is conserved between yeast and humans. However, there are likely differences in the detailed molecular mechanisms because the interaction network of phosphorylated CTD largely differs between species to accommodate variations in average numbers and lengths of introns. These observations provide further evidence that CTD phosphorylation plays roles in co-transcriptional splicing other than the regulation of transcription speed.

Although the kinase responsible for phosphorylating S5P in the gene body is still unclear, it is possible that both CDK9 and CDK7 affect Ser5 phosphorylation levels in the gene body among the several potential S5P kinases[36]. CDK9, a catalytic component of the P-TEFb complex, is already known as a Ser2 kinase but is also able to phosphorylate Ser5 *in vitro*[60]. BRD4, a bromo domain-containing protein that binds acetylated histone, is an activator of P-TEFb and stimulates CDK9 to phosphorylate Ser2[61,62]; however, an *in vitro* study revealed that the stimulating function of BRD4 on

P-TEFb is specific for Ser5 when Ser7 is pre-phosphorylated. This discrepancy may be partly explained by the abundant Ser5 phosphorylation at the promotor region by CDK7, which would mask the subtle, locus-specific change of Ser5 phosphorylation induced by CDK9 inhibition at the gene body. Indeed, the simultaneous pharmacological inhibition of both CDK7 and CDK9 significantly reduced S5P levels in the gene body[63,64], while neither the individual inhibition of CDK7 nor CDK9 could induce such a significant reduction. Furthermore, partial phosphorylation is observed when both CDK7 and CDK9 are inhibited, suggesting the existence of an additional Ser5 kinase at the gene body. DYRK1A is another interesting candidate for the Ser5 kinase because it is a known gene-specific kinase that phosphorylates both Ser2 and Ser5; however, the extent and generality of its role in Ser5 phosphorylation remain unclear[65].

Subcellular localization of transcription and splicing machineries

Studies of the subcellular localization of Pol II, spliceosomes, and mRNAs indicate that nuclear speckles connect co-transcriptional splicing and mRNA export[66–68]. Actively elongating hyperphosphorylated Pol II and pre-mRNA containing an intron are known to be co-localized with nuclear speckles[66,69,70]. Detection of the subcellular localization of the active spliceosome using phosphorylated SF3b155 as a marker revealed its localization at the periphery of nuclear speckles under normal conditions. When transcription is blocked by DRB treatment, phospho-SF3b155 was shown to relocate into the inner part of nuclear speckles[71], suggesting that post-transcriptional splicing occurs here. These results imply that nuclear speckles are spatially separated within their layered structure, although another study reported that post-transcriptional splicing can occur in the nucleoplasm[19].

Molecular networks around the S5P Pol II complex resemble nuclear speckles

Comprehensive analysis of the protein composition of the S5P Pol II complex obtained by mNET-seq showed that it contains many factors involved in the catalytically active spliceosome as well as those associated with the pre-catalytic complex (Figure 2). This suggests that S5P CTD functions as a scaffold from the assembly step to catalysis of the spliceosome [40], although U1 factors were poorly detected in contrast to a study using yeast[41]. The manually annotated protein list of nuclear speckles has already been reported[72,73]. Most proteins in this list were identified by MS analysis of purified inter chromatin granule clusters (IGCs), which are thought to be nuclear speckles observed by

electron microscopy. When comparing protein compositions of the S5P Pol II complex and nuclear speckles, about one third of S5P Pol II complex components are identified as nuclear speckles proteins, despite different purification methods. Notably, splicing factors are specifically enriched in the S5P Pol II complex (Figure 2). By contrast, 3′ end processing factors, transcription factors, and heterogeneous nuclear (hn)RNP proteins that are abundantly detected in nuclear speckles are absent from the S5P PolII complex. Interestingly, cyclophilin family peptidyl prolyl isomerases (PPIases) are enriched in the S5P complex compared with IGCs[40]. Cyclophilin family PPIL1, PPIE, PPIL2, PPIL3, and PPIG are core components of the spliceosome-related complex and were detected in the S5P complex (Figure 2). Additionally, PPIL4, another highly conserved cyclophilin known as Rct1 in *Schizosaccharomyces pombe*[74], Sig-7 in *Caenorhabditis elegans*[75], and Atcyp59 in *Arabidopsis thaliana*[76], was also detected in the S5P complex. Genetic and biochemical studies using *C. elegans* showed that the PPIL4 homolog Sig-7 interacts with CTD and functions in co-transcriptional splicing[75], suggesting an additional role for S5P in co-transcriptional splicing.

Collectively, both physical interaction and subcellular localization studies indicate that the periphery of nuclear speckles is the location of co-transcriptional splicing. Pol II may attach to nuclear speckles upon S5P phosphorylation and spliceosome assembly may begin at the periphery of nuclear speckles (as shown in Figure 4), which could explain the core-shell structure and transcription-dependent irregular shape of nuclear speckles[77–79]. If so, it remains to be determined how phosphorylated CTD is selectively recruited to the periphery of nuclear speckles.

Nuclear speckles as a liquid–liquid phase separation-oriented membraneless organelle

Liquid–liquid phase separation has recently been shown to provide force to assemble membraneless organelles (reviewed in [80–82]). Multivalent interactions with lower specificity through intrinsically disordered regions (IDRs) or low complexity domains (LCDs) with no ordered structure lead to the condensation of similar IDR- or LCD-containing proteins, resulting in the formation of characteristic droplets. The molecular property of IDRs and LCDs in droplet formation can be altered by post-translational modification[83–86], enabling cells to control the timing and composition of droplets. Splicing factors within nuclear speckles contain a high proportion of LCDs such as Arg–Ser dipeptide repeats (RS repeats) in SR family RNA binding proteins[72,73]. These repeats undergo extensive and cyclical phosphorylation and also recruit spliceosomes on their target RNA (reviewed in[87]). Because of their droplet-like behavior and protein

composition, nuclear speckles are thought to be a constitutively formed phase-separated subnuclear compartment[88]. However, the spatial pattern of their shape is highly dynamic, and responds to the cellular environment or the levels of RNA accumulation[79]. Because the CTD itself is also a large LCD, it is conceivable that multivalent interactions between CTD and IDR/LCD regions absorb Pol II onto the periphery of nuclear speckles, and that this may be enhanced by Ser5 phosphorylation.

Phase separation property of Pol II CTD

In support of the aforementioned idea, an enhancing role for CTD phosphorylation in droplet formation has been reported[89]. CycT1, a cyclin for CDK9, undergoes phase separation with CTD via its histidine-rich (his-rich) LCD, which allows CDK9 to efficiently phosphorylate CTD[89]. Hyperphosphorylated CTD has an increased ability to form droplets with CycT1 his-rich LCD. Because both CDK9 and CycT1 are known to be localized in nuclear speckles[90], P-TEFb-mediated CTD phosphorylation may occur at the periphery of nuclear speckles, thereby enhancing absorption of the Pol II complex into nuclear speckles by phase separation.

Interestingly, unphosphorylated CTD has also been shown to undergo phase separation. Pol II with its unphosphorylated CTD forms droplets with the mediator complex and many other transcription factors[91] on super-enhancers, and the protein partitioning required for transcriptional activation by phase separation is understood to form part of the molecular basis of the super-enhancer's activity[92–94]. This droplet formation containing an unphosphorylated CTD is repressed when Ser5 is phosphorylated by CDK7[95]. Furthermore, droplets formed by LCDs of FET proteins (FUS, EWS, and TAF15) also bind unphosphorylated CTD and the interactions are abrogated by hyper-phosphorylation that can be potentially mediated by CDK7/9[83,96]. These dephosphorylated state-dependent droplet formations of CTD are thought to occur during the pause function at the initiation phase of transcription, and CTD release from droplets by phosphorylation explains how CDK7 enables transcription initiation[45]. Taken together, these observations suggest that CDK7/9-mediated phosphorylation regulates which droplets CTD joins, as shown in Figure 3. Interestingly, an extension of the CTD repeat length in fly and human cells results in aberrant CTD granules that are more condensed with lower movement[97], suggesting that the length of CTD repeats is finely tuned for appropriate phase separation in the nuclear environment of each species.

