Short Communication

Establishing essential quality criteria for the validation of circular RNAs as biomarkers

Christina Pfafenrot, Christian Preußer*

Institute of Biochemistry, Justus Liebig University of Giessen, D-35392 Giessen, Germany

ABSTRACT

Non-coding RNAs were established in the last decade as a new valuable biomarker class for human diseases. Specifically, circular RNAs (circRNAs) were only recently discovered as a new large group of non-coding RNAs that, due to their circular configuration, are metabolically more stable compared to their linear counterparts and therefore highly suitable for biomarker use. Based on high-throughput sequencing, the catalogs of endogenous circRNAs with disease relevance and correlation continue to grow steadily. As a consequence, circRNAs emerged as novel and attractive biomarkers, indicated by numerous recent publications. Here we would like to stress the need of essential quality criteria for validation and characterization of circular RNAs. In addition to high-throughput sequencing, classical biochemical methods are essential and should be applied for the characterization of this special class of RNAs, in particular to convincingly confirm their circularity.

The rapid evolution of next generation sequencing technologies resulted in the nearly exponential development of putative RNA biomarkers in recent years. In particular, different classes of non-coding RNAs (ncRNA) appear promising for a wide range of clinical applications. This is also reflected in the so-called "Non-coding RNA revolution" [1] and the current focus in non-coding RNA research on investigating the underlying global RNA networks within cells, tissues, or organisms. As a relatively new class of ncRNA biomarkers, circular RNAs (circRNAs) have come more into focus within the last decade (Fig. 1). Although single examples of these particular RNA species had been known for more than forty years [2], circRNAs were established as a large RNA class only a few years ago, based on high-throughput sequencing and bioinformatics [3]. CircRNAs derived from pre-mRNAs (in some cases also from pre-ncRNAs) are characterized by a covalently closed loop structure, which is generated by a special mode of alternative splicing of pre-mRNAs, also called "backsplicing": A 5′ splice site is joined to an upstream instead of a downstream 3′ splice site (Fig. 2). Some studies also showed that constitutive splice signals as well as the canonical splicing machinery are involved in circularization [4,5]. The average size of a circRNAs can range from under 100 nts to over 4 kb in which the commonly observed circular RNAs consist of 2–3 exons with internal introns removed [6]. Furthermore, circRNA biogenesis can be promoted by *cis*-elements and *trans*-factors by bringing the corresponding splice sites in close proximity, resulting in circularization [7]. In addition to this major, exon-type circRNAs, there are also circular intronic RNAs (ciRNAs) and exon-intron circular RNAs (EIciRNAs) [6]. Although a defined function of circRNAs is still under debate, some studies revealed a microRNA sponge function of naturally occurring circRNAs like ciRS-7 and SRY [8,9]. In addition, several other, hypothetical roles have been discussed, such as templates for translation into peptides or proteins, protein sponging, allostery, scaffold functions in RNA-protein complex assembly, or antisense activity. CircRNAs are cell type-specifically expressed and, due to their structure, very stable compared to their linear counterparts [9]. Because of these biological properties and the correlation with various human diseases, including cardiovascular diseases [10,11], disorders of the nervous system [12,13], diabetic retinopathy [14] and cancer [15], it becomes clear that circRNAs may play an important role as specific biomarkers. Since circRNAs are also found in extracellular vesicles circulating in various body fluids such as blood and saliva this should greatly extend the potential of circRNAs as prognostic and diagnostic biomarkers, especially for liquid biopsies. Specifically, RNA-Seq reads spanning the backsplice junction allow the genome-wide bioinformatic identification of circRNAs. However, we observe a certain amount of different and facile approaches to unequivocally prove circularity of this important ncRNA class. This comprises all stages of circRNA identification, ranging from RNA-Seq approaches, bioinformatics, to biochemical validation. Here we would like to stress that stringent quality criteria have to be defined, which distinguish real circular configuration from possible artifacts.

---

* Corresponding author.
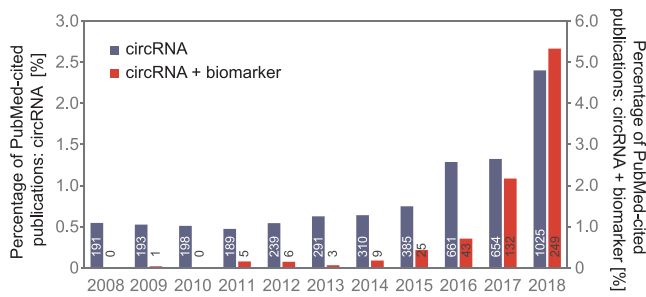 *E-mail address:* christian.preusser@chemie.bio.uni-giessen.de (C. Preußer).

