Toward Optimization of the Baculovirus Expression Vector System - Development of Genetic Tools to Improve Biologics Production

by

Mark R Bruder

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Bruno Gaillet Professor, Dept. of Chemical Engineering, Universite Laval
Supervisor:	Marc Aucoin Professor, Dept. of Chemical Engineering, University of Waterloo
Internal Member:	Valerie Ward Assistant Professor, Dept. of Chemical Engineering, University of Waterloo
Internal-External Member:	Brian Ingalls Professor, Dept. of Applied Mathematics, University of Waterloo
Internal-External Member:	Roderick Slavcev Associate Professor, School of Pharmacy, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

For Chapter 2, the author designed the experiments, conducted most of the data collection, and performed the analysis. Sadru-Dean Walji assisted with virus seed stock amplification and titrations via EPDA. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

For Chapter 3, the author designed the experiments, conducted the data collection, and performed the analysis. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

For Chapter 4, the author designed the experiments, carried out the data collection, and performed the analysis. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

For Chapter 5, the author designed the experiments, carried out the data collection and performed the bioinformatics analyses. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

Abstract

Small-molecule drugs have dominated the pharmaceutical industry and physician's prescription pads since the beginning of modern medicine. The most recent decades, however, are most aptly characterized as the era of biologics. Indeed, 7 of the top 10 revenue-generating therapeutics in 2020 were biologics, and although they are yet to be counted among the most commonly prescribed drugs, the drum beat of new product approvals and demand for biologic therapeutics is intensifying. Despite the potential of biologics as therapeutics, one of the factors that currently hinder their use is exorbitant cost of production. In preceding years, important developments in the optimization of media, feeding strategies, and downstream processing have led to significant improvement in the yield and decreased production cost of recombinant therapeutics. The most recent decade, however, has witnessed a revolution in the field akin to the systems biology approach toward rational design practiced for improving microbial production hosts; enabled by advancements in whole-genome sequencing, genome-scale models, and development of sophisticated genetic engineering tools, a new wave of engineered production hosts and platforms with improved or completely novel biochemical properties are being developed, leading to improved product yield and quality, and decreased production costs.

The baculovirus expression vector system (BEVS) is an established platform for the manufacture of recombinant proteins, viral vaccines, and gene therapy vectors. Despite the first recombinant protein being produced in the BEVS in the early 1980s, much of the intervening years has seen it utilized predominantly as a research tool in academic laboratories rather than as a commercial manufacturing platform. Consequently, relatively little attention has been devoted to its improvement as a production platform. Nevertheless, several BEVS-manufactured vaccines and therapeutics have recently been licensed for use in animals and humans, signifying that it may yet find mainstream use as a commercial manufacturing platform.

Although the BEVS boasts many features that make it attractive as a manufacturing platform, to realize its full potential, intrinsic limitations must be addressed: the lytic infection cycle and resulting short bioprocess duration can limit overall yield of recombinant proteins, and large amounts of progeny virus, cellular proteins, and debris from lysed cells are contaminants that necessitate extensive purification steps to achieve product purity. Additionally, genome instability remains a major barrier to scalability due to rapid loss of foreign gene expression. Although periodic reports in the literature describe strategies aimed at reducing contaminant progeny virions or improving yield and/or quality of the recombinant protein product by targeted deletion or addition of endogenous and heterologous genes in the baculovirus genome, genomes available commercially remain virtually unmodified. This thesis seeks to address these issues through the development of genetic tools aimed at optimization of the baculovirus genome.

We initiated this work by developing a platform for efficient scrutiny of baculovirus genes through targeted gene disruption and transcriptional repression using CRISPR-Cas9

technology. Using cell lines that were developed for constitutive expression of the Cas9 or dCas9 proteins, sequence-specific disruption or downregulation of target genes was achieved with efficiencies of up to 99%. The key factors affecting efficiency were choice of promoter for sgRNA expression and spacer sequence selection for gene targeting. CRISPR-mediated gene disruption was more effective than transcriptional repression in all cases. As a result of these findings, we confirmed sequence-specific disruption of the AcMNPV GP64 and VP80 structural proteins for recombinant protein production with reduced baculovirus contamination. Targeting these genes resulted in greater than 94% reduction in budded virus release. Importantly, production of the model cytoplasmic protein GFP and a model virus-like particle based on the HIV-1 Gag protein was not significantly affected by gene disruption, indicating that our approach could be more efficient than previously reported strategies.

Next, a microplate-based assay was developed to allow for efficient scrutiny of several baculovirus genes in parallel. The assay involved transfection of Sf9 cells constitutively expressing the Cas9 protein with a plasmid encoding a sgRNA, followed by infection with a recombinant BEV. Expression of a *gfp* reporter gene and analysis of infectious virus titer in cell culture supernatants were used as analogs for late gene expression and progeny budded virus release, thus providing insight toward the effect of various gene disruptions on the virus infection cycle. The critical factors established in the development of the assay included viable cell density, choice of transfection reagent, the amount of plasmid DNA transfected, the ratio of transfection reagent:plasmid DNA, the time interval between transfection and infection, virus multiplicity of infection, and the time interval between infection and harvest/analysis. The assay was used to scrutinize the effect of disrupting 13

AcMNPV genes, and the results agreed with those previously reported in all cases. Importantly, results could be realized in less than 2 weeks, which represented an improvement in efficiency of up to several months.

Finally, bioinformatics was used to select and evaluate baculovirus promoters with different expression characteristics than those routinely employed for foreign gene expression. We assesed the selected promoters by expressing model cytoplasmic and secreted proteins, and provided experimental evidence of the importance of promoter selection for foreign gene expression. We also examined sequence determinants that may be important for late gene transcription and translation initiation on a genome-wide scale.

Acknowledgements

First, I would like to extend my profound gratitude to my PhD advisor Marc Aucoin – you allowed me immense freedom to pursue my own research questions, make my own mistakes, and most importantly, time and patience so I could find success and learn from my experiences on my own. I have learned an incredible amount about being a scientist, leader, mentor, and person from you, for which I am grateful.

I would also like to thank Perry Chou, for initially getting me interested in the applied aspects of biology and helping instill in me the pragmatism and work ethic required to be successful in science during my MASc; Raymond Legge, for the many conversations we shared about teaching philosophy during all of the undergraduate courses I served as your Teaching Assistant; Stan Sokolenko, for teaching me the practical aspects of statistics, analysis, and encouraging me to learn R; and Aaron Frenette, for the many sessions of idea and experience sharing we've had. Your enthusiasm for science and life in general is truly infectious. I am particularly grateful to Professors Brian Ingalls, Roderick Slavcev, and Valerie Ward for sitting on my thesis advisory committee as well as Professor Bruno Gaillet for serving as an external examiner.

I have been privileged to share my PhD journey with my lab mates in the Aucoin lab. I have learned a lot from each of you, and deeply appreciate the many conversations, laughter, and constructive feedback sessions we've shared regarding our own research ideas and results.

Finally, I want to thank the National Science and Engineering Research Council for funding my research with the Doctoral Postgraduate Scholarship.

Dedication

I have had the profound privilege to share this journey – all of the highs and all of the lows – with one person. I am so thankful for you, and I can't wait to start the next chapter of our life together.

Table of Contents

List of Figures

List of Tables

Abbreviations

Note: only the most common abbreviations are represented here

AcMNPV : *Autographa californica* multiple nucleopolyhedrovirus

BEVS : baculovirus expression vector system

BV : budded virus

CRISPR: clustered regularly interspaced short palindromic repeats

Cas : CRISPR-associated protein

FP : fluorescent protein

HR : homologous recombination

IVT : infectious virus titer

KOV : knock out virus

MOI : multiplicity of infection

(M)NPV : (multiple) nucleopolyhedrovirus

ORF : open reading frame

RT-qPCR : reverse transcription quantitative polymerase chain reaction

Sf9 : ovarian cells of *Spodoptera frugiperda*, clone 9 (derived from Sf21)

TOI : time of infection

TR : transfection reagent

VCD : viable cell density

VLP : virus-like particle

Naming Nomenclature:

Gene names are referred to in italics (ex. *p10*; *polyhedrin* and *polh* commonly refer to the AcMNPV *polyhedrin* gene)

The first letter is capitalized for **protein names** (ex. P10)

Promoters are referred to in lower case, non-italics font (ex. p10)

"Δ" symbol refers to a DNA sequence deletion or truncation (ex. Δ*gp64*)

Introduction

Since Jonas Salk demonstrated efficient replication of poliomyelitis virus in cultured HeLa cells leading to the first effective polio vaccine more than 50 years ago, animal cell culture has played an important role in the development and production of medical treatments and therapeutics [1]. Previously deployed as virus replication hosts for production of 'native' human and veterinary vaccines, the development of recombinant DNA biotechnology in the 1970s and 1980s quickly catapulted animal cell culture to the forefront as a workhorse of the biopharmaceutical industry for production of vaccines and protein therapeutics [1, 2].

The first recombinant human protein licensed for therapeutic use was human insulin in 1982 by Genentech. Due to the structural simplicity of the insulin molecule, it was produced in the bacterium *Escherichia coli*, which is much more robust, fast growing, amenable to genetic modification, and has much less stringent growth requirements as compared to animal cells [2]. Expression of large proteins or proteins that require complex modifications for functionality, however, is problematic in prokaryotic hosts, and the cellular machinery required for glycosylation and other post-translational modifications necessary for bioactivity are only available in eukaryotic cells. This has prompted the

development of bioprocesses based on eukaryotic cell platforms for production of complex recombinant proteins [2, 3]. To this end, several eukaryotic cell culture systems have been developed as recombinant protein expression hosts, including yeast and filamentous fungi (e.g., *Saccharomyces cerevisiae* and *Aspergillus niger*), insect (e.g., *Spodoptera frugiperda*, *Drosophila melanogaster*, and *Trichoplusia ni*), and mammalian cell lines isolated from various organisms (e.g., human, monkey, hamster, rat, and mouse) [3]. Chief among these expression systems are mammalian cell lines derived from ovarian cells of the Chinese hamster (CHO) and human embryonic kidney (HEK 293) [3, 4]. As of 2018, 374 biopharmaceutical products (excluding vaccines) have gained regulatory approval in the US and EU. The ascendancy of mammalian production platforms is near absolute; approvals of biologics produced in mammalian systems has outpaced non-mammalian hosts 4:1 since 2014 and now represent over 50% of recombinant therapeutics approved for human or veterinary use to date [5]. Together, biologics generate over $150 billion USD annually in revenue [6, 7].

An alternative, non-mammalian, production platform that is nevertheless capable of producing large and complex recombinant proteins is the baculovirus expression vector system (BEVS). Despite the dominance of mammalian cells in biopharmaceutical production, the last decade has seen ~10 BEVS-manufactured products approved by the FDA for veterinary or human use, including the first recombinant adeno-associated virus (rAAV) gene therapy ever approved.

The BEVS platform consists of a recombinant *Autographa californica* baculovirus vector (AcMNPV) capable of infecting cultured cells isolated from the ovarian tissue of its natural replication hosts, the Fall Armyworm (*Spodoptera frugiperda*; Sf21 or Sf9 cells) or the

Cabbage Looper (*Trichoplusia ni*; High Five™ or Hi-5 cells) [8]. The recombinant AcMNPV expression vector (rBEV) is engineered to express the desired recombinant protein upon infection, typically under the control of the very strong, virus-derived, very late promoters p10 or polh, the latter of which is among the strongest promoters known in nature [8].

Despite having demonstrated successful expression of a non-native recombinant protein just a year after expression of human insulin in *E. coli* and predating the production of human tissue plasminogen activator (tPA) in CHO cells by 4 years [1], adoption of BEVS on a commercial scale has been slow [6, 8, 9]. Nevertheless, thousands of recombinant proteins have been expressed using this system and it has found broader adoption in the scientific research community, particularly in structural biology applications [10–12].

Historically, improved recombinant protein yields were afforded by improving the bioprocessing strategies employed; optimization of medium composition and feeding strategies, as well as the operating conditions of the bioreactor and downstream processing have been the primary drivers of these improvements [13]. A more recent emerging trend, however, is the progression toward improving yields through engineering the cell used for their production. Developments in 'omics' technologies - genomic, transcriptomic, proteomic, and metabolomics - have provided insights of the physiology and underlying cellular processes of animal cells that contribute to recombinant protein yield, and sophisticated genetic engineering technologies have enabled developing cell lines with improved specific productivity. However, genetic engineering tools available for manipulation of the BEVS remain underdeveloped and far out-paced by many prokaryotic and eukaryotic systems [2, 3, 8, 10].

The overall objective of this research is to broadly address this issue by establishing and validating genetic engineering strategies based on CRISPR-Cas9 technology for targeted gene disruption and transcriptional repression. A robust and sensitive system for gene disruption will enable efficient scrutiny of the rBEV genome toward its optimization by removing non-essential genes. For genes not amenable to complete disruption, targeted downregulation (CRISPRi) of genes may provide insight toward rewiring the rBEV to prolong the infection cycle and improve recombinant protein production. Finally, identifying promoters having different spatiotemporal expression profiles may allow for further optimization of expression for individual recombinant proteins to maximize yield and/or product quality. The thesis is divided into 5 chapters. Chapter 1 consists of a broad literature review of the relevant history, features, and applications of the BEVS, as well as an overview of CRISPR-Cas technology and its application in insect cell biotechnology. Chapter 2 presents the development of CRISPR-Cas9 for targeted gene disruption and transcriptional repression in the BEVS, while Chapter 3 presents the framework for an assay to efficiently scrutinize the effect of AcMNPV gene disruptions on late gene expression and rBEV replication. Chapter 3 confirms the site-specific gene disruption of an AcMNPV gene and examines expression of a model, complex, recombinant protein biologic in this system. Finally, Chapter 5 identifies and characterizes endogenous AcMNPV gene promoters with different expression profiles for use in the BEVS and examines sequence determinants that may impact late gene expression. Chapters 2-4 have been formatted for submission.

Chapter 1

Literature Review

1.1 General biology of *Autographa californica* multiple nucleopolyhedrovirus

1.1.1 Structural characteristics

Autographa californica multiple nucleopolyhedrovirus (AcMNPV) is the type species of the *Alphabaculovirus* genus of the Baculoviridae family of viruses, whose natural host range include invertebrate arthropods belonging to the orders Lepidoptera, Hymenoptera, Diptera, and Decapoda [14]. AcMNPV is a large, enveloped virus with a rod-shaped nucleocapsid approximately 300 nm in length and 50 nm in diameter [14]. Its circular, double-stranded DNA (dsDNA) chromosome is approximately 134 kbp in size, possesses an A+T content of 59% and encodes an estimated 154 open reading frames (ORFs) [14, 15]. Genome sequence data for AcMNPV has been available since the early 1990s, and an analysis of its transcriptome over the course of infection in High Five™ cells (*Trichoplusia*

ni) was recently reported [15, 16].

1.1.2 Pertinent gene expression of the AcMNPV infection cycle

The infection and replication cycle of AcMNPV is divided into two general phases: the early phase, during which viral genes are recognized and transcribed by the host cell RNA polymerase II, and the late phase, characterized by gene recognition and transcription by a virus-encoded RNA Polymerase. Both the early and late phases can be additionally subdivided into 'immediate early' and 'delayed early', and 'late' and 'very late' phases, respectively [16]. The immediate early and delayed early genes encode predominantly transactivators essential for subsequent viral gene expression and subversion of host cell gene expression and metabolism. The onset of the late phase of infection is marked by onset of viral DNA replication, and the activity of a virus-encoded RNA polymerase for expression of viral structural components necessary for the assembly of new nucleocapsids and envelope proteins [17]. The 'very late' class of genes, which is unique to baculoviruses, is comprised of the very strongly expressed *polyhedrin* (also referred to as *polh*) and *p10* genes, which encode Polyhedrin and P10 proteins. They form the major components of the matrix of occlusion bodies in the nucleus of infected cells, and are found associated with fibrillar structures in both the nucleus and cytoplasm [18]. Importantly, the very late class of genes is dispensable for progeny virus production [8, 10].

Baculovirus infection causes cell cycle arrest of insect cells at the G_2/M phase [19]. Host gene transcription and translation are significantly affected, with the majority of host genes down-regulated over the course of infection, and translation of host proteins is completely

abolished toward the late stages of infection [16, 17, 20–25]. Infection also triggers several stress responses by the host cell in an effort to prevent virus replication and dissemination, including the activation of apoptosis, DNA damage, and heat shock responses [17]. Among these, apoptosis, or programmed cell death, is an evolutionarily conserved mechanism among all multicellular organisms [26]. The initiation of the apoptotic pathway may be stimulated by baculovirus DNA replication, which triggers a host cell signal cascade involving a family of initiator and effector cysteine-dependent aspartate-directed protease (caspase) proteins. Caspase activation ultimately culminates in apoptotic body formation, cell shrinkage, membrane blebbing, condensation of chromatin and DNA fragmentation, and cellular death [26, 27]. In order to counteract this defense system, baculoviruses have evolved to block premature apoptosis in infected cells by encoding proteins that are direct inhibitors of caspases [27, 28]. Notably, the AcMNPV-encoded *p35* gene is transcribed in both the early and late stages of the infection cycle, and its protein product P35 strongly inhibits the activity of the effector or 'death' caspase, Caspase-1, of Sf9 cells (Sf-Caspase-1) [27, 29–31]. Other baculovirus anti-apoptosis proteins include P49, which inhibits the initiator caspase Sf-Caspase-X, and a class of proteins called inhibitor of apoptosis (IAPs) which target pro-Sf-Caspase-1 proenzyme and block its activation [27, 32, 33]. Although putative *iap* genes are present in almost all baculovirus genomes sequenced, only a small subset tested so far have shown anti-apoptotic activity, suggesting that either they may serve different roles in the infection process, or anti-apoptosis activity is restricted to certain scenarios or cell types [17]. Interestingly, the *iap* class of genes encoded by baculoviruses share homology with putative *iap* genes of insect cells, suggesting that baculovirus may have commandeered this strategy from their hosts [27]. AcMNPV encodes two putative

iap genes, however neither show anti-apoptosis activity [33, 34]. No homolog of P49 is present in the AcMNPV genome [15, 27].

In addition to late expression factor (lef) proteins stimulating host cell apoptosis, lef proteins also contribute to host protein translational arrest [23]. Although infected cells display a significant decline in host gene transcription, AcMNPV apparently up-regulates the expression of several host genes, particularly those related to metabolism and the tricarboxylic acid (TCA) cycle [22, 24, 25]. Additionally, members of the 70 kDa heat shock protein (HSP70) and heat shock cognate (HSC70) family are reportedly up-regulated in Sf9 cells during infection, ostensibly to support the productive replication of baculovirus DNA and assembly and maturation of progeny virions [17, 24, 25, 35, 36] (Figure 1.1).

1.2 Historical milestones leading to the 'invention' of BEVS

As introduced above, the preeminent scientific discoveries related to the Baculoviridae family and AcMNPV in particular that figured prominently in the development of BEVS as a tool for recombinant protein production occurred over the course of three decades during the middle of the past century [8]. Amazingly, however, observations of baculovirus appeared in the scientific literature dating as far back as 1856, 50 years before the derivation of the name 'polyhedrosis' and more than 100 years prior to the isolation of AcMNPV [37]. Over the course of those intervening 100 years, research focused around purification of polyhedra from infected insect tissue and recovery of infectious baculovirus virions from these polyhedra, studying the mechanism of infection of the midgut cells of the host or-

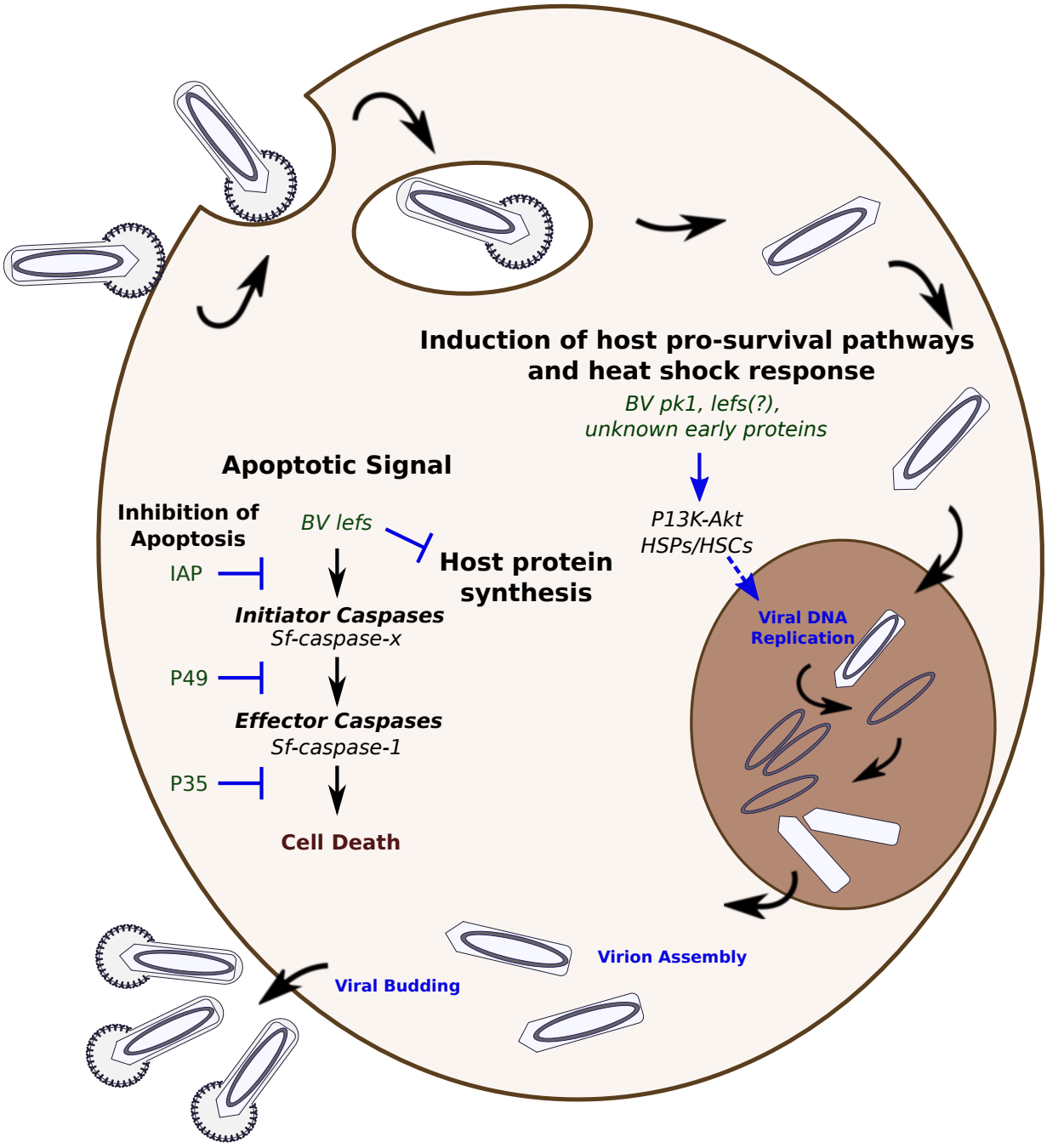


Figure 1.1: **Baculovirus gene expression and host cell response.** Important proteins produced by baculovirus shown in green, and their effect on the host cell is shown in black.

ganism, and demonstration of baculovirus infection in healthy, cultured insect tissue, *in vitro* [37].

The establishment of the first insect cell lines for *in vitro* culture in the early 1960s [38–40] was followed quickly by the demonstration of the presence of two infectious baculovirus forms in an infected insect; while virus found in the hemolymph of an infected insect could infect cultured tissues *in vitro*, virus purified from polyhedra could not [41]. This observation led to the understanding that occlusion-derived viruses (ODVs) found in the polyhedral-shaped occlusion bodies (OBs) were functionally and structurally distinct from non-ODVs found in hemolympth, which were termed budded viruses (BVs) [37, 38, 41]. Equipped with established, continuously-cultured Lepidopeteran cell lines in which baculovirus BVs were capable of replication [38, 40, 42, 43], the discovery and isolation of AcMNPV was reported in 1971 [44].

Owing to its wide host range and ability to infect cells *in vitro*, AcMNPV garnered interest in the research community. Thereafter, significant advancements surrounding its biology and biochemical tools for analysis were quickly established [37]. In addition to a purification procedure for viruses, procedures for isolation of viral proteins and DNA followed [45, 46]. The structural differences between ODV and BV forms were discerned and the plaque assay technique for isolation and quantification of AcMNPV was developed [47–50]. With the ability to isolate infectious foci from plaques, comparisons of the structural proteins and genomes of BV and ODV forms using SDS-PAGE and restriction endonuclease (REN) analyses were performed, and the presence of natural genetic variants was confirmed [51, 52]. The use of REN analysis also allowed the molecular cloning and mapping of individual genomic fragments and led to the first physical consensus map of the AcMNPV

genome and development of transfection methods for introducing baculovirus DNA into insect cells [53–55]. By analyzing protein patterns in infected cells, the temporal pattern of virus-specific protein expression was examined, and it was observed that expression was roughly divided into three distinct temporal classes [21, 56]. Significantly, production of Polyhedrin protein was reported toward the end of infection, and was proven to be virus-encoded and subsequently mapped to the genome and sequenced [57–59]. Translational maps of the AcMNPV genome also appeared, and a second, highly expressed late gene, *p10*, was reported [60–63].

Each of these studies contributed to form the extensive scaffold upon which the capstone discovery leading to BEVS was made: by constructing a series of plasmids with deletions within the *polyhedrin* ORF and co-transfecting them into insect cells along with wild-type AcMNPV DNA, mutant virus having an occlusion negative phenotype was observed and subsequently plaque purified. Characterization of this phenotype revealed deletions in the *polyhedrin* gene, confirming that the Polyhedrin protein was not essential for replication of virus in cell culture [64]. The coding sequence for the human beta interferon (IFN-β) gene was cloned under the control of the polh promoter and expressed, confirming AcMNPV as a tool for expression of recombinant proteins [65].

1.3 Establishing BEVS as a platform for production of complex biologics

1.3.1 Development of baculovirus vectors

The earliest recombinant baculovirus expression vectors (rBEVs) were isolated through a homologous recombination (HR) process. Wild-type AcMNPV DNA was co-transfected with a transfer plasmid containing the coding sequence of the foreign gene fused to the polh promoter. The co-transfection produced a mixture of parental and recombinant progeny, in which the coding sequence for the recombinant gene was inserted at the *polyhedrin* locus in the AcMNPV genome [10, 65, 66]. Progeny rBEVs could then be isolated by tedious plaque assays based on the OB-negative phenotype [64]. The HR process, however, was highly inefficient, with recombination frequencies of only ~0.1% [10]. The first major technological improvement for isolation of rBEVs occurred in 1990 when a unique *Bsu*36I REN site in the *polyhedrin* locus was used to linearize AcMNPV DNA prior to co-transfection. This development increased the frequency of HR to ~25-30% [67]. The selection of rBEVs was further improved to greater than ~90% by engineering an additional *Bsu*36I REN site within an essential gene (*orf1629*) of AcMNPV [68]. Digestion of wt AcMNPV DNA with *Bsu*36I resulted in linearization of the chromosome and concomitant deletion of the essential gene, with the resulting linearized AcMNPV DNA named BacPAK6 [68]. Co-transfection of the linear BacPAK6 DNA with a transfer vector containing the foreign gene in addition to the complementary *orf1629* coding sequence allowed for HR to restore the full-length *orf1629* gene and accompanying replication of rBEVs [68]. As with previous techniques, isolation

of rBEVs via plaque assay was still a requirement to remove any parental wt AcMNPV virus that escaped digestion with *Bsu*36I [10]. A method to select for rBEVs without the need for the plaque assay was developed in the same year, and was based on insertion of a foreign gene into the *polyhedrin* locus of AcMNPV through site-specific transposition into the baculovirus genome propagated in *E. coli* [69]. The AcMNPV genome was modified to contain a bacterial artificial chromosome (BAC), consisting of a mini-F replicon and antibiotic selectable marker, which permitted replication and selection of the 'bacmid' in *E. coli*, an *att*Tn7 site, which is the target sequence for the bacterial transposase Tn7, and a *lacZ* gene for blue-white screening [69]. By performing the generation of the rBEV genome and subsequent purification steps in *E. coli* rather than Sf9 cells, the need for the plaque assay was redundant [69]. This development greatly simplified the production of rBEVs, and as a result, bacmids are used widely for generating rBEVs for expression of recombinant proteins. Additionally, capitalizing on very efficient genetic engineering tools based on HR for bacteria, viruses with gene knockouts for function studies or with undesirable genes, such as those that affect the quality or stability of recombinant proteins, have also been developed [8]. Despite these advantages, bacmid systems may be limited to research or pre-clinical applications, as bacterial sequences and antibiotic resistance genes present in the viral DNA is undesirable. Bacmid-derived AcMNPV vectors are reportedly relatively unstable; non-infectious, defective interfering (DI) viruses accumulate upon repeated propagation in cell culture [70]. For commercial applications, rBEVS are still made by classical HR between linearized AcMNPV DNA and transfer plasmids harboring the heterologous gene flanked by homologous sequence required to restore the *orf1629* deletion on the viral genome [8]. Recently, several commercially-available systems have been

developed that combined the advantages of the bacmid but retained the HR-dependent production of rBEVs based on the triple-cut, linear DNA BacPAK6 system, including the *flash*BAC™ system licensed by Oxford Expression Technologies Ltd, and BacMagic™ (Novagen) [71, 72]. In these systems, the AcMNPV gene has a truncated *orf1629* gene, and a BAC containing a replicon and antibiotic resistance gene integrated in place of the *polh* gene to enable propagation of the circular viral genome in *E. coli* [72]. The viral genome can be isolated and purified from *E. coli*, and co-transfection with a transfer plasmid containing the foreign gene to insect cells will result in HR between the transfer plasmid and AcMNPV genome. In this way, integration of the foreign gene is coupled with the concomitant reconstitution of a full-length *orf1629* gene to permit viral replication and removal of bacteria-derived sequences [72]. Additionally, systems for generating rBEVs for expressing protein complexes comprising multiple subunits from a single baculovirus vector and at multiple different loci within the baculovirus genome have been developed [8, 73–75].

1.3.2 Engineering BEVS to improve recombinant protein expression

Deletion or silencing of endogenous AcMNPV genes

The field of recombinant protein expression has benefited immensely from advancements made in synthetic biology and recombineering technologies that enable targeted genome modification in a diverse range of cell types and organisms [76, 77]. Capitalizing on these advancements in genetic engineering tools, attention has been directed toward identifying and disrupting endogenous AcMNPV-encoded genes that may be detrimental for recom-

binant protein production quantity, quality, or stability [77]. Following the original disruptions of *polh* and *orf1629* genes that were fundamental to development of the BEVS [64, 68], the *v-cath* and *chiA* genes, encoding the cysteine protease cathepsin and chitinase, respectively, were identified as targets for disruption [73, 78, 79]. The chitinase and cathepsin proteins are involved in liquefaction of the infected insect host, and evidence suggested that their activity was contributing to proteolytic breakdown of recombinant proteins in cell culture [73, 78]. Production of recombinant proteins delivered with a Δ*chiA*Δ*v-cath* rBEV led to the host cells staying largely intact 72 hours post infection (hpi) and reduced proteolytic breakdown, indicating that these gene knockouts contribute significantly to recombinant protein stability [73, 78, 80]. Additionally, AcMNPV-encoded *p10*, *p26*, and *p74* genes have also been deleted from its genome [81]. These ORFs are located adjacent to each other in the genome, and encode proteins involved in infectivity of ODVs (P10 and P74), or have as yet undefined function (P26). All three genes, nevertheless, have been shown to be nonessential for infectivity of AcMNPV in cell culture [81]. Interestingly, while the decline in viability was slowest for the BacPAK6 AcMNPV genome, the Δ*chiA*Δ*v-cath* rBEV had higher viability at 72 hpi than the Δ*chiA*Δ*v-cath*Δ*p10*Δ*p26*Δ*p74* rBEV. This observation was contrary to a previous report that rBEVs with Δ*p10* disruptions improved viability [82], and potentially suggests that deletion of *chiA*, *v-cath*, *p26* and/or *p74* may have an effect on cell viability as well [81]. Nevertheless, the yield of 6 recombinant proteins was improved when the Δ*p10*Δ*p26*Δ*p74* deletions were included in addition to the Δ*chiA*Δ*v-cath* disruptions [81]. Despite the decrease in cell viability, each of the deletion mutants produced higher yields of recombinant protein, and at an earlier time point than the control, demonstrating that these gene disruptions improve quantity, quality, and

stability of recombinant protein expression [81].

Strategies aimed at eliminating the co-production of progeny virus have also been developed in order to reduce the burden of baculovirus removal on the downstream processing workflow. Typically, the target product is purified by a combination of filtration and chromatography techniques where significant losses can occur at each step. Moreover, rBEVs and enveloped virus-like particles (VLPs) are often similar in density and thus difficult to separate using ultracentrifugation and other chromatography-based techniques [83, 84]. To this end, AcMNPV structural proteins that are required for virion assembly can be targeted for disruption, resulting in abrogated assembly and/or budding of progeny virus. With this approach, a *trans*-complementing cell line engineered to express the deleted gene is required for production of infectious virus, which are subsequently used to infect cells not expressing the transgene for virus-free production of the target protein or VLP [83, 84]. The *vp80* gene, encoding the capsid-associated protein VP80, and *gp64*, which encodes the major envelope protein GP64, have been disrupted using this approach, leading to expression of recombinant enhanced green fluorescent protein (EGFP) and HIV-1 Gag VLPs with reduced baculovirus contamination in the supernatant, respectively [85, 86]. In both studies, however, it was noted that replication of the virus was impaired in the *trans*-complementing cell line, and overall expression of Gag may have been lower in the *gp64*-disrupted rBEV than the control [85, 86].

Genome instability is an intrinsic property of AcMNPV and possibly baculoviruses in general, when they are propagated in cell culture; spontaneous loss of large DNA fragments from the AcMNPV genome and enrichment of fragments containing origins of replication (ori) lead to accumulation of DI particles that are subsequently unable to propagate inde-

pendently [72, 87]. The accumuluation of DI particles results in reduction of the amount of infectious virus and corresponding loss in recombinant protein expression, as large DNA deletions often include the inserted foreign gene [70, 72, 88]. The genome of AcMNPV contains five homologous repeat (*hr; hr1-5*) regions dispersed throughout its genome that function as oris for DNA replication as well as transcriptional enhancers [72]. However, other non-*hr* oris have also been identified in the AcMNPV genome [70, 89–92]. To combat genome instability, the non-*hr* ori located within the *p94* open reading frame was deleted, resulting in stabilization of the genome [70]. However, recombinant protein expression was not prolonged, as excision of the bacteria-derived sequences (i.e. mini-F replicon, antibiotic resistence gene, and inserted foreign gene) was detected upon serial passage [70]. This was remedied by insertion of an extra copy of the *hr-1* sequence in the BAC adjacent to the foreign gene in an orientation-dependent fashion [93]. It is not yet clear whether this is as drastic an issue with the more recent developments in rBEV production systems such as *flash*BAC or BacMagic that do not contain BAC sequences [71, 72]. Nevertheless, another method for maintaining expression of foreign genes has been demonstrated, in which its expression is coupled to the expression of an essential AcMNPV gene using a bicistronic transcript with the two genes separated by an internal ribosome entry site (IRES) [93, 94].

Recently, an intensive genetic engineering project has been undertaken to improve BEVS by rewiring the entire BEV genome to enhance DNA stability and recombinant protein production [77]. Bioinformatics techniques were employed for comparative analysis with other members of the *Alphabaculovirus* genus to scrutinize gene synteny, in addition to extensive literature mining to identify putative essential and nonessential genes in the AcMNPV genome. Further, techniques were used to identify promoter motifs and repetitive

features in origins of replication that may be HR 'hot spots' [77]. The authors noted that the result of this analysis suggested that 62 genes could be nonessential, while 94 genes were classified as putative essential genes based on conservation or gene synteny, or whether they were unique to AcMNPV. However, it is also important to note the challenges this type of approach present; the Baculoviridae family is incredibly diverse in their genome content and organization, with genome size and annotated ORFs ranging from ~80-180 kbp and ~90-180, respectively [95]. Moreover, only ~30 genes appear to be conserved among all members, even as sequenced and annotated baculovirus genomes become more available [95]. As of a decade ago, only 60 of ~150 AcMNPV ORFs had a proven function whereas 72 were engimatic [96], and although functions have been assigned to additional ORFs since then, as of 2019 a significant proportion of putative gene annotations still relied on information from the closely related *B. mori* NPV (BmNPV) or inferred from homologs from other more distantly related viruses [97]. To this end, BmNPV is considered a close phylogenetic relative of AcMNPV, and while other group I *Alphabaculovirus* members such as *Plutella xylostella* MNPV and *Rachiplusia ou* MNPV appear in fact to be closer phylogenetically-related variants of AcMNPV, BmNPV has been the focus of considerably more research attention [95, 97]. Significantly, the BmNPV bacmid system was used to generate a library of single gene knockout viruses (KOVs) encompassing each of the 141 ORFs in its genome and evaluated for their infectivity phenotype [98]. The results revealed that 96 KOVs showed the ability to expand infections that were equivalent or slightly delayed compared to the wild-type control, while 37 KOVs expressed the GFP reporter but could not propagate infection, and only 10 showed neither GFP expression nor infectivity [98]. It was further noted that 45 of these KOVs had not been previously reported for

either BmNPV or AcMNPV, and perhaps more significantly, 9 genes that were previously identified as essential genes appeared to be nonessential according to this dataset. Since KOVs were obtained using conventional HR in insect cells prior to the development of the bacmid system for AcMNPV in the mid-1990s [69], it is possible that some ORFs were interpreted as essential due to failure of site-specific HR leading to a mutant AcMNPV genotype that differed from what was desired [98]. Caution must be taken when inferring AcMNPV gene function from BmNPV homologs, however, as virus-host interactions can have significant impacts on the phenotypes observed. For example, *Mamestra brassicae* NPV, which has a wide host range similar to AcMNPV, showed different protein expression profiles for both ODV and BVs in different infection hosts [99]. Similarly, the AcMNPV-encoded *hcf-1* gene, encoding host cell-specific factor 1, is dispensable for infectivity in Sf21 and Sf9 cells, but is required for DNA replication in High Five and Tn-368 cells. Further, expression of Hcf-1 allowed DNA replication of *Hyphantria cunea* MNPV, *Orgyia pseudotsugata* MNPV, and BmNPV in Tn-368 cells, which are normally not permissive to infection with any of these viruses [100, 101]. Finally, it appears that the phenotype of various AcMNPV mutants infecting Sf9 cells may be different than the phenotype observed for the homologous gene-disrupted BmNPV infecting BmN cells; for example, while a BmNPV virus with *tlp* (Bm68) deletion mutant was not essential for infectivity in BmN cells, disruption of the AcMNPV homologue Ac82 resulted in defective BV production and reduced or abolished very late gene transcription [98, 102, 103].

Targeted gene silencing or downregulation by RNA interference (RNAi) of AcMNPV or host cell genes has also shown promise for improving recombinant protein production in the BEVS. The canonical RNAi mechanism involves the production of RNA molecules

that inhibit gene expression by interacting with and neutralizing transcribed messenger RNA (mRNA) molecules through the RNA-induced silencing complex (RISC) pathway. Endogenously transcribed self-annealing short hairpin RNA (shRNA) molecules or exogenously introduced small interfering RNA (siRNA) or long double stranded RNA (dsRNA) molecules can lead to RISC-mediated gene silencing [104]. The endogenously transcribed shRNA, which mimics small, regulatory RNAs commonly expressed in animal and plant cells called precursor micro-RNAs (pre-miRNAs), must first be processed by the ribonuclease III (RNAse III) enzyme Drosha before being exported to the cytoplasm as a precursor shRNA (pre-shRNA). There, the RNA molecule (ie. pre-shRNA, siRNA, or dsRNA) is further processed by the RNAse III enzyme Dicer and loaded into the RISC [104, 105]. The major classical components of the RNAi pathway have been identified in Sf21 cells [105], and several studies have examined the phenotypic impact of gene silencing experiments using various RNAi approaches in several insects *in vivo* and cultured cells *in vitro* [104, 106]. Notably, targeting Caspase-1 in High Five, BmN, and Sf9 cells blocked the apoptotic pathway and resulted in improved recombinant protein production [107–111]. Similarly, silencing of the AcMNPV-encoded *v-cath*, *gp64*, *vp80*, or *orf34* genes resulted in enhanced heterologous gene expression, protein quality, or reduction in progeny virus [85, 112–114]. Finally, targeting the host cell *fdl* gene encoding β-*N*-acetylglucosaminidase in BmN cells improved complex-type *N*-linked glycosylation, while silencing the *cycE* gene encoding the cell cycle checkpoint protein Cyclin E resulted in arrest of High Five cells in G_1 phase and increased recombinant protein yield [115].

Expression of heterologous chaperone proteins

In addition to disruption of endogenous AcMNPV genes, strategies aimed at improving recombinant protein expression and stability through expression of heterologous genes either by the rBEV itself or from the host cell line have been employed [72]. Previous studies have suggested that mRNA transcript levels of endoplasmic reticulum-located (ER) molecular chaperones decrease in the early stages of infection [116], and subsequent studies have suggested that co-expression of chaperone or foldase proteins, which are crucial for folding proteins into their correct final three-dimensional conformation, may improve the quality of expressed recombinant protein [72]. Co-expression of mammalian folding enzymes such as calnexin, calreticulin, protein disulfide isomerase (PDI), immunoglobulin (IgG) heavy chain binding protein (BiP), heat shock protein 70 (Hsp70) and Hsp90 machinery have also shown promise for improving recombinant protein quality [117–123].

Improving post-translational processing

While insect cells are able to fold, modify, traffic, and assemble polypeptides to produce highly authentic, correctly processed, and soluble, recombinant proteins, the protein processing pathways are not necessarily equivalent to those of other higher eukaryotes such as mammalian cell lines [3, 10]. Notably, the insect cell *N*-glycosylation pathway does not produce complex products with terminal galactose or sialic acid residues; rather, the major insect cell end product is paucimannose *N*-glycans wherein the *N*-Acetylglucosamine residue is trimmed from the *N*-glycan, leaving bi-antennary mannose moeities [10, 124]. Extensive effort has been directed at developing insect cell lines expressing mammalian genes encod-

ing proteins with *N*-glycan processing activity [10]. Significantly, Sf9 cell lines expressing several mammalian genes for *N*-glycan processing have been derived, including five mammalian glycosyltransferases (SfSWT-1; human β1,2-*N*-acetylglucosaminyltransferase I and II, (GlcNAc-T I, GlcNAc-T II), bovine β1,4-galactosyltransferase (β4Gal-T I), rat α2,6-sialyltransferase (ST6Gal I), and mouse α2,3-sialyltransferase IV (αST3Gal IV)) [125]. SfSWT-1 cells could produce bi-antennary, terminally sialylated *N*-glycans [126]. This cell line was improved upon to encode an additional two mammalian (mouse) genes encoding sialic acid synthase (SAS) and CMP-sialic acid synthetase (CMAS) for CMP-sialic acid biosynthesis, and called SfSWT-3 [127]. Further, SfSWT-4 and SfSWT-5 cell lines express the minimal set of genes for production of terminally-sialylated *N*-glycans when cultured with the sialic acid precursor *N*-acetylmannosamine from constitutive or inducible promoters, respectively [128, 129]. These cell lines express the same genes as SfSWT-3 except GlcNAc-T I is omitted [128, 129]. SfSWT-4 was further engineered to SfSWT-6, which additionally expresses a human CMP-sialic acid transporter (hCSAT), which lead to higher levels of cell surface sialylation and also supported higher levels of recombinant glycoprotein sialylation when cultured in low concentrations of *N*-acetylmannosamine [128]. More recently, the *fdl* gene encoding β-*N*-acetylglucosaminidase was disrupted in Sf9 cells. This was an important achievement because FDL catalyzes the removal of a terminal *N*-acetylglucosamine residue from trimmed *N*-glycan-processing intermediates and thus antagonizes further elongation toward homogeneous, terminally sialylated *N*-glycans [130].

Promoter choices for recombinant protein production

While the p10 and polh promoters are able to generate an extraordinary abundance of mRNA transcripts, the burst of transcription occurs only during the very late stage of infection. This transcriptional burst primarily occurs when the host cellular machinery required for protein synthesis and post-translational processing are severly compromised, and host- and AcMNPV-produced proteases are present at high levels in the culture supernatant [11, 20]. As such, promoters active earlier in the infection cycle have been sought as alternatives to the very late promoters, leading to investigation of other endogenous AcMNPV promoters, constitutively active insect cell promoters, and engineered promoters thereof [87, 131]. The ie-1 promoter has shown promise for producing more active, correctly folded and processed secretory pathway eukaryotic proteins than the polh promoter [132]. Similarly, the gp64 promoter produces considerably less mRNA abundance than the polh promoter, but is active both early in the infection as well as late. Although the abundance of the HIV-1 GP41 was higher in cell lysates when expressed from the polh promoter, its expression from the gp64 promoter improved processing and glycosylation [133]. The basic protein (P6.9) promoter, which appears to be the most active late gene promoter in the AcMNPV genome [16], has been shown to produce superior yields of correctly assembled and processed Shaker potassium channels than the polh promoter [134]. Other promoters employed for recombinant protein expression or analyzed for transcriptional activity include viral white spot syndrome virus ie-1, cytomegalovirus immediate-early (CMV) promoter, v-cath, 39k, vp39, OpIE1, OpIE2, as well as tandem promoters such as vp39-polh (pcap-polh) and tandem ie-1 promoters. Hybrid promoters that combine transcription-enhancing

hr sequences and viral promoter sequences, including *hr3*-vp39, *hr5*-ie-1, *hr5*-OpIE2, and *hr5*-ie-1-p10 have also been developed [87, 135, 136]. Further, constitutive insect cell-derived promoters employed in Sf9 cells include hsp70, actin, elongation factor 1 (EF1), pB2, and enolase [135, 137]. Finally, fully synthetic [138] or endogenous AcMNPV promoters have been engineered for stronger or weaker transcription levels compared to their wild-type parent, including truncated ie-1 and p10 promoters [139, 140], and a stronger polh promoter [141, 142]. Despite the growing evidence that promoters active earlier in the infection cycle may lead to superior yields than the polh promoter despite not being as active, the search for promoters having higher activity than polh in other baculoviruses persists [143].

The role of the 3'UTR on gene expression has also been evaluated. Although it is clear that baculovirus pre-mRNA molecules are processed to generate 5'-capped and 3'-polyadenylated messengers, functional characterization of the process has remained elusive [144]. While evidence has suggested that the cannonical eukaryotic 'AAUAAA' polyadenylation signal is conserved in the 3'UTR of many baculovirus genes [16, 145], at least one study has suggested that polyadenylation of late mRNAs was different than early viral genes (and that of the host cell mRNAs). These researchers found uracil-rich regions within the ORF of late baculoviral genes that were transcribed and recognized as a terminantion signal analagous to the rho-dependent transcription termination of bacteria [146]. However, bioinformatic analysis performed in recent years appears to support processing of baculoviral pre-mRNAs by the host polyadenylosome complex [16, 144, 145]. Significantly, the 3'UTR of simian virus 40 (SV40) is routinely inserted downstream of the polh promoter in baculovirus transfer plasmids, despite evidence that including the 3'UTR of either the

polh or *p10* genes contributed to higher levels of gene expression than the SV40 3'UTR [147–150]. Further, addition of a synthetic AT-rich 21-bp sequence (Syn21) in the 5'UTR and replacing the SV40 3'UTR with the p10 3'UTR increased GFP yield and porcine circovirus type 2 (PCV2) VLPs 4.8-fold and 4.1-fold compared to the standard rBEV cassette consisting of the polh promoter and SV40 3'UTR [150].

The effect of 'competition', either between promoter elements for transcription, or mRNA molecules for translational machinery, has been reported previously in the literature [81, 151, 152]. Previous evidence has suggested that expression of proteins from the p10 promoter had a negative impact on the level of transcription [151] and translation [81] of genes expressed from the polh promoter in the same construct, however no differences in expression were apparent for genes transcribed from the p10 promoter either in the presence or absense of an active polh promoter. Similarly, observations of competition were also suggested for other polh-AcMNPV promoter pairs, including v-cath, gp64, basic, and ie-1 [152]. Contrarily, no competition was observed when 2, 3, and 5 subunit protein complexes were transcribed from polh and p10 promoters, perhaps indicating that the phenomenon of 'competition' may be more complex than just promoter choice [73].

1.3.3 Bioprocessing strategies for production of recombinant proteins and therapeutics

While animal cells have been cultured for the production of recombinant proteins for the majority of the last 100 years [2], development of large-scale, industrial bioprocesses lagged behind until the early 2000s [1]. Several strategies aimed at improving yields of recombi-

nant proteins and VLPs produced using BEVS have been explored, both at the lab-bench scale and industrial scale [153, 154]. Developments have included optimized medium compositions for improved growth and infectivity, and genetic engineering efforts to improve production and isolation of rBEVs as well as quantity, quality, and stability of produced recombinant proteins as outlined in the preceding sections [37, 72, 87]. Additionally, infection parameters such as multiplicity of infection (MOI) and time of infection (TOI), cell density at infection (CDI), monocistronic vs polycistronic rBEVs, and the effect of passage number of the virus and cells have been investigated [153]. Finally, process parameters surrounding bioreactor cultivation and downstream purification have also seen considerable investigation [153].

The effect of MOI on recombinant protein and VLP production has been investigated for both total MOI as well as MOI ratios for multiple rBEVs when multiple proteins are expressed from individual, co-infected, rBEVs [87, 153]. Notably, cells infected at high MOIs should be in the mid- to late-exponential phase, whereas lower CDI should be used to ensure peak cell density is reached when all cells are infected when low MOIs are used [155]. Several studies have also examined the effect of MOI on the formation of VLPs and multi-protein complexes. Co-infecting four rBEVs each encoding one rotavirus structural protein at an MOI of 10, 90-95% of particles were observed in their correct, triple-layered configuration, indicating efficient expression and self-assembly of rotavirus-like particles at high MOI [156]. Similarly, the amount of VP2 and VP6 expressed individually at an MOI of 5 matched the expression of these proteins when co-infected with a total MOI of 10, indicating that transcriptional/translational burden was not evident at an overall MOI of 10 [157]. However, it was observed that VP6 was produced in excess compared

to stoichiometric requirements, implying a waste of metabolic resources when equal MOIs were used [157]. This was corrected by infecting with MOI ratios of the rBEV ecoding VP6 to the rBEV encoding VP2 of between 0.2 and 0.6, resulting in an increase in the VP2 fraction of total recombinant protein to an approximately equal ratio with VP6 [157]. This approach was also employed for production of AAV vectors using the BEVS; rBEVs expressing the AAV replication and structural proteins (BacRep and BacCap, respectively) could be infected at equal MOIs, whereas a third rBEV encoding the AAV vector genome (BacITRGFP) could be used at a much lower MOI, resulting in only a slight decrease in infectious viral particle (IVP) titer compared to the high MOI condition [158]. Interestingly, at 96 hpi, only a slight decrease in IVP titer was observed for both the equal ratio (MOI = 5:5:5) BacITRGFP/BacRep/BacCap condition and the MOI = 1:9:9 condition as compared to the high MOI (9:9:9) condition, indicating that much lower amounts of viral stock could be used to produce approximately the same IVP titer [158].

Similarly, TOI has also been investigated as a parameter that may affect recombinant protein production when co-infection systems are employed [87, 153]. Particularly in cases where the native interactions between expressed foreign proteins is temporal in nature, such as VLP formation or proteins that require complicated post-translational processing, TOI represents additional degrees of freedom that can be optimized to improve production [87]. For production of rotavirus-like particles, TOI was effectively used to manipulate the stoichiometric ratio of VP2:VP6 protein, with a ~52% increase in VP2 production and concomitant ~95% decrease in VP6 when infection with the VP6-encoding rBEV was delayed 6 hours [157]. Staggering the TOI of viruses has also been examined in other studies, wherein the MOI of some of the rBEVs was below 1, while the MOI of others

was above 1, creating a delay in the delivery of the low-MOI rBEVs [87]. Consequently, only a small subset of the population will be infected initially with these viruses, with a later, secondary infection of viral progeny [87]. However, this strategy has not necessarily resulted in improved production, as using MOIs less than 1 resulted in lower IVP titer of AAV viral vectors [159], and delaying infection of certain rBEVs negatively affected production of bioactive AAV viral vectors even at high MOI [158].

While co-infection strategies may provide intrinsic advantages such as manipulating MOI and TOI of individual rBEVs in order to control timing, order and stoichiometric ratios of recombinant protein expression, these parameters may not be suitable for industrial-scale production [83, 87, 153]. Infecting with high MOI has been associated with shorter productive periods and lower volumetric yields, and additionally require large viral stocks for infection [83, 160]. In contrast, infecting with low MOI permits multiple cycles of infection which requires smaller viral stocks, but may lead to accumulation of DI particles [83, 160, 161]. Additionally, infecting at low MOI increases the duration of the bioreactor cultivation, which may compromise the integrity of the recombinant proteins due to the presence of proteases and cellular debris in the culture supernatant [160]. Nevertheless, the use of both co-infection strategies using multiple monocistronic rBEVs and co-expression strategies using a single, polycistronic rBEV have been employed successfully for production of simple recombinant proteins and complex, multiprotein complexes and VLPs [87]. Finally, as noted, mutant DI viruses tend to accumulate in viral stocks after multiple rounds of passaging, and cell productivity is affected by the physiological state of the cells [153]. Specifically, virus passage number, culture age, culture state and energy metabolism, as well as the depletion of nutrients in the medium all contribute to

the productivity of infected cells [155, 162–165].

In addition to optimization of various infection parameters for improving production of recombinant proteins and therapeutics using the BEVS, considerable attention has been given to process parameters surrounding bioreactor cultivation and downstream processing [153, 164]. Despite a multitude of bioreactors and cultivation modes having been investigated to varying degrees for the BEVS, the majority of cultivations are carried out in either batch or fed-batch mode using stirred-tank reactors (STR) [153, 154, 164]. Along with investigation of the effects of temperature, pH, osmolality, shear stress, dissolved oxygen, and carbon dioxide on insect cell cultivation, medium optimization, development of feeding strategies, and investigation of different bioreactor systems for cultivation and infection have been systematically examined and reviewed in detail elsewhere (for example, see [153, 154, 164]). Notably, insect cells are typically cultivated optimally at 27 °C, pH 6.2-6.4, 320-375 mOsm/kg osmolality, and no dissolved CO_2 [153]. Although insect cells tend to be fairly resilient to changes in pH, they are particularly sensitive to shear stress, osmotic shock, and lactate accumulation [153]. Glucose and the amino acid glutamine are the main sources of carbon in insect cell cultures, and Sf9 cells have rather efficient glucose metabolism, with ~80% of glucose being channeled into the tricarboxylic acid (TCA) pathway to produce pyruvate [166]. Unlike many mammalian cell lines, Sf9 cells do not produce significant amounts of lactate; rather, under conditions of glucose excess, the main waste metabolite is the amino acid alanine, whereas ammonia is produced under glucose limitation [164].

1.4 BEVS as a production platform for complex biologics and therapeutic applications

The BEVS platform has been used for an incredibly diverse set of applications, ranging from hosts for basic recombinant protein production and structural biology applications in research environments to commercial applications such as manufacturing of vaccines and virus-like particles (Figure 1.2). Owing to their ability to transduce mammalian cells (known as BacMam technology), the baculovirus itself has become important in drug discovery applications, and as a gene delivery vector for use in cell and gene therapies as well as an anti-cancer therapeutic. Subsequently, a number of BEVS-produced therapeutics have been approved for veterinary or human use by regulatory agencies in multiple countries (Table 1.1). The following section will describe a few of the applications that BEVS appears to be well suited as a platform technology.

1.4.1 Vaccines

The BEVS is widely used for recombinant protein production for subunit vaccines, but also for the production of VLPs to study virus assembly and as antigens for immunization [10]. Employing rBEVs for expression of heterologous viral proteins is particularly valuable for cases where cell culture-based viral replication systems are not available, such as human papillomavirus (HPV) and hepatitis C virus (HCV), and for production of vaccines targeted to serotypes of viruses that change frequently, such as Influenza virus [10]. As such, several enveloped and non-enveloped VLPs, recombinant proteins for subunit vaccines, and rBEVs displaying or expressing viral antigens have been studied for their utility as vaccine

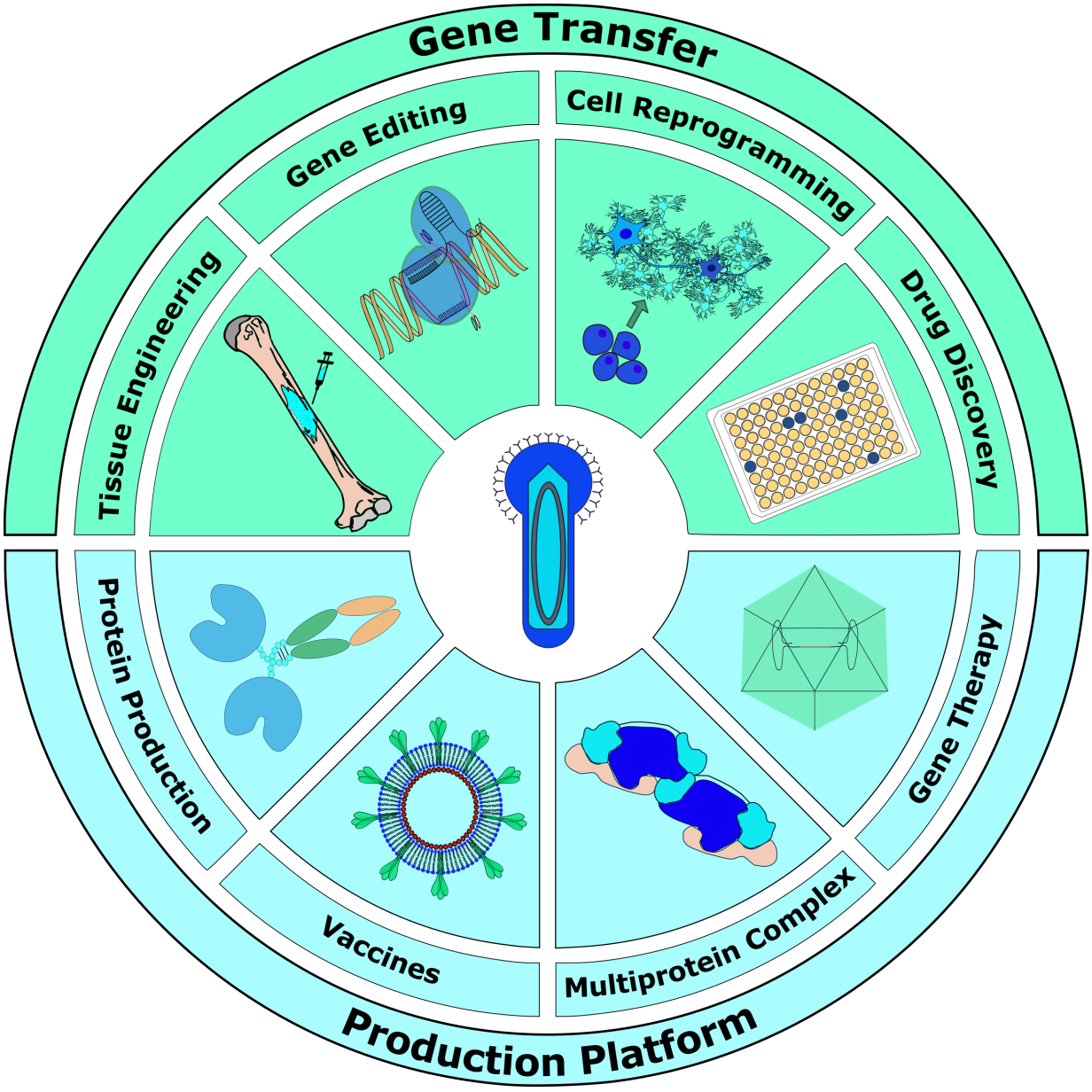


Figure 1.2: **BEVS as a manufacturing platform.** The BEVS has found utility as a gene transfer vector to support applications such as tissue engineering, gene editing, cell reprogramming, and drug discovery, and as a platform for production of recombinant proteins and complex biologics to support research and industrial biopharmaceuticals.

Table 1.1: Products approved for veterinary or human use

Product Name	Type	Product	Target Species	Indication	Manufacturer	Approved
Porcilis Pesti	Subunit Vaccine	E2	Pigs	Classical swine fever virus	MSD Animal Health	2000 (EMA)
Bayovac CSF E2	Subunit Vaccine	E2	Pigs	Classical swine fever virus	Bayer AG/Pfizer Animal Health	2001 (EMA)[a]
Circumvent PCV	VLP Vaccine	ORF2	Pigs	Porcine circovirus type 2	Merck Animal Health	2007 (FDA) 2009 (EMA)
Ingelvac CircoFLEX	VLP Vaccine	ORF2	Pigs	Porcine circovirus type 2	Boehringer Ingelheim Vetmedica	2006 (FDA) 2008 (EMA)
Porcilis PCV	VLP Vaccine	ORF2	Pigs	Porcine circovirus type 2	MSD Animal Health	2009 (EMA)
Flublok Quadrivalent	Subunit Vaccine	(Seasonal) HA	Humans	Influenza A & B	Sanofi Pasteur	2013 (FDA)
Cervarix	VLP Vaccine	HPV-16/18 L1	Humans	Human papillomavirus	GlaxoSmithKline	2009 (FDA) 2007 (EMA)
Provenge	Immunotherapy	PAP-GM-CSF	Humans	Prostate Cancer	Dendreon	2010 (FDA) 2013 (EMA)[a]
Glybera	Gene Therapy	rAAV	Humans	Lipoprotein lipase deficiency	uniQure	2012 (EMA)[a]

[a] : Removed from market

candidates. Several of these studies are summarized in Table 1.2. Significantly, several BEVS-produced vaccines are currently candidate vaccines in various phases of clinical trials, or currently commercialized and available worldwide, such as the HPV VLP Cervarix (GlaxoSmithKline) and the Influenza A vaccine Flublok (Protein Sciences) [6, 9, 167, 168]. Additionally, several companies are in various stages of clinical trials evaluating the efficacy of BEVS-produced vaccines for the current COVID-19 pandemic [169, 170]. Notably, after returning favourable results from combined Phase 1/2 clinical trials, the SARS-CoV-2 subunit vaccine NVX-CoV2373 (Novavax) is currently in Phase 3 clinical trials in the United States/Mexico and the United Kingdom [171, 172]. The BEVS-produced subunit vaccine, which consists of the full length, prefusion conformation S protein co-administered with Novavax' Matrix-M™ adjuvant, is only the fifth among more than 200 companies developing SARS-CoV-2 vaccines to reach Phase 3 clinical trials in the United States [169], and very recently reported a vaccine efficacy of ~89% at the primary endpoint of the trial in the UK. Regulatory agencies in multiple countries have started rolling reviews of NVX-CoV2373 for emergency use authorization, including the European Medicines Agency (EMA), US Food and Drug Administration (FDA), UK Medicines and Healthcare products Regulatory Agency (MHRA), and Health Canada.

As noted, production of Influenza vaccines produced using the BEVS is particularly attractive due to the seasonally-changing nature of prevalent Influenza serotypes [225]. As a result, new vaccines incorporating the relevant HA (and/or NA) serotypes must be produced yearly [225]. For enveloped viruses, the primary target for protective immunity is the envelope glycoproteins [167]. Proper protein glycosylation can affect protein folding, localization, solubility, and immunogenicity, and for many viruses, improper glycosylation can

Table 1.2: Selected vaccine antigens produced using IC-BEVS

	Virus	Expressed Proteins	Ref.
Enveloped VLPs	Ebola Virus	VP40, GP	[173, 174]
	Hepatitis B Virus (HBV)	Core, Surface antigen	[175]
	Human Immunodeficiency Virus (HIV)	Gag, Env, Pro	[176–180]
	Severe Acute Repiratory Syndrome Coronavirus (SARS-CoV)	S, E, M	[181, 182]
	Influenza Virus	HA, NA, M1, M2	[183–188]
	Japanese Encephalitis Virus	prM, E	[189, 190]
Non-Enveloped VLPs	Foot-and-Mouth Disease Virus	P12A, 3C	[191]
	Herpes Simplex Virus (HSV)	VP23, VP5, VP21/VP24, VP22a, VP26, VP19	[192, 193]
	Human Papillomavirus (HPV)	L1, L2	[194, 195]
	Polio Virus	VP0, VP1, VP3	[196, 197]
	Rotavirus	VP2, VP4, VP6, VP7	[156, 157, 198–200]
	Simian Virus 40 (SV40)	VP1, VP2, VP3	[201]
Subunit Vaccines	Middle East Respiratory Syndrome CoV (MERS-CoV)	S RBD (dimer)	[202–204]
	SARS-CoV	S	[205]
	SARS-CoV-2	S	[171]
	Japanese Encephalitis Virus	E	[206]
	Rabies Virus	G	[207, 208]
	Influenza	HA	[209–213]
Surface Display / Gene Transfer	Influenza	H5N1-HA VSV-G (display) H5N1-HA H5N2-HA (display/express) 3xM2e/LTB (display/express)	[214] [215] [216] [217]
	Zika Virus	HERV Env (display), prM, E	[218]
	Respiratory Syncytial Virus	M2 F (display/express), VISA (express)	[219] [220]
	Plasmodium falciparum (Malaria)	CS (display/express)	[221]
	Enterovirus 71	VP1 fused to SP, TMD/CTD of GP64	[222–224]

be highly detrimental to the antigenicity and immunogenicity of glycoproteins [226, 227]. As insect cells may not necessarily have the ability to produce glycoproteins with structurally authentic mammalian *N*- or *O*-glycans, concerns over antigenicity/immunogenicity of vaccines produced using the BEVS is potentially a serious issue [167]. For Influenza virus, however, the absense of structurally nonessential glycans on the envelope glycoproteins may actually be an effective strategy for vaccine design; truncation of the *N*-glycan structures on HA may increase sialic acid binding affinity [226] and facilitate the uptake of Influenza vaccine candidates by antigen-presenting cells for the adaptive immune response [228]. Influenza VLPs comprised of HA, NA, M1, and M2 (and several combinations thereof) and recombinant subunit vaccines have been generated using the BEVS, and have been demonstrated to induce protective immunity during preclinical and clinical studies [9, 167, 168]. Vaccines for pandemic Influenza strains have been of particular interest, and recently the interim results from a Phase 1 clinical trial investigating the immunogenicity of a chimeric HA-based universal influenza virus vaccine candidate produced using BEVS was reported (Table 1.3).

