BBA - Gene Regulatory Mechanisms xxx (xxxx) xxxx

Contents lists available at ScienceDirect

BBA - Gene Regulatory Mechanisms

journal homepage: www.elsevier.com/locate/bbagrm

Review

Pre-mRNA structures forming circular RNAs

Justin R. Welden, Stefan Stamm*

University of Kentucky, 741 South Limestone, Lexington, KY, 40536, United States of America

ARTICLE INFO

Keywords:
Circular RNAs
Back-splicing
RNA structure
Alu element

ABSTRACT

Circular RNAs are a recently discovered class of RNAs formed by covalently linking the 5' and 3' end of an RNA. Pre-mRNAs generate circular RNAs through a back-splicing mechanism. Whereas in linear splicing a 5' splice site is connected to a downstream 3' splice site, in back-splicing the 5' splice site is connected to an upstream 3' splice site. Both mechanisms use the spliceosome for catalysis. For back-splicing to occur, the back-splice sites must frequently be brought into close proximity, which is achieved through the formation of secondary structures in the pre-mRNA. In general, these pre-mRNA structures are formed by RNA base pairing between complementary sequences flanking the back-splicing sites. Proteins can abolish these RNA structures through binding to one of the complementary strands. However, proteins can also promote back-splicing without strong RNA structures through multimerization after binding to intronic regions flanking circular exons. In humans, *Alu*-elements comprising around 11% of the human genome are the best-characterized elements generating structures promoting circular RNA formation. Thus, intronic pre-mRNA structures contribute to the formation of circular RNAs.

1. Introduction

Circular RNAs (**circRNAs**) are covalently closed RNAs that are expressed in all branches of life [1]. Most circRNAs are generated through pre-mRNA back-splicing [2] where a downstream 5' splice site is connected to an upstream 3' splice site (Fig. 1A). Most circRNAs contain exons of mRNAs, but circRNAs composed of pre-mRNA introns have also been detected [3]. Other mechanisms to create circRNAs like tRNA splicing [4], self-splicing of tetrahymena rRNA [5], and the formation of circular viroid RNAs also exist [6,7], but are far less common.

In general, circRNAs have a much lower abundance than linear mRNAs and early reports considered them splicing artifacts [8]. As circRNAs lack a poly adenosine tail, they are diminished in libraries made by oligo dT priming. Their widespread expression became only apparent when next generation sequencing techniques were developed [9]. Currently it is estimated that, in general, circRNAs comprise 0.8–1% of mRNAs [9,10].

The overall function of circRNAs remains enigmatic. CircRNAs accumulate in the cytosol using an export pathway that depends on the RNA helicases UAP56 (DDX39B) and URH49 (DDX39A) [11], pointing towards functions in the cytosol. Some circRNAs have been shown to sequester miRNAs, acting as 'sponges' [12,13], which indirectly influences mRNA abundance. Numerous circRNAs are associated with ribosomes [14] and proof of principle experiments indicated that they can be translated [15–20]. In drosophila and humans, the highest diversity and expression of circRNAs has been found in the brain and increases during aging [21–23] suggesting tissue-specific roles for circRNAs.

The vast majority of exons in circRNAs are also present in their linear counterparts, suggesting a competition between linear splicing and back-splicing. In some cases, this competition can reduce linear mRNA expression [2,24,25], indicating a role of some circRNAs in mRNA expression.

2. Types of circular RNAs

The highest expressed individual circRNAs are generated by a single large exon through a back-splicing mechanism [26]. Overall most circRNAs contain multiple exons, mostly two exons. The number of circRNA transcripts decreases with the number of their forming exons [26]. Similar to linear mRNAs, circRNAs can be alternatively spliced [26,27]. Numerous circRNAs contain intronic sequences, but the majority of circRNAs lack intronic sequences, suggesting that multi-exon circRNAs are generated after parts of the pre-mRNA have been already spliced together, indicating that frequently linear splicing likely occurs prior to back-splicing [21].

Reflecting the low abundance and cell-type specific expression, there is a large discrepancy between circRNAs identified in different

* Corresponding author.
E-mail address: Stefan@stamms-lab.net (S. Stamm).

https://doi.org/10.1016/j.bbagrm.2019.194410
Received 21 May 2019; Received in revised form 16 July 2019; Accepted 17 July 2019