Fig. 1. Circular RNA as emerging novel biomarkers. The rapid development of circRNA research and circRNA in combination with biomarkers can be inferred from the percentage of PubMed-cited publications (https://www.ncbi.nlm.nih.gov/pubmed) using *circular RNA* (blue graphs) and *circular RNA + biomarker* (red graphs) as keywords in searches, of publications using the general terms *RNA* and *RNA + biomarker* as keywords, for publications between 2008 and 2018. The number inside the bars represents the absolute number of publications.
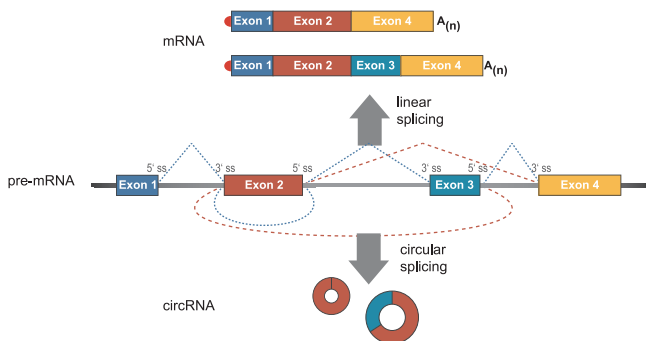

Fig. 2. Biogenesis of circular RNAs (circRNAs). Schematic representation of the biogenesis of linear mRNA and circular RNA. During canonical splicing, introns are removed and exons are joined with each other, generating either a single mRNA from the pre-mRNA (constitutive splicing) or several splice variants (alternative splicing). CircRNAs are generated by an alternative splicing mechanism, also referred to as backsplicing, in which a 5′ splice site (5′ ss) is joined to an upstream 3′ splice site (3′ ss) instead of a downstream 3′ ss. Different splice variants can be generated by alternative backsplicing resulting in single- or multi-exonic circular RNAs (circRNAs). Exons are depicted as colored boxes and introns as solid lines. The 5′ cap structure of mRNA is shown as red dots and the 3′ poly(A) tail as A(n). Constitutive (blue) and alternative splicing (red) is indicated by dashed lines.

Already the initial stage of identification of novel circRNAs requires the establishment and standardization of unbiased approaches. Several things have to be considered: The RNA integrity is critical, since neither RT-PCR might detect circRNAs molecules nor other methods such as Northern blotting, since the characteristic circular junction might get lost during partial degradation [16]. Even the choice of library preparation itself has a strong impact on the results. In particular, one of the first fundamental steps during RNA-Seq library preparation is crucial, namely the reverse transcription. Most of the reverse transcriptases used are known to be error prone and can frequently introduce mutations, which have been widely recognized [17]. As a result, aberrant *trans*-splicing products can be generated during reverse transcription, mainly based on a template-switching activity of some enzymes [18,19]. This phenomenon can also occur in the case of linear transcripts, but in particular for circRNAs can easily result in misinterpretations.

In addition, endogenous *trans*-spliced transcripts can lead to false positive events. In particular, duplicated exon sequences within one mRNA (intragenic *trans* splicing) or transcripts derived from different genes on different chromosomes (intergenic *trans* splicing) can be detected and result in false positive events [20]. Another critical step

during library preparation is that most of the commercially available RNA-Seq library kits are designed to analyze either the RNA profile of long RNAs, such as mRNA and lncRNAs, or of small ncRNAs, for example miRNA. Again, this general limitation applies also to circRNAs. Furthermore, the library preparation from circRNAs either requires initial fragmentation to generate accessible 5′ and 3′ ends, or relies on random hexamer priming, which may result in inefficiency and/or biased representation in RNA-Seq [21]. This turns out to be especially challenging for the analysis of RNA profiles from extracellular vesicles, which primarily contain short RNAs less than 500 nts, however longer transcripts were reported as well [22,23].

