Human Molecular Genetics, 2017, Vol. 26, No. 18 3564–3572

doi: 10.1093/hmg/ddx243
Advance Access Publication Date: 23 June 2017
Original Article

ORIGINAL ARTICLE

Circular RNA profiling reveals that circular RNAs from *ANXA2* can be used as new biomarkers for multiple sclerosis

Leire Iparraguirre[1,†], Maider Muñoz-Culla[1,2,†], Iñigo Prada-Luengo[1], Tamara Castillo-Triviño[1,2,3], Javier Olascoaga[1,2,3] and David Otaegui[1,2,*]

[1]Multiple Sclerosis Unit, Neurosciences Area, Biodonostia Health Research Institute, 20014, San Sebastián, Spain, [2]Spanish Network of Multiple Sclerosis, 08028, Barcelona, Spain and [3]Neurology Department, Donostia University Hospital, 20014, San Sebastián, Spain

*To whom correspondence should be addressed at: Neurosciences Area, Biodonostia Health Research Institute, Dr. Begiristain s/n, 20014, San Sebastián, Spain. Tel: +34 943006293; Fax: +34 943006250; Email: david.otaegui@biodonostia.org

Abstract

Multiple sclerosis is an autoimmune disease, with higher prevalence in women, in whom the immune system is dysregulated. This dysregulation has been shown to correlate with changes in transcriptome expression as well as in gene-expression regulators, such as non-coding RNAs (e.g. microRNAs). Indeed, some of these have been suggested as biomarkers for multiple sclerosis even though few biomarkers have reached the clinical practice. Recently, a novel family of non-coding RNAs, circular RNAs, has emerged as a new player in the complex network of gene-expression regulation. MicroRNA regulation function through a 'sponge system' and a RNA splicing regulation function have been proposed for the circular RNAs. This regulating role together with their high stability in biofluids makes them seemingly good candidates as biomarkers.

Given the dysregulation of both protein-coding and non-coding transcriptome that have been reported in multiple sclerosis patients, we hypothesised that circular RNA expression may also be altered. Therefore, we carried out expression profiling of 13.617 circular RNAs in peripheral blood leucocytes from multiple sclerosis patients and healthy controls finding 406 differentially expressed (P-value < 0.05, Fold change > 1.5) and demonstrate after validation that, circ_0005402 and circ_0035560 are underexpressed in multiple sclerosis patients and could be used as biomarkers of the disease.

Introduction

Multiple sclerosis is a common immune-mediated demyelinating disease of the CNS. It represents the leading cause of non-traumatic neurological disability in young adults between 20 and 40 years old (1,2). As a reflection of its complex pathophysiology, multiple sclerosis has a heterogeneous clinical presentation and course, so different clinical forms are distinguished. Relapsing–remitting multiple sclerosis (RR–MS), which represents around 85% of the cases, is the most prevalent, characterised by outbreaks of neurological disability symptoms lasting at least 24 h (relapses), followed by a complete or incomplete recovery phase (remission) (3,4).

The prevalence of multiple sclerosis is higher in women (ratio 2.6: 1) as with many autoimmune diseases (1). Indeed, gender effects have also been observed in other aspects of the disease including severity, clinical course, susceptibility and autoimmune response (5–7). Furthermore, transcriptome

[†]These authors contributed equally to this work. The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Received: April 27, 2017. **Revised:** June 13, 2017. **Accepted:** June 15, 2017

expression analysis has also revealed differences between female and male patients (8–11).

The broad phenotypic heterogeneity together with unpredictable clinical relapses and remissions make multiple sclerosis treatment and prognosis quite challenging. Thus, new tools such as precise and sensitive biomarkers are needed in order to help in the diagnosis, prognosis and decision-making in the treatment of the disease. In addition, biomarkers could also reflect pathogenic mechanisms of multiple sclerosis and thus could help to identify new therapeutic targets. For many years, intensive efforts have been directed towards the identification of biomarkers in body fluids, however, only a few have been incorporated into clinical practice (12,13). During recent years, the application of more advanced screening technologies such as gene expression and autoantibody arrays has opened up new categories of biomarkers.

Non-coding RNAs comprise one of the new class of biomarkers that have been proposed for several diseases (14). Among all the classes of non-coding RNAs (ncRNAs), microRNAs (miRNAs) are the most widely investigated. MiRNAs are single-stranded RNAs of approximately 22 nucleotides that regulate the expression of their target mRNAs in a post-transcriptional fashion (15). Hence, they are considered key players in the complex network of gene expression regulation and are known to be involved in several biological processes including development, differentiation and immunity (16). It has also been shown that they take part in several diseases such as cancer, neurodegeneration and autoimmune diseases (17–19).

Regarding multiple sclerosis, several gene expression studies have described a dysregulated network of both transcriptome and miRNAs in immune and neural cells from patients as well as in the experimental autoimmune encephalomyelitis (EAE) animal model (reviewed in (20) and (21)). Although the functional and pathological implications of the dysregulation are still not completely understood, the deregulated expression of certain cellular miRNAs has been proposed as a putative trigger in multiple sclerosis physiopathology (reviewed in (22)). In addition to cellular miRNA, miRNA expression studies of multiple sclerosis patients' blood and plasma have generated a significant amount of data about circulating dysregulated miRNAs. In a recent systematic review performed by Jagot *et al.*, they reported that there is current evidence of at least 62 circulating miRNAs significantly dysregulated in multiple sclerosis patient's blood, while profiling results are still continuously emerging (21). Some of those candidates have been suggested as biomarkers for disease diagnosis, severity or response to therapy, although further validation is needed to implement these candidate biomarkers in clinical practice (16,23–26). All things considered, the idea is that the regulatory network of the transcriptome is different in multiple sclerosis patients, which highlights the importance not only of the dysregulated mRNAs but also of the transcriptome regulators.

Recently, among an ever-growing list of RNA species, a novel class of ncRNAs has gained attention, circular RNAs (circRNAs). They were originally considered as secondary by-products of linear mRNA splicing and therefore, thought to be unlikely to play any important function in biological processes. However, the development of new RNA sequencing methodologies, and the growing production of datasets, as well as bioinformatics advances enabled discovering that large numbers of circRNAs are endogenous, abundant, stable and well conserved in mammalian cells (27–31). CircRNAs form through RNA splicing, resulting in a covalently closed loop structure characterized by a backspliced junction that makes them distinguisable from their linear counterparts. Although their functions remain unclear, several possibilities have been reported, with the strongest being their capacity to work as miRNA sponges (32,33), suggesting that they might also play an important role in the transcriptome regulation network. During the last 2 years circRNAs have been widely studied and related to some diseases such as cancer, vascular diseases, diabetes, osteoarthritis and neurological diseases (34–38).

Given the stability that circularity confers to this kind of RNAs, they have been reported as promising candidates for biomarkers of human diseases. Indeed, several circRNAs have already been suggested as biomarkers for a potential non-invasive diagnosis of several diseases such as vascular diseases, neurological diseases or cancers (reviewed in (39,40)).