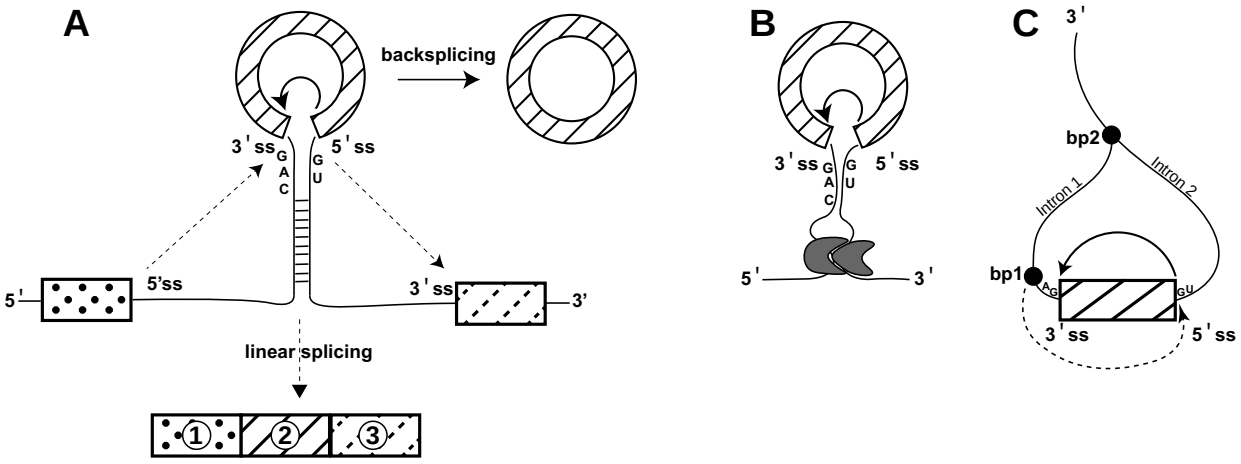

Fig. 1. Generation of circRNAs through back-splicing.
A. Back-splicing due to a pre-mRNA structure. A pre-mRNA containing exon 1, 2, and 3 is shown. The double stranded RNA structure formed by complementary regions in the pre-mRNA brings the 5' and 3' splice site of exon 2 into close proximity, allowing back-splicing (solid arrow) leading to the formation of a circRNA (striped circle). This process competes with the formation of linear mRNA, joining exons 1, 2 and 3, indicated by dashed lines with arrows.
B. Back-splicing can also occur when proteins (gray shapes) multimerize after binding to short recognition sequences that flank an exon [39], which brings the back-splicing sites into close proximity.
C. The presence of an exon in a lariat created during the splicing reaction can also lead to back-splicing of this exon. The exon is first present in a lariat formed by branch point 2 (bp2) in intron 2 of a three-exon substrate. Next a branch-point (bp1) in intron 1 forms a lariat with the guanosine of the 5' splice site (dashed line with arrow), allowing back-splicing to occur (solid line with arrow) [43].

experiments [10] and new circRNAs are being discovered [27]. CircRNA sequences are available from various databases: (CIRCpedia [27], circBase, [28] CircFunBase, [29], CircNet [30], circRNADb [31], and PlantcircBase [32]).

3. Mechanism of back-splicing

Back-splicing that creates circRNAs uses the same splicing machinery and most of the splicing enhancer and silencers as linear pre-mRNA splicing. Thus, exons present in circRNAs are surrounded by canonical splice sites in the pre-mRNA and the formation of circRNAs is sensitive to splicing inhibition [33]. siRNA mediated depletion of splicing factors increases formation of some circRNAs while the corresponding mRNAs were reduced, suggesting a coupling between circRNA and mRNA formation [34]. Minigene analyses showed that, similar to linear splicing, the recognition of exons in circRNAs is based on combinatorial control that takes into account splice site strength [33], the concentration of hnRNPs, SR-proteins, SR-protein kinases [35,36] and RNA-helicases [37]. For circRNAs made from single exons, the yield of circRNA formation increases with exon length [38] and in drosophila long flanking introns favor circRNA formation [21].

4. Recognition of back-splicing sites depends on the pre-mRNA secondary structure

For back-splicing to occur, the splice sites involved have to be brought into close contact, either through a secondary structure in the RNA (Fig. 1A) or through protein interactions (Fig. 1B). In most cases studied, these structures are generated through base-pairing of complementary RNA sequences within the same pre-mRNA molecule. Protein interactions have been shown for quaking I (QKI) that binds to short intronic recognition sites flanking the back-splicing sites resulting in circRNA formation [39]. Similarly, the muscle blind protein (MBL) binds to repeat sequences flanking back-splicing sites and also promotes circRNA formation [40]. Bridging of exons through intronic hnRNP binding sites have been shown for other proteins, such as hnRNPA1 [41] and PTB/hnRNPI [42], suggesting a more general role for this mechanism in circRNA formation. In lower eukaryotes that have less inverted repeats, circRNAs can be formed through a lariat precursor, where back-splicing occurs within the circular lariat [43], (Fig. 1C).

Due to its high expression, one of the first circRNAs identified was from the mouse SRY gene (sex determining region Y) [44]. This single-exon circRNA is located in a 2.7 kb long region flanked by a large inverted repeats (> 15.5 kb) that is necessary for the circularization [44,45], providing the first evidence that often genomic repeat elements flanking the back-splice sites generate the secondary structures necessary for proper alignment of the splice sites. In humans, these repeat elements are often provided by *Alu* elements [2] and their contribution to circRNA formation has been extensively studied experimentally [26,38]. Similarly, reverse complementary sequences not emanating from repeat elements have been detected in *C. elegans* [46]. However, a genome-wide screen in drosophila failed to detect direct repeats for the majority of circRNAs [21], although detailed experiments using model RNAs revealed that some highly expressed drosophila circRNAs, like laccase 2 (Fig. 3C) rely on repeats for their formation [35].

Genome wide cloning of human RNase resistant double stranded RNA revealed thousands of double stranded RNAs outside known repeats, further supporting the idea of RNA double stranded structures that occur independent of repeat elements [47]. Thus, despite the current research focus on inverted repeat elements, especially *Alu* elements, other RNA structures likely also facilitate back-splicing.

