

PRIMER

Circular RNAs: analysis, expression and potential functions

Steven P. Barrett¹ and Julia Salzman^{1,2,*}

ABSTRACT

Just a few years ago, it had been assumed that the dominant RNA isoforms produced from eukaryotic genes were variants of messenger RNA, functioning as intermediates in gene expression. In early 2012, however, a surprising discovery was made: circular RNA (circRNA) was shown to be a transcriptional product in thousands of human and mouse genes and in hundreds of cases constituted the dominant RNA isoform. Subsequent studies revealed that the expression of circRNAs is developmentally regulated, tissue and cell-type specific, and shared across the eukaryotic tree of life. These features suggest important functions for these molecules. Here, we describe major advances in the field of circRNA biology, focusing on the regulation of and functional roles played by these molecules.

KEY WORDS: CircRNA, Splicing, Regulation, Development, Non-coding RNA

Introduction

The past decades have seen an ever-growing list of diverse non-coding RNA species with functional capacity expressed in eukaryotic cells (Morris and Mattick, 2014). With the advent of next-generation sequencing, the catalog has grown more rapidly. Circular RNAs (circRNAs), the 3' and 5' ends of which are covalently linked, constitute a class of RNA recently discovered to be widespread and abundant (Salzman et al., 2012). CircRNAs are generally formed by alternative splicing of pre-mRNA (Fig. 1), in which an upstream splice acceptor is joined to a downstream splice donor in a process known as 'backsplicing' (Barrett et al., 2015; Schindewolf et al., 1996; Starke et al., 2015). Although there is still no consensus as to the function of circRNAs, a number of studies have revealed that circRNAs are expressed in a variety of eukaryotic organisms, demonstrate conservation across mammals, and are expressed in a regulated manner independent of their cognate linear isoforms. In this Primer, we outline progress in the circRNA field, describing historical examples of circRNAs and their discovery, methods of computational and experimental detection, putative and established functional roles for circRNAs, and the mounting evidence for circRNA regulation during mammalian development.

The discovery of circRNAs

The first examples of spliced circRNAs were serendipitously discovered following analyses of the human *DCC* gene, from which four potential circular isoforms were identified in an experiment that originally aimed to determine exon connectivity by RT-PCR amplification and sequencing (Nigro et al., 1991). Another early and

carefully studied example was the circRNA derived from the mouse *Sry* gene. In the developing genital ridge, the *Sry* gene produces a linear mRNA and its protein product plays a key role in sex determination during embryonic development. However, in adult mice testes, *Sry* is expressed exclusively as a 1.23-kb circular isoform (Capel et al., 1993). Despite extensive studies, no evidence of translation was found, leading researchers to question the functional significance of this isoform.

In later years, chance discoveries of low abundance circRNAs resulted in reports of a handful of other circRNA-producing genes, including the human *ETS-1* (*ETS1*) gene (Cocquerelle et al., 1992), human and rat cytochrome P450 genes (Zaphiropoulos, 1996, 1997), the rat androgen binding protein gene (*Shbg*) (Zaphiropoulos, 1997) and the human dystrophin gene (Surono et al., 1999). In addition, the *NCX1* gene in monkeys (Li and Lytton, 1999) and the *Drosophila muscleblind* (*mbl*) gene (Houseley et al., 2006) were posited to have highly abundant and regulated 'anomalous' RNA isoforms compatible with circRNA expression. Much more recently, a human non-coding RNA named ANRIL was found to have a low abundance transcript capable of circularizing (Burd et al., 2010), and it was also shown that the antisense transcript to the *CDRI* locus generated a highly abundant circular isoform that is now extensively studied (Hansen et al., 2011). But despite these few examples, researchers perceived circRNAs as rare, one-off occurrences with limited biological significance (Cocquerelle et al., 1993).

In 2012, a sea change occurred: in an attempt to find genomic rearrangements in cancers, the global expression of circRNAs in human pediatric acute lymphoblastic leukemia RNA-seq samples was serendipitously discovered, and it was shown that this phenomenon extended to leukocytes from healthy adults as well as to several other cancer and non-cancer cell lines and the mouse brain (Salzman et al., 2012). Two other bioinformatic analyses by other groups reaffirmed and extended many of these findings (Jeck et al., 2013; Memczak et al., 2013) and, together, these studies provided the foundation for a burgeoning new field.

The key findings from these publications are as follows: circRNAs are abundant – they are expressed in thousands of human genes and in some cases they demonstrate higher expression than their cognate linear isoforms (Jeck et al., 2013; Memczak et al., 2013; Salzman et al., 2012); circRNAs exhibit cell type-specific expression (Salzman et al., 2013); circRNAs demonstrate conservation between mouse and human (Jeck et al., 2013; Memczak et al., 2013); circRNAs are localized to the cytoplasm (Salzman et al., 2012) and are remarkably stable, with half-lives exceeding 48 h (Jeck et al., 2013); natural circRNAs do not appear to be translated (Guo et al., 2014; Jeck et al., 2013), although research has established that manufactured circRNAs designed with an internal ribosome entry site (IRES) can be translated *in vitro* and *in vivo* (Chen and Sarnow, 1995; Perriman and Ares, 1998); circRNAs are generally formed from longer-than-average exons and are normally flanked by longer-than-average introns in their associated pre-mRNAs (Jeck et al., 2013; Salzman et al., 2012),

¹Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, USA. ²Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94305, USA.

*Author for correspondence (Julia.salzman@stanford.edu)

 S.P.B., 0000-0002-7120-8365

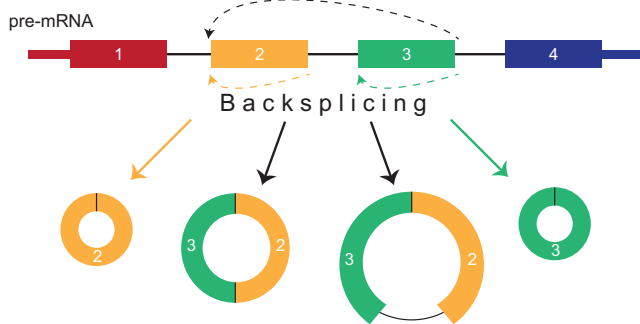


Fig. 1. circRNA splicing and isoform diversity. A great diversity of circular RNAs (circRNAs) can be generated from a single genomic locus. These circRNAs are created through a non-canonical splicing process known as 'backsplicing' in which a downstream splice donor is joined to an upstream splice acceptor. Such circRNAs can consist of one or more exons and can even contain unspliced intronic sequences. Colored bars, exons; black lines, introns.

which are enriched for complementary ALU elements thought to play a role in the biogenesis of many circRNAs in humans (Jeck et al., 2013). It should be noted that all of these findings are generalities, and exceptions can be found in almost every case.

The identification and characterization of circRNAs

Since the initial discovery of circRNAs, various biochemical tools have been developed to test for circularity, validate the existence of particular circRNAs, and ectopically express circRNAs. In addition, bioinformatic and statistical approaches have been developed to quantify the expression of circRNAs and identify new circRNAs with high confidence.

Biochemical methods for circRNA characterization

Biochemical experimental methods are important tools for validating and identifying circRNAs. Among the most basic tools to validate a circRNA is reverse transcription-PCR (RT-PCR). In this approach, the circRNA is converted to a cDNA via reverse transcription (RT). Because the cDNA is derived from a circRNA, the sequence contains a diagnostic exon-exon junctional sequence that is absent in the canonically spliced mRNA, and primers can thus be designed to specifically amplify and detect this diagnostic junction. These primers are known as 'inverse' or 'outward-facing' primers because, when aligned to the genome, their 3' ends face away from each other, unlike in standard PCR. This prevents the amplification of species that do not contain the diagnostic junction, such as mRNA or genomic DNA. Quantitative RT-PCR (qRT-PCR) can be used to assess quickly the relative abundance of the circRNA across a panel of samples.

