Chapter

Introductory Chapter: Gene Regulation, an RNA Network-Dependent Architecture

Payam Behzadi and Lernik Issakhanian

1. A historical background of genetics

Genetics is known as an old and ancient science that its origination goes back to at least 7000 years ago. Iranians are one of the earliest pioneers in genetics from ancient world. The brilliant Iranian (Persian) literature epic of *Shahnameh* edited by the Iranian shining star literate "Abolqasem Ferdowsi Toosi" is an invaluable evidence to prove this claim. By the time and progression in biology, the super-amazing molecule of DNA was discovered. Today, we know that the unique molecule of DNA involves the genetic and vital data within its bases as constitutional structures of nucleotides. Both eukaryotic and prokaryotic chromosomes are made up of DNA molecules. In addition to DNAs, the role and importance of RNAs are not lesser than DNAs [1].

In 1953, the interesting structure of DNA molecule with anti-parallel double-helix architecture was recognized by Watson and Crick. In 1958, the hypothesis of central dogma of molecular biology was published by Crick in which he described the translation of genetic language located on DNA into amino acid sequences of protein by the transient molecule of RNA (mainly messenger RNA (mRNA)). The primitive biological characteristics of mRNA were recognized in 1961, while these properties regarding ribosomal RNA (rRNA) and transfer RNA (tRNA) molecules were determined in the 1950s [2].

2. Central dogma

In accordance with central dogma of molecular biology, the genetic information hidden in the format of gene can be expressed throughout two vital processes of transcription (production of coding molecules of mRNA) and translation (in which nucleotides are replaced by amino acids). The central dogma's content involves both eukaryotes and prokaryotes. In this regard, the RNA polymerase (RNAP) contributes in the process of transcription to produce mRNA molecules from DNAs. In the process of translation, ribosomes are also contributed. As it is known, the eukaryotic processes of transcription and translation are separated; this separation is because of the presence of intron sequences among exons. Indeed, introns were discovered in 1977 in eukaryotes which resulted in the recognition of their localization among exon sequences. In other words, the arrangement of introns was revealed as noncoding mosaics of introns among the coding mosaics of exons. Normally introns, dependent on the eukaryotic cell's need, should be eliminated via the splicing process. The splicing process, which leads to the occurrence of the

spliceosome, separates the direct connection between transcription and translation stages, while the bacterial system misses splicing process. Therefore, the processes of transcription and translation couple together in bacteria. Furthermore, the coupling system of transcription-translation is achieved via the functional enzyme complex of RNAP and ribosome which is known as expressosome in bacteria [2–4].

Amazingly, the ribozyme domains (e.g. small self-cleaving RNAs) are present in noncoding sequences of introns and untranslated regions (UTRs) (mostly in 3′UTRs) [2, 5].

By the progression of knowledge and technology from the 1970s in the twentieth century, the importance of the role of RNA molecules increased more and more. Pre-mRNA molecules in eukaryotes which are known as heterogeneous nuclear RNA (hnRNA) have incredible role in gene regulation and expression. Different manners for the beginning of transcription, different patterns of splicing process, different lengths of poly-adenine tails and different patterns in RNA editing processes have direct effects on gene regulation and expression. Incredibly, the RNA molecules act as *trans*-acting factors that affect the processes of gene regulation and gene expression [6, 7].

These characteristics relating to RNAs support the RNA World Hypothesis (RWH). According to RWH which was proposed by Alex Rich, RNAs are bifunctional biopolymers which can appear in two different levels of bioactivities including enzymatic activities (ribozymes) and as informational genetic language. However, the Urzymes hypothesis challenges RWH. Urzymes are the earliest generation of enzymes originated from protein superfamilies with a conserved nucleus [6, 8–10].

The progression of the Internet, computers and online/offline software and tools, bioinformatics and computational biology and chemistry gives us a new horizon in association with gene, gene regulation, gene expression and gene structure. By the help of advanced knowledge and technology, now it is known that DNA molecules play their roles as CDs which are burned and contain cells' software tools as genetic information which encode into RNAs and proteins [1, 11].

3. RNA molecules, RNA network and gene regulation

In recent years throughout very strong evidences, it has revealed that the RNA molecules are not only temporal and transient transcripts which are obtained from the genes, but also they act as pivotal modulators that mediate pre-transcriptional and posttranscriptional steps and have direct effect on gene regulation and gene expression processes. Also, RNA molecules play a key role in producing and triggering vital signals within the wide web of genomic structure. Now, it is known that there are two main groups of coding RNA and noncoding RNA (ncRNA) molecules in which each of them has its own functions, structures and characteristics. The coding RNA molecules are those which can be translated into proteins, while the ncRNAs cannot be translated in proteins. According to previous studies, a large group of RNAs has no ability to be translated. And in mammals as an important group of eukaryotes, only less than 2% of the transcribed RNA molecules can be translated into protein molecules. So, the population of ncRNAs overpasses the population of coding RNAs [1, 2, 6, 12, 13].