In addition to its utility as a production platform for the manufacture of VLPs and subunit vaccines, the baculovirus itself has been studied extensively for its suitability as an adjuvant or antigen itself [237]. Owing to developments enabling the display of heterologous proteins on the surface of baculovirus particles [10, 238], several studies have shown that baculoviruses displaying antigenic proteins from pathogens such as Japanese encephalitis virus, influenza, and malaria can generate high titers of protective antibodies. Further, rBEVs equipped with a mammalian expression cassette can be combined with surface display to enhance the vaccine efficacy; vectors that both displayed and expressed

Table 1.3: Influenza vaccines in clinical development

Drug Name	Expressed Proteins	Product Type	Company/Sponsor	Approval Stage	US Clinical Trial Registry	Ref.
Panblok® H5	A/Indonesia/05/2005 H5N1-HA	Subunit	PSC[a]	Phase 1/2 (completed)	NCT01147068	[229]
Panblok® H7	A/Anhui/01/2013 H7N9-HA	Subunit	PSC[a]	Phase 1/2 (completed)	NCT02464163	[212]
N/A	A/California/04/2009 H1N1-HA	Subunit	PSC[a]	Phase 1/2 (completed)	N/A (Australia)	[230]
Panblok® H7	A/Anhui/01/2013 H7N9-HA	Subunit	BARDA[b]	Phase 2 (completed)	NCT03283319	N/A
N/A	A/California/04/2009 H1N1-HA, NA, M1	VLP	Novavax	Phase 2 (completed)	NCT0107299	[231]
NanoFlu™	Quadrivalent (Seasonal) HA stalk, head	Subunit	Novavax	Phase 3 (ongoing)	NCT04120194	[232]
N/A	A/Anhui/01/2013 H7H9-HA, NA, M1	VLP	Novavax	Phase 1/2 (completed)	NCT02078674	[233]
N/A	A/Indonesia/05/2005 H5N1-HA, NA, M1	VLP	Novavax	Phase 1/2 (completed)	NCT01594320	[234]
N/A	Chimeric HA-based universal	Subunit	PATH[c]	Phase 1 (completed)	NCT03300050	[235, 236]

[a]: Protein Sciences Corporation (now part of Sanofi Pasteur); [b]: Biomedical Advanced Research and Development Authority (U.S. Department of Health)
[c]: Program for Appropriate Technology in Health (nonprofit global health organization)

the *Plasmodium falciparum* circumsporozoite (CS) protein potentiated higher antibody titers and a more robust interferon-γ-producing T-cell response than either expressing or displaying-only vectors [221]. Similarly, baculovirus has shown promise as an effective adjuvant; immunization of mice with norovirus (NoV) VLPs or ovalbumin (OVA) formulated with live or inactivated baculovirus improved the immunogenicity of the NoV VLPs and allowed for a 10-fold lower dose of VLPs, and induced stronger IgG2a antibody and T cell responses to OVA protein than those responses to NoV VLP or OVA immunization alone [239]. More thorough reviews on this subject have been published recently [237, 240–242].

1.4.2 Cell and gene therapies

The BEVS has been used for production of viral vectors such as recombinant Adeno-Associated Virus (rAAV) for gene therapy applications [139]. Glybera, which was the first rAAV-based gene therapy approved by the FDA, was manufactured using BEVS [243],

and at least four BEVS-produced gene therapies for Hemophilia A and B are in various stages of clinical trials. BioMarin Pharmaceutical's Valoctocogene Roxaparvovec gene therapy for hemophilia A has returned favourable Phase III results and is expected to receive FDA approval for human use in the near future [244, 245]. Although the majority of rAAV production uses HEK293 cells as production host, recent improvements in the BEVS compare favourably with HEK293-manufactured rAAV, re-affirming the potential of the BEVS platform in modern medicine [246–248].

Aside from rAAV vectors, BEVS has been used to manufacture lentiviral vectors [249], re-program mouse embryo fibroblasts into neurons and induced pluripotent stem cells (iPSCs) [250, 251], and delivered zinc-finger nucleases (ZFN) to efficiently and site-specifically integrate the Yamanaka factors (*oct3/4*, *klf4*, *sox2*, and *c-myc*) at the AAVS1 locus to reprogram human fibroblasts to iPSCs [252]. Human adipose-derived stem cells, mesenchymal stem cells (MSCs), and rabbit bone marrow-derived MSCs have been transduced with baculovirus expressing the growth factors bone morphogenetic protein 2 (BMP-2), vascular endothelial growth factor (VEGF), or the human microRNA (miRNA) miR-148b to accelerate the remodeling and regeneration of bone [253–255]. Importantly, studies have suggested that rBEV-transduced MSCs feature a safety profile that appears to be appropriate to enable progression toward clinical testing [256, 257]. Moreover, baculovirus-delivered *Sleeping Beauty* transposon sustained transgene expression for at least 2 months in mouse eyes [258], potentially offering evidence of BEVS utility for manufacturing cell therapies such as chimeric antigen receptor T-cell (CAR-T) therapy [259].

1.4.3 Cancer therapies and drug discovery

Provenge (Dendreon) is an autologous cellular immunotherapy approved by the FDA in 2010 as a cellular vaccine for metastatic prostate cancer [260]. The therapy involves stimulation of CD52-enriched cells (which include antigen-presenting cells such as monocytes and dendritic cells (DC), T- and B-cells, and natural killer (NK) cells) *ex vivo* with BEVS-produced prostatic acid phosphatase-granulocyte-macrophage colony-stimulating factor (PAP-GM-CSF) [260]. In addition to Provenge, BEVS has been used for the manufacture of many as yet unapproved therapeutic biomolecules, including immunotherapies, cytokines, growth factors, and enzymes. Additionally, the rBEV itself has been used as the therapeutic modality itself: dendritic cells transduced with wildtype baculovirus increased surface expression of co-stimulatory molecules CD80, CD86, and major histocompatability complex (MHC) classes I and II, which are responsible for delivering secondary signals necessary for T-cell activation (CD80 and CD86), and processing and presenting antigens on the surface of the cell (MHC I and II), respectively. Additionally, these DCs secreted interferons and other proinflammatory cytokines to reduce the size of lung tumors caused by Lewis lung carcinoma in mouse model [261]. Further evaluation suggested that baculovirus-infected DCs induced a non-specific immune response and could be employed as an immunotherapeutic agent for malignancies in conjunction with other therapeutic drug regimens [262]. Similarly, intratumoral injection of a rBEV with a human endostatin-angiostatin (hEA) fusion protein expression cassette flanked by AAV inverted terminal repeats potently inhibited tumor growth and significantly improved survival in prostate cancer mouse models [263], and baculovirus delivery of the hEA expression cassette and

Sleeping Beauty transposon resulted in transgene integration, prolonged expression, and subsequent suppression of ovarian tumor xenografts in mice [264]. This system was further enhanced by displaying the decay accelerating factor (DAF; CD55) on the surface of the rBEV to inhibit the activation of the complement response and lead to higher transduction efficiency, ultimately delayed tumor growth and prolonged survival compared to the control [265]. Finally, an engineered rBEV consisting of a trigger-inducible RNA riboswitch sensitive to the expression signatures of miRNAs to enable selective expression of the pro-apoptotic human BCL2-associated X protein (hBax) transgene was developed. This system responded to miRNAs highly expressed in hepatocellular carcinoma (HCC; miRNA-196a) or normal (miR-126) cells to selectively kill HCC cells but not normal cells in *in vitro* mixed cell culture experiments [266].

The BEVS has also become an important platform technology in drug discovery applications. Targeting membrane proteins with agonist or antagonist molecules constitutes a large proportion of the drug portfiolo of pharmaceutical companies because they are biologically relevant and relatively simple to target. Further, these targets offer a large degree of chemical tractability; large libraries of chemical compounds can be screened to identify ones that modulate or interact with the target to elicit a desired biological effect [267]. Consequently, many cell-based assay formats have been developed to enable drug discovery [268], however the high protein expression requirements are often precluded by the toxicity of high target protein levels on the host cells thus making stable cell line generation labour intensive or ineffective [267]. Although transient transfection approaches can be effective alternative approaches, they are often costly and not particularly suited to high-volume and high-throughput screening methods that are often required; one such HTS campaign

initiated with the BacMam platform to identify agonists for Toll-like receptor 7 (TLR7) consisted of 150 384-well microtiter plates per day for 22 days (potentially ~1.3 million unique compounds screened) [267]. Additionally, due to the large cargo capacity of BEVS, large, multiprotein complexes including accessory proteins can be efficiently produced using BEVS [73, 267, 269]. In addition to TLR7, agonist and antagonists have been identified using BacMam technology for fatty acid receptors GPR40 and GPR120 [270], Histone H3 receptors [271], chemokine receptor CCR5 [272], various ion channels [273, 274], and G-protein-coupled receptors (GPCRs) such as neurokinin 3, prostanoid EP3, GPR103, and Urotensin II [267, 275].

1.5 CRISPR-Cas9 as a tool for sophisticated genetic engineering

1.5.1 Discovery and characterization of CRISPR-Cas

The first report of Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) short-sequence DNA repeats (SSRs) appeared in the literature in 1987, a serendipitous result of a comprehensive study of the *iap* gene involved in isozyme conversion of alkaline phosphatase in *E. coli* [276, 277]. These ~30bp, partially palindromic SSRs differed from most other repetitive elements present in sequenced genomes in that they were interspaced by non-repetitive sequences of approximately the same length [278, 279]. Significantly, CRISPR SSRs were found in >40% of sequenced bacterial genomes and >90% of archaea [278]. CRISPRs were originally hypothesized to be involved in regulation of gene expression

or DNA repair [278, 280]. However, the observations that the spacer sequences matched sequences in viral genomes or extra-chromosomal plasmids, and signature genes were identified as being well-conserved, transcriptionally-active, and typically clustered adjacent to the SSR elements, altered these theories. The acronyms CRISPR and cas were coined to describe these genetic elements, and a new hypothesis that CRISPR-Cas systems might be an adaptive defense mechanism against bacteriophage challenge was proposed [281–283]. This function was demonstrated experimentally in 2007; bacteriophage challenge of a wild-type strain of the bacterium *Streptococcus thermophilus* possessing CRISPR loci resulted in phage-resistant mutants that had acquired new spacers bearing complementary sequence to the invading phage DNA [284].

Initial analysis of the clusters of signature CRISPR associated *cas* genes identified three distinct types of CRISPR systems (Types I-III), each containing multiple subtypes, that are highly diverse in genetic structure and mechanism [285]. To more accurately describe the immense variability in genomic architecture and diversity of protein sequences found at CRISPR loci, this classification system has been recently expanded to encompass two distinct classes, five types, and 16 subtypes [286]. CRISPR-Cas Types I, II, and III are the most prevalent systems in both bacteria and archaea, and are differentiated by the presence of *cas3*, *cas9*, and *cas10* signature genes, respectively [285, 286]. Type I systems, comprised of six distinct subtypes (I-A to I-F) in addition to subtype I-U which is comprised of members yet to be characterized mechanistically, exhibit the greatest diversity (i.e. number of subtypes), and subtype 1-B is the most abundant CRISPR-Cas system found in nature [286, 287].

CRISPR-based immunity encompasses three distinct processes, termed adaptation, ex-

pression, and interference [286, 288]. Adaptation involves the acquisition of specific nucleotide sequence tags that act as the immune "memory" [289]. During periods of predation, nucleotide sequence tags, referred to as protospacers in their native context within the nucleotide sequence of invading genetic elements such as bacteriophage and plasmids [281–283, 289], are rapidly acquired and incorporated into the host genome by host Cas proteins into vast CRISPR arrays arranged of CRISPR SSRs flanking the foreign sequence, where they are subsequently termed spacers [289]. Up to 587 unique spacer sequences have been found in a single CRISPR array, exemplifying the exceptional level of attack experienced by microorganisms in nature [290, 291]. The Cas proteins involved in acquisition of new spacers are the only proteins conserved among all CRISPR-Cas subtypes [286]. During the expression phase, acquired spacer sequences are transcribed and, in conjunction with Cas proteins, provide resistance against invading genetic elements [286, 289]. CRISPR arrays are first transcribed into a single precursor CRISPR RNA (pre-crRNA), which is cleaved into repeat-spacer-repeat units by Cas6 (Type I and III systems) (116) or the ubiquitous housekeeping ribonuclease RNase III enzyme and a small trans-activating crRNA (tracrRNA) (Type II systems) [292] to yield mature crRNAs. The tracrRNA contains a 25 nucleotide sequence that is complementary to the repeat region of the pre-crRNA, and base-pairing between these two RNAs results in a double-stranded region that is recognized and processed by RNase III [289, 292]. The Cas9 protein is also required for crRNA processing in the Type II system [289, 292]. The processed crRNAs recruit and form stable ribonucleoprotein (crRNP) complexes with specific Cas proteins for target surveillance and nucleolytic attack of invader DNA during the interference phase of CRISPR immunity [289]. The crRNAs complex with the multiprotein 'Cascade' (Cas complex for antiviral

defence) in Type I systems and the Csm or Cmr complexes in Type III systems [289]. Although the proteins associated with Type I crRNPs are phylogenetically distinct from those of Type III, recent structural studies have shown striking architectural similarities between the two Types [293–295]. The Type II system, conversely, is much more simple and compact than Type I and III machinery; tracrRNA, crRNA, RNase III, and the Cas9 endonuclease are the sole determinants for target surveillance and interference [289, 296]. The interference phase is characterized by target surveillance by the crRNP complexes and nucleolytic attack of invading DNA elements. This sequence appears to proceed in a step-wise manner; first, scanning invader DNA to find a protospacer sequence that is complementary to the crRNA and discrimination of self from non-self, followed by base pairing between the seed region of the spacer and the complementary protospacer for strand displacement, and finally recruitment of a *trans*-acting nuclease or activation of intrinsic nuclease activity for interference [289]. Type I and II systems discriminate self from non-self by scanning DNA for a short nucleotide signature called a protospacer adjacent motif (PAM) [289]. The PAM is an antigenic signature found adjacent to the protospacer, and is not represented in the sequence of the spacer integrated into the host's CRISPR array during the adaptation phase [277, 289, 297]. In many organisms, the sequence of the PAM element is highly promiscuous, affording flexibility in recognition of invading protospacers [297]. The mechanism that Type III systems employ is as yet undefined, however evidence suggests discrimination follows a PAM-independent mechanism [289]. Nevertheless, base pairing between the crRNA and protospacer results in a conformational change of the crRNP complex, triggering the recruitment of Cas3 (Type I), Csm6 or Cmr4 (Type III) nucleases, or activating the intrinsic nucleolytic activity of Cas9 (Type II) for target

degradation [289].

1.5.2 Development of the CRISPR-Cas system of *S. pyogenes* for genetic engineering

Perhaps owing at least in part to the simplicity of the Type II system, CRISPR-Cas of the human pathogen *S. pyogenes* has received immense interest in the scientific research community. As such, phenomenal developments have led to improved understanding of fundamental CRISPR-Cas biology, as well as enabling CRISPR-Cas9 as a sophisticated tool for genetic engineering applications [277, 298–300]. The first demonstration that Cas9 could be programmed to produce double-stranded DNA breaks (DSBs) at precise locations *in vitro* was reported in 2012 [296], marking the beginning of the 'CRISPR revolution' [277, 301]. Initial proof-of-principle studies validating the efficacy of CRISPR-Cas9 *in vivo* followed in quick succession, including single and multiplexed targeting in bacteria [302], transformed human cancer and pluripotent stem cell lines in culture [303–305], yeast [306], the model organism zebrafish [307], and the model plants tobacco and thale cress [308]. Since then, hundreds of papers have been published in the literature describing the application of CRISPR-Cas9 for genetic engineering in a multitude of organisms *in vivo* and cell lines cultured *in vitro*, as well as further development of CRISPR-Cas9 technology through engineering its molecular components or developing Cas9 fusion proteins to enhance the effectiveness of the tool [298, 300, 301].

Although effective genome editing tools have been available for many bacteria for several years [309], the ability to introduce DSBs to induce small insertion and deletion (indel)

mutations via the non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathways in higher eukaryotes at multiple sites in parallel represented a major advantage over existing editing tools available such as zinc finger nucleases (ZFNs) or transcription activator-like effector nucleases (TALENs) [76]. Unlike ZFNs and TALENs, which recognize customizable DNA sequences through tedious construction of short protein DNA-binding domains concatenated together with a cleavage domain [76], Cas9 endonuclease is an RNA-guided nuclease (RGN) which uses simple DNA-RNA base-pairing between the crRNA and target DNA [300]. Targeting of Cas9 was further simplified by reducing the dual-RNA requirement of tracrRNA (required for target DNA binding) and crRNA (required for target DNA recognition) to a single, engineered, chimeric, guide RNA molecule (sgRNA) that retained the ~20-nucleotide sequence at the 5' end that determines DNA target site via Watson-Crick base-pairing, and the double-stranded structure at the 3' end that binds Cas9 [296]. This created a much simpler two-component system in which the sgRNA could be engineered to target Cas9 to any DNA target site through simple REN or PCR-based methods, greatly alleviating the complexity of re-targeting. Additionally, this has inspired the generation of large sgRNA libraries for multiplexing and facilitating genetic screens through generation of libraries of knockout mutants that are fundamental in elucidating the relationship between phenotype and underlying genotype [300, 303].

Bioinformatics-based analysis of *cas* genes revealed protein domains bearing structural homology to HNH and RuvC-type endonucleases [287, 310]. Cas9, containing domains homologous to both HNH and RuvC, was found to cleave one strand of dsDNA with each of the two domains; targeted mutation experiments confirmed that the HNH domain cleaves the complementary DNA strand, whereas the Cas9 RuvC-like domain cleaves

the non-complementary strand [296]. This led to mutants of Cas9 with D10A or H840A substitutions which rendered the HNH (H840A) or RuvC (D10A) domains catalytically inactive and Cas9 capable of cutting one strand rather than both strands of target DNA [296]. These Cas9 'nickases' (Cas9n) were employed alongside pairs of offset sgRNAs to target DNA for genome editing [300, 303, 304, 311]. Because DNA single-stranded breaks are repaired very efficiently via the endogenous base excision repair (BER) pathway [312], genome editing with Cas9n and pairs of offset sgRNAs to produce ssDNA nicks that mimic dsDNA breaks have shown promise in improving fidelity and reducing off-target or non-specific effects [300, 311, 313].

In addition to Cas9n mutants, a Cas9 variant with both HNH and RuvC domains catalytically inactivated by H840A and D10A substitutions, respectively, gave rise to an effective RNA-guided, DNA-binding protein (dCas9) [277, 314]. Employing this variant along with sgRNA was shown to be effective at interfering with transcriptional elongation, or as a steric block to binding of RNA polymerase (RNAP) or transcription factors (TF) in a system called CRISPR interference (CRISPRi) [314]. CRISPRi has been successfully implemented in a variety of prokaryotes with efficiencies greater than 90%, however, in eukaryotic systems repression efficiency appears to be generally not as robust [314–316]. However, tethering dCas9 to protein domains that recruit chromatin-modifying complexes or are involved in altering DNA methylation, such as the Krüppel-associated box (KRAB) of Kox1, WRPW domain of Hes1, mSin3 interaction domain (SID), histone deacetylases (HDAC), histone methyltransferases (HMT), or DNA methyltransferases (DNMT) have shown substantial promise as effective repressors, with repression efficiencies of greater than 75-80% often reported [315–319]. Correspondingly, tethering dCas9 to viral or eukaryotic

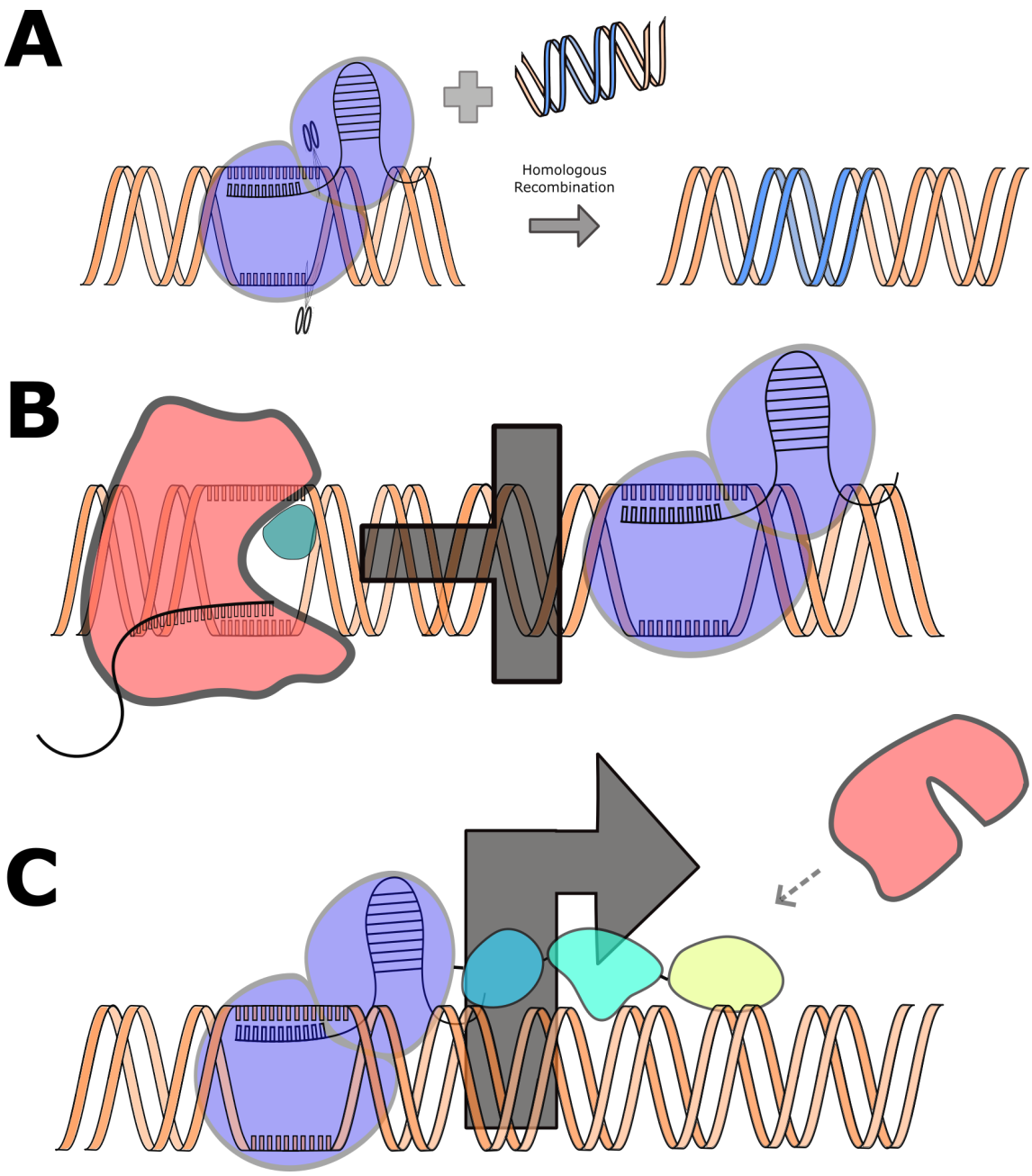


Figure 1.3: **CRISPR-Cas9 tools for genome editing and transcriptional control. A.** Genome editing with donor template, **B.** Transcriptional repression, and **C.** Transcriptional activation using VPR activation domains.

protein domains that enhance or activate transcription has similarly found success [320–326].

1.5.3 Application of CRISPR-Cas9 in insect cells and baculovirus

Genome editing in *D. melanogaster* and other insects *in vivo*

Despite the significant development of CRISPR-Cas9 as a prominent tool for genetic engineering in an immense array of model and less studied organisms, its adoption as a tool in insect cells in the broad academic community was slower than in other communities [327]. Nevertheless, a growing body of work is accumulating in the literature describing the application of CRISPR-Cas9 for genome editing and, more recently, transcriptional control [327]. As *Drosophila melanogaster* is one of the pre-eminent model organisms for biological studies in genetics and physiology, and a model for development, human disease, and evolution, it has predictably been the insect of choice for CRISPR-Cas9 engineering to date, both in the whole organism *in vivo* and cultured cells *in vitro* [327].

Typically, for genetic engineering studies *in vivo*, the Cas9 transgene and sgRNA are injected into germline or somatic cells of the insect in a variety of formats. Transient expression of the *cas9* gene and sgRNA may be achieved via injection of a plasmid, or delivered directly as *in vitro*-transcribed *cas9* mRNA and sgRNA or purified Cas9 protein and sgRNA as a ribonucleoprotein (RNP) complex [327]. Alternatively, the *cas9* gene and/or sgRNA can be integrated into a chromosome for stable expression and propagation of the gene to offspring [328]. The editing efficiencies achieved vary widely with the approach

used and the genomic loci targeted for editing, ranging from ~0.03-100% for knockouts (KO), and ~2-38% for knock-ins (KI), in which an editing template is also included to promote HDR [327]. The promoter chosen for expression of *cas9* and sgRNA also appears to have a significant impact on mutation rates; the nanos promoter, which induces strong and restricted expression in germ-line cells, has been employed for efficient heritable germ-line transmission of mutations without widespread targeting in somatic cells, whereas the actin5c and vasa promoters can induce mutations with high frequencies in germ-line and somatic cells alike [328–330]. The choice of U6 promoter for expression of the sgRNA also appears to influence mutation rates. A comparison of the U6:1, U6:2, and U6:3 promoters revealed that the U6:3 promoter was most effective for driving mutagenesis in both somatic and germ-line cells. It is worth noting that the majority of studies published have used the U6:2 promoter for sgRNA expression, however it appears to be the least efficient [327, 328, 330]. The Cas9 'nickase' variant along with dual sgRNA expression for generating offsetting ssDNA nicks was also successful at inducing mutation, however the efficiency was low compared to using wild-type Cas9 [330]. Additionally, simultaneous, multiplexed genome editing has been demonstrated by expressing dual sgRNAs each targeting a different gene. Each of the 120 flies screened had the dark cuticle consistent with mutation at the *e* locus, and 52% had one or two curled wings, consistent with mutation of the *curled* (*cu*) gene [330]. Finally, co-injection of plasmids encoding Cas9 controlled by the hsp70 promoter and sgRNA expressed by snRNA U6:2 promoter achieved mutations in the *yellow* (*y*) gene in the somatic cells of 62% of G_0 males, whereas co-injection of Cas9 mRNA and sgRNAs resulted in mutations in 88%, possibly indicating that the *in vitro*-transcribed platform (or RNP complexes directly) may be more effective than plasmid-based delivery [331, 332].

Following these initial studies in *D. melanogaster*, application of CRISPR technology was extended to several disease vectors, particularly mosquitoes of the genera *Anopheles* and *Aedes* [333–338], as well as a handful of members of Lepidoptera, including *Bombyx mori*, *Danaus plexippus*, *Helicoverpa armigera*, *Papilio xuthus*, *Spodoptera litura*, *S. littoralis*, and most recently in *S. exigua* [333, 339–350]. In the members of *Spodoptera*, *in vitro*-transcribed sgRNA and Cas9 mRNA (*S. litura*) or Cas9/sgRNA RNPs (*S. littoralis* and *S. exigua*) were employed to generate mutations in *SlitPBP3* and *Slabd-A* (*S. litura*), *Orco* (*S. littoralis*), or *PGE2R* genes (*S. exigua*), with reported mutation efficiencies ranging from ~10-88% of hatched larvae, and ~6.6-43.3% for germ-line transmission [345–347, 349]. Similarly, co-injection of *in vitro*-transcribed sgRNA and Cas9 mRNA to *B. mori* preblastoderm embryos resulted in 94-95% of hatched larvae with a mosaic translucent integument phenotype consistent with BmBLOS2 disruption, with 17% of mutants showing severe oily skin phenotype (OSP), suggesting complete loss-of-function mutations at the BmBLOS2 locus [343]. When 2 sgRNAs were co-injected, ~51% had severe OSP, and upon crossing of these mutants with wild-type *B. mori*, ~37% of G_1 egg batches had at least one individual showing completely translucent skin, indicative of effective germ-line transmission [343]. Using the same approach, ~17-35% mutation frequencies were observed for 4 candidate genes in the G_0 generation, and the heritable mutation transmission frequency was ~29% [342].

Finally, it is worth noting the significant attention of the research community aimed at developing strategies for the control of insect vector-borne diseases such as malaria, dengue fever, yellow fever, Zika virus, and West Nile Virus [351]. As strategies employing insecticides have largely failed to curb or control the transmission of disease from mosquito

to human, genetic-based strategies have long held the most promise [352]. Indeed, the potential for using genetic strategies for 'driving' anti-pathogen genes into populations of wild, highly distributed vectors for the disease date back to the late 1960s [353]. Various 'gene drive' strategies have been employed over the intervening years, however long lasting and self-sustaining population suppression has been elusive for a variety of reasons, most notably the development of resistance alleles that limit the spread of the 'drive' allele [351, 352]. The potential of CRISPR-Cas technology for 'gene drive' however, may represent a paradigm shift in this endeavour; indeed, a recent opinion published in the academic journal *Proceedings of the National Academy of Sciences of the United States of America* characterized CRISPR-based gene drive as a biocontrol silver bullet and went on to state that it may no longer be appropriate to question whether we *can*, but rather whether we *should* control organisms using gene drive [354]. Evidence of its potential is highlighted by the volume of reports in the literature detailing the implementation of CRISPR-Cas9 as an effective 'gene drive' technology in the last few years [355–363]. These recent developments underpin the importance and impact of CRISPR-Cas technology on insect cell biotechnology.

Genome editing of cultured insect cells *in vitro*

Genetic engineering based on CRISPR-Cas9 has also been extended to the *D. melanogaster* S2 cell line cultured *in vitro*. Using plasmid-based delivery of the *cas9* gene and sgRNA, indel mutations at the *y* locus were induced in approximately 11% of alleles 3 days post transfection, however this increased to 88% upon culturing the cells with puromcyin to

select for transfected cells [364]. To promote gene KI via the HDR pathway in S2 cells, dsDNA donor plasmids were constructed to insert heterologous DNA at the *y* and AG01 loci, respectively. Genome integration was successful at both loci, with ~2% of alleles at the *y* locus and ~4% at the AG01 locus positive for integration [364]. The efficiency of KI was increased to ~50% by using a PCR-generated KI substrate with homology arms as short as 29 bp in concert with RNA interference (RNAi) for depletion of the NHEJ enzyme DNA ligase 4 (lig4) [365]. Finally, CRISPR-Cas9 was employed to target the *fdl* gene encoding Fused Lobes (FDL) for knockout [366]. FDL is a Golgi, transmembrane (TM), β -*N*-acetylglucosaminidase enzyme involved in the *N*-glycosylation pathway to produce paucimannosidic *N*-glycans in insect cells [366]. Editing with CRISPR-Cas9 produced indel mutations in the TM domain of FDL and resulted in dramatic changes to the *N*-glycosylation profile of recombinant proteins expressed in S2 cells [366].

In the *in vitro*-cultured cell lines derived from *B. mori* (BmN and BmE), Cas9-mediated indel mutations have been introduced at a single genomic locus, large ~3 kb DNA deletions using 2 sgRNAs, and multiplexed genome editing using 10 sgRNAs targeted to 6 different genes each residing on a different chromosome, simultaneously ([341]. To promote HDR over NHEJ for more sophisticated genetic engineering involving KI of heterologous DNA, factors involved in the NHEJ pathway were targeted for disruption using CRISPR-Cas9, including Ku70, Ku80, LigIV, XLF, and XRCC4 [344, 348]. Homologous recombination efficiency was markedly increased over NHEJ in these cell lines, ranging from ~4-8-fold more efficient than control cells with none of the genes required for NHEJ knocked out [348]. Recently, a genome-wide screen to identify genes essential for cell viability was conducted using CRISPR-Cas9 in the BmE cell line [367]. In this study, a sgRNA library containing 94

000 sgRNAs targeting ~17 000 protein coding genes were used to generate stable knockouts using *piggyBac* transposition. Over 1000 genes were identified as being essential for cell viability under the growth conditions tested, of which ~82% had homologous counterparts that are essential genes in several other animal species. Importantly, this study represented the first CRISPR library screening platform for non-model (i.e. not *D. melanogaster*) insect cells [367].

For CRISPR-Cas applications in cultured cells, the *cas9* gene and sgRNA are often delivered and expressed via plasmid transfection, as opposed to delivery of *in vitro* transcribed *cas9* and sgRNA mRNA or Cas9/sgRNA RNPs which are more common for *in vivo* applications [327]. While the *cas9* gene is transcribed by RNA Polymerase II (RNAP II) to produce mature mRNA, the sgRNA is not polyadenylated and therefore promoters recognized by RNAP III, and in particular promoters from U6 and H1 genes, are preferentially employed for its expression in animal cells [368]. As mentioned, various U6 snRNAs and the RNAP III-specific promoters that enable their expression have been identified in *D. melanogaster*, *B. mori*, and various genera of mosquito, but U6 promoters had not been identified that displayed transcriptional activity in *S. frugiperda* or *T. ni* [130, 368–370]. This was remedied in 2017 with a significant advancement: U6 promoters were identified in Sf9 and High Five cells, enabling plasmid-based genome editing using CRISPR-Cas9 in both of these cell lines [130]. Additionally, a clonal Sf9 cell line was isolated in which the *fdl* gene encoding a processing β -*N*-acetylglucosaminidase was disrupted. FDL removes a terminal *N*-acetyglucosamine residue from trimmed *N*-glycan-processing intermediates, effectively eliminating the substrates for *N*-acetyglucosaminyltransferase II, which is essential for the elongation process to complex (mammalian-like) *N*-glycans [130]. This represented

a significant advancement in developing Sf9 cell lines capable of complex, mammalian-type *N*-glycosylation. Screens for isolating RNAP III promoters have been recently conducted to enable *in vitro* CRISPR-Cas9 experiments to more disease-relevant mosquito cell lines, and a mosquito-optimized set of CRISPR-Cas9 plasmids has been developed and demonstrated in *Ae. albopictus* and *Ae. aegypti* cells, which are prolific vectors for several viral pathogens [368, 371].

Implementation of dCas9 for transcriptional control and epigentic remodeling

Although CRISPR-Cas9 has been implemented as a tool for genome editing in several insects and insect-derived cell lines, it has so far been limited to *D. melanogaster* and more recently *B. mori* as a tool for transcriptional control. For transcriptional repression in *D. melanogaster*, dCas9 and sgRNA were targeted to the male-specific *rox1* and *rox2* RNAs, which encode long, non-coding RNAs (lncRNAs) that form complexes with the male-specific lethal (MSL) proteins to facilitate targeting of the complex to the X chromosome for hyper-activation of the X-linked genes in male flies [372]. Interestingly, transcriptional repression in cultured cells *in vitro* reached ~95% and ~90% reduction in RNA abundance for *rox1* and *rox2* targets, respectively, without tethering a repressive protein domain to dCas9 [372]. These reductions in RNA abundance represent repression efficiencies far superior to those achieved in human cells with dCas9 only, and even rivals repression efficacy with dCas9-KRAB fusion proteins [316]. Additionally, sgRNAs targeting both the non-template and template strand showed effectiveness, suggesting that the DNA strand bias for silencing observed in human cells may be absent in insect cells [315,

372]. For transcriptional activation, on the other hand, dCas9 fusions to the transcriptional activator domain VP64, which is composed of 4 tandem copies of the Herpes Simplex Virus (HSV) Protein 16 activation domain (AD), or VPR, which has p65 AD of NF-κ B and R transactivator (Rta) of Epstein-Barr Virus (EBV) appended to dCas9 in addition to VP64, have been employed in *D. melanogaster* S2 cells [373]. Activation of *Mtk* and *CecA1* with dCas9-VPR reached ~3300-33000-fold transcriptional activation compared to the control, while dCas9-VP64-mediated activation was substantially lower, reaching ~270-fold for *Mtk* [373]. This study was further extended to examination of two additional dCas9-AD fusions. The SAM (synergistic activation mediator) AD consists of dCas9-VP64 with a modified sgRNA containing 2 hairpins to bind MS2 bacteriophage MCP (MS2 Coat Protein) dimers to recruit the MCP-p65-hsf1 (human heat shock factor 1) activators [320]. Suntag, on the other hand, consists of dCas9 fused to a chain of 10 GCN4 peptide epitopes to recruit and bind single chain antibodies specific to the GCN4 epitope fused to VP64 ADs. The SAM and Suntag systems theoretically recruit 12 and 40 ADs, respectively [320, 322, 374]. The VPR, SAM, and Suntag transcriptional activators all achieved similar levels of gene induction in the S2 cell line [374]. Finally, the dCas9-VPR system was employed for single- and multiplexed activation of genes in *in vitro* cultured cells and fruit flies *in vivo* [375]. Robust activation was observed for single-target experiments, ranging from ~10-1000 fold mRNA change in S2 cells, and ~6-15 fold for multiplexed activation [375]. The activation level achieved for a given gene appeared to be inversely correlated with its basal level of expression (i.e. CRISPR-dCas9 mediated activation was most effective for genes that are expressed at low levels in a given cell type) [375]. For proof-of-principle dCas9-VPR activation *in vivo*, the Wg gene was targeted to drive ectopic *Wg* expression, producing a partial

duplication of the wing pouch and other patterning abnormalities ultimately resulting in death during the early pupal stages [375]. This represented the first *in vivo* demonstration of CRISPR-dCas9 mediated transcriptional activation in any multicellular organism [375].

Similar strategies have also been employed for transcriptional repression and activation in BmE cells, including fusion of dCas9 with KRAB, Hairy, SID, ERD, and SRDX transcriptional repressors [376], and VP64 and VPR transcriptional activators [377]. While repression efficiency appeared to vary substantially with the gene target and specific sgRNA targeted, ~50% reduction in RNA abundance was achieved for 3 different gene targets in this demonstration [376], and ~1.5-8000-fold enhancement of RNA for the dCas-VPR fusion [377]. These results are comparable with those reported in *D. melanogaster*. Finally, programmable demethylation of genomic DNA in cultured BmE cells was demonstrated based on dCas9-TET1 fusions [378]. Since tumor-suppressor genes are frequently methylated during cancer progression and promoter hypermethylation is frequently associated with gene silencing, targeted demethylation has been employed as a strategy to enhance gene transcription. In BmE cells, targeted demethylation resulted in the up-regulation of four genes ~1.2-1.7 fold compared to control experiments [378].

CRISPR-Cas9 technologoies in the BEVS

Despite the increasing development of CRISPR-Cas technologies in insects *in vivo* and cultured cells *in vitro*, CRISPR-Cas9 has seen extremely limited development in the BEVS. CRISPR-mediated gene disruption was recently applied as an antiviral strategy, in which sgRNA targeted to the key replication factor *ie-1* of BmNPV effectively inhibited virus

replication [333]. Using stable cell lines expressing *cas9* and sgRNA, antiviral efficiency reached ~99%, demonstrating the promising potential of the system to provide antiviral protection in *B. mori* cell lines [333]. Inhibition of BmNPV was extended to transgenic silkworms in which a virus inducible *cas9* gene and sgRNA were introduced via the *piggyBac* transposon. Innoculation of the transgenic silkworms with BmNPV activated expression of Cas9 to inhibit BmNPV replication in infected cells by targeting the *B. mori* $ATAD_3A$ for disruption, which is a mitochondrial protein commandeered by BmNPV to promote virus replication [379]. Importantly, the transgenic silkworms displayed substantially enhanced survival rates compared to the control in response to BmNPV OB infection. The authors claimed this development could represent an important approach for breeding disease-resistant insect lines [379]. Finally, Cas9/sgRNA RNPs were co-transfected with AcMNPV genomic DNA to generate targeted gene disruptions in AcMNPV rBEVs [380]. Isolation of individual viruses from the supernatant of transfected cells followed by screening for successful knockouts by PCR and Sanger sequencing produced isogenic viral stocks with the desired mutations. This was further extended to knock-in of a *gfp* expression cassette at the *chitanase* locus of AcMNPV. The editing rate for KO experiments ranged from 10-40% of the analyzed clones depending on target and specific sgRNA, and ~20% of the analyzed clones expressed GFP, for the KI experiments. Interestingly, different clones having the same disrupted gene appeared to produce different levels of GFP in infected *S. exigua* larvae, potentially indicating some clonal variability [380].

Chapter 2

Adapting CRISPR-Cas9 technology for the BEVS

As discussed in Chapter 1, disrupting or downregulating various AcMNPV-encoded or host cell genes appears to be a promising strategy for improving recombinant protein expression in the BEVS. The traditional approach for generating mutant AcMNPV genomes involves either conventional homologous recombination in insect cells between a wild-type BEV genome and a plasmid with homologous regions adjacent to the region to be deleted, or deleting the gene in a bacmid replicating in *E. coli*. Both of these strategies can be time-consuming and labour intensive, and *trans*-complementation of disrupted essential genes to produce infectious virus often results in impaired virus replication and transgene expression. Consequently, the function of many genes in the AcMNPV genome have yet to be confirmed experimentally, and AcMNPV genomes routinely used for expression of recombinant proteins remain virtually unmodified. Similarly, gene silencing through RNAi has also been successfully employed for improving recombinant protein production, however target sequence selection often requires extensive empirical validation to identify high

silencing efficiency targets. Even then, some experiments simply do not yield positive results. In fact, a review paper devoted in part to analyzing unsuccessful RNAi experiments in Lepidopteran insects has been published [106]. RNAi based on exogenously introduced dsRNA molecules are limited by transfection efficiency and poorly scalable due to expensive transfection reagents and synthesis of RNA.

The application of CRISPR-Cas9 technology for gene disruption and transcriptional repression would enable effective scrutiny of targets and offer a scalable and efficient platform technology. However, this work commenced when CRISPR-Cas9 had not yet been applied to Sf9 or High Five cells, nor had it been employed in AcMNPV. Further, expression of the sgRNA is generally conferred using a RNA polymerase III-dependent U6 promoter, however none had been identified in the Sf9 genome. Consequently, in addition to searching the Sf9 genome to identify an endogenous U6 promoter, testing U6 promoters from closely related species for activity in Sf9 cells was necessary. The results were submitted to the journal *Viruses*. The manuscript is presented below in its original form. Supporting information is provided in Appendix A.

It is important to note that while this work was ongoing, a U6 promoter was identified in the Sf9 genome and subsequently used for CRISPR-mediated editing of the Sf9 genome [130]. This promoter was in fact the same promoter that was identified in our own search, and while it was important to include these results in this thesis, these results were not included in the manuscript. As such, this data is included following the manuscript as supplemental results.

Comparison of CRISPR-Cas9 tools for transcriptional repression and gene disruption in the BEVS

Mark R. Bruder[1], Sadru-Dean Walji[1], and Marc G. Aucoin[1]

[1] Department of Chemical Engineering, University of Waterloo.

2.1 Abstract

The generation of knockout viruses using recombineering of bacmids has greatly accelerated scrutiny of baculovirus genes for a variety of applications. However, the CRISPR-Cas9 system is a powerful tool that simplifies sequence-specific genome editing and effective transcriptional regulation of genes compared to traditional recombineering and RNAi approaches. Here, the effectiveness of the CRISPR-Cas9 system for gene disruption and transcriptional repression in the BEVS was compared. Cell lines constitutively expressing the *cas9* or *dcas9* gene were developed, and recombinant baculoviruses delivering the sgRNA were evaluated for disruption or repression of a reporter *gfp* gene. Finally, the effect of disruption or downregulation of endogenous AcMNPV gene targets on viral gene expression and replication was measured. This study provides a proof-of-concept that CRISPR-Cas9 technology may be an effective tool for efficient scrutiny of baculovirus genes through targeted gene disruption and transcriptional repression.

2.2 Introduction

Baculoviruses are a diverse group of enveloped viruses that contain large, double-stranded DNA (dsDNA) genomes. The insect-specific infection cycle is complex and biphasic; primary infection of the insect requires occlusion-derived virus (ODVs), which are virions encapsulated in a proteinaceous matrix, while cell-cell transmission is established by the budded virus (BV) phenotype [95]. Initial interest in baculoviruses nearly 100 years ago was for their promise as safe and selective bioinsecticides [381], however a recombinant baculovirus *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV; rBEV) was used to produce human IFN-β in 1983, marking the inception of the baculovirus expression vector system (BEVS) as a platform for recombinant protein production [65]. Recombinant AcMNPV and its permissive cell lines Sf9 or High Five is the most common format of the BEVS [8]. A major milestone of BEVS biotechnology was the development of the bacmid system, which allowed site-specific integration of foreign genes in the AcMNPV genome propagated in *E. coli* [69]. Coupled with the λ-red or RecET recombineering systems [382, 383], mutant AcMNPV genomes with sequence-specific deletions (i.e., knockout viruses, KOVs) were efficiently obtained, allowing functional studies of many of the ~150 annotated open reading frames (ORFs) of AcMNPV [96, 384–388]. Studies identifying *per os* infectivity factors, genes with insecticidal properties, and gene disruptions that improved recombinant protein production have also been conducted using bacmid technology [78, 389–391].

Bacmid technology has undeniably propelled baculovirus biotechnology, however like all genetic engineering approaches, it has drawbacks; even with recombineering, targeted

gene disruption requires multiple experimental steps to achieve, and genetic instability has been reported during the 'hyper-recombination' state (i.e., during phage recombinase gene expression) resulting in intramolecular rearrangements in *E. coli* [392]. While this is largely abrogated by recombination of the bacmid and transfer vector in insect cells, isolation of the desired mutant rBEV can be challenging with this approach [72]. Evaluation of KOVs is typically conducted by transfecting Sf9 cells with the bacmid and observing the spread of infection, which can be impacted by transfection efficiency and it may fail to identify ORFs that are amenable to downregulation but not disruption. For example, disruption of the AcMNPV ORF34 ORF resulted in its categorization as an essential gene, however its downregulation by RNA interference (RNAi) not only permitted virus replication but improved heterologous gene expression [114]. Finally, trans-complementing cell lines may be required to enable production of infective virions when essential genes are targeted for disruption [85, 393].

Gene silencing via RNAi involves double-stranded RNA (dsRNA) molecules that inhibit gene expression by triggering the degradation of messenger RNA (mRNA) through the RNA-induced silencing complex (RISC). The dsRNA molecules are typically introduced by transfection of small interfering RNAs (siRNAs) directly or plasmids from which short hairpin RNAs (shRNA) are transcribed. While less prominently employed compared to gene disruption, RNAi has found utility in the BEVS [104]. Target sequence selection, however, often requires extensive empirical validation to identify high silencing efficiency targets, and the effectiveness may be limited by low transfection efficiency or cytotoxicity of the transfection reagent [394, 395].

The CRISPR-Cas9 system has seen development for precision genome editing and tar-

geted transcriptome engineering in a multitude of biological organisms over the past decade [298, 314, 396]. Genome editing was recently reported in Sf9 and High Five cells, and for genome editing of AcMNPV itself [130, 380]. Here, we have extended upon these recent advancements by developing engineering tools based on Cas9 and its nuclease deficient variant dCas9 for targeted gene disruption (CRISPRd) and transcriptional repression (CRISPRi), respectively. Using cell lines developed for expression of the *cas9* or *dcas9* genes and sgRNAs delivered by the rBEV, 4 virus-encoded genes were targeted for gene disruption and transcriptional repression. In each case, the expected phenotype was observed. Importantly, virus seed stocks could be produced in parental Sf9 cells, and the same rBEVs could be used for evaluating gene disruption and transcriptional repression. Target gene selection is achieved by exchanging the spacer sequence of the sgRNA using conventional PCR. Accordingly, this approach simplifies the generation of KOVs by reducing the experimental steps required and allows for investigation of gene function using gene disruption and transcriptional repression using the same rBEV.

2.3 Materials and methods

2.3.1 Cells

Sf9 cells were maintained in suspension culture in Gibco SF900 III serum free medium (Fisher Scientific, Whitby ON) as described previously [152]. Sf9 cells were transfected as adherent cultures in tissue culture treated 6-well plates (VWR, Mississauga ON) with Escort IV transfection reagent (Sigma-Aldrich, Oakville ON) according to manufacturer's

directions. To derive the transgenic Sf9-Cas9 and Sf9-dCas9 cell lines, parental Sf9 cells were transfected with the plasmid pOpIE2-Cas9-puro or pOpIE2-dCas9-puro, respectively. Approximately 48 hours post transfection (hpt), the growth medium was aspirated and replaced with fresh medium containing 5 μg/ml puromycin (Sigma-Aldrich). Resistant cells were pooled, adapted back to suspension culture, and maintained for ~10-15 additional passages under selective pressure prior to further analysis.

2.3.2 Plasmid construction

All plasmids used in this study were constructed using NEBuilder HiFi DNA Assembly master mix (New England Biolabs, Whitby ON) according to manufacturer's directions. Primers used for construction of all plasmids and retargeting sgRNAs were purchased from Integrated DNA Technologies (IDT; Coralville, IA) and are listed in Table A.1.

To construct plasmid pOpIE2-Cas9-puro, the cas9-T2A-pac region from pAc-sgRNA-Cas9 (Addgene #49330, Cambridge MA) [364] and a fragment containing the *Orgyia pseudotsugata* MNPV immediate-early 2 promoter (OpIE2) and 3' untranslated region (UTR) [397], origin of replication (ori), and ampicillin resistance gene (ampR) for propagation in *E. coli* were amplified via PCR and the 2 PCR fragments were used in a Gibson assembly reaction [398]. The resulting plasmid placed the cas9-T2A-puro expression cassette under the control of the constitutive OpIE2 promoter. To generate plasmid pOpIE2-dCas9-puro, the dcas9 ORF was amplified from pdCas9::BFP-humanized (Addgene #44247) [314] and used in a Gibson assembly reaction along with a PCR fragment containing the OpIE2 promoter, T2A-puro cassette, and OpIE2 3' UTR to place the dCas9 gene under the con-

stitutive control of the OpIE2 promoter.

For plasmids containing the OpIE2-GFP cassette and SfU6 sgRNA, the *mAzami-Green* gene (Addgene #54798; herein referred to as *gfp*) encoding a monomeric green-emitting fluorescent protein [399] gene was PCR amplified and placed between the OpIE2 promoter region and 3' UTR. Separately, the *S. frugiperda* U6-3 (SfU6) small nuclear RNA (snRNA) promoter [130] was synthesized as a gblock (IDT), and PCR amplified along with the single guide RNA (sgRNA) and transcriptional terminator from plasmid pCFD4-U6:1_U6:3tandemgRNAs (Addgene #49411) [330]. The OpIE2-GFP fragments were inserted along with the SfU6-sgRNA DNA fragment into pACUW51 to derive plasmid pOpIE2GFP-sgRNA.

To construct p10-GFP and p6.9-GFP-encoding CRISPR plasmids, first the coding region of the *p10* gene, including upstream and downstream sequences to include its endogenous promoter and 3' UTR, was amplified from AcMNPV genomic DNA and inserted into pACUW51. The p10 ORF was then replaced with the *gfp* gene, and the SfU6-sgRNA fragment was inserted downstream to derive p10GFP-sgRNA. Finally, the p6.9 promoter region was amplified from AcMNPV genomic DNA and inserted in place of the p10 promoter sequence in p10GFP-sgRNA to yield p6.9GFP-sgRNA.

The spacer sequences used to target Cas9 and dCas9 to specific AcMNPV genomic loci were selected using the sgRNA scorer 2.0 software [400]. Briefly, the coding sequence for the target gene was submitted to the sgRNA scorer 2.0 software which generated a list of putative target sites scored according to their predicted activity. For each gene target, 2-4 target sequences were selected based on two criteria: predicted activity and

the strand (template or nontemplate) the target sequence resided on. Inverse PCR was used to retarget sgRNA spacer sequences to the target of interest [401]. Two primers were designed to anneal to the cas9 handle of the sgRNA sequence and to the U6 promoter sequence and extend in opposite directions. The desired targeting spacer sequence was appended to these primer sequences, which were used to amplify the entire plasmid as a linear fragment. The spacer sequence served as the homologous sequence required for Gibson assembly to ligate and re-circularize the new plasmid. The spacer sequences used in this study are presented in Table 2.1.

Table 2.1: Protospacer sequences for sgRNA targets

rBEV	Target	Protospacer Sequence (5'-3')	PAM	Strand
Control	n/a	caccttgaagcgcatgaact	n/a	n/a
GFP1	gfp	gggcaagggcaacccctacg	agg	Template
GFP2	gfp	gtcgtaggcgaagggcaggg	ggg	Nontemplate
GFP3	gfp	gttgccgtactggaacacgg	tgg	Nontemplate
GFP4	gfp	ccgagggctaccactgggag	agg	Template
IE1	ie-1	accgtgtcggctccatccgggg	tgg	Nontemplate
IE2	ie-1	tgatatctgacagcgagactg	cgg	Template
VL1	vlf-1	acacggactcgaaccggggag	cgg	Nontemplate
VL2	vlf-1	ggcaacgatgcacgcccgacg	agg	Template
VP1	vp80	gcccgccgcaatcgccgccg	cgg	Template
VP2	vp80	gctggatgttacccgcgg	cgg	Nontemplate
VP3	vp80	tcgatgcggccaggtcgc	tgg	Nontemplate
VP4	vp80	gcggatcgctaaatgccg	tgg	Nontemplate

2.3.3 Virus generation, amplification, and quantification

Plasmids for homologous recombination at the *polyhedrin* locus in the AcMNPV genome were co-transfected with *flash*BACGOLD™ (Oxford Expression Technologies Ltd., Oxford UK) genomic DNA according to manufacturer's directions. Supernatant from each transfection was harvested 4-5 days post transfection and used to infect early-exponential phase ($\sim 1.5 - 2 \times 10^6$ cells/ml) suspension Sf9 cultures at low multiplicity of infection (MOI) to amplify the rBEVs for 3-4 days or until the viable cell density dropped to ~80%. After 2 sequential rounds of amplification, the rBEV titer was quantified using end-point dilution assay (EPDA). Briefly, Sf9 cells were diluted to $\sim 2.0 \times 10^5$ cells/ml, and 100 μl was used to seed each well of a 96-well plate (VWR). The virus was serially diluted (10^{-2} to 10^{-8}), and 10 μl of each dilution was added, in 12 replicates, to the 96-well plate. Plates were incubated for seven days at 27 °C, after which they were checked for green fluorescence using a fluorescence microscope. Results were converted from $TCID_{50}$ and reported as plaque forming units per ml (pfu/ml).

2.3.4 Infections

Sf9-dCas9 , Sf-Cas9, or the parental Sf9 cells were infected with rBEVs at a density of $\sim 1.5 - 2 \times 10^6$ cells/ml and MOI of 3. Samples were taken at 24, 48, and 72 hours post infection (hpi) and cells were fixed with 2% paraformaldehyde for ~30 minutes prior to further analysis.

2.3.5 Flow cytometry and analysis

Samples were analyzed using a FACSCalibur™ flow cytometer (BD Biosciences, San Jose CA) equipped with a 15-mW air-cooled argon-ion laser with an excitation frequency of 488 nm. Samples were run at the low flow setting (12 μl/min) and 10000 events were collected. Analysis of flow cytometry data was performed using FlowJo® V10 flow cytometry analysis software (FlowJo LLC, Ashland, OR). Briefly, after applying gates to remove debris and intrinsic cellular fluorescence from the analysis, median fluorescence intensity in FL1 was calculated.

2.3.6 Real-time reverse transcription polymerase chain reaction (RT-PCR)

Infected cells ($\sim 1.5 - 2 \times 10^6$ cells/ml, MOI = 3) were collected at 0, 24, 48, and 72 hpi by centrifugation at $1000 \times g$ for 10 min at 4 °C. RNA was extracted using the Geneaid Total RNA Mini kit (FroggaBio, Concord ON) and 500 ng was used as template for first-strand cDNA synthesis using the SensiFAST cDNA synthesis kit (FroggaBio) according to manufacturer's directions. Real-time PCR was performed using the SensiFAST SYBR Hi-ROX kit (FroggaBio) according to manufacturer's directions on an Applied Biosystems StepOnePlus™ Real-Time PCR System (Fisher Scientific). Primer pairs used for qPCR are given in Table A.1.

2.3.7 Western blot

Infected cells ($\sim 1.5 - 2 \times 10^6$ cells/ml, MOI = 3) were collected at 0, 24, 48, and 72 hpi by centrifugation at $1000 \times g$ for 10 min at 4 °C. The cells were lysed in RIPA buffer (Fisher Scientific), quantified by BCA assay (Fisher Scientific), and ~10 μg of protein was separated by electrophoresis in 10% TGX Stain-Free precast mini SDS-PAGE gels (Bio-Rad, Mississauga ON) according to manufacturer's directions. After transfer to PVDF membranes, western blot analysis was performed with anti-Cas9 (MAC133, Sigma-Aldrich) or anti-GP64 (AcV5, Fisher Scientific) as primary antibodies and goat anti-mouse IgG HRP secondary (Bio-Rad) and imaged on a ChemiDoc MP Imager (Bio-Rad). The Image Lab software was used for further image processing (Bio-Rad).

2.3.8 Quantification of baculovirus particles using flow cytometry

Sample preparation for analysis via flow cytometry was described previously [402]. Briefly, samples were diluted in D-PBS and fixed with paraformaldehyde for 1 hour, after which the samples were subjected to one freeze-thaw cycle followed by incubation with Triton X-100 to permeabilize the membrane. The nucleic acid stain SYBR Green I was added and incubated at 80 °C for 10 min in the dark to stain double stranded DNA. After cooling on ice, the samples were analyzed via flow cytometry. Flow-Set Fluorospheres (Beckman Coulter, Mississauga ON) were used for calibration and all samples were run in triplicate.

2.4 Results

2.4.1 Development of a Sf9 cell line for constitutive expression of Cas9 and dCas9

Expression of the Cas9 and dCas9 proteins were conferred via the development of transgenic Sf9 cell lines transfected with plasmids pOpIE2-Cas9-puro and pOpIE2-dCas9-puro, which include either the *cas9* or *dcas9* gene and the *pac* gene sequences separated by the viral T2A element [403]. After selection with puromycin for at least 2 weeks, resistant cells were pooled and maintained in suspension culture for an additional ~10-15 passages under selective pressure before removing puromycin from the medium. Although routine maintenance of these cell lines provided no evidence that ectopic expression of either Cas9 proteins had any effect on their growth, prior to performing any gene disruption or transcriptional repression experiments, the cell lines were characterized with infection experiments to determine whether there were any distinguishable differences between them and parental Sf9 cells. As shown in Figure 2.1A, transcription of the *gfp* reporter and the viral capsid protein *vp39* genes were similar, indicating that there were no discernable differences in progression of the infection. Similarly, the production of GFP protein from the viral late gene promoter p6.9 and progeny virus appeared unimpaired (Figure 2.1B & C). Interestingly, RT-PCR (Figure 2.1A) and western blot data (Figure A.1) indicated that transcription of *cas9* and *dcas9* were significantly downregulated by 24 hpi and protein was undetectable on western blot by 48 hpi.

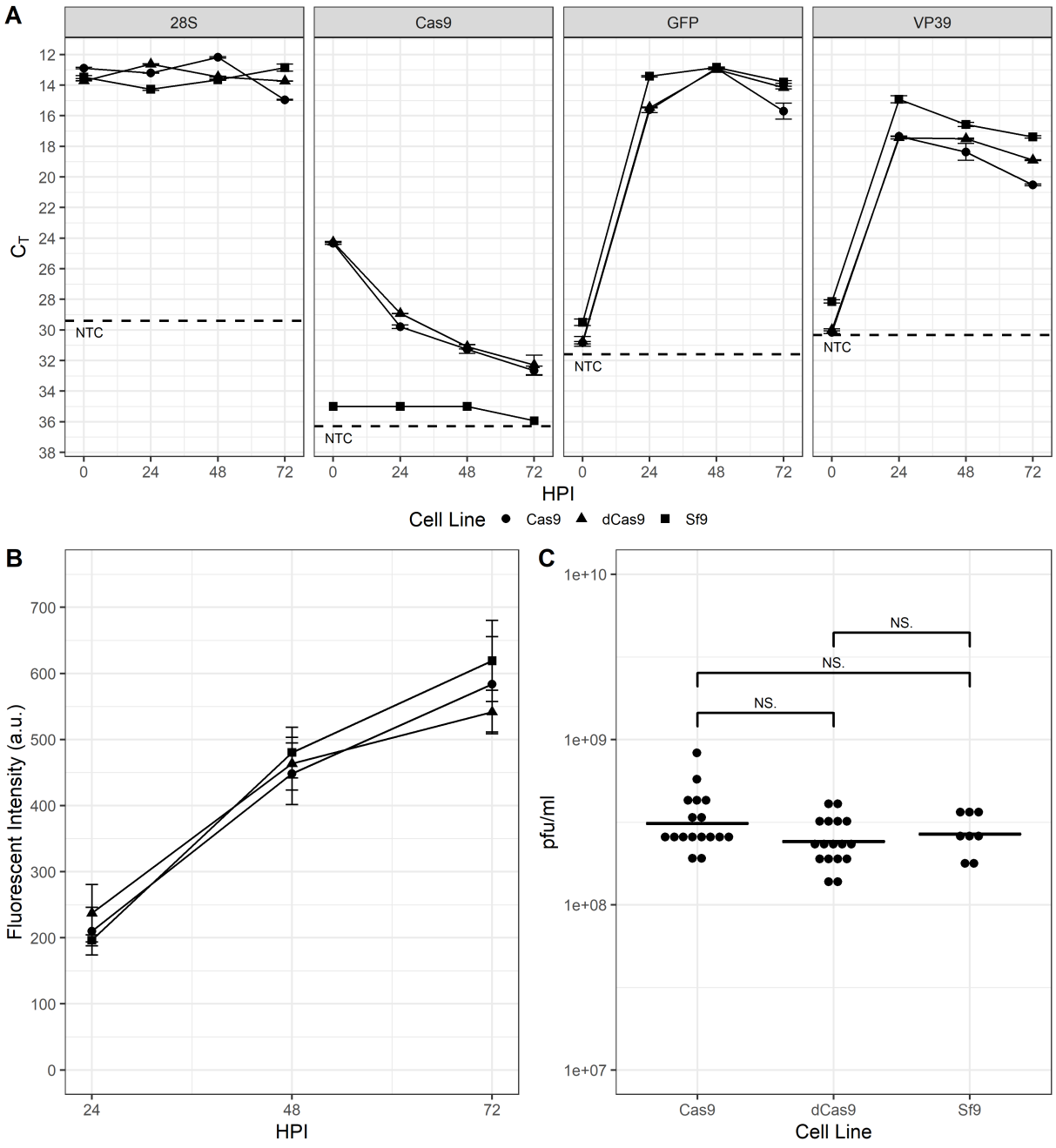


Figure 2.1: **Sf9-Cas9 and Sf9-dCas9 cells are indistinguishable from the parental Sf9 cell line. A.** RT-PCR expression analysis of virus-encoded *vp39* and *gfp* reporter gene are not affected by the presence of either *cas9* or *dcas9* expression. Both *cas9* and *dcas9* are downregulated in response to infection. **B.** GFP fluorescence intensity and **C.** progeny virus production are similar between all cell lines.

2.4.2 Evaluation of CRISPR-mediated repression and disruption on GFP production

Initial experiments sought to establish transcriptional repression of the rBEV-encoded *gfp* gene. Individual rBEVs with sgRNAs targeting the template (GFP1 and GFP4) and non-template (GFP2 and GFP3) strands within the *gfp* ORF were constructed, and repression of *gfp* transcribed from immediate early (OpIE2), late (p6.9) and very late (p10) promoters was assessed. Sf9-dCas9 cells infected with rBEVs encoding sgRNAs GFP2 and GFP3 showed a marked decrease in the proportion of GFP-positive cells and fluorescence intensity compared to the control at 48 hpi for OpIE2-GFP (Figure A.2A). For p6.9-GFP and p10-GFP, there appeared to be a slight (p6.9-GFP) or significant (p10-GFP) reduction in GFP-positive cells at 24 hpi, however there was no difference at 48 and 72 hpi (Figure 2.2A & A.2B). Nevertheless, the fluorescence intensity for GFP2 and GFP3 targets were reduced compared to the control at all time points for p6.9-GFP, and at 48 and 72 hpi for p10-GFP. Fluorescence intensity of rBEVs encoding sgRNAs GFP1 and GFP4 were indistinguishable from the control in all experiments, however, indicating potential strand bias for CRISPRi.

