MicroRNA biogenesis and variability

Jesús García-López, Miguel A. Brieño-Enríquez and Jesús del Mazo*


*Corresponding author: Jesús del Mazo. Dep. Cellular and Molecular Biology. Centro de Investigaciones Biológicas. CSIC. Ramiro de Maeztu, 9. 28040 Madrid. Spain. Telephone number: 34 918373112 #4324. Fax. 34 915360432. Email: jdelmazo@cib.csic.es

Running title: MicroRNA biogenesis
Abstract

MicroRNAs (miRNAs) are cell-endogenous small non-coding RNAs, which by mechanisms of RNA interference are involved in the post-transcriptional regulation of mRNAs. The biogenesis and function of miRNAs entail multiple elements with different alternative pathways. These, confer a high versatility of regulation and a high variability to generate different miRNAs and hence possess a broad potential to regulate gene expression. Here, we review the multiple, both canonical and non-canonical, mechanisms prevalent to generate miRNAs in animals. The “miRNome” panorama is enhancing our knowledge regarding the fine regulation of gene expression and is providing new insights concerning normal as opposed to pathological cell differentiation and development.

Key words: microRNAs, canonical biogenesis, non-canonical biogenesis, isomirs, editing

List of abbreviations:

(ADAR) Adenosine Deaminases Acting on RNA; (AGO) Argonae; (circRNAs) circular RNAs; (DGCR8) microprocessor complex subunit DiGeorge syndrome Critical Region 8; (DICER) cytoplasmic endoribonuclease with dsRNA "dicing" activity; (dsRNA) double-stranded RNA; (DROSHA) double-stranded RNA-specific endoribonuclease nuclear type III, also known as RNASEN; (endo-siRNAs) endogenous-small-interfering RNAs; (miRNAs) microRNAs; (PAZ-domain) protein domain contained in the PIWI, ARGONAUTE and ZWILLE proteins, involved in post-transcriptional gene silencing; (PTC) Papillary Thyroid Carcinoma; (piRNAs) piwi-
interacting RNAs; (pri-miRNAs) primary miRNAs; (pre-RNAs) precursor-miRNAs; (RISC) RNA-Induced Silencing Complex; (SNPs) Single Nucleotide Polymorphisms; (snoRNAs) small nucleolar RNAs; (TUDOR) Tudor-staphylococcal nuclease.

Introduction

MicroRNAs (miRNAs) are small non-coding RNAs (~21-23 nt) that act by RNA interference as post-transcriptional regulators of gene expression and play important roles in many developmental and cellular processes of eukaryotic organisms. Although the miRNA biogenesis pathways are similar in different phyla, there exist some differences between plants and animals that require a different study with respect to either biogenesis pathway [1]. We will focus in this review on the biogenesis of the miRNAs associated to the animal kingdom.

MicroRNAs regulate negatively gene expression via base-pairing with complementary mRNA sequences in a mechanism named RNA interference. The interaction between miRNAs and their mRNA targets blocks the regular translation or induces the degradation of targeted mRNAs. MicroRNAs might regulate about 60% of mammalian coding genes and show specific profiles of expression at the level of cell types and developmental stages. Genes coding for miRNAs are scattered throughout mammalian genomes in intra- and intergenic positions. At present, 1,600 miRNA precursor molecules that generate 2,042 mature forms have been identified in the human genome (as to miRBase release 19) [2, 3]. Mouse and rat genomes encode 1,281 and 723 mature miRNAs respectively [2, 3]. However, it is possible that the existing difference
with respect to the size of the human and rodent “miRNomes” results from a so far more exhaustive quest for human miRNAs. In fact, miRNAs are well conserved in eukaryotic organisms [4] and are thought to be the result of a vital and evolutionarily ancient component of genetic regulation.

The binding of mature miRNAs to their target mRNAs occurs through a specific miRNA region of about 6-8 nucleotides in length. This region is termed the ‘seed region’ and allows for each of these small RNAs to regulate the expression of hundreds of genes, generally by binding to the mRNA 3’ UTR region. MicroRNAs that share a similar ‘seed region’ belong to the same miRNA family. In general, members of a miRNA family regulate related genes or are involved in the regulation of similar biological events. Several families of miRNAs in mammals have been described, some of which are related to mechanisms of differentiation and self renewal and thus have been suggested as possible biomarkers for these types of processes. For example, the let-7 or miR-30 families promote differentiated cell fates while miR-290 family members are associated to self-renewal events.

