

## REVIEW

## CircRNAs in hematopoiesis and hematological malignancies

A Bonizzato<sup>1</sup>, E Gaffo<sup>2</sup>, G te Kronnie<sup>1</sup> and S Bortoluzzi<sup>2</sup>

Cell states in hematopoiesis are controlled by master regulators and by complex circuits of a growing family of RNA species impacting cell phenotype maintenance and plasticity. Circular RNAs (circRNAs) are rapidly gaining the status of particularly stable transcriptome members with distinctive qualities. RNA-seq identified thousands of circRNAs with developmental stage- and tissue-specific expression corroborating earlier suggestions that circular isoforms are a natural feature of the cell expression program. CircRNAs are abundantly expressed also in the hematopoietic compartment. There are a number of studies on circRNAs in blood cells, a specific overview is however lacking. In this review we first present current insight in circRNA biogenesis discussing the relevance for hematopoiesis of the highly interleaved processes of splicing and circRNA biogenesis. Regarding molecular functions circRNAs modulate host gene expression, but also compete for binding of microRNAs, RNA-binding proteins or translation initiation and participate in regulatory circuits. We examine circRNA expression in the hematopoietic compartment and in hematologic malignancies and review the recent breakthrough study that identified pathogenic circRNAs derived from leukemia fusion genes. CircRNA high and regulated expression in blood cell types indicate that further studies are warranted to inform the position of these regulators in normal and malignant hematopoiesis.

*Blood Cancer Journal* (2016) 6, e483; doi:10.1038/bcj.2016.81; published online 14 October 2016

## INTRODUCTION

Circular RNAs (circRNAs) are covalently closed RNA molecules, in which the 3'- and 5'-ends are linked in a non-collinear way by a process called back-splicing.<sup>1</sup> Unlike in linear RNA splicing, a splice donor site is joined to a splice acceptor site upstream in the primary transcript, yielding a circRNA.<sup>1</sup> CircRNAs can be formed by circularization of a single exon, two or more exons,<sup>2</sup> both exons and intron sequences (exon–intron circRNA, ElcirRNA)<sup>3</sup> or intron sequences only (circularized intron RNA; ciRNA)<sup>4</sup> (Figure 1). Several circular isoforms can be produced from a given gene, and different circRNAs from the same gene may show distinct expression profiles, as reported for *circSTAU2a* and *circSTAU2b*.<sup>5</sup>

Circularity confers specific properties to circRNA: they are highly stable, resistant to RNase R and have longer half-lives compared with linear RNAs<sup>6,7</sup> and tend to accumulate in cells with a low proliferation rate.<sup>8</sup> Detection of circRNA in human body fluids such as plasma<sup>9</sup> and saliva<sup>10</sup> indicates circRNAs as potential disease biomarkers.

The first description of circRNA dates back to several decades ago. Recently, circRNAs were relaunched by RNA-seq-based studies as an RNA species with high relevance for molecular biology and molecular oncology, and today over 10 000 human circRNAs have been identified.<sup>11–17</sup>

CircRNAs are non-poly-adenylated and coincidental discovery of circRNA in the past can be attributed to RNA extraction methods that mainly used polyA selection. Naturally occurring single-stranded covalently closed RNA molecules were first described in plant viroids<sup>18</sup> and were valued for their peculiar structure that allows for rolling circle replication.<sup>1</sup> A few studies in the nineties reported non-canonical splicing with scrambled exons of candidate tumor suppressor gene *DCC*,<sup>19</sup> 'missplicing' of *ETS1*

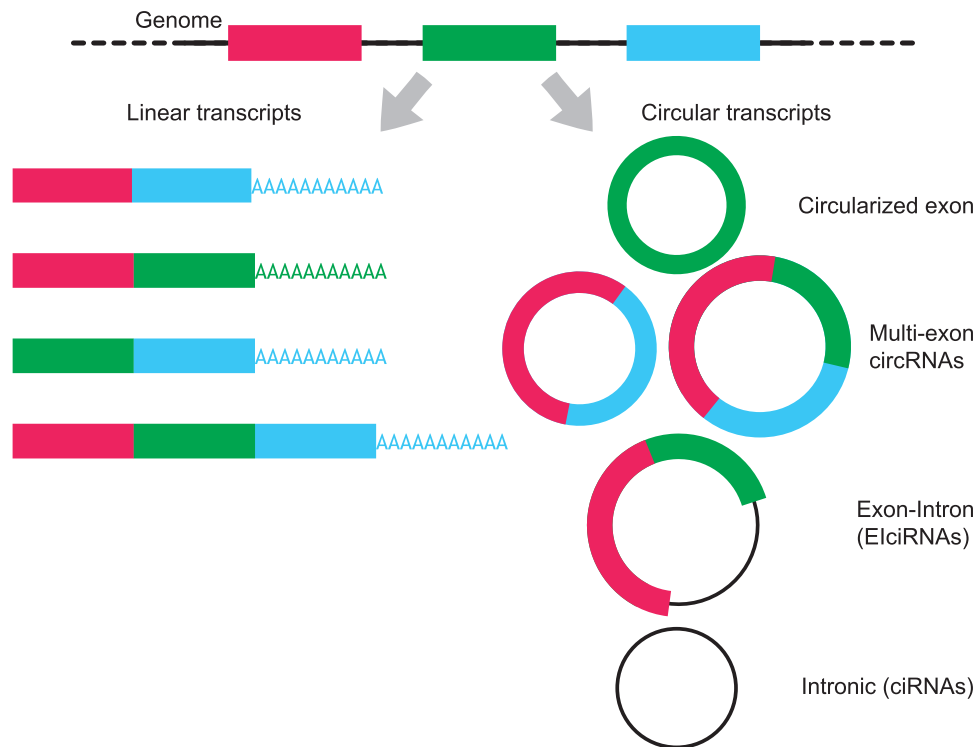
transcripts<sup>20</sup> and murine *Fmn*,<sup>21</sup> and exon circularization in nuclear extracts.<sup>22</sup> Moreover, whereas early in development the mRNA of the *SRY* is translated into the protein that triggers the sex-determining transcriptional cascade, in adults *SRY* transcripts are found as cytoplasmic circular *SRY* (*cSRY*) not particularly bound to polysomes<sup>23</sup> and later proven to efficiently sponge miR-138.<sup>24</sup> Other primary RNAs were found to be processed into circRNA isoforms such as *MLL (KMT2A)*,<sup>25</sup> *ETS1*,<sup>26</sup> *CYP2C18*,<sup>27</sup> *SLC8A1*<sup>(ref. 28)</sup> and dystrophin (*DMD*)<sup>29</sup> transcripts. Examples of circRNAs corresponding to linear noncoding RNA, as well as antisense RNA were also detected.<sup>30,31</sup>

Most of the above-mentioned studies postulated relevant biological functions for circular RNAs but were only confined to certain genes. In any case, shortly after publication of circular forms of an *INK4/ARF*-associated non-coding RNA<sup>30</sup> numerous studies embarked on transcriptome wide circRNA analysis showing developmental stage- and tissue-specific expression, and specific regulatory roles for circRNAs were suggested.<sup>2,13</sup> These new data triggered the interests of the scientific community resulting in the development of molecular methods to study circRNAs and of microarray platforms to measure expression levels of circRNAs (Figure 2), as well as the implementation of bioinformatics software to detect and discover circRNAs from RNA-seq data (Table 1) posing the basis for further experimental circRNA characterization, and for circRNA quantification and differential expression testing. Moreover, several available circRNA databases and web resources (Table 2) could be rather useful to explore putative circRNA interactions and functions.

Several studies on circRNAs in blood cell types and hematologic malignancies were recently conducted and will be discussed below.

<sup>1</sup>Department of Women's and Children's Health, University of Padova, Padova, Italy and <sup>2</sup>Department of Molecular Medicine, University of Padova, Padova, Italy. Correspondence: Dr G te Kronnie, Department of Women's and Children's Health, University of Padova, Via Giustiniani, 3, 35128 Padova, Italy. E-mail: traustekronnie@unipd.it

Received 22 July 2016; accepted 11 August 2016



**Figure 1.** Linear and circRNAs. CircRNAs are produced by backsplicing, and combinations of exons and introns give rise to different products, including single circularized exons, circRNAs formed by two or more exons, by exon and retained intron sequences (EI-ciRNAs) and by intronic sequences only.

## CIRC RNA BIOGENESIS

### Back-splicing

As anticipated, circRNA loops are generated by back-splicing from immature RNA, where ends are joined in a non-collinear way. CircRNAs are derived from Pol II transcripts just like linear transcripts. Back-splicing requires the spliceosomal machinery<sup>32–34</sup> as revealed by treatment of HeLa cells with a splice inhibitor followed by nascent RNA purification.

In the majority of cases, the generation of circRNA happens at the expense of their corresponding mRNA isoforms, and is characterized by the usage of canonical splice sites that precisely flank head-to-tail junctions of circular transcripts. In some cases, transcripts of specific genes are predominantly spliced into the circular isoform.<sup>11,12</sup> Ashwal-Fluss *et al.*<sup>32</sup> demonstrated that circularization and splicing of linear forms compete against each other. Kelly *et al.*<sup>35</sup> confirmed a direct correlation between exon skipping and circularization. Thus, circRNA biogenesis and regulation of mRNA production are tightly linked.

