ISJ 14: 271-281, 2017 ISSN 1824-307X

RESEARCH REPORT

Evaluation of a novel, short polyA signal from the *Bombyx mori* bidensovirus

M Wang, Q Yu, Y Li, Y Zhang, D Miao, Z Hu, Q Yao, K Chen

Institute of Life Sciences, Jiangsu University, 301 XueFu Road, Zhenjiang, Jiangsu 212013, China

Accepted August 7, 2017

Abstract

Bombyx mori bidensovirus (BmBDV) is the only virus which belongs to the new *Bidnaviridae* family established by the International Committee on Taxonomy of Viruses in 2012. The genes encoding the major capsid protein and DNA polymerase (pPolB) overlap partially at their 3′ untranslated regions, forming an extremely short, dual-function polyadenylation signal/stop codon (BmBDV polyA) to complete post-transcriptional modifications and terminate protein expression. Nevertheless, the functionality and usefulness of this signal regarding foreign genes remain unknown. To determine the effect of the BmBDV polyA on gene expression, the expression of the green fluorescent protein and firefly luciferase was evaluated with the BmBDV polyA, compared with the much larger SV40 polyA, under the control of the p5 or ie1 promoter. Fluorescence microscopy and dual luciferase assay both revealed enhanced expression of these proteins in the presence of the BmBDV polyA, meanwhile real-time qPCR also showed increased mRNA levels. Therefore, we conclude that the BmBDV polyA is a promising, characteristic polyA signal from an insect virus, and that it can promote gene expression at either the mRNA or protein levels. Additionally, this study also suggests that the BmBDV polyA potentially alleviates restrictions associated with the size of inserted fragments during constructing recombinant viruses.

Key Words: *Bombyx mori* bidensovirus; polyadenylation; protein expression assay; real-time qPCR

Introduction

Polyadenylation is an intermediate process between transcription and translation, which usually comprises the subsequent cleavage of the pre-mRNA and the addition of a polyA tail (Moore *et al*., 1985; Wickens, 1990). This process depends on recruiting cellular RNA processing factors, such as the cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulation factor (CstF), to the carboxyl-terminal domain of RNA polymerase II (Keller *et al*., 1991; Venkataraman *et al*., 2005). Then, CPSF and CstF recognize canonical AAUAAA motifs and GU-rich sequences in the transcripts, respectively (Murthy *et al*., 1995). Cleavage occurs approximately 10 - 30 nucleotides (nt) downstream of the AAUAAA motif. Subsequently, polyA polymerase adds a stretch of A residues to form the polyA tail, which is involved in mRNA stability and transportation from the nucleus to the cytoplasm (Beilharz *et al*., 2007). To the best of our knowledge, a typical polyA signal contains three elements of AAUAAA for CPSF binding and a downstream GU-rich region for CstF binding, as well as a seven-T element for potential transcription termination (Westwood *et al*., 1993; Jin *et al*., 2000).

Widely used polyA signals, such as the SV40 polyA and the HSVtk polyA, are usually derived from viruses associated with mammals. In the present study, the SV40 polyA was inserted downstream of the polyhedrin promoter in many baculovirus expression vector systems (BEVSs) (Salem *et al.*, 2015). Additionally, the SV40 polyA was inserted into a green fluorescent protein (GFP) gene expression cassette to test piggyBac-mediated germline transformation of the silkworm *B. mori* (Tamura *et al.*, 2000). The potential of being used for expressing exogenous products of polyA signal from insect viruses has not been studied very well.

BmBDV is the type species in the new genus *Bidensovirus*, which belongs to the *Bidnaviridae* family that was established by the International Committee on Taxonomy of Viruses in 2012 (Adams *et al.,* 2012). This virus can specifically infect the columnar cells of the midgut epithelium (Hu *et al*., 2013) of silkworms, leading to chronic densonucleosis disease. BmBDV has evolved from a

Corresponding authors:
Qin Yao
Keping Chen
Institute of Life Sciences
Jiangsu University
301 XueFu Road, Zhenjiang ,Jiangsu 212013, China
Email: yaoqin@ujs.edu.cn
Email: kpchen@ujs.edu.cn

Table 1 Primers used for small insertions synthesized by annealing in this study

Primer	One complementary strand sequence (5′-3′)	Ranges (nt)	Construction
BDVpA (63)	-ACGCGTTTGTTGGTTTTCAATAAATAAATGTATTAAAATTTAT AAATGTTTTATTCAATTACAACATCAGTCGAC-	2,896–2,958	pT-p5-gfp-BDVpA (63) pT-ie1-gfp-BDVpA (63) pGL_3-p5-BDVpA (63) pGL_3-ie1-BDVpA (63)
BDVpA (92)	-TCTAGATTGTTGGTTTTCAATAAATAAATGTATTAAAATTTAT AAATGTTTTATTCAATTACAACATCAGGATCC- -TCTAGAATAAATAAATGTATTAAAATTTATAAATGTTTTATTC AATTACAACATCATCCATATAATTAGAATACAATTCACCCTTA TAACCAAATGGTGTCGAC-	2,909–3,000	pGL_3-p5-BDVpA (92)
BDVpA (50)	-TCTAGAATAAATAAATGTATTAAAATTTATAAATGTTTTATTC AATTACAACATCAGTCGAC-	2,909–2,958	pGL_3-p5-BDVpA (50)
BDVpA (38)	-TCTAGAATAAATAAATGTATTAAAATTTATAAATGTTTTATTC AGTCGAC-	2,909–2,946	pGL_3-p5-BDVpA (38)
BDVpA (26)	-TCTAGAAAATGTTTTATTCAATTACAACATCAGTCGAC-	2,933–2,958	pGL_3-p5-BDVpA (26)
BDVpA (24)	-TCTAGAATAAATAAATGTATTAAAATTTATGTCGAC-	2,909–2,932	pGL_3-p5-BDVpA (24)

parvovirus ancestor from which it inherited a jelly-roll capsid protein and a superfamily 3 helicase (Krupovic *et al*., 2014). Just like other parvoviruses, BmBDV is a non-enveloped, linear DNA virus that has a spherical, icosahedral structure (Zhang *et al*., 2016). Although the virus is only 22 nm in diameter, its genome consists of two non-homologous single-stranded DNA molecules (VD1 and VD2; 6,543 and 6,024 nts, respectively), which are encapsidated into separate virions. The VD1 consists of four open reading frames (ORFs), which encode three nonstructural proteins [NS1, NS2 and DNA polymerase (pPolB)] (F. Wang *et al.*, 2011)and one structural protein (MCP, VP) (Li *et al*., 2009). The MCP transcript and the pPolB transcript overlap for 10 nts at their 3′ untranslated regions (UTRs). Additionally, the distance between the stop codons of VP and pPolB is only 24 nts. As mentioned above, this means that there is a short polyA signal (BmBDV polyA) in the overlap regions of these two genes. This characteristic is similar as SV40 polyA (contains two genes' polyA signal from complimentary strands), but the BmBDV polyA is much shorter in length.

