

A CRISPR-Cas9 and next-generation sequencing approach for late/very late AcMNPV gene disruption and comprehensive mutation analysis

by

Madhuja Chakraborty

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, 2025

© Madhuja Chakraborty 2025

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: Antonio Roldao
Head of Cell-based Vaccines Development Lab,
Animal Cell Technology Unit,
Instituto de Biología Experimental e Tecnológica

Supervisor: Marc G. Aucoin
Professor, Dept. of Chemical Engineering,
University of Waterloo

Internal Member: C. Perry Chou
Professor, Dept. of Chemical Engineering,
University of Waterloo

Internal Member: Christine Moresoli
Professor, Dept. of Chemical Engineering,
University of Waterloo

Internal-External Member: Andrew C. Doxey
Associate Professor, Dept. of Biology,
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 3, Madhuja Chakraborty and Marc G. Aucoin were involved in conceptualization. Investigation, formal analysis, and data curation were performed by Madhuja Chakraborty. Madhuja Chakraborty and Mark Bruder were involved in methodology. Visualization was done by Madhuja Chakraborty and Marc G. Aucoin. Madhuja Chakraborty wrote the original draft. Madhuja Chakraborty and Marc G. Aucoin were involved in the manuscript review and editing.

For Chapter 4, Madhuja Chakraborty, Lisa Nielsen, Delaney Nash, and Marc G. Aucoin were involved in conceptualization and methodology. Formal analysis and data curation were performed by Madhuja Chakraborty, Delaney Nash, and Marc G. Aucoin. Investigation was performed by Madhuja Chakraborty, Lisa Nielsen, and Marc G. Aucoin. Visualization was done by Madhuja Chakraborty and Marc G. Aucoin. Madhuja Chakraborty wrote the original draft. Madhuja Chakraborty, Lisa Nielsen, Delaney Nash, Jozef I. Nissimov, Trevor C. Charles, and Marc G. Aucoin were involved in the manuscript review and editing.

For Chapter 5, Madhuja Chakraborty, Lisa Nielsen, Delaney Nash, and Marc G. Aucoin were involved in conceptualization. Formal analysis and data curation were performed by Madhuja Chakraborty, Delaney Nash, and Marc G. Aucoin. Investigation was performed by Madhuja Chakraborty, Lisa Nielsen, and Marc G. Aucoin. Madhuja Chakraborty, Lisa Nielsen, Delaney Nash, Mark Bruder, and Marc G. Aucoin were involved in methodology. Visualization was done by Madhuja Chakraborty and Marc G. Aucoin. Madhuja Chakraborty wrote the original draft. Madhuja Chakraborty, Lisa Nielsen, Delaney Nash, Jozef I. Nissimov, Trevor C. Charles, and Marc G. Aucoin were involved in the manuscript review and editing.

Abstract

The recent global pandemic COVID-19 has taught us the importance of an efficient biologics manufacturing platform that is cost-effective, reliable, and has high product yield and quality. The baculovirus expression vector system (BEVS) has proven to be a promising platform for the production of recombinant proteins, vaccines, virus-like particles (VLPs), viral vectors, and/or other biologics. In the last two decades, many vaccines and therapeutics manufactured using BEVS have received licenses for animal and human use. The majority of the commercially available BEVS transfer plasmids have foreign genes under the viral polyhedrin (*polh*) or *p10* promoters. Although high gene expression can be achieved with the endogenous baculovirus promoters *p10* and *polh*, they are only active very late in the infection cycle when most of the host cellular machinery is turned off. Significant work has been done to identify native promoters and other regulatory elements with expression profiles higher than *polh*, as well as promoters weaker than *polh* to express secretory proteins that require extensive post-translational modifications (PTMs).

Certain regions, such as *polh*, *chiA*, and *v-cath*, in the baculovirus genome are not essential for their *in vitro* replication or foreign protein production in cell culture. Thus, it is possible that if there is an expression of genes not required for progeny virus and/or exogenous protein production in insect cell culture, the resources that are being used for their expression could be ‘an additional burden’, resulting in unnecessary depletion of cellular resources. Identifying and removing these genes would probably divert resources towards the production of foreign proteins and progeny viruses, which could improve the BEVS production platform. Moreover, it was previously demonstrated that there is a ‘competition effect’ among protein-coding genes for cellular resources when Sf9 cells are either coinfecting with two monocistronic recombinant baculovirus expression vectors (rBEVs) or infected with a dual-protein producing polycistronic rBEV. This work could point to a direction where competition can arise among baculovirus genes for the use of cellular resources, and the knockout of unnecessary genes could presumably lead to appropriate usage of the resources by the essential genes. Separately, the co-production of rBEVs and recombinant protein products in the supernatant complicates the downstream purification process. Disruption of genes essential for virion formation or production could prevent baculovirus contamination in the culture supernatant, thus reducing the burden on purification processes.

In the past, gene disruption or downregulation has been a fruitful strategy to improve the expression of foreign genes in the BEVS. However, the traditional methods used for mutant baculovirus genome generation are time-consuming, labor-intensive, and sometimes also produce wild-type viruses, hence requiring additional purification steps. Not until

recently has CRISPR-Cas9 gene editing technology been adapted to the Sf9 insect cells and the baculovirus *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV). It is believed to be an effective tool to scrutinize baculovirus genes by targeted gene disruption and transcription repression. A systematic study of the late and very late AcMNPV genes using a CRISPR-Cas9-based transfection-infection assay (T-I assay), disrupting the unnecessary sequences, and expressing exogenous gene(s) under a late promoter instead of very late promoters could extend the production time and improve biologics production. Moreover, in the final production stage, targeting AcMNPV genes that are required for progeny virus assembly or release but do not affect foreign protein production could minimize rBEV co-production. In this study, the T-I assay was used to probe late and very late AcMNPV genes for their essentiality. Based on the effect of individual gene disruptions on foreign protein (green fluorescent protein (GFP)) and budded virus (BV) production, 38 targeted AcMNPV genes were categorized as essential (reduced both GFP and BV production) and of special interest (reduced GFP production but did not lower BV production). While we identified 19 AcMNPV genes that are essential for BV production and GFP expression from the late p6.9 promoter, 19 other genes were identified as of special interest whose disruption only reduced GFP expression from the late p6.9 promoter.

While phenotypic changes were assessed using the CRISPR-Cas9-based T-I assay, investigating the genomes using whole-genome next-generation sequencing (NGS) revealed further information. First of all, shotgun sequencing was used to generate a consensus sequence of the p6.9GFP rBEV stock used in T-I assays, and this is the first report on whole-genome rBEV sequences to the best of our knowledge. This shotgun-sequenced rBEV served as the reference genome to identify mutations upon CRISPR-Cas9-mediated gene disruptions. We also provided a set of tiled-amplicon primers based on the reference genome and adapted a high-throughput tiled-amplicon sequencing assay to control and targeted rBEV genomes. This sequencing assay, combined with a bioinformatics pipeline for major species, was able to successfully detect mutations within the *gp64* gene when *gp64* targeting sgRNA was delivered to Sf9-Cas9 cells via a plasmid or rBEV. We further demonstrated that *gp64* disruption lowered BV levels without decreasing GFP production, thus reducing BV contamination in cell culture supernatant.

To probe the *gp64* gene further, we targeted it at six different locations using the T-I assay. Plasmids carrying one or two sgRNA targets were used to evaluate the impact of single and multiple targeting sites on virion and foreign protein production. *gp64* disruption with each of these sgRNA targets resulted in decreased infectious and total viral titers, whereas GFP production from the late p6.9 promoter was enhanced or remained similar to the control. Low-frequency genomic changes upon CRISPR-Cas9-mediated *gp64* disruptions were successfully assessed by the tiled-amplicon sequencing assay and a variant

calling pipeline based on the computational tool iVar. While the iVar tool was originally developed to investigate variants in wild-type virus populations, we adapted it to detect variants in a process system. We also demonstrated that variants can be preserved over viral propagation in cell culture, that is, variants present in the virus stock were also observed in the rBEV genomes recovered from the T-I assay, thus indicating that they are not detrimental to viral fitness.

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Prof. Marc Gordon Aucoin, for welcoming me into his research group and providing opportunities that enriched both my learning and teaching experiences. Thank you for your unwavering faith in me and for being the kind of mentor anyone would hope for. I am deeply grateful for your support, insightful critiques, constructive feedback, and the many discussions that have shaped my personal and professional growth.

I am grateful to everyone in the Aucoin Lab for the valuable insights and meaningful discussions we have shared over time. Special thanks to Dr. Mark Bruder for providing the virus vectors and four plasmids used in this thesis, and to Dr. Sara HK for training me in molecular work. I also thank Matt Patton and Alex Slavcev for their help with some of the molecular experiments. I am grateful to Prof. Trevor C. Charles for opening his lab space to us, and to our collaborators from the Charles Lab – Delaney, Alyssa, Samran, and Shiv – for their assistance with the sequencing work. My sincere thanks to Prof. C. Perry Chou, Prof. Christine Moresoli, and Prof. Andrew C. Doxey for being a part of my PhD committee, and Prof. Antonio Roldao for acting as my external examiner.

I would like to thank the Department of Chemical Engineering at the University of Waterloo and NSERC for funding this work. I am also grateful to the administrative and technical support staff of the Department of Chemical Engineering – Judy, Erene, Sarah, Nikki, Bert, and Rick – for their support throughout my time here.

To all my friends back home, thank you for always believing in me. Ishita and Ina – you two have been my unwavering support system since the day our paths crossed. To my labmates who became the friends I never knew I needed – Jackie, Sara HK, Chris S, Lisa, Reza, Chris, Scott, Faranak, CT, and Ken – thank you for standing by me, putting up with me, and for all the laughter and deep conversations along the way. To my best friend and partner, Prafful – I cannot thank you enough for being my support system. I genuinely do not know what I would have done without you here.

Lastly, I owe everything to my parents and my Bibi (granny). Thank you for your boundless support, for encouraging me in all my decisions, for trusting in my potential, and for loving me unconditionally. I am who I am because of you.

This work was carried out on the Haldimand Tract, the land granted to the Six Nations that includes six miles on each side of the Grand River.

Dedication

This is dedicated to my parents, Sanchita and Manash, and my grandmother, Ashoka.

Table of Contents

Examining Committee Membership	ii
Author's Declaration	iii
Statement of Contributions	iv
Abstract	v
Acknowledgements	viii
Dedication	ix
List of Figures	xv
List of Tables	xvii
List of Abbreviations	xix
1 Introduction	1
1.1 Hypotheses	4
1.2 Objectives	6
1.3 Thesis Outline	7

2	Literature Review	9
2.1	Baculovirus	9
2.1.1	Baculovirus Genome	10
2.1.2	Baculovirus Gene Transcription and Infection Cycle	11
2.2	Insect Cell Culture	13
2.3	BEVS and Its Advantages	15
2.4	Baculovirus Vector Development for Biologics Production	16
2.4.1	Generation of Recombinant Vectors	16
2.4.2	Genome Engineering of Recombinant Vectors	19
2.4.3	Recombinant Vector Infection Strategies	22
2.5	Production Platform for Biologics	25
2.5.1	Single and Multi-Subunit Recombinant Proteins	27
2.5.2	Gene Delivery in Mammalian Cells	28
2.5.3	Viral Vectors for Gene Therapy	30
2.5.4	Baculovirus Surface Display	31
2.6	CRISPR-Cas9 Genome Editing Tool	32
2.7	Next-Generation Sequencing	38
3	Probing late and very late AcMNPV genes using CRISPR-Cas9	42
3.1	Abstract	45
3.2	Introduction	46
3.3	Materials and Methods	49
3.3.1	Insect Cell Lines and Maintenance	49
3.3.2	sgRNA Plasmid Design and Construction	49
3.3.3	Recombinant Baculovirus Expression Vector Amplification	51
3.3.4	Quantification of Infectious Recombinant Virus Vector	51
3.3.5	CRISPR-Cas9-Based Transfection-Infection Assay	52
3.3.6	Flow Cytometry Analysis for Foreign Protein Production	53

3.3.7	Statistical Analysis	53
3.4	Results	54
3.4.1	Investigation of Controls for T-I Assay	54
3.4.2	CRISPR-Cas9-Mediated AcMNPV Gene Disruption	55
3.5	Discussion	61
3.5.1	Auxiliary AcMNPV Genes	63
3.5.2	DNA Replication AcMNPV Genes	67
3.5.3	Transcription AcMNPV Genes	69
3.5.4	Structural AcMNPV Genes	72
3.5.5	Unknown Function AcMNPV Genes	74
3.6	Conclusions	75
4	Adapting next-generation sequencing to recombinant baculovirus vectors	77
4.1	Abstract	79
4.2	Introduction	80
4.3	Materials and Methods	84
4.3.1	Cell Line and Maintenance	84
4.3.2	Baculovirus Amplification and Quantification	84
4.3.3	Plasmid Design and Construction	85
4.3.4	Shotgun Sequencing of the p6.9GFP rBEV	86
4.3.5	Transfection in Multi-Well Plates	87
4.3.6	Infection in Multi-Well Plates	88
4.3.7	Harvesting from Multi-Well Plates	88
4.3.8	Flow Cytometry Analysis	89
4.3.9	Tiled-Amplicon Sequencing Assay for rBEV Genomes	89
4.3.10	Bioinformatics Pipeline for Major Species	94
4.4	Results	97

4.4.1	Consensus Sequence of p6.9GFP rBEV from Shotgun Sequencing	97
4.4.2	Tiled-Amplicon Sequencing of p6.9GFP rBEV	105
4.4.3	Lowering Infectious Baculovirus from Sf9-Cas9 Cells Through Co-Expressing a sgRNA Targeting <i>gp64</i>	105
4.4.4	Tiled-Amplicon Sequencing of p6.9GFP rBEV upon T-I Assay	108
4.4.5	Lowering Infectious Budded Virus from Sf9-Cas9 Cells Through Plasmid-Based Delivery of sgRNA Targeting <i>gp64</i>	109
4.5	Discussion	111
4.5.1	Differences Compared to Reported Genomes from Shotgun Sequencing	113
4.5.2	Tiled-Amplicon Sequencing to Assess CRISPR-Cas9 Editing	114
4.5.3	Concluding Remarks	119
5	Detection of variants upon transient CRISPR-Cas9 targeting of <i>gp64</i>	120
5.1	Abstract	122
5.2	Introduction	123
5.3	Results	125
5.3.1	Effect of <i>gp64</i> Gene Disruption on Foreign Protein and Progeny Virus Production	125
5.3.2	Confirmation of CRISPR-Cas9-Mediated Gene Editing of <i>gp64</i>	127
5.3.3	Variant Generation or Conservation over Passages	131
5.3.4	Evaluation of Mutations Outside the Targeted <i>gp64</i> Gene	132
5.4	Discussion	135
5.4.1	CRISPR-Cas9-Mediated Targeted Disruption of <i>gp64</i> During Virus Propagation	137
5.4.2	Are CRISPR-Cas9 Off-Targets Observed in Our System?	139
5.4.3	Variant Conservation or Random Mutations Upon Virus Propagation in Cell Culture	140
5.5	Materials and Methods	142
5.5.1	Cell Line and Maintenance	142
5.5.2	Plasmid Design and Construction	143

5.5.3	Baculovirus Amplification and Quantification	144
5.5.4	Transfection-Infection Assay (T-I Assay)	145
5.5.5	Flow Cytometry Analysis of GFP upon <i>gp64</i> Gene Disruption	146
5.5.6	Total Baculovirus Quantification via Flow Cytometry	147
5.5.7	Tiled-Amplicon Sequencing Assay for rBEV Genomes	148
5.5.8	Bioinformatics Pipeline for Minor Species	149
5.6	Conclusions	150
6	Conclusions and Recommendations	151
	References	156
	APPENDIX	197
A	Chapter 3 supplementary	198
B	Chapter 4 supplementary	209
C	Chapter 5 supplementary	224

List of Figures

2.1	Baculovirus virion phenotypes.	10
2.2	CRISPR-Cas9 as a genome engineering tool.	36
2.3	Tiled-amplicon sequencing assay workflow.	41
3.1	Overview of transfection-infection assay.	44
3.2	Impact of disrupting auxiliary AcMNPV genes on (a) foreign protein and (b) progeny virus production.	56
3.3	Effect of targeting DNA replication AcMNPV genes on (a) GFP and (b) IVT.	57
3.4	Effect of transcription AcMNPV gene disruptions on (a) foreign protein and (b) BV production.	58
3.5	Impact of structural AcMNPV gene disruptions on (a) GFP and (b) IVT.	59
3.6	Effect of targeting unknown function AcMNPV genes on (a) foreign protein and (b) progeny virus production.	60
4.1	<i>in Process</i> CRISPR-Cas9 genome engineering of rBEVs.	78
4.2	Odd-numbered forward (L) and reverse (R) primers are combined to make Primer Pool 1 and even-numbered forward (L) and reverse (R) primers are combined to make Primer Pool 2.	91
4.3	Parameters used for the four Trimmomatic operations. The * here indicates a mandatory field.	96
4.4	Impact of sgRNA rBEV-mediated <i>gp64</i> gene disruption on (a) foreign protein production and (b) infectious virus titer.	106

4.5	Effect on (a) foreign protein production and (b) infectious virus titer upon CRISPR-Cas9 based T-I assay.	110
5.1	Effect of CRISPR-Cas9-mediated gene disruption of the AcMNPV <i>gp64</i> gene at single or dual locations on GFP production.	126
5.2	Impact of CRISPR-Cas9 targeting of the AcMNPV <i>gp64</i> gene using single or dual sgRNA targets on (a) infectious virus titer (IVT) and (b) total viral particles.	127
5.3	(a) Total and (b) filtered (excluding controls and conserved variants) mutation distribution in the different hrs of gDNA pools are highlighted here.	134
5.4	Overview of mutations in the pool of gDNAs outside the hrs and the CRISPR-Cas9 targeted <i>gp64</i> gene.	135
A.1	Flow cytometry profiles of controls for T-I assay	206
A.2	Ranked effect	207
A.3	Change in GFP or BV with respect to targeting efficiency	208
B.1	Visual representation of sequence alignments (Benchling).	222

List of Tables

3.1	Classification of AcMNPV gene disruptions by their effects on GFP (rows) and BV (columns). Each gene appears on a separate line within its cell. . .	63
4.1	Multiplex PCR setup.	92
4.2	Multiplex PCR amplification parameters.	92
4.3	Mutations from shotgun sequencing of p6.9GFP rBEV.	99
4.4	Changes observed in the gDNA of p6.9GFP_sgRNA_gp64+131 rBEV upon propagation in Sf9-Cas9 cells via tiled-amplicon sequencing assay.	108
4.5	Changes in the gDNA of p6.9GFP rBEV obtained from tiled-amplicon sequencing upon T-I assay.	109
4.6	Mutations within the <i>gp64</i> gene from tiled-amplicon sequencing of p6.9GFP rBEV upon <i>gp64</i> gene disruption.	111
5.1	Overview of true deletion mutations and corresponding variant frequencies at different locations within the <i>gp64</i> gene.	129
5.2	Conserved variants found in different samples outside the homologous repeat regions (hrs).	132
5.3	Spacer sequences used for control and AcMNPV <i>gp64</i> gene.	144
A.1	Spacer sequences used in this study for the control and AcMNPV genes. . .	199
B.1	Primers used to construct the scrambled control plasmid.	210
B.2	Positions of the tiled-amplicon primers in the p6.9GFP genome (shotgun-sequenced rBEV) as generated by Primal Scheme (Quick et al., 2017) and used in this study.	211

B.3	Tiled-amplicon primers generated by Primal Scheme (Quick et al., 2017) and used in this study.	216
B.4	Custom primers with unique i7 and i5 index pairs for each sample, used to construct MiSeq DNA libraries.	221
B.5	Mutations within the <i>gp64</i> gene from tiled-amplicon sequencing of the rBEVs p6.9GFP_sgRNA_gp64+131 and p6.9GFP upon <i>gp64</i> gene disruption. . . .	223
C.1	Custom primers (unique indexes with sequences appended to their 5' and 3' ends) used in this study to construct MiSeq DNA libraries.	225
C.2	Description of the column headers appearing in each .tsv output file obtained from running the variant calling pipeline with a reference genome and two technical replicates.	231

Abbreviations

<i>AcMNPV</i>	<i>Autographa californica</i> multiple nucleopolyhedrovirus
<i>BEVS</i>	baculovirus expression vector system
<i>BLT</i>	bead-linked transposomes
<i>BV</i>	budded virus
<i>BmNPV</i>	<i>Bombyx mori</i> nucleopolyhedrovirus
<i>CHO</i>	chinese hamster ovary cells
<i>DIPs</i>	defective interfering particles
<i>DNA</i>	deoxyribonucleic acid
<i>EPDA</i>	end point dilution assay
<i>EV</i>	extracellular virus
<i>EV71</i>	enterovirus 71
<i>FBS</i>	fetal bovine serum
<i>FDA</i>	food and drug administration
<i>GFP</i>	green fluorescent protein
<i>GV</i>	granuloviruses
<i>Gag – GFP</i>	retroviral Gag genetically tagged with GFP
<i>HDR</i>	homology-directed repair

<i>HEK293</i>	human embryonic kidney cells
<i>HIV</i> – 1	human immunodeficiency virus type 1
<i>HSC70</i>	heat shock cognate 70
<i>HSP70</i>	heat shock protein 70
<i>HSR</i>	heat shock response
<i>HT1</i>	hybridization buffer
<i>IAPs</i>	inhibitor of apoptosis proteins
<i>IDT</i>	integrated DNA technologies
<i>IPM</i>	integrated pest management
<i>IVT</i>	infectious virus titer
<i>KOVs</i>	knock-out viruses
<i>MDCK</i>	madin darby canine kidney cells
<i>MOI</i>	multiplicity of infection
<i>NGS</i>	next-generation sequencing
<i>NHEJ</i>	non-homologous end joining
<i>NOV</i>	non-occluded virus
<i>NPVs</i>	nucleopolyhedroviruses
<i>OB</i>	occluded bodies
<i>ORFs</i>	open reading frames
<i>PAM</i>	protospacer adjacent motif
<i>PBS</i>	phosphate buffered saline
<i>PIBs</i>	polyhedral inclusion bodies
<i>RNA</i>	ribonucleic acid

<i>SFM</i>	serum-free media
<i>SNPs</i>	single nucleotide polymorphisms
<i>Sf21</i>	<i>Spodoptera frugiperda</i> clonal isolate 21
<i>Sf9</i>	<i>Spodoptera frugiperda</i> clonal isolate 9
<i>T – I assay</i>	transfection-infection assay
<i>TCID50</i>	tissue culture infectious dose 50
<i>VLPs</i>	virus-like particles
<i>Vero</i>	african green monkey derived kidney cells
<i>WP10</i>	wild population 2010
<i>polh</i>	poyhedrin gene
<i>au</i>	arbitrary units
<i>eYFP</i>	enhanced yellow fluorescent protein
<i>hpi</i>	hours post infection
<i>hpt</i>	hours post-transfection
<i>hrs</i>	homologous repeat regions
<i>mAG</i>	monomeric Azami Green
<i>ori</i>	origin of replication
<i>pfu/mL</i>	plaque forming unit per mL
<i>rAAV</i>	recombinant adeno-associated virus vector
<i>rBEVs</i>	recombinant baculovirus expression vectors
<i>rpm</i>	revolutions per minute
<i>sgRNA</i>	single guide RNA

Chapter 1

Introduction

With the advantages of being more robust, faster growing, and requiring less strict growth conditions compared to eukaryotic cells, prokaryotic hosts, such as bacteria, are used for the large-scale production of recombinant proteins ([Butler, 2005](#)). However, while smaller proteins can be easily produced in prokaryotic cells, larger complex proteins require eukaryotic systems ([Demain, Vaishnav, 2009](#)). Moreover, with the enhancement of recombinant protein technology in the 1970s and 1980s, and the increasing demand for more complex biotherapeutics that require post-translational modifications (PTMs), eukaryotic hosts offer a more reliable and robust production platform ([Butler, 2005](#)).

Eukaryotic systems such as fungi, baculovirus expression vector system (BEVS), and mammalian cells are usually chosen for the production of recombinant glycosylated proteins. With high protein yields at low cost and the ability to perform glycosylation, fungi can be used for the production of proteins larger than 50 kDa ([Demain, Vaishnav, 2009](#)). However, amongst the different eukaryotic hosts, mammalian cells, such as

the Chinese hamster ovary (CHO) and human embryonic kidney (HEK293) cells, are predominantly used for biopharmaceutical production due to their capability for PTMs and human protein-like molecular structure assembly (Demain, Vaishnav, 2009; Zhu, 2012). The BEVS, though, is a middle ground between the different eukaryotic hosts, with its ability to carry out complex PTMs, proper protein folding, and high protein yields (Demain, Vaishnav, 2009).

The 1850s marked the earliest mention of the non-mammalian baculovirus in the literature (Summers, 2006). This was followed by research focused on polyhedra purification from infected insects, infectious virus recovery from polyhedra, and infection mechanism in insect midgut (Summers, 2006). The baculovirus has also been applied as an insecticide as a part of the integrated pest management (IPM) strategy (Haase et al., 2015), and the first baculoviral insecticide found its way to the market during the 1970s (Black et al., 1997). Later in 1983, the use of the baculovirus *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) for recombinant protein production in insect cell culture was first established (Smith et al., 1983b), thus marking the birth of the baculovirus expression vector system (BEVS). Briefly, Smith et al. (1983b) demonstrated the expression of the human interferon-beta gene under the control of the very strong polyhedrin promoter by infecting insect cells with a recombinant AcMNPV. Once it was known that the AcMNPV produces large amounts of the Polyhedrin protein and the corresponding gene was not essential for virus propagation in insect cell cultures, it largely contributed to the development of the BEVS (van Oers et al., 2015).

Although a majority of the biotherapeutic products use mammalian cells for production, it is well known from the literature that the BEVS is a promising platform for the

production of various proteins of interest owing to its ease of use and versatility. This platform comprises of a continuous insect cell line, typically Sf9, Sf21, and High-FiveTM cell lines, and a recombinant baculovirus expression vector (rBEV) carrying the gene(s) of interest, that is capable of infecting insect cell cultures (Palomares et al., 2015). The limited host range of baculovirus and high protein yields make the system a great choice for the production of recombinant proteins (George, 2016). Additionally, the ability of insect cells to perform complex mammalian-like PTMs, including glycosylation (James et al., 1996), phosphorylation (Héricourt et al., 2015), and disulfide bond formation (Hodder et al., 1996), makes BEVS a suitable platform for the production of foreign proteins that require such modifications (Reed, Muench, 1938). Furthermore, the lack of human adventitious viruses in insect cells (Summers, 2006) and the inability of baculovirus to grow in or infect mammalian cells (Sokolenko et al., 2012) makes it appropriate for the production of therapeutic proteins.

The advantages of the BEVS have made it a desirable platform for a large number of applications. Many higher-order proteins and protein complexes, such as antibodies (Zu Putlitz et al., 1990), viral vectors (Aucoin et al., 2006), and virus-like particles (VLPs) (Pushko et al., 2005, 2010), have been produced using this system. Moreover, the BEVS has been used as a surface display technology for displaying proteins with eukaryotic PTMs (Kost et al., 2005). Another application of the BEVS is the production of vectors for gene therapy. The first-ever vector to be approved for gene therapy by the Food and Drug Administration (FDA), a recombinant adeno-associated virus (rAAV) vector, is produced in the BEVS (Palomares et al., 2015; Felberbaum, 2015).

Despite the wide range of BEVS applications and advantages, the majority of the

work done with this system is restricted to academic and/or research laboratories. This could be because — the function(s) of many AcMNPV genes are not yet experimentally confirmed (Bruder, 2021); the majority of the commercial BEVS transfer plasmids have the gene(s) of interest under the control of the very late polyhedrin (polh) or p10 promoters (George, 2016); the co-production of virions with the recombinant products complicates the downstream purification processes (Marek et al., 2011; Lee et al., 2015; Bruder, Aucoin, 2023a); and the baculovirus is known to have inherent genomic instability (Kool et al., 1991; Pijlman et al., 2001). Moreover, commercially available BEVS usually have the complete genome or only a handful of non-essential/deleterious genes removed from their genome, which leads to the possible expression of many unnecessary genes. Additionally, AcMNPV gene deletion to improve foreign protein production is not well studied (Hitchman et al., 2010a). These drawbacks of the system could contribute to the limited industrial exposure of the BEVS. Nonetheless, it is gaining some popularity for industrial production with quite a few biotherapeutics being produced in this system in the last decade (Bruder, 2021).

1.1 Hypotheses

The driving hypothesis behind this research is that the large AcMNPV genome contains genes, driven by late and very late AcMNPV promoters, that are not required for infectious budded virus (BV) production or expression of foreign genes in insect cell culture, and thereby contribute to unnecessary usage of cellular resources. A systematic study of the AcMNPV genes active in the late and very late phases of the infection cycle using CRISPR-Cas9 and the subsequent disruption and identification of unnecessary sequences

would remove genetic burden from the recombinant AcMNPV genome and direct resources towards foreign protein or progeny virus production.

More specifically, by using a genetically engineered Sf9 cell line expressing the *cas9* gene, the essentiality of AcMNPV genes can be easily probed. Even more specifically, if a late promoter, such as p6.9, is used instead of very late promoters, such as polh and p10, which are active when most of the cellular machinery is turned off, sequences that are not necessary in the late and very late infection phases can be targeted to enhance biologics production. Thus, using a minimal rBEV genome and expressing foreign gene(s) under late promoters can shift production to an earlier point in the infection cycle when the host cells are not completely hijacked (as compared to the current utilization of polh or p10 promoters), as well as enhance the production of biologics.

Another hypothesis evaluated in this study is that there exist AcMNPV genes that affect BV but not foreign protein production from the late p6.9 promoter, and the reduction of budded baculovirus contamination in culture supernatants by minimizing or preventing BV formation will simplify the purification process for bio-industrial applications. Specifically, sequences that will reduce baculovirus production without compromising foreign protein production can be targeted by CRISPR-Cas9, while eliminating the need for creating a trans-complementing cell line, such that it reduces baculovirus co-production and facilitates downstream purification.

We further hypothesized that whole-genome next-generation sequencing (NGS) can confirm the CRISPR-Cas9-mediated mutations and detect any possible off-targets. Sequencing the rBEV genome before and after targeting is also necessary to determine whether the effect is due to gene disruption or if the mutation was already present in

the virus stock. Additionally, it was postulated that transient CRISPR-Cas9 disruption of rBEV genomes, as the virus undergoes its infection cycle, would result in variants in the genome pool.

1.2 Objectives

The overall objective of this work was to probe the essentiality of late and very late AcMNPV genes for BV and foreign protein production from the late p6.9 promoter, as well as to establish NGS pipelines to sequence the virus stock(s) and confirm CRISPR-Cas9-mediated mutations. It is to be noted that the flashBACTM GOLD AcMNPV vector is used as the backbone in this work, and the baculovirus *chitinase* (*chiA*) and *cathepsin* (*v-cath*) genes are already deleted from this genomic DNA (gDNA). The objectives of this study can be broken down into the following:

- (a) screen for genes active in the late and very late phases of the infection cycle to identify essential and of special interest AcMNPV sequences through targeted gene disruption using a previously developed CRISPR-Cas9-based transfection-infection assay (T-I assay). Briefly, these late and very late sequences are selected based on the presence of a 5'-TAAG-3' promoter motif and the literature review, and 2–3 sgRNAs are assessed for the disruption of each gene target;
- (b) use shotgun sequencing to generate a consensus sequence of the virus stock, rBEV carrying a green fluorescent protein (GFP) under the late p6.9 promoter (p6.9GFP rBEV), amplified in Sf9 cells. This shotgun-sequenced rBEV acts as the reference genome for plasmid-based delivery of sgRNA to Sf9-Cas9 cells for targeted gene disruption. Adapt a

tilted-amplicon sequencing assay to rBEVs by generating tiled-amplicon primers based on the reference genome and sequencing untargeted and targeted rBEV genomes. Apply the tiled-amplicon sequencing assay to sgRNA rBEV (carrying GFP and sgRNA for AcMNPV *gp64*) amplified in Sf9 cells and rBEV-based delivery of sgRNA to Sf9-Cas9 cells for targeted gene disruption;

(c) disrupt the AcMNPV *gp64* gene at multiple locations using the CRISPR-Cas9-based T-I assay and evaluate the impact of single (one spacer sequence) and dual (two spacer sequences) sgRNA targets on BV production and late exogenous protein production. Use the tiled-amplicon sequencing assay to sequence control and *gp64* targeted genomes and generate a variant calling pipeline to confirm targeted mutations and detect any off-targets.