Given the dysregulation of both protein-coding and non-coding transcriptome that have been reported in multiple sclerosis patients, we hypothesised that circRNA expression may be altered in patients. Therefore, we carried out circRNA expression profiling in peripheral blood leucocytes from multiple sclerosis patients and healthy controls and demonstrate that after validation experiments, circ_0005402 and circ_0035560 are underexpressed in multiple sclerosis patients, suggesting that they could be promising blood biomarkers for the disease.

Results

Differentially expressed circRNA profile by microarray

Out of the 13,617 probes interrogated in eight blood samples (four RR-MS untreated patients in remission and four healthy controls, discovery cohort), 10,183 (74.78%) are present in at least one sample. Among them, 406 are differentially expressed ($FC \geq 1.5$, $P \leq 0.05$) in patients, compared to healthy controls. 324 are downregulated and 82 upregulated (Fig. 1A). Hierarchical clustering performed with differentially expressed circRNAs, shows that this circRNA expression pattern is able to classify individuals according to their status (Multiple sclerosis or Controls) except one healthy control, who is grouped with multiple sclerosis patients (Fig. 1B).

Validation of candidate circRNAs by qPCR

Among the 406 differentially expressed circRNAs, the five upregulated circRNAs (circ_0000518, circ_0000517, circ_0000519, circ_0000520, circ_0001400) and the five downregulated circRNAs (circ_0056731, circ_0064644, circ_0005402, circ_0024892 and circ_0035560) showing the highest fold change were chosen in order to perform a technical validation of microarray results by qPCR (Table 1). Divergent primers were designed for each circRNA and 9/10 were successfully amplified, whereas hsa_circ_0064644 could not be amplified and was therefore excluded from this study.

The expression trend of five circRNAs is consistent with microarray results whereas in four circRNAs, it is opposite to that found in the microarray. After a second technical validation performed by Arraystar (Arraystar, Inc., USA), we find out that circ_0056731, circ_0005402, circ_0024892 and circ_0035560 are validated consistently by at least one of the validation qPCRs (Fig. 1C). Therefore, those four circRNAs are further selected for validation in a larger cohort of 20 RR-MS untreated patients in remission and 18 healthy controls (first validation cohort).

The Wilcoxon test shows that the expression level of both circ_0005402 ($FC = 0.395$, $P = 0.00007$) and circ_0035560 ($FC = 0.737$, $P = 0.0304$) is significantly lower in patients

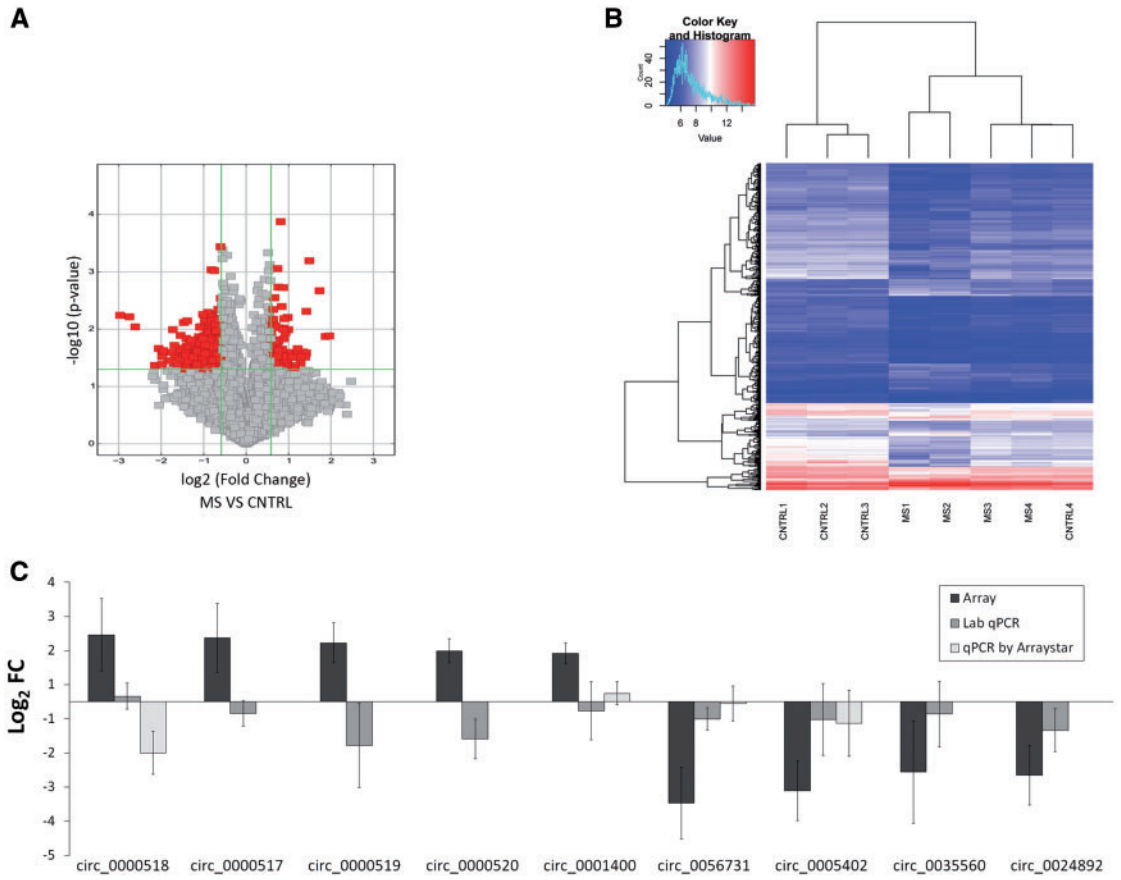

Figure 1. Differentially expressed circRNAs in multiple sclerosis. (**A**) Volcano plot showing the results for the 10.183 circRNAs expressed in at least one of the four samples studied in the array. The vertical lines correspond to 1.5-FC up and down, and the horizontal line represents a *P*-value of 0,05, so red dots depict the 82 upregulated and the 324 downregulated circRNAs with statistical significance. (**B**) Heatmap showing the differentially expressed circRNAs which quite efficiently clusters control vs patient samples. The colour scale reflects the log2 signal intensity and runs from blue (low intensity) to red (high intensity). The dendrogram shows the relationships among the expression levels of samples. (**C**) Comparison of circRNA expression levels between microarray and the technical validation results by qPCR in which four circRNAs were validated consistently by at least one of the validation qPCRs. Error bars represent the standard deviation for the patient group ($n = 4$) for which the fold change is calculated (MS= multiple sclerosis; CNTRL = control).