Despite of the fact that miRNAs have been for the best characterized small non-coding RNAs, some aspects of their biogenesis still remain to be uncovered. Recent studies have revealed alternative mechanisms of biogenesis and recycling of miRNAs [5-11]. This review will mainly focus on the mechanisms of miRNA biogenesis generated by both canonical and non-canonical pathways. Furthermore, we will take in hand the study of miRNA recycling mechanisms and posttranscriptional modifications, due to the alternative processing of precursors and editing processes. Recycling mechanisms
allow the production of new mature miRNA molecules without a need to generate new miRNA precursors.

The canonical miRNA biogenesis pathway

Typically, miRNA biogenesis is directed through a specific promoter or as part of a host gene in which miRNA is usually enclosed within intronic regions. Some miRNAs are closely located in a genome and are transcribed as part of a common transcript (cluster of miRNAs), similar to polycistronic units [12, 13]. RNA polymerase II transcribes miRNA genes as large polyadenilated RNA molecules named primary miRNAs (pri-miRNAs) [14-16] (Fig. 1). These pri-miRNAs contain sequence complementary regions capable of materializing stem loops which include bulges and mismatches. Primary miRNAs are processed in the nucleus by a protein complex containing the RNase III enzyme DROSHA and DGCR8 [17, 18]. As result of this processing, pri-miRNAs are cleaved into smaller double-stranded RNA (dsRNA) molecules known as pre-microRNAs. In mammals and other vertebrates, together with flies pre-miRNAs are exported to the cytoplasm by exportin-5 (XPO5) [19-21]. In the cytoplasm, pre-microRNAs are cleaved and despoiled of their loops by the RNase III enzyme DICER, which through interaction with the protein TRBP produces double stranded RNAs characterized by overhangs of 2-3 nucleotides at both ends [22, 23]. These processed products are known as miRNA duplexes, as functional mature miRNAs are single stranded. The last processing step requires the active participation of a ribonucleoprotein complex known as RNA-induced silencing complex (RISC) which is
able to unwind both strands. Although either strand of the miRNA duplex could potentially act as a mature miRNA, usually only one of the strands is incorporated into the RISC complex to induce the mRNA silencing [24, 25]. The strand selection process is still a subject of study full of recurrent controversies [24-29]. Specificity of either strand has been associated to different cell types and developmental stages [30]. Once loaded, RISC mediates in the recognition of the mRNA to be targeted.

Key components of the RISC complex are the Argonaute (AGO) family proteins [31-34]. Argonaute proteins bind to the different types of small non-coding interference RNAs such as: miRNAs, endogenous-small-interfering RNAs (endo-siRNAs) and piwi-interacting RNAs (piRNAs). In mammals, seven Argonaute family proteins have been described. These proteins have been classified in two subfamilies known as the AGO subfamily and the PIWI subfamily [31-34]. The proteins encompassed in the AGO subfamily (AGO1-4) are involved in the miRNA and the endo-siRNA pathways, nonetheless, only AGO2 displays endonuclease activity [35-37]. On the contrary, PIWI proteins are only involved in the piRNA pathway.