1.3 Thesis Outline

The first chapter of this thesis consists of a general introduction to the different biologics production platforms with a focus on the BEVS, some advantages, and applications of the BEVS, as well as the driving hypotheses and the objectives of this work. The second chapter is a relevant literature review of the research presented in this thesis. In Chapter 3, the effect of CRISPR-Cas9-mediated disruption of 38 endogenous AcMNPV genes on foreign gene expression from the late p6.9 promoter and BV production has been presented. Chapter 4 outlines the adaptation of NGS to rBEVs and the application of NGS to generate a virus stock consensus sequence and detect transient CRISPR-Cas9-mediated mutations. Identification of *gp64* variants upon targeted mutations and variant conservation upon viral propagation in cell culture, as well as reduction of virion co-production in cell culture,

have been presented in Chapter 5. Finally, Chapter 6 outlines overall conclusions and recommendations derived from this work. Chapter 3 has been formatted for submission, Chapter 4 has been published in *Viruses*, and Chapter 5 has been published in *International Journal of Molecular Sciences*.

Chapter 2

Literature Review

2.1 Baculovirus

Baculoviruses are enveloped DNA viruses belonging to the *Baculoviridae* virus family, with a very narrow host range within the arthropod invertebrates (Lu, Miller, 1997; Jehle et al., 2006). These viruses are divided into two major groups, nucleopolyhedroviruses (NPVs) and granuloviruses (GVs), based on the morphology of their occlusion bodies (Rohrmann, 2019c). While the former can form occlusion bodies composed of one or more virions encapsulated in a polyhedrin protein matrix, the latter can form occlusion bodies containing a single virion encased in a granulin protein matrix (Funk et al., 1997; Rohrmann, 2019c). The baculovirus virions can exist in one of the two forms during their life cycle: the occluded form, which is responsible for transmission between hosts, and the budded form for transmission within a host or cell culture (Figure 2.1) (Rohrmann, 2019c).

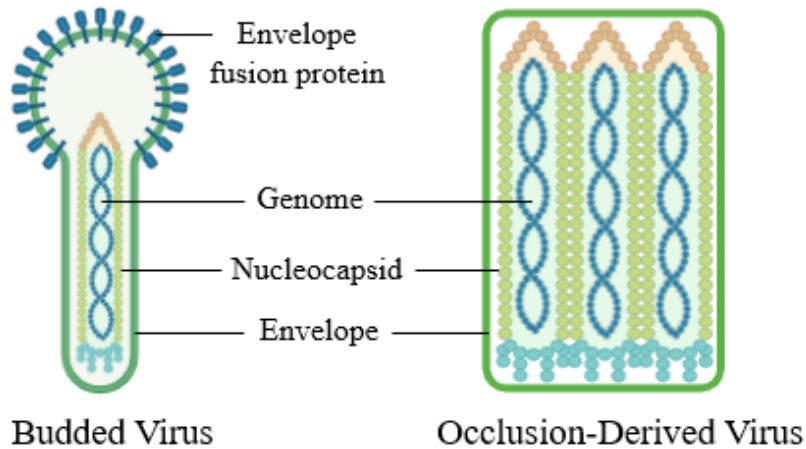


Figure 2.1: Baculovirus virion phenotypes.

2.1.1 Baculovirus Genome

The most widely studied baculovirus, AcMNPV, belongs to the *Alphabaculovirus* genus of the *Baculoviridae* family (Jehle et al., 2006; Chen et al., 2013). Along with a 134 kbp circular, supercoiled double-stranded DNA genome that has an overall A+T content of 59%, the baculovirus also features rod-shaped nucleocapsids approximately 300 nm in length and 50 nm in diameter, encasing the genome (Ayres et al., 1994; Jehle et al., 2006). The complete AcMNPV genome sequence data of the C6 strain was published in 1994 (Ayres et al., 1994) and can be accessed from NCBI (Reference Sequence: NC_001623.1). While it was originally reported that the AcMNPV encodes around 155 open reading frames (ORFs), upon adjustments made to the AcMNPV genome sequence it was realized that there are about 150 ORFs in the AcMNPV genome (Miele et al., 2011; Rohrmann, 2019a) and the average distance between the ORFs is 53 bp (Chen et al., 2013). Although the origin of baculovirus DNA replication is not known clearly, some reports suggest that the

homologous regions are the DNA replication initiation site (Kool et al., 1993), while others claim that initiation can occur at other sites (Habib, Hasnain, 2000). Nevertheless, DNA replication in baculovirus occurs either by rolling circle mechanism (Volkman, Openheimer, 1997) or by recombination (Okano et al., 2007) and requires several baculovirus gene products.

2.1.2 Baculovirus Gene Transcription and Infection Cycle

Baculovirus transmission in insect cell culture is mediated by adsorptive endocytosis, followed by nucleocapsid release into the cytoplasm by the baculovirus surface glycoprotein GP64 (Volkman, Goldsmith, 1985; Blissard, Wenz, 1992). After migrating to the nuclear pores, these nucleocapsids enter the nucleus via the pore complexes (Ohkawa et al., 2010) where destabilization of the capsid is thought to result in the release of the viral genome (Funk, Consigli, 1993). Subsequently, the uncoated viral genome in the nucleus is used for viral gene transcription and DNA replication (George, 2016). The baculovirus infection cycle is divided into three temporal phases: early, late, and very late. This temporal gene transcription ensures that gene products that are required for infection progression to the next class are available at desired times (Berretta et al., 2013). During the early phase, viral genes are recognized and transcribed by the host cell RNA polymerase II (Rohrmann, 2019c). This early phase, also known as the virus synthesis phase, occurs from 0.5 to 6 hours post-infection (hpi) (Chisholm, Henner, 1988; Kelly et al., 2007) and includes two sub-phases, immediate early and delayed early, during which the virus prepares the infected cells for viral DNA replication.

With the ability to encode their own RNA polymerase, baculoviruses do not depend on the host cell transcriptional system for viral late and very late gene(s) transcription (Rohrmann, 2019c). The cells start to produce extracellular virus (EV), also known as non-occluded virus (NOV) or BV, during the late or viral structural phase, between 6 to 12 hpi. Late genes coding for viral DNA replication and viral assembly are expressed with the peak release of EV between 18 to 36 hpi (Kelly et al., 2007). The very late phase, also known as the viral occlusion protein phase, marks the expression of the *polyhedrin* and *p10* genes related to the formation of occlusion bodies (OB)/polyhedral inclusion bodies (PIBs), and the beginning of cell lysis (Kelly et al., 2007; Chen et al., 2013). Nevertheless, the insect cell line and baculovirus strain can control the timing of the baculovirus infection cycle.

The transition from the host cell to viral translational machinery for protein synthesis could be to allow for better virus production (Nobiron et al., 2003). Cell cycle arrest of baculovirus-infected insect cells occurs at the G2/M phase (Braunagel et al., 1998). It is believed that to permit virus entry into the host cell nucleus, the virus deliberately causes cell arrest when nuclear membrane permeability is optimal (Braunagel et al., 1998). The majority of the host genes are down-regulated over the infection course with complete loss of host protein translation in the later phases of the infection cycle (Chen et al., 2013; Monteiro et al., 2012; Nobiron et al., 2003). Moreover, baculovirus is known to encode anti-apoptotic proteins such as P35 and inhibitor of apoptosis proteins (IAPs) to prevent premature cell death of infected cells by apoptosis, which would reduce virus replication levels (Birnbaum et al., 1994; Clem, Miller, 1993). Both P35 and IAPs function by blocking the initiator and effector caspases activity in the infected insect cells (Rohrmann, 2019b).

Another host cell stress response upon baculovirus infection besides apoptosis is the heat shock response (HSR) (Monteiro et al., 2012). AcMNPV up-regulates the expression of members of the 70 kDa chaperone family of heat shock protein (HSP70) and heat shock cognate (HSC70) in infected insect cells to help viral DNA replication as well as virion assembly and budding of progeny viruses (Monteiro et al., 2012; Salem et al., 2011; Lyupina et al., 2011, 2010). Moreover, the baculoviral replication-specific late expression factor (lef) proteins stimulate host cell apoptosis and cause translational arrest of host proteins during the late phase of the infection cycle (Schultz, Friesen, 2009). A detailed review of the baculovirus infection cycle can be found elsewhere (Rohrmann, 2019d).

2.2 Insect Cell Culture

As early as the 1930s, insect cell lines were initiated and used for the replication of insect baculoviruses and arboviruses (Arif, Pavlik, 2013). Insect cells usually do not carry any infectious human virus and especially retrovirus (Summers, 2006). Many continuous cell lines, such as *Spodoptera frugiperda* cells (Sf9 or Sf21 clones) and *Trichoplusia ni* cells (High-FiveTM clone), have been isolated and characterized for the production of biologics (Mena, Kamen, 2011). The cell lines grow well in suspension, thus allowing for recombinant protein production in large-scale bioreactors, and were rapidly developed for suspension culture in serum-free media (SFM) to minimize production costs (Maranga et al., 2003). The three most common insect cell lines, Sf9, Sf21, and High-FiveTM, used in research are derived from lepidopteran insects (Arunkarthick et al., 2017) with Sf9 and Sf21 cells being derived from pupal ovarian tissue of the fall armyworm, *Spodoptera frugiperda* (Wickham

et al., 1992) and High-FiveTM cells from the ovarian tissues of the cabbage looper, *Trichoplusia ni* (Hong et al., 2015). The ideal temperature for insect cell culture maintenance is between 26 to 28 °C, with a lower temperature resulting in decreased growth rate and a temperature higher than 30 °C leading to lower cell viability (Arunkarthick et al., 2017).

Grace's supplemented medium (TNM-FH), developed in the 1960s, has been a traditional medium of choice for *in vitro* insect cell culture (Grace, 1962). Other serum or hemolymph dependent media, such as IPL-41 and TC-100, and SFM, such as EXPRESS FIVE SFM, have been developed thereafter (Chan, Reid, 2016; Bauer, Schnapp, 2007). Optimized second generation SFM, Sf-900TM II and III SFM, with lot-to-lot consistency, reduced downstream processing complexity, and low cost, have facilitated large-scale production of recombinant proteins (Zhang et al., 1992). In serum-free and protein-free medium cultures, insect cells have a fast specific growth rate. Running batch and fed-batch bioreactors with cell density reaching a high level is quite straightforward and robust, till the culture conditions are tightly monitored and controlled. The ability of insect cells to grow in SFM is highly beneficial as it eliminates the chances for the occurrence of adventitious viruses, thus reducing the need for more intensive purification processes (Vicente et al., 2011). Furthermore, the requirement for annual adjustments to some virus vaccines, such as the influenza vaccine, makes BEVS suitable for the production of different commercial vaccines (Cox, Hollister, 2009).

2.3 BEVS and Its Advantages

The BEVS has become a versatile and powerful eukaryotic vector system for recombinant protein production, since its introduction in 1983 (Smith et al., 1983b). This platform, comprising of recombinant baculovirus and insect cell culture, involves the insertion of one or more gene(s) of interest into a baculovirus genome using a transfer plasmid, under the control of a baculovirus or insect promoter sequence and at a site that does not interfere with the baculovirus replication *in vitro* (George, 2016). The ideal sites for insertion of the gene(s) of interest are thought to be situated downstream of the very strong, very late baculoviral p10 or polyhedrin promoters, which leads to very high expression levels of the foreign genes (Smith et al., 1983c).

BEVS offers a number of advantages over other expression vector systems. The restricted host range and non-pathogenicity to mammals and plants make baculovirus an ideal choice (Fuxa, 1991). The requirement of helper cell lines or helper viruses is eliminated as all the genetic information can be contained in the baculovirus genome. The recombinant products produced using BEVS can be easily scaled up for large-scale production. Moreover, high levels of recombinant gene expression can be achieved using this system (Sokolenko et al., 2012) along with the mammalian-like post-translational modification capacity of insect cells.

The baculovirus has limited permissive cell lines for infection, and the BEVS has the advantage of lower risks of contamination with human adventitious agents over mammalian platforms and the ability to provide eukaryotic protein modifications, including N-glycosylation (Toth et al., 2014). While expression systems such as bacteria and yeast

are known to achieve higher yields, they do not match the degree of complexity, such as the expression of multiple or different gene(s) of interest required for VLPs, which can be achieved with BEVS (Vicente et al., 2011). Additionally, the product yields obtained are significantly higher using BEVS when compared to egg-based or mammalian cell-based VLP production, and the production time is also reduced to 12 weeks (James, 2009).

2.4 Baculovirus Vector Development for Biologics Production

2.4.1 Generation of Recombinant Vectors

AcMNPV is the most widely studied baculovirus that is extensively used for the production of recombinant proteins, VLPs, and other vaccine candidates (Jehle et al., 2006; Chen et al., 2013). Since the early 1980s, these viruses have been exploited as expression vectors (Smith et al., 1983b). The generation of a recombinant baculovirus vector was initially performed through homologous recombination (HR). The HR process generally involves co-transfection between the genomic AcMNPV DNA and a transfer plasmid DNA containing a foreign gene under the polh promoter (Kost et al., 2005; Smith et al., 1985). Upon co-transfection, a mixture of parental and recombinant progeny viruses is produced, and the progeny rBEVs are then resolved by tedious plaque assay based on the occlusion-negative plaque phenotypes of the recombinant clones (Kost et al., 2005; Smith et al., 1983a). However, it is to be noted that this HR resulted in a recombination frequency of only $\sim 0.1\%$ (Kost et al., 2005). In an effort to improve the recombinant virus recovery,

a unique Bsu36I site in the *polyhedrin* locus to linearize the AcMNPV DNA prior to co-transfection increased the HR frequency to $\sim 30\%$ (Kitts et al., 1990). Further improvement to rBEVs selection was made by engineering baculovirus DNA with multiple Bsu36I sites and specifically one within the essential *orf1629* AcMNPV gene (Kitts, Possee, 1993). This resulted in functional inactivation of the *orf1629* gene and linearization of the AcMNPV DNA — named BacPAK6, which, when co-transfected with a transfer vector containing the *orf1629* and foreign genes, restored the *orf1629* gene and produced infectious progeny viruses (Kitts, Possee, 1993). Although this approach achieved an HR frequency of over 90%, rBEVs isolation via plaque assay was still necessary (Kitts, Possee, 1993; Kost et al., 2005).

To improve rBEV generation and overcome the requirement of plaque assays for rBEV isolation, a bacmid system was developed based on the site-specific transposition of a foreign gene at the AcMNPV *polh* locus and a novel baculovirus vector could replicate as a plasmid in *E. coli* (DH10BAC) (Luckow et al., 1993). This AcMNPV shuttle vector (bacmid) contained a bacterial artificial chromosome (BAC) consisting of a mini F-replicon and a selectable antibiotic resistance marker, an *attTn7* site — bacterial transposase Tn7 target sequence, and a *lacZ* gene, inserted at the *polh* locus of AcMNPV (Luckow et al., 1993). The selection and identification of *E. coli* clones that received the bacmid DNA were made easy by the antibiotic resistance and *lacZ* markers, thus eliminating the need for plaque assays (Luckow et al., 1993; Kost et al., 2005). With the advantage of being a simplified method for rBEV generation, this system has been broadly used and was later commercialized as the Bac-to-Bac system (Kost et al., 2005). Nevertheless, it was observed that bacmid-derived parental and mutant AcMNPV vectors had a spontaneous deletion of

the BAC vector sequences upon repeated propagation in insect cells, and the instability of the parental bacmid led to the accumulation of defective interfering particles (DIPs) in cell culture (Pijlman et al., 2003).

In the last decade, other systems have been developed by combining the bacmid with HR based on the linearized BacPAK6 system in insect cells for rBEV generation, such as the flashBACTM system licensed by Oxford Expression Technologies Ltd. (van Oers, 2011). This system has the *chiA* gene removed from the AcMNPV genome; and has a truncated *orf1629* gene and a BAC containing a replicon and antibiotic resistance marker at the *polh* locus of the AcMNPV, for circular viral genome replication in *E. coli* (van Oers, 2011). Following viral genome isolation and purification from *E. coli*, the AcMNPV genome is co-transfected with a transfer plasmid for HR into insect cells, thus resulting in the integration of the foreign gene into the rBEV, restoration of the *orf1629* gene for virus propagation, and removal of the bacterial sequences in the final rBEV (van Oers, 2011; Bruder, 2021). Other variants of the flashBACTM system, such as the flashBACTM GOLD and flashBACTM ULTRA, follow the same principle and have the *chiA* and *v-cath* genes, and the *chiA*, *v-cath*, and *p26/p10/p74* genes deleted from their AcMNPV genomes, respectively (OET, 2019).

Additionally, another system based on Gibson assembly (Gibson et al., 2009) has been recently developed that involves rBEV generation by the assembly of several linear DNA fragments and is known as biGBac (Weissmann et al., 2016). Another study reported on the production of the first synthetic baculovirus genome, and it was a significant milestone in the BEVS field. This AcMNPV-based synthetic genome, known as AcMNPV-WIV-Syn1, was synthesized by a combination of PCR and transformation-associated recombination

(TAR) in yeast, and following transfection of Sf9 cells with AcMNPV-WIV-Syn1 DNA, the rescued progeny virus was found to be comparable to the parental virus (Shang et al., 2017).

2.4.2 Genome Engineering of Recombinant Vectors

While the AcMNPV genome remains largely unmodified, some key advances have been made in the genome engineering front by eliminating certain AcMNPV genes to improve recombinant protein production and stability. Disruption of the *chiA* and *v-cath* genes, whose protein products have been observed to cause liquefaction of infected insect host and proteolytic breakdown of recombinant proteins in infected cell culture, was presumed to improve recombinant protein stability (Hawtin et al., 1997; Berger et al., 2004; Kaba et al., 2004). Specifically, infection of host cells with a rBEV in which the *chiA* and *v-cath* genes were deleted, resulted in the cells staying intact till ~ 72 hpi along with a reduction in the proteolytic breakdown of the recombinant product (Berger et al., 2004; Kaba et al., 2004; Hitchman et al., 2010b). Moreover, it was shown that infections with double-deletion rBEVs ($\Delta chiA$ and $\Delta v-cath$) led to higher levels of secreted and nuclear or cytoplasmic recombinant proteins when compared to the wild type and single deletion ($\Delta chiA$ or $\Delta v-cath$) rBEVs (Hitchman et al., 2010b). Additionally, deletion of the dispensable AcMNPV genes — *p26*, *p10*, and *p74* with ORFs adjacent to each other in the AcMNPV genome improved the recombinant protein yield in infected cell culture; however, these deletions did not improve cell viability post-infection (Hitchman et al., 2010a). Nonetheless, higher levels of recombinant proteins at a time point earlier than the control were achieved with each of these modified rBEVs, thus confirming that the recombinant protein yield and

integrity could be enhanced by the introduction of these gene deletions in the AcMNPV genome (Hitchman et al., 2010a). More recently, overexpression of *lef5* driven by different promoters was shown to enhance the stability of exogenous genes and protein expressions during serial passages in Sf9 cells (Pei et al., 2025). Particularly, the stability of enterovirus 71 (EV71) VLPs was improved when *lef5* was overexpressed under the op166 promoter and *EV71-3CD* and *EV71-P1* were driven by the p6.9 and vp39 promoters, respectively, thus demonstrating its importance for complex protein production (Pei et al., 2025).

Furthermore, some studies have focused on reducing the burden on downstream purification of biologics by eliminating the co-production of recombinant proteins and progeny viruses. Despite various purification processes, such as chromatography, ultracentrifugation, and/or filtration, used to reduce baculovirus contaminants, some residual virions remain in the recombinant products (Lee et al., 2015), and these processes can also lead to significant product losses at each step (Bruder, 2021). Additionally, for enveloped VLP production, the similar size and density of the rBEVs and VLPs can complicate their separation via ultracentrifugation and chromatography (Fernandes et al., 2013; Yee et al., 2018). Thus, disruption of essential AcMNPV structural genes, that is, genes required for virion replication or assembly, would aid the downstream purification process by preventing rBEV contamination in culture supernatant (Yee et al., 2018). Nevertheless, the deletion of these viral genes would impair virus amplification to high-titer stocks required for protein production, and a trans-complementing insect cell line engineered to express the deleted gene is necessary for infectious virus production. These infectious rBEVs can then be used for infecting parental insect cells to produce recombinant proteins or VLPs without virion contaminants (Fernandes et al., 2013; Yee et al., 2018). Two groups assessed the disrup-

tion of the *vp80* or *gp64* genes expressing the AcMNPV capsid-associated protein VP80 or major envelope protein GP64, for the production of enhanced GFP (eGFP) or HIV-1 Gag VLPs along with a reduction in virus contamination, respectively (Marek et al., 2011; Chaves et al., 2018). It was observed that the absence of VP80 had no effect on very late gene expression and eliminated the co-production of virions (Marek et al., 2011), while deletion of the *gp64* gene resulted in reduced BV production and lower Gag expression in cells infected with $\Delta gp64$ rBEV (Chaves et al., 2018).

Almost a decade ago, one study identified essential and non-essential AcMNPV genes via extensive literature mining while performing comparative analysis with neighboring genomes of the *Alphabaculovirus* genus, along with scrutinizing gene synteny, promoter motifs, repetitive features in ori, and transposon integration sites (Vijayachandran et al., 2013). Out of the 156 ORFs shown in this genome map, 94 were designated as essential and 62 were non-essential based on gene synteny or uniqueness to the AcMNPV genome and function in oral infectivity, cell lysis, or apoptosis, respectively (Vijayachandran et al., 2013). However, given the diversity in genome content and organization of the *Baculoviridae* family, ~ 30 genes being conserved among all members, host-virus specific interactions resulting in an inconclusive assessment of gene function, and the lack of experimental confirmations for many AcMNPV gene functions contribute to the challenges that this study presents (Miele et al., 2011; Ono et al., 2012). More recently, two studies probed AcMNPV genes for their essentiality in BV production by generating gene-knockout bacmids and performing transfection and infection assays (Chen et al., 2021; Yu et al., 2023). While one study reported that 36 of 42 genes are dispensable for infectious BV production (Chen et al., 2021), another demonstrated 62 of 80 genes to be non-essential for progeny virus

production (Yu et al., 2023). Additionally, they also created theoretical minimal AcMNPV genome maps of ~ 68.3 or ~ 90.1 kbp by removing all the identified dispensable genes and most of the homologous repeat regions (hrs) or all the re-verified non-essential genes, respectively (Chen et al., 2021; Yu et al., 2023). Another study deleted 14 AcMNPV gene fragments, each containing at least two continuous known non-essential (for progeny virus) or unknown function genes, to investigate their impact on BV production and foreign gene expression from the very late p10 promoter (Zhang et al., 2023). While 4 DNA fragments containing 25 genes were found to be essential or influential for BV and foreign protein production, 1 fragment containing 3 genes was dispensable for BV but reduced exogenous protein production from the p10 promoter (Zhang et al., 2023).

2.4.3 Recombinant Vector Infection Strategies

Using the BEVS platform, a protein complex can be produced by expressing one or more foreign gene(s) via one of the three infection strategies: coinfection, coexpression, or a combination of both (Sokolenko et al., 2012). While coinfection, using multiple virus constructs — each carrying a gene of interest (monocistronic), provides an additional degree of flexibility that is beneficial for early exploratory work, coexpression, using a single virus construct carrying all the genes of interest (polycistronic), is preferred at the larger scale owing to the reduced number of virus constructs required for infection (George, 2016). Specifically, coexpression reduces the total amount of baculoviruses while decreasing the number of possible baculovirus combinations that can be found in a cell (Sokolenko et al., 2012). Moreover, a polycistronic baculovirus allows the desired production of all foreign proteins in every infected cell, and the expression ratios can be modulated by manipulating

the promoters that control the levels of protein expression (Road, 1995). Usually, when a baculovirus is used as an expression vector for *in vitro* infection, the naturally occurring *p10* or *polh* gene in a wild-type baculovirus is replaced with recombinant gene(s) of interest. A literature review revealed that single proteins or VLPs production was driven in the majority by p10 or polh promoters, with only a small fraction of them being driven by a promoter in tandem with a very late promoter and a handful used promoters other than p10 or polh (Sokolenko et al., 2012). Although the polh and p10 promoters are known to produce large amounts of recombinant protein(s), they are active in the very late phase of the baculovirus infection cycle when the cells are dying. The use of alternative promoters to control the timing and expression of foreign proteins to tailor expression ratios within the cell has been demonstrated previously (George, 2016; Bruder, Aucoin, 2022).

Many groups in the past have used coinfection as the primary expression strategy for the production of multi-protein complexes, owing to the ability of insect cells to be infected by multiple baculovirus constructs (Aucoin et al., 2006; Mena et al., 2010; Meghrouh et al., 2005). The crucial parameters for coinfection involve manipulation of the multiplicity of infection (MOI) and time of infection (TOI) to control the order, timing, and stoichiometric ratios of multi-protein complexes (Sokolenko et al., 2012; Bruder, 2021). Coinfection offers the advantage of investigating the ratio and timing of each gene necessary for the optimal production of VLPs or proteins that require extensive PTMs, by altering the MOI and TOI of individual rBEVs (Sokolenko et al., 2012). The production of AAV vectors using 3 rBEVs, expressing the AAV replication and structural proteins and the AAV genome, demonstrated how coinfection could lead to differences in yield depending on baculovirus ratios selected, thus highlighting the necessity to understand the relationship between

baculovirus constructs to optimize the coinfection strategy (Aucoin et al., 2006). Moreover, the inherent instability of large baculovirus constructs containing multiple genes makes the process scale-up unfavorable, thus projecting the importance of the coinfection strategy (Aucoin, 2007).

After repeated virus amplification in cell cultures at high MOI (> 3 pfu/cell), the production of recombinant protein and baculovirus is usually accompanied by the appearance and accumulation of DIPs (Wickham et al., 1991). These particles are known to interfere with the helper virus replication, lack considerable portions of the genome, and cause the passage effect. The phenomenon of passage effect, that is, the generation of defective particles upon recurrent passaging at high MOI, complicates the scale-up of baculovirus and recombinant protein *in vitro* (Pijlman et al., 2001). In a previous study, the presence of DIPs with a major genomic deletion upon serial passage of AcMNPV-E2 in Sf21 cells has been observed (Kool et al., 1991). Based on this work, another study demonstrated the presence of DIPs lacking 43% of the viral genome in low passage AcMNPV-E2 virus stocks and polyhedra, but not in AcMNPV isolate obtained prior to cell culture passaging (Pijlman et al., 2001). Thus, it has been postulated that the rapid generation of DIPs is an intrinsic property of baculovirus infection in insect cell culture, involving several recombination steps (Pijlman et al., 2001).

During a low MOI ($\ll 1$ pfu/cell) infection, not all cells are infected at once, and instead consist of multiple infection cycles. Hence, a low MOI infection results in some cells receiving the replicative virus, some receiving DIPs, and others not receiving any virus particles (Wickham et al., 1991). Several studies have shown the use of low MOI for virus amplification, thus reducing the passage effect. For the production of self-forming *Porcine*

parvovirus-like particles using the BEVS, a low MOI was used to avoid an additional virus amplification step and to minimize accumulation of DIPs (Maranga et al., 2003). Another study on the occurrence of DIPs within high passage AcMNPV, interference with recombinant protein, and infectious virus production was presented using three insect cell lines, and significant reductions in the specific productivity of a recombinant protein with MOI over 0.01 pfu/cell were observed (Wickham et al., 1991). Additionally, factors such as pH, temperature, and dissolved oxygen, and infection and harvest parameters (including MOI, TOI, and time of harvest (TOH)) are critical for optimal baculovirus vector production processes. To minimize the production of DIPs, which interfere with process productivity and baculovirus quality, a low MOI is preferred for virus amplification (Pijlman et al., 2003). Moreover, it has been reported that the accumulation of non-homologous repeat origin of replication regions (non-hrs ori) causes defective particle formation, and the deletion of these regions from the viral DNA can help to prevent the accumulation of DIPs and enhance genome stability (Pijlman et al., 2003). Although different genetic engineering techniques for vector development and/or enhancing protein yield and quality have been applied to address some drawbacks of the BEVS, there are potential areas that need to be improved for efficient production and advancement of the system (Possee et al., 2020).

2.5 Production Platform for Biologics

It is known from the literature that different platforms, such as bacteria, yeast, plant cells, mammalian cells, and insect cells, are available for recombinant protein production. While egg-based technology has been used for a long time for the production of inactivated and

live attenuated influenza vaccines, they are labor-intensive, have associated risks, and long production periods (Milián, Kamen, 2015). In the biopharmaceutical industry, however, the production of therapeutic proteins or vaccines using mammalian cell culture technology is well established (Le et al., 2010). Several mammalian cell lines, such as MDCK (Madin Darby canine kidney cells) (Rimmelzwaan et al., 1998), PER.C6 (human embryonic retinal cells) (Pau et al., 2001), Vero (African green monkey derived kidney cells) (Barrett et al., 1998), and HEK293 (human embryonic kidney cells) (Le et al., 2010), have been explored for the production of biologics. Optafu/Flucelvax[®] (Novartis) vaccine, which is a trivalent vaccine produced in the MDCK cells, was approved by the FDA in 2012 (Doroshenko, Halperin, 2009). Another licensed influenza vaccine, Perflucel[™], that is formulated with inactivated H1N1, H3N2, and influenza B virus is produced in the Vero cells (European Medicine Agency, 2012).

A recombinant baculovirus expression vector carrying the gene(s) of interest can be designed faster and thus result in shorter turnaround periods when compared to other production methods (Cox, Hashimoto, 2011). Moreover, a well-accepted safety profile of the processes employing BEVS (biosafety level I) has led to their consideration for commercial vaccine production (Vicente et al., 2011; Palomares et al., 2015). Since the establishment of the BEVS over 35 years ago for recombinant protein production, it has been used for a multitude of biotechnological applications (Possee et al., 2020). The applications of the BEVS as a production platform can be broadly classified into four categories, including the production of single or multi-subunit recombinant proteins, gene delivery vehicles for mammalian cells, viral vectors for gene therapy, and baculovirus particles with a surface display of antigens (van Oers et al., 2015).

2.5.1 Single and Multi-Subunit Recombinant Proteins

The BEVS has been used for the successful production of different recombinant proteins such as subunits, glycoproteins, and VLPs (van Oers et al., 2015). Additionally, the recombinant protein subunits and VLPs produced in the BEVS have found applications as vaccines (van Oers et al., 2015). The BEVS allows the use of multiple monocistronic rBEVs, each expressing a single protein or a single polycistronic rBEV expressing multiple proteins for the production of simple recombinant proteins, multiprotein complexes, and VLPs, via coinfection or coexpression, respectively (Sokolenko et al., 2012). Moreover, in the absence of cell culture-based systems for viral replication, such as human papillomavirus (HPV) and hepatitis C virus (HCV), the use of rBEVs for viral protein expression is of particular importance (Kost et al., 2005).

