mRNA *trans*-splicing in gene therapy for genetic diseases

Adeline Berger,[1] Séverine Maire,[2,3] Marie-Claude Gaillard,[2,3] José-Alain Sahel,[1,4,5,6] Philippe Hantraye[2,3] and Alexis-Pierre Bemelmans[2,3]*

Spliceosome-mediated RNA *trans*-splicing, or SMaRT, is a promising strategy to design innovative gene therapy solutions for currently intractable genetic diseases. SMaRT relies on the correction of mutations at the post-transcriptional level by modifying the mRNA sequence. To achieve this, an exogenous RNA is introduced into the target cell, usually by means of gene transfer, to induce a splice event in *trans* between the exogenous RNA and the target endogenous pre-mRNA. This produces a chimeric mRNA composed partly of exons of the latter, and partly of exons of the former, encoding a sequence free of mutations. The principal challenge of SMaRT technology is to achieve a reaction as complete as possible, i.e., resulting in 100% repairing of the endogenous mRNA target. The proof of concept of SMaRT feasibility has already been established in several models of genetic diseases caused by recessive mutations. In such cases, in fact, the repair of only a portion of the mutant mRNA pool may be sufficient to obtain a significant therapeutic effect. However in the case of dominant mutations, the target cell must be freed from the majority of mutant mRNA copies, requiring a highly efficient *trans*-splicing reaction. This likely explains why only a few examples of SMaRT approaches targeting dominant mutations are reported in the literature. In this review, we explain in details the mechanism of *trans*-splicing, review the different strategies that are under evaluation to lead to efficient *trans*-splicing, and discuss the advantages and limitations of SMaRT. © 2016 The Authors. *WIREs RNA* published by Wiley Periodicals, Inc.

*Correspondence to: alexis.bemelmans@cea.fr

[1]Centre de recherche Institut de la Vision, Sorbonne Universités, Université Pierre et Marie Curie UM80, Paris, France

[2]Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Département des Sciences du Vivant (DSV), Institut d'Imagerie Biomédicale (I2BM), Molecular Imaging Research Center (MIRCen), Fontenay-aux-Roses, France

[3]Centre National de la Recherche Scientifique (CNRS), Université Paris-Sud, Université Paris-Saclay, UMR 9199, Neurodegenerative Diseases Laboratory, Fontenay-aux-Roses, France

[4]Centre Hospitalier National d'Ophtalmologie des Quinze-Vingts, INSERM-DHOS, Paris, France

[5]Fondation Ophtalmologique Adolphe de Rothschild, Paris, France

[6]Institute of Ophthalmology, University College of London, London, UK

INTRODUCTION

Following a fluctuating development during the 1990's and the first decade of the 21st century, gene therapy is now enjoying a renewed interest both in the pharmaceutical industry and academic laboratories engaged in translational research. We are now witnessing the emergence of new techniques in medicine that are a direct consequence of the progress in biotechnology and genetic engineering.[1] A large number of gene therapy modalities have been devised to address specific different pathologies. The techniques range from the simplest, such as gene supplementation, to the more complex, such as genome editing using the CRISPR/Cas9 technology.[2] However, there remain cases where existing tools are either not applicable, or not effective enough, to expect a significant

therapeutic effect. This is true for the example of recessive genetic diseases, in which the size of the cDNA that would correct the phenotype is not compatible with the currently available gene transfer vectors. In this case, one would need a technology allowing to introduce into the vector only the part of the mutant transcript that need to be corrected. We may also evoke dominant genetic diseases where a tight regulation of healthy gene expression is essential, disqualifying the only use of mRNA silencing, and causing a suppression of the expression of the mutant allele or both alleles. It would be interesting to repair the mutant transcripts to make clean copies without altering expression level. In addition, in case of multiple dominant mutations of the same gene, development of a specific therapeutic vector for each of them can be hardly feasible from a logistical perspective. In other scenarios, the appropriate strategy would be to express an exogenous factor, by taking advantage of the regulation of the expression of an endogenous gene. For example, it would be useful to specifically express a toxin in tumor cells, or in the case of neurodegenerative diseases, to express on demand a trophic factor during inflammatory episodes when up-regulation of inflammatory genes occurs in glial cells.

One sees that in all these cases, intervention by genetic engineering to modify a given transcript in the target cell population could be beneficial. This is what proposes the technology of spliceosome-mediated RNA *trans*-splicing by diverting the spliceosome machinery to modify part of the exons of an endogenous transcript, using an artificial RNA brought by gene transfer into the cells to be corrected. In this review, we will return to the theory and the principles of operation of *trans*-splicing, we will present the latest techniques that allow its implementation using selected examples in the literature, and finally, we will discuss its main limitations.

PRINCIPLE, ADVANTAGES, AND APPLICATIONS OF SPLICEOSOME-MEDIATED RNA-*TRANS*-SPLICING

Splicing, a reaction occurring in the nucleus of eukaryotic cells, is catalyzed by the spliceosome, one of the largest ribonucleoproteic complex of the cell, and results in the elimination of intronic sequences from pre-mRNA. Splicing requires several *cis*-acting elements on the pre-mRNA: (1) 5′ donor and 3′ acceptor splice site consensus sequences that constitute the exon–intron boundaries; (2) a branch point, consisting of an adenosine, located in a consensus sequence

of the intron, 18–40 nucleotides upstream of the 3′ acceptor splice site, and (3) a polypyrimidine-rich sequence (PPT for polypyrimidine tract), just before the 3′ splice site, often composed of an uracil repetition. Progressively as it is transcribed in the nucleus, the pre-mRNA is handled by the spliceosome machinery and spliced.[3] To summarize the process, presplicing complexes composed of ribonucleoproteins are sequentially assembled on the pre-mRNA to form the active final spliceosome, a complex of about 12 MDa. Then, a sequence of two *trans*-esterifications occurs. The first one, involving the 5′ donor site and the branch point, leads to the formation of a lariat and the release of a 3′-OH free extremity at the end of the upstream exon. The second one involves this 3′-OH free extremity and the 3′ acceptor site, leading to the joining of the two exons and the release of the lariat. This reaction, universally present in eukaryotes, is called splicing, or more precisely *cis*-splicing as it involves a donor site, a branch point, and an acceptor site located on the same RNA molecule (Figure 1(a)).