### Cis regulatory features

CircRNA-forming exons are often flanked by particularly long introns, possibly reducing splicing efficiency.<sup>12,16</sup> Moreover, in humans these long introns are markedly enriched in ALU repeats,<sup>12</sup> and complementary sequences in introns are involved in specific folding of primary transcripts that favor circularization.<sup>32</sup> In the *Sry* gene, the activation of an upstream promoter triggers the synthesis of a primary transcript containing inverted repeats needed for circRNA production.<sup>36</sup> As *Drosophila* RNA circularization does not appear to be driven by structural complementarity of exon-flanking sequences but only determined by the length of exon-flanking introns, inverted repeats alone do not fully explain the production of circRNAs in eukaryotes. In addition, regulation of the dynamic expression of

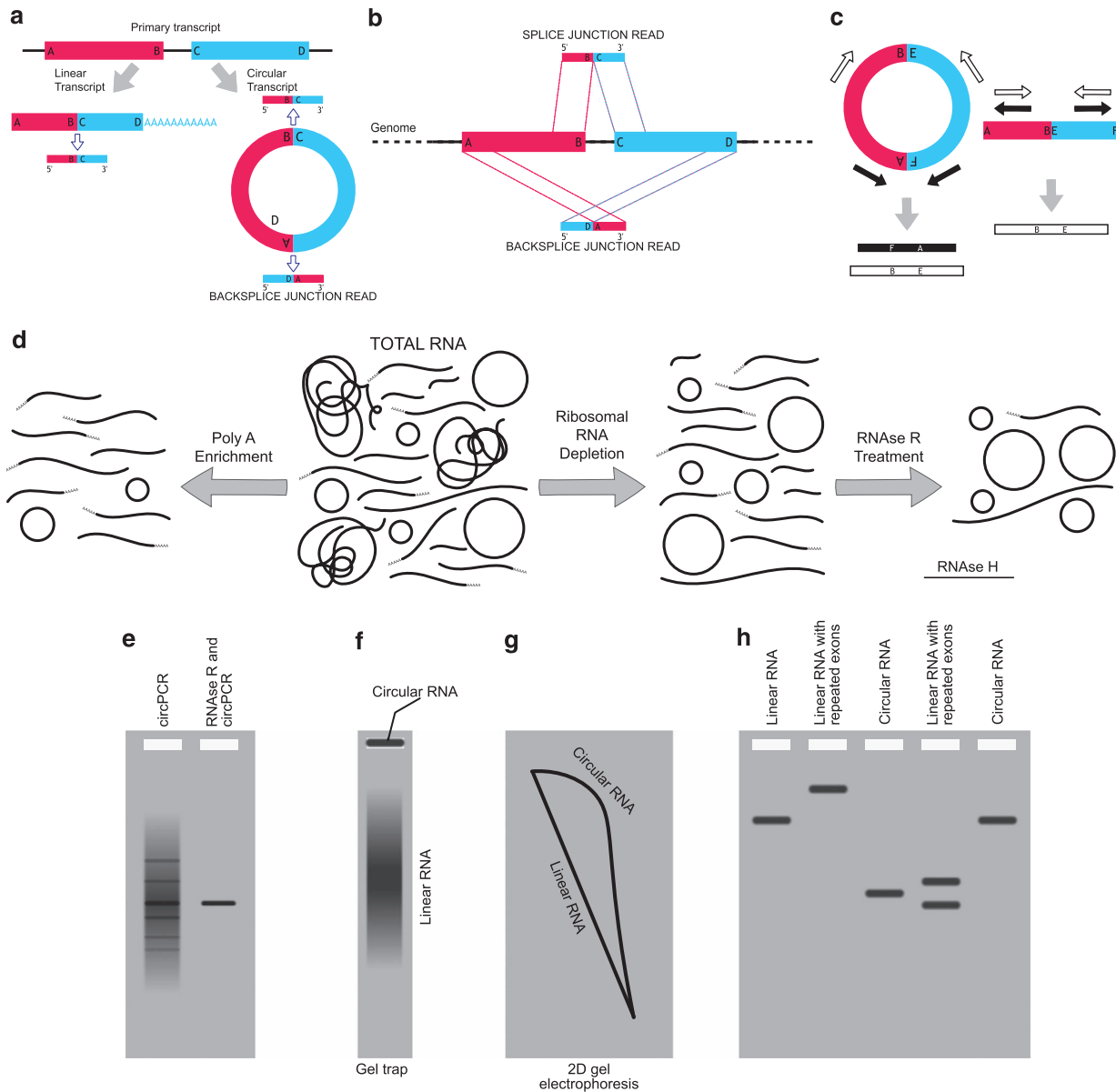
circRNAs in different cell types is likely also dependent on control by *trans*-acting factors.<sup>37</sup>

### Trans-acting factors

Besides the role of flanking sequence elements, introns encasing circRNAs are highly enriched in RNA A-to-I editing events.<sup>12</sup> In fact, knockdown of RNA-editing enzyme ADAR1 upregulated circRNA expression, favoring a mechanism of circRNA biogenesis whereby ADAR1 antagonizes circRNA expression by melting stems of RNA-RNA interactions within introns that putatively promote circularization.<sup>12</sup>

The Muscleblind (MBL) family of splicing factors was also shown to take part in the regulation of circRNA production by binding specific intronic sites flanking circularized exons.<sup>32</sup> Intriguingly, in the fly circRNA isoform expression of *MBL* itself is regulated by MBL protein. Decrease of circRNA expression after MBL knockdown supports a circRNA-promoting role for MBL proteins.<sup>32</sup>

In addition, RNA-binding protein (RBP) Quaking (QKI) regulates the formation of circular RNAs.<sup>38</sup> QKI dimers bind to specific bipartite sequences termed QKI response elements that are present in many RNAs, including coding mRNAs and primary miRNAs. Conn *et al.*<sup>38</sup> investigated the role of QKI in promoting circRNA biogenesis in transforming growth factor-beta-induced epithelial-mesenchymal transition of the epithelial HMLE cell line, demonstrating that the knockdown of the QKI-5 isoform specifically decreases the formation of circRNAs, and that insertion of synthetic QKI response elements in introns mediates circRNA formation. Metabolic tagging of nascent RNAs with 4-thiouridine has been used to study the link between nascent circRNA processing and transcription<sup>17</sup> showing that the efficacy of circRNA processing from primary transcripts is extremely low. This study also clarified that circRNAs are largely processed post-transcriptionally and confirmed that circRNAs are stable, being



**Sequence reconstruction**

CircRNAs length estimation can be obtained by Northern blot or PCR-based methods.

↓

Selective amplification and direct sequencing provide the actual circRNA structure.

**Expression quantification**

- Only RNA-seq reads derived from the backsplice junction are private of the circRNA.
- CircRNA expression estimation is based on the number of detected backsplice junction reads, since most reads fall in sequences shared with linear transcripts.

**Figure 2.** Molecular methods for circRNA detection, validation and study. **(a)** CircRNA detection from RNA-seq data grounds on the identification of sequence reads encompassing the backsplice junction; **(b)** Backsplice reads map to the genome in chiastic order (two segments of a single read align separately in reverse order) due to the backsplicing in circRNAs biogenesis. **(c)** Convergent primers (white arrows) designed in adjacent spliced exons amplify both linear and circular isoforms, whereas primers that are divergent in the linear transcripts (black arrows) can be used to specifically amplify the circular isoform; **(d)** PolyA enrichment protocols deplete circRNAs, whereas ribosome depletion and RNase R protocols enrich circRNAs; **(e)** RNase R digestion before reverse transcription-PCR lowers the amount of false-positive amplicons facilitating circRNA validation; **(f)** Gel Trap electrophoresis allows isolate the circular and linear fractions of the input RNA, as circRNAs are held in the well; **(g)** Two-dimensional acrylamide gel electrophoresis separates the circular RNA fraction in an off-diagonal curve; **(h)** RNA migration in agarose gel before and after a mild RNase H treatment resulting in a single cut per molecule shows that circular molecules bearing a ‘backsplice’ junction are discriminated from linear ones deriving from a duplication event, as only circRNA results in a single band after being cut once (**f-h** re-elaborated from<sup>56</sup>).