5. *Alu* elements form secondary structures promoting human circular RNA formation

Sequencing of human circRNAs revealed that they are often flanked by *Alu* elements in their genomic location [2,26]. *Alu* elements are about 300 nucleotides (nt) long short interspersed nuclear elements (SINE) [48–50] that were derived from the 7SL-RNA in the early primate lineage [51,52] (Fig. 2A, B). There are more than one million *Alu* element copies in the human genome that comprise about 11% of the human genome [53]. *Alu* elements continue to amplify through a polymerase III-derived RNA intermediate and it is estimated that there is about one new *Alu* insertion in 21 human births [54]. They are predominantly located in gene-rich regions [55], possibly because their small size does not interfere strongly with gene regulation [48]. *Alu* elements are subdivided into the J (Jurka, Jerzy), S (Smith, Temple)

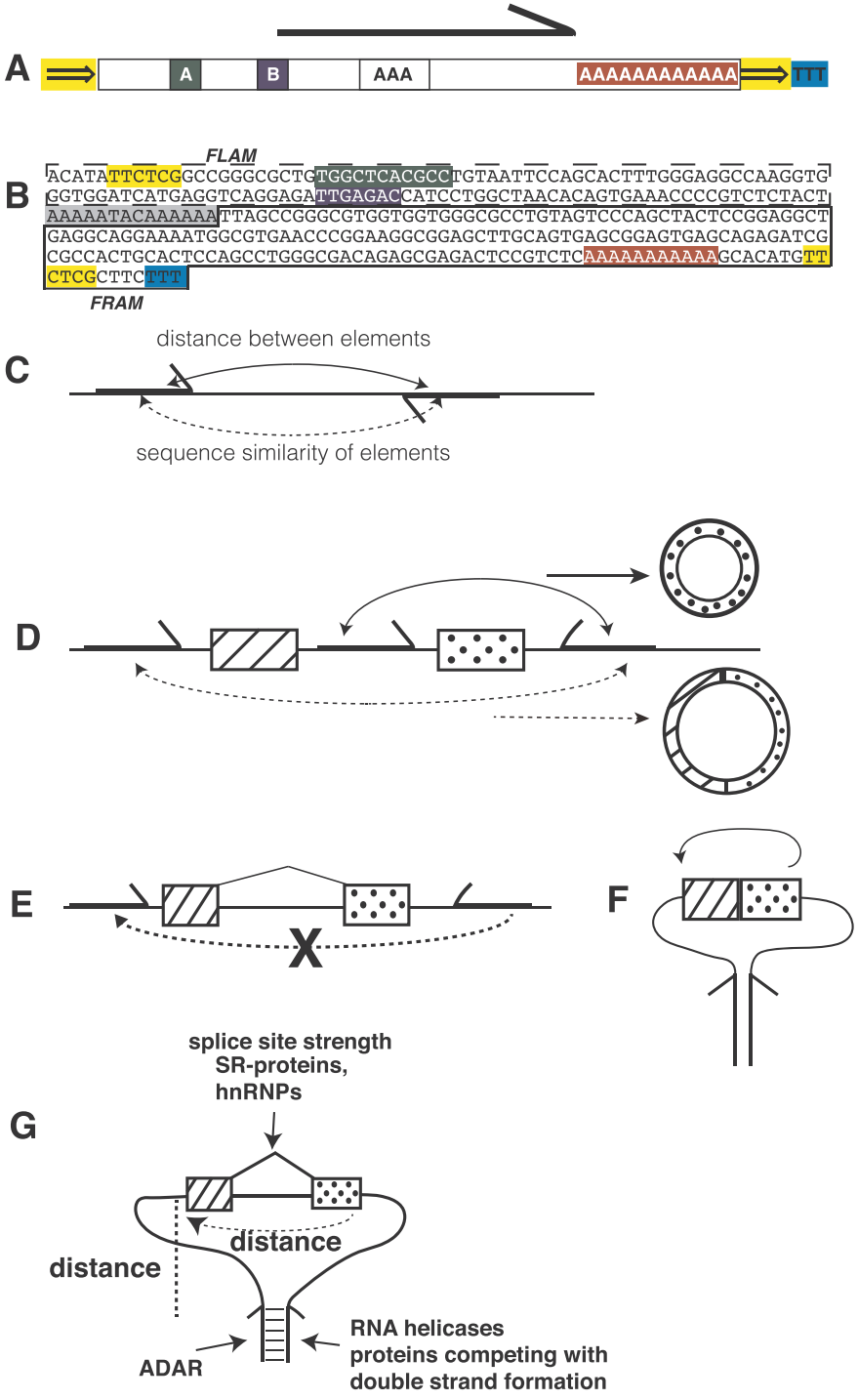

Fig. 2. *Alu* elements promote back-splicing through formation of structures in the pre-mRNA.

A. General sequence of an *Alu* element, "A, B": A and B boxes as recognition sites for DNA polymerase III. AAA: short mid A stretch; AAAA (red): terminal A stretch, TTT: downstream T-stretch, acting as a pol III termination site, arrows, yellow direct repeats (genomic Alu insertion site). The common depiction of an *Alu* element is an arrow that shows the direction from the A box to the terminal A stretch (arrowhead).

B. Sequence of an **AluJB** element with the elements highlighted from the human *MAPT* gene (hg38 chr17:46,014,583-46,014,913). FLAM: free left *Alu* monomer (dashed box), FRAM: free right Alu monomer (solid box). The coloring of the elements is similar to panel A.

C. Determinants of *Alu*-element mediated secondary structures. *Alu* elements oriented in opposite direction can form double stranded regions, measured through editing of the *Alu*-elements. The probability to form a double stranded structure increases exponentially when the *Alu*-elements are getting closer.

D. Arrangement of *Alu* elements in a pre-mRNA can promote alternative circRNA splicing. Depending on the interaction (dotted line) of the *Alu* elements a circle between exon 1 and 2 or a circle composed of exon 2 (dotted) can be formed.

