Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Genetics and Genomics 39 (2012) 503−520
JGG

METHOD

Effective Expression-Independent Gene Trapping and Mutagenesis Mediated by *Sleeping Beauty* Transposon

Guili Song [a,b], Qing Li [a], Yong Long [a], Perry B. Hackett [c], Zongbin Cui [a,*]

[a] *The Key Laboratory of Aquatic Biodiversity and Conservation, Institute of Hydrobiology, Chinese Academy of Sciences, 7 Donghu Rd., Wuhan, Hubei 430072, China*
[b] *Graduate University of the Chinese Academy of Sciences, Beijing 100049, China*
[c] *Genetics, Cell Biology, and Development, University of Minnesota, 321 Church St. SE, Minneapolis, MN 55455, USA*

Received 3 May 2012; revised 21 May 2012; accepted 28 May 2012
Available online 23 August 2012

ABSTRACT

Expression-independent gene or polyadenylation [poly(A)] trapping is a powerful tool for genome-wide mutagenesis regardless of whether a targeted gene is expressed. Although a number of poly(A)-trap vectors have been developed for the capture and mutation of genes across a vertebrate genome, further efforts are needed to avoid the 3′-terminal insertion bias and the splice donor (SD) read-through, and to improve the mutagenicity. Here, we present a *Sleeping Beauty* (SB) transposon-based vector that can overcome these limitations through the inclusion of three functional cassettes required for gene-finding, gene-breaking and large-scale mutagenesis, respectively. The functional cassette contained a reporter/selective marker gene driven by a constitutive promoter in front of a strong SD signal and an AU-rich RNA-destabilizing element (ARE), which greatly reduced the SD read-through events, except that the internal ribosomal entry site (IRES) element was introduced in front of the SD signal to overcome the phenomenon of 3′-bias gene trapping. The breaking cassette consisting of an enhanced splicing acceptor (SA), a poly(A) signal coupled with a transcriptional terminator (TT) effectively disrupted the transcription of trapped genes. Moreover, the *Hsp70* promoter from tilapia genome was employed to drive the inducible expression of SB11, which allows the conditional remobilization of a trap insert from a non-coding region. The combination of three cassettes led to effective capture and disruption of endogenous genes in HeLa cells. In addition, the Cre/LoxP system was introduced to delete the *Hsp70-SB11* cassette for stabilization of trapped gene interruption and biosafety. Thus, this poly(A)-trap vector is an alternative and effective tool for identification and mutation of endogenous genes in cells and animals.

KEYWORDS: Poly(A) trapping; *Sleeping Beauty* transposon; Insertional mutagenesis; HeLa cells; Zebrafish embryos

1. INTRODUCTION

Gene trapping is an efficient approach for insertional mutagenesis of genes in a target genome. A conventional gene-trap vector consists of a promoterless marker/reporter gene flanked by an upstream splice acceptor and a downstream poly(A) signal (Gossler et al., 1989; Stanford et al., 2001). When inserted into either an intron or an exon of expressed genes, the marker/reporter gene can be transcribed from the endogenous gene promoter and the transcription is terminated prematurely at the inserted poly(A) site. The nuclear pre-mRNA is then spliced to form a fusion transcript with the trapped exon immediately upstream of the insertion site. Since the processed fusion transcript encodes a truncated, often non-functional version of endogenous protein and the marker/reporter, gene trapping is employed to elucidate gene functions by disrupting expression of trapped genes across a target

* Corresponding author. Tel/fax: +86 27 6878 0090.
E-mail address: zbcui@ihb.ac.cn (Z. Cui).

http://dx.doi.org/10.1016/j.jgg.2012.05.010

genome (Stanford et al., 2001). During the past two decades, a number of gene-trap vectors have been successfully used for insertional mutagenesis and high-throughput screening of mutations in genes of embryonic stem cells (ESCs). The International Gene Trap Consortium (IGTC) has hold more than 650,000 ES cell lines and most of these cell lines were created predominantly through the use of various conventional gene-trap vectors, both plasmid- and retroviral-based (Skarnes et al., 2004; Guan et al., 2010). In addition, approximately 70% of the protein-encoding genes in the mouse genome have been disrupted by gene trap insertions and some of ESC lines harboring mutations in a single gene have been used for the generation of null mice (De-Zolt et al., 2009). However, the achievement of saturation mutagenesis in a target genome through conventional vectors is difficult because these vectors are not able to capture genes that are not or poorly expressed in undifferentiated cells.

