Review

Circular RNAs in Brain and Other Tissues: A Functional Enigma

Wei Chen$^{1,2,*}$ and Erin Schuman$^{3,*}$

Circular RNAs (circRNAs) are RNAs with a covalently closed loop structure that have recently regained the attention of biologists. Using deep RNA sequencing (RNA-seq) coupled with novel bioinformatic approaches, genome-wide studies have detected a large number of circRNAs, many of which are abundant, stable, and well conserved during evolution. With few exceptions, the function of most circRNAs remains elusive. Several recent studies have shown that circRNAs are more enriched in neuronal tissues and are often derived from genes specific for neuronal and synaptic function. Moreover, circRNA expression is regulated during neuronal development and by synaptic plasticity, suggesting specific neuronal functions. In this review, we discuss recent advances in the detection, biogenesis, and potential functions of circRNAs, with a particular focus on brain tissues.

CircRNA: A New Class of RNAs with Potential Regulatory Function

Over the past decade, genome research has fueled the discovery of an ever-growing list of novel RNA species. Beyond the classic tRNA, miRNA, and rRNA, there has emerged a striking diversity of additional RNA types, including miRNA, piwi-interacting (pi)RNA, small nucleolar (sno)RNA, small nuclear (sn)RNA, long noncoding (lnc)RNA, and other noncoding RNAs. More recently, in addition to these linear RNAs with distinct 5’ and 3’ ends, a group of circRNAs with covalently closed loop structures has recently gained attention. Initially discovered as plant viroids [1], hepatitis delta virus [2], or intermediate products from tRNA and rRNA processing ([3–5]; reviewed in [6]), circRNAs derived from either exons or introns were later identified in individual gene loci from eukaryotic cells [7–11]. CircRNAs derived from introns and exons represent two different RNA types with distinct biogenesis mechanisms and likely also exhibit different biological functions. Whereas the intron-derived circRNAs originate from intron lariats that escape debranching and degradation, accumulate in nucleus, and appear to function by regulating the expression of their host gene [10–13], exon-derived circRNAs are formed by back-splicing (i.e., the covalent joining of a downstream splice donor site with an upstream splice acceptor site from the same gene locus; Figure 1) [14–17]; exon-derived circRNAs are predominantly localized in cytoplasm and likely function via a variety of mechanisms [18] (see below). Although some of these initially discovered circRNAs are abundant and can represent the dominant form of an RNA derived from a particular genomic locus (e.g., the circRNA derived from the SRY locus in adult testis) [19], they were originally viewed as a transcription artifacts or splicing noise without important functions in biological processes [20,21]. However, this ‘artifact’ notion has been recently modified in light of several recent studies revealing a large number of circRNAs detected in various cell types of organisms ranging from Archaea to mammals [14–17,22–25]. Many of these circRNAs are abundant, stable, and well conserved during evolution [15,16,24,26,27]. Two seminal studies demonstrated that two previously identified circRNAs can serve as miRNA “sponges”, sequestering miRNAs and preventing their interactions with...
target mRNAs [15,28]. Here, we review recent advances in the detection, biogenesis, and potential functions of circRNAs, with a particular focus on brain tissues.

**CircRNA Detection in the Modern Era of RNA-Seq**

The global prevalence of circRNAs was underestimated until recently. CircRNAs were inaccessible to commonly used genome-wide RNA profiling techniques because of their closed structure (without free 5' or 3' ends) and the difficulty in distinguishing them from their linear RNA isoforms (derived from the same gene locus). From a detection standpoint, the key difference between a circRNA and its linear counterpart is the presence of specific back-splicing junction sequences. Using microarrays, the most popular technology for global RNA expression analysis before the advent of next-generation sequencing, with probes designed to recognize annotated genes or at least sequences present in the genome, it was impossible to detect circRNAs or distinguish the expression of circRNAs from their linear host genes. By contrast, RNA-seq could, in principle, detect any expressed RNAs and identify the ‘aberrant’ junction sequence, thereby signaling the presence of circRNAs [29]. Unfortunately, however, most early RNA-seq studies focused on the sequencing of RNAs that have polyA tails. As such, most studies began with a sample largely depleted of circRNAs owing to polyA RNA enrichment before the preparation of sequencing libraries. Any sporadic sequencing reads mapped to the circRNA junction were often considered to be experimental artifacts. It is only with the advent of sequencing of rRNA-depleted total RNA or nonpolyadenylated RNA (polyA-minus), as well as the development of novel bioinformatics tools, that the genome-wide discovery of circRNAs in different organisms and tissues has become possible [14–17,23,24].

