frontiers
in Molecular Biosciences

Gene Architecture Facilitates Intron-Mediated Enhancement of Transcription

Katherine Dwyer, Neha Agarwal, Lori Pile and Athar Ansari*

Department of Biological Science, Wayne State University, Detroit, MI, United States

Introns impact several vital aspects of eukaryotic organisms like proteomic plasticity, genomic stability, stress response and gene expression. A role for introns in the regulation of gene expression at the level of transcription has been known for more than thirty years. The molecular basis underlying the phenomenon, however, is still not entirely clear. An important clue came from studies performed in budding yeast that indicate that the presence of an intron within a gene results in formation of a multi-looped gene architecture. When looping is defective, these interactions are abolished, and there is no enhancement of transcription despite normal splicing. In this review, we highlight several potential mechanisms through which looping interactions may enhance transcription. The promoter-5′ splice site interaction can facilitate initiation of transcription, the terminator-3′ splice site interaction can enable efficient termination of transcription, while the promoter-terminator interaction can enhance promoter directionality and expedite reinitiation of transcription. Like yeast, mammalian genes also exhibit an intragenic interaction of the promoter with the gene body, especially exons. Such promoter-exon interactions may be responsible for splicing-dependent transcriptional regulation. Thus, the splicing-facilitated changes in gene architecture may play a critical role in regulation of transcription in yeast as well as in higher eukaryotes.

Keywords: transcription, splicing, intron, exon, gene architecture, gene regulation, gene looping

INTRODUCTION

Introns are intervening non-coding sequences in eukaryotic genes that are removed from the primary transcripts by the process of splicing (**Figure 1**). An elaborate splicing machinery is needed to remove introns to form the mature transcript (Padgett et al., 1986). The energy, time consumption, and complex nature of the spliceosome indicate that introns impose a large burden on eukaryotic organisms (Jo and Choi, 2015). However, in spite of all of these drawbacks, introns have been evolutionarily conserved, indicative of their having a fundamental and significant role in the cell (Carmel and Chorev, 2012; Rogozin et al., 2012). In fact, nearly half of all common genetic disorders in humans may be attributed to a disruption in the splicing process (Faustino and Cooper, 2003; Wang and Cooper, 2007; Padgett, 2012). An important question therefore is what is the physiological significance of introns in eukaryotes?

Research conducted during the last few decades has revealed novel biological functions of introns. An apparent advantage conferred by introns is their ability to increase proteomic complexity through the process of alternative splicing, meaning that a single gene can produce multiple isoforms of a protein dependent on cell type and environment (Nilsen and Graveley, 2010). This offers immense proteomic plasticity since the cell can modulate what protein isoform is produced depending on the developmental context, cell type and environmental cues (Marquez et al., 2015; Chaudhary et al., 2019). The presence of introns has also been shown to protect genomic integrity. The process of transcription is accompanied by formation of genotoxic R-loops, which are three stranded nucleic acid structures harboring an RNA-DNA hybrid (Niehrs and Luke, 2020). Intron-containing genes alleviate R-loop formation and in doing so confer genomic stability and protect cells from harmful stress responses (Bonnet et al., 2017). Recent studies have led to the identification of yet another novel function of introns. Some excised introns provide protection against environmental stress (Morgan et al., 2019; Parenteau et al., 2019). Introns are also the source of snoRNA, microRNA, and lncRNA (Carmel and Chorev, 2012). These non-coding RNA species regulate gene expression at the level of transcription and RNA stability. Outside of acting as a non-coding regulatory RNA molecule, introns possess the unique function of acting as mutational buffers that can protect coding regions from incurring deleterious mutations (Jo and Choi, 2015). A harmful mutation in the coding region may affect the function of the protein. In contrast, a mutation in introns, which are non-coding regions, has minimal chances of affecting the function of the protein. In human genes, introns make up the bulk of a gene and therefore are able to absorb detrimental mutations without affecting the protein function. Although introns have been implicated in a variety of functions in eukaryotes, not all introns are associated with every function described above.
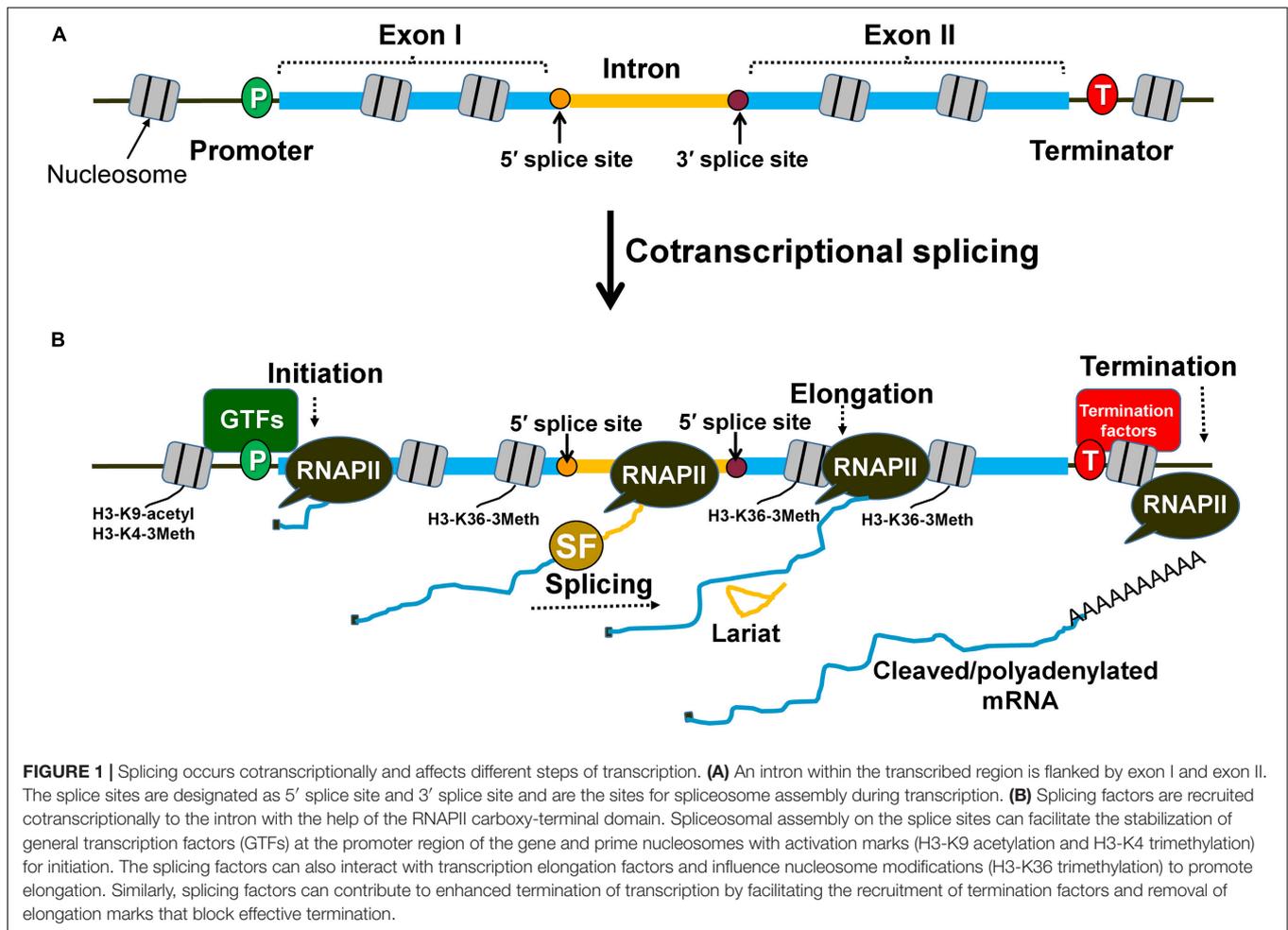
Of all the known functions of introns, one of the best known and evolutionarily conserved, is their ability to regulate the expression of genes that harbor them. Introns affect gene expression at multiple levels. They have been implicated in altering nucleosome positioning, transcription, RNA stability, nucleo-cytoplasmic export of mRNA, and translation efficiency (reviewed in Le Hir et al., 2003; Rose, 2008; Gallegos and Rose, 2015; Laxa, 2017; Shaul, 2017). The details regarding posttranscriptional regulation of gene expression by an intron have been covered in many recently published reviews (Gallegos and Rose, 2015; Laxa, 2017; Shaul, 2017). Here, we highlight the function of introns in regulating gene expression at the level of transcription. The transcription-enhancing potential of introns was first observed in cultured maize cells and transgenic mice (Callis et al., 1987; Brinster et al., 1988). Furthermore, cDNA of a number of human genes are not transcribed to the wild type level unless a promoter-proximal intron is included (Palmiter et al., 1991; Charron et al., 2007). Soon thereafter it was realized that introns play a general role in activating transcription in a variety of eukaryotes. Intron-mediated regulation of transcription has been observed in simple eukaryotes like yeast and Chlamydomonas as well as in higher

