

Design of a GFP reporter for splicing analysis in mammalian cells

Arthur T. Menezes

Universidade de Sao Paulo

Helder Y. Nagasse

Universidade de Sao Paulo

Patricia Pereira Coltri (✉ coltri@usp.br)

Universidade de São Paulo Instituto de Ciencias Biomédicas <https://orcid.org/0000-0001-8447-7516>

Research note

Keywords: pre-mRNA splicing, GFP, reporter system

Posted Date: April 8th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-20486/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective: The great majority of eukaryotic genes are formed by exons and introns. Pre-RNA transcripts are extensively co-transcriptionally processed, with the addition of a cap group at the 5' end, intron removal and exon ligation (splicing) followed by addition of a poly-A tail at 3' end. Splicing is performed by specialized macromolecular machinery named spliceosome, composed of five small ribonucleoprotein particles (snRNPs) and several proteins. The activity of this complex is highly accurate due to coordinated activity of its components. Altered splicing have already been related to the development of several diseases as amyotrophic lateral sclerosis and different types of cancer. Detailed understanding of splicing regulation in eukaryotic cells can be achieved using splicing reporter systems.

Results: We designed a new splicing reporter plasmid suitable for analysis in mammalian cells. Our reporter is based on splicing of the GFP sequence. The greatest advantages of this system are the ease of visualization of the splicing reaction, which needs only a fluorescence microscope, and the possibility of quantification of splicing efficiency using real-time PCR. The use of this system allows rapid and easy detection of the splicing reactions.

Introduction

Pre-mRNA splicing is an essential processing step in eukaryotes. Eukaryotic genes are formed by exons, which remain in the mature RNA sequence, and introns, which are removed during splicing. Splicing depends on the assembly of a complex machinery, the spliceosome, composed by 5 U-rich small nuclear RNAs (snRNAs) and associated proteins, forming the 5 U-snRNPs (small nuclear ribonucleoproteins: U1, U2, U4, U5 and U6). Besides these 5 U-snRNPs, the spliceosome is composed by more than 100 proteins, most of which are conserved from budding yeast to humans (1). During pre-RNA processing, the splice sites, which are conserved sequences within the intron and in the exon-intron borders, are recognized by spliceosome components and trigger the sequential assembly of this complex. U1 snRNP binds to the 5' splice site, located at the border of 5' exon and intron. Association of U2 snRNP to a branchpoint binding protein promotes its association to the branchpoint site, in the middle of the intron, leading to rearrangements and recruitment of the tri-snRNP particle (U5-U4/U6). After that, several RNA-protein rearrangements release U1 and U4 snRNPs and promote the association between U2 and U6 snRNAs, creating a catalytic core center (2). Two sequential trans-esterification reactions remove the intron and join the exons in a mature RNA. Besides the canonical splicing described above, most genes go through alternative splicing processes. In this case, different splice sites along exons or introns are recognized and result in an alternative transcript isoform (3, 4). The most common mechanism of alternative splicing is exon skipping, which results in the exclusion of an exon in the mature mRNA. Another mechanism is intron retention, in which mature mRNA maintains intronic sequences (5). In all cases, these transcripts will generate different proteins.

A major challenge in exploring splicing efficiency and catalysis is the difficulty to analyze it in a living cell. First, because different transcripts might have different splicing rates, especially due to differences in splice signal sequences (6). Second, splicing can be affected by the cellular environment, which includes the presence of splicing inhibitors and regulatory proteins (7). To analyze splicing efficiency, real-time RT-PCR can be the method of choice (8). It is simpler than methods that require autoradiography and cheaper than high-throughput methods, as for example RNA-seq, which spends high workflow for each experimental group tested (9). At the same time, the use of reporter plasmids based on fluorescent proteins, such as GFP, allows a simple and quick screening by fluorescence microscopy. To allow a precise characterization of splicing efficiency in cultured cells, we developed a reporter system based on GFP coding sequence. We also analyzed the effect of an intronic non-coding RNA (snR38A) on splicing efficiency. Our reporter transcript allows for visualization and quantification of splicing reactions in cultured cells by fluorescence microscopy and real-time RT-PCR.