**Table 1.** CircRNA discovery, characterization and quantification from RNA-seq data

Name	Aligner	Map steps	Annotation needed	Gives gene names	Only circRNA	Notes	References
CIRI	BWA-Mem	1	No (only for report)	Yes	Yes	Accounts for uncertainty of read mapping to the junction	67
Find_circ	Bowtie	2	No	No	Yes		13
CIRCexplorer	STAR, TopHat	2	Yes	Yes	Yes		15
Testrealign	Segemehl	1	No	No	No	Parses segemehl alignments	88
UROBORUS	TopHat, Bowtie	2	No	Yes	Yes	Do not underestimate expression; filters spurious alignments	89
NCLscan	BWA, Novoalign, BLAT	3	Yes	Yes	No	> 98% precision, test on poly(A)± libraries	90
MapSplice	MapSplice	1	Yes	Yes	No	Circular RNA explicit detection from MapSplice v2m (2/2013)	91
circRNA_finder	STAR	1	No	No	Yes		14
KNIFE	Bowtie	2	Yes	Yes	Yes	Statistical approach to enrich true positives	92
PTESFinder	Bowtie	1	No	No	No		93

Abbreviation: CircRNA, circular RNA. Computational methods available for circRNAs prediction. In the past reads mapping to the reference genome with non-collinear exon junctions were often considered artifacts and discarded; with the rise of interest on fusion transcripts produced from rearranged genomes, *trans*-splicing and circRNAs, RNA-seq aligners were improved to consider also these 'exotic' events, where transcript sequence parts correspond to regions that are remote in the normal reference genome (fusion and *trans*-spliced transcripts) or that map to the genome in chiasmic way (that is, two segments of a single read align separately in reverse order due to the backsplicing in circRNA biogenesis; see also Figure 2b). The table presents 10 computational methods that allow backsplice events identification using different strategies and read aligners; some approaches are specific to predict circRNAs, whereas others were developed with more general purposes, such as read alignment and/or detection of fusion events, and allow also circRNAs detection. For each method, the table indicates the read aligners and number of mapping steps implemented in the prediction strategy, and clarify whether the method requires genome annotations in input, whether it provides an annotation of predicted backsplice junctions in terms of overlapping genes and whether it has been designed explicitly for circRNA identification or is a more general purpose software; the last two columns report additional notes on specific software features and references. Recently, five circRNA prediction tools (circRNA finder, Find\_circ, CIRCexplorer, CIRI and MapSplice) have been compared by evaluating the levels of *bona fide* and false-positive circRNAs based on RNase R resistance data, showing that not in all cases the most abundant circRNAs are true positives, that circRNAs identified by a single method only are in general less reliable and that the combination of at least two methods might increase specificity.<sup>94</sup>

**Table 2.** Web resources dedicated to circRNAs

Web resource	Address	Main features	References
CircBase	<a href="http://www.circbase.org/">http://www.circbase.org/</a>	Simple circRNA database that provides a searchable table of circRNAs identified by several studies (five on human data) that can be useful to sort newly identified from known circRNAs	95
starBase v2.0	<a href="http://starbase.sysu.edu.cn/">http://starbase.sysu.edu.cn/</a>	Focuses on RNA–RNA and protein–RNA interaction networks inferred from CLIP-Seq data sets. Among others, in the miRNA–lncRNAs section, it includes predicted miRNA–circRNA interactions and can be searched to identify known circRNA that potentially sponge a specific miRNA	96
Circ2Traits	<a href="http://gyanxet-beta.com/circdb/">http://gyanxet-beta.com/circdb/</a>	Useful to explore potential associations of circRNAs with diseases based on predicted interactions of circRNAs with disease-associated miRNAs and on the overlap between disease-associated SNPs to circRNA loci	70
CircNet	<a href="http://circnet.mbc.nctu.edu.tw/">http://circnet.mbc.nctu.edu.tw/</a>	Provides expression profiles of circRNAs in 464 RNA-seq samples, with circRNA sequences and annotations in term of overlapping genes and interactions	97

The table describes four websites that store circRNA expression data and provide useful resources to explore possible circRNA functions.

thus abundant at a steady-state level and tending to accumulate particularly in cells with low proliferation rates.<sup>8</sup>

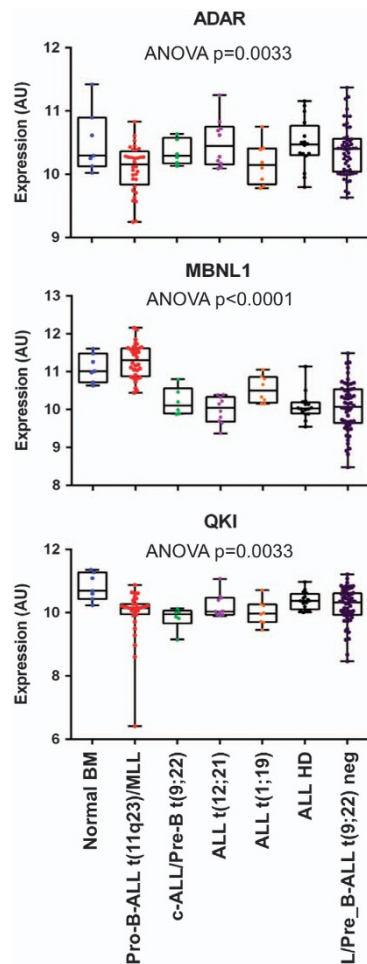
#### Backsplice as a new form of alternative splicing

Alternative RNA splicing is a complex tightly regulated phenomenon. Since the discovery of split genes robust knowledge was built on splicing prevalence, on complexity of splicing patterns and on molecular mechanisms that determine, regulate or change splicing, including RNA–protein interactions (splicing factors with *in cis* regulatory sites termed silencers or enhancers), RNA–RNA base-pairing interactions involving both *in trans* acting RNAs and *in cis* secondary structure formation, and also chromatin-based effects.<sup>39</sup>

Disease-causing mutations occur in splice sites or in regulatory elements, as well as in genes that encode splicing factors (*U2AF1*,

*SRSF2*, *SF3B1* and *ZRSR2*), and there is much interest in developing antisense oligonucleotides to control splicing patterns and using genome editing to correct disease-causing splicing defects. Alternative splicing is highly and commonly deregulated in cancer cells<sup>40–42</sup> and specifically impacts prognosis and disease course of myeloid malignancies, including chronic lymphocytic leukemia, acute lymphoblastic leukemia (ALL), acute myeloid leukemia, and myeloproliferative neoplasms.<sup>43–46</sup>

As circRNA biogenesis and splicing are interleaved processes, it can be hypothesized that mutations of splicing factors and/or alterations of regulatory elements have an impact on circRNA biogenesis. RBP involved in circRNA biogenesis might drive developmental regulation of circRNA formation and show deregulation in disease. Distinct expression levels of ADAR1, MBNL1 and QKI in normal bone marrow compared with B-cell leukemia subtypes (Figure 3) encourage investigations as also



**Figure 3.** Expression variation of enzymes involved in circRNA expression. Gene expression intensities of ADAR1, MBNL1 and QKI in samples of normal bone marrow and six B-cell leukemia subtypes carrying specific genetic aberrations (according to Haferlach *et al.*<sup>87</sup>); expression data obtained with HG-U133 Plus 2.0 (Affymetrix, Santa Clara, CA, USA).

subtle expression variations of ADAR1 were shown to be relevant for RNA circularization.<sup>5</sup>

Alternative splicing is a key mechanism through which fundamental processes during hematopoiesis are regulated,<sup>47</sup> posing the basis to interpret the consequences of genetic variation. Similarly, there is high demand to study circRNA in normal hematopoiesis, to connect biogenic mechanisms with biological functions of circRNAs, to accumulate fundamental knowledge needed to understand disease mechanisms and to inform strategies for therapeutic intervention.

#### CircRNA degradation

As circRNAs are endogenous cell products one might ask which endogenous mechanisms cells have to dispose of circRNA. In general, RNA is degraded by the exosome, a multiprotein complex that reminiscently of the proteasome forms a chamber with helicase activity, which unfolds and then degrades RNA. The degradation is prevalently exoribonucleolytic from the 3'-end, but the exosome catalytic subunit RRP44 also has endonuclease activity.<sup>48</sup>

According to available evidences circRNAs are not degraded by treatments (as tobacco acid pyrophosphatase plus terminator

5'-phosphate dependent exonuclease or highly processive 3'- to 5'-exoribonuclease RNase R digestion) that normally degrade linear RNA with free ends.<sup>31</sup> Regarding degradation, in general miRNAs can regulate cleavage of circRNAs. The better-characterized path toward degradation of a circRNA is that of *CDR1-as* (circular antisense transcript deriving from cerebellar degeneration-related protein 1 locus) that even presenting multiple miR-7 binding sites is completely resistant to miR-7-mediated degradation and also resistant to miR-769-mediated degradation, whereas the binding of miR-671 to *CDR1-as* directs Ago2-slicer-dependent cleavage.<sup>31</sup>

Undoubtedly, our understanding of the regulation of circRNA turnover and endogenous degradation mechanisms is limited. We can hypothesize that not only the deregulation of circRNA synthesis but also its degradation are biologically relevant. For instance, a lack of cleavage might result in undesired circRNA accumulation.

#### CIRC RNA CONSERVATION AND MOLECULAR FUNCTIONS

CircRNAs are evolutionarily conserved

CircRNAs were described in many eukaryotes from yeast to humans<sup>49</sup> and resulted very conserved at the nucleotide level: Memczak *et al.*<sup>13</sup> analyzed sequence conservation within circRNAs and showed that 223 human circRNAs with conserved circularization in mice were significantly more conserved in the third codon positions than exons not engaged in circular forms. CircRNAs are also depleted of polymorphisms in miRNA-binding sites.<sup>50</sup> Beyond apparent sequence conservation, both paralogous and orthologous gene pairs have been reported to express circular transcripts: human *HIPK2* and *HIPK3*, as well as murine *Hipk2/3* produce circRNAs.<sup>12</sup> Also conservation of circRNAs in terms of exonic sequences, bordering intronic sequences, precise backsplice junctions and expression patterns in mammals and to some extent in *Drosophila* has been recently reported.<sup>5</sup> The above indications of evolutionary preservation point to a central position of circRNAs in core biological processes.