For CRISPRd experiments, p6.9-GFP rBEVs encoding sgRNAs GFP2, GFP3, and GFP4 were used to infect Sf9-Cas9 cells. For all 3 sgRNAs, the proportion of GFP-positive cells was significantly reduced compared to the control at all time points. The GFP2 sgRNA resulted in the lowest GFP-positive phenotype compared to GFP3 and GFP4. Significantly, whereas the proportion of GFP-positive cells was higher at 24 hpi and increased by 48 hpi for GFP3 and GFP4, the rBEV encoding the GFP2 sgRNA was less

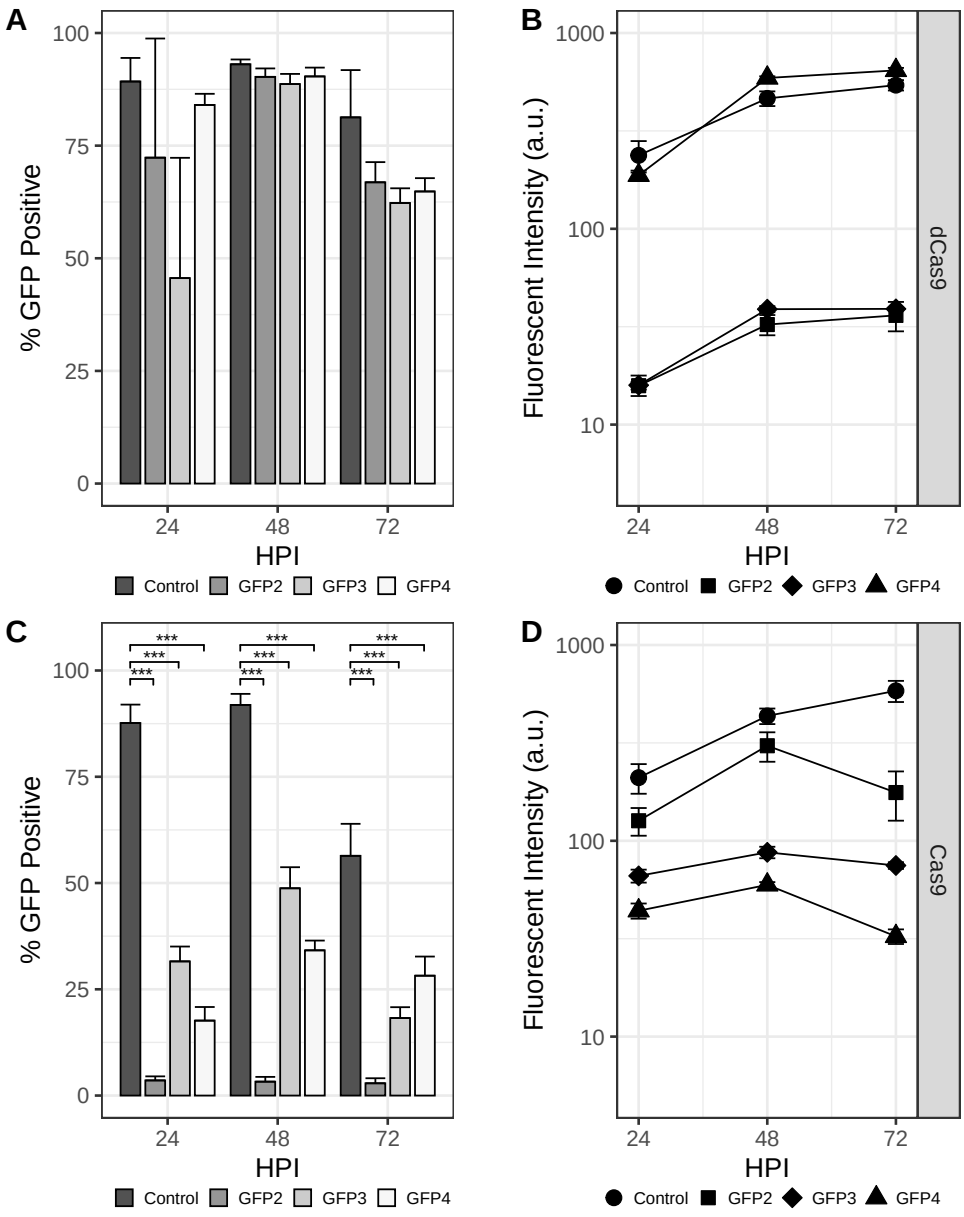


Figure 2.2: **CRISPR-mediated targeting of GFP transcribed from the late p6.9 promoter.** Percentage of population that is GFP-positive and fluorescence intensity of *gfp*-targeting rBEVS in Sf9-dCas9 cells (**A** & **B**), and Sf9-Cas9 cells (**C** & **D**), respectively.

than 10% GFP-positive at 24 hpi and did not increase as the infection progressed. For the fluorescence intensity measurements of the GFP-positive cells, though, the GFP2 sgRNA rBEV was only slightly reduced compared to the untargeted control rBEV. Conversely, the GFP3 and GFP4 sgRNAs had significantly reduced fluorescence intensity compared to both GFP2 and control (Figure 2.2C & D).

Importantly, parental Sf9 cells infected with each of the p10-GFP (data not shown) and p6.9-GFP rBEVs produced fluorescence intensity measurements that were both increased compared to the same infections in either Sf9-Cas9 or Sf9-dCas9 cells and were also all similar to the control (i.e., untargeted) rBEV. Finally, release of progeny BV was not statistically different for any of the viruses replicating in any cell line (Figure A.3). Taken together, these data indicate that production of GFP was influenced by the presence of both (d)Cas9 and sgRNA and was unaffected in the absence of either of these molecules.

2.4.3 Extension of CRISPRi and CRISPRd to endogenous AcMNPV *ie-1* and *vlf-1* genes

Next, the ability of the CRISPRi and CRISPRd systems to affect the expression of endogenous AcMNPV genes was assessed (Figure 2.3). Spacer sequences were selected to target the *ie-1* and *vlf-1* genes encoding immediate-early protein 1 (IE-1) and very late factor 1 (VLF-1), respectively. The *ie-1* gene encodes a transcriptional activator, and is essential for viral DNA replication, late gene expression, and subsequent progeny virus production. The *vlf-1* gene encodes a transcriptional activator for the very late class of genes but has no effect on late gene promoters. Production of progeny virus and GFP transcribed from the

p10 promoter was measured to assess the phenotypic impact of these targets in Sf9-dCas9 cells. Similar to previous experiments, the proportion of GFP-positive cells was reduced at 24 hpi for the IE1 sgRNA, however it increased by 48 hpi and was indistinguishable from the control. The proportion of GFP-positive cells was not affected for either IE2, VL1, or VL2. Reduced fluorescence for the rBEVs encoding nontemplate strand-targeting sgRNAs (IE1 and VL1), but not the sgRNAs targeting the template strand (IE2 and VL2) was observed. Finally, analysis of progeny virus production showed that the IE1 sgRNA reduced the infectious virus titer (IVT) ~90% compared to the control at 48 hpi. The difference in IVT for the other targets was not statistically significant (Figure 2.3A).

For CRISPRd in Sf9-Cas9 cells, a marked reduction in GFP-positive cells was observed for both IE1 and IE2 sgRNA rBEVs, but not VL1 or VL2, at all time points. Fluorescence intensity was also significantly reduced for each sgRNA compared to the untargeted control. Since VLF-1 stimulates transcription from very late promoters but has no effect on late genes, rBEVs encoding the VL1 and VL2 sgRNAs and the p6.9-GFP expression cassette were prepared and used to infect Sf9-Cas9 cells. Importantly, analysis of fluorescence indicated no difference compared to the control rBEV for both VL1 and VL2 (data not shown). Finally, the IVT was reduced by ~99% compared to the control for IE1 and IE2, and ~64% for VL2. The measured IVT for VL1 was not statistically different from the control (Figure 2.3B).

Parental Sf9 cells infected with each of the *ie-1* and *vlf-1*-targeted sgRNAs showed fluorescence and IVT levels that were consistent with control (i.e., non-disrupted/repressed) levels, indicating that both (d)Cas9 and sgRNA are required for disruption or downregulation of *ie-1* and *vlf-1* (Figure 2.3C). Additionally, since both fluorescence intensity and IVT

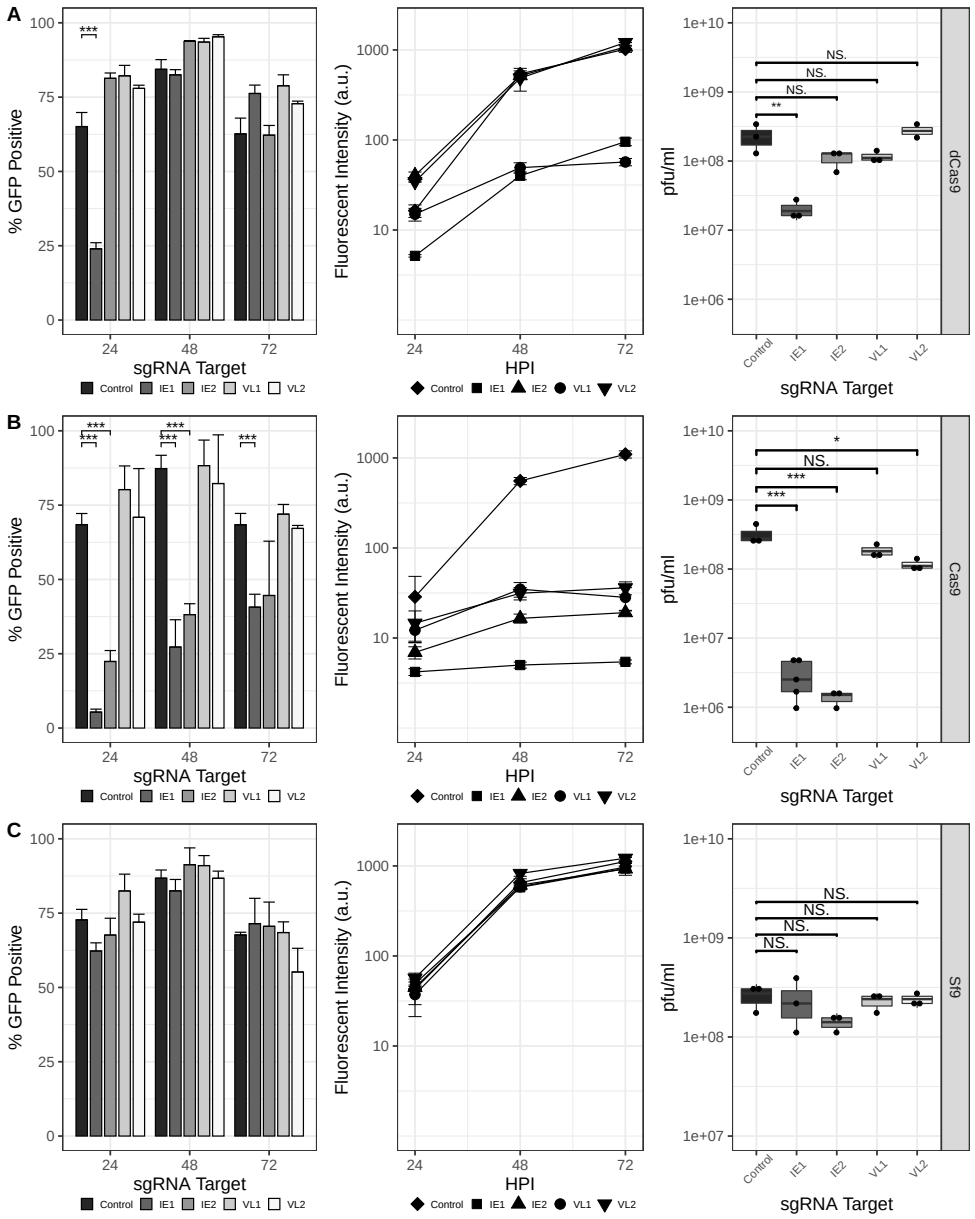


Figure 2.3: **CRISPR-mediated targeting of the AcMNPV *ie-1* and *vlf-1* genes.** Percent GFP-positive, fluorescence intensity, and IVT for rBEVs in **A.** Sf-dCas9, **B.** Sf9-Cas9, and **C.** parental Sf9 cells.

were similar for all of the rBEV infections in Sf9 cells, it is unlikely that the phenotypes observed in Sf9-Cas9 or Sf9-dCas9 cells are the result of genetic heterogeneity between the rBEVs or differences in viral gene expression and/or DNA replication.

2.4.4 CRISPRd is more effective than CRISPRi for obstructing progeny BV release

Finally, rBEVs with sgRNAs targeting the *vp80* gene were prepared and analyzed with CRISPRi and CRISPRd. The *vp80* gene encodes the capsid-associated protein VP80, and its disruption prevents capsid assembly but has no effect on late gene expression. The fluorescence intensity was similar for all *vp80*-targeting rBEVs compared to the untargeted control in each cell line, and showed no differences in either Sf9, Sf9-Cas9, or Sf9-dCas9 cells (Figure 2.4A). Similarly, the IVT of infected Sf9 cell culture supernatants for each rBEV was similar, indicating unimpaired BV release in the absense of Cas9 or dCas9. The nontemplate strand-targeting VP2, VP3, and VP4 rBEVs reduced IVT ~79%, ~68%, and ~57% compared to the control rBEV, respectively, in Sf9-dCas9 cells, while VP1 was similar to the control and to the IVT yielded from its infection of parental Sf9 cells. Infection with VP1 of Sf9-Cas9 cells, on the other hand, reduced the IVT by ~98% compared to the untargeted control, and VP2 by ~96% (Figure 2.4C). The latter result represents an ~85% improvement over the result in Sf9-dCas9 cells. Finally, flow cytometry analysis indicated a reduction in total number of particles in culture supernatant for VP80 targets as compared to non-targeted control experiments in Sf9-Cas9 cells (Figure 2.4B).

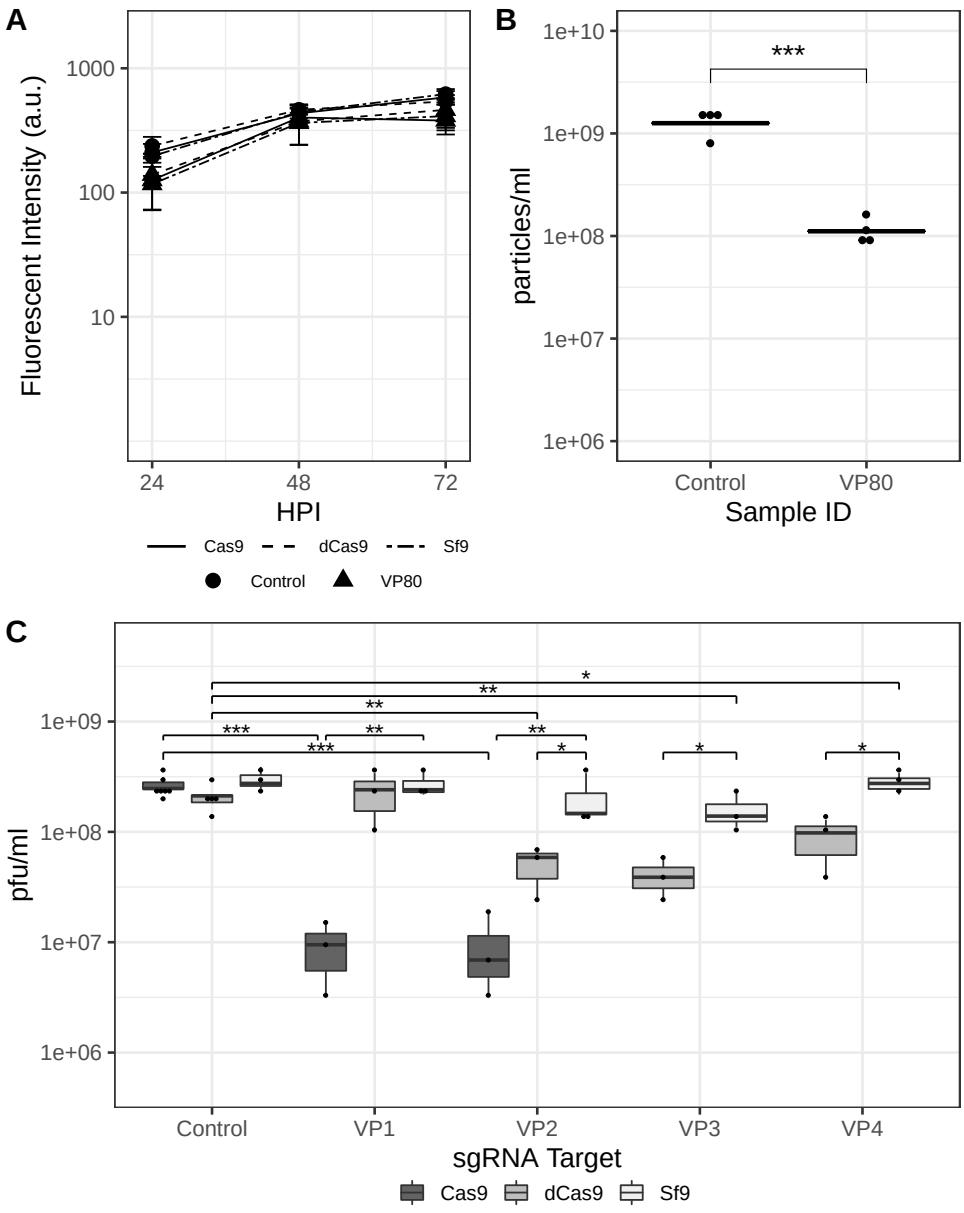


Figure 2.4: **CRISPR-mediated targeting of the AcMNPV *vp80* gene. A.** Mean fluroescence intensity for *vp80*-targeting and control rBEVs in Sf9-Cas9, Sf9-dCas9, and parental Sf9 cells. **B.** Total particles in culture supernatants of infected Sf9-Cas9 cells and **C.** IVT for control and *vp80*-targeting rBEVs in each cell line at 48 hpi.

2.5 Discussion

The present study sought to develop an efficient and robust technology for targeted genome engineering that would be capable of scrutinizing the effect of gene disruption or repression on viral gene expression and replication. To this end, transgenic Sf9 cell lines constitutively expressing the *cas9* or *dcas9* gene were developed. To test the efficacy of CRISPRd and CRISPRi using this approach, Sf9-Cas9 and Sf9-dCas9 cells were infected with rBEVs encoding both the sgRNA and target for disruption or repression. This strategy would ensure that every cell would receive the genetic code required for (d)Cas9 and sgRNA expression and thus present the highest probability for having the necessary resolution in the assay to observe the effects of the target gene disruption or repression.

Since downregulation of host protein expression due to infection with AcMNPV is a characteristic of the BEVS [404], experiments to assess expression of (d)Cas9 using RT-PCR and western blot were conducted. Consistent with prior studies, downregulation of *dcas9* and *cas9* was observed by 24 hpi, however the baculovirus immediate-early promoter supported transgene expression until 72 hpi [16, 405]. As the amount of dCas9 protein present could impact the effectiveness of CRISPRi, the repression of *gfp* transcribed from immediate-early (OpIE2), late (p6.9) and very late (p10) promoters were evaluated to establish the efficiency of repression for promoters differing in temporal and relative strength expression characteristics. Significant reduction in fluorescence intensity was observed for each promoter. Additionally, the possibility of a 'strand bias' was observed in the BEVS system, in which robust transcriptional repression can only be achieved by targeting the sgRNA/dCas9 complex to the nontemplate strand. This phenomenon has been observed in

various other prokaryotic and eukaryotic systems previously [314, 396]. Experiments conducted with Sf9-Cas9 cells (CRISPRd) showed decreased fluorescence intensity measurements compared to the control and the population of GFP-positive cells was significantly reduced. Notably, GFP2 was less than 5% GFP-positive, however fluorescence intensity was higher in Sf9-Cas9 cells than in Sf9-dCas9 cells. This observation is presumably due to the mechanisms by which CRISPRd and CRISPRi function; for CRISPRi, successful targeting blocks transcript elongation and leads to a reduction in mRNA produced and translated by the cell. This ultimately leads to an overall reduction in fluorescence intensity. Gene disruption mediated by CRISPRd results in indel mutations from dsDNA break repair, and protein expression is impacted by translation but not transcription [277, 314]. Any gene copies that are not successfully targeted or the indel mutation is silent would be transcribed and translated at wild-type levels.

Finally, endogenous AcMNPV *ie-1*, *vlf-1* and *vp80* genes were targeted for transcriptional repression and gene disruption. The IE-1 protein is the major transcriptional regulator of AcMNPV and is responsible for *trans*-activation of several known early genes [406, 407]. Importantly, it is one of several genes required for late gene expression and viral genome replication [407, 408]. Deletion of the *ie-1* gene results in loss of infectivity [384]. The VLF-1 protein is a regulator of very late gene transcription and is responsible for the 'transcriptional burst' observed for the very late class of genes; purified VLF-1 stimulated transcription of the very late polh promoter in a concentration depended fashion but had no apparent effect on the late 39k promoter [409, 410]. Complete deletion of the *vlf-1* gene may also impair assembly of BVs, although DNA replication and late gene transcription appeared to be reduced but permitted [385]. On the other hand, the *vp80* gene encodes a

capsid-associated protein that is essential for BV production but is not essential for viral late gene expression [85]. Selection of these endogenous genes provided the ability to observe the efficacy of CRISPRi and CRISPRd in several ways; repression/disruption of *ie-1* should impact the entire infection cycle of the rBEV, while targeting *vlf-1* should reduce expression from the very late p10 promoter but not the p6.9 promoter. Finally, disruption of *vp80* expression should impact the production of progeny BV but not inhibit late gene expression.

Infections with rBEVs encoding sgRNAs targeting each of these genes yielded the expected result in all three cases; significant reduction in both GFP and progeny virus production for IE-1 targets, reduced BV production but unimpaired late gene expression for VP80 targets, and targeting VLF-1 led to a reduction in fluorescence intensity for GFP expressed from the p10 promoter but not p6.9. Interestingly, although lower IVT was measured for VL1 and VL2 rBEVs in Sf9-Cas9 cells (~42% and ~64%, respectively) and for VL1 in Sf9-dCas9 cells (~50%), only the VL2 rBEV IVT was statistically different from the control. This could indicate that either the resolution in the assay is not sensitive enough to detect this difference or that enough VLF-1 was produced to permit replication and production of progeny virus to near wild-type levels. Nevertheless, these results agree with a previous study in which deletion of the *vlf-1* gene had no effect on late gene transcription but substantially reduced expression from the very late p10 promoter [385]. Unsurprisingly, the template-targeting sgRNA VP1 did not result in reduced progeny virus production in Sf9-dCas9 cells. Targeting the *vp80* gene with nontemplate-targeting sgRNAs VP2, VP3, and VP4, however, reduced IVT by ~79%, ~68%, and ~57%, respectively, however the result with VP4 was not statistically significant. Given that transcriptional repression ef-

ficiency has been observed to be inversely correlated to the distance of the target spacer sequence from the transcriptional start site [314], it may not be entirely surprising that this sgRNA was less effective.

In addition to the apparent strand bias in the experiments with Sf9-dCas9 cells, the proportion of cells displaying a GFP-positive phenotype at 24 hpi was substantially lower for the IE1 rBEV, while fluorescence was substantially reduced at all time points and for both targets in Sf9-Cas9 cells. Analysis of BV release at 48 hpi also showed ~90% decrease in the IVT for IE1 compared to controls for CRISPRi and ~99% for CRISPRd. This latter result is significant since a report in which transformed Sf9 cells expressing a ~470 bp dsRNA molecule targeting the AcMNPV *ie-1* gene exhibited strong viral repression at early stages of infection but subsequent recovery of viral proliferation was observed by the late stages of the infection cycle [411].

Deletion of the *vp80* gene has previously been shown to prevent BV production whilst permitting replication of viral DNA and transcription of viral late genes at or near wild-type levels [85]. Results presented here support these conclusions: production of GFP was similar for each virus in Sf9-dCas9, Sf9-Cas9, and the parental Sf9 cells, however production of progeny virus was decreased by >90% in Sf9-Cas9 cells. Interestingly, the supernatant from Sf9 cells infected with the Δ*vp80*-rBEV in that study appeared to have undetectable IVT [85]. Assessment of infected culture supernatants at 4, 8, and 12 hpi here and previously [152], however, revealed IVT ~$10^4 - 10^5$ pfu/ml at each time point (data not shown), indicating incomplete viral uptake before the onset of progeny BV release. Further, the *trans*-complementation strategy resulted in a ~25-fold decrease in BV seed production and constitutive expression of the *vp80* gene appeared unstable or toxic to Sf9 cells. Finally,

higher MOI (MOI = 10) was required in order to produce recombinant protein at the same level as the wild-type rBEV [85]. In this study, there was no difference in GFP production at MOI = 3 and each of the rBEVs displayed no indication of impaired replication in Sf9 cells. Taken together, this strategy may contribute to reduced downstream processing complexity by minimizing rBEV contamination. Nevertheless, targeted disruption of *vp80* reduced the IVT by ~98% and ~96% for VP1 and VP2, respectively. Compared to CRISPRi, these results indicate that CRISPRd may be more effective for reducing progeny virus production. Finally, to ensure that targeting *vp80* resulted in a reduction of particles released to the culture supernatant as opposed to the release of defective particles that are not infectious, flow cytometry was used to analyze supernatants from several control and VP80-targeted infections in Sf9-Cas9 cells. The results indicated ~90% reduction in particle concentration in the *vp80*-disrupted infections as compared to the control. Consistent with previous reports in which the ratio of total particles quantified using flow cytometry to IVT measured using EPDA ranged from 1 to 10 [402], the FC:IVT ratio in the samples analyzed was ~5-10 as well.

2.6 Concluding remarks

Taken together, the phenotypes observed in this report are consistent with disruption or repression of the endogenous AcMNPV *ie-1*, *vlf-1*, *vp80* genes. The results indicate that CRISPRd may be more effective than CRISPRi for total disruption of target gene expression, whereas CRISPRi allows transcription of the targeted gene at levels that are lower than wild-type, suggesting it may be more appropriate for targets that are not amenable

to deletion. Consequently, the CRISPRd tool developed here may be more useful for evaluating the essentiality of endogenous AcMNPV genes and reducing BV contamination in culture supernatants, whereas CRISPRi may be more effective for indentifying candidate targets whose differential expression may contribute to prolonging the infection cycle and accompanying bioprocess in order to increase yield of the target recombinant molecule. Importantly, the same rBEVs can be used for both CRISPRd and CRISPRi, which should allow for a more streamlined approach for scrutinizing baculovirus genes. Thus, these technologies are a valuable addition to the BEVS biotechnology 'tool-box'.

2.7 Supporting Results

2.7.1 Materials and methods

Bioinformatics

The *S. frugiperda* U6 small nuclear RNA (snRNA) promoter was identified by blastn queries of the *S. frugiperda* draft genome sequence (Genbank accession GCA_000753635.2) [412] using the U6-1 or U6-2 snRNA sequences of *Bombyx mori* [369] as the query sequence. Putative promoter elements were identified by comparing the consensus sequences of the *B. mori* or *Drosophila melanogaster* U6 promoters [370] to sequences upstream of the *S. frugiperda* sequences identified by BLAST. All multiple sequence alignments and consensus sequence determinations were performed using the *msa* package in the R programming environment [413].

Transient transfection, flow cytometry, and analysis

Sf9-dCas9 cells were seeded on tissue culture treated 6-well plates (~7.5×10^5 cells/well) and transfected with ~2 μg of plasmid using Escort IV transfection reagent according to manufacturer's directions. After incubating at 27 °C for 5-6 hours, the medium was aspirated and fresh SF900 III was added. Plates were incubated for an additional 48 hours post transfection (hpt), after which the cells were harvested, fixed in 2% paraformaldehyde (Fisher Scientific) prior to further analysis. Samples were analyzed using a BD FACScalibur flow cytometer and analyzed using FlowJo software, as above. Median fluorescence intensity of the control was used for gating and the percentage of the population with fluorescence intensity above and below that threshold was recorded.

2.7.2 Results and Discussion

Lepidopteran but not Dipteran U6 promoters display activity sufficient for CRISPRi

It was initially hypothesized that robust transcriptional repression would require strong expression of the sgRNA molecule. The predominance of employing endogenous U6 snRNA promoters for sgRNA expression in insect cells already in the literature [338, 341, 364, 366] implied that an endogenous RNA Polymerase III (RNAP III)-dependent promoter or from closely-related homologues would enable high level expression. However, no *S. frugiperda* RNAP III promoters had been reported as yet [412], and so experiments to test U6 snRNA promoters from the closely related lepidopteran *B. mori* and more dis-

tantly related dipteran*D. melanogaster* [370] were conceived, in addition to searching the *S. frugiperda* genome for its own U6 snRNA gene. The U6-2 promoter of *B. mori* (BmU6-2) and U6-3 promoter of *D. melanogaster* (DmU6-3) had previously been employed to produce sgRNA or shRNA molecules, and were chosen for analysis in Sf9 cells [330, 341, 344, 414]. A blastn search of the *S. frugiperda* genome using the sequence of the *B. mori* U6-1 or U6-2 snRNA [369] as the search query yielded 4 putative U6 snRNA sequences with very high query coverage (97-100%) and percent identity (90-97%), suggesting that the sequences could be a U6 snRNA gene (Figure 2.5). These sequences, along with ~500 nucleotides upstream to encompass the putative promoter region, were taken for further analysis. The upstream sequences were aligned with the consensus sequences of either the *B. mori* or *D. melanogaster* U6 promoter to identify promoter elements. Analysis of one of the putative hits identified sequence that resembled the proximal sequence element A (PSEA) and TATA box consensus sequence of *B. mori* and the interspecies insect PSEA/TATA consensus sequences for RNA Polymerase III-dependent promoters [370]. Importantly, this putative promoter sequence was located in the correct orientation adjacent to the putative snRNA sequence of *B. mori* U6 and (Figure 2.6). No similarity to the *D. melanogaster* U6 promoter sequences was found. The sequence immediately 5' to the putative snRNA sequence and ~400 nucleotides upstream was selected for analysis as a putative SfU6 for sgRNA expression in experiments along with BmU6 and DmU6.

To test the 3 U6 promoters (SfU6, BmU6, and DmU6), plasmids encoding GFP under the control of the OpIE2 promoter along with a sgRNA expressed from each of the U6 promoters were constructed. Spacer sequences to target dCas9 to the ORF of the *gfp* gene on either the template or nontemplate strand were selected (Figure 2.7A), and the

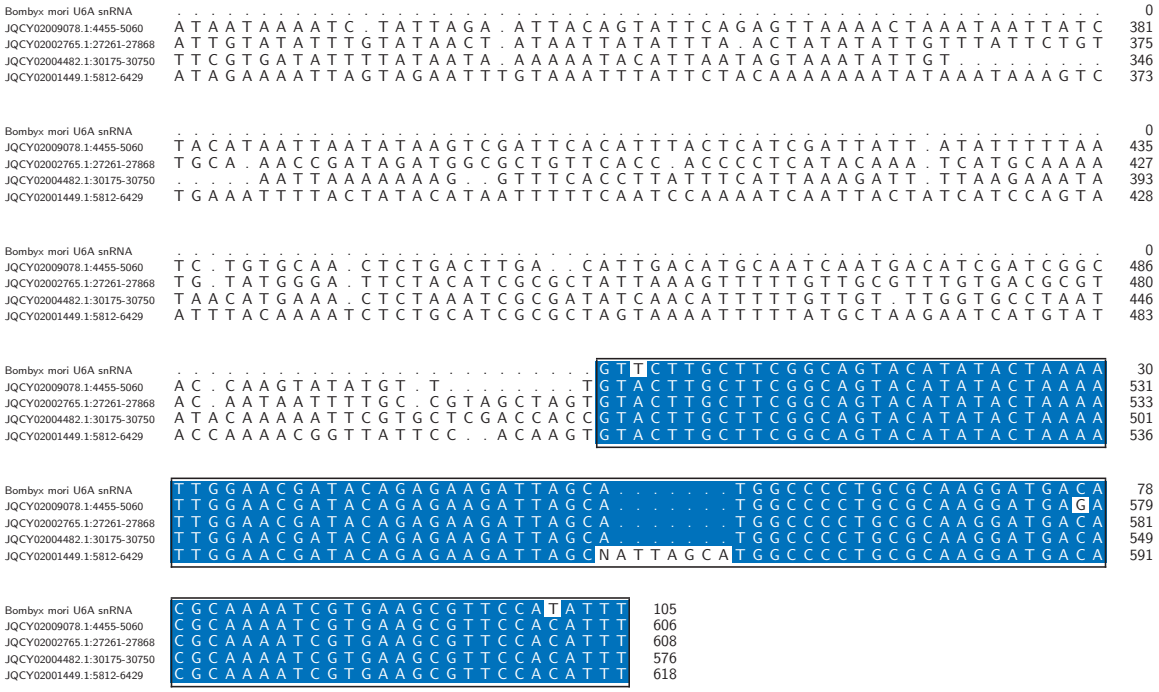


Figure 2.5: **Putative U6 snRNA genes present in the Sf9 genome.** Results from a blastn query of the Sf9 genome with the U6A snRNA gene of *B. mori* identified 4 putative results.

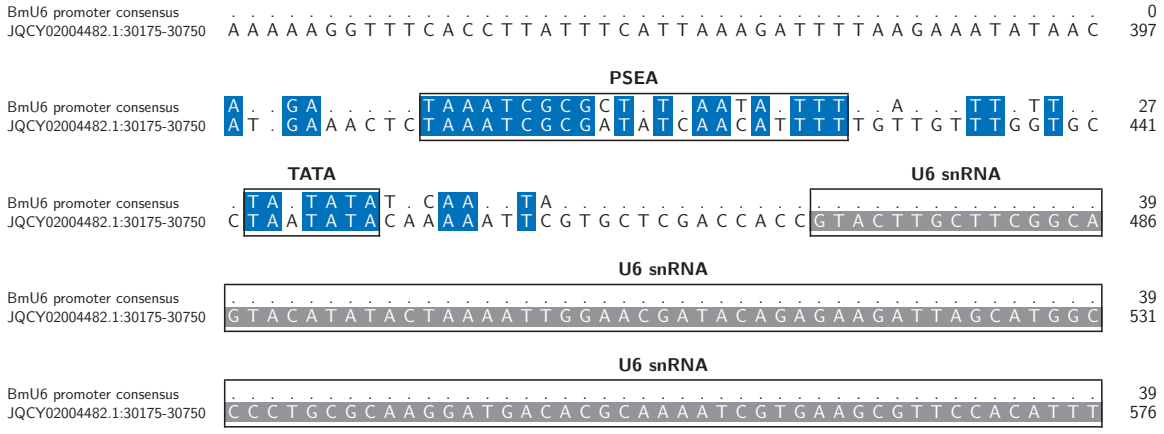


Figure 2.6: **Putative RNAP III-dependent promoter elements associated with a Sf9 U6 snRNA gene sequence.** The U6 promoter consensus sequence of *B. mori* was used to identify possible promoter elements located adjacent to sequences identified by blastn queries of the Sf9 genome with the *B. mori* U6A snRNA gene sequence.

resulting 6 plasmids were used to transfect Sf9-dCas9 cells. Fluorescence intensity was analyzed 48 hours post transfection by flow cytometry. The SfU6-driven sgRNA targeting the nontemplate strand at position 145 (GFP2) showed both a decreased proportion of GFP-positive cells (Figure 2.7B), and those that were GFP-positive were significantly lower in fluorescence intensity than the control (Figure 2.7C). The BmU6-driven GFP2 sgRNA also displayed a decrease in fluorescence intensity, however DmU6 GFP2 showed no difference as compared to the control. The template targeting spacer sequence (GFP1) expressed using all 3 U6 promoters, on the other hand, displayed fluorescence intensity distributions indistinguishable from the control transfections.

In light of this result, 2 additional spacer sequence targets were evaluated (GFP3 and GFP4) (Figure 2.8A). Once again, GFP3 showed both a decreased proportion of GFP-

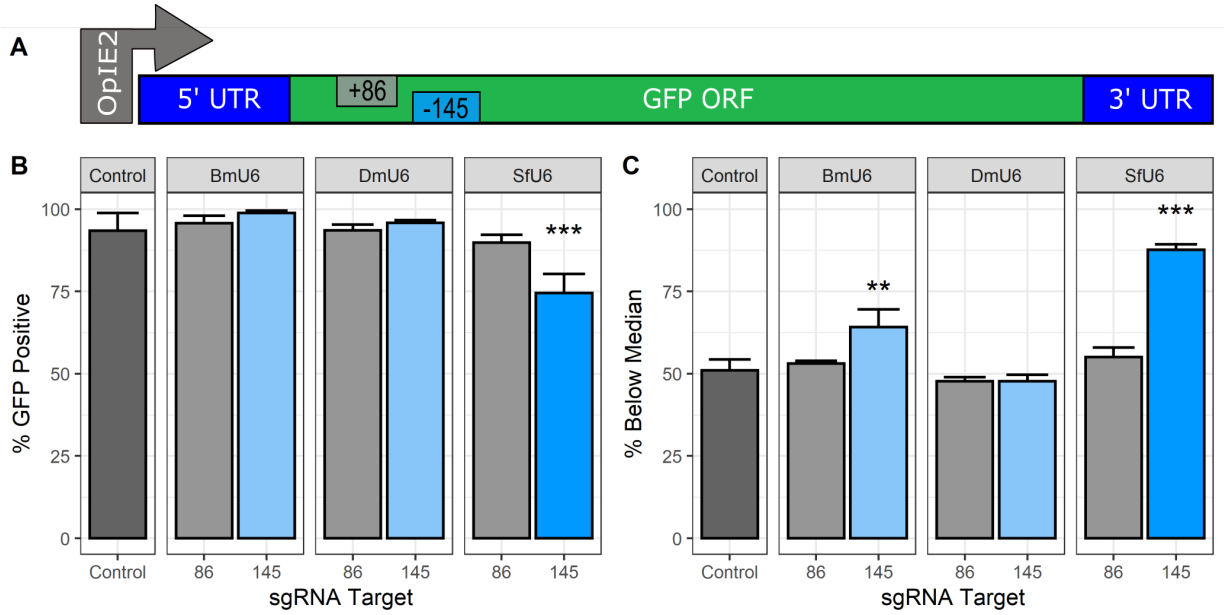


Figure 2.7: **Lepidopteran U6 promoters appear to permit transcriptional repression in Sf9-dCas9 cells. A.** Schematic representation of the location of template (GFP1) and nontemplate (GFP2) strand-targeting sgRNAs in the *gfp* ORF. **B.** Percentage of the population displaying GFP-positive phenotype and **C.** fluorescence intensity less than the median for the control plasmid encoding no sgRNA. Data is the average of at least 3 independent replicates at 48 hpt. ** denotes $p<0.05$, *** denotes $p<0.01$.

positive cells as well as an overall decrease in fluorescent intensity, however GFP4 was similar to the control (Figure 2.8B & C). To ensure that GFP2 and GFP3 did not show decreased fluorescence because of poor transfection efficiency, plasmids were transfected to parental Sf9 cells (i.e. not expressing dCas9). The results showed that both sgRNA targets displayed similar fluorescence intensity patterns compared to the control in Sf9 cells, suggesting that the decrease in fluorescence intensity is due to the presence of dCas9 and sgRNA (Figure 2.8C).

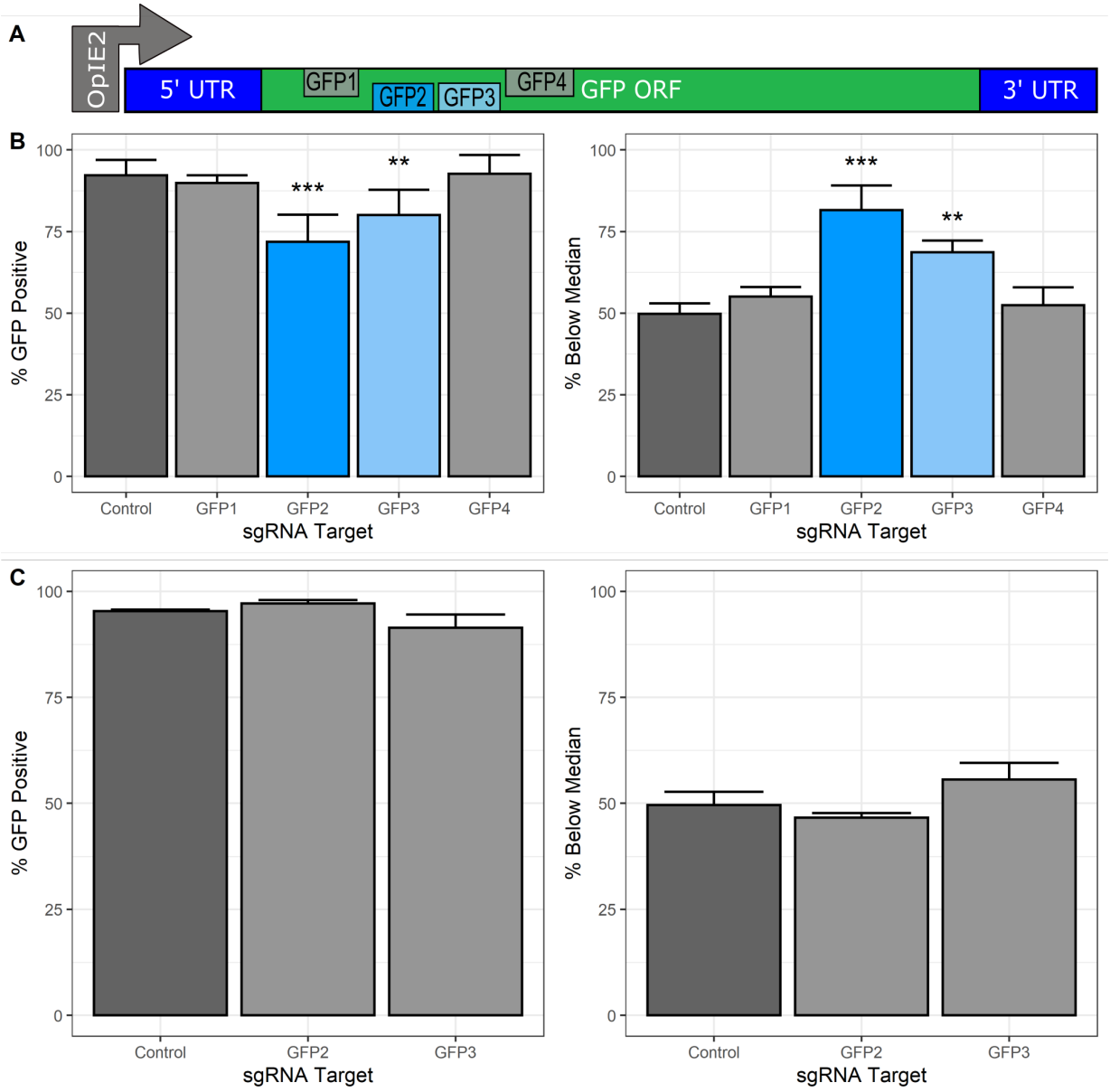


Figure 2.8: **sgRNAs targeting the nontemplate strand permit transcriptional repression but template strand targets do not. A.** Schematic representation of the 4 sgRNAs targeting the *gfp* ORF. **B.** Percentage of the population displaying GFP-positive phenotype (left panel) and fluorescence intensity lower than the median for the control (right panel) for in Sf9-dCas9 cells and. **C.** parental Sf9 cells. Data is the average of at least 3 independent replicates at 48 hpt. ** denotes $p<0.05$, *** denotes $p<0.01$.

Chapter 3

Efficient scrutiny of AcMNPV genes using CRISPR-Cas9

The development of CRISPR-mediated gene disruption technology described in Chapter 2 represented a substantial improvement for evaluating AcMNPV gene function over current platform technologies. First, our approach alleviated the burden of generating the gene knockout bacmid in *E. coli*, which requires several experimental steps to produce and select for the mutant bacmid followed by subsequent removal of the sequences used for selection. Since generation of the knockout virus is done in parental Sf9 cells (ie., not expressing Cas9), disrupting essential genes is also simplified; generating a *trans*-complementing cell line to allow initial production of infectious virus is also no longer required. However, the sgRNA required for sequence-specific gene disruption was delivered from the rBEV itself and represents an inefficiency in the strategy; new rBEVs are required to assess different sgRNA targets. Further improvement would require episomal expression of the sgRNA. Therefore, we devised a transfection-infection assay approach in which the sgRNA was delivered to the cell through transient transfection followed by infection with a rBEV.

In this way, assessing genomic targets required only a new plasmid, and results could be generated in less than 2 weeks as compared to more than 5 weeks with our previous strategy and potentially several months with the traditional approach. The results were submitted to the journal *ACS Synthetic Biology*. The manuscript is presented below in its original form. Supporting information is provided in Appendix B.

Development of a sensitive assay for efficient scrutiny of AcMNPV genes using CRISPR-Cas9

Mark R. Bruder[1] and Marc G. Aucoin[1]

[1] Department of Chemical Engineering, University of Waterloo.

3.1 Abstract

The baculovirus expression vector system is an established platform for production of recombinant proteins. In addition, owing to its large capacity for genetic insertions, it is increasingly used for transfer of large, multi-gene cargo to mammalian cells for a wide array of applications. Baculoviruses have very large genomes and previous studies have demonstrated improvements in recombinant protein production and genome stability through removal of some nonessential sequences. However, recombinant baculovirus expression vectors (rBEVs) in widespread use remain virtually unmodified. Traditional approaches for generating knockout viruses (KOVs) require several experimental steps to remove the

target gene prior to generation of virus. In order to optimize rBEV genomes by removing nonessential sequences, more efficient techniques for establishing and evaluating KOVs are required. Here, we have developed a sensitive assay utilizing CRISPR-Cas9-mediated gene targeting to examine the phenotypic impact of disruption of endogenous AcMNPV genes. For validation, 13 AcMNPV genes were targeted for disruption and evaluated for production of GFP and progeny virus. This assay represents an efficient strategy for identifying essential and nonessential AcMNPV genes through targeted disruption, and represents a valuable tool for developing an optimized rBEV genome.

3.2 Introduction

The Baculovirus Expression Vector System (BEVS) is an insect-specific manufacturing platform for expression of recombinant proteins for diverse applications in basic research and clinical therapeutics [84, 243]. The most commonly used BEVS format is recombinant *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV; rBEV) and cultured cells from a permissive host *Spodoptera frugiperda* (Sf9 cells) [84]. This system has been used to produce a multitude of recombinant proteins, antigens for vaccination, and viral vectors, and a growing number of BEVS products have received regulatory approval for use in human and veterinary applications [87, 238, 243]. Additionally, rBEVs have been increasingly used for gene transfer to mammalian cells [250, 252, 415].

A major milestone in BEVS biotechnology was the development of an AcMNPV genome that could be propagated as a bacterial artificial chromosome (BAC) in *E. coli*, known as bacmid technology [69]. This bacmid was modified to carry sequences necessary for

propagation and selection in *E. coli* and an attTn7 site to allow for insertion of DNA sequences by Tn7 transposition. Since then, various improvements have been made to the bacmid, including alternative or additional insertion sites to accept multiple transgenes [73, 416–418], and recently the construction and rescue of a synthetic AcMNPV produced by PCR and homologous recombination in yeast cells was reported [419]. Coupled with phage-derived recombination proteins [382], the AcMNPV bacmid could be manipulated to generate sequence-specific gene knockout viruses (KOVs) to study gene function [386, 388, 420, 421], identify gene deletions that enhanced recombinant protein quantity and quality [78], and improved genome stability [70]. These results indicate that optimization of the AcMNPV genome by removing nonessential sequences may be an attractive strategy for improving the BEVS platform [131]. Despite this, BEV genomes in widespread use remain virtually unmodified and their systematic optimization has remained a persistent yet elusive goal.

Bacmid technology has undeniably improved the efficiency of generating sequence-specific gene disruptions compared to traditional approaches [131]. However, generation of KOVs in *E. coli* requires multiple experimental steps, and is not particularly amenable to generating multiple different KOVs in parallel. As of 2012, KOVs had been constructed for only 40 (~25%) of AcMNPV genes, and a significant proportion of gene annotations still rely on information from *Bombyx mori* NPV (BmNPV), for which a comprehensive set of single-deletion KOVs exists [95–98, 422]. This strategy may be susceptible to errors, however, as virus-host interactions may lead to variability in the phenotypes observed. For example, the AcMNPV *hcf-1* gene is not essential for propagation in Sf9 cells, however its disruption impairs DNA replication in Tn-368 cells [100]. Similarly, *Mamestra brassicae*

NPV shows different protein expression profiles in different infection hosts [99]. Therefore, identification of nonessential baculovirus genes should be scrutinized experimentally directly in the desired permissive host.

The comprehensive KOV library for BmNPV includes 141 individual bacmids each having an unique single-gene disruption [98]. The growth properties of each KOV were evaluated and categorized into 4 phenotypes according to its ability to expand infection and express the *egfp* reporter gene. This study represents an attractive strategy for identifying nonessential genes, however its duplication for the up to ~125 AcMNPV genes for which KOVs have not been reported is a formidable task. Here, a transfection-infection assay was devised in which Sf9 cells expressing the *cas9* gene were transfected with a plasmid encoding a sgRNA to target a desired AcMNPV gene for disruption. These cells were subsequently infected with a rBEV, thus generating a KOV. The phenotypic impact of gene disruption was analyzed by evaluating the expression of *gfp* from the p6.9 or p10 promoter, and measuring infectious virus titer (IVT) in culture supernatants. To evaluate the approach, 13 candidate AcMNPV genes were selected as targets for disruption. Knockout mutants have been reported in the literature previously for each of these genes. This system correctly differentiated between nonessential and essential genes, consistent with prior reports. Significantly, since targeting any AcMNPV gene requires only the generation of a new plasmid for sgRNA expression, this strategy may be more efficient compared to the bacmid approach, and is amenable to scrutiny of many genes in parallel. Therefore, this strategy may be useful for efficient experimental scrutiny of AcMNPV genes and allow for annotation of the phenotypic impact of targeted gene disruption, leading to the further optimization of the rBEV genome. Finally, this approach should be broadly appli-

cable to other baculoviruses for which genome sequence data and methods for producing recombinant viruses exists [95, 98, 423–425].

3.3 Results and Discussion

The major goal of this study was to establish an efficient process for systematic optimization of rBEV genomes. Herein, we describe the successful application of CRISPR-Cas9 technology for efficient scrutiny of the effect of AcMNPV KOVs on viral late gene expression and progeny virus production. This technique is performed in multi-well plates, generates results in a few weeks and is amenable to screening many genes simultaneously. Additionally, this approach should be broadly applicable for scrutinizing gene function and genome optimization for the more than 60 baculoviruses for which genome sequence data exists [95].

3.3.1 Preliminary screening of transfection and infection conditions

Initial efforts established the process parameters that yielded consistent transfection efficiency and IVT. Parameters that were considered for the transfection included the transfection reagent (TR), TR concentration, and plasmid DNA concentration. For the infection, viable cell density (VCD), multiplicity of infection (MOI) of the rBEV, and time of harvest (TOH) were evaluated. Finally, the time interval between transfection and infection was optimized.

To establish the transfection reagent (Escort IV, FuGENE HD, or GenJet Plus) and transfection conditions (DNA:cell and TR:DNA) that yielded the highest percentage of GFP-positive cells, Sf9-Cas9 cells were seeded in tissue culture treated 6-well plates and transfected with 1.25 pg/cell plasmid (pOpIE2GFP) DNA with a 3:1 ratio of TR:DNA according to manufacturer's directions. At 4 hours post transfection (hpt), the transfection mixture was replaced with fresh medium in half of the replicate wells for each TR. Cells were harvested at 24 hpt and analyzed via flow cytometry. Transfections using Escort IV and GenJet Plus reagents ranged from ~30–50% GFP-positive with a mean fluorescence intensity of $\sim 1 \times 10^6$ arbitrary units (au), whereas the FuGENE HD reagent yielded 65-75% GFP-positive and $\sim 5 \times 10^6$ au mean fluorescence intensity (Figure 3.1). Exchanging the medium at 4 hpt did not significantly affect transfection efficiency or fluorescence intensity for any of the TRs. FuGENE HD was selected as the transfection reagent for further experiments.

To investigate whether further improvements in transfection efficiency could be realized, a face-centered central composite design of experiment (DoE) using FuGENE HD was conducted to analyze the effect of the variables held constant in the previous experiment. Three factor levels were chosen for both variables: 0.5-1.5 pg DNA:cell and 2:1-6:1 TR:DNA (Table 3.1). The lowest DNA:cell ratio yielded the lowest transfection efficiencies although there was slight improvement with increasing TR:DNA ratios. Similarly, the mid and high DNA:cell conditions appeared to follow this trend with increasing TR:DNA ratios however the mid and high TR:DNA conditions showed similar percent-positive and mean fluorescence intensity values, possibly indicating a maximum had been reached with the conditions tested or cytotoxicity associated with FuGENE HD itself.

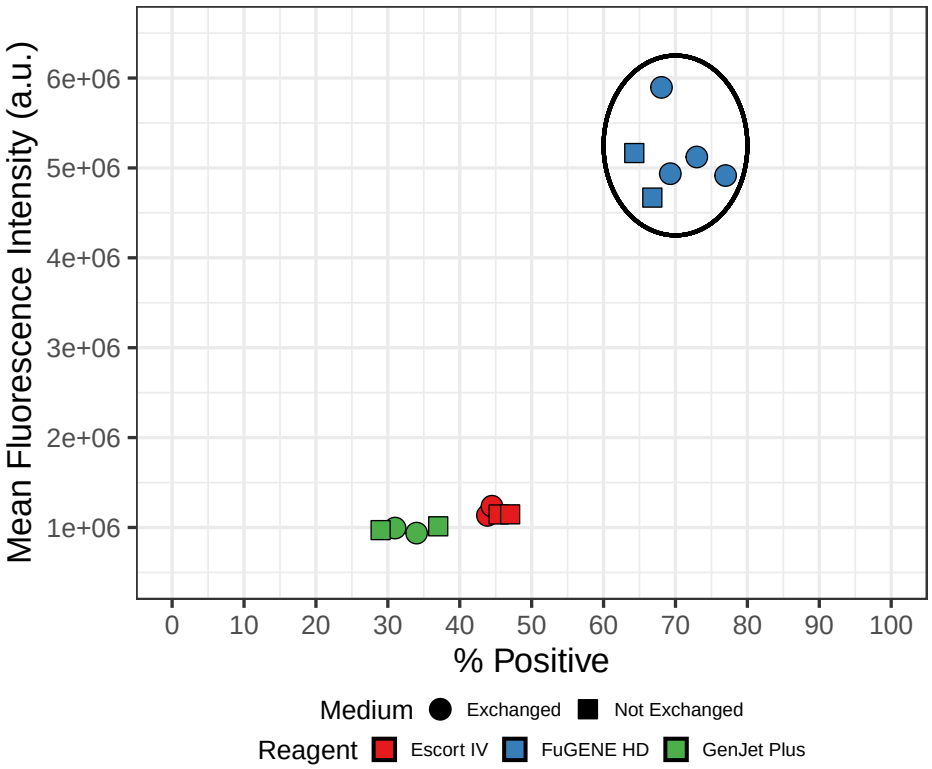


Figure 3.1: **Transfection of Sf9-Cas9 cells with FuGENE HD yields the highest transfection efficiency.** %-GFP positive vs Mean fluorescence intensity for Sf9-Cas9 cells transfected using different transfection reagents.

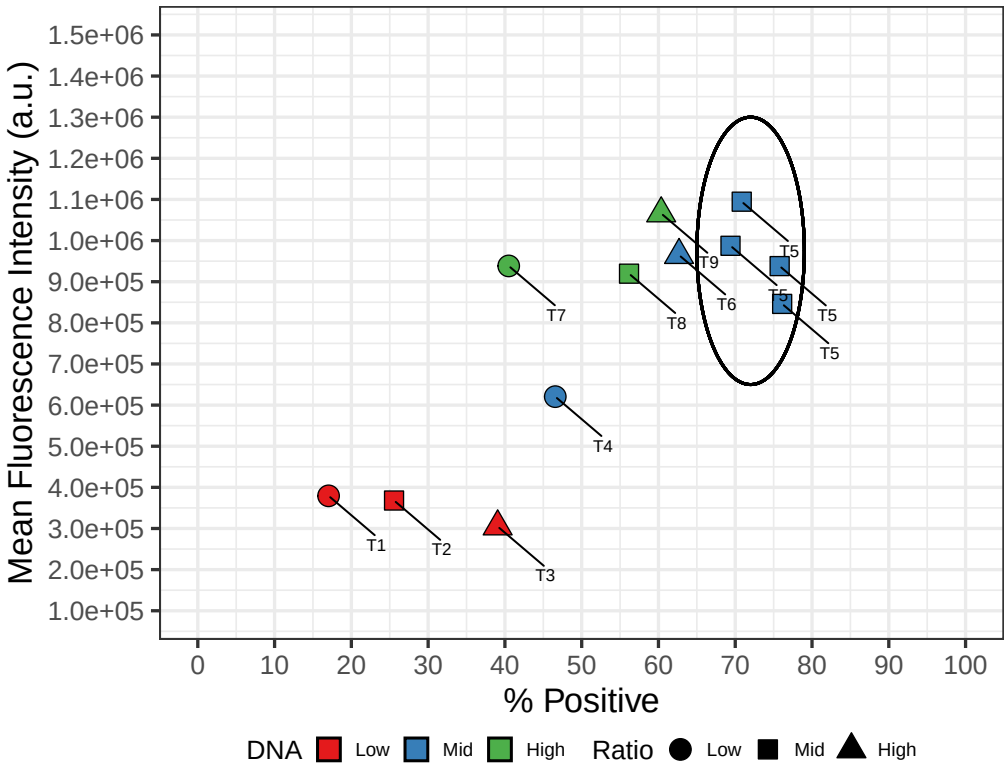


Figure 3.2: **Improvement of transfection efficiency with FuGENE HD.** %-GFP positive vs Mean fluorescence intensity for Sf9-Cas9 cells transfected with different amounts and ratios of DNA and FuGENE HD transfection reagent.

Table 3.1: Experimental conditions screened for selection of transfection protocol

Label	DNA:cell	TR:DNA
T1	0.5	2:1
T2	0.5	4:1
T3	0.5	6:1
T4	1.0	2:1
T5	1.0	4:1
T6	1.0	6:1
T7	1.5	2:1
T8	1.5	4:1
T9	1.5	6:1

To establish infection conditions, a Box-Behnken DoE approach was implemented to ensure reproducibility and ensure high yield of progeny virus ($> 2 \times 10^8$ pfu/ml infectious viral titer, IVT). The parameters selected for analysis were VCD ($\sim 1.5 - 3 \times 10^6$ vc/ml), MOI (3-7pfu/cell), and TOH (36-60 hpi) (Table 3.2). For the 1.5×10^6 vc/ml condition, the IVT ranged from $\sim 0.5 - 1.8 \times 10^8$ pfu/ml and increased with increasing MOI and TOH. Similar results were observed for the medium (2.25×10^6 vc/ml) and high (3×10^6 vc/ml) cell density conditions, however IVT surpassed $\sim 2 \times 10^8$ pfu/ml for all MOIs and TOH later than 48 hpi at these cell densities (Figure 3.3). The fitted second-order response surface identified both 1st and 2nd order terms as being significant for VCD and TOH (data not shown). Since the objective was to find conditions that yielded $> 2 \times 10^8$ pfu/ml rather than a maximum, the conditions that were chosen for further experiments were VCD = $\sim 2.5 \times 10^6$ vc/ml, TOH = 48 hpi, and MOI = 3 pfu/ml.

Table 3.2: Experimental conditions screened for selection of infection protocol

Label	VCD ($\times 10^6$)	TOH (hpi)	MOI (pfu/cell)
I1	1.5	36	5
I2	3	36	5
I3	1.5	60	5
I4	3	60	5
I5	1.5	48	3
I6	3	48	3
I7	1.5	48	7
I8	3	48	7
I9	2.25	36	3
I10	2.25	60	3
I11	2.25	36	7
I12	2.25	60	7
I13	2.25	48	5
I14	2.25	48	5
I15	2.25	48	5

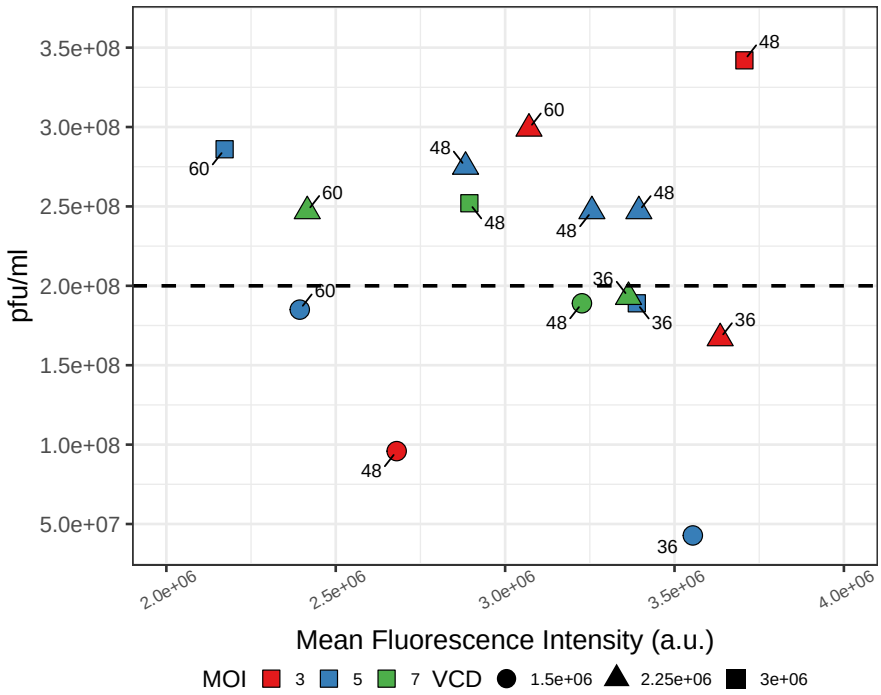


Figure 3.3: **Establishing infection parameters by DoE.** Mean fluorescence intensity vs IVT for Sf9-Cas9 cells infected with a GFP-producing rBEV to observe the effect of MOI, TOH, and VCD. The dashed line indicates the target IVT of $\sim 2 \times 10^8$ pfu/ml.

Finally, to ensure that the transfection itself did not have a significant effect on the infectivity of the cells, an experiment was conducted in which cells were transfected under the conditions described above with either a control plasmid encoding a sgRNA with a scrambled spacer sequence (Control), TR but no plasmid DNA (Mock), or with sterile water in place of both the TR and plasmid DNA (Infect). At 24 hpt the medium was aspirated and replaced with fresh medium containing rBEV (MOI = 3). The infected cells were harvested at 48 hpi and analyzed by flow cytometry and the IVT was measured using end-point dilution assay (EPDA). Both of the transfected conditions (Control and Mock) had similar mean fluorescence intensity and IVT values to the infected only condition (Figure 3.4), indicating that transfection did not have a negative impact on the infectivity of the cells. It is worth noting that visual inspection under the microscope indicated that there was more cell growth in the Infect condition than Mock or Control (data not shown), suggesting that the transfection reagent may have an impact on cell growth.

3.3.2 Evaluation of the time interval between transfection and infection and preliminary evaluation of assay performance

The final experimental condition to establish was the time interval between transfection and infection. Sf9-Cas9 cells were transfected with either a control sgRNA plasmid or a plasmid with the sgRNA targeted to the AcMNPV *ie-1* gene. The medium was aspirated in replicate wells after various time intervals (4, 8, 16, or 24 hpt) and replaced with fresh medium containing rBEV. Each condition was allowed to incubate with virus for an additional 48 hours after which cells and supernatant were harvested for analysis via flow

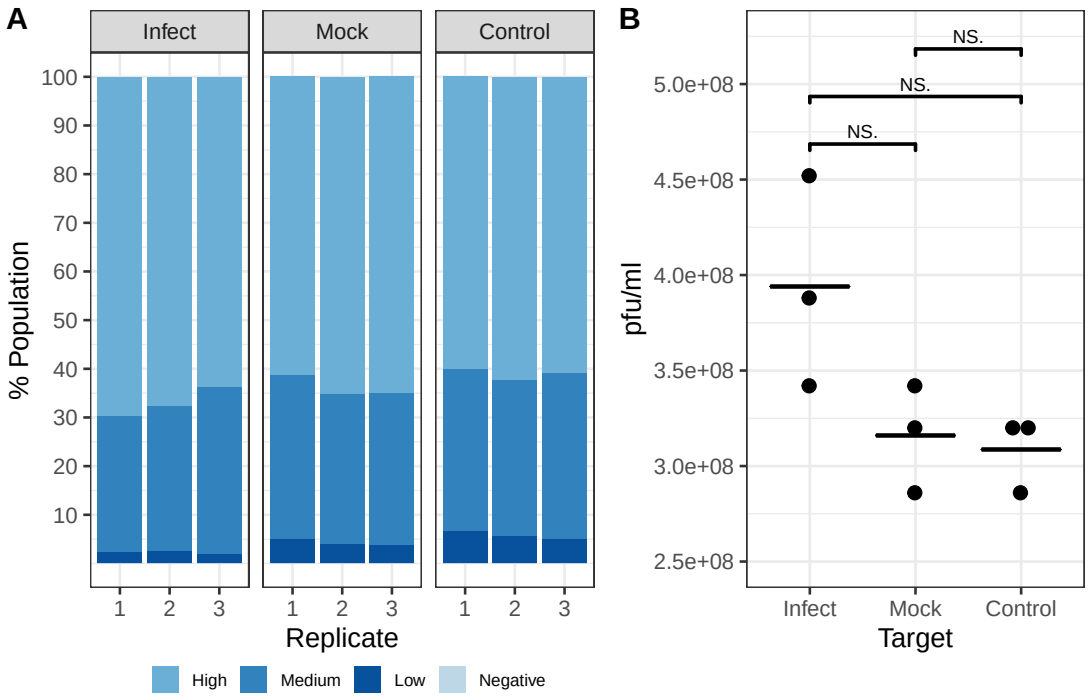


Figure 3.4: **Transfection does not have any impact on infection.** Sf9-Cas9 cells were transfected with reagent only (no DNA; Mock) or a control plasmid encoding an untargeted sgRNA (Control) and compared with cells that were not transfected (Infect). Following infection, **A.** fluorescence and **B.** IVT were analyzed and compared.

cytometry and EPDA, respectively. After applying gates to remove debris and intrinsic cellular fluorescence, observations were binned according to fluorescence intensity: negative (FL1-H < intrinsic cellular fluorescence); low (FL1-H < 10^5 au); medium (10^5 au $\leq$ FL1-H < 10^6 au) and high (FL1-H $\geq$ 10^6 au). The percentage of observations in the low bin was similar for the control sgRNA across each timepoint however there was an increase in the percentage of observations in the low bin and concomitant decrease in the high bin for the *ie-1* target that was apparent in the 16 hpt and 24 hpt conditions (Figure 3.5). Similarly, there was a statistically significant difference in progeny virus between control and *ie-1* targets in the 16 hpt and 24 hpt conditions. These results indicate that the time interval between transfection and infection should be 16-24 hours.