Complex enveloped and non-enveloped VLPs, such as influenza and rotavirus VLPs, comprised of multiple proteins, have been manufactured in the BEVS platform (Kang et al., 2009; Liu et al., 2013). The L1 capsid protein of the HPV is known to self-assemble into non-enveloped VLPs when expressed in the BEVS (Kirnbauer et al., 1992). Cervarix (GlaxoSmithKline), a licensed recombinant HPV VLP cancer vaccine composed of the HPV serotypes 16 and 18 L1 structural proteins, was successfully produced using the BEVS (Harper, 2008; Roldão et al., 2010). In addition, the BEVS is particularly valuable for the production of influenza vaccines due to the annually changing behaviour of influenza serotypes (Cox, Hollister, 2009). Both recombinant subunit and VLP influenza vaccines generated in the BEVS have exhibited protective immunity during preclinical and clinical trials (Roldão et al., 2010; Liu et al., 2013). The commercialized influenza A subunit vaccine FluBlok (Protein Sciences), consisting of recombinant hemagglutinin (rHA) protein

trimers, is manufactured in the BEVS, and this trivalent flu vaccine is annually updated to incorporate the relevant HA serotypes (Cox, Hollister, 2009; Kang et al., 2009). Another vaccine that is produced in the BEVS is a subunit marker vaccine containing a secreted form of the E2 glycoprotein to provide protection against classical swine fever virus (van Oers et al., 2015).

Moreover, the assembly of human severe acute respiratory syndrome coronavirus (SARS-CoV) VLPs expressing the spike (S), envelope (E), and membrane (M) structural proteins in the BEVS was first demonstrated in 2004 (Ho et al., 2004). In light of the recent COVID-19 pandemic, SARS-CoV-2 vaccines produced in the BEVS are in different stages of clinical trials (Liu et al., 2020) or have received regulatory approvals. Of particular interest is the NVX-CoV2373 (Novavax) SARS-CoV-2 subunit vaccine, comprising of the full-length prefusion conformation S protein and Novavax’s patented Matrix-MTM adjuvant (Keech et al., 2020; Tian et al., 2021), which is now approved by the FDA. Additionally, overall efficacies of 89.7% and 90.4% were reported at the primary endpoint of the Phase 3 trials for the NVX-CoV2373 vaccine in the UK and the US, respectively (Novavax, 2021).

2.5.2 Gene Delivery in Mammalian Cells

rBEVs can be used as gene delivery vehicles for the transient expression of recombinant proteins in different mammalian cells such as HeLa, CHO, 293, and bone marrow fibroblasts (Condreay et al., 1999). With the ability of AcMNPV to enter mammalian cells by transduction, rBEVs have been exploited to deliver transgenes of interest into these cells under mammalian expression cassettes (Hofmann et al., 1995; Condreay et al., 1999).

These baculovirus vectors, commonly referred to as BacMam vectors, are not capable of replicating in mammalian cells and thus cannot produce infectious progeny viruses ([Hofmann et al., 1995](#); [Possee et al., 2020](#)). The safety of baculoviruses, efficient mammalian cells transduction, and fast generation of baculovirus vectors, offered by the baculovirus technology, have made it a suitable platform for scalable virus production ([Lesch et al., 2008](#)). In addition, large foreign DNA stretches can be inserted in multiple segments into the viral genome ([van Oers et al., 2015](#)).

It was recently demonstrated that expressing vesicular stomatitis virus G-protein (VSV-G) under different AcMNPV promoters (ie1, gp64, p10) while driving a reporter gene under a CMV promoter, enhanced transduction efficiency in mammalian cells without compromising BV production ([Simonin et al., 2025](#)). Moreover, the highest transduction efficiency was observed when VSV-G was driven by the p10 promoter, thus exhibiting a promoter-tuning approach for improving vector production ([Simonin et al., 2025](#)). These gene delivery vehicles could also be used for high-throughput screening of gene function(s), drug testing, or transplantation therapies ([Kost et al., 2005](#); [Murguía-Meca et al., 2011](#)). Another useful application of this baculovirus technology is the production of large-scale replication-deficient lentiviruses in mammalian cells for clinical use ([Lesch et al., 2008](#)). At present, recombinant baculovirus vectors are used to deliver all the elements necessary for safe non-replicative lentivirus vector generation in 293T mammalian cells, with no detection of replication-competent lentivirus formation ([Lesch et al., 2008, 2011](#)).

2.5.3 Viral Vectors for Gene Therapy

The production of other viral vectors is a valuable application of the BEVS platform. With a single-stranded DNA genome, the non-enveloped adeno-associated virus (AAV) cannot replicate independently and requires a helper virus, such as an adenovirus or herpes simplex virus (HSV), for its replication to acceptable titers (Hoggan et al., 1966; Possee et al., 2020). However, it is known from the literature that BEVS can be used for the production of rAAV vectors for gene therapy applications (Urabe et al., 2002). The first rAAV-based gene therapy product, Glybera, approved by the FDA for human use to treat the rare autosomal-recessive genetic disorder lipoprotein lipase deficiency, was manufactured using the BEVS (Haddley, 2013; Felberbaum, 2015). Additionally, scalable high-yield BEVS-based rAAV production, comparable to that produced in HEK293 cells, has been achieved with advancements in the BEVS (Kondratov et al., 2017; Joshi et al., 2021).

Following the initial work by Urabe et al. on the production of functional rAAV vectors by coinfecting Sf9 cells with three rBEVs to deliver the essential genes *rep*, *cap*, and transgene flanked by inverted terminal repeats (ITRs), significant improvements have been made to reduce the number of coinfections (Urabe et al., 2002). While the improved Two-Bac system required coinfection with two rBEVs, given that the *rep* and *cap* genes were delivered by a polycistronic rBEV (Smith et al., 2009), the Mono-Bac system used a single rBEV to deliver the three essential genes (Galibert et al., 2021). Moreover, around four BEVS-based gene therapies for Hemophilia A and B are in different stages of clinical trials, with Valoctocogene Roxaparvovec gene therapy for Hemophilia A and Etranacogene Dezaparvovec gene therapy for Hemophilia B recently being approved by the FDA for human use (Dolgin, 2016; Pasi et al., 2020; Anguela, High, 2024; Symington et al., 2024).

2.5.4 Baculovirus Surface Display

The display of foreign proteins on the surface of a virus is a valuable technology for the biopharmaceutical industry. Foreign proteins are either displayed on the virus or the infected cell surface via fusion with the viral envelope glycoprotein GP64 (Tsai et al., 2020). This gene fusion with viral GP64 facilitates the efficient transport of the foreign antigens to the cell membrane for surface display on rBEV or infected cells (Grabherr et al., 1997). Additionally, since the rBEV comprises of the foreign proteins fused to the wild-type GP64, the infectivity of the virus is not compromised, and the rBEV can be propagated for high levels of foreign gene expression (van Oers et al., 2015; Tsai et al., 2020). While common practice for these modifications is to clone an additional copy of the entire *gp64* ORF downstream of the very late *polh* promoter, the use of only the C-terminal transmembrane domain of the *gp64* along with the early and late *gp64* promoter has also been explored (Grabherr et al., 1997).

It is known from the literature that for the surface display of eukaryotic proteins with complex PTMs and protein folding, the BEVS platform is a better alternative as compared to prokaryotic phage systems (Makela, Oker-Blom, 2008). Furthermore, the large insert capacity, ease of purification process for rBEV, and safe operating system make the BEVS an attractive platform for surface display of antigens (Tsai et al., 2020). With the ability to preserve the native multimeric fusion protein structure by the surface display technique, the utility of BEVS for the production of biopharmaceuticals has been expanded (Tsai et al., 2020). Moreover, the rBEVs displaying antigens of interest on their surface can be used as vaccine candidates and diagnostic platforms, as well as for drug screening, protein function studies, and gene therapies (Sergeeva et al., 2006; Makela, Oker-Blom, 2008; Tsai

[et al., 2020](#)). A detailed review of baculovirus surface display technology can be found elsewhere ([Grabherr et al., 2001](#); [Tsai et al., 2020](#)).

2.6 CRISPR-Cas9 Genome Editing Tool

With the development of recombinant DNA (rDNA) technology in the 1970s, several genome editing tools have emerged for the addition, deletion, and manipulation of genetic material at specific genome locations ([Hsu et al., 2014](#)). While the traditional gene engineering technologies, such as homologous recombination, have facilitated advances in biotechnology research and applications, these methods are time-consuming, labor-intensive, and often result in impaired virus replication along with wild-type virus production ([Hsu et al., 2014](#); [Bruder, 2021](#)). Nonetheless, recent successes in genetic engineering to overcome the challenges of the traditional methods have led to the development of CRISPR-Cas systems for precise genome editing.

Clustered regularly interspaced short palindromic repeat (CRISPR) short-sequence DNA repeats (SSRs) found in prokaryotic genomes constitute an array of highly conserved direct repeats that are interspaced by non-repetitive sequences called spacers, and these spacer sequences are known to match foreign genetic element sequences termed as proto-spacers ([Grissa et al., 2007](#); [Hsu et al., 2014](#)). Moreover, the signature CRISPR-associated (Cas) endonuclease genes clustered next to the SSRs are identified to be transcriptionally active and well-conserved ([Makarova et al., 2011](#); [Bruder, 2021](#)). This CRISPR-Cas system, comprising of the clustered *cas* genes, non-coding RNAs, and CRISPR array, is believed to be an adaptive immune system naturally occurring in many bacteria ([Ran et al., 2013](#)).

Three distinct CRISPR systems — Type I, II, and III with multiple subsystems have been identified and are generally differentiated by the presence of the signature *cas* genes — *cas3*, *cas9*, and *cas10*, respectively (Makarova et al., 2011, 2015). Additionally, another component of Type I and II systems, the protospacer adjacent motif (PAM), which is a 2–5 bp DNA sequence found adjacent to Cas targeted DNA sequence or one end of the protospacers, has functional importance for CRISPR-Cas immune mechanisms to separate self from non-self during target surveillance (Shah et al., 2013; Van Der Oost et al., 2014).

On the basis of the initial analysis, while it was hypothesized that the CRISPR system was a part of gene expression or DNA repair regulation (Makarova et al., 2002), it was later hypothesized to be involved in defense mechanisms following the observations that the spacer sequences matched viral and plasmid gene fragments (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). A three-stage process comprising of adaptation, expression, and interference, that is divided into two quasi-independent subsystems, is implemented by the CRISPR-based immunity system (Makarova et al., 2011). While the adaptation stage is included in the information processing subsystem, the expression and interference stages are included in the executive subsystem (Makarova et al., 2011). Short nucleotide sequence tags acting as the immune memory are acquired during the adaptation stage (Van Der Oost et al., 2014). Specifically, these nucleotide sequence tags, natively known as protospacers within the nucleotide sequence of invading bacteriophage and/or plasmids, are integrated into the host genome’s CRISPR loci by host Cas proteins, where they are referred to as spacers (Mojica et al., 2005; Bolotin et al., 2005; Pourcel et al., 2005). Following the adaptation phase, the CRISPR arrays are transcribed into a precursor CRISPR RNA (pre-crRNA) in the expression stage, which is cleaved into repeat-spacer-

repeat units and processed into short crRNAs (Deltcheva et al., 2011; Van Der Oost et al., 2014). Briefly, during this stage in the Type II system, a trans-activating crRNA (tracrRNA) containing a nucleotide sequence complementary to the pre-crRNA repeat region results in base-pairing between them, thus creating a double-stranded region that is recognized and processed by the housekeeping RNase III ribonuclease in the presence of Cas9 protein (Deltcheva et al., 2011; Van Der Oost et al., 2014). During the interference stage, the mature crRNA and specific Cas proteins assemble to form a stable ribonucleoprotein (crRNP) complex and then scan invader DNA for a nucleic acid sequence complementary to the crRNA and discriminate self from non-self. This is followed by base-pairing between the spacer and protospacer sequences and eventually degradation of the target by Cas nucleases (Van Der Oost et al., 2014).

Targeted and efficient modification of the eukaryotic genome has been made easier with the rapidly developing CRISPR-Cas9 genome editing technology (Hsu et al., 2014; Doudna, Charpentier, 2014). A profound interest in the CRISPR-Cas system led to developments that improved the basic understanding of this system and established it as a genetic engineering tool (Hsu et al., 2014; Doudna, Charpentier, 2014; Sander, Joung, 2014). Since the discovery of Cas9 as an RNA-programmable DNA endonuclease that is capable of introducing site-specific double-stranded DNA breaks (DSBs) in target DNA *in vitro* (Jinek et al., 2012), there has been a multitude of studies validating the efficacy as well as describing the applications of CRISPR-Cas9 for *in vivo* and *in vitro* genetic engineering (Doudna, Charpentier, 2014; Sander, Joung, 2014). Following the introduction of DSBs for targeted genome engineering, these DSBs are repaired by non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathways (Sander, Joung, 2014), and with the capability

of editing multiple sites in parallel in higher eukaryotes, it has an advantage over editing tools such as zinc finger nucleases (ZFNs) or transcription activator-like effector nucleases (TALENs) (Gaj et al., 2013). While NHEJ could result in efficient insertion-deletion (indel) mutations capable of disrupting reading frames, HDR could be used for introducing point mutations or desired sequences via recombination between the target locus and donor DNA templates (Sander, Joung, 2014). Moreover, the targeting of Cas9 was simplified by replacing the dual-RNA (tracrRNA crRNA) requirement with an engineered, chimeric single guide RNA (sgRNA) molecule containing a ~ 20 nucleotide target recognition sequence at the 5' end and a universal double-stranded hairpin structure at the 3' end that binds to Cas9, while retaining the Watson-Crick base-pairing between the tracrRNA and crRNA (Figure 2.2) (Jinek et al., 2012). Additionally, it was observed that Cas9 contained domains that are structurally homologous to both HNH and RuvC endonucleases (Haft et al., 2005; Makarova et al., 2006) and each Cas9 catalytic domain cleaved one strand of a dsDNA (Jinek et al., 2012). Specifically, while the Cas9 HNH domain cleaved the complementary strand, the RuvC-like domain cleaved the non-complementary strand (Jinek et al., 2012).

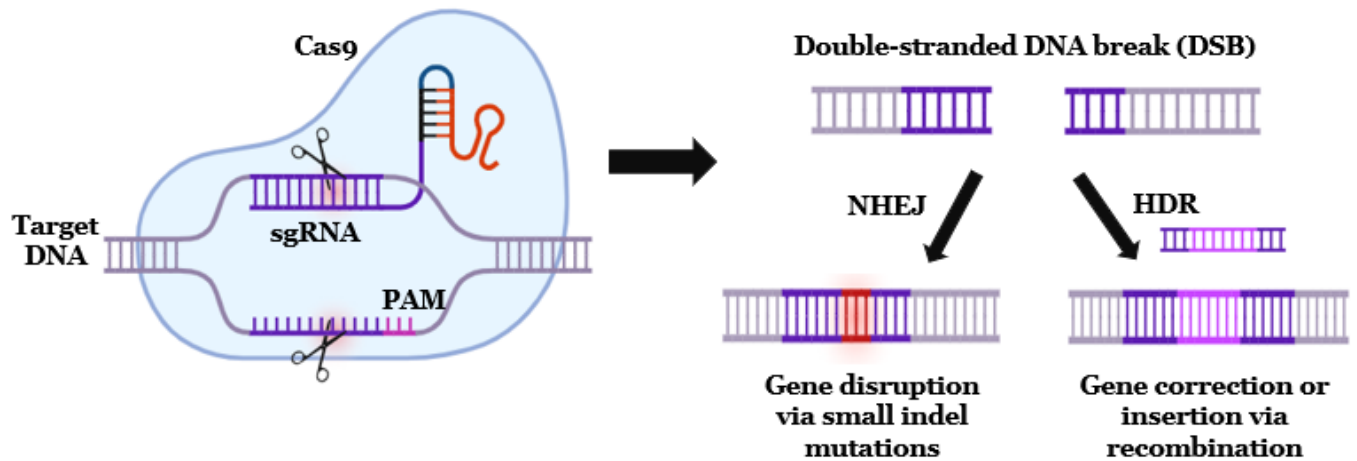


Figure 2.2: CRISPR-Cas9 as a genome engineering tool.

The application of CRISPR-Cas9 in insect cells began and is still predominantly performed with the model organism *Drosophila melanogaster* for both *in vivo* (whole insect) and *in vitro* (cell culture) genetic engineering (Reid, O’Brochta, 2016). For transient expression, the *cas9* gene and sgRNA could be injected via a plasmid, or *in vitro*-transcribed *cas9* mRNA and sgRNA or purified Cas9 protein and sgRNA as an RNP complex (Reid, O’Brochta, 2016), while for stable expression, the *cas9* gene and sgRNA could be integrated into a chromosome (Ren et al., 2013), for *in vivo* genome engineering. For *in vitro* CRISPR-Cas9 gene editing of the *Drosophila melanogaster* S2 cell line, the *cas9* gene and sgRNA were delivered via a plasmid, and indel mutations were induced in 88% of the alleles upon culturing the cells with puromycin (Bassett et al., 2014). Apart from this, CRISPR-Cas9-mediated indel mutations have also been introduced in the *Bombyx mori* derived cell line BmN (Liu et al., 2014). In addition, to promote HDR over NHEJ, CRISPR-Cas9 was used to disrupt factors of the NHEJ pathway in the silkworm (*Bombyx mori*) cells and HR efficiency was observed to be higher than NHEJ in these knocked out cell lines (Ma et al.,

2014; Zhu et al., 2015). Another study performed a genome-wide screening of the *Bombyx mori* BmE cell line using CRISPR-Cas9 to identify essential genes for cell viability and was the first CRISPR screening platform for insect cells other than the model *Drosophila melanogaster* (Chang et al., 2020). Nevertheless, while RNA polymerase II (RNAP II) is required for *cas9* transcription, sgRNA is transcribed by promoters recognized by RNAP III, particularly promoters from U6 and H1 genes, in animal cells (Anderson et al., 2020). Moreover, the U6 promoter for sgRNA expression was found to impact mutation rates and a comparison amongst the promoters U6:1, U6:2, and U6:3 led to the conclusion that U6:3 was the most efficient for mutagenesis in somatic and germ-line cells (Port et al., 2014). Different U6 snRNAs promoters have been recognized in *Drosophila melanogaster* and *Bombyx mori* (Hernandez et al., 2006; Smail et al., 2006); however, no U6 promoters with transcriptional activity had been identified in *S. frugiperda* and *T. ni* until 2017 and this discovery opened the door to plasmid-based CRISPR-Cas9 genome engineering in Sf9 and High-FiveTM cell lines (Mabashi-Asazuma, Jarvis, 2017).

Nonetheless, only limited progress has been made on the implementation of CRISPR-Cas9 genome editing technology in the BEVS (Bruder, 2021). A study aimed to inhibit virus replication used CRISPR-Cas9 gene editing to specifically disrupt the key replication factor *ie-1* in BmNPV using a *Bombyx mori* cell line stably expressing *cas9* and sgRNA, and this is the first work demonstrating CRISPR-Cas9 application in insect antiviral research (Dong et al., 2016). Subsequently, a virus-inducible CRISPR-Cas9 system in transgenic silkworms was established, in which *cas9* expression inhibiting BmNPV replication in infected cells by disrupting the *ATAD3A* gene, was activated upon inoculating the silkworms with BmNPV (Dong et al., 2018). A recent work also established an antiviral

strategy by applying a baculovirus-inducible CRISPR-Cas9 system triggered upon BmNPV infection in transgenic silkworms, for targeting the BmNPV *lef1* and *lef3* genes essential for viral DNA replication, by using the BmNPV 39k promoter for expressing Cas9 and the BmU6 promoter for driving the expression of 4 sgRNAs (Liu et al., 2022). CRISPR-Cas9 was eventually applied to AcMNPV and Sf21 cells for targeted gene disruption in AcMNPV rBEVs by co-transfecting AcMNPV genomic DNA with Cas9/sgRNA RNP complexes, and it is claimed to be the first report establishing the potential of CRISPR-Cas9 for baculovirus gene engineering, which could, in turn, improve the baculovirus as an expression vector and a biopesticide (Pazmiño-Ibarra et al., 2019). Recently, two studies focused on the application of CRISPR-Cas9 to optimize the BEVS for improved biologics production. Specifically, gene disruption and transcriptional repression in AcMNPV via CRISPR-Cas9 were compared, and a sensitive assay capable of efficiently scrutinizing the AcMNPV genome by targeted disruption was developed (Bruder et al., 2021; Bruder, Aucoin, 2023b). Two other studies utilized CRISPR-Cas9 to target the AcMNPV *gp64* or *vp80* genes without the need for generating trans-completing cell lines (Bruder, Aucoin, 2023a; Hausjell et al., 2023). Targeted disruption of these genes resulted in reduced BV production, thus facilitating downstream purification processes (Bruder, Aucoin, 2023a; Hausjell et al., 2023).

2.7 Next-Generation Sequencing

Commercially available AcMNPV DNA backbones for use in the BEVS are typically based on either the C6 or the E2 strains (Maghodia et al., 2014). The AcMNPV C6 strain is

the first baculovirus to be completely sequenced (Ayres et al., 1994). The sequencing data were used to predict AcMNPV's potential coding capacity, viral DNA replication initiation sites, early and late gene transcription regulation, and factors affecting translation efficiency (Ayres et al., 1994). Additionally, structural motifs were identified and functions of some genes were suggested (Ayres et al., 1994). On the other hand, the complete genome sequence of the AcMNPV E2 strain was determined only a decade ago (Maghodia et al., 2014). While the overall E2 sequence was found to be similar to the C6 sequence, it has an additional repeat in one of the hrs (hr2), which mainly contributes to the slightly larger E2 genome size (Cochran, Faulkner, 1983; Maghodia et al., 2014). Variations in the hr4b, several single-nucleotide polymorphisms (SNPs), and other mutations were also detected between the two strains (Maghodia et al., 2014). It was further revealed that the E2 strain aligned to partial C6 sequences instead of the C6 whole-genome sequence in many cases, thus demonstrating that the C6 whole-genome sequence is incorrect at certain locations (Maghodia et al., 2014; Rohrmann, 2019a).

With advances in sequencing methods, ultra-deep sequencing of a naturally occurring AcMNPV population (WP10) detected low-frequency mutations (0.025%) and revealed genomic variations, demonstrating its adaptive potential (Chateigner et al., 2015). This NGS method reported that the WP10 consensus sequence was 99.8% similar to the C6 genome sequence and suggested fusing the adjacent ORFs *AcOrf-20/21*, *AcOrf-58/59*, *AcOrf-106/107*, and *AcOrf-112/113* (Chateigner et al., 2015; Ayres et al., 1994; Rohrmann, 2019a). Additionally, some ORFs (*AcOrf-17*, *AcOrf-52*, *AcOrf-131*, *AcOrf-143*, and *AcOrf-145*) were found to be longer than what was previously reported due to the presence of SNPs and indel mutations (Chateigner et al., 2015; Ayres et al., 1994). Another study

used the Oxford Nanopore Technologies (ONT) MinION long-read sequencing platform to analyze the polyadenylated fraction of AcMNPV transcriptome using cDNA and direct RNA sequencing (Moldován et al., 2018). 132 novel transcripts and transcript isoforms were identified, and a complex meshwork of transcriptional overlaps with unknown function was revealed (Moldován et al., 2018). They also identified complex transcripts overlapping the hr1, thus indicating that transcription and replication machineries possibly interact with each other (Moldován et al., 2018). The ONT long-read sequencing platform was also used to perform a transcriptome-wide analysis of AcMNPV using cDNA, direct RNA, and Cap-selection sequencing methods (Boldogkői et al., 2018). All reads were mapped to the E2 sequence with 95953 cDNA sequencing reads, 2425 direct RNA sequencing reads, and 488847 Cap-selection sequencing reads aligned to the reference genome (Boldogkői et al., 2018). This large dataset can be used for deep analysis of AcMNPV and host cell transcriptomic and epitranscriptomic patterns (Boldogkői et al., 2018).

While genome sequencing has become a powerful tool for studying infectious diseases, metagenomic sequencing directly from clinical samples without isolation may result in insufficient viral reads (Quick et al., 2017). A comprehensive method, including an on-line primer design tool ‘Primal Scheme’, multiplex PCR for targeted enrichment of viral genomes, library preparation methods for ONT MinION and Illumina MiSeq platforms, and a consensus sequence generating pipeline, was developed to study Zika and other viral genomes directly from clinical samples (Figure 2.3) (Quick et al., 2017). This tiled-amplicon sequencing approach uses primer pools to amplify alternate regions separately to avoid overlapping of adjacent amplicons within a primer pool and preferentially generating short overlap amplicons covering the whole viral genome (Quick et al., 2017). High

coverage of Zika virus genomes from all clinical samples was obtained, thus demonstrating the sensitivity to generate near-complete genomes without isolation (Quick et al., 2017). In a subsequent study, an amplicon-based sequencing assay was paired with a bioinformatics pipeline to measure intrahost virus diversity using ONT and Illumina instruments (Grubaugh et al., 2019). A computational tool, iVar (intrahost variant analysis of replicates), was developed to specifically analyze viral amplicon-based sequencing data, and it was found that each sample should be prepared as technical duplicates and sequenced at a minimum depth of $400\times$ to confidently call variants in a sample (Grubaugh et al., 2019).

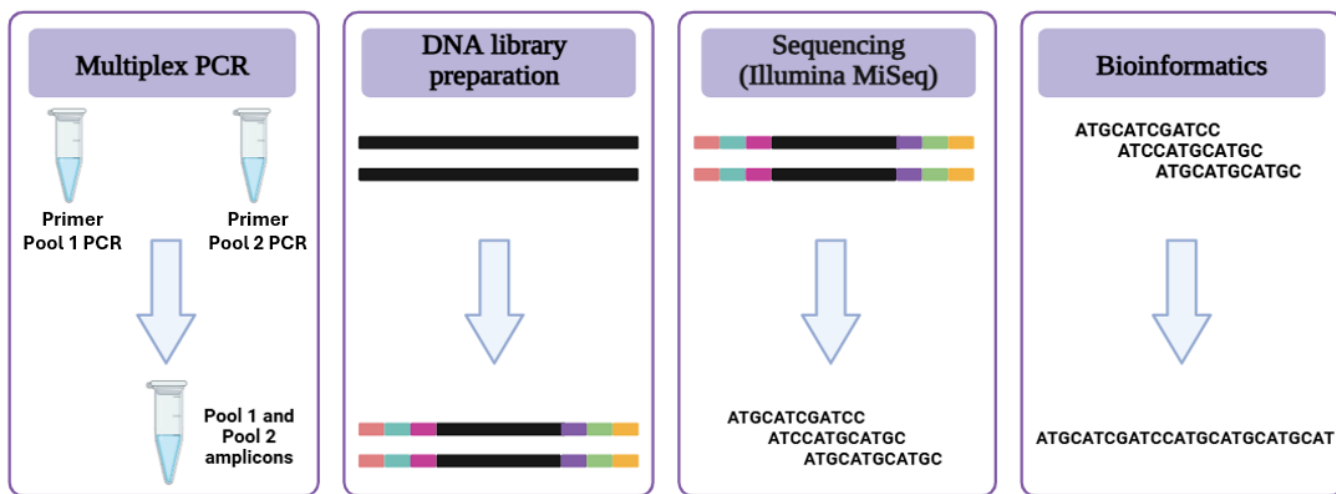


Figure 2.3: Tiled-amplicon sequencing assay workflow.

Chapter 3

Probing late and very late AcMNPV genes using CRISPR-Cas9

The BEVS platform typically employs the widely studied AcMNPV baculovirus for producing biologics in insect cell cultures. The AcMNPV has a large 134 kbp dsDNA genome coding for around 155 ORFs (Ayres et al., 1994; Miele et al., 2011). In cell cultures, baculoviruses exist in the BV form, which is responsible for transmission between cells (Jehle et al., 2006; Rohrmann, 2019c). It is believed that some baculovirus core genes and ODV-associated genes could be non-essential for BV and foreign protein production in cell culture (Rohrmann, 2019a; Chen et al., 2021). While some studies reported on AcMNPV gene essentiality, the primary focus was on BV production or foreign gene expression from very late promoters (Chen et al., 2021; Yu et al., 2023; Zhang et al., 2023). Here, we used a CRISPR-Cas9-based T-I assay to probe late and very late AcMNPV genes for their essentiality when the foreign gene was driven by a late promoter instead of the typically

used very late ones (Figure 3.1). The idea was to shift production to an earlier point in the infection cycle when the cells are healthier, and instead of using the cellular resources to produce viral proteins, they could be used to produce products of interest. In this study, 38 AcMNPV genes were selected based on the presence of a 5'-TAAG-3' promoter motif. It has been established in the literature that this promoter motif is associated with late and very late baculovirus gene transcription (Chen et al., 2013). Based on the effect of 38 endogenous AcMNPV gene disruptions on GFP expression from the late p6.9 promoter and BV production, we classified them as essential and of special interest. While 19 genes were found to be essential (reducing both GFP and BV), 19 others were identified as of special interest (reducing only GFP, with BV being enhanced or similar to the control). The research presented in this chapter has been formatted for submission.

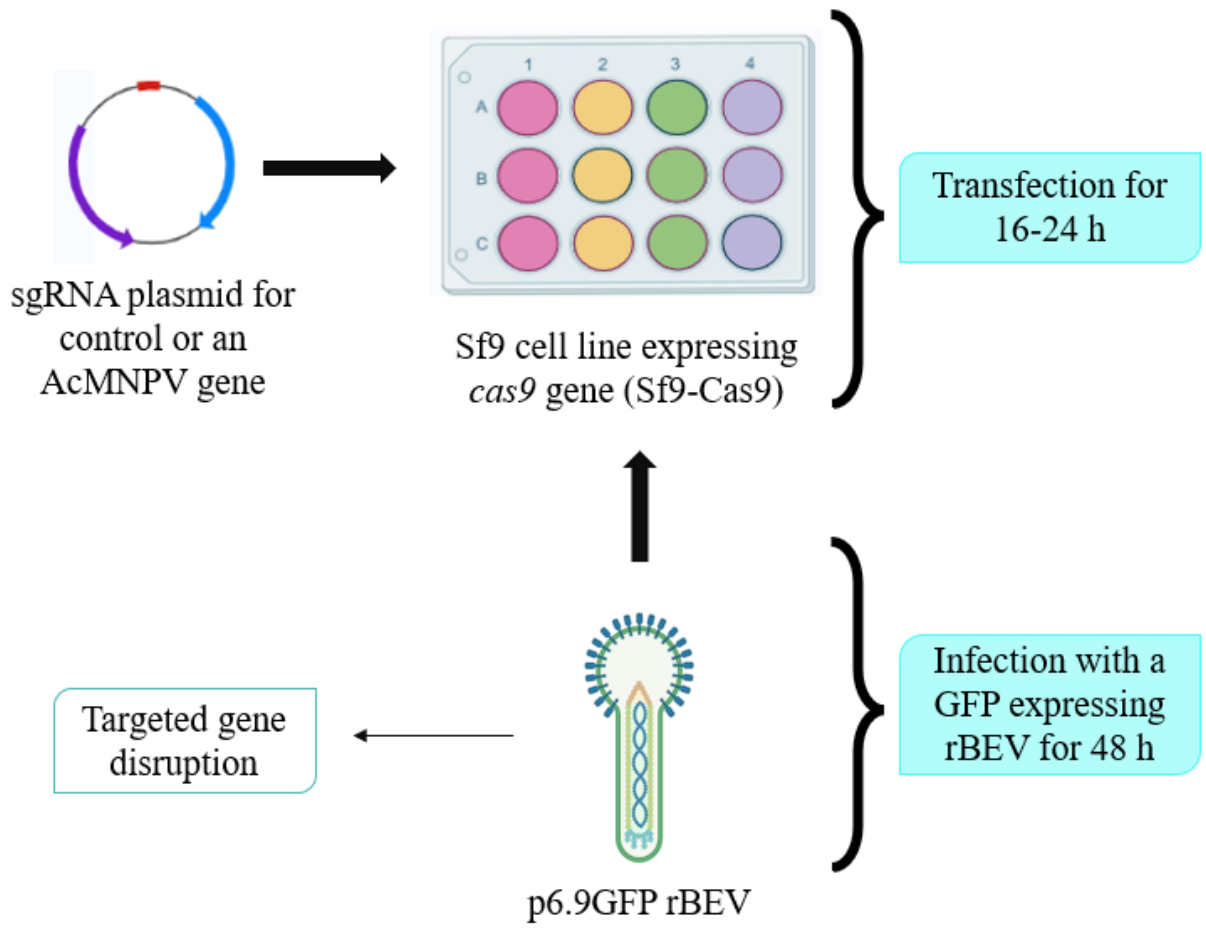


Figure 3.1: Overview of transfection-infection assay.