RNAs are important allosteric molecules that regulate the processes of gene regulation and gene expression. The allosteric property of RNAs enables them to switch different pathways by regulation of genes to be expressed or silenced. A great cooperation among rRNAs, tRNAs, mRNAs and small nuclear RNAs (snRNAs) in the processes of eukaryotic transcription and translation explains the depth of the RNA roles within a cell. In addition, eukaryotes encompass small nucleolar RNAs (snoRNAs), too. snoRNAs with 60–300 nucleotide lengths are known as

intermediate size of ncRNAs. They have different roles including targeting a diversity of RNA molecules such as mRNA. This property depicts the regulatory role of snoRNAs. In this regard, there is another group of snRNAs known as small Cajal body-specific RNAs (scaRNAs) located in Cajal bodies or coiled bodies (CBs). CBs are compact organelles which are recognized within the eukaryotic nuclei. The CBs are consisted of a mass of proteins including snRNPs and snoRNPs. scaRNAs and snoRNAs are significant molecules which contribute in processing and modification of other RNA molecules like mRNA, rRNA, tRNA and snRNA. Interestingly, rRNAs are remarkable molecules which constitute ribosomal peptidyl transferase site. This illustrates the important role of RNAs as ribozymes which contribute in activation of peptidyl transferring and constitution of peptide bonds within polypeptides. Today it is known that short ncRNAs (such as microRNAs (miRNAs), Piwi-interacting RNAs (piRNAs), transcription initiation RNAs (tiRNAs)), promoter-associated RNAs (PARs) (including promoter-associated small RNAs (PASRs), transcriptional start site-associated RNAs (TSSa-RNAs) and promoter upstream transcripts (PROMPTS)) and long noncoding RNAs (lncRNAs) (comprising large intergenic noncoding RNAs (lincRNAs), transcribed ultraconserved regions (T-UCRs), telomeric repeat-containing RNAs (TERRAs), etc.) have unique role in gene regulation and gene expression [1, 2, 6, 14–19].

In prokaryotes such as bacteria, a diversity of small RNAs (sRNAs) act as regulatory molecules. These molecules are contributed in gene regulation and gene expression. Moreover, the *cis*-acting sequences of regulatory RNAs which are known as riboswitches have an effective role on gene regulation. In addition, some special DNA sequences have been recognized in bacteria and archaea which are known as clustered regulatory interspaced short palindromic repeats (CRISPRs). CRISPERs are able to be transcribed. The processed CRISPR transcripts act as guide RNAs (gRNAs) which destruct the viral nucleic acids including RNA or DNA molecules [2].

4. The importance and the role of NCRNA molecules

The ncRNAs are divided into two groups of lncRNAs (with >200 nucleotide lengths) and small ncRNAs (sncRNAs) (with 20–35 nucleotide lengths). Among different types of sncRNAs, the miRNAs are well-studied RNAs. They belong to RNA interference (RNAi) members with the length of 19–24 nucleotides. Amazingly, >60% of the coding genes are regulated by miRNAs. In other words, these molecules contribute in posttranscriptional process by silencing the related genes by prevention of translation. Drosha and Dicer are key enzymes to manufacture a mature miRNA molecule. The linkage of a single-stranded miRNA with the members of Argonautes produces RNA-induced silencing complex (RISC). The RISC binds to its complementary sequence on the 3′UTR section of targeted mRNA. In addition to miRNA pathway, piRNA and siRNA pathways are members of RNAi pathways. piRNAs (with about 30 nucleotide lengths) and siRNAs resembling miRNAs bind to Argonautes to produce RISCs which contribute in gene silencing process. However, miRNAs have genomic origination, while the siRNAs may have exo- or endogenous origination. The piRNAs are free dicer pathways which bind to PIWI proteins. The PIWI proteins are known as the subset of Argonautes. The complex of PIWI proteins and piRNAs contributes in transposon repression via degradation of transcripts belonging to transposable elements (epigenetic regulation). Furthermore, the piRNAs mediate the DNA methylation too. The tiRNAs as members of sncRNAs with about 17 nucleotide lengths may act as transcriptional regulators [1, 13, 18, 20, 21].

snoRNAs contribute in rRNA modification within the nucleolus. These modifications involve pseudouridylation and 2'-O-methylation of rRNAs to promote the integrity and rRNA folding. According to characterizations of secondary structure and common motifs, the snoRNAs are divided into two main families of box C/D and box H/ACA. Interestingly, the majority of snoRNA molecules are mostly expressed from housekeeping gene introns. The presence of abundance of snRNAs and scaRNAs support the rate of snRNA modification [2, 16, 18].

PASRs (with 22–200 nucleotide lengths), PROMPTs (with lesser than 200 nucleotides in length) and TSSa-RNAs (with 20–90 nucleotide lengths) are categorized within the intermediate ncRNAs in size and have not been studied well [18].