The non-canonical biogenesis of miRNAs

In the past few years alternative sources of miRNAs have emerged. The biogenesis of such miRNAs entails some elements belonging to the canonical pathway but the genomic origin differs from that of the classical miRNAs (Fig. 1). Although most miRNAs are located in intergenic regions, some are derived from intragenic segments. A particular miRNA type derived from intra-transcribed loci is that known as mirtron
together with its variant called a sintrom. Mirtrons, localized in the intronic regions of mRNAs, can generate double stranded loop structures as do the regular miRNAs but are processed to pre-miRNAs by the spliceosome machinery (Fig 1). Their biogenesis pathway is DROSHA/DGCR8-independent as was demonstrated in Drosha mutants [38] and DGCR8 knockout cells [39, 40]. The existence of mirtrons, initially discovered in C. elegans and D. melanogaster [41, 42], has been demonstrated in different organisms from plants to mammals [43-45]. Following resolution of the intron lariat, the RNA product of splicing adopts a pre-miRNA like form, which can be further processed as a canonical pre-miRNA bound to XPO5, to then be transferred to the cytoplasm to continue with the canonical pathways (Fig. 1). The usual concept regarding a mirtron entails that the pre-miRNA-like mirtron is generated by the direct cleavage and excision within the splice donor and acceptor sites of the mRNA. That is, the spliceosome machinery generates both prime ends of the pre-mirtron. However, not in all pre-mirtrons do the ends of the double stranded regions coincide with the ends of the intron. Some pre-mirtrons retain at the 3’ or 5’ end a single stranded tail [39, 41, 46] which has to be processed before proceeding to the cytoplasmic export and cleavage by DICER. Such trimming is carried out by the action of the RNA exosome components, such as the Rrp6 nuclear exonuclease at the 3’ tail which stops its action at the secondary structure of the double-stranded stem of the pre-mirtron [46]. In vertebrates, 5’ end tailed mirtrons have been identified [39, 47, 48], while 3’ tailed pre-mirtrons have only been detected in Drosophila. Specific exonuclease trimming of the 5’ end tailed pre-mirtrons has not been identified yet. Identification by deep sequencing of mirtrons in
mammals [44] and the differences detected from those found in insects suggest a relatively late evolutionary diversification of mirtrons.

A recently discovered mirtron variants are the splicing-independent mirtron-like miRNAs, termed simtrons [49]. The main characteristics of this new variant of non-canonical miRNA biogenesis are the spliceosome-independent but DROSHA-dependent processing of the pre-simtrons together with an XPO5 lack of dependence for transport to the cytoplasm. Replacement of DICER processing by AGO2 slicing of miRNAs was reported as a non-canonical alternative to the miRNA biogenesis of particular miRNA [36, 50-52] However, simtron processing are independent through DICER or AGO2. This novel type of a mirtron-like microRNA enhances the mechanisms that control the regulation of miRNA biogenesis and as a consequence the functional role of specific miRNA originated from such alternative mechanisms.

Other sources of miRNAs bypassing canonical biogenesis have also been described. Some small nucleolar RNAs (snoRNAs) can render miRNAs [53, 54], which are similar to mirtrons substrates for processing in a DICER-dependent and DROSHA/DGCR-independent pathway [55]. Transfer RNAs can also generate miRNAs which have to be sliced by DICER [56]. We have also recently identified some piRNAs in oocyte and sperm mouse cells with potential mRNA targets, in other words, with functional roles similar to miRNAs (manuscript submitted) [57].

**Regulatory mechanisms involved in the miRNA biogenesis pathway**
The main regulatory mechanisms that operate in the canonical miRNA biogenesis pathway act on three levels: first, during the transcription of pri-microRNAs, second, by means of the editing mechanisms that can disrupt the processing of precursors and finally through the regulation of the processing machineries such as the DROSHA/DGCR8 complex or DICER.

With regard to the pri-microRNA regulation of transcription, in a manner similar to that of coding genes, it is mediated by transcription factors that are specific to cell types, respond to environmental stimuli or are necessary to trigger developmental pathways. For example, as to the transcription factors OCT4 and SOX2, which are involved in stem cell maintenance, these in turn also regulate the transcription of the *mmu-miR-302* miRNA cluster [58-60]. Both OCT4 and SOX2 proteins bind to the conserved promoter region of *miR-302* in a manner that *miR-302* and their cluster partners are expressed in the same cells or tissues and at the same time as Oct4 and Sox2 [61-63]. Another example regards the regulation of *pri-miRNA-34* transcription by the P53 protein [64, 65]. In response to DNA damage, P53 is activated and acts on the promoter of *miR-34* thus activating expression. It is thought that *miR-34* induces the stop signals of the cell cycle pathway. Additionally, similar as occurs in protein-coding genes, the methylation of promoter regions can affect miRNA expression. For example, miRNAs involved in tumor suppression such as *miR-148, miR-34b/c* or *miR-9* undergo aberrant hypermethylation patterns that have been associated to cancer disease [66-68]. The fact that the methylation of regulatory genome regions can also operate on the regulation of miRNA expression implies that the epigenetic modifications of the genome
at miRNA loci can induce additional check points with reference to epigenome regulatory mechanisms.