Table 1. Top ten deregulated circRNAs in multiple sclerosis. *P*-value was calculated from t-test

circRNA	Alias	Regulation	FC	*P*-value	circRNA type	Position	Best transcript	Gene
hsa_circRNA_000167	hsa_circ_0000518	up	3.89	0.0131	sense overlapping	chr14:20811404-20811554	NR_002312	RPPH1
hsa_circRNA_001678	hsa_circ_0000517	up	3.65	0.0134	sense overlapping	chr14:20811404-20811492	NR_002312	RPPH1
hsa_circRNA_002178	hsa_circ_0000519	up	3.31	0.0021	sense overlapping	chr14:20811436-20811534	NR_002312	RPPH1
hsa_circRNA_001846	hsa_circ_0000520	up	2.80	0.0006	sense overlapping	chr14:20811436-20811559	NR_002312	RPPH1
hsa_circRNA_103619	hsa_circ_0001400	up	2.67	0.0049	exonic	chr4:37633006-37640126	NM_001085399	RELL1
hsa_circRNA_056731	hsa_circ_0056731	down	7.82	0.0057	exonic	chr2:152698416-152717334	NM_000726	CACNB4
hsa_circRNA_103309	hsa_circ_0064644	down	6.72	0.0061	exonic	chr3:29910348-29941246	NM_014483	RBMS3
hsa_circRNA_101539	hsa_circ_0005402	down	6.11	0.0091	exonic	chr15:60648117-60674640	NM_004039	ANXA2
hsa_circRNA_024892	hsa_circ_0024892	down	4.44	0.0427	exonic	chr11:130749518-130749606	NM_014758	SNX19
hsa_circRNA_101541	hsa_circ_0035560	down	4.17	0.0217	exonic	chr15:60653139-60674640	NM_004039	ANXA2

compared to controls. For circ_0056731 and circ_0024892 the same downregulation trend is maintained, however the difference is not statistically significant (FC = 0.814, $P = 0.111$ and FC = 0.861, $P = 0.266$ respectively). When we measure the relative abundance of circRNAs compared to the linear transcripts derived from their respective host genes, only circ_0035560 (FC = 0.736, $P = 0.0412$) remains significantly deregulated while circ_0005402 is no longer significant (FC = 0.677, $P = 0.109$) (Fig. 2). Note that the expression of circ_0035560 is higher when compared to the expression of the linear RNA isoform.

Based on these results, the expression of circ_0005402 and circ_0035560 is studied in 21 relapse and remission paired samples in order to assess their potential implication in the biology of the disease. None of them passed the significance threshold in this second cohort (FC = 0.932, $P = 0.392$ and FC = 1.358, $P = 0.633$ respectively), neither do they in the analysis stratified

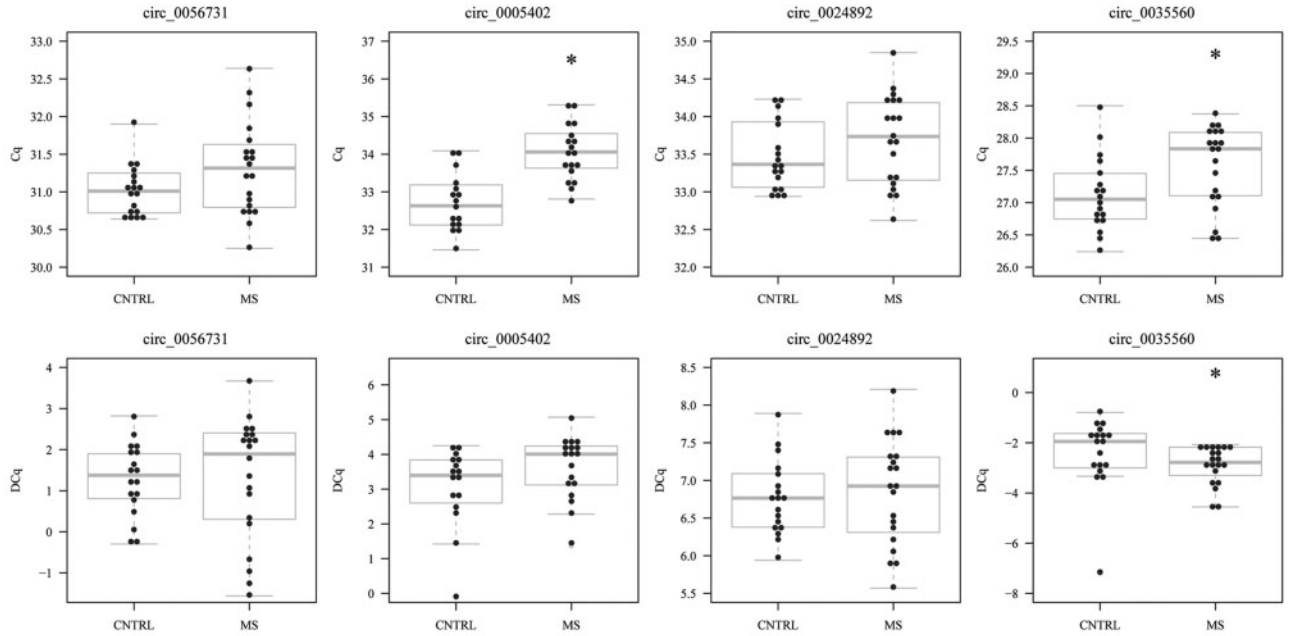

Figure 2. Candidate circRNAs expression in the first validation cohort. Both the expression of the circular isoform (above) and its expression after normalisation by the expression level of the corresponding linear counterpart (below) are depicted through dot and boxplots for the Multiple sclerosis and control groups. Note that the expression is represented by the Cq (above) and DCq (below) value obtained in the qPCR, where lower values correspond to higher expression values. *Indicates $P < 0.05$ assessed by Wilcoxon test. (MS= multiple sclerosis; CNTRL = control).

by sex (FC = 0.745, $P = 0.109$ and FC = 1.091, $P = 0.850$ for females and FC = 1.256, $P = 0.734$ and FC = 1.819, $P = 0.250$ for males) (Supplementary Material, Fig. S1). Moreover, to asses their early predictive value, two of those circRNAs were also studied in a third cohort of 10 patients with clinically isolated syndrome (CIS) for which no significant expression differences are found as assessed by the Wilcoxon test. Nevertheless, the same downregulation trend found for RR-MS patients is maintained (FC = 0.764, $P = 0.352$ and FC = 0.926, $P = 0.911$) (Supplementary Material, Fig. S2).

In order to assess whether the two differentially expressed circRNAs or even the linear form of the parental gene could be candidate biomarkers for the disease, the predictive value of circ_0005402, circ_0035560 and *ANXA2* is evaluated separately by employing the ROC curve analysis using the Cq value for each of them. The results show that the area under the curve (AUC) is significant ($p < 0.05$) for circ_0005402 with a value of 0.899 (95% CI: 0.747-0.976), a sensitivity of 94.4% and a specificity of 75.0%, for circ_0035560 with an AUC of 0.706 (95% CI:0.536-0.842), 55% sensitivity and 88.9% specificity, and also for the linear form of the gene for which the AUC is 0.831 (95% CI: 0.674-0.932) with a 80% sensitivity and 83.33% of specificity. (Fig. 3).