To test whether this novel polyA signal from an insect virus is potentially useful for expressing exogenous products in insect cell lines, an intact BmBDV polyA containing an AAUAAA motif and GU-rich regions was fused with a GFP-encoding gene (*gfp*) and a firefly luciferase-encoding gene under the control of the p5 or ie1 promoter, respectively. The p5 promoter (236 bp) is an early promoter of BmBDV, which is used to drive NS1 and NS2 gene expression (Zhu *et al.*, 2012). The ie1 promoter (632 bp) is a strong promoter that drives the expression of an immediate early gene in the baculovirus prototype *Autographa californica multiple nucleopolyhedrovirus* (Fu *et al.*, 2015) , and its activity is stronger than that of the p5 promoter. Then, the expression of GFP and firefly luciferase was assessed qualitatively by fluorescence microscopy and dual luciferase assay in different insect cell lines, respectively. To clarify whether the BmBDV polyA affects the transcription level of foreign genes, real-time qPCR was used to test the relative content of *gfp* mRNA obtained with the BmBDV polyA, compared with the SV40 polyA. Because the more widely used polyA signals have lengths of approximately 250 bp, the development of a much shorter, functional polyA signal that has the potential to increase the vector carrying capacity would be beneficial. Therefore, a dual luciferase reporter gene assay was performed to determine the core functional areas of the BmBDV polyA that might play significant roles in gene expression. Overall, determining the effect of the BmBDV polyA on the expression of foreign genes driven by different promoters in insect cells may provide a theoretical reference for its potential application to other viral expression vector systems.

Materials and Methods

Recombinant plasmid construction

Two different marker genes were used to test the roles of the BmBDV polyA in gene expression levels under the control of the p5 or ie1 promoters. Transfer vectors were constructed to generate many recombinants that express GFP- and firefly luciferase-encoding genes. The primers used for constructing the plasmids are listed in Table 1. To test the effect of the BmBDV polyA on *gfp*

expression levels driven by the p5 or ie1 promoters, we designed two 63-bp long primers (nt 2,896 - 2,958) that are derived from the genes encoding MCP and pPolB. Then, two restriction endonuclease sites (*Mlu*I and *Sal*I) were added at each end, and the small fragment of BDVpA(63) was synthesized by annealing. Subsequently, the small DNA was cleaved with *Mlu*I/*Sal*I to delete a redundant base to produce a sticky end. We had previously constructed two plasmids in our laboratory, pT-ie1-gfp-SV40 and pT-p5-gfp-SV40, each of which contains a *gfp* expression cassette. These two plasmids were digested with *Mlu*I/*Sal*I to delete the SV40 polyA, and they were ligated to the 63-bp *Mlu*I/*Sal*I fragment of BDVpA(63) to generate two new plasmids: pT-ie1-gfp-BDVpA(63) and pT-p5-gfp-BDVpA(63). For the pT-ie1-gfp and pT-p5-gfp constructs, a 720-bp *Bln*I/*Sal*I fragment carrying the *gfp* gene was cloned between the *Bln*I and *Mlu*I sites of pT-ie1-gfp-SV40 (the template was PCR-amplified using gene-specific primers: sense, 5′-CCTAGGATGGTGAGCAAGGG-3′; antisense, 5′-GTCGACTACTTGTACAGCTCGTCCA-3′). The plasmids pT-ie1-gfp-SV40 and pT-p5-gfp-SV40 were digested with *Bln*I/*Sal*I to retrieve pT-ie1 and pT-p5, and then ligated with the 720-bp *Bln*I/*Sal*I fragment to produce the transfer vectors pT-ie1-gfp and pT-p5-gfp. To detect the impact of the BmBDV polyA on firefly luciferase gene expression levels driven by the p5 or ie1 promoters, two small fragments of BDVpA(63) with different restriction endonuclease sites (*Xba*I/*Sal*I and *Xba*I/*Bam*HI) were synthesized by annealing, cleaved with *Xba*I/*Sal*I/*Bam*HI to produce a sticky end, and inserted downstream of the firefly luciferase gene in the long fragments pGL_3-p5-SV40$^-$ and pGL_3-ie1-SV40$^-$, which were retrieved from pGL_3-p5-SV40 and pGL_3-ie1-SV40 that are preserved in our laboratory, to produce two clones carrying the p5 or ie1 promoter upstream of the firefly luciferase gene with a downstream BDVpA(63) [pGL_3-p5-BDVpA(63) and pGL_3-ie1-BDVpA (63)]. To determine BmBDV polyA core functional areas, a series of signal fragments containing characteristic sequences (nt 2,909 - 3,000, 2,909 - 2,958, 2,909 - 2,946, 2,909 - 2,932 and 2,933 - 2,958) were synthesized by annealing, and cleaved with *Xba*I/*Sal*I to form sticky ends. Then, these small fragments were inserted into the long fragment pGL_3-p5-SV40$^-$, which was retrieved from pGL_3-p5-SV40 by digestion with *Xba*I/*Sal*I. Thus, these novel plasmids were named as follows: pGL_3-p5-BDVpA(92), pGL_3-p5-BDVpA(50), pGL_3-p5-BDVpA(38), pGL_3-p5-BDVpA(24) and pGL_3-p5-BDVpA(26). All transfer vectors were confirmed by restriction endonuclease and DNA sequence analyses.

Transfection of fusion plasmids and detection of GFP expression

A qualitative analysis of GFP production driven by the p5 and ie1 promoters in the presence of the BDVpA(63), compared to the SV40 polyA, was conducted as follows: The BmN cells used throughout this investigation were grown at 27 °C in 25-cm^2 flask cultures containing TC-100 (Life Technologies, Carlsbad, CA, USA) supplemented with 10 % fetal bovine serum (FBS; Gibco-BRL, Grand Island, NY, USA) and 1 % antibiotics (penicillin-streptomycin; HyClone Gibco-BRL). Hi5 and Sf21 cells were maintained at 27 °C in Grace's medium (Life Technologies) supplemented with 10% FBS and 1 % antibiotics. The cell confluency was approximately 80 % -95 % at the time of transfection. Six groups were set up in the transfection experiments. They included p5 promoter-regulated plasmids (pT-p5-gfp-SV40 and pT-p5-gfp-BDVpA(63)), ie1 promoter-regulated plasmids (pT-ie1-gfp-SV40 and pT-ie1-gfp-BDVpA(63)) and control plasmids (pT-p5-gfp and pT-ie1-gfp). These fragments were prepared by diagnostic restriction endonuclease digestions prior to transfection (usually 3 μg of DNA is needed to transfect one 25-cm^2 tissue culture flask). Transfections were performed with Cellfectin II Reagent liposomes (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions (as follows). Three micrograms of the experimental group plasmid and 200 μl of serum-free TC-100 medium were mixed. Then, the mixture was incubated for 15 min at room temperature. Then, 10 μl of Cellfectin II Reagent liposomes reagent and 200 μl of serum-free TC-100 medium were added and incubated for 15 min at room temperature. Finally, the DNA/Cellfectin II Reagent liposomes reagent mix was incubated for 30 min at room temperature. Meanwhile, the Sf9, Hi5, and BmN cells were washed three times with serum-free TC-100 culture medium. Then, the DNA/Cellfectin II Reagent liposomes mixture was added dropwise to the 25-cm^2 tissue culture flasks, which were placed in a 27 °C incubator. After 5 h, the medium containing the DNA/Cellfectin II Reagent liposomes reagent mixture was removed, and 10 ml of fresh Grace's complete medium was added to the Hi5 and Sf9 cells culture flasks, while TC-100 complete medium was added to the BmN cells culture flasks. Then, these cells were again placed in a 27 °C incubator. Green fluorescence emitted by GFP was observed by fluorescence microscopy 48 h after the transfections.

