

Chapter 2

Diversity, Overlap, and Relationships in the Small RNA Landscape

Michelle S. Scott

Abstract Rapidly evolving, of high abundance and great diversity, small RNAs are increasingly found to play central cellular regulatory roles, the extent of which we are only now starting to comprehend. The evolutionary association of diverse classes of small RNAs and transposable elements is offering clues about the origin, abundance, biogenesis pathways, and target acquisition mechanisms of small RNAs. And as well as a similar relationship with transposable elements, different types of small RNAs show commonalities in their processing pathways while displaying a wide degree of diversity and variation within their biogenesis pathways and amongst their precursors, likely allowing for flexible regulation. This book chapter examines the evolutionary relationship between small RNAs and transposable elements through the role transposable elements play in the expansion of small RNA classes as well as the acquisition of novel targets. The great diversity but also overlap in both the small RNA biogenesis pathways and functional entities are also explored.

Keywords Biogenesis pathways • microRNAs • transposable elements

2.1 Introduction and Overview

During the past decade, numerous members of diverse classes of small RNAs have been associated with transposable elements, offering clues about the origin, abundance, biogenesis pathways, target acquisition mechanisms, and rapid evolution of groups of small RNAs. Large numbers of small RNAs, and in particular miRNAs,

M.S. Scott (✉)

Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee, UK

Current address: Department of Biochemistry, University of Sherbrooke, Sherbrooke, Québec, Canada
e-mail: michelle.scott@usherbrooke.ca

have been recently found to originate from noncanonical precursors, display variation in their biogenesis pathways, or exert novel regulatory functionality, making clear-cut classification of these molecules increasingly difficult.

The great diversity of small RNA precursors, biogenesis pathways, and targeting mechanisms allows for combinatorial complexity and high flexibility in the regulation of multiple aspects of cellular function. Increased understanding of the evolutionary and regulatory relationships between these different noncoding RNAs will be central to unlocking the multiple layers of regulation underlying the cell's complexity. This understanding will also be important to determine how we can harness this knowledge to treat the numerous diseases likely to result from defects in the regulation of these pathways. For example, numerous miRNAs display deregulated expression and a functional involvement in cancers (reviewed in Garzon et al. 2006).

The first half of this chapter explores the relationships between transposable elements and small RNAs, examining in particular how they have served in the expansion of small RNA classes as well as the acquisition of novel targets, leading to both overlap and diversity. The rapid expansion of these elements has led to an RNA landscape displaying overlapping but also varied and diverse biogenesis pathways and functional entities which are explored, from a microRNA perspective, in the second half.

2.2 Evolutionary Relationship Between Small Noncoding RNAs and Transposable Elements

Members of diverse classes of small RNAs have strong ties with transposable elements (TEs), generating both small RNAs employed by the cell to suppress TE expression and transposition but also small RNAs which have acquired new functionality and serve other and diverse cellular roles, as described in Sect. 2.2.2. Also known as repeat elements and “jumping genes,” TEs have provided some small noncoding RNAs with a mechanism for expansion and acquisition of novel targets and functions, as explored in Sect. 2.2.3.

2.2.1 *Transposable Elements*

TEs are highly abundant genetic sequences which have the capacity of both moving and proliferating within and between genomes. It is estimated that between 30% and 50% of the sequence in mammalian genomes (45% in human), and even higher proportions in some plants, is derived from TEs, although most are currently inactive (reviewed in Cordaux and Batzer 2009; Mourier and Willerslev 2009; Tenaillon et al. 2010). As illustrated in Fig. 2.1, two main types of TEs have been

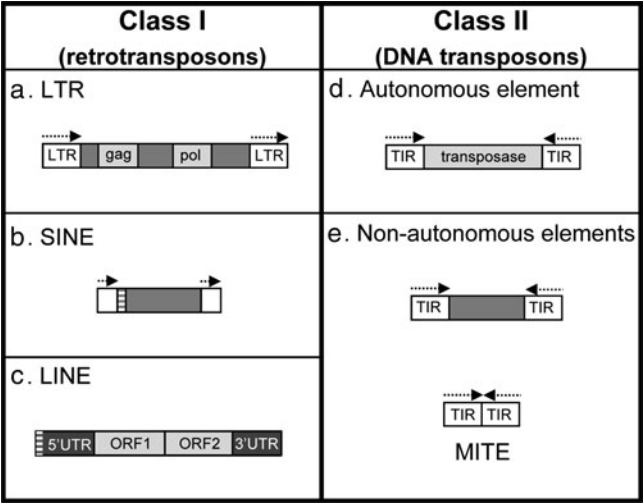


Fig. 2.1 *Classes and structures of TEs.* TEs can be classified into two groups based on their requirement of reverse transcription for transposition. LTRs (a) are flanked by direct repeats at their ends and encode proteins (gag and pol) which closely resemble retroviral proteins. SINEs (b) are flanked by direct repeats and encode an RNA polymerase III promoter, whereas LINEs (c) contain two open reading frames necessary for transposition (ORF1 and ORF2) which are flanked by untranslated regions (UTRs). Unlike LINEs, SINEs do not encode proteins and are believed to use the LINE retrotranscription machinery for reverse transcription. Class II elements do not utilize an RNA step and instead employ a transposase, which recognizes terminal inverted repeats (TIRs) for excision from the donor site and integration into an acceptor site. Autonomous DNA transposons (d) encode their own transposase, whereas nonautonomous DNA transposons (e) encode either a mutated version of the autonomous transposase gene, an unrelated portion of the host genome, or even a deleted version which consists simply of TIRs in a tail-to-tail orientation (MITE). Arrows represent direct or inverted repeats. Small striped boxes represent RNA polymerase promoters. Light gray boxes represent genes important for the transposition. Classes and characteristics of TEs are reviewed in Cordaux and Batzer (2009), Deininger and Batzer (2002), Richard et al. (2008), Slotkin and Martienssen (2007)

described, the retrotransposons and the DNA transposons, also referred to as class I and class II transposons respectively (reviewed in Cordaux and Batzer 2009; Slotkin and Martienssen 2007). DNA transposons can move to new genomic locations, either autonomously or nonautonomously, by excising themselves from their current location as a DNA molecule and inserting themselves elsewhere. In contrast, retrotransposons copy themselves using an RNA intermediate which is reversed transcribed and inserted back into the genome in a different location.

Three main subclasses of retrotransposons have been described (reviewed in Deininger and Batzer 2002; Richard et al. 2008; Cordaux and Batzer 2009):

- Long terminal repeats (LTRs)
- Long interspersed nuclear elements (LINEs)
- Nonautonomous retrotransposons including the abundant short interspersed nuclear elements (SINEs).

LTRs are very abundant in plant genomes. However, they have low activity in organisms like humans. In contrast, members of the LINEs and SINEs are believed to be currently active in humans (Lander et al. 2001; Mills et al. 2007). TEs range in length from less than one hundred to a few thousand nucleotides, and some TEs have been identified in very large copy numbers (e.g., over 1,000,000 copies of Alu elements, TEs of the SINE subclass of ~300 nucleotides in length, have been found in the human genome, representing approximately 10% of the genome) (Deininger and Batzer 2002; Richard et al. 2008; Cordaux and Batzer 2009). The very large copy number of different TEs in diverse organisms testifies to the great impact they had in shaping their host genomes.

While TEs can lead to genomic instability if inserted for example in functional genomic sequences such as protein coding or regulatory regions, they have also been found to cause the emergence of new regulatory features and genes, likely playing an important role in evolution and defining organism-specific characteristics (reviewed in Cordaux and Batzer (2009)). In addition to their role in the modification and duplication of protein-coding genes, TEs are emerging as drivers in the creation of novel noncoding genes in numerous organisms.

2.2.2 Association Between Small Noncoding RNAs and Transposable Elements

Various classes of small RNAs have been described as displaying an association with one or several types of TEs, as summarized in Table 2.1. Some small RNAs, such as germ line piRNAs (PIWI-interacting RNAs) and somatic endo-siRNAs (endogenous small interfering RNAs), derive from TEs and functionally interact with them, serving the purpose of suppressing TE expression and duplication (Saito and Siomi 2010; Ghildiyal et al. 2008; Siomi et al. 2008). In contrast, other small RNAs probably either evolved from TEs that subsequently acquired new functionality (a subset of miRNAs for example) or used TE transposition mechanisms for duplication from existing parental small RNA molecules such as has been described for some small nucleolar RNA (snoRNA) copies.