After sequencing, the large number of available circRNA prediction tools can introduce additional variability. Some of them require an established annotation and some knowledge about existing exon-intron structures, while others, such as CIRI [24], rely on *de novo* assembly and thus are able to identify new backsplice junctions. These different options of evaluation tools can result in a dramatic differences between the outputs of algorithms used [21]. Referring thereto, a recent study compared 11 different circRNA detection tools on stimulated and on real datasets and could show that CIRI, CIRCexplorer and KNIFE achieved better balanced performance between their precision and sensitivity compared to other methods [25]. Nevertheless, many of these algorithms show strongly aberrant results and a high degree of false positive results, for example in relation to their resistance to RNase R treatment. These studies show that CircRNA detection tools should be treated with care, and that a combination of tools is required to make reliable predictions [21]. Therefore one should consider the combination of more than one method for the bioinformatic analysis to minimize the number of false positives.

However, from our point of view, most critical is the often neglected stringent biochemical validation of the circRNAs, which should follow the bioinformatic evaluation. Unfortunately, in many recent publications, including many on biomarker detection, this important and crucial issue is restricted to a minimum. As a method of choice, RNA-Seq and the bioinformatic analysis in combination with RT-(q)PCR approaches are employed as the only assays for validation. RT-qPCR approaches have the great advantage – in particular digital PCR- to check genome-wide data sets in a quantitative manner and represents the simplest and fastest method to detect the circularity of circRNAs. By designing divergent primer pairs relative to the circularizing exon/s, PCR products can be generated (usually 100–200 nts), which specifically include the characteristic circular splice site. Beyond the scope of this point of view, this method offers certain limitations, which are particularly relevant in the validation of circRNAs and are mainly based on the problems caused by the transcriptases as already mentioned above. One possibility to circumvent this problem is to combine the RT-PCR analysis with RNase R treatment. RNase R is a 3′ to 5′ exoribonuclease that degrades linear RNA, whereas most circRNAs are resistant due to their circular configuration. Unfortunately, relatively large circRNAs tend to be not absolutely RNase R resistant and in addition, complex secondary structures and chemical modifications in linear RNAs can influence the RNase R activity [16]. Moreover, variations in RNases R activities among different batches and manufacturers, experimental settings, as well as secondary RNA structure, might affect the efficiency of linear RNA removal as well.

To unequivocally demonstrate a circular configuration, Northern blot analysis, which is RNase R independent, should be defined as the gold standard and also as a method of choice (for details on the method, see [26]). Although this classic type of RNA analysis is often neglected because it is somewhat more time-consuming and requires some practical experience, Northern blotting is greatly versatile and thereby ideally suitable for circRNA characterization. For Northern blot detection, probes are designed that either span the backsplice junction or hybridize to the total transcript. By simply choosing a suitable gel-electrophoresis system (agarose and/or polyacrylamide), the two possible configurations (circular or linear) can be clearly distinguished

from each other.

Furthermore, Northern blotting can also be combined with RNase R or RNase H ribonuclease digestion. In particular, the RNase H cleavage assay has emerged as an elegant method for validation of circRNAs [4,27]. In brief, antisense oligonucleotides are designed directed against a specific circRNA. RNase H recognizes DNA-RNA hybrids and subsequent cuts the RNA within the duplex. The characteristic shift of mobility during denaturing polyacrylamide gel electrophoresis from an aberrantly slow migration of the circRNA to the expected linear behavior, as well as the cleavage patterns of RNase H digestion, can then be analyzed by Northern blotting. Furthermore, new approaches should be considered for a biochemical validation of the circular configuration. Here, among other things, probe-based systems (*in situ* hybridization) could play a crucial role. Moreover, current developments in next-generation sequencing such as the nanopore technology (Oxford Nanopore Technologies) should likely provide new options in the future. MinION's for example could provide emerging direct RNA sequencing capabilities.

In sum, all the problems mentioned here can only be circumvented by combining several biochemical validation methods and by maintaining stringent quality standards. Especially in the field of biomarker development, complete and convincing circRNA characterization should be a top priority. The commonly mentioned low reproducibility and the variable analytical standards are often an obstacle during the development of potential new biomarkers. CircRNAs highlighted here, indeed harbor great biomarker potential, but there is a strong need for standardization and for setting up minimal requirements for validating this class of ncRNA. How could this be achieved? The growing circRNA community should define what is essential to establish a new circRNA species as such. In recent years this has been successfully implemented for other issues, such as the MIQE (<u>m</u>inimum <u>i</u>nformation for publication of <u>q</u>uantitative real-time PCR <u>e</u>xperiments) [28] and MISEV (<u>mi</u>nimal information for <u>s</u>tudies of <u>EVs</u>) [29], where the respective communities and their experts have provided recommendations on experimental methods and minimal information in reporting.