Dual luciferase reporter gene assay

To qualitatively analyze the influence of the BmBDV polyA on gene expression, a dual luciferase reporter gene assay was performed. BmN cells were grown at 27 °C in 24-well plates containing TC-100 supplemented with 10 % FBS and 1 % antibiotics. Hi5 cells were maintained at 27 °C in the Grace's medium supplemented with 10 % FBS and 1 % antibiotics. The cell confluency was approximately 80 % - 95 % at the time of transfection. Six groups were co-transfected with *Renilla* luciferase-expressing plasmids in the transfection experiments (the quantity of the firefly luciferase plasmids was 100 times that of the *Renilla* luciferase-expressing plasmids). They included p5 promoter-regulated plasmids [pGL_3-p5-SV40 and pGL_3-p5-BDVpA (63)], ie1 promoter-regulated recombinant plasmids [pGL_3-ie1-SV40 and pGL_3-ie1-BDVpA (63)], as well as control plasmids without a polyA signal (pGL_3-p5 and pGL_3-ie1). These plasmids were prepared by diagnostic restriction endonuclease digestions prior to

transfection (usually 1 μg of DNA is needed to transfect one well of a 24-well plate). Transfections were performed with Cellfectin II Reagent liposomes (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions (as mentioned above in the transfections with the *gfp* fusion plasmids). At 48 h after transfection, the complete medium in 24-well plates was removed. Then, the Hi5 and BmN cells were washed two times with 1×phosphate-buffered saline. Afterwards, 200 μl of passive lysis buffer was added and incubated for 15 min at room temperature to ensure that the cells were fully lysed. The mixture was collected in 1.5-ml Eppendorf tubes for the dual luciferase reporter gene assay. Then, 100 μl of cell lysate was added into an RNase-free Eppendorf tube and mixed with 50 μl of Luciferase Assay Reagent II (Promega, Madison, WI, USA). The value of the firefly luciferase activity, recorded as the *F* value, was measured by a Berthold Detection Systems detector (Berthold GmbH, Pforzheim, Germany). Then, the Eppendorf tube was removed, and 50 μl of Stop & Glo reagent (Promega, Madison, WI, USA) was added to terminate the firefly luciferase reaction. The value of the *Renilla* luciferase activity, recorded as the *R* value, was also determined by the Berthold Detection Systems detector. Finally, the fold activity was assessed in triplicate by a two-tailed Student's *t*-test.

Transcription level analysis

To test the differences in the numbers of *gfp* transcripts between recombinant plasmids with BDVpA(63) and the SV40 polyA, real-time qPCR was used. Total RNA was extracted from the different groups of transfected cells that were used to measure GFP protein expression (described above) at 48 h. The cells were scraped, collected in 50-ml conical tubes, and pelleted by centrifugation for 90 min at approximately 3,000 rpm at 4 °C in a bench-top clinical centrifuge. The supernatant was removed, and the cell pellet was collected. TRIzol reagent (Invitrogen) was used to extract the total RNA, which was confirmed by 1 % agarose gel electrophoresis. The extracted RNA was first quantified by spectrophotometry. Total RNA (500 ng) was treated with RNase-free DNase (Vazyme Biotech, Nanjing, China) to degrade potential DNA contaminants following the conditions recommended by the manufacture. Then, the DNA-free RNAs were used as templates for cDNA synthesis using two reverse primers(Pfaffl, 2001) , using the cDNA Synthesis Kit (Vazyme Biotech). The synthesized cDNA was diluted 100-fold with RNase-free water for a qPCR analysis of *gfp* transcript levels. In the qPCR, the primer pairs used to generate a 146-bp amplicon of the *gfp* gene were GFP-474F and GFP-619R. An internal reference amplicon generated with the primer pair rRNA-F and 28S rRNA-R that targets the *B. mori* 28S rRNA gene was used to normalize the reactions in Hi5 cells*(*Xue *et al*., 2010). Another internal reference amplicon generated with the primer pair *rpl3*-F and *rpl3*-R that targets the *B. mori rpl3* gene was used to normalize the reactions in BmN cells. *gfp* fusion plasmids without a polyA signal served as a control group. SYBR Green Super-mix kits (Vazyme Biotech) were used in the real-time qPCR. The amplification data were acquired by an ABI Prism 7300 system (Applied Biosystems, Foster City, CA, USA). The effects of BDVpA(63) on *gfp* transcript levels were expressed relative to the 28S rRNA (Xue & Cheng, 2010) or *rpl3* RNA levels. Each of the amplifications was run in triplicate to calculate the experimental variance in a statistical analysis by a two-tailed Student's *t*-test.

Identification of BmBDV polyA core functional areas

To determine BmBDV polyA core functional areas, a dual luciferase reporter gene assay was used. Hi5 and BmN cells in 24-well plates were co-transfected with *Renilla* luciferase-expressing plasmids and seven groups of firefly luciferase plasmids, including pGL_3-p5-BDVpA(92), pGL_3-p5-BDVpA(63), pGL_3-p5-BDVpA(50), pGL_3-p5-BDVpA(38), pGL_3-p5-BDVpA(26), pGL_3-p5-BDVpA(24) and the control plasmids without a polyA signal (pGL_3-p5). The co-transfections and dual luciferase reporter gene assay were performed as described above for the transfections with firefly luciferase fusion plasmids. The fold activity was evaluated in triplicate for statistical analysis.

Statistical analysis

Descriptive statistical analyses were performed using GraphPad Prism 5 software (GraphPad Software, Inc.). The data are all from at least three independent biological replicates per experimental variable and presented as means ± standard deviations. The differences between control and experimental groups were assessed by a two-tailed Student's *t*-test and accepted as significant at $p < 0.01$or $p < 0.001$.

Results

Construction of transient expression plasmids

The BmBDV genome contains a small number of genes whose distribution is very compact. Especially, the genes encoding MCP and pPolB overlap partially. The MCP transcript starts at nt 1,423 and ends at nt 2,931, and the pPolB transcript starts at nt 6,287 and ends at nt 2,922 (Wang *et al*., 2007). Thus, the two transcripts have a 10 nts overlap region at their 3′ UTRs. Meanwhile, the stop codon of VD1-ORF3 is located at nt 2,916, and the stop codon of VD1-ORF4 is situated at nt 2,940, which are only 24 nts apart. Thus, the 3′ UTRs containing the overlapping regions with characteristic sequences (an AAUAAA motif and GU-rich regions) were truncated and designated as BDVpA(63) (nt 2,896 - 2,958).

To investigate the effect of the BmBDV polyA on other genes, we constructed recombinant plasmids in which BDVpA(63) was inserted into expression vectors. BDVpA(63) was respectively fused to the 3′ terminal sequence of two reporters (genes encoding GFP and firefly luciferase) under the control of two different promoters, p5 or ie1 (Fig. 1). The modified variants pT-ie1-gfp-SV40 and pT-p5-gfp-SV40 were used as starting model systems. Then, the SV40 polyA was replaced by BDVpA(63) or deleted directly in these two plasmids. Thus, the new vectors were renamed pT-ie1-gfp-BDVpA(63), pT-ie1-gfp, pT-p5-gfp-BDVpA(63) and pT-p5-gfp (Fig. 1A).