Precursor miRNAs can suffer sequence modifications by ‘editing’ before they are processed. The elicitors of the editing mechanisms are the protein family of the Adenosine Deaminases Acting on RNA (ADAR) [69-72]. The members of this family act on dsRNAs, including miRNAs, by modifying their nucleotide sequences. The editing process involves the specific deamination of one or more adenosine nucleotides by transforming adenosine into inosine (Fig. 1 and 2) [69-72]. Inosine is recognized by the translation machinery and in the RNA-RNA binding as nucleotide guanosine. However, it has been reported that the edition can disrupt the recognition of pri- and pre-miRNA by the DROSHA/DGCR8 and DICER processing machineries [73, 74]. Alternatively, when edition affects the seed sequence of mature miRNA it can cause a phenomenon known as retargeting [75, 76]. The editing mechanism will be discussed in detail further on.

Finally, the activity of DROSHA/DGCR8 and DICER can be regulated during the biogenesis of some miRNAs. For example, the ribonucleoprotein hnRNPA1 binds to the loop region of pri-microRNA-18a to facilitate its processing by DROSHA/DGCR8 well ahead of their miRNA cluster partners (miR-17-92 cluster) [77, 78]. On the other hand, the activation of the ERK protein mediates the phosphorylation of TRBP by stabilizing the binding of pre-microRNAs with DICER [79]. As a result of this phosphorylation, the precursor molecule processing efficiencies are increased. Another example is the regulation of the let-7 biogenesis. The protein LIN28 represses both the processing of pri-microRNA to pre-microRNA of the let-7 microRNA as well as the
pre-microRNA-let-7 processing by DICER [80-82]. By a feedback mechanism, let-7 also regulates the translation of Lin28 mRNA [83].

**Bioavailability and miRNA recycling**

The global bioavailability of miRNAs is presumably related to the availability of the different elements of the miRNA biogenesis machinery. Changes in the levels of expression of the genes encoding for the biogenesis and function of miRNAs can be modulated during development or cell type differentiation [84]. Key element depletion in the biogenesis pathway could condition the production stop of new mature miRNAs. Such switch-off mechanisms can comprise functional roles during cell differentiation and development which drive the reprogramming of the miRNA-mediated gene expression. For example, the suppression of the maternal program to initiate the zygotic activation program is one of the biological processes that occur during the early stages of embryo development after fertilization. The suppression of the miRNA activity during these stages in mammals was suggested by studies in Dgcr8<sup>−/−</sup> mutant mice. Following this line, we have also reported on the global decay of expression of genes involved in the canonical biogenesis pathway from fertilization to blastocyst mouse embryos [5]. However, a global lack of new miRNA biogenesis does not necessarily imply an absence of specific miRNA availability. We have previously demonstrated, that in early embryo stages in which expression of genes encoding proteins involved in the biogenesis of miRNAs was dramatically down-regulated, specific miRNAs such as *mmu-miR-292-3p* and *mmu-miR-292-5p* could be preserved as double stranded molecules by a ‘protection’
through the binding to mRNA targets, pseudogenes, duplex passenger strands or other types of RNA-specific reservoirs [5] (Fig. 1). A similar miRNA protection by their heteroduplex cognate mRNA targets has also been reported in *C. elegans* [85]. Recent articles corroborate on the existence of such mature miRNA reservoirs in cells [7, 8, 86]. These reservoirs have the capacity to capture mature miRNA molecules to inhibit the miRNA silencing activities (Fig. 1). Some of these reservoirs are in fact RNA molecules that have a circular shape and can act as miRNA ‘sponges’ in which the miRNAs could be stored until their hypothetical recycling [6]. This new type of RNA molecule has been given the name of circular RNAs (circRNAs) [7, 8].

