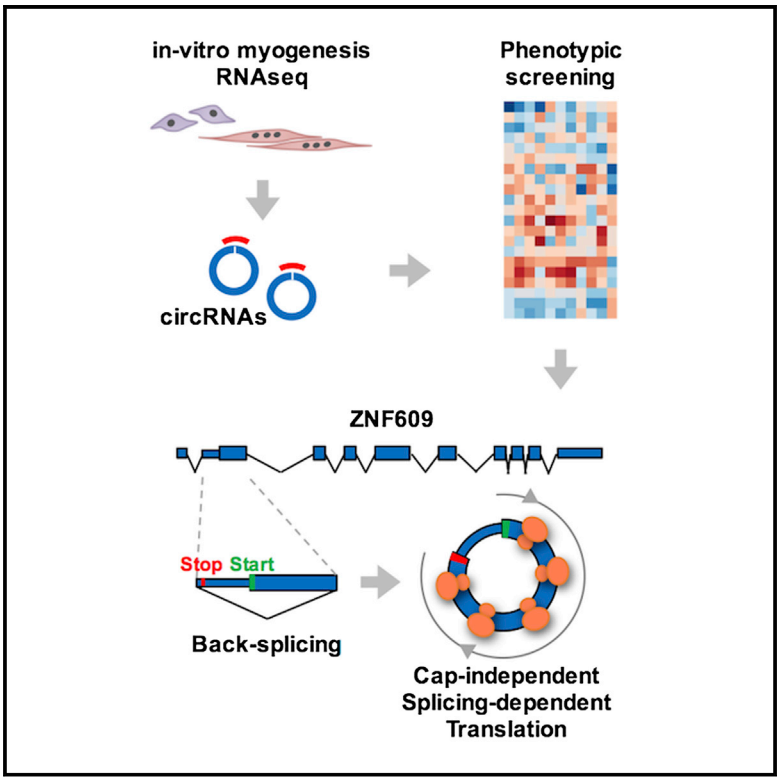
Article

Molecular Cell

Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis

Graphical Abstract



Authors

Ivano Legnini, Gaia Di Timoteo, Francesca Rossi, ..., Pietro Laneve, Nikolaus Rajewsky, Irene Bozzoni

Correspondence

irene.bozzoni@uniroma1.it

In Brief

Legnini et al. identified circ-ZNF609, a circular RNA expressed in murine and human myoblasts, which controls myoblast proliferation. Circ-ZNF609 contains an open reading frame and is translated into a protein in a splicing-dependent/cap-independent manner. Circ-ZNF609 translation can be modulated by stress conditions.

Highlights

- CircRNAs are conserved, abundant, and regulated in myogenesis
- High-throughput phenotypic screening reveals functional circRNAs
- Circ-ZNF609 regulates myoblast proliferation
- Circ-ZNF609 can be translated

Legnini et al., 2017, Molecular Cell 66, 22–37
April 6, 2017 © 2017 The Authors. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.molcel.2017.02.017

CellPress

Molecular Cell
Article

Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis

Ivano Legnini,[1] Gaia Di Timoteo,[1] Francesca Rossi,[1] Mariangela Morlando,[1] Francesca Briganti,[1] Olga Sthandier,[1] Alessandro Fatica,[1] Tiziana Santini,[2] Adrian Andronache,[3] Mark Wade,[3] Pietro Laneve,[2] Nikolaus Rajewsky,[4] and Irene Bozzoni[1,2,5,6,*]

[1]Department of Biology and Biotechnology Charles Darwin and IBPM, Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy
[2]Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161 Rome, Italy
[3]Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia (IIT), Via Adamello 16, 20139 Milan, Italy
[4]Berlin Institute for Medical Systems Biology, Max-Delbrück Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany
[5]Institut Pasteur Italy, Fondazione Cenci-Bolognetti, Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy
[6]Lead Contact
*Correspondence: irene.bozzoni@uniroma1.it
http://dx.doi.org/10.1016/j.molcel.2017.02.017

SUMMARY

Circular RNAs (circRNAs) constitute a family of transcripts with unique structures and still largely unknown functions. Their biogenesis, which proceeds via a back-splicing reaction, is fairly well characterized, whereas their role in the modulation of physiologically relevant processes is still unclear. Here we performed expression profiling of circRNAs during in vitro differentiation of murine and human myoblasts, and we identified conserved species regulated in myogenesis and altered in Duchenne muscular dystrophy. A high-content functional genomic screen allowed the study of their functional role in muscle differentiation. One of them, circ-ZNF609, resulted in specifically controlling myoblast proliferation. Circ-ZNF609 contains an open reading frame spanning from the start codon, in common with the linear transcript, and terminating at an in-frame STOP codon, created upon circularization. Circ-ZNF609 is associated with heavy polysomes, and it is translated into a protein in a splicing-dependent and cap-independent manner, providing an example of a protein-coding circRNA in eukaryotes.

INTRODUCTION

RNA molecules lacking 5′ and 3′ ends and shaped as covalently closed circular RNAs (circRNAs) were described decades ago in a variety of model systems, but they were considered rare events (Sanger et al., 1976; Capel et al., 1993). However, recent studies report that these molecules are commonly produced by thousands of genes (Salzman et al., 2012; Jeck et al., 2013; Memczak et al., 2013) and are usually generated by the joining of a 5′ splice site with the 3′ splice site of an upstream intron, in a so-called back-splicing reaction (Hansen et al., 2011; Ashwal-Fluss et al., 2014; Starke et al., 2015; Conn et al., 2015). It has been shown that circRNA expression is often highly conserved across species and particularly abundant in mammalian neuronal tissues (Rybak-Wolf et al., 2015). Moreover, the DNA that encodes circRNAs is more conserved than the DNA of flanking exons (Rybak-Wolf et al., 2015). Very little is known about their mechanism of action; some may be sponges for microRNAs (miRNAs), as shown for CDR1as and SRY in vertebrate neuronal tissues (Memczak et al., 2013; Hansen et al., 2013) and for circ-HIPK3 in human cancers (Zheng et al., 2016). Nevertheless, genome-wide studies have demonstrated that miRNA sponging activity cannot be generally applied (Memczak et al., 2013; You et al., 2015; Jeck and Sharpless, 2014), and other mechanisms have also been proposed, such as acting as platforms for protein interaction (Hentze and Preiss, 2013). This specifically pertains to circ-Mbl, which competes for the splicing of its linear counterpart by binding to Mbl (Ashwal-Fluss et al., 2014), and to circ-Foxo3, which blocks cell-cycle progression by interacting with Cdk (Du et al., 2016). Emerging evidence also suggests a potential role of circRNAs in different human diseases (Chen et al., 2016), with data indicating tumor-promoting properties in in vivo models (Guarnerio et al., 2016). Despite these data in favor of circRNAs' relevant functional roles, their impact on biological processes is still largely unexplored.

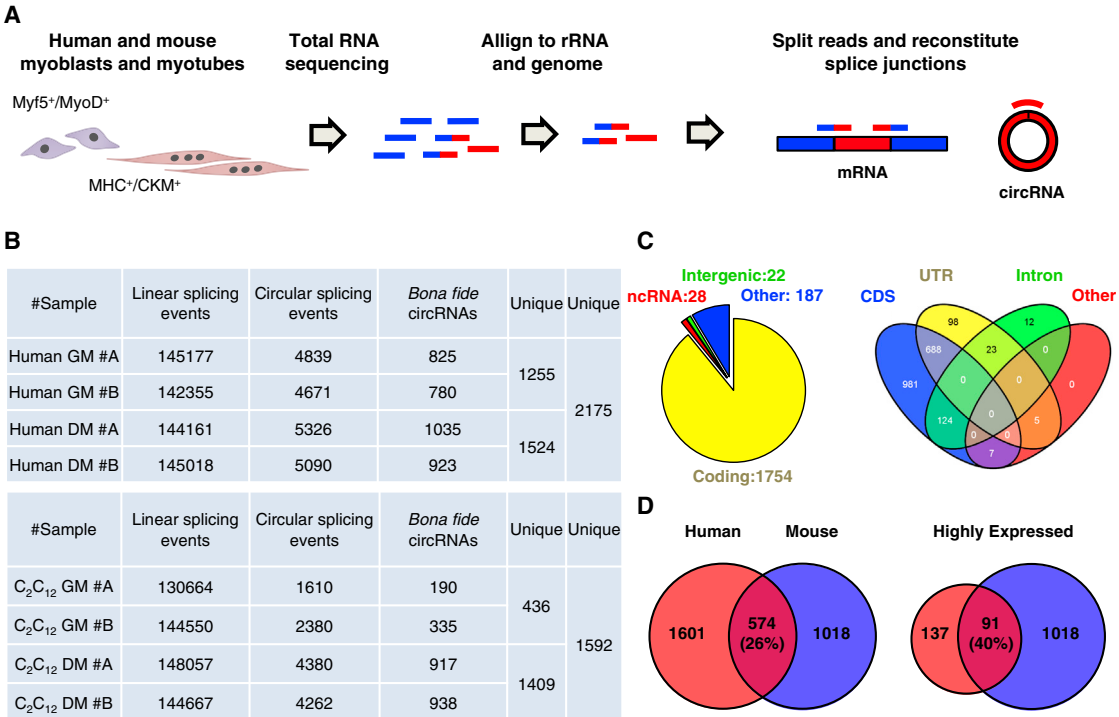
In this work, we have identified conserved circRNAs that are regulated during murine and human muscle differentiation and whose expression is altered in Duchenne muscular dystrophy (DMD) myoblasts. A dedicated knockdown strategy followed by a high-content phenotypic screening identified circRNAs' participation in the control of myogenesis; circ-ZNF609, selected on the basis of its ability to regulate myoblast proliferation, provided an interesting example of translation occurring on a circRNA.

RESULTS

CircRNAs Are Abundant, Conserved, and Highly Expressed

Total RNA was collected from two replicates of human and mouse (C_2C_{12}) myoblasts cultured in growth medium (GM) or

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).



#Sample	Linear splicing events	Circular splicing events	*Bona fide* circRNAs	Unique	Unique
Human GM #A	145177	4839	825	1255	2175
Human GM #B	142355	4671	780		
Human DM #A	144161	5326	1035	1524	
Human DM #B	145018	5090	923		

#Sample	Linear splicing events	Circular splicing events	*Bona fide* circRNAs	Unique	Unique
C_2C_{12} GM #A	130664	1610	190	436	1592
C_2C_{12} GM #B	144550	2380	335		
C_2C_{12} DM #A	148057	4380	917	1409	
C_2C_{12} DM #B	144667	4262	938		

Figure 1. Identification of CircRNAs in Mammalian Myoblast Differentiation

(A) Experimental and computational pipelines for the identification of circRNAs from human and mouse myoblasts (expressing MyoD and Myf5) and myotubes (expressing Myosin Heavy Chain [MHC] and Muscle Creatin Kinase [CKM]). Blue, rRNA and genomic reads; blue-red, reads across linear splice junctions; red, reads across back-splicing junctions.

(B) Number of detected linear and circular splicing events per sample with bona fide circRNAs passing selection filters (see the STAR Methods). Samples are human primary and mouse C_2C_{12} myoblasts (GM) and myotubes (DM) in two replicates (A and B).

(C) Left panel: genomic annotation of circRNAs. Right panel: structural annotation of circRNAs mapping to coding regions is shown.

(D) Overlap of unique circRNAs in human and mouse samples before and after (left and right) filtering for expression level higher than five unique reads in human samples.

See also Figure S1 and Table S1.

induced to differentiate into myotubes (differentiation medium, DM). Paired-end ribominus RNA sequencing (RNA-seq) was performed, and the FindCirc computational pipeline (Memczak et al., 2013) was applied in order to detect circRNAs. RNA-seq data were first processed for gene expression analysis: reads were mapped with TopHat and the transcriptome was reconstituted with Cufflinks (data summarized in Figure S1A). Differential gene expression analysis was performed in order to check the quality of the in vitro differentiation experiments. While fragments per kilobase of transcript per million mapped reads (FPKMs) of genes in technical replicates were almost perfectly correlated (>98% correlation; Figure S1B, upper panels), gene expression in myotubes compared to myoblasts was highly diverse, with >3,000 genes significantly upregulated or downregulated in one of the two conditions, in both human and mouse systems (Figure S1B, lower panels). When looking at the enrichment of gene products upregulated in myotubes versus myoblasts in terms of gene ontology (GO) biological process keywords, we found a consistent and significant enrichment for genes related to muscle differentiation and function (Figure S1C). Moreover, the expression of well-known specific markers of muscle differentiation was assessed by qRT-PCR, which confirmed the data obtained by RNA-seq (Figure S1D).

RNA-seq reads were then realigned to the reference genomes, the ones contiguously aligned were discarded, and the remaining reads were used as input for FindCirc (Figure 1A) to detect head-to-tail splicing sites. As described in Figure 1B, thousands of circular splicing events were found in both human and mouse samples, with almost 90% deriving from internal exons of protein-coding genes (Figure 1C). Almost 10% of circRNAs was expressed at similar or higher levels with respect to the linear counterpart (by comparing the number of reads mapping on circular versus linear splice junctions). Interestingly, almost 600 of these circRNAs were conserved between species of the ~2,100 and 1,600 circRNAs identified in human and mouse, respectively (Figure 1D, full list in Table S1). Our criterion for conservation was any circRNA in which the genomic location in human overlapped with the syntenic region in mouse. Considering that low-abundance species might have been missed by this analysis, the actual species overlap may be even higher. Indeed, when filtering human circRNAs for high expression level (more than five reads), the overlap with mouse circRNAs increased from ~25% to 40%.

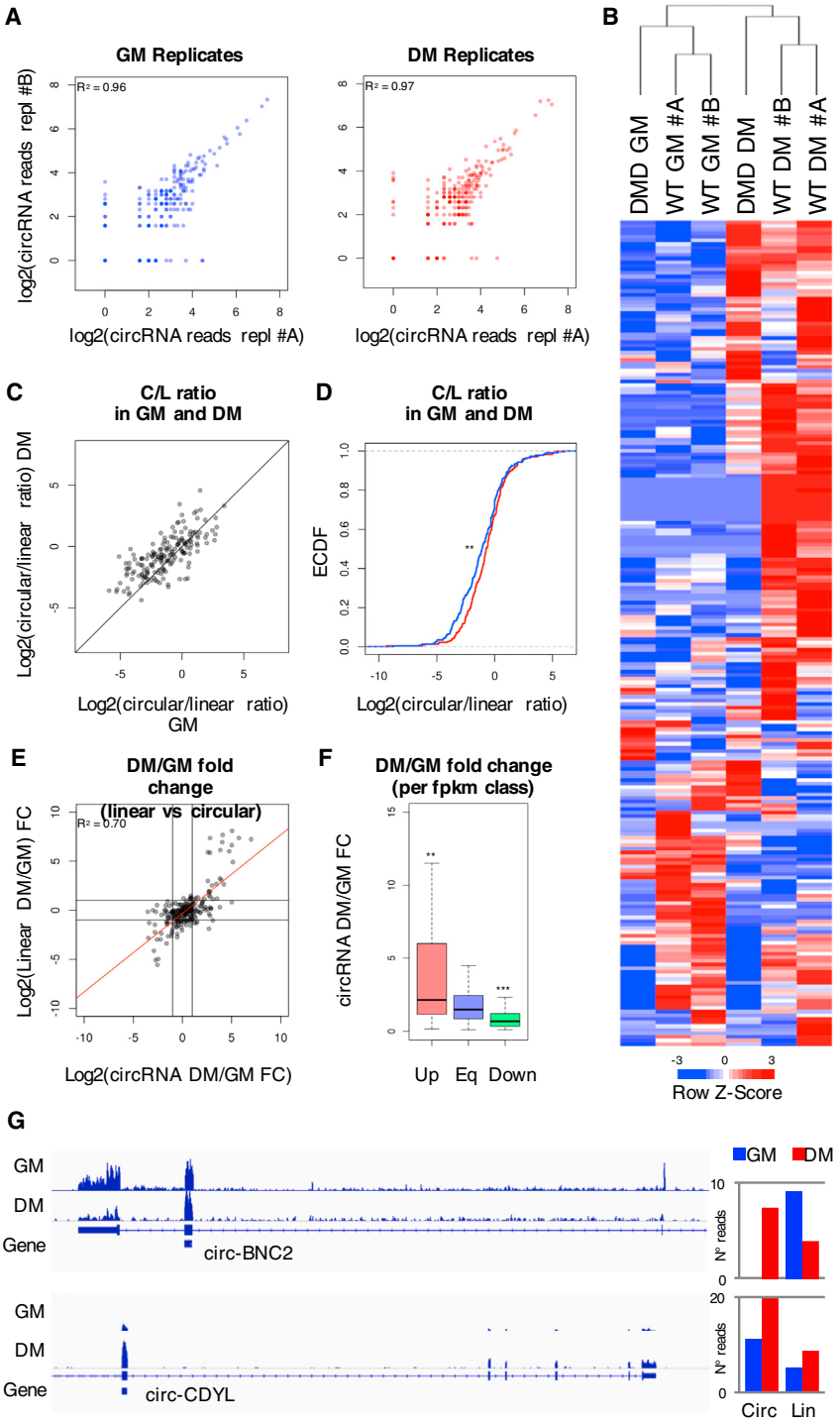


Figure 2. Differential CircRNA Expression in Human Myoblast Differentiation and Disease

(A) Scatterplots showing reads mapping to highly expressed circRNA junctions in two myoblast (GM, A and B, left) and myotube (DM, A and B, right) replicates.