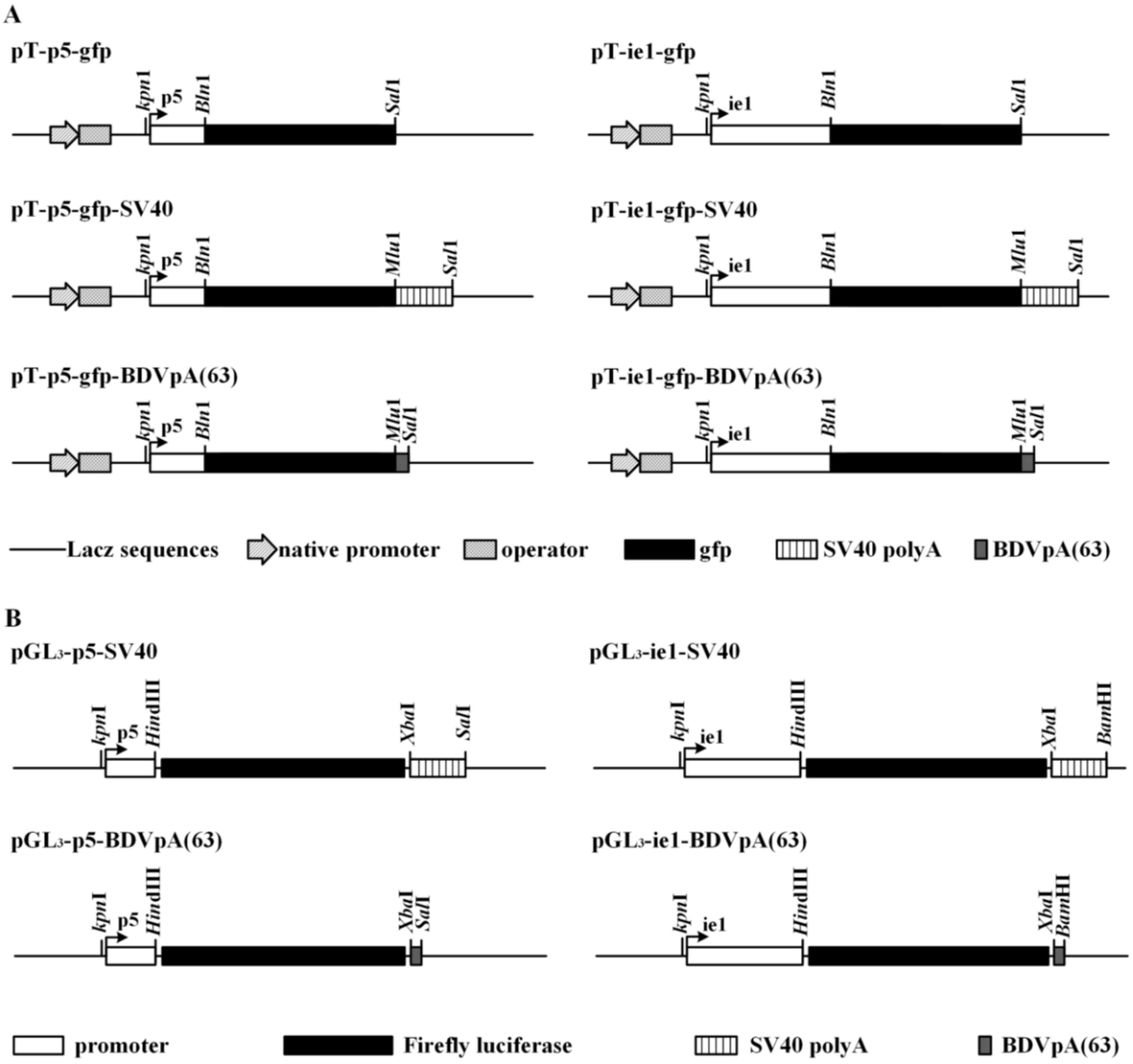

Fig. 1 Schematic diagrams of the constructs expressing two heterologous proteins**.** (A) Recombinant plasmids that were used to express the *gfp* gene under the control of two different promoters, p5 or ie1. (B) Firefly luciferase expression constructs that were used to test the activity of luciferase driven by the p5 or ie1 promoters.

To complement GFP fluorescence measurements, a dual luciferase reporter gene assay was used to visualize the difference in heterologous expression between each pair of the recombinant plasmids. First, the SV40 polyA was eliminated from the recombinant plasmids pGL$_3$-ie1-SV40 and pGL$_3$-p5-SV40, which were constructed previously in our laboratory, to form two novel plasmids named pGL$_3$-ie1-SV40 and pGL$_3$-p5-SV40$^-$, respectively. Then, BDVpA(63) was fused to the 3′ terminal sequence of the firefly luciferase reporter gene. The expression plasmids pGL$_3$-ie1-BDVpA(63) and pGL$_3$-p5-BDVpA(63) were constructed successfully (Fig. 1B).

Effect of the BmBDV polyA on GFP expression

To measure the level of foreign gene expression, various GFP fusion plasmids were respectively transfected into three insect cell lines (Hi5, BmN and Sf9). All the cells were observed at 48 h post-transfection by fluorescence microscopy. The *gfp* gene from recombinant plasmids with BDVpA(63) was expressed in the transfected cells. Compared with the SV40 polyA, more transfected cells produced green fluorescence, and greater fluorescence appeared in each cell (Fig. 2). In addition, the fluorescence intensity of the cells transfected with plasmids controlled by the ie1 promoter was slightly greater than that of plasmids driven by the p5 promoter, which is consistent with the fact that the activity of ie1 is stronger than that of p5. Moreover, almost equivalent fluorescence was displayed in both Hi5 and BmN cells that were transfected with the p5 promoter-based GFP fusion plasmids containing the SV40 polyA or BDVpA(63) (Figs 2A, B, E, F), while no GFP fluorescence was