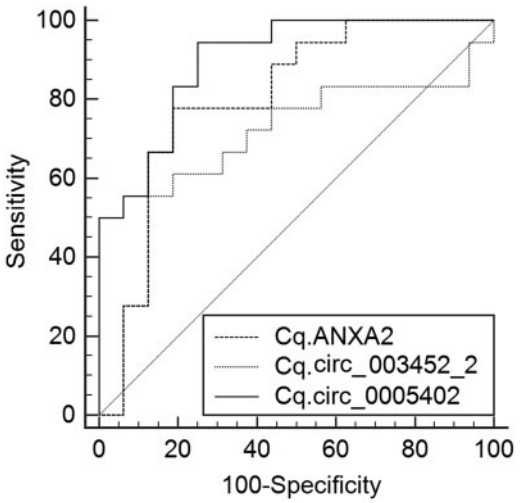

Figure 3. ROC curve for ANXA2 mRNA, circ_0003452_2 and circ_0005402 indicating their potential to be biomarkers. The AUC values are 0.831, 0.706 and 0.899 respectively.

Discussion

CircRNAs are a recently discovered class of ncRNA (33) with some special characteristics that make them good candidates to be biomarkers. They are highly conserved and are resistant to degradation by RNases, with a half-life over 48 h (30,41). In fact, in the last two years several papers have proposed a biomarker role for specific circRNAs in some diseases.

In this paper, we report a distinctive blood circRNA expression profile for RR-MS patients. To our knowledge this is the first study revealing a dysregulated profile of circRNAs in multiple sclerosis. During the preparation of our manuscript, Cardamone *et al*. presented a paper in which they studied the alternative splicing abnormalities in the *GSDMB* gene and found novel isoforms of this gene and a circRNA (hsa_circ_0106803) that is upregulated in RR-MS (42). In our dataset, the microarray analysis reveals the presence of 406 differentially expressed circRNAs, most of them (nearly 80%) showing a trend of downregulation in patients. From these 406, two circRNAs have been confirmed to be downregulated in the multiple sclerosis population after all the validations; circ_0005402, which is 6.11 times downregulated in arrays and 2.53 times in the first validation cohort, and circ_0035560, which is 4.17 times downregulated in arrays and 1.36 times in the first validation cohort. Moreover,

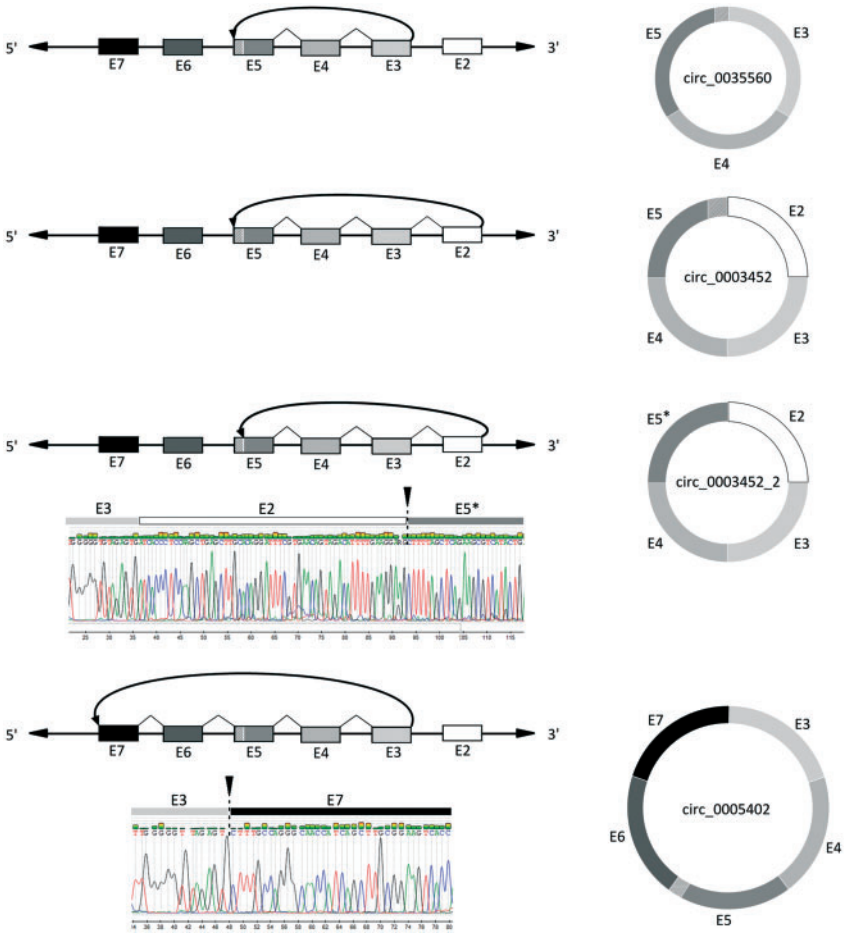

Figure 4. Schematic representation of four different backsplicing events of *ANXA2* leading to four circRNAs. Exons are indicated with E and the corresponding number numbered starting from the 3' since they are transcribed from the minus strand. E5* refers to shortened version of the exon five. For circ_0005402 and circ_0003452_2 the Sanger-Sequencing Electropherogram showing the backspliced junction is presented below the diagram.

this downregulation trend is also maintained for patients with CIS, pointing to an early dysregulation of their expression. Amplicons of both circRNAs have been sequenced to certify their circRNA condition.

Both of them are located in chromosome 15, inside the *ANXA2* gene (43) and have not been previously related to any pathology or biological process. Interestingly, the linear form of *ANXA2* is also downregulated. Even though the expression of these circRNAs might not be completely independent of the linear RNA isoform abundance, they are still notably downregulated as can be seen when the expression of circRNAs is normalised with *ANXA2* expression. This gene is predicted by circBase (http://circbase.org/; date last accessed June 27, 2017) to give rise to 30 circRNAs, from which only five have been included in the Arraystar platform (due to the fact that the rest did not fulfil the company's inclusion criteria). All of them are expressed in at least one sample, but only circ_0005402 and circ_0035560 are differentially expressed.