In RNA-seq data, the diagnostic feature of circRNAs is the short read mapped to a back-splicing (spliced in reverse order) site. To identify these junctional reads, a variety of bioinformatics
approaches have been developed that use the reads (single-end or paired-end) that cannot be directly mapped to the reference genome, split them, and align both ends separately [15,23,25,27,30–34]. The reads with the split alignments are further filtered. Depending on whether the alignment is performed on a list of candidate junctions constructed from known transcript models or de novo on the genome reference, the circRNA supporting reads need to span in a reverse order either known splicing junctions or novel ones with apparent splicing signals. cDNA sources could also produce such aberrant junction reads (e.g., chimeric cDNA due to template switching during the RT-PCR process, RNA trans-splicing, or genome rearrangement). A recent evaluation of five different circRNA prediction algorithms found dramatic differences between their outputs, emphasizing that the prediction results should be handled with care [35]. Therefore, for any candidate circRNAs identified, further independent experimental validation is always required [33]. For example, at the global level, because circRNAs do not have a poly(A) tail, their representation should be depleted in a poly(A)-enriched sequencing library. Thus, fewer candidate circRNA junction reads obtained from poly(A) RNA sequencing compared with RNA-depleted total RNA sequencing (of the same biological sample) serves as an indirect form of validation. Similarly, compared with linear RNAs, with few exceptions, circRNAs are generally endowed with a strong resistance to exonuclease RNase R [36]. Therefore, candidate circRNAs should be more enriched in the sequencing libraries prepared from RNA samples treated with RNase R [16]. For individual candidate circRNAs, more extensive validation can include the use of qPCR with ‘divergent’ primer pairs and Northern blot probing for the back-splicing junction sequences, probed with or without RNA linearization using RNase H, and with or without enrichment using RNase R [33]. More recently, as another independent validation approach, the divergent reverse transcription (RT)-PCR products were deep-sequenced for a subset of candidate circRNAs using the PacBio SMRT platform; here, the long reads allowed the authors to observe the cDNAs corresponding to the rolling circle RT products. The fact that these rolling circle RT sequences could only be observed for circRNAs, but not for any linear RNAs, serves as direct evidence for the circular nature of the circRNA structure [27]. Based on the sequences, circRNA isoforms were identified with the same head-to-tail junctions, but different internal exon composition. Together with other reports that multiple circRNAs with different junctions can form from the same gene loci [22,31] and that exons can be circularized with retained introns [13], these data demonstrate additional layers of circRNA diversity.

CircRNA Biogenesis
CircRNAs are likely formed by back-splicing. Similar to canonical splicing, back-splicing requires both a canonical splicing signal and the canonical spliceosome machinery [34,37], but is less efficient and largely occurs post-transcriptionally [38]. The biogenesis of circRNAs can be regulated by cis-elements and trans-factors (Figure 1). It has been demonstrated that both cis-elements and trans-factors can promote circRNA biogenesis by bringing the downstream donor and upstream acceptor sites into close proximity. Regarding cis-elements, it has been demonstrated that inverted Alu repeats are enriched in the introns flanking the circularized exons and nonrepetitive but complementary sequences can also serve to promote circularization both in vivo and in cell culture [16,30,31,39]. Concerning trans-factors, so far, at least three RNA-binding proteins (RBPs) have been reported to serve a regulatory role in circRNA biogenesis. For example, a known splicing factor, Muscleblind, can facilitate the circularization of its own second exon by binding to both of the flanking introns [24]. Similarly, another RBP, Quaking, known to have a role in pre-mRNA splicing, was shown to enhance the formation of many circRNAs by binding at its recognition elements within upstream and downstream introns [40]. Via a different mechanism, ADAR1, a double-strand RNA-editing enzyme, can negatively affect circRNA biogenesis [39]. It has been suggested that the A-to-I editing in the double-strand RNA region mediated by ADAR1 decreases the RNA-pairing potential across the flanking introns, thereby suppressing the back-splicing required for circRNA formation.
CircRNA Expression in Brain

So far, thousands of circRNAs have been discovered in eukaryotic cells. Their expression spans a broad dynamic range, from less than one copy to greater than several hundred copies per cell [15–17,23]. Compared with protein-coding mRNAs, the expression of circRNAs is more skewed towards the lower end, with most exhibiting low abundance. Although some circRNAs are more ubiquitously expressed, most of them show a complex tissue and/or cell type- and developmental stage-specific expression pattern [15,17,32].

