

# The Expanding Repertoire of Circular RNAs

Paul N Valdmanis<sup>1</sup> and Mark A Kay<sup>1</sup>

doi:10.1038/mt.2013.101

In recent years, reports suggesting a process of exon shuffling and the generation of circular RNAs have not gained wide support because they are inconsistent with the established dogma of linear exon splicing. This debate has now come full circle with evidence proving more definitively the existence of circular RNAs in two recent articles published in the 21 March issue of *Nature*. These articles describe an abundance of examples of circular RNAs in humans, mice, and *Caenorhabditis elegans* as well as a plausible mechanism for the function of at least some of these as “microRNA sponges” within a cell.<sup>1,2</sup> A microRNA (miRNA) sponge is a sequence that sequesters the available pool of a given miRNA so as to reduce its capacity to act on target genes. However, there is as yet relatively little information on the role that these circular RNAs play in cellular homeostasis or disease, such that one might question why these findings deserve the attention of readers of *Molecular Therapy*. One might have asked similar questions in the 1990s following the discovery of miRNAs, small interfering RNAs (siRNAs), and RNA interference (RNAi). Although miRNA pathways were once thought to be limited to plants and invertebrates, their importance in normal human development and disease becomes more apparent on almost a weekly basis. Moreover, delivery of short duplex siRNAs or DNA templates capable of producing RNA hairpins that are processed into siRNAs—eliciting robust RNAi-induced gene knockdown and RNAi-based therapies—is now in advanced clinical trials.<sup>3–5</sup> It is not too much of a stretch to

speculate that manipulating endogenous circular RNAs will someday provide novel approaches to treat human disease.

Circular RNAs were first described anecdotally in the 1990s, and the most well-studied example is that generated from the sex-determining Y (*SRY*) gene.<sup>6</sup> The significance of circular RNAs fell somewhat under the radar until last year, when a few large-scale high-throughput sequencing experiments brought them again to the fore in this age of “big data.”<sup>7–9</sup> These reports identified thousands of circular and nonlinear exon-shuffling events, derived from 1–2% of the genes transcribed in a cell. The circular RNAs were identified from multiple tissue types,<sup>9</sup> were resistant to RNase R treatment (which degrades linear RNA species), and had longer half-lives in cells relative to their linear RNA transcript counterparts.<sup>8</sup> Only the linear forms of these RNAs were present in heavy polyribosome fractions, suggesting that the circular forms are not translated.<sup>8</sup>

A few mechanisms for the formation of circular RNAs have been proposed (Figure 1), such as via circularization of a single exon, facilitated by the presence of adjacent repetitive sequence.<sup>6</sup> Alternatively, circular transcripts can form after either the choice of an incorrect splice acceptor or removal of several consecutive exons in an alternatively spliced transcript. The most basic form of circular RNA involves the looping-back of sequences to a previous exon within the transcript, for instance, a junction between the 3' end of exon 3 and the 5' end of exon 2 in a gene (Figure 1). The intronic sequence surrounding many circular transcripts is noticeably longer and enriched for repetitive elements such as ALU repeats, often in an inverted orientation.<sup>8</sup>

In the two new studies in *Nature*, the primary circular RNA of interest is an antisense transcript to *CDR1* termed *CDR1as* (*CDR1* antisense, used herein) or *ciRS-7* (circular RNA sponge for miR-7).<sup>1,2</sup> This

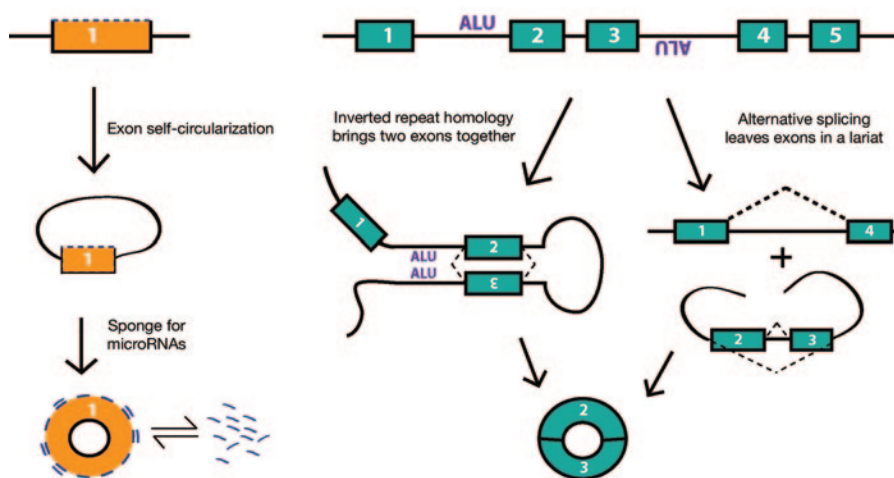
RNA was validated by rigorous bioinformatic and molecular studies. Seventy-four copies of a miR-7 binding site are present in this circular RNA owing to the presence of a tandem repeat in the transcribed sequence of *CDR1as*. These binding sites are conserved across eutherian (placental) mammals, whereas the intervening sequences generally are not. Other examples of miRNA sponges—for example, pseudogenes that act as competitive inhibitors of endogenous targets—often exist at an equivalent or lower stoichiometric abundance relative to their natural gene targets.<sup>10</sup> By contrast, the approximately 70 miR-7 targets present in the *CDR1as* transcript imply that the pool of cytoplasmic miR-7 would be rapidly sequestered following expression of *CDR1as*. Identification of circular RNAs specific for a miRNA of interest or engineering of the *CDR1as* sequences to include binding sites for the miRNA of interest could be a powerful mechanism to curb the effects of miRNAs that are overexpressed in a disease context, including many examples in oncogenesis.