Although RT-PCR is very simple and powerful, it also has biases and artifacts. Template switching during the RT step (Fig. 2A) can lead to a variety of artifactual templates that could give rise to PCR products even with inverse primers (Cocquet et al., 2006; Luo and Taylor, 1990; Roy et al., 2015). Thus, it is imperative that the PCR products are sequenced to verify the diagnostic scrambled junction. Even after this verification, it is possible that template switching (Kulpa et al., 1997), trans-splicing (splicing between two separate pre-mRNA molecules) (Agabian, 1990), or unexpected genomic duplication events might have produced a product precisely at the junctional boundary, thus yielding a linear transcript with the scrambled junctional sequence (Fig. 2A). Furthermore, as circRNAs have no defined end, it should be noted that strand-displacing RT enzymes can create long cDNAs consisting of

concatemers of exons (Fig. 2B). This phenomenon, known as rolling-circle RT, can lead to a laddering appearance on a gel after PCR (Barrett et al., 2015; You et al., 2015). This can be strong evidence for circularization but can also lead to an overestimation of circRNA by qPCR and sometimes by RNA-seq, depending on the protocol used for library preparation.

Northern blotting is also a simple and effective way to detect circRNA (Capel et al., 1993). Probes are designed to target the circularized exonic sequence and can even specifically target the diagnostic junctional sequence. To ensure specificity is not compromised, multiple probes targeting the circRNA can be used in separate blots. Northern blots have the advantage that the species being monitored is known precisely, as it has a particular mobility, whereas in PCR the monitored species need only contain the diagnostic sequence, which may arise from one or more sources of RNA or DNA (Fig. 2A).

RNase R, an exoribonuclease capable of processively degrading RNA from its 3' to 5' end, is a useful tool, often used in combination with those described above, to assess whether RNA is indeed circular and to enrich for circRNA in sequencing libraries (Suzuki et al., 2006; Vincent and Deutscher, 2006). However, the use of RNase R can introduce technical noise, as it can decay some circles, perhaps owing to endogenous nicking of the RNA or contaminating nucleases, and some linear molecules are resistant to digestion if their 3' ends are involved in base pairing (Vincent and Deutscher, 2006). Also, because this enzyme is not completely efficient, some fraction of linear RNA will remain even after treatment. For this reason, it is important to use a quantitative metric such as qRT-PCR to assess its efficacy, as other methods such as endpoint PCR might lead to signal saturation even after RNase treatment. A much less frequently used enzyme, XRNI, a 5'-to-3' exoribonuclease, can be used in a similar manner as RNase R, but mRNA must be decapped with tobacco acid pyrophosphatase (TAP) prior to treatment. Despite the imperfections described above, exonuclease resistance along with quantitative measures of resistance to the nuclease (e.g. qRT-PCR or northern blot) provides compelling evidence for circularity.

Another strong assay for circularity involves RNase H, an endoribonuclease that can cleave RNA at RNA-DNA hybrids. In this assay, two short DNA probes are annealed to the RNA of interest, and the RNA is cleaved by RNase H in both locations. When the RNA is run on a gel and blotted for the fragments, two bands should be observed if the RNA is circular or three if the RNA is linear (Capel et al., 1993). When RNA is cleaved with RNase H in the presence of each DNA probe separately, one band will be observed for a circular species and two for a linear. In addition, a change in migration can be observed for the circular species after it is linearized by a single cleavage event (Starke et al., 2015) because linear and circular species do not exhibit the same migration through a polyacrylamide gel matrix.

Two-dimensional denaturing polyacrylamide gel electrophoresis can also be used to distinguish between circRNAs and linear RNAs. In this technique, total RNA is run on a 2D gel containing different percentages of polyacrylamide in each dimension. Linear RNA migrates along the diagonal of the 2D gel according to its size, whereas circRNA travels in an arc owing to its anomalous migration (Tabak et al., 1988). This arc can be excised and sequenced using next generation sequencing (NGS) as an enrichment step (Awan et al., 2013). Alternatively, the 2D gel can be probed (via northern blotting) to quantify or identify particular circRNAs.

In addition to exoribonuclease treatment and 2D gels, ribosomal RNA (rRNA) depletion and polyA-depletion are common methods

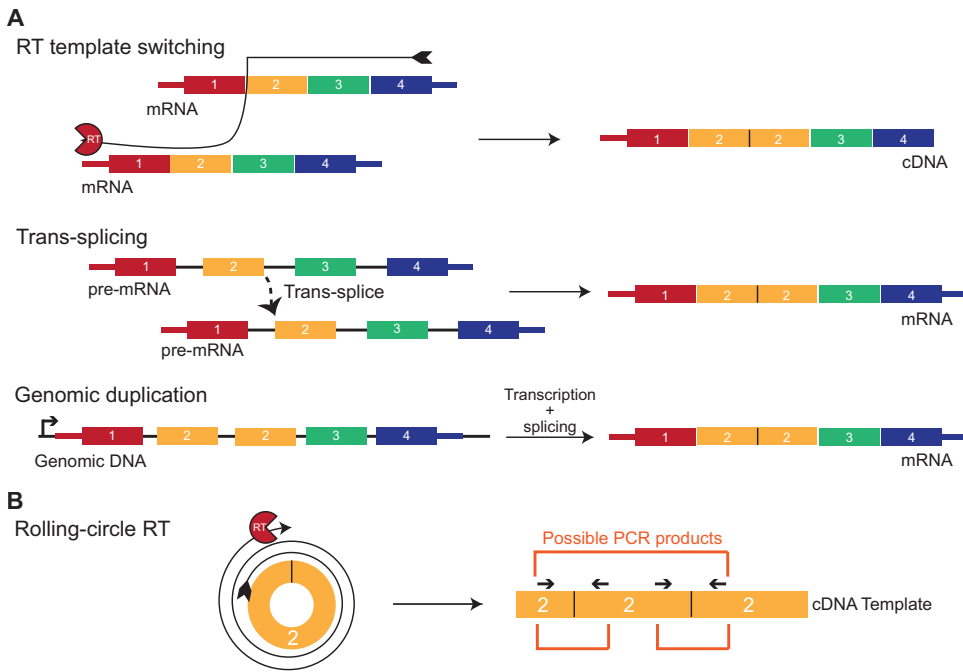


Fig. 2. circRNA detection: sources of backsplice junctional sequences.

Backsplice junctional sequences can arise from a variety of sources, resulting in false-positive detection and improper quantification of circRNAs. (A) Reverse transcriptase (RT) template switching, trans-splicing, and genomic duplication events can all give rise to linear products with spurious backsplice junctional sequences. These products can be amplified by PCR but are not derived from circRNAs. (B) Rolling-circle RT occurs when a short circRNA is converted to a long cDNA, leading to a concatamer of PCR-amplifiable sequences. Because the primers can bind to this concatameric sequence in several locations, many PCR products can be created from a single cDNA template.

used to enrich for circRNAs in sequencing libraries. Neither guarantees that the enriched sequences are exclusively circular, as many types of non-coding RNA will also survive these selections. Nonetheless, these biochemical methods can be combined with RNA-seq to assess genome-wide circRNA expression and, in principle, to distinguish artifactual circRNA expression from truly expressed circRNA. For example, intuitively, a comparison of circRNA detection rates in RNase R-treated samples versus ribosomal-depleted samples should provide evidence for the detection of a true circRNA. If such analyses are performed, it is important to normalize appropriately the RNA-seq datasets being compared (Salzman et al., 2011). However, biochemical treatment/purification before RNA-seq does not serve as a sufficient gold standard for genome-wide circRNA identification. For example, Jeck et al. reported depletion of the well-known circle derived from the *CDR1* antisense transcript after RNase R treatment (Jeck et al., 2013).