Several recent studies have shown that circRNAs are more enriched in neuronal tissues compared with other tissues [24–27]. Deep-sequencing of transcript from mouse brain, liver, heart, lung, and testis showed that, in brain, a significantly greater fraction of reads are circRNA junctional reads, a greater number of genes are hosts to circRNAs, and there is a greater number of tissue-specific circRNA hosts [27]. Indeed, 20% of the protein-coding genes in brain produce circRNAs. Strikingly similar results were obtained by Rybak-Wolf et al. in both mouse and human tissues [26]. In addition, some circRNAs exhibit differential expression in different brain areas, including the striatum, prefrontal cortex, olfactory cortex, cerebellum, and hippocampus [26]. A gene ontology analysis of the transcripts from which circRNAs are derived indicates one potential reason for the circRNA enrichment in brain: synaptic genes, including those encoding pre- and postsynaptic functional groups, are significantly enriched as circRNA host genes [27]. In addition, many circRNAs are enriched in synaptic fractions, including the neuropil and synaptosomes [26,27]. High-resolution in situ hybridization using probes that target the unique circRNA junctions indicate the clear and unequivocal localization of circRNAs in both the cell body and dendrites of cultured hippocampal neurons and hippocampal slices (Figure 2 and [27]).

Figure 2. Brain-expressed Circular (Circ)RNAs Are Enriched in the Synaptic Neuropil. High-resolution in situ hybridization experiments have been conducted in cultured hippocampal neurons or hippocampal slices using probe sets designed to detect the indicated circRNA (green) (A–F). In each case, many circRNA-positive particles are apparent not only in the cell bodies (nuclei stained with DAPI, blue), but also in the dendritic processes, detected using an anti-MAP2 antibody (red). A control (exon) probe designed to detect noncontiguous regions of two exons that could not form head-to-tail junction (see Methods) yielded just a few background particles (A). Scale bars = 20 (A), 50 (C), and 75 (F) μm.
Why Are CircRNAs Abundant in Brain?
Given that circRNA biogenesis can be regulated by cis-elements and trans-factors as discussed above, the higher abundance of circRNAs in brain might be attributed to this regulation. Indeed, as indicated above, many host genes that produce circular RNAs are expressed exclusively in brain, but not other tissues. In addition, neuronal genes often have long introns and it is known that circularized exons are more frequently flanked by longer introns [16,30,31,39]. Therefore, it is conceivable that the brain-specific genes may carry more sequence features that promote circRNA formation. Furthermore, on average, when a linear gene is expressed in brain as well as other tissue, more circRNAs are produced from the brain linear transcript [27]. Given that the same sequence is expressed in both brain and other tissues, the enhanced circRNA biogenesis observed in brain could, in principle, be attributed to the brain-biased expression of splicing factors and/or other RBPs that regulate circRNA biogenesis. However, the three currently known regulatory proteins (described above) do not fit this bill; in particular, the negative regulator of circRNA expression ADAR1 is known to have high activity in brain [41]. However, with more trans-acting factors to be discovered, it is tempting to speculate that many of those with biased expression in brain may act in concert and/or competition to promote circRNA formation. In addition to biogenesis, the high stability of circRNA also contributes to their accumulation in brain tissues comprising quiescent, post-mitotic neurons [38].