The study by Hansen *et al.*<sup>1</sup> showed that inverted repeats before and after the *CDR1as* transcript were essential for its circularization. Disabling these inverted repeats ensures that the *CDR1as* sequence remains linear and is subject to repression and eventual degradation by miR-7. Curiously, although the circular RNA transcript is resistant to miR-7-mediated degradation, the same group showed that a sequence with perfect complementarity to miR-671 also exists in *CDR1as* but is subject to Ago2-mediated cleavage and degradation.<sup>11</sup> Biotin-coupled miR-7 could pull down *CDR1as*, and fluorescence *in situ* hybridization analysis showed colocalization and compartmentalization of miR-7 and *CDR1as* to processing bodies dependent on the presence of miR-7. Many, if not most, of the proteins necessary for miRNA gene silencing are localized to processing bodies, which play fundamental roles in general messenger RNA (mRNA) decay, nonsense-mediated mRNA decay, AU-rich element-mediated mRNA decay, and miRNA-induced mRNA silencing. The two RNA species are also coexpressed in pyramidal neurons and interneurons. Stable but low and forced expression of *CDR1as* in HeLa cells was able to compete for miR-7 species and prevented repression of known miR-7 target 3' untranslated regions for the

<sup>1</sup>Departments of Pediatrics and Genetics, Stanford University, Stanford, California, USA

Correspondence: Mark A Kay, Departments of Pediatrics and Genetics, Stanford University, 269 Campus Drive, Room 2105, Stanford, California 94305, USA.

E-mail: markay@stanford.edu



**Figure 1 Potential mechanisms by which circular RNAs are generated.** (Left) Presumed mechanism of *CDR1as* generation and its microRNA sponge function. (Right) The functional significance of circularized RNAs generated by the two other mechanisms is not known.

genes *SNCA*, *EGFR*, and *IRS2*. Finally, some of the functional characteristics for *CDR1as* were observed for the well-known but still understudied *SRY* gene and its circular RNA product that harbors corresponding target sites for miR-138 (ref. 1).

Memczak *et al.*<sup>2</sup> focused on profiling the increasing numbers of circular RNAs they were able to cull from RNA sequencing data sets. They found nearly 2,000 circular RNAs each in humans and mice and approximately 700 in *C. elegans*. The vast majority of the identified circular RNAs overlapped with coding exons; 223 circular RNAs were fully exonic and conserved between humans and mice. Using these data, they could show that the third (wobble) nucleotide of codons in general was under greater selective pressure within these genes than in other genes—indicative of an evolutionary constraint for the nucleotide sequence of these circular RNAs. In addition, for circular RNAs containing multiple copies of the same miRNA seed, this seed sequence was much more conserved than the surrounding region. This, again, was most notable in the *CDR1as* transcript. The authors also validated *CDR1as* functionally, in this case by means of a zebrafish model. Zebrafish retain the ancestral miR-7 sequence but not the *CDR1as* transcript, which emerged following the evolution of placental mammals. Consistent with the role of miR-7 in neuronal function,<sup>1</sup> the authors found that antagonists to miR-7 caused

a reduction in midbrain size. Likewise, delivery of *CDR1as* in zebrafish, which presumably also sponges miR-7, led to a similar reduction in midbrain size in this model.