2.2.2.1 piRNAs and TEs

Present both in vertebrates and invertebrates, piRNAs were identified as small RNA interactors of PIWI proteins, a family originally characterized through genetic studies as playing a role in the maintenance of germ line integrity (Cox et al. 1998; Brennecke et al. 2007). piRNAs show a strong bias for uridine residues at their 5' end but no clear secondary structure features in flanking regions in the genome (O'Donnell and Boeke 2007). Most piRNAs map to a small number of position-conserved TE-rich clusters which express up to several thousand piRNAs

Table 2.1 Relationships between TEs and small RNAs

Small RNA class	Position in genome	Associated TEs	Relationship with TEs	References
piRNAs	TEs, repeats and piRNA clusters	Class I and class II	Transcribed from TEs, piRNAs serve in TE silencing	Kim et al. (2009), Saito and Siomi (2010), Siomi et al. (2008)
Endogenous siRNAs	TEs, repeats and endo-siRNA clusters	Class I and class II	Transcribed from TEs, endo-siRNAs serve in TE silencing	Ghildiyal et al. (2008), Kim et al. (2009), Saito and Siomi (2010), Siomi et al. (2008)
miRNAs	In introns of protein-coding and non-protein coding host genes. Others are encoded in intergenic transcription units	Class I and class II	TE expression and secondary structure appropriate for evolution into miRNA and generation of targets	Baskerville and Bartel (2005), Borchert et al. (2006), Kim et al. (2009), Lee et al. (2004), Rodriguez et al. (2004)
snoRNAs	Intergenic units or in introns of host genes. Can be clustered (frequent in plants) or individual (frequent in animals)	Class I (non-LTR retro-transposons)	Transposition machinery used for duplication	Dieci et al. (2009), Filipowicz and Pogacic (2002), Luo and Li (2007), Weber (2006)

displaying low sequence conservation (Malone and Hannon 2009; Brennecke et al. 2007). In both vertebrates and invertebrates, piRNAs have been detected almost uniquely in germ line cells, where they are believed to function in the silencing of TEs (Aravin et al. 2006; Brennecke et al. 2008; Das et al. 2008; Houwing et al. 2007). piRNAs are processed from single-stranded precursors derived from both sense and antisense TE transcripts (Brennecke et al. 2007; Saito et al. 2006). The biogenesis of piRNAs in flies and mammals has been proposed to involve primary and secondary processing in a mechanism referred to as the ping-pong cycle (illustrated in Fig. 2.2). This cycle involves primary piRNAs binding to their targets and recruiting PIWI family proteins, leading to target cleavage and TE transcript destruction and resulting in the production of secondary piRNAs, which can perpetuate the cycle (Brennecke et al. 2007; Gunawardane et al. 2007). In addition, piRNAs have been found to regulate the DNA methylation of TEs, thus also exerting epigenetic control over these elements (Brennecke et al. 2008; Kuramochi-Miyagawa et al. 2008).

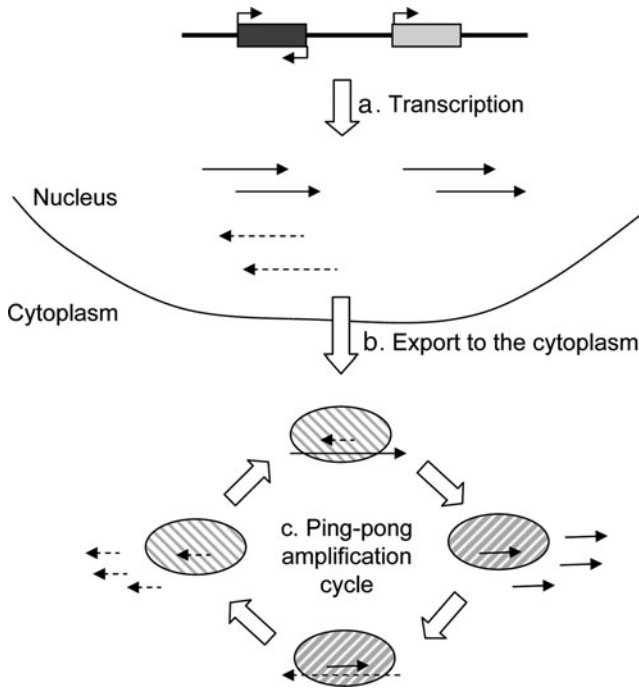


Fig. 2.2 *piRNA biogenesis and ping-pong amplification cycle.* (a) piRNA clusters (dark gray boxes) and functional TEs (light gray boxes) are transcribed in the nucleus. (b) Sense transcripts (solid arrows) and antisense transcripts (broken arrows) are exported to the cytoplasm. (c) Primary piRNAs complexed with PIWI family proteins (represented by lined ovals) bind to complementary TE and piRNA cluster transcripts leading to cleavage and amplification. These resulting short RNAs are bound by different PIWI family proteins and base pair to complementary antisense piRNA cluster transcripts resulting in antisense piRNAs, thus completing the proposed cycle. Block arrows represent progression to subsequent steps in the pathway. piRNA biogenesis is reviewed in Khurana and Theurkauf (2010), Kim et al. (2009)

2.2.2.2 Endogenous siRNAs and TEs

Originally observed during virus- and transgene-induced silencing in plants, canonical small interfering RNAs (siRNAs) originate from double-stranded RNAs (Carthew and Sontheimer 2009). More recently, endogenous siRNAs, including many derived from TEs, have been uncovered in a wide range of organisms including animals, plants, and fungi. Endogenous TE-derived siRNA precursors are believed to originate from diverse TEs including read-through transcription of DNA transposons as well as bidirectional transcription of LINE1 5' UTR and are found located in both intronic and intergenic regions (Slotkin and Martienssen 2007; Sunkar et al. 2005). As for piRNAs, endogenous siRNAs are believed to function in TE silencing, but they have been predominantly identified in somatic tissues (Ghildiyal et al. 2008; Kawamura et al. 2008; Chung et al. 2008). However,

some loci-producing piRNAs in flies have also been found to generate endogenous siRNAs (Kawamura et al. 2008). A cross talk between the piRNA and endogenous siRNA pathways has been described in worm (Das et al. 2008). The relationships between these two types of small RNAs and TEs are further described in Chap. 5.

2.2.2.3 MicroRNAs and TEs

Widely expressed in animals and plants, microRNAs (miRNAs) are ~22 nucleotide-long single-stranded RNAs that are processed out of hairpin precursors of variable length through an extensively characterized biogenesis pathway (see Sect. 2.3.1). Encoded in introns of protein-coding genes or independent transcription units, miRNAs are involved in gene silencing through the regulation of the stability and translation of target messenger RNAs (mRNAs), usually by base pairing with their 3' UTR (untranslated region) or their coding region (Bartel 2009; Lai 2005).

Numerous studies have reported TE-derived miRNAs in several different organisms. As early as 2002, miRNA-like molecules were described encoded in TEs in *Arabidopsis* (Llave et al. 2002). Following that report, through sequence analyses and computational searches, hundreds of previously identified as well as novel mammalian miRNAs were described as derived from TEs (Borchert et al. 2006; Piriyaopongsa et al. 2007; Smalheiser and Torvik 2005; Yuan et al. 2011). In mammals, a large proportion of miRNAs is found clustered in the genome, in regions highly enriched in TEs (Yuan et al. 2011). It has been proposed that the close proximity of TEs of similar sequence inserted in reverse orientations would result in structures resembling miRNA precursors which if expressed, might be recognized as substrates by the miRNA biogenesis pathway (Mourier and Willerslev 2009). Such a clustered TE region has been described on human chromosome 19 (referred to as C19MC) which encodes interspersed miRNAs and Alu elements. It was found that many of the Alu elements contain intact RNA polymerase III promoters which could ensure the expression of the miRNAs (Borchert et al. 2006). This suggests that TEs might represent not only a template from which small noncoding RNAs can evolve but also a mechanism to ensure their expression (Fig. 2.3).

The analysis of TE-derived miRNAs has revealed that miRNAs have evolved from all types of TEs described in the previous section. In human, while most frequently found associated with the L2 (from the LINE subclass) and MIR (from



Fig. 2.3 Three kilobase portion of the human C19MC cluster (Borchert et al. 2006; Kent et al. 2002). Light and dark gray boxes respectively represent miRNAs and Alu elements. The direction of transcription is indicated with arrows

the SINE subclass) elements, miRNAs have also been described as derived from LTRs and DNA transposons such as DNA mariners as well as other SINE and LINE elements including the B2, Alu, and L1 elements (Mourier and Willerslev 2009; Piriyaopongsa et al. 2007; Smalheiser and Torvik 2005). The proportion of miRNAs derived from different TE types varies depending on the organism (Yuan et al. 2011). A family of human miRNAs has also been described as originating from MITEs (see Fig. 2.1), leading to the formation of stable hairpins resembling miRNA precursors (Piriyaopongsa and Jordan 2007; Slotkin and Martienssen 2007). TE-derived miRNAs have recently been found to be significantly less conserved within mammals than miRNAs not derived from TEs, likely due to relatively recent acquisition (Yuan et al. 2011). The evolution of miRNAs from TEs is depicted in Fig. 2.5.

2.2.2.4 snoRNAs and TEs

snoRNAs are an ancient family of highly conserved and abundant small noncoding RNAs that predominantly function as guides for the chemical modification of ribosomal RNA (rRNA) (reviewed in Matera et al. 2007). Like miRNAs, snoRNAs are either encoded in introns of protein-coding genes or in independent transcription units. Two main types of snoRNAs have been described, the box C/D snoRNAs and the box H/ACA snoRNAs, which differ in terms of the chemical modification they catalyze. While several hundred snoRNAs, most of them highly conserved, have been described in mammalian organisms, computational searches have identified hundreds and even thousands of additional snoRNAs displaying TE characteristics (Weber 2006; Luo and Li 2007; Schmitz et al. 2008), described in Fig. 2.4. These TE-derived snoRNAs, referred to as snoRTs (snoRNA retroposons), result from the retroposition of existing (parental) snoRNA transcripts which employed LINE machinery to duplicate and transpose themselves to new genomic locations (Weber 2006).