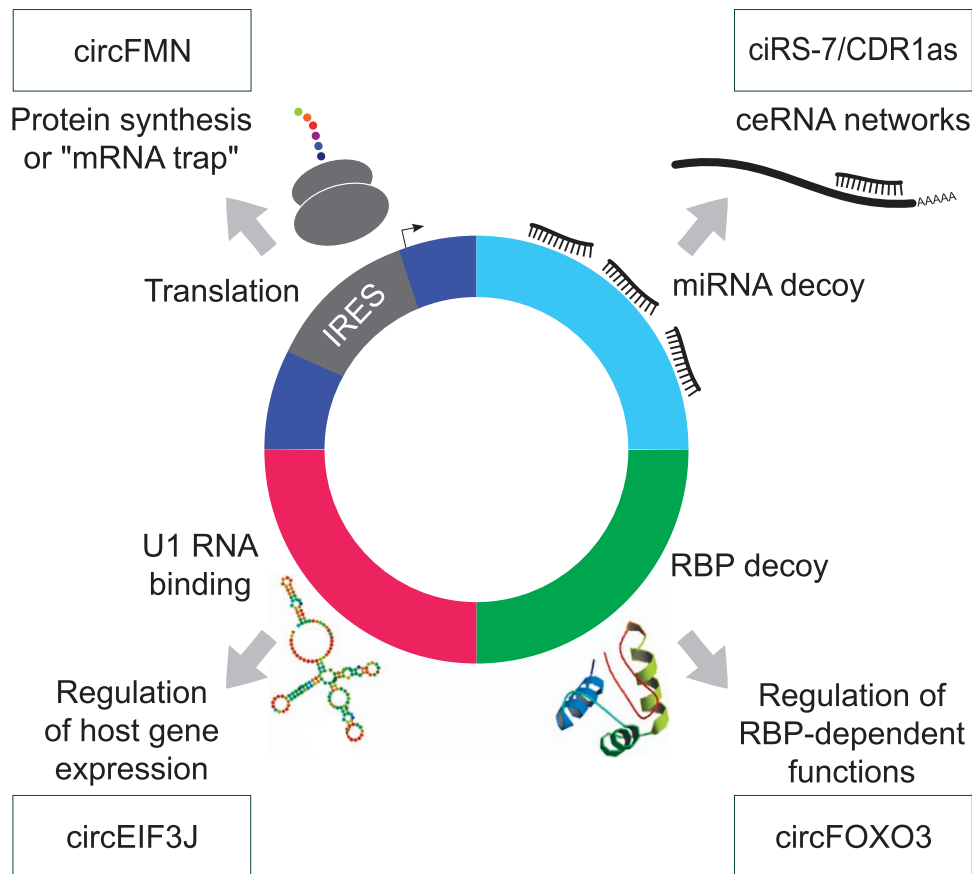
CircRNAs are seldom translated

It has previously been shown that eukaryotic ribosomes can initiate translation on circRNAs containing internal ribosome entry site elements (Figure 4) producing long-repeating polypeptides in the presence of a continuous open reading frame.<sup>51,52</sup> Efficient circRNA translation can occur in HEK293<sup>(ref. 34)</sup> and HeLa cells.<sup>53</sup> Intriguingly, a circRNA can be a sort of 'Möbius strip' with translation generating proteins either recurrently, or variably depending on whether or not the sequence length is a multiple of three nucleotides. A small viroid circRNA directly translated through three completely overlapping open reading frames shifting to a new reading frame at the end of each round has been reported as a natural supercompact 'nanogenome'.<sup>54</sup>

Even if in principle circRNAs can be translated, the majority of recently discovered and characterized circRNAs seem to have limited coding potential: seldomly associated neither with messenger ribonucleoprotein particles nor with translationally active polyribosomes, suggesting that circRNAs, as a species, are unlikely to be translated into peptides.<sup>55</sup> The fact that in the same study mass spectrometry failed to identify peptides encoded by backsplice junctions of circRNAs could be due to low sensitivity or to the position of open reading frames outside junctions and does not rule out that part of circRNAs can be translated in some cell types and/or conditions.

A circRNA produced by murine *Fmn*<sup>21</sup> contains an active translation start site not leading to protein synthesis. In this way, the circular form competes with the linear mRNA both impacting on the linear transcript abundance and providing an 'mRNA trap' that can sequester proteins of the translation initiation complex. Also, Jeck and Sharpless<sup>56</sup> uncovered that many single-exon





**Figure 4.** CircRNA functions. Elucidated circRNA functions include the ability to sponge miRNAs thus regulating the silencing of canonical targets (for example, *ciRS-7/CDR1-as* harbors 63 binding sites for miR-7) and participating to ceRNA networks; similarly circRNAs could decoy RBP ultimately regulating the functions in which RBP are implicated (for example, *circ-Foxo3* forms a ternary complex with p21 and CDK2 arresting cell cycle progression); circRNAs can also regulate *in cis* the expression of the gene from which they derive through interactions with the U1 RNA in the U1 RNP in the nucleus (for example, *circEIF3J*); moreover, circRNAs harboring an IRES could be translated to produce peptides or compete with mRNA translation (for example, *circFMN* contains an active translation start site not leading to the protein synthesis).

circRNAs contain a translation start site,<sup>57</sup> further exemplifying this mechanism of protein expression regulation by circRNAs.

CircRNAs are efficient miRNA sponges and participate to competing endogenous RNA networks

As demonstrated by several studies, circRNAs with multiple miRNA-binding sites are efficient miRNA sponges that participate in the regulation of specific cellular pathways (Figure 4).<sup>24,58</sup>

*CDR1-as* harbors 63 conserved binding sites for miR-7 displaying high miRNA-binding capacity and miRNA antagonist activity in the brain.<sup>13,24</sup> Following this functional description *CDR1-as* was renamed circular transcript *ciRS-7* (circular RNA sponge for miR-7) or called *CDR1-as/ciRS-7*. Notably, the circRNA is completely resistant to miR-7-mediated target destabilization and strongly suppresses miR-7 activity, resulting in increased levels of miR-7 targets, including *EGFR* and *IRS2*.

The previously mentioned *cSRY* decoys miR-138, for which it displays 16 target sites.<sup>12,24</sup> *Circ-ITCH*, a circRNA downregulated in carcinomas, was demonstrated to be a sponge of miR-7, miR-17 and miR-214, by increasing the level of *ITCH* and ultimately inhibiting the Wnt/ $\beta$ -catenin pathway.<sup>58,59</sup> Another circRNA (*hsa\_circ\_001569*) that acts as a sponge for miR-145 upregulates its targets enhancing cell proliferation and invasion of colorectal cancer.<sup>60</sup>

Recently, specific circRNA-miRNA axes have been shown to regulate cancer-related processes. CircRNAs can have both cancer-

promoting and -suppressing roles, depending on the molecular circuits in which they are involved and on the role of the interactors.<sup>58,59,61</sup> CircRNAs can exhibit anticancer effects: as synthetic circular sponges displayed superior anticancer activities compared with the linear sponges, RNA circles open new ways to deliver miRNA sponges with persistent effects.<sup>26</sup>

CircRNAs like linear isoforms can act as competing endogenous RNAs (ceRNAs) that decoy miRNAs and indirectly regulate protein-coding gene expression (Figure 4).<sup>62</sup> ceRNAs are implicated in the progression of cancer and impact on cancer hallmarks.<sup>57</sup> Being resistant to miRNA-mediated degradation circRNA can presumably also tether RISC components depriving the cellular pool of both miRNAs and RISC effectors.<sup>58</sup>

Following the observations of circRNAs acting as ceRNAs it has been asserted that such a mechanism may be common to all circRNAs. The latter finds support from previously mentioned shortage of polymorphisms in circRNAs' putative miRNA-binding sites. Other studies showed instead that only a minority of expressed circRNAs present multiple binding sites for specific miRNAs, and according to their observations circRNAs are, in general, not bound to miRNA-loaded Argonaute proteins.<sup>56,63</sup> In addition, argonaute co-immunoprecipitation experiments did not indicate an appreciable enrichment of circRNA-derived exons among argonaute family-bound transcripts, which would be expected if circRNAs were prevalently acting as ceRNAs. See Thomson *et al.*<sup>64</sup> for a review on evidences and open

questions on endogenous miRNA sponges. According to available data, supported by recent findings<sup>3,59</sup> we conclude that some circRNAs can act as ceRNAs, whereas others may be involved in a variety of other molecular mechanisms.

#### Interactions with RBPs

The decoy activity of circRNAs could be important also for RBPs (Figure 4). CircRNAs like linear RNAs may interact with RBPs in a sequence-specific and structural motif determined way. CircRNAs could function to store, sort or localize RBPs. Recently, the interaction between *Foxo3* circular RNA and specific proteins was shown to delay cell cycle progression.<sup>65</sup> *Foxo3* is a forkhead box O transcription factor and may behave as a tumour suppressor protein that limits cell proliferation and induces apoptosis and is frequently altered in cancer, shown to be deleted in lymphomas (diffuse large B-cell lymphoma), and translocated with MLL in leukemia.<sup>66</sup> In healthy cells high *circ-Foxo3* expression was found to be associated with cell cycle progression. Silencing endogenous *circ-Foxo3* promoted cell proliferation, whereas ectopic expression of *circ-Foxo3* repressed cell cycle progression by binding to cell cycle proteins cyclin-dependent kinase 2 (CDK2) and cyclin-dependent kinase inhibitor 1 (or p21). Normally, CDK2 interacts with cyclin A and cyclin E to facilitate cell cycle entry, while p21 inhibits these interactions and arrests cell cycle progression. The formation of the *circ-Foxo3*-p21-CDK2 ternary complex arrests the function of CDK2 and blocks cell cycle progression. This study identified an oncogenic function of a circRNA and indirectly demonstrated that circRNAs can have distinct functions with respect to that of protein products encoded by the same gene.<sup>65</sup>

#### *Cis* regulation of gene expression by circRNAs

Another line of evidence reported *cis*-regulatory roles for specific circRNAs. Exon-intron circRNAs derived from circularization of RNA with intron retention were identified as a subclass of ciRNAs, enriched in the nucleus, associated with Pol II.<sup>3</sup> Further analyses of two exon-intron circRNAs (*circEIF3J* and *circPAIP2*) showed interactions with Pol II, U1 snRNP and parental gene promoters through sequence complementarity between the U1 snRNA and an U1-binding site, which eventually promote the transcription of the gene from which they derived (host genes), triggering a positive-feedback loop.

