Elucidation of the upstream gene regulatory network to activate the Polyhedrin promoter in Bombyx mori nucleopolyhedrovirus

（カイコ核多角体ウイルスのポリヘドリンプロモーター活性化上流遺伝子の制御ネットワークの解明）

Unit of Applied Molecular Biology, Frontiers in Biosciences,

Graduate School of Agriculture, Hokkaido University

Toshiki Nakanishi

Contents

Preface

Gene expression in Bombyx mori nucleopolyhedrovirus (BmNPV) is dynamically regulated throughout the course of infection. The BmNPV genome encodes 143 genes of which expression has been annotated as immediate early, delayed early, late, and very late genes. Nucleopolyhedrovirses express the structural protein Polyhedrin at extremely high levels. Although molecular mechanism for the regulation of the Polyhedrin expression has been studied, the system-level understanding of the activating the *polyhedrin* promoter remains poorly understood. The first chapter describes development of a method to quantify time-resolved gene expression. The second chapter describes network modeling of the upstream gene regulatory network required for activating the *polyhedrin* promoter.

Chapter 1: Clinically approved chemical-controlled suppression of protein expression in BmN Cells

Abstract

Rapid control of protein abundance is crucial in both basic and applied research. Small molecule-controlled protein expression is promising tools for such rapid control. However, concerns arise regarding the residual chemicals in cells treated with these molecules, especially when recombinant proteins are intended for medical purposes. The small molecule-assisted shut-off (SMASh) utilizes inhibitors against hepatitis C virus (HCV) nonstructural protein 3 protease, enabling use of the Pharmaceuticals and Medical Devices Agency (PMDA)-approved, less concerning chemicals for the target protein control. In this study, I developed SMASh-tagged enhanced green fluorescent protein (EGFP) reporters to evaluate the efficacy of SMASh and confirmed its functionality in the Bombyx mori-derived ovary cell line, BmN. Upon comparing degron efficiencies, it was revealed that the HCV-derived degron in the SMASh tag was more effective at degrading the fused EGFP than the PEST sequence of mouse ornithine decarboxylase commonly used as a signal peptide for degradation. The activity of the SMASh tag was quantitatively tunable within the range of 10^{-3} to 1 μM of Asunaprevir in BmN cells. Higher concentrations of Asunaprevir, such as 5 μM, were toxic to the cells. It was also revealed the irreversibility of SMASh-mediated gene suppression in BmN cells. Our findings pave the way for achieving rapid and robust suppression of target protein expression with PMDA-approved drugs in silkworm biology.

Introduction

The manipulation of gene expression levels plays a pivotal role in basic biological research and the technical development of biotechnology. It enables the elucidation of gene function, unraveling regulatory pathways, and modifying cells and organisms for various applications. While gene expression can be controlled at the RNA or protein levels, transcriptional regulations such as controlling transcription rates using a transcriptional regulator(s) (Gossen and Bujard, 1992) or amount of mRNA using RNA interference (RNAi; Fire *et al*., 1998) often result in a delayed observable impact on function and phenotype. Systems that facilitate rapid modulation of gene expression are particularly valuable, as they can overcome the limitations associated with the inducible transcription or RNAi, which often suffer from delayed kinetics and sustained effects due to the durability of protein products (Weiss *et al*., 2007). For example, when the manipulation of gene expression gradually occurs, it is challenging to distinguish the primary and associated effects of the manipulation. Directly targeting protein levels for gene expression control offers more immediate approach, as protein abundance and activity of proteins are more closely linked to its functions than mRNA levels. Although recent advances have shed light on the determinants of translational efficiencies (Liu *et al*., 2021), the molecular mechanisms governing protein degradation are more thoroughly understood and readily applicable.

Control of protein degradation using small molecules offers quick manipulation of gene expression and the precision required for the detailed study of gene function, pathways, and higher regulatory organizations of biological processes. Additionally, the addition and subsequent removal of small molecules allow for flexible suppression and de-suppression of expression of the target protein (Banaszynski *et al*., 2006), and may be critical for proteins which expression must be tightly regulated, such as those toxics to the cells. In the field of molecular entomology, the chemical suppression of gene expression remains underdeveloped. The auxin-inducible degron (AID) system is the sole example of post-translational regulation applied to D. melanogaster, enabling auxin analog-induced degradation of target gene (Trost *et al*., 2016; Nishimura *et al*., 2009). This necessitates the introduction of the Skp, Cullin, F-box containing complex(SCF) E3 ubiquitin ligase together with the degron-tagged gene into the host organisms. Small molecule-mediated post-translational regulation of proteins by one component transgene was pioneered by Banaszynski *et al*. (2006) and developed using mammalian cells. Their system, which involves a destabilizing domain (DD) from human FKBP12 and the molecule Shield1, allows for the controlled degradation and stabilization of proteins with the absence and presence of the small molecule, respectively. The destabilized protein with DD is degraded by the ubiquitin proteasome system while supplement of Shield1 blocks the degradation by binding to DD. This principle is inverted in the ligand-induced target protein degradation systems such as AID, dTAG, and small molecule-assisted shut-off

(SMASh) (Yesbolatova *et al*., 2019). The degrons are targeted by the ubiquitin proteasome systems upon exposure to small molecules. AID, as described above, uses auxin-mediated protein degradation system (Nishimura *et al*., 2009). dTAG employs a FKBP12-fused protein, like the DD and Shield1 system, but the ligand bridges the target protein with cereblon, a component of the cullin-RING E3 ligase and leads the target protein for degradation (Winter *et al*., 2015). The SMASh tag is composed of the hepatitis C virus (HCV) nonstructural protein (NS) 3 protease domain, partial NS3 helicase, NS4A, and a cis-cleavage site (Chung *et al*., 2015). The partial NS3 helicase and NS4A-derived sequence functions as a degron. NS3 protease, the degron, and the NS3 cis-cleavage site are fused with a target protein for degradation controlled by the NS3 protease inhibitors. Proteins fused with the SMASh tag can be expressed unless the cleavage between the target protein and the SMASh tag is interrupted by an NS3 protease inhibitor.

Despite the progress in mammalian models, the implementation of such dynamic gene expression control systems in silkworms has been limited. Since silkworm has been studied as a host for recombinant protein production, developing a method to control gene expression in a manner compatible with downstream processes of recombinant proteins is crucial. This is particularly critical to produce recombinant proteins for medical purposes, where residual chemicals may pose concerns. Among the methods of post translationally controlling protein abundance, the SMASh system offers an advantage in post-translationally controlling protein abundance by utilizing the Pharmaceuticals and Medical Devices Agency (PMDA)-approved HCV NS3 protease inhibitors. This study explores the use of the SMASh tag and Asunaprevir, a PMDA-approved drug, for gene expression regulation in silkworm cells. The objective of this study is to assess and characterize the SMASh tag and Asunaprevir in BmN cells for silkworm research.

Materials and Methods

1. Cell line

BmN cells derived from silkworm (Bombyx mori) ovary were cultured at 26℃ in TC-100 medium (Applichem) supplemented with 10% Fetal Bovine Serum (FBS; Biosera) (see Appendix 1, Table 1).

2. Chemicals

Asunaprevir (BMS-650032) has been reported in 2012 (Pelosi *et al*., 2012) as an inhibitor of NS3 protease of the hepatitis C virus (Pelosi *et al*., 2012) and was approved by the PMDA in 2014. Asunaprevir (CAS:630420-16-5; Chemscene) was dissolved in dimetyl sulfoxide (DMSO) and used as indicated concentrations.

3. Basic Procedures for Genetic Experiments

i Preparation of Chemically Competent Cells

Chemical competent cells were made from *E. coli* DH5α strain. Colonies of *E. coli* were inoculated into 1.5 mL of LB broth (see Appendix 1, Table 2) in test tubes and shaken at 37 ℃ for 16 hours. Subsequently, the culture was added to 100 mL of LB broth and shaken at 37℃ until the OD600 reached 0.3-0.4. After rapid cooling in ice, the culture was divided into two Falcon tubes and subjected to centrifugation (3,000 rpm, 15 minutes, 4℃). The supernatant was partially removed (approximately 1 mL), and the bacterial pellet was gently resuspended. Tfb I (see Appendix 1, Table 3) was added (20 mL) and incubated for 10 minutes, followed by centrifugation (3,000 rpm, 15 minutes, 4℃). The supernatant was completely removed, and Tfb II (see Appendix 1, Table 4) was added (2 mL) to resuspend the bacterial pellet. The suspension was then transferred to pre-cooled 1.5 mL Eppendorf tubes (100 μL per tube) and stored at -80℃.

ii Plasmid DNA Extraction

Plasmid DNA was extracted from *E. coli* using the QIAprep Spin Miniprep Kit (Qiagen) according to the product manual. *E. coli* colonies were inoculated into 1.5 mL of LB broth with added antibiotics (Ampicillin, 150 μg/mL) and cultured at 37℃ for 16 hours with shaking. The culture was centrifuged (3,000 rpm, 3 minutes, 4℃) to remove the supernatant. 200μL of P1 buffer (Qiagen) with RNase A was added to the bacterial pellet, and mixed well using a vortex. Then, 200

μL of P2 buffer (Qiagen) was added, gently inverted, and stand at room temperature for 3 minutes. Subsequently, 300 μL of N3 buffer (Qiagen) was added, gently mixed, and then left in ice for 3 minutes before centrifugation (15,000 rpm, 10 minutes, 4℃). The supernatant was collected, loaded into the kit's column, and centrifuged (15,000 rpm, 1 minute, 4℃). PB buffer (Qiagen) and PE Buffer (Qiagen) were sequentially passed through the column, and centrifugation (15,000 rpm, 1 minute, 4℃) was performed to remove residual buffer. Finally, 50 μL of TE buffer (see Appendix 1, Table 5) was added to the column, left to stand for 1 minute, and centrifuged (10,000 rpm, 1 minute, room temperature) to recover the plasmid DNA.

iii Restriction Enzyme Treatment

Restriction enzyme treatment involved adding 1 μL of 10× Cut Smart Buffer (NEB) and 1 μL of the restriction enzyme to the solution containing DNA (less than 1 μg). The mixture was adjusted to a total volume of 10 μL with sterile distilled water and treated at 37℃ for 2 hours. When using I-SceI for restriction enzyme treatment, the treatment was carried out at 37℃ for 16 hours.

iv PCR

The primers used in this experiment are listed in Appendix 2. The reactions were carried out using PrimeSTAR HS DNA Polymerase (TaKaRa), with an extension time set at 1 minute per kilobase as the standard. After the reaction, the solution was purified by ethanol precipitation, and gel extraction was performed in the case when non-specific bands during electrophoresis was detected.

v SLiCE

SLiCE solution was prepared following the method by Motohashi (2015) using *E. coli* DH5α strain. *E. coli* was inoculated into 1.0 mL of LB broth and cultured with shaking at 37℃ and 200 rpm for 3 hours. Afterward, the culture was added to 50 mL of 2×TY broth (see Appendix 1, Table 6) and further cultured with shaking at 37℃ and 200 rpm until the OD600 reached 1.8 (approximately 5 hours). The culture was then transferred to a 50 mL Falcon tube and centrifuged (5,000×g, 10 minutes, 4℃) to collect the pellet. The collected pellet was washed with 50 mL of sterile water kept on ice, and centrifuged again (5,000×g, 5 minutes, 4℃). The pellet was dissolved in 1.5 mL of CelLytic B Cell Lysis Reagent (Sigma, B7435) and transferred to an Eppendorf tube. After incubating at room temperature for 10 minutes, the supernatant was transferred in 50 μl aliquots to

0.2 mL PCR tubes, mixed with an equal volume of 80% glycerol, and stored at -80℃. The SLiCE reaction was performed by adding 1 μl of the prepared SLiCE solution and 1 μl of 10× SLiCE buffer (see Appendix 1, Table 7) to a solution containing 20 ng of vector DNA and 60 ng of insert DNA. The solution was adjusted to 10 μl with sterile water, followed by incubation at 37℃ for 20 minutes. When performing transformation of *E. coli*, 1 μl of the reaction mixture was added to 100 μl of chemical competent DH5α cells, incubated on ice for 20 minutes, subjected to a 42℃ heat shock for 60 seconds, allowed to stand on ice for 2 minutes, and then 1 mL of SOC medium (see Appendix 1, Table 8) was added. The culture was shaken at 250 rpm for 1 hour at 37℃. After centrifugation (3,000 rpm, 3 minutes, 4℃), an appropriate amount of supernatant was removed, and the remaining supernatant was used to resuspend the *E. coli* cells. Each bacterial strain was seeded onto LB plates (see Appendix 1, Table 9) containing appropriate antibiotics based on the drug selection specific to each vector to form colonies.

vi Sequencing Reaction

Sequencing reactions were performed using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). The reaction mixture included 100 ng of template DNA, 10 pmol of oligonucleotide primer, 2 μL of 5× sequencing buffer, and 0.25 μL of premix, made up to a total volume of 10 μL with sterile distilled water. The reaction was carried out using the following cycling conditions: [95℃, 2 minutes / 98℃, 30 seconds / 96℃, 1 minute] × 1 cycle, [96℃, 10 seconds / 47℃, 40 seconds / 60℃, 4 minutes] × 8 cycles, [94℃, 5 minutes / 50℃, 30 seconds / 60℃, 4 minutes] × 10 cycles, [98℃, 19 seconds / 50℃, 20 seconds / 60℃, 4 minutes] × 7 cycles, and [60℃, 5 minutes] × 1 cycle. The reaction mixture was ethanol-precipitated, the pellet was dissolved in 20 μL of Hi-Di Formamide (Applied Biosystems), and the DNA suspension was heated at 95℃ for 2 minutes to denature the DNA. The mixture was rapidly cooled on ice, and the sample was analyzed using the ABI3130 sequence analyzer (ThermoFisherScientific). I used the DNA Sequencing Facility of the Graduate School of Agriculture.

4. Construction of SMASh reporter plasmids

The SMASh reporter plasmids to express tagged enhanced green fluorescent protein (EGFP; Cormack *et al*., 1996), pOpIE2p:EGFP::SMASh, pOpIE2p:EGFP::CS::FLAG, pOpIE2p:EGFP::CS::NS3pro::PEST::NS4Aβ, pOpIE2p:EGFP::CS::NS3pro::NS4Aβ::PEST, and NS3-cleavage site-mutated pOpIE2p:EGFP::SMASh^{RP}, were constructed with Seamless Ligation Cloning Extract (SLiCE; Motohashi, 2015) of PCR-amplified fragments amplified using the primers

listed in Table 10 (Fig. 1). The parental plasmids, pCS6-YFP-SMASh and pOpIE2p:EGFP::Luc were kindly provided from Dr. Michael Lin (Addgene plasmid # 68853 ; http://n2t.net/addgene:68853 ; RRID:Addgene_68853) and Dr. Takahiro Kusakabe, respectively. The NS4Aβstrand and PEST sequence used as a PCR template was synthesized by Eurofins Genomics. pOpIE2p:EGFP::SMASh^{RP} was constructed from pOPIE2p::EGFP::SMASh. The PCR reactions were performed using PrimeSTAR HS DNA Polymerase (TaKaRa). The sequences of the PCR-amplified regions in these plasmids were validated by sanger sequencing with ABI3130 sequence analyzer (Thermo Fisher Scientific) and the primers 15, 16, 17, and 18 (Table 10).

5. Transfection

The plasmids used in this study were transfected to BmN with FuGene HD (Promega) according to the manufacturer's instructions. Briefly, BmN cells were washed twice with serum-free TC-100 medium and seeded in a microplate (Iwaki) at a density of 5×10^4 cells per well and incubated at 26℃ for 2 hours. For transfection, 1 μg of DNA was mixed with TC-100 medium and transfection reagent to a final volume of 100 μL. This mixture was incubated at room temperature for 20 minutes, then added to cells in a microplate, and incubated for 6 hours at 26 ℃. Following this, the medium was replaced with 100 μL of TC-100 medium supplemented with 10 % FBS. The 'time point 0' for time course experiments was set to the completion of this procedure.

6. Fluorescence microscopy

The EGFP fluorescence and cell morphology of transfected and Asunaprevir-treated cells were examined using a MZ FL III T-HLW stereo fluorescence microscope equipped with a fluorescence filter GFP Plus (Leica), or TCS SP5 confocal microscope (Leica).

7. Quantitative measurement of fluorescence intensity

EGFP fluorescence intensity of transfected cells was measured using an infinite M200PRO (Tecan) at an excitation wavelength of 489 nm and a fluorescence wavelength of 520 nm. Intensity values were obtained at four points in each well, and their average value was used as the fluorescence intensity value for subsequent statistical analyses. To capture both weak fluorescence and the full dynamic range, fluorescence intensities were measured at five different gain levels of the excitation light. These intensities per well then were statistically integrated across the highest gain

and a low gain where no saturation occurred in any wells. This analysis was performed using a regression-based method (Dudley *et al*.,2002) implemented in a custom R script.

8. Western blotting

BmN cells were seeded at 2.0×10^5 cells per well in a 24-well plate and transfected with 0.4 μg of plasmid DNA. After 135 hours, cells were washed with phosphate-buffered saline and lysed in deionized water using a Sonifer II model 450 (Branson) while cooling on ice for 2 minutes (Duty Cycle 70%, Output control 4.5). Protein concentrations were determined by the Bradford method (Bradford, 1976). Equal amounts of protein (1 μg) were separated on a 10% SDS-PAGE and transferred to PVDF membranes using a semi-dry method. Membranes were blocked overnight with 3% N-Z amine (Sigma-Aldrich) in TBST at room temperature and incubated with the anti-GFP antibody ab290 (Abcam). After washing, membranes were incubated with the HRP-conjugated secondary antibody ab6721 (Abcam). Detection was performed using a chemiluminescent substrate and signals were captured with a luminescence imaging analyzer (ATTO).

9. Statistical analysis Dose-response curve analysis

The model fitting to estimate the dose-response curve was performed using R (version 3.5.1; R Core Team, 2003) and the drm package (version 3.0-1). The slope of GFP fluorescence profiles was estimated by fitting the following mixed effect linear model to the measurements:

$$F_{\mathrm{ijkl}} \sim -1+T_i : T_r : C_{\mathrm{k}} : W_{\mathrm{l}} + (R/Wl) + \varepsilon_{\mathrm{ijkl}}$$

where F, Ti, Tr, C, W, R, Wl and $\varepsilon_{\mathrm{ijkl}}$ indicate log2-transformed GFP fluorescence, time point, type of treatments, Asunaprevir concentration, washing cells, replicate, wells, and residuals, respectively. Ti, Tr, C, and W are fixed effects, and R, and Wl, and $\varepsilon_{\mathrm{ijkl}}$ are random effects. This model fitting was performed using the lme function in the nlme package (Pinheiro and Bates, 2000).

Results

HCV NS3 and NS4-derived degron function in BmN cells

The effectiveness of SMASh in mammals and yeast inspired us to explore its functionality in silkworms using the Bombyx mori-derived cell line BmN. The principle of the SMASh system is that the target protein fused with the SMASh tag can be degraded upon the cleavage between the target protein and the SMASh tag is inhibited by the drugs against an NS3 protease (Fig. 2A). In order to test functionality of the SMASh tag in BmN cells, reporter plasmids with FLAG epitope and/or SMASh tag fused to EGFP, under the control of the OpIE2 promoter (Theilmann and Stewart, 1992) and SV40 terminator (Fig. 2B), were constructed. BmN cells transfected with this plasmid were treated with either 2 μM Asunaprevir or DMSO for 72 hours. Post-treatment, GFP fluorescence (Fig. 2C) and the size of the anticipated cleaved protein (Fig. 2D) were assessed. A notable decrease in GFP fluorescence was observed following 2 μM Asunaprevir treatment (Fig. 2C). While weak fluorescence was detectable under a microscope, Western blotting with an anti-GFP antibody did not detect cleaved GFP (Fig. 2D). In cells treated with DMSO, a prominent band at approximately 30 kDa, the expected size for GFP, was observed. The SMASh system with Asunaprevir appeared to be functional in BmN cells. I then created a mutated SMASh-tagged EGFP plasmid pOpIE2p:EGFP::SMASh^{RP,} altering the NS3-cleavage site between the SMASh-tag and EGFP, by introducing amino acid substitutions to block cleavage by the NS3 in mammalian cells (Kolykhalov *et al*., 1994), to validate that NS3 protease functioned in BmN and compare the specificity of NS3 protease with that reported in mammalian cells (Fig. 2E). Western blot analysis of cells expressing this mutated protein without Asunaprevir revealed a faint band corresponding to the size of the uncleaved protein. On the other hand, the uncleaved protein was undetectable in cells treated with 2 μM Asunaprevir.. This indicates that the mutated protein was degraded without NS3 protease inhibition by Asunaprevir, likely due to impaired cleavage and the existence of fused NS4A-derived degron. Furthermore, the band representing the uncleaved protein was absent in cell extracts treated with 2 μM Asunaprevir, indicating additional effects of Asunaprevir beyond NS3 inhibition. These findings demonstrate that HCV NS3 protease in BmN cells can recognize and cleave the DEMEECSQHL sequence between EGFP and the SMASh tag, similar to its activity in mammalian cells. Moreover, the partial NS3 helicase and NS4-derived degron associated with the SMASh tag effectively degraded the uncleaved form of GFP in BmN cells.

The partial NS3 Helicase-NS4A outperforms PEST in BmN Cells

Following our observation that the degron in the SMASh tag effectively reduced the fused EGFP to undetectable levels in an Asunaprevir-dependent manner (Fig. 2D), I compared its

performance with the commonly used PEST degron (Rogers *et al*., 1986). The PEST degron has been shown to shorten the half-life of EGFP to less than 2 hours by fusing it to the C-terminal side of EGFP (Li *et al*., 1998), and it was shown to function as a degron sequence in Drosophila (He *et al*., 2019).For this comparison, the degron in our construct was substituted with PEST (pOpIE2p:EGFP::CS::NS3pro::PEST::NS4Aβ in Fig. 3A and B). The substitution with PEST resulted in a higher accumulation of both cleaved and uncleaved forms of EGFP compared to the original SMASh construct (Fig. 3B). While it was considered the possibility that the internal placement of the PEST sequence, which is typically located at the C-terminus for effective degradation, might not be fully functional, a modified version of the SMASh tag containing PEST (pOpIE2p:EGFP::CS::NS3pro::NS4Aβ::PEST in Fig. 3A and B) accumulated cleaved and uncleaved forms similar to those of the construct with internally positioned PEST. This indicated that PEST could function effectively in this internal region of the SMASh tag, and revealed that the degron derived from HCV NS3 helicase-NS4A is more effective at degrading GFP than the PEST degron in BmN cells. Microscopic observations revealed that the cells expressing the original SMASh-tagged GFP showed much lower GFP fluorescence than those from the modified SMASh-tagged ones upon treatment of 2 μM of Asunaprevir (Fig. 3C). However, the GFP fluorescence from the cells expressing the modified SMASh-tagged protein had fluorescence stronger than expected from the at the protein quantity of the cleaved GFP in the Western blotting, which may indicate the uncleaved form of the fusion proteins can emit fluorescence.

SMASh functions in BmN in a concentration-dependent manner

Next, the relationship between Asunaprevir concentration and GFP expression of SMASh-tagged EGFP in BmN cells was investigated. A range from 10^{-10} to 1 μM of Asunaprevir (Fig. 4A) was examined. GFP fluorescence quantification was performed using a fluorescent microplate reader. The fluorescence from cells treated with 10^{-10} to 10^{-4} μM Asunaprevir was comparable to that of DMSO-treated BmN cells. GFP fluorescence decreased in cells treated with concentrations higher than 10^{-4} μM Asunaprevir. The data were fitted to a dose-response curve to interpolate GFP fluorescence values at unmeasured concentrations. Using these estimates, the slopes of the GFP fluorescence profile ($\Delta\log_2$ [fluorescence intensity]) and their standard errors (Fig. 4B) were calculated. Differential slope values significantly different from zero, indicating the GFP fluorescence was decreasing with lowering Asunaprevir concentrations, were observed in the range of 6.3×10^{-4} to 2.51×10^{-1} μM Asunaprevir. The SMASh tag modulated GFP expression approximately 8-fold across a three-order-of-magnitude concentration range, under the control of the OpIE2 promoter and SV40 terminator. For higher Asunaprevir concentrations, 1 to 10 μM, expected to more strongly suppress GFP expression, effects of Asunaprevir concentrations on GFP expression

were investigated using a microscopy (Fig.4C). GFP fluorescence was undetectable in cells transfected with pOpIE2p:EGFP::SMASh and treated with these high Asunaprevir concentrations, while it was visible in cells transfected with pOpIE2p:EGFP::FLAG treated with up to 5 μM Asunaprevir. A reduction in both GFP fluorescence intensity and the number of GFP-expressing cells with 2 μM or higher Asunaprevir concentrations even in the cell transfected with the plasmid not carrying the SMASh tag was observed. This observation indicated negative effects on GFP expression and, probably, global gene expression. Cell viability also decreased with treatment at 5 μM or higher Asunaprevir concentrations. Therefore, it is likely that the high concentration of Asunaprevir affects global gene expression. Based on these findings, it was concluded that the optimal and effective Asunaprevir concentration range for BmN cells is 10^{-4} to 1μM under our experimental conditions. While 2 μM Asunaprevir effectively suppressed target protein expression, cytotoxicity in cells treated at this concentration was occasionally observed.

Shut-off by Asunaprevir in BmN cannot be reversed by exchanging culture media.

The potential to reverse inhibition by removing the reagent is a notable advantage of pharmaceutical approaches. In mammalian cells, exchanging culture media to remove Asunaprevir successfully aborted the degradation of tagged proteins (Chung *et al*., 2015). I aimed to determine if this approach could similarly reverse SMASh-mediated protein degradation and restore GFP expression in BmN cells. Initial experiments with BmN cells treated with 1μM Asunaprevir revealed no GFP fluorescence recovery following three washes with Asunaprevir-free media (data not shown). Then BmN cells transfected with pOpIE2p:EGFP::SMASh were treated with Asunaprevir at concentrations of 0.01, 0.1, and 1μM for 48 hours, focusing on whether low Asunaprevir concentrations could reverse suppression. After treatment, cells were washed, and GFP fluorescence was measured at one-hour intervals over 16 hours (Fig. 5A). Controls included non-treated BmN cells and those transfected with pOpIE2p:EGFP::Luc treated with Asunaprevir. In cells treated with 1 μM Asunaprevir, the increase in GFP fluorescence was significantly lower than in cells treated with lower concentrations of Asunaprevir, observed in both SMASh-tagged and luciferase-fused EGFP plasmids. Cells treated with 0.1 μM Asunaprevir also showed lower level of fluorescence. GFP fluorescence rates were comparable between washed and unwashed cells at all the tested concentrations. Statistical analysis of GFP fluorescence profiles showed no significant differences between these pairs under any condition (Fig. 5B), indicating that SMASh-mediated gene suppression in BmN cells is irreversible after Asunaprevir removal. This finding contrasts with mammalian cell data and may be due to differences in Asunaprevir exportation in BmN cells. Further investigation involved analyzing culture media from post-wash cells (96 hours post-transfection) used to treat new pOpIE2p:SMASh::EGFP-transfected cells (Figs. 5A and C).

Suppressed GFP fluorescence was noted in cells cultured with media from cells treated with 1 μM Asunaprevir and then washed. This observation suggests possible Asunaprevir release into the media, either through efficient export mechanisms or from dead cells. Considering Asunaprevir's negative impact on gene expression and cell viability of BmN cells (Fig. 4C), release from dead cells is plausible, even if an Asunaprevir exporter existed in the silkworm genome. These results indicate that in BmN cells, as opposed to mammalian cells, SMASh-mediated gene suppression remains irreversible.

Discussion

To elucidate the regulatory relationship among genes, it is required to perturb and/or control dynamic gene expression. However, gene expression control systems in silkworms have been limited. Therefore, there has not been researches on time-resolved regulation of BmNPV gene expression. Drug-controllable gene expression is one of the methods that can be used for fine-tuning gene expression. In this study, it was revealed that SMASh tag, initially developed for mammalian cells, also functions as a drug induced gene expression control system in BmN cells. While it was found that expression can be controlled in a drug concentration-dependent manner as in mammalian cells, it was also found that this control is irreversible in BmN cells, unlike mammalian cells. Despite the lack of the anticipated reversible effect of SMASh in BmN cells, however, our findings implied its potential as a robust gene expression shutdown of target protein in silkworm animals. Results in this study implied no significant Asunaprevir metabolization and degradation with the BmN cells over 96 hours. Although it is required further investigations of dynamics of Asunaprevir in silkworm animals, strong gene suppression may be possible in the silkworm body unless Asunaprevir is released through feces or decreased concentration due to rapid larval growth at later developmental stages.

In this study, the protein degradation of the degron-like sequence of SMASh tag was evaluated by constructing a new tag that replaced the degron-like sequence of SMASh tag with PEST and comparing the results. This is the first, to our knowledge, to report manipulation of degradation strength by modifying the SMASh tag. The degron sequence of SMASh tag reduced the fluorescence of non-cleaved GFP to an invisible level, and induced protein degradation before protein maturation, whereas PEST exhibited fluorescence of non-cleaved GFP. These results indicate that the ability of degron-like sequence of SMASh tag to induce protein degradation is higher than that of PEST. A fusion protein of EGFP and PEST had a fluorescent half-life of 2 hours in mammalian cells (Li *et al*., 1998), while GFP had a fluorescent half-life of ～26 hours (Corish and Tyler-Smith., 1999). Therefore, a fusion protein of degron-like sequence of SMASh and EGFP had a half-life shorter than 2 hours. Destabilization of fluorescent proteins is useful for increasing the temporal resolution of reporter analysis, and PEST has been used for this application (Li *et al*., 2019). The degron-like sequence of SMASh is also expected to be used in this application, but EGFP fluorescence was not observed with this sequence in this study, because the destabilization may have been too strong, inducing proteolysis before the maturation time, which is the time required for EGFP to show fluorescence after translation, had elapsed. Moreover, the improvement in time resolution caused by protein destabilization is a trade-off for the intensity of fluorescence (Li *et al*., 2019), so its use should be carefully considered.

Additionally, NS3 protease within these modified SMASh tags cleaved the fusion proteins into functional GFP, indicating that the HCV NS3 protease was likely to be potent in BmN cells without the NS4A TM helix and β-strand cofactor. The NS4A functions as a co-factor to increase the protease activity of NS3 (Landro *et al*., 1997), and was contained in the SMASh tag (Fig. 3A). Although it is demonstrated that NS3 alone functioned as a protease in mammalian cells (Jacob *et al*. 2018), this may not have been true in BmN, the cell line of nonhost species of HCV. Furthermore, since the BmN cells transfected with the original and modified SMASh tags grew without apparent side effects by expression of the protease, HCV NS3-based methods, such as StaPLs (Jacob *et al*. 2018), would likely be applicable in BmN cells as well.

In conclusion, our study demonstrates that (1) HCV NS3 protease and the partial NS3 helicase and NS4-derived degron are functional in BmN cells, enabling effective SMASh-tagged GFP degradation, (2) with a degradation efficiency surpassing that of PEST, (3) quantitatively across various Asunaprevir concentrations, and (4) in an irreversible manner. These findings open new avenues for exploring biological processes and developing biotechnological applications that require rapid and efficient gene suppression.

Chapter 2 : Elucidation of the regulatory system for *polyhedrin* promoter activity based of promoter activation profiles

Abstract

To understand the activation of the *polyhedrin* promoter during the very late stage of BmNPV infection, one requires understanding the gene regulatory network as the blueprint of the expression system. I aimed at understanding the regulatory network consisted of previously identified ten genes essential for the *polyhedrin* promoter activation. To this end, I constructed ten dual-reporter viruses to express two fluorescent proteins to monitor the *polyhedrin* promoter activity together with one of the ten genes. Combined the reporter viruses with a time-lapse imaging system, single-cell resolution measurement of the two promoter activities was achieved simultaneous. Genetic perturbations to the ten essential genes combined with the reporter imaging elucidated their regulatory relationships. This analysis identified *lef-8*, *lef-9*, and *lef-10* as marker genes of *polyhedrin* promoter activity and suggested the existence of at least two regulatory mechanisms. In addition, *p143* emerged as one of the primary genes controlling the regulatory mechanisms. Lastly, a genome-wide *p143* regulator screening revealed genes affecting *p143* promoter activity and possible association of DNA replication with the promoter regulation. From these results, *polyhedrin* promoter activity can be inferred from the activation profile of essential genes, suggesting the importance of the essential gene regulatory network in the early stages of infection for understanding the promoter activation system.

