



Published in final edited form as:

*Anal Biochem.* 2011 October 1; 417(1): 159–161. doi:10.1016/j.ab.2011.05.035.

## Importance of RNA analysis in interpretation of reporter gene expression data

**Victoria P. Belancio\***

Department of Structural and Cellular Biology, Tulane School of Medicine, Tulane Cancer Center, New Orleans, LA

### Abstract

Reporter gene assays have proven to be an important tool in analyzing cis and trans factors that influence gene expression. They have sometimes been adapted for studies in which they are not totally reliable. Modifications that change the RNA expressed from the reporter gene may result in regulation of reporter gene expression at multiple levels simultaneously. The data provided here illustrate the difficulties that may arise from post-transcription regulation in various reporter gene formats. This serves as a warning that further RNA studies may be necessary, if comparisons are to be made between reporter constructs whose RNA is not identical.

---

Reporter gene assays were originally developed as a powerful tool for studying traditional promoter and enhancer elements [1,2]. Because of their ease of use, further reporter constructs have been utilized to study the function of other critical sequence elements, including those influencing RNA stability [3], IRES elements [4], translation [5] and others [6]. In these types of experimental designs, contrary to the original usage of reporter genes, specific modifications to the reporter construct change sequence composition and/or size of the resulting mRNA. The majority of studies that involve identification of functional sequences in the context of reporter gene assay is designed to introduce point mutations, serial terminal or internal deletions or replacements of the sequences of interest within the 5', 3' UTRs, or IRES sequences [6,7]. The concerns that some of these manipulations, if performed without appropriate controls for their effects on mRNA production, can result in misleading interpretations have been raised previously [8]. While many of the experimental designs control for the fact that reporter gene constructs produce different mRNAs with changes that may cause altered protein production through multiple levels of regulation simultaneously, there are still some that overlook this important aspect of experimental design [7,9–11].

RNAs often contain cis-acting elements involved in RNA processing (such as splicing and polyadenylation) or RNA transport that can significantly influence their translational potential. Signals for RNA splicing and polyadenylation are composed of multiple sequences that, to function efficiently, have to be properly positioned with respect to one another [12]. Manipulations such as deletions or sequence substitutions that not only change the size of the 5' or 3' UTRs, but also their sequence composition can activate cryptic splice sites or splice enhancers that may alter splicing efficiency or result in the production of

---

© 2011 Elsevier Inc. All rights reserved.

\*Corresponding author: 1430 Tulane Ave., SL49, New Orleans, LA, 70112, Phone: 504-988-4506, vperpe@tulane.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

alternatively spliced non-functional RNAs [13]. Similar manipulations within the 3' UTR of a reporter gene may alter mRNA stability or transport [3,6]. Generation of alternative reporter gene mRNA products may skew the interpretation of the data obtained by assessing only the levels of the reporter protein. Similarly, even small insertions/deletions or point mutations can inactivate existing cis-acting elements, as well as activate cryptic or create functional cis-acting elements important for RNA processing. As a result, they can significantly change the quantity and/or spectrum of mRNA products [13]. The latter is particularly relevant when the effects of genomic polymorphisms or inactivating mutations are tested in the reporter gene assay because these point mutations can reside within a sequence with multiple functions [9].

Assessment of the consequences of point mutations by northern blot analysis is important when (i) multiple mRNA species can be generated by an expression vector creating a possibility that introduced point mutations may differentially affect production of these products and/or (ii) they reside within regulatory elements that affect transcription or translation. This is often a possibility when intron-containing genomic fragments are subcloned upstream of the reporter gene unit with a chimeric intron to generate a mini-gene reporter construct [14]. To demonstrate this point, reporter gene vectors containing a Long Interspersed Element-1 (L1) 5' untranslated region (5'UTR) upstream of the *Firefly* luciferase gene were used (Figure 1A). L1 elements provide a good experimental system to test the adequacy of the reporter gene assay because their 5'UTR contains a functional internal promoter with transcription factor binding sites as well as functional splice sites [13,15]. Further, a reporter gene approach is often used in L1-related studies [15,16]. This experimental design parallels methodology used in the reports that assess the presence of functional promoters within intronic regions and the effect of point mutations or polymorphisms within these sequences [11,17]. Figure 1B demonstrates northern blot analysis of mRNAs generated by a reporter gene construct that contains wild type (WT) L1 5'UTR or the same 5'UTR with a point mutation that eliminates a splice donor site [13] and modifies a binding site for RUNX3 transcription factor (Mut) [18]. As a result the construct containing the mutant 5'UTR does not generate any detectable spliced mRNA, a product that is produced by the wt construct. Additionally, it produces reduced levels of the unspliced mRNA compared to wt most likely due to the loss of RUNX3 binding [18]. Despite the differences in the mRNA species and levels of expression, both constructs generate equivalent amounts of luciferase activity, suggesting that this point mutation has no effect on protein expression. This example demonstrates that a single point mutation can not only change the amount of generated mRNA by altering promoter function, but also the spectrum of mRNAs produced by reporter gene constructs through complex RNA splicing. Importantly, these types of alterations can easily result from the conventional experimental approaches commonly used to manipulate reporter genes; therefore, they may produce misleading data by independently affecting protein production.