The spliceosome can also catalyze a *trans*-splicing reaction. Unlike the above-defined *cis*-splicing reaction, *trans*-splicing occurs between two different RNA molecules. The molecular process is exactly the same, except that the final mRNA is composed of the first exon(s) from one of the pre-mRNA and the following(s) exon(s) of the other one, creating a chimeric molecule. Although anecdotic when compared to *cis*-splicing, the occurrence of *trans*-splicing was initially described in the mid 80's using recombinant mRNAs in *in vitro* cell-free systems,[4,5] and it was quickly shown that this reaction was of physiological relevance among different lower eukaryotes, in particular the trypanosome.[6,7] Thereafter, evidence of *trans*-splicing has also been shown in mammals, with the demonstration that carnitine octanoyltransferase mRNAs are subject to *trans*-splicing in rat hepatocytes,[8] as are transcripts from the human estrogen receptor gene.[9] Moreover, recent refinements of RNAseq analysis helped to unambiguously identify physiologic *trans*-splicing events in human embryonic stem cells and demonstrated the role of a *trans*-spliced transcript in the maintenance of pluripotency.[10] It is thus established that *trans*-splicing exists naturally in a wide variety of organisms, but its physiological value remains to be elucidated. One possibility, suggested by Dixon et al.,[11] is that *trans*-splicing is at the origin of the exon-repetition phenomenon, which results in the tandem repeat of exons on a mRNA, in the absence of such repetition at the genomic level. In a 2007 study, these authors identified the existence of complementary
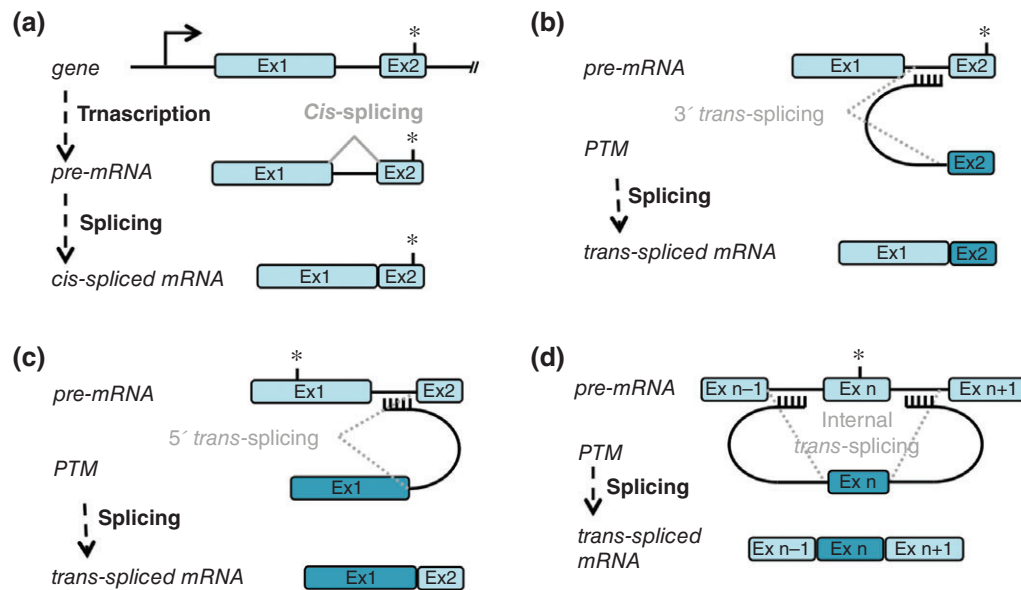
FIGURE 1 | Different modes of action of *trans*-splicing. (a) The classical splicing mechanism, or *cis*-splicing occurs within a single pre-mRNA molecule and leads to the elimination of the intronic sequences. (b, d) On the contrary, *trans*-splicing occurs between two splice sites located on two different pre-mRNAs. In SMaRT technology, an exogenous RNA, called pre-mRNA *trans*-splicing molecule (PTM) is used to replace one or several exons of an endogenous pre-mRNA. Depending of the orientation of the PTM, it is possible to replace 3′-exon(s) (b), 5′ exon(s) (c), or even internal exon(s) (d). Ex n: exon n; *: point mutation.

sequences in introns flanking the exons subjected to repetition, suggesting that exon-repetition occurs after *trans*-splicing between two pre-mRNA of the same gene, whose proximity is favored during transcription by the complementarity of intron sequences. This hypothesis is reinforced by the fact that exon-repetition operates strictly between transcripts encoded by the same allele,[12] and suggests that to promote *trans*-splicing between two pre-mRNA, it may be necessary to promote their closeness with complementary sequences allowing them to hybridize.

The discovery of these physiological mechanisms has prompted the use of *trans*-splicing for bioengineering purposes. Accordingly in 1999, using cell culture experiments, Puttaraju et al. demonstrated the feasibility of diverting *trans*-splicing for inducing the repair of an endogenous mRNA using exon exchange mediated by an artificial RNA capable of inducing *trans*-splicing.[13] Subsequently, they also showed the feasibility of using this strategy *in vivo*, leading to functional restoration of mutant cystic fibrosis trans-membrane conductance regulator (CFTR) in a human bronchial xenograft model system.[14] These studies paved the way for the use of spliceosome-mediated RNA *trans*-splicing, or SMaRT, as a gene therapy strategy. In this case, an artificial RNA, called pre-mRNA *trans*-splicing molecule (PTM), is engineered to specifically target an

endogenous pre-mRNA expressed in the target cell (Figure 1(b)–(d)). To be successful, the PTM must bind the target pre-mRNA and induce the *trans*-splicing reaction more efficiently than the *cis*-splicing one. For this purpose, a typical PTM is composed of: (1) a binding domain able to recognize the target intron on the endogenous pre-mRNA by base pairing, (2) an artificial intron to catalyze the splicing reaction, and (3) the cDNA containing the coding sequence of substitution (Figure 2). The precise order of these three elements in the PTM will depend on the location of the exon(s) to be replaced in the target mRNA, either in 5′ position or 3′ position (Figure 1 (b) and (c)). Replacement of an internal exon is also theoritically possible, although it relies on the occurrence of two *trans*-splicing events within the same pre-mRNA (Figure 1(d)). From its mode of action, the SMaRT technology presents several valuable benefits when considering gene therapy application:

- When targeting the first intron of the pre-mRNA and hence replacing all the subsequent sequence (depending of course of the cDNA size), only one PTM is necessary to repair numerous mutations. For allele-specific knock-down or silencing of dominant mutations, this unique PTM has a large benefit over CRISPR technology, ASO or shRNA molecules, which can target only one specific mutation at a time.
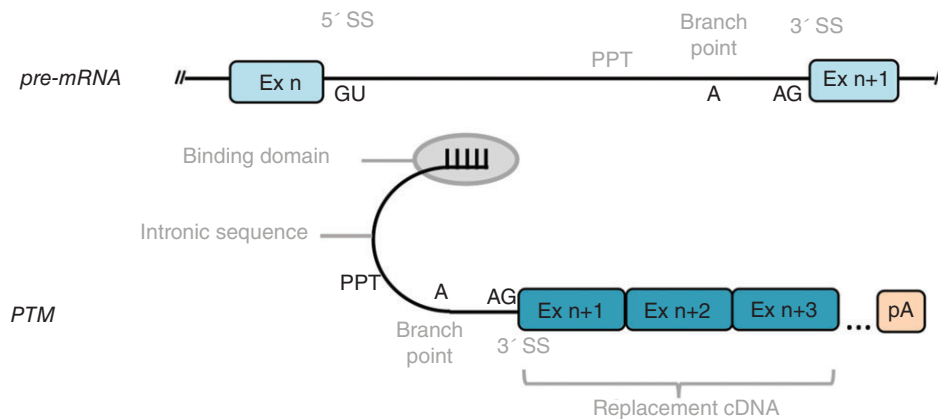
FIGURE 2 | Design of a PTM for 3′ *trans*-splicing. To achieve a 3′ replacement by *trans*-splicing, the PTM must be composed of, from 5′ to 3′: a binding domain, an artificial intron containing a polypyrimidine tract (PPT), a branch point, a 3′ acceptor site (3′ SS), and the replacement cDNA which contains the coding sequence of the exon(s) to be corrected. The necessary 5′ donor site (5′ SS) for the splice reaction is given by the endogenous pre-mRNA.