Functional bioavailability of specific miRNAs can also be modulated by the regulation of endogenous elements able to ‘sequester’ miRNA molecules. This hypothesis has been demonstrated by the regulated expression of pseudogenes which share the 3’ UTR of their corresponding coding genes allowing the binding of miRNAs that are naturally involved in the post-transcriptional negative regulation of the coding genes. This fact has been reported in the up-regulation of *PTEN* caused by the over expression of *PTENP1*, a known *PTEN* pseudogene, resulting in the suppression of cell growth. In turn, *PTEN* mRNA can act as a decoy for miRNAs that would down-regulate *PTENP1* transcripts [87].

A yet poorly characterized aspect of miRNA activity is the fate of these small RNA molecules after eliciting the mRNA silencing. Furthermore, how the miRNA-induced suppression finalizes, remains still an enigma. MicroRNAs have an unexpected long half-life and can reach high cell concentrations (Baccarini et al. 2011), even when key miRNA biogenesis elements such as DROSHA/DGCR8 or DICER are absent [88]
or show very low expression levels [89]. The existence of miRNA recycling pathways following the regulation of their targets facilitates their turnover.

One of the main features of miRNAs consists in their capacity to regulate multiple mRNA targets. However, the number of target transcripts in a cell often exceeds the number of miRNAs capable of regulating their transcripts [90]. It remains extremely difficult to explain how a relative low number of miRNAs are able to down-regulate a large number of transcripts. The constant recycling of miRNAs derived from loci where mRNA target degradation is actively being carried out could explain the efficiency of miRNAs with regard to their gene silencing activities. Recycling exponentially increases the capacity of each mature miRNA molecule to regulate even hundreds of mRNA molecules without requiring a new processing of the miRNA precursor molecule. Nevertheless, where and how these miRNAs are stored until use remains poorly understood. The existence of molecular miRNA reservoirs, as mentioned above, could explain some of the peculiarities of the functional miRNA dynamics. Previous to the discovery of circRNAs, the possibility had already been suggested that mRNA targets could act as active reservoirs of the miRNAs that regulate them [89].

In general, recent data has suggested that the interaction between miRNAs and their mRNA targets could be driven by dynamic ways. If the miRNA/mRNA interaction is the cause of target degradation, the miRNA could bind again to another, similar or different mRNA target molecule. Nonetheless, if the binding between the miRNA and its target inhibits translation without an mRNA cleavage, then the miRNA could keep on attached to the mRNA target and the mRNA would act as a reservoir of miRNAs. Finally, other types of RNAs such as pseudogene-coded mRNAs or circRNAs could
capture itinerant miRNAs, thus avoiding the interaction between miRNAs and their potential mRNA targets, which in response to demand could directly provide many functional and mature miRNAs.

“One for all and all for one”: enhancing versatility

Based on the principles of miRNA-target recognition, on average each miRNA can recognize about 100-200 potential target sites of the transcriptome [91-93], considering only the seed region of the miRNAs and the 3’UTRs of the mRNAs, although additional interaction sites occur [94]. In turn, each transcript has various predicted and functional miRNA sites [91, 95, 96]. More than 50% of the human protein-coding genes contain conserved miRNA targeting sites [97]. This implies that the number of combinatorial miRNA-mRNA interactions enormously enhance regulatory possibilities. Moreover, this potential variability is further expanded if we consider the alternative transcription of genes coding for mRNA targets (such as SNPs, alternative polyadenylation sites and alternative splicing products) along with all the different alternative modifications that result from the miRNA alternative biogenesis mechanisms.

MicroRNA editing

As dsRNAs, miRNA precursor molecules are targets of ADAR proteins. Recent evidences indicate that A-to-I editing of miRNA precursors (pri- and pre-miRNAs) affects both the miRNA function and biogenesis [98-100]. The editing of miRNA precursor molecules can alter the DROSHA/DGCR8 and DICER/TRBP recognition and
processing. Initially, evidence on miRNA editing was first observed during the processing of miR-22 [101] and was later also detected during the biogenesis of other miRNAs. Although, in the beginning it was suggested that miRNA precursor editing acted as a negative modulator of miRNA biogenesis, recently it has been demonstrated that the edited mature miRNAs can coexist with canonic mature miRNA molecules in the cell [74, 75, 102]. Actually, it has not been well established how the editing mechanisms trigger the elimination of miRNAs via TUDOR [99, 103] or, alternatively, modify miRNA precursors by changing their nucleotide sequence, to generate the alternative mature miRNA isoforms [75]. It seems that both routes act simultaneously. For example, research on miRNAs by deep sequencing in oocytes and zygotes revealed that mature miRNAs such as miR-376a, let7-g, miR-27a and miR-411, contained a considerable number of A-to-I sequence editions. In contrast, the edition levels of other miRNAs as for instance miR-151, miR-379 or miR-376b (-5p form) were low [75].