Methods

Construction of GFP reporter: The GFP coding sequence was amplified by PCR using pEGFP vector (Invitrogen) as DNA template in 2 separate fragments, named “exon 1” and “exon 2”. “Exon 1” was amplified using primers PC 20 (5' ACGCTGGATCCATGGTGAGCAAGGGCGAGG 3') and PC 21 (5' GGAGTGAATTCACCTCGGCGCGGGTCTTG 3'), flanked by BamHI and EcoRI restriction sites. “Exon 2” was amplified using primers PC 28 (5' GAAACTGCGGCCGCCACAGGTGAAGTTCGAGGGCGACACCC 3') and PC 29 (5' GCCGCTTCTAGACTTGTACAGCTCGTCCATGCC 3'), flanked by NotI and XbaI restriction sites. The intron sequence was amplified from AdML precursor, using plasmid pHMS388 (10). Primers used contained EcoRI and NotI restriction sites (PC26: 5' GAAACTGAATTCACCTCCCTCTCAAAGCGGGC 3', and PC 27: 5' GAAACTGCGGCCGCAAAAAAAGGGACAGGGTCAG 3').

“Exon 1”, “exon 2” and the intron amplification products were separately cloned into pGEM-T easy vector (Promega), according to manufacturer’s instructions, and sequenced to confirm their integrity. After confirmation of sequence integrity, the fragments were sub-cloned into pcDNA 3.1(+) vector (Invitrogen), using BamHI and XbaI enzymes, resulting in the splicing reporter pGFP-spl. The pcDNA 3.1(+) vector contains ampicillin resistance gene and pUC19 origin of replication, suitable for maintenance in *E. coli*. It also contains CMV promoter, for high-level expression in mammalian cells, and geneticin resistance marker, allowing for transient selection in cell culture.

To insert snR38A sequence in the pGFP-spl intron, this sequence was firstly amplified from HeLa genomic DNA using primers PC 1: 5' ATAGCCTCGAGCACAAGCCTATGATGG 3' and PC 2: 5' ATGATAAGCTTAAGCCTCAGAATAGA 3', flanked by XhoI and HindIII restriction sites. The PCR product was then cloned in pGEM-T easy vector, according to manufacturer’s instructions. snR38A fragment was then sub-cloned into the pGEM-intron, and this intron carrying the snR38A was interchanged with the intron in pGFP-spl, to finally construct pGFP-spl-snR38A plasmid.

Cell culture and transfection: HEK-293T cells were cultivated in 100 mm plates in 10 mL of DMEM / F12 with 10% FBS; 3.2 g / L sodium bicarbonate (NaHCO_3); in a humidified, controlled atmosphere incubator (5% CO_2) at 37 °C. Transfections were performed when cells reached 70–80% confluence. 1.5 μg of each plasmid were transfected into HEK293T cells using Lipofectamine 2000 (Invitrogen), according to manufacturer's instructions. pEGFP was used as a positive control. Transfected cells were selected using up to 1.6 mg/ml of geneticin (G418, Amresco). After selection, cells were collected and subjected to cell extract preparation.

Cell analysis by fluorescence microscopy: Transfected HEK293-T cells were analyzed by fluorescence microscopy to check for GFP expression. 48 h after transfection cells were visualized and photographed in a fluorescence microscope (Axio Vert. A1, Zeiss), at a 50X magnification. Non-transfected cells were used as a negative control. The photos were analyzed with the ImageJ software (1.46r, 32-bit version; Windows), to quantify the fluorescence intensity. For each cell group, 3 photos of different areas of the plates were analyzed, selecting the total area of each photo.