Zhang *et al.*<sup>4</sup> described circular intronic RNAs (ciRNAs) that were found to accumulate in human cells due to a failure in debranching and showed that knockdown of ciRNAs reduced expression of their host genes. One of these abundant RNAs, *ci-ankrd52*, largely accumulates to its sites of transcription, associates with the elongation Pol II machinery, and acts as a positive regulator of Pol II transcription.<sup>4</sup> Apparently also non-coding intronic segments of ciRNA transcripts can have a *cis*-regulatory role. *CDR1-as* stabilizes *CDR1* mRNA expression, probably with a sense-antisense-based feedback mechanism, where the antisense circRNA stimulates or stabilizes the sense mRNA with subsequent negative impact on antisense levels.<sup>31</sup>

The fairly well-characterized role of circRNAs to positively modulate host gene expression has foreseeable implications for stability of cell commitment choices such as in hematopoiesis.

In summary, circRNAs can regulate host gene expression but also participate to complex networks in which they compete for the binding of miRNAs (ceRNA networks) of RBPs or even for translation initiation. CircRNAs with different composition in terms of exon-intron inclusion result in a multitude of mechanisms that can affect transcriptome and proteome regulation. Through studies of specific tissues, cell types and conditions the evident versatility of circRNA is expected to reveal insight in all kinds of cellular processes.

## CIRC RNAs IN THE HEMATOPOIETIC COMPARTMENT

RNA-seq analyses showed dynamic expression of circular isoforms independent of linear transcript dynamics from the same gene<sup>5</sup> and cell- and differentiation stage-specific expression<sup>2,13</sup> prompting lines of research that focused on specific tissues. Knowing that circRNAs interfere in key cellular processes like self-renewal, proliferation and apoptosis there is growing interest to study of circRNA in the hematopoietic compartment. In the hematopoietic tissue, pioneering studies reported circRNA isoforms of key genes such as *MLL*<sup>25,67</sup> without receiving much resonance in the research community, probably due to the reported low expression of circular isoforms compared with abundant mRNAs encoding these key transcriptional regulators. In hindsight low expression of circRNAs of the above transcription factor can be understood from recent observations that in specific cell types highly expressed genes (in terms of expressed linear isoforms) give rise to relatively less circRNA compared with moderate- or low-expressed genes.<sup>5</sup>

CircRNA discovery by RNA-seq in hyperdiploid B-cell precursor-ALL and in sorted normal leukocyte cell populations

Aiming to discover new cancer-specific fusion transcripts in hyperdiploid B-lineage acute lymphoblastic leukemia, Salzman *et al.*<sup>11</sup> exploited RNA-seq and found many transcripts with permuted exon order, which they called 'scrambled exons' and attributed to circularized RNAs. In five samples of hyperdiploid B-cell precursor-acute lymphoblastic leukemia they detected hundreds of circRNA transcripts with >700 circular isoforms comprising more than 10% of all transcript isoforms produced from a comparable number of genes. These circRNAs were however not a specific feature of the leukemic cells; PCR verified that scrambled exons were also detected in remission samples of the patients, in HeLa cells and in normal primary human cells. Also in sorted cells populations, naive B cells (CD19+), hematopoietic stem cells (CD34+) and neutrophils, circRNA isoforms expressed by >800 genes were identified, with circRNA expression accounting for >10% of gene expression. This study showed that a particular gene can produce circRNAs in more than one leukocyte type, but single replicate-based preliminary estimations suggest quantitative differences among cell types. This first indication that circRNAs are expressed both in normal and malignant hematopoietic cells informs the number of circRNAs in immature and lineage-specific blood cells but does not provide a more specific and useful interpretation of circRNA relevance for hematopoietic cell functions and pathology. For instance the study reported and validated a few most abundant transcripts with scrambled exons (*ESYT2*, *FBXW4*, *CAMSAP1*, *KIAA0368*, *CLNS1A*, *FAM120A*, *MAP3K1*, *ZKSCAN1*, *MANBA*, *ZBTB46*, *NUP54*, *RARS* and *MGA*) but did not pay particular attention to circRNAs from numerous genes that are important for normal hematopoiesis and present genomic aberrations or deregulated expression in leukemia.

According to published data<sup>11</sup> we observed that circRNAs from genes related to B-cell differentiation and acute lymphoblastic leukemia (*JAK2*, *PAX5*, *IKZF1*, *ETV6* and *EBF1*) are prevalently present in hyperdiploid leukemia compared with normal leukocytes samples. The latter is likely an underestimation, as only 54 genes were screened and circRNAs supported by only 1 read were not considered.

The same authors also investigated circRNA expression in 15 different cancer and non-cancer cell lines, detecting around 47 000 circle-specific splice junctions from 8500 genes. The validation of 8 candidates confirmed that all were true circRNAs.<sup>2</sup> The study further specified that highly expressed circRNA showed cell-type-specific increase in expression that was not associated to an increase of the corresponding linear RNA. Notably, among others, the leukemia cell line K562 presented the

largest number of genes (16 559) with evidence of circular RNA expression.

A reanalysis of the same data set using the CIRI method<sup>67</sup> indicated that more universally shared circRNAs tend to have higher expression levels and verified that the expression patterns of linear transcripts of circRNA-encoding exons are more similar in cancer cells compared with non-cancer cell types, whereas cancer cells appear to have more diverse circRNA expression profiles, both considering exonic and intronic circRNAs.

Subsequently, a comparison of CD34+, CD19+, neutrophils and HEK293 (human embryonic kidney cells) considering only a single biological replicate per cell type (with sequencing depth around 20 million reads per samples) was reported.<sup>13</sup> The study detected 1950 circRNAs of which 939 are exclusively expressed in CD19+ cells, 333 in CD34+, 194 in neutrophils and 60 in HEK293 cells. Nineteen circRNAs resulted to be shared between these cell types. The emphasis of this study was on the demonstration that circRNAs are in part cell-type-specific and are expressed in a developmental stage-related manner.

CircRNAs in whole-blood samples from healthy individuals  
RNA-seq analysis of whole-blood samples<sup>68</sup> showed that whole blood was very rich in circRNAs, comparable to the cerebellum, with consistent data comparing two biological replicates. Also in this study the emphasis laid on the numeric evaluation and the demonstration that circRNAs are a natural component of the transcriptome. Whole blood is composed of a gamma of different cell types. Moreover, the plasma component may also be a sink for circulating circRNA of non-hematopoietic cell origin and can be in fact explored for disease biomarkers, as previously proposed for solid tumors.<sup>9</sup> We cannot exclude that plasma samples could also contain exogenous RNA. In malaria infection, thousands of very short circRNAs are produced by *Plasmodium falciparum*,<sup>69</sup> including dozens of circRNAs harboring >100 binding sites for a given human miRNA, pointing to highly versatile parasite–host interactions. Similarly, as already proposed for virus–host interactions, circRNAs of viral origin might sponge host miRNAs and vice versa.<sup>70</sup>

CircRNAs in platelets

Alhasan *et al.*<sup>71</sup> reported circRNA enrichment in platelets that they ascribed mainly to differential decay of linear RNA, considering the particular circRNA resistance to degradation. In the past integration of transcriptome and proteome data of platelets had given somewhat conflicting results.<sup>72</sup> Extensive degradation of linear RNA isoforms leaving circRNAs intact, which results in an extensive reduction of the translatable RNAs provides now a straightforward explanation for this apparent disparity. This study demonstrated that circRNAs are highly enriched not only in platelets but also in erythrocytes relative to nucleated cells finding that >3000 genes show 17- to 188-fold relative enrichment of circRNAs.