Probing late and very late AcMNPV baculovirus genes for essentiality in insect cell culture

Madhuja Chakraborty ¹ and Marc G. Aucoin ¹

¹ Department of Chemical Engineering, University of Waterloo

3.1 Abstract

Since the development of the baculovirus expression vector system (BEVS) in the 1980s, it has been widely used to produce biologics in insect cell cultures. The most commonly employed baculovirus, *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV), serves as a recombinant expression vector (rBEV) in BEVS. Prior studies have shown that deleting select sequences from the large AcMNPV genome can enhance budded virus (BV) and foreign protein production. However, these efforts have largely focused on BV titers or gene expression driven by very late promoters. In this study, we hypothesized that removing late/very late AcMNPV genes—specifically those not required for BV or foreign protein production—and using late promoters to drive transgene expression could extend production time and improve product yield and quality. We systematically screened 38 late/very late AcMNPV genes using a CRISPR-Cas9-based transfection-infection assay. Sf9-Cas9 cells were transfected with sgRNA plasmids and subsequently infected with a rBEV expressing a reporter gene under the late p6.9 promoter. The impact of each gene disruption on progeny virus and foreign protein production was assessed. Our screen

identified 19 essential genes (disruption reduced BV and foreign protein production) and 19 genes of special interest (disruption reduced foreign protein only). These findings support the development of a minimal genome rBEV. In particular, genes of special interest may be disrupted in the production of a minimal baculovirus genome destined for BacMam applications.

3.2 Introduction

Baculoviruses are insect-specific, enveloped double-stranded DNA (dsDNA) viruses with rod-shaped nucleocapsids (Jehle et al., 2006). Occlusion-derived virus (ODV) and budded virus (BV) are the two distinct baculovirus phenotypes, which are required for transmission between hosts and in cell culture, respectively (Jehle et al., 2006; Rohrmann, 2019c). *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) with a 134 kbp genome and 155 open reading frames (ORFs) is the most extensively-studied baculovirus (Ayles et al., 1994; Miele et al., 2011; Rohrmann, 2019a). The baculovirus expression vector system (BEVS), typically comprising of recombinant AcMNPV and insect cell culture, is a suitable production platform for complex proteins requiring mammalian-like post-translational modifications and has been used to produce numerous biologics of human and veterinary interest (Zhang et al., 2023; Yee et al., 2018; Felberbaum, 2015; Airene et al., 2013; van Oers et al., 2015).

The functions of many AcMNPV genes remain uncharacterized, yet baculovirus genome engineering presents a promising strategy to enhance recombinant baculovirus expression vectors (rBEVs) (Vijayachandran et al., 2013; Martínez-solís et al., 2019). While core

baculovirus genes are generally considered indispensable, some may be non-essential for BV production in insect cell culture (Chen et al., 2021). In addition, certain genes associated with ODV formation may also be dispensable for BV and recombinant protein production (Rohrmann, 2019a). To improve the BEVS platform, several studies have focused on deleting non-essential viral genes—such as *v-cath*, *chiA*, *p26*, *p10*, and *p74*—which has led to enhanced recombinant protein yield and stability (Kaba et al., 2004; Hitchman et al., 2010a). van Oers et al. (2015) proposed that constructing minimized baculovirus genomes, by eliminating unnecessary viral genes, could further improve production yields, allow for multiple gene insertions, and enhance overall product quality and safety (van Oers et al., 2015).

A previous study analyzed the effect of 42 AcMNPV genes on infectious BV production by creating knockout viruses (KOVs) and performing transfection and infection assays (Chen et al., 2021). They categorized 36 genes as dispensable, 3 as important, and 3 as essential, as well as proposed a minimal AcMNPV genome map with the dispensable genes and the majority of the homologous repeat regions removed (Chen et al., 2021). Another study re-validated 80 AcMNPV genes that are known to be dispensable for BV production (Yu et al., 2023). They constructed KOVs to analyze their impact on insect cell infection efficiency via fluorescent particles, identified 62 non-essential and 18 influential genes, and designed a theoretical minimal AcMNPV genome (Yu et al., 2023). It is to be noted that the different categorizations used in these studies were based on the effect of AcMNPV gene deletion on BV production in cell culture; however, their impact on foreign protein production was not investigated (Chen et al., 2021; Yu et al., 2023). A more recent study focused on the effect of KOVs on BV and foreign protein production from the very late

p10 promoter (Zhang et al., 2023). 14 KOVs were constructed by deleting DNA fragments consisting of at least two contiguous genes to analyze the impact of fragment deletion on viral propagation and exogenous gene expression (Zhang et al., 2023). 4 DNA fragments, *AcOrf-11* to *AcOrf-13*, *AcOrf-18* to *AcOrf-23*, *AcOrf-55* to *AcOrf-61*, and *AcOrf-114* to *AcOrf-122* were found to be essential or influential, while the rest were categorized as dispensable (Zhang et al., 2023). Nonetheless, the impact of individual AcMNPV gene deletion on foreign protein production was not examined, which might be necessary for efficiently creating minimal baculovirus genome(s) for use as a vector.

We hypothesized that AcMNPV contains late and very late genes that are not required for foreign protein or progeny virus production in cell culture, thus leading to a waste of cellular resources. A reduced rBEV genome expressing a gene of interest under a late promoter instead of the commonly used very late promoters would shift production to an earlier point in the infection cycle, when cells are healthier, and enhance biologics yield. In this study, we used a previously developed CRISPR-Cas9-based transfection-infection assay to screen AcMNPV genes without generating KOVs (Bruder, Aucoin, 2023b). Late and very late AcMNPV genes were disrupted to analyze their impact on BV levels and green fluorescent protein (GFP) production from the late p6.9 promoter. Briefly, *cas9* expressing Sf9 cells (Bruder et al., 2021) were transfected with sgRNA plasmids and then infected with a p6.9GFP rBEV for late exogenous gene expression during AcMNPV gene disruption. 2–3 sgRNAs were assessed for the disruption of each AcMNPV gene target. Upon CRISPR-Cas9-mediated disruption of 38 endogenous AcMNPV genes, they were divided into two major categories — essential and of special interest, based on the effect on late foreign protein and progeny virus production. Specifically, genes whose disruption

reduced both foreign protein and BV production were classified as essential, whereas genes whose disruption reduced only foreign protein but not BV production were classified as of special interest.

3.3 Materials and Methods

3.3.1 Insect Cell Lines and Maintenance

Previously developed transgenic Sf9-Cas9 cells carrying a Cas9-2A-puromycinR cassette ([Bruder et al., 2021](#)) and parental Sf9 cells were routinely maintained in Sf-900TM III serum-free media (SFM) (Gibco, Carlsbad, CA, USA) at 27 °C and 130 rpm. The cells were maintained in their exponential phase by passaging them every 3 to 4 days. To maintain *cas9* expression, 5 µg/mL puromycin was added to Sf9-Cas9 cells in alternating passages.

3.3.2 sgRNA Plasmid Design and Construction

All primers used in this study were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA). Control and re-targeting sgRNA plasmids were designed using the NEBuilder Assembly Tool and constructed using the NEBuilder 2× HiFi DNA assembly master mix (New England Biolabs, Whitby, ON) as previously described ([Chakraborty et al., 2024](#)). Briefly, to construct the pSfU6-sgRNA scrambled control plasmid, a backbone fragment and an insert were assembled in a two-fragment Gibson assembly reaction. While the backbone fragment was obtained by PCR-amplifying the ampicillin resistance gene

(ampR) and the pBR322 origin of replication (ori) from the pBR322-TIMER plasmid (Addgene # 103056) (Claudi et al., 2014), the SfU6-sgRNA insert was obtained by fusion PCR of two PCR-amplified fragments — an SfU6 promoter (Mabashi-Asazuma, Jarvis, 2017) gBlock gene fragment (synthesized by IDT) and a sgRNA fragment with a scrambled spacer sequence, a gRNA scaffold, and a transcriptional terminator at the 3' end (Addgene # 49411) (Port et al., 2014).

To construct re-targeting sgRNA plasmids, the pSfU6-sgRNA scrambled control plasmid was used as the template for inverse PCR (Bruder, Aucoin, 2023b). Entire re-targeting sgRNA plasmids were amplified as linear fragments using primers annealing either to the 3' end of the SfU6 promoter or the 5' end of the gRNA scaffold and altered spacer sequences, targeting a specific AcMNPV gene, appended to their 5' ends. The template DNA was then removed by DpnI restriction digestion, and the gel-extracted re-targeting sgRNA fragments were re-circularized by single-fragment Gibson assembly reactions. Spacer sequences targeting AcMNPV genes were designed using “CHOPCHOP” (<https://chopchop.cbu.uib.no/>) (Labun et al., 2019). Sequence selection criteria and primers have previously been reported (Chakraborty et al., 2024). Spacer sequences are listed in Supplementary Table A.1. Lastly, the mKate2 plasmid used in our control experiments consisted of the ampR and pBR322 ori backbone, along with *mKate2* in the multiple cloning site under the OpIE2 promoter. All Gibson assembled products were transformed into NEB 10- β chemically competent cells using a previously described heat shock protocol (Chakraborty et al., 2024). Transformants (transformed single colonies) were grown overnight in Terrific Broth with 100 μ g/mL ampicillin, and plasmids were extracted using the GeneJETTM Plasmid Midiprep Kit (Thermo Fisher, Mississauga, ON,

Canada) according to the manufacturer’s instructions.

3.3.3 Recombinant Baculovirus Expression Vector Amplification

In this study, we used a previously constructed rBEV carrying a GFP (monomeric Azami green) under the AcMNPV late p6.9 promoter (herein referred to as p6.9GFP rBEV) (Bruder, Aucoin, 2023b). The p6.9GFP rBEV was amplified at an MOI of ~ 0.1 pfu/cell in Sf9 cells and harvested when cell viability decreased between 80–90%. The resulting working virus stock was used for all subsequent experiments.

3.3.4 Quantification of Infectious Recombinant Virus Vector

An end-point dilution assay (EPDA) was used to quantify infectious virus titers (IVTs) as previously described (Reed, Muench, 1938; O’Reilly et al., 1992; Chakraborty et al., 2024). Briefly, 100 μ L of Sf9 cells diluted to 2×10^5 cells/mL in Sf-900TM III SFM were seeded to each well of 96-well tissue-culture treated plates (VWR International, Mississauga, ON, Canada). The plates were incubated at 27 °C for ~ 1 hour during which virus stock dilutions from 10^{-2} to 10^{-8} were prepared in Sf-900TM III SFM. 10 μ L of a dilution of a virus stock was added to each row of a 96-well plate. Virus-inoculated plates were incubated at 27 °C for 6–7 days, following which the wells were scored using a fluorescence microscope to determine which wells exhibited fluorescence. Finally, IVTs were reported as plaque-forming units per mL (pfu/mL) by multiplying the reciprocal of TCID₅₀ (expressed in mL of virus added) by 0.68 according to the Poisson distribution.

3.3.5 CRISPR-Cas9-Based Transfection-Infection Assay

A CRISPR-Cas9-based transfection-infection assay (T-I assay) was developed previously by [Bruder, Aucoin \(2023b\)](#) and used in this study. Briefly, tissue-culture-treated 12-well plates (VWR International, Mississauga, ON, Canada) were seeded with Sf9-Cas9 cells at a density of 0.9×10^6 cells/well or 1.12×10^6 cells/mL. Following ~ 1 hour incubation at 27°C for cell attachment, the cells were transfected with the scrambled control plasmid (n=3), the mKate2 plasmid (n=6), or re-targeting sgRNA plasmids (n=3 for each sgRNA plasmid) using FuGENE HD (Promega, Madison, WI, USA) transfection reagent according to the manufacturer's instructions. The transfection conditions for the T-I assay were optimized previously ([Bruder, Aucoin, 2023b](#)) and used here (transfection reagent:DNA 4:1 and DNA:cell 1 pg). Another 3 wells seeded with Sf9-Cas9 cells were left untransfected. After 16–24 hours post-transfection, the media was aspirated from each well, except 3 wells transfected with the mKate2 plasmid. Fresh Sf-900TM III SFM containing the p6.9GFP rBEV working virus stock was added to the wells to achieve an infection at an MOI of 3 pfu/cell. This yielded 3 replicates of the transfected+infected samples for the scrambled control, mKate2 control, and each re-targeting sgRNA plasmid; 3 replicates of the infected-only control; and 3 replicates of the transfected-only control. Finally, the cell cultures were harvested ~ 48 hours post-infection (hpi) by centrifugation at $800\times g$ for 15 minutes. The cell pellets were analyzed by flow cytometry, and the supernatants were stored at 4°C for further analysis by EPDA.

3.3.6 Flow Cytometry Analysis for Foreign Protein Production

To prepare the cell pellets for flow cytometry, they were first resuspended in fresh 2% paraformaldehyde prepared in phosphate-buffered saline (PBS). Subsequently, they were incubated at 4 °C for ~30 minutes, and then diluted in 1× PBS before running on a BD Accuri™ C6 Plus flow cytometer (BD Biosciences, San Jose, CA, USA). The cytometer was equipped with a 488 nm excitation frequency blue laser, and a non-standard 510/15 nm band-pass filter was placed in FL1 to capture the emission peak of the GFP used in this work. A low flow rate of 14 µL/min was used to run the samples, and 10000 events were collected for the analysis of each sample. Using licensed software, FlowJo™ V10 (Tree Star, Ashland, OR, USA), the acquired flow cytometry data was analyzed. Gates were applied to remove cell debris and to bin the observed fluorescence intensities. Specifically, a histogram of the gated population excluding the cell debris was used to bin the fluorescence intensities into high (FL1-H $\geq 10^6$ au) and low (FL1-H $< 10^6$ au) fluorescence bins. Finally, after applying these bins to all samples, they were exported to analyze and visualize the results further using the R programming language.

3.3.7 Statistical Analysis

In this study, data from three biological replicates of a sample are presented as the mean \pm 95% confidence interval. ANOVA with Dunnett's test was performed to examine significant differences between the scrambled control and each targeted sample. Dunnett's test is a post-hoc tool that compares multiple test groups to a single control group, rather than performing all possible pairwise comparisons. It is a conservative test that controls for type

I error (false positives). Since the aim was to compare the changes upon gene disruption to the scrambled control only, Dunnett's test was chosen as the statistical test.

3.4 Results

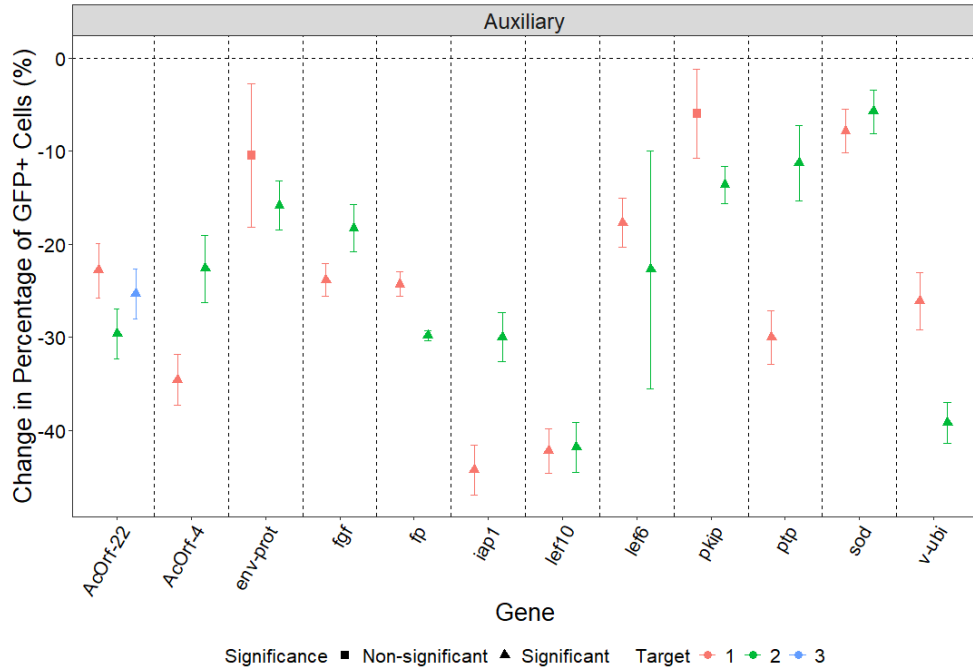
3.4.1 Investigation of Controls for T-I Assay

Four controls were investigated to phenotypically verify that gene disruption was due to targeted CRISPR-Cas9 genome editing. The scrambled control consisted of transfection with a plasmid carrying a spacer sequence designed not to target any p6.9GFP rBEV gene (scrambled targeting sequence), followed by infection with p6.9GFP rBEV. The mKate2 control involved transfection with an mKate2 plasmid not carrying any sgRNA, followed by infection with the p6.9GFP rBEV. The p6.9GFP rBEV infected-only control was used to assess whether transfection had a negative effect on foreign protein production. Finally, the transfected-only control, comprising of transfection with an mKate2 plasmid, was to verify the efficiency of transfection (Figure A.1a). From the population distribution of the transfected-only control, as captured by the FL3-H detector, it can be said that a minimum of 70–80% transfection efficiency was achieved (Figure A.1a). The green fluorescence intensity distributions of the scrambled, mKate2, and infected-only controls overlapped with each other (Figure A.1b). Additionally, the percentage of GFP-expressing cells in the high and low fluorescence bins was similar for all controls infected with the p6.9GFP rBEV (Figure A.1c). Based on these results, the scrambled control was chosen as the most appropriate negative control.

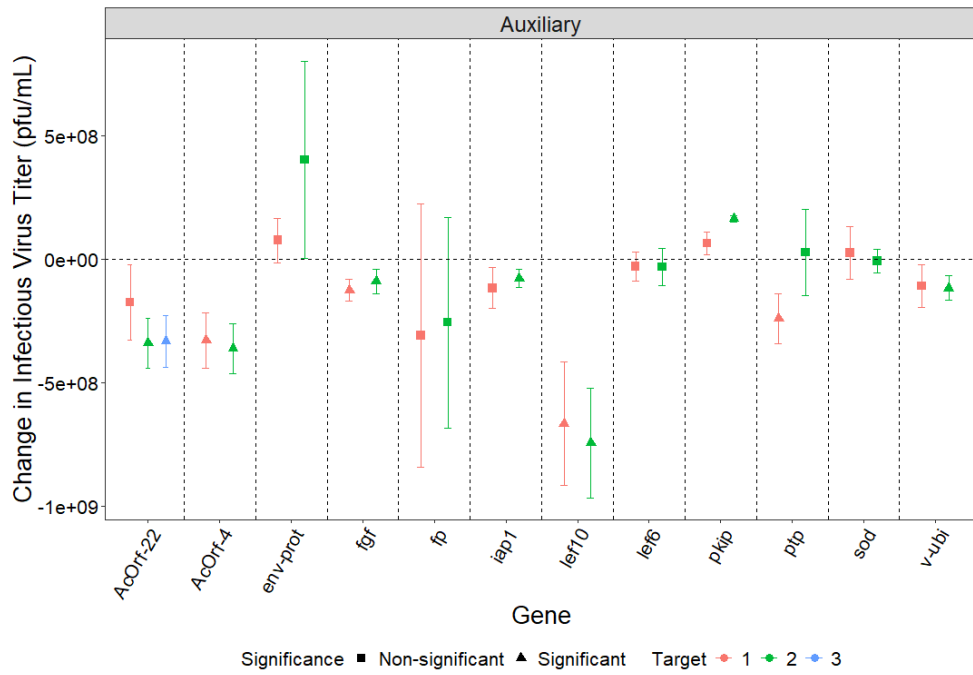
3.4.2 CRISPR-Cas9-Mediated AcMNPV Gene Disruption

In this study, 38 endogenous AcMNPV genes were targeted to study their effect on BV titers and foreign protein production from the late p6.9 promoter. To avoid bias, genes were distributed across different T-I assay runs without consideration of their known or predicted functions. For statistical purposes, each targeted gene was compared to the control within its respective experimental run to account for run-to-run variability and accurately determine its effect. Following the classification by [Cohen et al. \(2009\)](#), the 38 AcMNPV genes were grouped into five functional categories: auxiliary, DNA replication, transcription, structural, and unknown function, allowing visualization of the effects of targeted disruption across gene functions.

For each gene disruption, 2–3 sgRNAs were assessed to gain confidence in the observed effects. Targets 1–3 represent the different sgRNA targets for each gene, with increasing distance from the 5' end of a gene (Table [A.1](#)). Of the twelve auxiliary genes that were disrupted, five led to reductions in both GFP expression and BV production, while seven resulted in decreased GFP expression alone (Figure [3.2](#)). Disruption of DNA replication genes revealed that targeting *gta* and *lef1* lowered GFP levels, whereas disruption of *lef2* reduced both GFP expression and BV production (Figure [3.3](#)). Among the nine transcription genes targeted, five disruptions impaired both GFP expression and BV production, while four selectively decreased GFP expression only (Figure [3.4](#)). Disruption of five structural genes identified three that affected both GFP and BV production, and two that reduced only GFP expression (Figure [3.5](#)). Finally, of the nine unknown function genes, five decreased both GFP and BV levels, while four reduced GFP expression alone (Figure [3.6](#)).

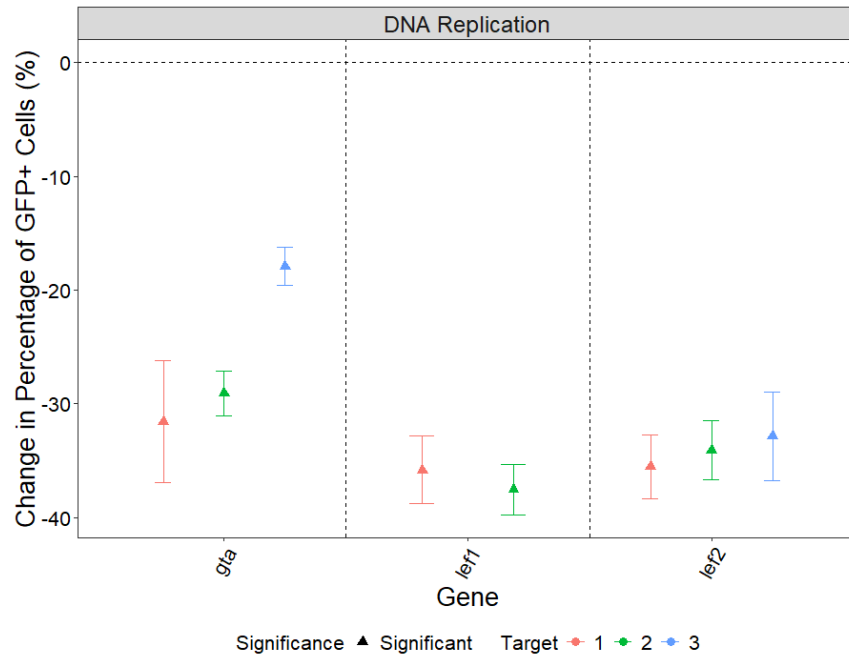


(a) GFP production

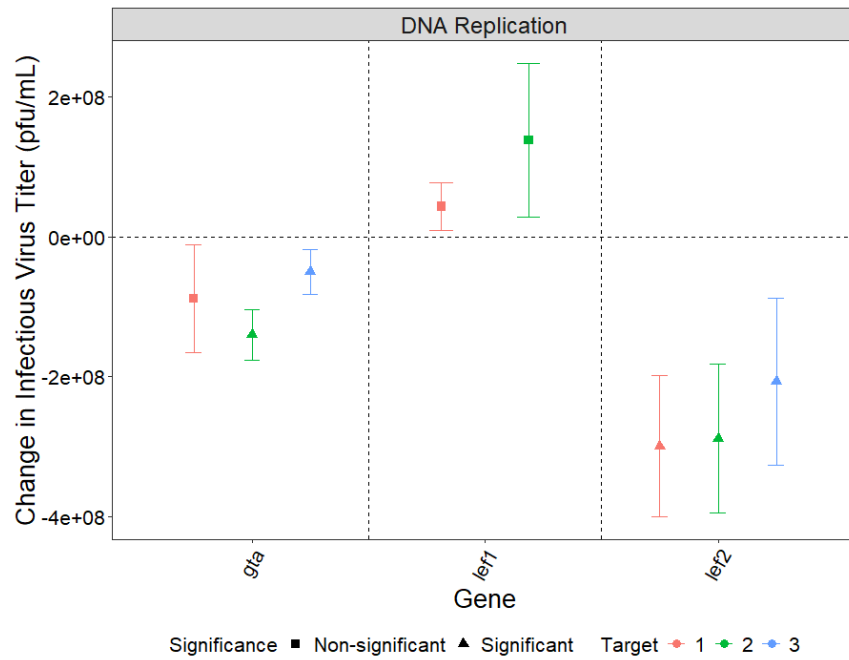


(b) IVT production

Figure 3.2: Impact of disrupting auxiliary AcMNPV genes on (a) foreign protein and (b) progeny virus production.

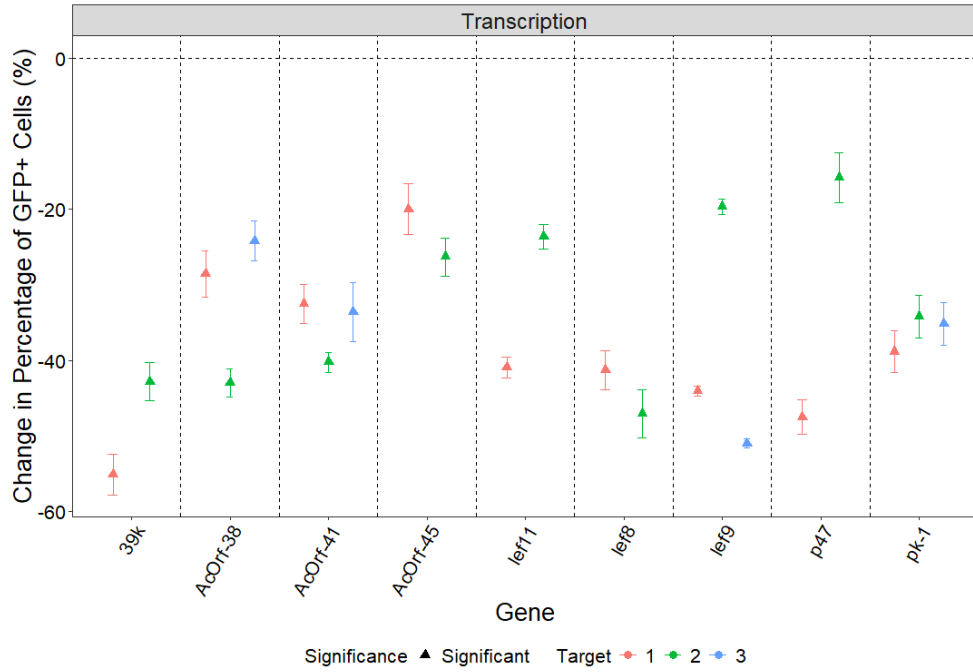


(a) GFP production

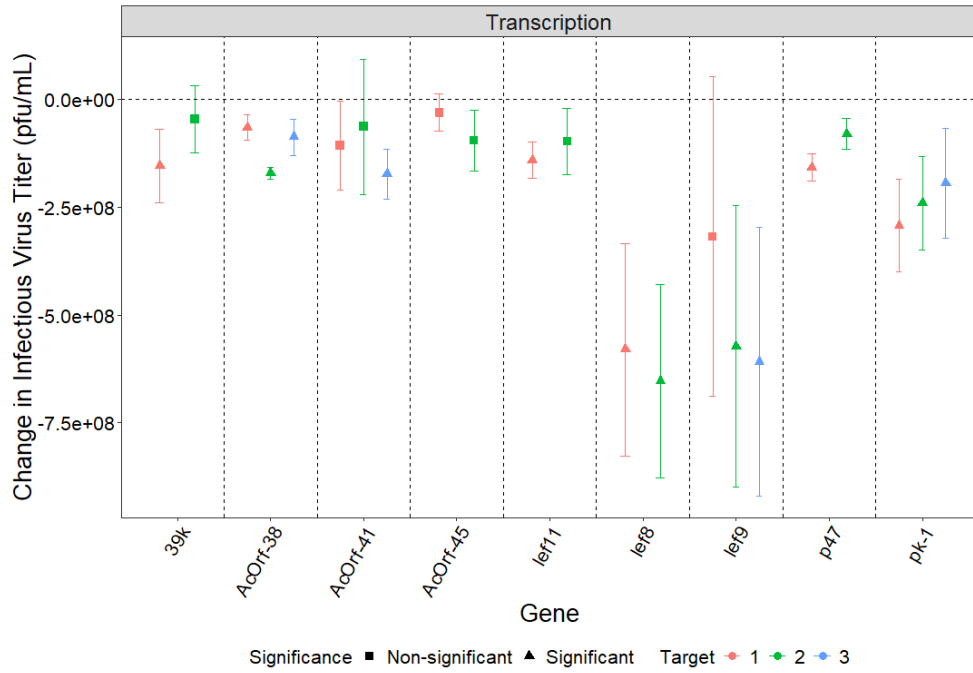


(b) IVT production

Figure 3.3: Effect of targeting DNA replication AcMNPV genes on (a) GFP and (b) IVT.