Regulation of CircRNA Expression by Neuronal Development and Plasticity
Several studies have shown that circRNA expression is regulated by various aspects of neuronal development. For example, the differentiation of neurons from cultured undifferentiated cells was associated with enhanced expression of a large population of circRNAs, while a smaller population exhibited decreased expression [26,38]. An analysis of circRNA expression in developing cultured hippocampal neurons (days E18, P1, P10, and P30) revealed a rather abrupt increase in circRNA levels at the P10–P30 transition, corresponding to the time of synapse formation. Those circRNAs that were upregulated at the P10–P30 transition were again most likely to derive from host transcripts that participate in synaptic function [27]. Whether the circRNA abundance is a simple function of the host transcript abundance also needs to be evaluated. However, during neuronal development, all possible relations between the circRNA and host transcript regulation exist [26,27]. In many cases, the circRNA regulation was independent or even opposite to that observed for the linear transcript, suggesting that, at least in some cases, the regulation of the circRNA is not a secondary consequence of the host transcript levels. Moreover, circRNAs are upregulated in neural tissues in an age-dependent manner in Drosophila [25]. Interestingly, circRNA expression also exhibits significant regulation by the expression of synaptic plasticity. Induction of homeostatic scaling in cultured neurons results in a significant alteration of many circRNAs, detected by both RNA-seq and high-resolution in situ hybridization [27].

CircRNA Function(s)
Several lines of evidence suggest that circRNAs have important regulatory functions. First, although most circRNAs are of low abundance, a significant population (10–100 or so, depending on the cell type) is expressed at a reasonable level; in many cases, the abundance of the circRNA exceeds that of the associated linear RNA isoform [14–16,23,26,27]. Second, the expression of circRNAs is often regulated in a cell type- and stage-specific manner [17]. As mentioned above, several studies have demonstrated the dynamic expression of circRNAs during development, neuronal differentiation, and upon neuronal stimulation and plasticity [26,27,32]. Often, the circRNA expression is regulated independent of its host gene, arguing against the idea that circRNAs are a mere by-product of canonical pre-mRNA splicing. Last, but not least, cross-species analyses have demonstrated the evolutionary conservation of circRNA expression. An analysis of circRNAs localized in rat and mouse hippocampus dendritic layers...
found that 23.6% of the circRNAs identified in mouse neuropil were also expressed in rat neuropil [27]. Another study showed that, of 15849 circRNAs expressed in mouse brain, 4522 were conserved in humans, with nearly identical splicing sites. The conservation of some circRNAs could even be traced back to Drosophila [26]. In addition to the expression, the sequences within circRNA were more conserved than the flanking exons. Compared with splicing sites from the same host genes that are not involved in forming head-to-tail junctions, the exonic sequences around head-to-tail junctions were more conserved [27]. Moreover, those sequences around common head-to-tail junctions detected in both mouse and rat were even more conserved, almost reaching the maximum PhastCons score, strongly suggesting their potential functional relevance.

What then are the known functions of circRNAs? So far, a clear function has only been demonstrated for a few circRNAs that can serve as miRNA ‘sponges’ [15,28,42], sequestering miRNAs and preventing their interactions with target mRNAs. However, bioinformatic analyses suggested these circRNAs are exceptional cases and few other circRNAs contain enough miRNA-binding sites to function as strong sponges [23,27]. In addition to miRNA regulation, circular RNAs could sequester RBPs and thereby could also regulate the intracellular store and transport of associated RBPs and/or mRNAs [18,33]. It has been shown, for example, that the circRNA derived from Muscleblind gene can bind the Muscleblind protein [24]. Given that the Muscleblind protein also enhanced the production of the circRNA, an inter-regulatory feedback loop potentially exists between the two, in which the Muscleblind-protein-enhanced production of circRNAs in turn would sponge and thereby mask the effect of protein. As another example, ectopically expressed circRNA derived from the Foxo3 gene (circ-Foxo3) can bind to the cell cycle proteins cyclin-dependent kinase 2 (CDK1) and cyclin-dependent kinase inhibitor 1 (p21). The formation of this circ-Foxo3-p21-CDK2 ternary complex arrested the function of CDK2 and blocked cell cycle progression [43]. So far, the binding sites have been experimentally determined for a limited set of RBPs. Based on binding sites for 38 RBPs predicted using binding sequence motifs deposited in the database RBPDB, the circRNAs expressed in mouse brain harbor fewer RBP binding sites and a lower RBP binding density compared with either the coding sequence or the 3’ untranslated region (UTR) of protein-coding genes, indicating that circRNAs, as a group, are no more likely to bind to RBPs than are linear mRNAs [27]. Thus, whether the RBP sponge is a predominant function of circRNAs awaits more targeted experimental studies.