3.3.3 Evaluation of assay sensitivity

After establishing reproducible transfection efficiency of ~70-75% and IVT ($> 2 \times 10^8$ pfu/ml), 13 endogenous AcMNPV genes for which prior detailed deletion mutant studies have been conducted were scrutinized (Table S3). Additionally, sgRNAs targeting the *gfp* reporter gene expressed from the rBEV and untargeted control plasmid were analyzed. Targeting the *gfp* gene resulted in decreased fluorescence but similar IVT to the control, while targeting nonessential genes *ac68*, *odv-e66*, and *p74* yielded similar fluorescence intensity distribution and IVT profiles as the control (Figure 3.6). Targeting the structural genes *gp64*, *orf132*, *vp39*, and *vp80*, on the other hand, yielded decreased progeny virus production compared to the control only (Figure 3.7). Conversely, *dnapol*, *ie-1* and *lef-3* targets resulted in decreased fluorescence intensity and IVT (Figure 3.8). To examine the

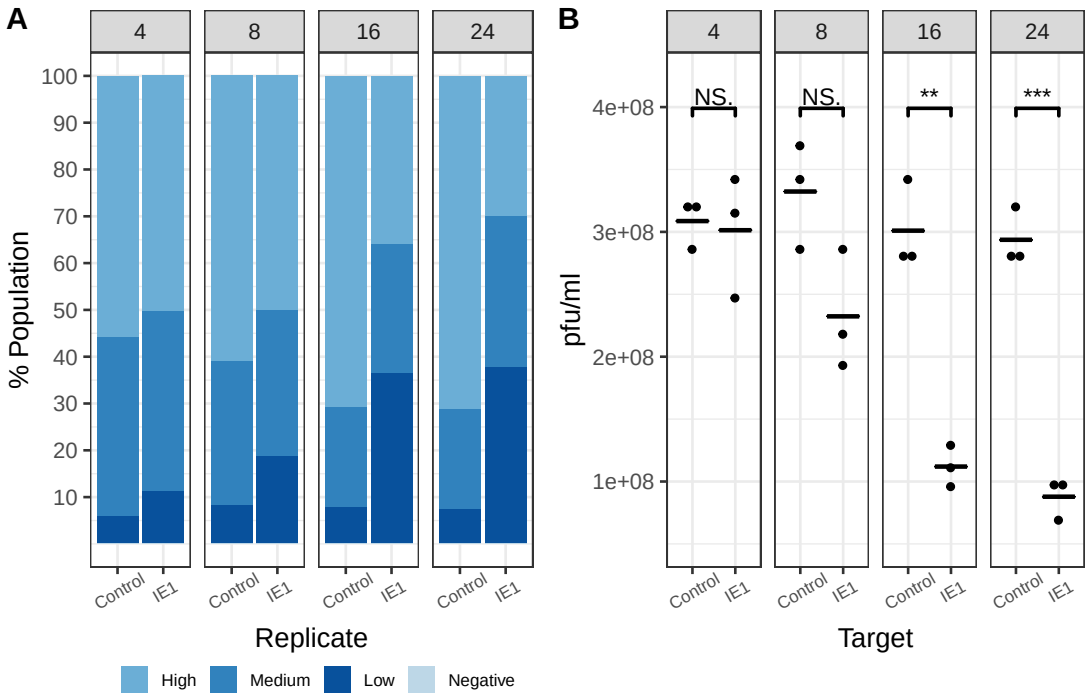


Figure 3.5: **The time interval between transfection and infection affects suppression of GFP and IVT.** Sf9-Cas9 cells were transfected with control and targeted sgRNA plasmids and infected with a GFP-producing rBEV following incubation for different time intervals. At 48 hpi, samples were harvested and analyzed for **A.** fluorescence and **B.** IVT.

effect of spacer sequence location on the KOV phenotype, 2 additional spacer sequences were selected to target the *lef-3* gene at increasing distance from the 5' end of the ORF. Interestingly, the phenotype observed was less distinguishable from the control or nonessential targets for spacer sequences closer to the 3' end of the *lef-3* ORF (Figure S1). This result is consistent with a previous study that reported that amino acids 1-125 of Lef-3 were required for efficient stimulation of DNA replication and late gene expression. Truncated Lef-3 protein that included the required amino acids also resulted in decreased expression compared to the full length protein [426].

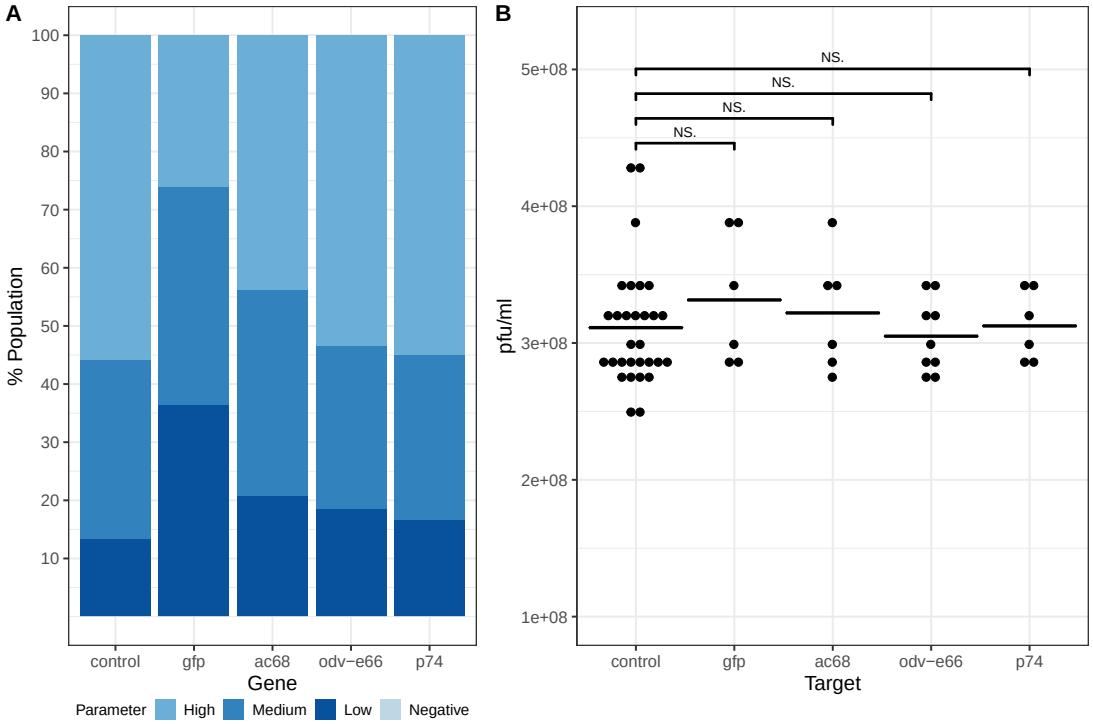


Figure 3.6: Nonessential gene targets show similar **A.** fluorescence intensity and **B.** IVT compared to the control while targeting *gfp* reduces fluorescence only.

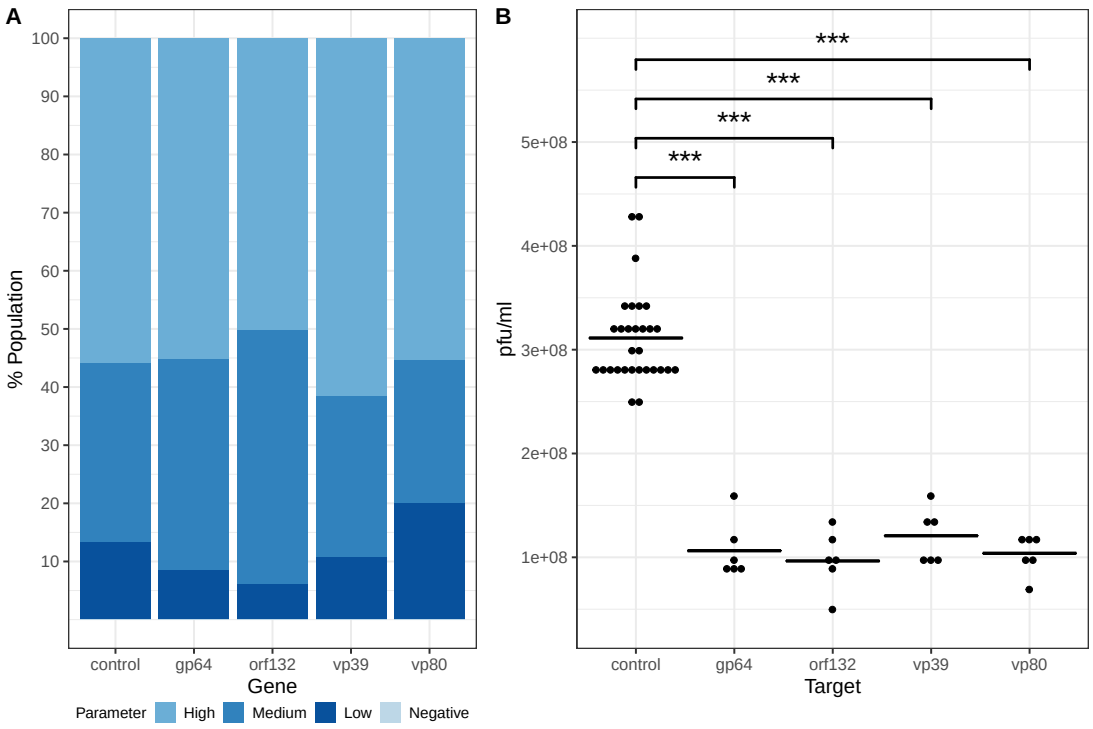


Figure 3.7: Targeting structural genes does not significantly affect **A**. fluorescence but reduces **B.** IVT

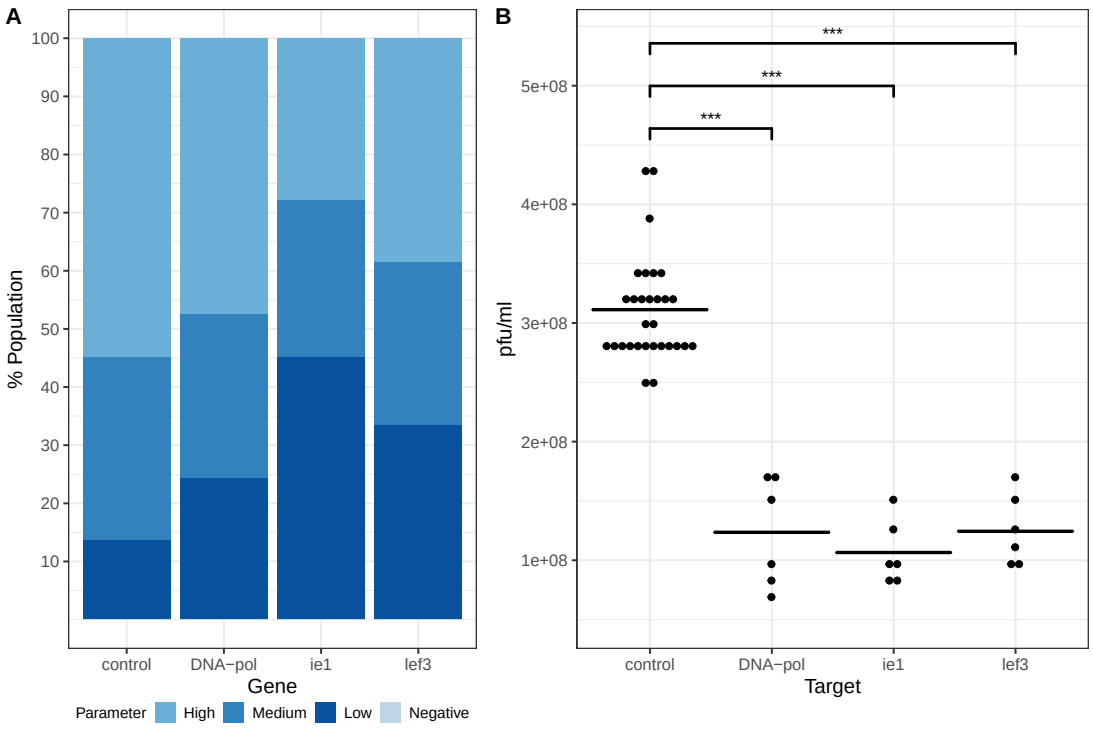


Figure 3.8: KOVs with disruptions in genes necessary for DNA replication reduces both **A.** fluorescence and **B.** IVT.

Next, spacer sequences were selected to target the *p6.9*, *pk-1*, and *vlf-1* genes. These targets were analyzed for their effect on progeny virus production and gene expression from late (p6.9) and very late (p10) promoters (Figure 3.9 and S2). Previous experiments have indicated that these genes have an impact on viral DNA replication, transcription, or are involved in the hyperexpression of the very late genes *p10* and *polyhedrin*. The Vlf-1 protein functions as a very late gene transcriptional activator, and its disruption resulted in impaired viral DNA replication and assembly into virions. Additionally, transcription of the very late genes was greatly reduced but the late class was not affected [385]. Our results indicated that targeting the *vlf-1* gene yielded results in broad agreement with this previous report: reduced fluorescence intensity with the p10GFP rBEV but the p6.9GFP rBEV was similar to control targets. Further, there was a statistically significant reduction in IVT compared to the control for both p6.9GFP and p10GFP rBEVs.

Similarly, the Pk-1 protein is required for progeny virus production and regulates the expression of the very late genes through its kinase activity and as a transcription initiation complex protein [427, 428]. Disruption of the *pk-1* gene did not impact DNA replication nor significantly impact late gene expression but severely reduced very late gene expression [428]. Our results indicate that *pk-1* disruption resulted in a significant impairment in expression from the p10 promoter and progeny virus production, however expression from the p6.9 promoter was also impaired in contrast to previous reports.

The P6.9 protein is involved in viral DNA packaging and virion formation and contributes to transcription regulation. Notably, P6.9 is one of the substrates of phosphorylation by PK-1 [429–431]. Our results indicated a moderate impact on *gfp* expression from both the late p6.9 promoter and very late p10 promoter and statistically significant

decrease in IVT for both rBEVs. These findings agree with those reported previously: although AcMNPV p6.9-null and p6.9 mutant deficient in Pk1-dependent hyperphosphorylation resulted in drastic reductions in IVT and very late (polh and p10) gene expression, only the p6.9-null mutant resulted in reduced expression of the late *vp39* gene compared to wildtype AcMNPV [429–431].

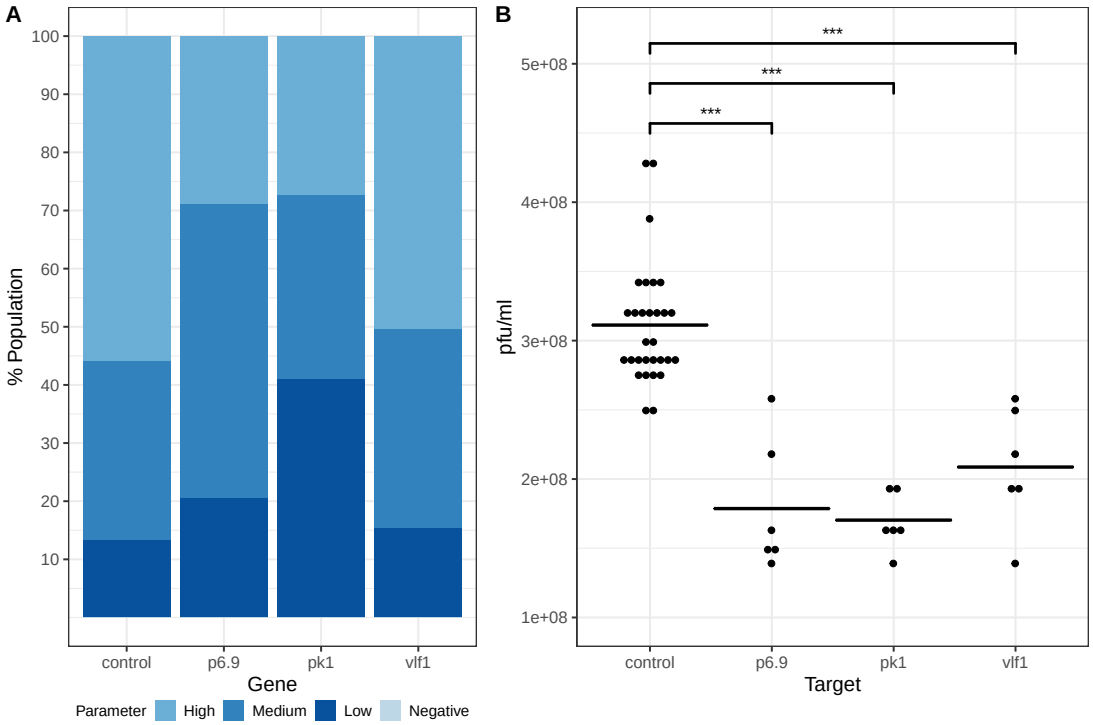


Figure 3.9: Disrupting non-structural auxillary genes primarily involved in very late gene expression and genome packaging has variable effects on **A.** GFP production from the late p6.9 promoter but results in significant reductions in **B.** IVT.

Finally, the assay was tested in parental Sf9 cells (ie., not expressing Cas9) with 10 randomly selected targeting plasmids in triplicate. The results showed fluorescence in-

tensity and IVT similar to the control (untargeted) sgRNA experiments in Sf9-Cas9 cells (Figure S3). Taken together, this indicates that reduction in fluorescence or IVT is due to the presence of both Cas9 and sgRNA, and there is no effect when either component are absent.

3.4 Conclusion

The focus of this study was to develop an efficient method for identifying essential and nonessential baculovirus genes. To this end, a microplate-based assay using CRISPR-Cas9 mediated gene disruption was developed to evaluate the effect of endogenous AcMNPV gene disruptions on late gene expression and progeny virus production. Several process parameters for transfection and infection conditions were assessed, and the assay was used to evaluate the effect of targeted disruption of AcMNPV genes for which there are previous reports of gene disruption in the literature. The results obtained were consistent with those reports for all targets. Importantly, this method is amenable to evaluating several genes with minimal experimental steps and in parallel. This assay is an efficient method for identifying essential and nonessential baculovirus genes, and should be broadly applicable to the study of other baculoviruses.

3.5 Methods

3.5.1 Cells and culture conditions

Sf9 and Sf9-Cas9 cells were maintained in suspension culture in Gibco SF900 III serum free medium (Fisher Scientific, Whitby ON) in borosilicate glass Erlenmeyer flasks (VWR, Mississauga ON) in a non-humidified 27 °C incubator equipped with an orbital shaker at 130 rpm. Puromycin (5μg/ml; Sigma-Aldrich, Oakville ON) was routinely added to the Sf9-Cas9 culture to ensure expression of the *cas9* gene.

3.5.2 Plasmid construction

All plasmids used in this study were constructed using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, Whitby ON) according to manufacturer's directions. Primers used for construction of all plasmids were synthesized by Integrated DNA Technologies (IDT; Coralville, IA) and are listed in Table S1.

To construct plasmid p10GFP, sequences upstream and downstream of the AcMNPV *p10* gene was amplified from AcMNPV genomic DNA and assembled in a Gibson assembly reaction with the coding sequence for a green fluorescent protein (*gfp*) [399]. To construct plasmid p6.9GFP, the promoter region of the AcMNPV *p6.9* gene was amplified from AcMNPV genomic DNA and inserted in place of the p10 promoter in plasmid p10GFP. Plasmid OpIE2GFP was constructed by replacing the *ie-2* ORF from plasmid pOpIE2E2.3 [397] with the (*gfp*) gene.

To construct plasmids expressing the sgRNAs, the SfU6-sgRNA was PCR-amplified from p6.9GFP-sgRNA and assembled with a PCR-amplified fragment consisting of an ampicillin resistance gene (*ampR*) and pBR322 origin of replication (ori) for propagation in *E. coli*. For retargeting of sgRNAs, primers encompassing the spacer sequence and either sequence homologous to the 5' end of the cas9 handle or 3' end of the SfU6 promoter sequence were used to re-generate the full length, linear pSfU6-sgRNA plasmid with altered spacer sequence via PCR. After treatment with *dpnI* restriction endonuclease (New England Biolabs) to remove template DNA, the fragment was gel extracted and re-circularized using Gibson assembly. Spacer sequences for AcMNPV gene targeting were designed using the CHOP-CHOP online platform (https://chopchop.cbu.uib.no/), and spacer sequences were selected based on distance from the 5' end of the ORF, predicted efficiency of targeting, and predicted off-target sites within the AcMNPV genome [432]. The spacer sequences used in this study are given in Table S2.

3.5.3 Transient transfection

Sf9-Cas9 cells were seeded on tissue culture treated 6-well plates (Fisher Scientific) at the desired density and transfected using Escort IV (Sigma Aldrich), FuGENE® HD (Promega, Madison WI), or GenJet™ Plus (FroggaBio, Concord ON) transfection reagents according to manufacturer's directions. After incubating at 27 °C, cells were harvested, fixed in 2% paraformaldehyde (Fisher Scientific) prior to further analysis.

3.5.4 Recombinant baculovirus generation, amplification, and quantification

Plasmids p6.9GFP or p10GFP were co-transfected with flashBACGOLD™ (Oxford Expression Technologies Ltd., Oxford UK) genomic DNA in to Sf9 cells seeded in tissue culture treated 6-well plates according to manufacturer's directions. Supernatant from each transfection was harvested 4-5 days post transfection and used to infect early-exponential phase ($\sim 1.5 - 2 \times 10^6$ cells/ml) suspension Sf9 cultures at low multiplicity of infection (MOI) to amplify the rBEVs for 3-4 days or until the viable cell density dropped to ~85-90%. After 2 sequential rounds of amplification, the rBEV titer was quantified using end-point dilution assay (EPDA). Briefly, Sf9 cells were seeded to each well of a 96-well plate (Fisher Scientific) at a density of 2.0×10^4 cells/ml. Separately, the rBEV was serially diluted (10^{-2} to 10^{-8}) in fresh SF900 III medium and 10μl of each dilution was added, in 12 replicates, to the 96-well plate. Plates were incubated for 6-7 days at 27 °C, after which wells were scored according to visualization of green fluorescence using a fluorescence microscope. Results were converted from $TCID_{50}$ and reported as plaque forming units per ml (pfu/ml).

3.5.5 Infections

Sf9-Cas9 cells were seeded on tissue culture treated 6-well plates at various densities and infected with the p6.9GFP rBEV diluted to the desired multiplicity of infection (MOI) in fresh SF900 III medium. After incubation at 27 °C for 36-60 hours post infection (hpi), cells were harvested and fixed with 2% paraformaldehyde prior to further analysis.

3.5.6 Transfection-infection assay

Sf9-Cas9 or parental Sf9 cells were seeded on tissue culture treated 6-well or 12-well plates (Fisher Scientific) at a density of ~2.5×10^6 cells/well (6-well) or ~9.0×10^5 cells/well (12-well). After allowing cells to attach for ~1 hour, the cells were transfected with Fugene HD transfection reagent according to manufacturer's directions. After incubation at 27 °C for 16-24 hours post transfection (hpt), the medium was aspirated and replaced with fresh SF900 III medium containing rBEV (MOI = 3). The infected wells were incubated for a further ~48 hours after which cells and supernatant were harvested for further analysis.

3.5.7 Flow cytometry and analysis

Samples were analyzed using a BD Accuri™ C6 Plus flow cytometer (BD Biosciences, San Jose CA) equipped with a blue laser with an excitation frequency of 488 nm and 510/15 nm band-pass filter. Samples were run at the low flow setting (14μl/min) and 10000 events were collected for analysis using the CSampler Plus autosampler. Analysis of flow cytometry data was performed using FlowJo® V10 flow cytometry analysis software (FlowJo LLC, Ashland, OR). For transfection-infection experiments, fluorescence intensity observations were binned and analyzed according to the proportion of events in each bin. The bins were defined as: negative (FL1-H < intrinsic cellular fluorescence); low (FL1-H < 10^5 au); medium (10^5 au $\leq$ FL1-H < 10^6 au) and high (FL1-H $\geq$ 10^6 au)

Chapter 4

CRISPR-mediated knockouts are target specific

The phenotypes observed in Chapter 2 were consistent with CRISPR-mediated transcriptional repression and gene disruption of the GFP reporter and target AcMNPV ORFs. However, due to the lack of commercially available antibodies for target AcMNPV proteins, the specificity of knockouts/knockdowns were not confirmed experimentally through analysis of the abundance of target proteins. There are, however, commercially available monoclonal antibodies for the AcMNPV envelope protein GP64. To provide further confirmation that generation of knockout viruses (KOVs) with CRISPR-Cas9 is target specific, rBEVs were constructed to target the *gp64* ORF. In addition to the expected phenotypes of reduced IVT but unaffected late gene expression, analysis of GP64 abundance in the membrane of Sf9 cells via immunofluorescence staining and total cell lysates via western blot was possible. Additionally, the manufacturing of recombinant proteins without contaminating virus in the supernatant is a strategy that has received some attention in the literature. A major drawback of the traditional approach for deleting essential genes,

however, is the relatively lower expression yields as compared to replicative rBEVs and requirement of a *trans*-complementing cell line expressing the essential gene to produce initial virus seed stocks. Results from the vp80-null KOV in Chapter 2 suggested that recombinant protein yield may be less effected with the CRISPR-mediated KOV approach, so rBEVs were constructed to express the HIV-1 Gag protein and target the *gp64* or *vp80* ORF to assess the production of Gag virus-like particles without contaminating rBEVs in the culture supernatant. The results are formatted for forthcoming submission.

Evaluation of virus-free manufacture of recombinant proteins using CRISPR-mediated gene disruption in baculovirus-infected insect cells

Mark R. Bruder[1] and Marc G. Aucoin[1]

[1] Department of Chemical Engineering, University of Waterloo.

4.1 Abstract

The manufacture and downstream processing of virus-like particles (VLPs) using the baculovirus expression vector system (BEVS) is complicated by the presence of large concentrations of baculovirus particles, which are similar in size and density to VLPs, and consequently are difficult to separate. To reduce the burden of downstream processing, CRISPR-Cas9 technology was used to introduce insertion-deletion (indel) mutations within the *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) *gp64* open reading

frame, which encodes the major envelope protein of AcMNPV. After comfirming the site-specific targeting of *gp64* leading to reduced budded virus (BV) release, the *gag* gene of human immunodeficiency virus type 1 was expressed to produce Gag VLPs. This approach was effective for producing VLPs using the BEVS whilst simultaneously obstructing BV release.

4.2 Introduction

Virus-like particles (VLPs) are an emerging class of biotherapeutic modality for delivery of therapeutic cargo such as chemotherapy, protein, and nucleic acid-based drugs, and as antigens for vaccination [433, 434]. VLPs are highly ordered structures that typically self-assemble from a single or multiple viral structural proteins to mimic the three-dimensional structure of the natural virus from which the structural proteins are derived. Additionally, VLPs may be enveloped or nonenveloped, and are replication/infection incompetent, as they lack the genetic material of the natural virus. Finally, the particulate structure of VLPs favours uptake by antigen presenting cells and can stimulate robust B cell and T cell-mediated adaptive and innate immune responses [434, 435].

The Baculovirus Expression Vector System (BEVS) has many features that make it an attractive platform for VLP production, including ease of manipulation and large capacity for foreign gene insertion which allows simultaneous expression of multiple proteins from the same recombinant BEV (rBEV) [87]. As such, the BEVS is a preferred platform for production of VLPs, and a multitude of studies have reported successful production of VLPs that mimic many enveloped and nonenveloped viruses [87]. Further, several BEVS-

produced VLPs have received regulatory approval for human or veterinary use, or are in various stages of clinical development [84, 234]. Nevertheless, significant process shortcomings must be addressed to realize the full potential of the BEVS for VLP production; large amounts of progeny virus, proteins, and cell debris resulting from the lytic infection cycle contaminate the supernatant, requiring extensive purification steps to achieve pharmaceutical-grade purity for clinical applications. In addition, enveloped VLPs and baculovirus are often similar in size, density, and have the same constituent membrane proteins, further complicating downstream processing [84].

To reduce the burden of baculovirus contamination on downstream processing, strategies have been devised wherein a gene encoding a baculovirus structural protein required for viral genome packaging, nucleocapsid assembly, or release of budded viruses (BV) is deleted from its genome. To enable initial production of infectious virus seed stocks, a *trans*-complementing cell line, in which the deleted gene is constitutively expressed, is required. The mutant rBEV is then used to infect parental cells (ie., not expressing the essential gene) for production of the recombinant protein/therapeutic. Using this approach, the AcMNPV *vp80* and *gp64* genes have been deleted to produce enhanced green fluorescent protein (EGFP) and HIV-1 Gag VLPs, respectively [85, 86]. Although these strategies were successful for reducing the contaminating baculovirus in the supernatant, initial propagation of the rBEV to generate the required viral seed stocks was impaired in both systems, and the overall yield of the recombinant protein from the knockout virus (KOV) may have similarly been affected [85, 86].

Here, a recently developed approach for generating rBEV KOVs using CRISPR-Cas9 was used to target the *gp64* gene for disruption. After confirming that targeting the

gp64 open reading frame (ORF) resulted in decreased GP64 abundance in infected cells, expression of the green fluorescent protein (GFP) reporter gene was assessed. Consistent with previous reports, disruption of *gp64* reduced progeny virus release but did not affect expression of GFP. Next, production of HIV-1 Gag VLPs was demonstrated with this approach. The yield of Gag VLPs was similar for all rBEVs in Sf9-Cas9 cells and Sf9 cells, further indicating that CRISPR-mediated disruption of structural genes may be an effective strategy for reducing BV release while maintaining high expression of foreign genes.

4.3 Materials and methods

4.3.1 Cells and culture conditions

Sf9 and Sf9-Cas9 cells were maintained in suspension culture in Gibco SF900 III serum free medium (Fisher Scientific, Whitby ON) in a non-humidified 27 °C incubator and shaken at 130 rpm on an orbital shaker. Puromycin (5μg/ml; Sigma-Aldrich, Oakville ON) was routinely added to the Sf9-Cas9 culture to select for expression of the *cas9* gene.

4.3.2 Recombinant baculovirus generation, amplification, and quantification

Transfer plasmids for rBEV generation were co-transfected with flashBACGOLD™(Oxford Expression Technologies Ltd., Oxford UK) genomic DNA to Sf9 cells using Escort IV transfection reagent (Sigma-Aldrich) according to manufacturer's directions. Supernatant from

each transfection was harvested 4-5 days post transfection and used to infect suspension Sf9 cultures (~1.5×10^6 cells/ml) at low multiplicity of infection (MOI) for 3-4 days to amplify the rBEV to higher infectious viral titer (IVT). Following one more round of amplification, the rBEV IVT was quantified using end-point dilution assay (EPDA). Briefly, Sf9 cells were diluted to a density of 2.0×10^5 cells/ml and 100 μl was seeded to each well of a 96-well plate (Fisher Scientific). Separately, the rBEV was serially diluted (10^{-2} to 10^{-8}) in fresh SF900 III medium and 10 μl of each dilution was added, in 12 replicates, to the 96-well plate. Plates were incubated for 6-7 days at 27 °C, after which wells were scored according to visualization of green fluorescence using a fluorescence microscope. Results were converted from $TCID_{50}$ and reported as plaque forming units per ml (pfu/ml).

4.3.3 Infections

Sf9-Cas9 or Sf9 cells were infected with rBEVs at a density of $\sim 1.5 - 2 \times 10^6$ cells/ml viable cells/ml at a MOI of 3 pfu/cell. Samples were harvested at the required times (hours post infection; hpi) wherein cells were centrifuged at $300 \times g$ for 10 minutes and resuspended in 2% paraformaldehyde diluted in phosphate buffered saline (PBS) for ~30 minutes prior to analysis by flow cytometry. The cell culture supernatant was kept at 4 °C and cell pellets for western blotting were frozen at −80 °C.

4.3.4 Western blot

Infected cells (~$1.5 - 2 \times 10^6$ cells/ml) were collected at ~20-24 hpi by centrifugation at $500 \times g$ for 10 min at 4 °C. The cells were lysed in RIPA buffer (Fisher Scientific),

quantified by Pierce BCA assay (Fisher Scientific), and ~10 μg of protein was separated by electrophoresis in 10% TGX Stain-Free precast mini SDS-PAGE gels (Bio-Rad, Mississauga ON) according to manufacturer's directions. After transfer to low fluorescence PVDF membranes, Western blot analysis was performed with anti-GP64 (AcV5, Fisher Scientific) primary antibody and goat anti-mouse IgG HRP secondary (Bio-Rad) and imaged on a ChemiDoc MP Imager (Bio-Rad). The Image Lab software (Bio-Rad) was used for further image processing.

4.3.5 Immunofluorescence

Infected cells (~1×10^6) were collected at ~12-15 hpi or ~48 hpi by centrifugation at $300 \times g$ for 10 min at 4 °C. The cells were washed twice with cold PBS + 0.5% Bovine Serum Albumin (PBS-BSA) and incubated with anti-gp64 (AcV1, Fisher Scientific) conjugated to APC diluted in PBS-BSA (1:1000) for ~30 min on ice. Cells were washed 3 times in PBS-BSA and resuspended finally in 200 μl PBS for analysis by flow cytometry.

4.3.6 Flow cytometry and analysis

Fluorescent cells were acquired using a BD Accuri™ C6 Plus flow cytometer (BD Biosciences, San Jose CA) equipped with 488 nm and 640 nm lasers. Samples were run at the low flow setting and 10000 events were collected and analyzed using FlowJo® V10 flow cytometry analysis software (FlowJo LLC, Ashland, OR).

4.3.7 Quantification of baculovirus particles using flow cytometry

Sample preparation for analysis via flow cytometry was described previously [402]. Briefly, samples were diluted in PBS and fixed with paraformaldehyde for 1 hour, subjected to one freeze-thaw cycle, and incubated with Triton X-100 to permeabilize the membrane. The nucleic acid stain SYBR Green I was added and incubated at 80 °C for 10 min in the dark to stain double stranded DNA (dsDNA). After cooling on ice, the samples were analyzed via flow cytometry. Flow-Set Fluorospheres (Beckman Coulter, Mississauga ON) were used for calibration and all samples were run in triplicate.

4.3.8 Quantification of Gag-VLPs with enzyme-linked immunosorbent assay (ELISA)

The supernatants of Sf9 and Sf9-Cas9 cells infected with Gag-expressing rBEVs were harvested by centrifugation at $1000 \times g$ for 10 min and filter sterilized with a 0.2 μm syringe filter. Gag-VLPs were quantified using the HIV-1 p24 ELISA Kit (Xpress Bio Life Science, Frederick MD) according to manufacturer's directions. The absorbance was measured using a Synergy 4 hybrid microplate reader (BioTek, Winooski, VT) at a wavelength of 450 nm. An HIV-1 p24 protein standard of known concentration was used to calculate the Gag concentration and estimate VLP yield.

4.4 Results

4.4.1 Targeting the *gp64* ORF is site specific

Initial experiments were conducted to confirm that sgRNAs designed to target the *gp64* gene were target-specific and resulted in the disruption of progeny virus release. Accordingly, the abundance of GP64 protein was analyzed by western blot and immunofluorescence staining in the cell membrane. Analysis of cell lysates from infected cells revealed that GP64 present in Sf9-Cas9 cells infected with rBEVs targeting the *gp64* ORF was reduced to ~1% compared to Sf9-Cas9 cells infected with control rBEVs. Parental Sf9 cells infected with the GP64-1 rBEV, on the other hand, showed GP64 levels indistinguishable from the control (Figure 4.1). Detection of GP64 in the plasma membrane of infected cells similarly revealed reduced fluorescence consistent with lower GP64 abundance in Sf9-Cas9 cells but not parental Sf9 cells (Figure 4.2). Taken together, these data indicate the sgRNAs designed to target the *gp64* ORF result in decreased abundance of GP64 protein.

4.4.2 Cas9-mediated disruption of *gp64* impacts progeny virus production but not late gene expression

Sf9-Cas9 cells infected with rBEVs encoding a *gfp* reporter gene transcribed from the late p6.9 gene promoter and sgRNAs targeting the *gp64* gene resulted in significant reduction of infectious viral titer (IVT) at 48 hpi compared to the untargeted control. Specifically, the mean IVT for control rBEVs in Sf9-Cas9 cells was ~2.65×10^8 pfu/ml whereas the IVT for the Δ*gp64* KOV was 4.03×10^6 pfu/ml. Conversely, Sf9 cells infected with the

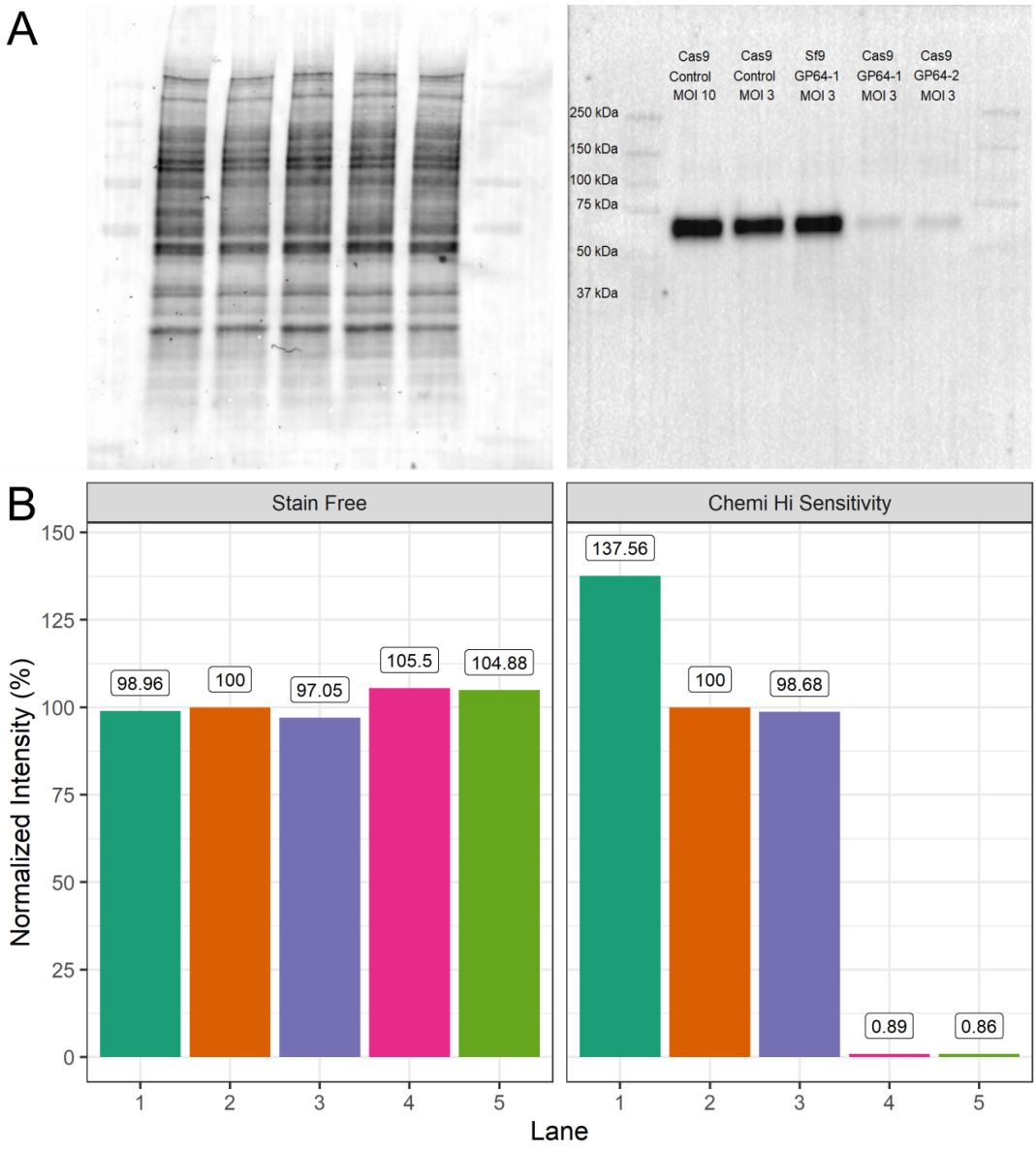


Figure 4.1: **CRISPR-mediated disruption of *gp64* is target specific. A.** Western blot and **B.** total protein normalization data for control and gp64-null KOVs.

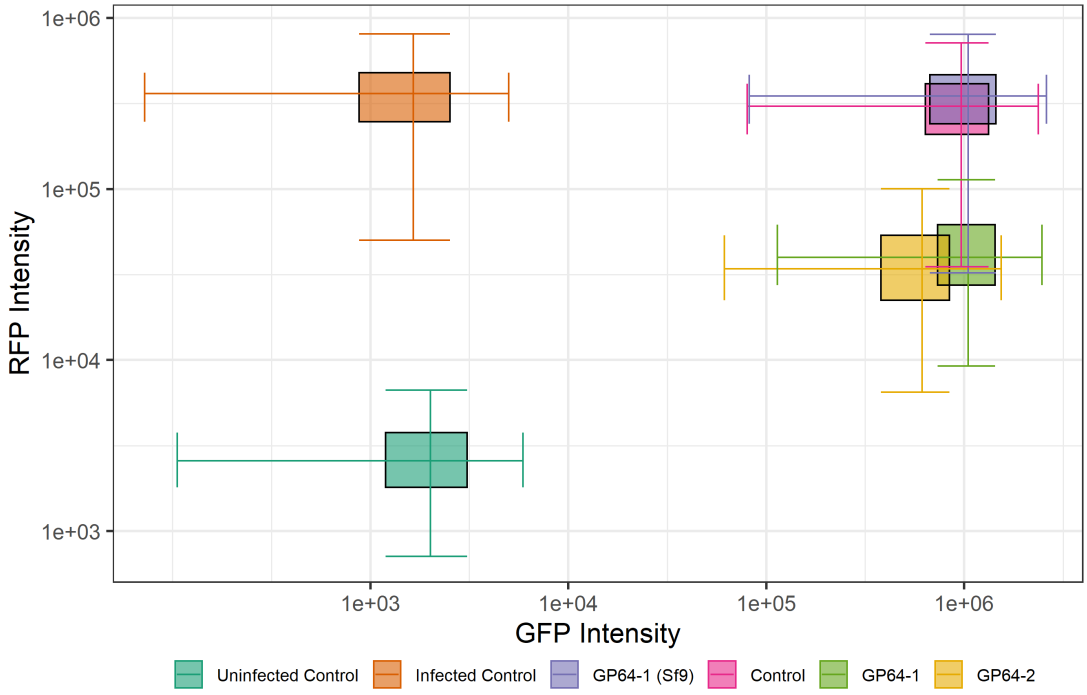


Figure 4.2: **CRISPR-mediated disruption of *gp64* reduces GP64 abundance in the membrane of Sf9-Cas9 cells.** GFP vs RFP fluorescence intensity for control and *gp64*-targeted rBEVs expressing the reporter GFP (x-axis) and stained with APC-conjugated anti-GP64 AcV1 antibody (y-axis). The width and height of the boxes and whiskers represent the IQR and $1.5 \times IQR$ boxplot for GFP and RFP distribution, respectively. Uninfected control: uninfected cells stained with AcV1 mAb; Infected Control: Infected with non-fluorescent control rBEV stained with AcV1; Control: Sf9-Cas9 cells infected with untargeted sgRNA, stained with AcV1; GP64-1/GP64-2: Sf9-Cas9 cells infected with *gp64*-targeted sgRNAs, stained with AcV1; GP64-1 (Sf9): parental Sf9 cells infected with GP64-1 rBEV and stained with AcV1.

same rBEVs yielded IVTs that were indistinguishable from each other (3.03×10^8 pfu/ml and 1.93×10^8 pfu/ml for control and *gp64*-targeting sgRNAs, respectively) and similar to the untargeted control rBEV in Sf9-Cas9 cells (Table 4.1). Analysis of cell culture supernatants at 8-12 hpi yielded IVT of $\sim 2.1 - 6.9 \times 10^4$ pfu/ml for all rBEVs in both cell lines, indicating virus uptake was similar for all rBEVs in Sf9-Cas9 and Sf9 cells (data not shown). Additionally, late gene expression appeared to be unaffected as there were small but insignificant differences in fluorescence intensity between control and *gp64*-disrupted rBEVs in both Sf9 and Sf9-Cas9 cells (Figure 4.3A, B). Finally, to confirm that this approach resulted in significant reduction of total particles in the supernatant as opposed to only IVT, analysis of cell culture supernatants by flow cytometry revealed that particle concentration was reduced ~90% compared to the untargeted control rBEV in Sf9-Cas9 cells (Figure 4.3C). This evidence suggests that CRISPR-mediated disruption of the *gp64* gene resulted in a reduction of particles in culture supernatants but does not significantly impact late gene expression.

Table 4.1: Summary of fluorescence intensity and virus quantification data for rBEVs in Sf9 and Sf9-Cas9 cells at 48 hpi.

rBEV	Sf9-Cas9			Sf9		
	FL. Intensity (au)	IVT (pfu/ml)	Particles/ml	FL. Intensity (au)	IVT (pfu/ml)	Particles/ml
Control	434 ± 3.96	$2.65 \pm 0.59 \times 10^8$	$1.47 \pm 0.76 \times 10^9$	443 ± 6.90	$3.03 \pm 0.74 \times 10^8$	-
GP64	367 ± 5.70	$4.03 \pm 1.89 \times 10^6$	$8.93 \pm 2.16 \times 10^7$	342 ± 13.70	$1.93 \pm 0.65 \times 10^8$	-

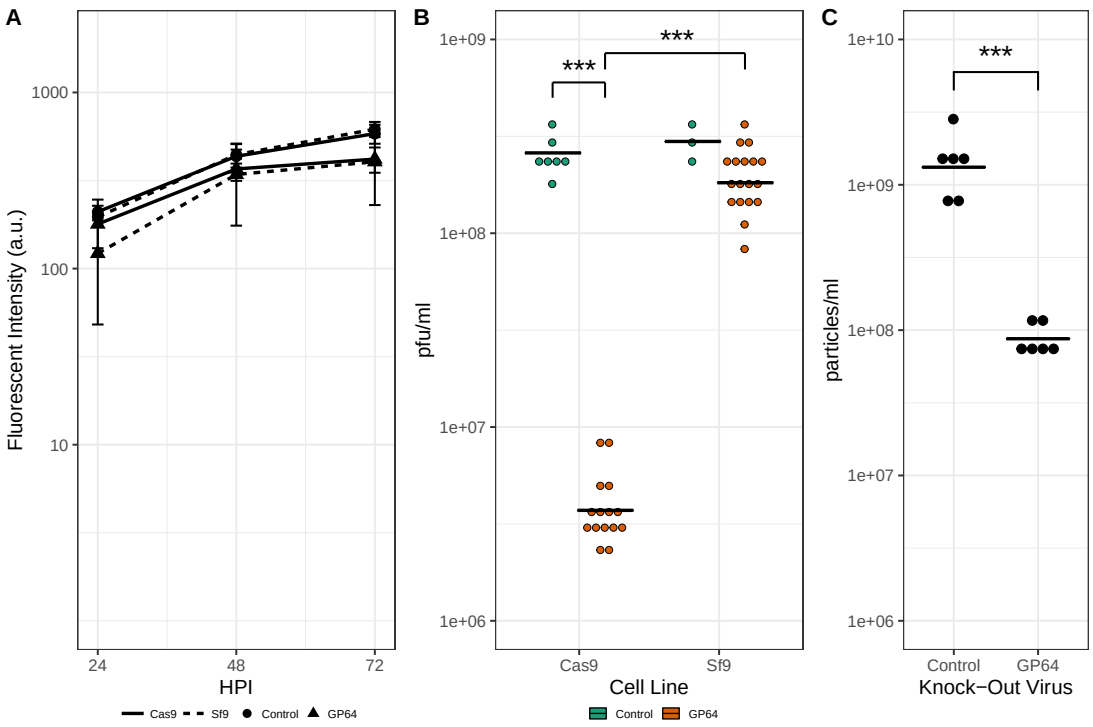


Figure 4.3: **GP64-disrupted KOVs show reduced IVT but unaffected late gene expression.** **A.** Production of GFP, **B.** IVT, and **C.** total particle concentration for control and *gp64*-targeting rBEVs in Sf9-Cas9 and parental Sf9 cells (A and B) or Sf9-Cas9 cells (C). Solid line: Sf9-Cas9 cells; dashed line: Sf9 cells; untargeted control (circles) and *gp64*-targeted (triangles) rBEVs

4.4.3 Production of HIV-1 Gag VLPs

In light of these results, the *gfp* reporter gene was replaced with the HIV-1 *gag* gene to investigate the production of Gag VLPs with this system. In addition to targeting the *gp64* ORF for disruption, rBEVs expressing *gag* and sgRNAs targeting the *vp80* ORF were also prepared. Infecting Sf9-Cas9 cells with rBEVs resulted in ~99% reduction of IVT for rBEVs targeting the *gp64* and ~94% for the *vp80* target (Figure 4.4A) compared to the same infections in Sf9 cells. Similarly, GP64 in the plasma membrane was reduced by ~99% for the *gp64*-targeting sgRNAs. Interestingly, targeting the *vp80* ORF resulted in ~65% reduction in fluorescence intensity compared to control infections in Sf9 cells, indicating that targeting the *vp80* ORF may have an impact on GP64 expression. Finally, quantification of Gag VLPs by ELISA indicated VLP yields of $\sim 2.9 - 4.6 \times 10^9$ particles/ml for all rBEVs in both Sf9 and Sf9-Cas9 cells (Table 4.2).

Table 4.2: Production of HIV-1 Gag VLPs in Sf9-Cas9 and Sf9 cells at 48 hpi.

rBEV	Particles/ml	
	Sf9-Cas9	Sf9
Control	$4.10 \pm 0.00 \times 10^9$	$4.61 \pm 0.00 \times 10^9$
GP64	$3.26 \pm 1.45 \times 10^9$	$2.90 \pm 0.77 \times 10^9$
VP80	$3.64 \pm 2.34 \times 10^9$	$3.49 \pm 2.59 \times 10^9$

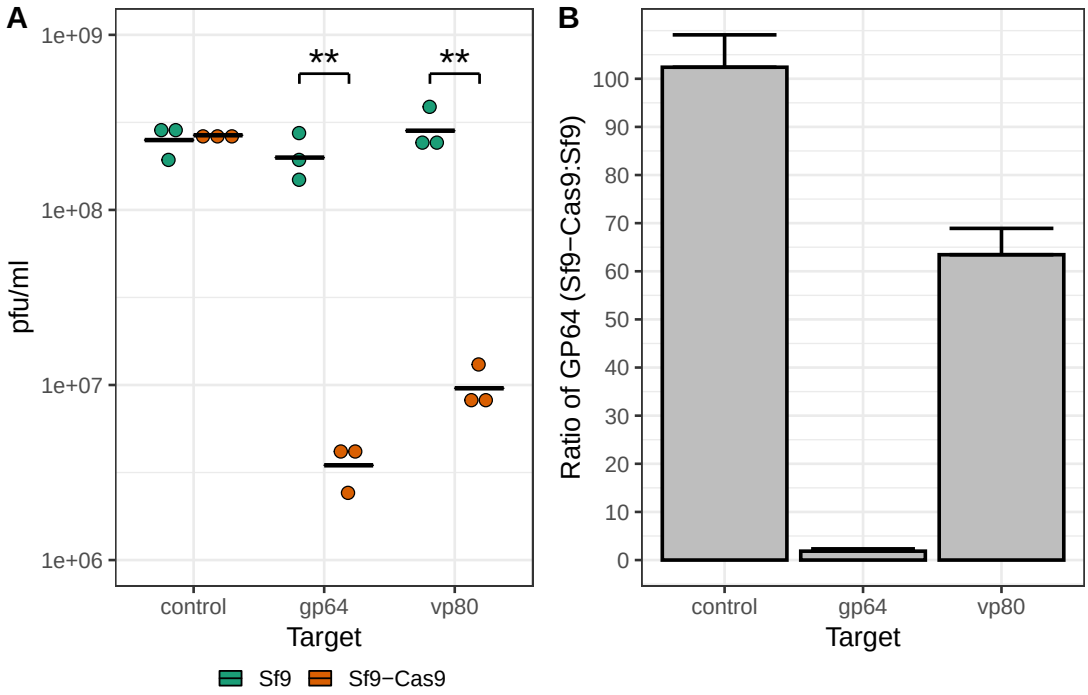


Figure 4.4: rBEVs targeting *gp64* and *vp80* have lower **A.** IVT and **B.** abundance of GP64 in infected cell membranes in Sf9-Cas9 cells than in Sf9 cells.

4.5 Discussion

Virus-like particles represent a promising platform for use as vaccine candidates because they lack the genome but mimic the structure of wildtype viruses. This affords VLPs attractive safety and immunogenicity profiles [436]. As such, VLPs comprised of structural antigens from a wide variety of viruses have been produced in several heterologous expression systems, including plant, mammalian, and insect hosts. Among these, mammalian and insect production platforms remain the most popular, particularly for enveloped VLPs, due to their ability to display authentic and correctly folded antigens with complex post-translational modifications (PTMs). While mammalian production hosts are more closely related to the natural host of the virus and capable of human-like PTMs, the cost of mammalian bioprocesses can be very high and the presence of adventitious agents may be a concern [437]. Insect hosts, and specifically the BEVS, is a scalable production system capable of exceptionally high expression levels, and as very few opportunistic pathogens can propagate in both insect and mammalian cells, safety concerns are generally minimized in this respect [437]. The presence of high concentrations of baculovirus particles that are co-produced along with VLPs in the culture supernatant, however, complicates and increases the cost of the downstream process required to purify the VLP [84].

To address this drawback, strategies have been devised to reduce or eliminate progeny baculovirus production through the targeted deletion of genes encoding structural proteins that are required for BV release, called knockout viruses (KOVs) [84–86]. This strategy requires the development of a *trans*-complementing cell line to enable replication of the rBEV. However, this approach may be less effective for rBEV seed production, and for-

eign gene expression and overall yield is reportedly lower than with conventional, wildtype rBEV systems [85, 86]. We recently developed a novel system for producing KOVs based on CRISPR-Cas9 mediated introduction of indel mutations in the AcMNPV genome (submitted article). This system was able to disrupt progeny BV release and/or reduce late gene expression through targeted disruption of several AcMNPV genes. Targeting the *vp80* gene, which encodes the nucleocapsid-associated protein VP80, with this approach resulted in reduced BV release but did not appear to significantly impact expression of the *gfp* reporter gene. To assess this strategy for its utility as an effective production platform for VLP production with concomitant reduced BV release, we targeted the AcMNPV *gp64* gene for disruption. After confirming successful obstruction of GP64 expression, production of HIV-1 Gag VLPs was assessed.

Initial efforts established and confirmed successful targeting of the *gp64* gene using CRISPR-Cas9. To this end, the abundance of GP64 in infected cell lysates and in the membrane of infected cells was measured. Our results indicated ~99% and ~90-95% reduction of GP64 in lysates and in the membrane of infected Sf9-Cas9 cells, respectively. Importantly, the abundance of GP64 in Sf9 cells infected with rBEVs targeting *gp64* was indistinguishable from control infections, indicating that disruption of GP64 expression was the result of CRISPR-mediated targeting of the *gp64* ORF.

Next, the effect of targeting *gp64* on late gene expression and progeny BV release was measured. Disruption of GP64 resulted in >98% and ~94% reduction of IVT and total particles/ml, respectively. This data is consistent with a previous report in which BV release was reduced by ~50-98% for different *gp64* gene truncations [438]. Similarly, GP64 appeared to be undetectable for the Δ*gp64* KOV via western blot, however direct quan-

tification of BV in the supernatant was not conducted in that report [86]. For late gene expression, our results indicated that expression of the *gfp* reporter gene was not significantly affected by *gp64* disruption. Although the median fluorescence intensity was slightly lower for *gp64*-targeting rBEVs compared to the control, this difference in expression was similar for both Sf9-Cas9 and Sf9 cell lines. This data could indicate that variability between individual virus stocks may have accounted for these differences as opposed to decreased late gene expression as a result of CRISPR-mediated targeting. Nevertheless, these differences were not statistically significant. This is a significant result, as previous reports indicated that high MOIs were required for similar EGFP yields between Δ*vp80* KOV and the control virus [85], and production of Gag VLPs appeared to be lower via western blot analysis between the Δ*gp64* KOV and the control [86].

Finally, we assessed the production of HIV-1 Gag VLPs with concomitant reduced BV contamination. The HIV-1 *gag* ORF encodes a 55 kDa polyprotein (Pr55 or Gag) that is processed into several proteins, including the 17 kDa matrix protein (p17 or MA), the 24 kDa capsid protein (p24 or CA), and the 7 kDa nucleocapsid protein (p7 or NC) [439]. Expression of Gag alone is sufficient for assembly and budding of VLPs, and several studies have demonstrated production of pseudotyped and non-pseudotyped Gag VLPs in the BEVS and in uninfected insect cells [86, 436, 440–443]. In addition to targeting *gp64*, rBEVs with sgRNAs targeting the *vp80* ORF were prepared in order to compare VLP production using both of these strategies. Similar to previous results, targeting the *gp64* ORF resulted in significant reduction of GP64 abundance in the plasma membrane of infected cells and IVT. The IVT of *vp80*-disrupted rBEVs was also significantly reduced compared to control infections in Sf9 cells. Unexpectedly, immunofluorescence staining

of GP64 in the plasma membrane of infected cells was observed to be lower in Sf9-Cas9 cells compared to Sf9 cells, suggesting that disruption of VP80 expression may impact GP64 production. Reduced GP64 was not observed by western blot analysis of cell lysates infected with a Δ*vp80* KOV previously [85], however staining of GP64 in the membrane of those cells was not conducted. On the other hand, analysis of VP39 by western blot indicated lower abundance in cells infected with the Δ*gp64* KOV [86]. The results here do not appear to be associated with off-site targeting of the Cas9/sgRNA ribonucleoprotein complex, as 2 other sgRNAs targeting the *vp80* ORF showed similar results (data not shown). Similarly, there were insignificant differences between GP64 measurements in the cell membranes infected with control or *vp80*/*gp64*-targeted rBEVs (data not shown). As such, this observation appears to be the result of a potential and as yet unreported interaction between *vp80* disruption and GP64 expression, and may require further scrutiny to assess this relationship. Nevertheless, both of these strategies were successful for producing Gag VLPs with concomitant reduction in rBEV contamination. Importantly, although the estimated yield of VLPs by p24 ELISA was lower compared to a control (ie., untargeted rBEV expressing the *gag* gene), yields of VLPs were similar in Sf9-Cas9 and Sf9 cells for all of the rBEVs, suggesting that these results might be due to variance among virus seed stocks as opposed to the strategy itself.

4.6 Concluding remarks

In this report, CRISPR-mediated disruption of the *gp64* gene was assessed. After confirming that this strategy resulted in target specific obstruction of GP64 and reduced BV

release, production of HIV-1 Gag VLPs was assessed and compared with a similar strategy in which the *vp80* ORF was targeted for disruption. Both strategies resulted in high level production of VLPs along with reduced rBEV contamination in culture supernatants. This strategy may be impactful for simplifying the purification of recombinant proteins and other complex biologics such as VLPs, and may be an improvement over previously reported strategies in which initial virus seed production was impaired and overall yield may be impacted.

Chapter 5

Alternative promoters for improving BEVS

Use of the very late polh or p10 promoters to control transcription of foreign genes in the BEVS is near ubiquitous. A review of the literature found more than 70 individual publications describing the production of single heterologous proteins or virus-like particles. Of these, only 3 proteins (of more than 50) were produced using promoters other than polh or p10 for the single protein studies, and p10 and polh were used for 97 of 105 viral structural proteins for producing VLPs. Further, 7 of the remaining 8 used a tandem vp39-polh (cappolh) promoter, leaving only 1 protein expressed by a promoter that did not include p10 or polh [87]. This strategy is not without reason; by 48 hpi, >30% of total cellular RNAs in baculovirus-infected cells are *polyhedrin* or *p10*-encoding mRNAs [16]. Despite this incredible abundance of polh and p10-transcribed mRNAs, several studies have demonstrated that alternative promoters may be more effective for producing recombinant proteins of higher quality and/or yield. Examples were discussed briefly in Chapter 1. Characterization of alternative viral promoters, however, has been generally limited to

those that are transcribed by the host RNA Polymerase II alone or in tandem with a very late promoter, such as the pIEx/Bac™ family of baculovirus transfer plasmids developed by Novagen. These vectors consist of the *hr5*-ie1 promoter in tandem with the p10 promoter. While ie-1 enables expression of the foreign gene early in the infection cycle, it is generally considered a 'weak' promoter and active transcription from the host RNAP II may be downregulated or completely subverted early in the infection cycle. The p10 and polh promoters are active only during the late stages of the infection cycle, potentially leaving the intervening time as relatively unproductive for foreign gene expression. Alternatively, expression of some foreign genes may be toxic to the cell and consequently have a negative effect on the production bioprocess. As such, use of weaker promoters may be necessary to avoid the accumulation of the foreign protein at a level that is toxic to the cell. To enable further optimization of recombinant protein production, characterization of alternative promoters that allow for optimization of the levels of recombinant protein produced is desirable.

Additionally, as discussed in Chapter 1, co-infection strategies have been used extensively for production of VLPs and multi-protein complexes. Co-infection allows for manipulation of MOI and TOI parameters of individual rBEVs to provide some control of timing, order, and stoichiometric ratios of multi-protein complexes. However, high MOIs of each individual rBEV are required to achieve uptake of each rBEV to every cell; in practice, the number of infected cells is generally lower than predicted using mathematical modelling [87]. Integrating multiple foreign genes into the same rBEV for co-expression at least in part mitigates this drawback, however it also removes the ability to modulate expression by manipulating MOI or TOI. Employing different promoters to allow for expression of

foreign genes that have different promoter 'strength' and/or temporal activation characteristics may improve co-expression of proteins that assemble to form multi-protein complexes such as VLPs with the correct relative abundance/stoichiometric ratios to improve overall yield and quality.

To improve upon the current 'catalogue' of promoters available for foreign gene expression in the BEVS, we used previously published transcriptome data [16] to select AcMNPV promoters having different relative transcription strengths. We assessed the selected promoters through expression studies using model cytoplasmic and secreted proteins. This work also prompted us to investigate if characteristics exist that directly result in the spatiotemporal properties of baculovirus promoters. For application, we expressed the adeno-associated virus *rep78* gene from each of the promoters to attempt to improve replication of the AAV genome. This work was conducted with an industrial partner over approximately 2 years. Due to the sensitive nature of the specific goals of the project, only a small subset of results pertaining to expression of the *rep78* gene are presented as additional results at the end of this chapter.

Utility of alternative promoters for foreign gene expression using the baculovirus expression vector system

Mark R. Bruder[1] and Marc G. Aucoin[1]

[1] Department of Chemical Engineering, University of Waterloo.

5.1 Abstract

The baculovirus expression vector system (BEVS) is a widely used platform for recombinant protein production for use in a wide variety of applications. Of particular interest is production of virus-like particles (VLPs), which consist of multiple viral proteins that self-assemble in strict stoichiometric ratios to mimic the structure of a virus but lacks its genetic material. While a significant amount of effort has been spent on optimizing expression ratios by co-infecting cells with multiple recombinant BEVs and modulating different process parameters, co-expressing multiple foreign genes from a single rBEV may offer more promise. However, there is currently a lack of promoters available with which to optimize co-expression of each foreign gene. To address this, previously published transcriptome data was used to identify promoters that have incrementally lower expression profiles and compared by expressing model cytoplasmic and secreted proteins. Bioinformatics was also used to identify sequence determinants that may be important for late gene transcription regulation, and translation initiation. The identified promoters and bioinformatics analyses may be useful for optimizing expression of foreign genes in the BEVS.

5.2 Introduction

The Baculovirus Expression Vector System (BEVS) is a versatile manufacturing platform for clinically important therapeutic proteins, viral vectors for gene therapy, and antigens for vaccination. In addition to being scalable, cost efficient, and capable of high product yields, the ability to rapidly produce a recombinant BEV (rBEV) carrying large foreign

DNA inserts has catapulted the BEVS platform toward mainstream acceptance in the biotechnology industry [243]. The current COVID-19 pandemic has underscored the importance of the BEVS for manufacturing antigens for vaccination: the SARS-CoV-2 vaccine NVX-CoV2373 developed by Novavax and produced using the BEVS platform is poised to be just the 4th vaccine approved for Emergency Use Authorization (EUA) by the US Food and Drug Administration (FDA), and would be the first protein subunit vaccine approved [444].

Although the BEVS has been used as an expression vector for thousands of recombinant proteins, the vast majority of research has focused on expression of a single foreign gene using the very strong, very late, polh or p10 promoters [8, 87]. Despite this, many studies have suggested that earlier promoters may be more effective for producing some foreign proteins, particularly those that require extensive post-translational processing and/or secretion, as the very late promoters are most active when host cell processes may be compromised [87, 131, 133, 134, 405]. For more complex biologics that are composed of multiple foreign proteins such as virus-like particles (VLPs), co-infection with multiple monocistronic baculoviruses is a strategy that is routinely employed [87]. Co-infection allows for optimization of process parameters such as multiplicity of infection (MOI) and time of infection (TOI) for each individual rBEV to modulate the timing of expression and stoichiometry of each protein to maximize self-assembly and overall yield. Substantial improvements in the production of adeno-associated virus (AAV) and rotavirus-like particles have been realized with this approach [87, 157, 158, 445]. Co-infection with multiple monocistronic baculoviruses does have potential disadvantages, however: various studies have suggested that as the number of viruses increases, the proportion of cells that are in-

fected with each virus (or in equal ratios) decreases, leading to efficiency loss in production of fully formed VLPs. Further, the burden of copying genetic material of multiple different viruses may lead to faster cell death [446]. Polycistronic baculoviruses, on the other hand, ensure that every protein necessary for the self-assembly of the VLP is expressed in each infected cell, and various studies have reported higher yields using co-expressed proteins from polycistronic baculoviruses than co-infections with multiple monocistronic vectors [87]. Due to the predominance of the p10 and polh promoters on commercially available transfer vectors, however, modulating expression parameters akin to MOI and TOI is virtually impossible.