- The preservation in time, space and quantity of the endogenous regulation of the target mRNA expression. As the PTM is in theory inert except for the *trans*-splicing reaction, the regulation of expression of the chimeric RNA is thus totally dependent on the natural regulation of the targeted pre-mRNA. This information is crucial in gene therapy, where the expression of a gene has to be constrained to a strict spatiotemporal pattern, and kept at a physiologic level in the case of dominant diseases.

- The dual capacity of the PTM to reduce mutated protein synthesis and to simultaneously promote the normal one in a single reaction. In contrast to most of other approaches that require the delivery of two molecules, *trans*-splicing only requires the delivery of the engineered PTM, the two other components of the reaction (targeted pre-mRNA and spliceosome) being naturally present in the target cell.

- As the PTM includes only a part of the cDNA sequence needed for repair, SMaRT technology relies on a small-sized restorative molecule while other gene therapy systems require full-length cDNA. This facilitates the choice of a viral vector for PTM delivery in targeted cells. The size of transgenic cassette is indeed often a major limitation in the choice of a viral vector leading for example to the inability of using AAV vectors. SMaRT technology could therefore enlarge viral vector choice for diseases involving genes that are usually considered too large to be included in an AAV vector.

After an initial study leading to the proof of concept that SMaRT can efficiently promote a functional protein *in vitro*,[15] *trans*-splicing has been developed using different strategies, which can be roughly classified in two groups of applications, the first one being the repair of mutations at the mRNA level, and the second one, the expression of a therapeutic factor under the regulation of an endogenous gene. Although the latter case has not been extensively explored, we should mention here the study by Wang et al. that demonstrates the possibility to use partial *trans*-splicing of an abundant mRNA to express a therapeutic factor in the target cell.[16] However, the most common use of SMaRT concerns the RNA repair strategy for genetic diseases and is summarized in Table 1. In the case of a recessive disease, it is likely that correcting only part of the mRNA will result in therapeutic benefit, because one ends up in the scenario of a heterozygous individual who presents no phenotype. SMaRT technology has thus been evaluated for a significant number of recessive pathologies, such as cystic fibrosis,[14,19,30] hemophilia A,[26] dysferlinopathies and titinopathies,[27] X-linked immunodeficiency with hyper-IgM,[25] spinal muscular atrophy,[17,18] severe combined immune deficiency,[21] Duchenne muscular dystrophy,[22,31] and epidermolysis bullosa.[20,24,32] All these studies have shown that it is possible to partially correct the cellular pool of mutated mRNA. However, when compared with each other, they also show that there is no consensus on how to orientate PTM design to reach a significant *trans*-splicing efficiency, nor to lead to a therapeutic effect, or even the way to test the PTM efficiency. Still, a characteristic that does emerge is

TABLE 1 | Summary of the Main Publications Reporting the Implementation of a *Trans*-Splicing Strategy for Gene Therapy of Genetic Diseases

Disease	Target Gene	Mutation	Authors	Models Vitro/Vivo	Efficiency
Cystic fibrosis	CFTR	Recessive	Mansfield et al.[15], Liu et al.[11,14]	Vitro	12%[15][1] 16%[1][12] 14.2%[14][2]
Duchenne muscular dystrophy	MDX	Recessive	Lorain et al.[17,18]	Vitro + vivo	29.6%[18][1]
Dysferlinopathies/Titinopathies	DYSF/TTN	Dominant/ recessive	Monjaret et al.[19]	Vitro + vivo	NA
Dystrophia myotonica type 1	DMPK	Dominant	Chen et al.[20]	Vitro	7.4%[1]
Epidermolysis bullosa simplex	COL7A1, K14	Recessive	Wally et al.[21], Murauer et al.[22],	Vitro	15.5%[21][1]
Frontotemporal dementia with parkinsonsm linked to chromosome 17	MAPT	Recessive	Rodriguez-Martin et al.[23]	Vitro	34%[1]
Hemophilia A	FVIII	Recessive	Chao et al.[16]	Vitro	13%[2] 6.3%[1]
Huntington's disease	HTT	Dominant	Rindt et al.[24]	Vitro	5.4%[1]
Retinitis pigmentosa	RHO	Dominant	Berger et al.[24]	Vitro + vivo	40%
Severe combined immune deficiency	DNA-PKcs	Recessive	Zayed et al.[25]	Vitro	15%[2]
Spinal muscular atrophy	SMN2	Recessive	Coady et al.[26–28], Shababi et al.[29]	Vitro + vivo	NA
X-linked immunodeficiency with hyper-IgM	CD40L	Recessive	Tahara et al.[25]	Vivo	10.4%[1]

NA, data not available.
[1] *trans*-Splicing efficiency determined by quantification of corrected mRNA or protein and expressed as percentage of normal level (direct quantification approach).
[2] *trans*-Splicing efficiency determined by quantification of the activity of the converted protein and expressed as percentage of normal level (indirect quantification approach).

the partial nature of *trans*-splicing. In the best case, the reported efficacy is approximately 20–40%, meaning that only 20–40% of the mutant mRNA pool of the target cell is converted to wild-type mRNA. This is probably why there are far fewer studies reporting application of *trans*-splicing for dominant mutations. In this context, it is necessary to reach a significant percentage of efficiency to expect a therapeutic effect. To our knowledge, there are only four studies that reported the implementation of *trans*-splicing in the case of a dominant mutation (Table 1). In 2009, Chen et al. reported the correction of *DMPK* pre-mRNA, the gene responsible for *dystrophia myotonica* type 1, with an efficiency ranging from approximately 1.5 to 8%.[33] In 2012, Rindt et al. demonstrated in cellular models of Huntington's disease that the mutated exon can be eliminated by 5′ replacement from the Huntingtin encoding mRNA.[23] Recently, Monjaret et al. described *trans*-splicing of the Titin pre-mRNA, whose mutations lead to muscular dystrophies.[27] In this study, the *trans*-spliced product was not accurately quantified and appeared clearly minor. Thus, the *trans*-splicing rate achieved in these studies, by the order of only a few percent, will likely not lead to

a therapeutic threshold in patients affected by these dominant conditions to be reached. Our team has recently reported a rate of *trans*-splicing of about 40% in a cell-line expressing a dominant mutation of rhodopsin, the gene responsible for retinitis pigmentosa.[34] Although encouraging, this rate is probably still insufficient for the point mutations that cause the most aggressive cases of retinitis pigmentosa. This highlights the fact that, to achieve successful SMaRT applications for dominant mutations, it is necessary to implement rational means for optimizing design and characterization of an effective PTM.