Ectopic expression of circRNAs: gaining insights into function and biogenesis

No methods currently exist to interrupt or induce expression of a specific circRNA from its endogenous locus. Thus, the function (Li et al., 2015; Memczak et al., 2013) and biogenesis (Ashwal-Fluss et al., 2014; Barrett et al., 2015; Kramer et al., 2015; Liang and Wilusz, 2014; Starke et al., 2015) of circRNAs is often studied using easily manipulable circRNA overexpression plasmids. In some cases, entire genes can be expressed on a plasmid (Barrett et al., 2015), although this is often not possible for mammalian genes. To circumvent this issue, mammalian vectors normally only contain the circularized exon(s) along with flanking splicing signals and intronic sequences, which harbor inverted repeats to facilitate their splicing into a circle (Fig. 3A) (Ashwal-Fluss et al., 2014; Hansen et al., 2013; Kramer et al., 2015; Li et al., 2015; Liang and Wilusz, 2014; Starke et al., 2015).

However, a common technical artifact arising from the use of these vectors occurs as a result of rolling circle transcription of the plasmid (Fig. 3B). For illustration, suppose the plasmid contains the circularized exon of a gene with flanking introns. If the transcription

termination signals in the vector are bypassed, the RNA polymerase will continue to transcribe around the entire plasmid, generating a concatamer of the RNA sequence contained in the plasmid. This results in a number of undesired transcripts, including a pre-mRNA containing tandem repeats of the circularized exon, which can then be spliced canonically, pairing upstream donors to downstream acceptors, and yielding a linear concatamer of the circularized exon. Importantly, this RNA will contain a scrambled junction and might appear identical to a bona fide circRNA depending on the experimental approach used for its detection. These artifactual products can lead to off-target effects on the cell and spurious circRNA quantification, especially by qRT-PCR. In some cases, care has been taken in vector design to minimize the expression of erroneous products (Kramer et al., 2015). However, for some applications, such as establishing translation or function of a circRNA, these artifacts must be completely eliminated.

Bioinformatic and statistical identification of circular RNAs

Although a number of experimental approaches (discussed above) can be used to identify circRNAs, more recent approaches have aimed to predict or detect the expression of circRNAs via computational methods. Whereas poly-A⁺ RNA-seq libraries are frequently employed for mRNA transcriptome profiling experiments, ribosomal RNA-depleted or total RNA libraries are most commonly used for circRNA profiling. The detection of circRNA in RNA-seq datasets is achieved by specifically searching for reads that are chimeric in the sense that the 5' sequence in the read is downstream of the 3' sequence with respect to transcription (Fig. 4). In addition, a variety of algorithms have been designed and are now available to detect circRNA. A recent article compared the performance of several such published algorithms, finding dramatic differences between sensitivity and specificity (Hansen et al., 2015). This study clearly demonstrates that the choice of algorithm used for the genome-wide detection of circRNA significantly impacts false positive and false negative rates of detection, and gold standards for assessing algorithm performance are needed. A number of computational and statistical considerations must also be taken into account during circRNA detection and quantification (Szabo

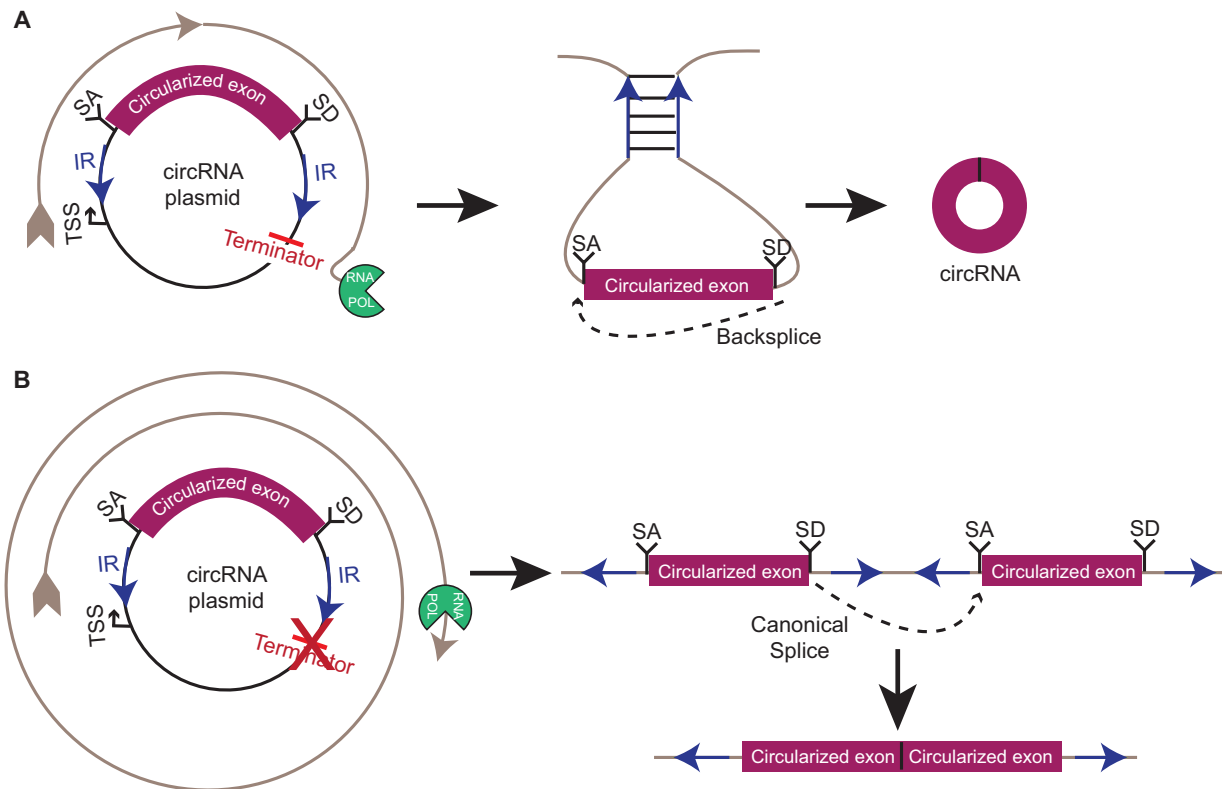


Fig. 3. circRNA overexpression vectors. circRNA overexpression vectors are often used as convenient tools for studying the function and biogenesis of circRNAs. (A) circRNAs are commonly overexpressed on plasmids using gene fragments under the control of a strong promoter (e.g. CMV in mammalian cell lines). Circularization is induced by inverted repeats (IR) flanking the circularized exon, which presumably bring the splice acceptor (SA) and splice donor (SD) into close proximity for backsplicing. (B) In some cases, RNA polymerase (RNA POL) can bypass termination signals, resulting in rolling-circle plasmid transcription. This can give rise to a linear transcript capable of being canonically spliced, yielding a spurious backsplice junctional sequence. TSS, transcription start site.

et al., 2015). As such, and given the homology between and within genes in almost all genomes, which can lead to read misalignments (Szabo et al., 2015), the highly precise and sensitive detection of RNA splicing into linear or circular molecules remains a significant challenge in the field (Engström et al., 2013; Peng et al., 2015). A user-friendly, searchable database of identified circRNAs is curated by the Rajewsky lab and can be found at circbase.org (Glažar et al., 2014).

The expression of circRNAs

Since the earliest examples of circRNAs derived from the *DCC* and *Sry* genes, there has been accumulating circumstantial evidence that circRNAs might play functional roles in eukaryotes. As we discuss below, circRNAs have been detected in a variety of developmental contexts and cell types, and their expression appears to be dynamic, indicating some level of regulation.

Cell-type and tissue specificity of circRNA expression

A number of circRNAs are expressed in a tissue-dependent manner. In the case of the *DCC* gene, researchers found that the expression of the circular isoform varied across a variety of human tissues and was not correlated with the expression of its cognate linear mRNA; in some cases, no circRNA was detected despite high levels of mRNA expression (Nigro et al., 1991). Recent studies report similar findings (Rybak-Wolf et al., 2015; Salzman et al., 2013; Szabo et al., 2015; You et al., 2015), demonstrating that this apparent regulation of circularization is a widespread phenomenon. Other studies also demonstrate that global levels of circRNA and mRNA do not correlate and that the diversity of circular isoforms from a

particular gene can vary across a panel of cell types (Salzman et al., 2013). Several groups have reported generally high levels of circRNA in the brains of pigs, humans and mice, with especially high levels in the cortex and cerebellum (Rybak-Wolf et al., 2015; Szabo et al., 2015; Venø et al., 2015; You et al., 2015). Furthermore, circRNAs expressed in fly heads or mouse brains are enriched in genes that code for neuronal proteins and synaptic factors, suggesting a potential role for circRNA in the central nervous system (Westholm et al., 2014; You et al., 2015).