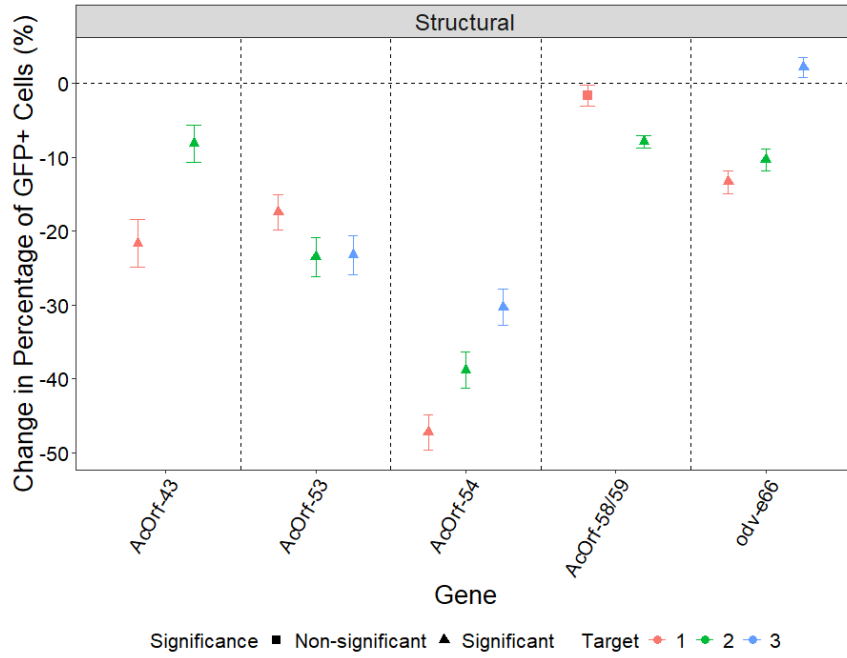


(a) GFP production

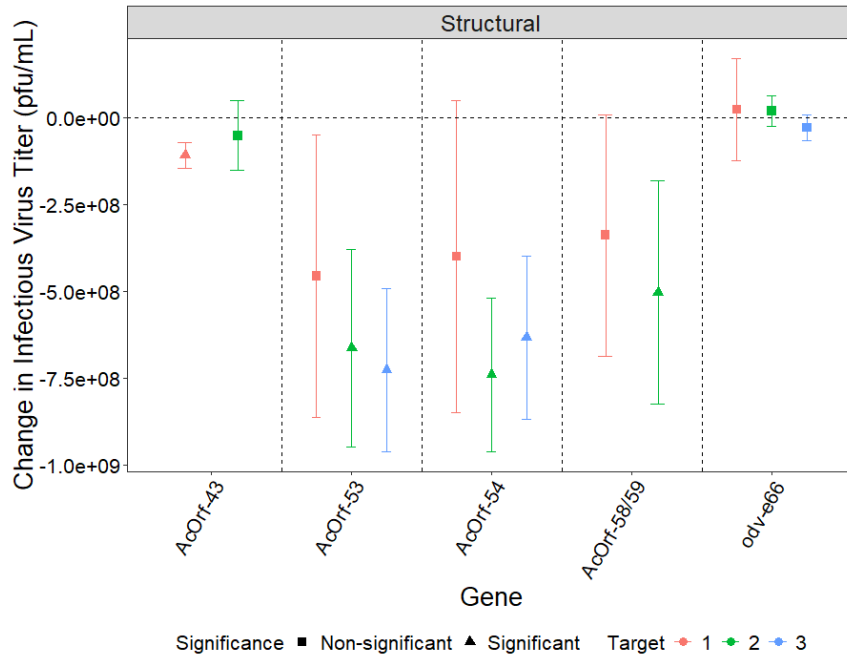


(b) IVT production

Figure 3.4: Effect of transcription AcMNPV gene disruptions on (a) foreign protein and (b) BV production.

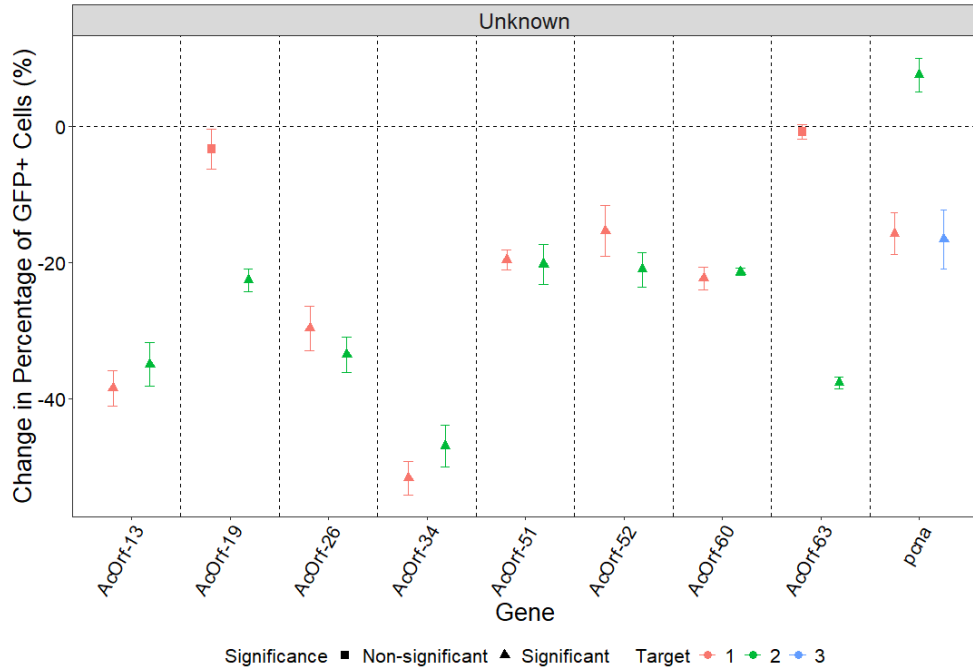


(a) GFP production

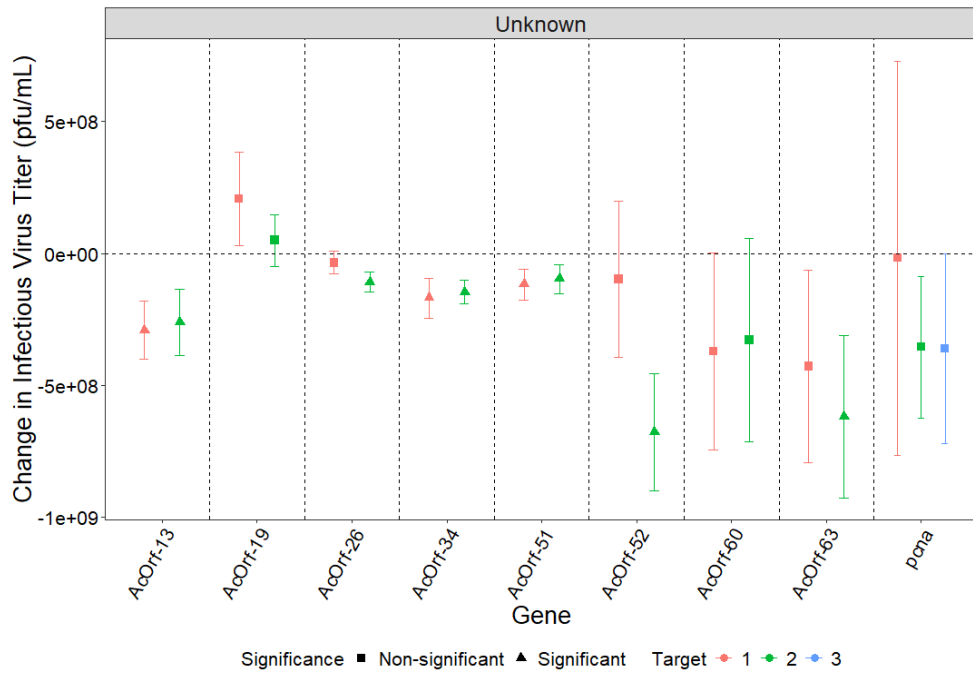


(b) IVT production

Figure 3.5: Impact of structural AcMNPV gene disruptions on (a) GFP and (b) IVT.



(a) GFP production



(b) IVT production

Figure 3.6: Effect of targeting unknown function AcMNPV genes on (a) foreign protein and (b) progeny virus production.

3.5 Discussion

Based on the presence of a 5'-TAAG-3' promoter motif in the baculovirus genome (Chen et al., 2013), indicative of gene expression in the late and very late phases of the infection cycle, 38 AcMNPV genes were chosen for disruption using a CRISPR-Cas9-based T-I assay (Bruder, Aucoin, 2023b). This T-I screening assay was used to evaluate the effect of endogenous AcMNPV gene disruption on the production of infectious baculovirus and foreign protein under a late promoter. Specifically, the aim was to identify late/very late AcMNPV genes whose disruption either reduced or increased (or remained neutral for) virus titers and/or foreign protein production from the late p6.9 promoter instead of the commonly used very late polh or p10 promoters thereby extending the production time by shifting it to an earlier phase of the infection cycle.

To ensure the robustness of our results, we included four controls in the T-I assay: scrambled, mKate2, infected-only, and transfected-only. The transfected-only control involving transfection with a plasmid encoding the *mKate2* gene under the OpIE2 promoter, confirmed that >70% of cells received the plasmid under assay conditions. All control groups infected with the p6.9GFP rBEV exhibited comparable levels of foreign protein expression, indicating that neither the transfection reagent nor the presence of a scrambled sgRNA adversely affected GFP production.

For statistical comparisons, the scrambled sgRNA condition was designated as the reference group. In this control, a plasmid encoding a scrambled sgRNA (based on a randomized version of a known targeting sequence, computationally verified to lack off-target sites in the baculovirus genome) was transfected into cells prior to p6.9GFP rBEV

infection. To assess statistical significance, we calculated p-values using a Dunnett’s test that is specifically designed for comparing multiple treatments against a common control and adjusts the significance threshold to control the family-wise error rate at a specified level (e.g., $\alpha = 0.05$). This conservative approach ensures high confidence in any identified differences.

As many as three different target sites were designed for each gene to examine the effect of the gene disruption. The locations of the sgRNA targets for each endogenous AcMNPV gene are in reference to the 5’ end of an ORF, with Target 1 being closest to the 5’ end and Target 3 being the farthest. We hypothesized that the greatest effects would be seen with those targets that were closest to the 5’ end and/or had the highest predicted targeting efficiency (as calculated by CHOPCHOP). The increasing distance from the N-terminal may not retain a functional domain, and thus, an effect might not be observed. Moreover, the sgRNA targets with higher targeting efficiencies were expected to induce a greater effect, irrespective of the location, as long as it was within a functional domain. We have collected some evidence to support these theories (Figure [A.2](#), [A.3](#)), albeit these trends are far from universal.

Table [3.1](#) gives an overview of the impact of gene disruption on BV production and exogenous protein production from the late p6.9 promoter. Based on the observed effects upon CRISPR-Cas9-mediated disruption, the AcMNPV genes were classified into two major types — essential (negative effect on BV and GFP) and of special interest (negative effect on GFP and similar effect on BV).

Table 3.1: Classification of AcMNPV gene disruptions by their effects on GFP (rows) and BV (columns). Each gene appears on a separate line within its cell.

GFP ↓ / BV →	Negative	Similar	Positive
Negative	AcOrf-4 AcOrf-13 AcOrf-22 AcOrf-34 AcOrf-38 AcOrf-51 AcOrf-52 AcOrf-53 AcOrf-54 AcOrf-58/59 AcOrf-63 lef2 lef8 lef9 lef10 ptp p47 pk-1 fgf	39k AcOrf-19 AcOrf-26 AcOrf-41 AcOrf-43 AcOrf-45 AcOrf-60 env-prot fp gta iap1 lef1 lef6 lef11 odv-e66 pcna sod v-ubi pkip	(none)
Similar	(none)	(none)	(none)
Positive	(none)	(none)	(none)

3.5.1 Auxiliary AcMNPV Genes

Twelve auxiliary genes (*AcOrf-22*, *AcOrf-4*, *env-prot*, *fgf*, *fp*, *iap1*, *lef10*, *lef6*, *pkip*, *ptp*, *sod*, and *v-ubi*) were disrupted to analyze their effect on IVT and GFP production from the late p6.9 promoter. Based on the effect, these twelve auxiliary AcMNPV genes were

categorized as essential or of special interest, as their disruption reduced BV and GFP production or did not reduce IVT while lowering GFP production from the p6.9 promoter in Sf9-Cas9 cell cultures, respectively.

3.5.1.1 Essential Auxiliary Genes

AcOrf-22 is a per os infectivity factor 2 gene present in all baculovirus genomes (Ohkawa et al., 2005; Rohrmann, 2019a). It has been previously shown to be essential for ODV but not BV production in cell culture (Ohkawa et al., 2005; Yu et al., 2023); however, we observed significant reductions in BV levels upon *AcOrf-22* disruption. Moreover, deletion of a gene fragment containing *AcOrf-22* resulted in reduced GFP expression from the very late p10 promoter (Zhang et al., 2023), which aligns with our observations when GFP is expressed from the late p6.9 promoter. *AcOrf-4* is identified as one of the six nuclear localization of actin (NLA) genes; however, its deletion has been shown not to impact actin translocation kinetics (Ohkawa et al., 2002; Gandhi et al., 2012). Deletion of *AcOrf-4* or of a DNA fragment containing *AcOrf-4* did not negatively impact BV production or GFP expression from the very late p10 promoter, respectively (Yu et al., 2023; Zhang et al., 2023). However, in this study, we have reported that *AcOrf-4* disruption significantly decreased IVT and GFP expression from the p6.9 promoter. The AcMNPV *fgf* gene codes for the fibroblast growth factor secreted protein that stimulates insect cell motility (Ayres et al., 1994; Detvisitsakun et al., 2005). Previous studies have demonstrated that deletion of *fgf* or of a gene fragment containing *fgf* is dispensable for BV production and GFP expression from the very late p10 promoter in cell culture (Detvisitsakun et al., 2006; Yu et al., 2023; Zhang et al., 2023). In contrast, our data indicate that targeting *fgf*

significantly reduces BV and GFP production from the late p6.9 promoter.

lef10 codes for the late expression factor 10 protein with homologs found in all Group I baculovirus genomes (Ayres et al., 1994; Rohrmann, 2019a). It is involved in viral late gene expression and is required for BV production in cell culture (Lu, Miller, 1995; Xu et al., 2016). Our data suggests that *lef10* disruption leads to a significant drop in both IVT and GFP levels. The AcMNPV *ptp* gene encoding a protein tyrosine phosphatase was found to be associated with BV phenotype (Ayres et al., 1994; Rohrmann, 2019a; Wang et al., 2010). It has RNA 5'-phosphatase activity that removes phosphate from the 5'-triphosphate end of RNA (Takagi et al., 1998). Previous studies have reported that deletion of DNA fragments containing *ptp* is dispensable for BV production and GFP expression from the very late p10 promoter (Yu et al., 2023; Zhang et al., 2023). However, our data indicate that disrupting *ptp* significantly decreases IVT levels and GFP production from the late p6.9 promoter.

3.5.1.2 Of Special Interest Auxiliary Genes

The *env-prot* gene codes for the envelope fusion (F) protein, which is not active in AcMNPV (Rohrmann, 2019a). While it is associated with budded viruses, its deletion does not impact infectious virion production in cell culture (Lung et al., 2003), which aligns with our observations. According to a previous study, deletion of a gene fragment containing *env-prot* resulted in reduced GFP expression from the very late p10 promoter; however, the effect of removing only *env-prot* was not investigated (Zhang et al., 2023). Our data indicates the same; disrupting *env-prot* leads to reduced GFP levels only from the late p6.9 promoter, thus deeming it of special interest. The AcMNPV *fp* gene encodes a 25 kDa protein, FP25, that is BV associated (Harrison et al., 1996; Wang et al., 2010). *fp* is

necessary for *polh* expression, and its deletion results in few polyhedra (Beames, Summers, 1989; Harrison et al., 1996). Deletion of *fp* has been previously shown to be dispensable for BV production in cell culture (Harrison et al., 1996; Yu et al., 2023), and our data suggests the same. On the other hand, deletion of a DNA fragment containing *fp* reduced GFP from the p10 promoter (Zhang et al., 2023), which is consistent with our observations when GFP is expressed from the p6.9 promoter. The *iap1* gene codes for an inhibitor of apoptosis protein found in Group I lepidopteran nucleopolyhedroviruses (NPVs) (Rohrmann, 2019a). Deletion mutants of *iap1* were found not to affect BV production in cell culture (McLachlin et al., 2001; Yu et al., 2023), and this is in line with what we observed upon targeting *iap1*. Moreover, our data indicate that disrupting *iap1* significantly reduces GFP production from the late p6.9 promoter.

The *lef6* gene encodes for the late expression factor 6 protein, and its homologs are found in all lepidopteran NPVs and granuloviruses (GVs) (Ayres et al., 1994; Rohrmann, 2019a). It is required for viral late and very late gene expression (Passarelli, Miller, 1994). Deletion of *lef6* has been shown to substantially delay the infection cycle and reduce BV production in insect cells (Lin, Blissard, 2002a). In this study, we have demonstrated that *lef6* disruption significantly reduces GFP production from the late p6.9 promoter, but BV production remains similar to the control. The *pkip* gene codes for a protein kinase-interacting protein that is known to stimulate viral pk-1 activity *in vitro* (Fan et al., 1998). It is required for the hyperexpression of viral very late genes, and its deletion has been shown to negatively impact progeny virus production (Lai et al., 2020; Fan et al., 1998; McLachlin et al., 1998). However, our data indicate that *pkip* disruption does not significantly impair BV production, while GFP expression from the late p6.9 promoter is

significantly reduced.

The viral *sod* gene encoding for superoxide dismutase is found in most of the lepidopteran baculovirus genomes (Tomalski et al., 1991; Rohrmann, 2019a). Previous studies reported that deletion of *sod* or of a DNA fragment containing *sod* does not negatively impact BV production or GFP expression from the very late p10 promoter (Tomalski et al., 1991; Yu et al., 2023; Zhang et al., 2023). However, in our work, targeting *sod* significantly reduced GFP levels from the late p6.9 promoter while BV production remained similar to the control. Finally, the auxiliary gene *v-ubi* encodes a viral ubiquitin protein that has been shown to be BV associated (Ayres et al., 1994; Wang et al., 2010). It was further reported to be attached to the inner surface of the viral membrane via a novel phospholipid anchor (Guarino et al., 1995). While a frameshift mutation within the coding region of this late gene resulted in a viable viral mutant, BV production was reduced by 5 to 10 fold in cell culture (Reilly, Guarino, 1996). Our data indicates that *v-ubi* disruption significantly decreases only GFP levels.

3.5.2 DNA Replication AcMNPV Genes

Three DNA replication genes (*gta*, *lef1*, and *lef2*) were disrupted to analyze their effect on IVT and GFP production from the late p6.9 promoter. While *gta* and *lef1* were categorized as of special interest (reduced only GFP), *lef2* was categorized as essential (reduced GFP and IVT), upon gene disruption.

3.5.2.1 Essential DNA Replication Gene

lef2 is a core baculovirus gene encoding for the late expression factor 2 protein acting as a DNA primase accessory factor (Ayres et al., 1994; Kool et al., 1994; Mikhailov, Rohrmann, 2002; Rohrmann, 2019a). *lef2* interacts with *lef1* and is required for transient DNA replication (Evans et al., 1997; Kool et al., 1994). Previous studies have demonstrated that *lef2* is necessary for BV production (Mikhailov, Rohrmann, 2002; Wu et al., 2010), which is in line with what we observed upon their disruption. We also demonstrated that their disruption significantly reduced GFP from the late p6.9 promoter.

3.5.2.2 Of Special Interest DNA Replication Genes

The AcMNPV *gta* gene encodes a global transactivator-like protein (Rohrmann, 2019a). It is believed to be involved in ATP-dependent unwinding of DNA, DNA recombination and repair, chromatin unwinding, as well as other functions (Rohrmann, 2019a). Deletion of *gta* or of a gene fragment containing *gta* has been previously shown to be dispensable for BV production in cell culture (Chen et al., 2021; Yu et al., 2023), which is consistent with our observation. However, disruption of *gta* led to a significant drop in GFP expression from the p6.9 promoter. *lef1* is a core baculovirus gene encoding for the late expression factor 1 protein with DNA primase activity (Ayres et al., 1994; Kool et al., 1994; Mikhailov, Rohrmann, 2002; Rohrmann, 2019a). *lef1* interacts with *lef2* and is believed to be required for transient DNA replication (Evans et al., 1997; Kool et al., 1994). Previous studies have demonstrated that *lef1* is necessary for BV production (Mikhailov, Rohrmann, 2002; Wu et al., 2010), which contradicts what we observed upon their disruption. Additionally, we

demonstrated that *lef1* disruption significantly reduced GFP from the late p6.9 promoter.

3.5.3 Transcription AcMNPV Genes

Nine transcription genes (*39k*, *AcOrf-38*, *AcOrf-41*, *AcOrf-45*, *lef11*, *lef8*, *lef9*, *p47*, and *pk-1*) were disrupted to analyze their effect on IVT and GFP production from the late p6.9 promoter. Based on the effect upon gene disruption, five transcription AcMNPV genes were identified as essential (reduced GFP and IVT), and four as of special interest (reduced only GFP).

3.5.3.1 Essential Transcription Genes

AcOrf-38 homologs are found in all lepidopteran NPVs and GVs, and they contain a conserved *Nudix* motif with highest homology to ADP-ribose pyrophosphatase (ADPRase) (Rohrmann, 2019a; Cohen et al., 2009; Ge et al., 2007). A previous study demonstrated that *AcOrf-38* has ADPRase activity and its deletion significantly impairs BV production (Ge et al., 2007). In this study, we report that targeting *AcOrf-38* significantly reduces BV and GFP production in cell culture. Homologs of the *lef8* gene coding for a late expression factor 8 protein are found in all baculoviruses (Ayres et al., 1994; Rohrmann, 2019a). *lef8*, containing a conserved motif GXKX4HGQ/NKG, is the largest subunit of the baculovirus DNA-dependent RNA-polymerase complex, which supports late and very late transcription (Guarino et al., 1998; Passarelli et al., 1994). A mutation in *lef8* has been shown to lower BV production and abolish very late transcription (Gauthier et al., 2012). In this study, *lef8* disruption significantly reduced IVT levels and GFP production

from the late p6.9 promoter. The *lef9* gene encoding a late expression factor 9 protein has homologs in all baculoviruses (Ayres et al., 1994; Rohrmann, 2019a). It is a subunit of the viral DNA-dependent RNA-polymerase complex and is required for late and very late gene expression (Guarino et al., 1998; Lu, Miller, 1994). *lef9* deletion was previously shown to be indispensable for BV production (Chen et al., 2021), which is consistent with our observations. Additionally, GFP production from the late p6.9 promoter was reduced significantly.

The *p47* gene with homologs present in all baculoviruses is another subunit of the baculovirus DNA-dependent RNA-polymerase complex and is required for transient late gene transcription (Rohrmann, 2019a; Guarino et al., 1998; Todd et al., 1995). A *p47* knocked-out mutant was found to be incapable of producing BVs in cell culture (Qi et al., 2023). In this study, we report that *p47* disruption significantly lowers IVT levels and GFP from the late p6.9 promoter. The AcMNPV *pk-1* gene codes for a protein kinase with strong similarity to serine-threonine protein kinases (Reilly, Guarino, 1994). PK-1 is a component of the viral very late transcription initiation complex involved in *lef8* phosphorylation, and it interacts with very late promoters to regulate their hyperexpression (Mishra et al., 2008; Liang et al., 2017). The kinase activity was found to be necessary for regulating viral propagation, and *pk-1* knockout was indispensable for BV production (Liang et al., 2013). Our data indicate that disrupting *pk-1* significantly decreases IVT and GFP from the late p6.9 promoter.

3.5.3.2 Of Special Interest Transcription Genes

The AcMNPV *39k/pp31* gene codes for a phosphoprotein that is found in the BV phenotype, with homologs present in genomes of all lepidopteran NPVs and GVs (Ayres et al., 1994; Wang et al., 2010; Rohrmann, 2019a). This late expression factor gene was reported to function at the transcription level, and purified 39K can bind to ssDNA and dsDNA with the same affinities (Lu, Miller, 1995; Guarino et al., 2002b). While *39k* deletion did not affect viral DNA replication, it reduced viral titers and transcript levels of many AcMNPV genes (Yamagishi et al., 2007). In contrast, our data indicate that *39k* disruption significantly reduces GFP levels, while IVT remains similar to the control. *AcOrf-41* encodes a late expression factor 12 protein that is required for transient late gene transcription in *Spodoptera frugiperda* cells but not for DNA replication (Rapp et al., 1998; Li et al., 1999; Guarino et al., 2002a). While *AcOrf-41* mutant viruses were viable in both Sf9 and High 5 cells, viral yields were reduced (Guarino et al., 2002a). However, in this study, we demonstrate that *AcOrf-41* disruption significantly lowers GFP expression from the late p6.9 promoter, while BV production remains similar to the control.

AcOrf-45 was reported to be necessary for the expression of *AcOrf-41* when provided in *cis* in a transient late transcription assay (Li et al., 1999). Deletion of DNA fragments containing *AcOrf-45* was found to be dispensable for BV production and GFP expression from the very late p10 promoter (Chen et al., 2021; Zhang et al., 2023). In contrast, we observed that disruption of *AcOrf-45* significantly reduces only GFP expression from the late p6.9 promoter. *lef11* codes for a late expression factor 11 protein that functions at the late gene transcription level (Lu, Miller, 1995). It is not required for transient DNA replication, but for late gene promoter activity (Lu, Miller, 1995; Todd et al., 1995).

lef11 deletion has been shown to negatively impact late gene transcription, viral DNA replication, and virus propagation in cell culture (Lin, Blissard, 2002b). Our data indicates that disrupting *lef11* results in decreased GFP levels only.

3.5.4 Structural AcMNPV Genes

Five structural genes (*AcOrf-43*, *AcOrf-53*, *AcOrf-54*, *AcOrf-58/59*, and *odv-e66*) were disrupted to analyze their effect on IVT and GFP production from the late p6.9 promoter. Upon targeted disruption of these structural AcMNPV genes, three were categorized as essential (reduced GFP and IVT), and two as of special interest (reduced only GFP).

3.5.4.1 Essential Structural Genes

AcOrf-53 is a highly conserved gene found in all baculoviruses and has been reported to be involved in nucleocapsid assembly (Garavaglia et al., 2012; Liu et al., 2008). A previous study demonstrated that deletion of *AcOrf-53* did not affect viral DNA replication; however, this *AcOrf-53* deletion mutant virus was not capable of producing BV in Sf9 cells (Liu et al., 2008). We have reported that *AcOrf-53* disruption significantly reduces IVT levels and GFP expression from the late p6.9 promoter. *AcOrf-54* codes for the structural protein VP1054 that is required for nucleocapsid assembly and is associated with both BV and ODV phenotypes (Olszewski, Miller, 1997; Marek et al., 2013). It was previously demonstrated that *AcOrf-54* deletion reduced both IVT levels and GFP expression from the very late p10 promoter in Sf9 cells (Marek et al., 2013), which is consistent with our data when GFP was driven by the late p6.9 promoter.

Re-sequencing of the AcMNPV C6 strain revealed that *AcOrf-58* and *AcOrf-59* could be combined into a single ORF (Harrison, Bonning, 2003; Chakraborty et al., 2025). *AcOrf-58/59* has a ChaB domain and was found to be associated with both ODV and BV phenotypes (Rohrmann, 2019a; Braunagel et al., 2003; Wang et al., 2010). Previous studies have demonstrated that *AcOrf-58/59* deletion did not negatively impact BV production, although deletion of an *AcOrf-55* to *AcOrf-61* DNA fragment reduced GFP expression from the p10 promoter (Chen et al., 2021; Zhang et al., 2023). However, in this study, we report that targeting *AcOrf-58/59* decreases both IVT levels and GFP production from the p6.9 promoter.

3.5.4.2 Of Special Interest Structural Genes

AcOrf-43 is a highly conserved gene whose homologs are found in all Group I and most Group II NPV genomes (Tao et al., 2013; Rohrmann, 2019a). Previous studies demonstrated that deletion of *AcOrf-43* or of a gene fragment containing *AcOrf-43* does not impair BV production in cell culture (Tao et al., 2013; Yu et al., 2023), which is consistent with our observations. Additionally, while no direct evidence of *AcOrf-43* knock-out effect on foreign protein production was found in the literature, we report that targeting this gene results in reduced GFP production from the p6.9 promoter. The *odv-e66* gene codes for the ODV-specific envelope protein ODV-E66 (Xiang et al., 2011). Deletion of *odv-e66* or of a DNA fragment containing *odv-e66* was shown not to affect BV or GFP production, respectively, in cell culture (Xiang et al., 2011); however, our data indicate that *odv-e66* disruption significantly reduces GFP from the late p6.9 promoter, while BV production is similar to the control.

3.5.5 Unknown Function AcMNPV Genes

Nine unknown function genes (*AcOrf-13*, *AcOrf-19*, *AcOrf-26*, *AcOrf-34*, *AcOrf-51*, *AcOrf-52*, *AcOrf-60*, *AcOrf-63*, and *pcna*) were disrupted to analyze their effect on IVT and GFP production from the late p6.9 promoter. Targeted disruption of these unknown function AcMNPV genes resulted in four of them being identified as of special interest (reduced GFP and not BV), and five as essential (reduced both GFP and BV).

3.5.5.1 Essential Genes of Unknown Functions

The functions of the AcMNPV *AcOrf-13*, *AcOrf-34*, *AcOrf-51*, *AcOrf-52*, and *AcOrf-63* genes are not known. It has been previously reported that while individual deletion of *AcOrf-13*, *AcOrf-34*, and *AcOrf-51* reduces BV production in cell culture, deletion of *AcOrf-52* or *AcOrf-63* is dispensable for BV production (Cai et al., 2012; Chen et al., 2021; Yu et al., 2023). Additionally, deletion of a DNA fragment containing *AcOrf-13* lowered GFP expression from the p10 promoter; however, deletion of a DNA fragment containing *AcOrf-63* was found to be dispensable for GFP production from the p10 promoter (Zhang et al., 2023). Our data indicate that disrupting these genes results in decreased IVT levels and GFP expression from the late p6.9 promoter, thus making them essential for cell culture.

3.5.5.2 Of Special Interest Genes of Unknown Functions

Although the exact roles of the AcMNPV *AcOrf-19*, *AcOrf-26*, *AcOrf-60*, and *pcna* genes are not known, previous studies have reported that individual deletion of these genes is

dispensable for BV production in cell culture (Chen et al., 2021; Yu et al., 2023). Additionally, deletion of DNA fragments containing *AcOrf-19* or *AcOrf-60* lowered GFP expression from the p10 promoter, and deletion of a DNA fragment containing *pcna* was found to be dispensable for GFP production from the p10 promoter (Zhang et al., 2023). Our data indicate that disrupting these genes reduces GFP from the late p6.9 promoter, while BV production remains similar to the control.

3.6 Conclusions

The majority of the work found in the literature mainly focused on the effect of baculovirus gene knock-out on virion production. In this study, we investigated the impact of AcMNPV gene disruption on both BV and foreign protein production. We have identified 19 AcMNPV genes that are essential for BV production and GFP expression from the late p6.9 promoter and 19 AcMNPV genes of special interest that are required for GFP production from the late p6.9 promoter but not for BV production. The choice of late or very late promoters to drive a foreign gene was found to play a role in the impact of AcMNPV gene disruption (Bruder, Aucoin, 2023b). While exogenous genes were typically expressed under the very late promoters in the literature, we used a late promoter for their expression, which could potentially contribute to the differences in the observed effects between our work and previous studies. We believe that deleting or disrupting the genes of special interest could improve the BacMam for expression in mammalian cells. Moving forward, we recommend disrupting different combinations of the genes of special interest to further probe their effect on progeny virus and foreign protein production. Moreover, the essential

genes should not be removed from the genome to maintain IVT levels and exogenous gene expression from late promoters.