Introduction

Gene expression regulation mechanisms in eukaryotic organisms play a crucial role in responding to environmental changes using a limited set of genes (Knight *et al*., 2005, Ho and Zhang., 2018). Changes in expression control mechanisms act as a driving force for diverse evolutions(Wray, 2007, De *et al*., 2013 , Bylino *et al*., 2020,). In eukaryotic organisms, the transcriptional regulation is a major system for gene expression regulation mechanisms. Transcription in eukaryotic organisms was initiated by the recognition and binding of promoter sequences which were upstream of open reading frames (ORFs) in the genome, by the basic transcription machinery comprised of RNA Polymerase II（Roeder, 1996, Buratowaki ,1994）. For example, there are TATA box with the common sequence "5'-TATAA-3'," and the CAAT box, which is the binding site for the RNA transcription factor (Smale and Baltimore 1989, Dynan and Tjian, 1983). Therefore, to understand the regulatory system of expression of a gene, it is important to identify the promoter sequence located upstream of the ORF of the gene and to understand the regulatory pattern of activation of the promoter. However, because the length of promoter sequences is diverse and it is difficult to define a promoter by analyzing motifs or its length, reporter assay systems using multiple candidate promoter sequences and plasmids with a candidate promoter sequence introduced upstream of the reporter gene have been widely used to identify the promoter sequence and unravel the gene expression regulation mechanisms.

Viruses that host eukaryotes involve an infection cycle that utilizes the host's transcription machinery. Viruses, like eukaryotes, undergo changes in their regulatory systems in response to their environment and evolve. On the other hand, viruses differ from eukaryotes in that they alter the host environment as the progression of infection and regulated the type and amount of genes expressed at each stage of infection. Therefore, plasmid-based reporter assay systems are not sufficient to understand the regulatory systems of viral gene expression, and it is necessary to understand each of the regulatory systems at each stage of infection in virus-infected cells.

Bombyx mori nucleopolyhedrovirus (BmNPV) is a type of baculovirus and double-stranded DNA virus. The hosts of this virus are silkworms, and cells derived from ovary of silkworms are widely utilized to understand the character of this virus. This virus has been shown to alter the host environment during infection (Gomi *et al*., 1999) . For example, baculovirus suppresses host gene expression as the progresses of infection (Masumoto., 2012) and this virus forms a structure called virogenic stroma (VS) which is a scaffold for DNA replication and gene expression (Nagamine *et al*., 2008). Additionally, in the early stages of infection, the virus utilizes host-derived RNA Polymerase II, while in later stages, genes are transcribed by a virus-specific RNA Polymerase complex consisting of *p47*, *lef-4*, *lef-8*, and *lef-9* subunits (Jin *et al*., 1998).

Therefore, in BmNPV, plasmid-based reporter assay is not sufficient for unrebelling the regulatory relationship of BmNPV genes, especially at the later stage of infection.

BmNPV gene expression during the infection is categorized into four stages: immediate early, early, late, and very late (Friesen and Miller, 1986). During immediate early stage of infection, ie1 functions as a transactivator and activates the expression of early genes (Carson *et al*., 1991; Guarino and Dong, 1991). ie1 forms foci in virus infected cells and their foci expand as the progression of infection and the localization of ie1 is thought to be matched with VS (Nagamine *et al*., 2008). ie1 foci began to be observed at 4 hours post-infection and expanded by 24 hours. During the early stage of infection, the early gens are transcribed in the ie1 foci (Okano *et al*., 1999; Kawasaki *et al*., 2004), and the essential genes for virus DNA replication and subunit genes of RNA polymerase are expressed. In addition, 19 genes required for late gene expression are expressed during this stage (Todd *et al*., 1996; Lu and Miller., 1994) and some of the 19 genes are called late expression factor (*lef*) genes. Early gene expression peaks between 6 and 12 h post infection, while DNA replication takes place between 6 and 18 h post infection (Friesen, 1997). It has been reported that during this infection phase, a switch occurs from transcription by host RNA polymerase II in the early stage of infection to transcription by virus-encoded RNA polymerase in the late stage of infection. (Morris *et al*., 1994). During the late stage of infection, the component protein genes of budded virus (BV) which is a progeny virus associated with viral capsids and cell-to-cell infection, are transcribed, and expressed by virus-derived RNA polymerase (Passarelli *et al*., 1994; Knebel *et al*., 2006). Additionally, *pk-1* and *vlf-1*, which are late gene, are involved the very late gene expression (Mclachlin and Miller, 1994; Fan *et al*., 1996). During the very late stage of infection, *polyhedrin*, which is a structure protein gene of occlusion derived virus (ODV) which is a progeny virus required for oral infectivity, is hyper expressed (Smith *et al*., 1982). Hyperexpression of *polyhedrin* during the very late stage of infection is activated by *polyhedrin* promoter. It is used industrially because it can express large numbers of foreign genes by introducing foreign genes downstream of the *polyhedrin* promoter in the viral genome and infecting cells or individuals (Maeda *et al*., 1985). This hyper expression accounts for 24 % of the total RNA in the host cell at the late stage of infection (Chen *et al*., 2013), yet sufficient expression is achieved by 69 base pair promoter sequence (Possee and Howard, 1987). However, this gene promoter alone is not activated. Previous studies revealed that at least 19 genes are necessary for activation in a transient expression system using plasmids (Todd *et al*., 1996, Lu and Miller., 1994). By Knocking out experiments of each gene of whole genes in BmNPV genome, 10 genes (*ie1, p143, lef-3, lef-4,lef-5, lef-8, lef-9, lef-10, lef-11*, and *p47*) have been identified as essential genes for *polyhedrin* promoter activity (Ono *et al*., 2012). In addition, the function of each essential gene was identified by individual validation experiments for each gene. IE1 promotes early viral gene expression and genome replication as a trans-activator (Guarino and Summers., 1986). P143 is associated with genome replication as a

DNA-helicase (McDougal and Guarino, 2000). LEF-3 is a single-stranded DNA binding protein (Hang *et al*., 1995). LEF-4, LEF-8, LEF-9, and P47 are RNA polymerase subunits (Guarino *et al*., 1998). LEF-5 is a TFIIs homolog associated with RNA elongation (Harwood *et al*., 1998). LEF-10 acts as a prion and aggregated state inhibits late gene expression (Xu *et al*., 2016, Nan *et al*., 2019). LEF-11 interacts with and upregulates both a host ATPase and HSPD1 proteins, likely facilitating viral DNA replication (Dong *et al*., 2017). While these 10 genes have been determined as essential gene for *polyhedrin* promoter activity, they are also required for virus DNA replication and transcription of the virus gene. Therefore, whether they are essential in the progression of infection stages, in the environmental formation for promoter activation, or their relation to *polyhedrin* promoter activity for each gene remains unclear. Furthermore, it is difficult to discuss this essentiality at the virus's functional level. For examples, genes related to DNA replication such as *p143*, *lef-3*, and *lef-11* were classified as essential genes, whereas other DNA replication-related genes like *lef-1* (Mikhailov and Rohrmann, 2002), *lef-2* (Sriram and Gopinathan, 1998), *DNA polymerase* (Vanarsdall *et al*., 2005) were not classified. To understand the *polyhedrin* hyper expression mechanism, it was necessary to determine the stages in the progression of infection that are essential for *polyhedrin* promoter activity, and understand why polyhedrin promoter activity was lost when each of the 10 genes was deleted.

In this study, I aimed to understand them by estimating the regulatory relationships among these 10 essential genes, which are the upstream genes of polyhedrin promoter activity, and by clarifying the correlation between the expression of each essential gene and polyhedrin promoter activity. First of all, I constructed dual-reporter bacmids to measure the dynamics both *polyhedrin* promoter activity and essential gene promoter activation simultaneously. The bacmids are the virus genome carrying elements to be replicated in *E. coli* (Luckow *et al.,* 1992*;* Ono *et al*., 2007). Next, by multi-point time-lapse observation of these dual-reporter virus-infected cells, the changes in promoter activation dynamics of each gene was able to be analyzed at the single-cell level. This allowed for analysis that accounted for infection stage progression, which could not be captured in previous population-level analyses, enabling the acquisition of a large amount of time-series data from corresponding dual fluorescent proteins. Furthermore, by combining the high-throughput knockdown system, Soaking RNAi (Mon *et al*,, 2013), diverse corresponding data sets, making it possible to analyze the relationship between essential genes and *polyhedrin* promoter activity were obtained. Analyzing the control relationships between essential genes also allowed us to elucidate the roles of each essential genes in virus infection stage progression. This experimental system serves as a reporter assay for capturing the changes in control relationships during virus infection stage progression and is a distinctive experiment compared to prior research, especially in handling a large amount of time-series data. This study capitalized on these features to unravel the dynamic control mechanisms with *polyhedrin* promoter activity placed downstream.

Materials & Methods

1. Cultured Cells

BmN cells derived from silkworm (*Bombyx mori*) ovarian tissue were utilized for virus bacmids transfection experiments. BmN cells were cultured at 26℃ in TC-100 medium (Applichem) containing 10% Fetal Bovine Serum (FBS) (biosera) (see Appendix 1, Table 1). BmN4-SIDI cells (Mon *et al*., 2012) were utilized for virus infection experiments. BmN4-SIDI cells were kindly provided from Dr. Hiroaki Mon and cultured at 26℃ in IPL -41 medium (Invitrogen) containing 10% Fetal Bovine Serum (FBS) (biosera).

2. Basic Procedures for Genetic Experiments

i. Preparation of Chemically Competent Cells

Chemical competent cells were made from *E. coli* DH5α strain. Colonies of *E. coli* were inoculated into 1.5 mLof LB broth (see Appendix 1, Table 2) in test tubes and shaken at 37℃ for 16 hours. Subsequently, the culture was added to 100 mL of LB broth and shaken at 37℃ until the OD600 reached 0.3-0.4. After rapid cooling in ice, the culture was divided into two Falcon tubes and subjected to centrifugation (3,000 rpm, 15 minutes, 4℃). The supernatant was partially removed (approximately 1 mL), and the bacterial pellet was gently resuspended. Tfb I (see Appendix 1, Table 3) was added (20 mL) and incubated for 10 minutes, followed by centrifugation (3,000 rpm, 15 minutes, 4℃). The supernatant was completely removed, and Tfb II (see Appendix 1, Table 4) was added (2 mL) to resuspend the bacterial pellet. The suspension was then transferred to pre-cooled 1.5 mL Eppendorf tubes (100 μL per tube) and stored at -80℃.

ii. Plasmid DNA Extraction

Plasmid DNA was extracted from *E. coli* using the QIAprep Spin Miniprep Kit (Qiagen) according to the product manual. *E. coli* colonies were inoculated into 1.5 mL of LB broth with added antibiotics (Ampicillin, 150 μg/mL) and cultured at 37℃ for 16 hours with shaking. When pFastBac-mScarlet-i-SV40T-I-*SceI*-nls-sfGFP-PEST was extracted, *E. coli* colonies were plated on 20 mL of LB plate with added antibiotics (Ampicillin, 150 μg/mL) and cultured at 37℃ for 16 hours without shaking. This was because culturing with shaking introduced the lack of gene from the plasmid. The colonies were suspended with 2 mL of sterile distilled water, and this was used as cultures. The culture was centrifuged (3,000 rpm, 3 minutes, 4℃) to remove the supernatant. 200μL

of P1 buffer (Qiagen) with RNase A was added to the bacterial pellet and mixed well using a vortex. Then, 200 μL of P2 buffer (Qiagen) was added, gently inverted, and stand at room temperature for 3 minutes. Subsequently, 300 μL of N3 buffer (Qiagen) was added, gently mixed, and then left in ice for 3 minutes before centrifugation (15,000 rpm, 10 minutes, 4℃). The supernatant was collected, loaded into the kit's column, and centrifuged (15,000 rpm, 1 minute, 4℃). PB buffer (Qiagen) and PE Buffer (Qiagen) were sequentially passed through the column, and centrifugation (15,000 rpm, 1 minute, 4℃) was performed to remove residual buffer. Finally, 50 μL of TE buffer (see Appendix 1, Table 5) was added to the column, left to stand for 1 minute, and centrifuged (10,000 rpm, 1 minute, room temperature) to recover the plasmid DNA.

iii. Restriction Enzyme Treatment

Restriction enzyme treatment involved adding 1 μL of 10× Cut Smart Buffer (NEB) and 1 μL of the restriction enzyme to the solution containing DNA (less than 1 μg). The mixture was adjusted to a total volume of 10 μL with sterile distilled water and treated at 37℃ for 2 hours. When using I-SceI for restriction enzyme treatment, the treatment was carried out at 37℃ for 16 hours.

iv. PCR

The primers used in this experiment are listed in Appendix 2. The reactions were carried out using PrimeSTAR HS DNA Polymerase (TaKaRa), with an extension time set at 1 minute per kilobase as the standard. After the reaction, the solution was purified by ethanol precipitation, and gel extraction was performed in the case when non-specific bands during electrophoresis was detected.

v. SLiCE

SLiCE solution was prepared following the method by Motohashi (2015) using *E. coli* DH5α strain. *E. coli* was inoculated into 1.0 mL of LB broth and cultured with shaking at 37℃ and 200 rpm for 3 hours. Afterward, the culture was added to 50 mL of 2×TY broth (see Appendix 1, Table 6) and further cultured with shaking at 37℃ and 200 rpm until the OD600 reached 1.8 (approximately 5 hours). The culture was then transferred to a 50 mL Falcon tube and centrifuged (5,000×g, 10 minutes, 4℃) to collect the pellet. The collected pellet was washed with 50 mL of

sterile water kept on ice, and centrifuged again (5,000×g, 5 minutes, 4℃). The pellet was dissolved in 1.5 mL of CelLytic B Cell Lysis Reagent (Sigma, B7435) and transferred to an Eppendorf tube. After incubating at room temperature for 10 minutes, the supernatant was transferred in 5 μl aliquots to 0.2 mL PCR tubes, mixed with an equal volume of 80% glycerol, and stored at -80℃. The SLiCE reaction was performed by adding 1 μl of the prepared SLiCE solution and 1 μl of 10× SLiCE buffer (see Appendix 1, Table 7) to a solution containing 20 ng of vector DNA and 60 ng of insert DNA. The solution was adjusted to 10 μl with sterile water, followed by incubation at 37℃ for 20 minutes. When performing transformation of *E. coli*, 1 μl of the reaction mixture was added to 100 μl of chemical competent DH5α cells, incubated on ice for 20 minutes, subjected to a 42℃ heat shock for 60 seconds, allowed to stand on ice for 2 minutes, and then 1 mL of SOC medium (see Appendix 1, Table 8) was added. The culture was shaken at 250 rpm for 1 hour at 37℃. After centrifugation (3,000 rpm, 3 minutes, 4℃), an appropriate amount of supernatant was removed, and the remaining supernatant was used to resuspend the *E. coli* cells. Each bacterial strain was seeded onto LB plates (see Appendix 1, Table 9) containing appropriate antibiotics based on the drug selection specific to each vector to form colonies.

vi. Gibson Assembly

Gibson Assembly was performed using the Gibson Assembly Master Mix (NEB). DNA solutions were mixed in a 1:3 molar ratio for vector fragments and insert fragments. A 5 μL mixture of these fragments and Gibson Assembly Master Mix (NEB) was incubated at 50℃ for 1 hour. The reaction mixture was diluted four times, and 4 μL of it was added to 100 μL of chemical competent DH5α cells. After placing it on ice for 20 minutes, a heat shock was performed at 42℃ for 60 seconds, followed by 2 minutes of incubation on ice. Subsequently, 1 mL of SOC medium (see Appendix 1, Table 8) was added, and the culture was shaken at 250 rpm for 1 hour at 37℃. After centrifugation (3,000 rpm, 3 minutes, 4℃), an appropriate amount of supernatant was removed, and the remaining supernatant was used to resuspend the *E. coli* cells. Each bacterial strain was seeded onto LB plates (see Appendix 1, Table 9) containing appropriate antibiotics based on the drug selection specific to each vector to form colonies.

vii. Sequencing Reaction

Sequencing reactions were performed using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). The reaction mixture included 100 ng of template DNA, 10 pmol of oligonucleotide primer, 2 μL of 5× sequencing buffer, and 0.25 μL of premix, made up to a total volume of 10 μL with sterile distilled water. The reaction was carried out using the following cycling conditions: [95℃, 2 minutes / 98℃, 30 seconds / 96℃, 1 minute] × 1 cycle, [96℃, 10 seconds / 47℃, 40 seconds / 60℃, 4 minutes] × 8 cycles, [94℃, 5 minutes / 50℃, 30 seconds / 60℃, 4 minutes] × 10 cycles, [98℃, 19 seconds / 50℃, 20 seconds / 60℃, 4 minutes] × 7 cycles, and [60℃, 5 minutes] × 1 cycle. The reaction mixture was ethanol-precipitated, the pellet was dissolved in 20 μL of Hi-Di Formamide (Applied Biosystems), and the DNA suspension was heated at 95℃ for 2 minutes to denature the DNA. The mixture was rapidly cooled on ice, and the sample was analyzed using the ABI3130 sequence analyzer (ThermoFisherScientific). I used the DNA Sequencing Facility of the Graduate School of Agriculture.

viii. Preparation of Electrocompetent Cells

Colonies of DH10B on LB plates were inoculated into 1.5 mL of LB broth and they were shaken at 37℃ for 16 hours. The culture was inoculated to 100 mL of LB broth and shaken at 37℃ until the OD600 reached 0.4. After rapid cooling on ice, the culture was divided into two 50 mL tubes and centrifuged (2,500 rpm, 15 minutes, 4℃). This entire process, from removing the supernatant to centrifugation, was performed twice. After washing with 10% glycerol twice, the bacterial pellet was resuspended in 1 mL of 10% glycerol, left to stand on ice for 5 minutes, and then distributed in 50 μL aliquots to 1.5 mL tubes that had been pre-cooled with dry ice. The tubes were stored at -80℃.

3. Donor Plasmid Construction

The donor plasmid for the creation of a dual reporter bacmid, pFastBac-DualReporter-I-SceI, was stepwise constructed using PCR-generated fragments and assembled via SLiCE or Gibson Assembly as outlined in Fig. 6A.

i Construction of pFastBac-mScarlet-I-SV40T

A vector fragment devoid of DsRed, derived from pFastBac-DsRed-SV40T (Ono, 2011), was a mplified using PCR with Primer 3 and Primer4 (Appendix 2, Table 11). A mScarlet-I fragment, ampl ified to contain 25 bp of vector sequence at both ends using Primer 1 and Primer 2, was derived from

puc19-mScarlet-I. These two fragments were ligated using SLiCE, resulting in the creation of pFast Bac-mScarlet-I-SV40T.

ii Construction of pFastBac-mScarlet-i-SV40T-I-SceI-nls-sfGFP-PEST (pFastBacDualReporter-I-SceI)

The vector fragment for SLiCE was obtained by I-SceI digestion of pFastBac-mScarlet-I-SV40T, followed by gel extraction and purification. The insert sequence, I-SceI-nls-sfGFP-PEST-p10T, was generated through a two-step Fusion PCR approach. The I-SceI recognition sequence was synthesized via template-independent PCR (Primer 11 and Primer 12). The nls-sfGFP fragment was amplified from pHR-scFv-GCN4-sfGFP-GB1-dWPRE (Addgene, 60907) using Primer 5 and Primer 6, and it was purified by ethanol precipitation. The PEST and p10T sequences were amplified with Primer 9, Primer 10, and Primer 10, Primer 11, respectively, using pUAST::IVS-Syn21-nlsGFP-PEST-P2A-nlsRFP-p10 (Li *et al*., 2019) as a template and purified by ethanol precipitation. Out of these four fragments, I-SceI and nls-sfGFP were combined through Fusion PCR (Primer 5 and Primer 12) and then amplified. PEST and p10T were combined through Fusion PCR (Primer 8 and Primer 10) and amplified separately. These two resulting fragments, I-SceI-nls-sfGFP and PEST-p10T, were further combined through Fusion PCR (Primer 8 and Primer 11), and the desired product, I-SceI-nls-sfGFP-PEST-p10T, was obtained by gel extraction. The vector fragment and insert fragment were ligated using SLiCE, resulting in the creation of pFastBacDualReporter-I-SceI. During the use of this plasmid in subsequent constructions, it was noted that regular shake culture resulted in the loss of the plasmid from *E. coli* during cultivation and introduced mutations in the sequence. To avoid this, plasmid DNA extraction was carried out using a shaking incubator at 37℃ for 16 hours, followed by detachment of colonies using sterile water, and the resulting bacterial suspension was used as a sample for plasmid DNA extraction.

iii Introduction of Baculovirus Promoters into pFastBacDualReporter-I-SceI

The promoter of genes essential for the activation of polyhedrin promoter were introduced using PCR (Primer 13 - 34). These promoters were amplified using primers specific to each promoter and a 10% SDS solution mixed at a 1:1 volume ratio with BmNPV bacmid virus supernatant. The mixture was incubated at 95℃ for 5 minutes and then centrifuged at 15,000 rpm for 15 minutes to obtain the supernatant. The promoter fragments were used as insert sequences for Gibson assembly, and pFastBacDualReporter-I-SceI was I-SceI digested at 37℃ for 15 hours. The gel-extracted fragments were used as the vector, and Gibson assembly was performed to introduce the baculovirus promoters into pFastBacDualReporter-I-SceI.

4. Construction of Dual Reporter Bacmids

Electrocompetent cells of *E. coli* DH10B, containing the T3 BmNPV bacmid (Ono *et al,* 2007) and the helper plasmid pMON7124 encoding the Tn7 transposase, were transformed with 500 ng of each donor plasmid used for the introduction of various promoters. Electroporation (200 Ω, 250 μFD, 2.5V) was used to induce transposition reactions. Subsequently, 1 mL of SOC medium was added, and the culture was shaken at 37℃ for more than 4 hours to induce transposition reactions. After shaking, 100 μL of the culture was mixed with 30 μL of 100 mg/mL IPTG solution and 30 μL of 0.4% X-Gal solution. The mixture was spread on LB-Kan/Gen/Tet plates (Kanamycin 50 μg/mL, Gentamycin 50 μg/mL, Tetracycline 50 μg/mL). The plates were incubated at 37℃ for 36 hours, and white colonies were selected. Transposition events were confirmed by PCR. Colonies were cultured, and plasmid purification was performed using the QIAGEN Plasmid Mini Kit (Qiagen) following the manufacturer's instructions, resulting in a mixture of various promoter-introduced dual reporter bacmids, donor plasmids, and pMON7124.

5. Basic Cell Experiment Procedures

i Transfection of BmN Cells

After washing the cultured cells twice with serum-free TC-100 medium, BmN cells were seeded in a 96-well plate (Iwaki) at a density of 5×10^4 cells per well and incubated at 26℃ for 2 hours. FuGENE HD Transfection Reagent (Promega) was mixed with TC-100 medium and 1 μg of DNA to achieve a final volume of 100 μL, and this mixture was incubated at room temperature for 20 minutes. Subsequently, this mixture was added to the cultured cells in the 96-well plate mentioned above, followed by a 6-hour incubation at 26℃. Afterward, all the medium was removed, and 100 μL of TC-100 medium containing 10% FBS was added to continue the incubation. The time point when this operation was completed was considered as the 0-hour transfection, and 120 hours after transfection, the medium from each well was collected to obtain the viral supernatant.

ii Titer Measurement of Viral Supernatant Using Plaque Assay

BmN cells were seeded in a 96-well plate (IWAKI) at a density of 5.0×10^4 cells per well after washing them twice with serum-free TC-100 medium. The cells were allowed to incubate for more than 2 hours at 26℃. The virus solution was serially diluted in serum-free TC-100 medium, and 45 μL of this diluted virus solution was added to each well, followed by a 1-hour incubation at 26℃.

Subsequently, the virus solution was removed, and a mixture of 5 mL of 2% SeqPlaque Agarose (Lonza) and 5 mL of 20% serum 2× TC100 medium was added at a rate of 120 µL per well. The cells were then incubated at 26℃, and after 4 days, they were stained with 0.15% Neutral Red. The Plaque Forming Unit (PFU) was calculated based on the number of plaques formed.

iii Baculovirus Infection of BmN Cells

BmN cells were seeded in a 96-well plate (Eppendorf Cell Imaging Plates, No. 0030741013) at a density of 1.0×10^4 cells per well after washing them twice with serum-free TC-100 medium. The virus solution was diluted with serum-free TC-100 medium to achieve a multiplicity of infection (MOI) of 1.0, 0.1, 0.01, and 0.001, and was added to each well. The cells were incubated for 1 hour at 26℃, after which the virus solution was removed, and a mixture of 5 mL of 2% SeqPlaque Agarose (Lonza) and 5 mL of 20% serum 2× TC100 medium was added at a rate of 120 µL per well. The time point when this medium completely solidified was defined as the infection time point 0.

6. Genome-wide screening

i. Construction of template plasmid for synthesis of double stranded RNA

Viral genome-wide screening experiments were performed by knockdown using the SoakingRNAi method, which requires dsRNA designed to be homologous to BmN4-SID1 cells and viral genes specifically. The dsRNA design for each viral gene was performed by primer3, which contains the genomic data of the viral gene and host gene in the database. Since some of the genes did not have homologous sequences to the host gene, dsRNA specific only for the viral gene was used for these genes. The region determined by design was amplified by PCR using the BmNPV genome as a template, and the PCR product was introduced between the mutually facing T7 promoters of a plasmid with pCR8 as the backbone (Fig 6B). The sequence was confirmed by Sanger sequencing, and the PCR product amplified by PCR using the created plasmid as template was purified by phenol chloroform extraction followed by ethanol precipitation and used as template for Invitro transcription.

ii. dsRNA synthesis

In vitro transcription was performed using the in vitro Transcription T7 Kit (TaKaRa) according to the product protocol. Purification was performed by phenol chloroform extraction and

ethanol precipitation, and quality check was performed by electrophoresis. Concentrations were measured by nanodrop.

iii. Soaking RNAi

Soaking RNAi was performed based on previous studies. 500 ng of dsRNA was added to each well of a 96-well plate and infection experiments were performed 3 days later for viral gene knockdown.

7. RT-qPCR

i. RNA extraction

RNA extraction from virus-infected cells was performed using TRIZOL® Reagent (Thermo Fisher Scientific); 154 μL of reagent per well of a 96-well plate was used to strip the cells by pipetting, which were then stored at -80℃. RNA was extracted according to the product protocol, adjusting the volume of other reagents to match the volume of reagent used. The extracted R N A was quantified by nanodrop and quality checked by electrophoresis.

ii. Reverse transcription

Genomic DNA was removed from the extracted RNA using DNase I, Amplification Grade (InvitorgenTM), and after heat inactivation of the enzyme, reverse transcription of the extracted RNA was performed using PrimeScriptTM Reverse Transcriptase. Product protocols were followed.

iii. qPCR

qPCR was performed using TB Green® Premix Ex Taq™ II (Tli RNaseH Plus) (TaKaRa) according to the product protocol. A ter primer capable of amplifying each gene specifically for viral and host genes was designed using primer3, with a Tm value of 60 degrees and an amplified product of about 100 bp. The number of cycles obtained for each viral gene was corrected by the number of cycles obtained targeting the host gene, Tubulin A gene, to obtain a ΔCt value, which was used as the viral gene expression level at each time point.

7. Imaging Analysis

i Optical System Used for Observation and Imaging Conditions

Time-lapse observation was conducted using a Spinning Disk Confocal Laser Microscope CSU-X1 (Microscope: Olympus IX71, Confocal Scanner Unit: CSU-X1) owned by NIBB. Laser sources at 640 nm (100 mW), 488 nm (60 mW), and 561 nm (50 mW) were employed, along with filters at 697/58, +520/35, and 617/73 for observing NucSpot Live Cell Nuclear Stains (Biotium) -stained cells, GFP fluorescence, and RFP fluorescence, respectively. Hamamatsu Photonics ImagEM 1K camera was used. Time-lapse imaging was controlled using the MetaMorph software (Molecular Devices), which managed stage positioning, timing, and auto-focusing. Imaging was performed using 96 Well Microplate (greiner, 655076), and 96 wells were imaged in a single session. Images were captured at six wavelengths for each well and each time point (W1: 640 nm with an exposure time of 10 ms, W2: 640 nm with an exposure time of 100 ms, W3: Bright field with an exposure time of 10 ms, W4: 488 nm with an exposure time of 1000 ms, W5: 561 nm with an exposure time of 50 ms, W6: 561 nm with an exposure time of 500 ms, The auto-focusing feature was applied to W1 at each well and each time point to capture the best focus slice as a single image. In W2 to W6, 11 Z-stack images were acquired at 10 μm intervals above and below the best focus slice. Acquired images were saved as a tiff format.

For imaging infected cells each time point, six combination of excitation/absorption filter sets and exposure times for image acquisition were established (Fig. 7B). In the first condition, we used the autofocus function of MetaMorph (Molucular devices), the control system for this observation. The autofocus function was used to search for the focal plane at each time point in each well by a two-step search: In the first step, fluorescence images were acquired every 50 μm within 150 μm of the focal plane at the previous time point in each well (at the first time point, the focal plane visually confirmed before the start of the observation), and the plane with the highest fluorescence intensity was selected as the candidate focal plane. In the second search, fluorescence images were acquired every 5 μm within a range of 50 μm up and down from the candidate focal plane, and the plane with the highest fluorescence intensity was designated as the focal plane at that time point for its well. NucSpot Live Cell Nuclear Stains (Biotium) fluorescent dye was used as the fluorescent dye to enable the observation of the best focal slice in each well even in the absence of fluorescence in the early stage of infection. This dye was excited by near-infrared light, which has minimal cell damage (Wan *et al*., 2019), and the fluorescence wavelength of this dye was distinct with one of other fluorescent proteins . When we used Hoechst 33342(Thermo Fisher Scientific) often used for live imaging, the decrease of *polyhedrin* promoter activity was confirmed and the effect of excitation with a short wavelength to cells couldn't be ignored (Fig. 7C). In the first condition, an exposure time was set to 10 ms, It was sufficient for auto-focus but not sufficient for cell-tracking. It was required to extend exposure time to conduct cell-tracking. However, it was difficult because the observation time at each timepoint was limited to one hour and if under the first condition, we try to achieve the two purposes, we would have to limit the search range of the autofocus, which could

reduce its accuracy. Therefore, observations were conducted under separate conditions in order to achieve these two purposes. In the second condition, exposure time was extended to capture images with sufficient fluorescence signals for the single cell tracking, even in the absence of virus-derived fluorescence during the early stages of infection. Z stack images, comprising 11 slices taken at 10 μm intervals from 50 μm above and below the best focal slice, were obtained for cell tracking due to potential plate distortion or cell morphology within the wells, which might cause inaccurate fluorescence in some cells. In the third condition, bright-field images were acquired. The light intensity of the microscope light source was increased, and the exposure time was minimized to reduce the observation time at a single time point. Z-stack images were acquired same as the second condition. A Python script using the Laplacian function was applied for images acquired with this condition to select the best focal slice of each cell for fluorescence intensity extraction. In the fourth condition, the fluorescence of sfGFP-PEST, expressed by the essential gene promoters, was observed. The fluorescent signals of this fluorescence protein were weaker due to fusion with the PEST sequence. This was the trade-off to reach the high temporal resolution. The promoter activity of essential genes in the early stages of infection was also weaker than that of the viral gene promoters in the late stages of infection. It was required long exposure time in primary infected cells, so the exposure time was stetted 1000 ms, which was the bottleneck in the measurement time for a single time point in this time-lapse observation. In the fifth and sixth conditions, the fluorescence of mScarlet-I, expressed by the *polyhedrin* promoter was observed. The difference of these two conditions was only in exposure time. In the fifth condition, I aimed to acquire images where the fluorescence signal did not saturate until the last time point. In the sixth condition, I aimed to acquire images with higher signal intensity to detect the early activation of the *polyhedrin* promoter.

ii Image Analysis Program for Single-Cell Tracking, and Extraction of Fluorescent Profiles of each cell

Time-lapse image data obtained during observations were processed using Python and the macro language of Fiji (ImageJ) (Schindelin et al., 2012) to acquire best-focus images for extracting fluorescence intensity, and the max intensity projection images for tracking. The scripts used to acquire these images were referred to Apprendix3. Fiji (ImageJ) plugin called TrackMate was utilized for single-cell level tracking analysis. When using TrackMate for single-cell tracking, the max intensity images were obtained from 11 Z-stack images in W2. The same process was applied for each time point, resulting in 73 images that were obtained as a new stack image. TrackMate was applied to the acquired stack images, parameters were defined, and XML-formatted data was obtained. Using R, the tracking results were exported as a CSV file, which included centroid

coordinates. At each time point, the region of interest (ROI) for each cell was defined as a 20x20 pixel square centered around the centroid coordinates of each cell at that time point. The median pixel value within this square was defined as the fluorescence intensity for each cell at each time point. Fluorescence intensities were acquired for each of the 11 Z-stack images for each wavelength. This resulted in 11 different fluorescence intensity values for each of the 5 wavelengths, totaling 55 different fluorescence intensity values for each cell at each time point. The sum of the 11 fluorescence intensity values obtained at each wavelength was compiled for 73 time points, creating a fluorescence profile, which was subsequently used in the statistical analysis.

iii Distinction between infected and uninfected cells

Among the fluorescence profiles of all cells, a group of cells obtained under non-infection conditions was defined as the non-infected cell group. For each non-infected cell, the following two values were calculated: 1. The sum of the differences in fluorescence intensity between adjacent time points within the fluorescence profile 2. The maximum value within the fluorescence profile. Histograms were plotted for the distributions of these two values, and the 95th percentile value was obtained for each distribution. Similarly, for all cells, including infected cells, these two values were obtained. Cells were classified as infected or non-infected based on whether either of these values exceeded the threshold defined for non-infected cells.