E, F. Removal of an intron through splicing can decrease the distance between dsRNA structure forming *Alu*-elements. E. shows the situation before splicing, where the *Alu*-elements are too far apart to allow back-splicing. F: Intron removal through splicing allows back-splicing and circRNA formation.

G. Elements controlling circRNA formation: hnRNPs and SR-proteins control general splice site recognition and influence competition between linear and circular splicing. Helicases influence the stability of the RNA structures promoting circular splicing, a larger distance between the back-splicing sites, i.e. the length of the circularized exon promotes back-splicing.

and Y (young) subfamilies, which reflect an alphabetical progression from the oldest (J) to the youngest (Y) *Alu* subtypes [56]. Since different subfamilies of *Alu* elements are highly similar in sequence, they can form secondary structures when inserted in opposite orientation in a pre-mRNA transcript, referred to as inverted repeated *Alu* structure (IRAlus) [48], (Fig. 2C, examples in Fig. 3A, B).

The ability of *Alu*-elements to form double-stranded secondary structures and thus their propensity to promote circRNA formation is reduced by adenosine to inosine (A- > I) editing, which is the most common post-transcriptional RNA modification in primates [57,58]. A- > I editing is performed by the adenosine deaminases acting on RNA (ADAR) family of enzymes that require RNA double strands as a template [59]. Thus, the editing of *Alu* elements can be used to measure whether an *Alu* element is in a double stranded conformation. The double strandedness, (i.e. the formation of a secondary structure) of an *Alu*-element mostly depends on the distance to the closest reversely oriented *Alu* element, which is around 800 nt. The *Alu*-element editability decreases exponentially with longer distances [57]. Other factors affecting double strandedness include the presence of additional competing *Alu* elements, which decrease double strandedness of a single *Alu* element; and the subfamily of the *Alu* element, where related subfamilies are edited stronger (Fig. 2C). ADAR knockdown increases circRNA expression in human cells [46], further suggesting a role of *Alu*-element modification in human circRNA formation. ADARs are not