RNA analysis and real-time RT-PCR: Cells were collected and washed in PBS buffer (3.2 mM Na_2HPO_4 , 0.5 mM KH_2PO_4 , 1.3 mM KCl, 135 mM NaCl, pH 7.4). Cell extracts were prepared using buffer A (KCl 10 mM, MgCl_2 1.5 mM, Tris-HCl 20 mM [pH 7.5], DTT 0.5 mM) followed by homogenization using Douncer. Total RNA was extracted from cellular extracts using Trizol reagent (Invitrogen), according to manufacturer's instructions. After Trizol extraction, samples were precipitated using 3 M sodium acetate pH 5.2 and 100% ethanol. After resuspension, samples were used for cDNA synthesis using the Superscript kit (Invitrogen) and random primers. 100 ng of these cDNAs were analyzed by real-time RT-PCR, using SYBR Green® reagent (Thermo) to verify splicing efficiency. Two pairs of primers were used: for the exon junction and exon 2 ("mature RNA") (primers PC 157: 5' GCGCCGAGGTGAAGTTC 3' and PC 156: 5' GATGCCCTTCAGCTCGATGC 3') and a pair annealing at both ends of the precursor ("total RNA") (primers PC 155: 5' GACGACGGCAACTACAAGAC 3' and PC 156: 5' GATGCCCTTCAGCTCGATGC 3'). The "mature RNA" amplicon was detected only when splicing reaction was successful. Importantly, splicing efficiency was measured as a ratio between mature mRNA and total RNA amplifications. Also, β -actin primers were used to normalize reactions (primers PC 168: 5' ACCTTCTACAATGAGCTGCG 3' and PC 169: 5' CCTGGATAGCAACGTACATGG 3').

Results And Discussion

AdML intron and GFP coding sequence were amplified with specific primers. GFP coding sequence was amplified in two separate exons (Fig. 1). The three fragments were first cloned into pGEM-T easy (Promega). After confirming sequence integrity by DNA sequencing, fragments were separated and ligated into pcDNA 3.1(+) vector. Correct fragment order was confirmed by restriction enzyme analysis, according to the plasmid map generated after GFP and intron cloning (Fig S1).

We then transfected the splicing reporter pGFP-spl into HEK-293T cells, as described in the Methods section, and proceeded to splicing analysis using fluorescence microscopy (Fig. 2). Fluorescence analysis showed untransfected cells did not show green fluorescence, as expected, and pEGFP showed

fluorescence due to EGFP expression. Transfection of our reporter plasmid, pGFP-spl, and also pGFP-spl-snR38A, resulted in fluorescence after 48 h. Quantification of fluorescence intensity showed pGFP-spl and the positive control pEGFP have similar fluorescence, indicating our reporter was successfully spliced and resulted in mature GFP protein (Fig. 2B). pGFP-spl-snR38A transfected cells showed a higher fluorescence intensity. This might indicate that the presence of a non-coding RNA inside the intron, such as snR38A, enhances splicing. Alternatively, it might indicate that the region in which we inserted snR38A facilitates intron identification by the spliceosome.

Formation of mature RNA was also confirmed by RT-PCR and by real-time RT-PCR (qPCR) (Fig. 3). In order to assess splicing efficiency, we amplified mature RNA and total RNA, with two different pairs of primers (as described on Methods). Importantly, "total RNA" includes both pre-mRNA and mRNA molecules, once primers are located on exon 1 and exon 2. RT-PCR revealed mature RNA was present in pGFP-spl and on the positive control, as expected. The same was observed for pGFP-spl-snR38A (Fig. 3A). Splicing efficiency was calculated after qPCR, by a ratio of cycle threshold numbers (Ct's) observed after mRNA and total RNA amplification. In the plot shown on Fig. 3, higher bars indicate lower splicing efficiency. As expected, positive control shows a ratio of 1, once it does not amplify the longer molecule of pre-RNA and has only mature GFP present. pGFP-spl showed a ratio of 1.4 of splicing efficiency, revealing mature RNA was successfully generated. Splicing of a precursor containing snR38A in the intron reduces splicing efficiency (rate around 1.6) but still retains splicing activity. At the same time, these samples still retain unspliced pre-RNA, which is expected considering pre-RNA splicing reactions dynamics (11) (Fig. 3B). Importantly, these results suggest our reporter is functional and performs splicing efficiently. The inclusion of a non-coding RNA in the intron reduced but did not abolish splicing.