8. Data availabilities

The code developed for data analysis and processing is available on GitHub at [https://github.com/ToshikiNakanishi/2023_DoctralThesis.git].

Results

In order to obtain a system-level understanding of the hyperactivation of the *polyhedrin* promoter, the gene regulatory network upstream of the promoter activation is crucial. Ono *et al*. (2012) identified ten genes essential for *polyhedrin* promoter activation, which include transcriptional activator, components for viral DNA replication or mRNA transcription. To elucidate the regulatory network consisting of these genes and relationship between the network and *polyhedrin* promoter activation, regulatory activities and resultant *polyhedrin* promoter activity have to be analyzed in a tightly associated manner. First, I developed a time-lapse imaging to monitor the activities using promoter-reporter viruses. Then, the regulatory network was inferred by measuring the promoter activities with genetic perturbations to the essential genes, and the gene playing the pivotal role in the regulatory network was pinpointed. Finally, a genome-wide RNAi screen was performed to further elucidate the system-level regulation of this pivotal gene.

Development of the experimental system for simultaneous monitoring activities of regulatory gene and *polyhedrin* promoter

The baculoviral gene regulation is divided into four stages in the temporal order of immediate early, delayed early, late and very late (Friesen and Miller., 1986). The genes essential for *polyhedrin* promoter activation are immediate or delayed early genes and *polyhedrin* is very late gene. In order to learn the regulatory mechanism of *polyhedrin* promoter activity from the essential gene activity, I determined to measure the expression level of the early genes and explore regulatory logics to activate polyhedrin promoter. Collection of measurements from individual cells is preferred to measurements from bulked cells for statistical learning. Therefore, I aimed to develop a middle-term time-lapse imaging for the upstream and *polyhedrin* promoter activities with a confocal fluorescent microscopy.

An overview of the viral constructs and imaging method developed in this study was shown in Figs. 7A and B. Ten dual-reporter BmNPV bacmids were constructed to express a red fluorescent protein mScarlet-I under control of the *polyhedrin* promoter and nls-sfGFP-PEST, a green fluorescent protein fused with a degradation signal and nuclear localization signal (Pe´delacq *et al*., 2006; Rogers *et al.*, 1986; Wang *et al*., 2003; Niopek *et al*., 2014) under control of promoter of one of the essential genes (*ie1, p47, p143, lef-3, lef-4, lef-5,lef-8,lef-9,lef-10,lef-11*) (Fig. 7A, Details of construct construction were described in Materials and Methods). The fluorescence proteins were selected for brightness and short maturation time for immediate reporting the promoter activities. PEST sequence was fused with sfGFP since short-lived reporter allowed for recapitulating the promoter activity accurately.

Microscopic setup was optimized for amount of data to collect and time to require. The total observation, covering all wells under all conditions for each time point, took approximately 52 minutes. Fig. 7D is an example of the acquired images of the dual-reporter virus-infected cells (*lef-5* dual-reporter). Green fluorescence expressed by the early gene promoter was observed in each cell, followed by the detection of fluorescence expressed by the late gene promoter.

Establishment of the fluorescence intensity extraction method for each cell to achieve a single-cell level of reporter analysis

I conducted the cell tracking during the time-lapse observation to construct the analyze system at the single cell level using the newly developed observation system. The Z-stack images of cells labeled with NucSpot Live Cell Nuclear Stains (Biotium) in the second condition of the time-lapse observation were converted to Max Intensity Projection (MIP) images and these MIP images for 73 time points in each well were converted into Time-Stack images. TrackMate (Tinevez *et al*., 2017 ;Ershov *et al*., 2022), an ImageJ plugin, was utilized to track the movement of each cell in these Time-Stack images. The precise parameter setting was required to conduct the precise cell-tracking for TrackMate. This parameter was set visually for each well, and the precise was evaluated by the movie exported after tracking which showed the trajectory of cells during the infection. The results of tracking for each cell were plotted as the trajectory (Fig. 8A). Each cell was moved randomly during the infection, so the observation system built in this study didn't affect the cell trajectory. The TrackMate results produced coordinate data of each cell and the fluorescent intensity of each cell from each slice of entire stack was extracted using these data with Python. The total fluorescent intensity of the entire stacks was computed from three condition (the 4^{th}- 6^{th} condition). The Time-series data for the fluorescent intensity of each cell for 72 hours during the infection were defined as fluorescent profiles in this study. Considering that the size of each BmN4-SID I cell is about 10 to 20 μm, this method to obtain fluorescence profiles had the possibility of inaccurate fluorescence intensity extraction for cells where the best focus plane was at 0 or 10, which are the edges of the Z stack. The change in the best-focus plane of each cell was plotted on a tile map to confirm the absence of such cells (Fig. 8B).. The results showed this fluorescent intensity extraction method demonstrated the capability to measure the fluorescence intensity of each cell accurately and allowed us to analyze the fluorescent profile at the single cell level.

Identification of infected cells and smoothing the fluorescence profile

NucSpot Live Cell Nuclear Stains-labeled cell tracking provided the advantage of recognizing all cells even in the absence of early-stage virus-derived fluorescence. However, there

were demerits that both of data from infected and uninfected cells were mixed in the obtained fluorescent profile data. To extract only the profile data of infected cells, two criteria were defined to distinguish between infected and uninfected cells. The first criterion involved calculating the changes in RFP fluorescence intensity between near time points for the 73 time points of the fluorescent profiles and summing these changes (referred to as the "Sum of difference between each timepoint" or "SumDiff"). Initially, using the fluorescent profiles obtained from the wells under virus not infected conditions, SumDiff for each cell were calculated and visualized as a histogram (Fig. 9A, upper right). The 95% quantile in this dataset was set as the threshold. Subsequently, a similar calculation was performed for all cells, and a histogram was shown (Fig. 9A, upper right). The threshold values for the non-infected cell data set were applied to the total cell data set, and cells with values higher than the threshold were defined as infected cells according to the first criterion. The second criterion utilized the maximum value from the RFP fluorescent profile of each cell. The maximum values of fluorescent profile of each cell from the well under virus not infected conditions were calculated, and a histogram was shown (Fig. 9A, bottom left). At the same as the first criterion, the 95% quantile was set as the threshold for this dataset. The threshold derived from the non-infected cell dataset was applied to the entire cell dataset (Fig. 9A, bottom right). Cells that showed values higher than the threshold were defined as infected cells according to the second criterion. From these two criteria, infected cells were identified as cells that were classified as infected by either criterion (Fig. 9B). In experiments that combined infection and knockdown, RFP fluorescence profiles used in these criteria was decreased due to the knockdown of essential gene for *polyhedrin* promoter activation. Therefore, it was impossible to distinct between infection and non-infection using this method under the knockdown conditions. Thus, the calculations were performed under non-knockdown conditions, and I defined infected cells by the two criteria and subsequently calculated the rate of infection using the number of virus infected cells divided by the total number of cells (infected cells/Total cells). Since the infection experiments were performed at the same MOI across all the knockdown experiments, a similar infection rate was assumed in the knockdown cells. Infected cells were defined in descending order of the RFP maximum value based on the infection rate calculated for each virus. The fluorescent profiles of cells defined as infected were subjected to a smoothing process using the Locally Weighted Scatterplot Smoothing method (loess) (Cleveland., 1979; Cleveland and Devlin., 1988) (Fig. 9C). These profiles were then utilized to analyze the characteristics of each promoter, the effects of gene knockdown, and the relationship between the activation of essential genes and *polyhedrin* gene promoters.

The activities of the viral promoters

The developed experimental system was used to characterize the activity of each gene promoter region selected for each dual reporter virus. For each virus, fluorescence images were acquired (Fig. 10A) and fluorescence profiles were extracted from individual cells. I averaged the fluorescence profiles at each timepoint per reporter virus to grasp better visualization of representative promoter kinetics (Fig. 10B). The promoters of the essential genes except for *ie1* and *p47* showed earlier peaks compared to that of the *polyhedrin* promoter. The onset of fluorescent profiles of these 8 genes were from 3 h.p.i and the maximum expression was observed by about 20 h.p.i. The promoter activities of *lef-8, lef-9, lef-10, p143* were stronger than those of *lef-3, lef-4, lef-5, lef-11*. The onset of mean *polyhedrin* promoter activation was at 10 h.p.i. and the expression level was plateaued by 30 h.p.i.. It was difficult to obtain the characteristics of *ie1* and *p47* promoter from the average profiles because each promoter of them showed peaks at later time points than *polyhedrin* while other eight genes were able to obtain the characteristics of each promoter. It might be due to the low activation ability of the promoters, with fluorescence intensity increase being more strongly influenced by viral genome replication and copy number than by the promoter activation ability.

Next, I compared these characteristics with those of each viral gene expression obtained by existing methods to verify the accuracy of the system and whether the system is suitable for gene expression kinetic analysis. The expression levels of each gene in virus-infected BmN4-SID1 cells at 0, 6, 12, 18, 24, 36, 48, and 72 hours were measured by RT-qPCR, and the expression profile of each gene was compared with the fluorescence profile of this test to clarify the relationship between the two profiles (Figure 10C). For all gene groups except ie1 and p47, there was a correlation between the two profiles. Both profiles did not differ in the dynamics of the increase in activity, but there were differences in the decay period of the profiles after the peak was reached . This was thought due to differences in mRNA and sfGFP-PEST degradation rates. Although differences were expected due to maturation time, sfGFP, which has a short maturation time, allows not to reveal differences in the profiles.

From these results, we concluded that the promoters in the dual-reporter viruses except for *ie1* and *p47* recapitulated mRNA expression changes of the corresponding endogenous genes. The consistency between the reporter activities and mRNA expression profiles would indicate that the reporter activities at the single-cell level were accurate.

Evaluation of essential gene perturbation system

I evaluated a perturbation system to combine with the reporter viruses. I chose Soaking RNAi method (Mon *et al*., 2015) to perform perturbations to all the combination of the essential gene reporter viruses and dsRNA against them in a high-throughput manner. The dsRNA-treated cells showed the targeted viral genes lower than the untreated cells 24 h.p.i (Fig. 11A). The reduction was ranged from 0.1% to 28%. dsRNA against the ten genes except for *lef-3* and *lef-5* reduced the target gene expression levels to 10% or less. The expression levels of *lef-3* and *lef-5* were reduced to 28% and 18%, respectively. These observations indicated soaking RNAi worked against viral genes, but the knockdown efficiency varied depending on the target genes. It was investigated whether the knockdown led to similar reduction in *polyhedrin* promoter activity as the complete removal of each of the essential genes abolished *polyhedrin* promoter activation (Ono *et al*. 2012). The *polyhedrin* promoter reporter activities of the reporter viruses upon RNAi treatment were reduced（Fig. 11B). Although the caution has to be required to interpret results obtained from *lef-3* knockdown experiments, these results indicated that soaking RNAi was effective for perturbing essential genes to analyze relationships between the essential genes and *polyhedrin* promoter activation.

Relationship between essential gene promoter activity and *polyhedrin* gene promoter activity

With RNAi, the activity of essential genes and polyhedrin promoters were monitored under perturbed and unperturbed conditions for one of the essential genes to determine the relationship between each essential gene and polyhedrin promoter activity. Reporter imaging with RNAi was performed for all combinations of dsRNA and reporter viruses. *Polyhedrin* promoter activity per cell was represented as sum of the RFP fluorescence intensities in individual cells over the time course. The summed RFP intensities were analyzed using LASSO regression (Tibshirani, 1996) as the response variable. GFP fluorescence profile for corresponding cells was used as explanatory variables. Using LASSO regression, the GFP fluorescence intensities at the time points were weighted to fit to explaining RFP intensity of the cell (Fig. 12A). The resultant weights calculated from the sparse modeling indicated importance of each timepoint in explaining *polyhedrin* promoter activation by the GFP profile and also indicate time points carrying redundant information. The weighted information for each timepoint of the model applied for each viral promoter is shown in the bar plot (Fig. 12A). In lef-8, *lef-9*, *lef-10*, the time points with coefficients of which absolute values were large were 0 to 25 hours and after 35 hours post infection. With considering that *polyhedrin* promoter was active 30 hours post-infection, indicative of very late stage of infection (Fig. 10B), the latter was likely to be essential gene activation in secondary infected

cells. The obtained coefficients by comparing the sums of observed RFP fluorescence intensity and ones predicted from the model per cell were validated (Fig. 12B). The correlation between measured and predicted values ranged between 0.35996 to 0.81681. Higher correlation was observed in *lef-8*, *lef-9*, *lef-10,* and *ie1* promoter reporters. These results indicated the relationship between expression profiles of *lef-8*, *lef-9*, and *lef-10* and *polyhedrin* gene expression was closer than other essential genes, and these three genes were determined as the quantitative marker genes for *polyhedrin* promoter activity.

Effect of essential gene knockdown on the activity of essential gene promoter

In order to explore regulatory network upstream of *polyhedrin* promoter activation, the essential genes regulating *lef-8*, *lef-9* and *lef-10* were explored. First, I inferred regulatory relationships among 10 essential genes using the perturbed reporter expression profiles. The effect of RNAi against one of the essential genes was estimated from the GFP fluorescent intensities under the target gene-knocked down using a linear model ($GFPE_{i,t} = Knockdown\ gene + \varepsilon$). The effects of the genetic perturbations on each reporter were summarized in Figs. 13. Representative knockdown effects on each promoter activity were shown in Fig. 13A. The effects of the perturbations to the essential genes revealed two properties of regulation by the genes: regulation of amplitude and timing of expression. For example, the amplitude of *lef-3* promoter activity was affected by RNAi against *ie1* and *lef-9* while RNAi against *p143* delayed the peak of the activation profile. In the activation profile of all essential gene promoters, knockdown of *ie1* had the largest reduction in amplitude on promoter activation, suggesting that *ie1* is the most upstream regulator of essential gene expression. In the activation profile of *lef-5*, *lef-4*, *lef-3*, *p143*, and *lef-11* promoters, knockdown of *p143* delayed the activation peak and increased the activation levels. On the other hand, in the activation profile of *lef-9*, *p47*, and lef-8 promoters, knockdown of *p143* lowered the activation levels.

Each essential gene knockdown showed statistically significant changes of the reporter expression profiles of the essential genes at least one time point (Figs. 13B and C). Fig. 13B showed the coefficients, with blue indicating reduced promoter activity by knockdown and orange indicating increased promoter activity. Fig. 13C showed the p-values for each coefficients. The patterns of differential expression of the reporters upon RNAi (marked by asterisks in Figs. 13B and C) differed in time points. This may indicate that the regulatory relationship among essential genes were changed as the progression of infection stages. These results indicated ten essential genes were in a regulatory relationship with each other and form a regulatory network and this network model was related to the *polyhedrin* promoter activity. To infer the essential gene regulatory network of

which topology changes over the infection, the k-means method was first used to cluster the timepoints based on the effect of knockdown of each essential gene on each essential gene promoter at all 73 timepoints, and then classified each timepoint into six clusters. As a result, all time points were classified as 1-7, 8-15, 16-22, 23-32, 33-47, and 48-73 (Fig. 14A). Since the activation peak of the *polyhedrin* promoter was at 30 hours post-infection (Fig. 10B), it was determined that the first three of these clusters represented the infection stage progression in the primary infection. Next, the regulatory network model at each stage of infection was estimated (Fig. 14B). At the stage of 1-7 h.p.i, knockdown of *ie1* caused the decrease of the genes containing *lef-3*, *lef-4*, *lef-5*, *lef-10*, and *lef-11*, promoter activity, which indicated *ie1* was activator of essential gene expressions and played a primary role as a hub of this network at the stage. This regulatory network model consisted only of genes that functioned as activators among genes. The regulation at this stage of infection initiated from the regulation of *ie1* to *lef-4* and *lef-5* at 3 h.p.i. (Fig. 14D), which suggested that *ie1* was the most upstream in the regulatory relationship of the ten essential genes. At the stage of 8-15 h.p.i., while *ie1* regulated the expression of each essential gene as a hub of the regulatory network, *p143* was found to function as a hub as well. Unlike 1-7h pi, the increase in fluorescence profile by knockdown, meaning that regulation as a repressor, was observed in *p143*, *lef-11*, *lef-10* and *lef-4*. In particular, *p143* functions as an activator for lef-8, *lef-9* and *lef-3*, whereas it functions as a repressor for *lef-11*, which suggested that its function is different from that of *ie1*. This function of essential gene as repressor was confirmed first at 10 h.p.i. in the regulation between *lef-10* and *lef-11* and the function of *p143* was confirmed first at 13 h.p.i.. At the stage of 16-22 h.p.i., while *ie1* regulated the expression of each essential gene as a hub of the regulatory network as with previous time points, the regulatory relationship among other essential genes were more complicated. Hubness of *lef-10*, *lef-11*, and *lef-5* was increased, and their regulation as repressors between essential genes was also increased. The results revealed that the 10 essential genes mutually regulated in a stage-dependent manner and thus composed a dynamic regulatory network.

Identification of essential genes regulating *lef-8, lef-9, and lef-10*

The primary regulators in the regulatory network of *lef-8*, *lef-9*, and *lef-10*, marker genes for polyhedrin promoter activation levels, were explored in the regulatory network. At the stage of 1-7 h.p.i., both *lef-9* and *lef-10* were activated by *ie1* (Fig. 14C). *lef-9* was also activated by other genes, while *lef-10* was activated only by *ie1*. At the stage of 8-15 h.p.i., three genes were activated by *ie1* (Fig. 14C). *lef-10* was activated only from *ie1* as in the previous infection stage, and *lef-9* was activated from the same gene as in the previous infection stage. On the other hand, *lef-9* was newly regulated by *lef-11*. *lef-8* was found to be regulated by *ie1*, *p143*, *lef-9*, and *lef-5* for the first time in this infection stage. At the stage of 16-22 h.p.i., three genes were activated by *ie1* as in the previous

infection stage (Fig. 14C). *lef-8* regulation was the same as in 8-15 h.p.i., but the regulatory relationship of *lef-10* and *lef-9* was changed. *lef-9* regulation from genes except *ie1* changed from activation regulation to repression regulation. This change was caused by a delay in the activation peak affected by knockdown. The regulation of *lef-10* from genes except *ie1* was first observed at this stage of infection and was repressive. This regulation was also caused by a delay in the activation peak due to knockdown. Of the three genes, *lef-8* was only regulated by activation, while *lef-10* was only regulated by all except *ie1* to accelerate the timing of the activation peak. *lef-9* was regulated by both. While *ie1* was responsible only for the regulation of activation, *p143* was responsible for both the regulation of activation and the acceleration of the timing of the activation peak. In addition, *lef-8* and *lef-9* were regulated by each other, while *lef-10* was not regulated by *lef-8* and *lef-9* (Fig. 14C, Fig. 14E). These results suggested that the common activation regulation between *lef-8* and *lef-9*, and the acceleration of the activation peak between *lef-9* and *lef-10* were regulated by independent mechanisms. The results suggested that *ie1* was responsible for only the activation regulation of the two regulations, while *p143* was responsible for both of them. In addition, *lef-10* is upstream of *lef-9* and *lef-10* in the regulation of the acceleration of the activation peak timing. These results suggested that understanding the two regulatory systems of *p143* is necessary to understand the regulatory systems of the three genes.

Genome-wide screening of genes that regulate *p143*

It was suggested that *lef-8*, *lef-9*, and *lef-10* have two independent regulatory systems in the essential gene regulatory network that varies depending on the infection stage, and *p143* was identified as a gene that functions in both of these two regulatory systems. In order to clarify which BmNPV system corresponds to each of the two regulatory systems, I screened all genes for genes that control *p143* and investigated the functions of the identified genes. I conducted *p143* dual-reporter bacmid infection experiments under each gene knockdown condition and analyzed the obtained GFP fluorescent profile data. I then performed linear modeling and calculated the effects of each gene knockdown on the activation dynamics of the *p143* promoter at each time point. Among the 143 genes, 54 genes showed a significant effect on *p143* during the infection (Fig. 15A). To elucidate the relationship between these 54 genes and p143, they were classified into six clusters based on the cosine similarity of the coefficients and profile distances at all time points.. The obtained dendrogram confirmed the clustering to be at six, and using the k-means method, I re-clustered the 54 genes into six categories (Fig. 15B). In cluster 1, a decline in the *p143* promoter activity peak was observed, with 22 genes classified (Fig. 15C). Cluster 2 revealed the delay in the *p143* activity peak and increased its intensity, with six genes, none of which were classified as essential genes but were reported to be involved in DNA replication. In Cluster 3, a decline in *p143*

promoter activity was observed at the early stage of infection. Cluster 4 showed a delay in the activation peak and a reduced rate in the activation rise. *p143* was classified in this cluster. Cluster 5 displayed no significant features. In Cluster 6, considerable changes in activity, including peak intensity from initiation, were observed. *ie1* was categorized in this cluster. From these results, it was suggested that *p143* was under the control of at least six patterns. The genes that caused only a delay in the activation peak were mainly related to DNA replication, which suggested that the acceleration of the activation peak in the independent regulatory systems to *lef-8*, *lef-9*, and *lef-10* could be related to DNA replication. The results also revealed that there are two types of regulatory patterns in activation regulation: a reduction in the activation peak and a reduction in the rate of activation onset.

Discussion

BmNPV hyperactivates *polyhedrin* promoter at extremely high levels and has been serving as recombinant expression vector using silkworms. Although the genes essential for activation of *polyhedrin* promoter were genetically identified in previous study (Ono *et al*., 2012), the roles of the essential genes remain elusive. In this study, I aimed at revealing the quantitative relationship between the essential genes and *polyhedrin* promoter activation and the regulatory network of the essential genes. To this end, the quantitative imaging system using promoter-reporter viruses and automation of the confocal laser microscopy were implemented. Additionally, the system in combination with RNAi, as a means of perturbation to the regulatory network, and statistical modeling were utilized. My findings involve (1) that expression profiles of *lef-8, lef-9*, and *lef-10* correlated with *polyhedrin* promoter activation most in the essential genes, (2) that *p143* as well as ie1 regulated the three genes during transitions of infection stages, and (3) that *p143* expression is regulated by potentially diverged functions including DNA replication.

The differential relationship between the essential genes and *polyhedrin* promoter activation and the infection stage-dependent change of the network topology strongly indicated a hierarchical and dynamic structure of the essential gene regulatory network. Suppose that *polyhedrin* promoter responds to a regulator in a quantitative manner, *polyhedrin* promoter is activated on the expression of the regulator and the correlation between these two events should be high. As the requirement of the essential genes in *polyhedrin* promoter activation was genetically evident (Ono *et al.*, 2012), the varied degree of correlations between the expression profiles and *polyhedrin* promoter activation indicated *polyhedrin* promoter activation levels and the expression of the essential genes were mathematically nonlinear for the essential genes other than *lef-8, lef-9*, and *lef-10*. The nonlinear relationships between them, in turn, indicate that the regulatory mechanism comprised of the essential genes is govern by a network rather than a sequential cascade. Perturbation analyses and statistical modeling indeed inferred a densely connected regulatory network (Fig. 14). And the topology of the network (connectivity, i.e. all the regulations among the genes at a time point) changed over the observed time points. Such dynamic behaviors in an upstream regulatory layer can generate nonlinearity in controlling downstream responses. On the other hand, relatively linear relationship between expression profiles of *lef-8*, *lef-9*, and *lef-10* and *polyhedrin* promoter activation levels might indicate that activities of the three genes are directly translated to *polyhedrin* promoter activation or that the three genes are *polyhedrin* share a common regulator. However, the regulatory network upstream of *polyhedrin* promoter activation seems not that simple. *lef-8* and *lef-9* are genes associated with transcription of late viral genes as subunits of RNA polymerase, and *lef-10* is a gene associated with DNA replication. It was shown that DNA replication and expression of viral RNA polymerase are the two key events required for the transition from early gene expression

to late genes in the baculovirus (Okano *et al*., 1999). If *polyhedrin* promoter activation is directly regulated by these functions, correlations of the expression profiles of other genes involved in these functions should have been high. For example, correlations of expression profiles of *lef-4* or *p143*, a subunit of RNA polymerase or DNA helicase, with *polyhedrin* promoter activation were 0.36 and 0.43, respectively. Although the possibility that each of *lef-8*, *lef-9, lef-10* and *polyhedrin* is regulated by a common regulator(s) cannot be ruled out, these reporter expression profiles are not so similar enough to assume regulation by a shared regulator. If this possibility is true, we need to hypothesize a regulator of which regulation is target gene dependent. In the perturbation analyses, regulators between *lef-10* and, *lef-8* and *lef-9*, were different, which suggested existence of two signaling pathways to *polyhedrin* promoter activation. Candidates for the common regulator were limited to *ie1* and *p143*. ie1 is an immediate early gene and is known to function as a transactivator of early genes. *ie1* was identified as an activator of other essential genes in this study except for *p47* of which expression levels was difficult to analyze due to its weak promoter strength. On the other hand, *p143* appeared to control downstream genes more dynamically. The effects of *ie1* regulation were observed throughout the primary infection, confirming that *ie1* regulation is extremely important in the progression of the viral infection stage and affects all subsequent infection stages, including early gene expression levels, DNA replication levels, and late gene expression levels. Compared to *ie1*, regulation of essential gene expression by *p143* changed with infection stage progression. In the early stages of infection, *p143* was regulated as an activator similar to *ie1*, but from the middle stage of infection, it was regulated to accelerate the peak of activation. The time point at which this change was observed (13h.p.i.) was applied to the time schedule of the infection stage progression in the previous study (Friesen, 1997), it was just at the time point when the gene expression peaked out in the early stage and the DNA replication peaked. This time point was also the time point at which the RNA polymerase used for transcription switched from host-derived RNA polymerase to virus-derived RNA polymerase (Morris and Miller, 1994). These findings suggested that the change in the regulatory system by *p143* could be caused by the progression of infection from early to late stage, or that the change in the regulatory system by *p143* could cause the progression of infection from early to late stage. Of the two regulatory systems by *p143*, *lef-8* regulated only activation, *lef-10* regulated only the acceleration of the activation peak, and *lef-9* was regulated by both. This means that these two regulatory systems are independent of each other. In other words, *p143* is responsible for the regulation of essential genes by different systems in the early infection period and in the transition period from early to late infection, respectively. These two independent regulatory systems were considered to be the key regulatory systems for activation of the *polyhedrin* promoter, since the strongest reduction in *polyhedrin* promoter activation was observed in the *p143* knockdown condition (Fig. 11B), which had a lower knockdown efficiency than *ie1* in this study(Fig. 11A). To further understand the two regulatory systems of p143 in detail, in this study, viral genes that control

p143 promoter activity were screened from all genes that BmNPV possesses. The results showed that when genes related to DNA replication were knocked down, a delay in the activation peak was observed, similar to that observed during *p143* knockdown. This result supported the earlier mentioned two possibilities in the regulatory system of *p143* during mid-infection, that the transition from early to late infection triggers this regulatory system. In other words, the progression of the infection stage from early to late, which DNA replication plays a role in, causes a switch in the *p143* regulatory system, which may lead to *polyhedrin* promoter activity. The hypothesis of the *polyhedrin* promoter activity system centered on the essential gene regulatory network model identified in this study is illustrated in Fig. 16. *Polyhedrin* promoter activity is initiated by the activation of essential gene expression by ie1, after which essential genes regulate each other. One checkpoint is reached during the transition from the early to the late stage. In the case of a successful transition, changes in the regulatory network system centered on *p143* occur, and the timing of *lef-9* and *lef-10* promoter activation is accelerated. The timing of these events and the timing and amount of *lef-8* expression determine *polyhedrin* promoter activity. This hypothesis provides the first link that DNA replication is essential for *polyhedrin* promoter activity and the relationship between the expression of virus-derived RNA polymers and the timing of DNA replication.

In summary, this study provided insights into the structure and dynamics of the regulatory network of genes essential for *polyhedrin* promoter activation in BmNPV infection. Although further detailed analyses are required to validate my hypotheses discussed above, this study unveiled the system-level control of *polyhedrin* promoter for the first time. The findings in this study would be served as a foundation to deepen understanding of baculoviruses for basic and applied biology.

General Discussion

In BmNPV, in which gene expression changes in accordance with the infection stage progression, an experimental system that enables temporal control of gene expression and a system that can acquire and analyze temporal information with high resolution is essential for unraveling the regulatory system among genes. In Chapter 2, I revealed that SMASh, a drug-dependent protein degradation induction system developed in mammalian cells as an experimental system that enables temporal control of gene expression, could be used in BmN cells as well. While concentration-dependent induction of proteolysis was possible as in mammalian cells, this induction was irreversible. I determined that it would be difficult to use this system to control gene expression at a specific stage of infection because this system cannot control the amount of protein cleaved by proteases prior to drug induction in addition. In Chapter 3, gene expression was controlled by knockdown. Soaking-RNAi method (Mon *et al*., 2013) used in this study can efficiently perturb the expression of many genes, so perturbation of the expression of 10 essential genes and perturbation of all viral genes could be easily performed. The results suggest that *ie1* and *p143* play an important role as hubs in the *polyhedrin* promoter activation system. In particular, regulation by *p143* was suggested to change depending on the infection stage.

Although SMASh was not used in this study, the possibility remains that SMASh could be used in temporal expression perturbations of *p143* and other early genes of the virus. The problem with the SMASh system, that proteins already cleaved by proteases cannot be controlled, may be ignored if the target protein is unstable. The half-life of a gene called Early Region 1A of herpesviruses is as short as 30 minutes (Spindler *et al*., 1984), and it is thought that SMASh can be used for such genes to accurately assess the effects of gene expression perturbation after drug addition. In baculoviruses, there are few examples of genes whose protein half-life has been examined, and this possibility needs to be verified. In baculoviruses, there are few examples of genes whose protein half-life has been examined, and this possibility needs to be verified, however, if it can be proved that the half-life of the early genes is short, as in herpesviruses, gene expression perturbation by the SMASh system may be used in the observation and analysis system constructed in Chapter 3. If this system could be used with *p143*, it would be possible to suppress only the acceleration control of the activation peak by adding the drug just after causing only the activation control, of the two types of control by *p143*. It would be possible to verify how much each of the two controls contributes to *polyhedrin* promoter activity by observing the effect of this on *polyhedrin* promoter activity.

The temporal analysis tool developed in this study is very useful for understanding the regulatory network of viruses at each stage of infection. In addition, this system using live imaging may be further developed when the dynamics of host cells can also be captured by imaging. If we

can create a time schedule of infection dynamics by adding the time information of post-infection events of the host to the data on the relationship between the progression of the infection stage and the time elapsed after infection obtained in this study, we can expect to be able to understand conventional RNA-seq and proteomic information by comparing it with the infection stage. I believe that by organizing the vast amount of information that has been obtained in previous studies on the basis of temporal information, we will be able to dramatically advance research for understanding the *polyhedrin* promoter activity system in baculoviruses and, subsequently, baculovirus research. I consider this study to be a research that has established a system that will serve as a foundation for this purpose.

Figures

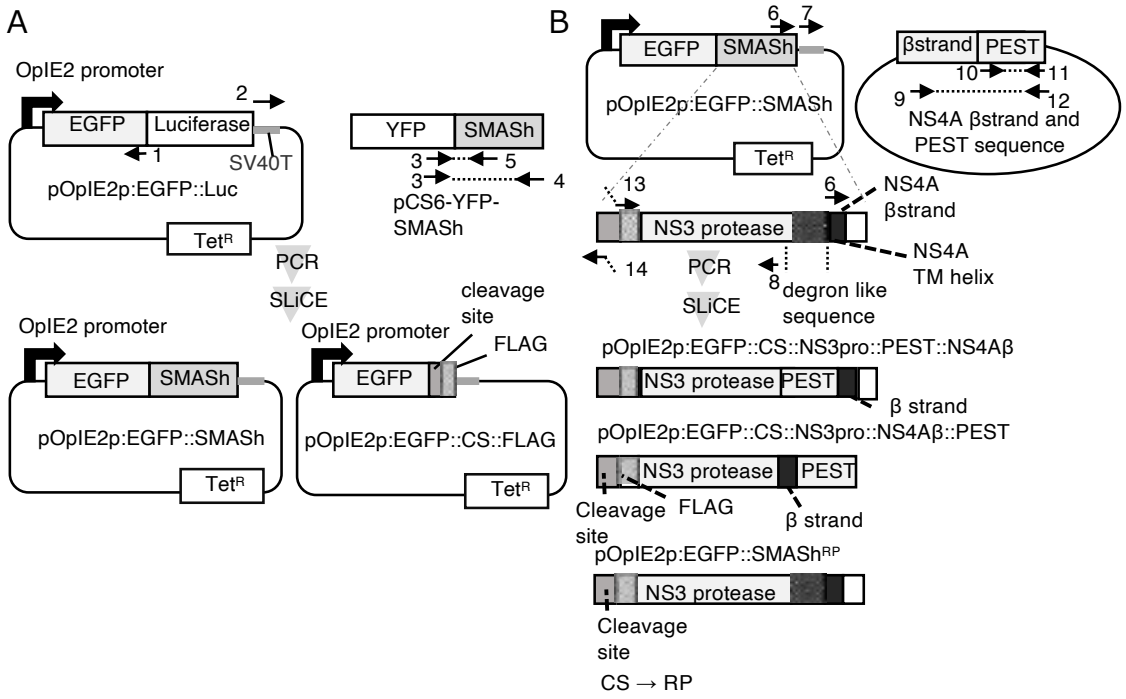

Figure 1. Scheme to construct SMASh reporter plasmids

The PCR amplified fragments to construct the indicated plasmids were assembled using Seamless Ligation Cloning Extract (SLiCE). The primers indicated by numbered arrows are listed in Table 10.