CTD phosphorylation directs the compartment where Pol II joins

Core components of the spliceosome have evolutionarily increased the content of IDRs. Interestingly, the IDR distribution in the spliceosomal protein is not equal; rather, early factors in spliceosome assembly contain more IDR[98]. Given that phosphorylated CTD stimulates an *in vitro* splicing reaction[6–8], functioning as a scaffold to mediate weak multivalent interactions of IDR may ensure flexible and efficient spliceosome assembly that processes largely varied introns in higher eukaryotes. Notably, the protein compositions of immunopurified U1 snRNP and Pol II particles from nuclear extracts largely overlap[99], but differ from that of the S5P elongating Pol II complex which lacks most U1 snRNP proteins[40] (Figure 2). Furthermore, FUS which associates with lower phosphorylated Pol II mediates interactions between U1 snRNP and Pol II[100], suggesting that phosphorylation-dependent conversion of interacting proteins plays a key role in early spliceosome assembly. Many exonic regions are marked and wrapped in nucleosomes. Given that BRD4 binds acetylated chromatin[101] and specifically upregulates CDK9 activity[102], CTD phosphorylation levels may increase near exonic regions through the action of P-TEFb.

The loose separation of droplets formed by phosphorylated or unphosphorylated CTD therefore appear to play a role in coupling or uncoupling splicing and other transcriptional processing steps (Figure 4). In consistent with this model, the recent study on this topic has shown that the CTD-unphosphorylated and -phosphorylated form of Pol II colocalize with the enhancer droplet and splicing droplet, respectively, and are separately built around the transcription site. This is the first evidence directly showing that CTD phosphorylation functions as a switch to select which droplet Pol II should join[94].

It remains unclear if S2P and S5P differ in physical properties in terms of phase separation and how the biased distribution of S5P is maintained in Pol II within co-transcriptional splicing. It will also be important to detect changes in S5P and S2P with spatiotemporally higher resolution. Considering these questions, elucidation of the dephosphorylation mechanism is important. Because phosphatases and PPIases functioning in CTD dephosphorylation have a specificity for S2P or S5P[103–107], it is tempting to examine if CTD localization can be affected by modulating not only phosphorylation but also dephosphorylation. Indeed, the overexpression of PPIase Pin1 that preferentially binds Ser5 induces both hyper phosphorylation of CTD and novel droplet formation of Pol II that does not overlap with nuclear speckles[108].

A caveat against the study of post-transcriptional splicing

Post-transcriptional splicing is used in response to environmental stress and neuronal activity or as a feedback mechanism to suppress gene expression noise[13–18]. Most post-transcriptional splicing occurs for a specific intron. Transcripts containing retained introns are observed as nuclear-tethered transcripts, meaning that the target transcript of post-transcriptional splicing must be stored in the nucleus until splicing. However, the mechanism of tethering these transcripts within the nucleus is unknown and it is unclear why the specific intron escapes co-transcriptional splicing. The difficulty of studying switching between co-transcriptional and post-transcriptional splicing occurs because no specific marker of post-transcriptional splicing exists as both modes yield the same product if it is a canonical intron. Thus, identification of specific marker RNA of post-transcriptional splicing will aid study of how the intron is relieved from co-transcriptional splicing in the phosphorylated CTD-dependent droplet.

Non-canonical splicing, a possible marker of post-transcriptional splicing?

Recently, notable progress has been made in *in silico* analysis of splicing from RNA-seq, leading to the identification of a wide variety of non-canonically spliced transcripts. Unexpectedly, non-canonical splicing, thought to be a rare and accidental event, is now recognized as a widespread and evolutionarily conserved occurrence. Furthermore, it is likely that cells actively use non-canonical splicing under strict regulation (reviewed in [109]). Notably, it has been proposed that particular types of splicing, such as back splicing which produces circular RNA (circRNA) and recursive splicing (RS) which removes very long introns, include post-transcriptional splicing because their reaction mechanisms or timing are inconsistent with co-transcriptional splicing[110,111].

Recursive splicing, a type of non-canonical splicing occurring in a post-transcriptional manner

RS is a type of non-canonical splicing that occurs at very long introns[112–115]. Normally, recursively spliced introns contain a tandem allele of 3′ and 5′ splice sites (ss) in the middle of the intron. When the upstream intron region is spliced out, 5′ ss become functional for the second splicing to enable removal of the long intron (Figure 5). Comprehensive analysis of introns removed by RS revealed several important features, including that over half of RS splice sites precede a cryptic exon (see “RS exon” in Figure 5)[116]. The role of the RS exon is thought to aid recognition of RS 3′ss by the exon definition mechanism. RS exons also have specific features to avoid inclusion[117].

Importantly, a time course study of RS splicing based on 4sUDRB-seq revealed that most RS splicing is not co-transcriptional[111]. Although the average transcription speed of human genes is around 2 to 2.5 kb/min, the transcription of RS genes does not take an hour despite the length of RS introns. For 90% of RS genes, junction reading indicating the beginning of RS splicing was detected 1 h after the restart of transcription, while the transcription of 44% of RS genes terminates within 30 min. This clearly shows that RS splicing occurs post-transcriptionally. Intriguingly, over 20,000 putative RS introns are identified by genomic features such as length and motif; however, only 378 actual introns indicate the marks of RS splicing[111]. It remains unclear why only a small portion of putative RS introns are processed by RS splicing. Interestingly, human RS splicing is cell type-specific when compared with that of the fly[111115], suggesting that additional regulatory mechanisms may exist for selecting RS splicing.

circRNA formation, another candidate for non-canonical and post-transcriptional splicing

Recently, progress has been made in identifying novel vertebrate circRNAs. As well as the first reported physiological function of circRNA as a molecular sponge for microRNA[118,119], circRNA is also known to function in various processes such as the modulation of transcription[120] and alternative splicing[121]. Furthermore, cytoplasmic circRNAs have been shown to be translated[122–124]. Existence of these physiologically functional circRNA suggests that circRNA production is a biologically desirable process to adjust to environmental changes or cell fate. Although circRNA production pathways are grouped into two classes[125] (Figure 6), circRNA formation always yields an irregular junction sequence that connects the 3′ end of an exon to the 5′ end of an upstream exon. Sequencing reads involving this region are not normally mapped to the genome and accumulate as unmapped reads. Recent improved algorithms with fewer false positives have expanded the repertoire of circRNAs, and the resulting collection allows the comprehensive analysis of cis-elements and trans-factors that enhance back splicing in circRNA formation[126–128].

An example of those features is the insertion of duplicated inverted repeat sequences that form secondary structures to maintain the splice site pair adjacent to each other, thereby enhancing the efficiency of back splicing (Figure 6)[129]. Intriguingly, as well as cis features that enhance back splicing, the absence of certain trans factors such as spliceosome factors, also enhances back splicing[110]. This suggests that a delay of co-transcriptional spliceosome assembly through a shortage of building blocks enhances circRNA formation, which is consistent with the fact that circRNA formation is a post-

transcriptional process. However, a biochemical fractionation study showed that circRNA is linked with chromatin-associated fractions[130]. Although it is unclear if chromatin fractions always contain nascent RNA, splicing of circRNA biogenesis may not be a single pathway. Nevertheless, the detection of cellular levels of circRNA may reflect the ability of the cell to carry out post-transcriptional splicing. Thus quantification of circRNA production may become a better way to test if the intranuclear environment is supportive for co-transcriptional splicing or if it switches to post-transcriptional splicing in the condition such that Ser5 phosphorylation of CTD is reduced in gene-body region, which would cause inefficient phase separation that enhances co-transcriptional splicing.