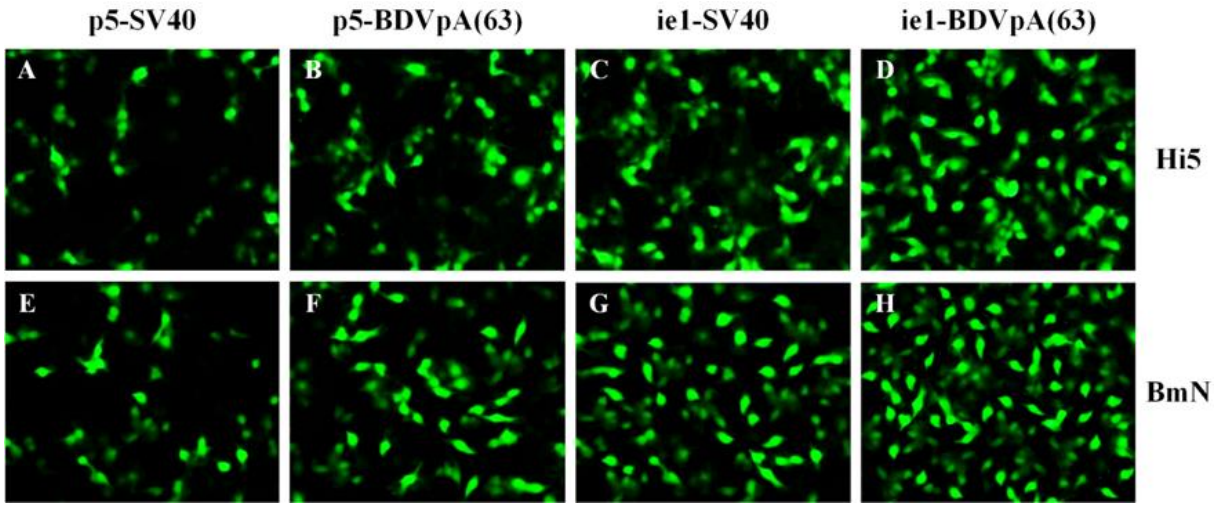

Fig. 2 Fluorescence microscopy images of cells displaying GFP expression at 48 h post-transfection. Hi5 cells were transfected with pT-p5-gfp-SV40 (A), pT-p5- gfp-BDVpA(63) (B), pT-ie1-gfp-SV40 (C), or pT-ie1-gfp-BDVpA(63) (D). BmN cells were transfected with pT-p5-gfp-SV40 (E), pT-p5-gfp-BDVpA(63) (F), pT-ie1-gfp-SV40 (G), or pT-ie1-gfp-BDVpA(63) (H). Cells were photographed at 10× magnification.

detected from the Sf9 cells transfected with either the pT-p5-gfp-SV40 or pT-p5-gfp-BDVpA(63) plasmids, which is the same as the results obtained from transfections with the control plasmid (pT-p5-gfp) (data not shown). The plasmids containing the ie1 promoter with two polyA signals both showed fluorescence in Hi5, BmN (Figs 2C, D, G, H) and Sf9 cells, but the fluorescence intensity in the Hi5 and BmN cells was much stronger than that in the Sf9 cells (data not shown). One of the possibilities is that the activity of BDVpA(63) may be relatively low in the Sf9 cells when the expression of *gfp* is driven by the p5 promoter. This finding laid the foundation for the selection of a cell line in the subsequent dual luciferase reporter gene assay. In brief, all these data suggested that BDVpA(63) in promoter-based vectors was more conducive to the expression of the *gfp* reporter gene, compared with the SV40 polyA, in insect cells.

Effect of the BmBDV polyA on firefly luciferase expression

The differences of firefly luciferase expression from the recombinant plasmids with BDVpA(63) or the SV40 polyA were detected by a dual luciferase reporter gene assay. Luciferase activity was determined 48 h after co-transfection with *Renilla* luciferase-expressing plasmids. The results revealed that the relative expression of firefly luciferase was higher when BDVpA(63) was used (Fig. 3). However, the increased levels of protein expression differed depending on the inserted polyA signal. Compared with the SV40 polyA, the greatest increase was detected from the plasmid pGL_3-ie1-BDVpA(63) in Hi5 cells (approximately 21.7-fold, Fig. 3B), followed by pGL_3-ie1-BDVpA(63) in BmN cells (up to 12.6-fold, Fig. 3D), pGL_3-p5-BDVpA(63) in Hi5 cells (approximately 6.5-fold, Fig. 3A) and pGL_3-p5-BDVpA(63) in BmN cells (approximately 2.7-fold, Fig. 3C). Hence, BDVpA(63) was more beneficial for the expression of the firefly luciferase gene than the SV40 polyA, which is consistent with the fluorescence microscopy results. All these data suggest that BDVpA(63) has the potential to promote the expression of different types of proteins in different insect cell lines.

Analysis of promoter-based gfp *transcription*

Because protein expression was remarkably enhanced in the presence of BDVpA(63), it was imperative to determine its role at the transcriptional level. *gfp* fusion plasmids were respectively transfected into Hi5 and BmN cells, which were subsequently cultured for 48 h. Total RNA was extracted from the transfected cells. The differences between the number of *gfp* transcripts were confirmed by real-time qPCR using the *B. mori* cellular 28S rRNA gene and the *rpl3* gene as an internal reference to normalize the reactions (Fig. 4). *gfp* fusion plasmids without a polyA signal served as a control group. The real-time qPCR results showed that the relative *gfp* mRNA content was higher in both types of insect cells when BDVpA(63) was used (Fig. 4). However, the increased levels of *gfp* transcripts differed depending on the types of promoters and cells. The greatest increase was detected in Hi5 cells when *gfp* expression was driven by the ie1 promoter (4.9-fold, Fig. 4B), followed by the ie1 promoter in BmN cells (3.6-fold, Fig. 4D), the plasmid pT-p5-gfp-BDVpA(63) in Hi5 cells (2.8-fold, Fig. 4A) and the plasmid pT-p5-gfp-BDVpA(63) in BmN cells (1.7-fold, Fig. 4C). Thus, the increased levels of *gfp* transcripts due to the insertion of BDVpA(63) was significantly higher than that of the SV40 polyA. These results also indicated that the integration of BDVpA(63) downstream of the *gfp* gene enhanced the level of gene transcription in insect cells.