(A) Construction of pOpIE2:EGFP::SMASh and pOpIE2:EGFP::CS::FLAG. "SV40T" and "TetR" indicate the SV40 terminator and tetracycline resistance gene, respectively.

(B) Construction of mutated SMASh reporter plasmids. Amplification of DNA fragments was carried out using these primer pair combinations: 6 with 8; 7 with 8; 9 with 12; 10 with 11; and 13 with 14.

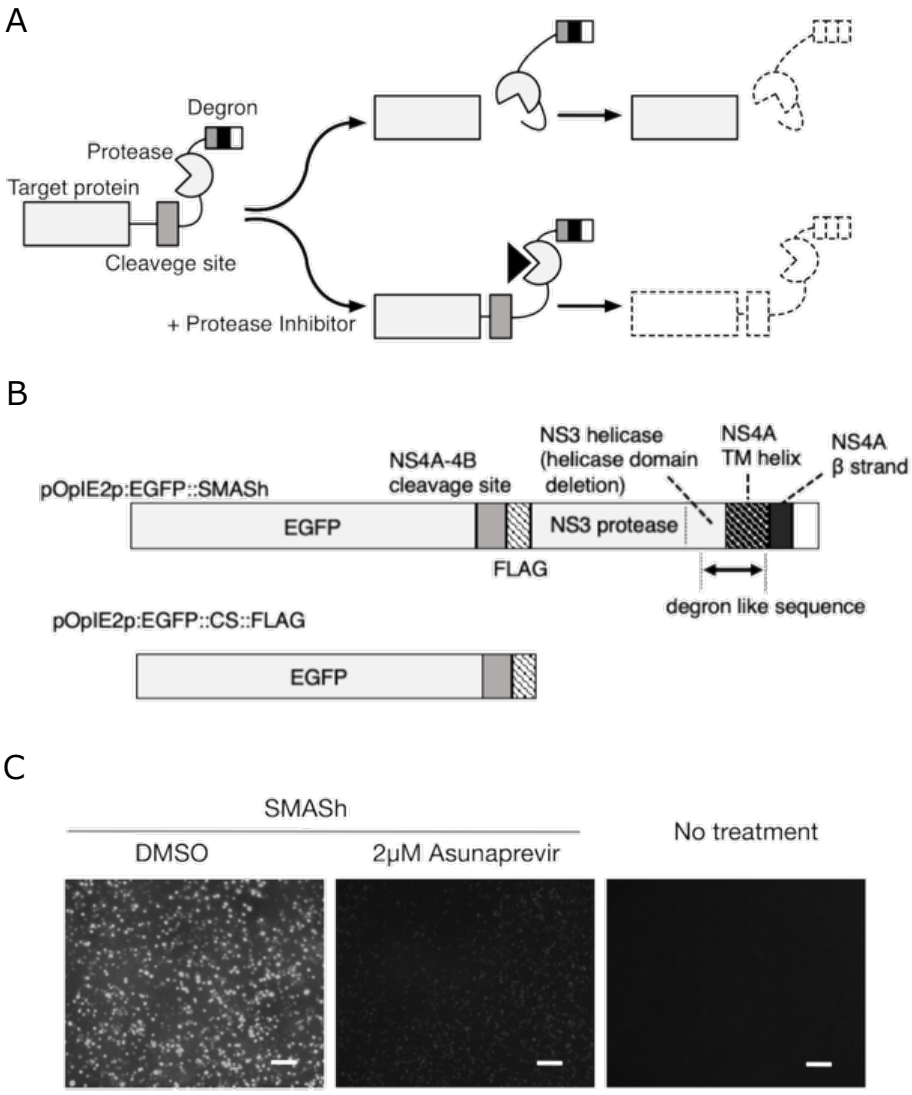

Figure 2. The SMASh tag functions in BmN Cells.

(A)Schematic representation of the SMASh system. The SMASh tag, consisting of protease and degron domains, is used in fusion with a target protein to enable shutoff of the newly translated target protein upon treatment with a protease inhibitor.

(B)Schematic of the constructs used in this study. "TM" denotes transmembrane domain. "OpIE2p:" indicates that the subsequent sequence is driven by the OpIE2 promoter, and “SV40T” indicates the SV40 terminator. "::" signifies translational fusion of the entities it separates.

(C)Substantial decrease in GFP fluorescence by the SMASh system. BmN cells transfected with pOpIE2p:EGFP::SMASh were treated with 2 μM Asunaprevir or DMSO for 72 hours and imaged. Non-transfected BmN cells were used as a negative control. Scale bars represent 200 μm.

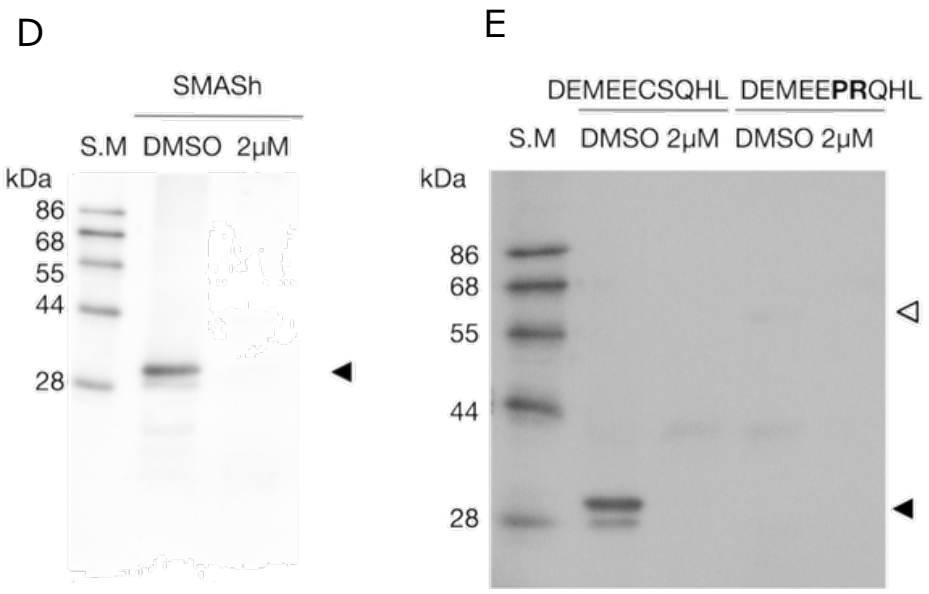

Figure 2.

(D)Protein degradation by the SMASh system in BmN cells. Western blot analysis was performed with an anti-GFP antibody on extracts from pOpIE2p:EGFP::SMASh-transfected BmN cells treated with 2 μM Asunaprevir or DMSO for 72 hours.

(E)Verification of NS3 protease cleavage at the recognition site in BmN cells. The cleavage sequence within the SMASh tag was altered to test NS3 protease cleavage specificity. Western blotting with an anti-GFP antibody was carried out on extracts from BmN cells transfected with either pOpIE2p:EGFP::SMASh (cleavage sequence: DEMEECSQHL) or pOpIE2p:EGFP::SMAShRP with a mutated cleavage sequence (DEMEERPQHL), treated with 2 μM Asunaprevir or DMSO for 135 hours. "S.M." stands for size marker. Filled and open arrows indicate the cleaved and uncleaved forms of EGFP::SMASh, respectively.

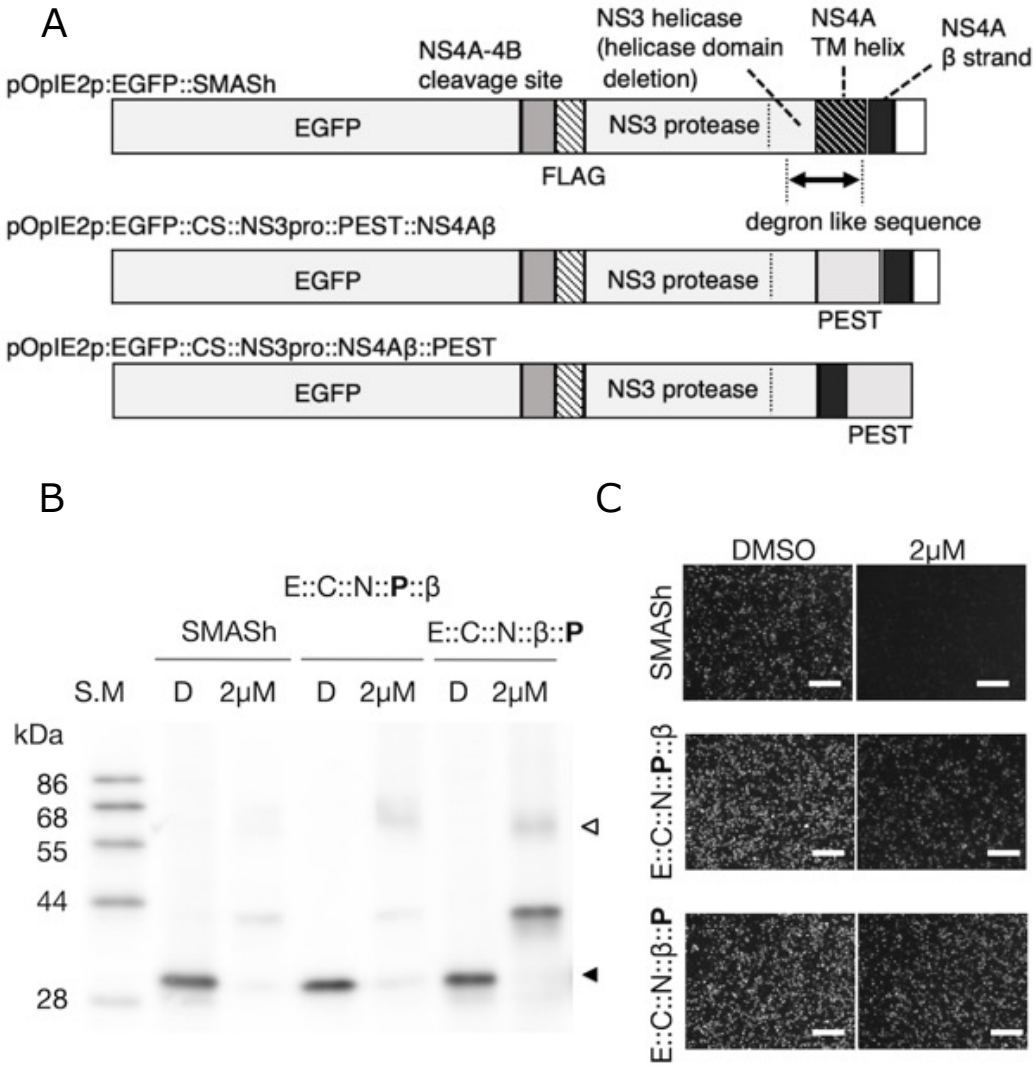

Figure 3. Degron in SMASh Tag Outperforms PEST in Protein Degradation Efficiency.
(A)Schematic of the SMASh constructs for degron performance comparison. These constructs express proteins under the OpIE2 promoter.

(B)Enhanced degradation by the SMASh tag degron. Western blot analysis was carried out using an anti-GFP antibody on cell lysates from BmN cells transfected with pOpIE2p:EGFP::SMASh (referred to as SMASh), pOpIE2p:EGFP::CS::NS3pro::PEST::NS4Aβ (abbreviated as E::C::N::P::β), or pOpIE2p:EGFP::CS::NS3pro::NS4Aβ::PEST (abbreviated as E::C::N::β::P). Cells were treated with either 2 μM Asunaprevir or DMSO for 135 hours. S.M., the size marker. “D” and “2μM” indicate cell extracts treated with DMSO and 2 μM Asunaprevir, respectively. Filled and open arrows identify the cleaved and uncleaved forms of fusion proteins. (C)Increased GFP fluorescence with substitution of the SMASh tag degron with PEST. BmN cells, transfected with the aforementioned plasmids, were treated with 2 μM Asunaprevir or DMSO for 72 hours before imaging. Scale bars represent 200 μm.

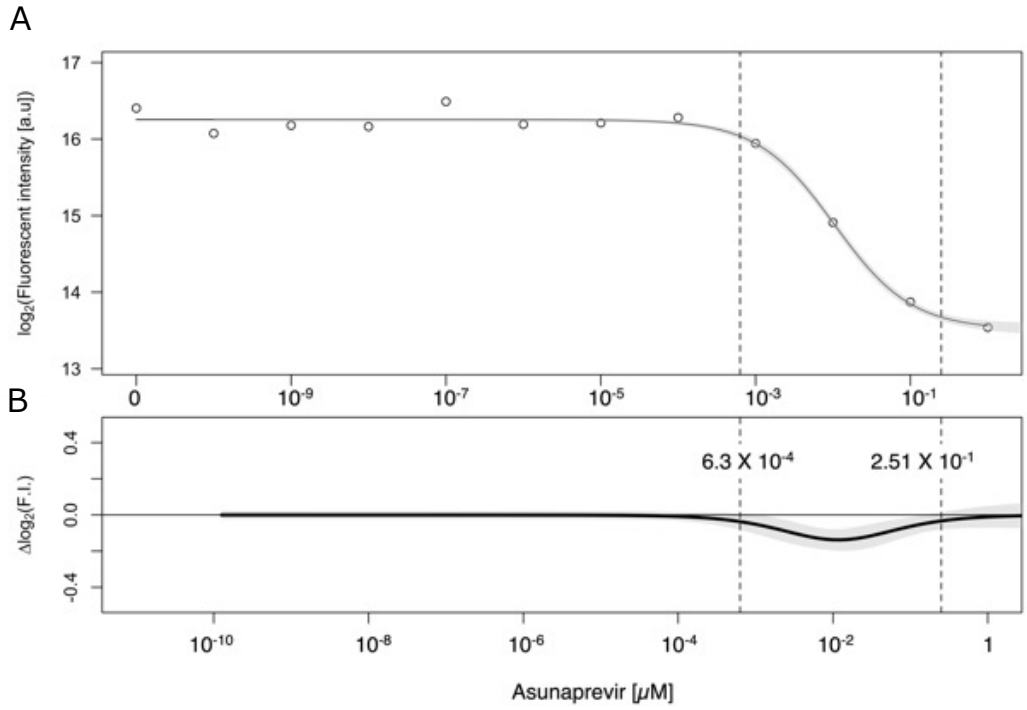

Figure 4. Concentration-dependent function of SMASh in BmN Cells.

BmN cells were transfected with pOpIE2p:EGFP::SMASh or pOpIE2p:EGFP::CS::FLAG and subjected to varying concentrations of Asunaprevir (10^{-10} to 10 μM) for 120 hours.

(A)Quantitative decrease in GFP fluorescence as a function of Asunaprevir concentration. Fluorescence intensity for cells transfected with pOpIE2p:EGFP::SMASh 120 hours post-transfection and treatment was quantified using a fluorescent microplate reader. The data were modeled to a dose-response curve, with measurements replicated across five wells for each condition, and the experiment was replicated independently twice.

(B) Identification of the effective Asunaprevir concentration range for modulating protein expression. Differential slope values from the GFP fluorescence in (A) were statistically analyzed using a t-test to evaluate the null hypothesis of no difference from zero.

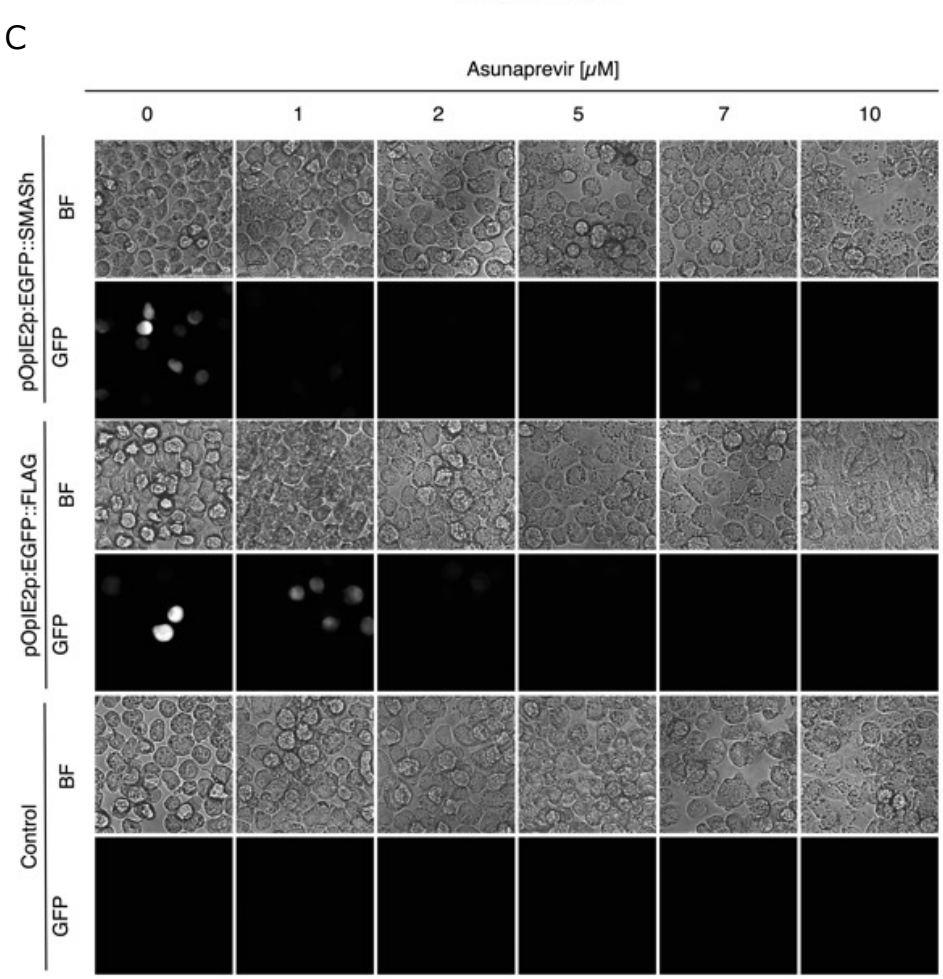

Figure 4.

(C) Cytotoxic effects at high concentrations of Asunaprevir in BmN cells.

Representative microscopic images of BmN cells transfected with pOpIE2p:EGFP::SMASh or pOpIE2p:EGFP::CS::FLAG after treatment with the indicated Asunaprevir concentrations. The scale bar represents 75 μm and is applicable to all the images. BF denotes bright field images.

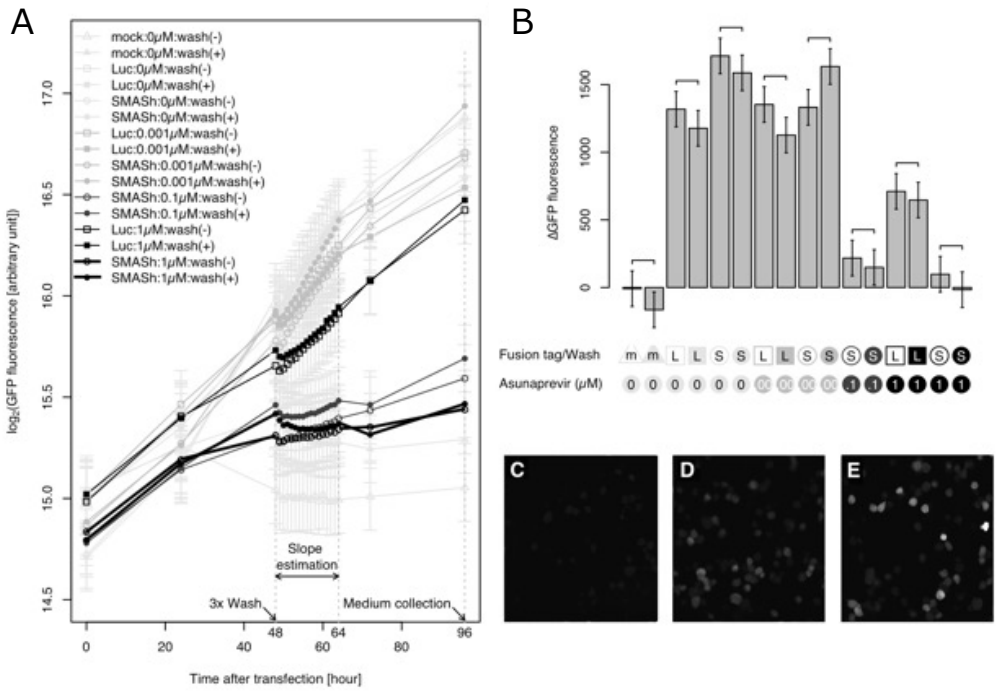

Figure 5. Irreversibility of shut-off in BmN cells post media exchange. BmN cells were transfected with pOpIE2p:EGFP::SMASh or pOpIE2p:EGFP::Luc (a fusion protein of EGFP and firefly luciferase) and treated with indicated Asunaprevir concentrations for 48 hours. Subsequently, the cells were washed and continued to be cultured in TC-100 medium supplemented with 10% FBS. GFP fluorescence was then measured at one-hour intervals for 16 hours to estimate the slope of fluorescence change ("Slope estimation"). At 96 hours post-transfection, which corresponds to 48 hours after medium exchange, the supernatant was collected for the analysis shown in I to I. (A)Experimental design and GFP fluorescence measurement. GFP fluorescence was quantified using a fluorescent microplate reader, and the data were fitted to a linear model. The lines and bars represent the estimated means and standard errors, respectively. Measurements were replicated across four to six wells per condition, with the entire experiment independently replicated twice. (B)Asunaprevir removal did not affect the slope of GFP fluorescence. Differential slope values derived from GFP fluorescence in washed and unwashed cells were statistically evaluated using a t-test, with p-values for all comparisons exceeding 0.05. Sample information is denoted as follows: triangle for untransfected cells (mock, m), rectangle for pOpIE2p:EGFP::Luc (Luc, L), and circle for pOoIE2p:EGFP::SMASh (SMASh, S). The symbols' color corresponds to Asunaprevir concentrations, with the code "0" for DMSO only, ".00" for 0.001μM, ".1" for 0.1μM, and "1" for 1μM. (C)-(E) GFP fluorescence in pOpIE2p:EGFP::SMASh-transfected cells cultured in media obtained from the cultures of Asunaprevir-treated, washed, and subsequently cultured cells. The collected media was diluted to half-strength with TC-100 medium supplemented with 10% FBS. Representative images for the cells cultured with the collected medium from the cells treated with 1μM (C) or 0.001 μM (D) Asunaprevir, or DMSO (E).

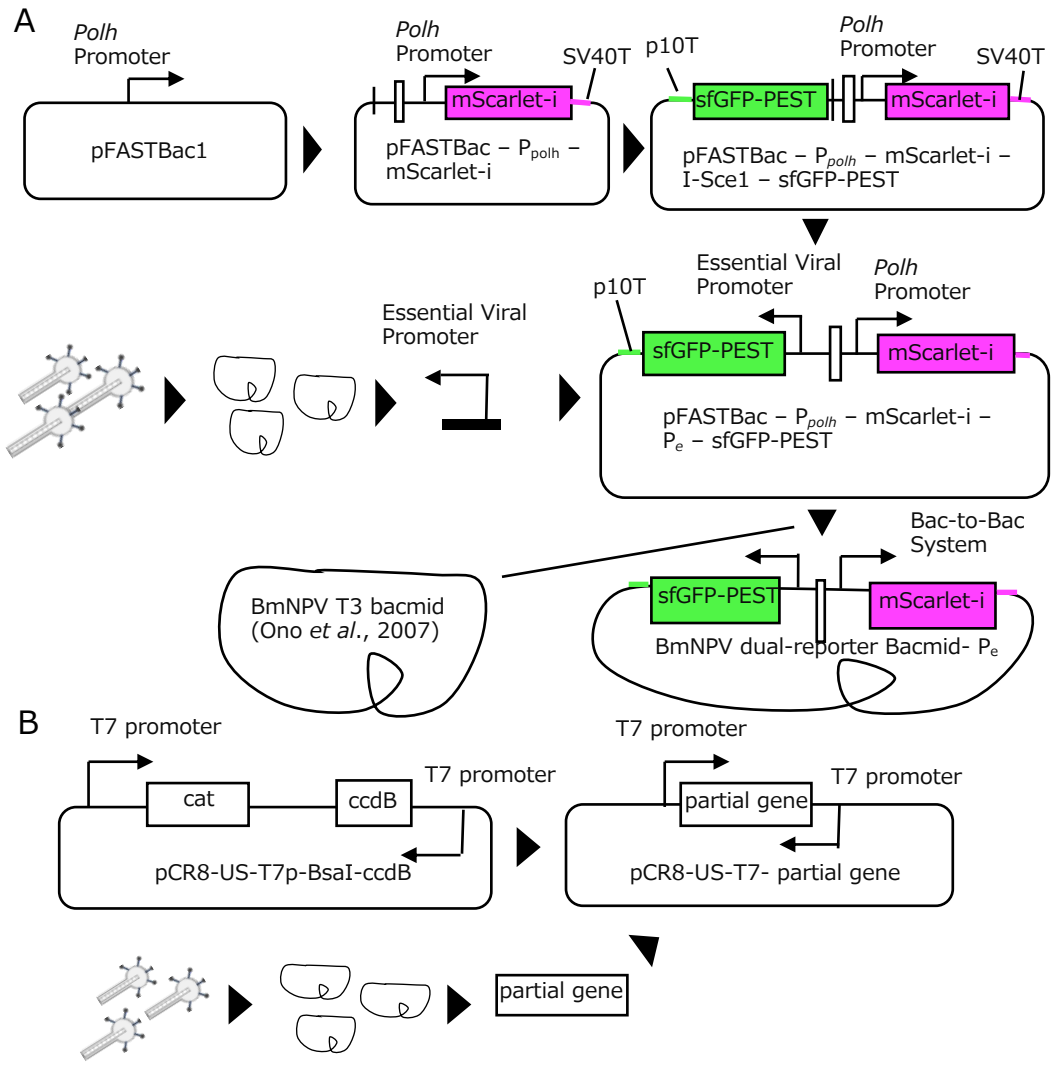

Figure 6. Construction of dual reporter bacmids and knockdown template DNA plasmids

(A)Construction of dual reporter bacmids. The dual reporter bacmid was constructed with the Bac-to-Bac system. First, pFastabac-P_{Polh}-m Scalet-I-I-Sce1-sgfGFP-PEST was constructed by cloning mScarlet-I and isce1-sgfGFP-PEST amplified by PCR, sequentially, into pFasTBac1 using SLiCE, respectively. This plasmid was treated by I-Sce1 and each essential gene promoter sequence amplified by PCR from the viral supernatant was cloned by SLiCE to construct pFastBac-P_{Polh}-mScalet-I-P_e-sfGFP-PEST. Finally, a dual reporter cassette was introduced into the bacmid by the Bac-to-Bac system. (B)Construction of dual reporter bacmids and knockdown template DNA plasmids**.** Knockdown template DNA plasmids were prepared by designing primers complementary to the T7 promoter of pCR8-US-T7p-BsaI-ccdB and cloning by the SLiCE method the template DNA sequences for the dsRNA of each gene amplified by PCR using primers designed to add the T7 promoter sequence at the end from the viral supernatant and the vector fragment that was amplified.

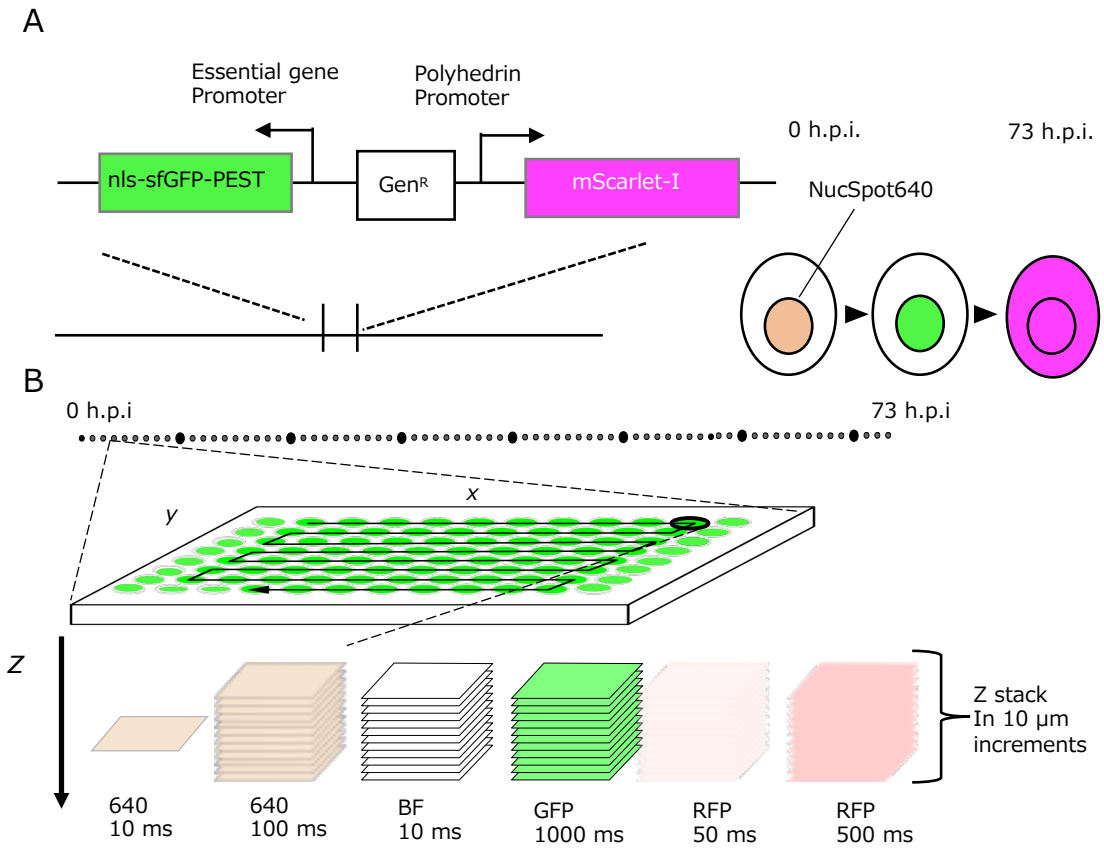

Figure 7. Construction of a new method for imaging analysis

(A) Schematic diagram of the dual reporter expression cassette in the dual reporter bacmid.

(B) Schematic diagram of the imaging conditions for the constructed time-lapse observation. Each infected cell was observed at 73 timepoints every hour for up to 72 hours after infection; images were acquired with six different excitation lights, filter sets and exposure time sets. A 96-well plate was used for observation.

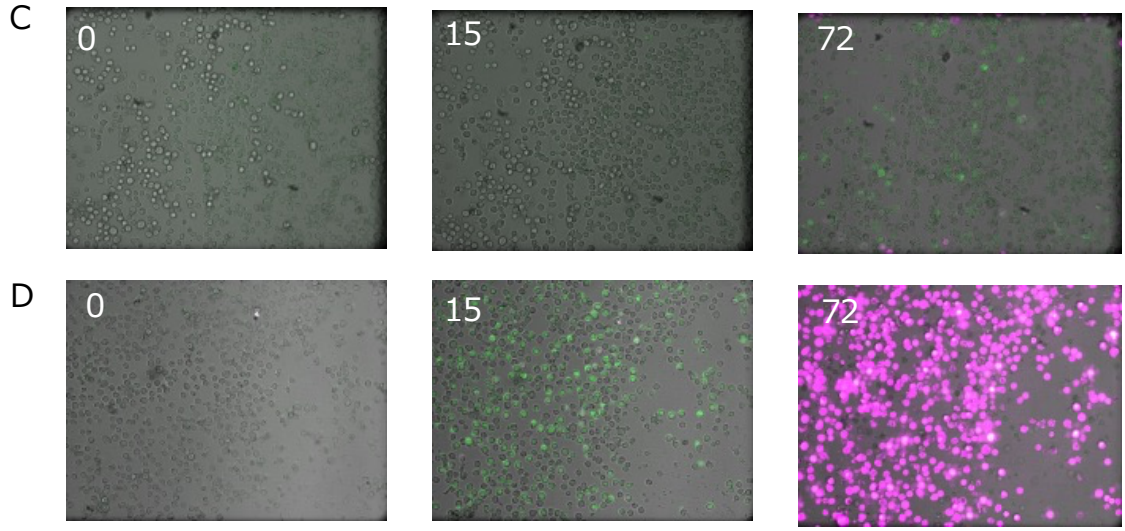

Figure 7.