eukaryotes like flies, worms, plants, and mammals including humans (Callis et al., 1987; Brinster et al., 1988; Bieberstein et al., 2012; Moabbi et al., 2012; Agarwal and Ansari, 2016; Baier et al., 2020). This enhancement property could be a very crucial function of introns in eukaryotes as a number of genes are dependent on introns for their normal transcription (Rose, 2019). In general, intron-containing genes in eukaryotes exhibit higher expression than their non-intronic counterparts (Ares et al., 1999; Juneau et al., 2006; Bieberstein et al., 2012; Gallegos and Rose, 2015; Ding and Elowitz, 2019; Baier et al., 2020). An intron may enhance transcription by a meager 2–3-fold, or it may augment mRNA output in the range of 10–100-fold or higher depending on the gene. Not all introns, however, can stimulate transcription. Some naturally occurring genes do not contain introns but are expressed efficiently, while some transgenes fail to express even in the presence of an intron (Pasleau et al., 1987; Malim et al., 1988; Rose, 2008). Introns are dispensable for enhancing transcription from a strong promoter (Huang and Liang, 1993). The transcription activation potential of introns though, is crucial for high expression of genes with a weak promoter. Some of the highly expressed genes like H2A and hepatitis B virus genes have cis-acting elements that appear to function like introns (Huang and Liang, 1993; Huang and Yen, 1995; Liu and Mertz, 1995). Despite some introns lacking the transcription enhancement potential, introns in general are emerging as an important component of the transcription regulatory machinery in eukaryotes (Rose, 2019).

Intron-mediated transcriptional regulation can be broadly divided into two categories: (1) splicing-independent, and (2) splicing-dependent regulation. Splicing-independent regulation is due to the presence of an enhancer or a promoter element within the intron (Kaneda et al., 1992; Bianchi et al., 2009; Beaulieu et al., 2011). Such introns can influence transcription even if their splicing function is compromised. In contrast, the splicing-dependent regulation requires a functional, splicing-competent intron within the transcribed region of the gene. Such introns cannot affect transcription if their splicing is inhibited by a mutation in the conserved sequences at the 5′ splice site, 3′ splice site and branchpoint, or if they are inserted in an anti-sense orientation (Furger et al., 2002; Moabbi et al., 2012; Agarwal and Ansari, 2016). This direct effect of a splicing-competent intron on transcription of a gene represents splicing-dependent regulation. It is often referred to as "intron-mediated enhancement" (IME). The focus of this review is to highlight key findings related to the possible mechanism of this enhancement with an emphasis on data indicating that genome architecture plays a critical role in the process.