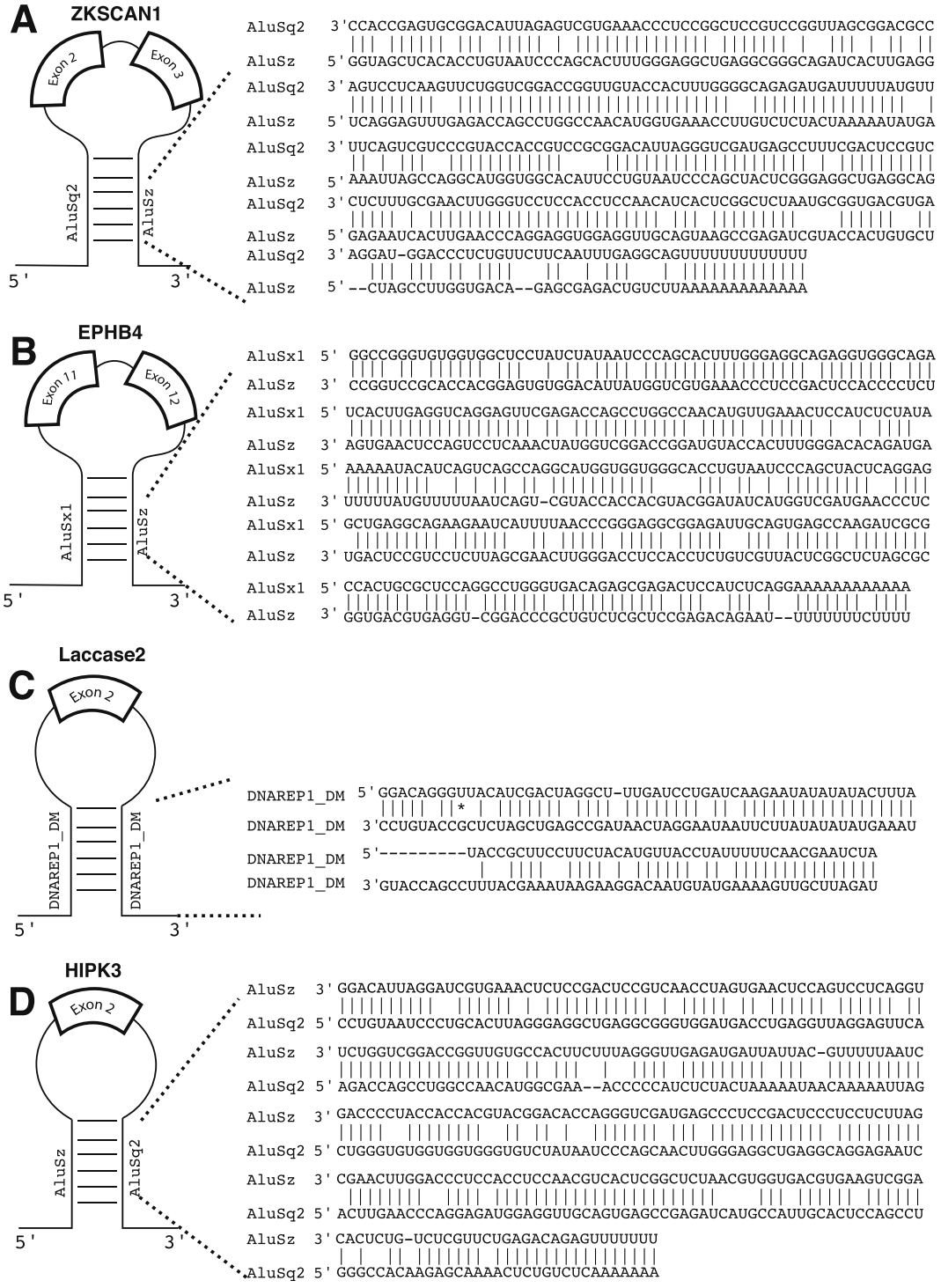

Fig. 3. Experimentally characterized examples of RNA structures leading to back-splicing.
Schematic structures and double stranded region of A. ZKSCAN1 [38], B. EPHB4 [38], C. LACCASE2 [35] and D. HIPK3 [38] genes are shown. Experimentally validated double stranded pre-mRNA regions are shown on the right. A "*" indicated G-U base pairing.

specific for double stranded RNAs formed by *Alu* elements, and thus ADAR knockdown in mouse promotes formation of some circRNAs as well [22].

Double stranded *Alu*-element structures can form within an intron as well as across an exon. Usually, only *Alu* elements localized across an exon in opposite directions promote circRNA formation [26]. Often, an *Alu*-element can base-pair with other distinct *Alu*-elements, leading to alternative circRNA formation (Fig. 2D). Since linear splicing reduces the distance between *Alu*-elements that flank exons it promotes circRNA formation when it occurs faster than back-splicing (Fig. 2E, F). Deletion analyses in several model systems showed that the full *Alu* element is not necessary for circRNA formation to occur [26]. In fact complementary sequences as small a 30–40 nt are sufficient and give stronger circRNA formation than longer elements [38] (Fig. 3A).

DHX9 is an RNA helicase that selectively binds to *Alu*-elements located in proximity (458 nt). DHX9 binds to an interferon-inducible form of ADAR, and knock down of DHX9 increases circRNA production [37,60], suggesting that DHX9 destabilizes pre-mRNA structures

necessary for circRNA formation.

There is emerging evidence that *Alu*-elements can also act in *trans*, i.e. between different pre-mRNA molecules, leading to trans-splicing [61], which could conceptually compete with circRNA splicing.

In summary, multiple factors control the occurrence of back-splicing and thus the formation of circRNAs. Most importantly, a double stranded pre-mRNA structure brings back-splicing sites together. Its formation is influenced by the availability of complementary sequences, which in turn are negatively influenced by RNA helicases, ADAR enzymes and single-stranded RNA binding proteins. Factors that influence intron splicing, such as splice site strength, the concentration of SR-proteins and hnRNPs as well as the availability of exon enhancers and silencers also impact on circRNA formation (Fig. 2G).

6. Experimental studies

The correlations between double stranded pre-mRNA structures and circRNA formation have been tested experimentally. Current reporter gene constructs and experimentally validated double stranded RNA structures are summarized in Fig. 3. The double stranded structures can be surprisingly small. For example the experimentally determined structure needed to form circRNAs forming from ZKSCAN1 exon 2/3 is only 36 nt long and contains 7 mismatches, although the natural repeat is much longer [38], (Fig. 3A). It is thus not surprising that circRNAs can form without known or identifiable repeats [33].

7. Outlook

Despite their low expression when compared to their linear counterparts, circRNAs could turn out to be biologically highly important. An increasing number of studies are showing now that circRNAs can form proteins [16,62]. CircRNAs are highly expressed in the brain, where they could encode new peptide hormones that are small, reflecting the short reading frames of circRNAs and work at small concentrations. For example, α-melanocyte-stimulating hormone (α-MSH), generated from POMC is a peptide composed of 13 amino acids that binds to the melanocortin 1 receptors with an affinity of 200 pM to 2 nM [63]. Thus, the generation of novel brain-specific signaling peptides could be an important function of circRNAs. It is notable that the development of the brain cortex in primates correlates with the expansion of *Alu* elements [64], which could result in so far unknown peptide hormones.

The formation of circRNAs depends on intronic elements that form double-stranded RNA structures. CircRNAs and possibly their encoded peptides could be an evolutionary force reflecting differences in intron sequences caused by the presence of species-specific repeat elements, such as primate-specific *Alu* elements and SINEC_Cf elements characteristic for various dog species [65]. Similarly, deep intronic mutations characteristic for cancer [66] and present in numerous hereditary diseases [67] could act by influencing circRNA expression.

So far, studies of pre-mRNA structures do not take RNA modifications other than A- > I editing into account. N6-Adenosine methylation (m6A) is another common mRNA modification that could influence RNA structures and thus circRNA formation. M6A RNA modification creates a strong binding site for YTH-domain proteins [68] that could bridge back-splicing sites. In addition m6A modifications create hnRNPG or hnRNPC binding sites [69] which abolishes RNA double stranded structures and was shown to change linear alternative splicing [69]. It is thus conceivable that circRNA formation is similarly influenced by m6A RNA modifications and could be thus under epigenetic control.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgments

This work was supported by the Department of Defense AZ180075 and a Jacqueline Noonan endowment (SS). JRW is a recipient of the College of Medicine excellence in graduate research award of the University of Kentucky.

