Biosc.Biotech.Res.Comm. Special Issue Vol 13 No 12 (2020) Pp-25-30

BBRC
Bioscience Biotechnology Research Communications

Progress in the Emerging Areas of RNA Biology and Implications in Life Sciences

K. Krishnam Raju
Centurion University of Technology and Management, Odisha, India

ABSTRACT

Recent advances in transriptomics such as Next gen sequencing and the discovery of novel mechanisms for control of gene expression such as microRNAs, RNAi and riboswitches gave a better insight on the complexity of the RNA function within the cell but also provides us with bigger opportunities for applying novel RNA-based engineering strategies in health care, agriculture and clinical diagnostics. Undestanding RNA biology in medical genomics offered new exciting opportunities for the development of novel therapeutic strategies in tackling various forms of cancer, viral infection diseases, deployment of novel antibiotics, and others. Likewise RNAi substantially contributed to crop improvement in obtaining improved cultivars with desirable traits such as biotic stress tolerance, toxin free plant types, seedless fruits, male sterile lines, delayed ripening of fruits, enhanced nutritional content, removal of secondary metabolites and altered floral characteristics.

KEY WORDS: EPITRANSCRIPTOMICS, POLYADENYLATION, ALTERNATIVE SPLICING, RNAI, CRISPR/CAS9.

INTRODUCTION

The various post-transcriptional modifications such as RNA editing, alternative splicing, and alternative polyadenylation greatly increase the transcript heterogeneity and biodiversity of proteins that can be encoded by the genome in humans. While the existence of post-transcriptional regulatory events enhances the scope and breadth of RNA functions at transcriptional, post-transcriptional, translational and post-translational levels and in conjunction with bioinformatic tools offers exciting opportunities for future genomic medicines and inducing plant resistance to various biotic stresses.

ARTICLE INFORMATION

Corresponding author email: k.krishnam@cutm.ac.in
Received 12th Oct 2020 Accepted after revision 28th Dec 2020
Print ISSN: 0974-6455 Online ISSN: 2321-4007 CODEN: BBRCBA

Thomson Reuters ISI Web of Science Clarivate Analytics USA and Crossref Indexed Journal

NAAS Journal Score 2020 (4.31) SJIF: 2020 (7.728)

Online Contents Available at: http//www.bbrc.in/

Epitranscriptomics: Recently a new field of study has emerged termed as epitranscriptomics which refers to the study of chemical modifications (mostly methyl groups) to RNA. In recent years it is becoming increasingly evident that RNA is also epigenetically modified. Next Generation Sequencing (NGS) technology revealed more than 170 kinds of biochemical ribonucleic acid (RNA) modifications in almost all living organisms and all kinds of RNAs, including ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), messenger RNAs (mRNAs), circular RNAs (circRNA) and long non coding RNAs. Such epigenetic RNA modifications include N6-methyladenosine (m6A), N6,2′O-dimethyladenosine (m6Am), 5-methylcytosine (m5C), 5-hydroxymethylcytosine (hm5C), N1-methyladenosine (m1A), inosine (I) and pseudouridine (ψ) at promoter sites as well regulate RNA metabolism by influencing RNA structure, RNA stability and splicing factors Lorna (2019). Special classes of proteins that add methyl groups to RNA are termed writers.

Classes of proteins that remove methyl groups from RNA are termed writers. The classes of proteins that enable reading of methyl groups on RNA are termed readers. The RNA modifications most commonly characterized

is N6-methyladenosine (m6A) when "writer" enzyme complex METTL3/14 adds a methyl group to adenosine to form m6A. Two different "eraser" enzymes ALKBH5 and FTO remove the methyl groups and restores adenosine. While the causal link between RNA modifications and disease development has not been established so far, but the elevated levels of METTL3 and ALKBH5 contributing to cancer progression makes this new field of RNA biology a future promising and potential approach in the development of cancer therapeutics vu et al. (2017) and Zhang et al. (2017)).