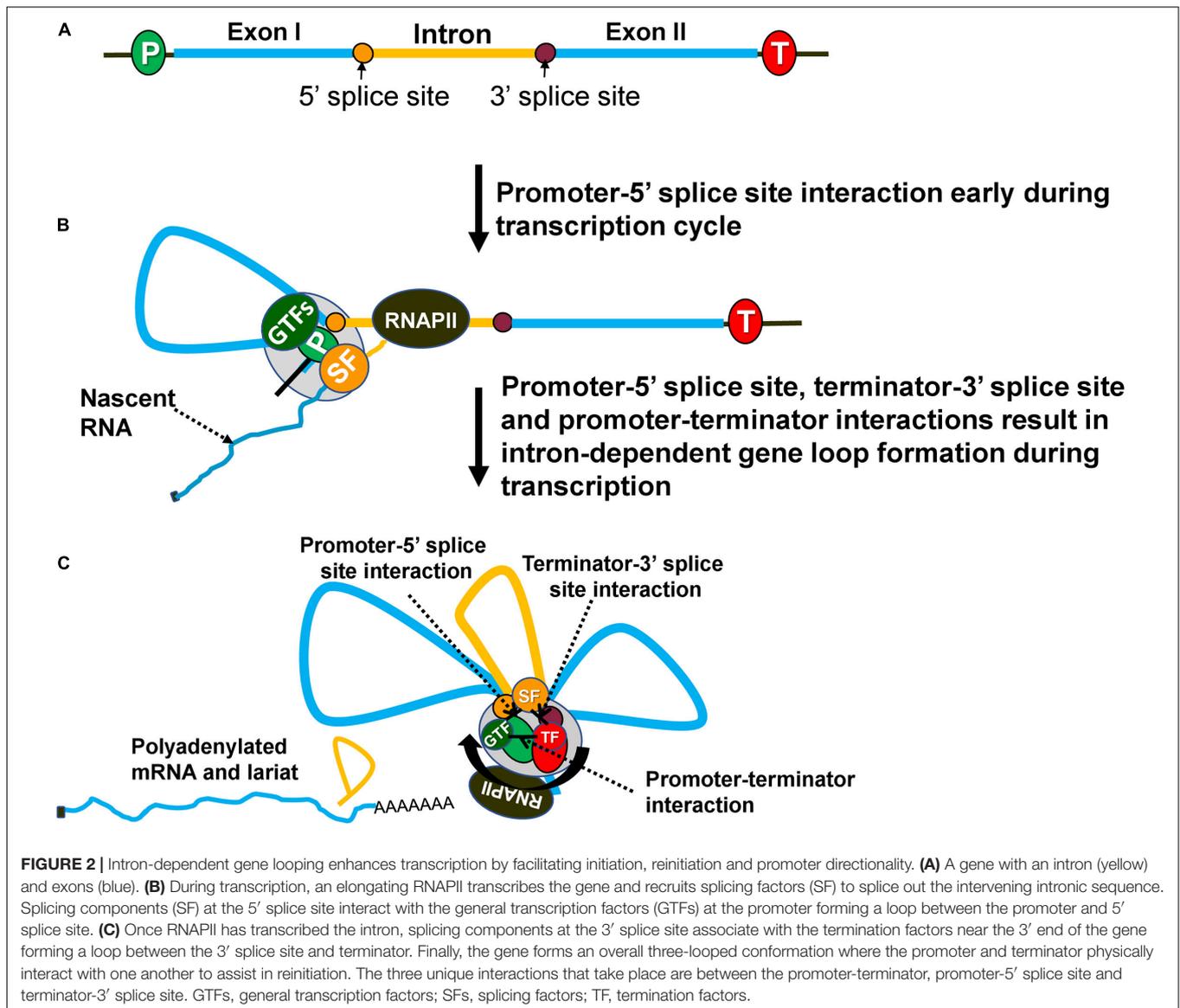## INTRON-MEDIATED ENHANCEMENT EFFECT MAY TARGET INITIATION, ELONGATION OR TERMINATION STEPS

To enhance transcription, the intron must be located within the transcribed region of the gene and should be spliced (**Figures 1A**, **2A**). The activation is maximum when an intron

FIGURE 1 | Splicing occurs cotranscriptionally and affects different steps of transcription. **(A)** An intron within the transcribed region is flanked by exon I and exon II. The splice sites are designated as 5′ splice site and 3′ splice site and are the sites for spliceosome assembly during transcription. **(B)** Splicing factors are recruited cotranscriptionally to the intron with the help of the RNAPII carboxy-terminal domain. Spliceosomal assembly on the splice sites can facilitate the stabilization of general transcription factors (GTFs) at the promoter region of the gene and prime nucleosomes with activation marks (H3-K9 acetylation and H3-K4 trimethylation) for initiation. The splicing factors can also interact with transcription elongation factors and influence nucleosome modifications (H3-K36 trimethylation) to promote elongation. Similarly, splicing factors can contribute to enhanced termination of transcription by facilitating the recruitment of termination factors and removal of elongation marks that block effective termination.

is present close to the promoter, and gradually decreases with increasing distance from the promoter (Callis et al., 1987; Rose, 2004; Bieberstein et al., 2012; Gallegos and Rose, 2017). There are, however, reports that in some genes, the activation potential of the intron is partially restored with increasing proximity to the terminator (Callis et al., 1987; Kim et al., 2011). Thus, the position of an intron within a gene is an important determinant of its transcription activation potential.

Although introns can impact any step of the transcription cycle to achieve higher mRNA output, the initiation step is the most frequent target. The transcription regulating ability of an intron is dependent on the cotranscriptional nature of splicing (**Figure 1B**). The carboxy-terminal domain (CTD) of RNA Polymerase II (RNAPII) is used as a recruitment and docking site for several splicing factors that help in coupling of transcription with splicing (Millhouse and Manley, 2005; Nojima et al., 2018). During cotranscriptional splicing, the spliceosome assembled on the elongating mRNA facilitates initiation/reinitiation by stabilizing assembly of the preinitiation complex (PIC) on the promoter (**Figure 2B**; Tian, 2001; Kwek et al., 2002; Das et al., 2007; Damgaard et al., 2008; Jobert et al., 2009). The 5′ splice site of nascent transcripts plays a crucial role in these processes. In fact, a 5′ splice site alone,

without being a part of an intron, can enhance transcription to some extent (Damgaard et al., 2008). The 5′ splice site enhances recruitment of TFIID, TFIIB, and TFIIH on the promoter (Das et al., 2007; Damgaard et al., 2008). Splicing factors like U1-snRNP and HNRNPU, which bind to the 5′ splice site, facilitate the recruitment of these general transcription factors (Tian, 2001; Das et al., 2007). U1-snRNP physically interacts with RNAPII during cotranscriptional splicing (Nojima et al., 2018). U1-snRNP also contacts the cyclin H subunit of TFIIH (Kwek et al., 2002). U1-snRNP may facilitate initiation and reinitiation of transcription by enabling recruitment of TFIIH, and probably TFIID and TFIIB as well (Kwek et al., 2002; Damgaard et al., 2008). HNRNPU similarly promotes initiation by enabling the recruitment of TFIIF to the pre-initiation complex (PIC) (Fiszbein et al., 2019). Furthermore, the promoter-proximal introns also affect histone H3-K9 acetylation and H3-K4 trimethylation near the 5′ end of genes (**Figure 1B**; Bieberstein et al., 2012). These two chromatin modifications facilitate the recruitment of the PIC on the promoter leading to the activation of genes, both in yeast and higher eukaryotes. Thus, an intron may affect initiation directly by facilitating the recruitment of general transcription factors on the promoter or indirectly by affecting the chromatin structure in the promoter region.