Despite the aforementioned utility of promoters that are active earlier in the infection cycle for foreign gene expression, previous reports have focused mainly on identifying endogenous or chimeric baculovirus promoters that achieve higher protein expression levels than the native polh or p10 promoters [131, 137, 142, 143, 447]. For complex VLPs, however, individual proteins may be required in very different amounts; for example, the abundance of the four distinct proteins that make up rotavirus-like particles ranges from 60-780 molecules per particle, and the ratios of the integral membrane proteins hemagglutinin (HA), neuraminidase (NA), and matrix (M2) of influenza A virus is approximately 4:1:0.04-0.4 (HA:NA:M2) [157, 448]. Moreover, for many non-enveloped VLPs such as rotavirus-like particles, a fixed stoichiometry between the four constituent proteins is required for properly assembled and stable particles. This presents a considerable challenge for co-expression of the required molecules from a polycistronic rBEV, as the expression levels of multiple proteins, each having unique characteristics, must be modulated to satisfy strict stoichiometric requirements. However, the catalogue of promoters that have been

reported so far for the BEVS includes very few that have lower expression profiles than p10/polh.

To improve upon the catalogue of rBEV promoters, we used previously reported transcriptome data [16] to select a series of native AcMNPV promoters with different expression characteristics. The promoters were evaluated and compared by expressing the model cytoplasmic protein green fluorescent protein (GFP) and secreted protein secreted alkaline phosphatase (SEAP). Finally, we aimed to evaluate whether sequence determinants were identifiable that govern the expression characteristics of late AcMNPV gene expression. In addition to adding new promoters to the rBEV expression catalogue, to our knowledge this is the first report of employing bioinformatics analyses for evaluation of AcMNPV promoters. It is our hope that this report may help improve production of complex biologics such as VLPs from polycistronic rBEVS by allowing for more modulation of expression of each foreign gene, as well as spur on more comprehensive evaluation of gene expression profiles to further our understanding of baculovirus gene expression and improve the BEVS as a biotechnology platform.

5.3 Materials and methods

5.3.1 Plasmid construction

All plasmids used in this study were constructed using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, Whitby ON) according to manufacturer's directions. Primers used for construction of all plasmids were synthesized by Integrated DNA Tech-

nologies (IDT; Coralville, IA) and are given in Table C.1. The genomic regions for each promoter are given in Table 5.2. The co-ordinates are based on the RefSeq entry for the AcMNPV genome (NC_001623.1) [15].

To construct the promoter-GFP transfer plasmids, the selected promoter regions were amplified from AcMNPV genomic DNA and inserted upstream of the *gfp* gene encoding green fluorescent protein (GFP) [399] in plasmid p6.9-GFP, described previously. For SEAP-expressing plasmids, the *gfp* in each of the promoter-GFP plasmids was replaced with the *seap* gene encoding secreted alkaline phosphatase (SEAP) amplified from pYSEAP (Addgene # 37326).

5.3.2 Recombinant baculovirus generation, amplification, and quantification

Transfer plasmids for recombinant baculovirus expression vector (rBEV) generation were co-transfected with *flash*BACGOLD™ (Oxford Expression Technologies Ltd., Oxford UK) genomic DNA to Sf9 cells using Escort IV transfection reagent (Sigma-Aldrich, Oakville ON) according to manufacturer's directions. Supernatant from each transfection was harvested 4-5 days post transfection and used to infect suspension Sf9 cultures (~1.5×10^6 cells/ml) at low multiplicity of infection (MOI) for 3-4 days to amplify the rBEV. Following two rounds of amplification, the rBEV infectious virus titer (IVT) was quantified using end-point dilution assay (EPDA). Briefly, Sf9 cells were seeded at a density of ~2.0×10^4 cells/well to each well of a 96-well plate (Fisher Scientific). Separately, the rBEV was serially diluted (10^{-2} to 10^{-8}) in fresh SF900 III medium and 10 μl of each dilution was

added in 12 replicates to the 96-well plate containing cells. Plates were incubated for 6-7 days at 27 °C, after which wells were scored according to visualization of green fluorescence using a fluorescence microscope or cytopathic effects. Results were converted from $TCID_{50}$ and reported as plaque forming units per ml (pfu/ml).

5.3.3 Infections

Sf9 cells in suspension were infected with rBEVs at a density of $\sim 1.5 - 2 \times 10^6$ cells/ml viable cells / ml at a MOI of ~3 pfu/cell. Samples were harvested at the required times (hours post infection; hpi) wherein cells were centrifuged at $300 \times g$ for 10 minutes and resuspended in 2% paraformaldehyde diluted in phosphate buffered saline (PBS) for ~30 min prior to analysis by flow cytometry where appropriate. The cell culture supernatant was kept at 4 °C for further analysis where appropriate.

5.3.4 Flow cytometry and analysis

Cells infected with GFP-producing rBEVs were analyzed using a FACSCalibur™ flow cytometer (BD Biosciences, San Jose CA) equipped with an argon-ion laser with an excitation frequency of 488 nm. Samples were run at the low flow setting (12 μl/min) and 10 000 events were collected. Analysis of flow cytometry data was performed using FlowJo® V10 flow cytometry analysis software (FlowJo LLC, Ashland, OR). Results are reported as the mean of at least 3 independent replicates.

5.3.5 SEAP activity assays

The supernatants of infections with SEAP-producing rBEVs were harvested by centrifugation at $1000 \times g$ for 10 min. SEAP activity was quantified using the SEAP Colorimetric Reporter Assay Kit (Novus Biologicals, Toronto ON) according to manufacturer's directions. The absorbance was measured using a Synergy 4 hybrid microplate reader (BioTek, Winooski, VT) at a wavelength of 405 nm. The SEAP concentration was determined using a calibration curve of known SEAP concentration standards. At least 3 independent replicates were run for each rBEV and each sample was run in triplicate for quantification of SEAP.

5.3.6 Real-time reverse transcription polymerase chain reaction (RT-PCR)

RNA was extracted from infected cells using the Geneaid Total RNA Mini kit (FroggaBio, Concord ON) and 500 ng was used as template for first-strand cDNA synthesis using the SensiFAST cDNA synthesis kit (FroggaBio) according to manufacturer's directions. Real-time PCR was performed using the SensiFAST SYBR Hi-ROX kit (FroggaBio) according to manufacturer's directions on an Applied Biosystems StepOnePlus™ Real-Time PCR System (Fisher Scientific). Primer pairs used for qPCR are given in Table C.1. Analysis was conducted in the R programming environment using the *pcr* package [449] using 28S rRNA as the internal reference gene for data normalization [450].

5.3.7 Bioinformatics

All bioinformatics analyses were conducted using the R programming environment using several Bioconductor packages, including *msa*, *genbankr*, and *ggbio* [413, 451, 452].

5.4 Results

5.4.1 Selection of AcMNPV promoters

The transcriptome data mapped to AcMNPV open reading frames (ORFs) [16] as reads per kilobase of transcript per million mapped reads (RPKM) values were used to select AcMNPV promoters with different expression characteristics. The ORFs were first divided into 'classes' based on the RPKM values (Table 5.1 and Figure C.1) to separate them into groups according to relative transcript abundance. Selection of individual promoters for evaluation was based on two main criteria: transcription should accelerate and reach steady-state levels early in the infection cycle and maintain expression levels until the later stages of infection. The promoters from ORFs having a variety of steady-state RPKM values were selected for evaluation (Figure 5.1 and Table 5.2). The vast majority of commercially available BEVS transfer plasmids include either the polh or p10 promoter, while the gp64 and *hr5*-ie1-p10 promoters are present in two (Table C.2). As the difference in expression levels between the gp64/ie1 and polh/p10 promoters is extremely large (Figure 5.1A), the promoters selected for further analysis were expected to provide intermediary levels of expression between these levels (Figure 5.1B).

Table 5.1: RPKM value ranges for AcMNPV transcript 'classes'

Class	RPKM Range	ORFs
Very High	RPKM $\geq$ 50000	5
High	20000 $\geq$ RPKM $<$ 50000	6
Medium	10000 $\geq$ RPKM $<$ 20000	17
Low	1000 $\geq$ RPKM $<$ 10000	65
Very Low	RPKM $<$ 1000	56

Table 5.2: AcMNPV genomic coordinates for promoters used in this study.

Promoter	Coordinates †	Predicted Promoter Motifs ‡	Direction	Class
polh	4428..4519	TAAG	+	Very Late
p6.9	86889..87204	CAGT, TAAG	−	Late
ctx	2246..2447	TAAG	−	Late
orf75	63528..63912	TAAG	−	Late
vp39	76578..77103	TAAG	−	Late
39k/pp31	30070..30398	TATA, CAGT, TAAG	−	Delayed Early
gp64	109718..110022	TATA, CAGT, TAAG	−	Early/Late
38k	85984..86276	TAAG	−	Late
Δp10	118635..118808	TAAG	+	Very Late

†: based on NCBI ref. NC_001623.1 [15]; ‡: based on ref. [16]

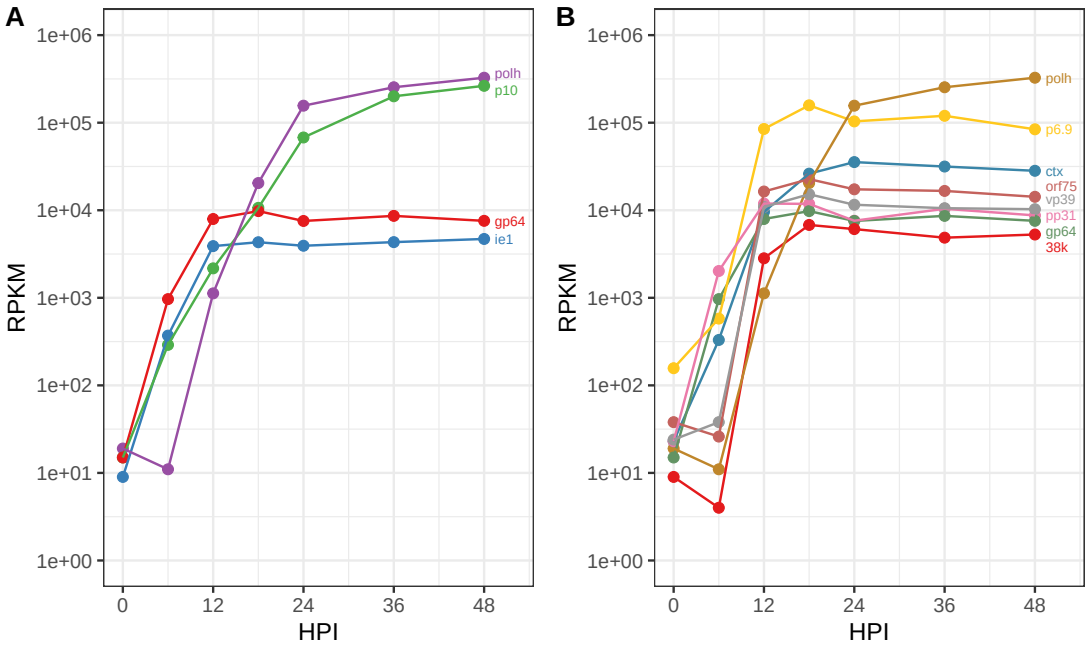


Figure 5.1: **The promoters included on commercially available transfer plasmids have drastically different transcription profiles. A.** The transcript abundance of AcMNPV polh, p10, gp64, and ie1 ORFs, which are among the only promoters available on commercial transfer plasmids for foreign gene expression. **B.** The transcript abundance profiles of AcMNPV ORFs selected for evaluation of the upstream promoter regions in this study. Promoters were selected for expression profiles between polh/p10 and gp64/ie1.

5.4.2 Evaluating the expression profile of AcMNPV promoters

Recombinant BEVs were prepared to express the model cytoplasmic protein GFP and secreted protein SEAP to assess the expression characteristics of the selected promoters. In addition to analysis of protein production, RNA was extracted from infected cells at 24, 48, and 72 hpi, and transcription of *gfp* or *seap* was analyzed using RT-PCR. Notably, for intracellular GFP production, the 39k promoter appeared to be the most active promoter during the earliest stages of the infection, with median fluorescence intensity of ~100 au by 12 hpi, whereas the GFP levels were less than 10 au for the other promoters. By 24 hpi, the p6.9 promoter had produced slightly higher levels of GFP compared to 39k. The polh promoter, on the other hand, was among the lowest in terms of fluorescence intensity until 24 hpi, after which expression increased rapidly and by 48 hpi the polh-GFP rBEV produced the highest level of GFP. Production of GFP from the vp39 promoter was similar to the 39k rBEV in the later stages of the infection, however the ctx and orf75 promoter rBEVs were lower than expected from their respective RPKM values (Figure 5.2A and Table 5.3). The Δp10, 38k, and gp64 promoter rBEVs yielded fluorescence intensity measurements that were lowest of the promoters tested, with Δp10-GFP rBEV reaching ~30 au by 48 hpi. RT-PCR data revealed that *gfp* transcript abundance from the p6.9 promoter was over two times greater than any other promoter at 24 hpi, and ~7× higher than the polh promoter. By 48 hpi, transcription from the polh promoter was only slightly lower than the p6.9 promoter and at 72 hpi was nearly 3× greater than p6.9 (Figure 5.2B).

Expression and secretion of SEAP followed a similar trend to that of GFP. By 24

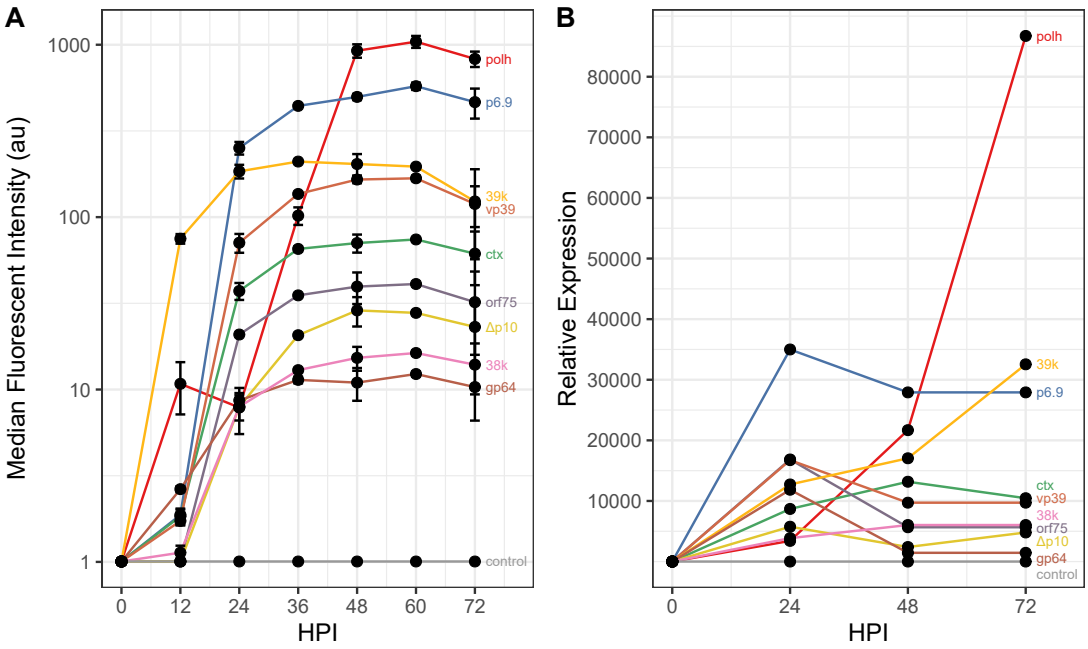


Figure 5.2: **Production of intracellular GFP from selected AcMNPV promoters. A.** Median fluorescence intensity measured using flow cytometry and **B.** relative transcript abundance measured using RT-qPCR at various times post infection.

hpi, SEAP activity corresponding to ~2-2.5 mg/l of SEAP protein had been secreted to the supernatant for the p6.9, 39k, and vp39 promoter rBEVs. Each of the other rBEVs had produced less than ~0.5 mg/l SEAP (Figure 5.3A). By 48 hpi, SEAP activity had more than doubled for the p6.9-SEAP rBEV to ~5.5 mg/l, whereas 39k and vp39 had produced ~50% more (~3-3.5 mg/l). The SEAP activity produced by the polh-SEAP rBEV, on the other hand, had increased by 5× in the same time interval and was similar to 39k and vp39. Despite a significant increase in transcription activity from the polh promoter between 48 and 72 hpi (and apparent sharp decrease in p6.9-*seap* transcription), SEAP activity reached ~6.5 mg/l and ~3.75 mg/l for p6.9-SEAP and polh-SEAP rBEVs at 72 hpi, respectively (Figure 5.3A). Significantly, SEAP production from the 39k and vp39 promoters was not substantially different from that of the polh promoter despite significant differences in transcriptional activity (Figure 5.3B). Similar to the GFP results, the ctx and orf75 promoters produced lower SEAP activity than expected, while the other promoter rBEVs produced SEAP levels in a similar ranking to GFP expression; production of SEAP from the gp64 and 38k rBEVs were lowest but their order swapped as compared to GFP expression. Interestingly, RT-PCR results suggested lower *seap* mRNA abundance was produced from the ctx promoter than orf75 (Table 5.3).

5.4.3 Bioinformatics exploration of sequence and genomic architecture determinants that contribute to promoter activity

The promoter activity results prompted the exploration of specific sequence or genome architecture determinants that may impact promoter activity. Along with RPKM values

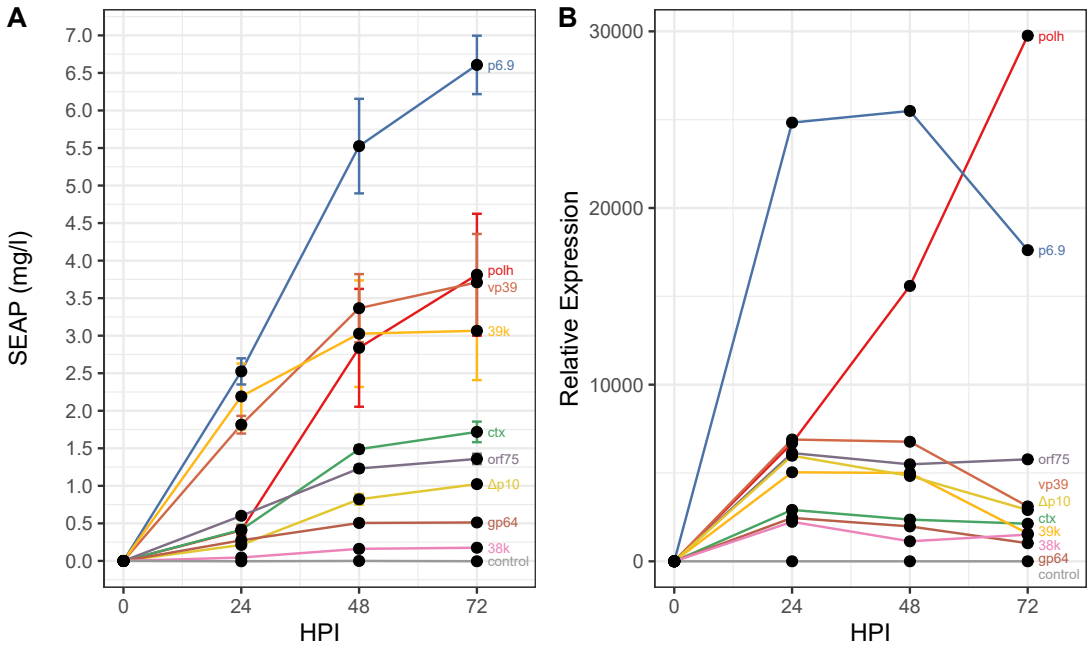


Figure 5.3: **Production of extracellular SEAP from selected AcMNPV promoters. A.** Yield of SEAP (mg/l) of culture supernatants measured using a colorimetric SEAP activity assay and **B.** relative transcript abundance measured using RT-qPCR at 24, 48, and 72 hours post infection.

Table 5.3: Promoters ranked by GFP and SEAP production at 48 hpi.

Promoter	Rank (48 hpi)				
	ORF	GFP		SEAP	
	RPKM	Protein	qPCR	Protein	qPCR
polh	1	1	2	2	2
p6.9	2	2	1	1	1
ctx	3	5	4	5	7
orf75	4	6	6	6	4
vp39	5	4	5	3	3
39k	6	3	3	4	5
gp64	7	9	9	8	8
38k	8	8	7	9	9
Δp10	n/a	7	8	7	6

for transcript abundance, putative transcription start sites (TSS) and promoter motifs were included in the transcriptome data set and were used for further analysis [16]. Additionally, sequences encompassing the 5' untranslated region (5'UTR), TSS, and upstream sequences including the putative promoter motifs were extracted from the AcMNPV genome and analyzed. Analysis was conducted using the ORFs divided into classes based on RPKM values. Unsurprisingly, the 'Very Low' class had the highest proportion of promoters with identified 'TATA' or 'CAGT' motifs, which indicate transcription from the host RNA Polymerase II (RNAP II). The 'Low' and 'Medium' classes had fewer RNAP II promoter motifs than 'Very Low', whereas the 'High' and 'Very High' classes had none (Figure C.2). Similarly, the 'TAAG' motif, which is required for initiation of transcription from the viral RNAP (vRNAP) was present in the promoter regions for all ORFs in the Medium, High,

and Very High classes, and was found less frequently in the promoters of ORFs in Low and Very Low classes. Interestingly, manual inspection of sequences around the TSS of several genes indicated deviation from the 'CAGT' motif. All of the examined ORFs with apparent deviations were designated as Low or Very Low transcript abundance. Inspection of the sequence surrounding the TAAG motif of several genes in each class found no deviation from this motif (data not shown). Next, the 5'UTRs were examined for differences between each class. The length of the 5'UTR (i.e. the number of nucleotides between the TSS and ATG start codon) was not statistically different between classes (Figure 5.4A). The $A+T$ content, on the other hand, was significantly higher for the Very High class compared to the other classes (Figure 5.4B).

Various studies in the literature have suggested that the homologous repeat (*hr*) regions found in baculovirus genomes have roles as origins of DNA replication (*oris*) and as transcriptional enhancers that may act in both *cis* and *trans*. To investigate whether there was a relationship between transcript abundance and distance from a *hr* region, the genome map of AcMNPV was colour-coded according to class (Figure 5.5) and the distance from the 5' end of the *hrs* to the start codon of each gene was calculated (Figure 5.6). There did not appear to be any discernible relationship between ORF transcript abundance and proximity to any of the *hrs*.

A previous study identified two octamers (5'-ATTGCAAG-3' and 5'-ATTAGGAA-3' herein referred to as upstream and downstream, respectively) located within the sequences upstream of both the *p6.9* and *vp39* TSSs [453]. The authors hypothesized that these could be protein binding sites for late gene transcription. To test this hypothesis, sequences upstream of each AcMNPV ORF (225 nucleotides upstream of the TAAG motif)

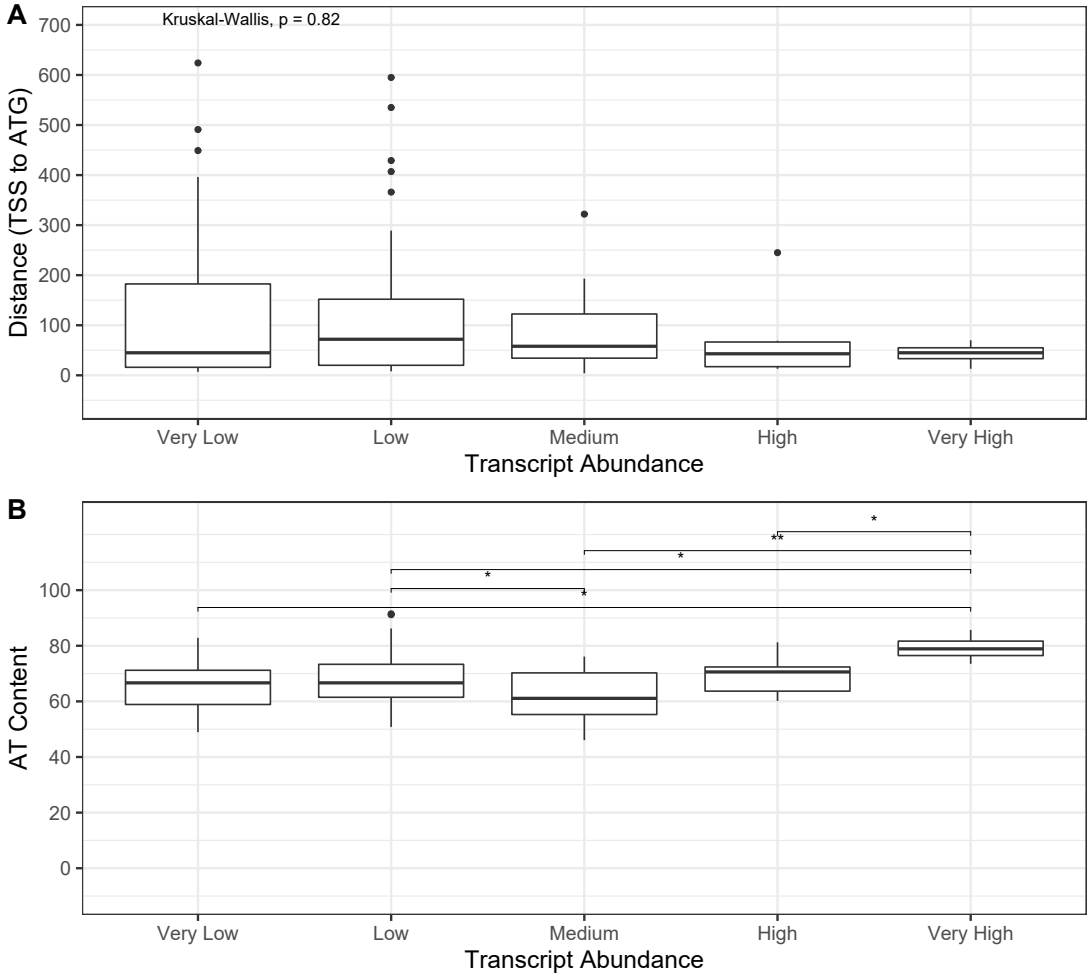


Figure 5.4: **Evaluation of the 5'UTRs of AcMNPV ORFs categorized according to transcript abundance. A.** Length (in nucleotides) and **B.** A/T content of the 5'UTR between the late gene promoter motif (5'-TAAG-3') and translation initiation codon (ATG).

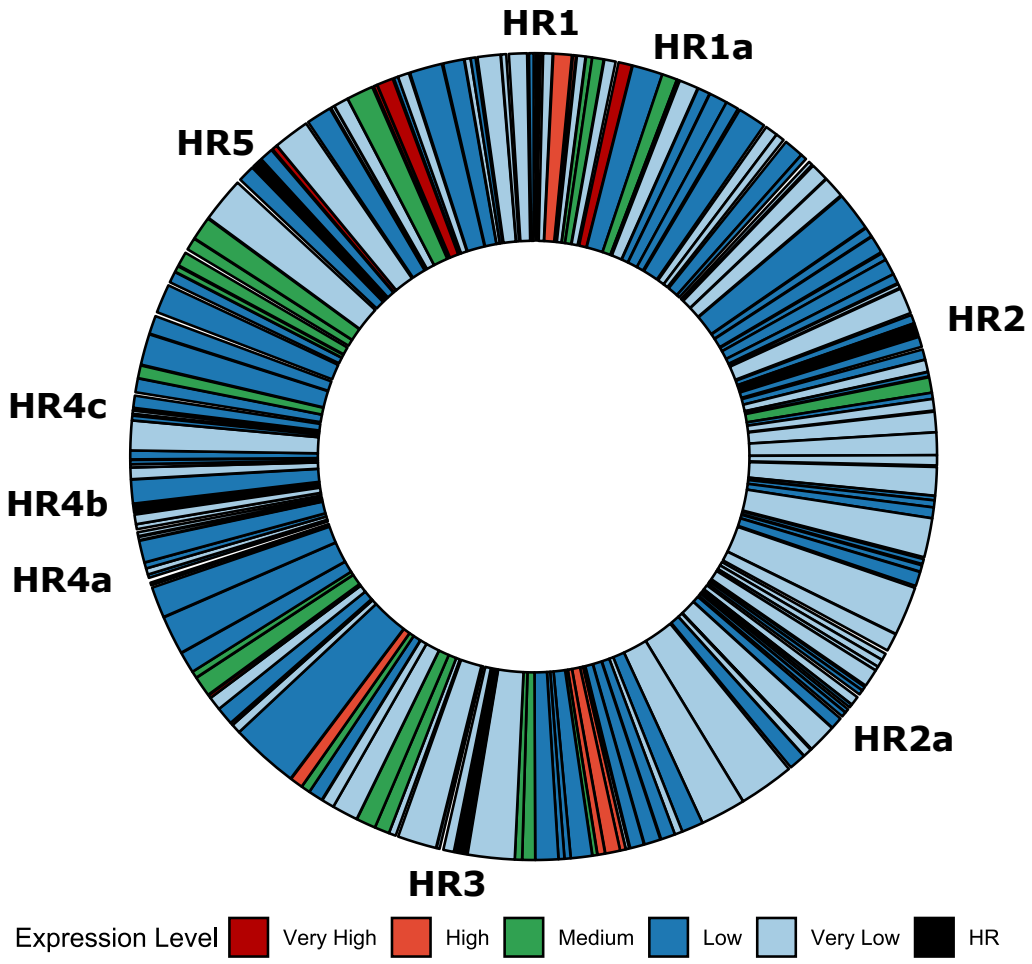


Figure 5.5: **Circular chromosome map of AcMNPV ORFs colour-coded according to transcript abundance.**

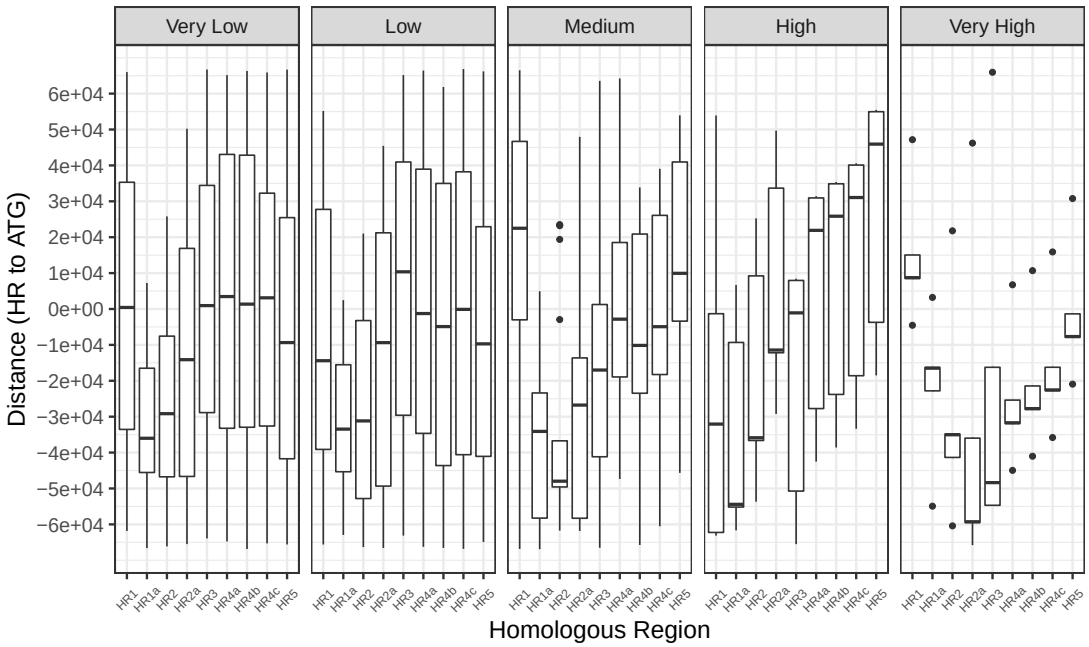


Figure 5.6: **Distance (in nucleotides) between the start codon and 5' end of each homologous region for every AcMNPV ORF categorized according to transcript abundance.** The distance was calculated by subtracting the genomic location of the 5' end of each AcMNPV ORF from the genomic location of the 5' end of each *hr*. Positive values represent ORFs located behind (clockwise) to the *hr* and negative values represent distances between ORFs that are located in front of (counterclockwise) the *hr*.

were searched using these octamer sequences as queries. Only two matching sequences (upstream of the *p6.9* and *vp39* TSSs) were found for both queries. A search of the entire AcMNPV genome only yielded 1 additional match for the downstream motif, which was located within the coding sequence of the *p74* gene, and 4 for the upstream motif, also found within coding sequences and not close to the TAAG motif for the adjacent downstream gene (data not shown). Two mismatches were required for both octamers to yield possible matches in the ~225 nucleotides upstream of the TAAG motif for the majority of AcMNPV ORFs, with multiple putative matches found for each (Figure 5.7). This same search was performed using the entire AcMNPV genome sequence: 1733 and 1439 matches were found for the upstream and downstream sequences, respectively, and were roughly divided equally between the coding and noncoding strands for both. Allowing for 2 mismatches in the upstream octamer yielded 52 potential matches in the 29 individual sequences belonging to ORFs in the Very High, High, and Medium classes. Only 7% (2 of 29) did not contain a putative match, whereas 38% and 31% had 1 or 2 matches. Additionally, 11 different octamer sequences were present in more than 1 upstream sequence. Although the sequences were dispersed throughout the length of the upstream regions (Figure 5.7), a multiple sequence alignment (MSA) revealed that the consensus sequence was 5'-ATTGCAAN-3'. For the downstream octamer, 47 potential matches were found however 34% of the sequences had no putative matches, and only 31% had either 1 or 2. Inspection revealed that 8 sequences were found in at least two upstream regions (Table C.4 & C.3). Similar to the upstream motif, although the consensus sequence was 5'-ATTAGGAA-3', these sequences were located randomly throughout the putative promoter regions. Manual inspection of the upstream regions that only had a single putative

match for either the upstream or downstream motifs revealed similar results.

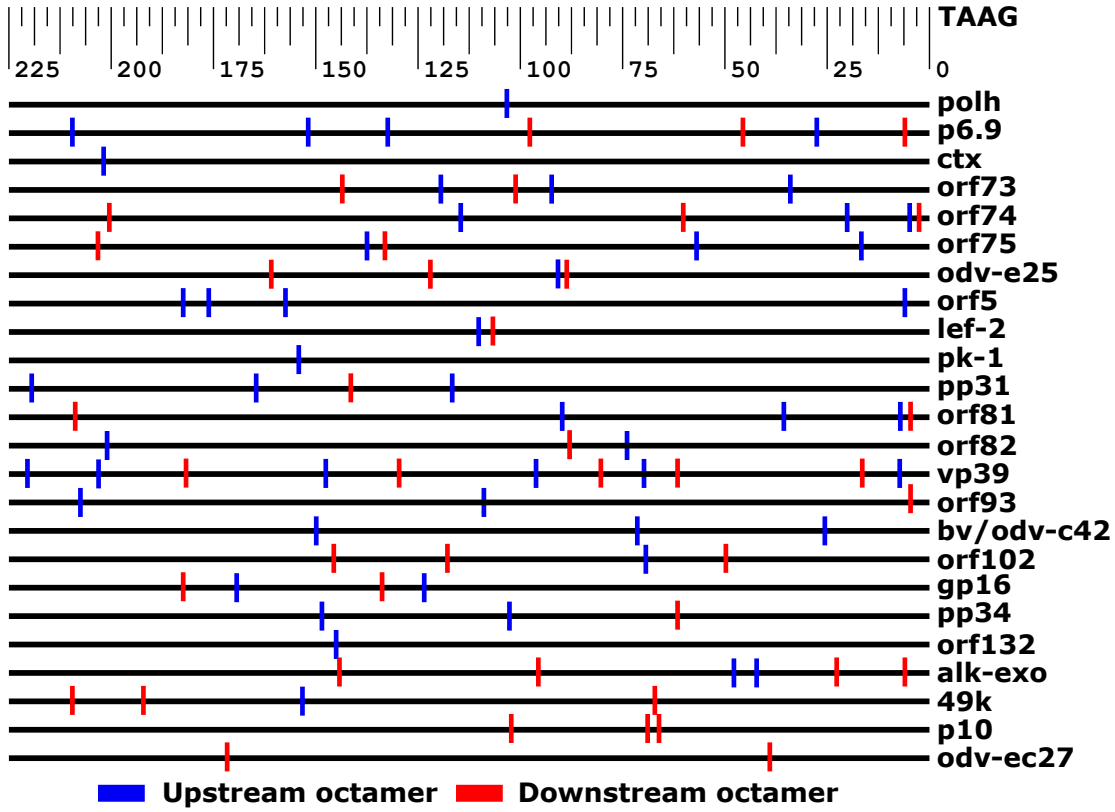


Figure 5.7: **The approximate positions of the upstream octamer (blue) and downstream octamer (red) in regions upstream of the late gene promoter motif for the most abundant AcMNPV ORFs.**

Finally, sequences surrounding the late gene TSS and translation initiation site (TIS) were extracted for each ORF and analyzed. Multiple sequence alignments were created for each class and the consensus sequence for each MSA was calculated for comparison. The consensus sequences for the nucleotides flanking the translation initiation site (TIS) and transcription initiation site (TSS) are given in Figure 5.8A and B, respectively. Addi-

tionally, the consensus sequence for the most abundantly transcribed ORFs excluding *polh* and *p10* was also included for the sequences surrounding the TSS. Conserved nucleotides are highlighted in green and the most conserved locations have a box placed around them. Nucleotide positions −2, −3, and −7 with respect to the TIS (A is +1) appear to be the most conserved sites flanking the TIS, whereas −4, +6, +9, +10, and +12 (+1 is the first A in the TAAG motif) are highly conserved sites flanking the TSS.

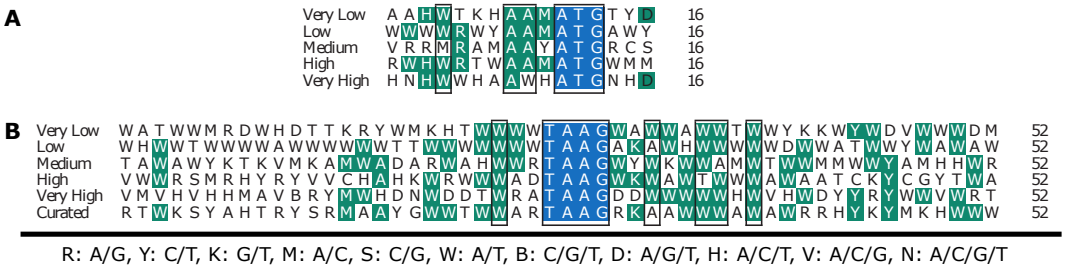


Figure 5.8: **Consensus sequences calculated from multiple sequence alignments. A.** Consensus sequence for the nucleotide sequences flanking the translation initiation site and **B.** the late gene promoter motif. Consensus sequences were calculated from multiple sequence alignments of sequences extracted from AcMNPV ORFs that were categorized according to transcript abundance.

5.5 Discussion

Since the first recombinant proteins were produced in the early 1980s, considerable research effort has been devoted to improving and optimizing expression of foreign genes in the BEVS [8]. The majority of this focus has revolved around increasing expression of

foreign genes over that achieved with the endogenous polh or p10 promoters. This goal has led to studies identifying novel native promoters, chimeric tandem promoters, and introducing additional regulatory elements in *cis* or in *trans* [131, 137, 142, 143, 447]. While this approach has yielded regulatory elements with higher expression profiles compared to polh, many of these reports have used the model cytoplasmic protein GFP for evaluation. Contrarily, several studies have suggested that weaker promoters than polh may improve production of proteins that require extensive post-translational processing and secretion [131, 133, 134, 405]. Moreover, some multi-protein complexes such as VLPs have strict stoichiometric requirements for proper assembly, and overproduction of proteins may lead to inefficiencies and wasteful accumulation of unassembled proteins and/or formation of incomplete or improperly assembled particles. These deficiencies may impact efficacy of the biologic or economic feasibility of the production process [157]. To improve production of some recombinant proteins and biologics in the BEVS, it may be necessary to identify additional promoters with transcription characteristics that allow for the most efficient expression of each molecule. To augment the current catalogue of promoters for the BEVS, previously published transcriptome data [16] was used to identify promoters with transcriptional activities incrementally lower than the polh promoter. While several of the promoters have been identified previously, to our knowledge at least 3 have not been previously evaluated. Nevertheless, none of those selected have been compared in any systematic way nor routinely used for recombinant protein production.

To evaluate and compare the promoters, rBEVs were prepared to express the *gfp* and *seap* genes. For both reporter proteins, expression levels were largely consistent with the expected ranking compared to the RPKM values for each ORF. However, production of

GFP and SEAP from the ctx and orf75 promoters were lower than expected, while expression from the 39k promoter was higher, particularly during the early stages of infection. This latter result agrees with a previous report in which *seap* mRNA transcribed from the 39k and p6.9 promoters were similar at 24 hpi [405]. Additionally, previous reports of novel promoters for the BEVS have shown that expression strength depended on the upstream sequence selected as the promoter; for example, a 120 bp sequence upstream of the *orf46* (pSeL120) gene of *Spodoptera exigua* MNPV (SeMNPV) produced nearly 2× more GFP fluorescence intensity in Sf21 cells compared to promoter sequences that were extended to either 140 bp (pSeL140) or 301 bp (pSeL). Deletion of the 25 nucleotides adjacent to the translation initiation site from pSeL, on the other hand, nearly abolished GFP production entirely [143]. Further, GFP expression from each promoter varied in different cell lines. In Se301 cells, pSeL140 and pSeL120 produced similar levels of GFP, while pSeL produced less than half the fluorescence intensity at 96 hpi. In Sf21 and Hi5 cells, however, pSeL120 remained high whereas pSeL and pSeL140 produced similar but lower levels of GFP. Interestingly, GFP produced from the AcMNPV polh and p131 promoters, which is the AcMNPV homolog of the SeMNPV *orf46* gene, produced similar levels of GFP in all 3 cell lines tested [143]. Given that the transcriptome data used for analysis here originated from AcMNPV infection of High Five cells and baculovirus protein expression profiles may differ between infection hosts [16, 99], the divergence between the expected and observed transcriptional strengths of the 39k, orf75, and ctx promoters in this report may be due to host specific factors or promoter sequence selection. Similar to the SeMNPV orf46 promoter, careful scrutiny of the specific sequences included may improve their transcriptional strength. Additionally, it is worth noting that the *ctx* gene is part of a polycistronic

transcript with the *ac-bro* gene, which may have an impact on its transcription [145].

Unsurprisingly, expression of GFP with the polh promoter increased sharply at 24 hpi and was nearly 2× higher than with the p6.9 promoter by 48 hpi. Despite a drastic increase in transcriptional activity between 48 and 72 hpi, median GFP fluorescence intensity increased only ~25%. However, by 72 hpi each of the infected cell populations had experienced precipitous declines in viability (data not shown). Increased permeability of intact dead/dying cells leading to leakage of GFP molecules out of the cell may account for the discrepancy between median fluorescence intensity and transcript abundance at 72 hpi, as has been observed previously [152]. Nevertheless, production of GFP from the other promoter rBEVs reached peak median fluorescence intensity by ~36 hpi and maintained this level until 60 hpi after which fluorescence intensity declined, consistent with the decline in cell viability and results from the p6.9 and polh rBEVs. Interestingly, fluorescence intensity measurements for the 39k and vp39 rBEVs reached very similar levels, however production of GFP from the 39k promoter increased sharply by 12 hpi and reached peak levels by 24 hpi. The Δp10 promoter, described previously [140], is a truncated AcMNPV p10 promoter in which the A/T-rich burst sequence for Vlf-1 binding between nucleotides +39 and +72 (relative to the +1 of the TAAG TSS) has been removed. Removal of the burst sequence from either the polh or p10 promoter regions strongly attenuates their transcriptional strength, however they initiate transcription at the same time and level as the full length promoters [409, 454]. Consistent with these observations, in this study the Δp10 and polh promoters had similar fluorescence intensity values until 24 hpi after which the polh promoter's transcriptional activity sharply increased whereas the Δp10 promoter increased GFP production very gradually to reach a maximum fluorescence intensity that

was ~2.5% that of polh at 48 hpi.

Expression of SEAP, on the other hand, was substantially higher from the p6.9 promoter than polh. Once again, despite a drastic increase in transcriptional activity between 48 and 72 hpi, SEAP activity from cell culture supernatants only increased by ~25% to ~3.5 mg/l for polh. This yield was only ~50% that of the p6.9 promoter rBEV, and was statistically indistinguishable from the yield from either the vp39 or 39k rBEVs. This latter result supports an earlier study in which SEAP activity from transformed insect cells was substantially higher when the *seap* gene was transcribed from the 39k promoter than polh. Important differences in that study may have had an effect on those results; in addition to transformed cell lines, each promoter was linked to the *hr5* ori, which acts as a transactivator of several early genes including *39K*, but may be detrimental for polh activity [405, 455, 456]. In that study, despite transcription of *seap* being highest with the polh promoter, a significant proportion of the SEAP protein produced was found in intracellular protein extracts, indicating that inefficient secretion may have played a large role in the lower activity observed. Significantly, a recently published comparative transcriptome analysis of baculovirus-infected cells revealed significant differences in gene expression between rBEVs expressing model intracellular (mCherry) and secreted (Hemagluttinin; HA) protein products [457]. Notably, although the proportion of mapped reads were significantly lower for mCherry transcripts compared to HA transcripts, western blot analysis indicated that more mCherry protein was produced in both the intracellular and extracellular fractions as compared to HA. Several host cell genes were regulated specifically in response to the expression of secreted HA protein; many of the differentially regulated genes were involved in the stress response to unfolded or misfolded proteins, providing further evidence that

protein folding and processing in the endoplasmic reticulum or Golgi apparatus is impaired or at its capacity when the polh promoter is employed [457].

Based on these results, further analysis was conducted aimed at identifying any genome architecture or sequence determinants that may impact transcriptional strength. The 5'UTR of the *polh* and *p10* genes are A/T-rich [409, 458], and previous studies have reported improved foreign gene expression by inserting A/T-rich leader sequences in the 5'UTR of the polh promoter [150, 459]. Additionally, productivity improvements have been reported by inserting the *hr1* or *hr5* homologous regions upstream of various promoters to enhance foreign gene expression in a *cis*-dependent manner [447, 456, 460–463]. However, these discoveries often require extensive experimentation and their effectiveness can vary widely depending on the promoters evaluated. For example, the effectiveness of including a 21 nucleotide sequence derived from the 5'UTR of a lobster *tropomyosin* cDNA sequence was based on extensive experience with expressing several variants of lobster Tropomyosin proteins [459]. Insertion of the *hr1* sequence upstream of promoter regions, on the other hand, significantly improved GFP production from the p10 promoter, but had no effect on either the p6.9 or polh promoters [447]. Interestingly, an earlier report suggested that inserting the *hr1* region downstream and in the reverse orientation in the *polyhedrin* locus contributed to hyperexpression of the foreign gene [462]. Another study reported that the *hr5* sequence strongly enhanced the 39k promoter but significantly impaired expression from the polh promoter [456]. We reasoned that if homologous regions could enhance transcription from any AcMNPV gene, the location and orientation of the ORF with respect to the nearest *hr* may provide insight toward chimeric promoter design. Additionally, sequences surrounding the late gene promoter motif, upstream region, and sequences ad-

jacent to the translation initiation ATG codon as well as the entire 5'UTR were analyzed for characteristics that may determine expression levels of each class.

Initially, the sequence of the 5'UTR for each AcMNPV ORF was extracted from the AcMNPV genome for further analysis of any differences that may be distinguishable between the aforementioned expression classes. The length of the 5'UTR (ie., the number of nucleotides between the TAAG motif and the ATG initiation codon) was not statistically different between classes, indicating that the length of 5'UTR sequence does not directly impact expression levels of AcMNPV genes. The A/T content of the 5'UTRs, however, was significantly higher for the Very High class compared to the other classes. This is consistent with previous analyses of the *p10* and *polyhedrin* 5'UTR sequences [458, 464], however in addition to the polh and p10 5'UTRs, the Very High class includes the *p6.9*, *odv-e18* and *odv-ec27* genes as well. Given that the p6.9 promoter was extremely effective for both intracellular and extracellular protein production, evaluation of the odv-e18 and odv-ec27 promoters may be pertinent. On the other hand, the A/T content between the Low and Medium classes were also significantly different, however the High class was not. It could reasonably be expected that if A/T content of the 5'UTR played a major role in gene expression, the High class would have higher A/T content than the Low and Very Low classes. Similar to the overall length of the 5'UTR, it appears that A/T content may also not be a significant determining factor in gene expression levels.

Next, the AcMNPV genome was annotated according to the expression class each ORF was classified as and the distance in nucleotides was calculated between the 5' end of each *hr* and the 5' end of each AcMNPV ORF. Cursory inspection of the colour coded genome map suggested that the most abundant ORFs were generally found in close proximity

to homologous regions, however no discernible patterns were identifiable for any class. For example, the *p10* ORF, which was found to be enhanced by the insertion of the *hr1* sequence upstream of the promoter, is located ~15 kilobases (kbp) from *hr1* in the *p10-hr1* orientation. The polh promoter, which was not influenced by *hr1* in the same orientation but contributed to hyperexpression of a foreign gene when it was placed downstream and in the reverse orientation of the expression cassette, is located ~4 kbp from the *hr1* region in the *hr1-polh* orientation. Similarly, the *39k* ORF is located ~46 kbp from *hr5*, which strongly enhances its activity. Although the *hr* regions can clearly enhance transcription of several promoters, there do not appear to be any obvious clues as to which specific promoters they may stimulate based on their genomic location or orientation.

The previously reported upstream and downstream octamer sequences were hypothesized to have a role in regulation of late gene expression [453]. The study noted that their spacing was similar in both sequence contexts; the upstream octamer was located 201 and 190 nucleotides upstream of the ATG initiation codon of the *vp39* and *p6.9* ORFs, respectively, whereas the downstream octamer was located 120 and 137 nucleotides from the initiation codons. The placement of these sequences with respect to the TSS was also similar: the 5' end of the upstream sequence is 143 and 148 nucleotides from the TAAG TSS motif (nearest to the initiation codon) for *vp39* and *p6.9* ORFs, respectively. It was reasoned that if these sequences were important for late transcription, they would be enriched at positions adjacent to TSSs in the genome, and the upstream regions of highly expressed ORFs would contain octamers with more optimal sequences and positions, whereas they may be sub-optimal for less expressed ORFs. While allowing for 2 nucleotide mismatches yielded putative matches in the majority of these sequences, the random positioning of

the octamers did not give us confidence that these may be sequences important for gene expression. Indeed, a genome wide search indicated that the upstream and downstream octamers with 2 mismatches appear in the genome (on either strand) at a rate of ~13 and 11 times per 1 kbp, respectively, or ~1.6 and ~1.3 per 250 bp on one DNA strand. These rates are nearly identical to the average number of matches for each octamer in the 225 nucleotide upstream sequences. Given this result and the random dispersion of their locations, it is unlikely that these octamers are late gene transcription regulatory elements.

Relatively little is understood about the sequence factors that govern translation initiation (TIS) of baculovirus mRNAs. Although several studies have demonstrated that the mammalian consensus TIS (ie., Kozak leader sequence) allows translation in the BEVs, the sequences flanking the most highly expressed AcMNPV genes differ significantly from the Kozak sequence [15]. While inserting the consensus Kozak leader sequence in the polh promoter did not improve expression of the human basic fibroblast growth factor, inclusion of bacterial, invertebrate, and A/T-rich synthetic leader sequences have resulted in substantial improvements in recombinant protein production [150, 459, 465, 466]. Interestingly, similar to the polh promoter, the L21 sequence contains an A-rich stretch 9 nucleotides upstream of the initiation codon. This stretch is followed by sequence that is virtually identical to the consensus Kozak sequence, potentially indicating that the Kozak sequence may influence translation initiation of some foreign genes. Aside from this similarity, however, these sequences have relatively few nucleotides in common [459]. Further, a previous study systematically introduced all possible single-nucleotide substitutions in the nucleotides flanking the initiation codon of the *gp64* gene and found that substitutions at only 2 positions within the *gp64* ORF (positions +4 and +5) significantly impacted trans-

lation efficiency. The authors noted, however, that more complex relationships involving multiple nucleotide positions may have larger, additive, effects on translation initiation [467]. Although gene expression is an inherently stochastic process and transcription and translation are independent and discrete events [468], we reasoned that genes that are transcribed less may also have less efficient translation initiation. Accordingly, sequences flanking the ATG initiation codon were extracted from each AcMNPV ORF, and consensus sequences were derived for MSAs from each expression class. Interestingly, although analysis of the *gp64* TIS suggested that nucleotides within the leader sequence upstream of the ATG had little impact on translation initiation, nucleotides at positions −2, −3, and −7 showed significant conservation, indicating they could be important for translation initiation. Although no clear patterns between classes emerged, there are differences that may be worthy of further exploration. For example, position −4 is a highly conserved A or less conserved W (A or T) for the Very High and High classes, respectively. The G to A substitution at the same position increased translation of GP64 by ~10% [467]. Similarly, substitution of G and T with any other nucleotide at positions +4 and +5 within the *gp64* ORF, respectively, increased expression by ~1.3-2.8 fold. While these positions within the *p6.9* ORF are 5'-GT-3', only the *ac-bro* ORF has G at position +4 and no other sequence in the Very High and High classes are occupied by T at position +5. While these observations could be coincidental due to the small number of ORFs in the Very High and High classes, they could be worthy of further experimental scrutiny by introducing these mutations in single substitutions and in combinations to measure their effect on foreign gene expression.

Finally, determining promoter motifs and the underlining mechanisms that control gene

transcription is a major goal of computational biology [469]. However, aside from the late gene TSS motif (5'-TAAG-3'), regulation of late gene expression of AcMNPV is not well understood [16]. Aside from the previously described octamer sequences found upstream of the *vp39* and *p6.9* ORFs, a few studies have identified nucleotide regions that have a large impact on transcription by using linker-scan mutation and deletion strategies to systematically introduce targeted mutations and truncations within the promoter region of AcMNPV promoters [141, 453, 454, 458, 464, 470–473]. Many of these studies have suggested that only ~15-20 nucleotides upstream and downstream of the TSS motif is required for full strength promoter activity, except for the polh and p10 5'UTRs which have binding sites for VLF-1 and are required for full activity [474]. However, each of these studies introduced multiple mutations simultaneously and among adjacent nucleotides, potentially obscuring complex and synergistic relationships between different nucleotide positions. For example, mutation of 7 or 13 nucleotides in the regions 9 to 18 nucleotides upstream or 6 to 19 nucleotides downstream, respectively, of the TSS of the *vp39* promoter resulted in ~50% reduction in transcriptional activity, while mutating 9/10 nucleotides between positions -2 and -11 reduced expression to ~10% compared to the control [454]. In addition to consensus sequences derived from MSAs from each expression class, a MSA and subsequent consensus sequence was calculated for a curated group of promoter sequences that included all of the promoter regions from the High and Very High classes with the exception of the 2 very late promoters, p10 and polh. Previous studies have suggested that the sequences surrounding the TAAG motif of these promoters have a lower affinity for the vRNAP and function as inefficient late gene promoters [454, 474]. Similar to the analysis of flanking TIS sequences, few clear patterns emerged, however some similarites may be

worthy of further exploration. For example, position +7 is A in every sequence in the Curated class, and is T for polh and p10 and W (A or T) for the other classes. Similarly, −11 and −12 are conserved 5'-AA-3' dinucleotide in the Curated class, however they are less conserved as promoter strength decreases. Interestingly, mutations that overlapped both of these sequences resulted in at least 50% reduction of vp39 promoter activity [454], and while these observations may be coincidental due to the small number of highly expressed AcMNPV ORFs, they may be worthy of experimental scrutiny to further examine their importance to late gene expression regulation.

5.6 Concluding remarks

In this study, promoters with transcriptional activity lower than the polh promoter were identified, characterized, and compared by evaluating expression of model cytoplasmic and secreted proteins. Although the polh promoter yielded the highest abundance of GFP, the p6.9 promoter produced nearly twice the amount of SEAP than polh, and the vp39 and 39k promoters yielded similar levels as polh. This adds further confirmation to previous reports in which weaker but earlier promoters resulted in higher yield and/or quality of recombinant proteins than the polh promoter, particularly those that require extensive post-translational processing and secretion. It is expected that the addition of these new promoters to BEVS arsenal may be useful for optimizing co-expression of individual protein constituents of complex biologics such as VLPs. Additionally, we used available transcriptomics and genomics data to scrutinize several determinants that have been previously hypothesized or suggested to be involved in late gene transcription regulation. As

high quality transcriptome and proteomic data becomes more available, this general workflow may be useful in elucidating sequence determinants governing late gene expression to optimize promoter and baculovirus genome design.

5.7 Supporting Results

As noted in the introduction to this chapter, the experiments and analysis presented were conducted as part of a larger project that was sponsored in part by an industrial partner. As such, the results presented here serve only to provide further evidence for the utility of the described promoters for optimizing expression of some foreign genes.

5.7.1 Materials and methods

Plasmid construction, recombinant baculovirus generation, amplification, and quantification

The adeno-associated virus type 2 *rep78* and *rep52* genes were amplified from plasmid pFBDLSR [139] and inserted into each of the promoter transfer plasmids described above to generate rBEVS by homologous recombination. Additional control plasmids consisting of the truncated *Orgyia pseudotsugata* NPV *ie-1* gene promoter upstream of the *rep78* gene and the polh-*rep52* cassette from pFBDLSR were constructed in the same manner. Recombinant baculovirus generation, amplification, and quantification of *rep78* and *rep52*-expressing rBEVS was carried out as described above.

Western blot

Infected cells ($\sim 1.5-2\times10^6$ cells/ml) were collected at ~72 hpi by centrifugation at $500\times g$ for 10 min at 4 °C. The cells were lysed in RIPA buffer (Fisher Scientific), quantified by Pierce BCA assay (Fisher Scientific), and ~20 μg of protein was separated by electrophoresis in 10% TGX Stain-Free precast mini SDS-PAGE gels (Bio-Rad, Mississauga ON) according to manufacturer's directions. After transfer to low fluorescence PVDF membranes, Western blot analysis was performed with anti-AAV2 Rep78 (catalog # 03-65169, American Research Products, Inc, Waltham MA) primary antibody and goat anti-mouse IgG HRP secondary (Bio-Rad) and imaged on a ChemiDoc MP Imager (Bio-Rad). The Image Lab software (Bio-Rad) was used for further image processing.

5.7.2 Results and Discussion

Production of the adeno-associated virus (AAV) gene therapy vector in insect cells using the baculovirus expression vector system was first reported in 2002 as an alternative production platform to the commonly used transient transfection based approach in mammalian cells [139]. The transient mammalian process requires transfection of multiple plasmids to each cell in order to produce AAV particles, which is inefficient and challenging to scale up. Development of cell lines stably replicating the *rep* genes and capable of inducible expression have been described, however cytotoxicity of the Rep proteins results in cell death induced by apoptosis or in low-level expression of *rep78* in these cell lines [475–477]. The BEVS is an attractive system for AAV production since it combines transient expression and scalability.

The first BEVS-based AAV production platform used a truncated ie-1 promoter (Δie1) from *Orgyia pseudotsugata* NPV for expression of the *rep78* gene and the AcMNPV polh promoter for expression of the *rep52* gene in a co-expression rBEV [139]. The Rep-producing baculoviruses produced lower titers than other recombinant baculoviruses, however, and the authors hypothesized that immediate-early expression of the Rep78 protein may be negatively affecting the yield of rBEVS. The Δie1 promoter was therefore replaced with a truncated AcMNPV p10 promoter (Δp10) promoter for low level expression of *rep78* late in the infection cycle. Using the Δp10 promoter yielded a 2-fold increase in the number of vector genomes per cell [140]. Separately, increasing the multiplicity of infection of the Rep-producing rBEV (using the truncated Δie-1 promoter for production of Rep78) improved AAV yields compared with modulating other process parameters, indicating that the Rep78 level is a critical factor in maximizing AAV particle yield [158]. Since the Δp10 promoter offers relatively low level of transcription, new rBEVs were prepared in which the *rep78* and *rep52* genes are transcribed from incrementally stronger promoters to find the optimal expression levels of Rep78 and Rep52 to improve AAV genome yield.

After amplification and quantification of each rBEV, cells were co-infected with each promoter-*rep78* rBEV (MOI = 2) and the polh-*rep52* rBEV (MOI = 2). The cell viability of infected cells at harvest (~66-72 hpi) is given in Figure 5.9A. Interestingly, the viability for cells infected with *rep78*-expressing rBEVs transcribed from 39k, vp39, ctx, and p6.9 promoters was significantly lower than the control (DLSR). After experiment 3, each rBEV was re-amplified and quantified from the original P1 stock to minimize virus degradation. While the viability of the control, Δie1, Δp10, 38k and orf75 rBEVs remained high for experiments 4-6, higher expression levels of Rep78 appeared to result in reduced viability

once again. Supernatant from each infection from experiment 5 was subsequently used to infect fresh cells, which were collected and analyzed by western blot for expression of the Rep proteins. Interestingly, no production of Rep78 was measured for 39k, ctx, and p6.9 promoter rBEVs (Figure 5.9B). It appears that high expression of Rep78 results in reduced viability and instability of the rBEV, leading to loss of transgene expression.

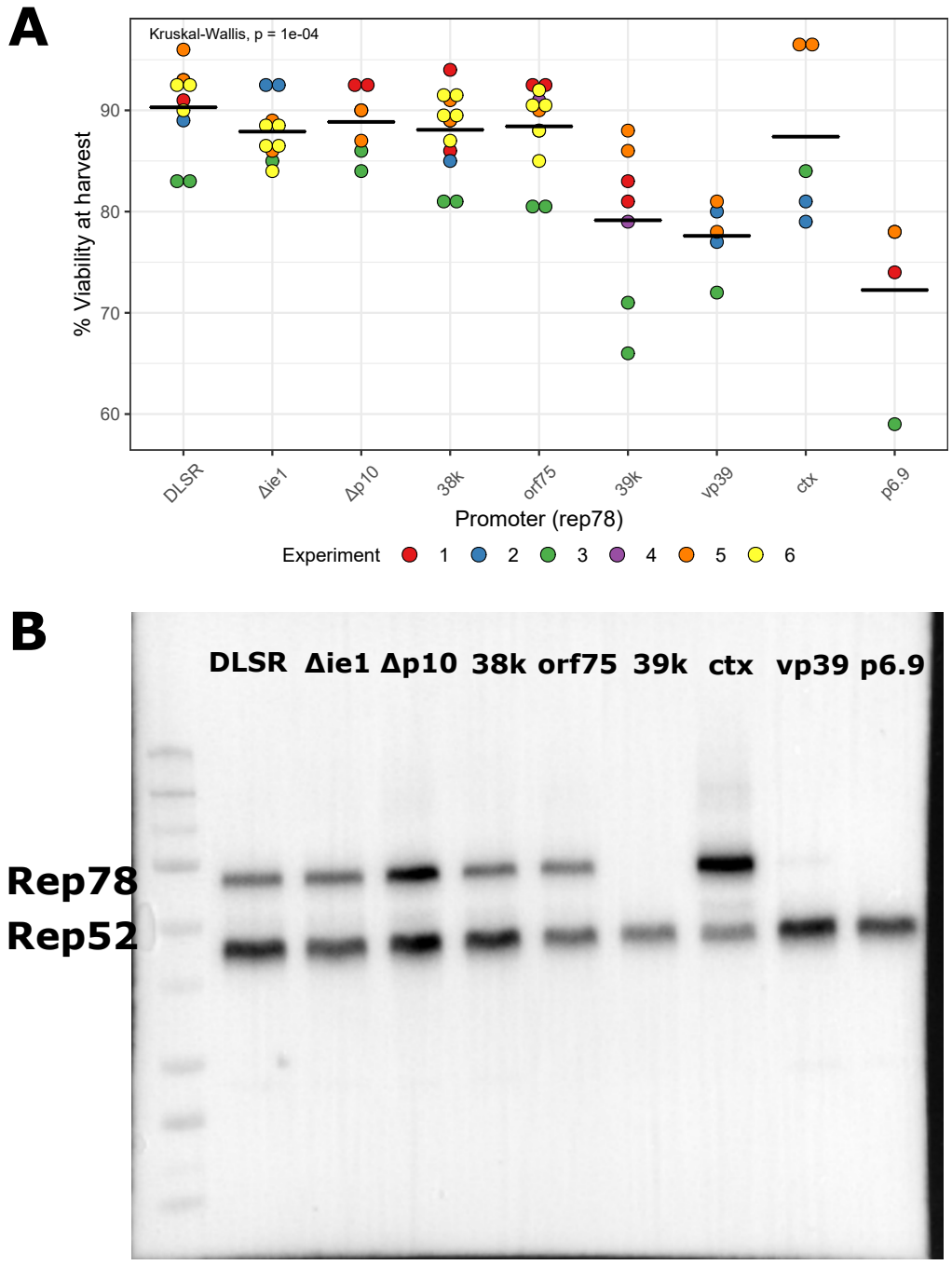


Figure 5.9: **High Rep78 levels appear to result in reduced cell viability. A.** Viability of cells infected with rBEVs expressing *rep52* and *rep78* at harvest. **B.** Western blot analysis of AAV Rep proteins from infected cell lysates at 72 hpi.