## FUSION-CIRC RNAs DERIVED FROM CHROMOSOMAL TRANSLOCATIONS HAVE ONCOGENIC ROLE

Fusion-circRNA discovery

Very recently fusion-circRNAs (f-circRNAs) derived from transcribed exons of chimeric genes generated by cancer-associated chromosomal translocation were discovered and proven to be oncogenic according to *in vitro* and *in vivo* experiments<sup>73</sup> (Figure 5). Instead of a discovery-driven approach this study used informed guessing to directly detect transcript circularization around the breakpoint/fusion region of two well-known recurrent leukemia-related translocations. The authors hypothesized that

juxtaposition of complementary sequences in introns at either side of the fusion regions could favor the formation of circRNAs and searched specifically for circRNAs expressed from fusion genes. f-circRNAs were thus detected by RT-PCR and then confirmed using RNA-seq and custom bioinformatics procedures, in promyelocytic leukemia with a PML/RAR $\alpha$  and acute myeloid leukemia with an MLL/AF9 fusion (Figure 5a). Both translocations gave rise to more than one f-circRNA characterized by different backsplice junctions, both in patient samples and in patient-derived cell lines. The discovery was also extended to solid tumors showing f-circRNAs transcription in translocated Ewing sarcoma and lung cancer.

f-circRNAs are proto-oncogenic and a requisite for leukemic cells viability

Guarnerio *et al.*<sup>73</sup> showed that f-circRNAs (*f-circPR* and *f-circM9*) expression in leukemic cells increases cell proliferation and clonogenicity and that f-circRNA silencing reverted the phenotype, demonstrating that these f-circRNAs are biologically active and exert pro-proliferative and proto-oncogenic activities (Figure 5b). Moreover, shRNA-based knockdown of *f-circM9* in leukemic THP1 cells resulted in increased apoptosis showing that f-circRNAs have an important role in maintaining the viability of leukemic cells.

*In vivo* study of f-circRNAs

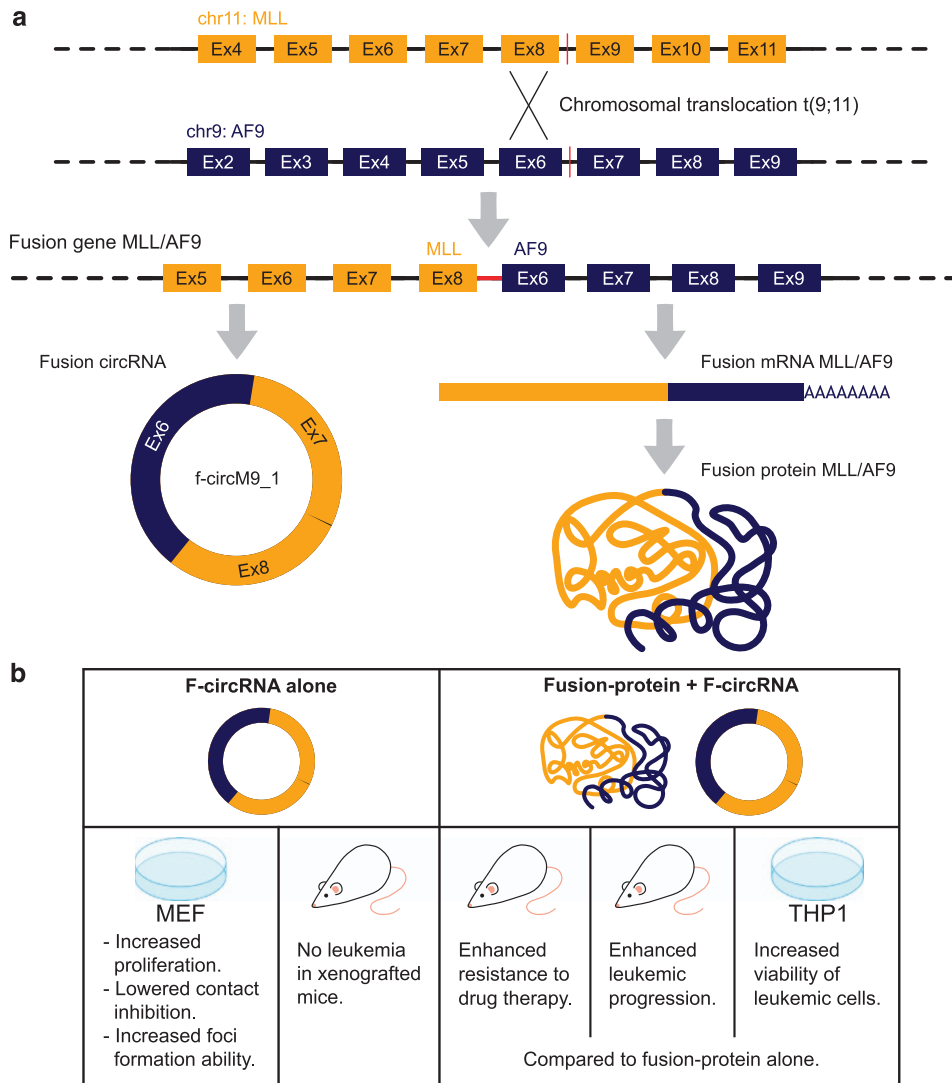
Human leukemic cells *in vivo* expressing f-circRNAs sustain disease progression in mouse. On the other hand, f-circRNAs alone did not trigger leukemia. Cells expressing *f-circM9* together with the MLL/AF9 fusion protein have an increased ability to proliferate and form colonies than cells expressing the fusion protein alone, strengthening the hypothesis that f-circRNAs contribute to the leukemogenic process *in vivo* (Figure 5b). Human leukemic cells *in vivo* expressing f-circRNAs sustain disease progression in mouse. Moreover, f-circRNAs expression could provide tumor cells protection to standard leukemia treatment, as arsenic trioxide, and confer survival advantage to leukemic cells in response in addition to standard-of-care leukemia treatment with cytarabine. Thus, according to these experiments in a pre-clinical setting *f-circM9* could impact therapeutic outcomes.

Interestingly, Guarnerio *et al.*<sup>73</sup> argued that the latency of leukemia development in animal models could be due to the absence of f-circRNAs in modeled expression of intron-less fusion genes. Even if it does provide neither data nor hypotheses on the mechanisms underlying the observed pathogenetic effect of f-circRNA, the study<sup>73</sup> is a breakthrough for leukemia research. In addition, as noticed by the authors, not only the expression of f-circRNA but also the reduction of circRNAs expressed from the non-translocated allele partner genes can contribute to the pathogenic effect.

## CONCLUSIONS AND OUTLOOK

CircRNA expression further challenges a simplistic definition of 'gene', reinforcing the concept that genes are complex transcriptional units, and that the sequence of a given genomic region is a sort of palimpsest (ancient parchment on which the original text was overwritten multiple times) that contains multiple, interleaved and overlapping information parcels.<sup>74</sup> Transcripts from the same locus use a common sequence in different ways, and perform distinct biological roles. In addition, circRNAs add new hints to our understanding of the alternative use and reuse of RNA sequences to produce different products and even small RNAs, as known in the case of miRtrons,<sup>75</sup> of tRNA- or snoRNA-derived miRNAs<sup>74,76</sup> and moRNAs,<sup>77–79</sup> that are expressed in blood cells,<sup>78,79</sup> and were shown to have pathogenic relevance in B-cell lymphomas.<sup>80</sup> Ultimately gene expression studies need to disentangle the





**Figure 5.** f-circRNAs derived from chromosomal translocations have oncogenic role. **(a)** Transcription of fusion genes generated by cancer-associated chromosomal translocation could generate both linear mRNA coding for oncogenic fusion proteins and f-circRNAs.<sup>73</sup> The figure depicts the example of *f-CircM9\_1* expressed in cells harboring the well-known acute myeloid leukemia MLL/AF9 t(9;11) translocation: *f-CircM9\_1* includes two sequences not present in the normal genome, the MLL exon 8 and AF9 exon 6 fusion junction derived from the chromosomal translocation, and the backsplice junction connecting MLL exon 7 with AF9 exon 6; **(b)** *f-CircM9* was demonstrated to be proto-oncogenic *in vitro* (increasing proliferation rate and foci forming ability in mouse embryonic fibroblasts, MEF), and required for leukemic cell (THP1) viability. *f-CircM9* alone resulted not sufficient to trigger leukemia *in vivo* when expressed in HSC xenografted in mice. Concurrent expression of *f-circM9* and MLL/AF9 fusion protein contributed to leukemia progression *in vivo* and *ex vivo* cells expressing *f-circM9* and MLL/AF9 displayed increased ability to proliferate and to form colonies. Furthermore, *f-circM9* expression in MLL/AF9 mouse model cells increased the resistance to leukemia treatments suggesting that *f-circM9* impacts to pre-clinical therapeutic outcome.

expression of linear and circRNAs expressed from each locus in order to dissect the distinct or complementary processes in which they participate.

A large body of data regarding molecular circuits that control cellular differentiation of the hematopoietic system is available and its deregulation in malignancies in association with genomic lesions is increasingly understood. In hematopoiesis, differentiated cell states are controlled by densely interconnected transcriptional circuits<sup>81</sup> in a seemingly hierarchical process of binary fate decisions, but the stiffness of cell fate may be more fluid<sup>82</sup> allowing for epigenetic regulation in response to mature blood cell demand. We envisage that circRNAs studies of hematopoietic cell stages will further elucidate how cell fate fluidity may depend on stably present circRNAs of key cell stage mRNAs.

Gene expression profiling of the protein-coding transcriptome has been very useful for the study of hematopoietic malignancies but will become more complete when integrated with circRNA expression data. Among other things it can be expected to elucidate discordant gene–protein expression often revealed for marker proteins of hematopoietic cell stages (for example, CD10, CD22 and CD38). Further, induced pluripotent stem cell modeling of hematopoietic diseases may also benefit from circRNA studies, hitherto nothing is known about the behavior of circRNAs in reprogramming procedures of induced pluripotent stem cell generation.