While many computationally identified snoRTs have not been found previously and have not been experimentally validated, others have previously been described as functional snoRNAs (Luo and Li 2007).

2.2.3 A Driving Force of Evolution

TEs have been described as the most nonconserved regions and represent the most lineage-specific elements in eukaryotic genomes (Lander et al. 2001). Their recent contribution to animal and plant genomes and their high abundance suggest that transposition is a common occurrence that is likely a strong driving force in evolution, providing a mechanism for the emergence of organism-specific regulatory elements. Recent studies have revealed that transposition likely leads to both the creation of small RNAs with new functionality as well as new targets for small RNAs as depicted in Fig. 2.5.

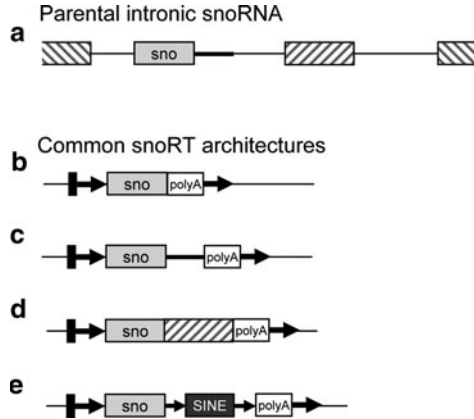


Fig. 2.4 *Common snoRT genomic architectures.* In all panels, snoRNAs and exons are represented respectively by *light gray boxes* and *diagonally lined boxes*, while the intronic region immediately flanking the parental snoRNA is depicted by a *thick black line*. snoRTs often display LINE characteristics including upstream L1 consensus recognition site (shown with a *thin black box*), poly A tails at their 3' end, and flanking direct repeats referred to as target-site duplications (TSDs, depicted by *black arrows*). (a) In mammals, parental snoRNAs are typically encoded in intronic regions of host genes. (b) The simplest form of snoRTs consists of a snoRNA followed by a poly A tail at its 3' end and flanked by TSDs. (c) Part of the intronic region downstream of the parental snoRNA can also be retroposed (represented by the *thick black line*). The intronic region flanking the 5' end of the parental snoRNA can also be retroposed (not shown here). (d) Exonic regions from the parental snoRNA host gene have been identified in some snoRTs. (e) Repeat elements of the SINE class, flanked by TSDs, are found downstream of the snoRNA sequence in some snoRTs. Examples of snoRTs are described in Luo and Li (2007), Weber (2006)

2.2.3.1 Creation of New Small RNAs

The integration of TEs in new genomic locations can lead to the creation of new small RNAs provided the TE is expressed and its transcript displays an appropriate structure that is recognized as a substrate in a small RNA biogenesis pathway. As described in the previous subsection, diverse small RNAs have been found to be TE-derived, generally following insertion into intergenic or intronic regions (illustrated in Fig. 2.5a, b). Some such TE-derived miRNAs display a high level of conservation, but many are lineage specific (Piriyapongsa et al. 2007; Smalheiser and Torvik 2005). Studies of the pattern of occurrence of TE-derived miRNA members of specific families throughout multiple organisms have led to the hypothesis that these families follow the birth-and-death model of evolution. As opposed to concerted evolution in which members of a family evolve in a similar, concerted way, in the birth-and-death model, new members of the family created by duplication either remain in the genome over long periods or are deleted or inactivated (Nei and Rooney 2005). Several miRNA families display gains and losses of members when closely related organisms are considered, suggesting a birth-and-death evolutionary process, functional diversification of the family, and a role in evolution and lineage-specific traits (Zhang et al. 2008; Yuan et al. 2010).

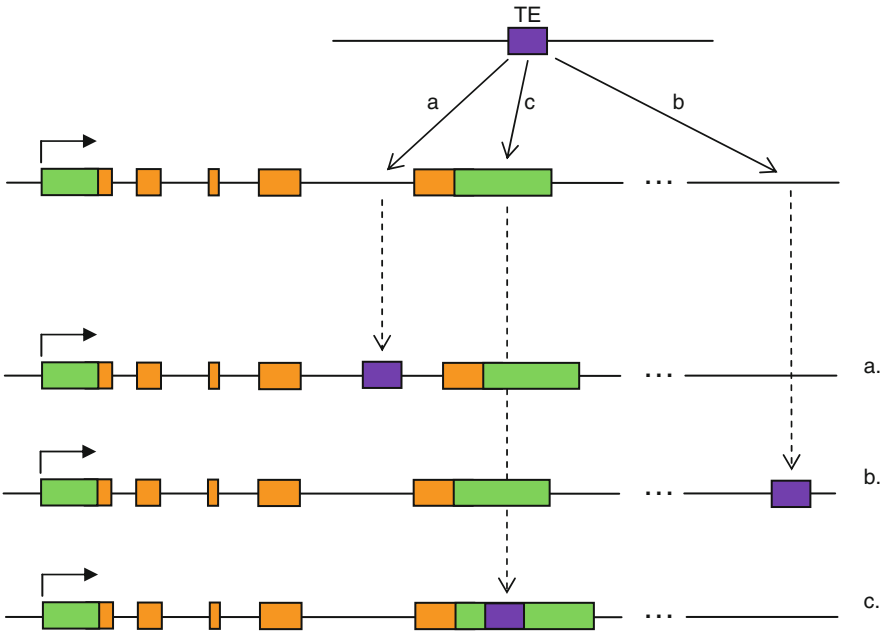


Fig. 2.5 *Transposition events generating new small RNAs or new small RNA targets.* The insertion of TEs can result in the emergence of novel features including intronic TE-derived small RNAs (a), intergenic TE-derived small RNAs (b), and TE-derived targets of small RNAs if inserted in reverse orientation into 3' untranslated regions (UTRs) (c), or coding regions (not shown), for example. TEs are depicted as purple boxes, UTRs as green boxes, and coding exons as orange boxes. Straight solid arrows represent transposition events and broken arrows show the genomic region depicted after the transposition

In addition to providing templates for the creation of new small RNAs, TEs have also been suggested to have allowed the expansion of specific small RNA families. The C19MC region of human chromosome 19 described in Sect. 2.2.2.3 and Fig. 2.3 consists of a cluster of alternating Alu elements and miRNAs of a primate-specific family. Alu elements from this cluster are proposed to have not only led to the creation of novel miRNAs through the reverse orientation of adjacent elements but are also believed to have facilitated the amplification of the region through a recombination event, leading to the generation of many additional copies of these miRNAs and the rapid expansion of this family. This large group of primate-specific miRNAs was found to be evolving rapidly, resulting in some nonfunctional miRNAs (pseudo-miRNAs) but also novel lineage-specific miRNAs (Zhang et al. 2008). Other families of miRNAs are also believed to have been amplified by Alu-mediated recombination events including placental-specific miRNAs derived from the MER53 DNA transposon (Yuan et al. 2010).

The duplication of functional small RNAs likely results in molecules that are under less evolutionary pressure to avoid mutation than a single-copy small RNA. In the case of snoRNAs, the numerous snoRTs originating from snoRNAs have

been proposed to serve two roles: safeguarding against mutation in parental copies and possibly also allowing for the rapid evolution of snoRNAs with novel targets (Weber 2006). In support of this hypothesis, numerous “orphan” snoRNAs and snoRTs have been identified, displaying typical snoRNA characteristics and features but with no known targets, possibly due to sequence diversification. Some such orphan snoRNAs are ubiquitously expressed in mammals suggesting they might have alternate functions and/or noncanonical targets (Bachellerie et al. 2002; Luo and Li 2007). And indeed, one large mammalian family of snoRNAs, the HBII-52 in human, displays a conserved region of complementarity to several transcripts including the serotonin 5-HT_{2C} receptor and has been found to regulate their alternative splicing through base pairing (Bachellerie et al. 2002; Kishore and Stamm 2006).

2.2.3.2 Creation of New Targets

Small RNAs generally exert their function by base pairing with their targets and bringing them in close proximity to effector proteins. TEs have been found to play an important role in generating not only novel small regulator RNAs but also novel targets for these molecules (Fig. 2.5c).

The large abundance of TEs and transposition events in genomes can likely rapidly generate regulatory networks if TEs with high sequence similarity to a TE-derived small RNA are inserted into protein-coding genes in reverse orientation, readily creating target sites for the small RNA. In human, many miRNAs display complementarity to conserved Alu elements in 3' UTRs of mRNAs (Smalheiser and Torvik 2006). Amongst all 3' UTR targets of human TE-derived miRNAs, approximately 10% were estimated to be TE-derived (Piriyapongsa et al. 2007). As discussed in the study, this is likely to be an underestimate because miRNA target-site prediction methods consider conservation in their prediction, and TEs are among the most lineage-specific elements of genomes. However, some TE-derived miRNAs were found to have up to 80% of their targets derived from TEs (Piriyapongsa et al. 2007). Other types of TEs have also been found to generate both miRNAs and their targets including LINE elements for which an example of conserved mammalian miRNA and its human-specific target mRNAs were described (Smalheiser and Torvik 2005).