(B) Heatmap showing expression levels of selected circRNAs in WT and DMD myoblasts (GM, A and B) and myotubes (DM, A and B). Values are normalized as row *Z* scores.

(C) Scatterplot showing circular/linear (C/L) ratio in GM versus DM, calculated as log2 of the number of reads mapping to a circular junction divided by the mean of reads mapping linearly at the same genomic coordinates.

(D) Empirical cumulative distribution of the C/L ratio in GM (blue) and DM (red).

(E) Scatterplot of fold change of reads mapping to circRNA junctions versus reads mapping to linear junctions at the same coordinates. Linear regression is shown in red. Black lines correspond to a fold change of 2.

(F) Bar plot representing fold change of circRNA expression in myotubes versus myoblasts (DM/GM) grouped by differential expression of host gene according to Cuffdiff FPKMs. Two and three asterisks indicate a Wilcoxon-Mann-Whitney test-derived p value below 0.02 and 0.01, respectively. Up, upregulated genes; Eq, non-differentially expressed genes; down, downregulated genes (fold change > 2).

(G) Left panels: coverage plots of genomic loci of circ-BNC2 and circ-CDYL. Right panels: bar plots show reads mapping to circular (Circ) and linear (Lin) splice sites for circ-BNC2 and circ-CDYL in GM (blue bars) and DM (red bars).

See also Figure S2 and Table S1.

CircRNA Expression Is Modulated in Muscle Differentiation and Disease

To perform quantitative analyses on circRNA expression and regulation, we took human samples and reduced further our list to a set of highly expressed molecules by requiring at least five unique reads mapping to the head-to-tail junction. With this threshold, while the technical replicates had an overall very good correlation (Figure 2A), the comparison between myoblasts and myotubes revealed a global change in circRNA expression (~45% circRNAs having ≥2-fold variation in one of the two conditions, Figure 2B), meaning that these molecules were modulated in response to cell differentiation. RNA-seq data from human myoblasts derived from DMD patients were also analyzed. Hierarchical clustering analysis of normal and dystrophic myoblasts and myotubes revealed that indeed DMD cells have a unique signature in terms of circRNA expression levels (Figure 2B): specific subsets of transcripts were differently abundant in DMD with respect to controls both in GM and in DM conditions. This is in agreement with the notion that Duchenne cultures have a much slower progression into the differentiation process (Cazzella et al., 2012). Notably, circRNA abundance tends in general to increase during differentiation (Figure 1B), as described for neuronal differentiation (Rybak-Wolf et al., 2015). The circular/linear ratio also followed this trend

(Figures 2C and 2D). This is possibly related to the high stability of these molecules: during differentiation, induction of circRNAs at the transcriptional level, combined with a slow turnover, could lead to their accumulation over time.

Then we addressed the question of whether modulation of circRNA expression in myogenesis was due to transcriptional regulation of the host gene or to post-transcriptional control, such as competitive biogenesis between the linear and the circular isoforms within the same gene (Ashwal-Fluss et al., 2014). Therefore, we analyzed the relationship between the fold change ratio of circular versus linear expression level in the two conditions tested (GM versus DM), and we found a positive correlation (Figure 2E). Additionally, according to FPKMs calculated with Cuffdiff, the abundance of circRNAs produced from upregulated, stable, or downregulated genes indicated that the induction of circRNAs coming from upregulated genes was significantly higher than that of stable and downregulated genes (Figure 2F). However, exceptions were detected, such as the BNC2 gene that produces mainly the linear mRNA in human and mouse myoblasts in growth conditions but predominantly the circRNA in differentiated cells. Figure 2G shows circ-BNC2 in comparison with circ-CDYL, which is an example of a concomitant increase in both circular and linear isoforms during myogenesis, probably because of transcriptional activation of the locus.

RNAi-Based CircRNA Functional Screening

Specific criteria were applied in order to restrict the number of candidates for further characterization: (1) conservation, (2) expression level, (3) modulation during differentiation, and (4) circular/linear ratio (see the STAR Methods). This filtering yielded 31 circRNAs, which are shown in Table 1. The selected species were initially validated by RT-PCR (summarized in Table 1, full results in Figures S2A and S2B, list of primers in Table S2). Of 31 candidates, only one, circ-MYL4, was not detected as a band of the expected size in human samples. Circ-TTTY16 was instead not detected in mouse samples, because it is located in the Y chromosome, which is not present in C_2C_{12} cells. For almost every circRNA, RT-PCR produced a band of the expected size and one or more larger products, possibly corresponding to concatemers generated by rolling circle retro-transcription. For a few circRNAs (CDYL, QKI, and ZNF609), we gel-extracted and sequenced those bands, confirming that they contained the head-to-tail junction and that they were concatemers (Figure S2C). The qRT-PCR analysis of a subset of six human circRNAs did not reveal a major second PCR product for any of them (Figure S2D), possibly because of the different enzyme and cycling conditions used with respect to non-quantitative PCR and opening to the possibility of using such technique for accurate measurement of relative circRNA expression (Figure S2E). Indeed, in all cases, the normalized RNA quantities measured by qRT-PCR reflected the levels observed by RNA-seq and RT-PCR (Figure S2E). For the same candidates, purification of total RNA with oligo dT revealed that they were preferentially recovered in the non-polyadenylated fraction with respect to their linear counterparts (Figure S2F). Similarly, comparison of dT- versus random examer-primed cDNA synthesis shows that circRNAs were retrotranscribed more efficiently with random examers than with dT with respect to linear RNAs (Figure S2G).

Resistance to the RNase R exonuclease was used to test whether the selected molecules were circular (Suzuki et al., 2006). Of 30 putative circRNAs, 29 were confirmed as non-accessible to the exonuclease. Only one, circ-NEB, was digested by RNase R as efficiently as its linear counterpart (summarized in Table 1, full results in Figures S2H and S2I). Moreover, nuclear/cytoplasmic fractionation indicated that all circular species were almost exclusively located in the cytoplasm, both in mouse and human cells (Figures S2J and S2K).

We then designed an image-based functional genetic screen to further characterize these molecules (summarized in Figure 3A). We tried to design at least two small interfering RNAs (siRNAs) targeting each of the 29 circRNAs, according to the following features: (1) binding site at the head-to-tail junction; (2) 7-mer seed not present in the linear host gene; and (3) chemical modifications within the seed region, in order to reduce the chance of an miRNA-like effect and passenger strand inactivation to direct strand-specific Ago loading. These modifications and their applications have been the objective of extensive studies (Jackson et al., 2006; Chen et al., 2008), and they have been here designed for customized targets (On-Target-plus by Dharmacon). Following these rules, it was possible to design two siRNAs for 20 circRNAs, one for five circRNAs, and none for the four remaining ones (Figure S3A; siRNAs listed in Table S3). Due to the poor transfectability of human primary myoblasts, we developed a reverse transfection protocol that allowed high-efficiency transfection of siRNAs in these cells (see the STAR Methods). The 45 siRNAs plus a negative control were applied in triplicate to myoblast cultures in a 96-well format, and, 24 hr post-transfection, cells were switched to DM for 48 and 96 hr. The relative abundance of the circular versus linear forms was tested in order to confirm the specific downregulation of the circRNA. This revealed that at least one siRNA for 17 different genes specifically downregulated the circular form without affecting the linear mRNA (Figure S3B).

Immunofluorescence analysis was performed for two common myogenic markers, Myogenin (MYOG) and Myosin Heavy Chain (MHC), together with DAPI staining (Figure S3C). Several readouts were analyzed in order to obtain a detailed picture of how knockdown of each circRNA affected cell differentiation and morphology. We included in this analysis 11 different phenotypes, including total cell number, the fraction of MYOG- or MHC-positive cells, and additional morphological parameters, such as the fusion index. From this screening, two strong and opposite phenotypes associated with siRNAs targeting circ-QKI and BNC2 emerged. The two siRNAs designed for circ-QKI (siRNA #1 and siRNA #2) both inhibited myoblast differentiation (Figure 3B). While siRNA #1 had a specific effect on circ-QKI and none on QKI mRNA, siRNA #2 did not alter circ-QKI levels but downregulated QKI mRNA (Figure S3B). This supports the hypothesis that both circ-QKI and QKI mRNA participate in the same process. To confirm that hypothesis, we raised a third siRNA specifically targeting QKI mRNA (si-QKI-L; Figure S3D). si-QKI-L, and additional experimental replicates with si-QKI #1 (Figures 3C, S3B, S3D, and S3E), had again a specific anti-myogenic effect, thus confirming that both the circRNA and the mRNA could be involved in promoting differentiation (Figure S3E, as in Li et al., 2003).

Table 1. The Selected 31 CircRNAs and Summary of Their Experimental Validation

Human CircRNAs in Human Myoblasts Cells

CircRNA	Chromosome	Start	End	Reads (GM A, GM B, DM A, DM B)	RT-PCR	RNase R	Cyt/Nuc
Circ-ACVR2A	chr2	148653869	148657467	0, 3, 10, 7	yes	yes	cyt
Circ-ADAMTS6	chr5	64747301	64769779	37, 34, 0, 5	yes	yes	cyt
Circ-ANKIB1	chr7	91924202	91957214	0, 3, 6, 6	yes	yes	cyt
Circ-ASH1L	chr1	155408117	155408859	6, 5, 2, 0	yes	yes	cyt
Circ-ASPH	chr8	62531536	62566219	3, 0, 2, 0	yes	yes	cyt
Circ-BACH1	chr21	30693541	30702014	11, 12, 9, 4	yes	yes	cyt
Circ-BNC2	Chr9	16435552	16437522	0, 0, 8, 7	yes	yes	cyt
Circ-CAMSAP1	Chr9	138773478	138774924	15, 8, 16, 17	yes	yes	cyt
Circ-CDYL	Chr6	4891946	4892613	10, 12, 18, 21	yes	yes	cyt
Circ-CRKL	Chr22	21288066	21288532	0, 2, 10, 8	yes	yes	cyt
Circ-DCBLD2	chr3	98568304	98600611	27, 31, 3, 4	yes	yes	cyt
Circ-HIPK3	Chr11	33307958	33309057	88, 82, 90, 93	yes	yes	cyt
Circ-MED13L	Chr12	116668337	116675510	10, 7, 5, 16	yes	yes	cyt
Circ-MEF2A	Chr15	100211527	100215663	0, 0, 2, 0	yes	yes	cyt
Circ-MYL4	Chr17	45297314	45297419	0, 0, 4, 8	no	N/A	N/A
Circ-NEB	chr2	152544139	152550950	0, 0, 18, 34	yes	no	N/A
Circ-PMS1	chr2	190656515	190682906	3, 0, 7, 3	yes	yes	cyt
Circ-PTP4A2	chr1	32381495	32385259	6, 12, 9, 15	yes	yes	cyt
Circ-QKI	chr6	163876310	163956157	0, 0, 7, 8	yes	yes	cyt
Circ-RTN4	chr2	55252221	55255356	0, 0, 4, 6	yes	yes	cyt
Circ-RTN4bis	chr2	55209650	55214834	10, 13, 22, 20	yes	yes	cyt
Circ-RUNX1	Chr21	36206706	36231875	0, 2, 9, 7	yes	yes	cyt
Circ-SEPT9	chr17	75398140	75398785	3, 0, 7, 3	yes	yes	cyt
Circ-SLC8A1	chr2	40655612	40657444	71, 53, 12, 13	yes	yes	cyt
Circ-SPECC1	chr17	20107645	20109225	7, 5, 12, 6	yes	yes	cyt
Circ-TMEFF1	chr9	103261046	103279053	3, 0, 5, 7	yes	yes	cyt
Circ-TTN	chr2	179557223	179558735	0, 0, 2, 3	yes	yes	cyt
Circ-ZFHX3	chr16	72984367	72994093	3, 2, 2, 0	yes	yes	cyt
Circ-ZNF292	chr6	87920169	87928449	8, 11, 11, 6	yes	yes	cyt
Circ-ZNF609	chr15	64791491	64792365	13, 4, 12, 8	yes	yes	cyt
Circ-TTTY16	chrY	7565120	7567528	2, 0, 0, 2	yes	yes	cyt

Mouse CircRNAs in C_2C_{12} Cells

CircRNA	Chromosome	Start	End	Reads (GM A, GM B, DM A, DM B)	RT-PCR	RNase R	Cyt/Nuc
Circ-Acvr2a	chr2	48870288	48873473	0, 0, 4, 6	yes	yes	cyt
Circ-Adamts6	chr13	104294446	104314383	0, 2, 2, 5	yes	yes	cyt
Circ-Ankib1	chr5	3747020	3772787	0, 4, 14, 13	yes	yes	cyt
Circ-Ash1l	chr3	89001772	89007877	0, 0, 5, 3	yes	yes	cyt
Circ-Asph	chr4	9583811	9610938	0, 0, 5, 4	yes	yes	cyt
Circ-Bach1	chr16	87715330	87722608	2, 0, 0, 2	yes	yes	cyt
Circ-Bnc2	chr4	84291692	84293662	2, 2, 12, 19	yes	yes	cyt
Circ-Camsap1	chr2	25965662	25966948	0, 0, 8, 7	yes	yes	cyt
Circ-Cdyl	chr13	35815908	35816575	7, 5, 19, 25	yes	yes	cyt
Circ-Crkl	chr16	17468892	17469358	0, 0, 1, 0	yes	yes	cyt
Circ_Dcbld2	chr16	58424560	58433463	5, 10, 2, 6	yes	yes	cyt
Circ-Hipk3	chr2	104470748	104471847	5, 7, 43, 42	yes	yes	cyt

(*Continued on next page*)