Recent RNA-seq data outlined transcriptional diversity in terms of (linear) alternative isoform-ratio variations among hematopoietic cells<sup>47,83</sup> and of non-coding RNA's impact in hematopoietic

lineage differentiation.<sup>84</sup> Specific miRNAs are expressed in a developmental-stage-specific manner.<sup>85,86</sup> miRNAs and other small RNAs are differentially expressed in disease.<sup>79</sup>

As in last years we abandoned the concept of the centrality of coding fraction of the transcriptome, the discovery of circRNAs made clear that the study of the linear RNAs only provides an incomplete picture of the cellular complexity. Focusing on linear transcription only we miss important elements, both for data interpretation and experimental design.

Today we have an increased appreciation of circRNA abundance, evolutionary conservation and diversity of functions and interactions. Specific data emerged of high and regulated circRNA expression in normal and malignant blood cells. The recent discovery and functional study of f-circRNAs provided important clues of the oncogenic role of this aberrant circRNA in leukemogenesis and of their relevance in modulating therapeutic outcome. Together these data clearly indicate that further studies of circular isoforms from different cell types and stages of the hematopoietic compartment and by rearranged or mutated genomes are warranted to better estimate the position of these new regulators in hematopoietic cell development and derived malignancies. The route toward elucidating circRNA biology is still long. Even a consistent nomenclature for circRNAs is sorely missed.

Definitely, circRNAs and their diverse molecular interactions participate to the circuitries that regulate the final cellular protein output adding to the richness and complexity of the underlying mechanisms. The particular stability of circRNAs may also make them valuable disease markers that can be identified in various body fluids and we envisage that a better understanding of circRNA biology will inform innovative therapeutic targets.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank for financial support to SB Fondazione Cassa di Risparmio di Padova e Rovigo Progetti di Eccellenza 2011/2012 by Ministero dell'Istruzione, dell'Università e della Ricerca PRIN 2010/11 (2010NYKNS7\_002) and University of Padova; to GteK Fondazione Cariplo, Ministero dell'Istruzione, dell'Università e della Ricerca FIRB RBAP11TF7Z and University of Padova. AB is recipient of a PhD fellowship from the Fondazione Città della Speranza. Silvia Bresolin, Luca Trentin, Andrea Bisognin, Giulia Anselmi, Elena Boldrin and Lueder H Meyer are acknowledged for fruitful discussions.

## AUTHOR CONTRIBUTIONS

Study design and concept: all authors; bibliographic review: AB, GteK and SB; bioinformatic method review: AB and EG; figure concepts and drawing: AB and SB; manuscript writing: GteK and SB; critical review and approval of the final manuscript: all authors.

## REFERENCES

- Lasda E, Parker R. Circular RNAs: diversity of form and function. *RNA* 2014; **20**: 1829–1842.
- Salzman J, Chen RE, Olsen MN, Wang PL, Brown PO. Cell-type specific features of circular RNA expression. *PLoS Genet* 2013; **9**: e1003777.
- Li Z, Huang C, Bao C, Chen L, Lin M, Wang X et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* 2015; **22**: 256–264.
- Zhang Y, Zhang X-O, Chen T, Xiang J-F, Yin Q-F, Xing Y-H et al. Circular intronic long noncoding RNAs. *Mol Cell* 2013; **51**: 792–806.
- Rybak-Wolf A, Stottmeister C, Glažar P, Jens M, Pino N, Giusti S et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol Cell* 2015; **58**: 870–885.
- Dean M, Fojo T, Bates S. Tumour stem cells and drug resistance. *Nat Rev Cancer* 2005; **5**: 275–284.

- Costa D, Gigoni A, Würth R, Cancedda R, Florio T, Pagano A. Metformin inhibition of neuroblastoma cell proliferation is differently modulated by cell differentiation induced by retinoic acid or overexpression of NDM29 non-coding RNA. *Cancer Cell Int* 2014; **14**: 59.
- Bachmayr-Heyda A, Reiner AT, Auer K, Sukhbaatar N, Aust S, Bachleitner-Hofmann T et al. Correlation of circular RNA abundance with proliferation-exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues. *Sci Rep* 2015; **5**: 8057.
- Li P, Chen S, Chen H, Mo X, Li T, Shao Y et al. Using circular RNA as a novel type of biomarker in the screening of gastric cancer. *Clin Chim Acta* 2015; **444**: 132–136.
- Bahn JH, Zhang Q, Li F, Chan T-M, Lin X, Kim Y et al. The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva. *Clin Chem* 2015; **61**: 221–230.
- Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 2012; **7**: e30733.
- Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 2013; **19**: 141–157.
- Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013; **495**: 333–338.
- Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P et al. Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep* 2014; **9**: 1966–1980.
- Zhang Z, Qi S, Tang N, Zhang X, Chen S, Zhu P et al. Discovery of replicating circular RNAs by RNA-seq and computational algorithms. *PLoS Pathog* 2014; **10**: e1004553.
- Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep* 2015; **10**: 170–177.
- Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang J-L et al. The Biogenesis of Nascent Circular RNAs. *Cell Rep* 2016; **15**: 611–624.
- Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci USA* 1976; **73**: 3852–3856.
- Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert J, Oliner JD et al. Scrambled exons. *Cell* 1991; **64**: 607–613.
- Cocquerelle C, Mascrez B, Hetuin D, Bailleul B. Mis-splicing yields circular RNA molecules. *FASEB J* 1993; **7**: 155–160.
- Chao CW, Chan DC, Kuo A, Leder P. The mouse formin (Fmn) gene: abundant circular RNA transcripts and gene-targeted deletion analysis. *Mol Med* 1998; **4**: 614.
- Pasman Z, Been M, Garcia-Blanco M. Exon circularization in mammalian nuclear extracts. *RNA* 1996; **2**: 603.
- Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell* 1993; **73**: 1019–1030.
- Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK et al. Natural RNA circles function as efficient microRNA sponges. *Nature* 2013; **495**: 384–388.
- Caldas C, So CW, MacGregor A, Ford AM, McDonald B, Chan LC et al. Exon scrambling of MLL transcripts occur commonly and mimic partial genomic duplication of the gene. *Gene* 1998; **208**: 167–176.
- Bailleul B. During in vivo maturation of eukaryotic nuclear mRNA, splicing yields excised exon circles. *Nat Struct Mol Biol* 1996; **24**: 1015–1019.
- Zaphiropoulos PG. 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* 1997; **17**: 2985–2993.
- Li X-F, Lytton J. A circularized sodium-calcium exchanger exon 2 transcript. *J Biol Chem* 1999; **274**: 8153–8160.
- Surono A, Takeshima Y, Wibawa T, Ikezawa M, Nonaka I, Matsuo M. Circular dystrophin RNAs consisting of exons that were skipped by alternative splicing. *Hum Mol Genet* 1999; **8**: 493–500.
- Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet* 2010; **6**: e1001233.
- Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ et al. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* 2011; **30**: 4414–4422.
- Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M et al. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 2014; **56**: 55–66.
- Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung L-H et al. Exon circularization requires canonical splice signals. *Cell Rep* 2015; **10**: 103–111.
- Wang Y, Wang Z. Efficient backsplicing produces translatable circular mRNAs. *RNA* 2015; **21**: 172–179.