2.2.3.3 A Prevalent Evolutionary Mechanism

Although often described as harmful and selfish elements in genomes, TEs have also been found to be beneficial to genomes, and mutualistic relationships between genomes and their colonizing TEs have been described (Faulkner and Carninci 2009; Cordaux and Batzer 2009; Malone and Hannon 2009). Genomes have evolved mechanisms to control and limit TE expansion, in great part through the use of small RNAs such as members of piRNAs and siRNAs (Malone and Hannon

2009) as described in Sect. 2.2.2. However, a diverse and growing body of evidence is now suggesting that transposition and TE activity are also a prevalent evolutionary mechanism in the de novo creation of small RNAs as well as in their expansion and generation of targets, providing additional regulatory layers in the control of cellular networks. As such, TE-derived miRNAs and snoRNAs represent a beneficial consequence of TE activity and expansion. As our understanding of the contribution of TEs to genome function and evolution increases, so too will our understanding of their contribution to small RNA evolution and organism-specific regulation.

2.3 Diversity and Overlap in the Small RNA Landscape, a miRNA Perspective

The best characterized small RNA biogenesis pathway is the miRNA biogenesis pathway which has been extensively investigated. In recent years however, numerous unrelated reports have identified miRNAs that deviate from the canonical pathway in terms of their precursors, biogenesis pathway, characteristics of the mature molecule, or in their functionality. Together with the contribution of TEs to miRNA and target diversity, variations from the canonical biogenesis pathway and diverse characteristics of the mature molecules add flexibility to miRNA regulatory networks.

2.3.1 Variation and Overlap in miRNA Biogenesis Pathways

In animals, most miRNAs are encoded in introns or within independent transcription units and are transcribed by the RNA polymerase II (RNA pol II) (Baskerville and Bartel 2005; Lee et al. 2004; Rodriguez et al. 2004). The resulting primary miRNA transcripts are then cleaved by the microprocessor complex which contains the nuclear RNase type III enzyme Drosha and the double-stranded RNA binding domain protein DGCR8 (Kim and Kim 2007; Lee et al. 2003). Processing of the primary miRNA transcript by the microprocessor complex generates the miRNA precursor hairpin, which is exported to the cytoplasm by exportin-5 where it is further processed by the RNase type III enzyme Dicer (Bernstein et al. 2001; Hutvagner et al. 2001; Ketting et al. 2001). Dicer releases the miRNA duplex, one strand of which is loaded onto specific argonaute (AGO) proteins forming the RNA-induced silencing complex (RISC) (Hutvagner and Zamore 2002; Mourelatos et al. 2002). The canonical miRNA biogenesis is reviewed in Bartel (2004) and Kim et al. (2009) and depicted in Fig. 2.6a. Although the majority of miRNAs in mammals have been found to be both DGCR8 and Dicer dependent (Babiarz et al. 2008), numerous examples of deviation from this canonical pathway have

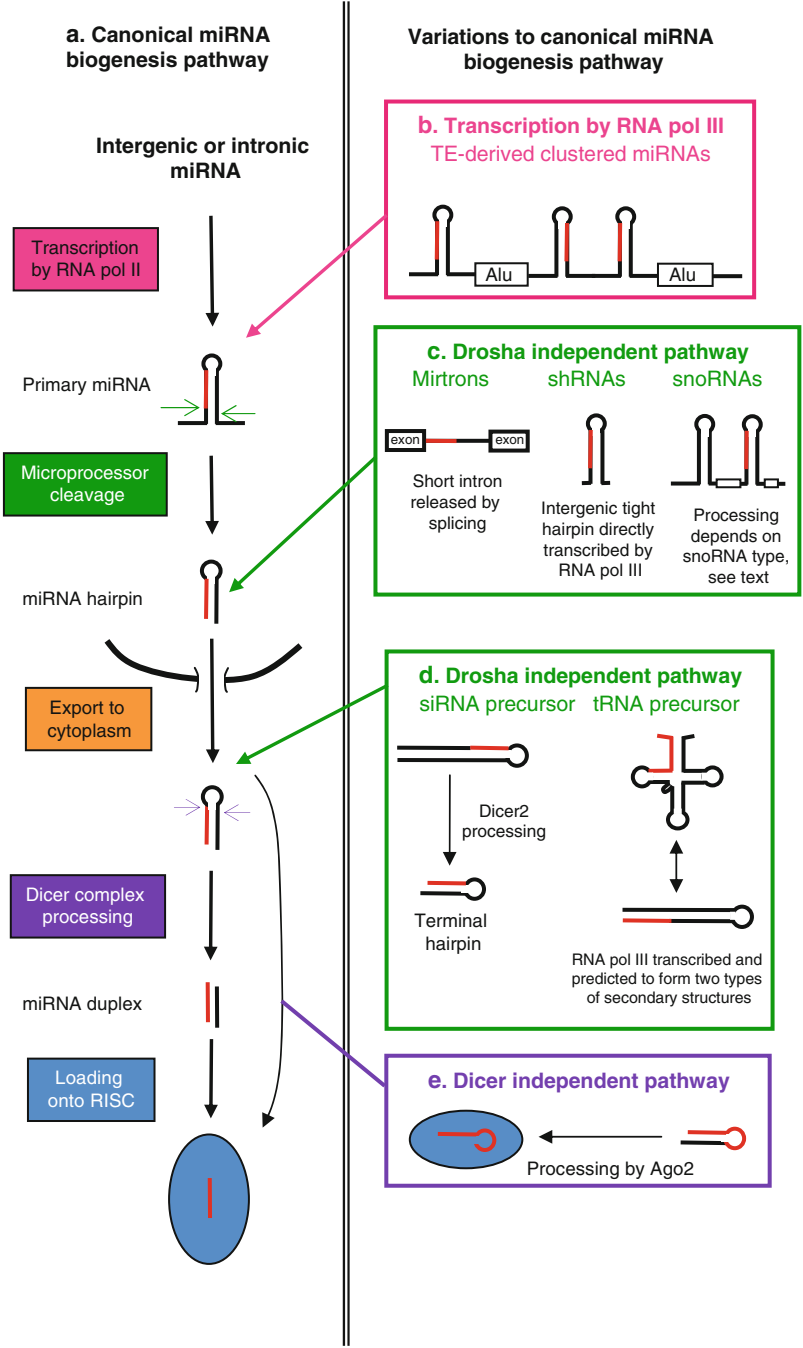


Fig. 2.6 Diversity and variation in the miRNA biogenesis pathway. The canonical miRNA biogenesis pathway (a) has been extensively investigated. Variations from the canonical pathway have been identified at most steps including at the level of transcription (b), Drosha processing (c and d), and Dicer processing (e)

been described, and numerous miRNAs have been shown to originate from diverse noncanonical precursors (illustrated in Fig. 2.6b–e).

2.3.1.1 Transcription of miRNA Genes by the RNA pol III

As described above, most animal miRNAs, both intronic and intergenic, are transcribed by the RNA pol II. However, the human chromosome 19 cluster (C19MC) of interspersed miRNAs and Alu elements described in Sect. 2.2.2.3 contains RNA pol III promoters, and miRNAs encoded within this region were found to be transcribed by the RNA pol III (Borchert et al. 2006), see Fig. 2.6b. Other TE-derived human miRNAs in similar genomic contexts are also likely to be transcribed by the RNA pol III (Borchert et al. 2006). In addition, viral miRNAs adjacent to a transfer RNA were also recently found to be transcribed by the RNA pol III as described further below (Bogerd et al. 2010). And shRNA-derived miRNAs were also found to be transcribed by the RNA pol III as described below (Babiarz et al. 2008). Thus, although a large majority of animal miRNAs are transcribed by the RNA pol II, other pathways are used for the generation of primary miRNA transcripts and are likely under the control of alternate regulatory networks.

2.3.1.2 Microprocessor-Independent miRNA Biogenesis

Subsets of noncanonical miRNA precursors have been found to bypass the microprocessor cleavage step. Included among them are mirtrons, small hairpin RNAs (shRNAs), and long hairpin siRNA precursors.

Mirtrons

Canonical intronic miRNAs are processed by Drosha prior to the host transcript intron splicing (Kim and Kim 2007). However, a group of short introns of size <150 nucleotides which fold into hairpins has been found to serve as precursors for miRNAs. Referred to as mirtrons, these precursors result from splicing and debranching of the intron lariat by the spliceosome, thus bypassing the Drosha cleavage step (reviewed in Winter et al. 2009). The resulting hairpin precursor is then exported to the cytoplasm and further processed following the canonical miRNA biogenesis pathway (Fig. 2.6c). Although not found in large numbers, mirtrons have been identified throughout the animal kingdom (Babiarz et al. 2008; Carthew and Sontheimer 2009; Winter et al. 2009).

Endogenous shRNA Precursors

A second type of miRNA precursors that was found to exhibit independence from Drosha is endogenous shRNAs, which form short tight stem loop structures (Fig. 2.6c). A subset of mammalian miRNAs was annotated as derived from shRNAs for three main reasons: they were found to be Dicer dependent, they are not encoded in introns and do not display splicing signals, and they show characteristic read position patterns within the full-length molecule. Such shRNA-derived miRNA precursors seem to be generated directly as short hairpins by RNA pol III transcription (Babiarz et al. 2008). In mammals, shRNA-derived miRNAs make up a much larger number of reads than mirtron-derived miRNAs (Babiarz et al. 2008).