To illustrate how reliance on the reporter gene output alone can produce false negative results a polyadenylation (PolyA or pA) site-containing fragment of a human L1 element [19] was used to test the function of this site in the reporter gene assay (experimental design is shown in Figure 2). A fragment of L1 sequence (1294–1729 nt of L1.3, L1wt) containing the functional polyA site or the same fragment containing a point mutation that inactivates this site (L1m) were cloned into the chimeric intron of the reporter gene vector (Figure 2A). The SV40 polyA signal was also cloned into the same position to serve as positive control for the presence of functional polyA site. Luciferase activity was determined in transiently transfected NIH 3T3 cells. Only control vector produced significant signal from luciferase activity, the other three constructs registered background levels of the luciferase activity (Fig. 2B). These data suggest that the polyA signal within the introduced L1 sequence is not the only functional factor that influences the readout of the assay because elimination of this

functional polyA site does not restore luciferase production. mRNA analysis of these constructs transiently transfected into NIH 3T3 cells demonstrated that L1m and control constructs produced a similar steady-state amount of mRNA, while SV40 and L1wt constructs generated undetectable or very low amounts of mRNA, respectively. Thus, mRNA analysis supports the hypothesis that the polyA site within the L1 sequence is functional. However, the majority of the mRNA produced from the L1m construct was unspliced demonstrating that introduced L1 sequence interferes with the proper splicing of the chimeric intron present within the reporter construct, which interferes with translation of luciferase protein. This example demonstrates that mRNA analysis can help to avoid making a false negative conclusion based on assessment of protein activity in the reporter gene assay. Additionally, it demonstrates, not surprisingly, that a single point mutation within a cis-acting sequence can have a very dramatic effect on mRNA production. This example mimics commonly used experimental design where genomic sequences composed of introns and exons are subcloned upstream of the chimeric intron followed by a reporter gene to generate minigenes for testing the effects of polymorphic changes or presence of functional promoters [9,11]. Because splicing efficiency is modulated by the cis-acting elements located both 5' and 3' relative to the splice donor and splice acceptor sites, sequence manipulations within introns as well as exonic regions can affect splicing efficiency [12].

The two examples shown here illustrate that point mutations introduced to assess a function of a particular sequence can also overlap with cis-acting elements controlling mRNA production and therefore have significant simultaneous effects on multiple aspects of gene expression that may not always be accurately detected by measurements of protein activity alone. Figure 2 also demonstrates that sequence manipulation within reporter construct can have profound effects on RNA processing. The same issues are likely to arise when serial truncations or internal deletions are introduced within sequences tested for function in reporter gene-based assays.

In conclusion, based on our data, it is necessary to include RNA analysis by northern blots into experimental designs that involve sequence manipulations in a reporter gene assay even if the changes are as small as point mutations as well as when new sequence compositions are generated in reporter gene constructs through deletions or insertions. Out of the two conventional approaches for RNA analysis, RT-PCR and northern blots, the latter can more reliably detect novel mRNA products and, therefore, it can prevent loss of data critical for making accurate conclusions about the outcome of a reporter-based experiment. Usage of polyA and splice site prediction programs ([http://rulai.cshl.edu/tools/polyadq/polyadq\\_form.html](http://rulai.cshl.edu/tools/polyadq/polyadq_form.html) and [http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)) [20,21] can be helpful for preliminary assessment of the potential impact point mutations or sequence combinations may have on the reporter construct expression.

## Materials and methods

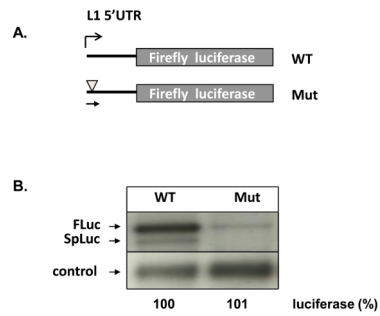
Culturing of NIH 3T3 cells, transfections, RNA harvest, polyA selection, northern blot analyses, and site directed mutagenesis were done as described [13,19]. Protein harvest and luciferase analysis were done according to the manufacture's protocol (Promega). pOPI3luc and pGL3 control vectors (Promega) were used for cloning.

## Acknowledgments

VPB is supported by P20RR020152, NIA 5K01AG030074-02 from the National Institutes of Health (NIH) and The Ellison Medical Foundation New Scholar in Aging award 547305G1