Conclusion

It appears that phosphorylation-based phase separation of CTD within nuclear speckles triggers efficient spliceosome assembly, thereby switching on co-transcriptional splicing. In this case, the definition of "co-transcriptional" may be more complicated than before because it is difficult to distinguish post-transcriptionally spliced introns by fractionation studies or pulse chase experiments if introns are spatially spliced in the same way but temporally uncoupled from transcription or *vice versa*. Nonetheless, phase separation-based coupling mechanisms of transcription and splicing allow us to imagine that the reduction of CTD phosphorylation increases post-transcriptional splicing by cancelling co-transcriptional assembly of the spliceosome. Recent progress in identifying non-canonical introns indicates that the unusual products yielded by post-transcriptional splicing may be valuable markers to determine if the inhibition of CDK9 or other CTD kinases affects the rates of post-transcriptional splicing. This may be a useful tool for future studies that elucidate how the cell regulates the co-transcriptional and post-transcriptional mode of splicing.

Acknowledgments

We thank Sarah Williams, PhD, and Sandra Cheesman, PhD, from Edanz Group (www.edanzediting.com) for editing a draft of this manuscript.This study was supported by JSPS KAKENHI Grant Number 16K08225 granted to H.M., and Toray Science and Technology Grant and JSPS KAKENHI Grant Number 17H03604 granted to S.N..

Reference

1. Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol*. 2011;3(7):1-2. doi:10.1101/cshperspect.a003707
2. Beyer AL, Osheim YN. Splice site selection, rate of splicing, and alternative

splicing on nascent transcripts. *Genes Dev*. 1988;2(6):754-765. doi:10.1101/gad.2.6.754

3. Baurén G, Wieslander L. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell*. 1994;76(1):183-192. doi:10.1016/0092-8674(94)90182-1
4. Jiménez-García LF, Spector DL. In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism. *Cell*. 1993;73(1):47-59. doi:10.1016/0092-8674(93)90159-N
5. Brugiolo M, Herzel L, Neugebauer KM. Counting on co-transcriptional splicing. *F1000Prime Rep*. 2013;5(April):1-6. doi:10.12703/p5-9
6. David CJ, Boyne AR, Millhouse SR, Manley JL. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev*. 2011;25(9):972-982. doi:10.1101/gad.2038011
7. Millhouse S, Manley JL. The C-terminal domain of RNA polymerase II functions as a phosphorylation-dependent splicing activator in a heterologous protein. *Mol Cell Biol*. 2005;25(2):533-544. doi:10.1128/MCB.25.2.533-544.2005
8. Hirose Y, Tacke R, Manley JL. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev*. 1999;13(10):1234-1239. doi:10.1101/gad.13.10.1234
9. Herzel L, Ottoz DSM, Alpert T, Neugebauer KM. Splicing and transcription touch base: Co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol*. 2017;18(10):637-650. doi:10.1038/nrm.2017.63
10. Hsin JP, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev*. 2012;26(19):2119-2137. doi:10.1101/gad.200303.112
11. De La Mata M, Alonso CR, Kadener S, et al. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell*. 2003;12(2):525-532. doi:10.1016/j.molcel.2003.08.001
12. Dujardin G, Lafaille C, de la Mata M, et al. How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Mol Cell*. 2014;54(4):683-690. doi:10.1016/j.molcel.2014.03.044
13. Boutz PL, Bhutkar A, Sharp PA. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev*. 2015;29(1):63-80. doi:10.1101/gad.247361.114
14. Braun CJ, Stanciu M, Boutz PL, et al. Coordinated Splicing of Regulatory

Detained Introns within Oncogenic Transcripts Creates an Exploitable Vulnerability in Malignant Glioma. *Cancer Cell*. 2017;32(4):411-426.e11. doi:10.1016/j.ccell.2017.08.018

15. Mauger O, Lemoine F, Scheiffele P. Targeted Intron Retention and Excision for Rapid Gene Regulation in Response to Neuronal Activity. *Neuron*. 2016;92(6):1266-1278. doi:10.1016/j.neuron.2016.11.032
16. Shalgi R, Hurt JA, Lindquist S, Burge CB. Widespread inhibition of posttranscriptional splicing shapes the cellular transcriptome following heat shock. *Cell Rep*. 2014;7(5):1362-1370. doi:10.1016/j.celrep.2014.04.044
17. Ninomiya K, Kataoka N, Hagiwara M. Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. *J Cell Biol*. 2011;195(1):27-40. doi:10.1083/jcb.201107093
18. Hansen MMK, Wen WY, Ingerman E, et al. A Post-Transcriptional Feedback Mechanism for Noise Suppression and Fate Stabilization. *Cell*. 2018;173(7):1609-1621.e15. doi:10.1016/j.cell.2018.04.005
19. Vargas DY, Shah K, Batish M, et al. Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell*. 2011;147(5):1054-1065. doi:10.1016/j.cell.2011.10.024
20. Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA Processingwith Transcription. *Cell*. 2002;108(4):501-512. doi:10.1109/JLT.2016.2552719
21. Martinez-Rucobo FW, Kohler R, van de Waterbeemd M, et al. Molecular Basis of Transcription-Coupled Pre-mRNA Capping. *Mol Cell*. 2015;58(6):1079-1089. doi:10.1016/j.molcel.2015.04.004
22. Millevoi S, Vagner S. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res*. 2009;38(9):2757-2774. doi:10.1093/nar/gkp1176
23. Bentley DL. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet*. 2014;15(3):163-175. doi:10.1038/nrg3662
24. Rodriguez J, Menet JS, Rosbash M. Nascent-Seq Indicates Widespread Cotranscriptional RNA Editing in Drosophila. *Mol Cell*. 2012;47(1):27-37. doi:10.1016/j.molcel.2012.05.002
25. Kaida D, Berg MG, Younis I, et al. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*. 2010;468(7324):664-668. doi:10.1038/nature09479
26. Berg MG, Singh LN, Younis I, et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell*. 2012;150(1):53-64.