The apparent regulation of circRNAs appears to be a general phenomenon that is conserved; even in fission yeast, some circRNAs exhibit changes in abundance that are independent of their linear isoform during nitrogen starvation (Wang et al., 2014). Together, these results strongly suggest that the expression of circRNAs is a regulated process.

The conservation of circRNA expression

circRNA expression across mammals also appears to be conserved. That is, the expression of circRNA isoforms from orthologous genes in mammals is greater than expected by chance. Several studies of circRNA conservation have focused on comparing circRNA expression in mice and humans. Estimates of the fraction of mouse circRNAs with human orthologs range widely from less than 5% to nearly 30% (Guo et al., 2014; Jeck et al., 2013; Memczak et al., 2013; Rybak-Wolf et al., 2015). A recent study of circRNAs in the porcine brain calculated that ~15–20% of the circRNAs produced in the mouse brain use splice sites that are orthologous to those used by circRNAs in the pig brain (Venø et al., 2015). Moderate conservation between pig and human brain circRNAs was

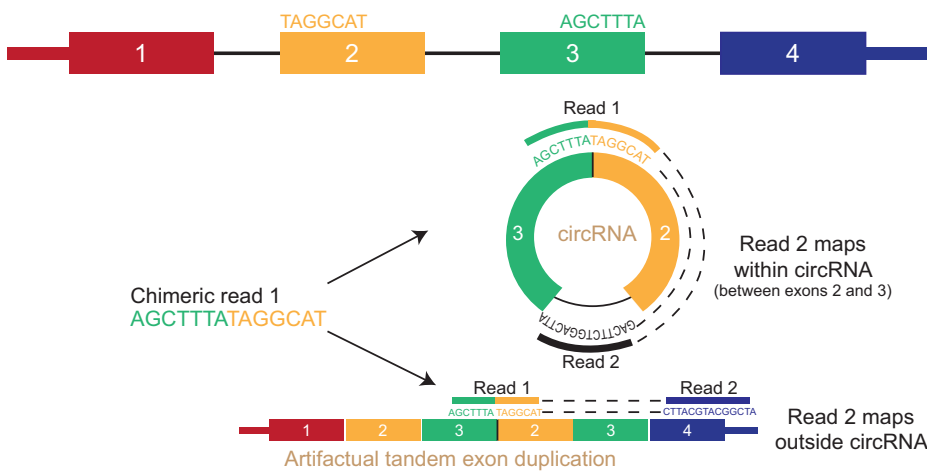


Fig. 4. Detecting circRNAs using bioinformatics. A chimeric read in which the 3' end of the read maps upstream of the 5' end with respect to the direction of transcription is the hallmark of circRNAs in RNA-seq data. In this example, the chimeric read appears to start in exon 3 and end upstream in exon 2. However, the presence of a chimeric read does not guarantee that it was derived from a circRNA (shown here and in Fig. 2A). If paired-end sequencing is employed, read 2 can inform the RNA source of the reads. For example, if read 2 maps outside the bounds of the circRNA (as defined by the chimeric read), then the read is likely to be derived from a linear species, perhaps of artificial origin.

also reported, with 5-10% of human brain circRNAs also being expressed in the pig brain; increased rates of conservation (~15%) emerged when considering only a subset of highly expressed circles. Another study estimated 23% conservation between mouse and rat circRNAs, and reported high conservation of circRNAs at the sequence level surrounding the circRNA scrambled junction compared with other splice site-proximal exonic sequences in the host gene (You et al., 2015).

The large variation between these results probably reflects differences in sequencing depth (as many circles have low expression and may escape detection in samples with low coverage), differences in the statistical definition of conservation, bioinformatic parameters used to analyze the datasets, and, importantly, differences in sources of tissue used for comparison. For example, studies determining rates of circRNA conservation between mouse and human samples stratified by tissue type and tested separately will, not surprisingly, yield different results than those performing the same test after collapsing across all organs (i.e. tests of conservation between any cataloged human or mouse sample). This effect is well-known in statistics and is coined Simpson's paradox.

In addition to their conservation across mammals, circRNAs have been, to our knowledge, detected in every eukaryotic species tested to date (Wang et al., 2014). Notably, circRNAs are expressed in microbial eukaryotes including *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*, two highly diverged yeasts with very limited examples of alternative splicing. CircRNA has also been detected in plants (*Arabidopsis thaliana*), protists (*Plasmodium falciparum* and *Dictyostelium dictyostelium*) and numerous animals from fly to human (Memczak et al., 2013; Salzman et al., 2013; Wang et al., 2014). These findings suggest that either circRNA is an ancient feature of gene expression, or it has independently arisen several times over the course of eukaryotic evolution.

The dynamic expression of circRNAs during development

circRNAs exhibit dynamic global changes in their expression levels during development. Studies in humans and flies, for example, have found a general induction in circRNA expression during embryonic development (Szabo et al., 2015; Westholm et al., 2014). In humans, induction was observed across a variety of tissues and was consistently observed for circRNAs spliced by both the major (U2) and minor (U12) spliceosome (Szabo et al., 2015). An interesting developmentally regulated circRNA is the conserved circRNA derived from the second exon of *NCX1* (*SLC8A1* – Human Gene

Nomenclature Database). In humans, this circRNA has the highest level of expression and induction of any circRNA expressed during human fetal development. Expressed primarily in the heart, the *NCX1* gene encodes a sodium/calcium exchanger responsible for transporting calcium out of the cardiomyocyte after contraction (Jordan et al., 2010). The developmental induction of *NCX1* circRNA expression was also recapitulated *in vitro* with experiments differentiating human embryonic stem cells to cardiomyocytes over the course of 2 weeks (Szabo et al., 2015). A putative function has yet to be ascribed to the *NCX1* circRNA isoform.

Many studies have focused specifically on neuronal development because of the high levels of circRNA expression in the brains of flies, mice, pigs and humans (Szabo et al., 2015; Venø et al., 2015; Westholm et al., 2014; You et al., 2015). In the most detailed study to date, You et al. found that circRNAs expressed in mouse brains were enriched in synapses compared with whole hippocampal homogenate and were largely derived from genes associated with synaptic function (You et al., 2015). Analysis of mouse brains from embryonic day (E) 18 to postnatal day (P) 30 revealed that the largest changes in circRNA abundance, both increases and decreases, occurred near the time of synapse formation (P10). Moreover, RNA fluorescence *in situ* hybridization (FISH) of circRNAs demonstrated localization to the cell body and dendrite, much like mRNAs (Cajigas et al., 2012) and regulatory RNAs such as microRNAs (miRNAs) (Tai and Schuman, 2006). A similar study focusing on pig embryonic brain development found peak circRNA expression at E60, which corresponds to a period of development with high levels of neurogenesis (Venø et al., 2015). Moreover, the circRNAs upregulated at this time point were associated with genes involved in Wnt signaling, axon guidance and TGFβ signaling. In flies, circRNAs were found to accumulate with age in the head and were also highly enriched in genes relating to development and signaling, neurogenesis, and neuronal morphology and function (Westholm et al., 2014).

In vitro experiments using both mouse and human cell lines (P19 cells and SH-SY-5Y cells, respectively) showed global increases in circRNA expression after differentiation to neurons using retinoic acid (Rybak-Wolf et al., 2015). Many circles increased with linear mRNA expression but rarely by the same factor, and some circRNAs even showed an inverse relationship with their cognate linear isoform, although how these changes relate to circRNA function during development requires further investigation.