Chapter 4

Adapting next-generation sequencing to recombinant baculovirus vectors

The BEVS platform has been used to produce different biologics; however, the majority of the commercially available baculovirus genomes remain virtually unmodified. It has been previously demonstrated that CRISPR-Cas9-mediated targeted *AcMNPV* gene disruption can be used to successfully study phenotypic changes ([Bruder et al., 2021](#); [Bruder, Aucoin, 2023b,a](#); [Hausjell et al., 2023](#)). In this study, transient or *in Process* CRISPR-Cas9 genome editing of rBEVs refers to CRISPR-mediated *AcMNPV* gene disruption while the recombinant virus undergoes its infection cycle in Sf9-Cas9 cells. There is a possible competition effect between virus replication and sgRNA/Cas9 targeting and cleaving; the Cas9 is available only for a limited time as it is expressed under the OpIE2 promoter, which is downregulated as infection progresses; and upon disruption, the targeted genomes can be repaired using the cell's repair mechanism via recombination. All these factors result in

a heterogeneous pool of viral genomes with untargeted, properly repaired, and mutated genomes in varying amounts; thus, sequencing in this system is not straightforward (Figure 4.1). Although genome sequences of native baculovirus backbones have been previously reported (Ayres et al., 1994; Maghodia et al., 2014; Chateigner et al., 2015), our study is the first to report on whole-genome sequencing of rBEVs. Here, we present a shotgun sequencing pipeline to generate a consensus sequence of a GFP-carrying rBEV (p6.9GFP rBEV) used in CRISPR-Cas9-based T-I assays, provide a set of tiled-amplicon primers based on the shotgun-sequenced rBEV, and adapt tiled-amplicon sequencing to rBEVs to investigate untargeted and targeted recombinant *AcMNPV* genomes. We also demonstrate that plasmid- and rBEV-based delivery of sgRNAs to Sf9-Cas9 cells for *gp64* disruption results in reduced BV levels while maintaining GFP production. The research presented in this chapter has been published in *Viruses*.

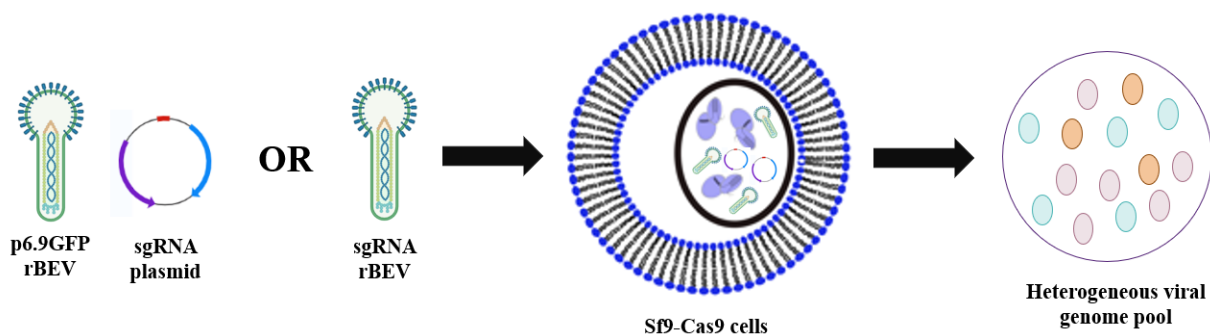


Figure 4.1: *in Process* CRISPR-Cas9 genome engineering of rBEVs.

Adapting Next-Generation Sequencing to *in Process* CRISPR-Cas9 Genome Editing of Recombinant *AcMNPV* Vectors: From Shotgun to Tiled-Amplicon Sequencing

Madhuja Chakraborty ¹, Lisa Nielsen ^{1,2}, Delaney Nash ², Jozef I. Nissimov ², Trevor C. Charles ² and Marc G. Aucoin ¹

¹ Department of Chemical Engineering, University of Waterloo

² Department of Biology, University of Waterloo

4.1 Abstract

The alphabaculovirus *Autographa californica* multiple nucleopolyhedrovirus (*AcMNPV*) is the most commonly used virus in the Baculovirus Expression Vector System (BEVS) and has been utilized for the production of many human and veterinary biologics. *AcMNPV* has a large dsDNA genome that remains understudied, and relatively unmodified from the wild-type, especially considering how extensively utilized it is as an expression vector. Previously, our group utilized CRISPR-Cas9 genome engineering that revealed phenotypic changes when baculovirus genes are targeted using either co-expressed sgRNA or transfected sgRNA into a stable insect cell line that produced the Cas9 protein. Here, we describe a pipeline to sequence the recombinant *AcMNPV* expression vectors using shotgun sequencing, provide a set of primers for tiled-amplicon sequencing, show that untargeted baculovirus vector genomes remain relatively unchanged when amplified in Sf9-Cas9 cells,

and confirm that *AcMNPV gp64* gene disruption can minimize baculovirus contamination in cell cultures. Our findings provide a robust baseline for analyzing *in process* genome editing of baculoviruses.

4.2 Introduction

Baculovirus vectors are versatile tools for producing recombinant products in insect cells. Together, they are commonly referred to as the Baculovirus Expression Vector System (BEVS). The BEVS platform utilizes a recombinant *Autographa californica* multiple nucleopolyhedrovirus (*AcMNPV*) baculovirus that can infect cultured cells derived from the ovarian tissue of its natural hosts, such as the Fall Armyworm (*Spodoptera frugiperda*; Sf21 or Sf9 cells) and the Cabbage Looper (*Trichoplusia ni*; High-Five™ or Hi-5 cells) (van Oers et al., 2015). The double-stranded DNA genome of the *AcMNPV* C6 strain is around 134 kbp and contains ~ 155 open reading frames (ORFs) or genes (Ayres et al., 1994; Miele et al., 2011; Rohrmann, 2019a). Although BEVS has gained popularity in laboratory settings, its application in industry remains limited. A bottleneck in the improvement of this system is that the functions of many *AcMNPV* genes have yet to be experimentally determined. In fact, most of the commercially available *AcMNPV* DNA backbones remain relatively unchanged (Bruder, Aucoin, 2023b). However, improvements in recombinant protein production and genome stability have been achieved by removing certain *AcMNPV* genes (Hitchman et al., 2010a,b; Pijlman et al., 2001). This shows that the BEVS platform could benefit from a systemic screening of genes to identify their essentiality. The removal or disruption of genes not essential for the production of foreign proteins

in cell culture has the potential to increase the production of foreign proteins (George et al., 2015). Using CRISPR-Cas9 gene editing technology, baculovirus genes can be targeted in the production phase without the need for gene-specific complementary cell lines.

CRISPR-Cas9 is a highly efficient tool for targeted *in vivo* and *in vitro* genetic engineering. Of *Drosophila melanogaster* S2 cells that gained resistance to puromycin from a plasmid that also contained the *cas9* gene and single guide RNA (sgRNA), 88% of the alleles were found to have indel mutations (Bassett et al., 2014). CRISPR has also been used to identify genes essential for *Bombyx mori* BmE cell viability through a genome-wide screen (Chang et al., 2020). CRISPR-Cas9 genome editing in insect antiviral research was first reported in 2016 and it showed that an inducible CRISPR-Cas9 system (p39 K-Cas9-sgRNA-mCherry), activated upon viral infection, could reduce the possibility of off-targets and produce transgenic silkworms resistant to *Bombyx mori* nucleopolyhedrovirus (*BmNPV*) (Dong et al., 2016). Two other studies constructed inducible CRISPR-Cas9 systems to introduce antiviral activity in transgenic silkworms by utilizing the *Bombyx mori* U6 promoter for expression of sgRNAs and the *BmNPV* 39k promoter for Cas9 expression only upon viral infection (Dong et al., 2018; Liu et al., 2022). One study created a transgenic line expressing Cas9 and sgRNA to knock out the host *BmATAD3A* gene and inhibit *BmNPV* replication (Dong et al., 2018), while the other used four sgRNAs targeting the *BmNPV* genes essential for viral replication, thus providing the transgenic silkworm with resistance to *BmNPV* infection (Liu et al., 2022).

U6 promoters with transcriptional activity for sgRNA expression in *S. frugiperda* and *T. ni* were not reported on until 2017 (Mabashi-Asazuma, Jarvis, 2017). This led to a study by Bruder et al. in 2021 that implemented the CRISPR-Cas9 tools in the BEVS,

specifically the baculovirus *AcMNPV* and Sf9 insect cells (Bruder et al., 2021). This work utilized engineered Sf9 cell lines expressing *cas9* or *dcas9* to compare targeted gene disruption (CRISPRd) or transcriptional repression (CRISPRi), respectively (Bruder et al., 2021). Building on this previous work, rBEVs with the reporter gene replaced with the human immunodeficiency virus type 1 (HIV-1) *gag* gene, also carrying sgRNAs targeting the *AcMNPV* *gp64* or *vp80* genes, were constructed to reduce baculovirus co-production in cell cultures (Bruder, Aucoin, 2023a). CRISPR-mediated disruption of the *gp64* or *vp80* genes resulted in $\sim 99\%$ or $\sim 94\%$ reduction in infectious virus titers (IVTs), respectively, without impacting Gag virus-like particle (VLP) production (Bruder, Aucoin, 2023a). An alternative strategy to minimize baculovirus contamination utilized an inducible CRISPR-Cas9 knockout system based on co-infection with two rBEVs carrying either the *cas9* endonuclease or a sgRNA for the *gp64* or *vp80* genes (Hausjell et al., 2023). While knocking out these genes resulted in a more than 90% reduction in virus titer, it also reduced the overall fluorescence intensity of the enhanced yellow fluorescent protein (eYFP) acting as a surrogate for recombinant protein production (Hausjell et al., 2023). Another study developed a CRISPR-Cas9 based transfection-infection assay to effectively scrutinize the baculovirus genes for essentiality by transfecting Sf9-Cas9 cells with sgRNA plasmid followed by infecting with a rBEV carrying green fluorescent protein (GFP) and evaluate its impact on foreign protein and IVT production (Bruder, Aucoin, 2023b).

Baculoviruses are prone to mutations during replication processes and inherently form defective interfering particles upon propagation in insect cell cultures (Miele et al., 2011; Pijlman et al., 2001), thus making genome sequencing an important tool to investigate the

baculovirus population being employed for various applications including, but not limited to, biologics production. Although phenotypic confirmation of CRISPR-Cas9 activity has previously been reported ([Bruder et al., 2021](#); [Bruder, Aucoin, 2023a,b](#)), an overall look at the entire rBEV genome upon targeted gene disruption has yet to be thoroughly examined. In addition, a deep understanding of the rBEV sequence is required prior to designing sgRNAs to target different genes of the rBEV. This paper addresses some of the current shortcomings and details the methodologies required to achieve these tasks.

In this work, we initially use shotgun sequencing on infected cell culture supernatant and describe the pipeline to generate a consensus sequence of the rBEV. While shotgun sequencing allows sequencing of unknown samples, it requires large quantities of DNA input; has the potential to introduce assembly errors, especially around repetitive regions; and would require additional processing to sequence intracellular baculovirus genomes. Thus, we decided to use a whole-genome tiled-amplicon sequencing assay that is robust, requires less DNA, provides high coverage, and is specific to genome(s) of interest ([Quick et al., 2017](#)). We adapted this tiled-amplicon approach ([Quick et al., 2017](#)) for the rBEV genomes recovered from the multi-well plate experiments, which uses Sf9-Cas9 cells as the host of baculovirus infection, and reported on the whole-genome sequences of these targeted and untargeted rBEV gDNAs (major species) and confirmed the targeted mutations. While the consensus sequence of the rBEV enabled us to confidently design sgRNAs to target specific *AcMNPV* genes, the more specific tiled-amplicon sequencing assay provided a foundation to study the whole *in process* rBEV genome and confirm the mutations upon CRISPR-Cas9 mediated gene targeting. This is especially important if you want to target genes whose disruption can reduce baculovirus contamination and facilitate the purification

process of your products, such as recombinant proteins or antigens for vaccines.

4.3 Materials and Methods

4.3.1 Cell Line and Maintenance

Parental Sf9 cells and Cas9 expressing Sf9 cells (Sf9-Cas9, consisting of a Cas9-2A-puromycinR gene cassette) (Bruder et al., 2021) were maintained in capped glass Erlenmeyer flasks in Sf-900TM III serum-free media (SFM) (Gibco, Carlsbad, CA, USA) at 27 °C and 130 rpm. The cells were passaged every 3–4 days when the viable cell density reached between 3×10^6 cells/mL and 5×10^6 cells/mL. Additionally, puromycin (Sigma-Aldrich, Oakville, ON, Canada) was added to the Sf9-Cas9 cells every other passage at a concentration of 5 µg/mL to ensure *cas9* expression.

4.3.2 Baculovirus Amplification and Quantification

Two recombinant baculoviruses were used in this work: a recombinant baculovirus expressing the GFP monomeric Azami green (mAG) under the *AcMNPV* late basic p6.9 promoter (Bruder, Aucoin, 2023b), herein referred to as p6.9GFP rBEV; and a baculovirus that co-expresses a sgRNA under a SfU6 promoter and mAG under p6.9 promoter (Bruder, Aucoin, 2023a), herein referred to as p6.9GFP_sgRNA_gp64+131 rBEV. The baculoviruses were amplified in Sf9 cells at a low MOI (~ 0.1 pfu/cell) until cell viability dropped between 80 and 90% to obtain a working virus stock for each rBEV.

Infectious virus titers were quantified by end-point dilution assay (EPDA), as previously

described (Reed, Muench, 1938; O’Reilly et al., 1992). Briefly, each well of 96-well tissue culture-treated plates (VWR International, Mississauga, ON, Canada) was seeded with 100 μ L of Sf9 cells diluted to 2×10^5 cells/mL and allowed to adhere for around 1 h at 27 °C. During the incubation period, virus stocks were serially diluted from 10^{-2} to 10^{-8} using Sf-900TM III SFM. The cells were then inoculated with 10 μ L of a virus dilution resulting in 12 replicates per dilution. The plates were incubated at 27 °C for 6–7 days and observed under a fluorescence microscope to determine green fluorescence. The reciprocal of TCID₅₀ expressed in mL of virus added (0.01 mL) was multiplied by 0.68 based on the Poisson distribution to obtain the virus titer in plaque-forming units per mL (pfu/mL).

4.3.3 Plasmid Design and Construction

All primers used in this study were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA) and can be found in Supplementary Table B.1. A previously constructed sgRNA plasmid with a scrambled spacer sequence was used as the control plasmid in this work (Chakraborty et al., 2024). Briefly, a PCR-amplified SfU6 promoter (Mabashi-Asazuma, Jarvis, 2017) gBlock gene fragment and gRNA scaffold along with a transcriptional terminator (Addgene # 49411) (Port et al., 2014) were amplified in a fusion PCR reaction to obtain the SfU6-sgRNA insert with a scrambled spacer sequence (5’-3’ cac-cttgaagcgcgatgaact). The backbone containing the ampicillin resistance gene (ampR) and the pBR322 origin of replication (ori) from the pBR322-TIMER plasmid (Addgene # 103056) (Claudi et al., 2014) was PCR-amplified separately to obtain the backbone fragment. Finally, the SfU6-sgRNA insert and the backbone fragment were assembled in a Gibson assembly reaction to construct the pSfU6-sgRNA scrambled control plasmid. Ad-

ditionally, the gp64+131 plasmid with a *gp64* targeting spacer sequence (5'-3' ggaaacgct-gcaaaaggacg) at the +131 location from the first nt within *gp64* was obtained from a previous study (Bruder, Aucoin, 2023b).

4.3.4 Shotgun Sequencing of the p6.9GFP rBEV

In this study, shotgun sequencing was used to sequence the genome of the p6.9GFP rBEV recovered from the supernatant of the infected Sf9 cells in suspension. Next, 15 mL of virus stock was filtered with a 0.22 μm PES filter and the gDNA was extracted using the AllPrep PowerViral DNA/RNA kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol without bead beating. To obtain enough starting material for sequencing sample preparation, multiple gDNA extractions were performed and the pooled extracts were concentrated with AMPure XP magnetic beads (Beckman Coulter, Mississauga, ON, Canada). The gDNA bound to the beads was eluted in a smaller volume to increase the starting material concentration, thus keeping the DNA yields high throughout the sample preparation steps. Using a dsFragmentase enzyme, gDNA was fragmented for short-read sequencing. The fragmented gDNA was then prepared for shotgun sequencing with the NEBNext[®] Ultra[™] II DNA Library Prep Kit for Illumina[®] (New England Biolabs, Whitby, ON, Canada) according to the manufacturer's instructions. Briefly, end-repair, 5' phosphorylation, and dA-tailing were performed on the fragmented gDNA. This was followed by the ligation of bell-shaped adapters to the fragments and then cleavage of the bell-shape by U-excision. The fragments were cleaned and size-selected, and PCR enrichment was used to add Illumina barcodes and sequencing adapters to the fragment ends. The prepared DNA library was sequenced on the MiSeq platform using 2×250 sequencing

with a MiSeq reagent kit v2 (Illumina, San Diego, CA, USA).

A best-guess sequence of the p6.9GFP rBEV, generated by combining the *AcMNPV* DNA backbone and a transfer vector carrying the GFP gene (Bruder, Aucoin, 2023b), was used as the reference genome to assemble our working viral p6.9GFP rBEV genome. Sequenced raw reads were processed with Trimmomatic (Bolger et al., 2014) to remove low-quality reads and index sequences. The short Illumina reads were then aligned to the reference genome of the p6.9GFP rBEV using BBtools (Bushnell et al., 2017) to produce a sam file of mapped, partially mapped, and unmapped reads. The sam file was converted to a bam file using samtools (Danecek et al., 2021) and Anvi'o (Eren et al., 2021). The bam file was then polished with Pilon (Walker et al., 2014) to resolve differences such as insertions, deletions, and single-nucleotide polymorphisms (SNPs) between the reads and the reference genome, which produced a FASTA file containing an improved representation of our working rBEV genome. Finally, using the MAFFT-7.0 alignment algorithm (Katoh et al., 2019), the reference genome and the newly assembled genome were aligned and the DNAdiff tool was used to identify differences and similarities between the two. The sequence of the newly assembled genome (hereby referred to as the shotgun-sequenced or working rBEV) was then used for all future work involving the p6.9GFP rBEV.

4.3.5 Transfection in Multi-Well Plates

A previously developed CRISPR-Cas9 based transfection-infection assay (T-I assay) was used in this study (Bruder, Aucoin, 2023b). Briefly, Sf9-Cas9 cells were seeded on a tissue-culture-treated 12-well plate (VWR International, Mississauga, ON, Canada) at

a density of 0.9×10^6 cells/well and allowed to adhere for around one hour at 27 °C. The cells were then transfected with the sgRNA scrambled control plasmid or the gp64+131 plasmid using FuGENE HD (Promega, Madison, WI, USA) transfection reagent following the manufacturer’s protocol.

4.3.6 Infection in Multi-Well Plates

For experiments conducted using the p6.9GFP_sgRNA_gp64+131 rBEV, Sf9-Cas9 cells were seeded on a tissue culture-treated 12-well plate (VWR International, Mississauga, ON, Canada) at a density of 0.9×10^6 cells/well and incubated overnight at 27 °C. The following day, the media were replaced with fresh Sf-900™ III media containing the p6.9GFP_sgRNA_gp64+131 rBEV or the p6.9GFP rBEV (infected-only control) to achieve synchronous infection. For experiments involving transfection, 16–24 h post-transfection (hpt), the media were aspirated from each transfected well, and fresh Sf-900™ III SFM containing p6.9GFP rBEV was added to achieve an infection at an MOI of 3 pfu/cell.

4.3.7 Harvesting from Multi-Well Plates

Approximately 48 h post-infection (hpi), the cell cultures were harvested by centrifugation at $800 \times g$ for 15 min and the cell pellets were analyzed by flow cytometry. The supernatants were stored at 4 °C for further analysis by EPDA and next-generation sequencing (NGS) and the cell pellets treated with the p6.9GFP_sgRNA_gp64+131 rBEV or the gp64+131 plasmid were stored at -80 °C for NGS analysis.

4.3.8 Flow Cytometry Analysis

Cell pellets were resuspended in 2% paraformaldehyde in phosphate-buffered saline (PBS) followed by incubation at 4 °C for 30 min. The samples were then diluted in 1× PBS and analyzed by a BD AccuriTM C6 Plus flow cytometer (BD Biosciences, San Jose, CA, USA) equipped with a 488 nm excitation frequency blue laser. Samples were run at a low flow rate (14 µL/min) and 10,000 events were collected for each sample. The acquired flow data were analyzed by FlowJoTM V10 software (Tree Star, Ashland, OR, USA). After applying gates to remove cell debris and intrinsic cellular fluorescence, the geometric mean fluorescence intensity captured by the FL1 detector was calculated. The R programming language was then used for downstream data analysis and visualization.

4.3.9 Tiled-Amplicon Sequencing Assay for rBEV Genomes

In this work, the tiled-amplicon sequencing assay was used to sequence the rBEV gDNAs recovered from multi-well plates. This whole genome NGS assay was adapted from the paper ‘Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples’ (Quick et al., 2017) and it uses a set of target-specific primers to amplify a genome. Primer sets bind across the genome to PCR-amplify specific regions, called amplicons, that make up the entire genome. Each successive primer pair overlaps with the next; thus, the overlapping PCR amplicons can be assembled to generate the entire genome.

4.3.9.1 Primer Design and Multiplex PCR

PrimalScheme (Quick et al., 2017) was used to design primers to amplify 5000 nt tiled segments of the rBEV gDNA extracted using the Wizard Genomic DNA Purification kit (Promega, Madison, WI, USA) following the manufacturer’s protocol. To design the primer scheme, the FASTA file of the shotgun-sequenced rBEV was selected as the reference genome and the amplicon length was set to 5000 nt. PrimalScheme recognized the input genome as linear; thus, to ensure 100% coverage of the circular genome, an additional amplicon that overlapped with the first and last amplicons was used to close the gap. Briefly, the first 3000 nt of the genome was added at the end of the last 2300 nt and it was run through PrimalScheme while setting the amplicon length to 5000 nt. Upon running the two primer schemes, the program returned files containing the primer locations and primer pairs (Supplementary Tables B.2 and B.3, respectively) which were subsequently ordered from IDT. The lyophilized tiled-amplicon primers were resuspended in nuclease-free water using the volumes suggested by IDT to achieve a stock concentration of 100 μ M. Primer pools were then prepared by pooling primer pairs producing odd-numbered amplicons (Primer Pool 1) separately from those producing even-numbered amplicons (Primer Pool 2) to a total concentration of 10.2 μ M for each pool and 0.34 μ M for each primer (Figure 4.2). This ensured primer pairs for alternate regions were pooled together and the amplicons overlapped between the pools. The tiled-amplicon primers generated for the shotgun-sequenced rBEV were utilized for all tiled-amplicon sequencing assays in this study.

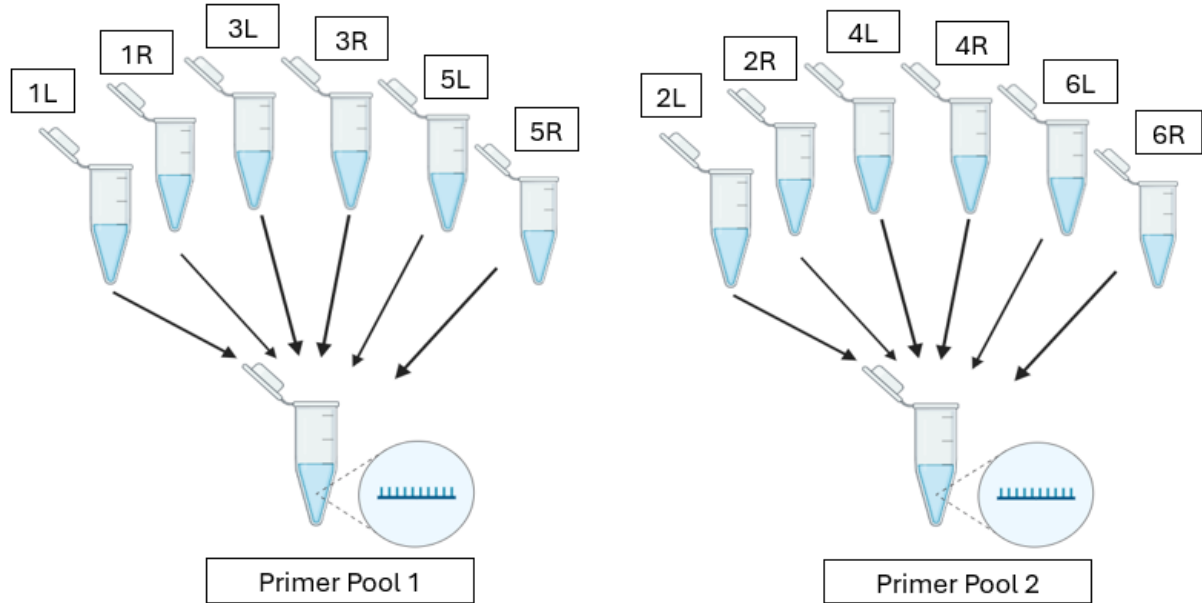


Figure 4.2: Odd-numbered forward (L) and reverse (R) primers are combined to make Primer Pool 1 and even-numbered forward (L) and reverse (R) primers are combined to make Primer Pool 2.

A multiplex PCR method generated 5000 nt amplicons with a short ~ 75 nt overlap between each PCR amplicon. This method allowed amplification of the whole genome with alternate rBEV genome regions amplified by Pool 1 or Pool 2 primers (Quick et al., 2017). Two PCR reactions (PCR Pool 1 and PCR Pool 2) were set up for each sample. The multiplex PCR setup and amplification parameters are shown in Tables 4.1 and 4.2, respectively.

Table 4.1: Multiplex PCR setup.

PCR Component	PCR Pool 1	PCR Pool 2
Q5 [®] High-Fidelity 2× Master Mix	12.5 µL	12.5 µL
Primer Pool 1 (10 µM)	2.5 µL	-
Primer Pool 2 (10 µM)	-	2.5 µL
DNA Template (1–10 ng) ¹	10 µL	10 µL

¹ Recovered rBEV gDNA used as the DNA template.

Table 4.2: Multiplex PCR amplification parameters.

PCR Step Number	PCR Step	Temperature	Duration
1	Initial denaturation	98 °C	30 s
2	Denaturation	95 °C	15 s
3	Annealing/extension	63 °C	5 min
N/A	Repeat steps 2 and 3 for 35 cycles	N/A	N/A
4	Hold	10 °C	Infinite

4.3.9.2 DNA Amplicon Cleanup and Quantification

First, 2 µL of the PCR products was run on 1% agarose gel to confirm the presence of 5000 nt amplicons and visually compare the Pool 1 and Pool 2 band intensities for each sample. Following that, Pool 1 and Pool 2 PCR products of a sample containing a similar amount of the 5000 nt amplicons were combined in a 96-well PCR plate. Next, 32 µL of room temperature AMPure XP magnetic beads (Beckman Coulter, Mississauga, ON, Canada) was added to the combined 40 µL PCR products for each sample and cleanup was performed according to the manufacturer's protocol. Finally, 20 µL of purified DNA amplicons, eluted in 22.5 µL Sox 5 (10 mM Tris-HCl, pH 8.5), was transferred to new wells in the PCR plate.

Quantification of 2 μ L of the purified amplicons was performed using the Qubit dsDNA high-sensitivity assay kit (Thermo Fisher Scientific, Waltham, MA, USA) with a Qubit 4 Fluorometer (Invitrogen by Thermo Fisher Scientific) in accordance with the manufacturer's instructions. All the concentrations were recorded for further sample preparation.

4.3.9.3 DNA Library Construction

Custom index and adapter primers were synthesized by IDT (Supplementary Table B.4) and were diluted to 10 μ M stocks stored at -20°C . Illumina DNA Prep, (M) Tagmentation kit (Illumina, San Diego, CA, USA) was used to prepare library DNA for sequencing according to the manufacturer's protocol. Briefly, bead-linked transposomes (BLTs) were used for DNA tagmentation, in which the DNA were fragmented and tagged with adapters (5'-3' CTGTCTCTTATACACATCT). This was followed by a cleanup step that involved stopping the tagmentation process and washing the adapter-tagged DNA. The tagmented DNA was then amplified using a reduced-cycle PCR that added index 1 (i7) adapters, index 2 (i5) adapters, and sequences necessary to generate sequencing-ready DNA fragments. Finally, the DNA library amplicons were purified using a double-sided bead purification process that excluded index primers and primer dimers. A Qubit dsDNA high-sensitivity assay kit (Thermo Fisher Scientific, Waltham, MA, USA) with a Qubit 4 Fluorometer (Invitrogen by Thermo Fisher Scientific) was then used to quantify the indexed amplicons as described by the manufacturer and median sizes of each DNA library (400 nt) were estimated by running them on 1% agarose gel.

4.3.9.4 Sequencing DNA Library Preparation

The concentration of the indexed amplicons in ng/ μ L was converted to nM using Equation (4.1) for sequencing cluster generation and each DNA library was diluted to 4 nM in Sox 5 solution for sequencing preparation.

$$\text{concentration in nM} = \frac{\text{concentration in ng}/\mu\text{L}}{660 \text{ g/mol} \times \text{average library size in bp}} \quad (4.1)$$

Each 4 nM DNA library was denatured using a freshly made 0.2 N NaOH solution. After 5 min of incubation at room temperature, the reaction was stopped using pre-chilled hybridization buffer (HT1) to obtain 1 mL of 20 pM denatured DNA library. Using HT1, the 20 pM denatured DNA library was diluted to the desired concentration. The 10 nM PhiX sequencing control (Illumina, San Diego, CA, USA) was also diluted to 4 nM in Sox 5 solution and the 4 nM PhiX library was then denatured and diluted as the DNA libraries. Finally, 588 μ L of denatured DNA library was combined with 12 μ L of denatured PhiX library and loaded onto the MiSeq reagent kit v3 (Illumina, San Diego, CA, USA) cartridge for sequencing.

4.3.10 Bioinformatics Pipeline for Major Species

Recombinant baculovirus tiled-amplicon sequencing reads were processed and analyzed using the Galaxy online platform (<https://usegalaxy.org/>, accessed on 28 November 2023). Reference genome(s), Illumina adapter sequences, and forward (R1) and reverse (R2) raw read files were uploaded to the server, followed by an assessment of read quality performed using the FastQC tool (Galaxy Version 0.74+galaxy1). The Trimmomatic

tool (Galaxy Version 0.39+galaxy2) was then used to trim the paired-end reads by selecting the ILLUMINACLIP step for custom adapter removal and performing 4 Trimmomatic Operations (SLIDINGWINDOW: 5:20, MINLEN: 175, LEADING: 15, TRAILING: 10) (Figure 4.3). The parameters of the Trimmomatic Operations can be changed if needed. Using FastQC, the quality of trimmed paired reads was analyzed and the files passed the adapter content criteria. The BWA-MEM2 tool (Galaxy Version 2.2.1+galaxy1) was used to build an index and map the forward and reverse trimmed paired reads to a genome from its history (reference genome). Optionally, samtools depth (Galaxy Version 1.15.1+galaxy2) could be run on the bam file to obtain the read depth at each position. An automated genome assembly tool, Pilon (Galaxy Version 1.20.1), then aligned the bam file (BWA-MEM2 alignment file) of one sample at a time with the reference genome to generate a Pilon-assembled genome. Following this, the DNAdiff tool (Galaxy Version 4.0.0+galaxy1) was used to evaluate similarities and differences between the Pilon assembled genome of each sample and the reference genome. Lastly, all the FASTA files were downloaded and pasted into a single file for visualization of genome alignments and identification of indel mutations performed by the MAFFT version 7 server (<https://mafft.cbrc.jp/alignment/server/>, accessed on 28 November 2023) (Katoh et al., 2019).

To further confirm the CRISPR-Cas9 mediated mutations within the targeted *gp64* gene, the Galaxy online platform was run from a different server (<https://usegalaxy.be/>, accessed on 15 February 2025) that provided a tool to analyze CRISPR-based mutations specifically. Briefly, after assessing the quality of the reads by FastQC, the Trimmomatic tool was used to remove adapters and low-quality reads (ILLUMINACLIP with custom

adapters, SLIDINGWINDOW: 4:20, MINLEN: 50, LEADING: 15, TRAILING: 10). Following this, FastQC was performed on the trimmed paired reads to ensure adapter removal and quality. Finally, the CRISPResso2 tool (Galaxy Version 0.1.1) (Pinello et al., 2016; Clement et al., 2019) was run utilizing the trimmed paired-end reads, a *gp64* amplicon sequence (the first 238 nucleotides of the *gp64* gene), the gp64+131 sgRNA sequence, and the following parameters—flexiguide homology: 60, minimum overlap length between reads: 5, maximum overlap length between reads: 251, and center of quantification window to use with respect to the 3' end of the provided sgRNA sequence: -3.