Functional similarities between *CDR1as* and the remaining circular RNAs that have been identified remain to be determined. Many of the latter are much shorter and have less obvious potential as miRNA sponges; moreover, at least a subset of them may simply reflect by-products of alternative splicing (Figure 1). So what are the other functions of these circular RNAs? Considering the historical evolution of our understanding of RNA biogenesis and function, it would be unwise to dismiss these RNAs as mere “junk” and move on. Yet the majority of circular transcripts comprise only a very few number of exons. Perhaps they can compete with exon splice enhancers or suppressors and thereby provide at least a partial explanation for the vast complexities and pervasiveness of alternative splicing. Or perhaps they act as sponges for other RNA-binding proteins. Several important questions remain; the mean length of the circular transcripts has not been established, nor the number of isoforms generated per coding transcript or how they are regulated or degraded. Answering these questions should provide some clue to what other cellular functions could be perturbed (intentionally or not) by delivering circular RNAs.

These circular RNAs may be relics of an ancestral world and/or may be related to existing viral RNA species that remain circular today, a notable example being the hepatitis D virus genome. This RNA virus replicates through a process of RNA-mediated RNA transcription despite the lack of a virus-encoded RNA-dependent RNA polymerase.<sup>12</sup> Could circular RNAs be copied in a similar manner? In any event, our experience in manipulating viral sequences may be a starting point for generating deliberately designed circular RNAs. Furthermore, minicircle DNA vectors that contain virtually no DNA between the 5' end of the promoter and the 3' end of the polyadenylation site are persistently transcribed and, unlike their canonical plasmid counterparts (which are transcriptionally silenced), remain bound to RNA polymerase II across the whole expression cassette.<sup>13</sup> Understanding the expression and epigenetic status of circular RNAs could help improve our understanding of epigenetic changes in gene expression persistence or silencing in the context of gene transfer vectors.

The large data sets that identify circular RNAs that are largely overlapping between independent studies,<sup>8</sup> along with evidence for potential function, indicate that their physiological impact can no longer be ignored, and the surprising revelations from these recent articles will provide new avenues for researching their function. Thus far, the following conclusions can be drawn about circular RNAs: they are conserved across tissues and appear in multiple species, and they are generally not translated and lack a poly(A) tail, yet are stable and resistant to exonuclease digestion. Some circular RNA species are large, but the vast majority appear to be three or fewer exons long. Much of this initial work has been performed using custom databases and enrichment procedures and thus probably underestimates the true abundance of these circular RNAs. It will be beneficial to establish consolidated protocols not only for the detection of these circular RNAs but also for their biogenesis from a biotechnology standpoint, thus adding another weapon to our arsenal for combating disease. In time, the potential for developing new nucleic acid therapeutics from circular RNAs will be realized, and lessons learned from current gene and oligonucleotide approaches may prove to be invaluable for their implementation.

## ACKNOWLEDGMENT

This work was supported by National Institutes of Health grant HL064274.

## REFERENCES

- Hansen, TB, Jensen, TI, Clausen, BH, Bramsen, JB, Finsen, B, Damgaard, CK *et al.* (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* **495**: 384–388.
- Memczak, S, Jens, M, Elefsinioti, A, Torti, F, Krueger, J, Rybak, A *et al.* (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**: 333–338.
- Kubowicz, P, Zelazarczyk, D and Pekala, E (2013). RNAi in clinical studies. *Curr Med Chem* **20**: 1801–1816.
- Martinez, T, Wright, N, López-Fraga, M, Jiménez, AI and Pañeda, C (2013). Silencing human genetic diseases with oligonucleotide-based therapies. *Hum Genet* **132**: 481–493.
- Ramachandran, PV and Ignacimuthu, S (2013). RNA interference—a silent but an efficient therapeutic tool. *Appl Biochem Biotechnol* **169**: 1774–1789.
- Capel, B, Swain, A, Nicolis, S, Hacker, A, Walter, M, Koopman, P *et al.* (1993). Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* **73**: 1019–1030.
- Al-Balool, HH, Weber, D, Liu, Y, Wade, M, Guleria, K, Nam, PL *et al.* (2011). Post-transcriptional exon shuffling events in humans can be evolutionarily conserved and abundant. *Genome Res* **21**: 1788–1799.
- Jeck, WR, Sorrentino, JA, Wang, K, Slevin, MK, Burd, CE, Liu, J *et al.* (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**: 141–157.
- Salzman, J, Gawad, C, Wang, PL, Lacayo, N and Brown, PO (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* **7**: e30733.
- Poliseno, L, Salmena, L, Zhang, J, Carver, B, Haveman, WJ and Pandolfi, PP (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**: 1033–1038.
- Hansen, TB, Wiklund, ED, Bramsen, JB, Villadsen, SB, Statham, AL, Clark, SJ *et al.* (2011). miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* **30**: 4414–4422.
- Haussecker, D, Cao, D, Huang, Y, Parameswaran, P, Fire, AZ and Kay, MA (2008). Capped small RNAs and MOV10 in human hepatitis delta virus replication. *Nat Struct Mol Biol* **15**: 714–721.
- Gracey Maniar, LE, Maniar, JM, Chen, ZY, Lu, J, Fire, AZ and Kay, MA (2013). Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level. *Mol Ther* **21**: 131–138.