Table 1. *Continued*

Human CircRNAs in Human Myoblasts Cells

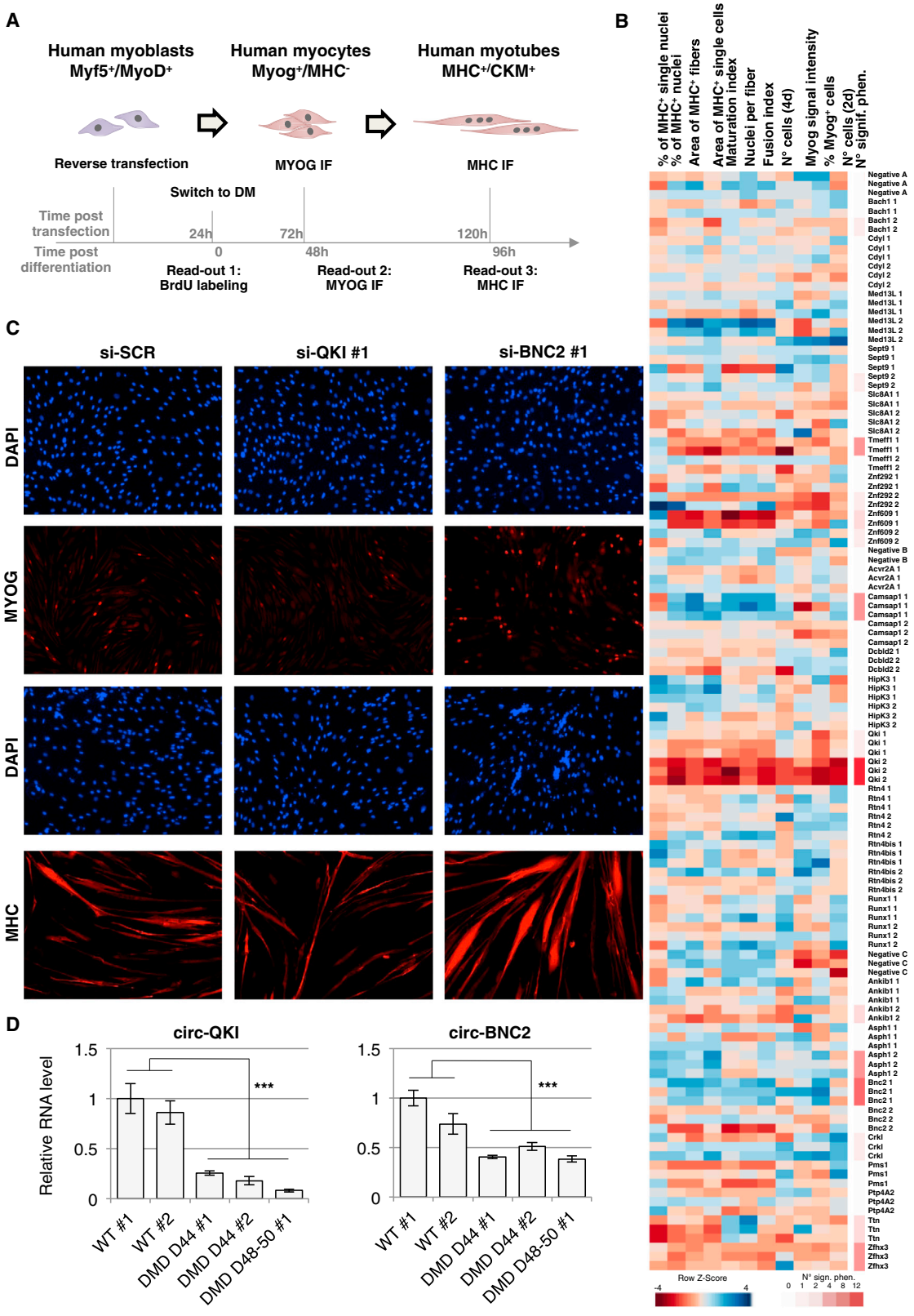
CircRNA	Chromosome	Start	End	Reads (GM A, GM B, DM A, DM B)	RT-PCR	RNase R	Cyt/Nuc
Circ-Med13l	chr5	118593332	118593570	6, 4, 17, 27	yes	yes	cyt
Circ-Mef2a	chr7	67264864	67268400	0, 0, 8, 3	yes	yes	cyt
Circ-Myl4	chr11	104584022	104584127	0, 0, 6, 3	yes	yes	cyt
Circ-Neb	chr2	52299476	52305391	0, 0, 39, 37	yes	no	N/A
Circ-Pms1	chr1	53256628	53282092	0, 4, 6, 8	yes	yes	cyt
Circ-Ptp4a2	chr4	129838977	129846539	2, 0, 0, 0	yes	yes	cyt
Circ-Qk	chr17	10238892	10282981	0, 0, 2, 0	yes	yes	cyt
Circ-Rtn4	chr11	29705534	29708770	2, 3, 6, 13	yes	yes	cyt
Circ-RTN4bis	chr11	29733632	29736498	0, 0, 0, 2	yes	yes	cyt
Circ-Runx1	chr16	92644153	92668812	0, 0, 5, 6	yes	yes	cyt
Circ-Sept9	chr11	117290397	117291036	2, 2, 9, 9	yes	yes	cyt
Circ-Slc8a1	chr17	81647808	81649638	0, 0, 12, 16	yes	yes	cyt
Circ-Specc1	chr11	62117699	62119282	0, 2, 0, 0	yes	yes	cyt
Circ-Tmeff1	chr4	48604580	48617331	0, 2, 6, 5	yes	yes	cyt
Circ-Ttn	chr2	76766305	76791371	0, 0, 3, 7	yes	yes	cyt
Circ-Zfhx3	chr8	108792201	108800691	2,3,4,3	yes	yes	cyt
Circ-Zfp292	chr4	34839331	34839986	0,0,12,14	yes	yes	cyt
Circ-Zfp609	chr9	65794621	65795495	0,0,10,7	yes	yes	cyt
Circ-Ttty16	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Columns contain: gene name, genomic coordinates (hg19 for human and mm10 for mouse) of the head-to-tail splice junction, number of reads mapped to the splice junction in human or mouse myoblasts (GM A and GM B) and myotubes (DM A and DM B), RT-PCR validation (yes if bands of the expected size were obtained and no if not), RNase R resistance (yes if the fraction of recovery after RNase R treatment of the circular form was significantly higher than for the linear form and no if not), and subcellular localization. "N/A" applies to the murine circ-Ttty16, where it was not detected because the host gene is located in the Y chromosome, which is missing in C_2C_{12} cells. cyt, cytoplasm; nuc, nucleus.

In the case of circ-BNC2, si-BNC2 #1 efficiently downregulated circ-BNC2 but had also a small effect on BNC2 mRNA, while si-BNC2 #2 targeted specifically circ-BNC2 and had no effect on the phenotype (Figure S3B). These data suggested that the observed phenotype was likely associated with BNC2 mRNA knockdown. We confirmed this by using a third siRNA, designed specifically for BNC2 mRNA (si-BNC2-L; Figures S3D and S3E), and we observed the same effect as with additional replicates of si-BNC2 #1 (Figures 3C, S3D, and S3E). Notably, during differentiation, a strong increase of the circular RNA paralleled the decrease of the linear BNC2 mRNA (Figures 2G, S2A, and S2B). Therefore, it is likely that biogenesis of circ-BNC2 is a mechanism for competing with the production of the anti-myogenic BNC2 mRNA when differentiation starts.

Interestingly, both circRNAs, upregulated during in vitro differentiation of control myoblasts, were downregulated in DMD conditions (Figure 3D), well correlating with the notion that dystrophic cells have altered progression into the differentiation process (Cazzella et al., 2012). For a different set of circRNAs highly expressed in myoblasts, we investigated the relationship between their abundance and proliferation capacity through a bromodeoxyuridine (BrdU)-labeling assay (Figure S3C). Among the circRNAs tested, we found that the knockdown of circ-ZNF609 (successfully accomplished with siRNA #1) reduced BrdU incorporation by ~80% (Figure 4A). In line with this, also markers of proliferation, such as CDK1 and cyclin A2, were significantly reduced after circRNA knockdown (Figure 4E), indicating a specific role of circ-ZNF609 in myoblast proliferation. Similar effects, even if at lower extents, were observed in murine cells (Figures S4A and S4B). The initial screening was subsequently validated by comparing the activity of three distinct siRNAs: si-circ (corresponding to siRNA #1 used in the screening), which targets the circ-ZNF609 head-to-tail junction; si-ex2, targeting both the circRNA and its linear counterpart; and si-mRNA targeting only the linear mRNA (Figure 4B). Both si-circ and si-ex2 reduced myoblast growth rate by 80% while si-mRNA, targeting only the linear mRNA, did not affect proliferation (Figures 4C and 4D). These results allowed us to exclude that the observed phenotype was due to an off-target effect, and they indicated that the decrease in proliferation was solely due to the circRNA activity. Both circular and linear ZNF609 RNA levels were measured after treatments by qRT-PCR, confirming the specificity of the three siRNAs (Figure 4E). Moreover, northern blot analysis confirmed that circ-ZNF609 was susceptible to si-circ treatment and resistant to RNase R (Figure 4F), while the linear ZNF609 mRNA was resistant to si-circ and susceptible to RNase R (Figure 4F). Notably, the circular form co-migrated with the product derived from an artificial construct, p-circ (see next paragraph), overexpressing circ-ZNF609 RNA (Figure 4F).

Also interestingly, circ-ZNF609, which is downregulated during myogenesis in control myoblasts, was instead found at



(legend on next page)

elevated levels in DMD conditions (Figure 4G). This is consistent with the delayed differentiation phenotype of dystrophic primary myoblasts and with the persistence of a high proportion of proliferating, non-fusing myoblasts (Cazzella et al., 2012). These data reinforce further the connection of circ-ZNF609 and myoblast proliferation.

Circ-ZNF609 Encodes a Protein

Circ-ZNF609 originates from the circularization of the second exon of its host gene. A 753-nt open reading frame (ORF) is present in circ-ZNF609, spanning from the putative AUG of the host gene to a STOP codon created 3 nt after the splice junction (Figure 5A). To determine whether this ORF (circORF) is functional, we first used sucrose gradient fractionation to test circ-ZNF609 loading onto polysomes. Figure 5B shows that a small but significant proportion of circ-ZNF609 sedimented with heavy polysome fractions while the rest remained in the non-bound fraction. The circular conformation of the polysome-bound circ-ZNF609 RNA was confirmed by RNase R resistance (Figures S5A and S5B). Moreover, we excluded the presence of possible linear concatemers by northern blot (Figure 4F; see also RT-PCR with primers in exons 1 and 3 of the ZNF609 gene in Figure S5C). In puromycin-treated myoblasts, circ-ZNF609 shifted to lighter polysomes, similarly to the behavior of the linear ZNF609 and HPRT mRNAs, suggesting that its sedimentation pattern was due to active translation. As a control, a circRNA with no ORF, circ-PMS1, remained on free RNA fractions. The association of circ-ZNF609 with heavy polysomes was also confirmed in mouse (Figure S5D).

To test the protein-coding ability of circ-ZNF609, we took advantage of an expression vector that was able to produce circular transcripts (Kramer et al., 2015). P-circ contains the second exon of the ZNF609 gene and is able to express circ-ZNF609 RNA at high levels (Figure 4F). Sequencing of the RNA produced from this construct revealed its perfect correspondence with the endogenous circ-ZNF609 RNA (data not shown). A construct was derived from p-circ that contained a 3×FLAG-coding sequence immediately upstream of the STOP codon, such that a flagged protein could be produced only upon formation of a circular template (p-circ3xF; Figure 6A). Due to the poor transfection efficiency of plasmid DNA in human and murine myoblasts, the following experiments were carried out in HeLa and N2A cells. The results indicate that in both cell types, in comparison to mock-treated cells, p-circ3xF gave rise to two flagged proteins (Figure 6B). Inspection of the sequence indicated that a second ATG codon was present 150 nt downstream of the first one and in the same frame. Mutants lacking either one of the two initiation codons (p-circΔ1 and p-circΔ2) showed that two bands observed by western blot corresponded to two isoforms: the upper band (ATG1) derived from translation from the first ATG, while the lower one (ATG2) originated from the second start codon (Figure 6B). Deletion of both ATGs (p-circΔ1-2) abolished both products. An additional mutant, with a TGA Stop codon replacing ATG2 (Figure 6A; STOP2), confirmed that translation of the flagged proteins required initiation at those precise sites (Figure S6C).

An additional construct, p-lin3xF, was raised carrying the same ORF present in circ-ZNF609 in a linear conformation. When transfected in HeLa and N2A cells, this plasmid induced the expression of the same two proteins observed with the circular substrate, even if with a much higher efficiency (Figures S6A and S6B). Moreover, the flagged protein resulted in nuclear localization (Figure S6C). Titration of the proteins produced by the linear versus the circRNAs, normalized for the amount of RNA produced, indicated that circRNA was translated at almost two orders of magnitude lower efficiency than the linear form (Figure S6B). These results are in line with previous observations showing that re-initiation and internal initiation of translation can be as low as 1%–10% the efficiency of cap-dependent initiation (Merrick, 2004). Moreover, while the circular template yielded approximately the same amount of the two isoforms, translation of the linear one was strongly biased in favor of the first ATG, indicating a substantial difference in the translation initiation mechanism between the circular and linear transcripts.

In cells expressing p-circ3xF and the different ATG mutant derivatives, we confirmed the production of circular transcripts by northern analysis (Figure S6D), with a specific control of RNase R resistance (Figure S6E). We also verified the integrity of the ORF by Sanger sequencing of the RT-PCR products (data not shown), and we further excluded the production of linear concatemers through RT-PCR (Figures S6F and S6G).