These epigenetic modifications termed as RNA epitranscriptomic marks are added by a series of writers (methyltransferases (METTL3 and METTL14), wilms tumour 1-associated protein (WTAP), KIAA1429, putative RNA binding proteins (RBM15/15B), and METTL16) and are removed by erasers (fat mass and obesity associated protein (FTO) and α-ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) (Lorna). Apart from regulating RNA metabolism, these RNA writers and erasers collectively termed as RNA-modifying proteins (RMPs) are also implicated in the pathological progression. This novel area study called epi-transcriptome is linked to the initiation and progression of cancers and other diseases. Accordingly, mutations in the writer or eraser have been associated with cancers such as hepatocellular carcinoma and acute myeloid leukemia (AML), with memory, fertility and metabolic phenotypes Lorna (2019). Currently RNA epigenomic writers and erasers are showing promise as diagnostic markers and future therapeutic strategies for oncology Zhen and Hongjuan (2016)

Non-coding RNAs (microRNAs and long non-coding RNAs) as diagnostic and therapeutic tools: Majority of the transcriptome is represented by noncoding RNAs (ncRNAs) such as ribosomal (rRNA), transfer (tRNA), Piwi-interacting (piRNAs), micro (miRNA), small nuclear (snRNA), small nucleolar (snoRNA) and other types of RNA. Long noncoding RNAs (lncRNAs) are located and transcribed from different genomic locations. Using this criteria lncRNAs are generally placed into five categories; sense, antisense, bidirectional, intronic, and intergenic Lina ey al. (2013).

a) sense lncRNA are transcribed from the sense strand of a protein coding gene and contains exons from protein coding genes. b) Antisense lncRNA are transcribed from the antisense strand of a protein coding gene. c) Bidirectional lncRNA sequence is located on the opposite strand from a protein coding gene whose transcription is initiated less than 1000 base pairs away, but share common promoters with protein coding genes d) Intronic lncRNA – These are transcribed entirely from within an intron of protein-coding genes. e) Intergenic lncRNA – These are transcribed from intergenic regions from both the strands. These RNA sequences are not located near any other protein coding loci.

Using large-scale complementary DNA (cDNA) sequencing projects such as FANTOM5 (Functional Annotation of Mammalian cDNA), more than 19,000 potentially functional long ncRNAs have been identified in various human sources suggesting that these RNAs constitute a heterogeneous group of diverse functions (Lina).

lncRNAs act to regulate chromatin remodeling and gene expression at the epigenetic, transcriptional and translational level. Long ncRNAs attract microRNAs and regulate expression level of transcripts containing common miRNA binding sites. LncRNAs can also form duplexes with target mRNAs and inhibit their translation or disrupting their stability. In addition, some long noncoding RNAs can modulate pre-mRNA splicing, protein localization, telomere replication, RNA interference, beyond transcription and translation regulation. Although few lncRNAs have been functionally characterized so far, it is now clear that alteration in their expression can contribute to various cancer types, neurodegenerative disorders, cardiovascular diseases, conditions associated with genome imprinting, aging, eye diseases, immune response, apoptosis and other pathologies. Since lncRNAs and miRNAs are dysregulated in diseased states, they have the potential to be used as prognostic markers and novel therapeutic targets Ling (2013).

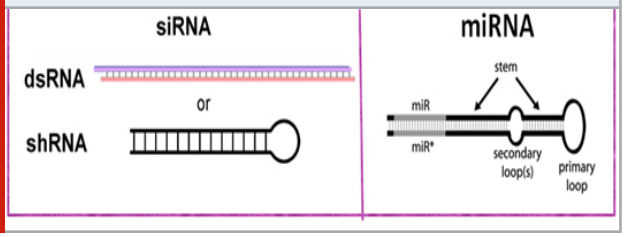
Difference between siRNA vs miRNA:

1. Unlike siRNAs (rarely conserved), many miRNAs are evolutionarily conserved in related organisms, which implies that they have important biological functions.
2. The major difference between siRNAs and miRNAs is that the former inhibit translation of a specific mRNA by degradation while the latter (miRNA) either degrade target mRNA (in plants) or inhibit the translation of multiple mRNAs because of imperfect base pairing (seen in animals).
3. miRNAs and siRNAs are generated in slightly different ways. miRNAs precursors are derived from distinct gene loci by RNA polymerase II transcription and subsequent processing into 23 nucleotide long miRNA by dicer-I enzyme. siRNAs do not have distinct gene loci and are derived from an exogeneous double-stranded RNA uptaken by the cell, and cleaved by dicer 2 enzyme. siRNAs are also encoded by transposable elements, viruses or heterochromatic DNA. Hence siRNAs are abundantly present in cells than miRNAs.
4. siRNAs are believed to be the most ancient form of RNA interference, with miRNAs being a later refinement.
5. miRNA can regulate multiple geens while siRNA mediate silencing of same genes from which they originate.
6. In terms of structural differences miRNA has a heteroduplex structure that includes an imperfect stem loop structure while the siRNA is a single duplex structure or extended hairpins Ian and Michael (2008) (Fig 1).