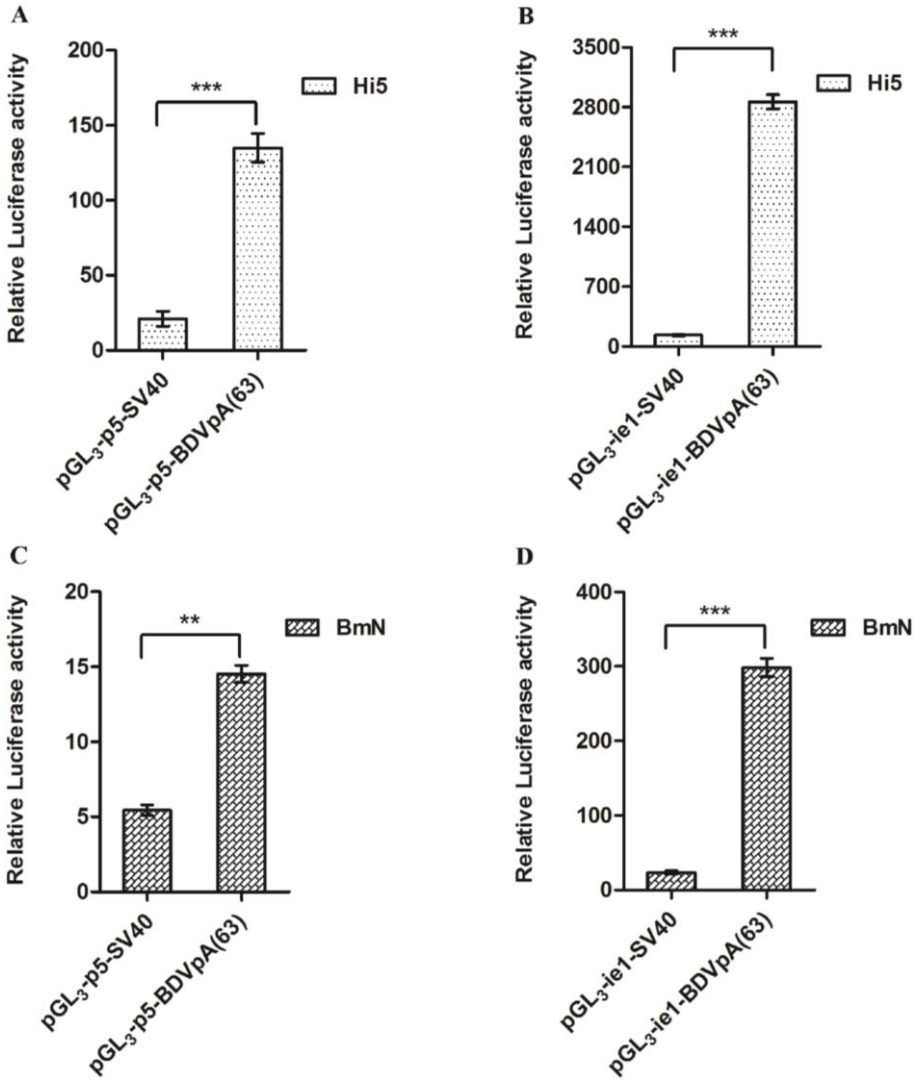

Fig. 3 Quantitative analysis of firefly luciferase expression to test the effects of the BDVpA(63), compared to the SV40 polyA, on protein expression. Hi5 (A, B) and BmN (C, D) cells were co-transfected separately with *Renilla* luciferase-expressing plasmids and the firefly luciferase constructs containing BDVpA(63) or the SV40 polyA, or the construct without a polyA signal. All the cells were harvested at 48 h post co-transfection and subjected to dual luciferase reporter gene assay in three independent cell transfections. The values are expressed as the ratios of firefly luciferase expression relative to those of the control plasmids pGL_3-p5 and pGL_3-ie1. Significance is indicated as ** $p < 0.01$ and *** $p < 0.001$. Error bars denote standard deviations.

Identification of BmBDV polyA core functional areas

To determine BmBDV polyA core functional areas that might play significant roles in gene expression, we performed a dual luciferase reporter gene assay. The characteristic sequences of the BmBDV polyA were truncated, and a series of firefly luciferase plasmids with BmBDV polyAs of different lengths were constructed, Therefore, these novel plasmids were named as follows: pGL_3-p5-BDVpA(92), pGL_3-p5-BDVpA(50), pGL_3-p5-BDVpA(38), pGL_3-p5-BDVpA(26) and pGL_3-p5-BDVpA(24) (Fig. 5A).

These five plasmids and the plasmid pGL_3-p5-BDVpA(63) were co-transfected separately into Hi5 and BmN cells with *Renilla* luciferase-expressing plasmids. Then, these cells were cultured for 48 h to confirm that the firefly luciferase gene was expressed. The discrepancies among the fold activity of the different forms of the BmBDV polyA were measured. The relative firefly luciferase activities were high in Hi5 (Fig. 5B) and BmN (Fig. 5C) cells that were transfected with plasmids containing BDVpA(92), BDVpA(63), BDVpA(50) and BDVpA(38), which all contained an AAUAAA motif and GU-rich regions. Interestingly, the highest relative firefly luciferase activity was observed in cells transfected with pGL_3-p5-BDVpA(50) (Figs 5B, C). Nevertheless, the relative firefly luciferase activities of cells transfected

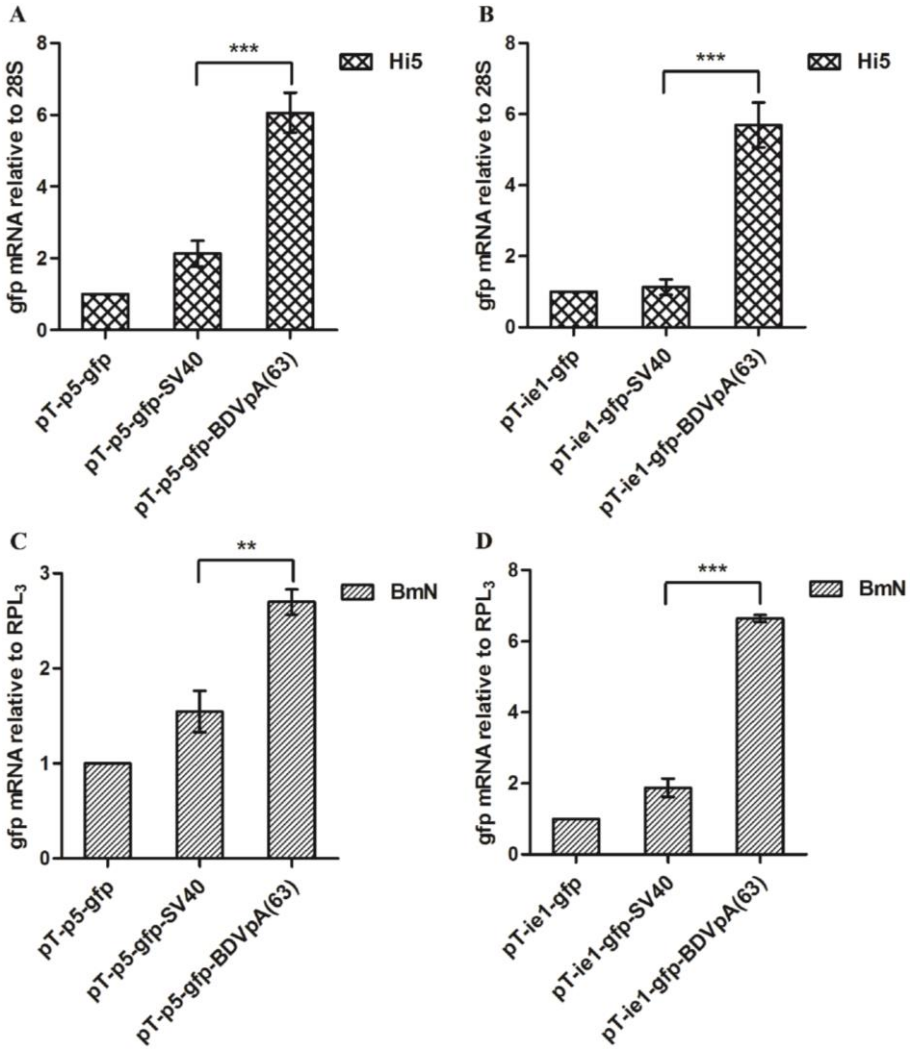

Fig. 4 Transcription analysis of *gfp* from different transfections**.** The constructs with BDVpA(63), the SV40 polyA, or those without a polyA signal were used to transfect Hi5 (A, B) and BmN (C, D) cells, and the transfected cells were harvested at 48 h for total RNA extraction. Equal amounts of RNA from different cell transfections were used as templates for amplification of a 146-bp amplicon of the *gfp* gene using the SYBR Green Dye RT-PCR Kit from Vazyme Biotech (Nanjing, China) for the real-time quantification of the *gfp* transcripts. The *B. mori* cellular 28S rRNA gene and *rpl3* gene were used as housekeeping genes in the real-time qPCR to normalize the reaction. Significant differences are indicated as $^{**}p < 0.01$ and $^{***}p < 0.001$. Error bars represent standard deviations from three independent cell transfections.

with the plasmids pGL_3-p5-BDVpA(24) and pGL_3-p5-BDVpA(26), which contained incomplete BDVpA, in which the AAUAAA motif and GU-rich regions that function independently, were reduced in both Hi5 and BmN cells (Figs 5B, C). Hence, BDVpA(50) comprising the AAUAAA motif, GU-rich regions, and 12 nts downstream was identified as the core functional region. It is noteworthy that a stop codon is contained within the AAUAAA motif of BDVpA(50). Thus, this core functional region acts as a combination stop codon and AAUAAA motif (polyA signal/stop codon) (Fig. 5A).