To further prove the protein-coding ability of the circ-ZNF609 RNA derived from the chromosomal gene, we designed a genome-editing approach for inserting the 3×FLAG-coding sequence in the endogenous ZNF609 locus. We used the CRISPR/Cas9 system in mouse embryonic stem cells in combination with a DNA donor designed for inserting the 3×FLAG-coding sequence in the same position as in p-circ3xF, so that a flagged protein could be produced only upon circularization of the second exon of ZNF609 (Figure 6D). After transfection, clonal selection, and genotypization (Figures S6H–S6J), we obtained one positive clone carrying two modified alleles: the first

Figure 3. CircRNA Knockdown followed by Phenotypic Screening in Human Primary Myoblasts

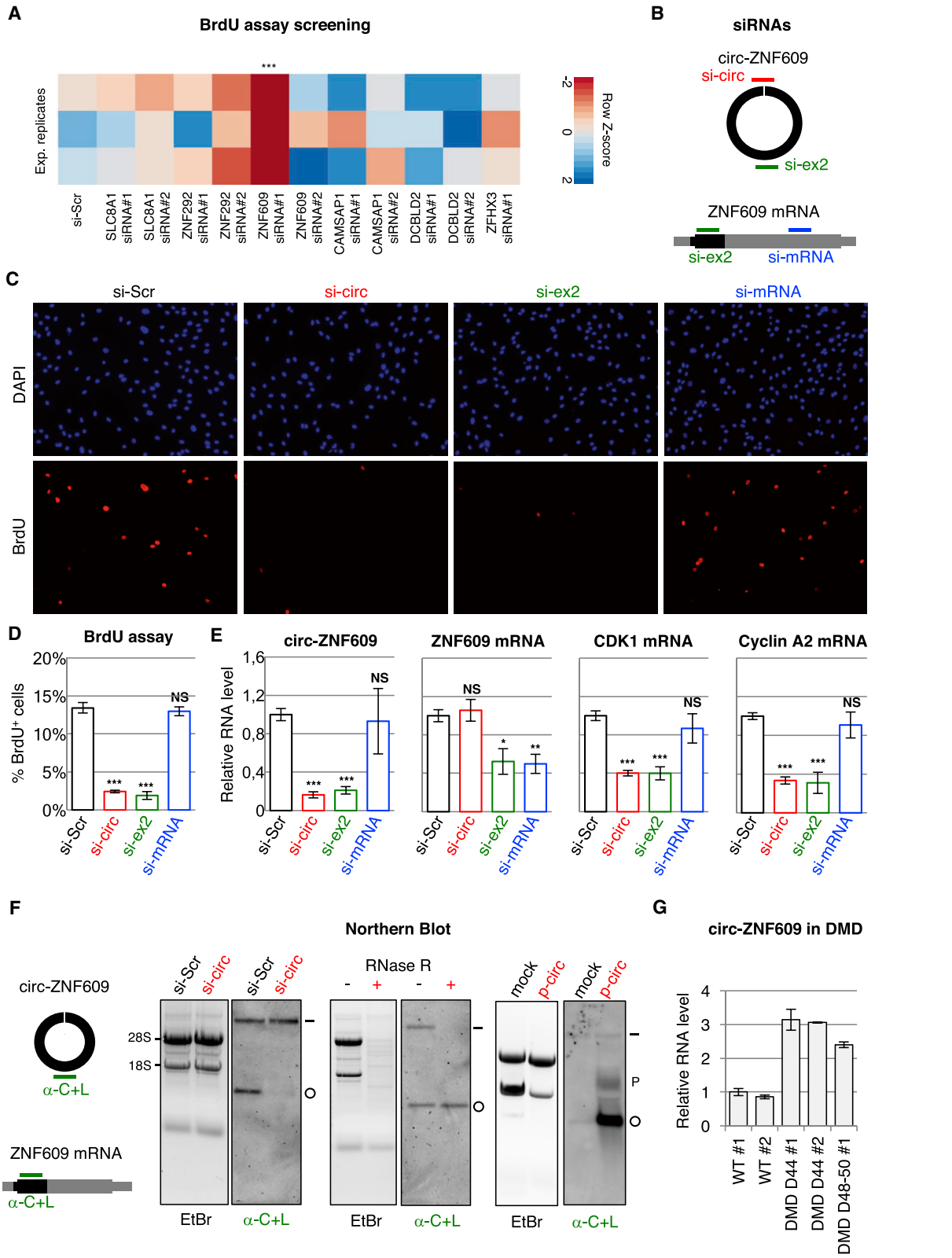
(A) Experimental workflow of the phenotypic screening strategy.

(B) Heatmap showing selected phenotypes in color scale after *Z* score normalization. High values are shown in blue while low values are shown in red. Each column indicates one phenotype and each row indicates one sample. The number of detected significant phenotypes is shown in red scale on the right.

(C) Human primary myoblasts were treated with either a control scramble siRNA (si-SCR) or with siRNAs against circ-QKI (si-QKI #1) and circ-BNC2 (si-BNC2 #1) and induced to differentiate for 48 (upper panels) and 96 hr (lower panels). A representative immunofluorescence for MYOG and MHC together with DAPI staining is shown.

(D) RNA levels measured by qRT-PCR of circ-QKI and circ-BNC2 in human control (WT) and Duchenne (DMD) myocytes. RNA measurements were performed from biological triplicates of two WT (light gray bars) and three DMD primary cell lines (dark gray bars).

See also Figure S3 and Table S3.



(legend on next page)

one contained the expected flagged allele, while the second carried a small deletion across the 3′ splice site of the first intron, which prevented circularization of the second exon (Figure S6J; F/D clone). Therefore, from this cell line only the flagged allele would be able to produce the circ-ZNF609 RNA, and the amount of circRNA produced would be half of that produced by a wild-type (WT) line. Since circ-ZNF609 is highly expressed in neural tissues (Rybak-Wolf et al., 2015), we applied an established neuronal differentiation protocol (Wichterle and Peljto, 2008), and we checked for the expression of the FLAG-tagged circ-ZNF609 RNA by northern blot in WT cells and in the F/D clone. As expected for their genomic context, F/D cells produced roughly half the circRNA of WT cells (Figure S6K); moreover, the circRNA was slightly larger than in the WT because of the inserted tag. RT-PCR and Sanger sequencing confirmed that the circ-ZNF609 RNA produced by F/D cells was identical to that made by the p-circ3xF construct. Due to the low translation efficiency and the high background in western analysis, we performed immunoprecipitation of F/D cells lysates with an anti-FLAG antibody, and then we subjected proteins (size selected between 30 and 40 kDa) to mass spectrometry on a polyacrylamide gel. As positive controls, we immunoprecipitated proteins from cells transfected with the constructs overexpressing either the linear or the circular flagged ORF. Figure 6E shows that numerous peptides, mapping to the circORF, were present in overexpression conditions with the linear construct (p-lin3xF), fewer with overexpression of the p-circ3xF construct, and one in the F/D sample. This is in agreement with the lower levels of RNA produced by the chromosomal gene, with respect to those obtained with the overexpression of both p-circ3xF (Figure 4F) and p-lin3xF, and to the lower translational efficiency from circular templates (Figures S6A and S6B). No ZNF609 peptides were instead detected in control WT cells.

As low initiation efficiency is commonly associated with cap-independent translation, which instead serves as a regulatory mechanism responding to various external stimuli, such as heat shock stress (Holcik and Sonenberg, 2005), we tested protein production in such condition. As shown in Figure 6F, heat shock induces a strong and significant increase in translation of the flagged circ-ZNF609 transcript.

IRES Activity of Circ-ZNF609 UTR

Since circRNAs are devoid of cap structure, translation should occur through sequences able to act as internal ribosome entry sites (IRESs). Notably, phylogenetic analysis of the 5′ and 3′ UTRs of ZNF609 in vertebrates indicated that the portion of the 5′ UTR present in circ-ZNF609 was significantly more conserved than the other UTR sequences situated in the linear ZNF609 mRNA (Figure S6L). To test whether it possessed IRES-like activity, the circ-ZNF609 UTR was inserted between the Firefly and Renilla luciferases of a bicistronic construct (UTR; Figure 6G). With respect to a vector containing a viral IRES sequence (EMCV), UTR displayed very little luciferase activity (Figure 6H). Remarkably, Renilla luciferase raised to levels even higher than those of EMCV when introns from ZNF609 (UTR-inZNF), HPRT (UTR-inHPRT), and β-globin (UTR-inβglob) genes were inserted into the UTR construct, reconstituting the original splice junction of circ-ZNF609. Furthermore, deletion of the 5′ splice site in UTR-inβglob (UTR-Δss) completely abolished luciferase levels (Figure 6H), indicating that Renilla luciferase (Rluc) activity did not originate from cryptic splice sites and ultimately proving the relevance of splicing in the induction of IRES activity. Finally, the circ-ZNF609 UTR was replaced with an unrelated sequence of similar length (MCS-inβglob). Notably, this construct was not able to trigger Rluc activity (Figures 6G and 6H), thus pointing out that IRES activity also relies on specific features of the UTR sequence. RT-PCR analysis indicated that correct splicing, where applicable, was obtained for all constructs (Figure S6M). Altogether, these data show that the UTR of circ-ZNF609 is able to work as an IRES in a splicing-dependent manner.

Moreover, synthetic FLAG-tagged circ-ZNF609 made by in vitro ligation (Figure S6N) was not translated upon transfection in comparison to the same molecule produced through back-splicing from plasmid p-circ3xF, further demonstrating the requirement

Figure 4. Circ-ZNF609 Controls Myoblast Proliferation

(A) Human myoblasts were treated with either a control siRNA (si-Scr) or with siRNAs against the circRNA indicated; cells were maintained in growth conditions. The color scale represents the percentage of BrdU-positive cells measured in triplicate.

(B) Schematic representation of the siRNAs used for circ-ZNF609 phenotype validation. The siRNA targeting specifically the circRNA is shown in red (si-circ), the siRNA targeting both the circRNA and the linear mRNA is shown in green (si-ex2), and the siRNA targeting only the mRNA is shown in blue (si-mRNA).

(C) Human primary myoblasts were treated with either a control scramble siRNA (si-Scr) or the siRNAs shown in (B). A representative immunofluorescence for BrdU is shown in red and DAPI staining is shown in blue.

(D) Percentage of BrdU incorporation levels in cells treated as in (C).

(E) RNA quantification by qRT-PCR of circ-ZNF609, ZNF609 mRNA, and two markers of cell-cycle progression (CDK1 and Cyclin A2) after treatment with the siRNAs as in (C).

(F) Left panel: schematic representation of the DIG-labeled probe used for northern blot analysis together with its location on the circular and linear forms of ZNF609 RNAs. Right panel: northern blots on 20 μg total RNA from primary myoblasts treated with either siRNAs (si-Scr and si-circ) or RNase R (– and +) are shown. The linear and the circular RNA forms are indicated aside the gels with the "–" and "o" symbols, respectively. As a control, northern blot was performed on 20 μg total RNA from myoblasts with the addition of 1 μg RNA from HeLa cells transfected with either an empty vector (mock) or a construct overexpressing circ-ZNF609 (p-circ). Each blot is shown in parallel to the EtBr staining of the agarose gel, where the migration of the 18S and 28S rRNAs is indicated.

(G) RNA levels measured by qRT-PCR of circ-ZNF609 in human wild-type (WT) and Duchenne (DMD) myoblasts upon differentiation. RNA measurements were performed from biological duplicates of two WT and three DMD primary myoblasts. All values derive from three independent experiments unless differently stated. One asterisk indicates a Student's t test-derived p value < 0.05, two asterisks indicate a p value < 0.02, and three asterisks a p value < 0.01. Bars show mean and SE.

See also Figure S4.

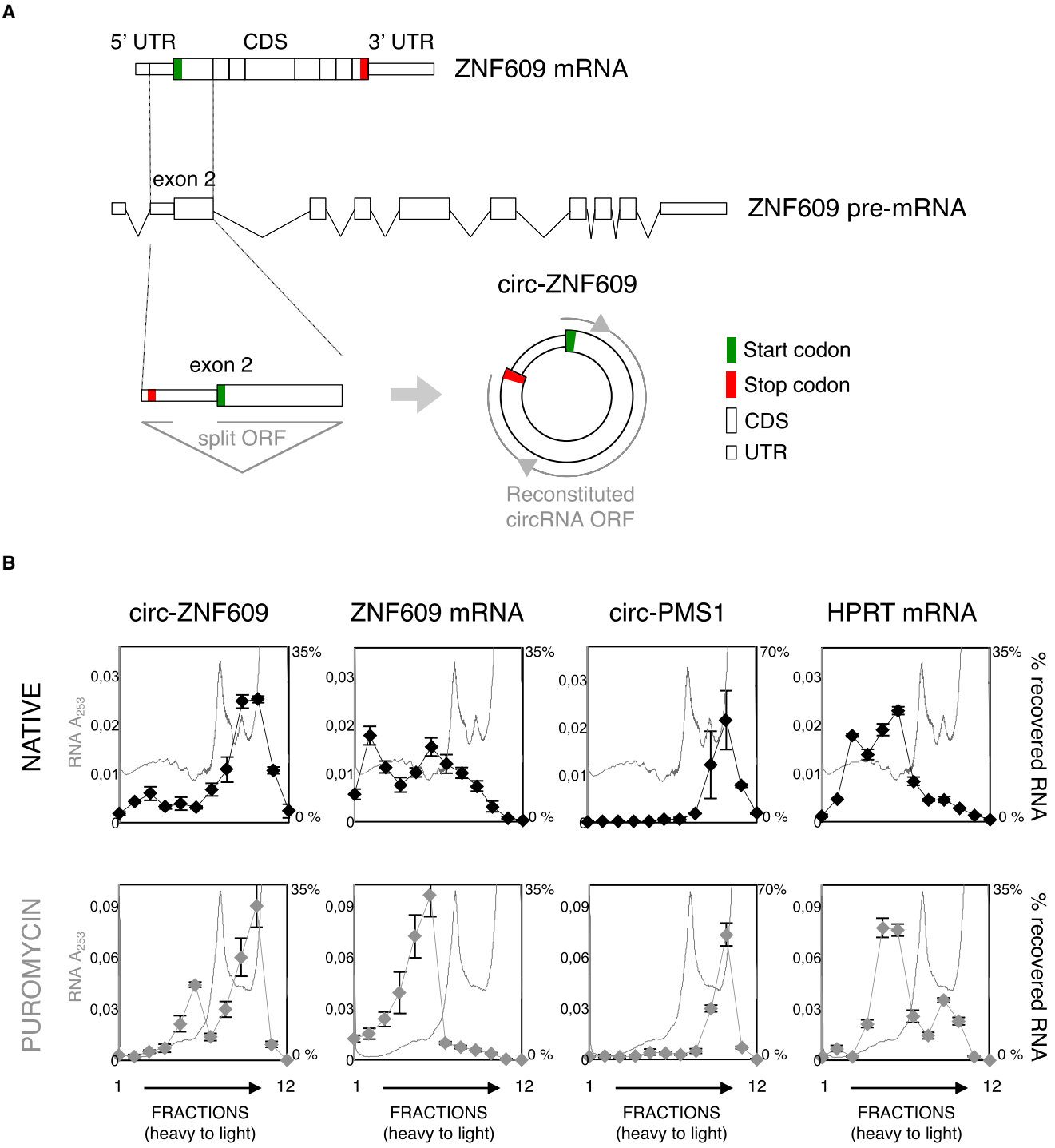


Figure 5. Circ-ZNF609 RNA Is Associated with Polysomes

(A) Schematic representation of the following: ZNF609 pre-mRNA (center), ZNF609 mRNA (above), and ZNF609 s exon (below) from which circ-ZNF609 (right) originates. The 5′ and 3′ UTRs are indicated by smaller boxes, while the coding region (CDS) is represented by larger boxes. Start and stop codons are shown in green and red, respectively.

(B) Cytoplasmic extracts from proliferating human myoblasts, either untreated (native, black) or treated with puromycin (gray), were loaded on 15%–50% sucrose gradients. Absorbance at 253 nm was measured and fractions were collected. Fraction density decreases from left to right; the panels show one representative profile of three independent biological replicates. Individual fractions were analyzed by qRT-PCR and are represented as a percentage of total RNA in each fraction. Profiles are shown for circ-ZNF609 and ZNF609 mRNA. Circ-PMS1 and HPRT mRNA represent negative and positive controls, respectively.

See also Figure S5.