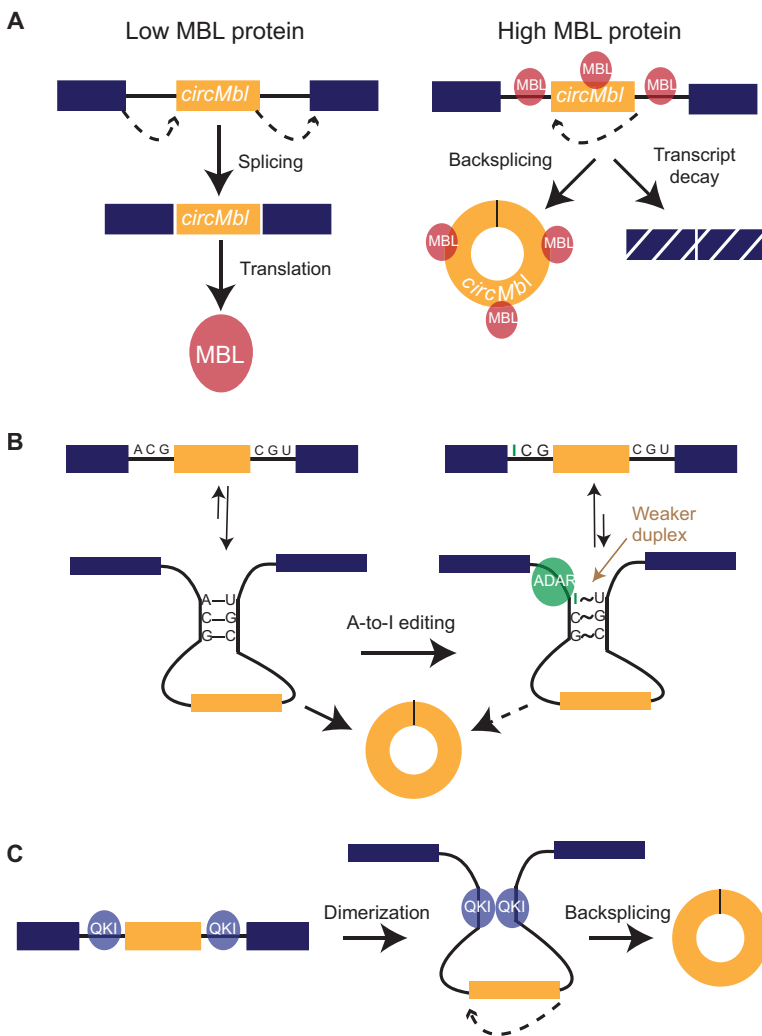


Fig. 5. Regulatory factors involved in circRNA production.

(A) Muscleblind (MBL, red circle) regulates the splicing of its own pre-mRNA into a circRNA (*circMbl*). In the presence of low amounts of MBL (left), the *mbl* transcript is canonically spliced to yield a translatable mRNA encoding the MBL protein. However, when MBL levels are high (right), MBL binds to the pre-mRNA and causes it to splice into a circRNA (*circMbl*), thereby preventing linear splicing and translation of the MBL protein. Furthermore, acting as an RBP sponge, *circMbl* can sequester MBL protein, lowering its free cellular concentration, thereby providing a feedback mechanism to regulate MBL levels. (B) Model for ADAR-mediated regulation of circRNA expression. ADAR family proteins (green circle) can edit and weaken RNA duplexes, decreasing the likelihood of circularization. (C) Model for QKI-mediated regulation of circRNA biogenesis. QKI (blue circle) binds to each intron flanking a circRNA and dimerizes to form a looped structure containing the exon(s). This looped structure promotes circularization of the exon(s).

Two key examples of highly expressed, neural-specific and conserved circRNAs that exhibit developmentally regulated expression are the circRNAs derived from the *RIMS2* gene (*circRIMS2*) and the circRNA derived from the *CDR1* antisense locus (*ciRS-7*) (Rybak-Wolf et al., 2015; Venø et al., 2015). These circRNAs are essentially exclusively expressed in the brain and are conserved across mouse, human and pig. Both circRNAs exhibit a general monotonic induction of expression during neuronal development *in vitro* (Rybak-Wolf et al., 2015), whereas they appear to exhibit peaks in expression during fetal development in pig (Venø et al., 2015).

Factors regulating circRNA abundance

The cell type-specific splicing and dynamic expression of circRNAs during development raise questions regarding the trans-acting factors that regulate circRNA biogenesis and decay. As RNA-binding proteins (RBPs) are the major trans-acting factors that regulate pre-mRNA splicing, attention has been focused on identifying those RBPs that might play a role in the splicing of circRNA.

The first identified factor regulating circRNA production was the Muscleblind protein (MBL) in *Drosophila* (Ashwal-Fluss et al., 2014). MBL is required for muscle development and the development of photoreceptor cells in the fly eye. It is expressed in embryonic muscle cells and lack of its expression is embryonic

lethal (Begemann et al., 1997). As a regulator of circRNA, MBL promotes the splicing of the second exon of its own pre-mRNA into a circRNA (*circMbl*), which is one of the most highly expressed circRNAs in the fly head (Ashwal-Fluss et al., 2014). It has been proposed that *circMbl* competes with *mbl* mRNA production, thereby decreasing the levels of MBL protein in a negative feedback-like mechanism (Fig. 5A). Enhancing this feedback, *circMbl* also contains several potential MBL-binding sites that are conserved from fly to human, which might serve to sequester the protein. However, MBL does not regulate the biogenesis of all tested circRNAs in *Drosophila*. Highlighting this, a study of the *laccase2* and *PlexA* genes found that circularization was under the combinatorial control of a number of heterogeneous nuclear ribonucleoproteins (hnRNPs) and SR proteins (Kramer et al., 2015). These factors may exert their effects directly on the pre-mRNA or indirectly by acting on other targets, which themselves are direct regulators.

In humans, members of the ADAR family of RBPs, which are best known for their role in RNA editing and convert adenine to inosine in RNA duplexes (Nishikura, 2010), have been implicated in circRNA production. ADAR is essential for mammalian development: *ADAR* knockout is embryonic or perinatal lethal, depending on the isoform that is absent (Higuchi et al., 2000; Wang et al., 2000). The reported enrichment of Alu elements in the introns flanking circularized exons, and the base-pairing interactions

implicated in circRNA biogenesis, generated the hypothesis that ADARs might regulate circRNA production. Specifically, it was proposed that high ADAR expression destabilizes the base-pairing interactions required for circRNA biogenesis, thereby decreasing production (Fig. 5B). This hypothesis was tested using short hairpin RNA (shRNA)-mediated knockdown of ADAR and revealed a subset of circRNAs with increased abundance after shRNA treatment, whereas other circRNAs exhibited unchanged levels (Ivanov et al., 2015; Rybak-Wolf et al., 2015). This effect is presumably due to direct interactions between ADAR and a pre-mRNA, given that A-to-I conversions were enriched nearby circRNA splice sites compared with controls, although future research is needed to test this hypothesis. Our current understanding of ADAR suggests that the likelihood of an interaction between ADAR and complementary sequences decays exponentially with distance (by a factor of e every 800 nt), suggesting that ADAR-mediated regulation may operate through indirect mechanisms in some cases (Bazak et al., 2014).

Quaking (QKI) is an additional RBP that regulates circRNA production (Conn et al., 2015). QKI is required for circulatory and neural development, including blood vessel formation (Noveroske et al., 2002) and myelination (Sidman et al., 1964), and, as a result, homozygous *QKI* knockouts are embryonic lethal. In a recent study, a fluorescent reporter derived from a *SMARCA5* gene fragment (exons 14–17) was used as a circRNA reporter to monitor splice isoform expression; this gene generates a circRNA from exons 15 and 16 that increases in expression upon induction of epithelial-to-mesenchymal transition (EMT) with TGF β (Conn et al., 2015). By RNA interference (RNAi)-based knockdown of a targeted panel of RBPs that also exhibited a change in their expression upon EMT and monitoring of *SMARCA5* splicing, QKI was identified as a regulator of circRNA biogenesis. This regulation depended on the presence of putative QKI-binding sites in the flanking introns of circularized exons, and these sites were sufficient to induce circRNA biogenesis, suggesting a simple mechanism that regulates the splicing of circRNA through dimerization of QKI between flanking introns (Fig. 5C).