Up to now, based on the literature, *ANXA2* gene had not been related to multiple sclerosis, neither in Genome Wide Association studies nor in expression studies. However, it has been related to several cancer types (44–46) and immune-mediated diseases such as arthritis or antiphospholipid syndrome (47,48). Thus, *ANXA2* presumably has an immune function, however, whether it has a pro or anti-inflammatory role could be context dependent (47,49,50). Interestingly, *ANXA2* has been proposed to play a role in the traversal across the blood-brain barrier, the barrier that protects the CNS (51). *ANXA2* has also been reported to be a target of miR-155, a critical miRNA in neuroinflammation at the blood-brain barrier and relevant in Th1 and Th17 cell differentiation and myeloid cell polarisation in multiple sclerosis. MiR-155 has also been shown to be significantly increased in both peripheral blood mononuclear cells and active lesions and to correlate with disease severity in multiple sclerosis patients and mice with EAE (52–54). Those findings could be in agreement with the downregulation of *ANXA2*, suggesting a possible complex interaction between miRNA, mRNA and circRNAs. Moreover, *ANXA2*, shows an important role in post-transcriptional regulation of gene expression and in the transport of mRNA and vesicles (55). These processes have been proposed as clues to understanding the network of transcriptome regulation in multiple sclerosis.

Circ_0005402 has a size of 480 bp and includes exons three, four, five, six and seven (referred to the exons of the NM_004039 mRNA transcript) (Fig. 4) and the presence of the backspliced junction between exons three and seven has been confirmed in our data by Sanger Sequencing. For circ_0035560, the structure proposed in circBase includes exons three, four and five (Fig. 4).

Table 2. Main clinical and demographical characteristics of the individuals enrolled in the study. A total of 45 multiple sclerosis patients and 26 healthy controls were studied after being separated into four different cohorts

		Sex	Age	EDSS	Evolut. Time	AOO
Discovery cohort	RR-MS ($n=4$)	F	43.8 (±6.8)	3.5 (2–4)	19.3 (±3.8)	24.5 (±5.1)
	Controls ($n=4$)	F	34.3 (±6.7)	–	–	–
1st validation cohort	RR-MS untreated ($n=20$)	F	48.5 (±14.3)	1.5 (0–6.5)	15.2 (±15.2)	33.3 (±10.9)
	Controls ($n=18$)	F	44.9 (±19.9)	–	–	–
2nd validation cohort	RR-MS relapse ($n=21$)	12F/9M	40.1 (±10.1)	2.5 (0–7.5)	9.4 (±8.8)	30.1 (±11.5)
	RR-MS remission ($n=21$)	12F/9M	40.4 (±10.3)	2.5 (0–7.5)	9.4 (±8.6)	
CIS cohort	CIS cases ($n=10$)	F	35.5 (±10.4)	1.6 (0–4)	0.5 (±0.9)	35.0 (±10.1)
	Controls ($n=10$)	F	36 (±9.0)	–	–	–

Abreviations: RR-MS: Relapsing-remitting multiple sclerosis; CIS: Clinically isolated syndrome; F: female; M: = male; Evolut. Time: Evolution time; AOO: age of onset; EDSS: Expanded Disability Status Scale. Sex, age, evolut. time and AOO data are presented as 'average (standard deviation)', EDSS data are shown as 'median(range)'.

However, when we checked the sequence of our amplicon, we found a junction that includes exon two between exons five and three, but not the proposed structure. CircBase already includes a circRNA formed by the exons two, three, four and five, hsa_circ_0003452 (which is not included in the Arraystar system). However, its sequence does not totally overlap with the amplicon we are obtaining since it includes an 11 nt shorter sequence of exon five than hsa_circ_0003452. In a nascent field such as circRNAs, nomenclature and databases need to be refined after new data is generated. Therefore, we propose calling this circRNA hsa_circ_0003452_2. So, from now on, we will refer to what we first called circ_0035560 as circ_0003452_2.

Through bioinformatic analysis of the sequence of both circRNA using *Circular RNA Interactome* tool (56) we found that circ_0005402 presents binding-sites for five RNA-binding proteins; eight sites for AGO2, four for EIF4A3, two for LIN28A and one for IGF2BP1 and PTB. On the other hand, it also presents a single binding site for 17 miRNAs and two binding sites for miR-1248 and miR-766. Due to the fact that circ_0003452_2 is also downregulated and shares part of the sequence with circ_0005402, we checked for common miRNA binding sites. Indeed 14 out of 25 miRNAs are targeted by both circRNAs, suggesting a cooperative regulation of these miRNAs. Nonetheless, we cannot exclude other functions for these circRNAs, such as regulation of parental gene expression or interaction with RNA-binding proteins (57).

Finally, our results confirm that both circRNA expression and even *ANXA2* expression could be used as biomarkers of RR-MS with good values of specificity and sensitivity. Indeed, we propose two particular circRNAs, circ_0005402 and circ_0003452_2, as new biomarkers in multiple sclerosis.

To sum up, we have studied, for the first time, the expression profile of more than 13.000 circRNAs in RR-MS patients. Our results report the implication of circRNAs in the disease and moreover, after validation, two circRNAs are proposed as biomarkers.

Materials and Methods

Blood sample collection

Whole blood (10 ml) was collected from a total of 45 patients with RR-MS, 10 CIS cases and 26 healthy donors in the Department of Neurology at Donostia University Hospital and these were distributed into four different cohorts.

It is known that gender affects the expression of the genes, and as the prevalence of multiple sclerosis is higher in women (around 65%), we decided to study only women in the discovery cohort and in the validation cohort, in order to avoid background noise coming from gender differences.

For the discovery cohort, we selected four RR-MS untreated patients in remission, while in the validation cohort 20 RR-MS untreated patients in remission were included. For each cohort four and 18 age matched healthy donors were selected respectively. Moreover, we included a second validation cohort of 21 RR-MS patients (twelve females and nine males). Two blood samples were drawn from those 21 patients in order to analyse the phenomena occurring during relapse: one during a relapse and another during a remission (10). A relapse was defined as an episode of new neurological symptoms of at least 24h of duration, not associated with fever or infection. Relapse blood samples were collected before giving any corticosteroid treatment. Finally, a third cohort of 10 patients in CIS and 10 age-matched controls was analysed (out of those 10 controls, 6 have been also included in the first validation cohort). CIS refers to a first episode of neurologic symptoms that may or may not go on to develop MS. In this case, all the patients in CIS selected for the study have been later diagnosed from RRMS in order to assess the predictive value of the circRNAs at this stage.

The main clinical and demographical characteristics of both patients and healthy donors are summarised in Table 2. Samples from all donors were collected after receiving written informed consent. The study was approved by the hospital's ethics committee and samples have been processed and stored at the Basque Biobank (http://www.biobancovasco.org; date last accessed June 27, 2017).

RNA isolation

Total RNAs, including small RNAs, was isolated from peripheral blood cells with the miRNeasy Mini Kit (Qiagen) following the manufacturer's instructions for the discovery, first validation and CIS cohorts. For the second validation cohort the LeukoLOCK kit (Ambion) was used for total RNA extraction, using the alternative protocol to capture small RNAs. RNA concentration was measured using a NanoDrop ND-1000 spectrophotometer.