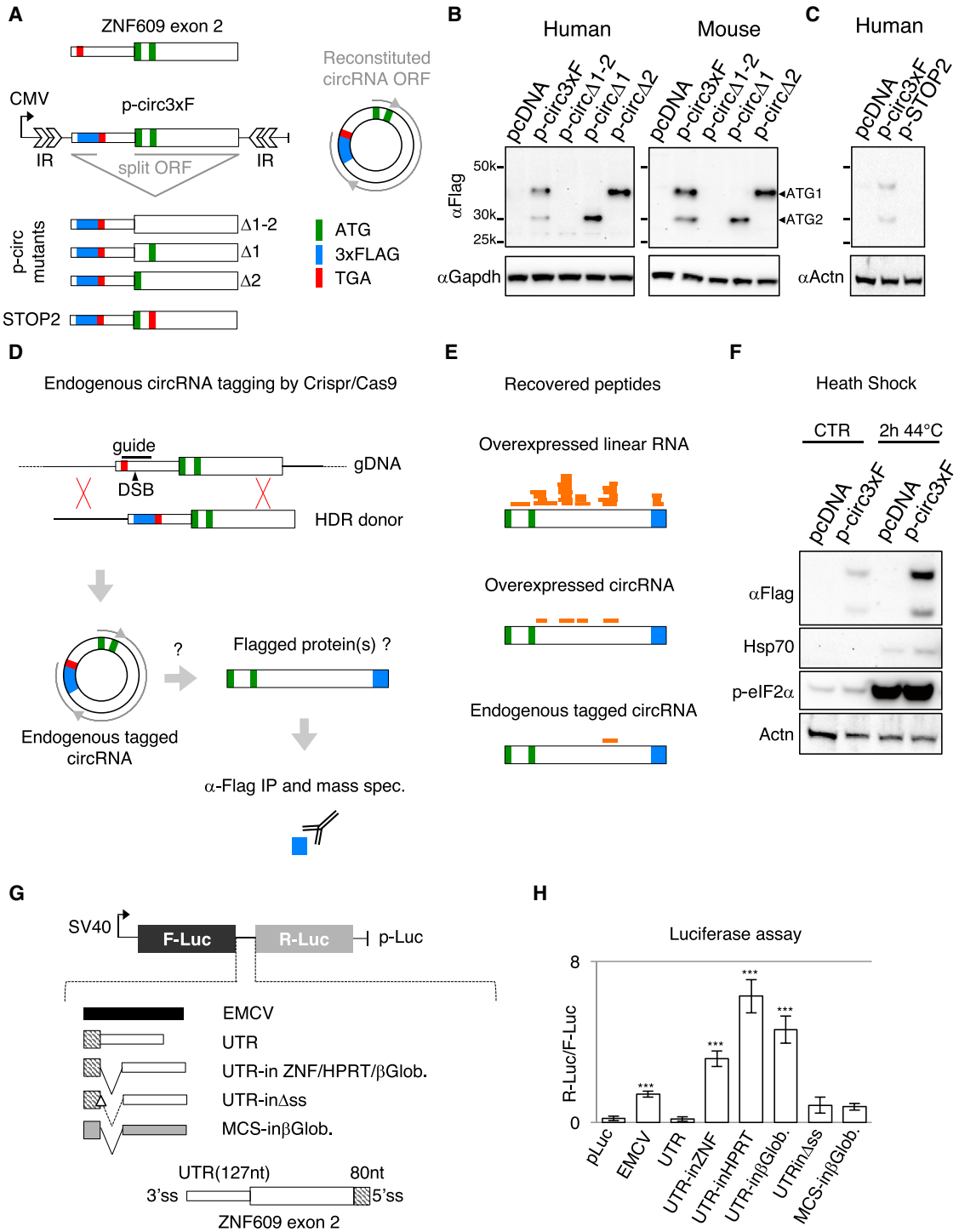


Figure 6. Circ-ZNF609 Has Protein-Coding Activity

(A) Schematic representation of the p-circ3xF construct. The second exon of the ZNF609 gene was cloned in a plasmid carrying inverted repeats (IRs) from the ZKSCAN1 gene (see the STAR Methods) in order to achieve circularization. A 3×FLAG tag was inserted upstream to the stop codon in order to obtain a flagged protein after circularization. Mutant derivatives with deleted ATGs are shown below. Start and stop codons are shown in green and red, respectively. A 3×FLAG-coding sequence is shown in blue.

(B) Western blot analysis with an anti-FLAG antibody on proteins from HeLa (human) and N2a (mouse) cells transfected with p-circ3xF and its mutant derivatives. GAPDH was used as a loading control.

(legend continued on next page)

of a splicing-dependent process for efficient circRNA translation (Figures S6N and S6O).

DISCUSSION

Due to several unique features, circRNAs are increasingly recognized as promising candidates for the identification of additional layers of gene expression control. For instance, circRNAs can pleiotropically modulate gene expression by competing for miRNA or protein binding (Hansen et al., 2013; Hentze and Preiss, 2013; Du et al., 2016), or they can compete with linear RNA production regulating the accumulation of full-length mRNA (Ashwal-Fluss et al., 2014). They have been shown to have differential tissue-specific expression (Memczak et al., 2013; Rybak-Wolf et al., 2015) and specific subcellular localization (Rybak-Wolf et al., 2015; You et al., 2015). Despite the large number of circRNAs identified in many different organisms, tissues, and developmental conditions, only a few examples are available that indicate a crucial role on cellular functions (Guarnerio et al., 2016; Yang et al., 2016). In part this is due to the difficulty in selective targeting of circRNAs, because their nature limits the choice of target region for selectivity and further constrains the parameters normally used for designing mRNA-targeting siRNAs.

Here, with a combination of bioinformatic tools, we obtained selective circRNA downregulation using modified siRNAs. These reagents were used in an image-based functional genetic screen of 25 circRNA species, conserved between mouse and human, which were differentially expressed during myogenesis. Such analysis provided one interesting case, circ-ZNF609, whose phenotype could be specifically attributed to the circular form and not to the linear counterpart. Notably, analysis of the circ-ZNF609 sequence revealed the presence of an ORF. We found that a fraction of circ-ZNF609 RNA is loaded onto polysomes and that, upon puromycin treatment, it shifted to lighter fractions, similar to canonical mRNAs. The coding ability of this circRNA was proved through the use of artificial constructs expressing circular tagged transcripts and by CRISPR/Cas9 tagging of the endogenous ZNF609 gene. Proteins derived from the translation of circRNAs were described a long time ago in Archaea, where homing endonucleases were produced by stable circularized introns (Kjems and Garrett, 1988; Dalgaard et al., 1993). Although there are reports of circRNA in vitro translation (Chen and Sarnow, 1995) and of proteins derived from artificial circRNA constructs (Wang and Wang, 2015; Abe et al., 2015), our data demonstrate that an eukaryotic endogenous circRNA may encode for a protein.

Generation of protein isoforms via back-splicing and circularization of selected exons from primary mRNA transcripts is an intriguing mechanistic alternate to conventional alternative splicing. Interestingly, termination codons present upstream of the AUG initiation codon in linear mRNAs can be converted into in-frame STOP codons upon circularization. Moreover, depending on the distance of the STOP codon from the newly formed junction, peptides with additional amino acids that are absent from the linear form can be produced. Therefore, these mechanisms enlarge the repertoire of protein domains that can be derived from a primary transcriptional unit, and they increase further the complexity of the genome output. This is particularly interesting in consideration of the fact that many circRNAs contain the canonical start codon of their host gene (735 circRNAs of 2,175 in our human dataset, well above the number expected by chance).

So far, we have no hints on the molecular activity of the proteins derived from circ-ZNF609 and as to whether they contribute to modulate or control the activity of the counterpart deriving from the linear mRNA. The zinc-finger protein 609 has been described as a transcriptional factor characterized by two zinc-finger domains. To date only limited information is available about its activity. It was detected as part of a protein interaction network dedicated to control embryonic stem cell (ESC) pluripotency in Nanog-associated complexes (Gagliardi et al., 2013) and as a direct interactor of Dax1 (Wang et al., 2006). ZNF609 was also shown to be required for Rag1 and Rag2 expression in developing thymocytes (Reed et al., 2013). The fact that the circ-ZNF609-encoded protein lacks the zinc-finger

(C) Western blot analysis with an anti-FLAG antibody on proteins from HeLa cells transfected with pcDNA, p-circ3xF, and its mutant derivative p-STOP2. ACTN was used as a loading control.

(D) Schematic representation of the genome-editing approach for tagging the endogenous ZNF609 locus. Mouse embryonic stem cells were edited with homology-directed repair of a CRISPR/Cas9-induced double-strand break (DSB) with a DNA donor carrying a 3×FLAG-coding sequence (in blue) immediately upstream of the stop codon in exon 2. The edited cells were cultured and harvested for IP with an anti-FLAG antibody. The eluted proteins were run on SDS-PAGE (size selected between 30 and 40 kDa) and subjected to mass spectrometry.

(E) Individual peptides detected by mass-spec are depicted in orange at their mapping position onto the protein sequence. Mass-spec data are shown for the overexpression of the linear ORF (see Figure S6A) by p-lin3xF (top), overexpression of the circular RNA from p-circ3xF (middle), and expression from the edited endogenous locus (bottom).

(F) Representative western blot analysis with an anti-FLAG antibody on proteins from HeLa cells transfected with either pcDNA or p-circ3xF in normal (CTR) or heat shock conditions. Circ-ZNF609 RNA levels were the same in both conditions (data not shown). As controls of the heat shock response, blots were hybridized with Hsp70 and phosphorylated eIF2 alpha antibodies. Actinin was used as a loading control.

(G) Schematic representation of the luciferase bicistronic constructs used. The SV40 promoter drives the transcription of Firefly (F-Luc) and Renilla (R-Luc) luciferase-coding regions in the control construct (p-Luc). The spacer of p-Luc was substituted with the IRES sequence of the Encephalomyocarditis virus (EMCV) or 80 nt from the 3′ portion of the second exon fused to the UTR of circ-ZNF609 (UTR). From UTR, additional constructs were raised carrying one intron from the ZNF609, HPRT, and β-globin (UTR-in ZNF/HPRT/βGlob.) genes. From UTR-in-βGlob. two other constructs were derived as follows: UTR-inΔss has a deletion of the 5′ splice site, while MCS-inβGlob. carries an unrelated sequence in place of the UTR.

(H) Luciferase activity derived from cells transfected with the constructs described in (G). Values from three independent experiments are expressed as ratio of Renilla versus Firefly activities. Three asterisks represent a Student's t test-derived p value < 0.01 between the value of each sample and the empty vector. Bars show mean and SE.

See also Figure S6.

domains would suggest different hypotheses on how it could interfere or cope with the activity of the full-length isoform: it could act as a dominant-negative competitor or as a modulator of alternative ZNF609 complex formation. Interestingly, the flagged protein derived from circ-ZNF609 exhibited mainly nuclear localization.

However, since the use of plasmids and other expression vectors as well as naked RNA produced a non-specific block of myoblast proliferation, we were unable to perform rescue experiments and univocally link the circ-ZNF609 phenotype to its protein-coding capacity.

Another relevant aspect about circRNAs, lacking the cap structure and the polyA tail, relates to the machinery and factors required for the assembly into translationally competent complexes. Cap-independent translation, driven by the so-called IRES, was initially described in viral RNAs (Pelletier and Sonenberg, 1988) and subsequently identified also in cellular RNAs (Weingarten-Gabbay et al., 2016). IRESs can act by various mechanisms, including the formation of RNA structures, Watson-Crick base pairing with the 18S ribosomal RNA, and often in cooperation with IRES-transacting factors (ITAFs) (Komar and Hatzoglou, 2011); moreover, putative cellular IRESs have been suggested to play crucial roles in different developmental conditions and in response to different types of stress (Holcik and Sonenberg, 2005).

In this work, we showed that the UTR element of circ-ZNF609, spanning from the termination to the initiation codons, can drive IRES-dependent translation but only if produced through a splicing event, suggesting that factors loaded on the transcript upon splicing might play a crucial role in ribosome recognition and translation initiation.

In this study, we have also observed that the circular template had a much lower translation activity when compared to the linear counterpart, a feature exhibited also when analyzing the cell line carrying an endogenous flagged allele obtained through CRISPR/Cas9 genome editing. Interestingly, heat shock induced significant activation of the flagged circ-ZNF609 translation. It is well known that cap-independent initiation is less efficient than the cap-dependent one (Merrick, 2004) and that regulation of cap-independent translation is often used to respond to several cellular stresses in order to allow immediate and selective changes in protein levels (Holcik and Sonenberg, 2005). Along this line, *cis*-acting elements have been also shown to respond to different cellular stresses and to contribute in the promotion of translation initiation through a cap-independent mechanism. This is the case of N6-methyladenosine (m6A) residues whose levels in 5′ UTRs are increased upon different cellular stresses (Zhou et al., 2015; Meyer et al., 2015). Notably, reanalysis of m6A cross-linking and immunoprecipitation (CLIP) data (Zhao et al., 2014) and m6A immunoprecipitation (IP) in myoblasts (Figures S6P and S6Q) revealed that circ-ZNF609 is highly methylated, suggesting a possible regulatory role of this modification in circ-ZNF609 translational activity.

In conclusion, circ-ZNF609 provides relevant material to identify the factors required for the translation of this class of transcripts devoid of cap and polyadenylated tails. Further work will be required in order to understand whether circRNA translation responds to specific stimuli or signals different from those regulating protein synthesis conducted by linear mRNAs.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture and treatments
- METHOD DETAILS
  - RNA isolation, treatments and analysis
  - Northern blots
  - RNaseq and circular RNA detection
  - Cell fractionation
  - Western blots
  - Immunofluorescence
  - Protein IP and mass spec
  - m6A IP
  - Genome editing with Crispr/Cas9
  - Plasmid construction
  - Preparation of synthetic circRNAs
  - Luciferase assay
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - RNaseq analyses
  - Image analysis
  - Statistical Analyses
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2017.02.017.

AUTHOR CONTRIBUTIONS

Conceptualization, I.L. and I.B.; Methodology, I.L., O.S., A.F., T.S., A.A., and M.W.; Software, I.L. and A.A.; Formal Analysis, I.L.; Investigation, I.L., G.D.T., F.R., M.M., F.B., A.F., T.S., and P.L.; Writing, I.B., I.L., and M.M.; Visualization, I.L.; Supervision, I.B. and N.R.; Project Administration and Funding Acquisition, I.B.

ACKNOWLEDGMENTS

We thank M. Marchioni from I.B.'s lab for technical support; F. Loreni and S. D'Amico for advice on polysome profiles; M.O. Rodriguez, P. Papavasileiou, S. Memczak, C. Stottemeister, and A. Rybak-Wolf from N.R.'s lab for advice, sharing protocols, and useful discussions; F. Ricci from M.W.'s lab for image acquisition; A. Birmingham of GE LifeSciences (Dharmacon) for collaboration on siRNA design; X.-H. Ho from G. Meister's lab for providing plasmids and for discussions; the Proteomic facility at the IGBMC of Strasbourg for mass spectrometry analyses; and the Telethon Network Genetic Biobanks (GTB12001F) and EuroBioBank for providing biological samples. This work was partially supported by grants from the following: ERC-2013 (AdG 340172-MUNCODD); Telethon (GGP16213); Epigen-Epigenomics Flagship Project (59377); Parent Project Italia (N/A); AFM-Telethon (17835); AriSLA full grant 2014 (ARCI), Fondazione Roma (N/A), and PRIN 2013 (20108XYHJS_003) to I.B.; and the Human Frontiers Science Program Award RGP0009/2014 to I.B. and N.R. I.L. was awarded an EMBO short-term fellowship (141-2014) for this work.

Received: August 1, 2016
Revised: November 28, 2016
Accepted: February 16, 2017
Published: March 23, 2017

REFERENCES

Abe, N., Matsumoto, K., Nishihara, M., Nakano, Y., Shibata, A., Maruyama, H., Shuto, S., Matsuda, A., Yoshida, M., Ito, Y., and Abe, H. (2015). Rolling Circle Translation of Circular RNA in Living Human Cells. Sci. Rep. *5*, 16435.

Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014). circRNA biogenesis competes with pre-mRNA splicing. Mol. Cell *56*, 55–66.

Ballard, D.H., and Brown, C.M. (1982). Computer Vision, First Edition (Prentice Hall).

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

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.