Previous works measured canonical and alternative splicing using fluorescence-based plasmids. A system based on an interrupted fluorescent protein gene to analyze alternative splicing was also constructed (12). In this system, the sequence of mCherry, which is translated in a red fluorescent protein, is interrupted by an intron. Depending on the alternative splicing pattern, the sequence that interrupts mCherry gene can be removed or skipped. The removal of this interrupting sequence allows for mCherry gene reconstitution and RFP expression. In another work alternative splicing efficiency of vascular endothelial growing factor A (VEGF-A) was analyzed (13). This gene has two isoforms, one of which excludes part of exon 8. With the use of a plasmid reporter based on dsRED sequence, they were able to detect alternative splicing of this exon. Zheng (9) developed a fluorescent-based plasmid to detect alternative splicing repressors. In this system, the alternative splicing of a given sequence can be studied based on the GFP or RFP expression. The sequence inclusion or exclusion during the splicing process determines which fluorescent protein will be expressed, allowing for a rapid detection of alternative isoforms. Fluorescence-based plasmids have been used to study a diversity of splicing transcripts, with different applications. In order to analyze the effects of different transcripts, our reporter system provides a new tool to analyze splicing in cultured cells, allowing for a characterization of effects that stimulate or inhibit splicing in a straightforward manner. Also, our system consists of yet another approach to study the effect of introns containing regulatory sequences, such as the snR38A, on splicing efficiency. Some

other possible applications would be the study of different elements that affect splicing dynamics, such as splice site mutations, miRNA carrying introns, and spliceosome targeting drugs.

Limitations

Our work describes a new approach to visualize and quantify splicing reactions using cultured cells. Despite the ease of visualization using the fluorescence microscope, we consider extremely important that analyses are performed along with a negative control, to overcome possible background fluorescence intensities of different cells. In order to confirm different amounts of unspliced and spliced products, RT-PCR should be performed with previously validated primers.

Abbreviations

snRNP - small ribonucleoprotein particles

GFP – green fluorescent protein

PCR – polymerase chain reaction

pre-mRNA – precursor messenger RNA

cDNA – complementary DNA

snRNA - small nuclear RNAs

RT-PCR – reverse transcriptase – polymerase chain reaction

RNA-seq – RNA sequencing

snR38A – small nuclear intronic non-coding RNA 38

AdML – adenovirus major league

DMEM / F12 – Dulbecco's modified Eagle's medium/ factor F12

FBS – fetal bovine serum

NaHCO₃ - sodium bicarbonate

G418 - geneticin

PBS – phosphate buffered saline

KH₂PO₄ – monopotassium phosphate

KCl – potassium chloride

NaCl – sodium chloride

MgCl₂ – magnesium chloride

Tris-HCl – acid Tris buffer

DTT - dithiothreitol

mRNA – messenger RNA

Ct – cycle threshold

RFP – red fluorescent protein

VEGF-A - vascular endothelial growing factor A

miRNA - microRNA

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The dataset used on this study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grant 2017/06994-9 and Conselho Nacional de Pesquisas Científicas (CNPq) grant 474672/2013-1.

Author's contributions: ATM and HN conceived and performed experiments; wrote the manuscript. PPC conceived the study and wrote the final manuscript. All authors read and approved the final manuscript.

Acknowledgements: We are grateful to Gisela T. Ramos for excellent technical support.

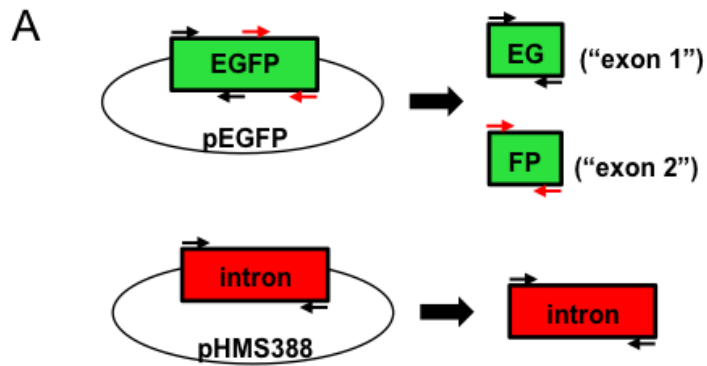
References

1. Hoskins AA, Moore MJ. The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem Sci.* 2012;37(5):179-88.
2. Zhang X, Yan C, Hang J, Finci LI, Lei J, Shi Y. An Atomic Structure of the Human Spliceosome. *Cell.* 2017;169(5):918-29 e14.

3. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol.* 2017;18(7):437-51.
4. Fu XD, Ares M, Jr. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat Rev Genet.* 2014;15(10):689-701.
5. Wang Y, Liu J, Huang BO, Xu YM, Li J, Huang LF, et al. Mechanism of alternative splicing and its regulation. *Biomed Rep.* 2015;3(2):152-8.
6. Alpert T, Herzog L, Neugebauer KM. Perfect timing: splicing and transcription rates in living cells. *Wiley interdisciplinary reviews RNA.* 2017;8(2).
7. Pai AA, Luca F. Environmental influences on RNA processing: Biochemical, molecular and genetic regulators of cellular response. *Wiley Interdiscip Rev RNA.* 2019;10(1):e1503.
8. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet.* 2018;59(3):253-68.
9. Zheng S. IRAS: High-Throughput Identification of Novel Alternative Splicing Regulators. *Methods in enzymology.* 2016;572:269-89.
10. Das R, Zhou Z, Reed R. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol Cell.* 2000;5(5):779-87.
11. Will CL, Luhrmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol.* 2011;3(7).
12. Schmidt HB, Barreau A, Rohatgi R. Phase separation-deficient TDP43 remains functional in splicing. *Nat Commun.* 2019;10(1):4890.
13. Stevens M, Star E, Lee M, Innes E, Li L, Bowler E, et al. The VEGF-A exon 8 splicing-sensitive fluorescent reporter mouse is a novel tool to assess the effects of splicing regulatory compounds in vivo. *RNA Biol.* 2019;16(12):1672-81.

Figures

Fig 1



B

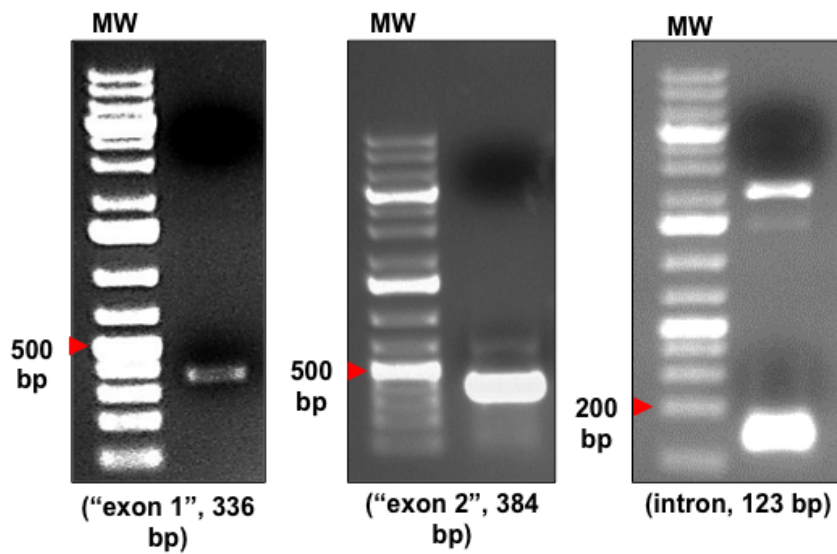
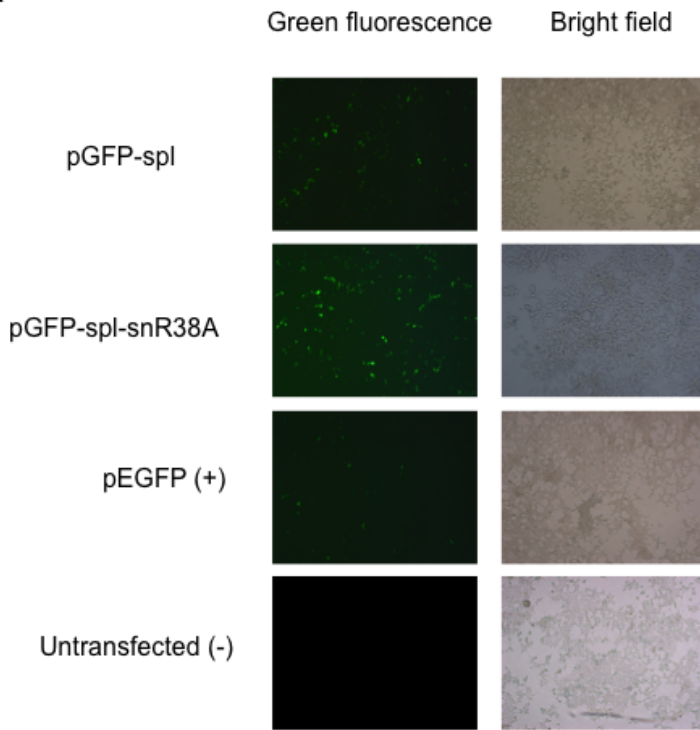