Functions of circRNAs

Given their dynamic and cell type-specific expression patterns during development, many studies have focused on the potential developmental roles and functions of circRNAs. Although some studies are beginning to point towards some generalized functions

for circRNAs, a unified explanation for the functions of the vast majority of circRNAs is lacking.

The function of specific circRNAs in development

Almost a decade ago, the *Drosophila muscleblind (mbl)* gene was found to give rise to developmentally regulated, highly abundant transcriptional products now known to be circRNAs (Ashwal-Fluss et al., 2014; Houseley et al., 2006). Because *mbl* is an essential gene in *Drosophila* and human (Artero et al., 1998; Begemann et al., 1997), future studies may reveal if and how the circular isoform contributes to this essentiality or plays a role in development. As mentioned above, *circMbl* might exert these effects by tuning the expression of the *mbl* transcript, competing with the production of mRNA and binding to MBL protein.

CircRNA from the mouse *Sry* gene is another early key example of developmentally regulated circRNA splicing. In this case, the linear *Sry* isoform is expressed in the developing genital ridge where it plays a fundamental role as a transcription factor in sex determination, whereas the circular isoform is expressed in adult testes (Capel et al., 1993). The circularization of the *Sry* transcript is dictated by promoter usage: the use of a promoter proximal to the coding region gives rise to a linear translated transcript, whereas the use of a distal promoter gives rise to a linear RNA containing long inverted repeats that is spliced to form a circRNA (Fig. 6) (Hacker et al., 1995). These complementary sequences are required for splicing the *Sry* circRNA (Dubin et al., 1995). Indeed, complementary sequence-mediated circularization appears to be a more general phenomenon; it has been corroborated for several circRNAs in recent years and is now the basis of many circRNA overexpression vectors (Kramer et al., 2015; Liang and Wilusz, 2014; Zhang et al., 2014). Although no function could be ascribed to the *Sry* circRNA at the time of its discovery, there is now evidence that it might function as a miRNA sponge for miR-138, binding up to 16 molecules of this miRNA per circRNA (Fig. 6; Fig. 7A) (Hansen et al., 2013).

Another circRNA named *ciRS-7*, or sometimes simply *CDRIas*, which is derived from the *Cdr1* antisense locus, also probably functions as a miRNA sponge (Hansen et al., 2013; Memczak et al., 2013). *ciRS-7* is very highly expressed in the mammalian brain, is induced during neuronal development (Rybak-Wolf et al., 2015) and has >70 potential miRNA-binding sites for miR-7, most of which are conserved across eutherian mammals (Hansen et al., 2013; Memczak et al., 2013). When *ciRS-7* is ectopically expressed

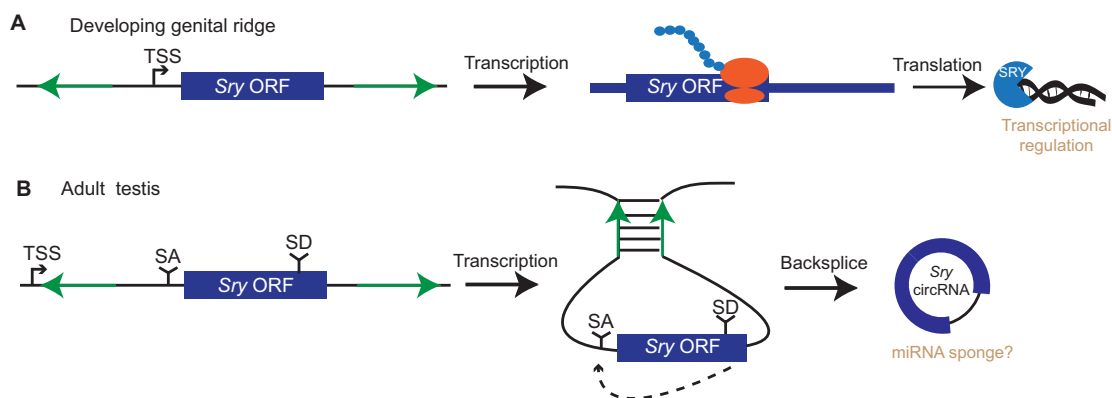


Fig. 6. Developmentally regulated expression of SRY. (A) In the genital ridge of the developing mouse embryo, the *Sry* transcription start site (TSS) occurs proximal to the open reading frame (ORF), yielding a translatable mRNA that gives rise to SRY protein, a transcription factor involved in sex determination. (B) In the adult testis, the TSS occurs far upstream, yielding a long transcript containing large inverted repeats (green arrows). This transcript is backspliced to form a circRNA, which might function as a miR-138 sponge. SA, splice acceptor; SD, splice donor.

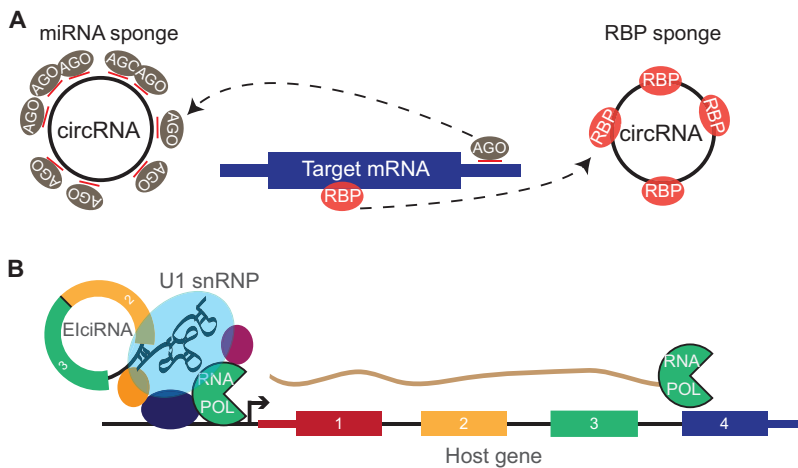


Fig. 7. circRNA functional classes. (A) circRNAs can function as a sponge containing several binding sites for a particular miRNA/RBP and can compete a miRNA/RBP away (dashed arrows) from its mRNA targets, altering gene expression. (B) Through an interaction with U1 snRNP, exon-intron circRNAs (EliciRNAs) can interact with transcription complexes at host genes to induce their transcription (adapted from Li et al., 2015).

in zebrafish, which normally do not express this circRNA but do express miR-7, defects in midbrain development are observed (Memczak et al., 2013), suggesting that this RNA might play a role in the development of the mammalian brain where ciRS-7 and miR-7 are co-expressed (Hansen et al., 2013).

Broad classes of circRNA function

The discovery that *ciRS-7* and *Sry* may serve as miRNA sponges (Fig. 7A) generated great excitement that circRNAs might play a general role in post-transcriptional regulation. However, the analysis of AGO2 crosslinking to circRNAs, as well as computational searches for enrichment of miRNA seed matches in exons contained in circRNA versus neighboring non-circularized exons, has revealed only a few other candidate circRNAs that might function in this manner, none of which has yet been validated (Guo et al., 2014; You et al., 2015). Indeed, both *Sry* and *ciRS-7* are exceptional in their primary sequence: both are circular RNAs hosted in genes with single exons, and both are derived from genomic regions with a highly repetitive sequence. The analysis of circRNA expression in organisms lacking siRNA pathways, namely, *S. cerevisiae* and *P. falciparum*, also supports additional functions for circRNA aside from a function as a miRNA sponge (Wang et al., 2014).