FIGURE 2 | Intron-dependent gene looping enhances transcription by facilitating initiation, reinitiation and promoter directionality. **(A)** A gene with an intron (yellow) and exons (blue). **(B)** During transcription, an elongating RNAPII transcribes the gene and recruits splicing factors (SF) to splice out the intervening intronic sequence. Splicing components (SF) at the 5′ splice site interact with the general transcription factors (GTFs) at the promoter forming a loop between the promoter and 5′ splice site. **(C)** Once RNAPII has transcribed the intron, splicing components at the 3′ splice site associate with the termination factors near the 3′ end of the gene forming a loop between the 3′ splice site and terminator. Finally, the gene forms an overall three-looped conformation where the promoter and terminator physically interact with one another to assist in reinitiation. The three unique interactions that take place are between the promoter-terminator, promoter-5′ splice site and terminator-3′ splice site. GTFs, general transcription factors; SFs, splicing factors; TF, termination factors.

Introns have also been found to stimulate transcription at the elongation as well as termination steps of transcription. Many independent studies have found U1-snRNP, SKIP and SC35, which are splicing factors, stimulating elongation by interacting with various transcription elongation factors (Fong and Zhou, 2001; Brès et al., 2005; Lin et al., 2008). Introns also influence elongation indirectly by affecting chromatin structure. The H3-K36 trimethylation mark, which is specifically associated with elongation of transcription, is enriched in intron-containing genes compared to the intron-less genes (**Figure 1B**; De Almeida et al., 2011). There are also a few reports of terminator-proximal introns influencing transcription by affecting termination of transcription either directly by helping in the recruitment of termination factors or indirectly by affecting the chromatin structure near the 3′ end of genes (Lutz et al., 1996; McCracken et al., 2002). The 3′ splice site of a terminator-proximal intron has been shown to enhance utilization of a downstream poly(A)

site by facilitating recruitment of the F 3′ end processing complex and poly(A) polymerase in mammalian cell lines (**Figure 1B**; Lutz et al., 1996; McCracken et al., 2002). The 3′ splice site also has an adverse effect on H3-K36 trimethylation, which is a transcription elongation mark that needs to be removed to facilitate termination of transcription (Kim et al., 2011). Although introns may target the elongation and termination steps to enhance transcription, the initiation step is emerging as the most frequent target.

INTRON-DEPENDENT LOOPED GENE ARCHITECTURE FACILITATES ENHANCEMENT OF TRANSCRIPTION

The classical view of transcriptional regulation by *cis*-acting regulatory sequences and *trans*-acting protein factors has

undergone radical changes due to research carried out during the last few decades (Papantonis and Cook, 2010; Ansari, 2019; Al-Husini et al., 2020). Genome topology or chromatin conformations formed by enhancer-promoter interactions and promoter-terminator interactions have been found to play a crucial role in regulation of transcription (Misteli, 2007). The physical interaction of the promoter and terminator regions of a gene during transcription results in the formation of a looped gene architecture (Ansari and Hampsey, 2005). Such gene loops are formed in an activator-dependent manner and have been observed in yeast as well as in higher eukaryotes (reviewed in Al-Husini et al., 2020). Activator-dependent gene looping has been shown to enhance transcription by facilitating direct transfer of polymerase from the terminator to the promoter for reinitiation, and by enhancing promoter directionality (Tan-Wong et al., 2012; Al Husini et al., 2013).

In budding yeast, the presence of an intron in a gene also results in a looped gene architecture (**Figure 2C**; Moabbi et al., 2012; Tan-Wong et al., 2012; Agarwal and Ansari, 2016). Intron-dependent gene looping, however, is mechanistically different from the activator-dependent looped conformation. The "Chromosome Conformation Capture" (3C) approach revealed that an intron-dependent loop is characterized by additional contacts of the promoter with the 5′ splice site and of the terminator with 3′ splice site (**Figures 2B,C**; Moabbi et al., 2012). The promoter-5′ splice site loop may not be observed in all genes as the first intron often is located very close to the promoter and can make direct contact with the promoter without loop formation. How the presence of an intron facilitates gene loop formation, however, is not yet clear. It has been proposed that the interaction of a 5′ splice site with the promoter and of a 3′ splice site with terminator bring the two ends of a gene in close physical proximity and facilitate the promoter-terminator contact (Al-Husini et al., 2020). Only a splicing-competent intron facilitates gene loop formation (Moabbi et al., 2012; Agarwal and Ansari, 2016). Mutation of either the 5′ or the 3′ splice sites abolishes intron-facilitated looped gene structure. Since the mutation of splice sites also adversely affects IME of transcription, it was proposed that intron-mediated gene looping may also enhance transcription in a manner similar to activator-dependent enhancement. Examination of the IME effect in the looping-defective mutants of yeast revealed that although splicing was normal, there was no enhancement of transcription (Moabbi et al., 2012; Agarwal and Ansari, 2016). The IME effect therefore is not due to splicing *per se* but due to the formation of a splicing-dependent looped gene architecture in budding yeast.

Evidence suggests that the three contact points in an intron-dependent looped gene structure; promoter-terminator contact, promoter-5′ splice site contact and terminator-3′ splice site (**Figure 2C**), are established by a protein-protein interaction of factors occupying promoter, terminator and intronic sites. A combination of ChIP and 3C approaches identified the crucial role of the general transcription factor TFIIB and CF1A termination complex in the promoter-terminator interaction (Medler et al., 2011). Similarly, the interaction of another general transcription factor TFIIH with U1-snRNP could lead to promoter-5′ splice site contact

(Kwek et al., 2002). The interactions contributing to terminator-3′ splice site contact are yet to be established. Splicing factors do not contact DNA directly, but proximity of splice sites on RNA with the corresponding DNA region during cotranscriptional splicing results in splicing factors getting crosslinked to the splice sites on DNA as well (Kotovic et al., 2003; Lin et al., 2008; Oesterreich et al., 2016; Minocha et al., 2018; Nojima et al., 2018). This evidence suggests that the intron-mediated gene loop formation is due to an interaction of the initiation and termination factors occupying distal ends of a gene with splicing factors bound to the intronic regions.