Figure 1

Amplification of GFP exons and AdML intron. (A) Schematic representation of EGFP amplification (upper part) and intron amplification (lower part), arrows show primers used. (B) Electrophoresis of PCR amplification products. Expected product lengths are shown in the bottom. MW, molecular weight ladder 1 kb (Thermo).

Fig 2

A



B

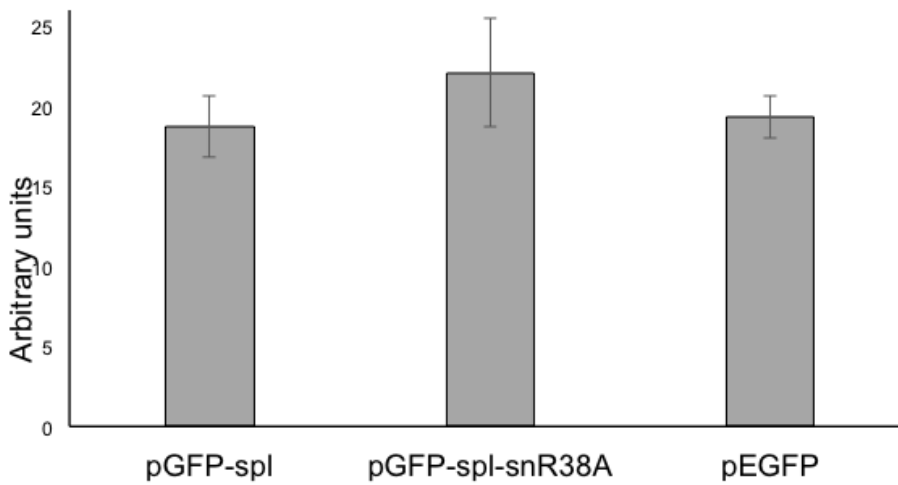


Figure 2

Fluorescence analysis of GFP expression after 48 hours of transfection on HEK293T cells. (A) Cells were transfected with pEGFP plasmid (Invitrogen) as a positive control. Untransfected cells were used as a negative control. Green fluorescence is shown on the left and bright field is on the right. Magnification 50X. (B) Fluorescence intensity quantification from 3 fields, standard deviation is shown. pEGFP, pGFP-

spl and pGFP-spl-snR38A intensity were compared. Y axis show arbitrary units after quantification on Image J software.

