Splicing-based Biocontainment Devices

Allison Taggart[1], **Miles Rogers**[1], **Jacob Beal**[1]
[1]Raytheon BBN Technologies, Cambridge, MA
allison.j.taggart@rtx.com,miles.rogers@rtx.com,jake.beal@rtx.com

1 INTRODUCTION

Endogenous RNA splicing is catalyzed by a large ribonucleoprotein complex, known as the spliceosome, and contains snRNAs that pair directly to the premRNA gene (Fig 1A, left). This pairing is critical for splice site recognition and accurate splice site pairing. Mutations to the splicing sequences involved in these interactions can disrupt recognition and result in incorrect mRNA processing. We propose to take advantage of these recognition events in RNA splicing to engineer biocontainment by making splicing-based containment devices that rely on a combination of native splicing machinery and artificial supplied splicing components.

Orthogonal splicing systems have been demonstrated in the lab in which premRNAs with mutated splicing sequences can be rescued with the application of artificial snRNAs containing the compensatory mutations [1] [2]. Our design inserts introns with mutations in one or more splice site regions and matching mutations in engineered snRNAs to restore complementarity (Fig 1A, right).

There are two strategies in which these splicing devices can be used for system containment. In the first strategy, the device can be inserted into one or more essential genes (Fig 1B). Here, when artificial splicing components are not supplied to the organism, the essential genes will not be correctly processed and the organism is no longer viable. In the second strategy, the device can be inserted into an engineered gene or gene of interest, such that in non-desired conditions the organism will persist but the gene containing the device will be nonfunctional (Fig 1C). There exist different implications and evolutionary pressures for escape mechanisms for each strategy, and the most appropriate strategy may depend on application.

2 MODELING SYSTEM EFFICACY REQUIREMENTS

The desired state, in which the engineered device is spliced and processed correctly in the presence of both native and synthetic spliceosomal components, can be computed as:

$d_{OS} = \alpha(f_{NE} + f_{EE})^k$

where α = transcription rate, f_{NE} = fraction of engineered introns spliced by native spliceosome, f_{EE} = fraction of engineered introns spliced by synthetic spliceosome, and k = number of synthetic introns. If d_{OS} is too low, the engineered device will not be processed and the system will die.

The nondesired state, in which biocontainment fails and splicing occurs in the absence of synthetic snRNAs, is:

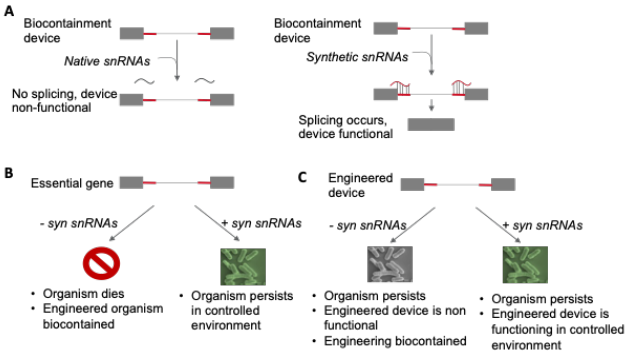

Figure 1: (A) Mutant introns block endogenous splicing (left), but splice with artificial snRNA (right). Two biocontainment strategies: (B) insert a artificial intron in an essential gene, or (C) into the engineered device.

$n_{OS} = \alpha(f_{NE})^k$

The efficacy e of the system is thus proportional to the ratio of desired splicing to nondesired splicing:

$e \propto \frac{(f_{NE}+f_{EE})^k}{(f_{NE})^k}$

Modeling system efficacy shows a relationship between the required efficiencies of recognition of both native and synthetic spliceosomal components to the overall efficiency of the system (Fig 2). The highest efficacy system would be a system in which f_{EE} (splicing in the presence of supplied synthetic RNAs) is 1, and f_{NE} (splicing in the absence of supplied synthetic RNAs) is 0. Overall, our results demonstrate that there is an inflection point, such that as long as we are able to maintain a low f_{NE}, f_{EE} only needs to achieve about 0.75 efficiency to achieve maximum system efficacy. Additionally, our modeling results suggest that adding multiple introns to the system can improve system efficiency. There exist multiple design strategies for increasing the number of synthetic introns to the system.