Long Hairpin siRNA Precursors

In *Drosophila*, miRNAs are typically processed by Dicer1, while endogenous siRNAs are processed by Dicer2 (reviewed in Miyoshi et al. 2010). However, it is believed that the terminal hairpins resulting from processing by Dicer2 of a subset of endogenous siRNAs can serve as miRNA precursors (Fig. 2.6d) which are recognized as substrates by Dicer1, producing a mature miRNA. Thus these siRNA precursors would generate both siRNAs and miRNAs as a result of sequential Dicer2 and Dicer1 processing (reviewed in Miyoshi et al. 2010).

tRNase Z-Derived Precursors

Several animal viruses encode viral miRNAs which display miRNA biogenesis features characteristic of canonical cellular miRNAs including transcription by the RNA polymerase II and processing by Drosha and Dicer. However, the murine gamma-herpesvirus 68 (MHV68) has recently been shown to encode miRNAs which employ noncanonical biogenesis pathways, including transcription by the RNA polymerase III of the miRNA precursors linked to an adjacent transfer RNA (tRNA). The resulting primary transcripts are cleaved by the tRNase Z, releasing the tRNA from the miRNA precursors and bypassing the Drosha cleavage step. The miRNA precursors are then further processed by Dicer to generate the mature miRNA (Bogerd et al. 2010).

2.3.1.3 Dicer-Independent miRNA Biogenesis

A small group of miRNAs processed in a Dicer-independent manner has been recently described in both mouse and zebra fish (reviewed in Suzuki and Miyazono 2011). Drosha has been shown to process the primary transcript of the well-conserved miR-451 miRNA resulting in a short hairpin precursor with a

17 nucleotide stem, which is shorter than the length required by Dicer for efficient processing. In the absence of Drosha, the mature form of the miRNA was highly reduced while its levels were not affected by the absence of Dicer (Cheloufi et al. 2010). As illustrated in Fig. 2.6e, following cleavage by Drosha, this noncanonical miRNA precursor is believed to be loaded directly onto the RISC complex where it is sliced by Ago2, thus bypassing Dicer processing (Suzuki and Miyazono 2011; Cheloufi et al. 2010). This suggests Dicer processing and RISC loading might not always be coupled (Babiarz et al. 2008; Miyoshi et al. 2010).

Thus over the past 3 years, diverse examples of alternate miRNA biogenesis pathways, typically with partial overlap with the canonical pathway, have been described, suggesting both that deviations from the canonical pathway are relatively common and that more such deviations will be found over the next few years. These examples demonstrate that diverse variations of the miRNA processing pathway exist and that no component of the canonical miRNA biogenesis pathway is essential for the biogenesis of all miRNAs. It should be noted however that most miRNAs are transcribed by the RNA pol II and require both the microprocessor and Dicer complexes for proper biogenesis (Babiarz et al. 2008).

2.3.1.4 Additional Noncanonical miRNA Precursors

The examples described above of variations to the canonical miRNA biogenesis pathway generally relate to miRNA precursors whose main function is the generation of miRNAs. However, subsets of two types of abundant cellular RNAs, small nucleolar RNAs (snoRNAs), and transfer (tRNAs), which play seemingly unrelated primary functions in the cell, have also been found to act as noncanonical miRNA precursors. Some long noncoding RNAs have also been found to serve as precursors to miRNAs.

Small Nucleolar RNAs (snoRNAs)

Several independent studies and lines of evidence suggest that snoRNAs can serve as precursors for miRNAs. Small RNAs derived from snoRNAs were identified in a deep sequencing dataset of small human RNAs associated with argonaute proteins (Ender et al. 2008). In particular, small RNAs of miRNA size were found derived from the box H/ACA snoRNA ACA45 and shown to display functional miRNA characteristics including gene silencing capabilities and endogenous targets. The processing of ACA45 was shown to depend on Dicer but not Drosha (Ender et al. 2008). Box C/D snoRNA-derived small RNAs were also identified and found to display miRNA characteristics including incorporation into RISC complexes and gene silencing capabilities in human and *Giardia lamblia* (Saraiya and Wang 2008; Brameier et al. 2011).

The processing of snoRNAs into small RNAs of size generally less than 30 nucleotides (referred to as sdRNAs for snoRNA-derived small RNAs) has recently

been shown to be widespread and conserved from most snoRNA loci in animal, *Arabidopsis*, and yeast genomes (Taft et al. 2009). The processing pathways responsible for the generation of sdRNAs were investigated, revealing that a subset of sdRNAs (those derived from box C/D snoRNAs) were only mildly downregulated in the absence of either DGCR8 or Dicer1, while some box H/ACA sdRNAs showed a pronounced response, displaying a downregulation in the absence of Dicer1 but an upregulation in the absence of DGCR8 (Taft et al. 2009). Thus, multiple processing pathways might be used for the generation of miRNAs and other small RNAs from snoRNAs (Fig. 2.6c).

While numerous examples of snoRNA-derived small RNAs and snoRNA-derived miRNAs have now been described, several reports have also identified known miRNA precursors with snoRNA-like features. Numerous reported, and in several cases extensively validated, miRNA precursors have been identified displaying sequence, structure, and functional snoRNA characteristics, suggesting a possible evolutionary relationship between subsets of miRNAs and snoRNAs (Ono et al. 2011; Scott et al. 2009). In addition, evidence of cross talk between the miRNA and snoRNA pathways has also been observed including core snoRNA-binding proteins in argonaute complexes (Hock et al. 2007).

Transfer RNAs

Several independent studies have recently reported small RNAs derived from transfer RNAs (tRNAs) and displaying miRNA processing characteristics (reviewed in (Pederson 2010; Suzuki and Miyazono 2011)). In mouse, miRNAs were identified originating from a tRNA-Ile gene which encodes a primary transcript with the capacity to form not only a mature tRNA cloverleaf secondary structure but also, alternatively, a long hairpin, as illustrated in Fig. 2.6d. The miRNAs originating from this precursor were found to be Drosha independent and Dicer dependent (Babiarz et al. 2008). In human, tRNA-derived small RNAs appear to be generated with a clear preference for 5' ends of the full-length molecule, indicating directed processing and accumulation as opposed to nonspecific degradation (Cole et al. 2009). Several studies found evidence of miRNA processing and binding characteristics for the tRNA-derived small RNAs such as processing by Dicer and binding to argonaute proteins (reviewed and discussed in Pederson 2010).

Long Noncoding RNAs

Long noncoding RNAs are generally defined as transcripts of size greater than 200 nucleotides that do not encode proteins. The imprinted and maternally expressed H19 long noncoding RNA was recently found to encode a previously reported miRNA in both mouse and human, providing another example of a noncanonical miRNA precursor. The primary transcript was found to be processed by Drosha

(Cai and Cullen 2007), but further investigation will be required to fully characterize its processing. In addition to this example, a computational analysis of tens of thousands of long messenger-like noncoding RNAs in mouse predicted they encode dozens of likely miRNA candidates, including 20 previously reported miRNAs (He et al. 2008). miRNA-encoding long noncoding RNAs were also described in *Arabidopsis* (Hirsch et al. 2006).

Thus, several types of noncanonical miRNA biogenesis pathways and precursor types exist and can lead to functional mature miRNAs. These diverse sources of miRNAs likely offer the possibility of variety and flexibility in their regulation.

2.3.2 Diversity of Targets and Targeting Mechanisms

Processing of miRNA precursor hairpins by Dicer generates short miRNA duplexes of approximately 22 base pairs. One strand preferentially associates with an argonaute protein, forming the core of the RISC complex (illustrated in Figs. 2.6a and 2.7a). Other small RNAs such as endogenous siRNAs are also loaded onto argonaute proteins (reviewed in Kim et al. 2009). The argonaute family member to which small RNAs are complexed depends on several factors including the cytoplasmic Dicer that processed the small RNA, as well as the extent of complementarity of the double-stranded precursor. The argonaute interactor plays an important role in determining the functionality of the complex. Argonaute loading and sorting is reviewed in Czech and Hannon (2011).

The RISC complex carries out gene silencing by base pairing to its messenger RNA (mRNA) targets. High complementarity of a miRNA to its target, which is typically seen in plants, results in mRNA cleavage, while lower complementarity, more characteristic of animals, promotes deadenylation of the mRNA followed by either maintenance in a translation-repressed state or decapping and 5'–3' decay (reviewed in Bartel 2004; Kim et al. 2009; Huntzinger and Izaurralde 2011). The target sequences of animal miRNAs typically lie in the 3' untranslated regions (3' UTRs) of mRNAs while in plants, they are predominantly found in coding regions (Carthew and Sontheimer 2009). However, mature miRNAs displaying atypical characteristics and functionality have been reported recently.