doi:10.1016/j.cell.2012.05.029
27. Funnell T, Tasaki S, Oloumi A, et al. CLK-dependent exon recognition and conjoined gene formation revealed with a novel small molecule inhibitor. *Nat Commun*. 2017;8(7):1-15. doi:10.1038/s41467-016-0008-7
28. Dye MJ, Gromak N, Proudfoot NJ. Exon tethering in transcription by RNA polymerase II. *Mol Cell*. 2006;21(6):849-859. doi:10.1016/j.molcel.2006.01.032
29. Gozani O, Potashkin J, Reed R. A Potential Role for U2AF-SAP 155 Interactions in Recruiting U2 snRNP to the Branch Site. *Mol Cell Biol*. 1998;18(8):4752-4760. doi:10.1128/mcb.18.8.4752
30. Chan SP, Kao DI, Tsai WY, Cheng SC. The Prp19p-associated complex in spliceosome activation. *Science (80-)*. 2003;302(5643):279-282. doi:10.1126/science.1086602
31. Mooney HA, Brown BG, Root TL, et al. Small Nuclear Ribonucleoprotein Remodeling During Catalytic Activation of the Spliceosome. *Science (80-)*. 2002;298(December):2205-2208.
32. Wahl MC, Lührmann R. SnapShot: Spliceosome Dynamics III. *Cell*. 2015;162(3):690-690.e1. doi:10.1016/j.cell.2015.07.033
33. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA*. 2013;4(1):49-60. doi:10.1002/wrna.1140
34. Fica SM, Nagai K. Cryo-electron microscopy snapshots of the spliceosome: Structural insights into a dynamic ribonucleoprotein machine. *Nat Struct Mol Biol*. 2017;24(10):791-799. doi:10.1038/nsmb.3463
35. Yan C, Wan R, Shi Y. Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. *Cold Spring Harb Perspect Biol*. 2019;11(1). doi:10.1101/cshperspect.a032409
36. Zaborowska J, Egloff S, Murphy S. The pol II CTD: New twists in the tail. *Nat Struct Mol Biol*. 2016;23(9):771-777. doi:10.1038/nsmb.3285
37. Chapman RD, Heidemann M, Hintermair C, Eick D. Molecular evolution of the RNA polymerase II CTD. *Trends Genet*. 2008;24(6):289-296. doi:10.1016/j.tig.2008.03.010
38. Jeronimo C, Collin P, Robert F. The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain. *J Mol Biol*. 2016;428(12):2607-2622. doi:10.1016/j.jmb.2016.02.006
39. Schüller R, Forné I, Straub T, et al. Heptad-Specific Phosphorylation of RNA Polymerase II CTD. *Mol Cell*. 2016;61(2):305-314. doi:10.1016/j.molcel.2015.12.003

40. Nojima T, Rebelo K, Gomes T, Grosso AR, Proudfoot NJ, Carmo-Fonseca M. RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. *Mol Cell*. 2018;72(2):369-379.e4. doi:10.1016/j.molcel.2018.09.004
41. Harlen KM, Trotta KL, Smith EE, Mosaheb MM, Fuchs SM, Churchman LS. Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Rep*. 2016;15(10):2147-2158. doi:10.1016/j.celrep.2016.05.010
42. Søgaard TMM, Svejstrup JQ. Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J Biol Chem*. 2007;282(19):14113-14120. doi:10.1074/jbc.M701345200
43. Jeronimo C, Robert F. Kin28 regulates the transient association of Mediator with core promoters. *Nat Struct Mol Biol*. 2014;21(5):449-455. doi:10.1038/nsmb.2810
44. Wong KH, Jin Y, Struhl K. TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape. *Mol Cell*. 2014;54(4):601-612. doi:10.1016/j.molcel.2014.03.024
45. Harlen KM, Churchman LS. The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat Rev Mol Cell Biol*. 2017;18(4):263-273. doi:10.1038/nrm.2017.10
46. Vos SM, Farnung L, Urlaub H, Cramer P. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature*. 2018;560(7720):601-606. doi:10.1038/s41586-018-0442-2
47. Marshall NF, Price DH. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem*. 1995;270(21):12335-12338. doi:10.1074/jbc.270.21.12335
48. Vos SM, Farnung L, Boehning M, et al. Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature*. 2018;560(7720):607-612. doi:10.1038/s41586-018-0440-4
49. Fischer WH, Fang S-M, Jones KA, Wei P, Garber ME. A Novel CDK9-Associated C-Type Cyclin Interacts Directly with HIV-1 Tat and Mediates Its High-Affinity, Loop-Specific Binding to TAR RNA. *Cell*. 1998;92:451-462. http://ac.els-cdn.com/S0092867400809393/1-s2.0-S0092867400809393-main.pdf?_tid=932d212a-dc81-11e4-8f42-00000aacb361&acdnat=1428341160_5fc92f25862e477fa4652bf3eb7c881f.

50. Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. P-TEFb Phosphorylates RD and Dissociates Negative Effectors from the Transactivation Response Element. *Mol Cell Biol*. 2004;24(2):787-795. doi:10.1128/MCB.24.2.787
51. Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell*. 2006;21(2):227-237. doi:10.1016/j.molcel.2005.11.024
52. Marshall NF, Peng J, Xie Z, Price DH, Chem JB. Control of RNA Polymerase II Elongation Potential by a Novel Carboxyl-terminal Domain Kinase * tiation activity of positive transcription elongation fac-. *October*. 1996;271(43):27176-27183.
53. Roberts GC, Gooding C, Mak HY, Proudfoot NJ, Smith CW. Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res*. 1998;26(24):5568-5572. http://www.ncbi.nlm.nih.gov/pubmed/9837984%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC148035.
54. Pandya-Jones A., Black D.L. Co-transcriptional splicing of constitutive and alternative exons. *Rna*. 2009;15(10):1896-1908. doi:10.1261/rna.1714509
55. Ip J, Schmidt D, Pan Q. Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome …*. 2011;(416):390-401. doi:10.1101/gr.111070.110.390
56. Fong N, Kim H, Zhou Y, et al. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev*. 2014;28(23):2663-2676. doi:10.1101/gad.252106.114
57. Muñoz MJ, Santangelo MSP, Paronetto MP, et al. DNA Damage Regulates Alternative Splicing through Inhibition of RNA Polymerase II Elongation. *Cell*. 2009;137(4):708-720. doi:10.1016/j.cell.2009.03.010
58. Barboric M, Lenasi T, Chen H, Johansen EB, Guo S, Peterlin BM. 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc Natl Acad Sci*. 2009;106(19):7798-7803. doi:10.1073/pnas.0903188106
59. Nojima T, Gomes T, Grosso ARF, et al. Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell*. 2015;161(3):526-540. doi:10.1016/j.cell.2015.03.027
60. Czudnochowski N, Bösken CA, Geyer M. Serine-7 but not serine-5

phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nat Commun*. 2012;3(May). doi:10.1038/ncomms1846

61. Moon KJ, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell*. 2005;19(4):523-534. doi:10.1016/j.molcel.2005.06.027
62. Yang Z, Yik JHN, Chen R, et al. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell*. 2005;19(4):535-545. doi:10.1016/j.molcel.2005.06.029
63. Glover-Cutter K, Larochelle S, Erickson B, et al. TFIIH-Associated Cdk7 Kinase Functions in Phosphorylation of C-Terminal Domain Ser7 Residues, Promoter-Proximal Pausing, and Termination by RNA Polymerase II. *Mol Cell Biol*. 2009;29(20):5455-5464. doi:10.1128/MCB.00637-09
64. Larochelle S, Amat R, Glover-Cutter K, et al. Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. *Nat Struct Mol Biol*. 2012;19(11):1108-1115. doi:10.1038/nsmb.2399
65. Vona CD, Bezdan D, Islam ABMMK, et al. Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. *Mol Cell*. 2015;57(3):506-520. doi:10.1016/j.molcel.2014.12.026
66. Dias AP, Dufu K, Lei H, Reed R. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun*. 2010;1(7):1-10. doi:10.1038/ncomms1103
67. Lamond AI, Spector DL. Nuclear speckles: A model for nuclear organelles. *Nat Rev Mol Cell Biol*. 2003;4(8):605-612. doi:10.1038/nrm1172
68. Spector DL, Lamond AI. Nuclear speckles. *Cold Spring Harb Perspect Biol*. 2011;3(2):1-12. doi:10.1101/cshperspect.a000646
69. Mortillaro MJ, Blencowe BJ, Wei X, et al. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc Natl Acad Sci*. 1996;93(16):8253-8257. doi:10.1073/pnas.93.16.8253
70. Bregman DB, Du L, Van Der Zee S, Warren SL. Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J Cell Biol*. 1995;129(2):287-298. doi:10.1083/jcb.129.2.287
71. Girard C, Will CL, Peng J, et al. Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat Commun*. 2012;3(994):1-12. doi:10.1038/ncomms1998