In summation, we hypothesize that the length of the BmBDV polyA may affect its activity. Meanwhile, both the AAUAAA motif and the GU-rich regions are crucial for the formation of polyadenylation, the subsequent processing of downstream sequences, as well as the efficiency of protein synthesis.

Discussion

Since increasing the expression of heterologous proteins in insect cells has research and pharmaceutical applications (Tani *et al.*, 2008; Kato *et al*., 2010), the SV40 polyA has been widely used in many commercial baculovirus expression vectors. In this study, we generated a novel, short, dual-functional polyA signal/stop codon from an insect virus. Our most intriguing discovery is that the BmBDV polyA promotes gene expression at both the

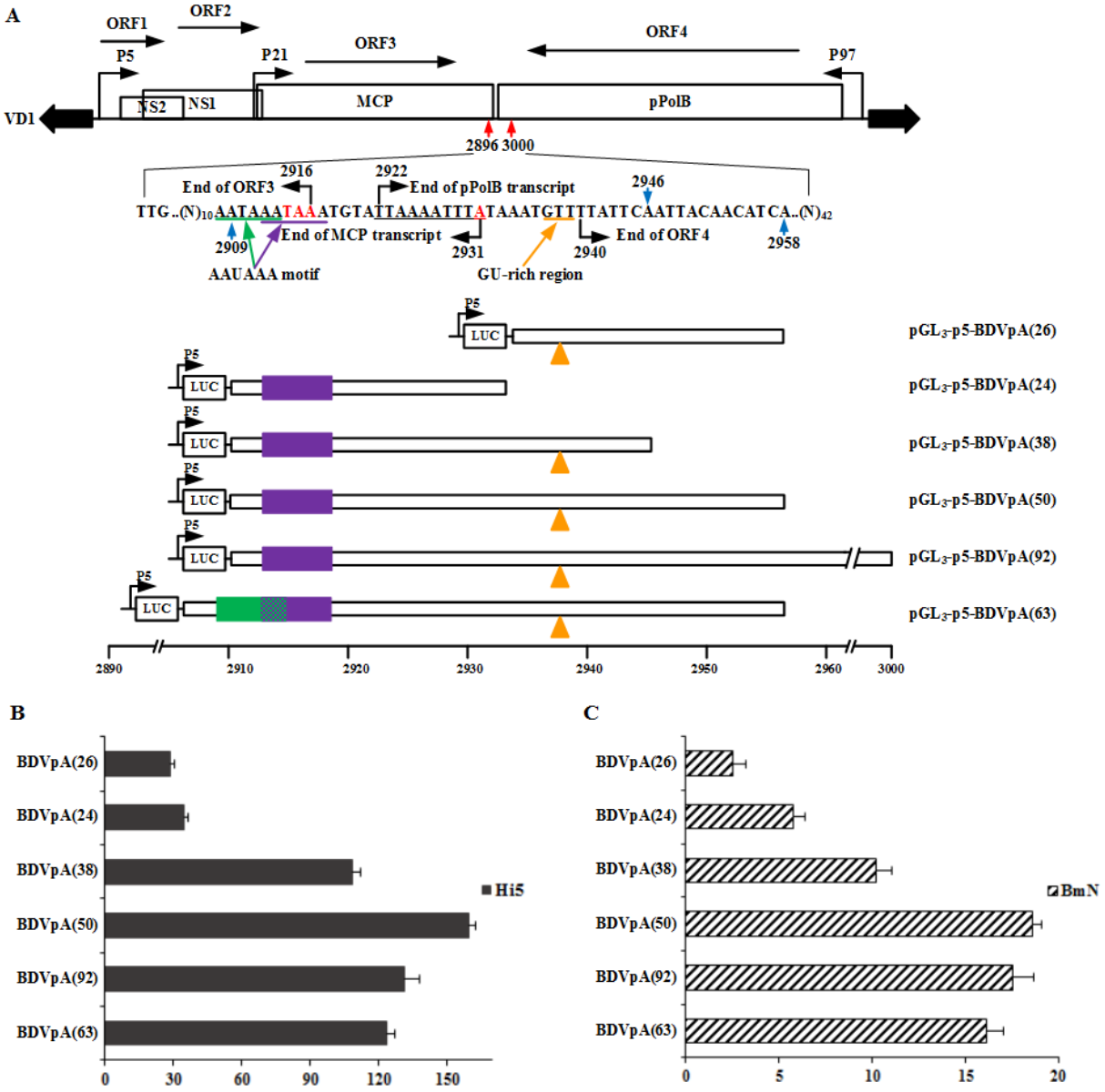

Fig. 5 The identification of BmBDV polyA core functional areas using dual luciferase reporter gene assay**.** (A) Schematic diagram of the constructs containing a series of BDVpA with different lengths. (B) Detection of firefly luciferase activity from transfected Hi5 cells. Hi5 cells were con-transfected with *Renilla* luciferase-expressing plasmids and the firefly luciferase constructs with a series of polyadenylation signal fragments of different lengths. All the cells were harvested at 48 h post co-transfection and subjected to a dual luciferase reporter gene assay in three independent cell transfections. (C) Measurement of the firefly luciferase activity in transfected BmN cells. BmN cells were co-transfected as described above. The values are expressed as the ratio of firefly luciferase activity relative to that of the control plasmid pGL_3-p5.The AAUAAA motif is denoted by purple and green boxes, and the GU-rich region is indicated by a yellow triangle. //indicates sequences omitted. Error bars signify standard deviations from three independent cell transfections.

mRNA and protein levels in diverse cell lines, which distinguishes it from the SV40 polyA, which increased mRNA levels, but decreased protein production in the BEVS (Salem *et al.*, 2015). These results also differ from an early report showing that the SV40 polyA driven by the strong promoter p10 from *Autographa californica multiple nucleopolyhedrovirus* reduces mRNA levels and protein synthesis (van Oers *et al*., 1999). Therefore, the new, dual-functional BmBDV polyA is a valuable for increasing mRNA production and protein synthesis in insect cells.