2.3.2.1 5' UTR Binding and Translation Activation

An affinity-based target-identification method revealed that a highly conserved murine miRNA, miR-10a, preferentially binds to ribosomal protein transcripts. Further investigations showed that regions in the 5' UTRs of these transcripts, and not in their 3' UTRs, are directly bound by the miRNA and that rather than translation repression, this binding leads to translation activation, thus resulting in stimulation of global protein synthesis (Orom et al. 2008). Similarly, a human

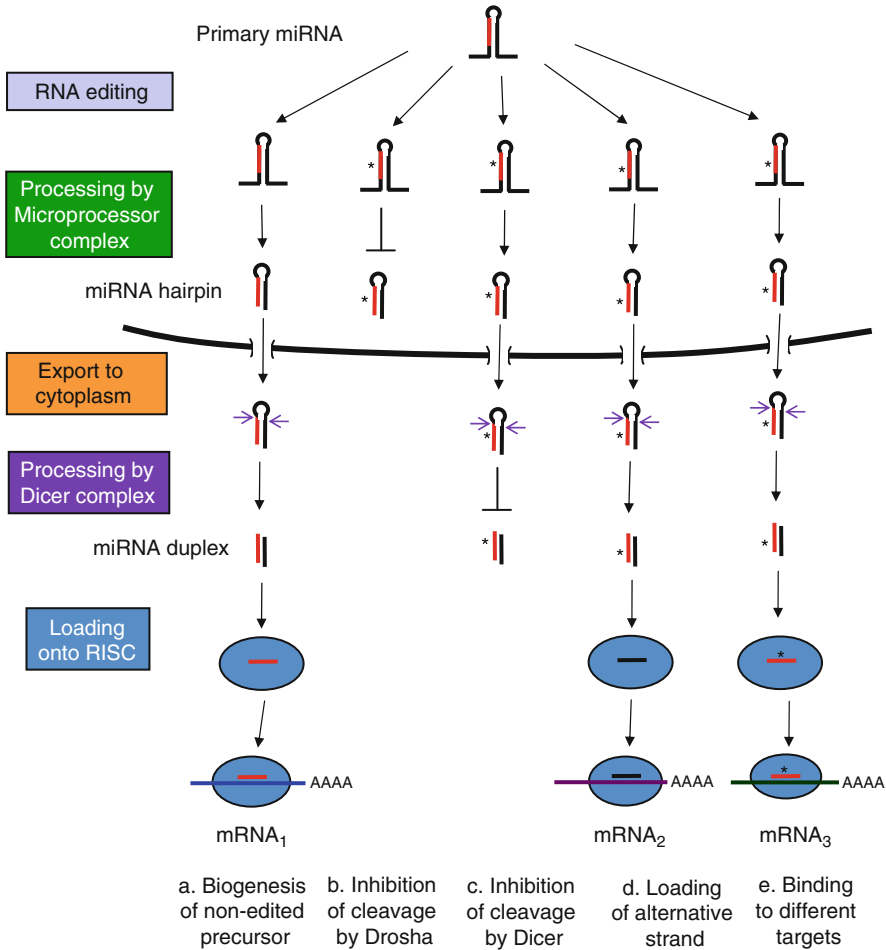


Fig. 2.7 *Diversity generated by RNA editing.* Primary miRNA transcripts can be modified by ADARs resulting in variation in the miRNA sequence (marked with a star). Such modifications can lead to diversity in the biogenesis pathway and in miRNA functionality

miRNA was found to bind to 5' UTRs in the hepatitis C virus genome resulting in translation activation of viral proteins (Henke et al. 2008). Several miRNAs have also been shown to bind targets in coding sequences in both mouse and human (Chi et al. 2009; Duursma et al. 2008; Forman et al. 2008; Tay et al. 2008).

In addition to variation in the position of targets within transcripts, a small number of human miRNAs have been also found to activate translation of their targets in a cell cycle-dependent manner (Vasudevan et al. 2007). Thus, a small number of miRNAs can cause the upregulation of translation rather than its repression.

2.3.2.2 Sequence Editing

RNA editing is the directed modification of specific positions in RNA transcripts resulting in molecules that differ from the template DNA. RNA editing is thus an important mechanism driving diversity in the small RNA landscape. Estimates of the proportion of human miRNA transcripts altered by RNA editing range from 6% to 16% (Blow et al. 2006; Kawahara et al. 2008). RNA editing of primary miRNA transcripts is carried out by ADARs (adenosine deaminases acting on RNA) and leads to the modification of adenosine (A) residues into inosines (I) (reviewed in Cai et al. 2009; Wulff and Nishikura 2010). RNA editing of miRNA transcripts can affect their biogenesis by blocking processing by Drosha or Dicer and thus plays a regulatory role in the biogenesis of miRNAs (Fig. 2.7b, c). RNA editing has also been proposed to cause a change in the miRNA duplex strand most predominantly chosen for incorporation in the RISC complex (Wulff and Nishikura 2010), see Fig. 2.7d. RNA editing can also lead to the binding of alternative targets when compared to the nonedited miRNA, thus extending the repertoire of targets that are regulated by a single miRNA locus (reviewed in Cai et al. 2009 and illustrated in Fig. 2.7e).

2.3.2.3 Nonprototypical miRNA Characteristics and Functions

Nuclear miRNAs

As described previously, the final steps in processing of miRNA precursors and release of mature miRNAs take place in the cytoplasm. However, a large number of human mature miRNAs have been detected in both the cytoplasm and nucleus, and some accumulate predominantly in the nucleus (Liao et al. 2010; Winter et al. 2009). In addition, many mammalian miRNA precursors as well as mature miRNAs have been detected in the nucleolus, some accumulating strongly in this compartment (Politz et al. 2009; Ono et al. 2011; Scott et al. 2009). The nuclear/nucleolar localization mechanism of miRNAs and their function in these compartments are currently not known. A subset of miRNAs might be processed by a noncanonical pathway and/or from noncanonical precursors (including snoRNAs, see Sect. 2.3.1.4) in the nucleus. Alternatively, they might localize in this compartment after processing in the cytoplasm for further modification including RNA editing or proper packaging into ribonucleoprotein complexes (RNPs) as discussed in Politz et al. (2009). However, evidence is also mounting that small RNAs are involved in a number of functions not directly related to translation regulation.

Functions Outside Posttranscriptional Regulation

Small RNAs identify their targets by base pairing to complementary sequences and wield their functionality by bringing effector proteins in close proximity to the

targets. Small RNAs thus have access to the whole spectrum of RNA molecules present in the cell to exert their regulatory roles and are not limited to the small range of specific targets that have been thus far extensively characterized. In recent years, small RNAs have been found to regulate numerous other cellular processes from chromatin structure to transcription and RNA processing (reviewed in Costa 2010; Carthew and Sontheimer 2009; Taft et al. 2010). Newly identified classes of small noncoding RNAs include promoter-associated short RNAs, transcription initiation RNAs (tiRNAs), and transcription-start-site-associated RNAs which regulate diverse aspects of the transcription of protein-coding genes or help to maintain these genes in an active state (Costa 2010). Different classes of small noncoding RNAs have also been proposed to play a role in the regulation of alternative splicing, X chromosome inactivation, and possibly the maintenance of telomeres (reviewed in Khanna and Stamm 2010; Taft et al. 2010). Some of these small RNAs with newly described functions derive from other well-characterized RNA molecules while others represent novel types. For example, small RNAs derived from specific snoRNAs have been found to regulate splicing of several transcripts (Kishore and Stamm 2006) while tiRNAs, derived from regions adjacent to transcription start sites in metazoans, have not been shown to produce other types of previously characterized small RNAs (Taft et al. 2009a).

2.4 Conclusions and Outlook

Rapidly evolving, abundant and highly diverse, small RNAs play fundamental cellular regulatory roles, the extent of which we are only now starting to grasp. Though only a small number of types have been extensively characterized, it is becoming clear that they are dynamic molecules, often displaying a strong association with transposable elements. Different types of small RNAs show commonalities in their origin and biogenesis pathways, making it often difficult to classify them in a clear-cut manner. In addition, a wide diversity and variation exist in small RNA biogenesis pathways, even when considering only one type of small RNAs. And though the canonical miRNA pathway has been extensively described, no component has been found essential for the biogenesis of all miRNAs. The diversity that exists in small RNA biogenesis pathways likely allows for differential regulation for molecules within the same class and in different cell types.

Small RNAs are central regulators in a steadily increasing number of fundamental cellular processes. They regulate a diverse and growing number of types of targets and are responsible for multiple layers in cellular regulatory networks. Understanding their functions and how they are themselves regulated will thus be central to a global comprehension of cellular networks and the numerous diseases caused by their deregulation.

Acknowledgments MSS is the recipient of a Caledonian Research Foundation Fellowship.