Microarray analysis

RNA samples from four RR-MS patients and four controls were prepared and hybridised to the Arraystar Human Circular RNA Microarray V2.0, which covers 13,617 circRNA probes, based on Arraystar's standard protocols. Briefly, total RNAs were treated

with RNase R (Ribonuclease R) (Epicentre, Inc.) to remove linear RNAs and enrich circular RNAs. They were then amplified and transcribed into fluorescent cRNA (complementary RNA) utilising a random priming method (Arraystar Super RNA Labeling Kit; Arraystar). The labelled cRNAs were hybridised onto the array and after having washed the slides they were scanned by the Agilent Scanner G2505C (Agilent Technologies).

After quantile normalisation of the raw data, low intensity filtering was performed, and the circRNAs that had at least one out of four samples with flags in 'P' or 'M', meaning Present (fluorescence intensity above the background) and Marginal (Fluorescence intensity above the background, but background intensity was not uniform), were retained for further analyses. Using Bioconductor gplots package a hierarchical clustering was created with differentially expressed circRNAs calculating Euclidean genetic distances and a complete linkage algorithm for clustering.

qPCR validation

Technical validation of a selection of differentially expressed circRNAs was performed in the same samples as those studied in the microarray analysis (discovery cohort) by real-time quantitative PCR (qPCR) using Power SYBRGreen Master Mix (Applied Biosystems). Five upregulated and five downregulated circRNAs, those showing the highest fold change, were selected for technical validation, namely, hsa_circ_0000518, hsa_circ_0000517, hsa_circ_0000519, hsa_circ_0000520, hsa_circ_0001400, hsa_circ_0056731, hsa_circ_0064644, hsa_circ_0005402, hsa_circ_0024892 and hsa_circ_0035560 (CircBase [http://circbase.org/]; date last accessed June 27, 2017).

In addition, four of those circRNAs (two upregulated: hsa_circ_0000518 and hsa_circ_0001400 and two downregulated: hsa_circ_0056731 and hsa_circ_0005402) were chosen for a second technical validation performed by Arraystar (Arraystar, Inc., USA) in order to confirm the reliability of our primer design.

PCR amplicons were subjected to Sanger Sequencing (ABIprism 3130), and checked for the presence of predicted backspliced junctions in order to test their circularity.

For the lab technical validation total RNA was reverse transcribed into cDNA with random primers using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Inc., USA) according to the manufacturer's instructions. CircRNA sequences were obtained from UCSC (University of California, Santa Cruz) Genome Browser (GRCh37/hg19 assembly) and divergent primers were designed to amplify the circular transcripts so that the qPCR amplicon spans the backspliced junction (58). Commonly used convergent primers detecting a sequence downstream of the circRNA producing exons, were also designed in order to amplify the linear transcripts. Primer3Plus software (59) and BLAST (Basic Local Alignment Search Tool from the National Centre of Biotechnology Information, NCBI) were used to assist the primer design and to ensure the target specificity of the primers respectively. Primer sequences are available in Supplementary material, Table S1.

Then, the cDNA was amplified using a 7900HT thermal cycler (Applied Biosystems, Inc., USA) or a CFX96/CFX384 Touch Real-Time PCR Detection System (Bio-Rad laboratories, Inc.) following the manufacturer's instructions. Each sample was run in triplicate. The presence of a single-peak in the melting curve indicated the specificity of the amplification. The expression level of the circular transcript, represented as Fold change (FC) was calculated using the $2^{-\Delta Cq}$ method, where ΔCq was calculated for the divergent amplicon (Cq Multiple sclerosis group- Cq Control group). In addition, we further calculated the FC after normalisation of the circular RNA expression by the expression level of its corresponding linear form (convergent amplicon), where ΔΔCq= [(divCq – conCq) Multiple sclerosis group - (divCq – conCq) Control group].

For the technical validation performed by Arraystar, divergent primers spanning across the back-spliced junction were used to amplify the circRNAs and β-actin was selected as the reference gene.

In addition, four circRNAs for which the trend of the FC was in agreement with the microarray results in at least one technical validation (circ_0005402, circ_0056731, circ_0024892 and circ_0035560) were further selected for validation in a second cohort of 20 untreated RR-MS patients in remission and 18 controls (first validation cohort) by qPCR as described above.

Moreover, the expression of two of these circRNA candidates (circ_0005402 and circ_0035560) was studied by qPCR in 21 matched samples of patients in relapse and remission (second validation cohort) and in a third cohort of 10 patients with CIS and 10 age-matched controls.

Receiver operating characteristic (ROC) curve

MedCalc 14.8 was employed to plot the ROC curve using DeLong *et al*. methodology (60).

Statistical analysis

For microarray data analysis, differentially expressed circRNAs between the disease and control groups were identified using FC and its statistical significance was estimated by t-test using the R software limma package. CircRNAs having $FC \geq 1.5$ and $P\text{-values} \leq 0.05$ were selected as significantly differentially expressed.

In order to compare the expression of circRNAs between groups in the qPCR validations, a Wilcoxon test was used (paired in relapse vs remission) due to the fact that some of the circRNAs do not follow normal distribution. This was assessed using the Shapiro-Wilk test. Analyses were performed in Excel and R 3.3.2 in RStudio v1.0.44. $P\text{-values} < 0.05$ were considered significant, exact *P*-values are indicated during the results section. qPCR results, plots were created in RStudio v1.0.44 with R 3.3.2 and 'beeswarm' package v0.2.3.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We want to thank all the patients that have donated samples to the study. We also thank the editors of HEDAPEN for improving the use of English in the manuscript and the Basque Foundation for Health Research and Innovation for funding this editing work.

Conflict of Interest statement. None declared.

Funding

Spanish Network of Multiple Sclerosis, Regional Council of Gipuzkoa [grant number DFG15/006] and Department of

Industry of the Basque Country [grant number ELKARTEK16/014]. LI was supported by the Department of Education of the Basque Government [grant number PRE_2016_1_0168].