The figure displays four screenshots of the Trimmomatic tool's parameter configuration interface, arranged in a 2x2 grid. Each screenshot shows a different operation selected for configuration.

- 1: Trimmomatic Operation**: Shows "Sliding window trimming (SLIDINGWINDOW)" selected. Parameters include "Number of bases to average across" set to 5 and "Average quality required" set to 20.
- 2: Trimmomatic Operation**: Shows "Drop reads below a specified length (MINLEN)" selected. The "Minimum length of reads to be kept" is set to 175.
- 3: Trimmomatic Operation**: Shows "Cut bases off the start of a read, if below a threshold quality (LEADING)" selected. The "Minimum quality required to keep a base" is set to 15. A note at the bottom states: "Bases at the start of the read with quality below the threshold will be removed".
- 4: Trimmomatic Operation**: Shows "Cut bases off the end of a read, if below a threshold quality (TRAILING)" selected. The "Minimum quality required to keep a base" is set to 10. A note at the bottom states: "Bases at the end of the read with quality below the threshold will be removed".

Figure 4.3: Parameters used for the four Trimmomatic operations. The * here indicates a mandatory field.

4.4 Results

4.4.1 Consensus Sequence of p6.9GFP rBEV from Shotgun Sequencing

The *AcMNPV* C6 strain sequence (NCBI accession number: NC_001623.1) (Ayres et al., 1994), modified to accommodate changes for use as an rBEV (the *chiA* and *v-cath* genes were removed and the promoter and reporter gene, i.e., p6.9GFP, were added) was used as a reference genome. To confirm the rBEV sequence prior to targeting different genes in the virus vector, shotgun sequencing was performed. Shotgun sequencing was selected for whole-genome sequencing of the p6.9GFP rBEV because it does not require specific knowledge of the DNA sequence. Twelve million raw reads were obtained, and around eight million processed reads passed the quality filters. The polished sequencing data were manually annotated and provided a better understanding of the genome by highlighting key differences between the assumed and the actual sequences of the rBEV. The shotgun-sequenced rBEV (131,545 bp) was ~400 bp shorter than the reference genome (131,944 bp) and included 15 insertions, 11 deletions, and 52 SNPs in various regions of the genome (Table 4.3). Among the mutations, a 37 bp deletion and a 7 bp SNP were seen in the 45 bp region between where *chiA* and *v-cath* genes were originally located. This region potentially served as the promoter for the bi-directional adjacent *chiA* and *v-cath* genes. Moreover, the homologous repeat regions (hrs) had five indel mutations ranging from 1 bp to 137 bp and seven SNPs of 1 bp each. While the *AcMNPV* genome has 9 hrs in different locations; specifically, hr1, hr1a, hr2, hr2a, hr3, hr4a, hr4b, hr4c, and hr5; the hrs mutations were confined to hr2, hr3, and hr5, as shown in Table 4.3. The repetitive nature of the

hrs could lead to sequencing or mapping errors and could be mutation hotspots (Boezen et al., 2022). The remaining mutations could be attributed to errors in the reference genome sequence (Maghodia et al., 2014) or differences in sequence properties of the original *AcMNPV* C6 strain (Ayles et al., 1994) and our rBEV stocks. A FASTA file of the shotgun-sequenced rBEV has been deposited with Borealis, The Canadian Dataverse Repository (<https://doi.org/10.5683/SP3/FIBEX4>, accessed on 20 February 2025). The shotgun-sequenced rBEV was used to design primers for tiled-amplicon sequencing and any sgRNA designed using the online tool CHOPCHOP (<https://chopchop.cbu.uib.no/>, accessed on 22 July 2023) (Labun et al., 2019) was matched with the sequenced genome to confirm the target locations.

Table 4.3: Mutations from shotgun sequencing of p6.9GFP rBEV.

	Mutation Region	Mutation Type	Mutation Length (bp)	AcMNPV C6 Sequence	Shotgun Sequence	^aAmino Acid Change	Position on Shotgun Sequence
	Homologous region 2 (hr2)	Deletion	49	ttaaact ¹	-	Not applicable	27,333
	Homologous region 2 (hr2)	Deletion	137	aactcgc ²	-	Not applicable	27,371
	Homologous region 3 (hr3)	Deletion	72	aatcgtg ³	-	Not applicable	71,257
	Homologous region 3 (hr3)	Deletion	72	caagtag ⁴	-	Not applicable	71,341
	Homologous region 2 (hr2)	Insertion	1	-	C	Not applicable	27,341
	Homologous region 2 (hr2)	SNPs	1	C	G	Not applicable	27,416
66	Homologous region 3 (hr3)	SNPs	1	A	C	Not applicable	71,255
	Homologous region 3 (hr3)	SNPs	1	A	G	Not applicable	71,294
	Homologous region 3 (hr3)	SNPs	1	A	G	Not applicable	71,306
	Homologous region 3 (hr3)	SNPs	1	C	G	Not applicable	71,378
	Homologous region 5 (hr5)	SNPs	1	T	C	Not applicable	115,201
	Homologous region 5 (hr5)	SNPs	1	T	C	Not applicable	115,250
	chiA/v-cath promoter	Deletion	37	tttaatt ⁵	-	Not applicable	105,599
	AcOrf-106 promoter	Insertion	1	-	A	Not applicable	94,054
	AcOrf-106 promoter	Insertion	1	-	A	Not applicable	94,124
	AcOrf-106 promoter	Insertion	1	-	C	Not applicable	94,172
	fgf promoter	SNPs	1	G	T	Not applicable	28,133

Table 4.3 Cont.

Mutation Region	Mutation Type	Mutation Length (bp)	AcMNPV C6 Sequence	Shotgun Sequence	^a Amino Acid Change	Position on Shotgun Sequence
AcOrf-57 promoter	SNPs	1	A	T	Not applicable	47,546
AcOrf-78 promoter	SNPs	1	G	C	Not applicable	65,943
chiA/v-cath promoter	SNPs	7	taaaaaa	cccgggc	Not applicable	105,601
‡lef2/AcOrf-603 recombination region	SNPs	2	CT	TC	Not applicable	3743
<i>pk-1</i>	Deletion	1	T	-	-	8157
<i>AcOrf-17</i>	Deletion	1	A	-	-	14,943
<i>AcOrf-21</i>	Deletion	1	C	-	-	17,024
<i>AcOrf-112</i>	Deletion	1	A	-	-	97,094
<i>ie-01</i>	Deletion	2	CG	-	-	122,767
<i>ie-01</i>	Deletion	1	A	-	-	123,928
<i>pk-1</i>	Insertion	1	-	C	His	8143
<i>egt</i>	Insertion	40	-	ctagaga ⁶	†LEISRDL*RSLEIS	12,428
<i>AcOrf-20</i>	Insertion	3	-	cgg	Pro	16,922
<i>AcOrf-52</i>	Insertion	1	-	G	Ala	44,907
<i>AcOrf-59</i>	Insertion	1	-	A	Asn	48,451
<i>AcOrf-106</i>	Insertion	2	-	CA	ThrAsn	94,260
<i>AcOrf-106</i>	Insertion	1	-	G	Glu	94,293

Table 4.3 Cont.

Mutation Region	Mutation Type	Mutation Length (bp)	AcMNPV C6 Sequence	Shotgun Sequence	^a Amino Acid Change	Position on Shotgun Sequence
<i>AcOrf-106</i>	Insertion	2	-	CG	Pro	94,343
<i>AcOrf-107</i>	Insertion	6	-	atttgg	IleTrp	94,422
<i>AcOrf-107</i>	Insertion	1	-	A	Arg	94,433
<i>pe/pp34</i>	Insertion	1	-	G	Thr	109,299
<i>AcOrf-1629</i>	SNPs	1	G	A	Phe	6371
<i>AcOrf-11</i>	SNPs	1	G	A	Asp	8685
<i>AcOrf-11</i>	SNPs	1	G	A	Trp	9434
<i>AcOrf-20</i>	SNPs	1	A	G	Leu	16,888
<i>AcOrf-20</i>	SNPs	2	GC	CA	Leu	16,925
<i>AcOrf-20</i>	SNPs	2	AG	GA	ProPro	16,930
<i>AcOrf-20</i>	SNPs	2	CG	GC	ProGlu	16,933
<i>env-prot</i>	SNPs	1	T	A	Val	21,007
<i>p47</i>	SNPs	1	A	C	Val	33,392
<i>gta</i>	SNPs	1	G	A	Asn	34,844
<i>gta</i>	SNPs	1	G	T	Phe	34,895
<i>odv-e66</i>	SNPs	1	C	G	Met	37,998
<i>odv-e66</i>	SNPs	1	T	C	Ser	38,304

Table 4.3 Cont.

Mutation Region	Mutation Type	Mutation Length (bp)	AcMNPV C6 Sequence	Shotgun Sequence	^a Amino Acid Change	Position on Shotgun Sequence
<i>AcOrf-51</i>	SNPs	1	A	T	Phe	44,182
<i>AcOrf-53</i>	SNPs	1	C	T	Phe	45,551
<i>AcOrf-54</i>	SNPs	1	T	G	Val	46,633
<i>AcOrf-57</i>	SNPs	1	C	T	Leu	47,770
<i>AcOrf-58</i>	SNPs	1	T	G	Arg	48,249
<i>lef9</i>	SNPs	1	G	A	Asn	50,209
<i>lef9</i>	SNPs	1	G	T	Ser	50,314
<i>tlp</i>	SNPs	1	G	T	Glu	68,066
<i>AcOrf-84</i>	SNPs	1	C	A	Arg	71,492
<i>AcOrf-93</i>	SNPs	1	C	G	Glu	79,916
<i>helicase</i>	SNPs	1	G	C	Leu	81,217
<i>AcOrf-106</i>	SNPs	1	A	C	Asn	94,263
<i>AcOrf-114</i>	SNPs	1	C	G	Gln	98,860
<i>AcOrf-120</i>	SNPs	1	A	G	Asp	102,644
<i>AcOrf-120</i>	SNPs	1	A	G	Leu	102,664
<i>pk-2</i>	SNPs	1	T	A	Phe	103,878
<i>94k</i>	SNPs	2	TA	CC	GluVal	113,302

Table 4.3 Cont.

Mutation Region	Mutation Type	Mutation Length (bp)	AcMNPV C6 Sequence	Shotgun Sequence	^a Amino Acid Change	Position on Shotgun Sequence
<i>p10</i>	SNPs	1	C	G	Ser	116,722
<i>p74</i>	SNPs	2	GT	TG	IleMet	116,950
<i>p74</i>	SNPs	1	A	G	Gln	118,441
<i>p74</i>	SNPs	1	A	T	Glu	118,509
<i>me53</i>	SNPs	1	A	G	Arg	120,034
<i>me53</i>	SNPs	1	A	G	Thr	120,036
<i>ie-0/ie-01</i>	SNPs	1	G	A	Gln	120,596
<i>ie-01</i>	SNPs	1	T	C	Ala	122,767
<i>ie-01</i>	SNPs	1	T	G	Ala	122,768
<i>pe38</i>	SNPs	1	G	C	Arg	131,047

^a Amino acid changes in the coding regions of the shotgun-sequenced rBEV. These are supposed changes as amino acids are not codon-optimized.

[†] The amino acid code LEISRDL*RSLEIS is LeuGluIleSerArgAspLeu**STOP**ArgSerLeuGluIleSer.

[‡] The mutation is not within the coding region of *lef2* or *AcOrf-603*.

Complete deletion and insertion mutation sequences greater than 7 bp are listed here:

¹ ttaaactcgctttacgagtagaaaattctacttgtaacgcatgatcaaggg;

² aactcgctttacgagtagaattctacttgtaacgcacgccaagggatgatgtcatttattgtgcaaagctgatgtcatctttgcacacgattataaacacaatcaaataatgactcatttgttttcaaactg;

³ aatcgtgcgttacaagtagaattctactcgtaaacgagttcgggtttgaaaaacaaatgacatcatttctt;

⁴ caagtagaattctactcgtaaagcgagtttagtttgaaaaacaaatgacatcatctcttgattatgttta;

⁵ ttaatttatcttaattttaagttgtaattatttat;

⁶ ctagagatctctagagatctctagagatctctagagatct.

4.4.2 Tiled-Amplicon Sequencing of p6.9GFP rBEV

Extracted p6.9GFP rBEV gDNA from the multi-well plate's infected-only control was sequenced using the tiled-amplicon sequencing assay, utilizing the shotgun-sequenced rBEV as the reference genome. This assay overcame some of the challenges faced with shotgun sequencing by being more specific and requiring a lower concentration of input DNA. A total of 475,872 read pairs were obtained and 90.97% of them passed the quality control, and 99.75% of the filtered reads aligned with the reference genome. The reads covered 100% of the genome with an average depth of 2006.12 \times . This confirmed that the designed tiled-amplicon primers covered the entire rBEV genome and could be used for future sequencing work. Compared to the reference genome, the more specific tiled-amplicon sequencing revealed a 35 bp deletion in the hr2 (aaatgatgtcattggatgagtcattgttttcaa). A visual representation of sequence alignment compared to the reference genome can be found in Supplementary Figure B.1. The FASTA file of the tiled-amplicon-sequenced p6.9GFP rBEV from the infected-only control has been deposited in Borealis, The Canadian Dataverse Repository (<https://doi.org/10.5683/SP3/FIBEX4>, accessed on 20 February 2025).

4.4.3 Lowering Infectious Baculovirus from Sf9-Cas9 Cells Through Co-Expressing a sgRNA Targeting *gp64*

Sf9-Cas9 cells were infected with the p6.9GFP_sgRNA_gp64+131 rBEV for targeted *gp64* gene disruption. Sf9-Cas9 cells infected with p6.9GFP rBEV served as the control here. Although the rBEVs are slightly different, they carry the same foreign protein, thus enabling a comparison between the targeted and untargeted rBEVs. Using the p6.9GFP_sgRNA_gp64+131

rBEV, the IVT was significantly reduced compared to the control (Figure 4.4b). On the other hand, foreign protein production, as measured by fluorescence intensity, was only slightly lower compared to the control (Figure 4.4a). These data indicate that disrupting *gp64* could minimize budded virus co-production in cell cultures.

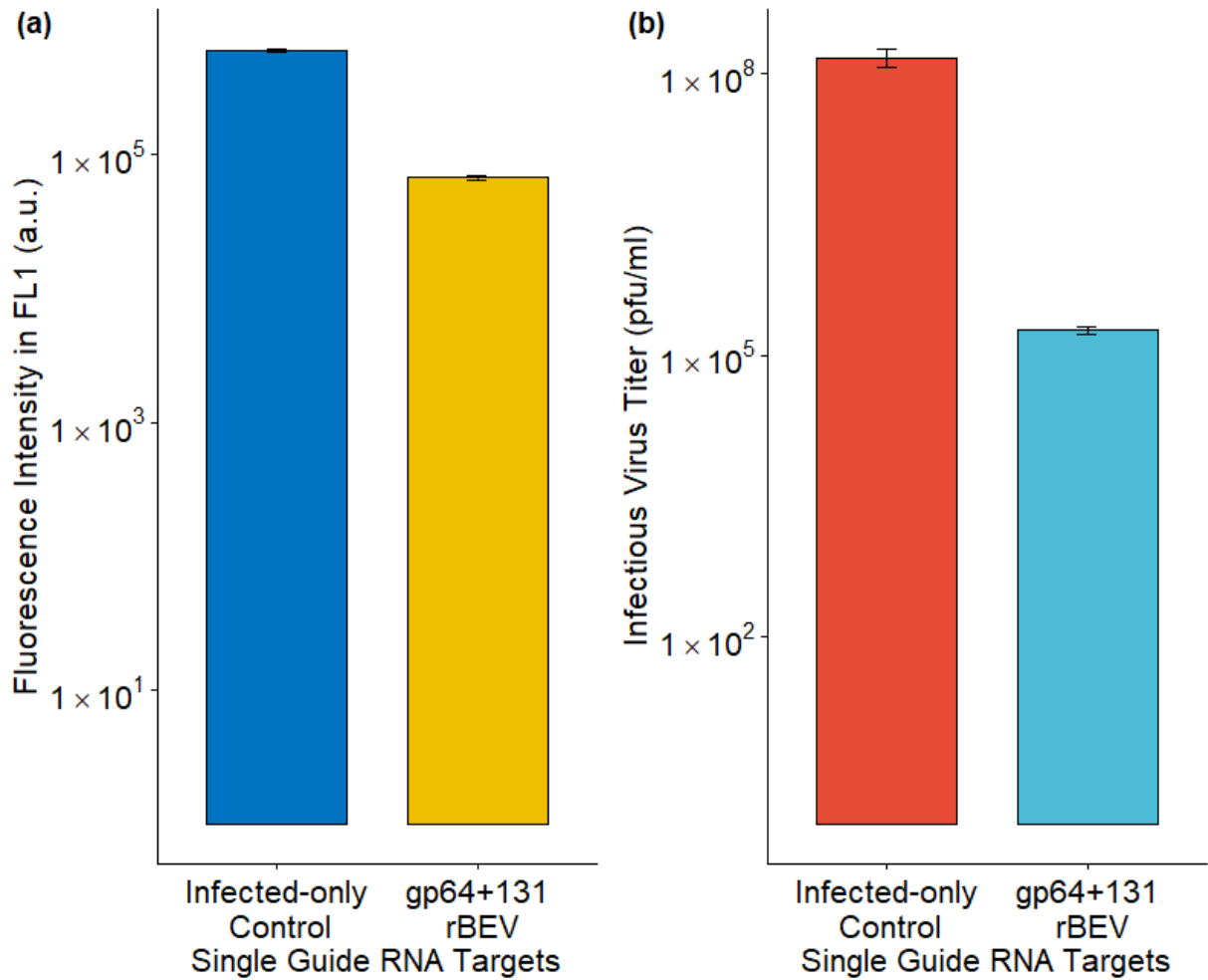


Figure 4.4: Impact of sgRNA rBEV-mediated *gp64* gene disruption on (a) foreign protein production and (b) infectious virus titer. The data from 3 replicates are presented here.

The gDNA of the p6.9GFP_sgRNA_gp64+131 rBEV amplified in Sf9 cells (where no

Cas9 was present) was sequenced by tiled-amplicon sequencing and used as the reference genome for analyzing mutations following the infection of Sf9-Cas9 cells with this rBEV. After *gp64* gene disruption, rBEV gDNA could only be recovered from the cell pellet fraction, thus emphasizing the need for tiled-amplicon sequencing to amplify the specific intracellular genome of interest. Compared to the reference genome, the consensus sequence of the p6.9GFP_sgRNA_gp64+131 rBEV following *gp64* targeting exhibited six SNPs and two indels in various regions of the gDNA (Table 4.4). The SNPs in *hr3*, *AcOrf-11*, and *AcOrf-84* and the two indel mutations within *pk-1* were also observed in the shotgun-sequenced rBEV. However, the CRISPResso2 sequencing pipeline was not able to detect any mutation within the *gp64* gene, even though phenotypic changes were observed. The FASTA files of the untargeted and targeted p6.9GFP_sgRNA_gp64+131 rBEV, and the detailed report generated by CRISPResso2 for the targeted rBEV can be accessed via Borealis, The Canadian Dataverse Repository (<https://doi.org/10.5683/SP3/FIBEX4>, accessed on 20 February 2025).

Table 4.4: Changes observed in the gDNA of p6.9GFP_sgRNA_gp64+131 rBEV upon propagation in Sf9-Cas9 cells via tiled-amplicon sequencing assay.

Mutation Region	Mutation Type	Mutation Length (bp)	Reference/Targeted Genome ¹	Position on Targeted rBEV Sequence
Homologous region 2	SNPs	2	CC/AT	28,491
Homologous region 3	SNPs	1	C/G	72,827
AcOrf-20 promoter	SNPs	1	C/G	17,961
<i>AcOrf-1629</i>	SNPs	1	A/G	7427
<i>pk-1</i>	Deletion	1	T/-	9210
<i>pk-1</i>	Insertion	1	-/C	9195
<i>AcOrf-11</i>	SNPs	1	G/A	9737
<i>AcOrf-84</i>	SNPs	1	C/A	72,941

¹ Changes in nucleotides between the p6.9GFP_sgRNA_gp64+131 rBEV reference genome (amplified in Sf9 cells) on the left and the p6.9GFP_sgRNA_gp64+131 rBEV targeted genome (amplified in Sf9-Cas9 cells) on the right.

4.4.4 Tiled-Amplicon Sequencing of p6.9GFP rBEV upon T-I Assay

The extracted p6.9GFP rBEV gDNA from the multi-well plate's scrambled control (cells transfected with a plasmid coding for a random sgRNA that did not target any baculovirus gene) was sequenced using the tiled-amplicon sequencing assay and compared to the shotgun-sequenced rBEV, which was used as the reference genome. A total of 966,391 read pairs were obtained, with 94.59% of the reads passing the quality filter thresholds while 99.65% of those reads aligned with the reference genome. The reads covered 100% of the genome with an average depth of 4063.722 \times . Compared to the reference genome, the tiled-amplicon sequencing revealed a 35 bp deletion in the hr2 (aaatgatgtcattggatgagtcattgttttcaa), a 20 bp deletion in the *egt* gene (ctagagatctctagagatct), and a 1 bp SNP in

the hr3 (A > G) (Table 4.5). Since the hrs have highly repetitive sequences, making these regions difficult to amplify and sequence, it is not conclusive whether the 1 bp SNP is a true mutation or a sequencing artifact. The FASTA file of the tiled-amplicon-sequenced p6.9GFP rBEV from the scrambled control can be accessed from Borealis, The Canadian Dataverse Repository (<https://doi.org/10.5683/SP3/FIBEX4>, accessed on 20 February 2025).

Table 4.5: Changes in the gDNA of p6.9GFP rBEV obtained from tiled-amplicon sequencing upon T-I assay.

Mutation Region	Mutation Type	Mutation Length (bp)	<i>AcMNPV</i> C6/Scrambled Sequence ¹
Homologous region 2	Deletion	35	aaatgat ² /-
Homologous region 3	SNPs	1	A/G
<i>egt</i>	Deletion	20	ctagaga ³ /-

¹ Changes in nucleotides between the *AcMNPV* C6 strain-based reference genome (shotgun-sequenced rBEV) on the left and the scrambled control on the right; ² aaatgatgtcattggatgagtcattgttttcaa, the complete 35 bp deletion sequence observed in the hr2; ³ ctagagatctctagatct, and the complete 20 bp deletion sequence observed in the *egt* gene.

4.4.5 Lowering Infectious Budded Virus from Sf9-Cas9 Cells Through Plasmid-Based Delivery of sgRNA Targeting *gp64*

The *AcMNPV* major glycoprotein gene, *gp64*, was disrupted by the CRISPR-Cas9-based T-I assay (Bruder, Aucoin, 2023b). While the fluorescence intensity, which represented GFP production, was similar to the control, the IVT was reduced upon *gp64* disruption (Figure 4.5).

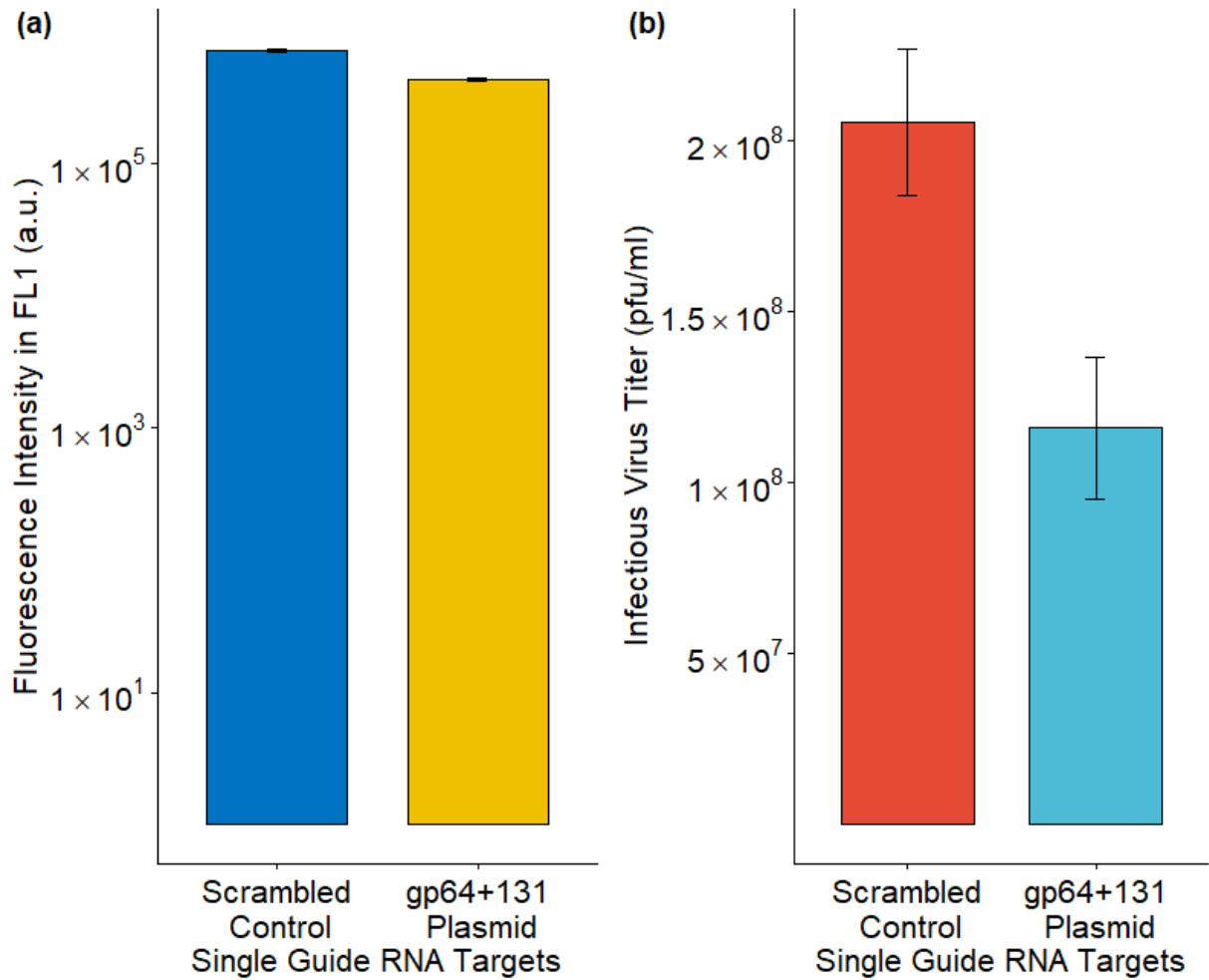


Figure 4.5: Effect on (a) foreign protein production and (b) infectious virus titer upon CRISPR-Cas9 based T-I assay. The data from three replicates are presented here.

The extracted gDNA of the p6.9GFP rBEV upon *gp64* gene disruption was sequenced by tiled-amplicon sequencing and compared to the shotgun-sequenced rBEV as the reference genome. The 35 and 20 bp deletions observed in the p6.9GFP rBEV genome from the scrambled control were also detected in the consensus sequence of the gp64+131 targeted p6.9GFP rBEV gDNA. Across the *gp64* gene, an average read depth of 156 was

obtained (with a minimum read depth of 100) using the Pilon/DNAdiff pipeline. The Pilon/DNAdiff pipeline alone did not reveal any mutations within *gp64*; however, utilizing the CRISPResso2 tool, 23 reads were aligned with the provided *gp64* reference amplicon sequence, and 4 (17.39%) of those reads revealed indel mutations within the *gp64* target region. Table 4.6 shows the mutations observed within *gp64* for each of the four reads (alleles). The FASTA file of the targeted p6.9GFP rBEV and the detailed report generated by CRISPResso2 have been deposited in Borealis, The Canadian Dataverse Repository (<https://doi.org/10.5683/SP3/FIBEX4>, accessed on 20 February 2025).

Table 4.6: Mutations within the *gp64* gene from tiled-amplicon sequencing of p6.9GFP rBEV upon *gp64* gene disruption.

Read	Mutation Region	Mutation Type	Mutation Length (bp)	Reference/Targeted <i>gp64</i> Sequence	Position on Targeted gDNA
Read 1	<i>gp64</i>	Deletion	12	gtccttttgcag/-	108,669
Read 2	<i>gp64</i>	Deletion	2	CC/-	108,671
Read 3	<i>gp64</i>	Deletion	1	C/-	108,671
Read 4	<i>gp64</i>	Deletion	2	GT/-	108,669

4.5 Discussion

AcMNPV is the most widely used baculovirus of the BEVS for the production of recombinant proteins, VLPs, vaccines, and other biologics (Pushko et al., 2005, 2010; Aucoin et al., 2010). While sequence data are available for the original *AcMNPV* C6 (Ayres et al., 1994) and E2 (Maghodia et al., 2014) strains, to the best of our knowledge, there is no

information on the whole-genome sequence of the rBEVs constructed with an *AcMNPV* DNA backbone and transfer plasmid. The major goal of this work was to generate the consensus sequence of the p6.9GFP rBEV and develop a pipeline to efficiently confirm the mutations upon CRISPR-Cas9 targeting. To achieve these tasks, two different sequencing methods were employed—shotgun sequencing and a tiled-amplicon sequencing assay. Although shotgun sequencing is fast and cost-effective compared to traditional approaches such as Sanger sequencing, it has limitations, including the requirement for a high concentration of starting material and difficulties in resolving repetitive regions, leading to assembly errors or gaps. On the other hand, the more specific tiled-amplicon sequencing assay with low DNA input requirements can encounter challenges, including dependence on a reference genome to design primers and the inability to amplify regions if mutations exist in primer-binding sites.

Baculoviruses with large dsDNA genomes have high mutation rates despite the known proofreading activity of its DNA polymerase (Tomalski et al., 1988; Boezen et al., 2022; Chateigner et al., 2015). A recent study was the first to report on the empirical estimates of mutation rates for *AcMNPV* after five serial passages in insect larvae and estimated a mutation rate of 1×10^{-7} s/n/r (mutation rate per base per strand copying) when the most stringent criteria for mutation calling were applied (Boezen et al., 2022). In this study, we used shotgun sequencing to confirm the sequence of the p6.9GFP rBEV generated in our laboratory: an rBEV constructed utilizing an *AcMNPV* DNA backbone derived from the C6 strain and a transfer plasmid carrying GFP under the p6.9 promoter (Bruder, Aucoin, 2023b). This data is especially important when targeted genome editing is used to probe the rBEV genome. Our shotgun-sequenced rBEV consensus genome is 99.7% similar to the

reference genome. The major differences with large indels between the reference genome and the shotgun-sequenced rBEV occurred mainly in the hrs, which are known to be prone to mutations (Chateigner et al., 2015; Boezen et al., 2022).