References

- Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M, Tuschl T (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 442(7099):203–207. doi:[nature04916](#) [pii] [10.1038/nature04916](#)
- Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, dicer-dependent small RNAs. *Genes Dev* 22(20):2773–2785. doi:[22/20/2773](#) [pii] [10.1101/gad.1705308](#)
- Bachellerie JP, Cavaille J, Huttenhofer A (2002) The expanding snoRNA world. *Biochimie* 84(8):775–790
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116(2):281–297. doi:[S0092867404000455](#) [pii]
- Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136(2):215–233. doi:[S0092-8674\(09\), 00008-7](#) [pii] [10.1016/j.cell.2009.01.002](#)
- Baskerville S, Bartel DP (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 11(3):241–247
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409(6818):363–366. doi:[10.1038/35053110](#)
- Blow MJ, Grocock RJ, van Dongen S, Enright AJ, Dicks E, Futreal PA, Wooster R, Stratton MR (2006) RNA editing of human microRNAs. *Genom Biol* 7(4):R27. doi:[gb-2006-7-4-r27](#) [pii] [10.1186/gb-2006-7-4-r27](#)
- Bogerd HP, Karnowski HW, Cai X, Shin J, Pohlers M, Cullen BR (2010) A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral microRNAs. *Mol Cell* 37(1):135–142. doi:[S1097-2765\(09\), 00920-4](#) [pii] [10.1016/j.molcel.2009.12.016](#)
- Borchert GM, Lanier W, Davidson BL (2006) RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* 13(12):1097–1101
- Brameier M, Herwig A, Reinhardt R, Walter L, Gruber J (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res* 39(2):675–686. doi:[gkq776](#) [pii] [10.1093/nar/gkq776](#)
- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128(6):1089–1103. doi:[S0092-8674\(07\), 00257-7](#) [pii] [10.1016/j.cell.2007.01.043](#)
- Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322(5906):1387–1392. doi:[322/5906/1387](#) [pii] [10.1126/science.1165171](#)
- Cai X, Cullen BR (2007) The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* 13(3):313–316. doi:[rna.351707](#) [pii] [10.1261/rna.351707](#)
- Cai Y, Yu X, Hu S, Yu J (2009) A brief review on the mechanisms of miRNA regulation. *Genom Proteom Bioinf* 7(4):147–154. doi:[S1672-0229\(08\), 60044-3](#) [pii] [10.1016/S1672-0229\(08\)60044-3](#)
- Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell* 136(4):642–655. doi:[S0092-8674\(09\)00083-X](#) [pii] [10.1016/j.cell.2009.01.035](#)
- Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ (2010) A Dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465(7298):584–589. doi:[nature09092](#) [pii] [10.1038/nature09092](#)
- Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460(7254):479–486. doi:[nature08170](#) [pii] [10.1038/nature08170](#)
- Chung WJ, Okamura K, Martin R, Lai EC (2008) Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. *Curr Biol* 18(11):795–802. doi:[S0960-9822\(08\), 00603-9](#) [pii] [10.1016/j.cub.2008.05.006](#)

- Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, Green PJ, Barton GJ, Hutvagner G (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* 15(12):2147–2160. doi:[rna.1738409](#) [pii] [10.1261/rna.1738409](#)
- Cordaux R, Batzer MA (2009) The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 10(10):691–703. doi:[nrg2640](#) [pii] [10.1038/nrg2640](#)
- Costa FF (2010) Non-coding RNAs: meet thy masters. *Bioessays* 32(7):599–608. doi:[10.1002/bies.200900112](#)
- Cox DN, Chao A, Baker J, Chang L, Qiao D, Lin H (1998) A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev* 12(23):3715–3727
- Czech B, Hannon GJ (2011) Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet* 12(1):19–31. doi:[nrg2916](#) [pii] [10.1038/nrg2916](#)
- Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, Ketting RF, Tavare S, Miska EA (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol Cell* 31(1):79–90. doi:[S1097-2765\(08\), 00392-4](#) [pii] [10.1016/j.molcel.2008.06.003](#)
- Deininger PL, Batzer MA (2002) Mammalian retroelements. *Genome Res* 12(10):1455–1465. doi:[10.1101/gr.282402](#)
- Dieci G, Preti M, Montanini B (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* 94(2):83–88. doi:[S0888-7543\(09\)00106-2](#) [pii]
- Duursma AM, Kedde M, Schrier M, le Sage C, Agami R (2008) miR-148 targets human DNMT3b protein coding region. *RNA* 14(5):872–877. doi:[rna.972008](#) [pii] [10.1261/rna.972008](#)
- Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G (2008) A Human snoRNA with microRNA-like functions. *Mol Cell* 32(4):519–528
- Faulkner GJ, Carninci P (2009) Altruistic functions for selfish DNA. *Cell Cycle* 8(18):2895–2900. doi:[9536](#) [pii]
- Filipowicz W, Pogacic V (2002) Biogenesis of small nucleolar ribonucleoproteins. *Curr Opin Cell Biol* 14(3):319–327
- Forman JJ, Legesse-Miller A, Collier HA (2008) A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proc Natl Acad Sci USA* 105(39):14879–14884. doi:[0803230105](#) [pii] [10.1073/pnas.0803230105](#)
- Garzon R, Fabbri M, Cimmino A, Calin GA, Croce CM (2006) MicroRNA expression and function in cancer. *Trends Mol Med* 12(12):580–587. doi:[S1471-4914\(06\), 00242-5](#) [pii] [10.1016/j.molmed.2006.10.006](#)
- Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, Zamore PD (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* 320(5879):1077–1081. doi:[1157396](#) [pii] [10.1126/science.1157396](#)
- Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 315(5818):1587–1590. doi:[1140494](#) [pii] [10.1126/science.1140494](#)
- He S, Su H, Liu C, Skogerbo G, He H, He D, Zhu X, Liu T, Zhao Y, Chen R (2008) MicroRNA-encoding long non-coding RNAs. *BMC Genom* 9:236. doi:[1471-2164-9-236](#) [pii] [10.1186/1471-2164-9-236](#)
- Henke JI, Goergen D, Zheng J, Song Y, Schuttler CG, Fehr C, Junemann C, Niepmann M (2008) MicroRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* 27(24):3300–3310. doi:[emboj2008244](#) [pii] [10.1038/emboj.2008.244](#)
- Hirsch J, Lefort V, Vankersschaver M, Boualem A, Lucas A, Thermes C, d'Aubenton-Carafa Y, Crespi M (2006) Characterization of 43 non-protein-coding mRNA genes in *Arabidopsis*, including the MIR162a-derived transcripts. *Plant Physiol* 140(4):1192–1204. doi:[pp.105.073817](#) [pii] [10.1104/pp.105.073817](#)

- Hock J, Weinmann L, Ender C, Rudel S, Kremmer E, Raabe M, Urlaub H, Meister G (2007) Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep* 8(11):1052–1060. doi:[7401088 \[pii\]](#) [10.1038/sj.embor.7401088](#)
- Houwing S, Kammenga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RH, Hannon GJ, Draper BW, Ketting RF (2007) A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* 129(1):69–82. doi:[S0092-8674\(07\), 00392-3 \[pii\]](#) [10.1016/j.cell.2007.03.026](#)
- Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* 12(2):99–110. doi:[nrg2936 \[pii\]](#) [10.1038/nrg2936](#)
- Hutvagner G, Zamore PD (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297(5589):2056–2060. doi:[10.1126/science.10738271073827 \[pii\]](#)
- Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293(5531):834–838. doi:[10.1126/science.10629611062961 \[pii\]](#)
- Kawahara Y, Megraw M, Kreider E, Iizasa H, Valente L, Hatzigeorgiou AG, Nishikura K (2008) Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res* 36(16):5270–5280. doi:[gkn479 \[pii\]](#) [10.1093/nar/gkn479](#)
- Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H (2008) Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* 453(7196):793–797. doi:[nature06938 \[pii\]](#) [10.1038/nature06938](#)
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. *Genome Res* 12(6):996–1006
- Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15(20):2654–2659. doi:[10.1101/gad.927801](#)
- Khanna A, Stamm S (2010) Regulation of alternative splicing by short non-coding nuclear RNAs. *RNA Biol* 7(4):480–485. doi:[12746 \[pii\]](#)
- Khurana JS, Theurkauf W (2010) piRNAs, transposon silencing, and Drosophila germline development. *J Cell Biol* 191(5):905–913. doi:[jcb.201006034 \[pii\]](#) [10.1083/jcb.201006034](#)
- Kim YK, Kim VN (2007) Processing of intronic microRNAs. *EMBO J* 26(3):775–783
- Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* 10(2):126–139. doi:[nrm2632 \[pii\]](#) [10.1038/nrm2632](#)
- Kishore S, Stamm S (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2 C. *Science* 311(5758):230–232
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW, Hata K, Li E, Matsuda Y, Kimura T, Okabe M, Sakaki Y, Sasaki H, Nakano T (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev* 22(7):908–917. doi:[22/7/908 \[pii\]](#) [10.1101/gad.1640708](#)
- Lai EC (2005) miRNAs: whys and wherefores of miRNA-mediated regulation. *Curr Biol* 15(12):R458–R460
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrum J, Mesirov JP, Miranda C, Morris W, Naylor J et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921
- Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425(6956):415–419. doi:[10.1038/nature01957](#) [nature01957 \[pii\]](#)
- Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23(20):4051–4060
- Liao JY, Ma LM, Guo YH, Zhang YC, Zhou H, Shao P, Chen YQ, Qu LH (2010) Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular

- distribution of miRNAs and tRNA 3' trailers. *PLoS One* 5(5):e10563. doi:[10.1371/journal.pone.0010563](https://doi.org/10.1371/journal.pone.0010563)
- Llave C, Kasschau KD, Rector MA, Carrington JC (2002) Endogenous and silencing-associated small RNAs in plants. *Plant Cell* 14(7):1605–1619
- Luo Y, Li S (2007) Genome-wide analyses of retrogenes derived from the human box H/ACA snoRNAs. *Nucleic Acids Res* 35(2):559–571
- Malone CD, Hannon GJ (2009) Small RNAs as guardians of the genome. *Cell* 136(4):656–668. doi:[S0092-8674\(09\), 00127-5 \[pii\] 10.1016/j.cell.2009.01.045](https://doi.org/10.1016/j.cell.2009.01.045)
- Matera AG, Terns RM, Terns MP (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* 8(3):209–220. doi:[nrm2124 \[pii\] 10.1038/nrm2124](https://doi.org/10.1038/nrm2124)
- Mills RE, Bennett EA, Iskow RC, Devine SE (2007) Which transposable elements are active in the human genome? *Trends Genet* 23(4):183–191. doi:[S0168-9525\(07\), 00059-5 \[pii\] 10.1016/j.tig.2007.02.006](https://doi.org/10.1016/j.tig.2007.02.006)
- Miyoshi K, Miyoshi T, Siomi H (2010) Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. *Mol Genet Genom* 284(2):95–103. doi:[10.1007/s00438-010-0556-1](https://doi.org/10.1007/s00438-010-0556-1)
- Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* 16(6):720–728. doi:[10.1101/gad.974702](https://doi.org/10.1101/gad.974702)
- Mourier T, Willerslev E (2009) Retrotransposons and non-protein coding RNAs. *Brief Funct Genom Proteomic* 8(6):493–501. doi:[elp036 \[pii\] 10.1093/bfpg/elp036](https://doi.org/10.1093/bfpg/elp036)
- Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* 39:121–152. doi:[10.1146/annurev.genet.39.073003.112240](https://doi.org/10.1146/annurev.genet.39.073003.112240)
- O'Donnell KA, Boeke JD (2007) Mighty Piwis defend the germline against genome intruders. *Cell* 129(1):37–44. doi:[S0092-8674\(07\), 00396-0 \[pii\] 10.1016/j.cell.2007.03.028](https://doi.org/10.1016/j.cell.2007.03.028)
- Ono M, Scott MS, Yamada K, Avolio F, Barton GJ, Lamond AI (2011) Identification of human miRNA precursors that resemble box C/D snoRNAs. *Nucleic Acids Res*. doi:[gkq1355 \[pii\] 10.1093/nar/gkq1355](https://doi.org/10.1093/nar/gkq1355)
- Orom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* 30(4):460–471. doi:[S1097-2765\(08\), 00328-6 \[pii\] 10.1016/j.molcel.2008.05.001](https://doi.org/10.1016/j.molcel.2008.05.001)
- Pederson T (2010) Regulatory RNAs derived from transfer RNA? *RNA* 16(10):1865–1869. doi:[rna.2266510 \[pii\] 10.1261/rna.2266510](https://doi.org/10.1261/rna.2266510)
- Piriyaopongsa J, Jordan IK (2007) A family of human microRNA genes from miniature inverted-repeat transposable elements. *PLoS One* 2(2):e203
- Piriyaopongsa J, Marino-Ramirez L, Jordan IK (2007) Origin and evolution of human microRNAs from transposable elements. *Genetics* 176(2):1323–1337
- Politz JC, Hogan EM, Pederson T (2009) MicroRNAs with a nucleolar location. *RNA* 15(9):1705–1715. doi:[rna.1470409 \[pii\] 10.1261/rna.1470409](https://doi.org/10.1261/rna.1470409)
- Richard GF, Kerrest A, Dujon B (2008) Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol Mol Biol Rev* 72(4):686–727. doi:[72/4/686 \[pii\] 10.1128/MMBR.00011-08](https://doi.org/10.1128/MMBR.00011-08)
- Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res* 14(10A):1902–1910
- Saito K, Siomi MC (2010) Small RNA-mediated quiescence of transposable elements in animals. *Dev Cell* 19(5):687–697. doi:[S1534-5807\(10\), 00464-8 \[pii\] 10.1016/j.devcel.2010.10.011](https://doi.org/10.1016/j.devcel.2010.10.011)
- Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev* 20(16):2214–2222. doi:[gad.1454806 \[pii\] 10.1101/gad.1454806](https://doi.org/10.1101/gad.1454806)
- Saraiya AA, Wang CC (2008) snoRNA, a novel precursor of microRNA in *Giardia lamblia*. *PLoS Pathog* 4(11):e1000224. doi:[10.1371/journal.ppat.1000224](https://doi.org/10.1371/journal.ppat.1000224)

- Schmitz J, Zemann A, Churakov G, Kuhl H, Grutzner F, Reinhardt R, Brosius J (2008) Retroposed SNOfall – a mammalian-wide comparison of platypus snoRNAs. *Genome Res* 18 (6):1005–1010. doi:[gr.7177908](#) [pii] [10.1101/gr.7177908](#)
- Scott MS, Avolio F, Ono M, Lamond AI, Barton GJ (2009) Human miRNA precursors with box H/ACA snoRNA features. *PLoS Comput Biol* 5(9):e1000507. doi:[10.1371/journal.pcbi.1000507](#)
- Siomi MC, Saito K, Siomi H (2008) How selfish retrotransposons are silenced in *Drosophila* germline and somatic cells. *FEBS Lett* 582(17):2473–2478. doi:[S0014-5793\(08\), 00508-5](#) [pii] [10.1016/j.febslet.2008.06.018](#)
- Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8(4):272–285. doi:[nrg2072](#) [pii] [10.1038/nrg2072](#)
- Smalheiser NR, Torvik VI (2005) Mammalian microRNAs derived from genomic repeats. *Trends Genet* 21(6):322–326
- Smalheiser NR, Torvik VI (2006) Alu elements within human mRNAs are probable microRNA targets. *Trends Genet* 22(10):532–536. doi:[S0168-9525\(06\), 00260-5](#) [pii] [10.1016/j.tig.2006.08.007](#)
- Sunkar R, Girke T, Zhu JK (2005) Identification and characterization of endogenous small interfering RNAs from rice. *Nucleic Acids Res* 33(14):4443–4454. doi:[33/14/4443](#) [pii] [10.1093/nar/gki758](#)
- Suzuki HI, Miyazono K (2011) Emerging complexity of microRNA generation cascades. *J Biochem* 149(1):15–25. doi:[mvq113](#) [pii] [10.1093/jb/mvq113](#)
- Taft RJ, Glazov EA, Cloonan N, Simons C, Stephen S, Faulkner GJ, Lassmann T, Forrest AR, Grimmond SM, Schroder K, Irvine K, Arakawa T, Nakamura M, Kubosaki A, Hayashida K, Kawazu C, Murata M, Nishiyori H, Fukuda S, Kawai J, Daub CO, Hume DA, Suzuki H, Orlando V, Carninci P, Hayashizaki Y, Mattick JS (2009a) Tiny RNAs associated with transcription start sites in animals. *Nat Genet* 41 (5):572–578. doi:[ng.312](#) [pii]
- Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS (2009) Small RNAs derived from snoRNAs. *RNA* 15(7):1233–1240. doi:[rna.1528909](#) [pii] [10.1261/rna.1528909](#)
- Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS (2010) Non-coding RNAs: regulators of disease. *J Pathol* 220(2):126–139. doi:[10.1002/path.2638](#)
- Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I (2008) MicroRNAs to nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455(7216):1124–1128. doi:[nature07299](#) [pii] [10.1038/nature07299](#)
- Tenaillon MI, Hollister JD, Gaut BS (2010) A triptych of the evolution of plant transposable elements. *Trends Plant Sci* 15(8):471–478. doi:[S1360-1385\(10\), 00096-8](#) [pii] [10.1016/j.tplants.2010.05.003](#)
- Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318(5858):1931–1934
- Weber MJ (2006) Mammalian small nucleolar RNAs are mobile genetic elements. *PLoS Genet* 2(12):e205
- Winter J, Jung S, Keller S, Gregory RI, Diederichs S (2009) Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* 11(3):228–234. doi:[ncb0309-228](#) [pii] [10.1038/ncb0309-228](#)
- Wulff BE, Nishikura K (2010) Substitutional A-to-I RNA editing. *WIREs RNA* 1(1):90–101. doi:[10.1002/wrna.10](#)
- Yuan Z, Sun X, Jiang D, Ding Y, Lu Z, Gong L, Liu H, Xie J (2010) Origin and evolution of a placental-specific microRNA family in the human genome. *BMC Evol Biol* 10:346. doi:[1471-2148-10-346](#) [pii] [10.1186/1471-2148-10-346](#)
- Yuan Z, Sun X, Liu H, Xie J (2011) MicroRNA genes derived from repetitive elements and expanded by segmental duplication events in mammalian genomes. *PLoS One* 6(3):e17666. doi:[10.1371/journal.pone.0017666](#)
- Zhang R, Wang YQ, Su B (2008) Molecular evolution of a primate-specific microRNA family. *Mol Biol Evol* 25(7):1493–1502. doi:[msn094](#) [pii] [10.1093/molbev/msn094](#)



<http://www.springer.com/978-3-642-22516-1>

Regulatory RNAs

Basics, Methods and Applications

Mallick, B.; Ghosh, Z. (Eds.)

2012, X, 554 p., Hardcover

ISBN: 978-3-642-22516-1