References

1. Amur, S., Parekh, A. and Mummaneni, P. (2012) Sex differences and genomics in autoimmune diseases. *J Autoimmun*, **38**, J254–J265.
2. Greer, J.M. and McCombe, P.A. Role of gender in multiple sclerosis: clinical effects and potential molecular mechanisms. *J. Neuroimmunol.*, **234**, 7–18.
3. Tullman, M.J. (2013) Overview of the epidemiology, diagnosis, and disease progression associated with multiple sclerosis. *Am. J. Manag. Care*, **19**, S15–S20.
4. Fawaz, C.N., Makki, I.S., Kazan, J.M., Gebara, N.Y., Andary, F.S., Itani, M.M., El-Sayyed, M., Zeidan, A., Quartarone, A., Darwish, H. *et al.* (2015) Neuroproteomics and microRNAs studies in multiple sclerosis: transforming research and clinical knowledge in biomarker research. *Expert Rev. Proteomics*, **12**, 1–14.
5. Ramien, C., Taenzer, A., Lupu, A., Heckmann, N., Engler, B., Patas, K., Friese, M.A. and Gold, S.M. (2015) Sex Effects on Inflammatory and Neurodegenerative Processes in Multiple Sclerosis. *Neurosci. Biobehav. Rev.*, 10.1016/j.neubiorev.2015.12.015.
6. Tremlett, H., Zhao, Y., Joseph, J. and Devonshire, V. (2008) Relapses in multiple sclerosis are age- and time-dependent. *J. Neurol. Neurosurg. Psychiatry*, **79**, 1368–1374.
7. Voskuhl, R.R. and Gold, S.M. (2012) Sex-related factors in multiple sclerosis susceptibility and progression. *Nat. Rev. Neurol.*, **8**, 255–263.
8. Menon, R., Di, D.M., Cordiglieri, C., Musio, S., La, M.L., Milanese, C., Di Stefano, A.L., Crabbio, M., Franciotta, D., Bergamaschi, R. *et al.* (2012) Gender-based blood transcriptomes and interactomes in multiple sclerosis: involvement of SP1 dependent gene transcription. *J.Autoimmun.*, **38**, J144–J155.
9. Achiron, A. and Gurevich, M. (2009) Gender effects in relapsing-remitting multiple sclerosis: correlation between clinical variables and gene expression molecular pathways. *J. Neurol. Sci.*, **286**, 47–53.
10. Irizar, H., Munoz-Culla, M., Sepulveda, L., Saenz-Cuesta, M., Prada, A., Castillo-Trivino, T., Zamora-Lopez, G., de Munain, A.L., Olascoaga, J. and Otaegui, D. (2014) Transcriptomic profile reveals gender-specific molecular mechanisms driving multiple sclerosis progression. *PLoS One*, **9**, e90482.
11. Muñoz-Culla, M., Irizar, H., Sáenz-Cuesta, M., Castillo-Triviño, T., Osorio-Querejeta, I., Sepúlveda, L., López de Munain, A., Olascoaga, J. and Otaegui, D. (2016) SncRNA (microRNA &snoRNA) opposite expression pattern found in multiple sclerosis relapse and remission is sex dependent. *Sci. Rep.*, **6**, 20126.
12. Comabella, M. and Montalban, X. (2014) Body fluid biomarkers in multiple sclerosis. *Lancet Neurol.*, **13**, 113–126.
13. D'Ambrosio, A., Pontecorvo, S., Colasanti, T., Zamboni, S., Francia, A. and Margutti, P. (2015) Peripheral blood biomarkers in multiple sclerosis. *Autoimmun. Rev.*, **14**, 1097–1110.
14. Esteller, M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.*, **12**, 861–874.
15. Eulalio, A., Huntzinger, E. and Izaurralde, E. (2008) Getting to the root of miRNA-mediated gene silencing. *Cell*, **132**, 9–14.
16. Muñoz-Culla, M., Irizar, H., Castillo-Triviño, T., Sáenz-Cuesta, M., Sepúlveda, L., Lopetegi, I., De Munain, A.L., Olascoaga, J., Baranzini, S.E. and Otaegui, D. (2014) Blood miRNA expression pattern is a possible risk marker for natalizumab-associated progressive multifocal leukoencephalopathy in multiple sclerosis patients. *Mult. Scler. J.*, **20**, 1851–1859.
17. Ceribelli, A., Satoh, M. and Chan, E.K. (2012) MicroRNAs and autoimmunity. *Curr.Opin.Immunol.*, **24**, 686–691.
18. Calin, G.A. and Croce, C.M. (2006) MicroRNA signatures in human cancers. *Nat.Rev.Cancer*, **6**, 857–866.
19. Weinberg, M.S. and Wood, M.J. (2009) Short non-coding RNA biology and neurodegenerative disorders: novel disease targets and therapeutics. *Hum.Mol.Genet.*, **18**, R27–R39.
20. Munoz-Culla, M., Irizar, H. and Otaegui, D. (2013) The genetics of multiple sclerosis: review of current and emerging candidates. *Appl. Clin. Genet.*, **6**, 63–73.
21. Jagot, F. and Davoust, N. (2016) Is it worth considering circulating microRNAs in multiple sclerosis? *Front. Immunol.*, **7**, 1–11.
22. Huang, Q., Xiao, B., Ma, X., Qu, M., Li, Y., Nagarkatti, P., Nagarkatti, M. and Zhou, J. (2016) MicroRNAs associated with the pathogenesis of multiple sclerosis. *J. Neuroimmunol.*, **295**, 148–161.
23. Regev, K., Healy, B.C., Khalid, F., Paul, A., Chu, R., Tauhid, S., Tummala, S., Diaz-Cruz, C., Raheja, R., Mazzola, M.A. *et al.* (2017) Association Between Serum MicroRNAs and Magnetic Resonance Imaging Measures of Multiple Sclerosis Severity. *JAMA Neurol.*, **2115**, 1–11.
24. Regev, K., Paul, A., Healy, B., von Glenn, F., Diaz-Cruz, C., Gholipour, T., Mazzola, M.A., Raheja, R., Nejad, P., Glanz, B.I. *et al.* (2016) Comprehensive evaluation of serum microRNAs as biomarkers in multiple sclerosis. *Neurol. Neuroimmunol. neuroinflammation*, **3**, e267.
25. Fenoglio, C., De Riz, M., Pietroboni, A.M., Calvi, A., Serpente, M., Cioffi, S.M.G., Arcaro, M., Oldoni, E., Scarpini, E. and Galimberti, D. (2016) Effect of fingolimod treatment on circulating miR-15b, miR23a and miR-223 levels in patients with multiple sclerosis. *J. Neuroimmunol.*, **299**, 81–83.
26. Quintana, E., Ortega, F.J., Robles-Cedeño, R., Villar, M.L., Buxó, M., Mercader, J.M., Alvarez-Cermeño, J.C., Pueyo, N., Perkal, H., Fernández-Real, J.M. *et al.* (2017) miRNAs in cerebrospinal fluid identify patients with MS and specifically those with lipid-specific oligoclonal IgM bands. *Mult. Scler. J.*, 10.1177/1352458516684213.
27. Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., Herzog, M. *et al.* (2015) Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell*, **58**, 870–885.
28. 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.
29. Guo, J.U., Agarwal, V., Guo, H. and Bartel, D.P. (2014) Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.*, **15**, 409.
30. 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.
31. 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.