Cazzella, V., Martone, J., Pinnarò, C., Santini, T., Twayana, S.S., Sthandier, O., D'Amico, A., Ricotti, V., Bertini, E., Muntoni, F., and Bozzoni, I. (2012). Exon 45 skipping through U1-snRNA antisense molecules recovers the Dys-nNOS pathway and muscle differentiation in human DMD myoblasts. Mol. Ther. *20*, 2134–2142.

Chen, C.Y., and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. Science *268*, 415–417.

Chen, P.Y., Weinmann, L., Gaidatzis, D., Pei, Y., Zavolan, M., Tuschl, T., and Meister, G. (2008). Strand-specific 5′-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. RNA *14*, 263–274.

Chen, Y., Li, C., Tan, C., and Liu, X. (2016). Circular RNAs: a new frontier in the study of human diseases. J. Med. Genet. *53*, 359–365.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, 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.

Dalgaard, J.Z., Garrett, R.A., and Belfort, M. (1993). A site-specific endonuclease encoded by a typical archaeal intron. Proc. Natl. Acad. Sci. USA *90*, 5414–5417.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201–206.

Du, W.W., Yang, W., Liu, E., Yang, Z., Dhaliwal, P., and Yang, B.B. (2016). Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. Nucleic Acids Res. *44*, 2846–2858.

Gagliardi, A., Mullin, N.P., Ying Tan, Z., Colby, D., Kousa, A.I., Halbritter, F., Weiss, J.T., Felker, A., Bezstarosti, K., Favaro, R., et al. (2013). A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. EMBO J. *32*, 2231–2247.

Guarnerio, J., Bezzi, M., Jeong, J.C., Paffenholz, S.V., Berry, K., Naldini, M.M., Lo-Coco, F., Tay, Y., Beck, A.H., and Pandolfi, P.P. (2016). Oncogenic Role of Fusion-circRNAs Derived from Cancer-Associated Chromosomal Translocations. Cell *165*, 289–302.

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.

Hentze, M.W., and Preiss, T. (2013). Circular RNAs: splicing's enigma variations. EMBO J. *32*, 923–925.

Holcik, M., and Sonenberg, N. (2005). Translational control in stress and apoptosis. Nat. Rev. Mol. Cell Biol. *6*, 318–327.

Jackson, A.L., Burchard, J., Leake, D., Reynolds, A., Schelter, J., Guo, J., Johnson, J.M., Lim, L., Karpilow, J., Nichols, K., et al. (2006). Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. RNA *12*, 1197–1205.

Jeck, W.R., and Sharpless, N.E. (2014). Detecting and characterizing circular RNAs. Nat. Biotechnol. *32*, 453–461.

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.

Kjems, J., and Garrett, R.A. (1988). Novel splicing mechanism for the ribosomal RNA intron in the archaebacterium Desulfurococcus mobilis. Cell *54*, 693–703.

Komar, A.A., and Hatzoglou, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle *10*, 229–240.

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.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Legnini, I., Morlando, M., Mangiavacchi, A., Fatica, A., and Bozzoni, I. (2014). A feedforward regulatory loop between HuR and the long noncoding RNA linc-MD1 controls early phases of myogenesis. Mol. Cell *53*, 506–514.

Li, Z., Takakura, N., Oike, Y., Imanaka, T., Araki, K., Suda, T., Kaname, T., Kondo, T., Abe, K., and Yamamura, K. (2003). Defective smooth muscle development in qkI-deficient mice. Dev. Growth Differ. *45*, 449–462.

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.

Merrick, W.C. (2004). Cap-dependent and cap-independent translation in eukaryotic systems. Gene *332*, 1–11.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5′ UTR m(6)A Promotes Cap-Independent Translation. Cell *163*, 999–1010.

Pelletier, J., and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature *334*, 320–325.

Reed, N.P., Henderson, M.A., Oltz, E.M., and Aune, T.M. (2013). Reciprocal regulation of Rag expression in thymocytes by the zinc-finger proteins, Zfp608 and Zfp609. Genes Immun. *14*, 7–12.

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., 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.

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

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.

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.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. *31*, 46–53.

Wang, Y., and Wang, Z. (2015). Efficient backsplicing produces translatable circular mRNAs. RNA *21*, 172–179.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364–368.

Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., Gritsenko, A.A., Stern-Ginossar, N., Yakhini, Z., Weinberger, A., and Segal, E. (2016). Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science *351*, aad4939.

Wichterle, H., and Peljto, M. (2008). Differentiation of mouse embryonic stem cells to spinal motor neurons. Curr. Protoc. Stem Cell Biol. *Chapter 1*, 1H.1.1–1H.1.9.

Yang, P., Qiu, Z., Jiang, Y., Dong, L., Yang, W., Gu, C., Li, G., and Zhu, Y. (2016). Silencing of cZNF292 circular RNA suppresses human glioma tube formation via the Wnt/β-catenin signaling pathway. Oncotarget *7*, 63449–63455.

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.

Zhao, X., Yang, Y., Sun, B.F., Shi, Y., Yang, X., Xiao, W., Hao, Y.J., Ping, X.L., Chen, Y.S., Wang, W.J., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res. *24*, 1403–1419.

Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., Chen, B., Luo, Y., Lyu, D., Li, Y., Shi, G., et al. (2016). Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. Nat. Commun. *7*, 11215.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B.K. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature *526*, 591–594.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-Flag M2	Sigma-Aldrich	Cat# F1804-1MG
anti-Actinin (H-300)	Santa Cruz Biotechnology	Cat# sc-15335; RRID: AB_2223809
anti-GAPDH (6C5)	Santa Cruz Biotechnology	Cat# sc-32233; RRID: AB_627679
anti-myogenin	Santa Cruz Biotechnology	Cat# sc12732; RRID: AB_627980
anti-myosin	This paper	N/A
anti-mouse Cy3/AlexaFluor488	Jackson ImmunoResearch	Cat# 115-165-003
anti-m6A	SYSY	Cat# 202 003; RRID: AB_2279214
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Sigma-Aldrich	Cat# D6546
FBS	Sigma-Aldrich	Cat# F7524
L-Glutamine	Sigma-Aldrich	Cat# G7513
EGF	Corning	Cat# 354052
Penicillin/Streptomycin	Sigma-Aldrich	Cat# P0781
FGFb	Millipore	Cat# 01-106
Human Skeletal Muscle DM	Promocell	Cat# C-23061
Insulin	Sigma-Aldrich	Cat# 11376497001
Collagen Type I	Corning	Cat# 354236
Optimem	GIBCO	Cat# 31985-047
RNase R	Epicenter	Cat# MRNA092
VILO Superscript	Thermo Fisher Scientific	Cat# 11754050
SYBR Mastermix	QIAGEN	Cat# 208052
bis-tris-acrylamide gel	Thermo Fisher Scientific	Cat# WG1402BOX
Embryonic Stem Medium	Merck Millipore	Cat# SLM-220-B
GlutaMAX	Thermo Fisher Scientific	Cat# 35050-061
LIF	Merck Millipore	Cat# ESG1107
Advanced D-MEM/F12	Thermo Fisher Scientific	Cat# 12634010
Neurobasal	Thermo Fisher Scientific	Cat# 211030049
KnockOut Serum Replacement	Thermo Fisher Scientific	Cat# 10828028
B-27	Thermo Fisher Scientific	Cat# 17504-044
NE Aminoacids	Merck Millipore	Cat# TMS-001-C
Retinoic Acid	Sigma-Aldrich	Cat# R2625
Glyoxal loading dye	Ambion	Cat# AM8551
Mytaq DNA polymerase	Bioline	Cat# BIO-21105
Cycloheximide	Sigma-Aldrich	Cat# 01810
RNase guard	Thermo Fischer Scientific	Cat# N8080119
Qiazol reagent	QIAGEN	Cat# 79306
Hybond N+ membrane	GE Healthcare	Cat# RPN303B
T7 RNA polymerase	Promega	Cat# P2077
Proteinase K	Thermo Fischer Scientific	Cat# 25530049
Puromycin	SERVA	Cat# 33835.02
Paraformaldehyde	Electron microscopy science	Cat# 157-8
Goat serum	Thermo Fischer Scientific	Cat# 16210072
ProLong Diamond reagent	Thermo Fischer Scientific	Cat# P36930

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DAPI	Sigma-Aldrich-Aldrich	Cat# D9542
FastDigest EcoRI	Thermo Fischer Scientific	Cat# FD0274
FastDigest NheI	Thermo Fischer Scientific	Cat# FD0973
BbsI	Neb	Cat# R0539S
FastDigest XbaI	Thermo Fischer Scientific	Cat# FD0684
NorthernMax buffer	Ambion	Cat# AM8671
PBS	Sigma-Aldrich	N/A
LDS sample buffer	Invitrogen	Cat# NP0007
DNase	Thermo Fischer Scientific	Cat# EN0525
Alkaline Phosphatase	Sigma-Aldrich	Cat# 000000010713023001
RNase inhibitors	Thermo Fischer Scientific	Cat# EO0384
T4 Polynucleotide kinase	BioLabs	Cat# M0236S
DMSO	Sigma-Aldrich	Cat# D4540
T4 RNA Ligase 1	BioLabs	Cat# M0204S
Lipofectamine 2000	Thermo Fischer Scientific	Cat# 11668019
Dharmafect lipid reagent	Dharmacon	Cat# T-2001-01
Critical Commercial Assays		
Ribominus technology	Thermo Fisher Scientific	Cat# A1083708
Real Genomics Genomic DNA extraction kit	RBC Bioscience	Cat# YGB300
miRNeasy spin columns	QIAGEN	Cat# 217004
DIG-luminescence detection kit	Sigma-Aldrich	Cat# 11363514910
DIG-RNA labeling mix	Sigma-Aldrich	Cat# 000000011277073910
In-Fusion HD Cloning Kit	Clontech	Cat# 639649
Dynabeads Protein G Immunoprecipitation kit	Invitrogen	Cat# 10007D
5-Bromo-2′-deoxy-uridine Labeling and Detection kit I	Sigma-Aldrich	Cat# 11296736001
Dual Glo luciferase assay	Promega	Cat# E2920
Deposited Data		
RNA sequencing raw data	This paper	GEO: GSE70389
Experimental Models: Cell Lines		
Mouse myoblasts C_2C_{12}	ATCC	Cat# CRL-1772
Human male myoblasts (WT)	Telethon Biobank	N/A
Human male myoblasts (DMD)	Telethon Biobank	N/A
HeLa	ATCC	Cat# CCL-2
N2A	ATCC	Cat# CCL-131
Mouse embryonic stem cells	This paper	N/A
Mouse embryonic fibroblasts	Millipore	Cat# 92590
Oligonucleotides		
DNA oligonucleotides used in this work are listed in Table S2		
siRNAs used in this work are listed in Table S3		
Recombinant DNA		
pcDNA3.1(-)	Thermo Fisher Scientific	N/A
p-circ (name used in this work)	Hung Ho Xuan, Gunter Meister Lab	N/A
p-circ-3xF (name used in this work)	Hung Ho Xuan, Gunter Meister Lab	N/A
circRNA mini vector ZKSCAN1	Addgene	#60649
p-circ-3xF Δ1	This paper	N/A
p-circ-3xF Δ2	This paper	N/A
p-circ-3xF Δ1-2	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
p-STOP2	This paper	N/A
p-lin-3XF	This paper	N/A
px330 vector	Addgene	#42230
pGL3-Control vector	Promega	N/A
pRL-TK vector	Promega	N/A
p-Luc empty vector	This paper	N/A
UTR+In.β Glob.D5′ss	This paper	N/A
UTR vector	This paper	N/A
pIRES2-EGFP plasmid	Clontech	N/A
pLuc-EMCV-IRES	This paper	N/A
UTR+In. βGlob vector	This paper	N/A
UTR+In.-ZNF vector	This paper	N/A
UTR+In.HPRT vector	This paper	N/A
MCS+In.β Glob vector	This paper	N/A
Software and Algorithms		
Trimmomatic	Bolger et al. (2014)	N/A
FastQC	N/A	N/A
Tophat2	Trapnell et al. (2009)	N/A
Cufflinks	Trapnell et al. (2010)	N/A
Cuffdiff	Trapnell et al. (2013)	N/A
Bowtie2	Langmead and Salzberg (2012)	N/A
FindCirc	Memczak et al. (2013)	N/A
http://crispr.mit.edu	N/A	N/A
Acapella software development/run-time environment	Perkin Elmer	N/A
Custom image analysis scripts	This paper	N/A
circRNA annotation and data parsing scripts	This paper	N/A

CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Irene Bozzoni (irene.bozzoni@uniroma1.it).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture and treatments

Mouse myoblasts (C_2C_{12}, ATCC) were cultured in Growth Medium (GM, DMEM with 20% FBS, L-Glutamine 2 mM and Penicillin/Streptomycin) and induced to differentiate in DM (FBS reduced to 0.5%). Human myoblasts (WT and DMD) were cultured in Growth Medium (GM, DMEM with 10% FBS, L- Glutamine 2 mM, insulin 50 mg/ml, FGFb 25 ng/ml, EGGF 1 ng/ml, Penicillin/Streptomycin) and induced to differentiate with Human Skeletal Muscle DM (Promocell).

Reverse transfection of human myoblasts with siRNAs was performed as follows: 96-well optical plates (Corning) were treated overnight with 350 mg of Collagen Type I in water and 0.5% acetic acid, washed with PBS and UV-sterilized. Twenty microliters of siRNA transfection mix (150 nM siRNA, 0.2 μL Dharmafect lipid reagent from Dharmacon, 20ml transfection medium: DMEM with L-glutamine 2mM and insulin 50 mg /ml) was seeded in each well, left at RT for 20 min and then diluted with 30 μL of transfection medium. Cells (12000 per well) were added after resuspension in double complete medium (GM with 2X concentration of serum, antibiotics, FGFb and EGF). Cells were switched to DM after 24h and fixed or collected for analysis at 48 hr and 96 hr.

HeLa and N2A cells were cultured in DMEM with 10% FBS, L-Glutamine 2 mM and Penicillin/Streptomycin 1X. HeLa and N2A cells were transfected with 2 μg of plasmid DNA with Lipofectamine 2000. Two μl of lipofectamine were added to 300 μL of Optimem with DNA, mixed and pipetted in a 35-mm plate with 200000 cells. Medium was replaced 4 hr later and cells harvested for protein and RNA analyses after 48 hr. For heat shock experiments, p-circ3xF-tranfected HeLa cells were cultured at 44°C for three hours.

METHOD DETAILS

RNA isolation, treatments and analysis

Total RNA in this study was extracted with Qiazol reagent according to the manufacturer's protocol (QIAGEN). RNA from 96-well plates and nuclear/cytoplasmic fractions was purified with miRNEasy spin columns according to the manufacturer's specifications (QIAGEN).

RNase R treatment was performed as follows: 1-10 μg of total RNA was diluted in 20 μL of water with 4u RNase R/μg unless differently stated and 2 μL of enzyme buffer (Epicenter), then incubated 15 min at 37°C and purified by phenol chloroform extraction. One pg of a DNA spike- in molecule was added to each reaction for qPCR normalization. DNA spike-in was produced from the multiple cloning site in pcDNA3.1(-) (Life technologies).

Retrotranscription of RNA in this study was performed with VILO Superscript (Life Technologies): up to 2 μg of RNA were retrotranscribed in a 10 μL reaction mix with 1 μL of enzyme and 2 μL of buffer, incubated 10 min at 25°C, 60 min at 42°C and 5 min at 85°C. qRT-PCR analyses in this study were performed as follows: cDNA (2-10 ng equivalent) was added to a reaction composed of 7.5 μL 2X SYBR Mastermix (QIAGEN), 1.5 μL of 5 mM primers and water to a final volume of 15 μl. DNA amplification (40 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 30 s, followed by melting curve analysis) was monitored with an ABI 7500 Fast qPCR instrument.