72. Saitoh N, Spahr CS, Patterson SD, Bubulya P, Neuwald AF SD. Proteomic Analysis of Interchromatin Granule Clusters. *Mol Biol Cell*. 2004;15(August):3876-3890. doi:10.1091/mbc.E04
73. Galganski L, Urbanek MO, Krzyzosiak WJ. Nuclear speckles: Molecular organization, biological function and role in disease. *Nucleic Acids Res*. 2017;45(18):10350-10368. doi:10.1093/nar/gkx759
74. Chang AY, Castel SE, Ernst E, Kim HS, Martienssen RA. The Conserved RNA Binding Cyclophilin, Rct1, Regulates Small RNA Biogenesis and Splicing Independent of Heterochromatin Assembly. *Cell Rep*. 2017;19(12):2477-2489. doi:10.1016/j.celrep.2017.05.086
75. Ahn JH, Rechsteiner A, Strome S, Kelly WG. A Conserved Nuclear Cyclophilin Is Required for Both RNA Polymerase II Elongation and Co-transcriptional Splicing in Caenorhabditis elegans. *PLoS Genet*. 2016;12(8):1-27. doi:10.1371/journal.pgen.1006227
76. Gullerova M, Barta A, Lorković ZJ. AtCyp59 is a multidomain cyclophilin from Arabidopsis thaliana that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. *Rna*. 2006;12(4):631-643. doi:10.1261/rna.2226106
77. Mintz PJ, Spector DL. Compartmentalization of RNA processing factors within nuclear speckles. *J Struct Biol*. 2000;129(2):241-251. doi:10.1006/jsbi.2000.4213
78. Hall LL, Smith KP, Byron M, Lawrence JB. Molecular anatomy of a speckle. *Anat Rec - Part A Discov Mol Cell Evol Biol*. 2006;288(7):664-675. doi:10.1002/ar.a.20336
79. Fei J, Jadaliha M, Harmon TS, et al. Quantitative analysis of multilayer organization of proteins and RNA in nuclear speckles at super resolution. *J Cell Sci*. 2017;130(24):4180-4192. doi:10.1242/jcs.206854
80. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: Organizers of cellular biochemistry. *Nat Rev Mol Cell Biol*. 2017;18(5):285-298. doi:10.1038/nrm.2017.7
81. Mitrea DM, Kriwacki RW. Phase separation in biology; Functional organization of a higher order Short linear motifs - The unexplored frontier of the eukaryotic proteome. *Cell Commun Signal*. 2016;14(1):1-20. doi:10.1186/s12964-015-0125-7
82. Courchaine EM, Lu A, Neugebauer KM. Droplet organelles? *EMBO J*. 2016;35(15):1603-1612. doi:10.15252/embj.201593517
83. Kwon I, Kato M, Xiang S, et al. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*.

2013;155(5):1049. doi:10.1016/j.cell.2013.10.033
84. Aumiller WM, Keating CD. Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nat Chem*. 2016;8(2):129-137. doi:10.1038/nchem.2414
85. Kato M, Yang YS, Sutter BM, Wang Y, McKnight SL, Tu BP. Redox State Controls Phase Separation of the Yeast Ataxin-2 Protein via Reversible Oxidation of Its Methionine-Rich Low-Complexity Domain. *Cell*. 2019;177(3):711-721.e8. doi:10.1016/j.cell.2019.02.044
86. Yang YS, Kato M, Wu X, et al. Yeast Ataxin-2 Forms an Intracellular Condensate Required for the Inhibition of TORC1 Signaling during Respiratory Growth. *Cell*. 2019;177(3):697-710.e17. doi:10.1016/j.cell.2019.02.043
87. Zhou Z, Fu XD. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma*. 2013;122(3):191-207. doi:10.1007/s00412-013-0407-z
88. Marzahn MR, Marada S, Lee J, et al. Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles. *EMBO J*. 2016;35(12):1254-1275. doi:10.15252/embj.201593169
89. Lu H, Yu D, Hansen AS, et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature*. 2018;558(7709):318-323. doi:10.1038/s41586-018-0174-3
90. Herrmann CH, Mancini M a. The Cdk9 and cyclin T subunits of TAK/P-TEFb localize to splicing factor-rich nuclear speckle regions. *J Cell Sci*. 2001;114(Pt 8):1491-1503. http://www.ncbi.nlm.nih.gov/pubmed/11282025.
91. Boija A, Klein IA, Sabari BR, et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*. 2018;175(7):1842-1855. doi:10.1016/j.cell.2018.10.042
92. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A Phase Separation Model for Transcriptional Control. *Cell*. 2017;169(1):13-23. doi:10.1016/j.cell.2017.02.007
93. Cho WK, Spille JH, Hecht M, et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science (80-)*. 2018;361(6400):412-415. doi:10.1126/science.aar4199
94. Guo YE, Manteiga JC, Henninger JE, et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*. 2019;572:543-548. doi:10.1038/s41586-019-1464-0
95. Boehning M, Dugast-Darzacq C, Rankovic M, et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol*

Biol. 2018;25(9):833-840. doi:10.1038/s41594-018-0112-y

96. Burke KA, Janke AM, Rhine CL, Fawzi NL. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell*. 2015;60(2):231-241. doi:10.1016/j.molcel.2015.09.006
97. Lu F, Portz B, Gilmour DS. The C-Terminal Domain of RNA Polymerase II Is a Multivalent Targeting Sequence that Supports Drosophila Development with Only Consensus Heptads. *Mol Cell*. 2019;73(6):1232-1242.e4. doi:10.1016/j.molcel.2019.01.008
98. Korneta I, Bujnicki JM. Intrinsic Disorder in the Human Spliceosomal Proteome. *PLoS Comput Biol*. 2012;8(8). doi:10.1371/journal.pcbi.1002641
99. Chi B, O'Connell JD, Yamazaki T, Gangopadhyay J, Gygi SP, Reed R. Interactome analyses revealed that the U1 snRNP machinery overlaps extensively with the RNAP II machinery and contains multiple ALS/SMA-causative proteins. *Sci Rep*. 2018;8(1):1-11. doi:10.1038/s41598-018-27136-3
100. Yu Y, Reed R. FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP. *Proc Natl Acad Sci*. 2015;112(28):8608-8613. doi:10.1073/pnas.1506282112
101. Dey A, Chitsaz F, Abbasi A, Misteli T, Ozato K. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci*. 2003;100(15):8758-8763. doi:10.1073/pnas.1433065100
102. Itzen F, Greifenberg AK, Bösken CA, Geyer M. Brd4 activates P-TEFb for RNA polymerase II CTD phosphorylation. *Nucleic Acids Res*. 2014;42(12):7577-7590. doi:10.1093/nar/gku449
103. Krishnamurthy S, He X, Reyes-Reyes M, Moore C, Hampsey M. Ssu72 is an RNA polymerase II CTD phosphatase. *Mol Cell*. 2004;14(3):387-394. doi:10.1016/S1097-2765(04)00235-7
104. Yeo M, Lin PS, Dahmus ME, Gill GN. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J Biol Chem*. 2003;278(28):26078-26085. doi:10.1074/jbc.M301791200
105. Chambers RS, Dahmus ME. Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J Biol Chem*. 1994;269(42):26243-26248.
106. Mosley AL, Pattenden SG, Carey M, et al. Rtr1 Is a CTD Phosphatase that Regulates RNA Polymerase II during the Transition from Serine 5 to Serine 2 Phosphorylation. *Mol Cell*. 2009;34(2):168-178. doi:10.1016/j.molcel.2009.02.025