DESIGN OF HIGHLY EFFICIENT PTM

The first step in designing powerful PTMs is the implementation of methods for accurate and reproducible quantification of *trans*-splicing efficiency. These methods can be divided into two main categories, i.e., direct and indirect approaches. The former consist of direct quantification of *cis* and/or *trans*-splicing products detected either at the mRNA or the protein level. Spliced mRNAs are usually quantified using end-point or quantitative RT-PCR

with specific primers and probes to discriminate *cis*- and *trans*-spliced products. This kind of method estimates the relative amount of repaired mRNAs establishing proof that *trans*-splicing occurs while assessing PTM activity.[14,24,26,30,32,35–39] An alternative approach of quantification at the mRNA level consists in the addition of a silent mutation in the replacement cDNA of the PTM in order to produce different restriction profiles between *cis* and *trans*-spliced mRNAs. Thus, amplification of *cis*- and *trans*-spliced products is achieved in a single RT-PCR reaction using the same primer set, and precise quantification of PTM activity is accomplished by restriction analysis of the amplicon.[23,34] The essential advantage of this approach is to avoid the bias of comparing two different PCR amplifications, one for the *cis*-spliced product, and one for the *trans*-spliced, reflecting respective level of *cis*- and *trans*-spliced events. For *trans*-splicing assessment at the protein level, western-blot analysis is the most frequently used technique, and quantification is performed by measuring level of normal protein expression alone or compared with mutated proteins with size discrimination parameters,[24,30] or using specific antibodies.[14,37] Another direct way to quantify *trans*-splicing is to use an artificial reporter system. In this case, the target intron is introduced into the open reading frame of a reporter gene, which can encode a functional reporter protein only after *trans*-splicing. The first systems described were based on the use of the LacZ gene and measuring the ß-galactosidase activity generated by *trans*-splicing.[28,35] More recently, the team of J. Bauer described a reporter system based on fluorescent proteins, which allows the efficient pre-screening of the PTM target on the intron to be *trans*-spliced.[29] *Trans*-splicing efficiency may also be measured indirectly. For instance, as *trans*-splicing gene therapy aims to provide phenotype improvements, PTM activity may be assessed by measuring therapeutic indexes in *in vitro* or *in vivo* models. Several methods have been used for that purpose, such as immunofluorescence assay enabling for example the assessment of normal protein localization[34] or expression[24], cell death measurement in the case of anti-tumor *trans*-splicing,[39,40] and functional assessment of ion channel by electrophysiology for CFTR repair.[14,19]

Effective *trans*-splicing requires to favor *trans*-splicing over *cis*-splicing. That can be achieved by a fine-tuning of the PTM design. In their initial study, Puttaraju et al. demonstrated that *trans*-splicing requires the presence of several motifs in the PTM molecule, such as a functional branch point, a PPT and the AG consensus motif as splice acceptor site at the 3′ end for a 3′ replacement strategy (Figure 2). Other factors are potentially important for triggering an efficient *trans*-splicing reaction, their influence on the *trans*-splicing rate was therefore investigated by different teams. Firstly, the localization of the PTM binding domain on the target intron plays a significant role in this challenge and could be determinant when targeting sequences involved in splicing. For instance, Murauer et al. tried to block *cis*-splicing by designing a binding domain complementary to the acceptor site located just downstream the target intron.[29] They described that blocking the endogenous 3′ splice site led to an increase of *trans*-splicing efficiency, although this is not a general rule.[34] Secondly, the choice of the target intron is likely to modulate the rate of *trans*-splicing event and in particular the potency of 5′ and 3′ splice site. Exon and intron definition models, which depict the recognition, binding and stability pattern of spliceosome first step assembly, might be of importance.[41,42] It is known that eukaryotic genes are characterized by short exons and long introns, preventing recognition across introns in favor of the exons. Accordingly, Philippi et al. tested whether the strength of 3′ splice sites is a parameter to take into account to favor the *trans*-splicing reaction.[38] This strength can be estimated by calculation of the MaxEnt (maximum entropy) score, which assigns a score to the splicing sites based on their homology to the canonical sequence.[43] 5′ and 3′ splice sites have consensus motifs, which in theory modulate splicing efficiencies based on their capacity to bind splicing factors. Thus, the splice site strength is highly relevant in the choice of the target intron, as it represents the splicing dynamic in terms of splicing factor interaction. Strong splice sites associated with high MaxEnt scores should promote splicing more efficiently than weak splice sites associated with low MaxEnt scores. Applied to *trans*-splicing, this assumption underlies that weak 3′ splice site in the target intron tends to enhance *trans*-splicing events since this splice site is in competition with a stronger PTM splice site.[38] Third, Coady et al.[44] and Shababi et al.[45] have demonstrated that the combination of PTM with antisense oligonucleotides (ASO) capable of blocking the splice-site targeted by the PTM increases the *trans*-splicing rate, most likely by inhibition of *cis*-splicing by the ASO. Coady et al. developed a single vector for ASO and PTM delivery and showed its efficiency *in vivo* in a mouse model of spinal muscular atrophy.[17] Similarly, Wang et al. suggested that splicing efficiency is correlated to the time of intron synthesis.[46] This mechanism is certainly due to the competition between *cis*-splicing, which requires the synthesis of both endogenous

splice sites, and *trans*-splicing which only needs the endogenous 5′ splice site. Fourthly, another way to increase *trans*-splicing efficiency is to insert an intron in the replacement cDNA of the PTM.[22,28,34] In this case, the addition of intronic splicing enhancer (ISE) in the PTM sequence may promote *trans*-splicing events. It is currently difficult to predict the effect of an ISE according to its sequence, its location related to which exon to include, and its potential to recruit splicing enhancer factors. However, new methods are currently being developed to better locate ISEs as well as splicing repressor sequences and should be useful to increase PTM efficiency.[47,48] Splicing is highly dependent on RNA/protein interactions which promote or prevent pre-mRNA availability to the spliceosome.[49] This is true also for PTMs which have to reach the target intron despite complex pre-mRNA conformations resulting from RNA/protein interactions. These interactions are a clue for PTM optimization and better knowledge of this aspect is needed to understand its real impact on *trans*-splicing. Finally, to maximize *trans*-splicing efficiency, the delivery mode of the PTM has to be taken into account. We have shown that this efficiency is directly correlated to the amount of PTM molecules available in the target cell.[34] For *in vivo* applications, it will thus be important to use the most efficient gene transfer vector to transduce target cells.