(C)Cytotoxicity was confirmed when cells were stained with Hoechst 33342. Cytotoxicity was observed when Hoechst 33342 was used for cell tracking and short-wavelength excitation was applied during time-lapse observation, and red fluorescence expressed by the *polyhedrin* promoter was not observed in cells that exhibited GFP fluorescence. Images showed BmN4-SID1 cells infected with *lef-5* dual reporter bacmid with MOI=1
(D)Staining with Nucspot 640 was suitable for time-lapse observation of infected cells. No cytotoxicity was observed when Nucspot 640 was used for cell tracking and excitation by long wavelengths. Images showed BmN4-SID1 cells infected with *lef-5* dual reporter bacmid with MOI=1

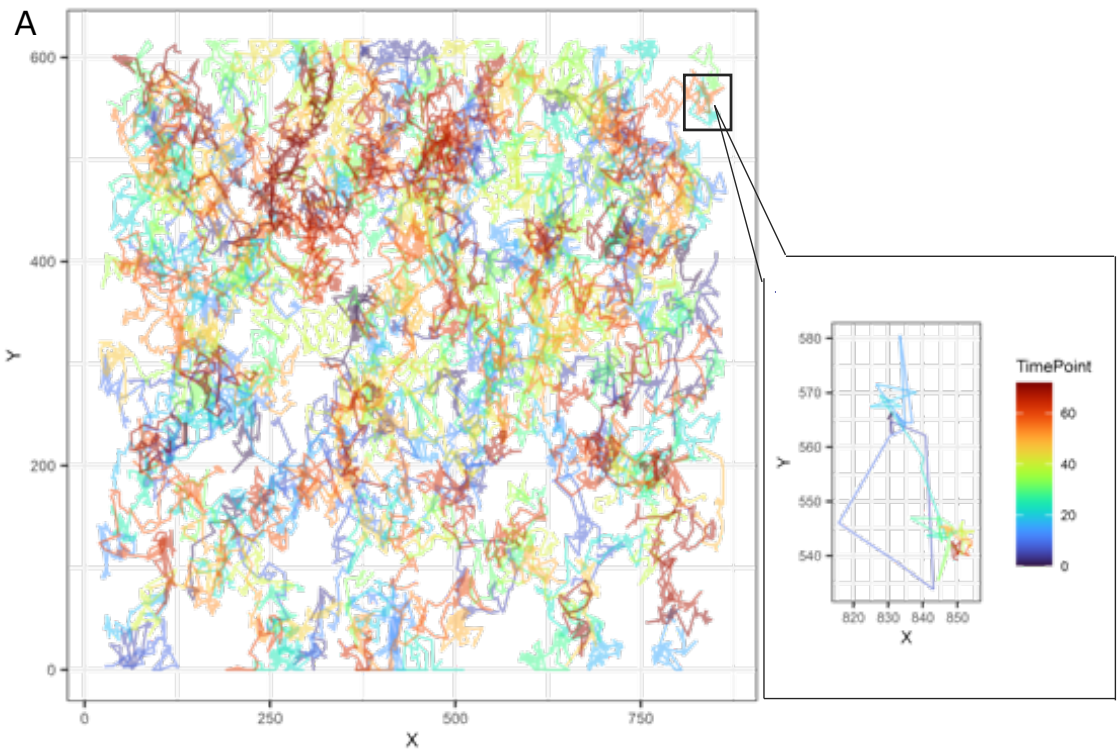

Figure 8. Verification of accuracy of single-cell tracking

(A) Accurate observation and cell tracking without microscope-derived bias was successfully achieved**.** The figure summarizes the coordinate data for each time point obtained by TrackMate and showed the trajectory of each cell at all time points. Each color indicates the result in each cell. In addition, the result for one of the cells is shown zoomed in. In this figure, elapsed time is indicated by color

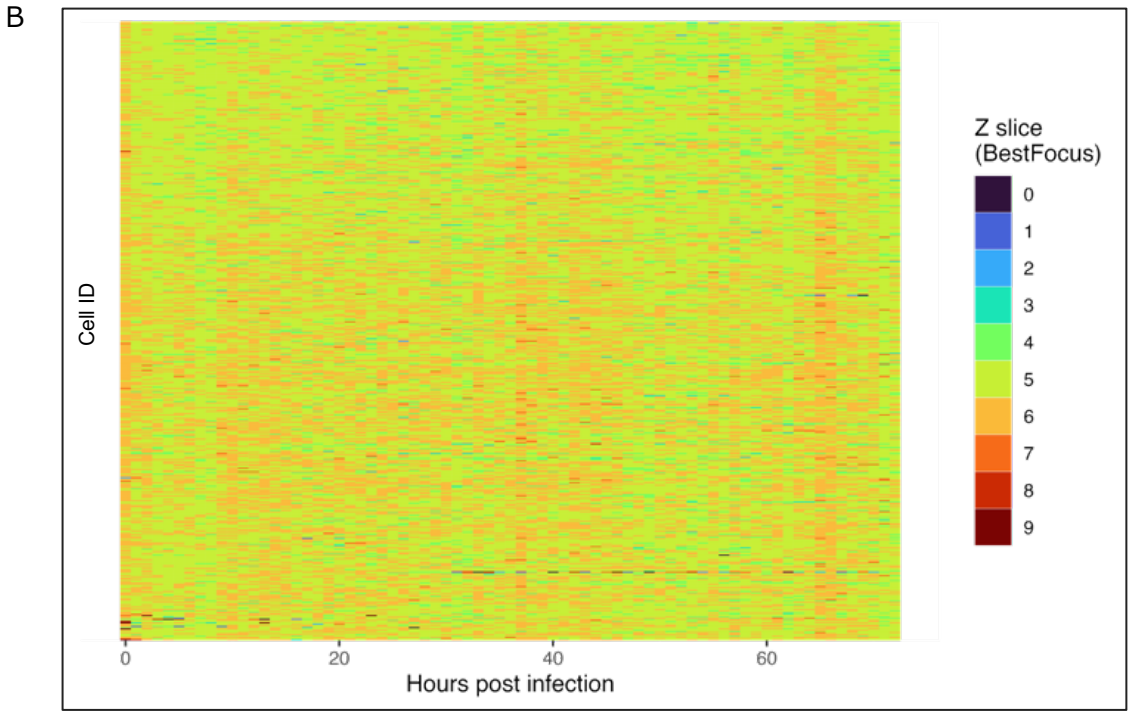

Figure 8.

(B)The fluorescence intensities of all cells were accurately obtained by acquiring Z-stack images. Heat map showing changes in the best-focus slice of each cell. The vertical axis indicates each cell and the horizontal axis indicates the time elapsed since infection. The best focus surface was calculated by a script using a Laplacian function and obtained for each cell and each time point.

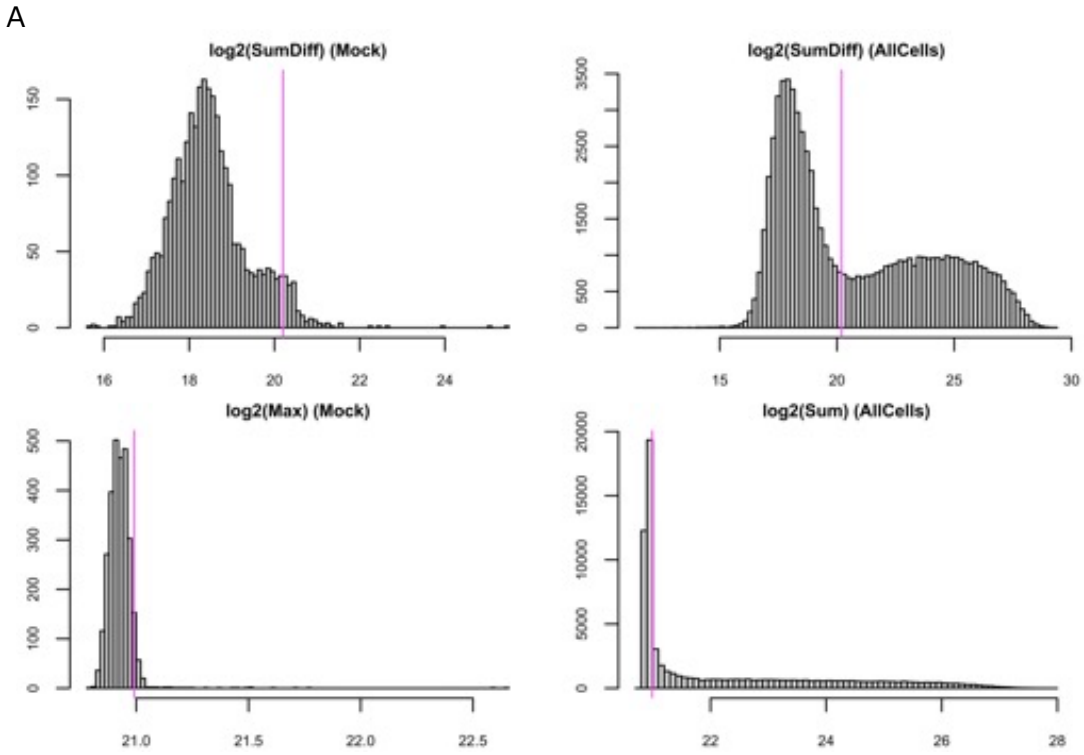

Figure 9. Identification of infected cells and extraction of fluorescent profile per each cell

(A)The tow criterion used to identify infected cells. Histogram of only uninfected cells and each threshold and its application to the histogram of all cells to obtain the two thresholds defined for differentiating infected and uninfected cells, respectively. Histograms. The top two histograms are the histograms obtained by adding up the differences between adjacent timepoints in the fluorescence profile calculated by the first criterion. The left histogram is for non-infected cells only. The right histogram is for all cells, including infected cells. The bottom two histograms show the maximum fluorescence profile calculated by the second criterion. The histogram on the right is for non-infected cells. The histogram on the right shows the histogram for all cells including infected cells. In all graphs, the magenta vertical lines indicate the threshold values calculated for each criterion.

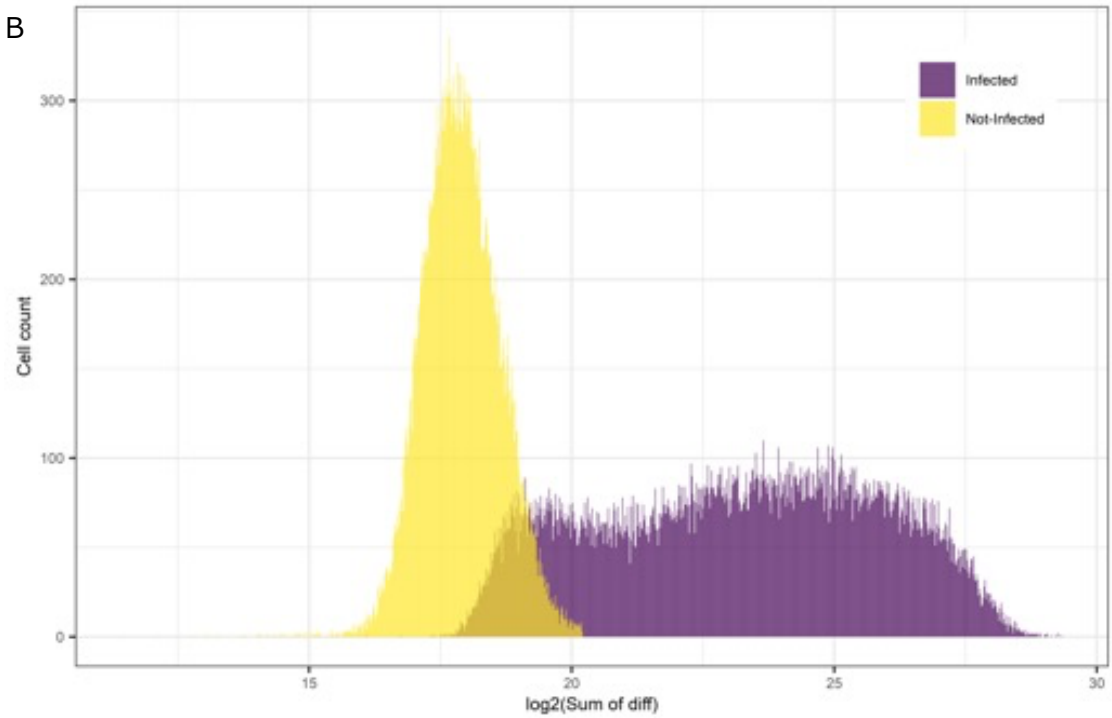

Figure 9.

(B)The two criteria were combined to identify infected cells. Results for the distinction between infected and uninfected cells. Infected and uninfected cells were color-coded.

C

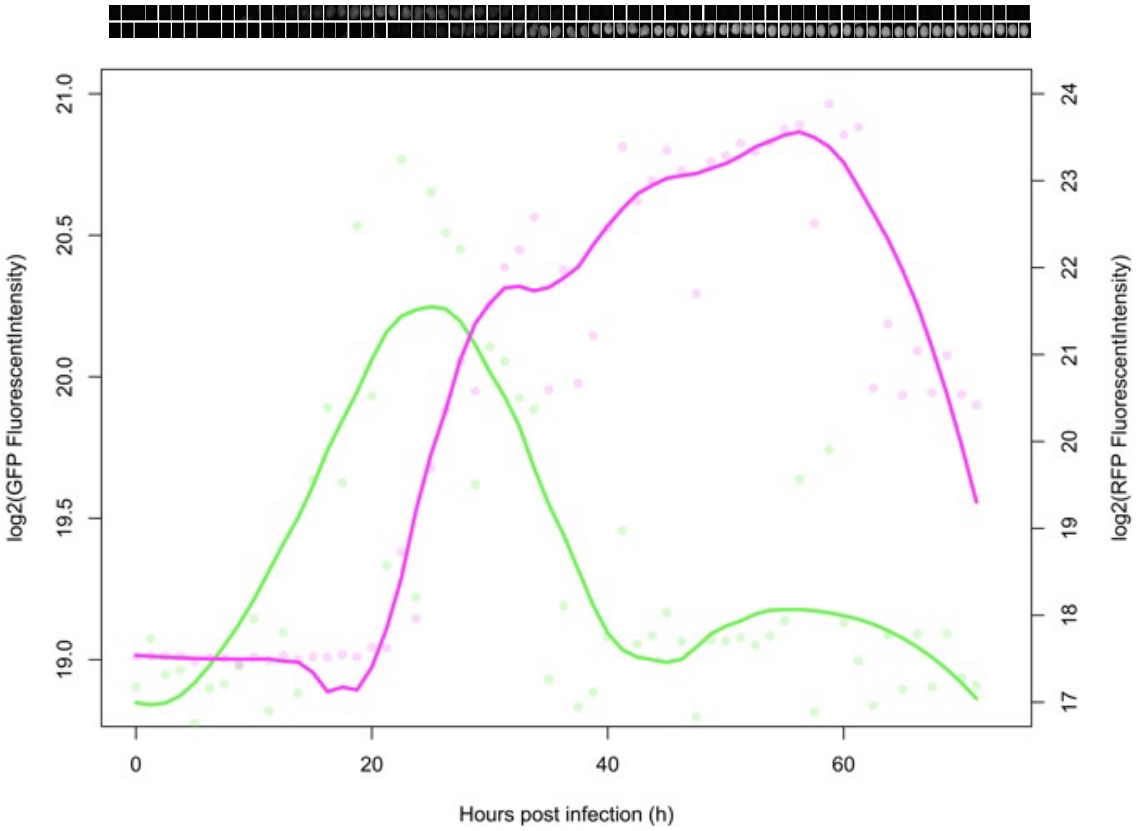

Figure 9.

(C)Fluorescence intensity at each time point was acquired for each infected cell and smoothed to obtain a fluorescence profile. Examples of the fluorescence profile of infected cells and the fluorescence profile smoothed by the loess process. The cell images at the corresponding timepoints are placed on top of the figure. Top image obtained under GFP imaging conditions; bottom image obtained under RFP imaging conditions.

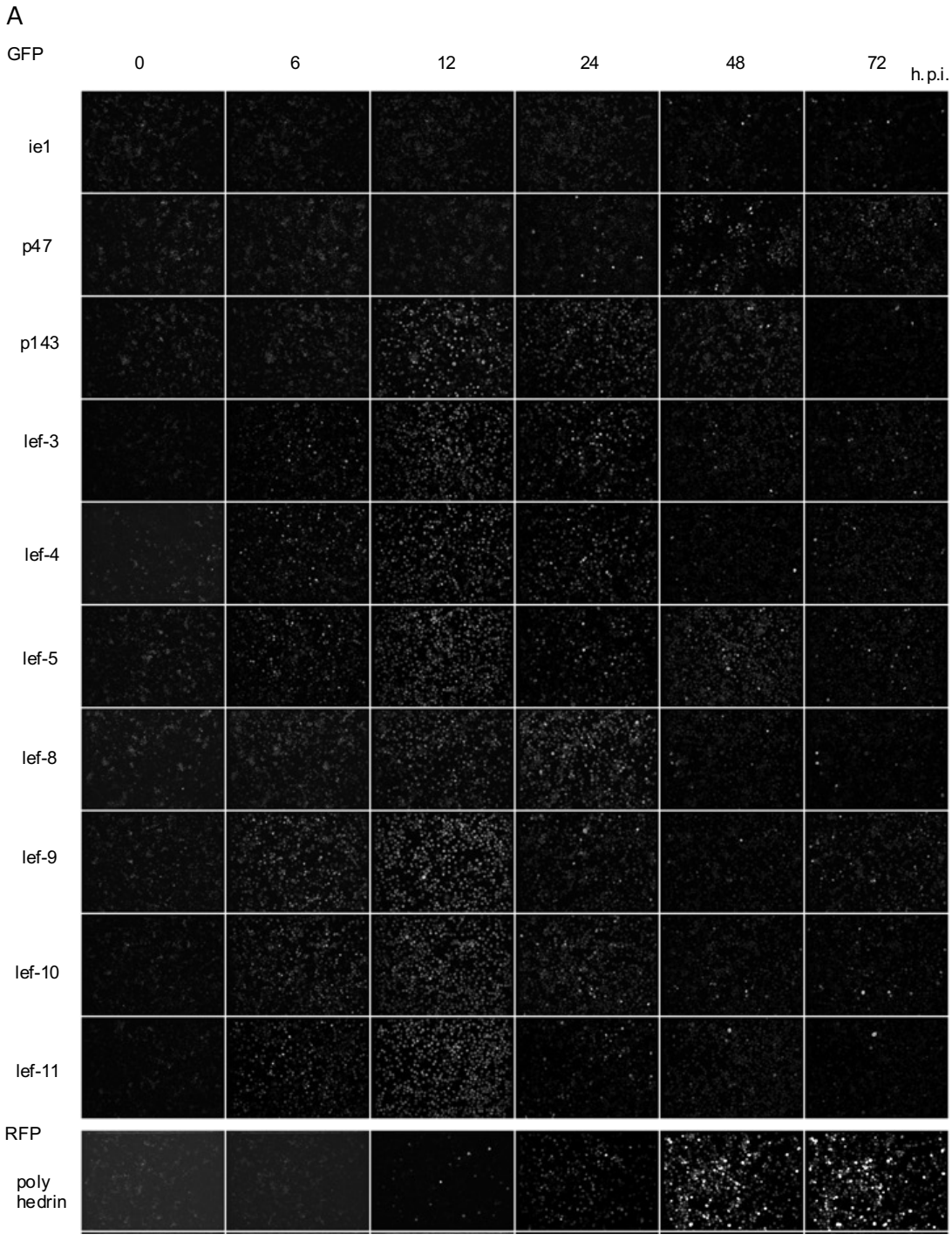

Figure 10. The character of each viral promoter and comparison with conventional methods

(A) Examples of observation results of infection experiments of 10 essential gene dual reporter bacmids produced. Examples of results from each essential transgenic dual reporter bacmids infection experiment were shown in the image. Each image of essential genes showed GFP fluorescence signal and only *polyhedrin* showed RFP fluorescence signal.

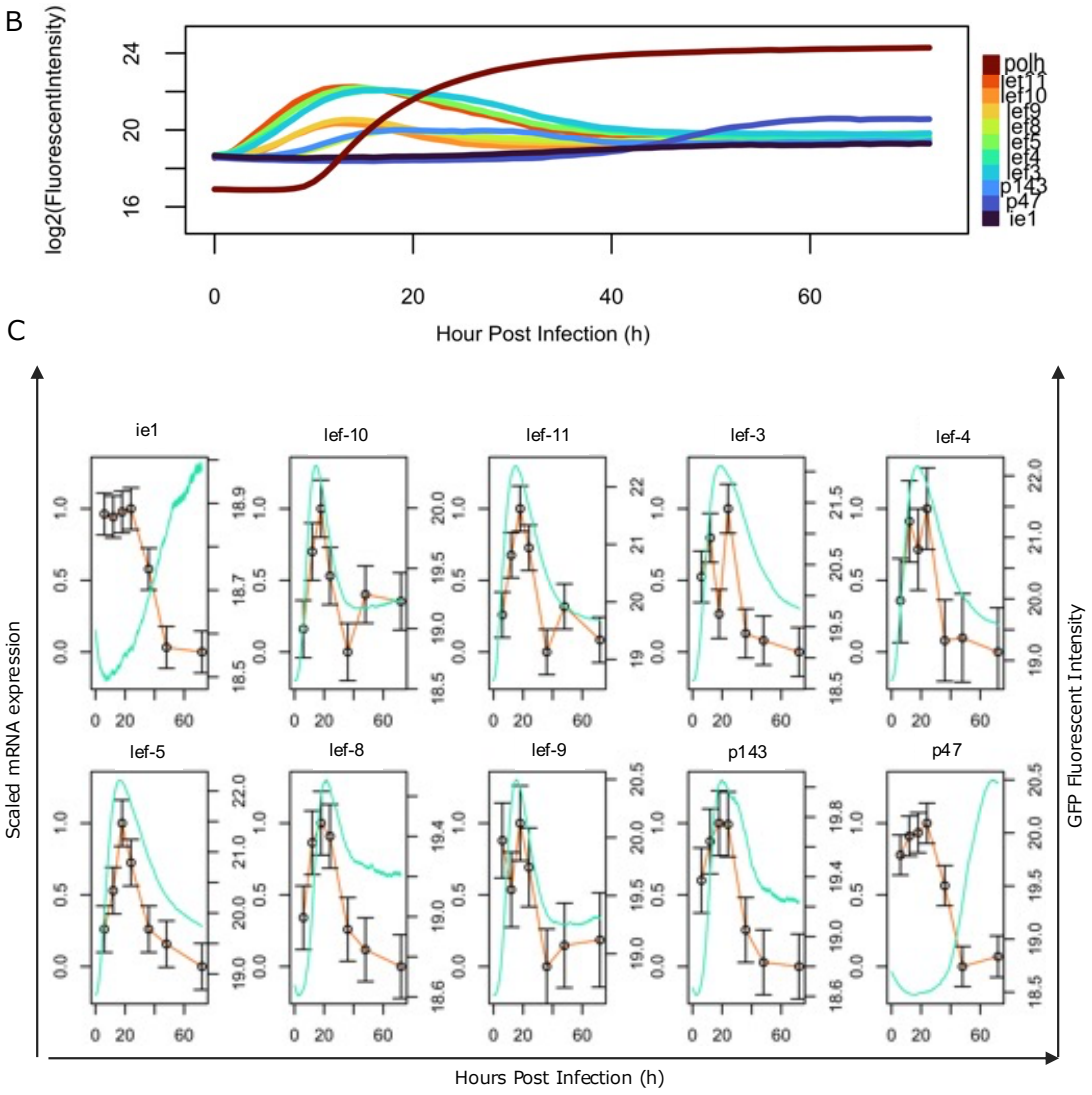

Figure 10.

(B) Activation profile of each essential gene promoter. The fluorescence profiles of each cell acquired were summarized for each promoter used and the mean values were plotted. Standard errors were shown in the ribbon plot. The results for each promoter are color-coded.

(C) Comparison of fluorescence profiles and RNA data. Expression levels of each gene were measured by RT-qPCR and plotted. The expression levels of each gene were expressed relative to that of silkworm Tubulin A, which was used as an intermediary control. Timepoints with no points indicate timepoints where the expression levels were below the detection limit. The fluorescence profile of each gene was also plotted at the same time. The left axis shows the relative expression levels of each gene in the RNA data, and the right axis shows the fluorescence profile values. Orange represents the RNA data and cyan represents the fluorescence profile.

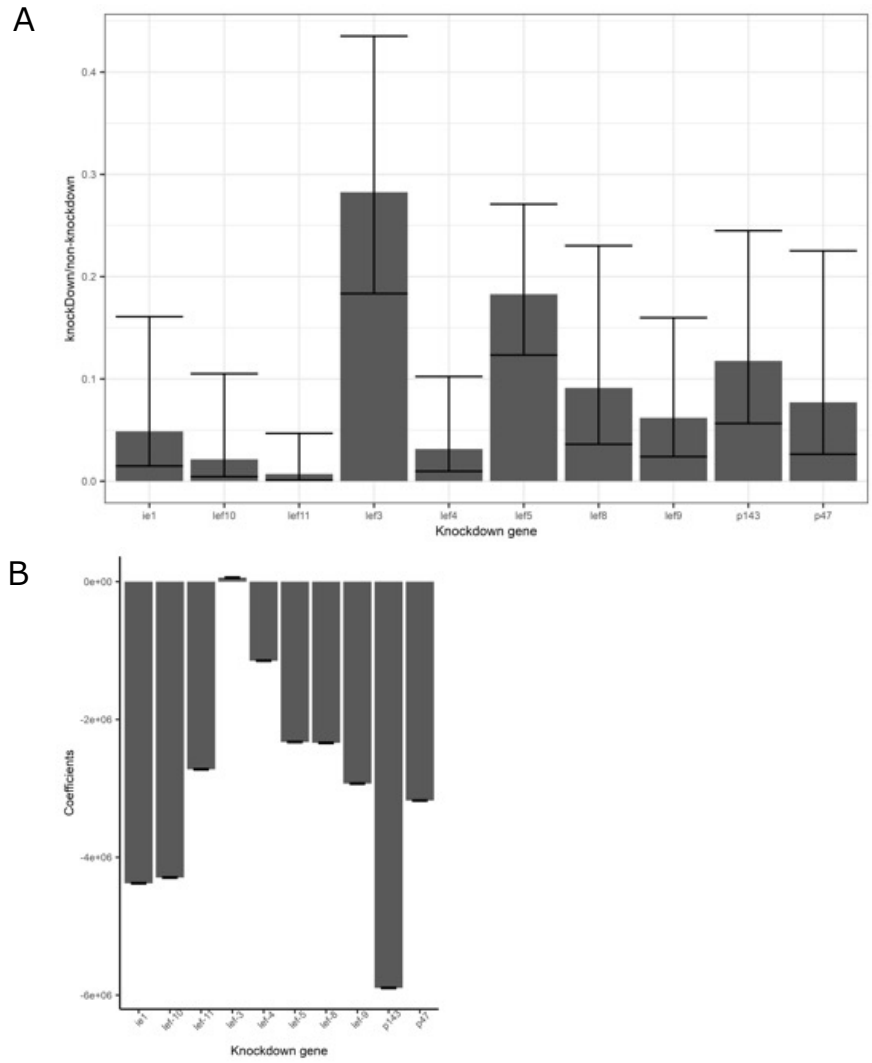

Figure 11. Evaluation of the efficiency of essential gene knockdown

(A)Verification of each essential gene knockdown efficiency. Each essential gene knockdown-treated BmN4-SID1 cell was infected with BmNPV 72 hours after treatment, and RNA was collected from the cells 24 hours after infection and RT-qPCR was conducted. ΔCt values were calculated using TublinA as reference under knockdown and non-knockdown conditions, respectively, and ΔΔΔCt values (ΔCt value for knockdown condition – ΔCt value for non-knockdown condition) were calculated and expressed as a real number. Error bars indicated standard errors.

(B)Effect of each essential gene knockdown on *polyhedrin* promoter activity. The effect of essential gene knockdown on *polyhedrin* promoter activity was elucidated as coefficients that were calculated by fitting a linear model(*RFP Fluorescent Intensity= Knockdown + Timepoint* + ε) to RFP fluorescent Intensity. Error bars indicated standard errors of the coefficients.

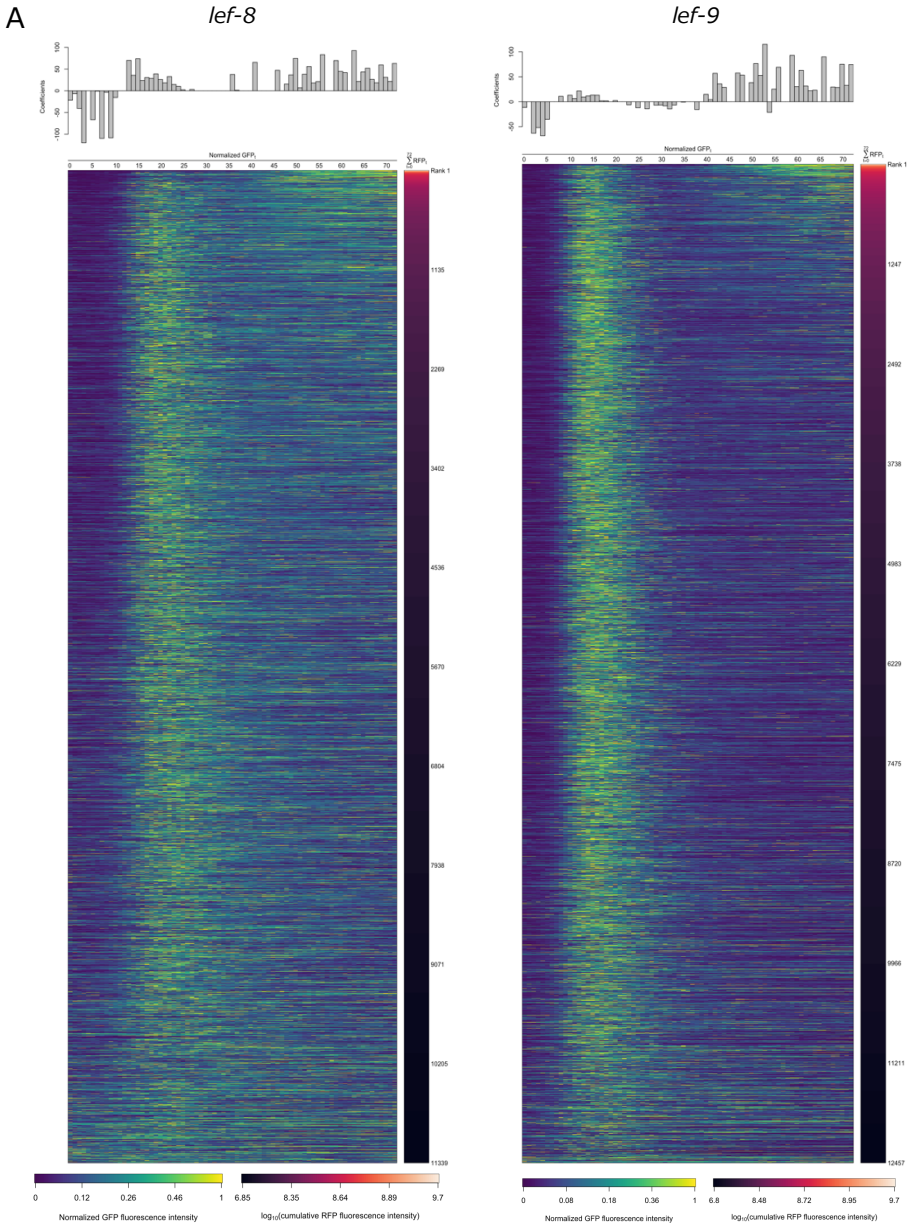

Figure 12. Identifying essential genes with promoter activation profiles related to *polyhedr in* promoter activity.