107. Mayfield JE, Burkholder NT, Zhang YJ. Dephosphorylating eukaryotic RNA polymerase II. *Biochim Biophys Acta - Proteins Proteomics*. 2016;1864(4):372-387. doi:10.1016/j.bbapap.2016.01.007
108. Xu YX, Manley JL. Pin1 modulates RNA polymerase II activity during the transcription cycle. *Genes Dev*. 2007;21(22):2950-2962. doi:10.1101/gad.1592807
109. Sibley CR, Blazquez L, Ule J. Lessons from non-canonical splicing. *Nat Rev Genet*. 2016;17(7):407-421. doi:10.1038/nrg.2016.46
110. Liang D, Tatomer DC, Luo Z, et al. The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting. *Mol Cell*. 2017;68(5):940-954.e3. doi:10.1016/j.molcel.2017.10.034
111. Zhang XO, Fu Y, Mou H, Xue W, Weng Z. The temporal landscape of recursive splicing during Pol II transcription elongation in human cells. *PLoS Genet*. 2018;14(8):1-24. doi:10.1371/journal.pgen.1007579
112. Hatton A, Subramaniam V, Lopez A. Generation of Alternative Ultrabithorax Isoforms and Stepwise Removal of a Large Intron by Resplicing at Exon–Exon Junctions. 1998;2:787-796.
113. Burnette JM, Miyamoto-Sato E, Schaub MA, Conklin J, Lopez AJ. Subdivision of large introns in Drosophila by recursive splicing at nonexonic elements. *Genetics*. 2005;170(2):661-674. doi:10.1534/genetics.104.039701
114. Sibley C, Emmett W, Blazquez L, et al. Recursive splicing in long vertebrate genes. *Nature*. 2015;521:371-375. doi:10.1038/nature14466
115. Duff MO, Olson S, Wei X, et al. Genome-wide identification of zero nucleotide recursive splicing in Drosophila. *Nature*. 2015;521(7552):376-379. doi:10.1038/nature14475
116. Joseph B, Kondo S, Lai EC. Short cryptic exons mediate recursive splicing in Drosophila. *Nat Struct Mol Biol*. 2018;25(5):365-371. doi:10.1038/s41594-018-0052-6
117. Blazquez L, Emmett W, Faraway R, et al. Exon Junction Complex Shapes the Transcriptome by Repressing Recursive Splicing. *Mol Cell*. 2018;72(3):496-509.e9. doi:10.1016/j.molcel.2018.09.033
118. Memczak S, Jens M, Elefsinioti A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 2013;495(7441):333-338. doi:10.1038/nature11928
119. Hansen TB, Jensen TI, Clausen BH, et al. Natural RNA circles function as efficient microRNA sponges. *Nature*. 2013;495(7441):384-388.

doi:10.1038/nature11993
120. Li Z, Huang C, Bao C, et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol*. 2015;22(3):256-264. doi:10.1038/nsmb.2959
121. Conn VM, Hugouvieux V, Nayak A, et al. A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation. *Nat Plants*. 2017;3:17053. https://doi.org/10.1038/nplants.2017.53.
122. Legnini I, Di Timoteo G, Rossi F, et al. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell*. 2017;66(1):22-37.e9. doi:10.1016/j.molcel.2017.02.017
123. Pamudurti NR, Bartok O, Jens M, et al. Translation of CircRNAs. *Mol Cell*. 2017;66(1):9-21.e7. doi:10.1016/j.molcel.2017.02.021
124. Yang Y, Fan X, Mao M, et al. Extensive translation of circular RNAs driven by N 6 -methyladenosine. *Cell Res*. 2017;27(5):626-641. doi:10.1038/cr.2017.31
125. Li X, Yang L, Chen LL. The Biogenesis, Functions, and Challenges of Circular RNAs. *Mol Cell*. 2018;71(3):428-442. doi:10.1016/j.molcel.2018.06.034
126. Tarrero LC, Ferrero G, Miano V, et al. Luminal breast cancer-specific circular RNAs uncovered by a novel tool for data analysis. *Oncotarget*. 2018;9(18):14580-14596. doi:10.18632/oncotarget.24522
127. Smid M, Wilting SM, Uhr K, et al. The circular RNome of primary breast cancer.
128. Chen CY, Chuang TJ. NCLcomparator: Systematically post-screening non-co-linear transcripts (circular, trans-spliced, or fusion RNAs) identified from various detectors. *BMC Bioinformatics*. 2019;20(3):1-11. doi:10.1186/s12859-018-2589-0
129. Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev*. 2014;28(20):2233-2247. doi:10.1101/gad.251926.114
130. Ashwal-Fluss R, Meyer M, Pamudurti NR, et al. CircRNA Biogenesis competes with Pre-mRNA splicing. *Mol Cell*. 2014;56(1):55-66. doi:10.1016/j.molcel.2014.08.019

Figure Legends

Figure 1. Schematic of the spliceosome assembly pathway. (left) Assembly model based on “intron definition”. U1 and U2 snRNP first recognize both ends of an intron. (right) Assembly by the “exon definition” model is superimposed on a cartoon of elongating Pol II. The exonic region of nascent RNA is tethered on Pol II CTD by U1 and U2 snRNPs. Nucleosomes often contain exonic regions, and the reduction of elongation rates at the

nucleosome may induce deposition of both U1 and U2 snRNP on exon ends leading to tethering of the exon. At this time, the direction of U1 and U2 is opposite to the intron definition model. This complex further accepts tri-snRNP to form a B-like complex that can be converted to a catalytically active spliceosome when paired to the upstream or downstream exon.

Figure 2. Comparison of spliceosome factors detected in the S5P complex or nuclear speckles. Gray boxes indicate that the protein is in the list. Proteins colored by light green or orange are specifically listed in nuclear speckles or S5P complex, respectively. The classification of spliceosome factors is according to reference 27.

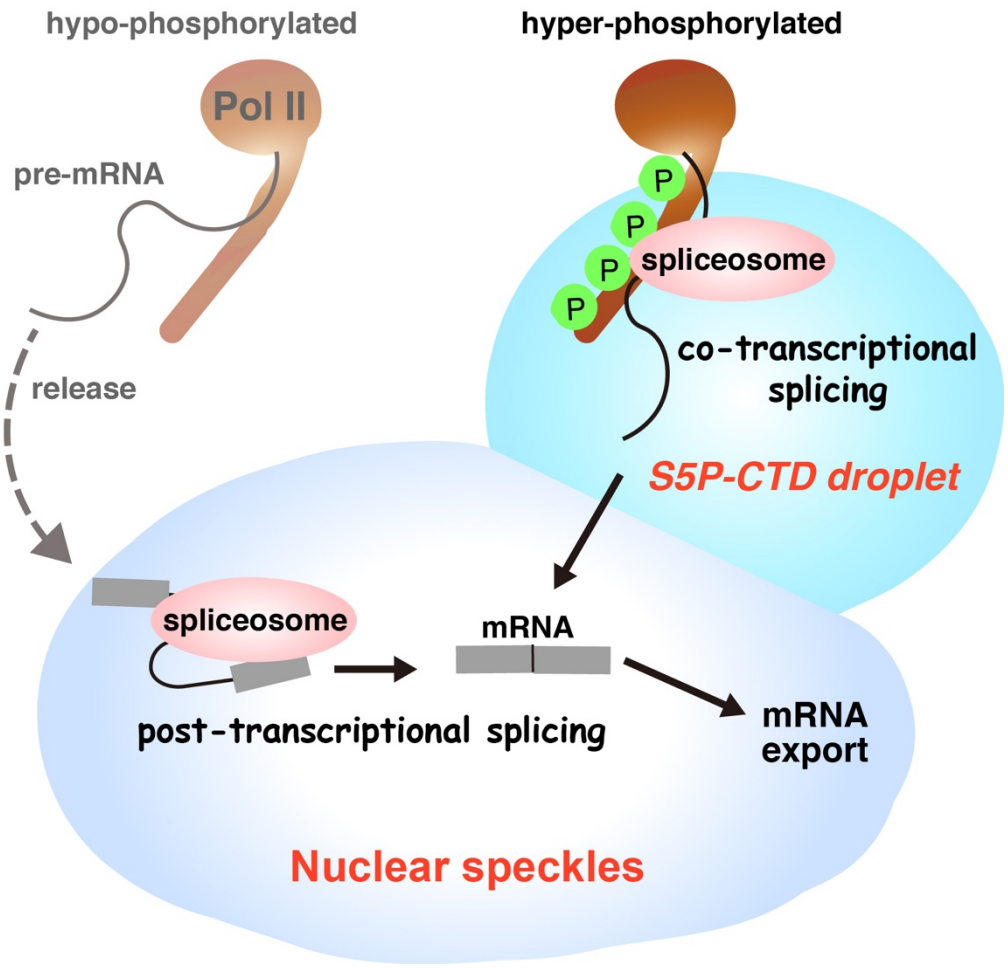
Figure 3. Depiction of independent droplet formation by phosphorylated and unphosphorylated CTD. Phosphorylation of CTD triggers its movement to different classes of phase separated droplets.