- 35 Kelly S, Greenman C, Cook PR, Papanonis A. Exon skipping is correlated with exon circularization. *J Mol Biol* 2015; **427**: 2414–2417.
- 36 Dolci S, Grimaldi P, Geremia R, Pesce M, Rossi P. Identification of a promoter region generating Sry circular transcripts both in germ cells from male adult mice and in male mouse embryonal gonads. *Biol Reprod* 1997; **57**: 1128–1135.
- 37 Salzman J. Circular RNA expression: its potential regulation and function. *Trends Genet* 2016; **32**: 309–316.
- 38 Conn SJ, Pillman KA, Toubia J, Conn VM, Salamanidis M, Phillips CA *et al*. The RNA binding protein quaking regulates formation of circRNAs. *Cell* 2015; **160**: 1125–1134.
- 39 Lee Y, Rio DC. Mechanisms and regulation of alternative pre-mRNA splicing. *Annu Rev Biochem* 2015; **84**: 291–323.
- 40 Liu S, Cheng C. Alternative RNA splicing and cancer. *Wiley Interdiscip Rev RNA* 2013; **4**: 547–566.
- 41 Bisognin A, Pizzini S, Perilli L, Esposito G, Mocellin S, Nitti D *et al*. An integrative framework identifies alternative splicing events in colorectal cancer development. *Mol Oncol* 2014; **8**: 129–141.
- 42 Yoshida K, Ogawa S. Splicing factor mutations and cancer. *Wiley Interdiscip Rev RNA* 2014; **5**: 445–459.
- 43 Rossi D, Brusca G, Spina V, Rasi S, Khiabani H, Messina M *et al*. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* 2011; **118**: 6904–6908.
- 44 Martínez-Avilés L, Besses C, Álvarez-Larrán A, Camacho L, Pairet S, Fernández-Rodríguez C *et al*. Mutations in the RNA splicing machinery genes in myelofibrotic transformation of essential thrombocythaemia and polycythaemia vera. *Br J Haematol* 2014; **164**: 605–607.
- 45 Tefferi A, Lasho T, Finke C, Knudson R, Ketterling R, Hanson C *et al*. CALR vs JAK2 vs MPL-mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. *Leukemia* 2014; **28**: 1472–1477.
- 46 Hou H-A, Liu C-Y, Kuo Y-Y, Chou W-C, Tsai C-H, Lin C-C *et al*. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. *Oncotarget* 2016; **7**: 9084–9101.
- 47 Chen L, Kostadima M, Martens JH, Canu G, Garcia SP, Turro E *et al*. Transcriptional diversity during lineage commitment of human blood progenitors. *Science* 2014; **345**: 1251033.
- 48 Makino DL, Halbach F, Conti E. The RNA exosome and proteasome: common principles of degradation control. *Nat Rev Mol Cell Biol* 2013; **14**: 654–660.
- 49 Chen L-L, Yang L. Regulation of circRNA biogenesis. *RNA Biol* 2015; **12**: 381–388.
- 50 Thomas LF, Sætrum P. Circular RNAs are depleted of polymorphisms at microRNA binding sites. *Bioinformatics* 2014; **30**: 2243–2246.
- 51 Bretscher MS. Translocation in protein synthesis: a hybrid structure model. *Nature* 1968; **218**: 675–677.
- 52 Chen C-Y, Sarnow P. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* 1995; **268**: 415.
- 53 Abe N, Matsumoto K, Nishihara M, Nakano Y, Shibata A, Maruyama H *et al*. Rolling circle translation of circular RNA in living human cells. *Sci Rep* 2015; **5**: 16435.
- 54 AbouHaidar MG, Venkataraman S, Golshani A, Liu B, Ahmad T. Novel coding, translation, and gene expression of a replicating covalently closed circular RNA of 220 nt. *Proc Natl Acad Sci USA* 2014; **111**: 14542–14547.
- 55 You X, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G *et al*. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat Neurosci* 2015; **18**: 603–610.
- 56 Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nat Biotechnol* 2014; **32**: 453.
- 57 Qu S, Yang X, Li X, Wang J, Gao Y, Shang R *et al*. Circular RNA: a new star of noncoding RNAs. *Cancer Lett* 2015; **365**: 141–148.
- 58 Li F, Zhang L, Li W, Deng J, Zheng J, An M *et al*. Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/β-catenin pathway. *Oncotarget* 2015; **6**: 6001–6013.
- 59 Huang G, Zhu H, Shi Y, Wu W, Cai H, Chen X. cir-ITCH plays an inhibitory role in colorectal cancer by regulating the Wnt/β-catenin pathway. *PLoS One* 2015; **10**: e0131225.
- 60 Xie H, Ren X, Xin S, Lan X, Lu G, Ss Y *et al*. Emerging roles of circRNA001569 targeting miR-145 in the proliferation and invasion of colorectal cancer. *Oncotarget* 2016; **7**: 26680–26691.
- 61 Wang K, Long B, Liu F, Wang J-X, Liu C-Y, Zhao B *et al*. A circular RNA protects the heart from pathological hypertrophy and heart failure by targeting miR-223. *Eur Heart J* 2016; **37**: 2602–2611.
- 62 Ala U, Karreth FA, Borgia C, Pagnani A, Taulli R, Léopold V *et al*. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc Natl Acad Sci USA* 2013; **110**: 7154–7159.
- 63 Guo JU, Agarwal V, Guo H, Bartel DP. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* 2014; **15**: 409.
- 64 Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet* 2016; **17**: 272–283.
- 65 Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nat Struct Mol Biol* 2016; **44**: 2846–2858.
- 66 Zuna J, Burjanivova T, Mejstrikova E, Zemanova Z, Muzikova K, Meyer C *et al*. Covert preleukemia driven by MLL gene fusion. *Genes Chromosomes Cancer* 2009; **48**: 98–107.
- 67 Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol* 2015; **16**: 4.
- 68 Memczak S, Papavasileiou P, Peters O, Rajewsky N. Identification and characterization of circular RNAs as a new class of putative biomarkers in human blood. *PLoS one* 2015; **10**: e0141214.
- 69 Broadbent KM, Broadbent JC, Ribacke U, Wirth D, Rinn JL, Sabeti PC. Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics* 2015; **16**: 454.
- 70 Ghosal S, Das S, Sen R, Basak P, Chakrabarti J. Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits. *Front Genet* 2013; **4**: 283.
- 71 Alhasan AA, Izougu OG, Al-Balool HH, Steyn JS, Evans A, Colzani M *et al*. Circular RNA enrichment in platelets is a signature of transcriptome degradation. *Blood* 2016; **127**: e1–e11.
- 72 Londin ER, Hatzimichael E, Loher P, Edelstein L, Shaw C, Delgrosso K *et al*. The human platelet: strong transcriptome correlations among individuals associate weakly with the platelet proteome. *Biol Direct* 2014; **9**: 3.
- 73 Guarnerio J, Bezzi M, Jeong JC, Paffenholz SV, Berry K, Naldini MM *et al*. Oncogenic role of fusion-circRNAs derived from cancer-associated chromosomal translocations. *Cell* 2016; **165**: 289–302.
- 74 Tuck AC, Tollervey D. RNA in pieces. *Trends Genet* 2011; **27**: 422–432.
- 75 Schamberger A, Sarkadi B, Orbán TI. Human mirtrons can express functional microRNAs simultaneously from both arms in a flanking exon-independent manner. *RNA Biol* 2012; **9**: 1177–1185.
- 76 Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, dicer-dependent small RNAs. *Genes Dev* 2008; **22**: 2773–2785.
- 77 Bortoluzzi S, Biasiolo M, Bisognin A. MicroRNA–offset RNAs (moRNAs): by-product spectators or functional players? *Trends Mol Med* 2011; **17**: 473–474.
- 78 Bortoluzzi S, Bisognin A, Biasiolo M, Guglielmelli P, Biomonte F, Norfo R *et al*. Characterization and discovery of novel miRNAs and moRNAs in JAK2V617F-mutated SET2 cells. *Blood* 2012; **119**: e120–e130.
- 79 Guglielmelli P, Bisognin A, Saccoman C, Mannarelli C, Coppe A, Vannucchi AM *et al*. Small RNA sequencing uncovers new miRNAs and moRNAs differentially expressed in normal and primary myelofibrosis CD34+ cells. *PLoS One* 2015; **10**: e0140445.
- 80 Maute RL, Schneider C, Sumazin P, Holmes A, Califano A, Basso K *et al*. tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci USA* 2013; **110**: 1404–1409.
- 81 Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, McConkey ME *et al*. Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 2011; **144**: 296–309.
- 82 Notta F, Zandi S, Takayama N, Dobson S, Gan OI, Wilson G *et al*. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* 2016; **351**: aab2116.
- 83 Shi L, Lin Y-H, Sierant MC, Zhu F, Cui S, Guan Y *et al*. Developmental transcriptome analysis of human erythropoiesis. *Hum Mol Genet* 2014; **23**: 4528–4542.
- 84 Morlando M, Ballarino M, Fatica A. Long non-coding RNAs: new players in hematopoiesis and leukemia. *Front Med (Lausanne)* 2015; **2**: 23.
- 85 O'Connell RM, Chaudhuri AA, Rao DS, Gibson WSJ, Balazs AB, Baltimore D. MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output. *Proc Natl Acad Sci USA* 2010; **107**: 14235–14240.
- 86 Zhang L, Sankaran VG, Lodish HF. MicroRNAs in erythroid and megakaryocytic differentiation and megakaryocyte-erythroid progenitor lineage commitment. *Leukemia* 2012; **26**: 2310–2316.
- 87 Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Bene MC *et al*. Clinical utility of microarray-based gene expression profiling in the diagnosis and sub-classification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. *J Clin Oncol* 2010; **28**: 2529–2537.
- 88 Otto C, Stadler PF, Hoffmann S. Lacking alignments? The next-generation sequencing mapper segemehl revisited. *Bioinformatics* 2014; **30**: 1837–1843.
- 89 Song X, Zhang N, Han P, Moon B-S, Lai RK, Wang K *et al*. Circular RNA profile in gliomas revealed by identification tool UROBORUS. *Nucleic Acids Res* 2016; **44**: e87.

- 90 Chuang T-J, Wu C-S, Chen C-Y, Hung L-Y, Chiang T-W, Yang M-Y. NCLscan: accurate identification of non-co-linear transcripts (fusion, trans-splicing and circular RNA) with a good balance between sensitivity and precision. *Nucleic Acids Res* 2015; **44**: e29.
- 91 Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, Savich GL *et al*. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nat Struct Mol Biol* 2010; **38**: e178.
- 92 Szabo L, Morey R, Palpant NJ, Wang PL, Afari N, Jiang C *et al*. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. *Genome Biol* 2015; **16**: 126.
- 93 Izuogu OG, Alhasan AA, Alafghani HM, Santibanez-Koref M, Elliott DJ, Jackson MS. PTESFinder: a computational method to identify post-transcriptional exon shuffling (PTES) events. *BMC Bioinformatics* 2016; **17**: 1–11.
- 94 Hansen TB, Venø MT, Damgaard CK, Kjems J. Comparison of circular RNA prediction tools. *Nucleic Acids Res* 2015; **44**: e58.
- 95 Glažar P, Papavasileiou P, Rajewsky N. circBase: a database for circular RNAs. *RNA* 2014; **20**: 1666–1670.
- 96 Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nat Struct Mol Biol* 2013; **42**: D92–D97.
- 97 Liu Y-C, Li J-R, Sun C-H, Andrews E, Chao R-F, Lin F-M *et al*. CircNet: a database of circular RNAs derived from transcriptome sequencing data. *Nucleic Acids Res* 2015; **44**: D209–D215.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016