Chapter 6

Conclusions and Recommendations

The era of biologics is upon us. While significant productivity improvements were realized in past decades through optimization of growth medium, feeding strategies, and downstream purification workflows, attention must now turn to engineering the production host itself in order to keep up with demand for biologics and develop new classes with novel properties. A common theme presented in this thesis is the relatively little attention directed toward improving the baculovirus expression vectors themselves; significantly, the genome remains virtually wild-type despite demonstrations of significant genome instability during routine passaging and deletion of non-essential genes that improved expression of some foreign genes. Additionally, the use of the polh and p10 promoters are virtualy absolute despite demonstrations that promoters active earlier in the infection cycle improved the yield and/or quality of many recombinant proteins. It is our contention that a general lack of sophisticated genetic engineering technologies for the BEVS is a major contributor to this underdevelopment. For example, the traditional approach for scrutinizing the function of AcMNPV genes involves deletion of the target gene in a bacmid propagating

in *E. coli*, and the development of a *trans*-complementing cell line for generation of initial virus seed stocks. This process involves multiple tedious and time consuming steps that may take several months before infectious virus seed stocks are produced.

The major contribution of this thesis is the development of advanced genetic engineering tools for targeted gene disruption and transcriptional repression. A microplate-based assay was developed to enable efficient scrutiny of the effect of AcMNPV gene deletions on late gene expression and progeny infectious budded virus release, with results achievable in 2 weeks. This assay should be used to identify AcMNPV genes that are not essential for late gene expression or infectious progeny virus, as well as sequences that contribute to genome instability. It is expected that optimization of the BEV genome through removal of non-essential genes and sequences detrimental to stability will enhance foreign gene expression. Following the development of a 'minimal' BEV genome, targeted transcriptional repression can be used to survey its effect on the progression of the infection cycle. It is proposed that delaying or slowing down the progression of the infection cycle might result in delays in the viability drop of cells, thus prolonging the bioprocess and resulting in increased recombinant protein yields. After identifying targets for reduced expression, it may be possible to modulate their expression through mutating sequences involved in transcription regulation. Specific sequences that are responsible for late gene transcription regulation, however, remains poorly understood. While this was not a major goal of this thesis, a framework that may be useful for identifying specifc sequence determinants that may be involved in transcriptional regulation was presented. Several nucleotide positions that appear to be well conserved among the most active promoters of AcMNPV were identified that may be worthy of further analysis.

Finally, optimization of process parameters such as MOI and TOI have resulted in improved production of multiprotein complexes such as VLPs. This strategy generally requires introduction of each foreign gene on multiple monocistronic rBEVs due to a limited number of promoters available for optimization of expression on polycistronic rBEVs. Co-infection strategies, however, suffer from losses in efficiency due to decreases in the proportion of cells that are infected by each rBEV. To address this, several promoters were selected and evaluated to augment the current catalogue of available promoters with expression profiles incrementally weaker than the polh and p10 promoters. This promoter library can be used to optimize the expression of each protein constituent of multiprotein complexes to improve assembly, quality, and overall yield of the target biologic. Design of experiments could be used to develop models to predict the promoter/expression level required to optimize these parameters to reduce the experimental burden of the optimization process. The bioinformatics analyses presented may also be useful for further optimization of transcription or translation of target foreign genes.

References

1. Kretzmer G. Industrial processes with animal cells. *Applied Microbiology and Biotechnology* **59,** 135–142. doi:10.1007/s00253-002-0991-y (2002).
2. Butler M. Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. *Applied Microbiology and Biotechnology* **68,** 283–291. doi:10.1007/s00253-005-1980-8 (2005).
3. Demain A. L. & Vaishnav P. Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances* **27,** 297–306. doi:10.1016/j.biotechadv.2009.01.008 (2009).
4. Zhu J. Mammalian cell protein expression for biopharmaceutical production. *Biotechnology Advances* **30,** 1158–1170. doi:10.1016/j.biotechadv.2011.08.022 (2012).
5. Walsh G. Biopharmaceutical benchmarks 2018. *Nature Biotechnology* **36,** 1136–1145. doi:10.1038/nbt.4305 (2018).
6. Walsh G. Biopharmaceutical benchmarks 2014. *Nature Biotechnology* **32,** 992–1000. doi:10.1038/nbt.3040 (2014).
7. Lai T., Yang Y. & Ng S. Advances in Mammalian Cell Line Development Technologies for Recombinant Protein Production. *Pharmaceuticals* **6,** 579–603. doi:10.3390/ph6050579 (2013).
8. Van Oers M. M., Pijlman G. P. & Vlak J. M. Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. *The Journal of General Virology* **96,** 6–23. doi:10.1099/vir.0.067108-0 (2015).
9. Cox M. M. J. Recombinant protein vaccines produced in insect cells. *Vaccine* **30,** 1759–1766. doi:10.1016/j.vaccine.2012.01.016 (2012).
10. Kost T. A., Condreay J. P. & Jarvis D. L. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotechnology* **23,** 567–575. doi:10.1038/nbt1095 (2005).

11. Assenberg R., Wan P. T., Geisse S. & Mayr L. M. Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology* **23,** 393–402. doi:10.1016/j.sbi.2013.03.008 (2013).

12. Nettleship J. E., Assenberg R., Diprose J. M., Rahman-Huq N. & Owens R. J. Recent advances in the production of proteins in insect and mammalian cells for structural biology. *Journal of Structural Biology* **172,** 55–65. doi:10.1016/j.jsb.2010.02.006 (2010).

13. Lim Y., Wong N. S. C., Lee Y. Y., Ku S. C. Y., Wong D. C. F. & Yap M. G. S. Engineering mammalian cells in bioprocessing – current achievements and future perspectives. *Biotechnology and Applied Biochemistry* **55,** 175–189. doi:10.1042/ba20090363 (2010).

14. Jehle J. A., Blissard G. W., Bonning B. C., Cory J. S., Herniou E. A., Rohrmann G. F., Theilmann D. A., Thiem S. M. & Vlak J. M. On the classification and nomenclature of baculoviruses: A proposal for revision. *Archives of Virology* **151,** 1257–1266. doi:10.1007/s00705-006-0763-6 (2006).

15. Ayres M. D., Howard S. C., Kuzio J., Lopez-Ferber M. & Possee R. D. The complete DNA sequence of Autographa californica nuclear polyhedrosis virus. *Virology* **202,** 586–605. doi:10.1006/viro.1994.1380 (1994).

16. Chen Y.-R., Zhong S., Fei Z., Hashimoto Y., Xiang J. Z., Zhang S. & Blissard G. W. The transcriptome of the baculovirus Autographa californica multiple nucleopolyhedrovirus in Trichoplusia ni cells. *Journal of Virology* **87,** 6391–6405. doi:10.1128/jvi.00194-13 (2013).

17. Monteiro F., Carinhas N., Carrondo M. J. T., Bernal V. & Alves P. M. Toward system-level understanding of baculovirus-host cell interactions: from molecular fundamental studies to large-scale proteomics approaches. *Frontiers in Microbiology* **3,** 391. doi:10.3389/fmicb.2012.00391 (2012).

18. Van Oers M. M., Malarme D., Jore J. & Vlak J. M. Expression of the Autographa californica nuclear polyhedrosis virus p10 gene: effect of polyhedrin gene expression. *Archives of Virology* **123,** 1–11. doi:10.1007/bf01317134 (1992).

19. Braunagel S. C., Parr R., Belyavskyi M. & Summers M. D. Autographa californica nucleopolyhedrovirus infection results in Sf9 cell cycle arrest at G2/M phase. *Virology* **244,** 195–211. doi:10.1006/viro.1998.9097 (1998).

20. Nobiron I., O'Reilly D. R. & Olszewski J. A. Autographa californica nucleopolyhedrovirus infection of Spodoptera frugiperda cells: a global analysis of host gene regulation during infection, using a differential display approach. *Journal of General Virology* **84,** 3029–3039. doi:10.1099/vir.0.19270-0 (2003).

21. Carstens E. B., Tjia S. T. & Doerfler W. Infection of Spodoptera frugiperda cells with Autographa californica nuclear polyhedrosis virus I. Synthesis of intracellular proteins after virus infection. *Virology* **99,** 386–398. doi:10.1016/0042-6822(79)90017-5 (1979).

22. Yu Q., Xiong Y., Gao H., Liu J., Chen Z., Wang Q. & Wen D. Comparative proteomics analysis of Spodoptera frugiperda cells during Autographa californica multiple nucleopolyhedrovirus infection. *Virology Journal* **12,** 115–11. doi:10.1186/s12985-015-0346-9 (2015).

23. Schultz K. L. W. & Friesen P. D. Baculovirus DNA replication-specific expression factors trigger apoptosis and shutoff of host protein synthesis during infection. *Journal of Virology* **83,** 11123–11132. doi:10.1128/jvi.01199-09 (2009).

24. Salem T. Z., Zhang F., Xie Y. & Thiem S. M. Comprehensive analysis of host gene expression in Autographa californica nucleopolyhedrovirus-infected Spodoptera frugiperda cells. *Virology* **412,** 167–178. doi:10.1016/j.virol.2011.01.006 (2011).

25. Chen Y.-R., Zhong S., Fei Z., Gao S., Zhang S., Li Z., Wang P. & Blissard G. W. Transcriptome responses of the host Trichoplusia ni to infection by the baculovirus Autographa californica multiple nucleopolyhedrovirus. *Journal of Virology* **88,** 13781–13797. doi:10.1128/jvi.02243-14 (2014).

26. Kerr J. F., Wyllie A. H. & Currie A. R. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *British Journal of Cancer* **26,** 239–257. doi:10.1038/bjc.1972.33 (1972).

27. Clem R. J. Baculoviruses and apoptosis: the good, the bad, and the ugly. *Cell Death and Differentiation* **8,** 137–143. doi:10.1038/sj.cdd.4400821 (2001).

28. Ikeda M., Yamada H., Hamajima R. & Kobayashi M. Baculovirus genes modulating intracellular innate antiviral immunity of lepidopteran insect cells. *Virology* **435,** 1–13. doi:10.1016/j.virol.2012.10.016 (2013).

29. Mehrabadi M., Hussain M., Matindoost L. & Asgari S. The Baculovirus Antiapoptotic p35 Protein Functions as an Inhibitor of the Host RNA Interference Antiviral Response. *Journal of Virology* **89,** 8182–8192. doi:10.1128/jvi.00802-15 (2015).

30. Bump N. J., Hackett M., Hugunin M., Seshagiri S., Brady K., Chen P., Ferenz C., Franklin S., Ghayur T. & Li P. Inhibition of ICE family proteases by baculovirus antiapoptotic protein p35. *Science* **269,** 1885–1888. doi:10.1126/science.7569933 (1995).

31. Clem R. J., Fechheimer M. & Miller L. K. Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* **254,** 1388–1390. doi:10.1126/science.1962198 (1991).

32. Zoog S. J., Schiller J. J., Wetter J. A., Chejanovsky N. & Friesen P. D. Baculovirus apoptotic suppressor P49 is a substrate inhibitor of initiator caspases resistant to P35 in vivo. *The EMBO Journal* **21,** 5130–5140. doi:10.1038/sj.emboj.7594736 (2002).

33. Seshagiri S. & Miller L. K. Baculovirus inhibitors of apoptosis (IAPs) block activation of Sf-caspase-1. *Proceedings of the National Academy of Sciences of the United States of America* **94,** 13606–13611. doi:10.1073/pnas.94.25.13606 (1997).

34. McLachlin J. R., Escobar J. C., Harrelson J. A., Clem R. J. & Miller L. K. Deletions in the Ac-iap1 gene of the baculovirus AcMNPV occur spontaneously during serial passage and confer a cell line-specific replication advantage. *Virus Research* **81,** 77–91. doi:10.1016/S0168-1702(01)00362-8 (2001).

35. Lyupina Y. V., Dmitrieva S. B., Timokhova A. V., Beljelarskaya S. N., Zatsepina O. G., Evgen'ev M. B. & Mikhailov V. S. An important role of the heat shock response in infected cells for replication of baculoviruses. *Virology* **406,** 336–341. doi:10.1016/j.virol.2010.07.039 (2010).

36. Lyupina Y. V., Zatsepina O. G., Timokhova A. V., Orlova O. V., Kostyuchenko M. V., Beljelarskaya S. N., Evgen'ev M. B. & Mikhailov V. S. New insights into the induction of the heat shock proteins in baculovirus infected insect cells. *Virology* **421,** 34–41. doi:10.1016/j.virol.2011.09.010 (2011).

37. Summers M. D. *Milestones Leading to the Genetic Engineering of Baculoviruses as Expression Vector Systems and Viral Pesticides* doi:10.1016/s0065-3527(06)68001-9 (2006).

38. Grace T. D. Establishment of four strains of cells from insect tissues grown in vitro. *Nature* **195,** 788–789. doi:10.1038/195788a0 (1962).

39. Martignoni M. E. & Scallion R. J. Multiplication in vitro of a Nuclear Polyhedrosis Virus in Insect Amoebocytes. *Nature* **190,** 1133–1134. doi:10.1038/1901133a0 (1961).

40. Gaw Z. Y., Liu N. T. & Zia T. U. Tissue culture methods for cultivation of virus grasserie. *Acta Virologica* **3(Suppl),** 55–60 (1959).

41. Vaughn J. L. & Faulkner P. Susceptibility of an insect tissue culture to infection by virus preparations of the nuclear polyhedrosis of the silkworm (Bombyx mori L.) *Virology* **20,** 484–489. doi:10.1016/0042-6822(63)90098-9 (1963).

42. Hink W. F. Established insect cell line from the cabbage looper, Trichoplusia ni. *Nature* **226,** 466–467. doi:10.1038/226466b0 (1970).

43. Goodwin R. H., Vaughn J. L., Adams J. R. & Louloudes S. J. Replication of a nuclear polyhedrosis virus in an established insect cell line. *Journal of Invertebrate Pathology* **16,** 284–287. doi:10.1016/0022-2011(70)90072-8 (1970).

44. Vail P. V., Sutter G., Jay D. L. & Gough D. Reciprocal infectivity of nuclear polyhedrosis viruses of the cabbage looper and alfalfa looper. *Journal of Invertebrate Pathology* **17,** 383–388. doi:10.1016/0022-2011(71)90013-9 (1971).

45. Kawanishi C. Y. & Paschke J. D. Density-gradient centrifugation of the virions liberated from Rachiplusia ou nuclear polyhedra. *Journal of Invertebrate Pathology* **16,** 89–92. doi:10.1016/0022-2011(70)90211-9 (1970).

46. Summers M. D. & Anderson D. L. Granulosis virus deoxyribonucleic acid: a closed, double-stranded molecule. *Journal of Virology* **9,** 710–713. doi:10.1128/jvi.9.4.710-713.1972 (1972).

47. Henderson J. F., Faulkner P. & MacKinnon E. A. Some Biophysical Properties of Virus Present in Tissue Cultures Infected with the Nuclear Polyhedrosis Virus of Trichoplusia ni. *The Journal of General Virology* **22,** 143–146. doi:10.1099/0022-1317-22-1-143 (1974).

48. Summers M. D. & Volkman L. E. Comparison of biophysical and morphological properties of occluded and extracellular nonoccluded baculovirus from in vivo and in vitro host systems. *Journal of Virology* **17,** 962–972. doi:10.1128/jvi.17.3.962-972.1976 (1976).

49. Ramoska W. A. & Hink W. F. Electron microscope examination of two plaque variants from a nuclear polyhedrosis virus of the alfalfa looper, Autographa californica. *Journal of Invertebrate Pathology* **23,** 197–201. doi:10.1016/0022-2011(74)90184-0 (1974).

50. Hink W. F. & Vail P. V. A plaque assay for titration of alfalfa looper nuclear polyhedrosis virus in a cabbage looper (TN-368) cell line. *Journal of Invertebrate Pathology* **22,** 168–174. doi:10.1016/0022-2011(73)90129-8 (1973).

51. Smith G. E. & Summers M. D. Analysis of baculovirus genomes with restriction endonucleases. *Virology* **89,** 517–527. doi:10.1016/0042-6822(78)90193-9 (1978).

52. Lee H. H. & Miller L. K. Isolation of genotypic variants of Autographa californica nuclear polyhedrosis virus. *Journal of Virology* **27,** 754–767. doi:10.1128/jvi.27.3.754-767.1978 (1978).

53. Burand J. P., Summers M. D. & Smith G. E. Transfection with baculovirus DNA. *Virology* **101,** 286–290. doi:10.1016/0042-6822(80)90505-X (1980).

54. Vlak J. M. & Smith G. E. Orientation of the Genome of Autographa californica Nuclear Polyhedrosis Virus: a Proposal. *Journal of Virology* **41,** 1118–1121. doi:10.1128/jvi.41.3.1118-1121.1982 (1982).

55. Lübbert H., Kruczek I., Tjia S. & Doerfler W. The cloned EcoRI fragments of Autographa californica nuclear polyhedrosis virus DNA. *Gene* **16,** 343–345. doi:10.1016/0378-1119(81)90092-5 (1981).

56. Dobos P. & Cochran M. A. Protein synthesis in cells infected by Autographa californica nuclear polyhedrosis virus (Ac-NPV): the effect of cytosine arabinoside. *Virology* **103,** 446–464. doi:10.1016/0042-6822(80)90203-2 (1980).

57. Vlak J. M., Smith G. E. & Summers M. D. Hybridization Selection and In Vitro Translation of Autographa californica Nuclear Polyhedrosis Virus mRNA. *Journal of Virology* **40,** 762–771. doi:10.1128/jvi.40.3.762-771.1981 (1981).

58. Hooft van Iddekinge B. J. L., Smith G. E. & Summers M. D. Nucleotide sequence of the polyhedrin gene of Autographa californica nuclear polyhedrosis virus. *Virology* **131,** 561–565. doi:10.1016/0042-6822(83)90522-6 (1983).

59. Van der Beek C. P., Saaijer-Riep J. D. & Vlak J. M. On the origin of the polyhedral protein of Autographa californica nuclear polyhedrosis virus Isolation, characterization, and translation of viral messenger RNA. *Virology* **100,** 326–333. doi:10.1016/0042-6822(80)90523-1 (1980).

60. Esche H., Lübbert H., Siegmann B. & Doerfler W. The translational map of the Autographa californica nuclear polyhedrosis virus (AcNPV) genome. *The EMBO Journal* **1,** 1629–1633. doi:10.1002/j.1460-2075.1982.tb01365.x (1982).

61. Smith G. E., Vlak J. M. & Summers M. D. Physical Analysis of Autographa californica Nuclear Polyhedrosis Virus Transcripts for Polyhedrin and 10,000-Molecular-Weight Protein. *Journal of Virology* **45,** 215–225. doi:10.1128/jvi.45.1.215-225.1983 (1983).

62. Rohel D. Z., Cochran M. A. & Faulkner P. Characterization of two abundant mRNAs of Autographa californica nuclear polyhedrosis virus present late in infection. *Virology* **124,** 357–365. doi:10.1016/0042-6822(83)90352-5 (1983).

63. Smith G. E., Vlak J. M. & Summers M. D. In Vitro Translation of Autographa californica Nuclear Polyhedrosis Virus Early and Late mRNAs. *Journal of Virology* **44,** 199–208. doi:10.1128/jvi.44.1.199-208.1982 (1982).

64. Smith G. E., Fraser M. J. & Summers M. D. Molecular engineering of the Autographa californica nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene. *Journal of Virology* **46,** 584–593. doi:10.1128/jvi.46.2.584-593.1983 (1983).

65. Smith G. E., Summers M. D. & Fraser M. J. Production of Human Beta Interferon in Insect Cells Infected with a Baculovirus Expression Vector. *Molecular and Cellular Biology* **3,** 2156–2165. doi:10.1128/mcb.3.12.2156 (1983).

66. Smith G. E., Ju G., Ericson B. L., Moschera J., Lahm H. W., Chizzonite R. & Summers M. D. Modification and secretion of human interleukin 2 produced in insect cells by a baculovirus expression vector. *Proceedings of the National Academy of Sciences of the United States of America* **82,** 8404–8408. doi:10.1073/pnas.82.24.8404 (1985).

67. Kitts P. A., Ayres M. D. & Possee R. D. Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucleic Acids Research* **18,** 5667–5672. doi:10.1093/nar/18.19.5667 (1990).

68. Kitts P. A. & Possee R. D. A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* **14,** 810–817 (1993).

69. Luckow V. A., Lee S. C., Barry G. F. & Olins P. O. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. *Journal of Virology* **67,** 4566–4579. doi:10.1128/jvi.67.8.4566-4579.1993 (1993).

70. Pijlman G. P. Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells. *Journal of General Virology* **84,** 2669–2678. doi:10.1099/vir.0.19438-0 (2003).

71. Hitchman R. B., Possee R. D. & King L. A. Baculovirus expression systems for recombinant protein production in insect cells. *Recent Patents on Biotechnology* **3,** 46–54. doi:10.2174/187220809787172669 (2009).

72. Van Oers M. M. Opportunities and challenges for the baculovirus expression system. *Journal of Invertebrate Pathology* **107 Suppl,** S3–15. doi:10.1016/j.jip.2011.05.001 (2011).

73. Berger I., Fitzgerald D. J. & Richmond T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nature Biotechnology* **22,** 1583–1587. doi:10.1038/nbt1036 (2004).

74. Raj D. B. T. G., Vijayachandran L. S. & Berger I. OmniBac: universal multigene transfer plasmids for baculovirus expression vector systems. *Methods in Molecular Biology* **1091,** 123–130. doi:10.1007/978-1-62703-691-7_7 (2014).

75. Kanai Y., Athmaram T. N., Stewart M. & Roy P. Multiple large foreign protein expression by a single recombinant baculovirus: A system for production of multivalent vaccines. *Protein Expression and Purification* **91,** 77–84. doi:10.1016/j.pep.2013.07.005 (2013).

76. Gaj T., Gersbach C. A. & Barbas C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* **31,** 397–405. doi:10.1016/j.tibtech.2013.04.004 (2013).

77. Vijayachandran L. S., Raj D. B. T. G., Edelweiss E., Gupta K., Maier J., Gordeliy V., Fitzgerald D. J. & Berger I. Gene gymnastics. *Bioengineered* **4,** 279–287. doi:10.4161/bioe.22966 (2014).

78. Kaba S. A., Salcedo A. M., Wafula P. O., Vlak J. M. & van Oers M. M. Development of a chitinase and v-cathepsin negative bacmid for improved integrity of secreted recombinant proteins. *Journal of Virological Methods* **122,** 113–118. doi:10.1016/j.jviromet.2004.07.006 (2004).

79. Hawtin R. E., Zarkowska T., Arnold K., Thomas C. J., Gooday G. W., King L. A., Kuzio J. A. & Possee R. D. Liquefaction of Autographa californica Nucleopolyhedrovirus-Infected Insects Is Dependent on the Integrity of Virus-Encoded Chitinase and Cathepsin Genes. *Virology* **238,** 243–253. doi:10.1006/viro.1997.8816 (1997).

80. Hitchman R. B., Possee R. D., Siaterli E., Richards K. S., Clayton A. J., Bird L. E., Owens R. J., Carpentier D. C. J., King F. L., Danquah J. O., Spink K. G. & King L. A. Improved expression of secreted and membrane-targeted proteins in insect cells. *Biotechnology and Applied Biochemistry* **56,** 85–93. doi:10.1042/ba20090130 (2010).

81. Hitchman R. B., Possee R. D., Crombie A. T., Chambers A., Ho K., Siaterli E., Lissina O., Sternard H., Novy R., Loomis K., Bird L. E., Owens R. J. & King L. A. Genetic modification of a baculovirus vector for increased expression in insect cells. *Cell Biology and Toxicology* **26,** 57–68. doi:10.1007/s10565-009-9133-y (2010).

82. Thomas C. J., Brown H. L., Hawes C. R., Lee B. Y., Min M.-K., King L. A. & Possee R. D. Localization of a Baculovirus-Induced Chitinase in the Insect Cell Endoplasmic Reticulum. *Journal of Virology* **72,** 10207–10212. doi:10.1128/jvi.72.12.10207-10212.1998 (1998).

83. Fernandes F., Teixeira A. P., Carinhas N., Carrondo M. J. T. & Alves P. M. Insect cells as a production platform of complex virus-like particles. *Expert Review of Vaccines* **12,** 225–236. doi:10.1586/erv.12.153 (2013).

84. Yee C. M., Zak A. J., Hill B. D. & Wen F. The Coming Age of Insect Cells for Manufacturing and Development of Protein Therapeutics. *Industrial & Engineering Chemistry Research* **57,** 10061–10070. doi:10.1021/acs.iecr.8b00985 (2018).

85. Marek M., van Oers M. M., Devaraj F. F., Vlak J. M. & Merten O.-W. Engineering of baculovirus vectors for the manufacture of virion-free biopharmaceuticals. *Biotechnology and Bioengineering* **108,** 1056–1067. doi:10.1002/bit.23028 (2011).

86. Chaves L. C. S., Ribeiro B. M. & Blissard G. W. Production of GP64-free virus-like particles from baculovirus-infected insect cells. *The Journal of General Virology* **99,** 265–274. doi:10.1099/jgv.0.001002 (2018).

87. Sokolenko S., George S., Wagner A., Tuladhar A., Andrich J. M. S. & Aucoin M. G. Co-expression vs. co-infection using baculovirus expression vectors in insect cell culture: Benefits and drawbacks. *Biotechnology Advances* **30,** 766–781. doi:10.1016/j.biotechadv.2012.01.009 (2012).

88. Kool M., Voncken J. W., van Lier F. L., Tramper J. & Vlak J. M. Detection and analysis of Autographa californica nuclear polyhedrosis virus mutants with defective interfering properties. *Virology* **183,** 739–746. doi:10.1016/0042-6822(91)91003-Y (1991).

89. Kool M., Goldbach R. W. & Vlak J. M. A putative non-hr origin of DNA replication in the HindIII-K fragment of Autographa californica multiple nucleocapsid nuclear polyhedrosis virus. *The Journal of general virology* **75,** 3345–3352. doi:10.1099/0022-1317-75-12-3345 (1994).

90. Kool M., Voeten J. T., Goldbach R. W., Tramper J. & Vlak J. M. Identification of seven putative origins of Autographa californica multiple nucleocapsid nuclear polyhedrosis virus DNA replication. *Journal of General Virology* **74 (Pt 12),** 2661–2668. doi:10.1099/0022-1317-74-12-2661 (1993).

91. Pijlman G. P., Vermeesch A. M. G. & Vlak J. M. Cell line-specific accumulation of the baculovirus non-hr origin of DNA replication in infected insect cells. *Journal of Invertebrate Pathology* **84,** 214–219. doi:10.1016/j.jip.2003.10.005 (2003).

92. Habib S. & Hasnain S. E. Differential Activity of Two Non-hr Origins during Replication of the Baculovirus Autographa californica Nuclear Polyhedrosis Virus Genome. *Journal of Virology* **74,** 5182–5189. doi:10.1128/JVI.74.11.5182-5189.2000 (2000).

93. Pijlman G. P., Roode E. C., Fan X., Roberts L. O., Belsham G. J., Vlak J. M. & Oers M. M. v. Stabilized baculovirus vector expressing a heterologous gene and GP64 from a single bicistronic transcript. *Journal of Biotechnology* **123,** 13–21. doi:10.1016/j.jbiotec.2005.10.022 (2006).

94. Pelletier J. & Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334,** 320–325. doi:10.1038/334320a0 (1988).

95. Miele S. A. B., Garavaglia M. J., Belaich M. N. & Ghiringhelli P. D. Baculovirus: Molecular Insights on Their Diversity and Conservation. *International Journal of Evolutionary Biology* **2011,** 379–424. doi:10.4061/2011/379424 (2011).

96. Cohen D. P. A., Marek M., Davies B. G., Vlak J. M. & Oers M. M. v. Encyclopedia of Autographa californica nucleopolyhedrovirus genes. *Virologica Sinica* **24,** 359–414. doi:10.1007/s12250-009-3059-7 (2009).

97. Rohrmann G. F. *Baculovirus Molecular Biology* 4th ed. (National Center for Biotechnology Information, Bethesda (MD), 2019).

98. Ono C., Kamagata T., Taka H., Sahara K., Asano S.-i. & Bando H. Phenotypic grouping of 141 BmNPVs lacking viral gene sequences. *Virus Research* **165,** 197–206. doi:10.1016/j.virusres.2012.02.016 (2012).

99. Hou D., Chen X. & Zhang L.-K. Proteomic Analysis of Mamestra Brassicae Nucleopolyhedrovirus Progeny Virions from Two Different Hosts. *PLoS ONE* **11,** e0153365. doi:10.1371/journal.pone.0153365 (2016).

100. Lu A. & Miller L. K. Species-specific effects of the hcf-1 gene on baculovirus virulence. *Journal of Virology* **70,** 5123–5130. doi:10.1128/jvi.70.8.5123-5130.1996 (1996).

101. Tachibana A., Hamajima R., Tomizaki M., Kondo T., Nanba Y., Kobayashi M., Yamada H. & Ikeda M. HCF-1 encoded by baculovirus AcMNPV is required for productive nucleopolyhedrovirus infection of non-permissive Tn368 cells. *Scientific Reports* **7,** 3807. doi:10.1038/s41598-017-03710-z (2017).

102. Iwanaga M., Kurihara M., Kobayashi M. & Kang W. Characterization of Bombyx mori Nucleopolyhedrovirus orf68 Gene That Encodes a Novel Structural Protein of Budded Virus. *Virology* **297,** 39–47. doi:10.1006/viro.2002.1443 (2002).

103. Gauthier D., Thirunavukkarasu K., Faris B. L., Russell D. L. & Weaver R. F. Characterization of an Autographa californica multiple nucleopolyhedrovirus dual mutant: ORF82 is required for budded virus production, and a point mutation in LEF-8 alters late and abolishes very late transcription. *Journal of General Virology* **93,** 364–373. doi:10.1099/vir.0.037028-0 (2012).

104. Chavez-Pena C. & Kamen A. A. RNA interference technology to improve the baculovirus-insect cell expression system. *Biotechnology Advances* **36,** 443–451. doi:10.1016/j.biotechadv.2018.01.008 (2018).

105. Ghosh S., Kakumani P. K., Kumar A., Malhotra P., Mukherjee S. K. & Bhatnagar R. K. Genome wide screening of RNAi factors of Sf21 cells reveal several novel pathway associated proteins. *BMC Genomics* **15,** 775. doi:10.1186/1471-2164-15-775 (2014).

106. Terenius O. *et al.* RNA interference in Lepidoptera: An overview of successful and unsuccessful studies and implications for experimental design. *Journal of Insect Physiology* **57,** 231–245. doi:10.1016/j.jinsphys.2010.11.006 (2011).

107. Hebert C. G., Valdes J. J. & Bentley W. E. In vitro and in vivo RNA interference mediated suppression of Tn-caspase-1 for improved recombinant protein production in High Five™ cell culture with the baculovirus expression vector system. *Biotechnology and Bioengineering* **104,** 390–399. doi:10.1002/bit.22411 (2009).

108. Hebert C. G., Valdes J. J. & Bentley W. E. Investigating apoptosis: characterization and analysis of Trichoplusia ni-caspase-1 through overexpression and RNAi mediated silencing. *Insect Biochemistry and Molecular Biology* **39,** 113–124. doi:10.1016/j.ibmb.2008.10.009 (2009).

109. Lai Y.-K., Hsu J. T.-A., Chu C.-C., Chang T.-Y., Pan K.-L. & Lin C.-C. Enhanced recombinant protein production and differential expression of molecular chaperones in sf-caspase-1-repressed stable cells after baculovirus infection. *BMC Biotechnology* **12,** 83–1. doi:10.1186/1472-6750-12-83 (2012).

110. Wang Q., Zhou Y., Chen K. & Ju X. Suppression of Bm-Caspase-1 Expression in BmN Cells Enhances Recombinant Protein Production in a Baculovirus Expression Vector System. *Molecular Biotechnology* **58,** 319–327. doi:10.1007/s12033-016-9931-4 (2016).

111. Zhang X., Xu K., Ou Y., Xu X. & Chen H. Development of a baculovirus vector carrying a small hairpin RNA for suppression of sf-caspase-1 expression and improvement of recombinant protein production. *BMC Biotechnology* **18,** 24. doi:10.1186/s12896-018-0434-1 (2018).

112. Kim E. J., Kramer S. F., Hebert C. G., Valdes J. J. & Bentley W. E. Metabolic engineering of the baculovirus-expression system via inverse "shotgun" genomic analysis and RNA interference (dsRNA) increases product yield and cell longevity. *Biotechnology and Bioengineering* **98,** 645–654. doi:10.1002/bit.21353 (2007).

113. Lee H. S., Lee H. Y., Kim Y. J., Jung H. D., Choi K. J., Yang J. M., Kim S. S. & Kim K. Small interfering (Si) RNA mediated baculovirus replication reduction without affecting target gene expression. *Virus Research* **199,** 68–76. doi:10.1016/j.virusres.2015.01.015 (2015).

114. Salem T. Z., Zhang F. & Thiem S. M. Reduced expression of Autographa californica nucleopolyhedrovirus ORF34, an essential gene, enhances heterologous gene expression. *Virology* **435,** 225–238. doi:10.1016/j.virol.2012.10.022 (2013).

115. Wu H.-C., Hebert C. G., Hung C.-W., Quan D. N., Carter K. K. & Bentley W. E. Tuning cell cycle of insect cells for enhanced protein production. *Journal of Biotechnology* **168,** 55–61. doi:10.1016/j.jbiotec.2013.08.017 (2013).

116. Yun E.-Y., Goo T.-W., Kim S.-W., Choi K.-H., Hwang J.-S., Kang S.-W. & Kwon O.-Y. Changes in cellular secretory processing during baculovirus infection. *Biotechnology Letters* **27,** 1041–1045. doi:10.1007/s10529-005-8108-1 (2005).

117. Ailor E. & Betenbaugh M. J. Modifying secretion and post-translational processing in insect cells. *Current Opinion in Biotechnology* **10,** 142–145. doi:10.1016/S0958-1669(99)80024-X (1999).

118. Ailor E. & Betenbaugh M. J. Overexpression of a cytosolic chaperone to improve solubility and secretion of a recombinant IgG protein in insect cells. *Biotechnology and Bioengineering* **58,** 196–203. doi:10.1002/(sici)1097-0290(19980420)58:2/3<196::aid-bit12>3.0.co;2-b (1998).

119. Tate C. G., Whiteley E. & Betenbaugh M. J. Molecular chaperones improve functional expression of the serotonin (5-hydroxytryptamine) transporter in insect cells. *Biochemical Society Transactions* **27,** 932–936. doi:10.1042/bst0270932 (1999).

120. Zhang L., Wu G., Tate C. G., Lookene A. & Olivecrona G. Calreticulin Promotes Folding/Dimerization of Human Lipoprotein Lipase Expressed in Insect Cells (Sf21). *Journal of Biological Chemistry* **278,** 29344–29351. doi:10.1074/jbc.m300455200 (2003).

121. Culina S., Lauvau G., Gubler B. & Endert P. M. v. Calreticulin promotes folding of functional human leukocyte antigen class I molecules in vitro. *Journal of Biological Chemistry* **279,** 54210–54215. doi:10.1074/jbc.m410841200 (2004).

122. Hsu T. A. & Betenbaugh M. J. Coexpression of molecular chaperone BiP improves immunoglobulin solubility and IgG secretion from Trichoplusia ni insect cells. *Biotechnology Progress* **13,** 96–104. doi:10.1021/bp960088d (1997).

123. Teng C.-Y., Chang S.-L., van Oers M. M. & Wu T.-Y. Enhanced Protein Secretion From Insect Cells by Co-Expression of the Chaperone Calreticulin and Translation Initiation Factor eIF4E. *Molecular Biotechnology* **54,** 68–78. doi:10.1007/s12033-012-9545-4 (2013).

124. Jarvis D. L., Kawar Z. S. & Hollister J. R. Engineering N-glycosylation pathways in the baculovirus-insect cell system. *Current Opinion in Biotechnology* **9,** 528–533. doi:10.1016/s0958-1669(98)80041-4 (1998).

125. Hollister J., Grabenhorst E., Nimtz M., Conradt H. & Jarvis D. L. Engineering the protein N-glycosylation pathway in insect cells for production of biantennary, complex N-glycans. *Biochemistry* **41,** 15093–15104. doi:10.1021/bi026455d (2002).

126. Hollister J. R. & Jarvis D. L. Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian B1, 4-galactosyltransferase and a2, 6-sialyltransferase genes. *Glycobiology* **11,** 1–9. doi:10.1093/glycob/11.1.1 (2001).

127. Aumiller J. J., Hollister J. R. & Jarvis D. L. A transgenic insect cell line engineered to produce CMP-sialic acid and sialylated glycoproteins. **13,** 497–507. doi:10.1093/glycob/cwg051 (2003).

128. Mabashi-Asazuma H., Shi X., Geisler C., Kuo C.-W., Khoo K.-H. & Jarvis D. L. Impact of a human CMP-sialic acid transporter on recombinant glycoprotein sialylation in glycoengineered insect cells. *Glycobiology* **23,** 199–210. doi:10.1093/glycob/cws143 (2013).

129. Aumiller J. J., Mabashi-Asazuma H., Hillar A., Shi X. & Jarvis D. L. A new glycoengineered insect cell line with an inducibly mammalianized protein N-glycosylation pathway. *Glycobiology* **22,** 417–428. doi:10.1093/glycob/cwr160 (2012).

130. Mabashi-Asazuma H. & Jarvis D. L. CRISPR-Cas9 vectors for genome editing and host engineering in the baculovirus-insect cell system. *Proceedings of the National Academy of Sciences of the United States of America* **114,** 9068–9073. doi:10.1073/pnas.1705836114 (2017).

131. Martínez-Solís M., Herrero S. & Targovnik A. M. Engineering of the baculovirus expression system for optimized protein production. *Applied Microbiology and Biotechnology* **103,** 113–123. doi:10.1007/s00253-018-9474-7 (2019).

132. Jarvis D. L., Weinkauf C. & Guarino L. A. Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expression and Purification* **8,** 191–203. doi:10.1006/prep.1996.0092 (1996).

133. Grabherr R., Ernst W., Doblhoff-Dier O., Sara M. & Katinger H. Expression of foreign proteins on the surface of Autographa californica nuclear polyhedrosis virus. *Biotechniques* **22,** 730–735. doi:10.2144/97224rr02 (1997).

134. Higgins M. K., Demir M. & Tate C. G. Calnexin co-expression and the use of weaker promoters increase the expression of correctly assembled Shaker potassium channel in insect cells. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1610,** 124–132. doi:10.1016/s0005-2736(02)00715-0 (2003).

135. Bleckmann M., Fritz M. H. Y., Bhuju S., Jarek M., Schürig M., Geffers R., Benes V., Besir H. & Heuvel J. v. d. Genomic Analysis and Isolation of RNA Polymerase II Dependent Promoters from Spodoptera frugiperda. *PLoS ONE* **10,** e0132898–16. doi:10.1371/journal.pone.0132898 (2015).

136. Bleckmann M., Schürig M., Chen F.-F., Yen Z.-Z., Lindemann N., Meyer S., Spehr J. & Heuvel J. v. d. Identification of Essential Genetic Baculoviral Elements for Recombinant Protein Expression by Transactivation in Sf21 Insect Cells. *PLoS ONE* **11,** e0149424. doi:10.1371/journal.pone.0149424 (2016).

137. López-Vidal J., Gómez-Sebastián S., Sánchez-Ramos I. & Escribano J. M. Characterization of a Trichoplusia ni hexamerin-derived promoter in the AcMNPV baculovirus vector. *Journal of Biotechnology* **165,** 201–208. doi:10.1016/j.jbiotec.2013.03.012 (2013).

138. Blissard G. W., Kogan P. H., Wei R. & Rohrmann G. F. A synthetic early promoter from a baculovirus: Roles of the TATA box and conserved start site CAGT sequence in basal levels of transcription. *Virology* **190,** 783–793. doi:10.1016/0042-6822(92)90916-d (1992).

139. Urabe M., Ding C. & Kotin R. M. Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. *Human Gene Therapy* **13,** 1935–1943. doi:10.1089/10430340260355347 (2004).

140. Urabe M., Nakakura T., Xin K.-Q., Obara Y., Mizukami H., Kume A., Kotin R. M. & Ozawa K. Scalable generation of high-titer recombinant adeno-associated virus type 5 in insect cells. *Journal of Virology* **80,** 1874–1885. doi:10.1128/jvi.80.4.1874-1885.2006 (2006).

141. Rankin C., Ooi B. G. & Miller L. K. Eight base pairs encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin gene expression. *Gene* **70,** 39–49. doi:10.1016/0378-1119(88)90102-3 (1988).

142. Kato T., Manohar S. L., Kanamasa S., Ogata M. & Park E. Y. Improvement of the transcriptional strength of baculovirus very late polyhedrin promoter by repeating its untranslated leader sequences and coexpression with the primary transactivator. *Journal of Bioscience and Bioengineering* **113,** 694–696. doi:10.1016/j.jbiosc.2012.01.010 (2012).

143. Martínez-Solís M., Gómez-Sebastián S., Escribano J. M., Jakubowska A. K. & Herrero S. A novel baculovirus-derived promoter with high activity in the baculovirus expression system. *PeerJ* **4,** e2183. doi:10.7717/peerj.2183 (2016).

144. Peros I. G., Cerrudo C. S., Pilloff M. G., Belaich M. N., Lozano M. E. & Ghiringhelli P. D. Advances in the Bioinformatics Knowledge of mRNA Polyadenylation in Baculovirus Genes. *Viruses* **12,** 1395. doi:10.3390/v12121395 (2020).

145. Moldován N., Tombácz D., Szűcs A., Csabai Z., Balázs Z., Kis E., Molnár J. & Boldogkői Z. Third-generation Sequencing Reveals Extensive Polycistronism and Transcriptional Overlapping in a Baculovirus. *Scientific Reports* **8,** 8604. doi:10.1038/s41598-018-26955-8 (2018).

146. Jin J. & Guarino L. A. 3'-end formation of baculovirus late RNAs. *Journal of Virology* **74,** 8930–8937. doi:10.1128/JVI.74.19.8930-8937.2000 (2000).

147. Salem T. Z., Seaborn C. P., Turney C. M., Xue J., Shang H. & Cheng X.-W. The Influence of SV40 polyA on Gene Expression of Baculovirus Expression Vector Systems. *PLoS ONE* **10,** e0145019. doi:10.1371/journal.pone.0145019 (2015).

148. Shang H., Garretson T. A., Kumar C. S., Dieter R. F. & Cheng X.-W. Improved pFastBac™ donor plasmid vectors for higher protein production using the Bac-to-Bac® baculovirus expression vector system. *Journal of Biotechnology* **255,** 37–46. doi:10.1016/j.jbiotec.2017.06.397 (2017).

149. Van Oers M. M., Vlak J. M., Voorma H. O. & Thomas A. A. M. Role of the 3′ untranslated region of baculovirus p10 mRNA in high-level expression of foreign genes. *Journal of General Virology* **80,** 2253–2262. doi:10.1099/0022-1317-80-8-2253 (1999).

150. Liu Y., Zhang Y., Yao L., Hao H., Fu X., Yang Z. & Du E. Enhanced production of porcine circovirus type 2 (PCV2) virus-like particles in Sf9 cells by translational enhancers. *Biotechnology Letters* **37,** 1765–1771. doi:10.1007/s10529-015-1856-7 (2015).

151. Chaabihi H., Ogliastro M. H., Martin M., Giraud C., Devauchelle G. & Cerutti M. Competition between baculovirus polyhedrin and p10 gene expression during infection of insect cells. *Journal of Virology* **67.** doi:10.1128/jvi.67.5.2664-2671.1993 (1993).

152. George S., Jauhar A. M., Mackenzie J., Kieβlich S. & Aucoin M. G. Temporal characterization of protein production levels from baculovirus vectors coding for GFP and RFP genes under non-conventional promoter control. *Biotechnology and Bioengineering* **112,** 1822–1831. doi:10.1002/bit.25600 (2015).

153. Drugmand J.-C., Schneider Y.-J. & Agathos S. N. Insect cells as factories for biomanufacturing. *Biotechnology Advances* **30,** 1140–1157. doi:10.1016/j.biotechadv.2011.09.014 (2012-09).

154. Contreras-Gómez A., Sánchez-Mirón A., García-Camacho F., Molina-Grima E. & Chisti Y. Protein production using the baculovirus-insect cell expression system. *Biotechnology Progress* **30,** 1–18. doi:10.1002/btpr.1842 (2014).

155. Wong K. T., Peter C. H., Greenfield P. F., Reid S. & Nielsen L. K. Low multiplicity infection of insect cells with a recombinant baculovirus: The cell yield concept. *Biotechnology and Bioengineering* **49,** 659–666. doi:10.1002/(sici)1097-0290(19960320)49:6<659::aid-bit7>3.0.co;2-n (1996).

156. Crawford S. E., Labbe M., Cohen J., Burroughs M. H., Zhou Y. J. & Estes M. K. Characterization of virus-like particles produced by the expression of rotavirus capsid proteins in insect cells. *Journal of Virology* **68,** 5945–5952. doi:10.1128/jvi.68.9.5945-5952.1994 (1994).

157. Palomares L. A., López S. & Ramírez O. T. Strategies for manipulating the relative concentration of recombinant rotavirus structural proteins during simultaneous production by insect cells. *Biotechnology and Bioengineering* **78,** 635–644. doi:10.1002/bit.10243 (2002).

158. Aucoin M. G., Perrier M. & Kamen A. A. Production of adeno-associated viral vectors in insect cells using triple infection: Optimization of baculovirus concentration ratios. *Biotechnology and Bioengineering* **95,** 1081–1092. doi:10.1002/bit.21069 (2006).

159. Meghrous J., Aucoin M. G., Jacob D., Chahal P. S., Arcand N. & Kamen A. A. Production of recombinant adeno-associated viral vectors using a baculovirus/insect cell suspension culture system: from shake flasks to a 20-L bioreactor. *Biotechnology Progress* **21,** 154–160. doi:10.1021/bp049802e (2005).

160. Radford K. M., Cavegn C., Bertrand M., Bernard A. R., Reid S. & Greenfield P. F. The indirect effects of multiplicity of infection on baculovirus expressed proteins in insect cells: secreted and non-secreted products. *Cytotechnology* **24,** 73–81. doi:10.1023/a:1007962903435 (1997).

161. Zwart M. P., Erro E., van Oers M. M., Visser J. A. G. M. d. & Vlak J. M. Low multiplicity of infection in vivo results in purifying selection against baculovirus deletion mutants. *The Journal of General Virology* **89,** 1220–1224. doi:10.1099/vir.0.83645-0 (2008).

162. Krell P. J. Passage effect of virus infection in insect cells. *Cytotechnology* **20,** 125–137. doi:10.1007/bf00350393 (1996).

163. Jäger V. Perfusion bioreactors for the production of recombinant proteins in insect cells. *Cytotechnology* **20,** 191–198. doi:10.1007/bf00350399 (1996).

164. Ikonomou L., Schneider Y. J. & Agathos S. N. Insect cell culture for industrial production of recombinant proteins. *Applied Microbiology and Biotechnology* **62,** 1–20. doi:10.1007/s00253-003-1223-9 (2003).

165. Carinhas N., Bernal V., Monteiro F., Carrondo M. J. T., Oliveira R. & Alves P. M. Improving baculovirus production at high cell density through manipulation of energy metabolism. *Metabolic Engineering* **12,** 39–52. doi:10.1016/j.ymben.2009.08.008 (2010).

166. Bernal V., Carinhas N., Yokomizo A. Y., Carrondo M. J. T. & Alves P. M. Cell density effect in the baculovirus-insect cells system: a quantitative analysis of energetic metabolism. *Biotechnology and Bioengineering* **104,** 162–180. doi:10.1002/bit.22364 (2009).

167. Liu F., Wu X., Li L., Liu Z. & Wang Z. Use of baculovirus expression system for generation of virus-like particles: Successes and challenges. *Protein Expression and Purification* **90,** 104–116. doi:10.1016/j.pep.2013.05.009 (2013).

168. Roldão A., Mellado M. C. M., Castilho L. R., Carrondo M. J. T. & Alves P. M. Virus-like particles in vaccine development. *Expert Review of Vaccines* **9,** 1149–1176. doi:10.1586/erv.10.115 (2010).

169. Liu C., Zhou Q., Li Y., Garner L. V., Watkins S. P., Carter L. J., Smoot J., Gregg A. C., Daniels A. D., Jervey S. & Albaiu D. Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Central Science* **6,** 315–331. doi:10.1021/acscentsci.0c00272 (2020).

170. Shin M. D., Shukla S., Chung Y. H., Beiss V., Chan S. K., Ortega-Rivera O. A., Wirth D. M., Chen A., Sack M., Pokorski J. K. & Steinmetz N. F. COVID-19 vaccine development and a potential nanomaterial path forward. *Nature Nanotechnology* **15,** 646–655. doi:10.1038/s41565-020-0737-y (2020).

171. Keech C. *et al.* Phase 1–2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. *New England Journal of Medicine* **383,** 2320–2332. doi:10.1056/nejmoa2026920 (2020).

172. Tian J.-H. *et al.* SARS-CoV-2 spike glycoprotein vaccine candidate NVX-CoV2373 immunogenicity in baboons and protection in mice. *Nature Communications* **12,** 372. doi:10.1038/s41467-020-20653-8 (2021).

173. Sun Y., Carrion R., Ye L., Wen Z., Ro Y. T. & Brasky K. Protection against lethal challenge by Ebola virus-like particles produced in insect cells. *Virology* (2009).

174. Ye L., Lin J., Sun Y., Bennouna S., Lo M., Wu Q., Bu Z., Pulendran B., Compans R. W. & Yang C. Ebola virus-like particles produced in insect cells exhibit dendritic cell stimulating activity and induce neutralizing antibodies. *Virology* **351,** 260–270. doi:10.1016/j.virol.2006.03.021 (2006).

175. Takehara K., Ireland D. & Bishop D. H. Co-expression of the hepatitis B surface and core antigens using baculovirus multiple expression vectors. *The Journal of general virology* **69,** 2763–2777. doi:10.1099/0022-1317-69-11-2763 (1988).

176. Overton H. A., Fujii Y., Price I. R. & Jones I. M. The protease and gag gene products of the human immunodeficiency virus: Authentic cleavage and post-translational modification in an insect cell expression system. *Virology* **170,** 107–116. doi:10.1016/0042-6822(89)90357-7 (1989).

177. Wang B.-Z., Liu W., Kang S.-M., Alam M., Huang C., Ye L., Sun Y., Li Y., Kothe D. L., Pushko P., Dokland T., Haynes B. F., Smith G., Hahn B. H. & Compans R. W. Incorporation of high levels of chimeric human immunodeficiency virus envelope glycoproteins into virus-like particles. *Journal of Virology* **81,** 10869–10878. doi:10.1128/jvi.00542-07 (2007).

178. Buonaguro L., Tornesello M. L., Tagliamonte M., Gallo R. C., Wang L. X., Kamin-Lewis R., Abdelwahab S., Lewis G. K. & Buonaguro F. M. Baculovirus-derived human immunodeficiency virus type 1 virus-like particles activate dendritic cells and induce ex vivo T-cell responses. *Journal of Virology* **80,** 9134–9143. doi:10.1128/jvi.00050-06 (2006).

179. Sailaja G., Skountzou I., Quan F.-S., Compans R. W. & Kang S.-M. Human immunodeficiency virus-like particles activate multiple types of immune cells. *Virology* **362,** 331–341. doi:10.1016/j.virol.2006.12.014 (2007).

180. Buonaguro L., Buonaguro F. M., Tornesello M. L., Mantas D., Beth-Giraldo E., Wagner R., Michelson S., Prevost M. C., Wolf H. & Giraldo G. High efficient production of Pr55(gag) virus-like particles expressing multiple HIV-1 epitopes, including a gp120 protein derived from an Ugandan HIV-1 isolate of subtype A. *Antiviral Research* **49,** 35–47. doi:10.1016/S0166-3542(00)00136-4 (2001).

181. Ho Y., Lin P.-H., Liu C. Y. Y., Lee S.-P. & Chao Y.-C. Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochemical and Biophysical Research Communications* **318,** 833–838. doi:10.1016/j.bbrc.2004.04.111 (2004).

182. Mortola E. & Roy P. Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system. *FEBS Letters* **576,** 174–178. doi:10.1016/j.febslet.2004.09.009 (2004).

183. Latham T. & Galarza J. M. Formation of Wild-Type and Chimeric Influenza Virus-Like Particles following Simultaneous Expression of Only Four Structural Proteins. *Journal of Virology* **75,** 6154–6165. doi:10.1128/jvi.75.13.6154-6165.2001 (2001).

184. Pushko P., Tumpey T. M., Hoeven N. V., Belser J. A., Robinson R., Nathan M., Smith G., Wright D. C. & Bright R. A. Evaluation of influenza virus-like particles and Novasome adjuvant as candidate vaccine for avian influenza. *Vaccine* **25,** 4283–4290. doi:10.1016/j.vaccine.2007.02.059 (2007).

185. Bright R. A., Carter D. M., Daniluk S., Toapanta F. R., Ahmad A., Gavrilov V., Massare M., Pushko P., Mytle N., Rowe T., Smith G. & Ross T. M. Influenza virus-like particles elicit broader immune responses than whole virion inactivated influenza virus or recombinant hemagglutinin. *Vaccine* **25,** 3871–3878. doi:10.1016/j.vaccine.2007.01.106 (2007).

186. Wen Z., Ye L., Gao Y., Pan L., Dong K., Bu Z., Compans R. W. & Yang C. Immunization by influenza virus-like particles protects aged mice against lethal influenza virus challenge. *Antiviral Research* **84,** 215–224. doi:10.1016/j.antiviral.2009.09.005 (2009).

187. Kang S.-M., Yoo D.-G., Lipatov A. S., Song J.-M., Davis C. T., Quan F.-S., Chen L.-M., Donis R. O. & Compans R. W. Induction of long-term protective immune responses by influenza H5N1 virus-like particles. *PLoS ONE* **4,** e4667. doi:10.1371/journal.pone.0004667 (2009).

188. Haynes J. R., Dokken L., Wiley J. A., Cawthon A. G., Bigger J., Harmsen A. G. & Richardson C. Influenza-pseudotyped Gag virus-like particle vaccines provide broad protection against highly pathogenic avian influenza challenge. *Vaccine* **27,** 530–541. doi:10.1016/j.vaccine.2008.11.011 (2009).

189. Yamaji H., Nakamura M., Kuwahara M., Takahashi Y., Katsuda T. & Konishi E. Efficient production of Japanese encephalitis virus-like particles by recombinant lepidopteran insect cells. *Applied Microbiology and Biotechnology* **97,** 1071–1079. doi:10.1007/s00253-012-4371-y (2013).

190. Du R., Yin F., Wang M., Hu Z., Wang H. & Deng F. Glycoprotein E of the Japanese encephalitis virus forms virus-like particles and induces syncytia when expressed by a baculovirus. *Journal of General Virology* **96,** 1006–1014. doi:10.1099/vir.0.000052 (2015).

191. Subramanian B. M., Madhanmohan M., Sriraman R., Reddy R. V. C., Yuvaraj S., Manikumar K., Rajalakshmi S., Nagendrakumar S. B., Rana S. K. & Srinivasan V. A. Development of foot-and-mouth disease virus (FMDV) serotype O virus-like-particles (VLPs) vaccine and evaluation of its potency. *Antiviral Research* **96,** 288–295. doi:10.1016/j.antiviral.2012.09.019 (2012).

192. Tatman J. D., Preston V. G., Nicholson P., Elliott R. M. & Rixon F. J. Assembly of Herpes Simplex virus Type 1 Capsids Using a Panel of Recombinant Baculoviruses. *The Journal of General Virology* **75,** 1101–1113. doi:10.1099/0022-1317-75-5-1101 (1994).

193. Thomsen D. R., Roof L. L. & Homa F. L. Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. *Journal of Virology* **68,** 2442–2457. doi:10.1128/jvi.68.4.2442-2457.1994 (1994).

194. Kirnbauer R., Taub J., Greenstone H., Roden R., Dürst M., Gissmann L., Lowy D. R. & Schiller J. T. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. *Journal of Virology* **67,** 6929–6936. doi:10.1128/jvi.67.12.6929-6936.1993 (1993).

195. Volpers C., Schirmacher P., Streeck R. E. & Sapp M. Assembly of the major and the minor capsid protein of human papillomavirus type 33 into virus-like particles and tubular structures in insect cells. *Virology* **200,** 504–512. doi:10.1006/viro.1994.1213 (1994).

196. Urakawa T., Ferguson M., Minor P. D., Cooper J., Sullivan M., Almond J. W. & Bishop D. H. L. Synthesis of immunogenic, but non-infectious, poliovirus particles in insect cells by a baculovirus expression vector. *Journal of General Virology* **70,** 1453–1463. doi:10.1099/0022-1317-70-6-1453 (1989).

197. Bräutigam S., Snezhkov E. & Bishop D. H. L. Formation of Poliovirus-like Particles by Recombinant Baculoviruses Expressing the Individual VP0, VP3, and VP1 Proteins by Comparison to Particles Derived from the Expressed Poliovirus Polyprotein. *Virology* **192,** 512–524. doi:10.1006/viro.1993.1067 (1993).

198. Conner M. E., Zarley C. D., Hu B., Parsons S., Drabinski D., Greiner S., Smith R., Jiang B., Corsaro B., Barniak V., Madore H. P., Crawford S. & Estes M. K. Virus-like particles as a rotavirus subunit vaccine. *The Journal of Infectious Diseases* **174,** S88–92. doi:10.1093/infdis/174.Supplement_1.S88 (1996).

199. O'Neal C. M., Crawford S. E., Estes M. K. & Conner M. E. Rotavirus virus-like particles administered mucosally induce protective immunity. *Journal of Virology* **71,** 8707–8717. doi:10.1128/jvi.71.11.8707-8717.1997 (1997).

200. Labbe M., Charpilienne A., Crawford S. E., Estes M. K. & Cohen J. Expression of rotavirus VP2 produces empty corelike particles. *Journal of Virology* **65,** 2946–2952. doi:10.1128/jvi.65.6.2946-2952.1991 (1991).

201. Inoue T., Kawano M.-a., Takahashi R.-u., Tsukamoto H., Enomoto T., Imai T., Kataoka K. & Handa H. Engineering of SV40-based nano-capsules for delivery of heterologous proteins as fusions with the minor capsid proteins VP2/3. *Journal of Biotechnology* **134,** 181–192. doi:10.1016/j.jbiotec.2007.12.006 (2008).

202. Dai L., Zheng T., Xu K., Han Y., Xu L., Huang E., An Y., Cheng Y., Li S., Liu M., Yang M., Li Y., Cheng H., Yuan Y., Zhang W., Ke C., Wong G., Qi J., Qin C., Yan J. & Gao G. F. A Universal Design of Betacoronavirus Vaccines against COVID-19, MERS, and SARS. *Cell* **182,** 722–733.e11. doi:10.1016/j.cell.2020.06.035 (2020).

203. Lan J., Yao Y., Deng Y., Chen H., Lu G., Wang W., Bao L., Deng W., Wei Q., Gao G. F., Qin C. & Tan W. Recombinant Receptor Binding Domain Protein Induces Partial Protective Immunity in Rhesus Macaques Against Middle East Respiratory Syndrome Coronavirus Challenge. *EBioMedicine* **2,** 1438–1446. doi:10.1016/j.ebiom.2015.08.031 (2015).

204. Lan J., Deng Y., Chen H., Lu G., Wang W., Guo X., Lu Z., Gao G. F. & Tan W. Tailoring Subunit Vaccine Immunity with Adjuvant Combinations and Delivery Routes Using the Middle East Respiratory Coronavirus (MERS-CoV) Receptor-Binding Domain as an Antigen. *PLoS ONE* **9,** e112602. doi:10.1371/journal.pone.0112602 (2014).

205. He Y., Li J., Heck S., Lustigman S. & Jiang S. Antigenic and Immunogenic Characterization of Recombinant Baculovirus-Expressed Severe Acute Respiratory Syndrome Coronavirus Spike Protein: Implication for Vaccine Design. *Journal of Virology* **80,** 5757–5767. doi:10.1128/jvi.00083-06 (2006).

206. Kuwahara M. & Konishi E. Evaluation of Extracellular Subviral Particles of Dengue Virus Type 2 and Japanese Encephalitis Virus Produced by Spodoptera frugiperda Cells for Use as Vaccine and Diagnostic Antigens. *Clinical and Vaccine Immunology* **17,** 1560–1566. doi:10.1128/cvi.00087-10 (2010).

207. Ramya R., Subramanian B. M., Sivakumar V., Senthilkumar R. L., Rao K. R. S. S. & Srinivasan V. A. Expression and Solubilization of Insect Cell-Based Rabies Virus Glycoprotein and Assessment of Its Immunogenicity and Protective Efficacy in Mice. *Clinical and Vaccine Immunology* **18,** 1673–1679. doi:10.1128/cvi.05258-11 (2011).

208. Fu Z. F., Rupprecht C. E., Dietzschold B., Saikumar P., Niu H. S., Babka I., Wunner W. H. & Koprowski H. Oral vaccination of racoons (Procyon lotor) with baculovirus-expressed rabies virus glycoprotein. *Vaccine* **11,** 925–928. doi:10.1016/0264-410x(93)90379-c (1993).

209. Lakey D. L., Treanor J. J., Betts R. F., Smith G. E., Thompson J., Sannella E., Reed G., Wilkinson B. E. & Wright P. F. Recombinant Baculovirus Influenza A Hemagglutinin Vaccines Are Well Tolerated and Immunogenic in Healthy Adults. *The Journal of Infectious Diseases* **174,** 838–841. doi:10.1093/infdis/174.4.838 (1996).

210. Dunkle L. M., Izikson R., Patriarca P. A., Goldenthal K. L., Cox M. & Treanor J. J. Safety and Immunogenicity of a Recombinant Influenza Vaccine: A Randomized Trial. *Pediatrics* **141,** e20173021. doi:10.1542/peds.2017-3021 (2018).

211. Dunkle L. M., Izikson R., Patriarca P. A., Goldenthal K. L., Muse D. & Cox M. M. J. Randomized Comparison of Immunogenicity and Safety of Quadrivalent Recombinant Versus Inactivated Influenza Vaccine in Healthy Adults 18–49 Years of Age. *The Journal of Infectious Diseases* **216,** 1219–1226. doi:10.1093/infdis/jix478 (2017).

212. Stadlbauer D., Rajabhathor A., Amanat F., Kaplan D., Masud A., Treanor J. J., Izikson R., Cox M. M., Nachbagauer R. & Krammer F. Vaccination with a Recombinant H7 Hemagglutinin-Based Influenza Virus Vaccine Induces Broadly Reactive Antibodies in Humans. *mSphere* **2,** e00502–17. doi:10.1128/msphere.00502-17 (2017).

213. Treanor J. J., Sahly H. E., King J., Graham I., Izikson R., Kohberger R., Patriarca P. & Cox M. Protective efficacy of a trivalent recombinant hemagglutinin protein vaccine (FluBlok®) against influenza in healthy adults: A randomized, placebo-controlled trial. *Vaccine* **29,** 7733–7739. doi:10.1016/j.vaccine.2011.07.128 (2011).

214. Prabakaran M., Velumani S., He F., Karuppannan A. K., Geng G. Y., Yin L. K. & Kwang J. Protective immunity against influenza H5N1 virus challenge in mice by intranasal co-administration of baculovirus surface-displayed HA and recombinant CTB as an adjuvant. *Virology* **380,** 412–420. doi:10.1016/j.virol.2008.08.002 (2008).

215. Wu Q., Fang L., Wu X., Li B., Luo R., Yu Z., Jin M., Chen H. & Xiao S. A pseudotype baculovirus-mediated vaccine confers protective immunity against lethal challenge with H5N1 avian influenza virus in mice and chickens. *Molecular Immunology* **46,** 2210–2217. doi:10.1016/j.molimm.2009.04.017 (2009).

216. Chen C.-Y., Liu H.-J., Tsai C.-P., Chung C.-Y., Shih Y.-S., Chang P.-C., Chiu Y.-T. & Hu Y.-C. Baculovirus as an avian influenza vaccine vector: Differential immune responses elicited by different vector forms. *Vaccine* **28,** 7644–7651. doi:10.1016/j.vaccine.2010.09.048 (2010).

217. Zhang J., Fan H.-Y., Zhang Z., Zhang J., Zhang J., Huang J.-N., Ye Y. & Liao M. Recombinant baculovirus vaccine containing multiple M2e and adjuvant LTB induces T cell dependent, cross-clade protection against H5N1 influenza virus in mice. *Vaccine* **34,** 622–629. doi:10.1016/j.vaccine.2015.12.039 (2016).

218. Choi H., Jang Y., Cho H., Shin H. Y. & Kim Y. B. Development of Zika Virus DNA Vaccine Using Envelope Modified Baculoviral Gene Delivery System. *Proceedings* **50,** 85. doi:10.3390/proceedings2020050085 (2020).

219. Lee J.-Y. & Chang J. Recombinant baculovirus-based vaccine expressing M2 protein induces protective CD8+ T-cell immunity against respiratory syncytial virus infection. *Journal of Microbiology* **55,** 900–908. doi:10.1007/s12275-017-7306-6 (2017).

220. Zhang Y., Qiao L., Hu X., Zhao K., Zhang Y., Chai F. & Pan Z. Baculovirus vectors expressing F proteins in combination with virus-induced signaling adaptor (VISA) molecules confer protection against respiratory syncytial virus infection. *Vaccine* **34,** 252–260. doi:10.1016/j.vaccine.2015.11.027 (2016).