Acknowledgement

References

[1] T.R. Cech, J.A. Steitz, The noncoding RNA revolution-trashing old rules to forge new ones, Cell 157 (2014) 77–94, https://doi.org/10.1016/j.cell.2014.03.008.

[2] H.L. Sanger, G. Klotz, D. Riesner, H.J. Gross, A.K. Kleinschmidt, Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures, Proc. Natl. Acad. Sci. U. S. A. 73 (1976) 3852–3856.

[3] L.L. Chen, The biogenesis and emerging roles of circular RNAs, Nat. Rev. Mol. Cell Biol. 17 (2016) 205–211, https://doi.org/10.1038/nrm.2015.32.

[4] S. Starke, I. Jost, O. Rossbach, T. Schneider, S. Schreiner, L.H. Hung, A. Bindereif, Exon circularization requires canonical splice signals, Cell Rep. 10 (2015) 103–111, https://doi.org/10.1016/j.celrep.2014.12.002.

[5] R. Ashwal-Fluss, M. Meyer, N.R. Pamudurti, A. Ivanov, O. Bartok, M. Hanan, N. Evantal, S. Memczak, N. Rajewsky, S. Kadener, circRNA biogenesis competes with pre-mRNA splicing, Mol. Cell 56 (2014) 55–66, https://doi.org/10.1016/j.molcel.2014.08.019.

[6] E. Lasda, R. Parker, Circular RNAs: diversity of form and function, RNA 20 (2014) 1829–1842, https://doi.org/10.1261/rna.047126.114.

[7] W.R. Jeck, J.A. Sorrentino, K. Wang, M.K. Slevin, C.E. Burd, J. Liu, W.F. Marzluff, N.E. Sharpless, Circular RNAs are abundant, conserved, and associated with ALU repeats, RNA 19 (2013) 141–157, https://doi.org/10.1261/rna.035667.112.

[8] T.B. Hansen, T.I. Jensen, B.H. Clausen, J.B. Bramsen, B. Finsen, C.K. Damgaard, J. Kjems, Natural RNA circles function as efficient microRNA sponges, Nature 495 (2013) 384–388 https://doi.org/10.1038/nature11993.

[9] S. Memczak, M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S.D. Mackowiak, L.H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M. Landthaler, C. Kocks, F. le Noble, N. Rajewsky, Circular RNAs are a large class of animal RNAs with regulatory potency, Nature 495 (2013) 333–338, https://doi.org/10.1038/nature11928.

[10] L.M. Holdt, A. Stahringer, K. Sass, G. Pichler, N.A. Kulak, W. Wilfert, A. Kohlmaier, A. Herbst, B.H. Northoff, A. Nicolaou, G.F. Gäbel, F. Beutner, M. Scholz, J. Thiery, K. Musunuru, K. Krohn, M. Mann, D. Teupser, Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans, Nat. Commun. 7 (2016) 12429, https://doi.org/10.1038/ncomms12429.

[11] L.M. Holdt, A. Kohlmaier, D. Teupser, Molecular roles and function of circular RNAs in eukaryotic cells, Cell. Mol. Life Sci. 75 (2018) 1071–1098, https://doi.org/10.1007/s00018-017-2688-5.

[12] M. Piwecka, P. Glažar, L.R. Hernandez-Miranda, S. Memczak, S.A. Wolf, A. Rybak-Wolf, A. Filipchyk, F. Klironomos, C.A. Cerda Jara, P. Fenske, T. Trimbuch, V. Zywitza, M. Plass, L. Schreyer, S. Ayoub, C. Kocks, R. Kühn, C. Rosenmund, C. Birchmeier, N. Rajewsky, Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function, Science. 357 (2017), https://doi.org/10.1126/science.aam8526 pii: eaam8526.

[13] L. Errichelli, S. Dini Modigliani, P. Laneve, A. Colantoni, I. Legnini, D. Capauto, A. Rosa, R. De Santis, R. Scarfò, G. Peruzzi, L. Lu, E. Caffarelli, N.A. Shneider, M. Morlando, I. Bozzoni, FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons, Nat. Commun. 8 (2017) 14741, https://doi.org/10.1038/ncomms14741.