To summarize, several aspects of the PTM's design, target, environment and delivery can be modulated and are constantly improved leading to better understanding. However, an obvious clue in PTM design is that a great part of the success remains in the binding domain. Most of the groups working on *trans*-splicing have tried to modify the size and the location of this sequence with direct consequences on *trans*-splicing efficiency and specificity.[15,20,28,34,50] The presence of *trans*-splicing hotspots in the target intron has been demonstrated. In these hotspots, a shift on the target intron of about 30 base pairs (bp) applied to a 150-bp long binding domain can lead to significant variability in PTM efficiency.[34] Nevertheless, predicting an effective binding domain is hard to achieve. Even if bioinformatics analyses based on secondary structures of PTMs and their targets are used to perform a splice site availability prediction, the process still needs to be improved, as some aspects of splicing sequence remain unclear.

Therefore, screening strategies are needed to select effective binding domains. To date, very few binding domains have been tested for each PTM reported in the literature, because the techniques of PTM synthesis and efficiency assessment are cumbersome, thus limiting the discovery of highly efficient PTMs. For this reason, a larger PTM screening and efficient reporter system suitable for automated analysis needs to be implemented. The team of J. Bauer has described a PTM screening system in which *trans*-splicing of an artificial minigene generates a fluorescent protein, thus enabling an easy and fast detection of *trans*-splicing events.[29,51] To achieve this, the target introns are introduced into a minigene, suitable for testing every kind of *trans*-splicing orientation. This system consists of two elements. The first is a reporter minigene including the target intron flanked by two exons, one of which is the partial ORF of a fluorescent protein, placed upstream or downstream of the intron depending on whether one wishes to promote 5′ or 3′ *trans*-splicing. The second one is a library of PTMs that differ only in the binding domain to the target intron and whose cDNA can reconstruct the complete ORF of the fluorescent protein, allowing the detection of *trans*-splicing. For generating a high diversity of binding domain, it is possible to proceed by fragmentation of the target intron using one or more high-frequency cutting restriction enzymes. In this case, however, the fragments are not generated in a totally random manner, and parts of the intron, such as GC-rich or repetitions containing regions can escape fragmentation. Alternatively, the use of physical methods may allow higher fragment diversity. Thanks to the rise of next generation sequencing, sonication-based DNA fragmentation methods that can generate reproducibly fragments of specific lengths are now available. For generating the library of PTMs once the target intron has been fragmented, it is cloned into a vector containing the unvarying part of the PTM. This library can then be screened by cotransfection with the minigene that generates a fluorescent protein following *trans*-splicing. Although attractive, this screening strategy requires changing the genomic environment: the target intron is no longer surrounded by endogenous exons but by artificial ones and this change may affect the PTM activity. Thus, in a recent experiment we aimed to assess the reliability of this fluorescent reporter system to isolate efficient binding domains regardless of the genomic context of the target intron. We constructed a mini-gene encoding fluorescent proteins and containing the first intron of the human rhodopsin gene. Using this system, we showed that binding domains previously characterized on the full-length human rhodopsin gene induce the same rate of *trans*-splicing on the mini-gene (Figure 3), demonstrating that the genomic environment of the target intron has no major influence on the binding domain efficiency.
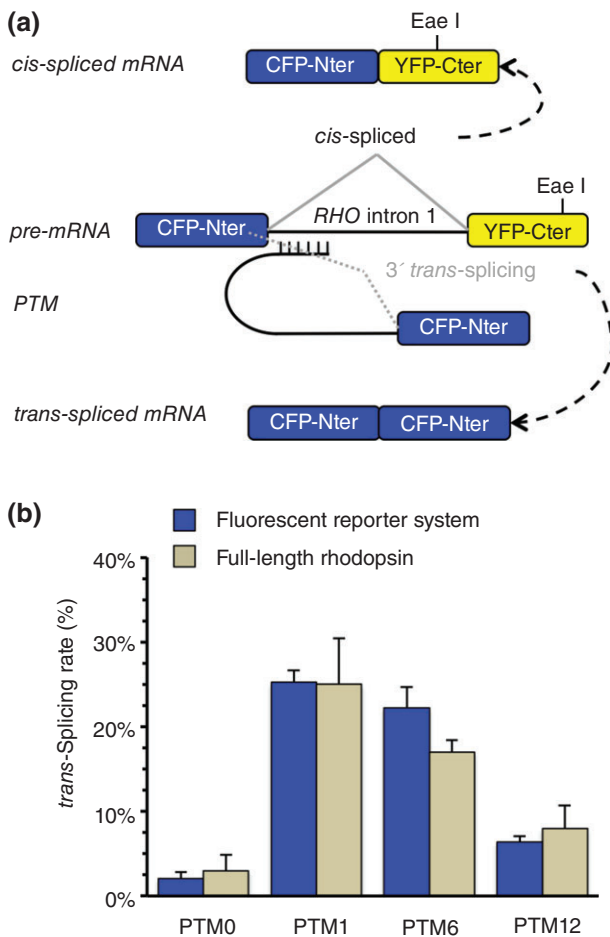
FIGURE 3 | Assessment of a fluorescent reporter system for PTM screening. (a) A fluorescent system was adapted to intron 1 of the rhodopsin gene (RHO), and the EaeI restriction site allowed quantifying *trans*-splicing using the same methodology described in Berger et al.[34] To that aim, RHO intron 1 was subcloned between sequences coding respectively the N-terminal part of CFP (CFP-Nter) and the C-terminal part of YFP (YFP-Cter), while the PTM contained the C-terminal part of CFP (CFP-Cter). With this system, *cis*-splicing leads to an mRNA encoding a chimeric protein fluorescing minimally in green, and *trans*-splicing leads to full-length CFP, characterized by a bright blue fluorescence. (b) Three binding domains, previously characterized in the context of the full rhodopsin pre-mRNA (for details, see Berger et al.[34]), were tested with the fluorescent reporter system using the end-point PCR amplification followed by enzymatic restriction method, and led to the same *trans*-splicing rate, thus showing that *trans*-splicing efficiency is independent of the genomic context of the binding domain. PTM0: negative control without binding domain. Location of binding domains of PTM1, 6, and 12 on RHO intron 1 have been described previously.[32] Bars represent mean ± SD; n = 6 for fluorescent reporter system; for full-length rhodopsin, n = 9 for PTM0 and 1, and n = 3 for PTM6 and 12.

DRAWBACKS AND LIMITATIONS OF THE SMART TECHNOLOGY

One of the main limitations in the development of SMaRT technology is linked to the lack of information regarding the mechanism involved in *trans*-splicing. While the *cis*-splicing mechanism is now largely understood and documented in the literature, *trans*-splicing as a natural event in human has raised only a few publications and most of the iceberg remains subsurface.[9,10,52,53] This lack of data is undoubtedly due to the difficulty to isolate *trans*-splicing products compared to those of *cis*-splicing, chromosomal rearrangements or technical artifacts. Because the *trans*-splicing reaction mechanism still presents some black holes, it is currently difficult or impossible to predict the efficiency of a given PTM. Notably the importance of the 3D-structure of both target and PTM RNA are to date insufficiently taken into consideration. Despite several tools developed to predict the secondary structures of pre-mRNA (ViennaRNA Package[a]; RNAfold[b]; mfold[c]...), it remains quite difficult to estimate the final shape of RNA and thus to design binding sequences based on this information. Interestingly, a recent report suggests that it is also important to consider the sequence of the target intron, and particularly the homology of the splice sites to the consensus sequences.[38] Based on this characteristic, the authors showed that it is possible to rank the different introns of a target pre-mRNA to determine which are the most likely to lead to an efficient *trans*-splicing reaction. To progress on this purpose, we can count on progress of bioinformatic tools in terms of 3D structure prediction, and, as discussed above, PTM screening using randomly-generated libraries. In addition to this specific limitation, like all molecular mechanisms used to modify gene expression, SMaRT technology still presents some demonstrated or theoretical drawbacks, which are covered below.