To circumvent this problem, a number of expression-independent or poly(A)-trap vectors have been developed and successfully utilized for the identification of differentially expressed genes (Chen et al., 2004; Lin et al., 2006; Tsakiridis et al., 2009). Basic poly(A)-trap vector contains are porter/selectable marker gene flanked by an upstream constitutive promoter and a downstream splice donor (SD). Integration of a trap cassette upstream of a functional poly(A) sequence from endogenous genes leads to the generation of a stable pre-mRNA and the proper splicing between the trap SD with a splicing acceptor (SA) downstream of the insertion site gives rise to a fusion transcript, which encodes the reporter/selectable marker and an N-terminal truncated version of endogenous protein. Theoretically, poly(A)-trapping is suitable for the capture of genes almost equally regardless of their transcriptional status in target cells (Matsuda et al., 2004). Nevertheless, later findings indicate that these basic poly(A)-trap vectors have a bias toward the last introns of trapped genes because of the activation of nonsense-mediated mRNA decay (NMD) mechanism (Baker and Parker, 2004). NMD promotes the selection of trapping events in the 3′-most intron of target genes as it triggers the degradation of the selectable marker's transcript based on the presence of a premature termination codon. This disadvantage appears to be resolved in the UPATrap vector through insertion of an internal ribosomal entry site (IRES) and three initiation codons in front of its SD signal (Shigeoka et al., 2005). However, the majority of poly(A)-trap vectors still suffer from an NMD-based handicap and attempts are still undergoing to overcome other problems such as the background SD read-through events. In addition, gene-trap vectors available appear to each have their own insertional 'hot spots' (Nord et al., 2007). Thus, the saturation mutagenesis of a target genome can be achieved most economically through the use of a wider range of vectors and approaches (Skarnes, 2005).

Pseudotyped retroviral vectors are now extensively utilized for insertional mutagenesis in vertebrates (Medico et al., 2001; Ellingsen et al., 2005). These vectors can be extremely efficient in generating mutant, but have three major disadvantages: limitation in packaging size, mediation of gene silencing and strong insertional biases (Ellis, 2005; Uren et al., 2005). Recently, several transposon-based vectors have been developed as alternatives for elucidation of gene functions in mouse and zebrafish (Kawakami et al., 2004; Jenkins et al., 2005; Largaespada et al., 2005; Rad et al., 2010; Clark et al., 2011). In comparison with viral vehicles, transposons can carry a large DNA fragment up to 10 kb (Geurts et al., 2003; Zayed et al., 2004; Huang et al., 2010). However, transposon-based vectors have a caveat that less than 10 copies of transposons are usually found in the genome of transgenic animals (Clark et al., 2004; Kawakami et al., 2004). Moreover, over produced transposases appear to be harmful to cellular growth and proliferation (Huang et al., 2010; Galla et al., 2011). Therefore, new strategies were urgently needed to generate novel transposon-based vectors for efficient gene trapping and insertional mutagenesis in vertebrates.

Sleeping Beauty (SB) transposon system is a synthetic member of the Tc1/mariner superfamily, which can transpose in an autonomous and horizontal manner (Plasterk et al., 1999). Sequences related to Tc1/marinar transposons are found in a wide range of host genomes and an active SB transposase gene was resurrected from an extinct DNA sequence in the genomes of salmonid fish through the correction of accumulated mutations (Ivics et al., 1997). The SB system consists of two components: a transposase gene and a DNA-transposon. SB transposons insert into TA dinucleotide sites in a recipient sequence *via* a cut-and-paste mechanism (Liu et al., 2005). Accordingly, it is used for long-term expression in transgenesis (Hackett et al., 2005) and insertional mutagenesis in vertebrates (Jenkins et al., 2005; Largaespada et al., 2005). In addition, an analysis of 1336 inserts in primary and cultured mammalian cells has demonstrated that SB integration exhibits less preference for transcriptional units and/or their promoters compared with retroviruses (Yant et al., 2005). Therefore, the SB system appears to be an ideal non-viral vehicle for production of transgenic animals and potential genome-wide mutagenesis due to its relative high activity, large cargo capacity and less integration site preference.

In this study, we described the generation of an effective poly(A)-trap vector through a combination of three functional cassettes. In the gene-finding cassette, an IRES from the encephalomyocarditis virus (EMCV) was inserted downstream of the selectable marker/reporter gene to suppress the NMD-mediated mRNA degradation of trapped genes and an AU-rich RNA-destabilizing element (ARE) was introduced into an intron in the SD to enrich for authentic trapping events (Xu et al., 1997). In the gene-breaking cassette, an exon containing a strong SA from the carp (*Cyprinus carpio*) *β-actin* gene (Liu et al., 1990) and an efficient poly(A) signal followed by a transcriptional terminator element from the human *gastrin* gene were used to disrupt the expression of trapped genes. The tilapia *HSP70* promoter (Molina et al., 2001) was employed to drive the expression of SB11 transposase, which allows the remobilization of integrated traps from non-coding sites to new locations and thus increases the opportunity of trapping and mutating endogenous genes. Activities of all components