CircRNAs mostly comprise protein-coding exons and many have reasonable open reading frames. In addition, synthetic circRNAs that contain internal ribosome entry sites can be translated in vitro [44,45]. However, evidence for the translation of endogenous circRNAs is lacking. Furthermore, based on polysome profiling data, it was shown that circRNAs are enriched in the nonribosomal RNA fraction and strongly depleted in the polysome-bound fractions, indicating that endogenous circRNAs, as a group, are unlikely to be translated into peptides [27].

Together, these results suggested that circRNAs represent a heterogeneous group of transcripts that affect cellular function via as yet undiscovered and likely diverse mechanisms. In addition to the potential function of circRNA transcripts as trans-regulators, the production of circRNA could also regulate the transcriptional output of its host gene in cis by competing with the canonical splicing [24]. Moreover, two examples from a specific class of circRNAs with retained introns (ELciRNAs) have been shown to predominately localize in the nucleus, bind to U1 snRNP and thereby promote the expression of their host genes [13]. Finally, Dong et al. recently demonstrated that circRNAs can be retrotranscribed and ultimately inserted back into the host genome as processed pseudogenes, which might hold the potential to change the genome structure and thereby regulate gene expression [46].
CircRNA Stability: A Clue to Function?

Compared with linear RNAs, circRNAs are more stable, likely due to their resistance to RNA exonucleases [15,16,20]. This high stability suggests that the apparent concentration of many circRNAs is dominated by their slow turnover rather than by their production [38]. Therefore, in quiescent and postmitotic cells, such as neurons, circRNAs could accumulate, resulting in higher concentrations than the linear RNAs even though the relative rate of their production remains constant. This fits with observations that circRNAs accumulate during human fetal development, neuronal differentiation, and synaptic development, and in the aging fly brain [25–27,32]. The high RNase resistance of circRNAs endows them with high stability in blood and other body fluids, thereby making them potential disease biomarkers for various diseases. Therefore, it is tempting to speculate that the relative longevity of circRNAs may be a clue to their, thus far, largely elusive function. In neurons, stable molecules are often posited as potential repositories of information storage. The long-lived nature of circRNAs as well as their localization in dendrites, near synapses, could allow them to bear modifications or harbor proteins or RNAs that represent the local history of the synapse. As such, the presence or accumulation of circRNAs in situ could represent a marker for synapses that share histories.

Concluding Remarks

CircRNAs have recently regained attention as noncoding RNA molecules with a potential regulatory function. CircRNAs are highly abundant in brain and are often derived from genes specific for neuronal and synaptic function. CircRNA expression is regulated during neuronal development and by synaptic plasticity, and often such regulation is independent of that of the host linear transcripts. With few exceptions, the function of most circRNAs remains elusive and, as a heterogeneous group, they might regulate different aspects of cellular functions via as yet undiscovered and likely diverse mechanisms (see Outstanding Questions).

Acknowledgments

We thank members of the Chen and Schuman labs for data, analyses, and discussion. We specifically thank Mantian Wang for Figures 1 and 2.

References

8. Cocquerelle, C. et al. (1990) Splicing with inverted order of exons occurs proximal to large introns. EMBO J. 11, 1095–1098

Outstanding Questions

How are circRNAs generated in the brain? Their prevalence in brain tissue implies that brain-specific or enriched splicing factors play an important role. Is there a common biogenesis pathway for all circRNAs?

How is circRNA expression regulated in the brain? A few studies have indicated that circRNAs are regulated by synaptic development and homeostatic plasticity, but specific regulation by discrete molecular or cellular events is missing.

How do circRNAs function in the brain? This is the biggest outstanding question in the entire field. A few functions have been proposed so far, including serving as a sponge or platform for miRNAs or RBPs, but bioinformatic analysis of brain circRNAs indicates that these cases may represent exceptions and there must be additional, as yet undiscovered, functions.

 Might circRNAs serve as ‘information carriers’, given their long lifetime? circRNAs are resistant to cleavage, making them very long-lived relative to other RNA molecules. Might long-lived molecules, such as circRNAs, serve as memory encoders?