A critical issue is how intron-dependent looped gene architecture brings about enhancement of transcription. All three of the physical interactions in an intron-facilitated gene loop have the potential to enhance transcription. The promoter-5′ splice site interaction can facilitate initiation/reinitiation by stabilizing the assembly of the PIC (**Figure 2B**). U1-snRNP, which binds the 5′ splice site and exhibits an interaction with TFIIH, may play a crucial role in this regard (Kwek et al., 2002). U1-snRNP may directly help in the recruitment of TFIIH, and possibly TFIIB and TFIID on the promoter if the 5′ splice site is located in close proximity to the promoter (Damgaard et al., 2008). The 5′ splice sites, however, may be located several hundred nucleotides away from the promoter. In such a scenario, a loop formed by promoter-5′ splice site interaction may play a critical role in assembly or stabilization of the PIC on the promoter (**Figure 2B**). The net result will be enhanced initiation or reinitiation of transcription. This may explain why a functional 5′ splice site alone could bring about an increase in transcription of HIV-1 and β-globin genes (Damgaard et al., 2008). The enhancement of transcription elicited by a 5′ splice site alone, however, was much lower (75% less) compared to that brought about by a full-length intron. The interaction of the 5′ splice site with the promoter is therefore not sufficient to achieve enhancement of transcription fully. Contacts of the promoter with the terminator and of the 3′ splice site with the terminator may contribute significantly to the IME effect (**Figure 2C**). Thus, all three interactions in an intron-dependent gene loop may play a role in the enhancement of transcription.

The promoter-terminator interaction is especially crucial as it can enhance transcription by facilitating reinitiation and by conferring promoter directionality (Tan-Wong et al., 2012; Al Husini et al., 2013). It is conceivable that proximity of promoter and terminator in the intron-dependent gene loop may similarly contribute to the enhancement of transcription by a similar mechanism (**Figure 2C**). Juxtaposition of the terminator and promoter facilitates release of polymerase from the terminator region near the promoter. The polymerase is then recycled back to the juxtaposed promoter for reinitiation of transcription (Al Husini et al., 2013). Such a coupling of termination to reinitiation, with a concomitant increase in the transcriptional activity, has been observed for RNAPIII, RNAPI, mitochondrial polymerase and archaeal polymerase (Dieci and Sentenac, 1996; Jansa et al., 2001; Martin et al., 2005; Spitalny and Thomm, 2008). The promoter-terminator interaction is emerging as a critical player in regulation of transcription in eukaryotes.

In budding yeast, juxtaposition of the promoter and terminator also confers promoter directionality, which is the enhancement of promoter-initiated transcription of mRNA while keeping the upstream antisense transcription in check (Tan-Wong et al., 2012; Al Husini et al., 2013). The presence of an intron also enhances promoter directionality of yeast genes (Agarwal and Ansari, 2016). In the absence of a splicing-competent intron, mRNA synthesis exhibits a decline, while uaRNA (upstream antisense RNA) transcription increases. In a looping-defective strain, despite the presence of a splicing-competent intron and normal splicing, promoter directionality is adversely affected. The proximity of the promoter and terminator in the intron-mediated gene loop allows the termination factors bound to the 3′ end to contact the 5′ end of a gene and bring about termination of uaRNA transcription (Agarwal and Ansari, 2016). The net result is higher transcription of the gene by enhancing promoter directionality.

Although, the terminator-3′ splice site interaction can promote cotranscriptional recruitment of termination factors, leading to efficient termination of transcription, its role in IME effect needs further exploration. It is, however, clear from the studies described above that intron-dependent looped gene architecture likely plays a crucial role in enhancement of transcription in budding yeast.

SPLICING-DEPENDENT GENE LOOPING IN HIGHER EUKARYOTES

In budding yeast, merely 4% of genes contain introns, but these few genes produce more than 25% of total cellular mRNA (Ares et al., 1999). On average, yeast intron-containing genes produce 3.7 times more mRNA than their non-intronic counterparts (Juneau et al., 2006). In higher eukaryotes, a far higher proportion of genes harbor introns, and splicing-dependent activation of transcription is prevalent in higher eukaryotes as well (reviewed in Rose, 2008; Laxa, 2017; Shaul, 2017; Rose, 2019). This raises the question if splicing-associated changes in gene architecture also occur in higher eukaryotes, and if they contribute to enhancement of transcription.

Chromatin interaction analyses have revealed the presence of physical interactions of promoters with their gene body in mammalian systems. 3C analysis of the human BRCA1 gene revealed an interaction of the promoter and terminator regions of the gene with the intronic regions in a transcription-dependent manner (Tan-Wong et al., 2008). Intragenic gene loops formed by the interaction of exons with cognate promoters have been identified in human cell lines on a genomewide scale using the ChIA-PET (Chromatin Interaction Analysis with Paired-End Tag) approach (Mercer et al., 2013). Hi-C analysis corroborated the presence of intragenic chromatin loops formed by the interaction of promoters with exons in the human genome (Ruiz-Velasco et al., 2017). This genomewide study also demonstrated intragenic gene loops between exons and the 3′ ends of genes. This is reminiscent of the intron-terminator interaction observed in budding yeast and the human BRCA1 gene (Tan-Wong et al., 2008; Moabbi et al., 2012). ChIA-PET identified CTCF

(CCCTC-binding factor) as the protein that facilitates the interaction of promoter with exons (Mercer et al., 2013). It was further shown that the CTCF-mediated promoter-exon loops are prevalent in genes coding for proteins involved in cell signaling and response to stimuli. The promoter-exon loop formation is accompanied by trimethylation of histone H3-K4, acetylation of histone H3-K27 and trimethylation of histone H3-K36 in the coding region of the gene (Ruiz-Velasco et al., 2017). These three histone modifications are associated with transcriptionally active chromatin (Bannister and Kouzarides, 2011). Modification of histone H3-K4 and H3K9 by methylation and acetylation have been implicated in enhancement of transcription by the promoter-proximal intron in mammalian cells thereby suggesting that the promoter-exon interaction may enhance transcription by affecting chromatin structure (Bieberstein et al., 2012). The promoter-exon interaction may be the mammalian structural equivalent of yeast intron-mediated gene loops.