RT-PCR for circRNA detection was performed with 0.2 μL of Mytaq DNA polymerase (Bioline) in 25 μL water with 5 μL reaction buffer, 15 ng of cDNA, 2.5 μL 5mM primers. Reaction was carried out for 28 cycles at 95°C for 25 s, 55°C for 25 s, 72°C for 25 s. Five microliters of PCR were run on 2% agarose gels. All primers used in this study are reported in Table S2. RT-PCR for circular and linear RNA from the same locus was carried out for the same number of cycles.

Northern blots

RNA (10-20 μg for detection of endogenous circRNA and 1-2 μg for detection of overexpressed circRNA) was denatured with one volume of glyoxal loading dye (Ambion) for 30' at 50°C and loaded on 1.2% agarose gel. Electrophoresis was carried out for 2-3 hr at 60 V. RNA was transferred on Hybond N+ membrane (GE Healthcare) by capillarity overnight in 10X SSC. Transferred RNA was cross-linked with UV at 1200 × 100 μJ/cm2 and the membrane was washed in 50 mM Tris pH 8.0 for 20 min at 45°C. Prehybridization and hybridization were performed in NorthernMax buffer (Ambion) at 68°C (30 min and overnight respectively). 500 ng of DIG-labeled probe in 10 mL were used for hybridization. The membrane was then washed with 2X SSC 0.1% SDS twice 30 min, then once 30 min and once 1 hr with 0.2X SSC 0.1% SDS at hybridization temperature. The membrane was the processed for DIG detection (hybridization with anti-DIG antibody, washing and luminescence detection) with the DIG luminescence detection kit (Roche), according to the manufacturer instructions. DIG-labeled probes were produced by in vitro transcription with DIG-RNA labeling mix (Roche) of PCR templates produced with the oligonucleotides ZNF609_C_L_FW, ZNF609_C_L_T7_RV used with both human and mouse cDNAs. Transcription with T7 RNA polymerase (Promega) was carried out overnight and the RNA was then loaded on 6% denaturing polyacrylamide gel, run for 1 hr in TB 1X at 100 V. The gel was stained with EtBr 1:10000 for 10 min, washed in ddH2O for 10 min and the bands were cut and resuspended in 500 μL of Na Acetate 0.3 M, 0.1 mM EDTA, 0.2% SDS pH 5.5. Elution was carried out overnight on a rotating wheel and the eluted RNA was purified by PCA and precipitated with 1 μL glycogen in 1 mL 100% Ethanol.

RNaseq and circular RNA detection

Illumina TruSeq library preparation was performed by Istituto di Genomica Applicata (Udine, Italy) from 1 μg of total RNA, depleted of ribosomal RNA with Ribominus technology (Life Technologies). For each sample, 20-40 million 100bp long paired-end reads were collected for each sample (GEO: GSE70389), then trimmed and analyzed as reported in QUANTIFICATION AND STATISTICAL ANALYSIS, section "*RNAseq analyses*."

Cell fractionation

Nuclear/cytoplasmic fractionation of human and mouse cells was performed according to Legnini et al. (2014) as follows: cells (1 × 106) were washed with PBS and immediately lysed on ice with 500 μL of cytoplasmic lysis buffer (140 mM NaCl, 1.5 mM MgCl2, 10 mM Tris pH 7.5, 0.5% NP40) for 5 min. After scraping the cells and placing them in a 2 mL collection tube, 500 μL of a sucrose cushion (same as lysis buffer with 12% Sucrose) was gently pipetted to the bottom of the tube. Sample was centrifuged 10 min at 12000 x g. Supernatant and pellet were then separately resuspended in 1 mL of Qiazol, while sucrose was discarded. RNA extraction was carried out as described above.

Cytoplasm fractionations on sucrose gradients were performed as follows: 5x106 cells were lysed 5 min on ice with 500 μL of TNM (10 mM Tris pH 7.5, 10mM NaCl, 10 mM MgCl2) lysis buffer supplemented with 10 mg/ml cycloheximide, 1X PIC (Complete, EDTA free, Roche) and 1X RNase guard (Thermo Scientific), then spinned 2 min at 15000 x g for discarding nuclear pellets. Samples were centrifuged on 9 mL 15%–45% sucrose gradient at 38000 rpm with a SW41 rotor (Beckman) for 1 hr 30 min at 4°C. Fractions were collected with a Bio-logic LP (Biorad), 1 pg of a spike-in DNA as internal control was added to each fraction, together with 10 μg of glycogen, 100 μg of Proteinase K (Roche) and 40 μL of SDS. Samples were left 1 hr at 37°C and RNA was purified by PCA extraction and precipitated with isopropanol. DNA spike-in was produced amplifying the multiple cloning site from pcDNA3.1(-) (Life technologies).

For puromycin treatments, medium was replaced 3 hr before lysis with 1 mM puromycin containing- medium.

Western blots

Cells were harvested with 50-100 μL of Protein Extraction Buffer (100 mM Tris pH 7.5, EDTA 1 mM, SDS 2%, PIC 1X (Complete, EDTA free, Roche) and incubated 20 min on ice, centrifuged at 15000 x g for 15 min at 4°C. Proteins (15-30 μg) were loaded on 4%–12% bis-tris-acrylamide gel (Life technologies) and transferred to a nitrocellulose membrane. The membrane was blocked in 5% milk and hybridized with the following antibodies, depending on the experiment.

FLAG protein sequence was detected by WB through F1804-1MG monoclonal anti-Flag M2 (Sigma-Aldrich).

Actinin and GAPDH were detected by WB through polyclonal anti-Actinin (H-300) sc-15335 (Santa Cruz Biotechnology) and monoclonal anti-GAPDH (6C5) sc-32233 (Santa Cruz Biotechnology), respectively.

Immunofluorescence

Cells were washed twice in 100 mL PBS, fixed with 100 μL of 4% paraformaldehyde for 15 min at room temperature, washed twice with 100 μL of PBS, permeabilized and blocked with 100 μL of PBS with 0.2% Triton X-100 and 1% goat serum, incubated with 50 μL of in-house made anti-myogenin/anti-myosin antibody (available upon request), washed twice in PBS, incubated with 50 μL anti-mouse Cy3/AlexaFluor488 secondary antibody diluted 1:300 in PBS-DAPI and finally washed twice in PBS. Acquisition for image-based phenotypic analysis was performed with a widefield Nikon TiE Microscope equipped with a Lumencor SpectraX LED light source and PerfectFocus 3; 4 × and 10 × objectives were used (Nikon Instruments). For image analysis see QUANTIFICATION AND STATISTICAL ANALYSIS, section "*Image Analysis*."

For BrdU staining, we used 5-Bromo-2′-deoxy-uridine Labeling and Detection kit I (Roche). Labeling reagent was added to the cell culture medium (diluted 1:1000) 3 hr before fixation (human primary myoblasts) or 15' before fixation (C_2C_{12}), then cells were fixed with a solution made of 7 parts of ethanol and 3 parts of 50mM glycine at pH 2.0 for 20 min at –20°C. Cells were then washed twice with PBS and incubated for 1 hr at room temperature with 30 μL of antibody anti-BrdU diluted 1:10 in PBS. Cells were again washed twice and incubated for 30 min with anti-mouse Cy3/AlexaFluor488 secondary antibody diluted 1:300 in PBS-DAPI and finally washed twice in PBS. We took 4 pictures of each well with a Zeiss AXIO Observer A.1 with a 10X objective, both in the DAPI and Cy3 channels. For BrdU image analysis see QUANTIFICATION AND STATISTICAL ANALYSIS, section "*Image Analysis*."

For immunofluorescence of p-lin3xF-transfected cells, they were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) in PBS for 20 min at 4°C. Cells were then washed in PBS and permeabilized and blocked with 0.2%Triton X-100 /0,2% BSA in PBS for 20 min at room temperature. Subsequently, cells were processed for immunostaining with anti-FLAG M2 primary antibody (Sigma Aldrich) diluted 1:1000 in 0,2% BSA/1% Goat Serum/PBS for 16h at 4°C. Cells were then incubated with secondary antibody (Goat anti-Mouse IgG Cy3-conjugated, Jackson ImmunoResearch, West Grove, PA) diluted 1:500 in 1% Goat Serum/PBS for 45 min at room temperature. Finally, after extensive washing, DAPI (Sigma-Aldrich) was used to label nuclei and the coverlips were mounted in ProLong Diamond reagent (ThermoFischer Scientific). Samples were imaged on an Olympus IX73 inverted microscope equipped with Confocal Imager CREST X-LIGHT spinning disk, a CoolSNAP Myo CCD camera (Photometrics) and a Lumencor Spectra X LED illumination. The images were acquired using a LUCPlanFLN 20X objective (NA 0,45) and a UPLANSApo 60 X oil objective (NA 1.35) and collected with MetaMorph software (Molecular Devices). Stacks of images were taken automatically with 0,2 microns between the Z-slices. All Z stacks were processed with ImageJ software and merged with maximum intensity projection.

Protein IP and mass spec

Cells were trypsinized and pelleted at 500 g for 5 min, washed in PBS and lysed in an equal volume to the cell pellet of lysis buffer (250 mM NaCl, 50 mM Tris pH 8, 5 mM EDTA, 0.8% NP40, 1 mM DTT, 5% glycerol, 50 mM NaF) with 20 strokes in a dounce homogenizer. Lysates were left on ice for 5 min and centrifuged at 15000 g for 10 min. Protein concentration of the supernatant was then measured by Bradford assay and a volume containing 5-10 mg of proteins was diluted 1:2 in dilution buffer (50 mM Tris pH 8, 5% glycerol). Fifty μl of protein G-coupled dynabeads (immunoprecipitation kit, Invitrogen) were washed in binding buffer (form the kit), resuspended in 200 μL of binding buffer containing 10 μg of anti- FLAG M2 antibody (Sigma-Aldrich Aldrich) and left rotating on a wheel for 1 hr at 4°C. Beads were then washed in binding buffer and incubated with the diluted lysates in a final volume of 600-800 μl. The binding was carried out at 4°C with the samples rotating on a wheel. Beads were then washed 3 times with low salt wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1mM EDTA, 0.25% NP40, 5% glycerol), 2 times in high salt wash buffer (same as low salt with 300 mM NaCl) and once in PBS. Samples were eluted with 20 □l Elution buffer (from the kit) and 10 μL LDS sample buffer (Invitrogen). Proteins were run on a 4%–12% polyacrylamide gel for 1 hr at 160 V and bands were cut according to the protein marker between 30 and 40 KDa. Bands were stored at 4°C in 300 μL H2O and sent for mass spectrometry analysis to the proteomic platform at IGBMC - Strasbourg, where they were digested with trypsin and analyzed according to their standard mass spectrometry pipeline.

m6A IP

Cells were washed twice with PBSand lysed with 2ml of Qiazol (QIAGEN). The RNA was extracted according to Qiazol protocol. The m6A RIP was performed according to Dominissini et al. (2012) as follow: a volume containing 100 μg of RNA was diluted in a final volume of 300 mL with IPP buffer (10 mM Tris pH 7.4,150 Mm NaCl, 0.1% NP-40) containing 10 μg of anti-m6A antibody and incubated for 2 hr at 4°C.

The mixture was then immunoprecipitated by incubation with 50 μL of G-coupled dynabeads (immunoprecipitation kit, Invitrogen) at 4°C for an additional 2 hr.

Beads were then washed 5 times with IPP buffer and resuspended in 500 μL of Qiazol.

RNA was extracted according to Qiazol protocol and analyzed by qRT-PCR. Beads-only and isotypic IgG-coupled-beads were used as negative controls.

Genome editing with Crispr/Cas9

The insertion of a 3xFLAG coding sequence at position 3 of the Zfp 609 exon 2, immediately before the stop codon, was obtained in mouse embryonic stem cells by homology-directed repair of a Crispr/Cas9-induced DNA break. A guide RNA targeting the beginning of Zfp 609 exon 2 was designed with CRISPR design (http://crispr.mit.edu) and cloned in a px330 vector (see *Plasmid Construction* section). As a DNA donor, both a ssDNA oligo (ssDNA donor) and a plasmid donor carrying the same sequence plus 800 nt of homology upstream and downstream were used. Transfection of 500 ng of donor DNA plus 500 ng of px330 was performed in mES cells with 2 μL of lipofectamine 2000. After 48 hr, cells were split and a half was used for gDNA extraction and genotyping. gDNA was extracted with Real Genomics Genomic DNA extraction kit. Oligos crispr-A, B and C were used for amplifying the WT and recombinant alleles (Figure S6D) on control cells and transfected cells. PCR products were Sanger sequenced and the plasmid donor-transfected cells resulted positive for the insertion. We diluted trypsinized cells at 0.5 cells/well in two 96 wells, and recovered a total of 20 colonies with the plasmid donor-transfected cells. Colonies were split after 10 days and half of each was grouped producing 3 pools, which were separately genotyped. Pool #2 appeared positive, therefore the single colonies were further scrutinized and only one positive clone was identified (named ''F/D,'' Figures S6E and S6F).

Plasmid construction

p-circ and p-circ-3xF were kindly provided by Hung Ho Xuan and Gunter Meister. Briefly, p-circ was produced by cloning the circZNF609 sequence in the circRNA mini vector ZKSCAN1 (Kramer et al., 2015), addgene #60649) and p-circ-3xF was derived by inserting a 3xFLAG-coding sequence (GACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGA CGATGA CAAG) immediately upstream of the TAA codon present at position 3 in the circRNA sequence. We obtained the Δ1 derivative through inverse PCR on the p-circ-3XF vector with noATG1-ZNF609_FW and noATG1-ZNF609_RV oligonucleotides, while the Δ2 with noATG2-ZNF609_FW and noATG2- ZNF609_RV, the p-STOP2 with STOP2_fw and noATG2-ZNF609_RV on the p-circ-3xF template. The Δ1-2 was produced with an inverse PCR with noATG1-ZNF609_FW and noATG1- ZNF609_RV on the Δ2 template.

The p-lin-3xF vector was produced by inserting an amplicon obtained with pZNF_FW and pZNF_RV oligonucleotides from C_2C_{12} gDNA using the In-Fusion HD Cloning Kit (Clontech) into pcDNA3.1- (Invitrogen), digested with EcoRI and NheI. The 3xFLAG tag was then inserted into the resulting vector through inverse PCR with 3xFlag_1/2FW and 3xFlag_1/2RV oligonucleotides. The oligonucleotides (Flag_zfp_guide2_FW and Flag_zfp_guide2_RV) encoding the Cas9 guide RNA were annealed and ligated in a BbsI-digested px330 vector (addgene #42230, (Cong et al., 2013)).