Similarly, the function of *circMbl* in *Drosophila* to sequester the MBL protein might also be exceptional (Ashwal-Fluss et al., 2014). Supporting the function of circRNAs as an RBP sponge, an analysis of mined photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) data from 20 RBPs revealed slightly higher cluster density for circularized exons than for a control cohort of neighboring exons (Guo et al., 2014). However, a bioinformatic analysis of 38 RBP sequence motifs found that circularized exons contained a lower RBP-binding density than did the coding sequence or 3' UTRs of mRNAs (You et al., 2015). It is important to note that these results are not necessarily contradictory; as circRNAs are not translated, RBPs may not be easily displaced, accounting for high experimentally observed cluster densities despite bioinformatic predictions. Still, in order to have an appreciable effect on the concentration of an RBP in the cell without an enrichment of RBP-binding sites, a circRNA would have to be very highly expressed, making an RBP or miRNA sponge function unlikely for most circRNAs (Denzler et al., 2014), but, of course, rare cases may exist.

Another potential function of circRNAs could be to compete with the splicing of an mRNA, as in the case of *mbl* (Fig. 5A)

(Ashwal-Fluss et al., 2014). As circRNAs almost always consist of exons that are also included in mRNA, the production of a circRNA would be expected to interrupt or compete with the splicing of the linear mRNA in most cases (unless a stable exon-skipped transcript can be formed). Whether this 'function' is merely a by-product of circRNA biogenesis remains to be tested. Although not a strict requirement, a feedback mechanism between the gene product and the splicing of its pre-mRNA would argue in favor of such a function.

Some circRNAs have also been implicated in transcriptional or post-transcriptional gene regulation of their host genes. The *CDRIas* circRNA is purported to promote the expression of *CDRI* sense mRNA, but the precise mechanism by which this is achieved is unknown (Hansen et al., 2011). More recently, a class of regulatory circRNAs, named exon-intron circRNAs (EliciRNAs), has been identified and appears to play a role in transcriptional regulation (Li et al., 2015). Such EliciRNAs are multiexon circRNAs containing one or more unspliced intervening introns. Unlike most circRNAs, some EliciRNAs have been shown to be localized to the nucleus, and through an interaction with the U1 small nuclear ribonucleoprotein (snRNP), a spliceosomal component, can promote transcription of their parental genes (Fig. 7B). In this way, circRNA might function as a scaffold for RBPs regulating transcription. This example suggests a provocative hypothesis for a potential broader role for circRNAs as stable molecular scaffolds, much like some long noncoding RNAs (e.g. HOTAIR) (Tsai et al., 2010; Yoon et al., 2013).

Conclusions

Although the number of circRNAs with known functions is expanding, there are still thousands of circRNAs for which the functions remain unknown. It is possible that the majority of circRNAs have a single as yet unknown function or act together to serve one unified role. Still, it is possible that a large fraction of expressed circRNAs are non-functional and merely 'noisy' by-products of splicing; circRNAs appear to be regulated and conserved, but may simply piggyback on the regulatory factors involved in linear mRNA splicing, which could have conserved binding patterns. Because the expression of a circRNA is related to the expression of its host gene, it is difficult to probe these functional questions using standard techniques. A deeper understanding of circRNA biogenesis may allow us to specifically knock out circRNAs using genome-editing tools. This would open the door to testing for functional consequences of circRNA

expression. Furthermore, ectopic circRNA expression plasmids could be used to overexpress these molecules, which might be useful in functional knockout rescue experiments.

Regardless of their specific functional roles, circRNAs provide fodder for many basic cell biological questions regarding their biogenesis, nuclear export and decay, and they may even prove useful as biomarkers of cellular states owing to their stability and dynamic expression.

Acknowledgements

We would like to acknowledge the members of the Salzman lab and the Stanford Biochemistry Department for useful discussions as we prepared this manuscript.

Competing interests

The authors declare no competing or financial interests.

Funding

This work was supported by the National Cancer Institute; the National Institute of General Medical Sciences; a McCormick-Gabilan Fellowship (Stanford University School of Medicine); a Donald E. and Delia B. Baxter Foundation Faculty Scholars Award (to J.S.); a Lucille P. Markey Biomedical Research Fellowship; and the National Science Foundation Graduate Research Fellowship Program (to S.P.B.). J.S. is an Alfred P. Sloan fellow in Computational & Evolutionary Molecular Biology, and S.P.B. is a fellow of the Stanford ChEM-H Chemistry/Biology Interface Pre-doctoral Training Program. Deposited in PMC for release after 12 months.