3 DESIGN VERSUS EVOLUTIONARY ESCAPE

Splicing-based biocontainment can be implemented using a number of design strategies, outlined in Fig 3., all of which use an unspliced biocontainment intron to interrupt gene function. Inserting a biocontainment intron with a length not a multiple of three will shift codons out of frame for

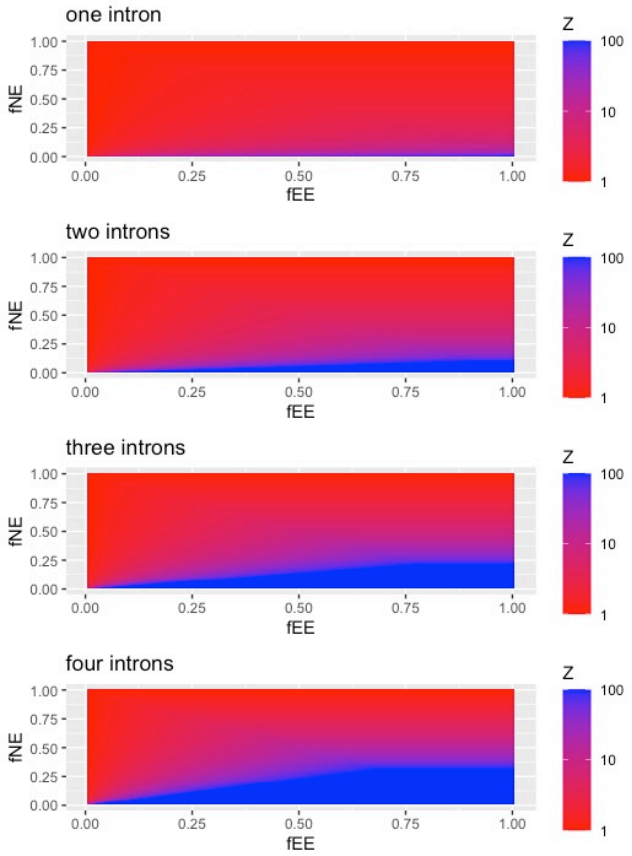

Figure 2: Modeling estimate for biocontainment efficiency from various numbers of introns and range of splicing parameters, coloring efficacy on a log scale.

the remainder of the transcript, likely destroying function (design 1). Similarly, inserting a biocontainment intron containing a stop codon can produce a non-functional truncated protein (design 2). Another design category is to insert a biocontainment intron in a known activity domain or relevant protein structure that is necessary for protein function. When unspliced, the biocontainment intron will interrupt these domains and could interfere with function (design 3).

There are additional strategies in which a biocontainment intron can be used to introduce targeting sites, such as cleavage, RNA binding protein, or siRNA targeting sites, which can all be used to block gene processing (designs 4,5). It is also possible to design a strain with an essential small RNA contained within the biocontainment, in which processing of the small RNA requires splicing and correct intron lariat formation (design 6). Lastly, biocontainment introns can be designed to contain destabilizing sequences or structures, such as hairpins, that when unspliced interfere with function (Design 7).

Several classes of evolutionary escape and biocontainment failure must be considered, including point mutations,

1	Unspliced intron puts mRNA out of frame
2	Unspliced intron introduces premature stop codon
3	Unspliced intron interrupts an activity domain or causes a misfolding change
4	Unspliced intron contains a cleavage site that can be targeted for mRNA turnover
5	Unspliced intron contains a site that recruits RBPs or siRNAs that target transcript and block processing
6	Unspliced intron contains required small RNA
7	Unspliced intron contains destabilizing mRNA or protein sequence or nonfunctional structure

Figure 3: Potential splicing-based biocontainment designs

recombination and deletion, horizontal gene transfer, and alternate pathways for essential genes. Point mutations may restore splice site sequences, affecting all designs, so it may be beneficial to have multiple mutations in the splice site sequences or spread multiple biocontainment introns across one or multiple genes. Point mutations may also restore the reading frame (Design 1) or remove a premature stop codon (Design 2), which could similarly be mitigated by insertion of multiple introns. Interruption of an activity domain (Design 3) is likely to be more resistant to point mutations, and may prove to be an attractive design choice. Recombination and deletion may remove biocontainment introns, but is predicted to be a lesser threat due to the precision required at the intron boundaries. Horizontal gene transfer may be a bigger concern, but is dependent on application, chassis, and mating patterns. Similarly the degree of concern about alternate pathways depends on organism characterization. Both horizontal gene transfer and alternate pathways can be mitigated by spreading biocontainment introns across multiple genes. Leaky gene expression is likely to be a challenge for Designs 4 and 5, which require functional external systems to be recruited to the site for containment. Finally, Design 6 may be challenging to implement, due to ambiguities with intron lariat stability and processing, but these may be overcome with well-characterized introns.

In sum, the use of an orthogonal splicing system appears to be a viable approach to biocontainment, with a number of potential designs, some of which are more resilient to escape than others. Ongoing work in our laboratory aims to demonstrate these approaches in practice.

REFERENCES

[1] Smith, D. J., Konarska, M. M., and Query, C. C. Insights into branch nucleophile positioning and activation from an orthogonal pre-mRNA splicing system in yeast. *Molecular Cell 34*, 3 (May 2009), 333–343.

[2] Zhuang, Y., and Weiner, A. M. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes & Development 3*, 10 (Oct. 1989), 1545–1552.