The promoter-exon interaction may be responsible for recently reported internal exon-mediated activation of transcription in mammalian cells (Fiszbein et al., 2019). Internal exons are small exons generally less than 300 nucleotides in length and are flanked by at least one exon on the 5′ side and one exon on the 3′ side. They are responsible for transcriptional regulation of thousands of mammalian genes. The phenomenon is called EMAT (exon-mediated activation of transcription). Activation of transcription by internal exons is not dependent on the sequence of the exon, but on the splice sites flanking the intron. Mutation of either the 5′ or 3′ splice site completely abrogated transcription activation potential of an internal exon (Fiszbein et al., 2019). Furthermore, exon-mediated activation involved recruitment of the general transcription factor TFIIF on the promoter by the splicing factor HNRNPU (heterogeneous nuclear ribonucleoprotein U), which is known to interact with TFIIF through its N-terminal domain (Kim and Nikodem, 1999). In addition, the exon-promoter interaction is accompanied by trimethylation and acetylation of histone H3, which may also result in enhanced transcription by affecting both transcription initiation and elongation (Bieberstein et al., 2012; Ruiz-Velasco et al., 2017).

The exon-mediated transcriptional regulation in mammalian cells exhibits striking similarities to the intron-mediated regulation in budding yeast: (1) internal exon or intron must be located within one kbp of the promoter to activate transcription; (2) transcription activation by both occurs in a splicing-dependent manner; and (3) transcriptional regulation in both cases involves splicing-dependent recruitment of general transcription factors on the promoter. The possibility of exon-mediated transcriptional activation through intragenic gene loops formed by the interaction of a promoter with the gene body therefore cannot be ruled out. Both intron-mediated and exon-mediated enhancement of transcription are in fact splicing-mediated regulations. Gene architecture playing a general role in splicing-mediated regulation of transcription is an attractive possibility. The studies from yeast strongly suggest the involvement of gene looping in the splicing-mediated regulation, but evidence from higher eukaryotes are still preliminary and needs further investigation. Nevertheless, it is clear that introns

are not merely junk coding sequence but are an important regulator of cellular functions.

AUTHOR CONTRIBUTIONS

AA conceptualized. KD and NA prepared the original draft. KD made figures and helped with editing. AA and LP edited the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

Agarwal, N., and Ansari, A. (2016). Enhancement of transcription by a splicing-competent intron is dependent on promoter directionality. *PLoS Genet.* 12:e1006047. doi: 10.1371/journal.pgen.1006047

Al Husini, N., Kudla, P., and Ansari, A. (2013). A role for CF1A 3′ end processing complex in promoter-associated transcription. *PLoS Genet.* 9:e1003722. doi: 10.1371/journal.pgen.1003722

Al-Husini, N., Medler, S., and Ansari, A. (2020). Crosstalk of promoter and terminator during RNA polymerase II transcription cycle. *Biochim. Biophys. Acta Gene Regulat. Mechanis.* 1863:194657. doi: 10.1016/j.bbagrm.2020.194657

Ansari, A. (2019). Recent Trends in Eukaryotic Transcription: Crucial Role of Gene Architecture in Transcriptional Regulation. *J. Cytol. Mol. Biol.* 4:2.

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3′-end processing machinery in RNAP II-dependent gene looping. *Genes Dev.* 19, 2969–2978. doi: 10.1101/gad.1362305

Ares, M., Grate, L., and Pauling, M. H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* 5, 1138–1139. doi: 10.1017/s1355838299991379

Baier, T., Jacobebbinghaus, N., Einhaus, A., Lauersen, K. J., and Kruse, O. (2020). Introns mediate post-transcriptional enhancement of nuclear gene expression in the green microalga Chlamydomonas reinhardtii. *PLoS Genet.* 16:e1008944. doi: 10.1371/journal.pgen.1008944

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22

Beaulieu, E., Green, L., Elsby, L., Alourfi, Z., Morand, E. F., Ray, D. W., et al. (2011). Identification of a novel cell type-specific intronic enhancer of macrophage migration inhibitory factor (MIF) and its regulation by mithramycin. *Clin. Exp. Immunol.* 163, 178–188. doi: 10.1111/j.1365-2249.2010.04289.x

Bianchi, M., Crinelli, R., Giacomini, E., Carloni, E., and Magnani, M. (2009). A potent enhancer element in the 5′-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. *Gene* 448, 88–101. doi: 10.1016/j.gene.2009.08.013

Bieberstein, N. I., Oesterreich, F. C., Straube, K., and Neugebauer, K. M. (2012). First exon length controls active chromatin signatures and transcription. *Cell Rep.* 2, 62–68. doi: 10.1016/j.celrep.2012.05.019

Bonnet, A., Grosso, A. R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S. C., et al. (2017). Introns protect eukaryotic genomes from transcription-associated genetic instability. *Mol. Cell* 67, 608–621. doi: 10.1016/j.molcel.2017.07.002

Brès, V., Gomes, N., Pickle, L., and Jones, K. A. (2005). A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev.* 19, 1211–1226. doi: 10.1101/gad.1291705

Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E., and Palmiter, R. D. (1988). Introns increase transcriptional efficiency in transgenic mice. *Proc. Natl. Acad. Sci.* 85, 836–840. doi: 10.1073/pnas.85.3.836

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. *Genes Dev.* 1, 1183–1200. doi: 10.1101/gad.1.10.1183

Carmel, L., and Chorev, M. (2012). The function of introns. *Front. Genet.* 3:55. doi: 10.3389/fgene.2012.00055

Charron, M., Chern, J. Y., and Wright, W. W. (2007). The cathepsin L first intron stimulates gene expression in rat sertoli cells. *Biol. Reproduct.* 76, 813–824. doi: 10.1095/biolreprod.106.057851

Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A. S., Byrne, L. J., Wilson, C. M., et al. (2019). Alternative splicing and protein diversity: plants versus animals. *Front. Plant Sci.* 10:708. doi: 10.3389/fpls.2019.00708

Damgaard, C. K., Kahns, S., Lykke-Andersen, S., Nielsen, A. L., Jensen, T. H., and Kjems, J. (2008). A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol. Cell* 29, 271–278. doi: 10.1016/j.molcel.2007.11.035

Das, R., Yu, J., Zhang, Z., Gygi, M. P., Krainer, A. R., Gygi, S. P., et al. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* 26, 867–881. doi: 10.1016/j.molcel.2007.05.036

De Almeida, S. F., Grosso, A. R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., et al. (2011). Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat. Struct. Mol. Biol.* 18:977. doi: 10.1038/nsmb.2123

Dieci, G., and Sentenac, A. (1996). Facilitated recycling pathway for RNA polymerase III. *Cell* 84, 245–252. doi: 10.1016/s0092-8674(00)80979-4

Ding, F., and Elowitz, M. B. (2019). Constitutive splicing and economies of scale in gene expression. *Nat. Struct. Mol. Biol.* 26, 424–432. doi: 10.1038/s41594-019-0226-x

Faustino, N. A., and Cooper, T. A. (2003). Pre-mRNA splicing and human disease. *Genes Dev.* 17, 419–437. doi: 10.1101/gad.1048803

Fiszbein, A., Krick, K. S., Begg, B. E., and Burge, C. B. (2019). Exon-mediated activation of transcription starts. *Cell* 179, 1551–1565. doi: 10.1016/j.cell.2019.11.002

Fong, Y. W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature* 414, 929–933. doi: 10.1038/414929a

Furger, A., O'Sullivan, J. M., Binnie, A., Lee, B. A., and Proudfoot, N. J. (2002). Promoter proximal splice sites enhance transcription. *Genes Dev.* 16, 2792–2799. doi: 10.1101/gad.983602

Gallegos, J. E., and Rose, A. B. (2015). The enduring mystery of intron-mediated enhancement. *Plant Sci.* 237, 8–15. doi: 10.1016/j.plantsci.2015.04.017

Gallegos, J. E., and Rose, A. B. (2017). Intron DNA sequences can be more important than the proximal promoter in determining the site of transcript initiation. *Plant Cell* 29, 843–853. doi: 10.1105/tpc.17.00020

Huang, J., and Liang, T. J. (1993). A novel hepatitis B virus (HBV) genetic element with Rev response element-like properties that is essential for expression of HBV gene products. *Mol. Cell. Biol.* 13, 7476–7486. doi: 10.1128/mcb.13.12.7476

Huang, Z. M., and Yen, T. S. (1995). Role of the hepatitis B virus posttranscriptional regulatory element in export of intronless transcripts. *Mol. Cell. Biol.* 15, 3864–3869. doi: 10.1128/mcb.15.7.3864

Jansa, P., Burek, C., Sander, E. E., and Grummt, I. (2001). The transcript release factor PTRF augments ribosomal gene transcription by facilitating reinitiation of RNA polymerase I. *Nucleic Acids Res.* 29, 423–429. doi: 10.1093/nar/29.2.423

Jo, B. S., and Choi, S. S. (2015). Introns: the functional benefits of introns in genomes. *Genomics Informat.* 13:112. doi: 10.5808/gi.2015.13.4.112

Jobert, L., Pinzon, N., Van Herreweghe, E., Jady, B. E., Guialis, A., Kiss, T., et al. (2009). Human U1 snRNA forms a new chromatin-associated snRNP with TAF15. *EMBO Rep.* 10, 494–500. doi: 10.1038/embor.2009.24

Juneau, K., Miranda, M., Hillenmeyer, M. E., Nislow, C., and Davis, R. W. (2006). Introns regulate RNA and protein abundance in yeast. *Genetics* 174, 511–518. doi: 10.1534/genetics.106.058560

Kaneda, S., Horie, N., Takeishi, K., Takayanagi, A., Seno, T., and Ayusawa, D. (1992). Regulatory sequences clustered at the 5′ end of the first intron of the human thymidylate synthase gene function in cooperation with the promoter region. *Somat. Cell Mol. Genet.* 18, 409–415. doi: 10.1007/bf01233080

Kim, M. K., and Nikodem, V. M. (1999). hnRNP U inhibits carboxy-terminal domain phosphorylation by TFIIH and represses RNA polymerase II elongation. *Mol. Cell. Biol.* 19, 6833–6844. doi: 10.1128/mcb.19.10.6833

Kim, S., Kim, H., Fong, N., Erickson, B., and Bentley, D. L. (2011). Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proc. Natl. Acad. Sci.* 108, 13564–13569. doi: 10.1073/pnas.1109475108

Kotovic, K. M., Lockshon, D., Boric, L., and Neugebauer, K. M. (2003). Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol. Cell. Biol.* 23, 5768–5779. doi: 10.1128/mcb.23.16.5768-5779. 2003

Kwek, K. Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., et al. (2002). U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat. Struct. Biol.* 9, 800–805.

Laxa, M. (2017). Intron-mediated enhancement: a tool for heterologous gene expression in plants? *Front. Plant Sci.* 7:1977. doi: 10.3389/fpls.2016.01977

Le Hir, H., Nott, A., and Moore, M. J. (2003). How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* 28, 215–220. doi: 10.1016/s0968-0004(03)00052-5

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X. D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* 15:819. doi: 10.1038/nsmb.1461

Liu, X., and Mertz, J. E. (1995). HnRNP L binds a cis-acting RNA sequence element that enables intron-dependent gene expression. *Genes Dev.* 9, 1766–1780. doi: 10.1101/gad.9.14.1766

Lutz, C. S., Murthy, K. G., Schek, N., O'Connor, J. P., Manley, J. L., and Alwine, J. C. (1996). Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev.* 10, 325–337. doi: 10.1101/gad.10.3.325

Malim, M. H., Hauber, J., Fenrick, R., and Cullen, B. R. (1988). Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* 335, 181–183. doi: 10.1038/335181a0

Marquez, Y., Höpfler, M., Ayatollahi, Z., Barta, A., and Kalyna, M. (2015). Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. *Genome Res.* 25, 995–1007. doi: 10.1101/gr.186585.114