The first miRNA called lin-4 was discovered in Caenorhabditis elegans and seven years later lin-7 was identifies in the same organism Li and Kowdley

(2012). With the advancement in the areas of functional genomics and bioinformatics, novel miRNAs were implicated in various human diseases and already made a transition from laboratories to clinical research. miRNAs are stable in cells, tissues and body fluids making them as an ideal choice as biomarkers to study disease progression in various human diseases such as cancer, neurodegenerative, cardiovascular and liver disease and viral infection Szelenberger (2019).

Figure 1: Structural difference between siRNA vs miRNA



Engineering Plant Metabolic Pathways through RNAi: The recent understanding of functional genomics in non-coding RNAs that has substantially contributed to crop improvement is RNA interference (RNAi). Small interfering RNAs (siRNAs) and microRNAs (miRNAs) activate RNAi machinery inside the cells, abrogates the specific gene function without affecting other agronomic traits. Several desirable traits improved by RNAi gene silencing were biotic stress tolerance, removal of toxins from improved cultivars, development of seedless fruits, development of male sterile lines, delayed ripening of fruits, nutritional content, removal of secondary metabolites and altered floral characteristics Jagtap et al. (2011). RNAi has been used to modify plant metabolic pathways to enhance nutrient content and reduced toxin production (Table 1). The technique takes advantage of the heritable and stable RNAi phenotypes in plants.

Therapeutic application of alternative splicing: During the final step in formation of a mature, functional mRNA, the introns are removed from precursor mRNA and exons are joined together in a process called mRNA splicing. Each intron usually starts with a (5') GU and ends with an (3') AG (called GU-AG rule). Splicing occurs at these two Short, Conserved Sequences in Pre-mRNAs via Two Transesterification Reactions. In the first transesterification reaction, the ester bond is formed between the 5′ phosphorus of the intron-I and the 2′ OH of the branch-site Adenosine residue. In the second trans-esterification reaction, the ester bond is formed between the 5′ phosphorus of exon 2 and the 3′OH of exon 1, releasing the intron in a loop called lariat structure and joining the two exons. In each reaction, one phosphate-ester bond is exchanged for another. Since the number of phosphate-ester bonds in the molecule is not changed in either reaction, no energy is consumed. The net result of these two transesterification reactions is that two exons are ligated and the intervening intron is released as a branched lariat structure Verma and agarwal (2005).

In certain instances, the use of alternative splicing or alternative RNA splicing, or differential splicing of hnRNA, is a regulated and an evolutionary conserved process during gene expression that results in a single gene coding for multiple proteins. In this process, particular exons of a gene may be included within or excluded from the final, processed messenger RNA (mRNA) produced from that gene. Consequently, the proteins translated from alternatively spliced mRNAs will contain differences in their amino acid sequence and in their biological functions. Notably, alternative splicing allows the human genome to direct the synthesis of many more proteins than would be expected from its 20,000 protein-coding genes. Alternative patterns of RNA splicing result from tissue-specific adaptive and developmental control mechanisms. There are numerous modes of alternative splicing observed, of which the most common is exon skipping. In this mode, a particular exon may be included in mRNAs under some conditions or in particular tissues, and omitted from the mRNA in others.

Alternative splicing is a combinatorial process because many features are involved in the generation of transcript diversity such as splice site strength, intron-exon architecture, RNA secondary structure, splicing regulatory elements, promoter use and transcription speed by RNA polymerase and the presence of post-transcriptional nucleotide modifications. A comprehensive understanding of the factors that influence alternative splicing decisions is necessary to predict disease associated mis-splicing events Hossein and Hertel (2019). It has been estimated that approximately 15% of all mutations that cause genetic diseases result in defective splicing of pre-mRNA. Defects in splicing has been shown to be associated with genetic disorders such as β-thalassemia, cystic fibrosis, Duchenne muscular dystrophy, etc. the defects in alternative splicing contributes to cancer and also induce drug resistance Jhung-Chun (2018) Riboswitches are in vivo RNA aptamers that regulate both gene expression and alternative splicing by binding to small-molecule ligands Kenneth et al. (2009).

Therapeutic application of polyadenylation: Processing of the primary transcript involves the addition of a cap to the 5'-end, removal of introns, and polyadenylation (the addition of a poly(A) tail to the 3'-end) to produce the mature mRNA. Pre-mRNAs are first cleaved at specific 3′ Sites and subsequently polyadenylated. Both cleavage and polyadenylation of a pre-mRNA is carried out by by an enzyme, called poly (A) polymerase, utilizing ATP as a substrate. This step is called polyadenylation. The 3' end of most protein-coding genes and long non-coding RNAs is cleaved and polyadenylated. Control of polyadenylation is mediated by a number of sequence elements such as the polyadenylation site itself, but also a series of upstream (U and UGUA rich) and downstream (U and GU rich) elements. Similar to 5' capping, the 3' poly(A) tail is important for the export of mRNA from nucleus to cytoplasm, protein translation and stability of mRNA. The oligoadenylated RNAs are

efficiently decapped and degraded in 5' to 3' manner than polyadenylated and unadenylated RNAs.