References

[1] P.L. Wang, Y. Bao, M.C. Yee, S.P. Barrett, G.J. Hogan, M.N. Olsen, J.R. Dinneny, P.O. Brown, J. Salzman, Circular RNA is expressed across the eukaryotic tree of life, PLoS One 9 (2014) e90859.
[2] W.R. Jeck, J.A. Sorrentino, K. Wang, M.K. Slevin, C.E. Burd, J. Liu, W.F. Marzluff, N.E. Sharpless, Circular RNAs are abundant, conserved, and associated with ALU repeats, Rna 19 (2013) 141–157.
[3] Y. Zhang, X.O. Zhang, T. Chen, J.F. Xiang, Q.F. Yin, Y.H. Xing, S. Zhu, L. Yang, L.L. Chen, Circular intronic long noncoding RNAs, Mol. Cell 51 (2013) 792–806.
[4] Z. Lu, G.S. Filonov, J.J. Noto, C.A. Schmidt, T.L. Hatkevich, Y. Wen, S.R. Jaffrey, A.G. Matera, Metazoan tRNA introns generate stable circular RNAs in vivo, RNA 21 (2015) 1554–1565.
[5] A.J. Zaug, P.J. Grabowski, T.R. Cech, Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction, Nature 301 (1983) 578–583.
[6] H.L. Sanger, G. Klotz, D. Riesner, H.J. Gross, A.K. Kleinschmidt, Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures, Proc. Natl. Acad. Sci. U. S. A. 73 (1976) 3852–3856.
[7] T. Toptan, B. Abere, M.A. Nalesnik, S.H. Swerdlow, S. Ranganathan, N. Lee, K.H. Shair, P.S. Moore, Y. Chang, Circular DNA tumor viruses make circular RNAs, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) E8737–E8745.
[8] C. Cocquerelle, B. Mascrez, D. Hetuin, B. Bailleul, Mis-splicing yields circular RNA molecules, FASEB J. 7 (1993) 155–160.
[9] J. Salzman, C. Gawad, P.L. Wang, N. Lacayo, P.O. Brown, Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types, PLoS One 7 (2012) e30733.
[10] W.R. Jeck, N.E. Sharpless, Detecting and characterizing circular RNAs, Nat. Biotechnol. 32 (2014) 453–461.
[11] C. Huang, D. Liang, D.C. Tatomer, J.E. Wilusz, A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs, Genes Dev. 32 (2018) 639–644.
[12] T.B. Hansen, T.I. Jensen, B.H. Clausen, J.B. Bramsen, B. Finsen, C.K. Damgaard, J. Kjems, Natural RNA circles function as efficient microRNA sponges, Nature 495 (2013) 384–388.
[13] A.C. Panda, Circular RNAs act as miRNA sponges, Adv. Exp. Med. Biol. 1087 (2018) 67–79.
[14] C. Ragan, G.J. Goodall, N.E. Shirokikh, T. Preiss, Insights into the biogenesis and potential functions of exonic circular RNA, Sci. Rep. 9 (2019) 2048.
[15] N. Abe, K. Matsumoto, M. Nishihara, Y. Nakano, A. Shibata, H. Maruyama, S. Shuto, A. Matsuda, M. Yoshida, Y. Ito, H. Abe, Rolling circle translation of circular RNA in living human cells, Sci. Rep. 5 (2015) 16435.
[16] I. Legnini, G. Di Timoteo, F. Rossi, M. Morlando, F. Briganti, O. Sthandier, A. Fatica, T. Santini, A. Andronache, M. Wade, P. Laneve, N. Rajewsky, I. Bozzoni, Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis, Mol. Cell 66 (2017) 22–37 e29.
[17] X. Zheng, L. Chen, Y. Zhou, Q. Wang, Z. Zheng, B. Xu, C. Wu, Q. Zhou, W. Hu, C. Wu, J. Jiang, A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling, Mol. Cancer 18 (2019) 47.
[18] M. Zhang, K. Zhao, X. Xu, Y. Yang, S. Yan, P. Wei, H. Liu, J. Xu, F. Xiao, H. Zhou, X. Yang, N. Huang, J. Liu, K. He, K. Xie, G. Zhang, S. Huang, N. Zhang, A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma, Nat. Commun. 9 (2018) 4475.
[19] X.F. Li, J. Lytton, A circularized sodium-calcium exchanger exon 2 transcript, J. Biol. Chem. 274 (1999) 8153–8160.
[20] A. Costello, N.T. Lao, N. Barron, M. Clynes, Continuous translation of circularized mRNA improves recombinant protein titer, Metab. Eng. 52 (2019) 284–292.
[21] J.O. Westholm, P. Miura, S. Olson, S. Shenker, B. Joseph, P. Sanfilippo, S.E. Celniker, B.R. Graveley, E.C. Lai, Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation, Cell Rep. 9 (2014) 1966–1980.
[22] A. Rybak-Wolf, C. Stottmeister, P. Glazar, M. Jens, N. Pino, S. Giusti, M. Hanan, M. Behm, O. Bartok, R. Ashwal-Fluss, M. Herzog, L. Schreyer, P. Papavasileiou, A. Ivanov, M. Ohman, D. Refojo, S. Kadener, N. Rajewsky, Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed, Mol. Cell 58 (2015) 870–885.
[23] M.T. Veno, T.B. Hansen, S.T. Veno, B.H. Clausen, M. Grebing, B. Finsen, I.E. Holm, J. Kjems, Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development, Genome Biol. 16 (2015) 245.
[24] C.W. Chao, D.C. Chan, A. Kuo, P. Leder, The mouse formin (Fmn) gene: abundant circular RNA transcripts and gene-targeted deletion analysis, Mol. Med. 4 (1998) 614–628.
[25] F. Gualandi, C. Trabanelli, P. Rimessi, E. Calzolari, L. Toffolatti, T. Patarnello, G. Kunz, F. Muntoni, A. Ferlini, Multiple exon skipping and RNA circularisation contribute to the severe phenotypic expression of exon 5 dystrophin deletion, J.