The main drawback of SMaRT technology is common to most nucleic acids-targeting engineering tools: the specificity of the molecule. One cannot imagine the PTM, generated to replace a mutated sequence, as a source of production of aberrant proteins due to off-target *trans*-splicing with random pre-mRNAs. Although these potential hybrid mRNAs should be processed by nonsense mediated decay or nonstop decay,[54] it is essential to validate the specificity of the PTM for the target sequence and limit nonspecific events. Several approaches should be considered for this purpose. (1) One must

determine the right size of the binding site that might limit off-target binding. Puttaraju et al. showed that increasing the length of the binding domain sequence up to 153 bases significantly enhanced the efficiency of the PTM.[15] A binding sequence of approximately 150 bases dramatically decreases the probability of finding the entire and exact corresponding sequence in a human genome, thus ensuring in theory a highly specific homology of the designed sequence. However, long binding sequences may only partially bind to nonspecific targets, decreasing the specificity of the molecule. The design of the PTM has to take in consideration these parameters and find a balance to obtain a good *trans*-splicing efficiency with a high specificity. (2) Take into account the importance of the 3′ splice site, which can induce more or less off-target *trans*-splicing depending on its potency.[50] (3) Increase the specificity of the PTM delivery to the targeted cells. By using an AAV vector for example, the adjustment in the capsid choice is crucial to introduce the PTM in the right cells, thus limiting potential nonspecific *trans*-splicing in cells that do not express the target pre-mRNA. (4) Similarly, take into consideration the expression level of the targeted mRNA, and adapt the PTM expression appropriately. We have shown that this efficiency is directly correlated to the amount of PTM molecules available in the target cell,[34] but an overproduction of PTM would be useless, and could potentially increase the off-target effect, or even be potentially toxic for the cell as can be overexpression of exogenous small RNA.[55,56]

The issue of binding domain specificity also raises the following question: what happens if the PTM does not bind to the specific sequence? Unfortunately, the production of truncated proteins due to the translation of PTM designed for 3′-replacement has already been observed in the absence of *trans*-splicing events.[15,27] Although 3′-replacement PTMs do not possess the first exon(s) of the endogenous mRNA, they are transcribed in the cell by a RNA polymerase II, and thus will be naturally capped and stabilized.[57] They are therefore likely to possess all the necessary elements for promoting translation, i.e., a 5′-cap, an alternative AUG start codons and a poly-adenylation sequence. Moreover, while the great majority of transcripts are translated from the first AUG of the KOZAK sequence, a secondary AUG can also be used to initiate translation. Although to our knowledge, there are no data available in the literature, this problem may be avoided in 5′-replacement PTMs for which the poly-adenylation signal is missing. In this case, this signal is given by the target pre-mRNA and is therefore excluded from the PTM design.

The last main drawback of SMaRT technology is inevitably linked to the efficiency of the reaction. In the context of an aggressive dominant disease, the level of mutated protein synthetized is determinant for the evolution of the disease. Even a low level of expression of a protein which leads to a dominant-negative effect will induce the disease. However, to date, nobody has been able to reach a 100% *trans*-splicing efficiency. Besides playing with the design of the PTM, the reaction efficiency could be ameliorated by increasing the likelihood of the target pre-mRNA coming into contact with the PTM, promoting more *trans*- than *cis*-splicing events. One major obstacle to achieve this goal is the coupling of mRNA transcription and splicing. Although this coupling is not required for the splice event to take place, the recruitment of elements making up the splicing machinery by the carboxy-terminal domain of the RNA polymerase II often makes the two processes simultaneous. Because it appears difficult to play on the transcription mechanism, the only way to intervene before *cis*-splicing is by accelerating the speed of action of the PTM, playing on its availability and its proximity with the target pre-mRNA. This availability will be defined by the quantity of PTM. As previously discussed, it is necessary to find a balance between providing enough PTM molecules, and avoiding a potential toxicity. Regarding the proximity of the two molecules, the reaction involves a closed localization of the PTM to the transcription site of the target, to bind and *trans*-splice the pre-mRNA before *cis*-splicing occurs. For now the PTM is simply delivered by conventional gene transfer means, either transfection for *in vitro* studies, or viral-mediated gene transfer *in vivo*, and lost in the huge pool of nucleic acids of the nucleus of the target cell. Increasing the likelihood of PTM and target pre-mRNA meeting could thus dramatically change the reaction efficiency. To that end, we could imagine the future development of a viral vector that allows targeted integration in the chromatin of the target cell (recently reviewed by Ott de Bruin et al.[58]). This development would be useful to introduce the PTM sequence near its target gene, or even in an intron of the target gene, ensuring proximity between the two RNA during and/or after transcription.

CONCLUSION

In the vast arena of human genetic diseases, advances in screening and sequencing technologies has allowed the identification of numerous genes which, when mutated, induce a specific condition. In both

recessive and dominant diseases, gene therapy is one of the most promising treatments for now, by cDNA supplementation in the first case and selective inhibition of the mutated form in the second one. In both cases, *trans*-splicing can bring a valuable alternative when technological barriers prohibit the use of more conventional approaches. In the context of cDNA supplementation for recessive conditions, the main limitation is the size of the expression cassette that needs to be introduced in the appropriate gene transfer vector. *Trans*-splicing overcomes this limit by allowing only a portion of the cDNA to be included in the vector: the PTM can target any of the intronic sequences of a given gene, thereby permitting the adaptation of the cDNA size to the capacity of the vector. In case of a very long pre-mRNA with a mutation in the middle, the internal *trans*-splicing approach can even be considered. Finally, this advantage facilitates the choice of the gene transfer vector, which can thus be dictated only by its tropism for the target cell. In the case of a dominant disease, despite the impressive progress made in the development of molecular tools in this field, including RNAi, nucleases and CRISPR systems, no technology is currently able to provide a universal solution, with broad application, high efficiency and no side-effect. The SMaRT technology could bring the answer to many if not all of these obstacles. Multiple advantages of SMaRT technology, that are unique to it and make it a major competitor for the other molecular tools, include: the requirement of only one exogenous RNA (the other components, i.e., the spliceosome, being provided by the cell), the possibility to use the same PTM for multiple mutations in the same gene, the reduction/inhibition of the mutated protein synthesis, the promotion of the normal protein synthesis at the same time, and finally the preservation of the endogenous gene expression regulation. Nevertheless, *trans*-splicing has not been widely developed so far, and still needs to be explored and better understood, allowing the development of more powerful PTMs. One of the main lacks concerns the prediction of the 3D structures of both target pre-mRNA and PTM, which are most likely determinant for *trans*-splicing efficiency. We and others have already shown that any change in the sequence of the PTM, not only in the binding domain but also in the cDNA replacement sequence, can dramatically modify this reaction efficiency. Despite this common observation, there are no clear and strict rules to define how to design the perfect PTM. All recommendations are based on each personal experience, and the best way to select an efficient PTM is for now to screen as many molecules as possible. In addition to complete all unexplored and unknown *trans*-splicing natural events that occur in all species including humans, a detailed and profound study on *trans*-splicing mechanism should enlighten us on the design of the PTM. This idea opens a door regarding the potential of *trans*-splicing for gene therapy. The knowledge to define the rules to achieve 100% reaction efficiency, while ensuring the absence of off-target effects, would make of the SMaRT technology an incredibly effective and powerful gene therapy tool.