In general, the outcome of ADAR editing varies according to where the edited nucleotides are located. It has been established that the editing can impinge on the processing by DROSHA/DGCR8 and DICER/TRBP [100, 103]. However, is not yet clear whether edition results in a change of any recognition sequence into miRNA precursor molecules or if on the contrary the spatial conformation of miRNA molecules is changed. As dsRNA molecules, the pri- and pre-microRNA adopt tridimensional structures that could determine the accessibility of proteins involved in miRNA biogenesis. Secondary structures are also determined by the nucleotide sequence. In this manner, A-to-I editions can modify the secondary structure of precursor molecules,
hampering DROSHA/DGCR8 and DICER/TRBP recognition and in consequence hindering miRNA precursor processing.

The fate of the edited precursor molecules which are not processed has so far not been firmly established. Hypothetically, edited miRNA precursors could be maintained into cell reservoirs until needed for use as occurs with mature miRNAs. Nevertheless, the degradation of the edited-non-processed miRNA precursor molecules is so far the only mode reported [103]. For example, the editing of *miR-142* precursor molecules blocks the DROSHA/DGCR8 pri-microRNA cleavage [100]. It has been observed that *pri-miRNA-142* sequences displayed multiple A-to-I changes. These hyper-edited miRNA precursors were degraded via TUDOR-SN. This endonuclease recognizes inosine residues in dsRNAs and induces their cleavage [100, 103, 104]. Consequently, ADAR proteins could mark the miRNA precursor molecules that must be cleaved by hyper-editing their adenosines. It has not been well established how many A-to-I changes are necessary to promote the degradation of precursor miRNA molecules by TUDOR-SN. However, recently it was suggested that this degradation could occur for specific cytoplasm domains that have been named T-bodies (TUDOR-SN bodies) [75].

Nevertheless, editing can also modify the mature sequence of miRNAs. Such editions probably alter the function instead of the miRNA biogenesis. Due to the precise recognition requirement between the miRNA seed region and its specific target, A-to-I editions which affect the seed region can generate miRNA retargeting. However, both in mice and humans, editing and other polymorphisms are not promiscuously detected in the seed regions [75, 105], suggesting a strong selective constraint on the miRNA/mRNA recognition site. Finally, ADAR proteins can act directly on mRNAs
by modifying the recognition-site where the miRNA potentially bind [106]. This could also have an effect on the binding between the miRNA and their targets with regard to mRNA.

*IsomiRs*

Usually, miRNAs are annotated as a single defined sequence [107]. Using RNA-sequencing, it has been observed that a pre-miRNA very often gives rise to more than one mature miRNA sequence [107-109]. These variants are named isomiRs and almost all of these molecules were initially considered to be artifacts [107, 110-112]. Nowadays, the capacity of isomiRs to be associated to RISC and to the translational machinery of polysomes has been demonstrated, further indicating that they could also interact with mRNAs [110, 113]. These miRNA variants can encompass substitutions, insertions or deletions (polymorphic isomiRs), 3´-isomirs and 5´-isomirs. 5´and 3´ isomiRs include non-template additions and 5´ - 3´ cleavage variations [107, 110] (Fig. 2).

In terms of the number of miRNAs and their overall abundance, the most common type of isomiR in animals and plants consist of the 3´ isomiRs [113-116]. As in regular miRNAs, isomiRs vary among different cells or tissues and according to specific biological stimuli [114-117]. This suggests that the presence of some isomiRs could be regulated according to different cell functions.