(A) GFP fluorescence profiles were transformed by variance stabilizing normalization and shown in a heat map. The horizontal axis represents hours post infection, and the vertical axis represents the results for each cell. The vertical axis was sorted from highest to lowest total fluorescence intensity of RFP. The total fluorescence profile of RFP corresponding to the GFP fluorescence profile of each infected cell was shown. The bars on the heat map showed the weighting of each time point in the fluorescence profiles obtained by LASSO regression fitted into the profile data. The vertical axis is coefficients, with a positive value indicating a positive correlation to RFP expression intensity and a negative value indicating a negative correlation. The results of lef-8,lef-9,lef-10 and ie1 were shown as exmples.

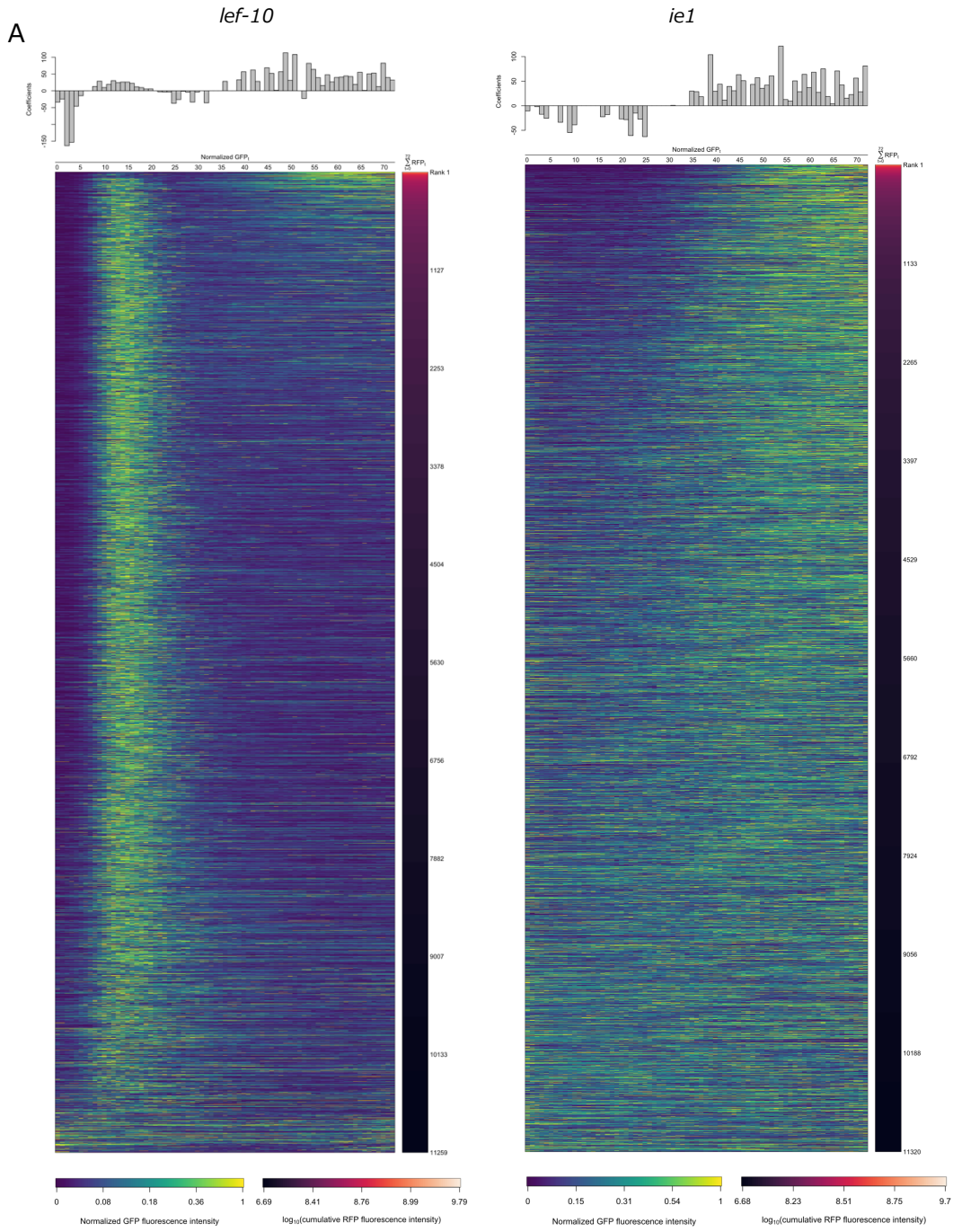
A
lef-10
ie1
Coefficients
Normalized GFP$_i$
RFP$_i$
Rank 1
Normalized GFP fluorescence intensity
log$_{10}$(cumulative RFP fluorescence intensity)

Figure 12. (A)

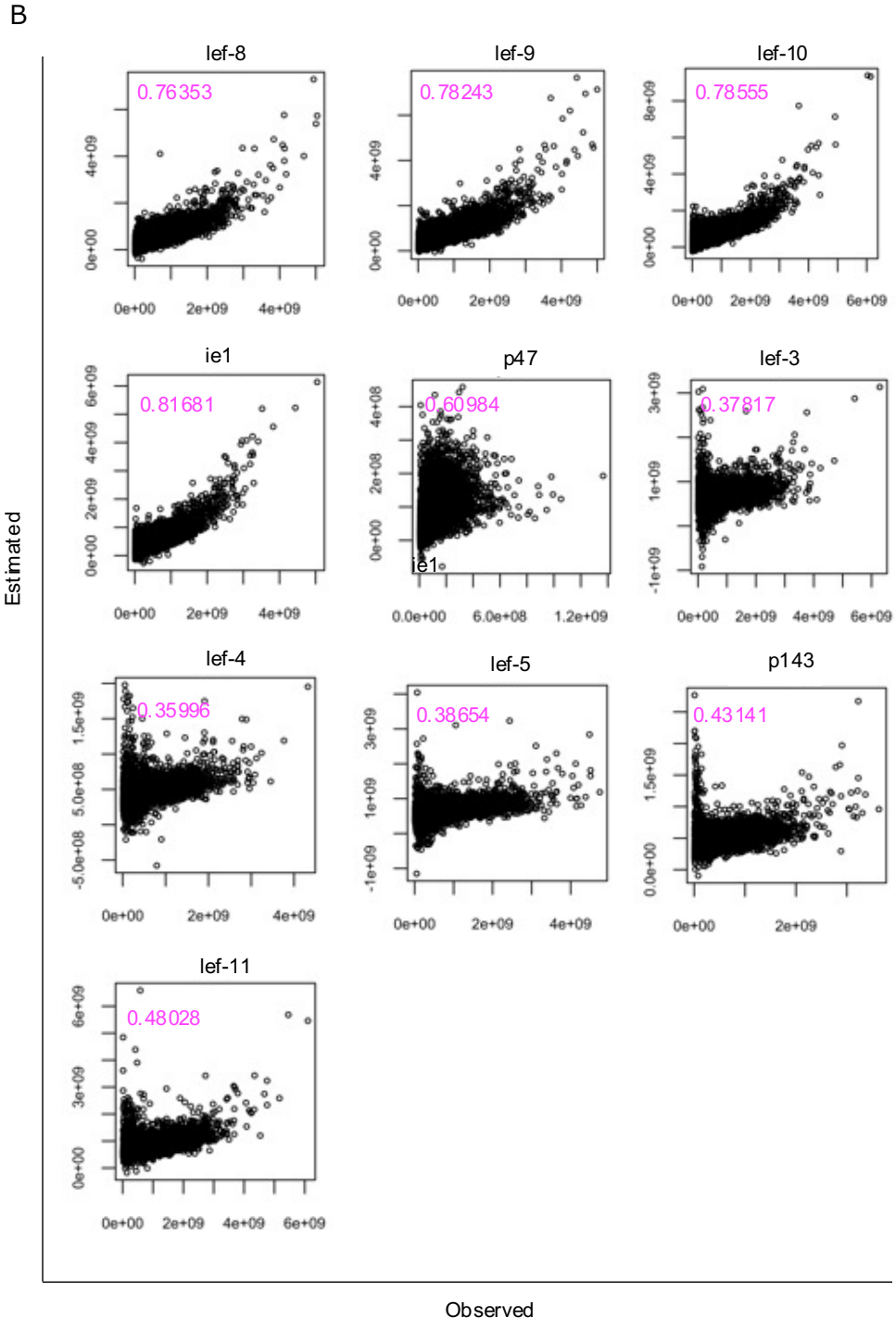

Figure 12.

(B) Comparison of predicted values based on LASSO regression and observed values. Scatterplot showing the estimated total *polyhedrin* promoter activation intensity obtained by substituting the GFP fluorescence profile into the LASSO regression of each essential gene promoter and the observed value in the same cells. The coefficients of two values were shown in the upper left corner.

A

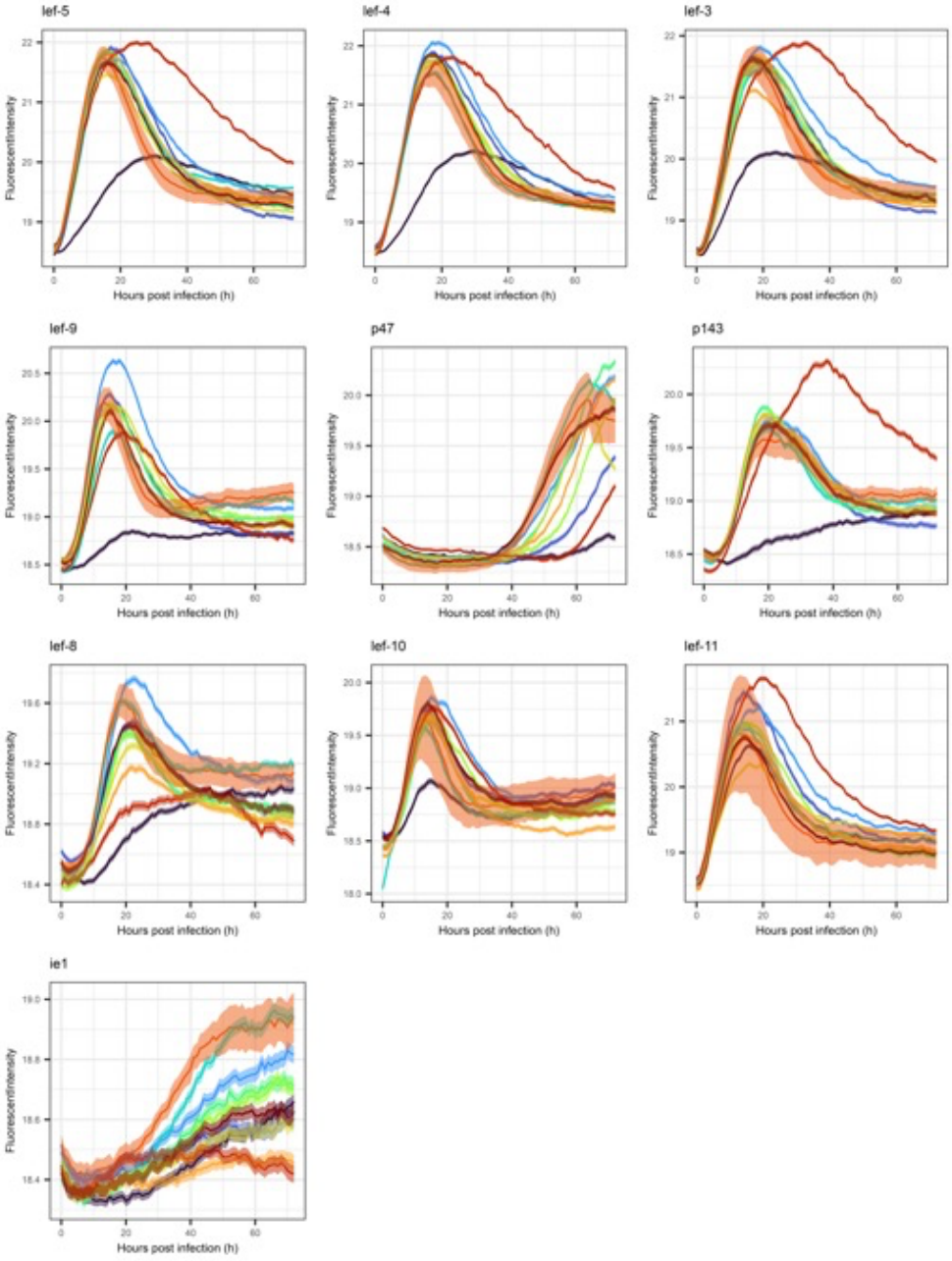

Figure 13. Effect of essential gene knockdown on the activation intensity of each essential gene

(A) Each essential gene promoter activity under the each essential gene knockdown condition. The activation profile of each essential gene promoter under each essential gene knockdown condition was estimated by adding the intercept calculated under the non-knockdown condition to the coefficients calculated under each essential gene knockdown condition. Regulatory relationships among essential genes in the immediate early stages of infection. The figure on the left is a network diagram showing the first control relationships seen after infection. Each node showed each essential gene and each edge showed significant regulation. Orange indicates when knockdown increased promoter activity and blue indicates when knockdown decreased promoter activity.

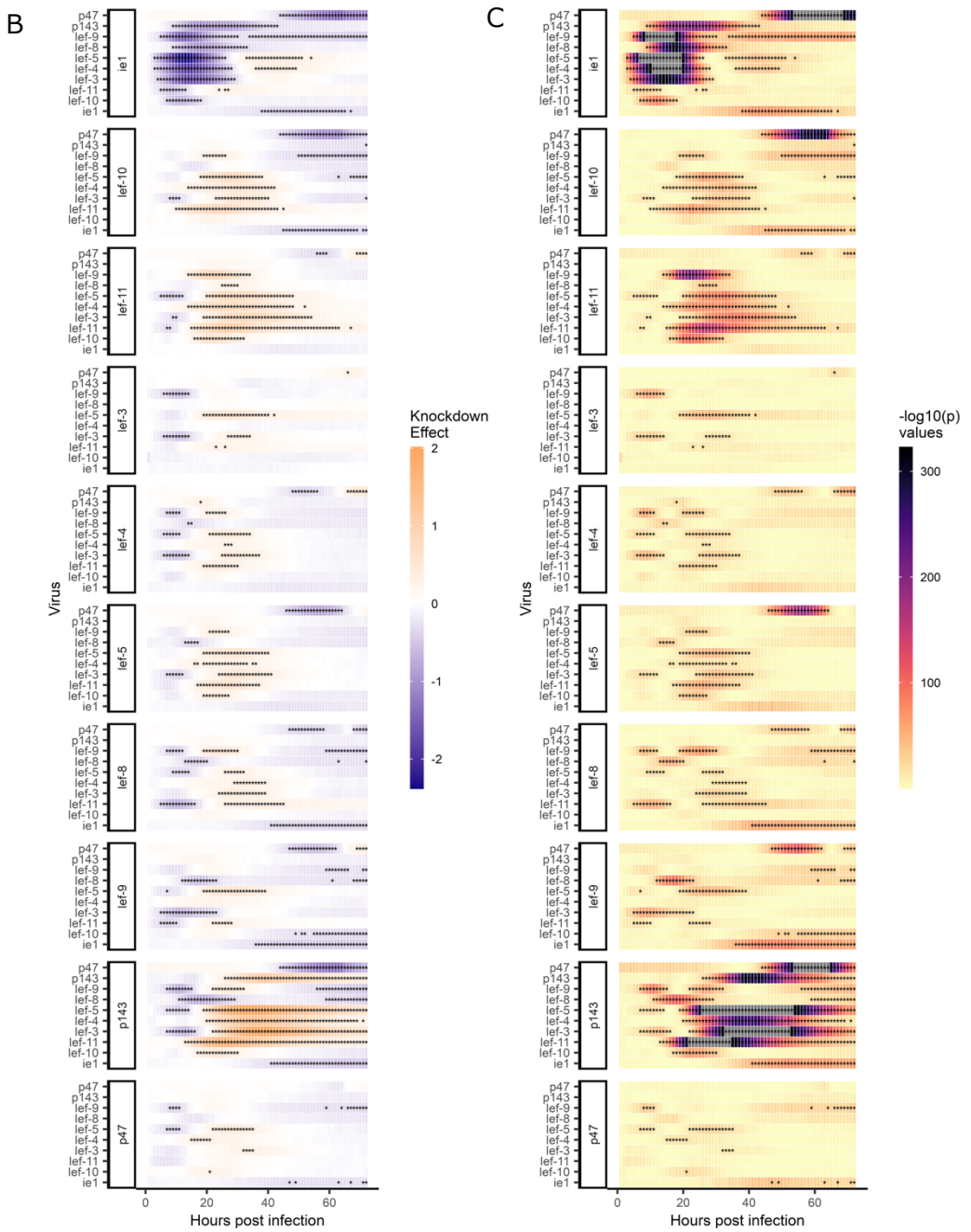

Figure 13.

(B)Coefficients of each essential gene knockdown condition on each essential gene promoter activity at each timepoint of infection. The fluorescence profiles under knockdown and non-knockdown conditions were compared, and the calculated Coefficient values are shown for each gene knockdown. The horizonal axis represented hours post infection. The vertical axis represented each essential gene promoter condition. Time points with significant differences were marked with an asterisk.

(C)p-values of each essential gene knockdown condition on each essential gene at each timepoint of infection

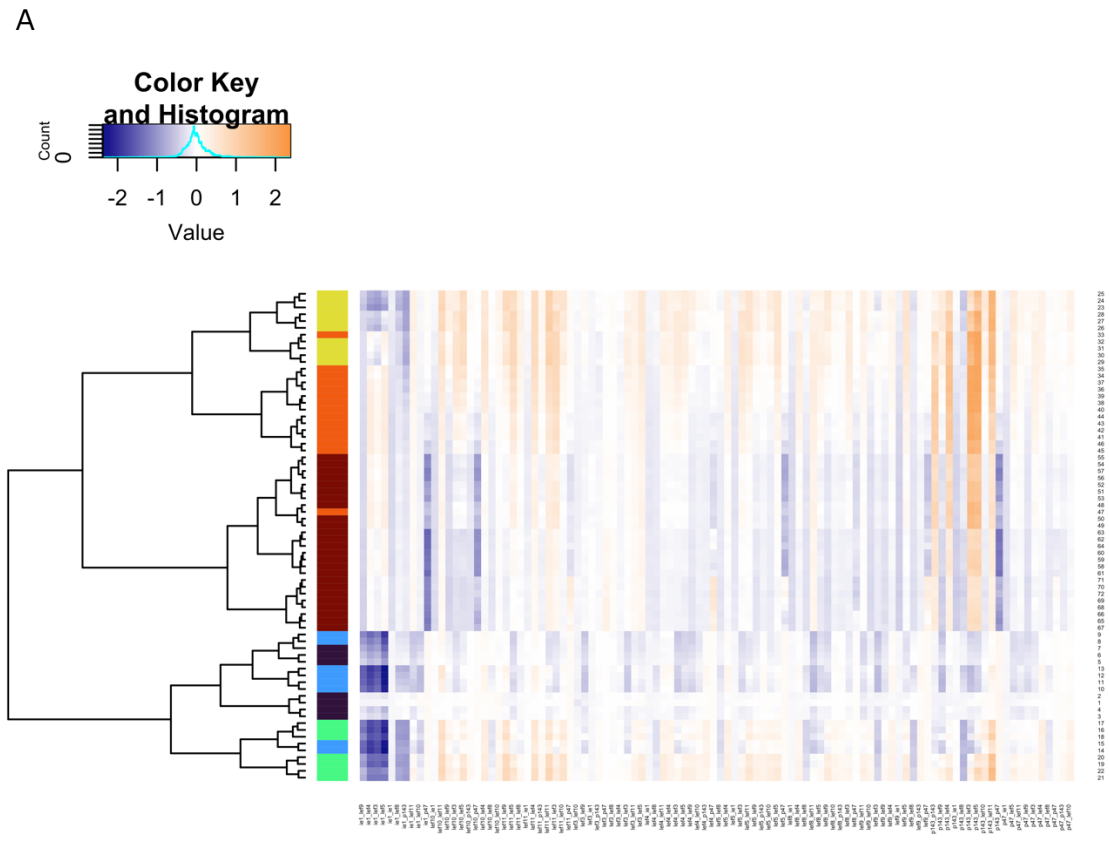

Figure 14. Estimation of essential gene regulatory network model in primary infected cells

(A)Estimation of infection stages with characteristic regulation by clustering of temporal information. The data shown in Fig. 13A were clustered by the k-means method in the time direction. The clustering resulted in sequential timepoints classified into identical clusters, totaling six clusters. Based on the *polyhedrin* promoter activity profiles in previous studies and in this study, three clusters up to 22 h.p.i. were determined to be the infection stages that showed primary infection. The heat map shows timepoints on the vertical axis and the knockdown gene/promoter combination on the horizontal axis. Clustering results were indicated by color next to the heat map.

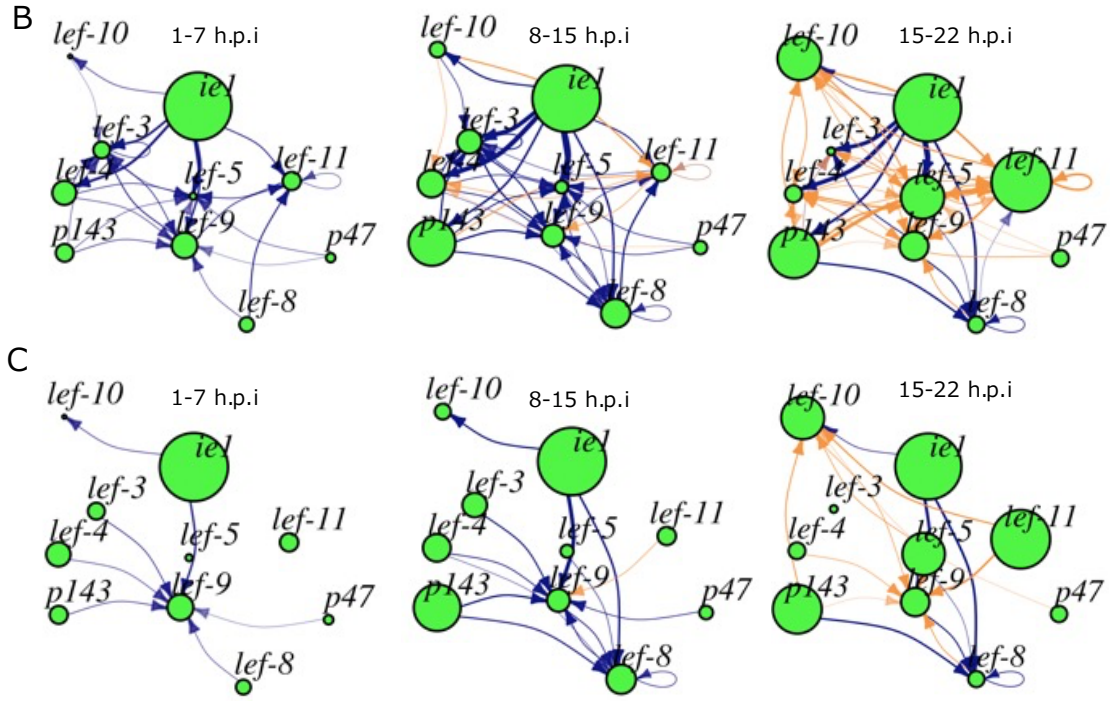

Figure 14.

(B) Estimation of essential gene regulatory network models at each infection stage
Edges were tied between genes that were significantly regulated by knockdown at each infection stage. The color of the edge indicated the positive or negative effect of knockdown, and the size of the edge indicated the magnitude of the value. The size of the node indicated the hubness of each gene in the network.

(C)Estimation of essential gene regulatory subnetwork models at each infection stage. Subnetwork model extracting only regulatory relationships to *lef-8*, *lef-9*, and *lef-10* from the control network model in B

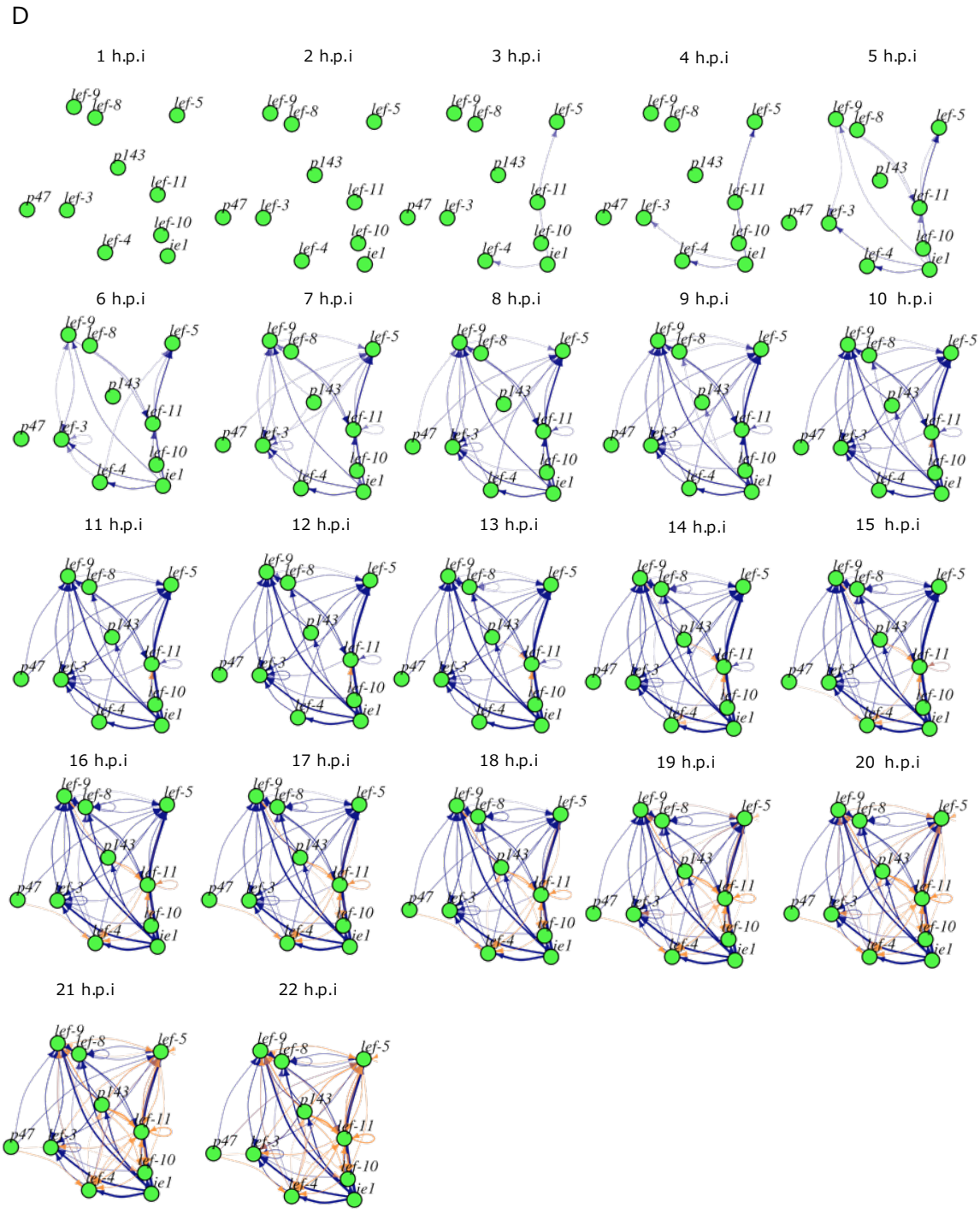

Figure 14.

(D) Estimation of essential gene regulatory network models at each time point of infection

E

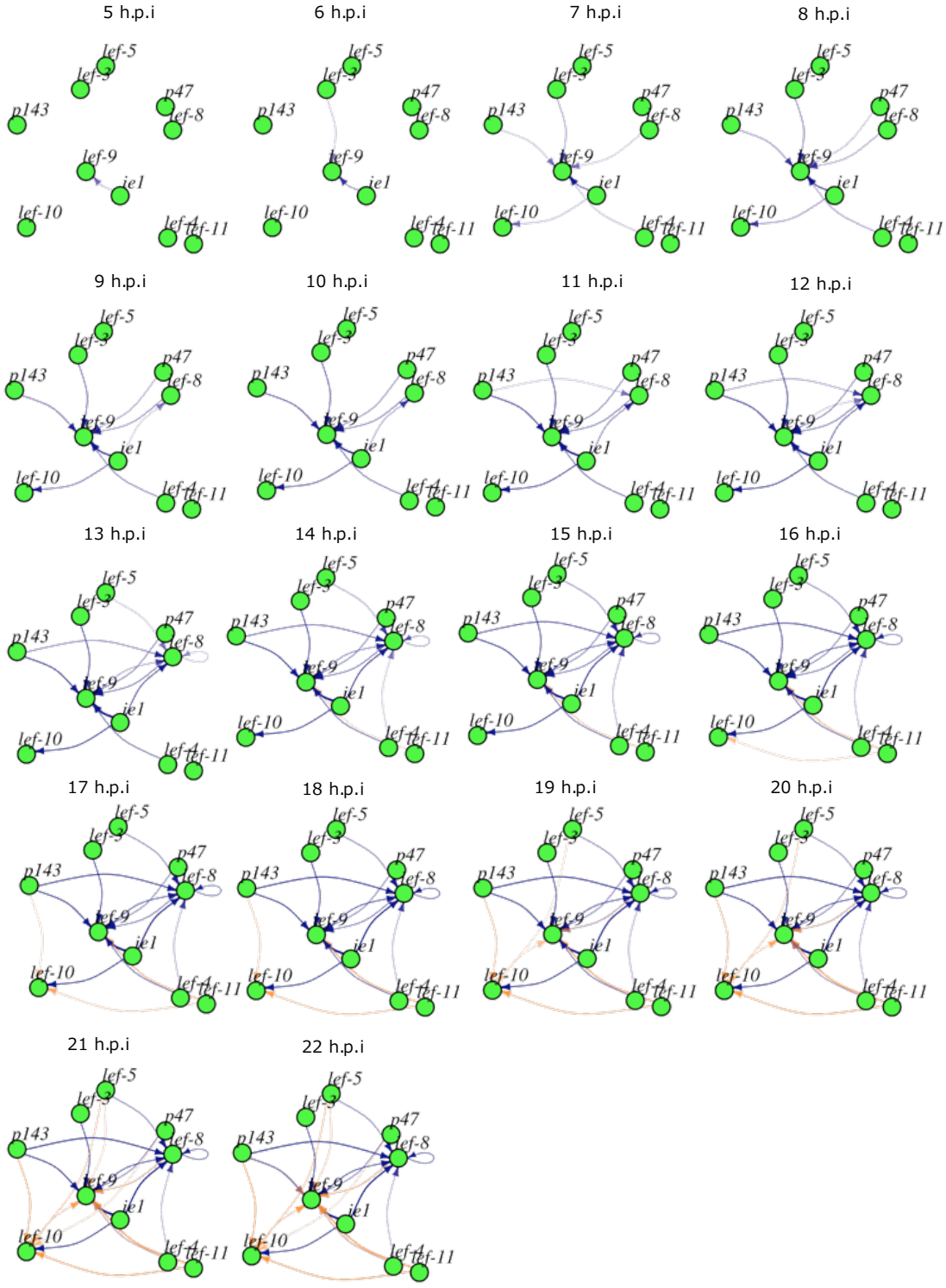

Figure 14.

(E) Estimation of essential gene regulatory subnetwork models at each time point of infection

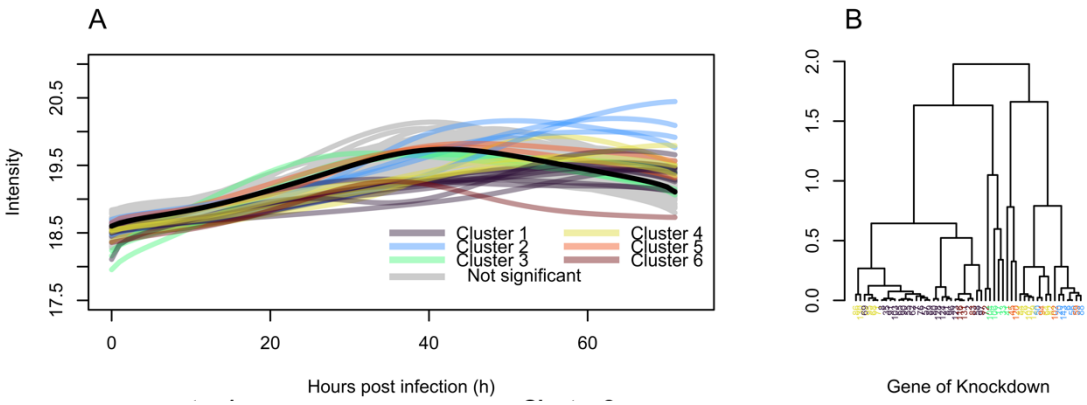

Figure 15. Effects of knockdown of each of the 143 genes on *p143* promoter activity

(A) *p143* fluorescence profile for each gene knockdown.