Fig 3

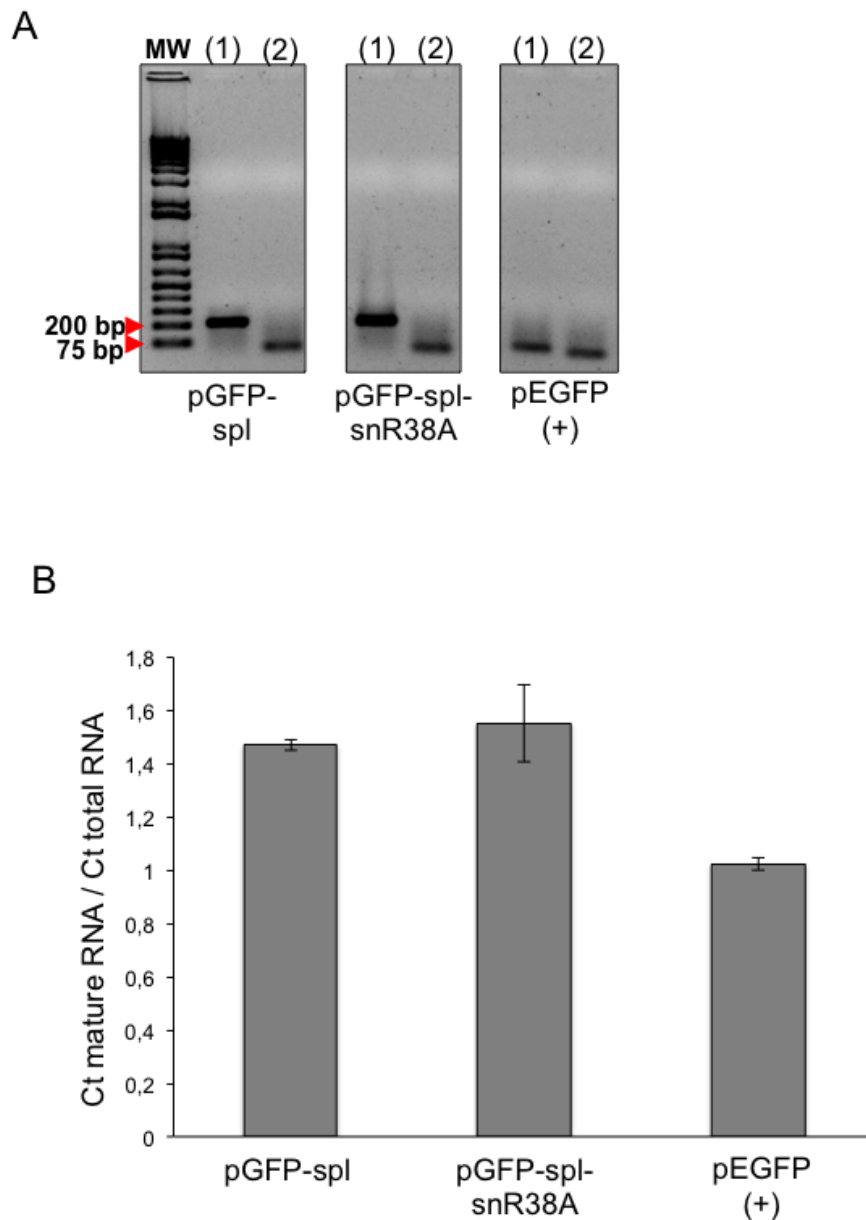


Figure 3

RT-PCR and real-time RT-PCR (qPCR) analysis. Cells were transfected with pGFP-spl, pGFP-spl-snR38A, and pEGFP (Invitrogen) as a positive control. Reactions were performed using primers to detect pre-RNA (1) and mRNA (2). (A) Electrophoresis of RT-PCR amplification products using pGFP-spl [expected lengths

(1): 204bp; (2): 59bp], pGFP-spl-snR38A [expected lengths (1): 276bp; (2): 59bp], and positive control [expected lengths (1): 81bp; (2): 59bp]. MW. Molecular weight ladder 1 kb (Thermo). (B) qPCR using samples of pGFP-spl, pGFP-spl-snR38A, and positive control, as indicated. The plot shows relative splicing efficiency of each sample by generating a ratio with the number of Ct's found for mature RNA and number of Ct's for total RNA (y-axis).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigS1legend.docx](#)
- [FigS1.tiff](#)