*5´ and 3´ template modifications*

It has been assumed that DICER needs a defined pre-miRNA distance of 22nt from the 5´ to the 3´end in order to be able to cleave both strands to produce a miRNA
duplex [118-121]. However, a high proportion of the identified isomiRs containing 5’ and 3’ modifications were derived from processing variations in the cleavage position of the precursor molecules by the DROSHA or DICER enzymes [107, 122, 123]. The most abundant isomiRs have been found to differ only by 1 or 2 nt at the 5’-or 3’-end of their sequences [123]. These results have indicated that DROSHA is more specific than DICER during the cleavage process. In consequence, more variation with regard to cleavage sites occurs near to the loop as compared to the base of the stem in pre-miRNA hairpins. DICER’s cleavage imprecision can be explained by the simultaneous DICER protein recognition of the RNA-recognition PAZ domain and the RNAse III domain. DICER’s dual recognition permits to adopt a relatively flexible structure that accommodates pre-miRNA substrates, whereas the RNA recognition associated to DROSHA cleavage is provided only by protein DGCR8 [124, 125].

Changes of miRNA templates could also be associated to exoribonucleases. MicroRNAs bound to AGO proteins could be modified by nucleolytic trimming at the 3’ ends to generate isomiRs. For example, in the Drosophila miRNA biogenesis pathway, Liu et al., (2011) [126] found that mir-34 displays multiple isoforms that differ at the 3’ end. These isoforms are produced by the action of the 3’ → 5’ exoribonuclease CG9247/NIBBLER. Trimming of miRNA at the 3’ ends occurs after removal of the passenger miRNA strand from the pre-RISC and may be the final step of the RISC assembly, ultimately enhancing the target [127].

5’ and 3’ non-template modifications
The 3′ ends of mature miRNAs are highly heterogeneous, whereas the 5′ ends which correspond to the seed regions are relatively invariable. The patterns and sources of heterogeneity seem to vary depending on the miRNA species and according to the cell type. The 3′ ends often contain 1 – 3 extra nucleotides that do not match with the genomic DNA sequences. Sequence alterations of miRNAs can occur by the addition of non-template nucleotides to the miRNA termini [111, 128]. The first description of a 3′ end modification of small RNA was rendered in the hen1 mutant of Arabidopsis [129]. HEN1 is a methyl transferase that adds a methyl group to the 2′-OH at the 3′ end of the RNA [130]. The addition of nucleotides is performed by a group of nucleotidyl transferases [131]. In humans there are twelve nucleotidyl transferases, seven of them are implicated in isomiR generation (reviewed by [107]). In both humans and mice, these enzymes have uridytransferase and/or adenyltransferase activity which explain the most frequent non-template modifications such as insertions of single or multiple U (uridylation) or A (adenylation) [132]. Uridylation plays a significant role in the control of miRNA biogenesis. In mammalian embryonic stem cells, let-7 biogenesis is suppressed by the LIN28 protein which binds to the terminal loop of let-7 precursors [133, 134]. Of special interest remains the fact that LIN28 induces the 3′ uridylation of pre-let-7 by recruiting the terminal nucleotidyl transferase TUT4 [135]. The oligo U-tail added by TUT4 blocks the DICER processing mechanism and facilitates the decay of pre-let-7. In the case of the mammalian miR-122, which is adenylated by cytoplasmic poly (A) polymerase GLD-2 (or also TUTase2), 3′ end adenylation is also implicated in its stabilization [136]. Deep sequencing of the AGO-associated small RNAs has shown
that adenylated miRNAs are relatively depleted in the AGO2 and AGO3 complexes, whereby suggesting that adenylation may interfere with AGO loading.