The lncRNAs are consisted of heterogeneous ncRNAs with the length of >200 nucleotides. They are known as the dominant mammalian ncRNAs of transcriptome. lncRNAs are recognized as effective eukaryotic ncRNAs with a wide range of functions. lncRNAs are involved in gene expression and gene regulation by regulating the posttranscriptional processes, epigenetic modifications, etc. [2, 18, 22].

In conclusion, we can claim that RNA molecules have a key role in molecular biology. In recent years, the sciences of bioinformatics, computational biology and computational chemistry have offered us a new promise regarding the importance of RNAs in association with gene regulation.

Acknowledgements

We have special thanks to Lada BOZIC, the author service manager of InTechOpen Company for her excellent collaboration, management and arrangement for preparing this valuable book.

Conflict of interest

The authors declare no conflicts of interest.

Author details

Payam Behzadi[1*] and Lernik Issakhanian[2]

1 Department of Microbiology, College of Basic Sciences, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran

2 Independent Researcher, PCR Specialist, Yerevan, Armenia

*Address all correspondence to: behzadipayam@yahoo.com

References

[1] Behzadi P, Behzadi E, Ranjbar R. Basic Modern Molecular Biology. 1st ed. Tehran: Persian Science & Research Publisher; 2014

[2] Morris KV, Mattick JS. The rise of regulatory RNA. Nature Reviews. Genetics. 2014;**15**(6):423

[3] Klaholz BP. The ribosome holds the RNA polymerase on track in bacteria. Trends in Biochemical Sciences. 2017;**42**(9):686-689

[4] Kohler R, Mooney R, Mills D, Landick R, Cramer P. Architecture of a transcribing-translating expressome. Science. 2017;**356**(6334):194-197

[5] De La Peña M, García-Robles I. Intronic hammerhead ribozymes are ultraconserved in the human genome. EMBO Reports. 2010;**11**(9):711-716

[6] Licatalosi DD, Darnell RB. Resolving RNA complexity to decipher regulatory rules governing biological networks. Nature Reviews. Genetics. 2010;**11**(1):75

[7] Bass BL. RNA editing by adenosine deaminases that act on RNA. Annual Review of Biochemistry. 2002;**71**(1):817-846

[8] Neveu M, Kim H-J, Benner SA. The "strong" RNA world hypothesis: Fifty years old. Astrobiology. 2013;**13**(4):391-403

[9] Li L, Francklyn C, Carter CW. Aminoacylating urzymes challenge the RNA world hypothesis. Journal of Biological Chemistry. 2013;**288**(37):26856-26863

[10] Carter CW. Urzymology: Experimental access to a key transition in the appearance of enzymes. Journal of Biological Chemistry. 2014;**289**(44):30213-30220

[11] Omoto CK, Lurquin P, Lurquin PF. Genes and DNA: A Beginner's Guide to Genetics and its Applications. New York, United States: Columbia University Press; 2004

[12] Fatica A, Bozzoni I. Long non-coding RNAs: New players in cell differentiation and development. Nature Reviews. Genetics. 2014;**15**(1):7

[13] Bunch H. Gene regulation of mammalian long non-coding RNA. Molecular Genetics and Genomics. 2018;**293**(1):1-15

[14] Bandyra K, Luisi B. Central dogma alchemy. RNA. 2015;**21**(4):558-559

[15] Ronchetti D, Todoerti K, Tuana G, Agnelli L, Mosca L, Lionetti M, et al. The expression pattern of small nucleolar and small Cajal body-specific RNAs characterizes distinct molecular subtypes of multiple myeloma. Blood Cancer Journal. 2012;**2**(11):e96

[16] Meier UT. RNA modification in Cajal bodies. RNA Biology. 2017;**14**(6):693-700

[17] Xie J, Zhang M, Zhou T, Hua X, Tang L, Wu W. Sno/scaRNAbase: A curated database for small nucleolar RNAs and cajal body-specific RNAs. Nucleic Acids Research. 2006;**35**(suppl_1):D183-D1D7

[18] Esteller M. Non-coding RNAs in human disease. Nature Reviews. Genetics. 2011;**12**(12):861

[19] Darzacq X, Jády BE, Verheggen C, Kiss AM, Bertrand E, Kiss T. Cajal body-specific small nuclear RNAs: A novel class of 2′-O-methylation and pseudouridylation guide RNAs. The EMBO Journal. 2002;**21**(11):2746-2756

[20] Wilson RC, Doudna JA. Molecular mechanisms of RNA interference.

Annual Review of Biophysics. 2013;**42**:217-239

[21] Ku H-Y, Lin H. PIWI proteins and their interactors in piRNA biogenesis, germline development and gene expression. National Science Review. 2014;**1**(2):205-218

[22] Bánfai B, Jia H, Khatun J, Wood E, Risk B, Gundling WE, et al. Long noncoding RNAs are rarely translated in two human cell lines. Genome Research. 2012;**22**(9):1646-1657