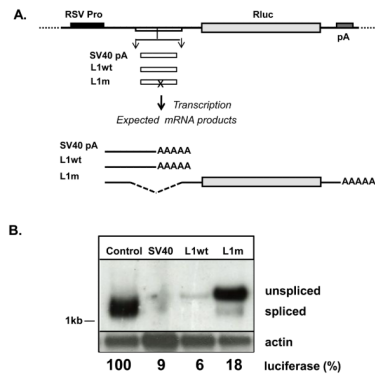
## References

1. Gorman CM, Merlino GT, et al. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. *Proc Natl Acad Sci U S A*. 1982; 79:6777–6781. [PubMed: 6294651]
2. Gorman CM, Moffat LF, Howard BH. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol Cell Biol*. 1982; 2:1044–1051. [PubMed: 6960240]
3. Li F, Hu DY, et al. RNA-binding protein HuR regulates RGS4 mRNA stability in rabbit colonic smooth muscle cells. *Am J Physiol Cell Physiol*. 2010; 299:C1418–C1429. [PubMed: 20881234]
4. Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*. 1988; 334:320–325. [PubMed: 2839775]
5. Qiao H, Lu N, et al. Rare codons in uORFs of baculovirus p13 gene modulates downstream gene expression. *Virus Res*. 2011; 155:249–253. [PubMed: 20970467]
6. Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell*. 2009; 35:467–478. [PubMed: 19716791]
7. Yokoyama NN, Pate KT, et al. A role for YY1 in repression of dominant negative LEF-1 expression in colon cancer. *Nucleic Acids Res*. 2010; 38:6375–6388. [PubMed: 20525792]
8. Kozak M. Rethinking some mechanisms invoked to explain translational regulation in eukaryotes. *Gene*. 2006; 382:1–11. [PubMed: 16859839]
9. Kankova K, Benes P, Kuchrickova S. Functional analysis of the common haplotype in the receptor for advanced glycation end-products gene previously identified as a susceptibility factor for diabetic nephropathy. *Exp Clin Endocrinol Diabetes*. 2010; 118:93–95. [PubMed: 20112183]
10. Shih PA, Wang L, et al. Peptide YY (PYY) gene polymorphisms in the 3'-untranslated and proximal promoter regions regulate cellular gene expression and PYY secretion and metabolic syndrome traits in vivo. *J Clin Endocrinol Metab*. 2009; 94:4557–4566. [PubMed: 19820027]
11. Takahashi S, Inoue I, et al. A promoter in the novel exon of hPPARgamma directs the circadian expression of PPARgamma. *J Atheroscler Thromb*. 2010; 17:73–83. [PubMed: 20093779]
12. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol*. 2009; 10:741–754. [PubMed: 19773805]
13. Belancio VP, Hedges DJ, Deininger P. LINE-1 RNA splicing and influences on mammalian gene expression. *Nucleic Acids Res*. 2006; 34:1512–1521. [PubMed: 16554555]
14. Wang H, Molfenter J, et al. Promotion of exon 6 inclusion in HuD pre-mRNA by Hu protein family members. *Nucleic Acids Res*. 2010; 38:3760–3770. [PubMed: 20159993]
15. Swergold GD. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol*. 1990; 10:6718–6729. [PubMed: 1701022]
16. Han JS, Szak ST, Boeke JD. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature*. 2004; 429:268–274. [PubMed: 15152245]
17. Bellini I, Pitto L, et al. DeltaN133p53 expression levels in relation to haplotypes of the TP53 internal promoter region. *Hum Mutat*. 2010; 31:456–465. [PubMed: 20127977]
18. Yang N, Zhang L, et al. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Research*. 2003; 31:4929–4940. [PubMed: 12907736]
19. Perepelitsa-Belancio V, Deininger P. RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nat Genet*. 2003; 35:363–366. [PubMed: 14625551]
20. Reese MG, Eeckman FH, et al. Improved splice site detection in Genie. *J Comput Biol*. 1997; 4:311–323. [PubMed: 9278062]
21. Tabaska JE, Zhang MQ. Detection of polyadenylation signals in human DNA sequences. *Gene*. 1999; 231:77–86. [PubMed: 10231571]



**Figure 1. The effect of point mutations on the reporter gene mRNA**

(A) Two reporter gene constructs containing either wild type 5'UTR sequence of the human L1 element (WT) or the same sequence containing a point mutation that mutates a splice donor site (T to C at position 99 of L1.3 [13]) and a RUNX binding site [18] (Mut, light grey triangle). Black horizontal arrow indicates the position of the probe used to detect mRNA produced by these constructs. (B) Northern blot analysis. Fluc indicates full-length mRNA produced by both WT and Mut constructs; SpLuc marks spliced mRNA produced by the wt construct. Control denotes mRNA for neomycin resistance gene encoded by the reporter gene and used as transfection and loading control. The numbers at the bottom demonstrate relative levels of reporter gene activity (% percent of WT).



**Figure 2. False negative outcomes can be prevented by northern blot analysis**

(A) Renilla luciferase (Rluc, light grey box) reporter gene construct. RSV promoter (RSV Pro, black box), Intron (white box), TK polyadenylation signal (pA, dark grey box). DNA fragments subcloned into the intron of the vector: SV40 pA site, L1 sequence containing functional polyA site (L1wt), and the same L1 sequence with mutated pA signal (L1m, AATAAA at position 1788–1793 of L1.3 to GATCAA [19]). (B) Northern blot analysis of polyA-selected mRNA from NIH 3T3 cells transiently transfected with the constructs described above and probed with probes to Rluc or actin as described [13,19]. Spliced and unspliced mRNA produced by the reporter construct is indicated. 1kb is a size marker. The numbers at the bottom indicate relative levels of reporter gene activity.