Based on linear modeling, the *p143* fluorescence profile compared to the non-knockdown condition is shown based on the coefficient calculated for each gene knockdown condition and each time point. *p143* fluorescence profiles are color-coded for each gene based on the clustering results in B. Gray genes represent genes that did not have a statistically significant change in *p143* at all time points. Fluorescence profiles in the non-knockdown condition are shown in black.

(B) Dendrogram based on Cosine similarity between Coefficient at each gene knockdown

The Coefficient values at each time point for each gene were compiled into a profile indicating the effect of gene knockdown, and a tree diagram was created based on the Cosine similarity between these profiles. From these results, the number of clusters was determined to be six, and clustering was performed again using non-hierarchical clustering, and the results are indicated by color. The numbers correspond to the BmNPV gene numbers at NCBI.

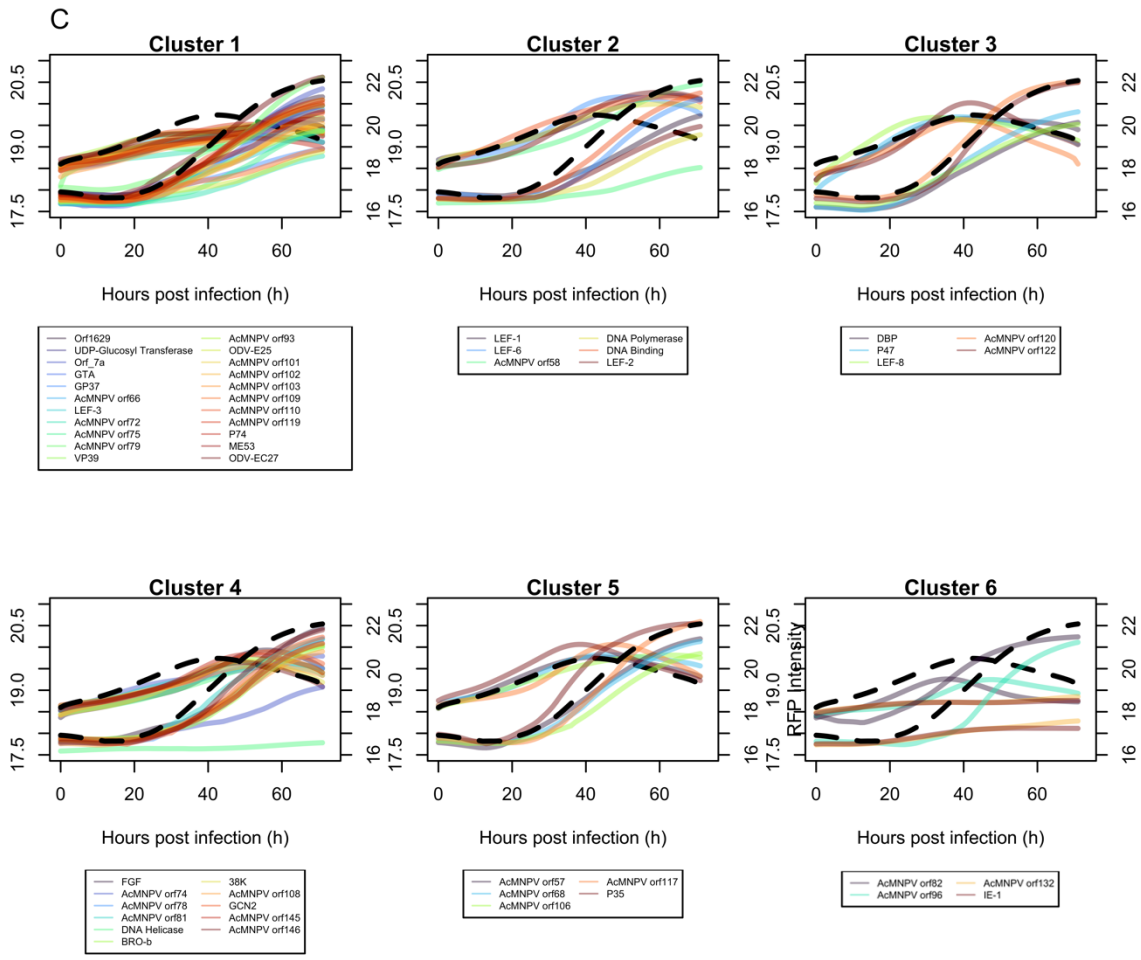

Figure 15.

(C) *p143* promoter activity profile and *polyhedrin* promoter activity profile by cluster

The effects on *p143* promoter activity and *polyhedrin* promoter activity during knockdown of each gene classified in each cluster are shown. The left axis label shows the fluorescence profile of *p143* and the right axis label shows the fluorescence profile of *polyhedrin*. Each gene is color-coded for each cluster, and the gene name corresponding to each color is shown in the legend.

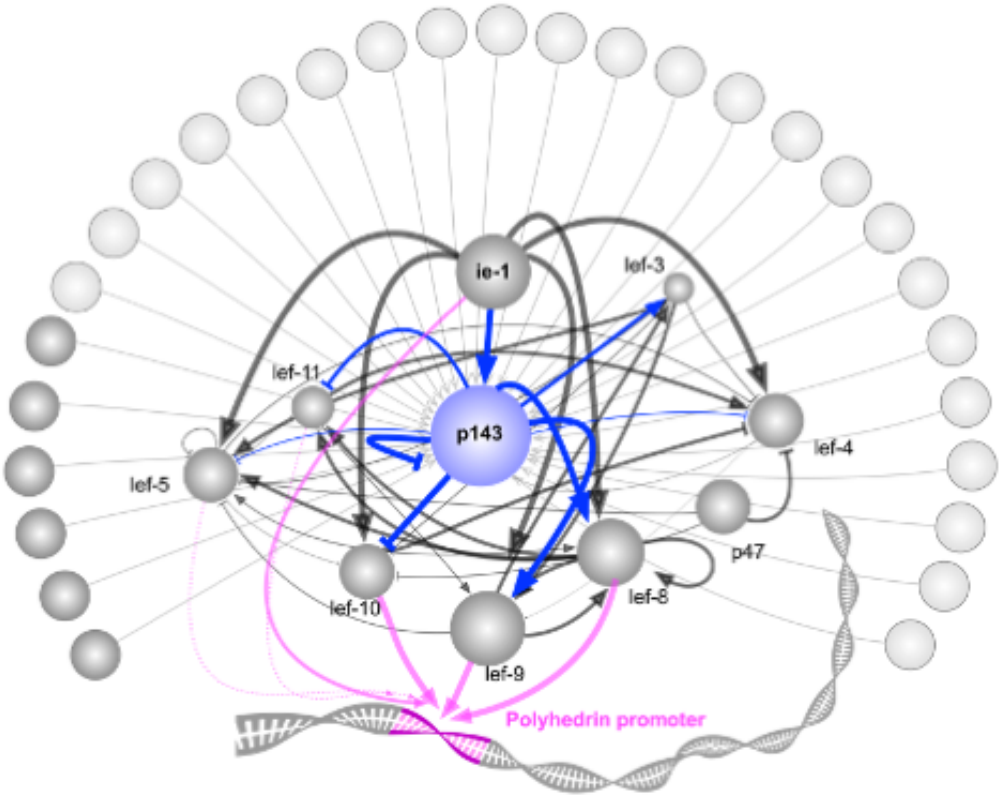

Figure 16. A polyhedrin promoter activation model centered on essential genes that synchronizes with the progression of infection stages inferred from this study

References

Banaszynski, L.A., Chen, L., Maynard-Smith, L.A., Ooi, A.G.L., and Wandless, T.J. (2006)
A Rapid, Reversible, and Tunable Method to Regulate Protein Function in Living Cells Using Synthetic Small Molecules.
Cell, **126**, 995–1004.

Bradford, M.M. (1976)
A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding.
Anal. Biochem., **72**, 248-254

Buratowski, S. (1994)
The basics of basal transcription by RNA polymerase II.
Cell, **77,** 1–3.

Bylino, O.V., Ibragimov, A.N., and Shidlovskii, Y.V. (2020)
Evolution of Regulated Transcription.
Cells, **9**, 1675.

Carson, D.D., Summers, M.D., and Guarino, L.A. (1991)
Molecular analysis of a baculovirus regulatory gene.
Virology, **182**, 279–286.

Chen, Y.-R., Zhong, S., Fei, Z., Hashimoto, Y., Xiang, J.Z., Zhang, S., and Blissard, G.W. (2013)
The Transcriptome of the Baculovirus Autographa californica Multiple Nucleopolyhedrovirus in Trichoplusia ni Cells.
J. Virol., **87**, 6391–6405.

Chung, H.K., Jacobs, C.L., Huo, Y., Yang, J., Krumm, S.A., Plemper, R.K., Tsien, R.Y., and Lin, M.Z. (2015)
Tunable and reversible drug control of protein production via a self-excising degron.
Nat. Chem. Biol., **11,** 713–720.

Cleveland, W.S. (1979)
Robust Locally Weighted Regression and Smoothing Scatterplots.
J. Am. Stat. Assoc., **74(368)**, 829-836

Cleveland, W.S., and Devlin, S.J. (1988)
Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting.
J. Am. Stat. Assoc., **83(403)**, 596-610

Corish, P., and Tyler-Smith, C. (1999)
Attenuation of green fluorescent protein half-life in mammalian cells.
Protein Eng. Des. Sel., **12**, 1035–1040.

De, M., A., Sebé-Pedrós, A., Šestak, M.S., Matejčić, M., Torruella, G., Domazet-Lošo, T., and Ruiz-Trillo, I. (2013)
Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages.
Proc. Natl. Acad. Sci. U.S.A., **110.**

Dong, Z.-Q., Hu, N., Dong, F.-F., Chen, T.-T., Jiang, Y.-M., Chen, P., Lu, C., and Pan, M.-H. (2017)
Baculovirus LEF-11 Hijack Host ATPase ATAD3A to Promote Virus Multiplication in Bombyx mori cells.
Sci. Rep. **7**, 46187.

Dudley, A.M., Aach, J., Steffen, M.A., and Church, G.M. (2002)
Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range.
Proc. Natl. Acad. Sci. U.S.A **99**, 7554–7559.

Dynan, S. W. and Tijia, T., (1983)
Isolation of Transcription Factors That Discriminate between Different Promoters Recognized by RNA Polymerase II.
Cell, **32**, 669-680

Ershov, D., Phan, M.-S., Pylvänäinen, J.W., Rigaud, S.U., Le Blanc, L., Charles-Orszag, A., Conway, J.R.W., Laine, R.F., Roy, N.H., Bonazzi, D., et al. (2022)
TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines.
Nat. Methods, **19**, 829–832.

Fan, X., Thirunavukkarasu, K., and Weaver, R.F. (1996)
Temperature-Sensitive Mutations in the Protein Kinase-1 (pk-1) Gene of theAutographa californicaNuclear Polyhedrosis Virus That Block Very Late Gene Expression.
Virology, **224,** 1–9.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998).
Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*.
Nature, **391**, 806-811

Friesen, P. D. (1997)
Regulation of Baculovirus Early Gene Expression.
The Baculoviruses, 141-170

Gomi, S., Majima, K., and Maeda, S. (1999)
Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
J. Gen. Virol., **80**, 1323–1337.

Gossen, M., and Bujard, H. (1992)
Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.
Proc. Natl. Acad. Sci. U.S.A., **89,** 5547–5551.

Guarino, L.A., and Dong, W. (1991)
Expression of an enhancer-binding protein in insect cells transfected with the Autographa californica nuclear polyhedrosis virus IE1 gene.
J. Virol., **65**, 3676–3680.

Guarino, L.A., and Summers, M.D. (1986)
Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene.
J. Virol., **57**, 563–571.

Guarino, L.A., Gonzalez, M.A., and Summers, M.D. (1986)
Complete Sequence and Enhancer Function of the Homologous DNA Regions of *Autographa californica* Nuclear Polyhedrosis Virus.
J. Virol., **60,** 224–229.

Guarino, L.A., Xu, B., Jin, J., and Dong, W. (1998)
A Virus-Encoded RNA Polymerase Purified from Baculovirus-Infected Cells.
J. Virol. ***72,*** 7985–7991.

Hang, X., Dong, W., and Guarino, L.A. (1995)
The lef-3 gene of Autographa californica nuclear polyhedrosis virus encodes a single-

stranded DNA-binding protein.
J. Virol., **69,** 3924–3928.

Harwood, S.H., Li, L., Ho, P.S., Preston, Å.K., and Rohrmann, G.F. (1998)
AcMNPV Late Expression Factor-5 Interacts with Itself and Contains a Zinc Ribbon Domain That Is Required for Maximal Late Transcription Activity and Is Homologous to Elongation Factor TFIIS.
Virology, **250,** 118–134.

He L., Binari, R., Huang, J., Falo-Sanjuan, J., Perrimon, N. (2019)
In vivo study of gene expression with an enhanced dual-color fluorescent transcriptional timer.
elife, **8**, e46181.

Ho, W.-C., and Zhang, J. (2018)
Evolutionary adaptations to new environments generally reverse plastic phenotypic changes.
Nat. Commun. **9,** 350.

Jacob, F., and Monod, J. (1961)
Genetic regulatory mechanisms in the synthesis of proteins.
J. Mol. Biol., **3**, 318-356

Jacobs, C.L., Badiee, R.K., and Lin, M.Z. (2018)
StaPLs: versatile genetically encoded modules for engineering drug-inducible proteins.
Nat. Methods, **15,** 523–526.

Jin, J., Dong, W., and Guarino, L.A. (1998).
The LEF-4 Subunit of Baculovirus RNA Polymerase Has RNA 5′ -Triphosphatase and

ATPase Activities.
J. Virol., **72**, 10011–10019.

Kawasaki, Y., Matsumoto, S., and Nagamine, T. (2004)
Analysis of baculovirus IE1 in living cells: dynamics and spatial relationships to viral structural proteins.
J. Gen. Viol., 85, 3575–3583.

Knebel-Mörsdorf, D., Quadt, I., Li, Y., Montier, L., and Guarino, L.A. (2006)
Expression of Baculovirus Late and Very Late Genes Depends on LEF-4, a Component of the Viral RNA Polymerase Whose Guanyltransferase Function Is Essential.
J. Virol., **80**, 4168–4173.

Knight, C.A., Molinari, N. A., Petrov, D. A., (2005).
The Large Genome Constraint Hypothesis: Evolution, Ecology and Phenotype.
Ann. Bot.,.**95,** 177–190.

Kolykhalov, A.A., Agapov, E.V., and Rice, C.M. (1994)
Specificity of the hepatitis C virus NS3 serine protease: effects of substitutions at the 3/4A, 4A/4B, 4B/5A, and 5A/5B cleavage sites on polyprotein processing.
J. Virol., **68**, 7525–7533.

Landro, J.A., Raybuck, S.A., Luong, Y.P.C., O'Malley, E.T., Harbeson, S.L., Morgenstern, K.A., Rao, G., and Livingston, D.J. (1997)
Mechanistic Role of an NS4A Peptide Cofactor with the Truncated NS3 Protease of Hepatitis C Virus: Elucidation of the NS4A Stimulatory Effect via Kinetic Analysis and Inhibitor Mapping.
Biochemi., **36**, 9340–9348.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.-C., and Kain, S.R. (1998).
Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter.
J. Biol. Chem., **273,** 34970–34975.

Liu, Y., Yang, Q., and Zhao, F. (2021)
Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding.
Annu. Rev. Biochem. **90,** 375–401.

Lu, A., and Miller, L.K. (1995)
The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication.
J. Virol., **69,** 975–982.

Luckow, V.A., Lee, S.C., Barry, G.F., and Olins, P.O. (1993)
Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli.
J. Virol., **67,** 4566–4579.

Maeda, S., Kawai, T., Obinatat, M., and Fujiwara, H. (1985)
Production of human a-interferon in silkworm using a baculovirus vector.
Nature, **315,** 13

Masumoto, M., Ohde, T., Shiomi, K., Yaginuma, T., and Niimi, T. (2012)
A Baculovirus Immediate-Early Gene, ie1, Promoter Drives Efficient Expression of a Transgene in Both Drosophila melanogaster and Bombyx mori.
PLoS ONE, **7,** e49323.

McDougal, V.V., and Guarino, L.A. (2000)
The Autographa californica Nuclear Polyhedrosis Virus p143 Gene Encodes a DNA Helicase.
J. Virol., **74**. 11

McLachlin, J.R., and Miller, L.K. (1994)
Identification and characterization of vlf-1, a baculovirus gene involved in very late gene expression.
J. Virol .,**68,** 7746–7756.

Mikhailov, V.S., and Rohrmann, G.F. (2002)
Baculovirus Replication Factor LEF-1 Is a DNA Primase.
J. Virol., **76,** 2287–2297.

Mon, H., Kobayashi, I., Ohkubo, S., Tomita, S., Lee, J., Sezutsu, H., Tamura, T., and Kusakabe, T. (2012)
Effective RNA interference in cultured silkworm cells mediated by overexpression of Caenorhabditis elegans SID-1.
RNA Biol., **9,** 40–46.

Morris, T.D., and Miller, L.K. (1994)
Mutational analysis of a baculovirus major late promoter.
Gene, **140,** 147–153.

Motohashi, K. (2015).
A simple and efficient seamless DNA cloning method using SLiCE from Escherichia coli laboratory strains and its application to SLiP site-directed mutagenesis.
*BMC Biotechnol., **15,*** 47.

Nagamine, T., Kawasaki, Y., Abe, A., and Matsumoto, S. (2008)
Nuclear Marginalization of Host Cell Chromatin Associated with Expansion of Two Discrete Virus-Induced Subnuclear Compartments during Baculovirus Infection.
J. Virol., **82,** 6409–6418.

Nan, H., Chen, H., Tuite, M.F., and Xu, X. (2019)
A viral expression factor behaves as a prion.
*Nat. Commun.,***10,** 359.

Niopek, D., Benzinger, D., Roensch, J., Draebing, T., Wehler, P., Eils, R., and Di Ventura, B. (2014).
Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells.
Nat. Commun., **5,** 4404.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009)
An auxin-based degron system for the rapid depletion of proteins in nonplant cells.
Nat. Methods **6,** 917–922.

Okano, K., Mikhailov, V.S., and Maeda, S. (1999).
Colocalization of Baculovirus IE-1 and Two DNA-Binding Proteins, DBP and LEF-3, to Viral Replication Factories.
J. Virol., **73**, 110–119.

Okano, K., Mikhailov, V.S., and Maeda, S. (1999)
Colocalization of Baculovirus IE-1 and Two DNA-Binding Proteins, DBP and LEF-3, to Viral Replication Factories.
J. Virol., **73,** 110–119.

Ono, C., Kamagata, T., Taka, H., Sahara, K., Asano, S., and Bando, H. (2012) Phenotypic grouping of 141 BmNPVs lacking viral gene sequences.
Virus Res., **165,** 197–206.

Ono, C., Nakatsukasa, T., Nishijima, Y., Asano, S., Sahara, K., and Bando, H. (2007) Construction of the BmNPV T3 Bacmid System and Its Application to the Functional Analysis of BmNPV he65.
J. Insect Biotechnol. Sericology, **76**, 161-167.

Passarelli, A.L., Todd, J.W., and Miller, L.K. (1994)
A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases.
J. Virol., **68**, 4673–4678.

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006)
Engineering and characterization of a superfolder green fluorescent protein.
Nat. Biotechnol., **24**, 79–88.

Pelosi, L.A., Voss, S., Liu, M., Gao, M., and Lemm, J.A. (2012)
Effect on Hepatitis C Virus Replication of Combinations of Direct-Acting Antivirals, Including NS5A Inhibitor Daclatasvir.
Antimicrob. Agents. Chemother., **56**, 5230–5239.

Pinheiro, J. C. and Bates, D. M. (2000)
In *Mixed-Effects Models in S and S-PLUS*.
Springer, New York, NY

Possee, R.D., and Howard, S.C. (1987)
Analysis of the polyhedrin gene promoter of the *Autographa californica* nuclear polyhedrosis virus.
Nucl.Acids. Res., **15,** 10233–10248.

R Core Team (2003).
R: A language and environment for statistical computing.
R Foundation for Statistical Computing.

Roeder, R.G. (1996).
The role of general initiation factors in transcription by RNA polymerase II.
Trends Biochem. Sci., **21(9)**: 327-35.

Rogers, S., Wells, R., and Rechsteiner, M. (1986)
Amino Acid Sequences Common to Rapidly Degraded Proteins: The PEST Hypothesis.
Science, **234,** 364–368.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012)
Fiji: an open-source platform for biological-image analysis.
Nat. Methods, **9**, 676–682.

Smale, S.T., and Baltimore, D. (1989)
The “initiator” as a transcription control element.
Cell, **57,** 103–113.

Smith, G.E., Vlak, J.M., and Summers, M.D. (1983)
Physical Analysis of *Autographa californica* Nuclear Polyhedrosis Virus Transcripts for

Polyhedrin and 10,000-Molecular-Weight Protein.
J. Virol., **45,** 215–225.

Spindler, K.R., Rosser, D.S., and Berk, A.J. (1984)
Analysis of adenovirus transforming proteins from early regions 1A and 1B with antisera to inducible fusion antigens produced in Escherichia coli.
J. Virol, **49**, 132–141.

Sriram, S., and Gopinathan, K.P. (1998)
The Potential Role of a Late Gene Expression Factor, Lef2, fromBombyx moriNuclear Polyhedrosis Virus in Very Late Gene Transcription and DNA Replication.
Virology, **251,** 108–122.

Tibshirani, R. (1996)
Regression Shrinkage and Selection Via the Lasso.
J. R. Statist. Soc. B, **58,** 267–288.

Tinevez, J.-Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017)
TrackMate: An open and extensible platform for single-particle tracking.
Methods, **115,** 80–90.

Todd, J.W., Passarelli, A.L., Lu, A., and Miller, L.K. (1996)
Factors regulating baculovirus late and very late gene expression in transient-expression assays.
J. Virol., **70,** 2307–2317.

Trost, M., Blattner, A.C., and Lehner, C.F. (2016)
Regulated protein depletion by the auxin-inducible degradation system in Drosophila

melanogaster.
*Fly ,***10,** 35–46.

Vanarsdall, A.L., Okano, K., and Rohrmann, G.F. (2005)
Characterization of the replication of a baculovirus mutant lacking the DNA polymerase gene.
Virology, **331,** 175–180.

Wan, M., Zhu, Y., and Zou, J. (2019)
Novel near-infrared fluorescent probe for live cell imaging.
Exp. Ther. Med., **19(2),** 1213-1218

Wang, N., Chen, W., Linsel-Nitschke, P., Martinez, L.O., Agerholm-Larsen, B., Silver, D.L., and Tall, A.R. (2003)
A PEST sequence in ABCA1 regulates degradation by calpain protease and stabilization of ABCA1 by apoA-I.
J. Clin. Invest., **111,** 99–107.

Weiss, W.A., Taylor, S.S., and Shokat, K.M. (2007)
Recognizing and exploiting differences between RNAi and small-molecule inhibitors.
Nat. Chem. Biol., **3,** 739–744.

Winter, G.E., Buckley, D.L., Paulk, J., Roberts, J.M., Souza, A., Dhe-Paganon, S., and Bradner, J.E. (2015)
Phthalimide conjugation as a strategy for in vivo target protein degradation
Science, **348,** 1376–1381.

Wray, G.A. (2007)
The evolutionary significance of cis-regulatory mutations.
Nat. Rev. Genet., **8**, 206–216.

Xu, X., Zhou, X., Nan, H., Zhao, Y., Bai, Y., Ou, Y., and Chen, H. (2016)
Aggregation of AcMNPV LEF-10 and Its Impact on Viral Late Gene Expression
PLOS ONE, **11**, e0154835.

Yesbolatova, A., Tominari, Y., and Kanemaki, M.T. (2019)
Ligand-induced genetic degradation as a tool for target validation.
Drug Discov. Today, ***31,*** 91–98.

Appendix 1 Methods of preparation of reagents used in this study.

- Table 1 TC-100 medium composition

TC-100	20.57 g
$NaHCO_3$	0.35 g
Adjust to pH 6.2 with NaOH	
Fill up to 1 L with dH2O	
Filtrated with 0.2 μm nitrocellulose membrane	

- Table 2 LB broth composition

Bacto tryptone	10 g
Yeast extract	5 g
NaCl	10 g
Fill up to 1 L with dH2O and autoclaved	

- Table 3 Tfb I composition

KOAc	30 mM
RbCl	100 mM
$CaCl_2$	10 mM
$MnCl_2 \cdot 4H_2O$	60 mM
Glycerol 15%	
Fill up to 100 mℓ with dH2O	
Adjust to pH5.8 with CH_3COOH	
Filtrated with 0.2 μm nitrocellulose membrane	

- Table 4 Tfb II composition

MOPS	10 mM
$CaCl_2 \cdot 2H_2O$	75 mM
RbCl	10 mM
Glycerol 15%	
Fill up to 100 mL with dH_2O	

Adjust to pH6.5 with KOH	
Filtrated with 0.2 μm nitrocellulose membrane	

- Table 5 TE buffer composition

Tris-Cl（pH 8.0）	100mM
EDTA	10mM
Autoclave to sterilize	
Tris-Cl（1 M）	
Tris base	121.1 g
Adjust to pH 8.0 with HCl	
Fill up to 1 L with dH2O and autoclaved	

- Table 6 2×TY medium composition

dH_2O	900 mL
Bacto tryptone	16 g
Yeast extract	10 g
NaCl	5 g
Adjust to pH7.0 with 5N NaOH	
Fill up to 1 L with dH2O and autoclaved	

- Table 7 10×SLiCE buffer

Tris-HCl pH7.5	500 mM
$MgCl_2$	100 mM
ATP	10 mM
Dithiothreitol(DTT)	10 mM
Filtrated with 0.2 μm nitrocellulose membrane	

- Table 8 SOC medium composition

sol. A	
Bacto tryptone	2 g
Yeast extract	0.5 g

2M KCl	125 μℓ
5M NaCl	200 μℓ
Fill up to 93 mL with dH2O and autoclaved	
sol. B	
1 M $MgCl_2$	1 mℓ
1 M $MgSO_4$	1 mℓ
0.4 M Glucose	5 mℓ
Filtrated with 0.2 μm nitrocellulose membrane	
Add sol. B to sol.A	

- Table 9 LB-plate composition

Bacto tryptone	10 g
Yeast extract	5 g
NaCl	10 g
1.5% agar	
Fill up to 1 L with dH2O and autoclaved	

Appendix 2 Oligonucleotide primers used in this study

- Table 10 Oligonucleotide primers used for the construction of SMASh plasmid

ID	Sequence	Pair
1	CTTGTACAGCTCGTCCATGCCGAG	2
2	AATGTAACTGTATTCAGCGATGACGAAATTCTTAG	1
3	ATCACTCTCGGCATGGACGAGCTG	4,5
4	GGAGGATTACAATAGCTAAGAATTTCGTCATCGCTGAATACAGTTACATTGCCTCAGTAGAGA ACCTCCCTGTCAG	3
5	GGAGGATTACAATAGCTAAGAATTTCGTCATCGCTGAATACAGTTACATTTCATGTCCCGGAA GAGCCCTTATCGTCG	3
6	TGAGGCAATGTAACTGTATTCAGCG	8
7	GGCTGCGTGGTCATAGTGGGCAGG	8
8	CGTCAGGGTGACTGCTGGTGGAG	6,7
9	AACTCCTCTCCACCAGCAGTCACCCTGACGTGCGTGGTCATAGTGGGCAGGATCG	12
10	AACTCCTCTCCACCAGCAGTCACCCTGACGTCACACGGCTTCCCACCAGAGG	10
11	GACGATCCTGCCCACTATGACCACGCAGCCGACGTTGATCCTGGCTGAGGCGCAG	11
12	GTCATCGCTGAATACAGTTACATTGCCTCAGACGTTGATCCTGGCTGAGGCGCAG	9
13	CCAAGGCAGCACTTACCCGGCGCCGGCAGTAGTGGCGATATCATGGATTACAAGG	14
14	ACTGCCGGCGCCGGGTAAGTGCTGCCTTGGCTCTTCCATCTCATCCCCGGGCTTG	13
15	AACCACTACCTGAGCACCCAGTCCG	-
16	CGTCACCACTTTGTACAAGAAAGCTGGG	-
17	GACCGGACGAGTGTTGTCTT	-
18	GCGTAGAATCGAGACCGAGG	-

- Table 11 Oligonucleotide primers used for the construction of dual reporter virus bacmids

ID	Name	Sequence
1	mScarlet-I amplification_FW	GATTATTCATACCGTCCCACCATCGGGCGCATGGTGAGCAAGGGCGAGGCAG
2	mScarlet-I amplification_RV	TCTACAAATGTGGTATGGCTGATTATGATCCTAGTACAGCTCGTCCATGCCGCCG
3	pFastBac fragment amplification for SLiCE with mScarlet-I_FW	GATCATAATCAGCCATACCACATTTGTAG
4	pFastBac fragment amplification for SLiCE with mScarlet-I_RV	GCGCCCGATGGTGGGACGGTATGAA
5	sfGFP amplication_FW	CCCAAGAAGAAGCGCAAGGTGAGCAAAGGAGAAGAACTTTTCACTGGAG
6	sfGFP amplication_RV	TTTGTAGAGCTCATCCATGCCATGTG
7	p10 amplication_FW	AATGAATCGTTTTTAAAATAACAAA
8	p10 amplication_RV	GGATAACCGTATTACCGCCTTTGAGTGAGCATCGATCCCCGCGGCGTTAACTCGAATCGCTATCCAAGC
9	PEST amplication_FW	ATTACACATGGCATGGATGAGCTCTACAAATACAAGAAGCTTAGCCATGGCTTCCCG
10	PEST amplication_RV	TTATAAAACAATTGATTTGTTATTTTAAAAACGATTCATTTCACACATTGATCCTAGCAGAAGCACAG
11	I-SceI recognition site amplication_FW	TGGTACCGAGATCGAGGCCTGTCTAGAGAAGCCCGCCACCCCGCGGTAGGGATAACAGGGTAATATG
12	I-SceI recognition site amplication_RV	AAAGTTCTTCTCCTTTGCTCACCTTGCGCTTCTTCTTGGGCATATTACCCTGTTATCCCTACCGCGG
13	ie1 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATAGTCGTTTGGTTGTTCACGATCG
14	ie1 Promoter amplication_RV	TGCATGGAGCCGGGCCACCTCGACCTATCAATGTCTTTGTGATGCGCG
15	vp39 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATATTGTTGCCGTTATAAATATGGACC
16	vp39 Promoter amplication_RV	TGCATGGAGCCGGGCCACCTCGACCAAAAAAATTGACCAAAGCTTTTCTG
17	lef8 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATCGTAAAGCGATTATTGCACACTAAT
18	lef8 Promoter amplication_RV	TGCATGGAGCCGGGCCACCTCGACCGAAGCGTTTCCATTTTCCAACAAAG
19	p47 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATAATTATGGAAAATAACGCGGAATTC
20	p47 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAAATGTTTGTAGCTTGTTTCTGAAAAA
21	lef11 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATCGTTGTTGTCTATCTTTTTTAGAGT

22	lef11 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAATCTAGAAAAAATTTCCATACCACGA
23	lef9 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATATTACGTCTGTACACGAACACGTAT
24	lef9 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAACGGCCACAAATATTTTTACGGGCCC
25	lef3 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATATTGCTGTTGTTGTCAATATGTGGG
26	lef3 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAATTGGGCGTTTGTCAAATAATTTTGA
27	lef4 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATATTTATAACGGCAACAATATGGCGC
28	lef4 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAATCTAGAATGCTTCTTGTAGTTGCGT
29	p143 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATTTTGGCTATCGTGTTTATATTTTTCG
30	p143 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAATGTCCAAGTCCACGTAGCCATTG
31	lef5 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATGTTCTCGTTTTAAGCGAGTACGCAG
32	lef5 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAACGCATCGAGTGTGCTACATGATC
33	lef10 Promoter amplication_FW	CTTCTTGGGCATATTACCCTGTTATGATTTGTACAACTTTGGTACGTG
34	lef10 Promoter amplication_RV	GAGATCGGCTTCCCGGTAGGGATAAGATTGCATTTTAAAAAAATTCC