Figure 4. CTD undergoes phosphorylation-dependent loose separation in nuclear speckles. Pol II may be incorporated into droplets depending on its phosphorylation levels. If Ser5 is highly phosphorylated, the droplet in which CTD is incorporated preferentially recruits spliceosome factors compared with other processing factors. In the paused state, less-phosphorylated Pol II may be trapped in U1 and FUS-enriched droplets which protect nascent RNA. Nuclear speckles store many RNA-binding proteins so spliced RNA and released transcripts containing an intron may be tethered deep inside nuclear speckles until export.

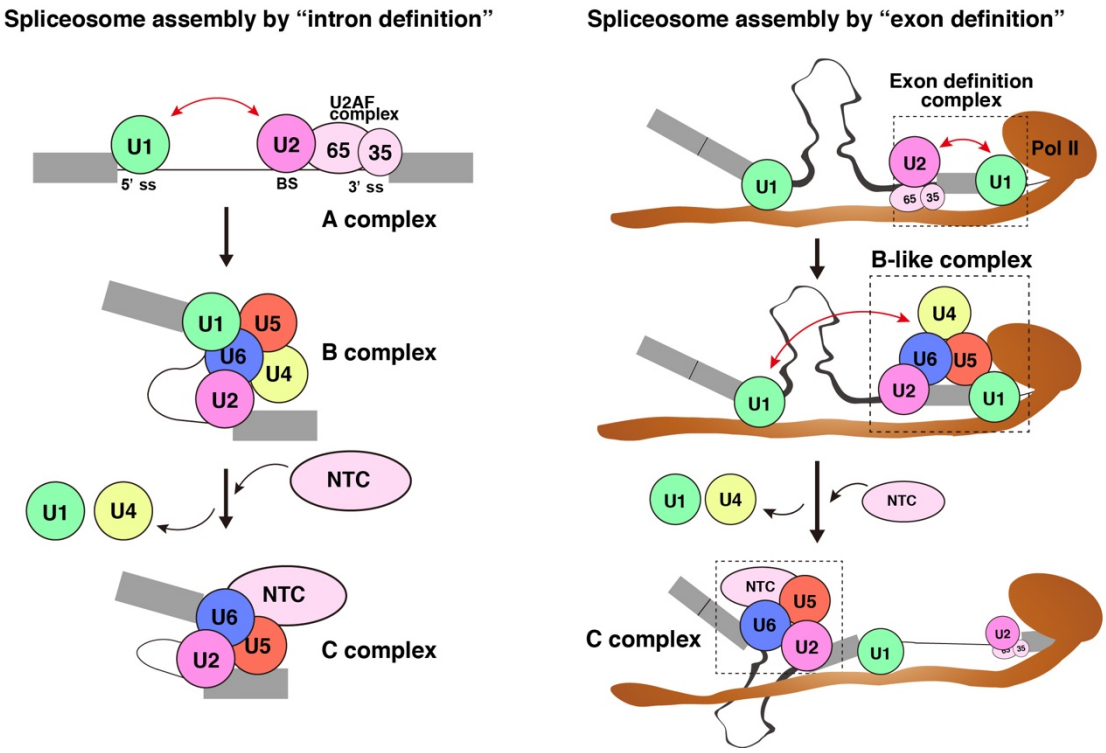
Figure 5. Schematic of recursive splicing. Recursive splicing is divided into two steps. The RS exon is important in the first step but is removed at the second. RS exons have several features compared with canonical exons.

Figure 6. Schematic of two possible circRNA formation pathways. circRNA formation occurs in independent pathways. In model A, a cis-element and an RNA-binding protein cause incorrect spliceosome assembly. In model B, re-splicing of lariat introns containing skipped exons causes back splicing.