See page 1160

## T Lymphocytes Are Not Immune

Gianpietro Dotti<sup>1</sup>

doi:10.1038/mt.2013.102

Administration of genetically modified T lymphocytes to cancer patients has increased exponentially over the past 15 years, and these cells have been given to many hundreds of patients. These studies have often used retroviral vectors, and, in sharp contrast to clinical trials using hematopoietic stem cells as targets, gene-modified T cells have not subsequently undergone malignant transformation. This has led to the impression that the more differentiated state of T lymphocytes may offer protection from retroviral mutagenesis. In this issue of *Molecular Therapy*, however, Heinrich and colleagues document the occurrence of T-cell

lymphoma in RAG<sup>-/-</sup> mice transplanted with OT-1 transgenic T cells genetically modified with a gammaretroviral vector encoding the green fluorescent protein, suggesting that, although T cells may be relatively insensitive to malignant transformation by gammaretroviruses, they are not immune.<sup>1</sup>

Early gene transfer studies focused primarily on detecting T-cell persistence and trafficking to the tumor site by the expression of nonfunctional genes.<sup>2,3</sup> The herpes simplex thymidine kinase is one of the earliest examples of a therapeutic gene expressed in T lymphocytes to pharmacologically induce their death in case of graft-versus-host disease in patients who received an allogeneic hematopoietic stem cell transplant.<sup>4</sup> More recently, the expression of T-cell receptors and chimeric antigen receptors to redirect the antigen specificity of polyclonal T lymphocytes set off a new wave of enthusiasm for adoptive immunotherapy in cancer patients.<sup>5–7</sup> Because T lymphocytes are rapidly dividing cells, all these studies have used vectors

that produce genomic integration of the transgene of interest, which is critical to ensure its long-term functional expression in progeny cells. Most of these studies over the past 20 years have used gammaretroviral vectors for this purpose—in particular, vectors derived from the Moloney murine leukemia retrovirus.<sup>8</sup>

A drawback of this powerful system, however, is the potential to induce genotoxic events by deregulating the expression of oncogenes, ultimately leading to neoplastic transformation of the transduced cells. Although this phenomenon was well recognized in murine models, genotoxicity related to insertion of gammaretroviral vectors in human subjects remained a theoretical risk until 2003, when integration of a retroviral vector encoding the common  $\gamma$ -chain of the interleukin-2 receptor caused deregulated expression of the *LMO2* gene in hematopoietic stem cells and resulted in T-cell leukemia in patients with X-linked severe combined immunodeficiency who had been cured of their disease by gene transfer.<sup>9</sup> These malignant T cells were derived from the transduced stem cell population. Subsequently, neoplastic transformation was also reported in patients treated for chronic granulomatous disease.<sup>10</sup>

Although similar retroviral vectors have been used to modify mature human T lymphocytes, genotoxicity has not been evident in patients who have received these cells. Human T lymphocytes, like hematopoietic stem cells, are susceptible to insertional mutagenesis, but these events do not necessarily evolve to a malignant phenotype in this cell subset. T lymphocytes are generally stimulated by crosslinking of the T-cell receptor and cytokines to allow retroviral integration, and elegant analyses have shown that integrated proviruses are located predominantly around promoters of genes that are active in T cells at the time of transduction, probably because the chromatin configuration of these active genes makes them more accessible.<sup>11</sup> The insertional gene activation in T cells is determined predominantly by the characteristics of the transcriptional regulatory elements carried by the vector rather than by the vector type or design.<sup>12</sup> Vector integration can cause deregulated expression of the nearby genes in T lymphocytes,<sup>11</sup> but,

<sup>1</sup>Center for Cell and Gene Therapy and Departments of Immunology and Medicine, Baylor College of Medicine, Methodist Hospital and Texas Children's Hospital, Houston, Texas, USA  
Correspondence: Gianpietro Dotti, Center for Cell and Gene Therapy, Baylor College of Medicine, 6621 Fanin Street, MC 3-3320, Houston, Texas 77030, USA.  
E-mail: gdotti@bcm.edu