Similar to the SV40 polyA, the BmBDV polyA is situated at the regions of genes in two complementary strands, and it contains at least two elements: an AAUAAA sequence for CPSF binding and a downstream GU-rich region for CstF binding. One interesting observation in this study is that intact

BmBDV polyAs of different lengths (BDVpA(92), BDVpA(63), BDVpA(50) and BDVpA(38)) resulted in different protein synthesis efficiencies. Especially, the fold activity of BDVpA(50), comprising the AAUAAA motif, GU-rich regions, and 12 downstream nt, was the highest among all the polyA signals. A comparison of the four intact BmBDV polyAs showed that BDVpA(63) has two partial overlapping AAUAAA motifs. The second AAUAAA motif is followed by a dinucleotide with a guanosine (G), whereas no G nucleotide is present in the dinucleotide right after the first AAUAAA motif. The other three BmBDV polyA signals (BDVpA(92), BDVpA(50) and BDVpA(38)) only have the second AAUAAA motif. These data also indicate that additional sequences other than the canonical AAUAAA motif are needed for the efficient processing of the 3′ end of promoter-controlled transcripts during insect cell transfections. Additionally, the evidence obtained from this study confirms that the lack of a G in the dinucleotide after the AAUAAA motif of 3′ UTRs might be responsible for the weak 3′ end processing (Salem *et al*., 2015). Because of diverse truncated fragments right after GU-rich region, BDVpA(50), BDVpA(92) and BDVpA(38) resulted in different firefly luciferase activities. Nevertheless, the exact reason for this remains unclear. One plausible explanation could be that sequences downstream of the GU-rich region have different secondary structures, thereby influencing the binding of cellular RNA processing factors to the polyA signal of transcripts and interfering with translation initiation, which is required for efficient protein synthesis.

When the BmBDV polyA was inserted downstream of GFP- and firefly luciferase-encoding genes under the control of two different promoters, increased transcript levels and protein synthesis were both detected, compared with the SV40 polyA. Hence, the incorporation of the BmBDV polyA into expression plasmids is very useful. It is unknown whether it has the same effect in insect transgenic vectors and BEVSs. As Sf9 cells are a high transfection efficiency line, these cells are commonly used for the production of recombinant proteins. In the measurement of GFP expression, the fluorescence intensity in the Hi5 and BmN cells was much stronger than that in the Sf9 cells, the reason of which may be related to the specificity of BmBDV polyA for cell lines. It is also unknown whether the increased mRNA and protein levels can be achieved in mammalian systems. Therefore, it is interesting to consider the detailed aspects that could affect the use of this new polyA signal in different cell lines in future research. Additionally, this short polyA signal may overcome the obstacle of improving the carrying capacity of plasmid or viral vectors, thereby making it an alternative to the use of small promoters and intramolecular recombination, which requires two separate virions to infect the same cell (Flotte, 2000; Shevtsova *et al*., 2005). In summation, the BmBDV polyA has the potential to be used for the effective and stable expression of heterologous RNAs and proteins in insect cells. Furthermore, our study provides a reference for constructing plasmids with a small insertion, as well as for researching the function of other polyA signals.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant Nos. 31570150 and 31272507) and the Jiangsu Provincial Natural Science Foundation (Youth Project, grant No. BK20160507).

References

Adams MJ, Carstens EB. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2012). Arch Virol. 157: 1411-1422, 2012.

Beilharz TH, Preiss T. Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. RNA 13: 982-997, 2007.

Flotte TR. Size does matter: overcoming the adeno-associated virus packaging limit. Respir. Res. 1: 16-18, 2000.

Fu Y, Li X, Du J, Zheng, S, Liang A. Regulation analysis of AcMNPV-mediated expression of a Chinese scorpion neurotoxin under the IE1, P10 and PH promoter in vivo and its use as a potential bio-insecticide. Biotechnol. Lett. 37: 1929-1936, 2015.

Hu ZY, Li GH, Li GT, Yao Q, Chen KP. *Bombyx mori* bidensovirus:The type species of the new genus Bidensovirus in the new family Bidnaviridae. Sci. Bull. 58: 4528-4532, 2013.

Jin J, Guarino LA. 3'-end formation of baculovirus late RNAs. J. Virol. 74: 8930-8937, 2000.

Kato T, Kajikawa M, Maenaka K, Park EY. Silkworm expression system as a platform technology in life science. Appl. Microbiol. Biotechnol. 85: 459-470, 2010.

Keller W, Bienroth S, Lang KM, Christofori G. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. EMBO J. 10: 4241-4249, 1991.

Krupovic M, Koonin EV. Evolution of eukaryotic single-stranded DNA viruses of the Bidnaviridae family from genes of four other groups of widely different viruses. Sci Rep. 4: 5347, 2014.

Li G, Sun C, Zhang J, He Y, Chen H, Kong J, *et al*. Characterization of *Bombyx mori* parvo-like virus non-structural protein NS1. Virus Genes 39: 396-402, 2009.

Moore CL, Sharp PA. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell 41: 845-855, 1985.

Murthy KG, Manley JL. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. Genes Dev. 9: 2672-2683, 1995.

Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29(9):e45, 2001.

Salem TZ, Seaborn CP, Turney CM, Xue J, Shang H, Cheng XW. The Influence of SV40 polyA on Gene Expression of Baculovirus Expression Vector Systems. PLOS ONE 10(12), e0145019, 2015.

Shevtsova Z, Malik JM, Michel U, Bahr M, Kugler S. Promoters and serotypes: targeting of adeno-associated virus vectors for gene transfer in the rat central nervous system in vitro and in vivo. Exp. Physiol. 90: 53-59, 2005.

Tamura T, Thibert C, Royer C, Kanda T, Abraham E, Kamba M, *et al*. Germline transformation of the silkworm *Bombyx mori* L. using a piggyBac transposon-derived vector. Nat. Biotechnol. 18: 81-84, 2000.

Tani H, Abe T, Matsunaga TM, Moriishi K, Matsuura Y. Baculovirus vector for gene delivery and vaccine development. Future Virol. 3: 35-43, 2008.

van Oers MM, Vlak JM, Voorma HO,Thomas AA. Role of the 3' untranslated region of baculovirus p10 mRNA in high-level expression of foreign genes. J. Gen. Virol. 80 (Pt 8): 2253-2262, 1999.

Venkataraman K, Brown KM,Gilmartin GM. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. Genes Dev. 19: 1315-1327, 2005.

Wang F, Hu Z, He Y, Li G, Kong J, Cao J, *et al*. The non-structural protein NS-2 of *Bombyx mori* parvo-like virus is localized to the nuclear membrane. Curr. Microbiol. 63: 8-15, 2011.

Wang YJ, Yao Q, Chen KP, Wang Y, Lu J, Han, X. Characterization of the genome structure of *Bombyx mori* densovirus (China isolate). Virus Genes 35: 103-108, 2007.

Westwood JA, Jones IM, Bishop DH. Analyses of alternative poly(A) signals for use in baculovirus expression vectors. Virology 195: 90-99, 1993.

Wickens M. How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. 15: 277-281, 1990.

Xue JL, Cheng XW. Using host 28S ribosomal RNA as a housekeeping gene for quantitative real-time reverse transcription-PCR (qRT-PCR) in virus-infected animal cells. Curr. Protoc. Microbiol. Chapter 1, Unit1D.2., 2010.

Xue JL, Salem TZ, Turney CM, Cheng XW. Strategy of the use of 28S rRNA as a housekeeping gene in real-time quantitative PCR analysis of gene transcription in insect cells infected by viruses. J. Virol. Methods 163: 210-215, 2010.

Zhang PP, Miao D, Zhang YH, Wang MZ, Hu Z. Lu P, *et al.* Cloning and rescue of the genome of *Bombyx mori* bidensovirus, and characterization of a recombinant virus. Virol. J. 13: 126, 2016.

Zhu S, Li G, Hu Z, Chen K, Li G, Guo X, *et al*. Characterization of the promoter elements of *Bombyx mori* bidensovirus nonstructural gene 1. Curr. Microbiol. 65: 643-648, 2012.