32. Lasda, E. and Parker, R. (2014) Circular RNAs: diversity of form and function. *RNA*. **20**, 1829–1842.
33. Kulcheski, F.R., Christoff, A.P. and Margis, R. (2016) Circular RNAs are miRNA sponges and can be used as a new class of biomarker. *J. Biotechnol.*, **238**, 42–51.
34. Shang, X., Li, G., Liu, H., Li, T., Liu, J., Zhao, Q. and Wang, C. (2016) Comprehensive circular RNA profiling reveals that hsa_circ_0005075, a new circular RNA biomarker, is involved in hepatocellular carcinoma development. *Medicine (Baltimore)*, **95**, e3811.
35. Fan, X., Weng, X., Zhao, Y., Chen, W., Gan, T. and Xu, D. (2017) Circular RNAs in cardiovascular disease: an overview. *Biomed. Res. Int.*, **2017**, 5135781.
36. Xu, H., Guo, S., Li, W. and Yu, P. (2015) The circular RNA Cdr1as, via miR-7 and its targets, regulates insulin transcription and secretion in islet cells. *Sci. Rep.*, **5**, 12453.
37. Liu, Q., Zhang, X., Hu, X., Dai, L., Fu, X., Zhang, J. and Ao, Y. (2016) Circular RNA Related to the chondrocyte ECM regulates MMP13 expression by functioning as a MiR-136 'Sponge' in human cartilage degradation. *Sci. Rep.*, **6**, 22572.
38. Li, T.-R., Jia, Y.-J., Wang, Q., Shao, X.-Q. and Lv, R.-J. (2016) Circular RNA: a new star in neurological diseases. *Int. J. Neurosci.*, **7454**, 1–18.
39. Abu, N. and Jamal, R. (2016) Circular RNAs as promising biomarkers: A mini-review. *Front. Physiol.*, **7**, 1–6.
40. Lu, D. and Xu, A.D. (2016) Mini Review: Circular RNAs as potential clinical biomarkers for disorders in the central nervous system. *Front. Genet.*, **7**, 1–5.
41. Hwang, J.M., Chang, B.L. and Park, S.S. Leber's hereditary optic neuropathy mutations in Korean patients with multiple sclerosis. *Ophthalmologica*, **215**, 398–400.
42. Cardamone, G., Paraboschi, E.M., Rimoldi, V., Duga, S., Soldà, G. and Asselta, R. (2017) The characterization of GSDMB splicing and backsplicing profiles identifies novel isoforms and a circular RNA that are dysregulated in multiple sclerosis. *Int. J. Mol. Sci.*, **18**, 576.
43. Glažar, P., Papavasileiou, P. and Rajewsky, N. (2014) circBase: a database for circular RNAs. *RNA*, **20**, 1666–1670.
44. Yamamoto, T., Kudo, M., Peng, W.-X., Takata, H., Takakura, H., Teduka, K., Fujii, T., Mitamura, K., Taga, A., Uchida, E. *et al.* (2016) Identification of aldolase A as a potential diagnostic biomarker for colorectal cancer based on proteomic analysis using formalin-fixed paraffin-embedded tissue. *Tumor Biol.*, **37**, 13595–13606.
45. El-Abd, N., Fawzy, A., Elbaz, T. and Hamdy, S. (2016) Evaluation of annexin A2 and as potential biomarkers for hepatocellular carcinoma. *Tumor Biol.*, **37**, 211–216.
46. Kling, T., Ferrarese, R., Ó hAilín, D., Johansson, P., Heiland, D.H., Dai, F., Vasilikos, I., Weyerbrock, A., Jörnsten, R., Carro, M.S. *et al.* (2016) Integrative modeling reveals Annexin A2-mediated epigenetic control of mesenchymal glioblastoma. *EBioMedicine*, **12**, 72–85.
47. Cañas, F., Simonin, L., Couturaud, F. and Renaudineau, Y. (2015) Annexin A2 autoantibodies in thrombosis and autoimmune diseases. *Thromb. Res.*, **135**, 226–230.
48. Pianta, A., Drouin, E.E., Crowley, J.T., Arvikar, S., Strle, K., Costello, C.E. and Steere, A.C. (2015) Annexin A2 is a target of autoimmune T and B cell responses associated with synovial fibroblast proliferation in patients with antibiotic-refractory Lyme arthritis. *Clin. Immunol.*, **160**, 336–341.
49. Swisher, J.F.A., Khatri, U. and Feldman, G.M. (2007) Annexin A2 is a soluble mediator of macrophage activation. *J. Leukoc. Biol.*, **82**, 1174–1184.
50. Chao, P.-Z., Hsieh, M.-S., Cheng, C.-W., Hsu, T.-J., Lin, Y.-T., Lai, C.-H., Liao, C.-C., Chen, W.-Y., Leung, T.-K., Lee, F.-P. *et al.* (2015) Dendritic cells respond to nasopharygeal carcinoma cells through annexin A2-recognizing DC-SIGN. *Oncotarget*, **6**, 159–170.
51. Fang, W., Fa, Z.-Z., Xie, Q., Wang, G.-Z., Yi, J., Zhang, C., Meng, G.-X., Gu, J.-L. and Liao, W.-Q. (2017) Complex roles of Annexin A2 in host blood–brain barrier invasion by *Cryptococcus neoformans*. *CNS Neurosci. Ther.*, 10.1111/cns.12673.
52. Lopez-Ramirez, M.A., Wu, D., Pryce, G., Simpson, J.E., Reijerkerk, A., King-Robson, J., Kay, O., De Vries, H.E., Hirst, M.C., Sharrack, B. *et al.* (2014) MicroRNA-155 negatively affects blood-brain barrier function during neuroinflammation. *faseb J.*, **28**, 2551–2565.
53. Zhang, J., Cheng, Y., Cui, W., Li, M., Li, B. and Guo, L. (2014) MicroRNA-155 modulates Th1 and Th17 cell differentiation and is associated with multiple sclerosis and experimental autoimmune encephalomyelitis. *J. Neuroimmunol.*, **266**, 56–63.
54. Moore, C.S., Rao, V.T.S., Durafourt, B.A., Bedell, B.J., Ludwin, S.K., Bar-Or, A. and Antel, J.P. (2013) MiR-155 as a multiple sclerosis-relevant regulator of myeloid cell polarization. *Ann. Neurol.*, **74**, 709–720.
55. Hagiwara, K., Katsuda, T., Gailhouste, L., Kosaka, N. and Ochiya, T. (2015) Commitment of Annexin A2 in recruitment of microRNAs into extracellular vesicles. *FEBS Lett.*, **589**, 4071–4078.
56. Dudekula, D.B., Panda, A.C., Grammatikakis, I., De, S., Abdelmohsen, K. and Gorospe, M. (2016) CircInteractome: a web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biol.*, **13**, 34–42.
57. Chen, I., Chen, C.-Y. and Chuang, T.-J. (2015) Biogenesis, identification, and function of exonic circular RNAs. *Wiley Interdiscip. Rev. RNA*, **6**, 563–579.
58. Jeck, W.R. and Sharpless, N.E. (2014) Detecting and characterizing circular RNAs. *Nat. Biotechnol.*, **32**, 453–461.
59. Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A.M. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.*, **35**, 71–74.
60. DeLong, E.R., DeLong, D.M. and Clarke-Pearson, D.L. (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, **44**, 837–845.