Med. Genet. 40 (2003) e100.
[26] X.O. Zhang, H.B. Wang, Y. Zhang, X. Lu, L.L. Chen, L. Yang, Complementary sequence-mediated exon circularization, Cell 159 (2014) 134–147.
[27] X.O. Zhang, R. Dong, Y. Zhang, J.L. Zhang, Z. Luo, J. Zhang, L.L. Chen, L. Yang, Diverse alternative back-splicing and alternative splicing landscape of circular RNAs, Genome Res. 26 (2016) 1277–1287.
[28] P. Glazar, P. Papavasileiou, N. Rajewsky, circBase: a database for circular RNAs, RNA 20 (2014) 1666–1670.
[29] X. Meng, D. Hu, P. Zhang, Q. Chen, M. Chen, CircFunBase: a database for functional circular RNAs, Database (Oxford) (2019) 2019.
[30] Y.C. Liu, J.R. Li, C.H. Sun, E. Andrews, R.F. Chao, F.M. Lin, S.L. Weng, S.D. Hsu, C.C. Huang, C. Cheng, C.C. Liu, H.D. Huang, CircNet: a database of circular RNAs derived from transcriptome sequencing data, Nucleic Acids Res. 44 (2016) D209–D215.
[31] X. Chen, P. Han, T. Zhou, X. Guo, X. Song, Y. Li, circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations, Sci. Rep. 6 (2016) 34985.
[32] Q. Chu, X. Zhang, X. Zhu, C. Liu, L. Mao, C. Ye, Q.H. Zhu, L. Fan, PlantcircBase: a database for plant circular RNAs, Mol. Plant 10 (2017) 1126–1128.
[33] S. Starke, I. Jost, O. Rossbach, T. Schneider, S. Schreiner, L.H. Hung, A. Bindereif, Exon circularization requires canonical splice signals, Cell Rep. 10 (2015) 103–111.
[34] D. Liang, D.C. Tatomer, Z. Luo, H. Wu, L. Yang, L.L. Chen, S. Cherry, J.E. Wilusz, The output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting, Mol. Cell 68 (2017) 940–954 e943.
[35] M.C. Kramer, D. Liang, D.C. Tatomer, B. Gold, Z.M. March, S. Cherry, J.E. Wilusz, Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins, Genes Dev. 29 (2015) 2168–2182.
[36] J.R. Welden, J. van Doorn, P.T. Nelson, S. Stamm, The human MAPT locus generates circular RNAs, Biochim. Biophys. Acta Mol. Basis Dis. (2018) 2753–2760.
[37] E.W. Ottesen, D. Luo, J. Seo, N.N. Singh, R.N. Singh, Human survival motor neuron genes generate a vast repertoire of circular RNAs, Nucleic Acids Res. 47 (2019) 2884–2905.
[38] D. Liang, J.E. Wilusz, Short intronic repeat sequences facilitate circular RNA production, Genes Dev. 28 (2014) 2233–2247.
[39] S.J. Conn, K.A. Pillman, J. Toubia, V.M. Conn, M. Salmanidis, C.A. Phillips, S. Roslan, A.W. Schreiber, P.A. Gregory, G.J. Goodall, The RNA binding protein quaking regulates formation of circRNAs, Cell 160 (2015) 1125–1134.
[40] R. Ashwal-Fluss, M. Meyer, N.R. Pamudurti, A. Ivanov, O. Bartok, M. Hanan, N. Evantal, S. Memczak, N. Rajewsky, S. Kadener, circRNA biogenesis competes with pre-mRNA splicing, Mol. Cell 56 (2014) 55–66.
[41] F.U. Nasim, S. Hutchison, M. Cordeau, B. Chabot, High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism, RNA 8 (2002) 1078–1089.
[42] R. Lamichhane, G.M. Daubner, J. Thomas-Crusells, S.D. Auweter, C. Manatschal, K.S. Austin, O. Valniuk, F.H. Allain, D. Rueda, RNA looping by PTB: evidence using FRET and NMR spectroscopy for a role in splicing repression, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 4105–4110.
[43] S.P. Barrett, P.L. Wang, J. Salzman, Circular RNA biogenesis can proceed through an exon-containing lariat precursor, Elife 4 (2015) e07540.
[44] B. Capel, A. Swain, S. Nicolis, A. Hacker, M. Walter, P. Koopman, P. Goodfellow, R. Lovell-Badge, Circular transcripts of the testis-determining gene Sry in adult mouse testis, Cell 73 (1993) 1019–1030.
[45] R.A. Dubin, M.A. Kazmi, H. Ostrer, Inverted repeats are necessary for circularization of the mouse testis Sry transcript, Gene 167 (1995) 245–248.
[46] A. Ivanov, S. Memczak, E. Wyler, F. Torti, H.T. Porath, M.R. Orejuela, M. Piechotta, E.Y. Levanon, M. Landthaler, C. Dieterich, N. Rajewsky, Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals, Cell Rep. 10 (2015) 170–177.
[47] M. Shen, E. Eyras, J. Wu, A. Khanna, S. Josiah, M. Rederstorff, M.Q. Zhang, S. Stamm, Direct cloning of double-stranded RNAs from RNase protection analysis reveals processing patterns of C/D box snoRNAs and provides evidence for widespread antisense transcript expression, Nucleic Acids Res. 39 (2011) 9720–9730.