Luciferase bicistronic reporter constructs were obtained from pGL3-Control vector (Promega) and pRL-TK vector (Promega). In order to get p-Luc empty vector, FLuc gene was isolated from pGL3-Control vector through the digestion with the restriction enzymes NheI and XbaI. This fragment was ligated to pRL-TK Vector digested with NheI. ZNF609 5′UTR was amplified with FW_5′UTR-circZNF609 and REV_5′UTR-circZNF609 oligonucleotides and cloned between the two luciferase genes in p-Luc empty vector, linearized by inverse PCR using FW_inversePCR-1 and RV_inversePCR-1 oligonucleotides. In order to obtain the UTR vector, eighty-four nucleotides upstream the 5′ splice junction of ZNF609 s exon were amplified with Fw_84ntZNF609 and Rev_84ntZNF609-2 oligonucleotides. The fragment was then inserted in the vector containing the ZNF609 5′UTR, linearized by inverse PCR with FW_inversePCR-2 and RV_inversePCR-1 oligonucleotides. In order to get UTR+In.β Glob vector, eighty- four nucleotides upstream the 5′ splice junction of ZNF609 s exon were amplified with Fw_84ntZNF609 and Rev_84ntZNF609 oligonucleotides and the second intron of Hbb-bs gene was amplified with Fw_b-globin-intr2 and Rev_b-globin-intr2 oligonucleotides. The vector containing only ZNF609 5′UTR was linearized as previously described, then the amplicons were cloned upstream ZNF609 5′UTR to obtain UTR+In.β Glob vector. ZNF609 UTR, 84-nucleotide fragment and Hbb-bs second intron were amplified from C_2C_{12} gDNA by PCR. UTR+In.βGlob.D5′ss vector was obtained by inverse PCR from UTR+In.βGlob vector, using Luc_DeltaSplicing_FW and Luc_DeltaSplicing_RV oligonucleotides. In order to obtain UTR+In.-ZNF vector, 485 nucleotides from ZNF609 Intron 1-2 were amplified from mouse gDNA with FW_pLuc_In.Znf and RV_pLuc_In.Znf oligonucleotides. FW_pLuc_In.Znf contains forty nucleotides from the 5′ of ZNF609 Intron 2-3 right upstream the Intron 1-2 annealing sequence. This amplicon was inserted into UTR vector, which was previously linearized by inverse PCR, using FW_inversePCR-2 and RV_inversePCR-3 oligonucleotides. UTR+In.HPRT vector was obtained inserting HPRT Intron 8-9 into UTR vector, which was linearized by inverse PCR as previously described. HPRT Intron was amplified from mouse gDNA, using FW_pLuc_In.HPRT and RV_pLuc_In.HPRT oligonucleotides. In order to get MCS+In.βGlob vector, p-Luc empty vector was linearized by inverse PCR, as previously described. The second intron of Hbb-bs gene was amplified from UTR+In.βGlob vector, using FW_pLuc_In.bGlob_noUtr and RV_pLuc_In.bGlob_noUtr oligonucleotides. This amplicon was cloned in p-Luc vector. Then, this construct was linearized by inverse PCR, using FW_inversePCR-1 and RV_inversePCR-4 oligonucleotides. MCS sequence from pcDNA3.1(+) (Thermo Fisher Scientific) was amplified using FW_pLuc_MCS_noUtr and RV_pLuc_MCS_noUtr oligonucleotides. This amplicon was then cloned in the previous construct, in order to obtain MCS sequence downstream the βGlobin intron. EMCV IRES was amplified from pIRES2-EGFP plasmid (Clontech) with FW_IRES and RV_IRES oligonucleotides and cloned

between luciferase genes (pLuc-EMCV-IRES) in p-Luc empty vector, linearized as previously described. UTR vector, UTR+In.βGlob vector, UTR+In.-ZNF vector, UTR+In.HPRT vector, MCS+In.βGlob vector and EMCV-IRES vector were realized with In-Fusion HD Cloning Kit (Clontech).

Preparation of synthetic circRNAs

The linear version of the flagged circZNF609 transcript was obtained by in vitro transcription from a PCR-generated template (using p-circ-3xF as template and primers T7_ZNF_Flag_Fw and ZNF_Flag_Rv listed in Table S2) in presence of T7 polymerase (Promega) following the manufacturer's instructions. Upon DNase treatment, The RNA was purified through polyacrylamide gel electrophoresis and the 5′-terminal triphosphate removed by treatment with Alkaline Phosphatase (Roche) in presence of RNase inhibitors. The RNA was then subjected to Ethanol precipitation upon the addition of 10 μg of glycogen (Roche). A phosphate group was then attached to the 5′-OH using ATP and T4 Polynucleotide kinase (BioLabs); the linear transcript carrying now 5′-phosphate and 3′-OH ends was subjected to Ethanol precipitation in presence of 10 μg of glycogen (Roche). Ligase reaction was carried out in a final volume of 100 μl, incubating the linear transcript at 95°C for 2 min followed by 5 min at 75°C in presence of 10% DMSO and equimolar amount of splint DNA (oligo in Table S2). T4 RNA Ligase 1 (BioLabs), 1X T4 RNA Ligase Buffer, 10mM ATP and RNase inhibitor will be then added and the reaction was carried out at 16°C for 16 hr. The circularized RNA product was separated and purified from the linear transcript by polyacrylamide gel electrophoresis. Thirty nanograms of both circular and linear ZNF609 transcripts were subjected to RNase R treatment followed by qRT-PCR to assess the circularity of the gel-purified RNA molecules.

Luciferase assay

pLuc empty vector, pLuc-UTR vector, pLuc-intr.-UTR vector, pLuc-EMCV-IRES vector and pLuc-intr.-UTR-D-SpI. vector were transfected in C_2C_{12} cells as follows: 1 μL of lipofectamine 2000 was added to 100 μL of Optimem together with 500 ng of DNA, mixed and pipetted onto 50,000 pre-seeded cells in a 12 well plate. Medium was replaced 4 hr later and cells harvested for subsequent analyses after 48 hr.

Cells were harvested with PLB lysis buffer (Promega) and RLuc and FLuc activities were measured by Dual Glo luciferase assay (Promega) according to the manufacturer's protocol. A part of cell lysate was kept for RNA extraction and checked by RT-PCR with FLuc-pGL3control_FW and RLuc-pRLTK_RV oligonucleotides.

QUANTIFICATION AND STATISTICAL ANALYSIS

RNaseq analyses

For standard gene expression analysis, reads were mapped with TopHat, transcriptomes defined by Cufflinks and differential expression analyzed with Cuffdiff. For circRNA detection, paired-end reads were split and used separately as input to the following pipeline: reads were aligned to rRNA with Bowtie2, the remaining ones were mapped to hg19 or mm10 genome assemblies with Bowtie2. Reads that were left after these alignments were used by FindCirc, Memczak et al., 2013) for finding circRNAs. For quantitative analyses of circRNA regulation in human muscle differentiation, circRNAs with at least 5 reads in one sample were used. The total number of reads mapping to the head-to-tail splice junction was used for measuring circRNA expression level, while the mean of total reads mapping linearly to the same splice junction was used for measuring the abundance of the linear isoform. For comparing WT and DMD samples, circRNA expression level was normalized by dividing the reads mapping to the head-to-tail junctions by the total number of spliced reads identified for each sample by FindCirc.

After the circRNA detection process, all human circRNA events were annotated using bedtools to the Ensembl GRCh37 gene set. Annotation for gene structure features (CDS, UTR, Intron etc.) was performed with the R package CiRcus. For experimental validation, a subset of circRNAs were selected according to the following criteria: human circRNAs were first filtered for expression (at least 5 unique reads in one sample, at least 2 unique reads in its biological replicate), then sorted by the average fold change ratio between myotubes and myoblasts, then filtered for conservation (at least two unique reads in one sample in mouse of an overlapping circRNA). The ones with an absolute fold change ratio > 2 were then sub-divided in 6 classes, according to the expression pattern of their cognate mRNA: (circRNA upregulated > 2, downregulated < 0.5) x (mRNA upregulated > 2, downregulated < 0.5, non regulated). Two or 3 of the top candidates were picked among each class, yelding a list of 16 circRNAs (ACVR2A, ANKIB1, ASH1L, ASPH1, BNC2, DCBLD2, MEF2A, MYL4, NEB, PMS1, PTP4A2, RTN4, RUNX1, SEPT9, SLC8A1, TMEFF1, TTN). Additional candidates were added according to their high expression level (HIPK3, ZNF292, BACH1, SPECC1, RTN4bis) and their high circular/linear ratio (ADAMTS6, CAMSAP1, CDYL, MED13L, ZFHX3, ZNF609, TTTY16). To such list, we added CRKL and QKI for the known role of their host genes in muscle differentiation. For each circRNA selection after sorting for fold change, expression level, circular/linear ratio, we manually inspected the RNaseq coverage and isoforms expression, and we manually excluded those situations that seemed too complicated for a validation experiment (too many isoforms, unexpected coverage, apparent *trans*-splicing events etc.). More details about bioinformatics analyses, all files and all scripts are available upon request.

Image analysis

The acquired images were analyzed using two custom image analysis algorithms (available upon request), developed and executed in the Acapella software development/run-time environment (Perkin Elmer). Both algorithms used a Perkin Elmer proprietary algorithm on the DAPI staining for nuclei detection.

DAPI and MYOG staining at 48 hr were used to extract (for each siRNA treatment) the total cell number, the percentage of MYOG+ cells, and the average MYOG signal intensity. The DAPI&MYOG algorithm workflow can be summarized as follows: 1) nuclei detection (on DAPI channel); 2) MYOG signal intensity estimation (from Cy3 channel) and 3) MYOG ± nuclei classification. In the absence of a cytoplasmic marker, following the segmentation on DAPI, Voronoi tessellations were used to define local regions for each detected nucleus. Thus, foreground and background masks were defined to represent the nuclear area and its local background. These masks were further used to extract average DAPI and MYOG intensities for every nucleus. The standard deviations of MYOG intensity in the nuclei local background were estimated. The DAPI/MYOG signal intensity of every nucleus was defined as the difference between the DAPI/MYOG intensities of its local foreground and background. This local approach compensates for both intra- and inter-image spatial variability of illumination and minimizes possible errors in estimating DAPI/MYOG signal intensities over different wells (and fields). Following a similar local approach, and this time compensating for different inter-image noise level/content, the MYOG-positive and -negative classification was developed to use a local image adaptive threshold to apply on the MYOG signal intensity. For each analyzed image, the threshold was defined as 1.5*median of all standard deviations of MYOG signal intensities of the local backgrounds of all nuclei detected in the current image. Then, each nucleus was classified as MYOG+ or MYOG- based on the comparison of its MYOG signal intensity and the current local image threshold (see the example Figure showing the MYOG-positive classification).

DAPI and MHC staining at 96 hr were used to extract (for each siRNA treatment) the total number of cells, the percentage of MHC-positive nuclei, and additional morphological parameters such as the fusion index and the number of nuclei per fiber. The algorithm workflow can be summarized as follows: 1) nuclei detection (on DAPI channel); 2) nuclei count correction procedure; 3) cytoplasm segmentation (on Red channel); 4) MHC signal intensity estimation (from the Red channel); 5) MHC-positive and -negative nuclei classification; 6) Red Cell/Fiber MHC+ nuclei classification. Prior to all analyses, a circular mask was applied on the images of all channels in order to remove vignetting.

Following DAPI segmentation of the nuclei, we applied a nuclei count correction procedure. This was required to overcome the issue of contiguous, 'syncytial-like' nuclei that were present in myotubes. The nuclei count correction procedure was developed based on a combination of some morphological features of the detected nuclei: the normalized nuclear area (with respect to the median of all nuclei area of the current image) and the nuclear axial length ratio (nucleus width/length). We could distinguish three classes of nuclei. First, all nuclei having a normalized area smaller than 1, which count as 1 nucleus each. A second class was defined by taking the integer of values derived after applying the floor function to nuclei with a normalized area greater than 1 (for example, nuclei with values between 2.0 and 2.99 were scored as 2, while those between 3.0 and 3.99 were scored as 3). Third, we defined a subclass with all nuclei having a normalized area between 1.5 and 2 and with an axial length < 0.5 were counting as 2 nuclei each.

We used a Perkin Elmer proprietary algorithm to define the cytoplasmic area for each nucleus. The cytoplasmic area is assigned to the nuclei based on the signal in the Red channel. This was associated with two problems. First, because of the peculiar morphology of cells and fibers (cells have elongated shape, fibers appear stitched/clumpy and are multinucleated), individual cells/fibers were erroneously subdivided. Second, in some cases, the method attempted to assign cytoplasm to cells that, on visual inspection, had no signal in the Red channel.

Thus, we defined a population of 'cell' objects including a nuclear region (with one or more nuclei) and a cytoplasmic region. All the detected 'cell' objects that had a cytoplasmic/nuclear area smaller than one were removed from the cytoplasmic segmentation result, since they were potentially MHC-negative.

From the remaining population of cells, we defined a cellular mask (cytoplasm and nuclear areas) on which MHC-positive/negative nuclei classification was performed. All nuclei with at least one quarter of their area overlapping the cellular mask were classified as MHC-positive. All other nuclei were labeled as MHC-negative.

The Red Cell/Fiber classification was implemented using a work-around of the cytoplasm segmentation. The definition of a fiber was as follows: stitched/clumpy cells with multiple and short distanced nuclei (clusters). Red Cells were defined as elongated and mono-nucleate. We used the presence of nuclei clusters to differentiate between Red Cells and Fibers. Specifically, nuclei that were within 100 μm of each other were defined as being part of the same cluster. Thus, we used the morphological dilation operator (circular structuring element of 30 pixels diameter) on the nuclear mask given by all MHC-positive nuclei. The resulting 'dilated' mask was then intersected ('AND' operator) with the cellular mask to separate nuclei of neighboring Red Cells. This resulting binary image was used to define nuclei clusters by using connected component labeling (Ballard and Brown, 1982) and selecting only those components containing more than one MHC-positive nucleus. Finally, all MHC-positive nuclei belonging to nuclei clusters were labeled as 'Fiber nuclei'. Similarly, all those MHC-positive nuclei that were not part of a nuclear cluster are labeled as Red Cell nuclei. The result of this classification was transferred to the nuclei-associated 'cell' objects, and thus labeled as Fibers and Red Cells.

For each well, this algorithm reports a set of 8 parameters. The first four parameters are based on nuclei, namely, total number of (detected) nuclei, percentage of MHC-positive nuclei, percentage of MHC-positive Red-Cell nuclei, average number of nuclei per Fiber (estimated by the average number of nuclei per nuclei cluster). The following two parameters are cell area descriptors: the fraction of Red Cell area and the fraction of Fibers area. These fractions were estimated as the total area of Red Cells/Fibers with respect

to the entire area of the analyzed image. Two additional parameters were implemented to better describe cellular phenotypes. The fusion index was defined as the fraction of total number of Fiber nuclei to the total number of detected nuclei. The maturation index was defined as the fraction of the total number of Fiber nuclei to the total number of MHC-positive nuclei. All raw images, segmentation results and scripts are available upon request.

BrdU images were analyzed manually by using ImageJ: tiff files were converted into binary images by setting a fixed threshold for both DAPI and BrdU signal intensity, merged edged of close nuclei were split through the Watershed plug-in and nuclei count was performed with the built-in particle analysis plug-in. Accepted nuclei dimension were set in a range corresponding to the second and third quartiles of all particles detected.

Statistical Analyses

All experiments were repeated for at least 3 times unless differently stated in the legend; arithmetic means together with standard error are reported in all plots shown in this study unless differently stated in the legend. Statistical significance for comparisons of means was assessed by Student's t test for qRT-PCRs, luciferase assays and immunofluorescence-based screenings. Dedicated statistics were applied to RNaseq data analysis as specified in the legends. P values below 0.05 were marked by 1 asterisk, while 2 asterisks indicate a p value < 0.02 and 3 asterisks a p value < 0.01. When multiple comparisons were performed, p values were scaled according to the Bonferroni correction and then filtered for corrected p value < 0.1.

DATA AND SOFTWARE AVAILABILITY

All software used in this manuscript is listed in the Key Resources Table. Additional dedicated scripts developed for this work are available upon request. The accession number for the sequences reported in this paper is GEO: GSE70389.