221. Strauss R., Hüser A., Ni S., Tuve S., Kiviat N., Sow P. S., Hofmann C. & Lieber A. Baculovirus-based Vaccination Vectors Allow for Efficient Induction of Immune Responses Against Plasmodium falciparum Circumsporozoite Protein. *Molecular Therapy* **15,** 193–202. doi:10.1038/sj.mt.6300008 (2007).

222. Meng T., Kolpe A. B., Kiener T. K., Chow V. T. K. & Kwang J. Display of VP1 on the Surface of Baculovirus and Its Immunogenicity against Heterologous Human Enterovirus 71 Strains in Mice. *PLoS ONE* **6,** e21757. doi:10.1371/journal.pone.0021757 (2011).

223. Premanand B., Kiener T. K., Meng T., Tan Y. R., Jia Q., Chow V. T. & Kwang J. Induction of protective immune responses against EV71 in mice by baculovirus encoding a novel expression cassette for capsid protein VP1. *Antiviral Research* **95,** 311–315. doi:10.1016/j.antiviral.2012.05.017 (2012).

224. Kiener T. K., Premanand B. & Kwang J. Immune responses to baculovirus-displayed enterovirus 71 VP1 antigen. *Expert Review of Vaccines* **12,** 357–364. doi:10.1586/erv.13.18 (2014).

225. Kang S.-M., Song J.-M., Quan F.-S. & Compans R. W. Influenza vaccines based on virus-like particles. *Virus Research* **143,** 140–146. doi:10.1016/j.virusres.2009.04.005 (2009).

226. Wang C.-C., Chen J.-R., Tseng Y.-C., Hsu C.-H., Hung Y.-F., Chen S.-W., Chen C.-M., Khoo K.-H., Cheng T.-J., Cheng Y.-S. E., Jan J.-T., Wu C.-Y., Ma C. & Wong C.-H. Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 18137–18142. doi:10.1073/pnas.0909696106 (2009).

227. Dowling W., Thompson E., Badger C., Mellquist J. L., Garrison A. R., Smith J. M., Paragas J., Hogan R. J. & Schmaljohn C. Influences of glycosylation on antigenicity, immunogenicity, and protective efficacy of ebola virus GP DNA vaccines. *Journal of Virology* **81,** 1821–1837. doi:10.1128/jvi.02098-06 (2007).

228. Pan Y.-S., Wei H.-J., Chang C.-C., Lin C.-H., Wei T.-S., Wu S.-C. & Chang D.-K. Construction and Characterization of Insect Cell-Derived Influenza VLP: Cell Binding, Fusion, and EGFP Incorporation. *Journal of Biomedicine and Biotechnology* **2010,** 1–11. doi:10.1155/2010/506363 (2010).

229. Treanor J. J., Wilkinson B. E., Masseoud F., Hu-Primmer J., Battaglia R., O'Brien D., Wolff M., Rabinovich G., Blackwelder W. & Katz J. M. Safety and immunogenicity of a recombinant hemagglutinin vaccine for H5 influenza in humans. *Vaccine* **19,** 1732–1737. doi:10.1016/s0264-410x(00)00395-9 (2001).

230. Gordon D. L., Sajkov D., Woodman R. J., Honda-Okubo Y., Cox M. M., Heinzel S. & Petrovsky N. Randomized clinical trial of immunogenicity and safety of a recombinant H1N1/2009 pandemic influenza vaccine containing Advax™ polysaccharide adjuvant. *Vaccine* **30,** 5407–5416. doi:10.1016/j.vaccine.2012.06.009 (2012).

231. López-Macías C., Ferat-Osorio E., Tenorio-Calvo A., Isibasi A., Talavera J., Arteaga-Ruiz O., Arriaga-Pizano L., Hickman S. P., Allende M., Lenhard K., Pincus S., Connolly K., Raghunandan R., Smith G. & Glenn G. Safety and immunogenicity of a virus-like particle pandemic influenza A (H1N1) 2009 vaccine in a blinded, randomized, placebo-controlled trial of adults in Mexico. *Vaccine* **29,** 7826–7834. doi:10.1016/j.vaccine.2011.07.099 (2011).

232. Shinde V., Cho I., Plested J. S., Agrawal S., Fisk J., Cai R., Zhou H., Pham X., Zhu M., Cloney-Clark S., Wang N., Zhou B., Lewis M., Price-Abbott P., Patel N., Massare M. J., Smith G., Keech C., Fries L. & Glenn G. M. Comparison of the Safety and Immunogenicity of a Novel Matrix-M-adjuvanted Nanoparticle Influenza Vaccine with a Quadrivalent Seasonal Influenza Vaccine in Older Adults: A Randomized Controlled Trial. *medRxiv,* 1–33. doi:10.1101/2020.08.07.20170514 (2020).

233. Fries L. F., Smith G. E. & Glenn G. M. A Recombinant Virus-like Particle Influenza A (H7N9) Vaccine. *The New England Journal of Medicine* **369,** 2564–2566. doi:10.1056/nejmc1313186 (2013).

234. Khurana S., Wu J., Verma N., Verma S., Raghunandan R., Manischewitz J., King L. R., Kpamegan E., Pincus S., Smith G., Glenn G. & Golding H. H5N1 Virus-Like Particle Vaccine Elicits Cross-Reactive Neutralizing Antibodies That Preferentially Bind to the Oligomeric Form of Influenza Virus Hemagglutinin in Humans. *Journal of Virology* **85,** 10945–10954. doi:10.1128/jvi.05406-11 (2011).

235. Nachbagauer R., Liu W.-C., Choi A., Wohlbold T. J., Atlas T., Rajendran M., Solórzano A., Berlanda-Scorza F., García-Sastre A., Palese P., Albrecht R. A. & Krammer F. A universal influenza virus vaccine candidate confers protection against pandemic H1N1 infection in preclinical ferret studies. *npj Vaccines* **2,** 26. doi:10.1038/s41541-017-0026-4 (2017).

236. Bernstein D. I. *et al.* Immunogenicity of chimeric haemagglutinin-based, universal influenza virus vaccine candidates: interim results of a randomised, placebo-controlled, phase 1 clinical trial. *The Lancet Infectious Diseases* **20,** 80–91. doi:10.1016/s1473-3099(19)30393-7 (2020).

237. Lin S.-Y., Chung Y.-C. & Hu Y.-C. Update on baculovirus as an expression and/or delivery vehicle for vaccine antigens. *Expert Review of Vaccines* **13,** 1501–1521. doi:10.1586/14760584.2014.951637 (2014).

238. Airenne K. J., Hu Y.-C., Kost T. A., Smith R. H., Kotin R. M., Ono C., Matsuura Y., Wang S. & Ylä-Herttuala S. Baculovirus: an insect-derived vector for diverse gene transfer applications. *Molecular Therapy* **21,** 739–749. doi:10.1038/mt.2012.286 (2013-04).

239. Heinimäki S., Tamminen K., Malm M., Vesikari T. & Blazevic V. Live baculovirus acts as a strong B and T cell adjuvant for monomeric and oligomeric protein antigens. *Virology* **511,** 114–122. doi:10.1016/j.virol.2017.08.023 (2017).

240. Hu Y.-C., Yao K. & Wu T.-Y. Baculovirus as an expression and/or delivery vehicle for vaccine antigens. *Expert Review of Vaccines* **7,** 363–371. doi:10.1586/14760584.7.3.363 (2014).

241. Premanand B., Wee P. Z. & Prabakaran M. Baculovirus Surface Display of Immunogenic Proteins for Vaccine Development. *Viruses* **10,** 298. doi:10.3390/v10060298 (2018).

242. Fabre M. L., Arrías P. N., Masson T., Pidre M. L. & Romanowski V. Baculovirus-derived vectors for immunization and therapeutic applications. *Emerging and Reemerging Viral Pathogens,* 197–224. doi:10.1016/b978-0-12-814966-9.00011-1 (2020).

243. Felberbaum R. S. The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors. *Biotechnology Journal* **10,** 702–714. doi:10.1002/biot.201400438 (2015).

244. Dolgin E. Early clinical data raise the bar for hemophilia gene therapies. *Nature Biotechnology* **34,** 999–1001. doi:10.1038/nbt1016-999 (2016).

245. Pasi K. J., Rangarajan S., Mitchell N., Lester W., Symington E., Madan B., Laffan M., Russell C. B., Li M., Pierce G. F. & Wong W. Y. Multiyear Follow-up of AAV5-hFVIII-SQ Gene Therapy for Hemophilia A. *New England Journal of Medicine* **382,** 29–40. doi:10.1056/nejmoa1908490 (2020).

246. High K. H., Nathwani A., Spencer T. & Lillicrap D. Current status of haemophilia gene therapy. *Haemophilia* **20,** 43–49. doi:10.1111/hae.12411 (2014).

247. Kondratov O., Marsic D., Crosson S. M., Mendez-Gomez H. R., Moskalenko O., Mietzsch M., Heilbronn R., Allison J. R., Green K. B., Agbandje-McKenna M. & Zolotukhin S. Direct Head-to-Head Evaluation of Recombinant Adeno-associated Viral Vectors Manufactured in Human versus Insect Cells. *Molecular Therapy* **25,** 2661–2675. doi:10.1016/j.ymthe.2017.08.003 (2017).

248. Joshi P. R., Venereo-Sanchez A., Chahal P. S. & Kamen A. A. Advancements in Molecular Systems Design and Bioprocessing of Recombinant Adeno-associated Virus Gene Delivery Vectors using the Insect-Cell Baculovirus Expression Platform. *Biotechnology Journal,* 2000021. doi:10.1002/biot.202000021 (2020).

249. Lesch H. P., Laitinen A., Peixoto C., Vicente T., Makkonen K.-E., Laitinen L., Pikkarainen J. T., Samaranayake H., Alves P. M., Carrondo M. J. T., Ylä-Herttuala S. & Airenne K. J. Production and purification of lentiviral vectors generated in 293T suspension cells with baculoviral vectors. *Gene Therapy* **18,** 531–538. doi:10.1038/gt.2010.162 (2011).

250. Mansouri M., Bellon-Echeverria I., Rizk A. e. l., Ehsaei Z., Cosentino C. C., Silva C. S., Xie Y., Boyce F. M., Davis M. W., Neuhauss S. C. F., Taylor V., Ballmer-Hofer K., Berger I. & Berger P. Highly efficient baculovirus-mediated multigene delivery in primary cells. *Nature Communications* **7,** 1–13. doi:10.1038/ncomms11529 (2016).

251. Takata Y., Kishine H., Sone T., Andoh T., Nozaki M., Poderycki M., Chesnut J. D. & Imamoto F. Generation of iPS Cells Using a BacMam Multigene Expression System. *Cell Structure and Function* **36,** 209–222. doi:10.1247/csf.11008 (2011).

252. Phang R.-Z., Tay F. C., Goh S.-L., Lau C.-H., Zhu H., Tan W.-K., Liang Q., Chen C., Du S., Li Z., Tay J. C.-K., Wu C., Zeng J., Fan W., Toh H. C. & Wang S. Zinc Finger Nuclease-Expressing Baculoviral Vectors Mediate Targeted Genome Integration of Reprogramming Factor Genes to Facilitate the Generation of Human Induced Pluripotent Stem Cells. *STEM CELLS Translational Medicine* **2,** 935–945. doi:10.5966/sctm.2013-0043 (2013).

253. Lin C.-Y., Chang Y.-H., Lin K.-J., Yen T.-C., Tai C.-L., Chen C.-Y., Lo W.-H., Hsiao I.-T. & Hu Y.-C. The healing of critical-sized femoral segmental bone defects in rabbits using baculovirus-engineered mesenchymal stem cells. *Biomaterials* **31,** 3222–3230. doi:10.1016/j.biomaterials.2010.01.030 (2010).

254. Lin C.-Y., Chang Y.-H., Kao C.-Y., Lu C.-H., Sung L.-Y., Yen T.-C., Lin K.-J. & Hu Y.-C. Augmented healing of critical-size calvarial defects by baculovirus-engineered MSCs that persistently express growth factors. *Biomaterials* **33,** 3682–3692. doi:10.1016/j.biomaterials.2012.02.007 (2012).

255. Chen C.-Y., Wu H.-H., Chen C.-P., Chern S.-R., Hwang S.-M., Huang S.-F., Lo W.-H., Chen G.-Y. & Hu Y.-C. Biosafety Assessment of Human Mesenchymal Stem Cells Engineered by Hybrid Baculovirus Vectors. *Molecular Pharmaceutics* **8,** 1505–1514. doi:10.1021/mp100368d (2011).

256. Liao Y.-H., Chang Y.-H., Sung L.-Y., Li K.-C., Yeh C.-L., Yen T.-C., Hwang S.-M., Lin K.-J. & Hu Y.-C. Osteogenic differentiation of adipose-derived stem cells and calvarial defect repair using baculovirus-mediated co-expression of BMP-2 and miR-148b. *Biomaterials* **35,** 4901–4910. doi:10.1016/j.biomaterials.2014.02.055 (2014).

257. Chuang C.-K., Wong T.-H., Hwang S.-M., Chang Y.-H., Chen G.-Y., Chiu Y.-C., Huang S.-F. & Hu Y.-C. Baculovirus Transduction of Mesenchymal Stem Cells: In Vitro Responses and In Vivo Immune Responses After Cell Transplantation. *Molecular Therapy* **17,** 889–896. doi:10.1038/mt.2009.30 (2009).

258. Turunen T. A. K., Laakkonen J. P., Alasaarela L., Airenne K. J. & Ylä-Herttuala S. Sleeping Beauty–baculovirus hybrid vectors for long-term gene expression in the eye. *The Journal of Gene Medicine* **16,** 40–53. doi:10.1002/jgm.2756 (2014).

259. Dunbar C. E., High K. A., Joung J. K., Kohn D. B., Ozawa K. & Sadelain M. Gene therapy comes of age. *Science* **359,** 4672. doi:10.1126/science.aan4672 (2018).

260. Walsh G. Biopharmaceutical benchmarks 2010. *Nature Biotechnology* **28,** 917–924. doi:10.1038/nbt0910-917 (2010).

261. Suzuki T., Chang M. O., Kitajima M. & Takaku H. Induction of antitumor immunity against mouse carcinoma by baculovirus-infected dendritic cells. *Cellular & Molecular Immunology* **7,** 440–446. doi:10.1038/cmi.2010.48 (2010).

262. Suzuki T., Chang M. O., Kitajima M. & Takaku H. Baculovirus activates murine dendritic cells and induces non-specific NK cell and T cell immune responses. *Cellular Immunology* **262,** 35–43. doi:10.1016/j.cellimm.2009.12.005 (2010).

263. Luo W.-Y., Shih Y.-S., Lo W.-H., Chen H.-R., Wang S.-C., Wang C.-H., Chien C.-H., Chiang C.-S., Chuang Y.-J. & Hu Y.-C. Baculovirus vectors for antiangiogenesis-based cancer gene therapy. *Cancer Gene Therapy* **18,** 637–645. doi:10.1038/cgt.2011.35 (2011).

264. Luo W.-Y., Shih Y.-S., Hung C.-L., Lo K.-W., Chiang C.-S., Lo W.-H., Huang S.-F., Wang S.-C., Yu C.-F., Chien C.-H. & Hu Y.-C. Development of the hybrid Sleeping Beauty-baculovirus vector for sustained gene expression and cancer therapy. *Gene Therapy* **19,** 844–851. doi:10.1038/gt.2011.129 (2012).

265. Wang Z., Li M., Ji Y., Yang M., Yang W., Wang J. & Li W. Development of a novel bivalent baculovirus vectors for complement resistance and sustained transgene expression and its application in anti-angiogenesis gene therapy. *Biomedicine & Pharmacotherapy* **123,** 109765. doi:10.1016/j.biopha.2019.109765 (2020).

266. Lin M.-W., Tseng Y.-W., Shen C.-C., Hsu M.-N., Hwu J.-R., Chang C.-W., Yeh C.-J., Chou M.-Y., Wu J.-C. & Hu Y.-C. Synthetic switch-based baculovirus for transgene expression control and selective killing of hepatocellular carcinoma cells. *Nucleic Acids Research* **46,** gky447–. doi:10.1093/nar/gky447 (2018).

267. Kost T. A., Condreay J. P., Ames R. S., Rees S. & Romanos M. A. Implementation of BacMam virus gene delivery technology in a drug discovery setting. *Drug Discovery Today* **12,** 396–403. doi:10.1016/j.drudis.2007.02.017 (2007).

268. Thomsen W., Frazer J. & Unett D. Functional assays for screening GPCR targets. *Current Opinion in Biotechnology* **16,** 655–665. doi:10.1016/j.copbio.2005.10.008 (2005).

269. Bieniossek C., Imasaki T., Takagi Y. & Berger I. MultiBac: expanding the research toolbox for multiprotein complexes. *Trends in Biochemical Sciences* **37,** 49–57. doi:10.1016/j.tibs.2011.10.005 (2012).

270. Briscoe C. P., Peat A. J., McKeown S. C., Corbett D. F., Goetz A. S., Littleton T. R., McCoy D. C., Kenakin T. P., Andrews J. L., Ammala C., Fornwald J. A., Ignar D. M. & Jenkinson S. Pharmacological regulation of insulin secretion in MIN6 cells through the fatty acid receptor GPR40: identification of agonist and antagonist small molecules. *British Journal of Pharmacology* **148,** 619–628. doi:10.1038/sj.bjp.0706770 (2006).

271. Medhurst A. D. *et al.* GSK189254, a Novel H3 Receptor Antagonist That Binds to Histamine H3 Receptors in Alzheimer's Disease Brain and Improves Cognitive Performance in Preclinical Models. *Journal of Pharmacology and Experimental Therapeutics* **321,** 1032–1045. doi:10.1124/jpet.107.120311 (2007).

272. Watson C., Jenkinson S., Kazmierski W. & Kenakin T. The CCR5 Receptor-Based Mechanism of Action of 873140, a Potent Allosteric Noncompetitive HIV Entry Inhibitor. *Molecular Pharmacology* **67,** 1268–1282. doi:10.1124/mol.104.008565 (2005).

273. Goehring A., Lee C.-H., Wang K. H., Michel J. C., Claxton D. P., Baconguis I., Althoff T., Fischer S., Garcia K. C. & Gouaux E. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nature Protocols* **9,** 2574–2585. doi:10.1038/nprot.2014.173 (2014).

274. Wright P. D., Veale E. L., McCoull D., Tickle D. C., Large J. M., Ococks E., Gothard G., Kettleborough C., Mathie A. & Jerman J. Terbinafine is a novel and selective activator of the two-pore domain potassium channel TASK3. *Biochemical and Biophysical Research Communications* **493,** 444–450. doi:10.1016/j.bbrc.2017.09.002 (2017).

275. Ames R. S., Fornwald J. A., Nuthulaganti P., Trill J. J., Foley J. J., Buckley P. T., Kost T. A., Wu Z. & Romanos M. A. BacMam Recombinant Baculoviruses in G Protein–Coupled Receptor Drug Discovery. *Receptors and Channels* **10,** 99–107. doi:10.3109/10606820490514969 (2011).

276. Ishino Y., Shinagawa H., Makino K., Amemura M. & Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology* **169,** 5429–5433. doi:10.1128/jb.169.12.5429-5433.1987 (1987).

277. Hsu P. D., Lander E. S. & Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157,** 1262–1278. doi:10.1016/j.cell.2014.05.010 (2014).

278. Mojica F. J., Díez-Villaseñor C., Soria E. & Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology* **36,** 244–246. doi:10.1046/j.1365-2958.2000.01838.x (2000).

279. Van Belkum A., Scherer S., Alphen L. v. & Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. *Microbiology and Molecular Biology Reviews* **62,** 275–293. doi:10.1128/MMBR.62.2.275-293.1998 (1998).

280. Makarova K. S., Aravind L., Grishin N. V., Rogozin I. B. & Koonin E. V. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research* **30,** 482–496. doi:10.1093/nar/30.2.482 (2002).

281. Mojica F. J. M., Díez-Villaseñor C., García-Martínez J. & Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution* **60,** 174–182. doi:10.1007/s00239-004-0046-3 (2005).

282. Bolotin A., Quinquis B., Sorokin A. & Ehrlich S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151,** 2551–2561. doi:10.1099/mic.0.28048-0 (2005).

283. Pourcel C., Salvignol G. & Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151,** 653–663. doi:10.1099/mic.0.27437-0 (2005).

284. Barrangou R., Fremaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D. A. & Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712. doi:10.1126/science.1138140 (2007).

285. Makarova K. S., Haft D. H., Barrangou R., Brouns S. J. J., Charpentier E., Horvath P., Moineau S., Mojica F. J. M., Wolf Y. I., Yakunin A. F., Oost J. v. d. & Koonin E. V. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology* **9,** 467–477. doi:10.1038/nrmicro2577 (2011).

286. Makarova K. S., Wolf Y. I., Alkhnbashi O. S., Costa F., Shah S. A., Saunders S. J., Barrangou R., Brouns S. J. J., Charpentier E., Haft D. H., Horvath P., Moineau S., Mojica F. J. M., Terns R. M., Terns M. P., White M. F., Yakunin A. F., Garrett R. A., Oost J. v. d., Backofen R. & Koonin E. V. An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* **13,** 722–736. doi:10.1038/nrmicro3569 (2015).

287. Haft D. H., Selengut J., Mongodin E. F. & Nelson K. E. A guild of forty-five CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology* **1,** e60. doi:10.1371/journal.pcbi.0010060 (2005).

288. Van der Oost J., Jore M. M., Westra E. R., Lundgren M. & Brouns S. J. J. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences* **34,** 401–407. doi:10.1016/j.tibs.2009.05.002 (2009).

289. Oost J. v. d., Westra E. R., Jackson R. N. & Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nature Reviews Microbiology* **12,** 479–492. doi:10.1038/nrmicro3279 (2014).

290. Bhaya D., Davison M. & Barrangou R. CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. *Annual Review of Genetics* **45,** 273–297. doi:10.1146/annurev-genet-110410-132430 (2011).

291. Grissa I., Vergnaud G. & Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research* **35,** W52–W57. doi:10.1093/nar/gkm360 (2007).

292. Deltcheva E., Chylinski K., Sharma C. M., Gonzales K., Chao Y., Pirzada Z. A., Eckert M. R., Vogel J. & Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607. doi:10.1038/nature09886 (2011).

293. Spilman M., Cocozaki A., Hale C., Shao Y., Ramia N., Terns R., Terns M., Li H. & Stagg S. Structure of an RNA silencing complex of the CRISPR-Cas immune system. *Molecular Cell* **52,** 146–152. doi:10.1016/j.molcel.2013.09.008 (2013).

294. Rouillon C., Zhou M., Zhang J., Politis A., Beilsten-Edmands V., Cannone G., Graham S., Robinson C. V., Spagnolo L. & White M. F. Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Molecular Cell* **52,** 124–134. doi:10.1016/j.molcel.2013.08.020 (2013).

295. Zhang J., Rouillon C., Kerou M., Reeks J., Brugger K., Graham S., Reimann J., Cannone G., Liu H., Albers S.-V., Naismith J. H., Spagnolo L. & White M. F. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Molecular Cell* **45,** 303–313. doi:10.1016/j.molcel.2011.12.013 (2012).

296. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A. & Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337,** 816–821. doi:10.1126/science.1225829 (2012).

297. Shah S. A., Erdmann S., Mojica F. J. M. & Garrett R. A. Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biology* **10,** 891–899. doi:10.4161/rna.23764 (2013-05).

298. Doudna J. A. & Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346,** 1258096–1258096. doi:10.1126/science.1258096 (2014).

299. Xu T., Li Y., Nostrand J. D. V., He Z. & Zhou J. Cas9-Based Tools for Targeted Genome Editing and Transcriptional Control. *Applied and Environmental Microbiology* **80,** 1544–1552. doi:10.1128/aem.03786-13 (2014).

300. Sander J. D. & Joung J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* **32,** 347–355. doi:10.1038/nbt.2842 (2014).

301. Barrangou R. Cas9 targeting and the CRISPR revolution. *Science* **344,** 707–708. doi:10.1126/science.1252964 (2014).

302. Jiang W., Cox D., Zhang F., Bikard D. & Marraffini L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology* **31,** 233–239. doi:10.1038/nbt.2508 (2013).

303. Cong L., Ran F. A., Cox D., Lin S., Barretto R., Habib N., Hsu P. D., Wu X., Jiang W., Marraffini L. A. & Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–823. doi:10.1126/science.1231143 (2013).

304. Mali P., Yang L., Esvelt K. M., Aach J., Guell M., DiCarlo J. E., Norville J. E. & Church G. M. RNA-Guided Human Genome Engineering via Cas9. *Science* **339,** 823–826. doi:10.1126/science.1232033 (2013).

305. Jinek M., East A., Cheng A., Lin S., Ma E. & Doudna J. RNA-programmed genome editing in human cells. *eLife* **2,** e00471. doi:10.7554/elife.00471 (2013).

306. DiCarlo J. E., Norville J. E., Mali P., Rios X., Aach J. & Church G. M. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Research* **41,** 4336–4343. doi:10.1093/nar/gkt135 (2013-04).

307. Hwang W. Y., Fu Y., Reyon D., Maeder M. L., Tsai S. Q., Sander J. D., Peterson R. T., Yeh J.-R. J. & Joung J. K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology* **31,** 227–229. doi:10.1038/nbt.2501 (2013-03).

308. Li J.-F., Norville J. E., Aach J., McCormack M., Zhang D., Bush J., Church G. M. & Sheen J. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nature Biotechnology* **31,** 688–691. doi:10.1038/nbt.2654 (2013-08).

309. Sharan S. K., Thomason L. C., Kuznetsov S. G. & Court D. L. Recombineering: a homologous recombination-based method of genetic engineering. *Nature Protocols* **4,** 206–223. doi:10.1038/nprot.2008.227 (2009).

310. Makarova K. S., Grishin N. V., Shabalina S. A., Wolf Y. I. & Koonin E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct* **1,** 7. doi:10.1186/1745-6150-1-7 (2006).

311. Ran F. A., Hsu P. D., Lin C.-Y., Gootenberg J. S., Konermann S., Trevino A. E., Scott D. A., Inoue A., Matoba S., Zhang Y. & Zhang F. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154,** 1380–1389. doi:10.1016/j.cell.2013.08.021 (2013).

312. Dianov G. L. & Hübscher U. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Research* **41,** 3483–3490. doi:10.1093/nar/gkt076 (2013).

313. Hsu P. D., Scott D. A., Weinstein J. A., Ran F. A., Konermann S., Agarwala V., Li Y., Fine E. J., Wu X., Shalem O., Cradick T. J., Marraffini L. A., Bao G. & Zhang F. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* **31,** 827–832. doi:10.1038/nbt.2647 (2013).

314. Qi L. S., Larson M. H., Gilbert L. A., Doudna J. A., Weissman J. S., Arkin A. P. & Lim W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152,** 1173–1183. doi:10.1016/j.cell.2013.02.022 (2013).

315. Gilbert L. A., Horlbeck M. A., Adamson B., Villalta J. E., Chen Y., Whitehead E. H., Guimaraes C., Panning B., Ploegh H. L., Bassik M. C., Qi L. S., Kampmann M. & Weissman J. S. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–661. doi:10.1016/j.cell.2014.09.029 (2014).

316. Gilbert L. A., Larson M. H., Morsut L., Liu Z., Brar G. A., Torres S. E., Stern-Ginossar N., Brandman O., Whitehead E. H., Doudna J. A., Lim W. A., Weissman J. S. & Qi L. S. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154,** 442–451. doi:10.1016/j.cell.2013.06.044 (2013).

317. Konermann S., Brigham M. D., Trevino A., Hsu P. D., Heidenreich M., Cong L., Platt R. J., Scott D. A., Church G. M. & Zhang F. Optical control of mammalian endogenous transcription and epigenetic states. *Nature,* 1–17. doi:10.1038/nature12466 (2013).

318. McDonald J. I., Celik H., Rois L. E., Fishberger, G F. T., Rees R., Kramer A., Martens A., Edwards J. R. & Challen G. A. Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. *Biology Open* **5,** 866–874. doi:10.1242/bio.019067 (2016).

319. Vojta A., Dobrinić P., Tadić V., Bočkor L., Korać P., Julg B., Klasić M. & Zoldoš V. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research* **44,** 5615–5628. doi:10.1093/nar/gkw159 (2016).

320. Konermann S., Brigham M. D., Trevino A. E., Joung J., Abudayyeh O. O., Barcena C., Hsu P. D., Habib N., Gootenberg J. S., Nishimasu H., Nureki O. & Zhang F. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583–588. doi:10.1038/nature14136 (2014).

321. Mali P., Aach J., Stranges P. B., Esvelt K. M., Moosburner M., Kosuri S., Yang L. & Church G. M. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology* **31,** 833–838. doi:10.1038/nbt.2675 (2013).

322. Tanenbaum M. E., Gilbert L. A., Qi L. S., Weissman J. S. & Vale R. D. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell* **159,** 635–646. doi:10.1016/j.cell.2014.09.039 (2014-10).

323. Xu X., Tao Y., Gao X., Zhang L., Li X., Zou W., Ruan K., Wang F., Xu G.-l. & Hu R. A CRISPR-based approach for targeted DNA demethylation. *Cell Discovery* **2,** 16009. doi:10.1038/celldisc.2016.9 (2016).

324. Maeder M. L., Angstman J. F., Richardson M. E., Linder S. J., Cascio V. M., Tsai S. Q., Ho Q. H., Sander J. D., Reyon D., Bernstein B. E., Costello J. F., Wilkinson M. F. & Joung J. K. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nature Biotechnology* **31,** 1137–1142. doi:10.1038/nbt.2726 (2013).

325. Kearns N. A., Pham H., Tabak B., Genga R. M., Silverstein N. J., Garber M. & Maehr R. Functional annotation of native enhancers with a Cas9–histone demethylase fusion. *Nature Methods* **12,** 401–403. doi:10.1038/nmeth.3325 (2015).

326. Perez-Pinera P., Kocak D. D., Vockley C. M., Adler A. F., Kabadi A. M., Polstein L. R., Thakore P. I., Glass K. A., Ousterout D. G., Leong K. W., Guilak F., Crawford G. E., Reddy T. E. & Gersbach C. A. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods* **10,** 973–976. doi:10.1038/nmeth.2600 (2013).

327. Reid W. & O'Brochta D. A. Applications of genome editing in insects. *Current Opinion in Insect Science* **13,** 43–54. doi:10.1016/j.cois.2015.11.001 (2016).

328. Ren X., Sun J., Housden B. E., Hu Y., Roesel C., Lin S., Liu L.-P., Yang Z., Mao D., Sun L., Wu Q., Ji J.-Y., Xi J., Mohr S. E., Xu J., Perrimon N. & Ni J.-Q. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proceedings of the National Academy of Sciences of the United States of America* **110,** 19012–19017. doi:10.1073/pnas.1318481110 (2013).

329. Ren X., Yang Z., Xu J., Sun J., Mao D., Hu Y., Yang S.-J., Qiao H.-H., Wang X., Hu Q., Deng P., Liu L.-P., Ji J.-Y., Li J. B. & Ni J.-Q. Enhanced Specificity and Efficiency of the CRISPR/Cas9 System with Optimized sgRNA Parameters in Drosophila. *Cell Reports* **9,** 1151–1162. doi:10.1016/j.celrep.2014.09.044 (2014-11).

330. Port F., Chen H.-M., Lee T. & Bullock S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **111,** E2967–76. doi:10.1073/pnas.1405500111 (2014).

331. Gratz S. J., Cummings A. M., Nguyen J. N., Hamm D. C., Donohue L. K., Harrison M. M., Wildonger J. & O'Connor-Giles K. M. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194,** 1029–1035. doi:10.1534/genetics.113.152710 (2013-08).

332. Bassett A. R., Tibbit C., Ponting C. P. & Liu J.-L. Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. *Cell Reports* **4,** 220–228. doi:10.1016/j.celrep.2013.06.020 (2013).

333. Dong Z.-Q., Chen T.-T., Zhang J., Hu N., Cao M.-Y., Dong F.-F., Jiang Y.-M., Chen P., Lu C. & Pan M.-H. Establishment of a highly efficient virus-inducible CRISPR/Cas9 system in insect cells. *Antiviral Research* **130,** 50–57. doi:10.1016/j.antiviral.2016.03.009 (2016).

334. Kistler K. E., Vosshall L. B. & Matthews B. J. Genome engineering with CRISPR-Cas9 in the mosquito Aedes aegypti. *Cell Reports* **11,** 51–60. doi:10.1016/j.celrep.2015.03.009 (2015).

335. Basu S., Aryan A., Overcash J. M., Samuel G. H., Anderson M. A. E., Dahlem T. J., Myles K. M. & Adelman Z. N. Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in Aedes aegypti. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 4038–4043. doi:10.1073/pnas.1502370112 (2015).

336. Hall A. B., Basu S., Jiang X., Qi Y., Timoshevskiy V. A., Biedler J. K., Sharakhova M. V., Elahi R., Anderson M. A. E., Chen X. G., Sharakhov I. V., Adelman Z. N. & Tu Z. A male-determining factor in the mosquito Aedes aegypti. *Science* **348,** 1268–1270. doi:10.1126/science.aaa2850 (2015).

337. Hammond A., Galizi R., Kyrou K., Simoni A., Siniscalchi C., Katsanos D., Gribble M., Baker D., Marois E., Russell S., Burt A., Windbichler N., Crisanti A. & Nolan T. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. *Nature Biotechnology* **34,** 78–83. doi:10.1038/nbt.3439 (2015).

338. Gantz V. M., Jasinskiene N., Tatarenkova O., Fazekas A., Macias V. M., Bier E. & James A. A. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. *Proceedings of the National Academy of Sciences of the United States of America* **112,** E6736–E6743. doi:10.1073/pnas.1521077112 (2015).

339. Markert M. J., Zhang Y., Enuameh M. S., Reppert S. M., Wolfe S. A. & Merlin C. Genomic access to monarch migration using TALEN and CRISPR/Cas9-mediated targeted mutagenesis. *G3: Genes Genomes Genetics* **6,** 905–915. doi:10.1534/g3.116.027029 (2016).

340. Li X., Fan D., Zhang W., Liu G., Zhang L., Zhao L., Fang X., Chen L., Dong Y., Chen Y., Ding Y., Zhao R., Feng M., Zhu Y., Feng Y., Jiang X., Zhu D., Xiang H., Feng X., Li S., Wang J., Zhang G., Kronforst M. R. & Wang W. Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. *Nature Communications* **6,** 8212–10. doi:10.1038/ncomms9212 (2015).

341. Liu Y., Ma S., Wang X., Chang J., Gao J., Shi R., Zhang J., Lu W., Liu Y., Zhao P. & Xia Q. Highly efficient multiplex targeted mutagenesis and genomic structure variation in Bombyx mori cells using CRISPR/Cas9. *Insect Biochemistry and Molecular Biology* **49,** 35–42. doi:10.1016/j.ibmb.2014.03.010 (2014).

342. Wei W., Xin H., Roy B., Dai J., Miao Y. & Gao G. Heritable Genome Editing with CRISPR/Cas9 in the Silkworm, Bombyx mori. *PLoS ONE* **9,** e101210. doi:10.1371/journal.pone.0101210 (2014).

343. Wang Y., Li Z., Xu J., Zeng B., Ling L., You L., Chen Y., Huang Y. & Tan A. The CRISPR/Cas System mediates efficient genome engineering in Bombyx mori. *Cell Research* **23,** 1414–1416. doi:10.1038/cr.2013.146 (2013).

344. Ma S., Chang J., Wang X., Liu Y., Zhang J., Lu W., Gao J., Shi R., Zhao P. & Xia Q. CRISPR/Cas9 mediated multiplex genome editing and heritable mutagenesis of BmKu70 in Bombyx mori. *Scientific Reports* **4,** 1–6. doi:10.1038/srep04489 (2014).

345. Bi H.-L., Xu J., Tan A.-J. & Huang Y.-P. CRISPR/Cas9-mediated targeted gene mutagenesis in Spodoptera litura. *Insect Science* **23,** 469–477. doi:10.1111/1744-7917.12341 (2016).

346. Koutroumpa F. A., Monsempes C., François M.-C., Cian A. d., Royer C., Concordet J.-P. & Jacquin-Joly E. Heritable genome editing with CRISPR/Cas9 induces anosmia in a crop pest moth. *Scientific Reports,* 1–9. doi:10.1038/srep29620 (2016).

347. Zhu G.-H., Xu J., Cui Z., Dong X.-T., Ye Z.-F., Niu D.-J., Huang Y.-P. & Dong S.-L. Functional characterization of SlitPBP3 in Spodoptera litura by CRISPR/Cas9 mediated genome editing. *Insect Biochemistry and Molecular Biology* **75,** 1–9. doi:10.1016/j.ibmb.2016.05.006 (2016).

348. Zhu L., Mon H., Xu J., Lee J. M. & Kusakabe T. CRISPR/Cas9-mediated knockout of factors in non-homologous end joining pathway enhances gene targeting in silkworm cells. *Scientific Reports* **5,** 18103. doi:10.1038/srep18103 (2015).

349. Kim Y., Ahmed S., Baki M. A. A., Kumar S., Kim K., Park Y. & Stanley D. Deletion mutant of PGE2 receptor using CRISPR-Cas9 exhibits larval immunosuppression and adult infertility in a lepidopteran insect, Spodoptera exigua. *Developmental & Comparative Immunology* **111,** 103743. doi:10.1016/j.dci.2020.103743 (2020).

350. Ye Z.-F., Liu X.-L., Han Q., Liao H., Dong X.-T., Zhu G.-H. & Dong S.-L. Functional characterization of PBP1 gene in Helicoverpa armigera (Lepidoptera: Noctuidae) by using the CRISPR/Cas9 system. *Scientific Reports* **7,** 8470. doi:10.1038/s41598-017-08769-2 (2017).

351. Hay B. A., Oberhofer G. & Guo M. Engineering the Composition and Fate of Wild Populations with Gene Drive. *Annual Review of Entomology* **66,** 407–438. doi:10.1146/annurev-ento-020117-043154 (2021).

352. Sinkins S. P. & Gould F. Gene drive systems for insect disease vectors. *Nature Reviews Genetics* **7,** 427–435. doi:10.1038/nrg1870 (2006).

353. Curtis C. F. Possible Use of Translocations to fix Desirable Genes in Insect Pest Populations. *Nature* **218,** 368–369. doi:10.1038/218368a0 (1968).

354. Webber B. L., Raghu S. & Edwards O. R. Opinion: Is CRISPR-based gene drive a biocontrol silver bullet or global conservation threat? *Proceedings of the National Academy of Sciences of the United States of America* **112,** 10565–10567. doi:10.1073/pnas.1514258112 (2015).

355. Champer J., Wen Z., Luthra A., Reeves R., Chung J., Liu C., Lee Y. L., Liu J., Yang E., Messer P. W. & Clark A. G. CRISPR Gene Drive Efficiency and Resistance Rate Is Highly Heritable with No Common Genetic Loci of Large Effect. *Genetics* **212,** genetics.302037.2019. doi:10.1534/genetics.119.302037 (2019).

356. Schmidt H., Collier T. C., Hanemaaijer M. J., Houston P. D., Lee Y. & Lanzaro G. C. Abundance of conserved CRISPR-Cas9 target sites within the highly polymorphic genomes of Anopheles and Aedes mosquitoes. *Nature Communications* **11,** 1425. doi:10.1038/s41467-020-15204-0 (2020).

357. Champer J., Yang E., Lee E., Liu J., Clark A. G. & Messer P. W. A CRISPR homing gene drive targeting a haplolethal gene removes resistance alleles and successfully spreads through a cage population. *Proceedings of the National Academy of Sciences of the United States of America* **117,** 24377–24383. doi:10.1073/pnas.2004373117 (2020).

358. Yan Y. & Finnigan G. C. Development of a multi-locus CRISPR gene drive system in budding yeast. *Scientific Reports* **8,** 17277. doi:10.1038/s41598-018-34909-3 (2018).

359. Nash A., Urdaneta G. M., Beaghton A. K., Hoermann A., Papathanos P. A., Christophides G. K. & Windbichler N. Integral gene drives for population replacement. *Biology Open* **8,** bio.037762. doi:10.1242/bio.037762 (2018).

360. Adolfi A., Gantz V. M., Jasinskiene N., Lee H.-F., Hwang K., Terradas G., Bulger E. A., Ramaiah A., Bennett J. B., Emerson J. J., Marshall J. M., Bier E. & James A. A. Efficient population modification gene-drive rescue system in the malaria mosquito Anopheles stephensi. *Nature Communications* **11,** 5553. doi:10.1038/s41467-020-19426-0 (2020).

361. Li M., Yang T., Kandul N. P., Bui M., Gamez S., Raban R., Bennett J., C H. M. S., Lanzaro G. C., Schmidt H., Lee Y., Marshall J. M. & Akbari O. S. Development of a confinable gene drive system in the human disease vector Aedes aegypti. *eLife* **9,** e51701. doi:10.7554/elife.51701 (2020).

362. Kandul N. P., Liu J., C. H. M. S., Wu S. L., Marshall J. M. & Akbari O. S. Transforming insect population control with precision guided sterile males with demonstration in flies. *Nature Communications* **10,** 84. doi:10.1038/s41467-018-07964-7 (2019).

363. Raban R. R., Marshall J. M. & Akbari O. S. Progress towards engineering gene drives for population control. *Journal of Experimental Biology* **223,** jeb208181. doi:10.1242/jeb.208181 (2020).

364. Bassett A. R., Tibbit C., Ponting C. P. & Liu J. L. Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9. *Biology Open* **3,** 42–49. doi:10.1242/bio.20137120 (2014).

365. Böttcher R., Hollmann M., Merk K., Nitschko V., Obermaier C., Philippou-Massier J., Wieland I., Gaul U. & Förstemann K. Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. *Nucleic Acids Research* **42,** e89. doi:10.1093/nar/gku289 (2014).

366. Mabashi-Asazuma H., Kuo C.-W., Khoo K.-H. & Jarvis D. L. Modifying an Insect Cell N-Glycan Processing Pathway Using CRISPR-Cas Technology. *ACS Chemical Biology* **10,** 2199–2208. doi:10.1021/acschembio.5b00340 (2015).

367. Chang J., Wang R., Yu K., Zhang T., Chen X., Liu Y., Shi R., Wang X., Xia Q. & Ma S. Genome-wide CRISPR screening reveals genes essential for cell viability and resistance to abiotic and biotic stresses in Bombyx mori. *Genome Research* **30,** 757–767. doi:10.1101/gr.249045.119 (2020).

368. Anderson M. A. E., Purcell J., Verkuijl S. A. N., Norman V. C., Leftwich P. T., Harvey-Samuel T. & Alphey L. S. Expanding the CRISPR Toolbox in Culicine Mosquitoes: In Vitro Validation of Pol III Promoters. *ACS Synthetic Biology* **9,** 678–681. doi:10.1021/acssynbio.9b00436 (2020).

369. Smail S. S., Ayesh K., Sierra-Montes J. M. & Herrera R. J. U6 snRNA variants isolated from the posterior silk gland of the silk moth Bombyx mori. *Insect Biochemistry and Molecular Biology* **36,** 454–465. doi:10.1016/j.ibmb.2006.03.004 (2006).

370. Hernandez G., Valafar F. & Stumph W. E. Insect small nuclear RNA gene promoters evolve rapidly yet retain conserved features involved in determining promoter activity and RNA polymerase specificity. *Nucleic Acids Research* **35,** 21–34. doi:10.1093/nar/gkl982 (2006).

371. Rozen-Gagnon K., Yi S., Jacobson E., Novack S. & Rice C. M. A selectable, plasmid-based system to generate CRISPR/Cas9 gene edited and knock-in mosquito cell lines. *Scientific Reports* **11,** 736. doi:10.1038/s41598-020-80436-5 (2021).

372. Ghosh S., Tibbit C. & Liu J.-L. Effective knockdown of Drosophila long non-coding RNAs by CRISPR interference. *Nucleic Acids Research* **44,** e84. doi:10.1093/nar/gkw063 (2016).

373. Chavez A., Scheiman J., Vora S., Pruitt B. W., Tuttle M., Iyer E. P. R., Lin S., Kiani S., Guzman C. D., Wiegand D. J., Ter-Ovanesyan D., Braff J. L., Davidsohn N., Housden B. E., Perrimon N., Weiss R., Aach J., Collins J. J. & Church G. M. Highly efficient Cas9-mediated transcriptional programming. *Nature Methods* **12,** 326–328. doi:10.1038/nmeth.3312 (2015).

374. Chavez A., Tuttle M., Pruitt B. W., Ewen-Campen B., Chari R., Ter-Ovanesyan D., Haque S. J., Cecchi R. J., Kowal E. J. K., Buchthal J., Housden B. E., Perrimon N., Collins J. J. & Church G. Comparison of Cas9 activators in multiple species. *Nature Methods* **13,** 563–567. doi:10.1038/nmeth.3871 (2016).

375. Lin S., Ewen-Campen B., Ni X., Housden B. E. & Perrimon N. In Vivo Transcriptional Activation Using CRISPR/Cas9 in Drosophila. *Genetics* **201,** 433–442. doi:10.1534/genetics.115.181065 (2015).

376. Wang X., Ma S., Liu Y., Lu W., Sun L., Zhao P. & Xia Q. Transcriptional repression of endogenous genes in BmE cells using CRISPRi system. *Insect Biochemistry and Molecular Biology* **111,** 103172. doi:10.1016/j.ibmb.2019.05.007 (2019).

377. Wang X.-G., Ma S.-Y., Chang J.-S., Shi R., Wang R.-L., Zhao P. & Xia Q.-Y. Programmable activation of Bombyx gene expression using CRISPR/dCas9 fusion systems. *Insect Science* **4,** 220. doi:10.1111/1744-7917.12634 (2018).

378. Liu Y., Ma S., Chang J., Zhang T., Chen X., Liang Y. & Xia Q. Programmable targeted epigenetic editing using CRISPR system in Bombyx mori. *Insect Biochemistry and Molecular Biology* **110,** 105–111. doi:10.1016/j.ibmb.2019.04.013 (2019).

379. Dong Z., Huang L., Dong F., Hu Z., Qin Q., Long J., Cao M., Chen P., Lu C. & Pan M.-H. Establishment of a baculovirus-inducible CRISPR/Cas9 system for antiviral research in transgenic silkworms. *Applied Microbiology and Biotechnology* **102,** 9255–9265. doi:10.1007/s00253-018-9295-8 (2018).

380. Pazmiño-Ibarra V., Mengual-Martí A., Targovnik A. M. & Herrero S. Improvement of baculovirus as protein expression vector and as biopesticide by CRISPR/Cas9 editing. *Biotechnology and Bioengineering* **116,** 2823–2833. doi:10.1002/bit.27139 (2019).

381. Moscardi F. Assessment of the application of baculoviruses for control of Lepidoptera. *Annual Review of Entomology* **44,** 257–289. doi:10.1146/annurev.ento.44.1.257 (1999).

382. Datsenko K. A. & Wanner B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 6640–6645. doi:10.1073/pnas.120163297 (2000).

383. Muyrers J. P. P., Zhang Y., Testa G. & Stewart A. F. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Research* **27,** 1555–1557. doi:10.1093/nar/27.6.1555 (1999).

384. Stewart T. M., Huijskens I., Willis L. G. & Theilmann D. A. The Autographa californica multiple nucleopolyhedrovirus ie0-ie1 gene complex is essential for wild-type virus replication, but either IE0 or IE1 can support virus growth. *Journal of Virology* **79,** 4619–4629. doi:10.1128/jvi.79.8.4619-4629.2005 (2005).

385. Vanarsdall A. L., Okano K. & Rohrmann G. F. Characterization of a baculovirus with a deletion of vlf-1. *Virology* **326,** 191–201. doi:10.1016/j.virol.2004.06.003 (2004).

386. Lin G. & Blissard G. W. Analysis of an Autographa californica nucleopolyhedrovirus lef-11 knockout: LEF-11 is essential for viral DNA replication. *Journal of Virology* **76,** 2770–2779. doi:10.1128/jvi.76.6.2770-2779.2002 (2002).

387. Fang M., Dai X. & Theilmann D. A. Autographa californica Multiple Nucleopolyhedrovirus EXON0 (ORF141) Is Required for Efficient Egress of Nucleocapsids from the Nucleus. *Journal of Virology* **81,** 9859–9869. doi:10.1128/jvi.00588-07 (2007).

388. McCarthy C. B. & Theilmann D. A. AcMNPV ac143 (odv-e18) is essential for mediating budded virus production and is the 30th baculovirus core gene. *Virology* **375,** 277–291. doi:10.1016/j.virol.2008.01.039 (2008).

389. Zhu S., Wang W., Wang Y., Yuan M. & Yang K. The baculovirus core gene ac83 is required for nucleocapsid assembly and Per Os infectivity of Autographa californica Nucleopolyhedrovirus. *Journal of Virology* **87,** 10573–10586. doi:10.1128/jvi.01207-13 (2013).

390. Boogaard B., Evers F., van Lent J. W. M. & van Oers M. M. The baculovirus Ac108 protein is a per os infectivity factor and a component of the ODV entry complex. *Journal of General Virology* **100,** 669–678. doi:10.1099/jgv.0.001200 (2019).

391. Ros V. I. D., Houte S., Hemerik L. & van Oers M. M. Baculovirus-induced tree-top disease: how extended is the role of egt as a gene for the extended phenotype? *Molecular Ecology* **24,** 249–258. doi:10.1111/mec.13019 (2015).

392. Westenberg M., Soedling H. M., Mann D. A., Nicholson L. J. & Dolphin C. T. Counter-selection recombineering of the baculovirus genome: A strategy for seamless modification of repeat-containing BACs. *Nucleic Acids Research* **38,** e166–e166. doi:10.1093/nar/gkq596 (2010).

393. Monsma S. A., Oomens A. G. & Blissard G. W. The GP64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. *Journal of Virology* **70,** 4607–4616. doi:10.1128/jvi.70.7.4607-4616.1996 (1996).

394. Zhao H.-F., L'Abbé D., Jolicoeur N., Wu M., Li Z., Yu Z. & Shen S.-H. High-throughput screening of effective siRNAs from RNAi libraries delivered via bacterial invasion. *Nature Methods* **2,** 967–973. doi:10.1038/nmeth812 (2005).

395. Xu J., Nagata Y., Mon H., Li Z., Zhu L., Iiyama K., Kusakabe T. & Lee J. M. Soaking RNAi-mediated modification of Sf9 cells for baculovirus expression system by ectopic expression of Caenorhabditis elegans SID-1. *Applied Microbiology and Biotechnology* **97,** 5921–5931. doi:10.1007/s00253-013-4785-1 (2013).

396. Bruder M. R., Pyne M. E., Moo-Young M., Chung D. A. & Chou C. P. Extending CRISPR-Cas9 Technology from Genome Editing to Transcriptional Engineering in the Genus Clostridium. *Applied and Environmental Microbiology* **82,** 6109–6119. doi:10.1128/aem.02128-16 (2016).

397. Theilmann D. A. & Stewart S. Molecular analysis of the trans-activating IE-2 gene of Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus. *Virology* **187,** 84–96. doi:10.1016/0042-6822(92)90297-3 (1992).

398. Gibson D. G., Young L., Chuang R.-Y., Venter J. C., Hutchison C. A. I. & Smith H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6,** 343–U41. doi:10.1038/nmeth.1318 (2009).

399. Karasawa S., Araki T., Yamamoto-Hino M. & Miyawaki A. A Green-emitting fluorescent protein from Galaxeidae coral and its monomeric version for use in fluorescent labeling. *Journal of Biological Chemistry* **278,** 34167–34171. doi:10.1074/jbc.m304063200 (2003).

400. Chari R., Yeo N. C., Chavez A. & Church G. M. sgRNA Scorer 2.0: A Species-Independent Model To Predict CRISPR/Cas9 Activity. *ACS Synthetic Biology* **6,** 902–904. doi:10.1021/acssynbio.6b00343 (2017).

401. Stemmer W. P. & Morris S. K. Enzymatic inverse PCR: a restriction site independent, single-fragment method for high-efficiency, site-directed mutagenesis. *Biotechniques* **13,** 214–20 (1992).

402. Shen C. F., Meghrous J. & Kamen A. Quantitation of baculovirus particles by flow cytometry. *Journal of Virological Methods* **105,** 321–330. doi:10.1016/s0166-0934(02)00128-3 (2002).

403. Wang Y., Wang F., Wang R., Zhao P. & Xia Q. 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. *Scientific Reports* **5,** 16273. doi:10.1038/srep16273 (2015).

404. Ooi B. G. & Miller L. K. Regulation of Host RNA Levels during Baculovirus Infection. *Virology* **166,** 515–523. doi:10.1016/0042-6822(88)90522-3 (1988).

405. Lin C.-H. & Jarvis D. L. Utility of temporally distinct baculovirus promoters for constitutive and baculovirus-inducible transgene expression in transformed insect cells. *Journal of Biotechnology* **165,** 11–17. doi:10.1016/j.jbiotec.2013.02.007 (2013).

406. Guarino L. A. & Summers M. D. Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene. *Journal of Virology* **57,** 563–571. doi:10.1128/jvi.57.2.563-571.1986 (1986).

407. Passarelli A. L. & Miller L. K. Three baculovirus genes involved in late and very late gene expression: ie-1, ie-n, and lef-2. *Journal of Virology* **67,** 2149–2158. doi:10.1128/jvi.67.4.2149-2158.1993 (1993).

408. Rapp J. C., Wilson J. A. & Miller L. K. Nineteen baculovirus open reading frames, including LEF-12, support late gene expression. *Journal of Virology* **72,** 10197–10206. doi:10.1128/JVI.72.12.10197-10206.1998 (1998).

409. Yang S. & Miller L. K. Activation of baculovirus very late promoters by interaction with very late factor 1. *Journal of Virology* **73,** 3404–3409. doi:10.1128/JVI.73.4.3404-3409.1999 (1999).

410. Mistretta T.-A. & Guarino L. A. Transcriptional activity of baculovirus very late factor 1. *Journal of Virology* **79,** 1958–1960. doi:10.1128/jvi.79.3.1958-1960.2005 (2005).

411. Kanginakudru S., Royer C., Edupalli S. V., Jalabert A., Mauchamp B., Chandrashekaraiah, Prasad S. V., Chavancy G., Couble P. & Nagaraju J. Targeting ie-1 gene by RNAi induces baculoviral resistance in lepidopteran cell lines and in transgenic silkworms. *Insect Molecular Biology* **16,** 635–644. doi:10.1111/j.1365-2583.2007.00753.x (2007).

412. Kakumani P. K., Malhotra P., Mukherjee S. K. & Bhatnagar R. K. A draft genome assembly of the army worm, Spodoptera frugiperda. *Genomics* **104,** 134–143. doi:10.1016/j.ygeno.2014.06.005 (2014).

413. Bodenhofer U., Bonatesta E., Horejš-Kainrath C. & Hochreiter S. msa: an R package for multiple sequence alignment. *Bioinformatics* **31,** 3997–3999. doi:10.1093/bioinformatics/btv494 (2015).

414. Tanaka H., Fujita K., Sagisaka A., Tomimoto K., Imanishi S. & Yamakawa M. ShRNA expression plasmids generated by a novel method efficiently induce gene-specific knockdown in a silkworm cell line. *Molecular Biotechnology* **41,** 173–179. doi:10.1007/s12033-008-9108-x (2009).

415. Zhu H., Zhang L., Tong S., Lee C. M., Deshmukh H. & Bao G. Spatial control of in vivo CRISPR–Cas9 genome editing via nanomagnets. *Nature Biomedical Engineering* **3,** 126–136. doi:10.1038/s41551-018-0318-7 (2019).

416. Maghodia A. B., Geisler C. & Jarvis D. L. A New Bacmid for Customized Protein Glycosylation Pathway Engineering in the Baculovirus-Insect Cell System. *ACS Chemical Biology.* doi:10.1021/acschembio.0c00974 (2021).

417. Pijlman G. P., Grose C., Hick T. A. H., Breukink H. E., Braak R. v. d., Abbo S. R., Geertsema C., van Oers M. M., Martens D. E. & Esposito D. Relocation of the attTn7 Transgene Insertion Site in Bacmid DNA Enhances Baculovirus Genome Stability and Recombinant Protein Expression in Insect Cells. *Viruses* **12,** 1448. doi:10.3390/v12121448 (2020).

418. Palmberger D., Wilson I. B. H., Berger I., Grabherr R. & Rendic D. SweetBac: A New Approach for the Production of Mammalianised Glycoproteins in Insect Cells. *PLoS ONE* **7,** e34226. doi:10.1371/journal.pone.0034226 (2012).

419. Shang Y., Wang M., Xiao G., Wang X., Hou D., Pan K., Liu S., Li J., Wang J., Arif B. M., Vlak J. M., Chen X., Wang H., Deng F. & Hu Z. Construction and Rescue of a Functional Synthetic Baculovirus. *ACS Synthetic Biology* **6,** 1393–1402. doi:10.1021/acssynbio.7b00028 (2017).

420. Lin L., Wang J., Deng R., Ke J., Wu H. & Wang X. ac109 is required for the nucleocapsid assembly of Autographa californica multiple nucleopolyhedrovirus. *Virus Research* **144,** 130–135. doi:10.1016/j.virusres.2009.04.010 (2009).

421. Li G., Chen H., Tang Q., Huang G., Deng R., Wang J. & Wang X. Effect of ac68 knockout and lef3 leading sequence disruption on viral propagation. *Current Microbiology* **62,** 191–197. doi:10.1007/s00284-010-9691-5 (2011).

422. Roy P. & Noad R. Use of Bacterial Artificial Chromosomes in Baculovirus Research and Recombinant Protein Expression: Current Trends and Future Perspectives. *ISRN Microbiology* **2012,** 1–11. doi:10.5402/2012/628797 (2012).

423. Wang H., Deng F., Pijlman G. P., Chen X., Sun X., Vlak J. M. & Hu Z. Cloning of biologically active genomes from a Helicoverpa armigera single-nucleocapsid nucleopolyhedrovirus isolate by using a bacterial artificial chromosome. *Virus Research* **97,** 57–63. doi:10.1016/j.virusres.2003.07.001 (2003).

424. Hilton S., Kemp E., Keane G. & Winstanley D. A bacmid approach to the genetic manipulation of granuloviruses. *Journal of Virological Methods* **152,** 56–62. doi:10.1016/j.jviromet.2008.05.015 (2008).

425. Pijlman G. P., Dortmans J. C. F. M., Vermeesch A. M. G., Yang K., Martens D. E., Goldbach R. W. & Vlak J. M. Pivotal Role of the Non-hr Origin of DNA Replication in the Genesis of Defective Interfering Baculoviruses. *Journal of Virology* **76,** 5605–5611. doi:10.1128/jvi.76.11.5605-5611.2002 (2002).

426. Yu M. & Carstens E. B. Identification of a Domain of the Baculovirus Autographa californica Multiple Nucleopolyhedrovirus Single-Strand DNA-Binding Protein LEF-3 Essential for Viral DNA Replication. *Journal of Virology* **84,** 6153–6162. doi:10.1128/jvi.00115-10 (2010).

427. Mishra G., Chadha P. & Das R. H. Serine/threonine kinase (pk-1) is a component of Autographa californica multiple nucleopolyhedrovirus (AcMNPV) very late gene transcription complex and it phosphorylates a 102kDa polypeptide of the complex. *Virus Research* **137,** 147–149. doi:10.1016/j.virusres.2008.05.014 (2008).

428. Liang C., Li M., Dai X., Zhao S., Hou Y., Zhang Y., Lan D., Wang Y. & Chen X. Autographa californica multiple nucleopolyhedrovirus PK-1 is essential for nucleocapsid assembly. *Virology* **443,** 349–357. doi:10.1016/j.virol.2013.05.025 (2013).

429. Peng Y., Li K., Pei R.-j., Wu C.-c., Liang C.-y., Wang Y. & Chen X.-w. The protamine-like DNA-binding protein P6.9 epigenetically up-regulates Autographa californica multiple nucleopolyhedrovirus gene transcription in the late infection phase. *Virologica Sinica* **27,** 57–68. doi:10.1007/s12250-012-3229-x (2012).

430. Wang M., Tuladhar E., Shen S., Wang H., van Oers M. M., Vlak J. M. & Westenberg M. Specificity of baculovirus P6.9 basic DNA-binding proteins and critical role of the C terminus in virion formation. *Journal of Virology* **84,** 8821–8828. doi:10.1128/jvi.00072-10 (2010).

431. Li A., Zhao H., Lai Q., Huang Z., Yuan M. & Yang K. Posttranslational Modifications of Baculovirus Protamine-Like Protein P6.9 and the Significance of Its Hyperphosphorylation for Viral Very Late Gene Hyperexpression. *Journal of Virology* **89,** 7646–7659. doi:10.1128/jvi.00333-15 (2015).

432. Labun K., Montague T. G., Krause M., Torres Cleuren Y. N., Tjeldnes H. & Valen E. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research* **47,** W171–W174. doi:10.1093/nar/gkz365 (2019).

433. Rohovie M. J., Nagasawa M. & Swartz J. R. Virus-like particles: Next-generation nanoparticles for targeted therapeutic delivery. *Bioengineering & Translational Medicine* **2,** 43–57. doi:10.1002/btm2.10049 (2017).

434. Buonaguro L., Tagliamonte M., Tornesello M. L. & Buonaguro F. M. Developments in virus-like particle-based vaccines for infectious diseases and cancer. *Expert Review of Vaccines* **10,** 1569–1583. doi:10.1586/erv.11.135 (2011).

435. Lua L. H., Connors N. K., Sainsbury F., Chuan Y. P., Wibowo N. & Middelberg A. P. Bioengineering virus-like particles as vaccines. *Biotechnology and Bioengineering* **111,** 425–440. doi:10.1002/bit.25159 (2014).

436. Haynes J. R. Influenza virus-like particle vaccines. *Expert Review of Vaccines* **8,** 435–445. doi:10.1586/erv.09.8 (2014).

437. Fuenmayor J., Gòdia F. & Cervera L. Production of virus-like particles for vaccines. *New Biotechnology* **39,** 174–180. doi:10.1016/j.nbt.2017.07.010 (2017).

438. Oomens A. G. P. & Blissard G. W. Requirement for GP64 to drive efficient budding of Autographa californica multicapsid nucleopolyhedrovirus. *Virology* **254,** 297–314. doi:10.1006/viro.1998.9523 (1999).

439. Wagner R., Fließbach H., Wanner G., Motz M., Niedrig M., Deby G., Brunn A. v. & Wolf H. Studies on processing, particle formation, and immunogenicity of the HIV-1gag gene product: a possible component of a HIV vaccine. *Archives of Virology* **127,** 117–137. doi:10.1007/bf01309579 (1992).

440. Gheysen D., Jacobs E., Foresta F. d., Thiriart C., Francotte M., Thines D. & Wilde M. D. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* **59,** 103–112. doi:10.1016/0092-8674(89)90873-8 (1989).

441. Garnier L., Ravallec M., Blanchard P., Chaabihi H., Bossy J. P., Devauchelle G., Jestin A. & Cerutti M. Incorporation of pseudorabies virus gD into human immunodeficiency virus type 1 Gag particles produced in baculovirus-infected cells. *Journal of Virology* **69,** 4060–4068. doi:10.1128/jvi.69.7.4060-4068.1995 (1995).

442. Puente-Massaguer E., Gòdia F. & Lecina M. Development of a non-viral platform for rapid virus-like particle production in Sf9 cells. *Journal of Biotechnology* **322,** 43–53. doi:10.1016/j.jbiotec.2020.07.009 (2020).

443. Puente-Massaguer E., Lecina M. & Gòdia F. Application of advanced quantification techniques in nanoparticle-based vaccine development with the Sf9 cell baculovirus expression system. *Vaccine* **38,** 1849–1859. doi:10.1016/j.vaccine.2019.11.087 (2020).

444. Creech C. B., Walker S. C. & Samuels R. J. SARS-CoV-2 Vaccines. *JAMA* **325,** 1318–1320. doi:10.1001/jama.2021.3199 (2021).

445. Palomares L. A., Mena J. A. & Ramírez O. T. Simultaneous expression of recombinant proteins in the insect cell-baculovirus system: Production of virus-like particles. *Methods* **56,** 389–395. doi:10.1016/j.ymeth.2012.01.004 (2012).

446. Roldão A., Vieira H. L., Alves P. M., Oliveira R. & Carrondo M. J. Intracellular dynamics in rotavirus-like particles production: Evaluation of multigene and monocistronic infection strategies. *Process Biochemistry* **41,** 2188–2199. doi:10.1016/j.procbio.2006.06.019 (2006).

447. Gómez-Sebastián S., López-Vidal J. & Escribano J. M. Significant productivity improvement of the baculovirus expression vector system by engineering a novel expression cassette. *PLoS ONE* **9,** e96562. doi:10.1371/journal.pone.0096562 (2014).

448. Bouvier N. M. & Palese P. The biology of influenza viruses. *Vaccine* **26,** D49–D53. doi:10.1016/j.vaccine.2008.07.039 (2008).

449. Ahmed M. & Kim D. R. pcr: an R package for quality assessment, analysis and testing of qPCR data. *PeerJ* **6,** e4473. doi:10.7717/peerj.4473 (2018).

450. Xue J.-L., Salem T. Z., Turney C. M. & Cheng X.-W. Strategy of the use of 28S rRNA as a housekeeping gene in real-time quantitative PCR analysis of gene transcription in insect cells infected by viruses. *Journal of Virological Methods* **163,** 210–215. doi:10.1016/j.jviromet.2009.09.019 (2010).

451. Huber W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods* **12,** 115–121. doi:10.1038/nmeth.3252 (2015).

452. Yin T., Cook D. & Lawrence M. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biology* **13,** R77. doi:10.1186/gb-2012-13-8-r77 (2012).

453. Thiem S. M. & Miller L. K. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus Autographa californica nuclear polyhedrosis virus. *Journal of Virology* **63,** 2008–2018. doi:10.1128/jvi.63.5.2008-2018.1989 (1989).