NOTES

[a] http://www.tbi.univie.ac.at/RNA/

[b] http://nhjy.hzau.edu.cn/kech/swxxx/jakj/dianzi/Bioinf4/miRNA/miRNA1.htm

[c] http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form

REFERENCES

1. Collins M, Thrasher A. Gene therapy: progress and predictions. *Proc Biol Sci* 2015, 282:20143003.

2. Riordan SM, Heruth DP, Zhang LQ, Ye SQ. Application of CRISPR/Cas9 for biomedical discoveries. *Cell Biosci* 2015, 5:33.

3. Matera AG, Wang Z. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* 2014, 15:108–121.

4. Konarska MM, Padgett RA, Sharp PA. Trans splicing of mRNA precursors in vitro. *Cell* 1985, 42:165–171.

5. Solnick D. Trans splicing of mRNA precursors. *Cell* 1985, 42:157–164.

6. Murphy WJ, Watkins KP, Agabian N. Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for trans splicing. *Cell* 1986, 47:517–525.

7. Sutton RE, Boothroyd JC. Evidence for trans splicing in trypanosomes. *Cell* 1986, 47:527–535.

8. Caudevilla C, Serra D, Miliar A, Codony C, Asins G, Bach M, Hegardt FG. Natural trans-splicing in carnitine octanoyltransferase pre-mRNAs in rat liver. *Proc Natl Acad Sci USA* 1998, 95:12185–12190.

9. Flouriot G, Brand H, Seraphin B, Gannon F. Natural trans-spliced mRNAs are generated from the human estrogen receptor-alpha (hER alpha) gene. *J Biol Chem* 2002, 277:26244–26251.

10. Wu CS, Yu CY, Chuang CY, Hsiao M, Kao CF, Kuo HC, Chuang TJ. Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency. *Genome Res* 2014, 24:25–36.

11. Dixon RJ, Eperon IC, Samani NJ. Complementary intron sequence motifs associated with human exon repetition: a role for intragenic, inter-transcript interactions in gene expression. *Bioinformatics* 2007, 23:150–155.

12. Rigatti R, Jia JH, Samani NJ, Eperon IC. Exon repetition: a major pathway for processing mRNA of some genes is allele-specific. *Nucleic Acids Res* 2004, 32:441–446.

13. Puttaraju M, Jamison SF, Mansfield SG, Garcia-Blanco MA, Mitchell LG. Spliceosome-mediated RNA trans-splicing as a tool for gene therapy. *Nat Biotechnol* 1999, 17:246–252.

14. Liu X, Jiang Q, Mansfield SG, Puttaraju M, Zhang Y, Zhou W, Cohn JA, Garcia-Blanco MA, Mitchell LG, Engelhardt JF. Partial correction of endogenous DeltaF508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA trans-splicing. *Nat Biotechnol* 2002, 20:47–52.

15. Puttaraju M, DiPasquale J, Baker CC, Mitchell LG, Garcia-Blanco MA. Messenger RNA repair and restoration of protein function by spliceosome-mediated RNA trans-splicing. *Mol Ther* 2001, 4:105–114.

16. Wang J, Mansfield SG, Cote CA, Jiang PD, Weng K, Amar MJ, Brewer BH Jr, Remaley AT, McGarrity GJ, Garcia-Blanco MA, et al. Trans-splicing into highly abundant albumin transcripts for production of therapeutic proteins in vivo. *Mol Ther* 2009, 17:343–351.

17. Coady TH, Lorson CL. Trans-splicing-mediated improvement in a severe mouse model of spinal muscular atrophy. *J Neurosci* 2010, 30:126–130.

18. Coady TH, Shababi M, Tullis GE, Lorson CL. Restoration of SMN function: delivery of a trans-splicing RNA re-directs SMN2 pre-mRNA splicing. *Mol Ther* 2007, 15:1471–1478.

19. Liu X, Luo M, Zhang LN, Yan Z, Zak R, Ding W, Mansfield SG, Mitchell LG, Engelhardt JF. Spliceosome-mediated RNA trans-splicing with recombinant adeno-associated virus partially restores cystic fibrosis transmembrane conductance regulator function to polarized human cystic fibrosis airway epithelial cells. *Hum Gene Ther* 2005, 16:1116–1123.

20. Wally V, Brunner M, Lettner T, Wagner M, Koller U, Trost A, Murauer EM, Hainzl S, Hintner H, Bauer JW. K14 mRNA reprogramming for dominant epidermolysis bullosa simplex. *Hum Mol Genet* 2010, 19:4715–4725.

21. Zayed H, Xia L, Yerich A, Yant SR, Kay MA, Puttaraju M, McGarrity GJ, Wiest DL, McIvor RS, Tolar J, et al. Correction of DNA protein kinase deficiency by spliceosome-mediated RNA trans-splicing and sleeping beauty transposon delivery. *Mol Ther* 2007, 15:1273–1279.

22. Lorain S, Peccate C, Le Hir M, Griffith G, Philippi S, Precigout G, Mamchaoui K, Jollet A, Voit T, Garcia L. Dystrophin rescue by trans-splicing: a strategy for DMD genotypes not eligible for exon skipping approaches. *Nucleic Acids Res* 2013, 41:8391–8402.

23. Rindt H, Yen PF, Thebeau CN, Peterson TS, Weisman GA, Lorson CL. Replacement of huntingtin exon 1 by trans-splicing. *Cell Mol Life Sci* 2012, 69:4191–4204.

24. Wally V, Klausegger A, Koller U, Lochmuller H, Krause S, Wiche G, Mitchell LG, Hintner H, Bauer JW. 5' trans-splicing repair of the PLEC1 gene. *J Invest Dermatol* 2008, 128:568–574.