References

- Agabian, N. (1990). Trans splicing of nuclear pre-mRNAs. *Cell* **61**, 1157-1160.
- Artero, R., Prokop, A., Paricio, N., Begemann, G., Pueyo, I., Mlodzik, M., Perez-Alonso, M. and Baylies, M. K. (1998). The muscleblind gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by Dmef2. *Dev. Biol.* **195**, 131-143.
- Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., Evtantal, N., Memczak, S., Rajewsky, N. and Kadener, S. (2014). circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* **56**, 55-66.
- Awan, A. R., Manfredo, A. and Pleiss, J. A. (2013). Lariat sequencing in a unicellular yeast identifies regulated alternative splicing of exons that are evolutionarily conserved with humans. *Proc. Natl. Acad. Sci. USA* **110**, 12762-12767.
- Barrett, S. P., Wang, P. L. and Salzman, J. (2015). Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife* **4**, e07540.
- Bazak, L., Levanon, E. Y. and Eisenberg, E. (2014). Genome-wide analysis of Alu editability. *Nucleic Acids Res.* **42**, 6876-6884.
- Begemann, G., Paricio, N., Artero, R., Kiss, I., Perez-Alonso, M. and Mlodzik, M. (1997). muscleblind, a gene required for photoreceptor differentiation in *Drosophila*, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development* **124**, 4321-4331.
- Burd, C. E., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z. and Sharpless, N. E. (2010). Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.* **6**, e1001233.
- Cajigas, I. J., Tushev, G., Will, T. J., tom Dieck, S., Fuerst, N. and Schuman, E. M. (2012). The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453-466.
- Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., Goodfellow, P. and Lovell-Badge, R. (1993). Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* **73**, 1019-1030.
- Chen, C. Y. and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268**, 415-417.
- Cocquerelle, C., Daubersies, P., Majérus, M. A., Kerckaert, J. P. and Bailleul, B. (1992). Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* **11**, 1095-1098.
- Cocquerelle, C., Mascrez, B., Hetuin, D. and Bailleul, B. (1993). Mis-splicing yields circular RNA molecules. *FASEB J.* **7**, 155-160.
- Cocquet, J., Chong, A., Zhang, G. and Veitia, R. A. (2006). Reverse transcriptase template switching and false alternative transcripts. *Genomics* **88**, 127-131.
- Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmandis, M., Phillips, C. A., Roslan, S., Schreiber, A. W., Gregory, P. A. and Goodall, G. J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell* **160**, 1125-1134.
- Denzler, R., Agarwal, V., Stefano, J., Bartel, D. P. and Stoffel, M. (2014). Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol. Cell* **54**, 766-776.
- Dubin, R. A., Kazmi, M. A. and Ostrer, H. (1995). Inverted repeats are necessary for circularization of the mouse testis *Sry* transcript. *Gene* **167**, 245-248.
- Engström, P. G., Steijger, T., Sipos, B., Grant, G. R., Kahles, A., Rättsch, G., Goldman, N., Hubbard, T. J., Harrow, J., Guigó, R. et al. (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat. Methods* **10**, 1185-1191.
- Glažar, P., Papavasileiou, P. and Rajewsky, N. (2014). circBase: a database for circular RNAs. *RNA* **20**, 1666-1670.
- Guo, J. U., Agarwal, V., Guo, H. and Bartel, D. P. (2014). Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* **15**, 409.
- Hacker, A., Capel, B., Goodfellow, P. and Lovell-Badge, R. (1995). Expression of *Sry*, the mouse sex determining gene. *Development* **121**, 1603-1614.
- Hansen, T. B., Wiklund, E. D., Bramsen, J. B., Villadsen, S. B., Statham, A. L., Clark, S. J. and Kjems, J. (2011). miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **30**, 4414-4422.
- Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K. and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384-388.
- Hansen, T. B., Venø, M. T., Damgaard, C. K. and Kjems, J. (2015). Comparison of circular RNA prediction tools. *Nucleic Acids Res.* **44**, e58.
- Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R. and Seeburg, P. H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78-81.
- Houseley, J. M., Garcia-Casado, Z., Pascual, M., Paricio, N., O'Dell, K. M. C., Monckton, D. G. and Artero, R. D. (2006). Noncanonical RNAs from transcripts of the *Drosophila* muscleblind gene. *J. Hered.* **97**, 253-260.
- Ivanov, A., Memczak, S., Wylter, E., Torti, F., Porath, H. T., Orejuela, M. R., Piechotta, M., Levanon, E. Y., Landthaler, M., Dieterich, C. et al. (2015). Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* **10**, 170-177.
- Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., Marzluff, W. F. and Sharpless, N. E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**, 141-157.
- Jordan, M. C., Henderson, S. A., Han, T., Fishbein, M. C., Philipson, K. D. and Roos, K. P. (2010). Myocardial function with reduced expression of the sodium-calcium exchanger. *J. Card. Fail.* **16**, 786-796.
- Kramer, M. C., Liang, D., Tatomer, D. C., Gold, B., March, Z. M., Cherry, S. and Wilusz, J. E. (2015). Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev.* **29**, 2168-2182.
- Kulpa, D., Topping, R. and Telesnitsky, A. (1997). Determination of the site of first strand transfer during Moloney murine leukemia virus reverse transcription and identification of strand transfer-associated reverse transcriptase errors. *EMBO J.* **16**, 856-865.
- Li, X.-F. and Lytton, J. (1999). A circularized sodium-calcium exchanger exon 2 transcript. *J. Biol. Chem.* **274**, 8153-8160.
- Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., Zhong, G., Yu, B., Hu, W., Dai, L. et al. (2015). Exon-intron circular RNAs regulate transcription in the nucleus. *Nat. Struct. Mol. Biol.* **22**, 256-264.
- Liang, D. and Wilusz, J. E. (2014). Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.* **28**, 2233-2247.
- Luo, G. X. and Taylor, J. (1990). Template switching by reverse transcriptase during DNA synthesis. *J. Virol.* **64**, 4321-4328.
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S. D., Gregersen, L. H., Munschauer, M. et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333-338.
- Morris, K. V. and Mattick, J. S. (2014). The rise of regulatory RNA. *Nat. Rev. Genet.* **15**, 423-437.
- Nigro, J. M., Cho, K. R., Fearon, E. R., Kern, S. E., Ruppert, J. M., Oliner, J. D., Kinzler, K. W. and Vogelstein, B. (1991). Scrambled exons. *Cell* **64**, 607-613.
- Nishikura, K. (2010). Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **79**, 321-349.
- Noveroske, J. K., Lai, L., Gaussin, V., Northrop, J. L., Nakamura, H., Hirschi, K. K. and Justice, M. J. (2002). Quaking is essential for blood vessel development. *Genesis* **32**, 218-230.
- Peng, Z., Yuan, C., Zellmer, L., Liu, S., Xu, N. and Liao, D. J. (2015). Hypothesis: artifacts, including spurious chimeric RNAs with a short homologous sequence, caused by consecutive reverse transcriptions and endogenous random primers. *J. Cancer* **6**, 555-567.
- Perriman, R. and Ares, M. (1998). Circular mRNA can direct translation of extremely long repeating-sequence proteins in vivo. *RNA* **4**, 1047-1054.
- Roy, C. K., Olson, S., Graveley, B. R., Zamore, P. D. and Moore, M. J. (2015). Assessing long-distance RNA sequence connectivity via RNA-templated DNA-DNA ligation. *Elife* **4**, e03700.
- Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R. et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* **58**, 870-885.
- Salzman, J., Jiang H. and Wong, W. H. (2011). Statistical modeling of RNA-Seq data. *Stat. Sci.* **26**, 62-83.
- Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. and Brown, P. O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* **7**, e30733.

- Salzman, J., Chen, R. E., Olsen, M. N., Wang, P. L. and Brown, P. O. (2013). Cell-type specific features of circular RNA expression. *PLoS Genet.* **9**, e1003777.
- Schindewolf, C., Braun, S. and Domdey, H. (1996). In vitro generation of a circular exon from a linear pre-mRNA transcript. *Nucleic Acids Res.* **24**, 1260-1266.
- Sidman, R. L., Dickie, M. M. and Appel, S. H. (1964). Mutant mice (Quaking and Jimpy) with deficient myelination in the central nervous system. *Science* **144**, 309-311.
- Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L.-H. and Bindereif, A. (2015). Exon circularization requires canonical splice signals. *Cell Rep.* **10**, 103-111.
- Surono, A., Takeshima, Y., Wibawa, T., Ikezawa, M., Nonaka, I. and Matsuo, M. (1999). Circular dystrophin RNAs consisting of exons that were skipped by alternative splicing. *Hum. Mol. Genet.* **8**, 493-500.
- Suzuki, H., Zuo, Y., Wang, J., Zhang, M. Q., Malhotra, A. and Mayeda, A. (2006). Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* **34**, e63.
- Szabo, L., Morey, R., Palpant, N. J., Wang, P. L., Afari, N., Jiang, C., Parast, M. M., Murry, C. E., Laurent, L. C. and Salzman, J. (2015). Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. *Genome Biol.* **16**, 126.
- Tabak, H. F., Van der Horst, G., Smit, J., Winter, A. J., Mul, Y. and Groot Koerkamp, M. J. A. (1988). Discrimination between RNA circles, interlocked RNA circles and lariats using two-dimensional polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **16**, 6597-6605.
- Tai, H.-C. and Schuman, E. M. (2006). MicroRNA: microRNAs reach out into dendrites. *Curr. Biol.* **16**, R121-R123.
- Tsai, M.-C., Manor, O., Wan, Y., Mosammamaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E. and Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689-693.
- Venø, M. T., Hansen, T. B., Venø, S. T., Clausen, B. H., Grebing, M., Finsen, B., Holm, I. E. and Kjems, J. (2015). Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development. *Genome Biol.* **16**, 245.
- Vincent, H. A. and Deutscher, M. P. (2006). Substrate recognition and catalysis by the exoribonuclease RNase R. *J. Biol. Chem.* **281**, 29769-29775.
- Wang, Q., Khillan, J., Gadue, P. and Nishikura, K. (2000). Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science* **290**, 1765-1768.
- Wang, P. L., Bao, Y., Yee, M.-C., Barrett, S. P., Hogan, G. J., Olsen, M. N., Dinneny, J. R., Brown, P. O. and Salzman, J. (2014). Circular RNA is expressed across the eukaryotic tree of life. *PLoS ONE* **9**, e90859.
- Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S. E., Graveley, B. R. and Lai, E. C. (2014). Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* **9**, 1966-1980.
- Yoon, J. H., Abdelmohsen, K., Kim, J., Yang, X., Martindale, J. L., Tominaga-Yamanaka, K., White, E. J., Orjalo, A. V., Rinn, J. L., Kreft, S. G. et al. (2013). Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat. Commun.* **4**, 2939.
- You, X., Vlatkovic, I., Babic, A., Will, T., Epstein, I., Tushev, G., Akbalik, G., Wang, M., Glock, C., Quedenau, C. et al. (2015). Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* **18**, 603-610.
- Zaphiropoulos, P. G. (1996). Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. *Proc. Natl. Acad. Sci. USA* **93**, 6536-6541.
- Zaphiropoulos, P. G. (1997). Exon skipping and circular RNA formation in transcripts of the human cytochrome P-450 2C18 gene in epidermis and of the rat androgen binding protein gene in testis. *Mol. Cell. Biol.* **17**, 2985-2993.
- Zhang, X.-O., Wang, H.-B., Zhang, Y., Lu, X., Chen, L.-L. and Yang, L. (2014). Complementary sequence-mediated exon circularization. *Cell* **159**, 134-147.