**Single nucleotide polymorphisms (SNPs)**

Genetic variations range from large chromosomal anomalies to single nucleotide changes. SNPs are the most frequently identified variants of DNA sequences. SNPs affecting miRNAs could potentially affect the maturation process of miRNAs, the silencing machinery, the structure or the expression level of mature miRNA and the base pairing to target sites. Polymorphisms at a level of single nucleotides may also have functional roles in relation to miRNA-mediated gene regulation [137, 138]. Screening by numerous bioinformatics analyses has so far demonstrated only a very low density of SNPs in miRNA profiles [105, 139]. Polymorphisms in pre-miRNA may have an effect on miRNA maturation and thereby modulate miRNA expression. Several groups have tried to identify SNPs within or flanking pre-miRNA sequences using experimental or bioinformatics approaches. In one study, 173 human pre-miRNAs pertaining to 96 Japanese individuals were sequenced, identifying 10 SNPs in 10 pre-miRNA hairpins [140]. Yet another study, identified 12 SNPs located within 227 human pre-miRNA sequences [141]. Borel et al. (2012) [142] detected sixty-five SNPs in 49 pre-miRNAs but only in the case of 3 pre-miRNAs (hsa-miR-125a, hsa-miR-627 and hsa-miR-662) were the SNPs located within the seed region [142], thus demonstrating that SNPs within miRNA seed regions are very rare (~6%). Furthermore, a G/U polymorphism, located at the eighth nucleotide within the mature sequence of miR-125a, has been functionally characterized to block the processing of pri-miRNA into pre-miRNA.
altering the translation suppression of the mRNA target: Lin28 [141]. All of the above comprise good examples concerning the importance of the miRNA related SNP.

Recently, the role of SNPs has been analyzed with reference to miRNA processing of cancer cells. For example, the $hsa$-$miR-146a$ SNP (rs2910164) within the $pre-miR146a$ sequence consists of one of the most thoroughly studied cases [137]. This SNP reduces both the amount of $pre-miR146a$ and mature $hsa$-$miR-146a$ affecting ultimately the DROSHA/DGCR8 cropping step, with consequent multiple associations to papillary thyroid carcinoma (PTC), familial/sporadic breast cancer, ovarian cancer, prostate cancer, and hepatocellular carcinoma [143-148]. In addition, SNPs of mRNA located in the 3’ UTR region of the specific binding site of miRNAs induce regulation pattern changes of the corresponding transcript. An SNP located in the $KRAS$ 3’ UTR induces the over expression of KRAS, which has been correlated to an increased risk to acquire non-small cell lung cancer due to the lack of miRNA let-7 binding [149]. All of these studies have further contributed to the knowledge about miRNA binding site SNPs and cancer susceptibility; nonetheless, the need remains to carry out systematic studies to understand the role that SNPs play in connection to biological processes.

**Expert opinion**

The continuous progress in the quest to understand gene expression regulation mediated by small non-coding RNA, including miRNAs, is deeply revolutionizing our concepts in relation to the dynamics of functional genomics. The ubiquitous presence of these actors in the panorama of most biological processes, including pathologies, is
offering new visions, new experimental approaches and new perspectives with regard to the study of cellular and developmental biology. The multiple variants of the ‘miRNome’ are enhancing the world of fine regulation, the diversity, as well as the response to environmental changes experienced by biological systems. However, both the biogenesis and the real functionality of some of the described and annotated miRNAs deposited in the miRNA databases (miRBase) will need a further re-definition as recently has already been proposed [150]. Comprehension of the functional roles and possibilities of gene regulation by miRNAs has only just initiated.

Acknowledgements

We thank R.M. Fratini for the English revision of the manuscript. Research on microRNAs carried out in del Mazo’s laboratory was supported by the following grants: CEFIC-LRi, MEDDTL (11-MRES-PNRPE-9-CVS-072), France; and CSIC (PIE 201020E016), Spain.

References:


23. Daniels SM, Melendez-Pena CE, Scarborough RJ, Daher A, Christensen HS, El Far M, Purcell DF, Laine S, Gatignol A. Characterization of the TRBP domain


70. Heale BS, Keegan LP, O'Connell MA. ADARs have effects beyond RNA editing. Cell Cycle 2009; 8:4011-4012.


84. Gonzalez-Gonzalez E, Lopez-Casas PP, del Mazo J. The expression patterns of genes involved in the RNAi pathways are tissue-dependent and differ in the


**Figure legends**

Figure 1: MicroRNA biogenesis, degradation, storage and recycling. The figure illustrates the different pathways to generate functional miRNAs.

Figure 2: Schematic representation of the different types of possible modifications to generate variations with respect to a hypothetical canonical sequence (22nt).
Different types of miRNA variations

a) 3′ non-template modifications

b) 3′ template modifications (cleavage)

c) 5′ non-template modifications

d) 5′ template modifications (cleavage)

e) Deletions

f) Insertions

g) SNP’s

h) A to I edition