Many genes contain more than one polyadenylation site. The use of alternative polyadenylation (APA) sites from a single transcript results in mRNA heterogeneity by generating mRNA isoforms that differ either in their coding sequence or in their 3' untranslated regions. Differential use of polyadenylation sites may also have impacts on function, stability, localization, and translation efficiency of target mRNAs. APA misregulation has been identified in human diseases. High levels of proliferation and differentiation capacity of cells correlates with increased relative expression of shorter 3'UTR isoforms and indicate UTR-based mRNA regulation Zheng and Tian (2014). Growing lines of evidence have shown that RNA-binding proteins (RBPs) play important roles in regulation of APA. Some RBPs are part of the machinery for cleavage and polyadenylation; others influence polyadenylation choice through binding to adjacent regions. Patterns of alternative polyadenylation are also regulated by differential binding of RNA binding proteins; cleavage stimulation factor (CSTF2) and cleavage factor I (CFI) of the main polyadenylation machinery have been shown to have effects on relative expression of alternatively polyadenylated isoforms Bin and Graber (2012).

Table 1. Improvement of quality traits engineered through RNAi Kamthan et al. (2015)

Trait	Target Gene	Host	Application
Enhanced nutrient content	Lyc	Tomato	Increased concentration of lycopene (carotenoid antioxidant)
	DET1	Tomato	Higher flavonoid and b-carotene contents
	omega-3 fatty acid desaturase [an enzyme converting linoleic acid (18:2) to alpha-linolenic acid (18:3)]	soyabean	Reduced alpha-linoleic acid content (1–3%) compared to non-transgenic soybean seed (7–10%).
	Starch branching enzymes (SBEIIa and SBEIIb)	Wheat, Sweet potato, Maize	Increased in amylose content for glycemic management and digestive health
	FAD2	Canola, Peanut, Cotton	Increased oleic acid content
	SAD1	Cotton	Increased stearic and oleic acid content
	22-kDa maize zein storage proteins (ZLKR/SDH)	Maize	Lysine-fortified maize
	Sucrose phos- phatase (SPP)	Potato	Reduction in conversion of sucrose to hexose sugars
Reduced alkaloid production	CaMXMT1 COR	Coffee Opium poppy	Decaffeinated coffee Production of non-narcotic alkaloid, instead of morphine
	CYP82E4	Tobacco	Six fold reduced levels of the carcinogen nornicotine in cured leaves
Heavy metal accumulation	ACR2	Arabidopsis	Arsenic hyperaccumulation for phytoremediation
Reduced polyphenol production	delta-cadinene synthase gene	Cotton	Lower gossypol levels in cotton seeds, for safe consumption
Ethylene sensitivity	LeETR4 ACC oxidase gene	Tomato Tomato	Early ripening tomatoes Longer shelf life because of slow ripening

Reduced allergenicity	Arah2	Peanut	Allergen-free peanuts
	Lolp5	Ryegrass	Reduced allergen in rye grass pollen
Reduced production of lachrymatory factor synthase (LFS)	Lyce3	Tomato	Reduced in tomato peel
	α-globulin and β-glyoxalase	Rice	mutated α-amylase/ trypsin inhibitor rice line
	lachrymatory factor synthase gene	Onion	Tearless onion

The cytoplasmic polyadenylation element binding protein 1 (CPEB1), an RNA-binding protein that regulates mRNA translation, also controls alternative 3'-UTR processing. CPEB1-mediated 3'-UTR shortening correlates with cell proliferation and tumorigenesis Ran et al. (2013). Other RNA binding proteins involved in regulation of cleavage and polyadenylation are RNA splicing factor hnRNP H1. Since RBPs influence the 3' end processing of alternatively polyadenylated transcripts, they can potentially be used as novel therapeutic agents Zheng and Tian (2014).

CRISPR/Cas9 genome editing: Genome editing technologies have become vital genetic tools in delivering pathogen resistance in plants. Due to higher success rate, easier to implement and less expensive nature of clustered regularly interspaced short palindrome repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) technology, it has overtaken other site directed modification methods such as meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs). This review focuses on the recent advances in plant protection using CRISPR/Cas9 technology in model plants and crops in response to viral, fungal and bacterial diseases. As regards to the development of viral disease resistance, integration of CRISPR-encoding sequences in Arabidopsis and Nicotiana benthamiana genome has been the main approach to prevent viral multiplication. With regards to fungal and bacterial disease resistance in crop species such as rice, tomato, wheat, and citrus, the strategies were based on targeted modification of susceptibility genes using CRISPR/Cas9 Borelli et al., (2018).