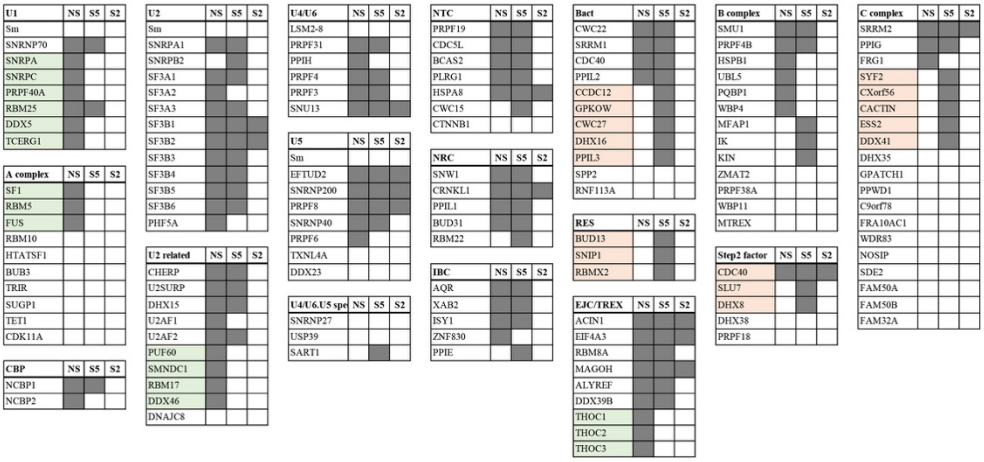
Maita and Nakagawa, Visual summary



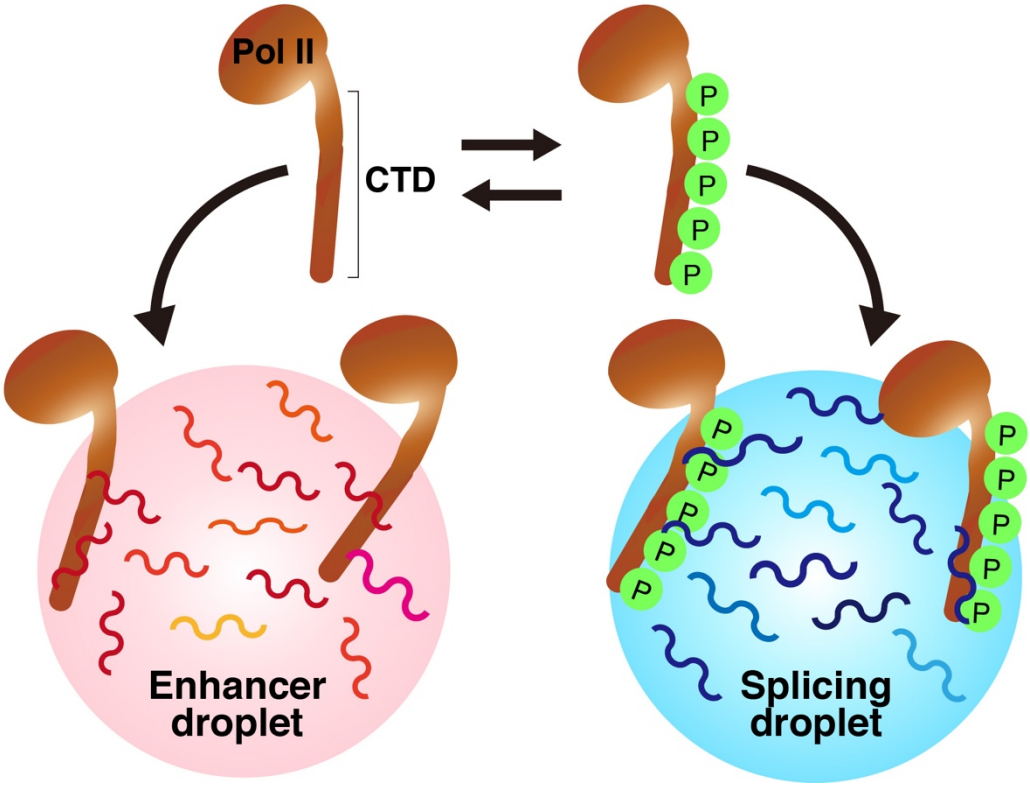
Maita and Nakagawa, Figure 1



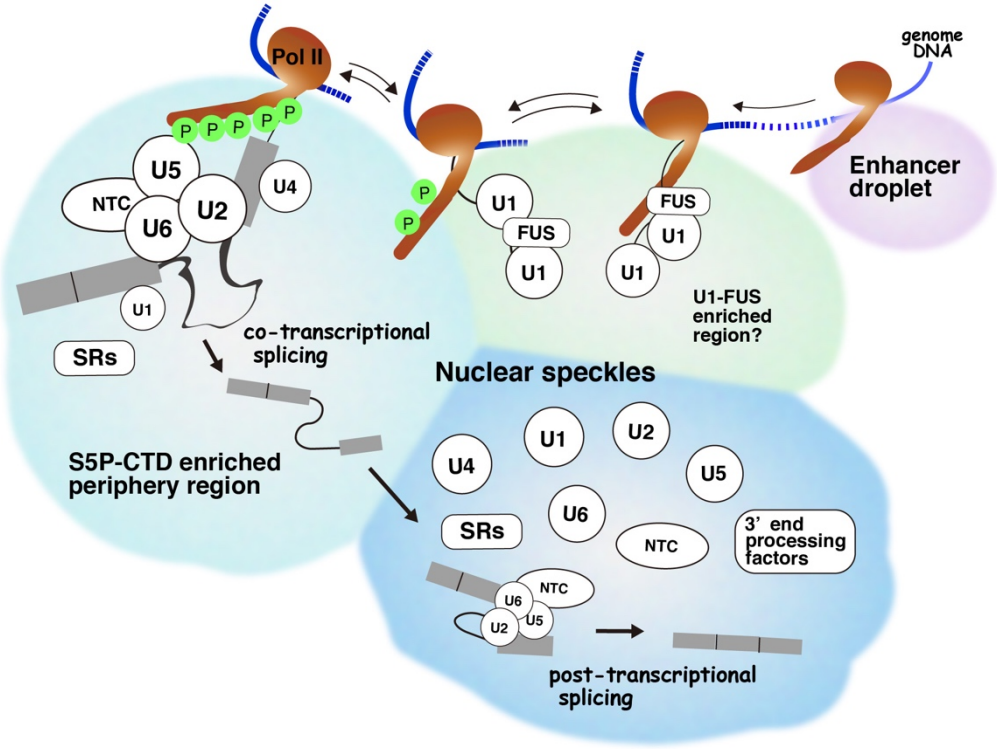
Maita and Nakagawa, Figure 2



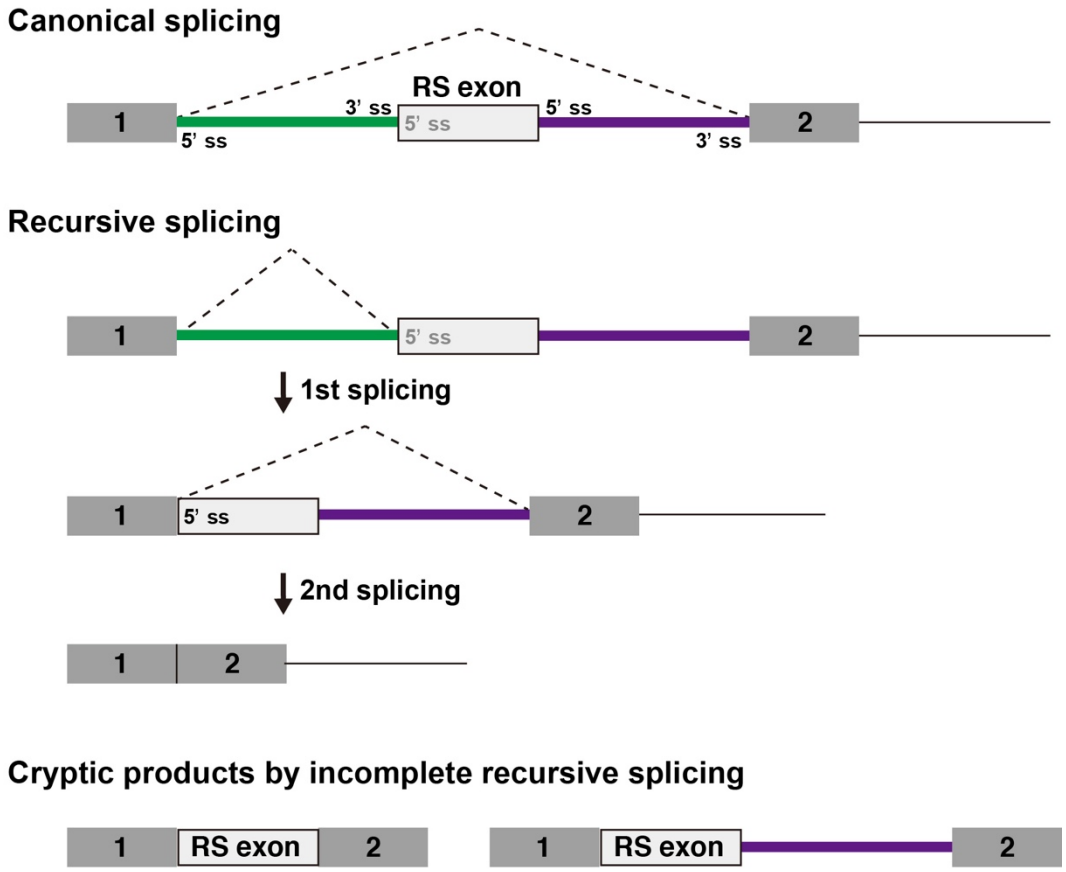
Maita and Nakagawa, Figure 3



Maita and Nakagawa, Figure 4



Maita and Nakagawa, Figure 5



Maita and Nakagawa, Figure 6