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

455. Sokal N., Nie Y., Willis L. G., Yamagishi J., Blissard G. W., Rheault M. R. & Theilmann D. A. Defining the roles of the baculovirus regulatory proteins IE0 and IE1 in genome replication and early gene transactivation. *Virology* **468-470,** 160–171. doi:10.1016/j.virol.2014.07.044 (2014).

456. Lee J.-H., Gwak W.-S., Bae S.-M., Choi J.-B., Han B.-K. & Woo S.-D. Increased productivity of the baculovirus expression vector system by combining enhancing factors. *Journal of Asia-Pacific Entomology* **21,** 1079–1084. doi:10.1016/j.aspen.2018.08.003 (2018).

457. Koczka K., Peters P., Ernst W., Himmelbauer H., Nika L. & Grabherr R. Comparative transcriptome analysis of a Trichoplusia ni cell line reveals distinct host responses to intracellular and secreted protein products expressed by recombinant baculoviruses. *Journal of Biotechnology* **270,** 61–69. doi:10.1016/j.jbiotec.2018.02.001 (2018).

458. Weyer U. & Possee R. D. Functional analysis of the p10 gene 5′ leader sequence of the Autographa californica nuclear polyhedrosis virus. *Nucleic Acids Research* **16,** 3635–3653. doi:10.1093/nar/16.9.3635 (1988).

459. Sano K.-I., Maeda K., Oki M. & Maéda Y. Enhancement of protein expression in insect cells by a lobster tropomyosin cDNA leader sequence. *FEBS Letters* **532,** 143–146. doi:10.1016/s0014-5793(02)03659-1 (2002).

460. Tiwari P., Saini S., Upmanyu S., Benjamin B., Tandon R., Saini K. S. & Sahdev S. Enhanced expression of recombinant proteins utilizing a modified baculovirus expression vector. *Molecular Biotechnology* **46,** 80–89. doi:10.1007/s12033-010-9284-3 (2010).

461. Lo H.-R., Chou C.-C., Wu T.-Y., Yuen J. P.-Y. & Chao Y.-C. Novel baculovirus DNA elements strongly stimulate activities of exogenous and endogenous promoters. *Journal of Biological Chemistry* **277,** 5256–5264. doi:10.1074/jbc.m108895200 (2002).

462. Venkaiah B., Viswanathan P., Habib S. & Hasnain S. E. An Additional Copy of the Homologous Region (hr1) Sequence in the Autographa californica Multinucleocapsid Polyhedrosis Virus Genome Promotes Hyperexpression of Foreign Genes. *Biochemistry* **43,** 8143–8151. doi:10.1021/bi049953q (2004).

463. Rodems S. M. & Friesen P. D. The hr5 transcriptional enhancer stimulates early expression from the Autographa californica nuclear polyhedrosis virus genome but is not required for virus replication. *Journal of Virology* **67,** 5776–5785. doi:10.1128/jvi.67.10.5776-5785.1993 (1993).

464. Possee R. & Howard S. Analysis of the polyhedrin gene promoter of the Autographa californica nuclear polyhedrosis virus. *Nucleic Acids Research* **15,** 10233–10248. doi:10.1093/nar/15.24.10233 (1987).

465. Peakman T. C., GiCharles I., Sydenham M. A., Gewert D. R., Page M. J. & Makoff A. J. Enhanced expression of recombinant proteins in insect cells using a baculovirus vector containing a bacterial leader sequence. *Nucleic Acids Research* **20,** 6111–6112. doi:10.1093/nar/20.22.6111 (1992).

466. Hills D. & Crane-Robinson C. Baculovirus expression of human basic fibroblast growth factor from a synthetic gene: role of the Kozak consensus and comparison with bacterial expression. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1260,** 14–20. doi:10.1016/0167-4781(94)00171-x (1995).

467. Chang M.-J., Kuzio J. & Blissard G. W. Modulation of Translational Efficiency by Contextual Nucleotides Flanking a Baculovirus Initiator AUG Codon. *Virology* **259,** 369–383. doi:10.1006/viro.1999.9787 (1999).

468. Paulsson J. Models of stochastic gene expression. *Physics of Life Reviews* **2,** 157–175. doi:10.1016/j.plrev.2005.03.003 (2005).

469. McLeay R. C. & Bailey T. L. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* **11,** 165. doi:10.1186/1471-2105-11-165 (2010).

470. Ooi B. G., Rankin C. & Miller L. K. Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. *Journal of Molecular Biology* **210,** 721–736. doi:10.1016/0022-2836(89)90105-8 (1989).

471. Weyer U. & Possee R. D. Analysis of the promoter of the Autographa californica nuclear polyhedrosis virus p10 gene. *Journal of General Virology* **70,** 203–208. doi:10.1099/0022-1317-70-1-203 (1989).

472. Scheper G. C., Vries R. G., Broere M., Usmany M., Voorma H. O., Vlak J. M. & Thomas A. A. Translational properties of the untranslated regions of the p10 messenger RNA of Autographa californica multicapsid nucleopolyhedrovirus. *Journal of General Virology* **78,** 687–696. doi:10.1099/0022-1317-78-3-687 (1997).

473. Qin J., Liu A. & Weaver R. F. Studies on the Control Region of the p10 Gene of the Autographa californica Nuclear Polyhedrosis Virus. *Journal of General Virology* **70,** 1273–1279. doi:10.1099/0022-1317-70-5-1273 (1989).

474. Yang S. & Miller L. K. Control of Baculovirus Polyhedrin Gene Expression by Very Late Factor 1. *Virology* **248,** 131–138. doi:10.1006/viro.1998.9272 (1998).

475. Yang Q., Chen F. & Trempe J. P. Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins. *Journal of Virology* **68,** 4847–4856. doi:10.1128/jvi.68.8.4847-4856.1994 (1994).

476. Hölscher C., Hörer M., Kleinschmidt J. A., Zentgraf H., Bürkle A. & Heilbronn R. Cell lines inducibly expressing the adeno-associated virus (AAV) rep gene: requirements for productive replication of rep-negative AAV mutants. *Journal of Virology* **68,** 7169–7177. doi:10.1128/jvi.68.11.7169-7177.1994 (1994).

477. Schmidt M., Afione S. & Kotin R. M. Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *Journal of Virology* **74,** 9441–9450. doi:10.1128/jvi.74.20.9441-9450.2000 (2000).

478. Li G., Wang J., Deng R. & Wang X. Characterization of AcMNPV with a deletion of ac68 gene. *Virus Genes* **37,** 119–127. doi:10.1007/s11262-008-0238-9 (2008).

479. Vanarsdall A. L., Okano K. & Rohrmann G. F. Characterization of the replication of a baculovirus mutant lacking the DNA polymerase gene. *Virology* **331,** 175–180. doi:10.1016/j.virol.2004.10.024 (2005).

480. Chen G., Yan Q., Fang Y., Wu L., Krell P. J. & Feng G. The N Terminus of Autographa californica Multiple Nucleopolyhedrovirus DNA Polymerase Is Required for Efficient Viral DNA Replication and Virus and Occlusion Body Production. *Journal of Virology* **92,** e00398–18. doi:10.1128/jvi.00398-18 (2018).

481. Lung O., Westenberg M., Vlak J. M., Zuidema D. & Blissard G. W. Pseudotyping Autographa californica Multicapsid Nucleopolyhedrovirus (AcMNPV): F Proteins from Group II NPVs Are Functionally Analogous to AcMNPV GP64. *Journal of Virology* **76,** 5729–5736. doi:10.1128/jvi.76.11.5729-5736.2002 (2002).

482. Vanarsdall A. L., Okano K. & Rohrmann G. F. Characterization of the role of very late expression factor 1 in baculovirus capsid structure and DNA processing. *Journal of Virology* **80,** 1724–1733. doi:10.1128/jvi.80.4.1724-1733.2006 (2006).

483. Xiang X., Chen L., Hu X., Yu S., Yang R. & Wu X. Autographa californica multiple nucleopolyhedrovirus odv-e66 is an essential gene required for oral infectivity. *Virus Research* **158,** 72–78. doi:10.1016/j.virusres.2011.03.012 (2011).

484. Yang M., Wang S., Yue X. L. & Li L.-L. Autographa californica multiple nucleopolyhedrovirus Orf132 encodes a nucleocapsid-associated protein required for budded-virus and multiply enveloped occlusion-derived virus production. *Journal of Virology* **88,** 12586–12598. doi:10.1128/jvi.01313-14 (2014).

485. Li Y., Wang J., Deng R., Zhang Q., Yang K. & Wang X. vlf-1 deletion brought AcMNPV to defect in nucleocapsid formation. *Virus Genes* **31,** 275–284. doi:10.1007/s11262-005-3242-3 (2005).

486. Bai H., Hu Y., Hu X., Li J., Mu J., Zhou Y., Chen X. & Wang Y. Major capsid protein of Autographa californica multiple nucleopolyhedrovirus contributes to the promoter activity of the very late viral genes. *Virus Research* **273,** 197758. doi:10.1016/j.virusres.2019.197758 (2019).

487. Gomi S., Majima K. & Maeda S. Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus. *Journal of General Virology* **80,** 1323–1337. doi:10.1099/0022-1317-80-5-1323 (1999).

Appendix A

Supplementary Information for Chapter 2

A.1 Supplementary tables

Table A.1: Primers used in this study

Plasmid Construct	Sequence (5'-3')	Use (template)
pOpIE2-dCas9-puro	ctgtcattatcttagtttgtattgtcatg gctcatgttgtgtcgctg	OpIE2 5'/3' UTR (pOpIE2E2.3)
	tacagcgacacaacatgagcatggacaagaagtattctatcggactg ccctcaagcttgcctgatccacctaccttgcgcttcttcttg	dCas9 (pdCas9:BFP)
	ggatcaggcaagcttgagg acaaactaagataatgacagctggatccctcgagtcagg	T2A/pac (pAc-sgRNA-Cas9)
pOpIE2-Cas9-puro	tacagcgacacaacatgagcatggacaagaagtattctatcggactg	Cas9-puro (pAc-sgRNA-Cas9)
rBEV transfer plasmids	cactaacctaggtagctgagcgc cgctggactggcatgaac	rBEV ORF603/ORF1629
OpIE2GFP-sgRNA	cgaagttcatgccagtccagcggcaaggcgattaagttgggta gctcatgttgtgtcgctg	OpIE2 5' UTR
	tacagcgacacaacatgagcatggtgagcgtgatcaagc caaactaagataatgacagttacttggcctggctgggcagc	GFP CDS
	ctgtcattatcttagtttgtattgtcatg gattaagttgggtaacgccagttc	OpIE2 3' UTR
	gaactggcgttacccaacttaatcaatctgcttagggttaggcg ctcagctacctaggttagtggagagtgcaccatatgcggt	SfU6-sgRNA)
pACUW51-p10	gttctagtggttggctacgtatactccgcgttggagtcttgtgtgcta tgctttatttgtaaccattataagctgctcgctatacactcgcatggag	p10 region (AcMNPV genome)
	gcagcttataatggttacaaataaagc ggagtatacgtagccaaccactagaac	pACUW51 backbone
p10GFP-sgRNA	cgaagttcatgccagtccagcggatgaagtggttcgcatcct ggtgatggtgcatgattgtaaataaaatgtaatttacag	p10 5'UTR
	atgaatcgtttttaaaataacaaatcaattgttttataa gatgattaagttgggtaacgccagttcgacatgataagatacattg	p10 3'UTR
	atcatgcaccatcaccaccatcatatggtgagtgtgattaaaccag gttatttaaaaacgattcatggcgcgccttacttggcctggctgg	replace p10 ORF with *gfp* gene
	gaactggcgttacccaacttaatcatctgcttagggttaggcgttttg ctcagctacctaggttagtggagagtgcaccatatgcggt	SfU6-sgRNA
p6.9GFP-sgRNA	ccgaagttcatgccagtccagcgaaattccgttttgcgacg ccatatgatggtggtgatggtgcatgtttaaattgtgtaatttatg	p6.9 5' UTR to replace p10 5' UTR
Retarget sgRNAs	gttttagagctagaaatagcaagttaaaataagg cggtggtcgagcacga	retarget sgRNA† (Cas9 handle) retarget sgRNA† (SfU6)
qPCR primers	cgacgttgctttttgatcct gcaacgacaagccatcagta	28S
	gacgatcgctaggcatttag gcgtgttgcttgtgaaac	VP39
	gacaacagggagaagattgag ggaggttcttatcgaagttagtc	(d)Cas9
	tctacgacatcaggttcgacgg tccttcttggccttgtaggtgg	GFP

235

†: spacer sequence appended to 5' end of sequence

A.2 Supplementary figures

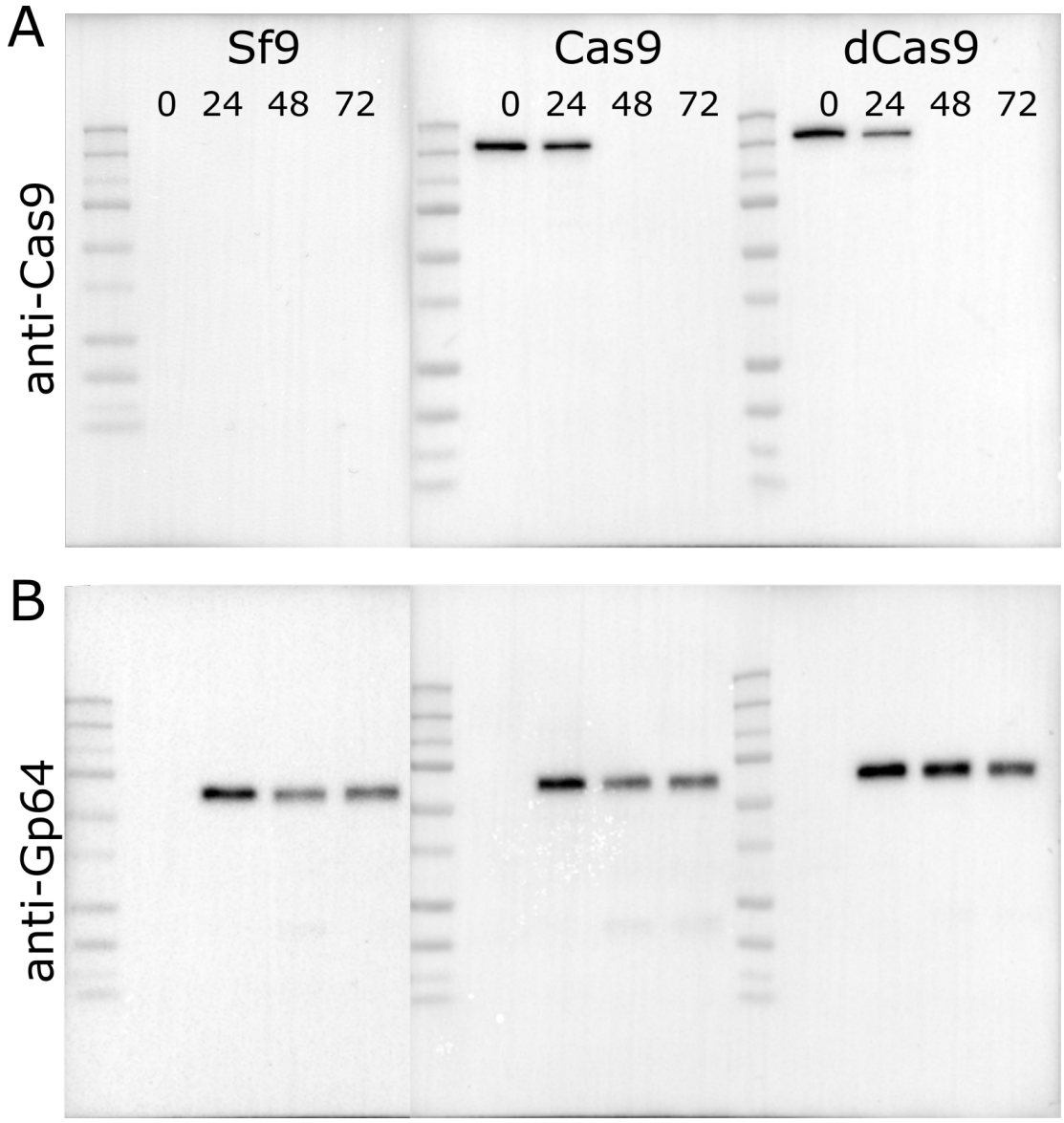


Figure A.1: **Expression of Cas9 and dCas9 is obstructed by infection.** Western blot analysis of infected Sf9, Sf9-Cas9, and Sf9-dCas9 cells for production of **A.** (d)Cas9 and **B.** GP64.

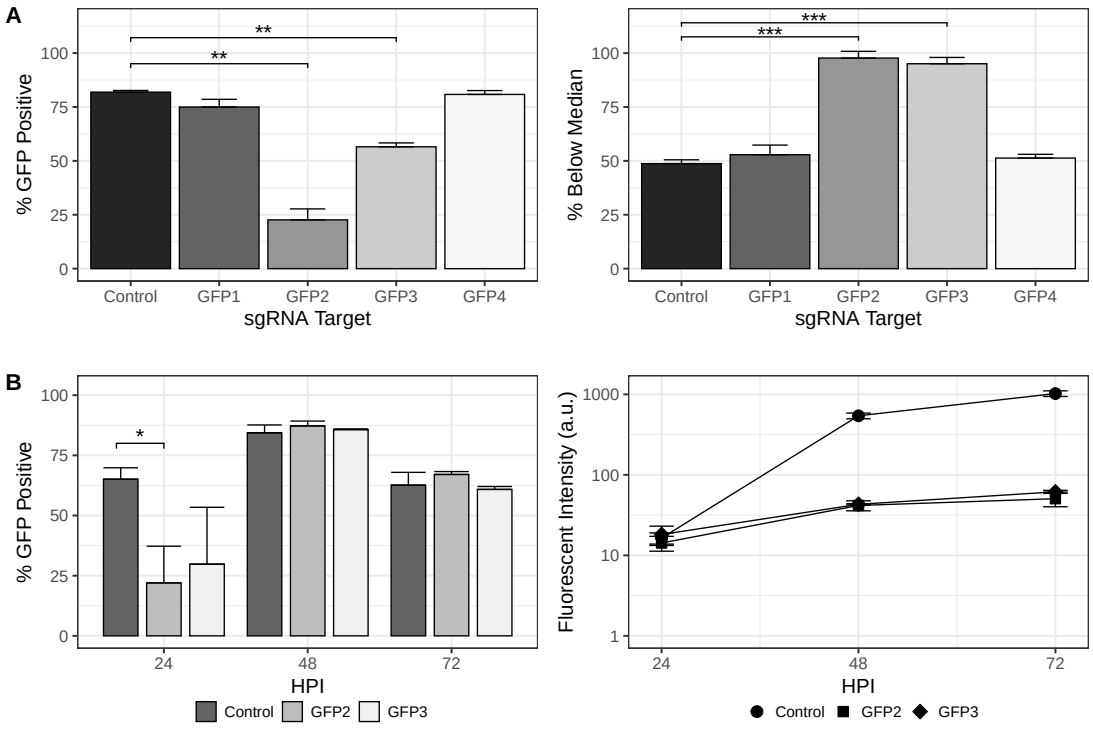


Figure A.2: **CRISPRi-mediated repression of *gfp* transcribed from OpIE2 or p10 promoters. A.** Percent GFP-positive cells and proportion with fluorescent intensity below the median of the control for OpIE2GFP-sgRNA rBEVs and **B.** Percent GFP-positive and median fluorescent intensity for p10GFP-sgRNA rBEVs with sgRNAs targeting the *gfp* gene.

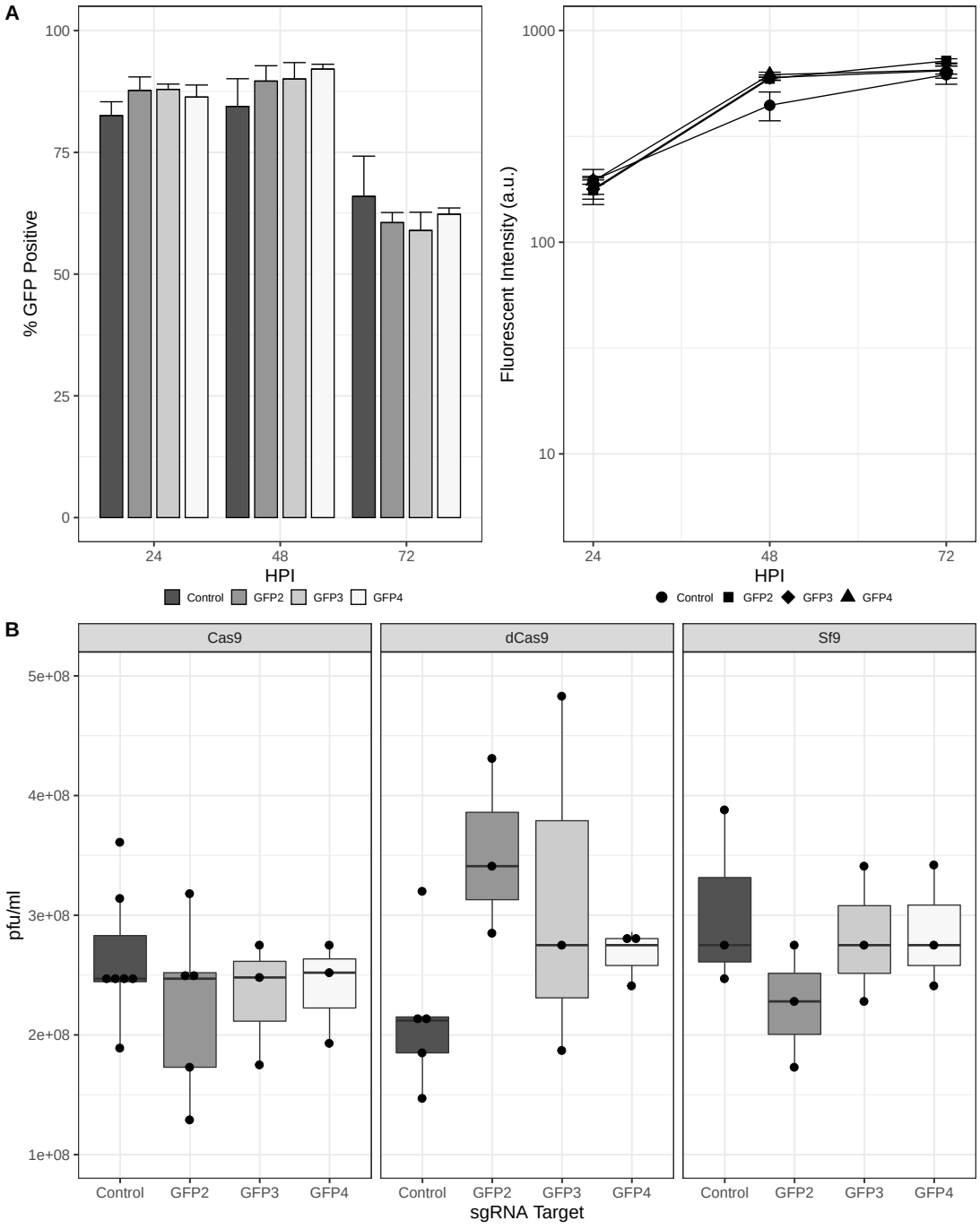


Figure A.3: **p6.9GFP-sgRNA rBEVs A.** Control infections of p6.9GFP-sgRNA rBEVs GFP2, GFP3, and GFP4 in Sf9 cells and **B.** IVT for GFP2, GFP3, and GFP4 from culture supernatants from infected Sf9-Cas9, Sf9-dCas9, and parental Sf9 cells.

Appendix B

Supplementary Information for Chapter 3

B.1 Supplementary tables

Table B.1: Primers used in this study

Plasmid Construct	Sequence (5'-3')	Use (template)
pOpIE2GFP	ctgtcattatcttagtttgtattgtcatg gttgtgtcgctgtaacagatgc	OpIE2 5'/3' UTR
	cagcatctgttacagcgacacaacatggtgagcgtgatcaagcccg caatacaaactaagataatgacagttacttggcctggctgggcagc	gfp ORF
p10GFP	ccgaagttcatgccagtccagcggcgttggagtcttgtgtgcta tcgggcttaatcacactcaccatgattgtaaataaaatgtaatttacag	p10 5' UTR
	atggtgagtgtgattaagcccg ttacttggcctggctgggcagc	gfp ORF
	gcccagccaggccaagtaaatgaatcgtttttaaaataacaaatcaattg tcgctatacactcgcatggag	p10 3' UTR
	ctccatgcgagtgtatagcgagcagcttataatggttacaaataaagc cgctggactggcatgaac	plasmid backbone
p6.9GFP	ccgaagttcatgccagtccagcgaaattccgttttgcgacg cgggcttaatcacactcaccatgtttaaattgtgtaatttatgtagc	p6.9 5' UTR
	gctacataaattacacaatttaaacatggtgagtgtgattaagcccg cgctggactggcatgaac	gfp ORF & backbone
pU6sgRNA	gctttcgctaaggatgatttctggaattctaaagatctgcttaggg cgctacggcgtttcacttctgaggagagtgcaccatatgcggtg	SfU6-sgRNA‘
	ctcagaagtgaaacgccgtag ctttagaattccagaaatcatccttagc	plasmid backbone
Retarget sgRNAs	gttttagagctagaaatagcaagttaaaataagg cggtggtcgagcacga	retarget sgRNA† (fwd primer) retarget sgRNA† (rev primer)

†: spacer sequence appended to 5' end of sequence

Table B.2: Protospacer sequences for CRISPR targets

Gene	Target	Protospacer Sequence (5'-3')	PAM	Strand
GP64	+131	GGAAACGCTGCAAAAGGACG	TGG	Template
GP64	-160	GTTGTAGTCCGTCTCCACGA	TGG	Nontemplate
GP64	+384	TTTCGCGACAACGAGGGCCG	CGG	Template

Table B.3: Comparison of AcMNPV and BmNPV KOV phenotypes

Target		Phenotype			
Gene	ORF No.†	AcMNPV KOV Description from literature	Ref.	BmNPV‡	This Study
ac68	Ac68/Bm56	No Impairment for BV	[421, 478]	A	A
dnapol	Ac65/Bm53	no DNA replication or BV release; late expression not measured	[479, 480]	C	C
gp64	Ac128/Bm105	no BV release	[481, 482]	C	C
ie-1	Ac147/Bm123	Impaired DNA replication & late gene expression	[384, 407]	D	D
lef-3	Ac67/Bm55	Impaired DNA replication & late gene expression	[421, 426]	D	D
odv-e66	Ac46/Bm37	No Impairment for BV	[483]	A	A
orf132	Ac68/Bm109	Impaired nucleocapsid assembly & BV release	[484]	C	C
p6.9	Ac100/Bm84	Impaired nucleocapsid assembly, BV release, & very late gene expression	[429, 431]	C	C
p74	Ac138/Bm115	No Impairment for BV	[81]	A	A/B
pk-1	Ac10/Bm3	Impaired nucleocapsid assembly, BV release,	[428, 431]	C	C/D
vlf-1	Ac77/Bm63	Impaired nucleocapsid assembly, BV release, & very late gene transcription	[385, 485]	C	C
vp39	Ac89/Bm72	Impaired BV release & very late gene transcription	[85, 486]	C	C
vp80	Ac104/Bm88	Impaired nucleocapsid maturation & BV release	[85]	C	C

†: ORF numbers refer to previous reports ([15, 487]), ‡: Taken from [98]

B.2 Supplementary figures

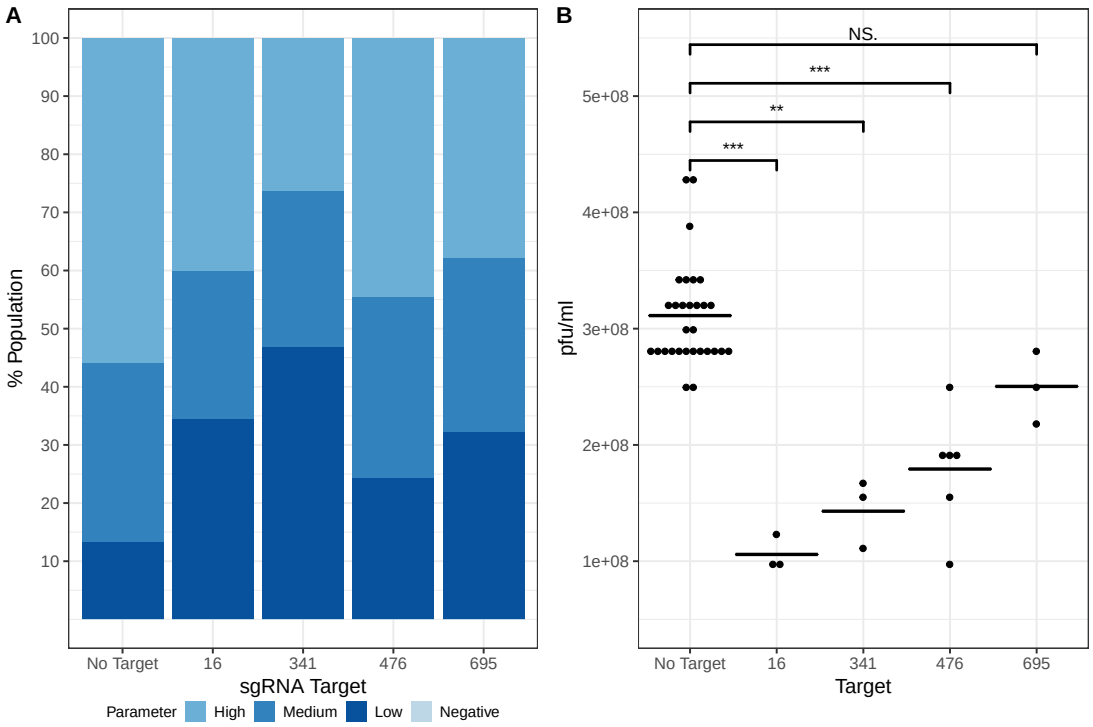


Figure B.1: **The location of the targeting spacer sequence within the *lef-3* ORF may impact the observed phenotype.** The impact of KOVs generated with different sgRNAs on **A.** GFP production and **B.** IVT.

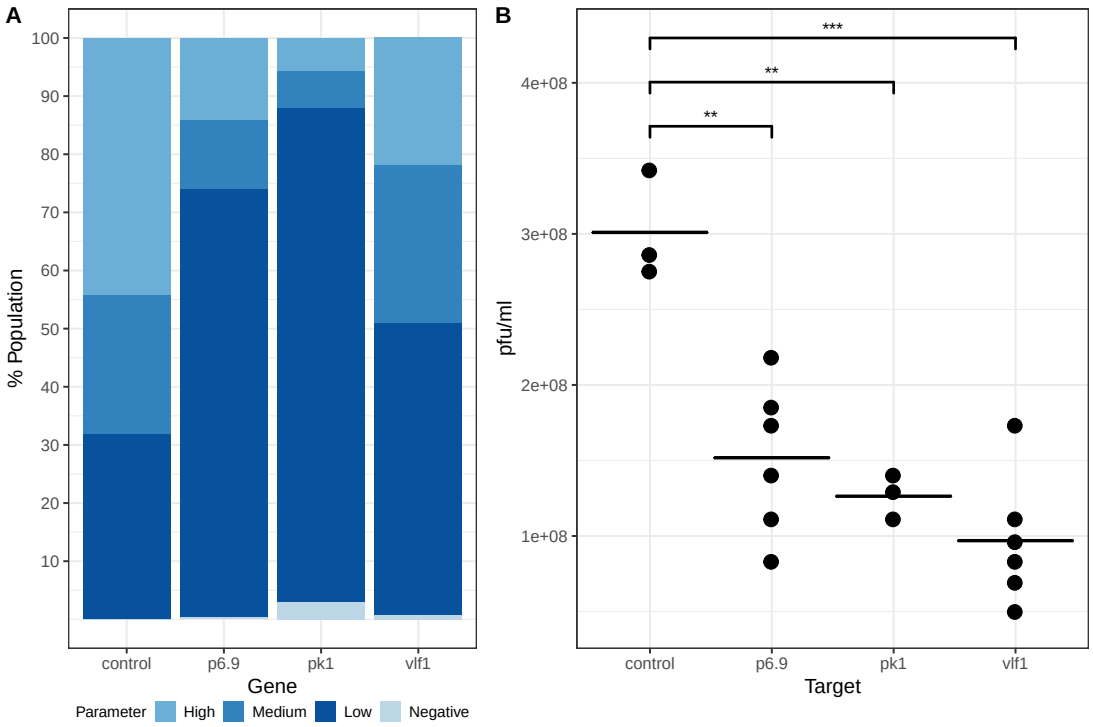


Figure B.2: Disrupting non-structural auxillary genes primarily involved in very late gene expression and genome packaging has significant effects on **A.** GFP production from the very late p10 promoter and reductions in **B.** IVT.

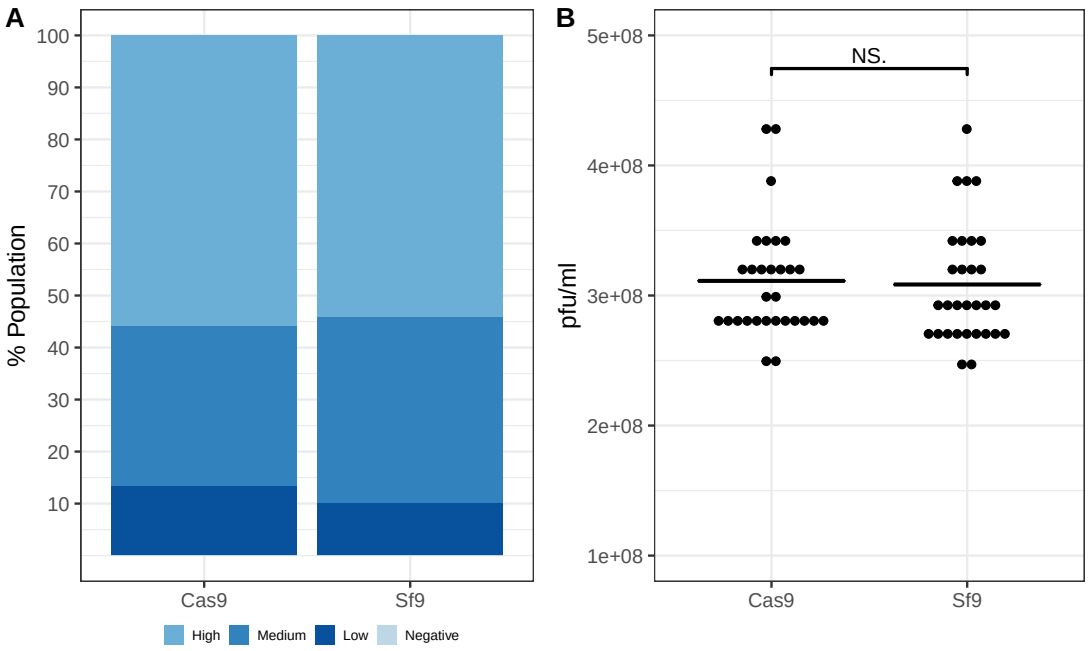


Figure B.3: **The presence of Cas9 and sgRNA are both necessary for CRISR-mediated gene disruption resulting in reduction of fluorescence or IVT. A.** Fluorescence and **B.** IVT for untargeted sgRNAs in Sf9-Cas9 cells and 10 targeted sgRNAs in Sf9 cells are indistinguishable.

Appendix C

Supplementary Information for Chapter 5

C.1 Supplementary tables

Table C.1: Primers used in this study

Plasmid Construct	Sequence (5'-3')	Description
Promoter-GFP	atggtgagtgtgattaagccc cgctggactggcatgaac	pHR-GFP (promoterless)
	caccgaagttcatgccagtccagcg ccacttgcgagttttgcag cttcatctcgggcttaatcacactcaccat attgttgccgttataaatatgg	vp39 promoter
	caccgaagttcatgccagtccagcg cagcgaagaggagaacaaca cttcatctcgggcttaatcacactcaccat tttaacaaacacttatcaatctattgagc	orf75 promoter
	caccgaagttcatgccagtccagcg tcgtacagctcagtgtacagtttg gatcttcatctcgggcttaatcacactcaccat aatgtacaaaaatggaccagttacg	38k promoter
	caccgaagttcatgccagtccagcg ccccccaaaaattgcac gatcttcatctcgggcttaatcacactcaccat gtttgcttcttgtaaacctttgaaac	39k promoter
	caccgaagttcatgccagtccagcg tcgcccgacgatca cttcatctcgggcttaatcacactcaccat attaacggcgatttttttaaattatc	ctx promoter
	caccgaagttcatgccagtccagcg ggttagttccagatagccatcg cttcatctcgggcttaatcacactcaccat cttgcttgtgtgttccttattga	gp64 promoter
	gtatattaattaaaatac atggtgagtgtgattaagccc ggcttaatcacactcaccat gtattttaattaatatacaaatgatttgataataattc	Δp10 promoter
	caccgaagttcatgccagtccagcg aaattccgttttgcgacg catctcgggcttaatcacactcaccat gtttaaattgtgtaatttatgtagctgta	p6.9 promoter
	caccgaagttcatgccagtccagcgcgtaggcctttgaattccg ggcttaatcacactcaccat atttataggttttttttattacaaaactgttacgaaaacag	polh promoter
Promoter-SEAP	ggcgcgccatgaatcgtttttaaaataac	pHR-promoter fwd
	atgcttctcttattgctgctgctgggcctgag gttatttttaaaaacgattcatggcgcgcc ttatgtctgctcgaagcggcc	SEAP gene
	caggcccagcagcagcaataagagaagcat attgttgccgttataaatatgg	vp39 rev
	caggcccagcagcagcaataagagaagcat tttaacaaacacttatcaatctattgagc	orf75 rev
	caggcccagcagcagcaataagagaagcat aatgtacaaaaatggaccagttacg	38k rev
	caggcccagcagcagcaataagagaagcat gtttgcttcttgtaaacctttgaaac	39k rev
	caggcccagcagcagcaataagagaagcat attaacggcgatttttttaaattatc	ctx rev
	caggcccagcagcagcaataagagaagcat cttgcttgtgtgttccttattga	gp64 rev
	caggcccagcagcagcaataagagaagcat gtattttaattaatatacaaatgatttg	Δp10 rev
	caggcccagcagcagcaataagagaagcat gtttaaattgtgtaatttatgtagctgta	p6.9 rev
	caggcccagcagcagcaataagagaagcat atttataggttttttttattacaaaactg	polh rev
qPCR primers	cgacgttgctttttgatcct gcaacgacaagccatcagta	28S
	tctacgacatcaggttcgacgg tccttcttggccttgtaggtgg	GFP
	agtacccagatgactacagc ggatctcgtatttcatgtctcc	SEAP

Table C.2: Promoters on commercially available BEVS transfer plasmids

System	Family	Plasmid	Expression	Promoter(s)	polyA
Transposition	pFastbac™	pFastbac-1, HT	single	polh	SV40
		pFastbac Dual	dual	polh/p10	SV40, HSV TK
	MultiBac™	pIDC, pIDK, pIDS	single‡‡	polh/p10	SV40, HSV TK
		pFL, pKL, pSPL, pUCDM	dual	polh/p10	SV40, HSV TK
		pACEBac1, pACEBac2	single‡‡	polh/p10	SV40, HSV TK
Homologous Recombination	pBAC™	pBAC-1/2/3	single	polh	n/a
		pBAC-4x	multi	polh/p10	synthetic
		pBAC-5	single	gp64	n/a
		pBAC-6†	single	gp64	n/a
		pBACsurf-1‡	single	polh	n/a
	pIEx/Bac™	pIEx/Bac-1/3/4/5	single	*hr5*-ie1-p10	ie1
	pTriEx™	pTriEx-1.1/2/3/4/5/6/7	single	p10	rabbit β-globin
	pAB™	pAB-6xHis/GST/MBP	single	polh	n/a
		pAB-bee/bee-8xHis/bee-FH†	single	polh	n/a

†: secretion signal included; ‡: gp64 fusion for surface display; ‡‡: contains multiplication element for multigene compatability

Table C.3: Position and sequence of putative upstream octamer matches in relation to TAAG motif.

ORF	Start	End	Sequence
polh	109	102	ATTGTAAT
p6.9	209	202	ATTACAAT
p6.9	152	145	ATTGCAAG
p6.9	131	124	ATTACAAT
p6.9	27	20	AATGCAAA
Ac-bro	203	196	ATTGCCAC
ctx	203	196	ATTGCCAC
orf73	120	113	ATTGAAAC
orf73	90	83	ATTGAAAA
orf73	34	27	ATTGCAAA
orf74	115	108	ATTGCATA
orf74	22	15	ACTGCCAG
orf74	7	0	ATAGTAAG
orf75	61	54	ATCGCAAT
orf75	19	12	ATCGCAAC
orf75	135	128	ATTATAAG
odv-e25	96	89	ATTGCGAA
orf5	185	178	TTTGCATG
orf5	181	174	CATGCAAG
orf5	160	153	ATTGCGAT
orf5	7	0	CTTGTAAG
lef-2	115	108	ATTGTAAT
pk-1	153	146	TTGGCAAG
pp31	222	215	ATTGCAGG
pp31	165	158	ATTGCACG
pp31	116	109	AATACAAG
orf81	89	82	AATGCAAT
orf81	36	29	ATTTCAAA
orf81	7	0	ATAGTAAG
orf82	200	193	ATTTCATG
orf82	75	68	ATTTCAAT
vp39	203	196	CTTGCGAG
vp39	72	65	ATTTCAAT
vp39	7	0	CTTGTAAG
vp39	145	138	ATTGCAAG
vp39	223	216	CTTGTAAG
vp39	97	90	ATTGCAAG
orf93	215	208	AGTGCATG
orf93	110	103	GTTGCAAG
bv/odv-c42	150	143	GTCGCAAG
bv/odv-c42	72	65	GTTGCAAA
bv/odv-c42	159	152	GTTGCAAA
bv/odv-c42	27	20	GATGCAAG
orf102	71	64	ATTGAAAT
gp16	170	163	ATAGCAAC
gp16	124	117	GTTGCAAG
pp34	147	140	CTTGCAAA
pp34	103	96	ATGGCAAA
orf132	142	135	TTTGCTAG
alk-exo	52	45	AATGCAAT
alk-exo	46	39	ATTGGAAC
49k	155	148	AATGCAAT

Table C.4: Position and sequence of putative downstream octamer matches in relation to TAAG motif.

ORF	Start	End	Sequence
p6.9	99	92	ATTAGGAA
p6.9	47	40	ATTTGGGA
p6.9	8	1	ATTAATAA
p10	105	98	ATTCAGAA
p10	72	65	ACTATGAA
p10	64	57	ATTATGCA
odv-ec27	170	163	AGTAGTAA
odv-ec27	37	30	TTTATGAA
orf73	141	134	TTTAGAAA
orf73	100	93	ATAAGGAC
orf74	208	201	AGTAGAAA
orf74	64	57	TTTAGCAA
orf74	8	1	AATAGTAA
orf75	203	196	AAAAGGAA
orf75	132	125	ATAAGTAA
orf75	129	122	AGTAAGAA
odv-e25	160	153	ATTATGTA
odv-e25	119	112	GTTCGGAA
odv-e25	96	89	ATTGCGAA
odv-e25	21	14	ATTGGGAA
lef-2	102	95	TTTACGAA
pp31	141	134	ATTCGGAC
orf81	209	202	ATTATCAA
orf81	8	1	AATAGTAA
orf82	90	83	TTTAAGAA
vp39	182	175	GTTTGGAA
vp39	130	123	AATAGGTA
vp39	64	57	ATTAGGAA
vp39	134	127	GTTTGGAA
vp39	82	75	AATAGGTA
vp39	16	9	ATTAGGAA
orf93	8	1	ATAAGTAA
orf102	147	140	AGTTGGAA
orf102	115	108	ATGAGCAA
orf102	49	42	ATTTGTAA
gp16	183	176	ATTTTGAA
gp16	132	125	ATTAACAA
pp34	60	53	ATTAACAA
alk-exo	143	136	ATCAAGAA
alk-exo	136	129	ACTAAGAA
alk-exo	92	85	GTTGGGAA
alk-exo	47	40	AATTGGAA
alk-exo	22	15	ATTAGGTC
alk-exo	9	2	ATTTGGTA
49k	210	203	ATTAATAA
49k	192	185	GTTATGAA
49k	60	53	ATGTGGAA

C.2 Supplementary figures

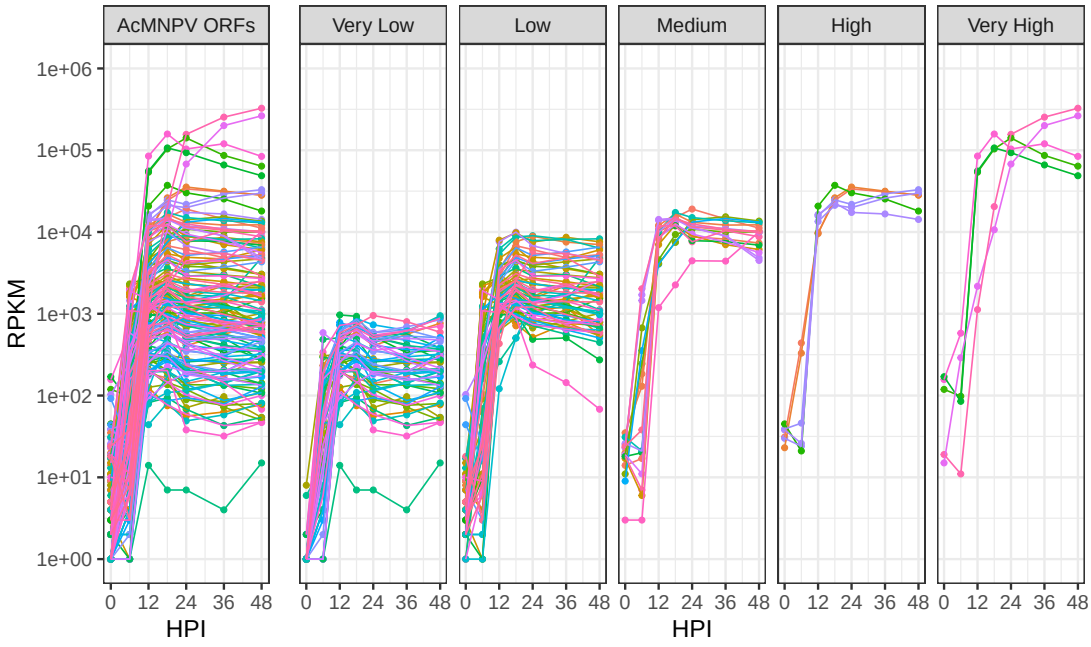


Figure C.1: Each AcMNPV ORF was categorized according to transcript abundance.

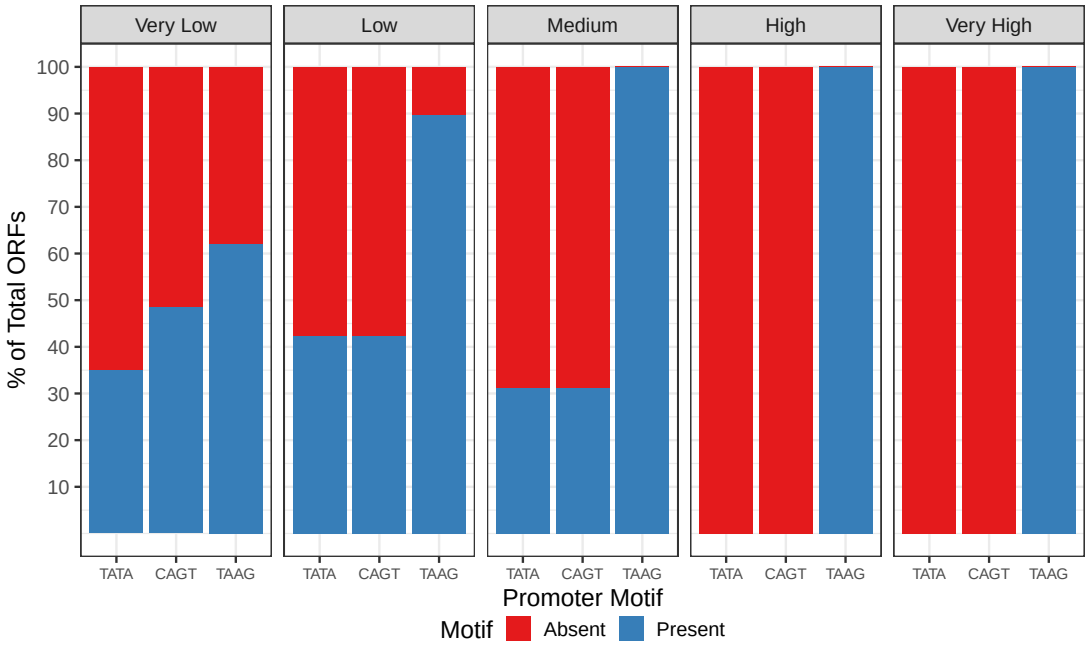
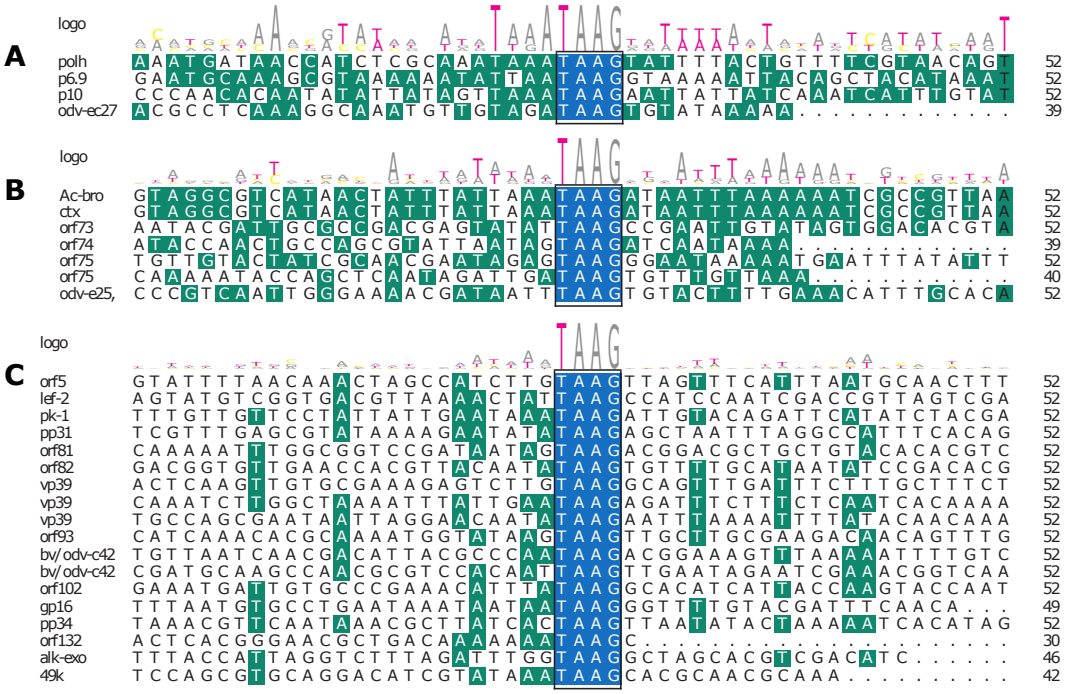


Figure C.2: **Proportion of AcMNPV ORFs with different promoter motifs, categorized according to transcript abundance.** TATA and CAGT motifs are recognized and transcribed by the host RNAP II whereas the TAAG motif is recognized and transcribed by the viral RNAP.


A
logo
polh AAATGATAACCATCTCGCAAATAAATAAGTATTTTACTGTTTTCGTAACAGT 52
p6.9 GAATGCAAAGCGTAAAAAATATTAATAAGGTAAAAATTACAGCTACATAAAT 52
p10 CCCAACACAATATATTATAGTTAAATAAGAATTATTATCAAATCATTTGTAT 52
odv-ec27 ACGCCTCAAAGGCAAATGTTGTAGATAAGTGTATAAAAA. 39
B
logo
Ac-bro GTAGGCGTCATAACTATTTATTAAATAAGATAATTTAAAAAAATCGCCGTTAA 52
ctx GTAGGCGTCATAACTATTTATTAAATAAGATAATTTAAAAAAATCGCCGTTAA 52
orf73 AATACGATTGCGCCGACGAGTATATTAAGCCGAATTGTATAGTGGACACGTA 52
orf74 ATACCAACTGCCAGCGTATTAATAGTAAGATCAATAAAA. 39
orf75 TGTTGTACTATCGCAACGAATAGAGTAAGGGAATAAAAATGAATTTATATTT 52
orf75 CAAAAATACCAGCTCAATAGATTGATAAGTGTTTGTTAAA. 40
odv-e25, CCCGTCAATTGGGAAAACGATAATTTAAGTGTACTTTGAAACATTTGCACA 52
C
logo
orf5 GTATTTTAACAAACTAGCCATCTTGTAAGTTAGTTTCATTTAATGCAACTTT 52
lef-2 AGTATGTCGGTGACGTTAAAACTATTAAGCCATCCAATCGACCGTTAGTCGA 52
pk-1 TTTGTTGTTCCTATTATTGAATAAATAAGATTGTACAGATTCATATCTACGA 52
pp31 TCGTTTGAGCGTATAAAAGAATATATAAGAGCTAATTTAGGCCATTTCACAG 52
orf81 CAAAAATTTGGCGGTCCGATAATAGTAAGACGGACGCTGCTGTACACACGTC 52
orf82 GACGGTGTTGAACCACGTTACAATATAAGTGTTTGCATAATATCCGACACG 52
vp39 ACTCAAGTTGTGCGAAAGAGTCTTGTAAGGCAGTTTGATTTCTTTGCTTTCT 52
vp39 CAAATCTTGGCTAAAATTTATTGAATAAGAGATTCTTTCTCAATCACAAAA 52
vp39 TGCCAGCGAATAATTTAGGAACAATATAAGAATTTAAAATTTTATACAACAAA 52
orf93 CATCAAACACGCAAAATGGTATAAGTAAGTTGCTTGCGAAGACACAGTTTG 52
bv/odv-c42 TGTTAATCAACGACATTACGCCCAATAAGACGGAAAGTTAAAATTTTGTC 52
bv/odv-c42 CGATGCAAGCCAACGCGTCCACAATTAAGTTGAATAGAATCGAAACGGTCAA 52
orf102 GAAATGATTGTGCCCGAAACATTTATAAGGCACATCATTACCAAGTACCAAT 52
gp16 TTTAATGTGCCTGAATAAATAATAATAAGGGTTTGTACGATTTCAACA. . . 49
pp34 TAAACGTTCAATAAACGCTTATCACTAAGTTAATATACTAAAAATCACATAG 52
orf132 ACTCACGGGAACGCTGACAAAAAATAAGC. 30
alk-exo TTTACCATTAGGTCTTTAGATTTGGTAAGGCTAGCACGTCGACATC. 46
49k TCCAGCGTGCAGGACATCGTATAAATAAGCACGCAACGCAAA. 42

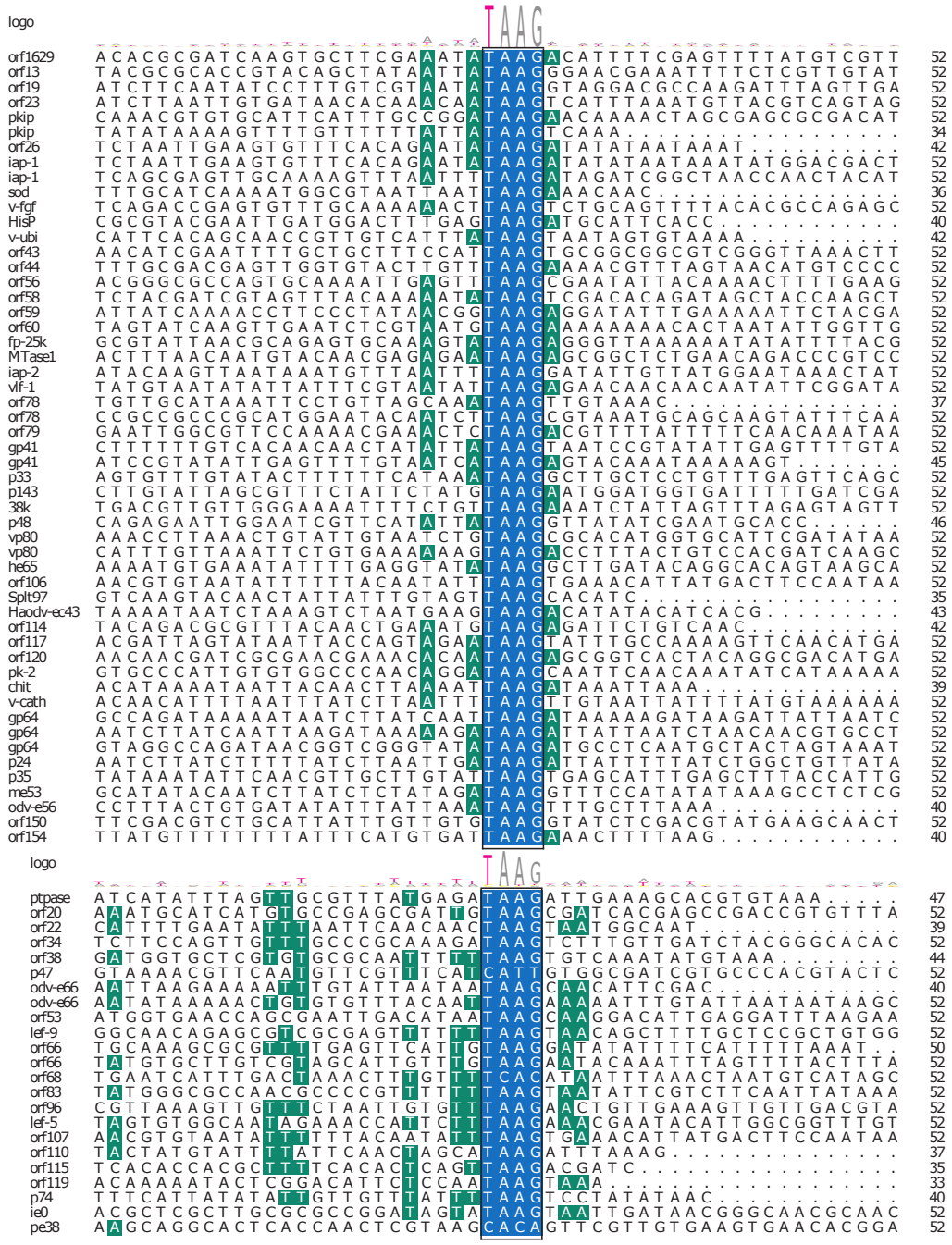
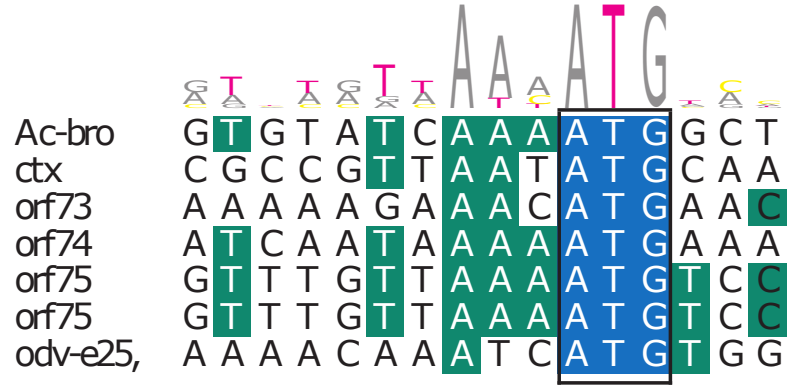


Figure C.3: Sequences flanking the late gene promoter motif for **A.** Very High, **B.** High, **C.** Medium, **D.** Low, and **E.** Very Low classes.

A

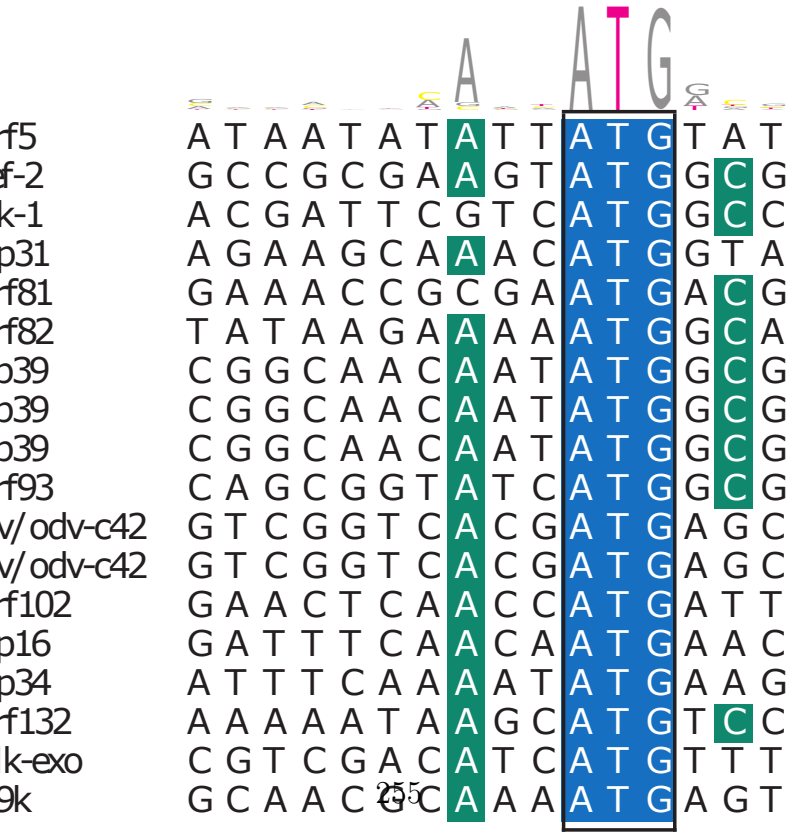
polh	A C C T A T A A A T A T G C C G	16
p6.9	C A A T T T A A A C A T G G T T	16
p10	A T T T A C A A T C A T G T C A	16
odv-ec27	T G T A T A A A A A T G A A A	16

B



Ac-bro	G T G T A T C A A A A T G G C T	16
ctx	C G C C G T T A A T A T G C A A	16
orf73	A A A A A G A A A C A T G A A C	16
orf74	A T C A A T A A A A A T G A A A	16
orf75	G T T T G T T A A A A T G T C C	16
orf75	G T T T G T T A A A A T G T C C	16
odv-e25,	A A A A C A A A T C A T G T G G	16

C



orf5	A T A A T A T A T T A T G T A T	16
lef-2	G C C G C G A A G T A T G G C G	16
pk-1	A C G A T T C G T C A T G G C C	16
pp31	A G A A G C A A A C A T G G T A	16
orf81	G A A A C C G C G A A T G A C G	16
orf82	T A T A A G A A A A A T G G C A	16
vp39	C G G C A A C A A T A T G G C G	16
vp39	C G G C A A C A A T A T G G C G	16
vp39	C G G C A A C A A T A T G G C G	16
orf93	C A G C G G T A T C A T G G C G	16
bv/odv-c42	G T C G G T C A C G A T G A G C	16
bv/odv-c42	G T C G G T C A C G A T G A G C	16
orf102	G A A C T C A A C C A T G A T T	16
gp16	G A T T T C A A C A A T G A A C	16
pp34	A T T T C A A A A T A T G A A G	16
orf132	A A A A A T A A G C A T G T C C	16
alk-exo	C G T C G A C A T C A T G T T T	16
49k	G C A A C G C A A A A T G A G T	16

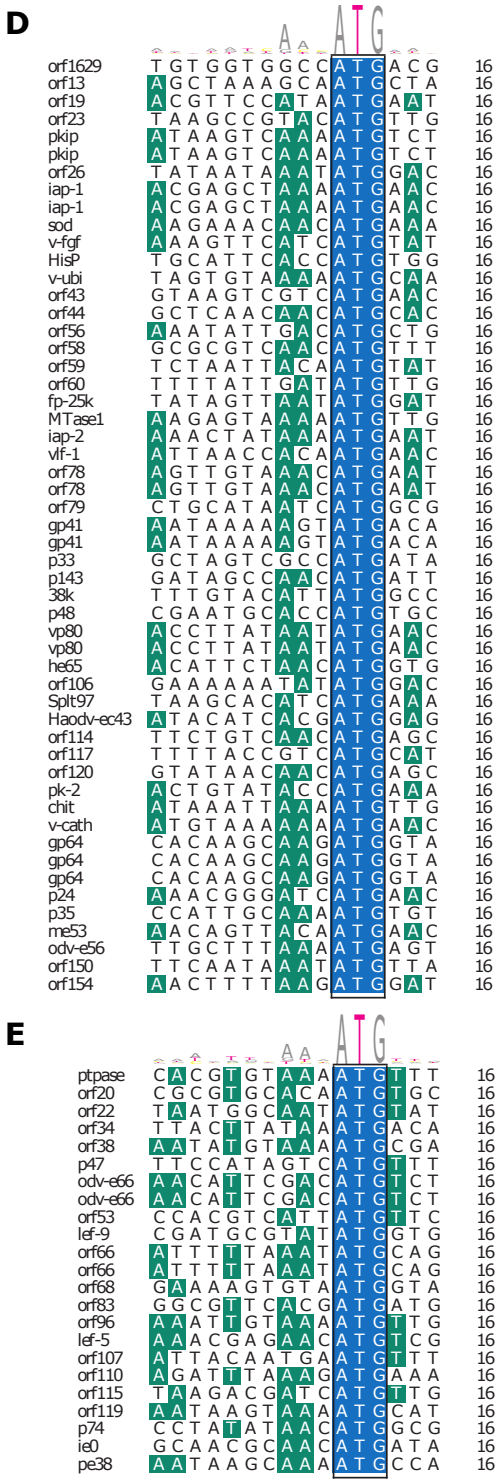


Figure C.4: Sequences flanking the translation initiation site for **A.** Very High, **B.** High, **C.** Medium, **D.** Low, and **E.** Very Low classes.