[48] L.L. Chen, L. Yang, ALUternative regulation for gene expression, Trends Cell Biol. 27 (2017) 480–490.
[49] J. Hasler, T. Samuelsson, K. Strub, Useful 'junk': Alu RNAs in the human transcriptome, Cell. Mol. Life Sci. 64 (2007) 1793–1800.
[50] P. Deininger, Alu elements: know the SINEs, Genome Biol. 12 (2011) 236.
[51] Y. Quentin, Fusion of a free left Alu monomer and a free right Alu monomer at the origin of the Alu family in the primate genomes, Nucleic Acids Res. 20 (1992) 487–493.
[52] Y. Quentin, Origin of the Alu family: a family of Alu-like monomers gave birth to the left and the right arms of the Alu elements, Nucleic Acids Res. 20 (1992) 3397–3401.
[53] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrinos, M.J. Morgan, Initial sequencing and analysis of the human genome, Nature 409 (2001) 860–921.
[54] J. Xing, Y. Zhang, K. Han, A.H. Salem, S.K. Sen, C.D. Huff, Q. Zhou, E.F. Kirkness, S. Levy, M.A. Batzer, L.B. Jorde, Mobile elements create structural variation: analysis of a complete human genome, Genome Res. 19 (2009) 1516–1526.
[55] R. Versteeg, B.D. van Schaik, M.F. van Batenburg, M. Roos, R. Monajemi, H. Caron, H.J. Bussemaker, A.H. van Kampen, The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes, Genome Res. 13 (2003) 1998–2004.
[56] M.A. Batzer, P.L. Deininger, U. Hellmann-Blumberg, J. Jurka, D. Labuda, C.M. Rubin, C.W. Schmid, E. Zietkiewicz, E. Zuckerkandl, Standardized nomenclature for Alu repeats, J. Mol. Evol. 42 (1996) 3–6.
[57] L. Bazak, E.Y. Levanon, E. Eisenberg, Genome-wide analysis of Alu editability, Nucleic Acids Res. 42 (2014) 6876–6884.
[58] A. Athanasiadis, A. Rich, S. Maas, Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome, PLoS Biol. 2 (2004) e391.
[59] K. Nishikura, A-to-I editing of coding and non-coding RNAs by ADARs, Nat. Rev. Mol. Cell Biol. 17 (2016) 83–96.
[60] T. Aktas, I. Avsar Ilik, D. Maticzka, V. Bhardwaj, C. Pessoa Rodrigues, G. Mittler, T. Manke, R. Backofen, A. Akhtar, DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome, Nature 544 (2017) 115–119.
[61] T.J. Chuang, Y.J. Chen, C.Y. Chen, T.L. Mai, Y.D. Wang, C.S. Yeh, M.Y. Yang, Y.T. Hsiao, T.H. Chang, T.C. Kuo, H.H. Cho, C.N. Shen, H.C. Kuo, M.Y. Lu, Y.H. Chen, S.C. Hsieh, T.W. Chiang, Integrative transcriptome sequencing reveals extensive alternative trans-splicing and cis-backsplicing in human cells, Nucleic Acids Res. 46 (2018) 3671–3691.
[62] N.R. Pamudurti, O. Bartok, M. Jens, R. Ashwal-Fluss, C. Stottmeister, L. Ruhe, M. Hanan, E. Wyler, D. Perez-Hernandez, E. Ramberger, S. Shenzis, M. Samson, G. Dittmar, M. Landthaler, M. Chekulaeva, N. Rajewsky, S. Kadener, Translation of CircRNAs, Mol. Cell 66 (2017) 9–21 (e27).
[63] W. Siegrist, F. Solca, S. Stutz, L. Giuffre, S. Carrel, J. Girard, A.N. Eberle, Characterization of receptors for alpha-melanocyte-stimulating hormone on human melanoma cells, Cancer Res. 49 (1989) 6352–6358.
[64] J.S. Mattick, M.F. Mehler, RNA editing, DNA recoding and the evolution of human cognition, Trends Neurosci. 31 (2008) 227–233.
[65] W. Wang, E.F. Kirkness, Short interspersed elements (SINEs) are a major source of canine genomic diversity, Genome Res. 15 (2005) 1798–1808.
[66] S. Diederichs, L. Bartsch, J.C. Berkmann, K. Frose, J. Heitmann, C. Hoppe, D. Iggena, D. Jazmati, P. Karschnia, M. Linsenmeier, T. Maulhardt, L. Mohrmann, J. Morstein, S.V. Paffenholz, P. Ropenack, T. Ruckert, L. Sandig, M. Schell, A. Steinmann, G. Voss, J. Wasmuth, M.E. Weinberger, R. Wullenkord, The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations, EMBO Mol. Med. 8 (2016) 442–457.
[67] R. Vaz-Drago, N. Custodio, M. Carmo-Fonseca, Deep intronic mutations and human disease, Hum. Genet. 136 (2017) 1093–1111.
[68] K.D. Meyer, S.R. Jaffrey, Rethinking m(6)A readers, writers, and erasers, Annu. Rev. Cell Dev. Biol. 33 (2017) 319–342.
[69] N. Liu, K.I. Zhou, M. Parisien, Q. Dai, L. Diatchenko, T. Pan, N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein, Nucleic Acids Res. 45 (2017) 6051–6063.