[14] C. Liu, M.D. Yao, C.P. Li, K. Shan, H. Yang, J.J. Wang, B. Liu, X.M. Li, J. Yao, Q. Jiang, B. Yan, Silencing of circular RNA-ZNF609 ameliorates vascular endothelial dysfunction, Theranostics. 7 (2017) 2863–2877, https://doi.org/10.7150/thno.19353.

[15] S. Qu, Z. Liu, X. Yang, J. Zhou, H. Yu, R. Zhang, H. Li, The emerging functions and roles of circular RNAs in cancer, Cancer Lett. 414 (2018) 301–309, https://doi.org/10.1016/j.canlet.2017.11.022.

[16] A.C. Panda, S. De, I. Grammatikakis, R. Munk, X. Yang, Y. Piao, D.B. Dudekula, K. Abdelmohsen, M. Gorospe, High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs, Nucleic Acids Res. 45 (2017) e116, https://doi.org/10.1093/nar/gkx297.

[17] A. Sebastián-Martín, V. Barrioluengo, L. Menéndez-Arias, Transcriptional inaccuracy threshold attenuates differences in RNA-dependent DNA synthesis fidelity between retroviral reverse transcriptases, Sci. Rep. 8 (2018) 627, https://doi.org/10.1038/s41598-017-18974-8.

[18] J. Cocquet, A. Chong, G. Zhang, R.A. Veitia, Reverse transcriptase template switching and false alternative transcripts, Genomics. 88 (2006) 127–131, https://doi.org/10.1016/j.ygeno.2005.12.013.

[19] J. Houseley, D. Tollervey, Apparent non-canonical trans-splicing is generated by reverse transcriptase in vitro, PLoS One 5 (2010) e12271, , https://doi.org/10.1371/journal.pone.0012271.

[20] C. Preußer, A. Bindereif, Exo-endo trans splicing: a new way to link, Cell Res. 23 (2013) 1071–1072, https://doi.org/10.1038/cr.2013.105.

[21] T.B. Hansen, M.T. Venø, C.K. Damgaard, J. Kjems, Comparison of circular RNA prediction tools, Nucleic Acids Res. 44 (2016) e58, https://doi.org/10.1093/nar/gkv1458.

[22] C. Preußer, L.H. Hung, T. Schneider, M. Hardt, A. Moebus, S. Santoso, A. Bindereif, Selective release of circRNAs in platelet-derived extracellular vesicles, J. Extracell. Vesicles 7 (2018) 1424473, , https://doi.org/10.1080/20013078.2018.1424473.

[23] R. Crescitelli, C. Lässer, T.G. Szabó, A. Kittel, M. Eldh, I. Dianzani, E.I. Buzás, J. Lötvall, Distinct RNA profiles in subpopulations of extracellular vesicles: apoptotic bodies, microvesicles and exosomes, J. Extracell. Vesicles 2 (2013), https://doi.org/10.3402/jev.v2i0.20677.

[24] Y. Gao, J. Wang, F. Zhao, CIRI: an efficient and unbiased algorithm for de novo circular RNA identification, Genome Biol. 16 (2015) 4, https://doi.org/10.1186/s13059-014-0571-3.

[25] X. Zeng, W. Lin, M. Guo, Q. Zou, A comprehensive overview and evaluation of circular RNA detection tools, PLoS Comput. Biol. 13 (2017) e1005420, , https://doi.org/10.1371/journal.pcbi.1005420.

[26] T. Schneider, S. Schreiner, C. Preußer, A. Bindereif, O. Rossbach, Northern blot analysis of circular RNAs, Methods Mol. Biol. 1724 (2018) 119–133, https://doi.org/10.1007/978-1-4939-7562-4_10.

[27] W.R. Jeck, N.E. Sharpless, Detecting and characterizing circular RNAs, Nat. Biotechnol. 32 (2014) 453–461, https://doi.org/10.1038/nbt.2890.

[28] S.A. Bustin, V. Benes, J.A. Garson, et al., The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, Clin. Chem. 55 (2009) 611–622, https://doi.org/10.1373/clinchem.2008.112797.

[29] J. Lötvall, A.F. Hill, F. Hochberg, E.I. Buzás, D. Di Vizio, C. Gardiner, Y.S. Gho, I.V. Kurochkin, S. Mathivanan, P. Quesenberry, S. Sahoo, H. Tahara, M.H. Wauben, K.W. Witwer, C. Théry, Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles, J. Extracell. Vesicles 3 (2014) 26913, https://doi.org/10.3402/jev.v3.26913.