Martin, M., Cho, J., Cesare, A. J., Griffith, J. D., and Attardi, G. (2005). Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. *Cell* 123, 1227–1240. doi: 10.1016/j.cell.2005.09.040

McCracken, S., Lambermon, M., and Blencowe, B. J. (2002). SRm160 splicing coactivator promotes transcript 3′-end cleavage. *Mol. Cell. Biol.* 22, 148–160. doi: 10.1128/mcb.22.1.148-160.2002

Medler, S., Al Husini, N., Raghunayakula, S., Mukundan, B., Aldea, A., and Ansari, A. (2011). Evidence for a complex of transcription factor IIB (TFIIB) with Poly (A) polymerase and cleavage factor 1 subunits required for gene looping. *J. Biol. Chem.* 286, 33709–33718. doi: 10.1074/jbc.m110.193870

Mercer, T. R., Edwards, S. L., Clark, M. B., Neph, S. J., Wang, H., Stergachis, A. B., et al. (2013). DNase I–hypersensitive exons colocalize with promoters and distal regulatory elements. *Nat. Genet.* 45, 852–859. doi: 10.1038/ng.2677

Millhouse, S., and Manley, J. L. (2005). The C-terminal domain of RNA polymerase II functions as a phosphorylation-dependent splicing activator in a heterologous protein. *Mol. Cell. Biol.* 25, 533–544. doi: 10.1128/mcb.25.2.533-544.2005

Minocha, R., Popova, V., Kopytova, D., Misiak, D., Hüttelmaier, S., Georgieva, S., et al. (2018). Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Res.* 46, 9749–9763. doi: 10.1093/nar/gky640

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787–800. doi: 10.1016/j.cell.2007.01.028

Moabbi, A. M., Agarwal, N., El Kaderi, B., and Ansari, A. (2012). Role for gene looping in intron-mediated enhancement of transcription. *Proc. Natl. Acad. Sci.* 109, 8505–8510. doi: 10.1073/pnas.1112400109

Morgan, J. T., Fink, G. R., and Bartel, D. P. (2019). Excised linear introns regulate growth in yeast. *Nature* 565, 606–611. doi: 10.1038/s41586-018-0828-1

Niehrs, C., and Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol.* 21, 167–178. doi: 10.1038/s41580-019-0206-3

Nilsen, T. W., and Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457–463. doi: 10.1038/nature08909

Nojima, T., Rebelo, K., Gomes, T., Grosso, A. R., Proudfoot, N. J., and Carmo-Fonseca, M. (2018). RNA polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. *Mol. Cell* 72, 369–379. doi: 10.1016/j.molcel.2018.09.004

Oesterreich, F. C., Herzel, L., Straube, K., Hujer, K., Howard, J., and Neugebauer, K. M. (2016). Splicing of nascent RNA coincides with intron exit from RNA polymerase II. *Cell* 165, 372–381. doi: 10.1016/j.cell.2016.02.045

Padgett, R. A. (2012). New connections between splicing and human disease. *Trends Genet.* 28, 147–154. doi: 10.1016/j.tig.2012.01.001

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986). Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* 55, 1119–1150.

Palmiter, R. D., Sandgren, E. P., Avarbock, M. R., Allen, D. D., and Brinster, R. L. (1991). Heterologous introns can enhance expression of transgenes in mice. *Proc. Natl. Acad. Sci.* 88, 478–482. doi: 10.1073/pnas.88.2.478

Papantonis, A., and Cook, P. R. (2010). Genome architecture and the role of transcription. *Curr. Opin. Cell Biol.* 22, 271–276. doi: 10.1016/j.ceb.2010.03.004

Parenteau, J., Maignon, L., Berthoumieux, M., Catala, M., Gagnon, V., and Abou Elela, S. (2019). Introns are mediators of cell response to starvation. *Nature* 565, 612–617. doi: 10.1038/s41586-018-0859-7

Pasleau, F., Leung, F., and Kopchick, J. J. (1987). A comparison of bovine growth hormone expression directed by bGH genomic or intronless DNA in transiently transfected eukaryotic cells. *Gene* 57, 47–52. doi: 10.1016/0378-1119(87)90 175-2

Rogozin, I. B., Carmel, L., Csuros, M., and Koonin, E. V. (2012). Origin and evolution of spliceosomal introns. *Biol. Direct* 7, 1–28.

Rose, A. B. (2004). The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. *Plant J.* 40, 744–751. doi: 10.1111/j.1365-313x.2004.02247.x

Rose, A. B. (2008). Intron-mediated regulation of gene expression. *Nuclear Pre mRNA Process. Plants* 326, 277–290. doi: 10.1007/978-3-540-76776-3_15

Rose, A. B. (2019). Introns as gene regulators: a brick on the accelerator. *Front. Genet.* 9:672. doi: 10.3389/fgene.2018.00672

Ruiz-Velasco, M., Kumar, M., Lai, M. C., Bhat, P., Solis-Pinson, A. B., Reyes, A., et al. (2017). CTCF-mediated chromatin loops between promoter and gene body regulate alternative splicing across individuals. *Cell Syst.* 5, 628–637. doi: 10.1016/j.cels.2017.10.018

Shaul, O. (2017). How introns enhance gene expression. *Int. J. Biochem. Cell Biol.* 91, 145–155. doi: 10.1016/j.biocel.2017.06.016

Spitalny, P., and Thomm, M. (2008). A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. *Mol. Microbiol.* 67, 958–970. doi: 10.1111/j.1365-2958.2007.06084.x

Tan-Wong, S. M., French, J. D., Proudfoot, N. J., and Brown, M. A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc. Natl. Acad. Sci.* 105, 5160–5165. doi: 10.1073/pnas.0801048105

Tan-Wong, S. M., Zaugg, J. B., Camblong, J., Xu, Z., Zhang, D. W., Mischo, H. E., et al. (2012). Gene loops enhance transcriptional directionality. *Science* 338, 671–675. doi: 10.1126/science.1224350

Tian, H. (2001). RNA ligands generated against complex nuclear targets indicate a role for U1 snRNP in co-ordinating transcription and RNA splicing. *FEBS Lett.* 509, 282–286. doi: 10.1016/s0014-5793(01)03188-x

Wang, G. S., and Cooper, T. A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.* 8, 749–761. doi: 10.1038/nrg2164