25. Tahara M, Pergolizzi RG, Kobayashi H, Krause A, Luettich K, Lesser ML, Crystal RG. Trans-splicing repair of CD40 ligand deficiency results in naturally regulated correction of a mouse model of hyper-IgM X-linked immunodeficiency. *Nat Med* 2004, 10:835–841.

26. Chao H, Mansfield SG, Bartel RC, Hiriyanna S, Mitchell LG, Garcia-Blanco MA, Walsh CE. Phenotype correction of hemophilia A mice by spliceosome-mediated RNA trans-splicing. *Nat Med* 2003, 9:1015–1019.

27. Monjaret F, Bourg N, Suel L, Roudaut C, Le Roy F, Richard I, Charton K. Cis-splicing and translation of the pre-trans-splicing molecule combine with efficiency in spliceosome-mediated RNA trans-splicing. *Mol Ther* 2014, 22:1176–1187.

28. Mansfield SG, Clark RH, Puttaraju M, Kole J, Cohn JA, Mitchell LG, Garcia-Blanco MA. 5' exon replacement and repair by spliceosome-mediated RNA trans-splicing. *RNA* 2003, 9:1290–1297.

29. Murauer EM, Koller U, Hainzl S, Wally V, Bauer JW. A reporter-based screen to identify potent 3' trans-splicing molecules for endogenous RNA repair. *Hum Gene Ther Methods* 2013, 24:19–27.

30. Mansfield SG, Kole J, Puttaraju M, Yang CC, Garcia-Blanco MA, Cohn JA, Mitchell LG. Repair of CFTR mRNA by spliceosome-mediated RNA trans-splicing. *Gene Ther* 2000, 7:1885–1895.

31. Lorain S, Peccate C, Le Hir M, Garcia L. Exon exchange approach to repair Duchenne dystrophin transcripts. *PLoS One* 2010, 5:e10894.

32. Murauer EM, Gache Y, Gratz IK, Klausegger A, Muss W, Gruber C, Meneguzzi G, Hintner H, Bauer JW. Functional correction of type VII collagen expression in dystrophic epidermolysis bullosa. *J Invest Dermatol* 2011, 131:74–83.

33. Chen HY, Kathirvel P, Yee WC, Lai PS. Correction of dystrophia myotonica type 1 pre-mRNA transcripts by artificial trans-splicing. *Gene Ther* 2009, 16:211–217.

34. Berger A, Lorain S, Josephine C, Desrosiers M, Peccate C, Voit T, Garcia L, Sahel JA, Bemelmans AP. Repair of rhodopsin mRNA by spliceosome-mediated RNA trans-splicing: a new approach for autosomal dominant retinitis pigmentosa. *Mol Ther* 2015, 23:918–930.

35. Dallinger G, Puttaraju M, Mitchell LG, Yancey KB, Yee C, Klausegger A, Hintner H, Bauer JW. Development of spliceosome-mediated RNA trans-splicing (SMaRT) for the correction of inherited skin diseases. *Exp Dermatol* 2003, 12:37–46.

36. Rodriguez-Martin T, Garcia-Blanco MA, Mansfield SG, Grover AC, Hutton M, Yu Q, Zhou J, Anderton BH, Gallo JM. Reprogramming of tau alternative splicing by spliceosome-mediated RNA trans-splicing: implications for tauopathies. *Proc Natl Acad Sci USA* 2005, 102:15659–15664.

37. Avale ME, Rodriguez-Martin T, Gallo JM. Trans-splicing correction of tau isoform imbalance in a mouse model of tau mis-splicing. *Hum Mol Genet* 2013, 22:2603–2611.

38. Philippi S, Lorain S, Beley C, Peccate C, Precigout G, Spuler S, Garcia L. Dysferlin rescue by spliceosome-mediated pre-mRNA trans-splicing targeting introns harbouring weakly defined 3' splice sites. *Hum Mol Genet* 2015, 24:4049–4060.

39. He X, Liu F, Yan J, Zhang Y, Yan J, Shang H, Dou Q, Zhao Q, Song Y. Trans-splicing repair of mutant p53 suppresses the growth of hepatocellular carcinoma cells in vitro and in vivo. *Sci Rep* 2015, 5:8705.

40. He X, Liao J, Liu F, Yan J, Yan J, Shang H, Dou Q, Chang Y, Lin J, Song Y. Functional repair of p53 mutation in colorectal cancer cells using trans-splicing. *Oncotarget* 2015, 6:2034–2045.

41. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA* 2013, 4:49–60.

42. Robberson BL, Cote GJ, Berget SM. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol* 1990, 10:84–94.

43. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004, 11:377–394.

44. Coady TH, Baughan TD, Shababi M, Passini MA, Lorson CL. Development of a single vector system that enhances trans-splicing of SMN2 transcripts. *PLoS One* 2008, 3:e3468.

45. Shabasi M, Lorson CL. Optimization of SMN trans-splicing through the analysis of SMN introns. *J Mol Neurosci* 2012, 46:459–469.

46. Takahara T, Tasic B, Maniatis T, Akanuma H, Yanagisawa S. Delay in synthesis of the 3' splice site promotes trans-splicing of the preceding 5' splice site. *Mol Cell* 2005, 18:245–251.

47. Wang Y, Ma M, Xiao X, Wang Z. Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules. *Nat Struct Mol Biol* 2012, 19:1044–1052.

48. Wang Y, Xiao X, Zhang J, Choudhury R, Robertson A, Li K, Ma M, Burge CB, Wang Z. A complex network of factors with overlapping affinities represses splicing through intronic elements. *Nat Struct Mol Biol* 2013, 20:36–45.

49. Martinez-Contreras R, Fisette JF, Nasim FU, Madden R, Cordeau M, Chabot B. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol* 2006, 4:e21.

50. Kikumori T, Cote GJ, Gagel RF. Promiscuity of pre-mRNA spliceosome-mediated trans splicing: a problem for gene therapy? *Hum Gene Ther* 2001, 12:1429–1441.

51. Koller U, Wally V, Bauer JW, Murauer EM. Considerations for a successful RNA trans-splicing repair of genetic disorders. *Mol Ther Nucleic Acids* 2014, 3:e157.

52. Finta C, Zaphiropoulos PG. Intergenic mRNA molecules resulting from trans-splicing. *J Biol Chem* 2002, 277:5882–5890.

53. Romani A, Guerra E, Trerotola M, Alberti S. Detection and analysis of spliced chimeric mRNAs in sequence databanks. *Nucleic Acids Res* 2003, 31:e17.

54. Wagner E, Lykke-Andersen J. mRNA surveillance: the perfect persist. *J Cell Sci* 2002, 115:3033–3038.

55. Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 2006, 441:537–541.

56. Boudreau RL, Martins I, Davidson BL. Artificial microRNAs as siRNA shuttles: improved safety as compared to shRNAs in vitro and in vivo. *Mol Ther* 2009, 17:169–175.

57. McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* 1997, 11:3306–3318.

58. Ott de Bruin LM, Volpi S, Musunuru K. Novel genome-editing tools to model and correct primary immunodeficiencies. *Front Immunol* 2015, 6:250.