FUTURE PROSPECTS

Post transcriptional events such as alternative splicing, alternative polyadenylation and epigenetic modification of transcriptomes contribute to transcript heterogeneity and protein diversity, any misregulation in these events leads to human diseases. This is an exciting era of geneomic medicine where the future RNA-based therapeutics target post-transcriptional regulatory circuits at the interface of genes and proteins. Additionally development of siRNA constructs that can target multiple endogenous genes can usher development of novel plant types with altered phenotypic capabilities. By selectively suppression gene expression enhance the genetic value of crop plants that can be used for breeding for superior plant types that is impossible to be generated by conventional plant breeding methods. RNAi shortens breeding in woody trees by grafting transgenic RNAis on non-transgenic cutting for transmitting trait in a short time.

REFERENCES

Bin T., and H.G. Graber. 2012. Signals for pre-mRNA cleavage and polyadenylation. Wiley interdisciplinary reviews RNA. 3: 385–396.

Borrelli VMG, Brambilla V, Rogowsky P, Marocco A, Lanubile A. 2018. The Enhancement of Plant Disease Resistance Using CRISPR/Cas9 Technology. Front Plant Sci. 24: 1-20.

Hossein S., and K.J Hertel. 2019. Combinatorial regulation of alternative splicing. Biochimica et Biophysica Acta. 1862: 11–12.

Jagtap UB., R.G Gurav, V.A. Bapat. 2011. Role of RNA interference in plant improvement. Naturwissenschaften. 98: 473-92.

Jung-Chun L. 2018. Therapeutic Applications of Targeted Alternative Splicing to Cancer Treatment. International Journal of Molecular Sciences. 19: 75.

Kamthan A., A. Chaudhur, M. Kamthan and A. Datta. 2015. Small RNAs in plants: recent development and application for crop improvement. Frontiers in Plant Scinces. 6:208.

Kenneth JD., G. Veronica, G. Shikha, J.E. Shively, Y. Yen, and R.K. Gaur. 2009. Alternative splicing as a therapeutic target for human diseases. Methods in molecular biology (Clifton, N.J.) 555: 127–144.

Li Y., & K.V. Kowdley. 2012. MicroRNAs in common human diseases. Genomics, proteomics & bioinformatics. 10: 246–253.

Lina Ma., B. Vladimir, M. Bajic, and Zhang. 2013. On the classification of long non-coding RNAs. RNA Biology. 10: 924–933.

Ling W., W. Xingwu, L. Liyan, Y. Zheng, N. Zhang, and M. Yang. 2019. The emerging role of noncoding RNAs in colorectal cancer chemoresistance. Cellular Oncology. 42: 757–768.

Lorna, W. H. 2019. RNA biology provides new therapeutic targets for human disease. Frontiers in

Genetics10: 205.
Ran E., A.P. Ugalde and A. Reuven. 2013. Alternative cleavage and polyadenylation: extent, regulation and function. Nature Reviews Genetics. 14: 496–506.
Szelenberger R., M. Kacprzak, J. Saluk-Bijak, M. Zielinska and M. Bijak. 2019. Plasma MicroRNA as a novel diagnostic. Clin Chim Acta. 499: 98-107.
Verma, P.S. and Agarwal, V.K. 2005. Cell Biology, Genetics, Molecular Biology, Evolution and Ecology. 56-58. S. Chand & Company LTD. New Delhi.
Vu LP., B.F Pickering, Y. Cheng, S. Zaccara, D. Nguyen, G. Minuesa et al. 2017. The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. Nature Medicine. 23: 1369-1376.
Zhang S., B.S. Zhao, A. Zhou, K. Lin, S. Zheng, Z. Lu et al. 2017. Sustaining FOXM1 Expression and Cell Proliferation Program. Cancer Cell. 10: 591-606.e6.
Zhen D., and C. Hongjuan. 2016. The Emerging Roles of RNA Modifications in Glioblastoma. Cancers 12: 1-27.
Zheng D., and B. Tian. 2014. RNA-Binding Proteins in Regulation of Alternative Cleavage and Polyadenylation. In: Yeo G. (eds) Systems Biology of RNA Binding Proteins. Advances in Experimental Medicine and Biology 825: 97-125. Springer, New York, NY.