- Table 12 Oligonucleotide primers used for the genome-wide screening

ID	Name	Sequence
35	pCR8 amplification product from pCR8-ccdB_FW	CCCTATAGTGAGTCGTATTACGCCATC
36	pCR8 amplification product from pCR8-ccdB_RV	CCCTATAGTGAGTCGTATTACGCCCTCAG
37	egfp_FW	GCGTAATACGACTCACTATAGGGGGCAAGCTGACCCTGAAGTT
38	Bmnpvgp001_FW	GCGTAATACGACTCACTATAGGGAAGCGCAAGAAGCACCTAGT
39	orf1629_FW	GCGTAATACGACTCACTATAGGGACAGCGAAGTGTCTCGCTAC
40	Bmnpvgp003_FW	GCGTAATACGACTCACTATAGGGAAGCACTCGATCGCGTGTAT
41	Orf_4_FW	GCGTAATACGACTCACTATAGGGGATCGGCAGTTGCACCTTTG
42	Orf_5_FW	GCGTAATACGACTCACTATAGGGGGCAACAGGTTTGAGCAGTT
43	lef-1_FW	GCGTAATACGACTCACTATAGGGGATTGCGTACAACGACAGCC
44	egt_FW	GCGTAATACGACTCACTATAGGGGTTTAGAAAGCGCGGAGTGG
45	Orf_7a_FW	GCGTAATACGACTCACTATAGGGTTTTTGGTTTGTGCGACTGC
46	bv/odv-e26_FW	GCGTAATACGACTCACTATAGGGAAGAAGACGACGACAACGCA
47	Orf_9_FW	GCGTAATACGACTCACTATAGGGGGAAGAGCCCAATTGCGTTG
48	Orf_10_FW	GCGTAATACGACTCACTATAGGGCGACGACGACGACTACAAAA
49	Orf_11_FW	GCGTAATACGACTCACTATAGGGCGACGCAACACGACTACACT
50	arif-1_FW	GCGTAATACGACTCACTATAGGGGATCACTGGGCTGTGTTGGA
51	Orf_13_FW	GCGTAATACGACTCACTATAGGGCTTTGACCGCATTATGCCCG
52	Orf_14_FW	GCGTAATACGACTCACTATAGGGTCCTCGCATTGGAGCAACAT
53	pkip_FW	GCGTAATACGACTCACTATAGGGCAAAATGTTGTGCATAGCGG
54	dbp=Orf_16_FW	GCGTAATACGACTCACTATAGGGGCCAAAAGCAACTTTGGCCA
55	Orf_17_FW	GCGTAATACGACTCACTATAGGGGGACGGCTCTGTTGTTAGCA
56	iap1_FW	GCGTAATACGACTCACTATAGGGGCGCGTATGCCAACAAGATT
57	lef-6_FW	GCGTAATACGACTCACTATAGGGAAAGAGTTTTTGATTCATATTGCGCC
58	Orf_20_FW	GCGTAATACGACTCACTATAGGGTGTTCTCAAAGAATTACAACGCA
59	Orf_21_FW	GCGTAATACGACTCACTATAGGGCGGCGTTTAACCTGATGCTG
60	bro-a_FW	GCGTAATACGACTCACTATAGGGCTCATCACCAAGGAGGGTGT
61	Orf_22a_FW	GCGTAATACGACTCACTATAGGGTGCAAATGGCTTTCATGTTG
62	sod_FW	GCGTAATACGACTCACTATAGGGGAGGTTTGCACGGCTTTCAC

63	fgf_FW	GCGTAATACGACTCACTATAGGGCGGCGGTTTTTGGCAGTAAA
64	Orf_25_FW	GCGTAATACGACTCACTATAGGGACGATAAATTGTTTCGCTGTCGA
65	Bmnpvgp028_FW	GCGTAATACGACTCACTATAGGGAAAACATTGACGGGCAAAAC
66	39k_FW	GCGTAATACGACTCACTATAGGGTAAGATCAAGCAGCCCGAGT
67	lef-11_FW	GCGTAATACGACTCACTATAGGGAAAATTGCACGCACTTAGGC
68	Orf_29_FW	GCGTAATACGACTCACTATAGGGCGAAACGCTGCAGGATTGTT
69	p43_FW	GCGTAATACGACTCACTATAGGGACGTGCCAATTCTCGCAAAC
70	p47_FW	GCGTAATACGACTCACTATAGGGAAAAACAAATTTGCGTTTGAAAGCG
71	Orf_32_FW	GCGTAATACGACTCACTATAGGGCCACAATGATGAAACGAACG
72	gta_FW	GCGTAATACGACTCACTATAGGGTGTACGATAAGTTAAAATGTGAATCGG
73	Orf_34_FW	GCGTAATACGACTCACTATAGGGTGAACACCCGATACGCTACT
74	Orf_35_FW	GCGTAATACGACTCACTATAGGGCAAGCGCTTCGAGTTTTGGC
75	Orf_36_FW	GCGTAATACGACTCACTATAGGGTCTGCAAAAGTCAACGAACG
76	odv-e66_FW	GCGTAATACGACTCACTATAGGGCGCCAAGTTTGCGGAAATGA
77	ets_FW	GCGTAATACGACTCACTATAGGGTGATAGAACGCACAGTAACTCG
78	lef-8_FW	GCGTAATACGACTCACTATAGGGCGGCAGATTTTTGAGTTTTCCCA
79	Orf_40_FW	GCGTAATACGACTCACTATAGGGAAAATGCGCGTTTAGTGCTT
80	Orf_41_FW	GCGTAATACGACTCACTATAGGGAATTGCGACGCATTTTGCCT
81	Orf_42_FW	GCGTAATACGACTCACTATAGGGAAGATTATTGGCCAACGTGC
82	lef-10_FW	GCGTAATACGACTCACTATAGGGCGGACGTCAACCTCATCAAT
83	vp1054_FW	GCGTAATACGACTCACTATAGGGCACTTGTTGCGCAACGATCA
84	Orf_44_FW	GCGTAATACGACTCACTATAGGGAAAAAGTAGCGCTTGGAAAAA
85	Orf_45_FW	GCGTAATACGACTCACTATAGGGATGTTGCGATCAATCATACGAC
86	Orf_46_FW	GCGTAATACGACTCACTATAGGGTGAAACTTCCCGTTTACGCT
87	Orf_47_FW	GCGTAATACGACTCACTATAGGGGCAAAGCCGCCTACGAGATA
88	Orf_48_FW	GCGTAATACGACTCACTATAGGGTCAAAACCTTCCCTATAACGG
89	fp_FW	GCGTAATACGACTCACTATAGGGAAACGCAAATCGACGAAAAC
90	lef-9_FW	GCGTAATACGACTCACTATAGGGATCTGGTTGACATCTCGGGC
91	Orf_51_FW	GCGTAATACGACTCACTATAGGGGTAGCCGATGAAATTAGGCG
92	gp37_FW	GCGTAATACGACTCACTATAGGGGCTCTCACGGCTATCTGTCC

93	Bmnpvgp056_FW	GCGTAATACGACTCACTATAGGGGATCGACAACGCGCATGTTT
94	Orf_54_FW	GCGTAATACGACTCACTATAGGGCGAGCATACTTAATAGCGCGG
95	lef-3_FW	GCGTAATACGACTCACTATAGGGGTGCGAGGCTAAAGAAAACG
96	Orf_56_FW	GCGTAATACGACTCACTATAGGGGTTAGCGTGGAGGGAATTGA
97	Orf_57_FW	GCGTAATACGACTCACTATAGGGCATGTTTGCCGACACGTTGA
98	iap2_FW	GCGTAATACGACTCACTATAGGGCCACAGACGGTCGATTCAGA
99	Orf_58a_FW	GCGTAATACGACTCACTATAGGGAACCTGCTGCAAAGTAAAACTTG
100	Orf_59_FW	GCGTAATACGACTCACTATAGGGCAAAGCTCATTCGTTTGCAG
101	Orf_60_FW	GCGTAATACGACTCACTATAGGGAATACCCGATGAAAACGATCCT
102	Orf_61_FW	GCGTAATACGACTCACTATAGGGCCGTGGTCAAAACCACTCTT
103	Orf_62_FW	GCGTAATACGACTCACTATAGGGATATTTGTTGTTGGGCGCACT
104	vlf-1_FW	GCGTAATACGACTCACTATAGGGGGTTCGAGTCCGTGTTCGAT
105	Orf_64_FW	GCGTAATACGACTCACTATAGGGTGGACGTGCCATACTATCGG
106	Orf_65_FW	GCGTAATACGACTCACTATAGGGGGCGACGACTCTGTACACCA
107	gp41_FW	GCGTAATACGACTCACTATAGGGTGACATGTGCATCGACACGA
108	Orf_67_FW	GCGTAATACGACTCACTATAGGGACGTTTCAGCGCCTAAGAAA
109	Orf_68_FW	GCGTAATACGACTCACTATAGGGCCAGACATCATTGTGAACGC
110	p95_FW	GCGTAATACGACTCACTATAGGGGACAAGGACGACGTGACCAT
111	p15_FW	GCGTAATACGACTCACTATAGGGTTTATGAATGCTTTGGGCTTG
112	p30_FW	GCGTAATACGACTCACTATAGGGGTTAGACGACGCGGAAGACA
113	vp39_FW	GCGTAATACGACTCACTATAGGGAACGAAGAGGCGGTTAACGT
114	lef-4_FW	GCGTAATACGACTCACTATAGGGGACAAGATGGCCCCTCGAAA
115	Orf_74_FW	GCGTAATACGACTCACTATAGGGCTACACCTCCACCACCACCT
116	Orf_75_FW	GCGTAATACGACTCACTATAGGGACTCAATTTGCGTTCGTGCC
117	Orf_76_FW	GCGTAATACGACTCACTATAGGGTCTTACGGACGACCAATGCG
118	odv-e25_FW	GCGTAATACGACTCACTATAGGGAAATTCCTTAACCGAGGCGT
119	Bmnpvgp082_FW	GCGTAATACGACTCACTATAGGGTCGTACAGAGAAACCGACATCA
120	Orf_79_FW	GCGTAATACGACTCACTATAGGGACACATTGTTGTTTGGAGCGT
121	bro-b_FW	GCGTAATACGACTCACTATAGGGGGCGTTATACAGCTGTTCATGC
122	bro-c_FW	GCGTAATACGACTCACTATAGGGTTTGTTTGCCAACGAAATGA
123	38k_FW	GCGTAATACGACTCACTATAGGGTGTGCTTAGCAACAAACCGC

124	lef-5_FW	GCGTAATACGACTCACTATAGGGCGTCTACCGGCCATCTGTTT
125	Bmnpvgp088_FW	GCGTAATACGACTCACTATAGGGCCACATATGGTTTGACACGC
126	Orf_85_FW	GCGTAATACGACTCACTATAGGGTTGGAGATGCCTTCGACCAC
127	Orf_86_FW	GCGTAATACGACTCACTATAGGGGGCAACAACATGTCGCAAAC
128	Orf_87_FW	GCGTAATACGACTCACTATAGGGCAACCATCGGATTGCGCCAA
129	vp80_FW	GCGTAATACGACTCACTATAGGGGACGCCAATATTTCCACGGC
130	he65_FW	GCGTAATACGACTCACTATAGGGTGCTGTTCAACGTAGGCGAT
131	Orf_90_FW	GCGTAATACGACTCACTATAGGGCGAACAAAATACATGGACGAT TCG
132	Orf_91_FW	GCGTAATACGACTCACTATAGGGTTATATCAAGAGTGACGGGCG
133	Orf_92_FW	GCGTAATACGACTCACTATAGGGTGTGCTAGTGTACAGGCACG
134	Orf_92a_FW	GCGTAATACGACTCACTATAGGGTGTACGCGATGTATTTTTGC
135	Orf_93_FW	GCGTAATACGACTCACTATAGGGTACAACAACAACCGCAAACC
136	Orf_94_FW	GCGTAATACGACTCACTATAGGGCAAACCTCGCTCAATCACCA
137	Orf_95_FW	GCGTAATACGACTCACTATAGGGGCCACCAACGCTTTGCTAAA
138	Orf_95a_FW	GCGTAATACGACTCACTATAGGGTATTTCACGTCCCGCTTTTT
139	Orf_96_FW	GCGTAATACGACTCACTATAGGGGTTCCTAACGCGCTCAAAAA
140	Orf_97_FW	GCGTAATACGACTCACTATAGGGATAACGCCGACACCGAAACT
141	Orf_98_FW	GCGTAATACGACTCACTATAGGGACGTTGTAGAAGCGTGCGAT
142	Orf_98a_FW	GCGTAATACGACTCACTATAGGGTTGAATCAATTTGTCTTCGTGT T
143	Orf_99_FW	GCGTAATACGACTCACTATAGGGTATCCTGCTGTTTGTCGTGG
144	gcn2_FW	GCGTAATACGACTCACTATAGGGTCCTGCAGATTTTGGCTTTT
145	Orf_101_FW	GCGTAATACGACTCACTATAGGGTGGTGTTTTCGGACAATCAA
146	lef-7_FW	GCGTAATACGACTCACTATAGGGATGTCGAGCGTTACAAAGCG
147	chitinase_FW	GCGTAATACGACTCACTATAGGGAGGACGGCTGTTCCATAAGC
148	Bmnpvgp111_FW	GCGTAATACGACTCACTATAGGGTTTGACTGGCGTCGTCTCAA
149	gp64/67_FW	GCGTAATACGACTCACTATAGGGACTAGTAAATCAGTCATACCAA GGCT
150	p24_FW	GCGTAATACGACTCACTATAGGGGACATCGAACACGCACAACT
151	gp16_FW	GCGTAATACGACTCACTATAGGGGAACTTTTGGGCCACGTTTA
152	pp34_FW	GCGTAATACGACTCACTATAGGGTGCCTTTGCAAGCGTTTCAA
153	Orf_109_FW	GCGTAATACGACTCACTATAGGGATCCCATATTAAATACGCCTCTT GA
154	alk-exo_FW	GCGTAATACGACTCACTATAGGGGGTGGCGGTTTCTGTGCC

155	Orf_110a_FW	GCGTAATACGACTCACTATAGGGGTTTTGGACGCTGTGGAGTG
156	Orf_111_FW	GCGTAATACGACTCACTATAGGGCTTTGCATTCAAACGCTCTG
157	p35_FW	GCGTAATACGACTCACTATAGGGAAATCGACGTGTCCCAGACG
158	p26_FW	GCGTAATACGACTCACTATAGGGAGTTTCGGTTGTGACGACGT
159	p10_FW	GCGTAATACGACTCACTATAGGGCCATTGCGGAAACTAACACA
160	p74_FW	GCGTAATACGACTCACTATAGGGAGATGGCGCACTAGGTTTCC
161	me53_FW	GCGTAATACGACTCACTATAGGGCCCAACGAGCGCATTCAAAC
162	ie-0_FW	GCGTAATACGACTCACTATAGGGAAAGAAATTGAAGGCGCGTA
163	Orf_118_FW	GCGTAATACGACTCACTATAGGGTGGATTGGAATGGTGTGCGA
164	odv-e18_FW	GCGTAATACGACTCACTATAGGGTACTGGCGCTACGACTAGCA
165	odv-ec27_FW	GCGTAATACGACTCACTATAGGGTCAGCACCCTGGAAATGACC
166	Orf_121_FW	GCGTAATACGACTCACTATAGGGCTCATCTAAAGTGTCACAGCG A
167	Orf_122_FW	GCGTAATACGACTCACTATAGGGTCCGCCGACATACAATGTTA
168	ie-1_FW	GCGTAATACGACTCACTATAGGGTTAACGCGTCGTACACCAGT
169	odv-e56_FW	GCGTAATACGACTCACTATAGGGGCCGGTTTCACAAATGTGCT
170	Orf_125_FW	GCGTAATACGACTCACTATAGGGTCTAATCATGTTGGGTTTGAAC A
171	Orf_126_FW	GCGTAATACGACTCACTATAGGGTGTTCAAACCCAACATGATTAG A
172	ie-2_FW	GCGTAATACGACTCACTATAGGGCGTATTCAAGCAGTCGCAGC
173	pe38_FW	GCGTAATACGACTCACTATAGGGACAGACAGTTGGAAGACGCT
174	Orf_129_FW	GCGTAATACGACTCACTATAGGGAATAGATGTAAAATACCATATG CCG
175	ptp_FW	GCGTAATACGACTCACTATAGGGCGCGTTGGCACAACTATTTA
176	bro-d_FW	GCGTAATACGACTCACTATAGGGAGGACAAGCAGATTGAGGCC
177	bro-e_FW	GCGTAATACGACTCACTATAGGGGGCAATCCGCTGTATTTGCA
178	Orf_133_FW	GCGTAATACGACTCACTATAGGGTTGACTCGATTGCGCGATCT
179	Orf_134_FW	GCGTAATACGACTCACTATAGGGACGTCAAAAATTAACAATGCG CG
180	lef-2_FW	GCGTAATACGACTCACTATAGGGTGTCGTCGACATGCTGAACA
181	egfp_RV	GCGTAATACGACTCACTATAGGGTTCTGCTTGTCGGCCATGAT
182	Bmnpvgp001_RV	GCGTAATACGACTCACTATAGGGTCGTTGTTCATGCCCACGTA
183	orf1629_RV	GCGTAATACGACTCACTATAGGGTTAGGCACGGAAGAAGGTGG
184	Bmnpvgp003_RV	GCGTAATACGACTCACTATAGGGAGTACACAAAGTCACGAGCGT

185	Orf_4_RV	GCGTAATACGACTCACTATAGGGGTTGAACACGTTGGCGATGT
186	Orf_5_RV	GCGTAATACGACTCACTATAGGGTGCCAACAAGGGATTTGGGT
187	lef-1_RV	GCGTAATACGACTCACTATAGGGACGGCCTCTCTCACACAATG
188	egt_RV	GCGTAATACGACTCACTATAGGGTCCGTCATCAAGTTCGCCTT
189	Orf_7a_RV	GCGTAATACGACTCACTATAGGGACATAAGCGCGTGTGAACAG
190	bv/odv-e26_RV	GCGTAATACGACTCACTATAGGGGCTAGCAATCTCGTCCGGAT
191	Orf_9_RV	GCGTAATACGACTCACTATAGGGTCACTTTACGATCGGCGACA
192	Orf_10_RV	GCGTAATACGACTCACTATAGGGAAGGGTCGTTTCATAATGCG
193	Orf_11_RV	GCGTAATACGACTCACTATAGGGTGTTGAATCGTAACCCCGTC
194	arif-1_RV	GCGTAATACGACTCACTATAGGGTCCAGCATTGCACAGCTAGT
195	Orf_13_RV	GCGTAATACGACTCACTATAGGGACATAGAGGTCCTGTCCCCC
196	Orf_14_RV	GCGTAATACGACTCACTATAGGGGCAGCGTGTCGTCGTCTATA
197	pkip_RV	GCGTAATACGACTCACTATAGGGTTCACTTGCTGTTTTTCATCCA
198	dbp=Orf_16_RV	GCGTAATACGACTCACTATAGGGCGTCAATGGTGCCCATCATG
199	Orf_17_RV	GCGTAATACGACTCACTATAGGGAGCACGTTTGGGTCGTTTTG
200	iap1_RV	GCGTAATACGACTCACTATAGGGAACATTGCGGGTTTTCGGTG
201	lef-6_RV	GCGTAATACGACTCACTATAGGGTGCCAAAACATATCCGTCCA
202	Orf_20_RV	GCGTAATACGACTCACTATAGGGTTTCCATATTCTGCAGCTCG
203	Orf_21_RV	GCGTAATACGACTCACTATAGGGAGTTCAACAGCCCCTCTTCG
204	bro-a_RV	GCGTAATACGACTCACTATAGGGACGAACCCAAGCGTTTAATG
205	Orf_22a_RV	GCGTAATACGACTCACTATAGGGCTTGGTTAAGAAGGGTCCCG
206	sod_RV	GCGTAATACGACTCACTATAGGGCCAAACGGCCGTCAGAATTG
207	fgf_RV	GCGTAATACGACTCACTATAGGGCTAAACTTGCCCAGCTTGCG
208	Orf_25_RV	GCGTAATACGACTCACTATAGGGCGTATGCGTTTGTGTCCAGG
209	Bmnpvgp028_RV	GCGTAATACGACTCACTATAGGGCCCTCCTCGTAATCGTAACA
210	39k_RV	GCGTAATACGACTCACTATAGGGATCGAAGTGGGCACGTATTC
211	lef-11_RV	GCGTAATACGACTCACTATAGGGTAAACCTTTGAAACGACCCG
212	Orf_29_RV	GCGTAATACGACTCACTATAGGGGTGTACGTGTTGGGTTTGGC
213	p43_RV	GCGTAATACGACTCACTATAGGGTGATTGATCTCGAGCCACCG
214	p47_RV	GCGTAATACGACTCACTATAGGGTCAGTATAACGACTGGTCAAA ATTGA
215	Orf_32_RV	GCGTAATACGACTCACTATAGGGTTTTGCACAAAATACGGCAA

216	gta_RV	GCGTAATACGACTCACTATAGGGATTCCACCCATTGCGAAACC
217	Orf_34_RV	GCGTAATACGACTCACTATAGGGGTTGGGGTAAATTGGAATCG
218	Orf_35_RV	GCGTAATACGACTCACTATAGGGCGTTTACCGTTTCGTCGCAA
219	Orf_36_RV	GCGTAATACGACTCACTATAGGGATTTTTGTCCAACCACTCGC
220	odv-e66_RV	GCGTAATACGACTCACTATAGGGGAATTGGGACTCCGAACGGT
221	ets_RV	GCGTAATACGACTCACTATAGGGAACTGTTTAATTTTAATGCTAA CGC
222	lef-8_RV	GCGTAATACGACTCACTATAGGGTGTTGTTTGTGAAGCGTGTTGT
223	Orf_40_RV	GCGTAATACGACTCACTATAGGGGGCTTTACAATGTTTTCGGC
224	Orf_41_RV	GCGTAATACGACTCACTATAGGGTGTGTCTGTTGTCCGTCGAG
225	Orf_42_RV	GCGTAATACGACTCACTATAGGGATCGAGTTCAATGTCCAGGG
226	lef-10_RV	GCGTAATACGACTCACTATAGGGTACGTGGACGCGTTACTTTG
227	vp1054_RV	GCGTAATACGACTCACTATAGGGCTTTGCCCGCTTTCACTATCG
228	Orf_44_RV	GCGTAATACGACTCACTATAGGGATTGTAATCTCGCTTGCCGA
229	Orf_45_RV	GCGTAATACGACTCACTATAGGGACACAGGGTAAAATAGGGCAG A
230	Orf_46_RV	GCGTAATACGACTCACTATAGGGAAAACGTTTCGCACTTTTGG
231	Orf_47_RV	GCGTAATACGACTCACTATAGGGTTCGTCCTCATCTTCGCTGG
232	Orf_48_RV	GCGTAATACGACTCACTATAGGGGCAATTTCGTCTTCGGTTTC
233	fp_RV	GCGTAATACGACTCACTATAGGGGTAAAACGGCAACAGAGCGT
234	lef-9_RV	GCGTAATACGACTCACTATAGGGACCGTCGCAGTCTGTGTTAG
235	Orf_51_RV	GCGTAATACGACTCACTATAGGGTTCAAACATGTCTTCAAATTCG T
236	gp37_RV	GCGTAATACGACTCACTATAGGGAGGTTCGTCCATTCCGCTTT
237	Bmnpvgp056_RV	GCGTAATACGACTCACTATAGGGGCTGTAATCATGTCGACTTCAT T
238	Orf_54_RV	GCGTAATACGACTCACTATAGGGCGGTGCCGTTAAGTTCTCTG
239	lef-3_RV	GCGTAATACGACTCACTATAGGGCCAGCAGCATTGAGATTTGA
240	Orf_56_RV	GCGTAATACGACTCACTATAGGGTGCAACATTTTTACGCTTCG
241	Orf_57_RV	GCGTAATACGACTCACTATAGGGCTTGTTCGTTTTCGCGTCCA
242	iap2_RV	GCGTAATACGACTCACTATAGGGTGGCATCCGGAACAACGTAA
243	Orf_58a_RV	GCGTAATACGACTCACTATAGGGATTACCGTCGCTGTCGTCTT
244	Orf_59_RV	GCGTAATACGACTCACTATAGGGCATATGCTGATGCTGTTGTGAA
245	Orf_60_RV	GCGTAATACGACTCACTATAGGGACGCGACGTATTTTGATTTGAG T

246	Orf_61_RV	GCGTAATACGACTCACTATAGGGTGCTTCGTCCAGTTTTGTGA
247	Orf_62_RV	GCGTAATACGACTCACTATAGGGTCAATCTATTGAGCTGGTATTTTTG
248	vlf-1_RV	GCGTAATACGACTCACTATAGGGAGCGTTTCTTTGGTGACCGA
249	Orf_64_RV	GCGTAATACGACTCACTATAGGGAGGAGAATCAGTGTCAATACTATCCG
250	Orf_65_RV	GCGTAATACGACTCACTATAGGGGCTAACAGAAATTTATGCAACAAAA
251	gp41_RV	GCGTAATACGACTCACTATAGGGCACGTCACCGATTCCGCC
252	Orf_67_RV	GCGTAATACGACTCACTATAGGGTCGCAATTTTTAAAGGCCAC
253	Orf_68_RV	GCGTAATACGACTCACTATAGGGGTTTCTTAGGCGCTGAAACG
254	p95_RV	GCGTAATACGACTCACTATAGGGCGTCTGGCCTGTTTTCACAC
255	p15_RV	GCGTAATACGACTCACTATAGGGTGCCTCTTGTTTATTTTCACACC
256	p30_RV	GCGTAATACGACTCACTATAGGGTCGTTACTCAATCTGGCGTT
257	vp39_RV	GCGTAATACGACTCACTATAGGGTTCCTCCGTGTCGATTTGCA
258	lef-4_RV	GCGTAATACGACTCACTATAGGGCCACGGACACGGTCATGTAA
259	Orf_74_RV	GCGTAATACGACTCACTATAGGGGGATAGAACGGGCAGTCTGA
260	Orf_75_RV	GCGTAATACGACTCACTATAGGGTCTCCCCATGTCATGCGTTC
261	Orf_76_RV	GCGTAATACGACTCACTATAGGGTTATCGTTTTCCCAATTGACGGG
262	odv-e25_RV	GCGTAATACGACTCACTATAGGGATATCAAAGTTTGCGGTGCC
263	Bmnpvgp082_RV	GCGTAATACGACTCACTATAGGGGTCAGGCGGCAAGTTTAACG
264	Orf_79_RV	GCGTAATACGACTCACTATAGGGTCGAGAACGAGCGTGTGATC
265	bro-b_RV	GCGTAATACGACTCACTATAGGGCTCGGTCTGCAGCACGTAC
266	bro-c_RV	GCGTAATACGACTCACTATAGGGTGTGCTTGGCTTTGAACTTG
267	38k_RV	GCGTAATACGACTCACTATAGGGCGGTGGAACCTTCGCTGATA
268	lef-5_RV	GCGTAATACGACTCACTATAGGGGCGTACACCCAGTGTTGGTA
269	Bmnpvgp088_RV	GCGTAATACGACTCACTATAGGGGTGTTCTGTAACTTCGGCGG
270	Orf_85_RV	GCGTAATACGACTCACTATAGGGAACTTGCAATTGGACACGGC
271	Orf_86_RV	GCGTAATACGACTCACTATAGGGTGTCGTAAACGCTCCCCAAA
272	Orf_87_RV	GCGTAATACGACTCACTATAGGGTGTTTTCGTCCGAATACTTGCT
273	vp80_RV	GCGTAATACGACTCACTATAGGGCTCCAGCATCAGCATCGCTA
274	he65_RV	GCGTAATACGACTCACTATAGGGCGGGTCCGCGTTCATTTAAC
275	Orf_90_RV	GCGTAATACGACTCACTATAGGGGCTGATGGCATTAGTGTTGGT

276	Orf_91_RV	GCGTAATACGACTCACTATAGGGTGGTTTAACAAGATCACAACCAA
277	Orf_92_RV	GCGTAATACGACTCACTATAGGGCACAAGCAATCGGTCGCAAA
278	Orf_92a_RV	GCGTAATACGACTCACTATAGGGTTCAATAGGGTATCGGGCAC
279	Orf_93_RV	GCGTAATACGACTCACTATAGGGACCGCGATTGTAAAACTTGTG
280	Orf_94_RV	GCGTAATACGACTCACTATAGGGAGACTGTTAACAAACATCACGTTGC
281	Orf_95_RV	GCGTAATACGACTCACTATAGGGAAATTGGCCAAGCGATTCGG
282	Orf_95a_RV	GCGTAATACGACTCACTATAGGGGAAGAGTGTGGTGTGACGCA
283	Orf_96_RV	GCGTAATACGACTCACTATAGGGGCATAACAATTCAACGTTGTCAC
284	Orf_97_RV	GCGTAATACGACTCACTATAGGGCGGGCAGTAACAGGTCAGAG
285	Orf_98_RV	GCGTAATACGACTCACTATAGGGTTTTGTTTAGCATGACTGCCA
286	Orf_98a_RV	GCGTAATACGACTCACTATAGGGTTAAAAACTACGGCACATTTTGG
287	Orf_99_RV	GCGTAATACGACTCACTATAGGGTCAATTTGTCTTCGTGTTTGTTG
288	gcn2_RV	GCGTAATACGACTCACTATAGGGTGGCTTTGGGATAAACTTGC
289	Orf_101_RV	GCGTAATACGACTCACTATAGGGAATGTTAGAGAACGGTGGCG
290	lef-7_RV	GCGTAATACGACTCACTATAGGGCCGATTCAACGTGACGATGC
291	chitinase_RV	GCGTAATACGACTCACTATAGGGACGATCGTTGCAGAGCTTCA
292	Bmnpvgp111_RV	GCGTAATACGACTCACTATAGGGGCCATAGGAATAGGGCCGAC
293	gp64/67_RV	GCGTAATACGACTCACTATAGGGGTTGGGATCCAGCGAGCC
294	p24_RV	GCGTAATACGACTCACTATAGGGAAGTTGAAATGGCCTCGTTG
295	gp16_RV	GCGTAATACGACTCACTATAGGGGCACGTTGTTAACTACCCCG
296	pp34_RV	GCGTAATACGACTCACTATAGGGTTGCCTGGGAGACGATGAAC
297	Orf_109_RV	GCGTAATACGACTCACTATAGGGTGATGGATTTTCTTAGTTCTTGATCTT
298	alk-exo_RV	GCGTAATACGACTCACTATAGGGGGTATGCTCGCGTTTTAAAACA
299	Orf_110a_RV	GCGTAATACGACTCACTATAGGGCGTTGAAACGTGTTCGTCAA
300	Orf_111_RV	GCGTAATACGACTCACTATAGGGCGTGTAATTTTCGGGACCAT
301	p35_RV	GCGTAATACGACTCACTATAGGGTCAACACGCTGCCATTTTGG
302	p26_RV	GCGTAATACGACTCACTATAGGGTTAAAGGTCCCGGCATCCTC
303	p10_RV	GCGTAATACGACTCACTATAGGGAGCAGTGTCACCGGTCAATA
304	p74_RV	GCGTAATACGACTCACTATAGGGCTGTTGTCGGGCGTGTATCT
305	me53_RV	GCGTAATACGACTCACTATAGGGTGCAATAATTGTTAATTGTAGGTCTCG

306	ie-0_RV	GCGTAATACGACTCACTATAGGGATTCGCAACATTCTTTTGGC
307	Orf_118_RV	GCGTAATACGACTCACTATAGGGACCGTTCATAGCGGGGTTTT
308	odv-e18_RV	GCGTAATACGACTCACTATAGGGGAGGCGTGTTCATAAAGGGA
309	odv-ec27_RV	GCGTAATACGACTCACTATAGGGAAACGTCTTGGCTATCCGCA
310	Orf_121_RV	GCGTAATACGACTCACTATAGGGGTTTCTATACATGCGAGCCG
311	Orf_122_RV	GCGTAATACGACTCACTATAGGGAAAATGCACTGACACGTTGC
312	ie-1_RV	GCGTAATACGACTCACTATAGGGCAACAGGCAGCTCAAGGGAT
313	odv-e56_RV	GCGTAATACGACTCACTATAGGGGCAGTGCGTTTTGAACACCT
314	Orf_125_RV	GCGTAATACGACTCACTATAGGGGCGCCATGTTGGTAAAATCT
315	Orf_126_RV	GCGTAATACGACTCACTATAGGGTGTTTTGGTTAGCAGTACATCCA
316	ie-2_RV	GCGTAATACGACTCACTATAGGGTCTAAACCTCGCTGCAGAGC
317	pe38_RV	GCGTAATACGACTCACTATAGGGGCACAATGGACAGCACACAA
318	Orf_129_RV	GCGTAATACGACTCACTATAGGGTGATGCAAGAATTGTATGTTGTTG
319	ptp_RV	GCGTAATACGACTCACTATAGGGGCCGGGACACTTTTCTGTAA
320	bro-d_RV	GCGTAATACGACTCACTATAGGGGTTGGCAGCGTCGATGTTAC
321	bro-e_RV	GCGTAATACGACTCACTATAGGGGGCCGCAATTCAGAGAAACC
322	Orf_133_RV	GCGTAATACGACTCACTATAGGGAAATTCCCTCCGGCGTTGAT
323	Orf_134_RV	GCGTAATACGACTCACTATAGGGATTTCGCGGCTTCTTGCAC
324	lef-2_RV	GCGTAATACGACTCACTATAGGGAGCAAACGGACAGAGCTTGT