4.5.1 Differences Compared to Reported Genomes from Shotgun Sequencing

Small indels and SNPs that were observed in genes/ORFs, such as *AcOrf-17*, *AcOrf-20*, *AcOrf-21*, *AcOrf-58*, *AcOrf-59*, *AcOrf-106*, *AcOrf-107*, *AcOrf-112*, and *pe/pp34*, within our shotgun-sequenced rBEV were also seen in the *AcMNPV* WP10 (Wild Population 2010) consensus sequence compared to the *AcMNPV* C6 genome (Chateigner et al., 2015). Our shotgun data also revealed that only when the original ORFs, *AcOrf-20/21*, *58/59*, *106/107*, and *112/113* were combined into single ORFs, and the original ORFs, *AcOrf-17*, *AcOrf-52*, and *AcOrf-131* (*pe/pp34*) were stretched to a greater length, did stop codons appear at the appropriate locations. This is in line with reports that these ORFs should be combined and are not consistent with the original ORF classification (Ayres et al., 1994; Rohrmann, 2019a; Harrison, Bonning, 2003).

A 40 bp insertion region in *egt*, consisting of four ten-base repeats, included a premature stop codon and resulted in a frameshift that caused multiple stop codons to be detected downstream of the insertion. It was previously reported that *egt* deletions commonly occurred upon baculovirus propagation in cell culture and are not essential for in vitro replication (Kumar, Miller, 1987; Cory et al., 2004). Other SNPs and small indels in our shotgun-sequenced rBEV are possibly uniquely related to the virus stock in question,

emphasizing the need for whole-genome sequencing before targeted gene editing. Shotgun sequencing provided a consensus sequence for the p6.9GFP rBEV used in this study.

4.5.2 Tiled-Amplicon Sequencing to Assess CRISPR-Cas9 Editing

The CRISPR-Cas9 genome-editing technology has gained traction in the last decade. Targeted gene disruption in *AcMNPV* rBEVs was achieved by co-transfection of *AcMNPV* gDNA with Cas9/sgRNA RNP complexes and is the first report to establish the potential of CRISPR-Cas9 for baculovirus gene editing to improve the baculovirus as an expression vector and a biopesticide (Pazmiño-Ibarra et al., 2019). A more recent study focused on adapting the CRISPR-Cas9 technology to *AcMNPV*/Sf9 cells and comparing gene disruption and transcriptional repression in *AcMNPV* vectors via CRISPR-Cas9/dCas9 (Bruder et al., 2021). In another study, a sensitive CRISPR-Cas9-based transfection-infection assay capable of efficiently scrutinizing the *AcMNPV* genome by targeted gene disruption was developed to aid optimization of the BEVS for improved biologics production (Bruder, Aucoin, 2023b). Previous studies have portrayed that CRISPR-Cas9 mediated targeting of *gp64* reduced the co-production of baculovirus in cell cultures (Bruder, Aucoin, 2023a; Hausjell et al., 2023). In our work, we observed that disruption of *gp64* by either infection of Sf9-Cas9 cells with the p6.9GFP_sgRNA_gp64+131 rBEV or transfection of Sf9-Cas9 cells with the gp64+131 plasmid followed by infection with the p6.9GFP rBEV resulted in significantly greater reduction in IVT compared to the effect on GFP production. These data confirmed that *gp64* gene disruption has the potential to minimize baculovirus con-

tamination in cell cultures.

Although CRISPR-Cas9 can be effectively applied for targeted gene editing (Bruder, Aucoin, 2023b,a), changes in phenotypes alone cannot confirm the cut sites or any off-target effects. Some studies have utilized the previously described T7 endonuclease I assay to detect mutations only in the target regions (Liu et al., 2014; Xiang et al., 2024). Other studies confirmed CRISPR-Cas9 mutations by PCR-amplifying the target sites and ligating them into cloning vectors which are then sequenced by M13 primers (Dong et al., 2016, 2018). However, these methods do not provide a comprehensive genome-wide analysis of any unintended effects. Whole-genome sequencing can compare the targeted genomes with the reference genome to screen for genome-wide off-targets and identify SNPs and indel mutations. In this work, we adapted the tiled-amplicon sequencing assay (Quick et al., 2017) to *AcMNPV* vectors to sequence the whole rBEV genome upon CRISPR-Cas9 targeting. We postulated that this sequencing protocol could be widely utilized for rBEVs, and sequenced the p6.9GFP and the p6.9GFP_sgRNA_gp64+131 rBEV genomes from the multi-well plate experiments and recovered in the cell pellet.

4.5.2.1 Sequencing Infected-Only and Scrambled Controls

A 100% coverage of the p6.9GFP rBEV genomes from the infected-only and scrambled controls was achieved with the tiled-amplicon primers generated for the reference genome (shotgun-sequenced rBEV) using PrimalScheme (Quick et al., 2017). Compared to the reference genome, the p6.9GFP rBEV genomes from the two controls had a 35 bp deletion in one of the hrs (hr2). Previous studies have shown that deletion of individual hrs or combinations of up to 5 hrs does not significantly affect virus replication in cell cultures (Carstens,

Wu, 2007; Bossert, Carstens, 2018). Thus, this 35 bp deletion in the hrs presumably does not affect rBEV replication. This deletion is believed to have not been detected by shotgun sequencing due to a reference genome assembly error. The tiled-amplicon sequencing assay, being more precise, was able to detect these deletions in the p6.9GFP rBEV genomes recovered from infected-only and scrambled controls. Additionally, the consensus sequence of the gp64+131 targeted p6.9GFP rBEV also had the same deletions seen in the controls, providing additional evidence that the deletion was only detected using the tiled-amplicon sequencing. The scrambled control differed from the infected-only control, with a shorter insertion in *egt*. Whereas the shotgun sequencing and the infected-only control showed a 40 bp insertion consisting of four ten-base repeats (5'-3' ctagagatct), the scrambled control insertion consisted of only two of the same ten-base repeats. The consensus sequence of the gp64+131 targeted p6.9GFP rBEV also only had the 20 bp insertion of the same repeats.

4.5.2.2 CRISPR-Cas9 Mutations Using p6.9GFP_sgRNA_gp64+131 rBEV in Sf9-Cas9 Cells

Phenotypically, the assay had a significant effect on the production of baculovirus when the p6.9GFP_sgRNA_gp64+131 rBEV was propagated in Sf9-Cas9 cells; however, as a result, the quantity of baculovirus in the supernatant was reduced, complicating viral gDNA recovery and sequencing. On the other hand, genomes recovered from the cell pellet underwent sequencing and revealed no mutations in *gp64* using Pilon/DNAdiff or CRISPResso2. It is not clear what happens to viral gDNA that does not undergo DNA repair once cleaved with a double-stranded break. The DNA damage response might signal the degradation of the viral genome via nucleases; however, we provide no evidence of this at this time.

In this system, there is competition between the cleaving of the viral DNA and the replication and packaging of genomes. It is possible that the rate of replication is faster than the rate of targeting and cleaving of the viral genomes. The level of Cas9 in this system decreases with time post-infection because Cas9 is driven by an *OpIE2* promoter, which is naturally downregulated as the infection progresses (Bruder et al., 2021). Furthermore, there will be a significant pool of genomes that the cell can use for homologous recombination, which could avoid mutations yet slow down the replication and packaging of viral genomes. In an attempt to find possible mutations, less stringent Trimmomatic parameters were used (SLIDINGWINDOW: 4:20, MINLEN: 170, LEADING: 15, TRAILING: 10) along with Pilon/DNAdiff on the sequencing data for both the reference and targeted p6.9GFP_sgRNA_gp64+131 rBEV genomes. The result of this analysis showed a duplication of the gene segments from *AcOrf-119* to *pe/pp34*, (which includes the *gp64* gene) in the assembled genomes. Within the first segment in p6.9GFP_sgRNA_gp64+131 rBEV, a large deletion (>50 bp) within the gp64+131 target region was found. Within the repeated region, and in p6.9GFP_sgRNA_gp64+131 rBEV, a large insertion (>50 bp) within the gp64+131 target region was observed. Although the duplication was seen in both the p6.9GFP_sgRNA_gp64+131 rBEV genome and the reference genome, the deletion and insertion were unique to the p6.9GFP_sgRNA_gp64+131 rBEV genome (propagated in the Sf9-Cas9 cells). Using the less stringent Trimmomatic parameters is not recommended, and should not be taken as hard evidence; however, in the absence of any other evidence, it points to the effect of the CRISPR-Cas9 editing.

4.5.2.3 CRISPR-Cas9 Mutations Using a gp64+131 Plasmid in Sf9-Cas9 Cells with the T-I Assay

CRISPR-Cas9-mediated targeting can be utilized to identify genes that are not required for foreign protein production in cell culture and their disruption or removal can improve foreign protein yields (Bruder, Aucoin, 2023b). Consistent with previous findings, using sgRNA plasmid transfection followed by infection yields a detectable effect that is not as large as what is observed if the sgRNA is provided by the baculovirus vector itself. This is understandable because uptake through transfection is not as efficient as with baculovirus vectors, nor do the plasmids replicate within the cell. It is therefore expected that some cells may not carry any sgRNA (even though 3.38×10^5 plasmids per cell are used). Compared to CRISPR genomic editing of a single target within the cell, similar to the case of using the p6.9GFP_sgRNA_gp64+131 rBEV, we target a baculovirus gene while the virus is going through a productive infection cycle. All these factors combined make it difficult to investigate targeted genomes in a pool of viral gDNAs that either (1) bypassed CRISPR-mediated targeting by infecting cells that remained untransfected (no sgRNA); or (2) were repaired by homologous recombination using the untargeted genomes as a template. We still attempted to analyze CRISPR-mediated mutations within the targeted *gp64* gene using the CRISPResso2 pipeline. It is to be noted that our raw reads came from whole-genome sequencing of the rBEV and not just the regions flanking the target sequence. CRISPResso2 preprocessed 96.09% of the raw reads that passed the quality control. Overall, 82.09% of the reads also passed the CRISPResso2 filters (preprocessing) before being aligned with the reference amplicon. Since the tool was run from the Galaxy server, some parameters, such as the minimum alignment score (default: 60), were

masked and could not be modified. The stringent filtering criteria applied by CRISPResso2 probably filtered out reads that could otherwise be aligned with the reference amplicon. This analysis resulted in a low abundance of reads (23 reads) that aligned with the *gp64* reference amplicon sequence, with deletions detected within the *gp64* target region for 4 (17.39%) of those reads. A detailed information file (info.pickle) and running log of this tool can be accessed via the CRISPResso2_p6.9GFP_rBEV_sgRNAPlasmid_gp64 report, which can be obtained from Borealis.

4.5.3 Concluding Remarks

Our findings suggest that having a comprehensive knowledge of the sequence of the rBEV is essential to ensure specific sequences (genes) within its genome are being targeted, CRISPR-Cas9-mediated *gp64* gene disruption can be used as an effective tool to reduce budded virus contamination in cell cultures, and the bioinformatics pipeline can detect CRISPR-Cas9-mediated mutations within the target region. Nonetheless, the shotgun sequencing pipeline outlined here can be applied to generate a consensus sequence of other recombinant viruses for which only a best-guess sequence has been identified. Additionally, the tiled-amplicon sequencing assay pipeline can be applied to other viruses for which tiled-amplicon primers can be confidently designed; that is, the genome sequence is available. This sequencing assay being more specific and requiring a lower starting material concentration makes it easier to investigate the genomes of different viruses upon gene editing or amplification in cell cultures.

Chapter 5

Detection of variants upon transient CRISPR-Cas9 targeting of *gp64*

The baculovirus *gp64* gene codes for the envelope fusion protein, GP64, that aids virus entry into cells (Pearson, Rohrmann, 2002; Rohrmann, 2019a). A drawback of the BEVS is the co-production of rBEVs with products of interest in the cell culture supernatant. Strategies have been devised to minimize BV contamination in cell culture by targeting or eliminating baculovirus structural genes, such as *gp64* (Marek et al., 2011; Bruder, Aucoin, 2023a,b; Chakraborty et al., 2025; Hausjell et al., 2023). As previously mentioned, transient or *in Process* CRISPR-Cas9 genome engineering of rBEVs, which involves CRISPR-Cas9-mediated AcMNPV gene disruption while the rBEV undergoes its infection cycle in Sf9-Cas9 cells, makes sequencing of the targeted genomes difficult. We still decided to investigate both phenotypic and intracellular and extracellular genomic changes upon targeted AcMNPV gene disruption. Here, we use a previously developed CRISPR-Cas9-

based T-I assay to target *gp64* at six different locations and report reduced infectious and total viral titers while maintaining or enhancing GFP production from the late p6.9 promoter. We also use tiled-amplicon sequencing to investigate the control and targeted rBEV genomes recovered from the cell pellet and supernatant fractions upon T-I assay, and describe a variant calling pipeline to successfully detect low-frequency mutations. Furthermore, we report that variants present in the virus stock can be conserved over viral propagation in cell culture. The occurrence of these mutations as variants (instead of the consensus sequence) upon virus amplification suggests that they are not detrimental to viral fitness. The research presented in this chapter has been published in *International Journal of Molecular Sciences* and the structure is as per the journal requirements.

Baculovirus Variant Detection from Transient CRISPR-Cas9-Mediated Disruption of *gp64* at Different Gene Locations

Madhuja Chakraborty¹, Lisa Nielsen^{1,2}, Delaney Nash², Mark R. Bruder¹; Jozef I. Nissimov², Trevor C. Charles² and Marc G. Aucoin¹

¹ Department of Chemical Engineering, University of Waterloo

² Department of Biology, University of Waterloo

5.1 Abstract

The Baculovirus Expression Vector System (BEVS) is an important protein and complex biologics production platform. The baculovirus GP64 protein is the major envelope glycoprotein that aids in virus entry and is required for cell-to-cell transmission in cell culture. Several studies have developed strategies around *gp64* gene disruption in an attempt to minimize baculovirus co-production. Here, we investigate the result of transiently targeting the baculovirus *gp64* gene with CRISPR-Cas9 during infection. Because not all genomes are effectively disrupted, we describe a variant calling methodology that allows the detection of the targeted mutations in *gp64* even though these mutations are not the dominant sequences. Using a transfection-infection assay (T-I assay), the AcMNPV *gp64* gene was targeted at six different locations to evaluate the effects of single and multiple targeting sites, and we demonstrated a reduction in the levels of baculovirus vectors while maintaining or enhancing foreign protein production when protein was driven by a p6.9 promoter.

Viral genomes were subsequently isolated from the supernatant and cell pellet fractions, and our sequencing pipeline successfully detected indel mutations within *gp64* for most of the single-guide RNA (sgRNA) targets. We also observed that 68.8% of variants found in the virus stock were conserved upon virus propagation in cell culture, thus indicating that they are not detrimental to viral fitness. This work provides a comprehensive assessment of CRISPR-Cas9 genome editing of baculovirus vectors, with potential applications in enhancing the efficiency of the BEVS.

5.2 Introduction

The baculovirus expression vector system (BEVS), consisting of an insect cell line and recombinant baculovirus expression vectors (rBEVs), was first utilized for the production of biologically active human interferon β (Smith et al., 1983b) in the 1980s. This study paved the path for numerous applications of the BEVS, including the production of recombinant proteins, commercial vaccines, virus-like particles (VLPs), adeno-associated virus-vectors, and even its use as a BacMam viral vector for gene therapy (Urabe et al., 2002; Kost et al., 2005; Airene et al., 2013; Felberbaum, 2015; van Oers et al., 2015; Yee et al., 2018). The ability of insect cells to perform mammalian-like post-translational modifications (PTMs), the inherent biosafety of baculoviruses, the ability of baculoviruses to accept large DNA insertions, and the resulting overexpression of genes inserted in the *polyhedrin* locus of the rBEV, make the BEVS an attractive platform for laboratory research and industrial production (Zhang et al., 2023; van Oers et al., 2015). Among baculoviruses, *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) is the most widely used baculovirus

in the BEVS (Rohrmann, 2019a). The AcMNPV C6 strain was the first baculovirus to be fully sequenced (Ayres et al., 1994). It has a double-stranded DNA (dsDNA) genome of approximately 134 kbp in size and 155 open reading frames (ORFs) (Ayres et al., 1994; Miele et al., 2011). Despite the widespread application and numerous advantages of the BEVS, the co-production of recombinant baculoviruses in insect cell cultures complicates the separation and purification of many desired products.

The budded virus (BV) phenotype is responsible for baculovirus propagation in cell cultures (Wang et al., 2010) and the GP64 glycoprotein, found in Group I nucleopolyhedroviruses (NPVs) such as the AcMNPV, is used as the membrane fusion protein for virus entry (Pearson, Rohrmann, 2002; Rohrmann, 2019a). This low pH-activated protein is one of the most abundant proteins in the AcMNPV BV envelope, and its deletion impairs the budding of progeny viruses from cells (Blissard, Wenz, 1992; Guo et al., 2024; Wang et al., 2010; Oomens, Blissard, 1999; Monsma et al., 1996). Strategies for eliminating GP64 using a transcomplementing cell line strategy (Marek et al., 2011) has not seen widespread use, likely due to problems in generating high titer stocks needed for the infection process. Recently, CRISPR-Cas9 disruption of *gp64* has been used to reduce progeny virus production in cell cultures, improving the quality of the resulting downstream processing feed (Bruder, Aucoin, 2023a,b; Chakraborty et al., 2025; Hausjell et al., 2023).

More recently, we reported on the consensus sequence of an rBEV that contains a green fluorescent protein (GFP) gene that is used in CRISPR-Cas9 gene-targeting studies (Chakraborty et al., 2025). However, genomes transiently targeted during an infection do not result in changes to a majority of the genomes. This may be in part due to an imbalance in the kinetics of baculovirus replication and single-guide RNA (sgRNA)-

Cas9 targeting-cleaving, or may be more directly related to reduced *cas9* expression post-infection because the expression is driven by an OpIE2 promoter (Bruder et al., 2021). Transiently targeted baculovirus genomes are therefore more likely to behave like virus variants in human populations. Here, we aim to demonstrate (through sequencing) that the phenotypic changes observed are due to CRISPR-Cas9-mediated targeted disruption and that the *gp64* gene can be targeted at multiple locations. We also describe a variant calling bioinformatics pipeline that successfully confirms mutations within the targeted *gp64* gene. Additionally, while investigating the variants, we observed that there are certain regions, such as the homologous repeat regions (hrs), where mutations are consistently present; that is, they are carried over as variants between viral passages.

5.3 Results

5.3.1 Effect of *gp64* Gene Disruption on Foreign Protein and Progeny Virus Production

The CRISPR-Cas9-based transfection-infection assay (T-I assay) (Bruder, Aucoin, 2023b) was used to disrupt the AcMNPV *gp64* gene at different locations using single or dual sgRNA targets, with one or two *gp64* targeting spacer sequences, respectively. Disruption of the *gp64* gene resulted in similar fluorescence for all six targets compared to the scrambled control. While the percentage of cells expressing GFP was enhanced for most of the targets (*gp64*+131, *gp64*-160, *gp64*+378, *gp64*+418, and *gp64*+131/384), it was slightly lower (not significant) for the *gp64*+278 target as compared to the control (Figure 5.1).

On the other hand, disrupting *gp64* reduced the infectious virus titer (IVT) for all the targets as expected, with the maximum decrease observed for the dual sgRNA target (Figure 5.2a). Additionally, the total viral particles, including infectious and non-infectious virions, dropped upon *gp64* disruption compared to the scrambled control (Figure 5.2b).

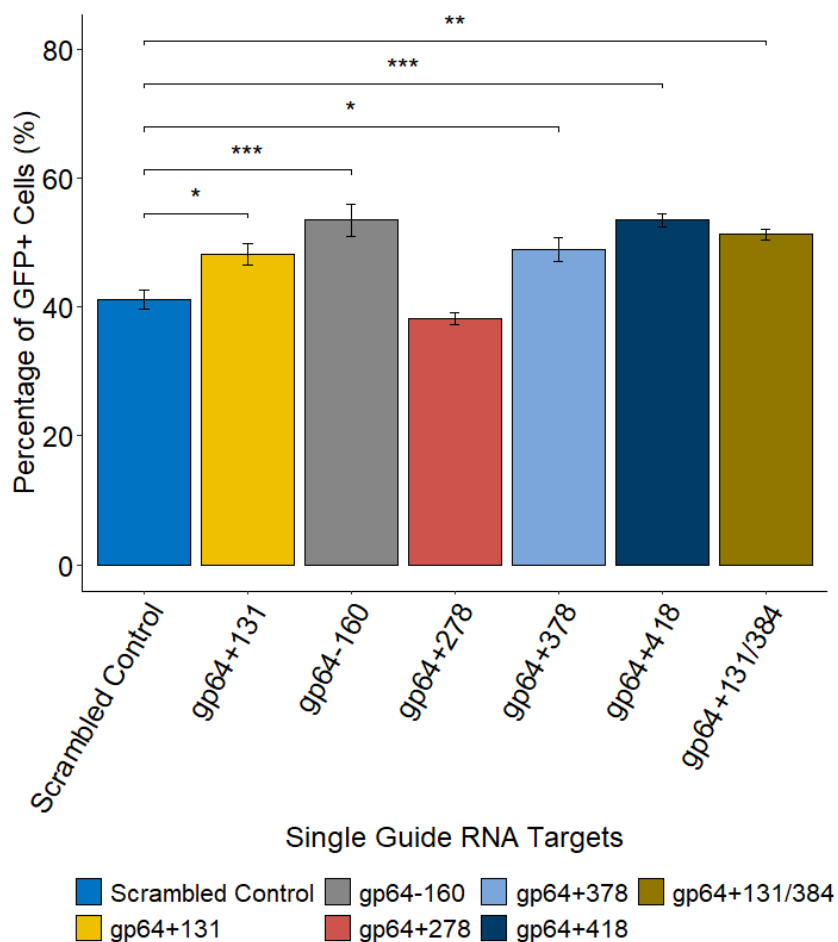


Figure 5.1: Effect of CRISPR-Cas9-mediated gene disruption of the AcMNPV *gp64* gene at single or dual locations on GFP production. The high fluorescence bin data from 3 biological replicates for each target and the scrambled control are shown, where the error bars represent the mean \pm standard error (within groups). ANOVA with Dunnett's test was performed to compare each target to the scrambled control. ***, p -value < 0.001 ; **, p -value < 0.01 ; and *, p -value < 0.05 .

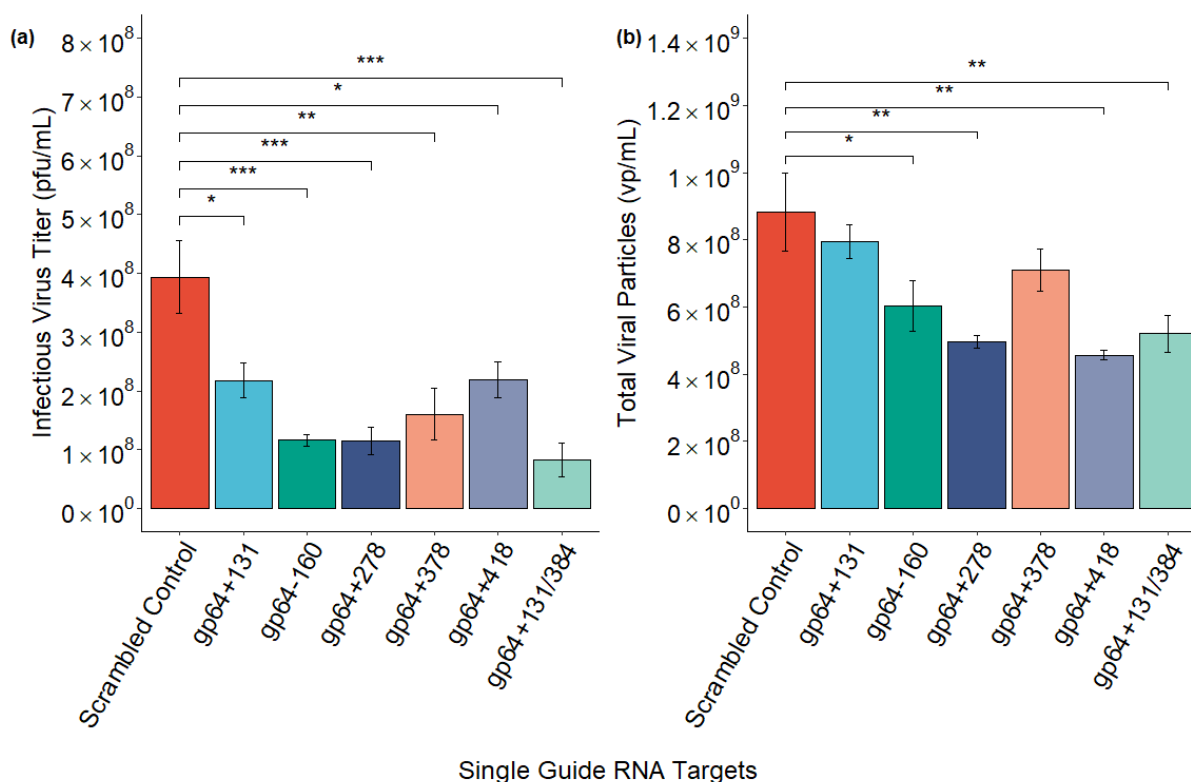


Figure 5.2: Impact of CRISPR-Cas9 targeting of the AcMNPV *gp64* gene using single or dual sgRNA targets on (a) infectious virus titer (IVT) and (b) total viral particles. Data from 3 biological replicates for the scrambled control and each target are presented with error bars for the mean \pm standard error (within groups). ANOVA with Dunnett's test compared each target to the scrambled control. ***, p -value < 0.001; **, p -value < 0.01; and *, p -value < 0.05.

5.3.2 Confirmation of CRISPR-Cas9-Mediated Gene Editing of *gp64*

CRISPR-Cas9 targeted rBEV genomes were analyzed to confirm the cut sites and mutations within the *gp64* gene. A shotgun-sequenced rBEV genome (Chakraborty et al.,

2025) served as the reference genome for data analysis. Table 5.1 gives an overview of the estimated and observed indel positions in *gp64* as well as the observed *gp64* mutations and variant frequencies. While observed mutation positions fell within the expected range for targets gp64+131, gp64-160, gp64+378, and gp64+418, no mutations were detected within the target region for gp64+278 and the dual gp64+131/384 (Table 5.1). Interestingly, for target gp64-160, true mutations (Fisher’s exact test p -value ≤ 0.05) were identified at multiple sites in the estimated range, which could potentially be variants of this gene disruption.

Specifically, targeted disruption of *gp64* resulted in deletion mutations within the *gp64* gene for most of the sgRNA targets (Table 5.1). Importantly, no mutations were observed in the *gp64* gene for both scrambled and infected-only controls. A close look at the deletion mutations within *gp64* revealed that the same mutations occurred in the intracellular (cell pellet) and extracellular (supernatant) fractions for each target except gp64+131 and gp64-160 (Table 5.1). The gp64-160 target led to three additional deletions within *gp64* when the gDNA was extracted from the supernatant compared to the cell pellet, thus resulting in a total of five possible variants. For the gp64+131 target, a deletion appeared in the cell pellet fraction, but no mutation was detected in the supernatant fraction by our sequencing pipeline. Table 5.1 also shows the frequencies at which each of these observed mutations occurred in the cell pellet and supernatant fractions of the *gp64* targeted genomes. While the lowest detected frequency was 1.43%, the maximum frequency was 8.47%, both corresponding to variants obtained from targeting *gp64* with the gp64-160 target.

Table 5.1: Overview of true deletion mutations (Fisher’s exact test p-value ≤ 0.05) and corresponding variant frequencies at different locations within the *gp64* gene of the p6.9GFP rBEV genomes recovered from the CRISPR-Cas9-based T-I assay.

Target ¹	Estimated position (bp) ²	Observed position (bp) [mutation] (Frequency) ³	
		Cell pellet	Supernatant
gp64+131	107,222 - 107,227	107,222 [-C] (6.62%)	Not observed
gp64-160	107,199 - 107,204		107,199 [-CTCCA] (3.05%)
			107,200 [-TCCA] (3.48%)
		107,201 [-CCA] (6.30%)	107,201 [-CCA] (5.44%)
		107,202 [-CA] (8.47%)	107,202 [-CA] (6.34%)
			107,204 [-C] (1.43%)
gp64+278	107,075 - 107,080	Not observed	Not observed
gp64+378	106,975 - 106,980	106,975 [-T] (4.51%)	106,975 [-T] (4.04%)
gp64+418	106,935 - 106,940	106,935 [-CA] (4.24%)	106,935 [-CA] (2.17%)
gp64+131/384	107,222 - 107,227	Not observed	Not observed
	106,969 - 106,974		

¹ sgRNA targeting *gp64* at a specific location within the *gp64* gene; ² Position of 1-5 bp indel mutations in *gp64* that could occur 3-4 bp upstream of the PAM upon CRISPR-Cas9 targeting; ³ Position of true mutations within *gp64*, true deletion mutations within *gp64*, and frequencies at which those true mutations occur in both cell pellet and supernatant fractions, as obtained from tiled-amplicon sequencing.

5.3.3 Variant Generation or Conservation over Passages

To investigate whether variants are a product of virus amplification in cell culture or are conserved over passages, we ran the variant calling pipeline for the p6.9GFP rBEV amplified in Sf9 cells in suspension, utilizing the consensus sequence of the same rBEV (shotgun-sequenced rBEV) (Chakraborty et al., 2025) as the reference genome. Compared to the reference genome, 141 true mutations were detected when the rBEV was analyzed for variants (hereby referred to as virus stock variants). It is to be noted that upon variant analysis of the rBEV, most of the mutations were in the three hrs, hr1, hr2, and hr3, with no mutations detected within the *gp64* gene. While 44 mutations were unique to the virus stock, 97 mutations were carried over during the T-I assay, as observed in the genomes recovered from the T-I assay, which were sequenced by the tiled-amplicon sequencing assay and analyzed by the variant calling pipeline. Specifically, the majority of the conserved mutations in the T-I assay variants were in the different hrs (hr1, hr2, and hr3), with only 14 of them residing outside these repetitive regions (Table 5.2). The common variants observed in both the shotgun sequenced virus stock and the tiled-amplicon-sequenced gDNAs from the T-I assay can be considered variants that are preserved upon virus propagation in Sf9-Cas9 cells, and they may not be detrimental to virus fitness, as they were observed again as variants.

Table 5.2: Conserved variants found in different samples outside the homologous repeat regions (hrs).

Mutation Region	Mutation Type	Virus Stock Consensus/ Virus Stock Variant/ T-I Assay Variant ¹	Mutation Position
AcOrf-84 promoter	SNPs	G/T/T	71,443
fgf 3' UTR	SNPs	A/G/G	27,502
fgf 3' UTR	SNPs	G/A/A	27,505
fgf 3' UTR	SNPs	G/A/A	27,506
fgf 3' UTR	SNPs	G/A/A	27,509
<i>AcOrf-603</i>	SNPs	C/T/T	3960
<i>AcOrf-1629</i>	SNPs	G/A/A	6375
<i>lef10</i>	SNPs	G/A/A	45,761
<i>AcOrf-91</i>	SNPs	T/A/A	78,627
<i>AcOrf-91</i>	SNPs	A/T/T	78,666
<i>AcOrf-1629</i>	Insertion	-/GATC/GATC	7304
<i>AcOrf-51</i>	Insertion	-/A/A	44,073
<i>egt</i>	Deletion	² CTAGAGA/-/-	12,427
<i>AcOrf-91</i>	Deletion	TAT/-/-	78,926

¹ Changes in nucleotides between the reference genome on the left, the variants in the shotgun sequenced virus stock in the middle, and the variants in the tiled-amplicon sequenced rBEV genomes recovered from the T-I assay on the right; ² CTAGAGATCTCTAGAGATCT, the complete 20 bp deletion sequence observed in the *egt* gene.

5.3.4 Evaluation of Mutations Outside the Targeted *gp64* Gene

Mutations in the viral genomes were also detected outside the targeted *gp64* gene following the CRISPR-Cas9-based T-I assay. However, as we are assessing variants in a heterogeneous pool of gDNAs, these untargeted mutations do not necessarily exist in the same genome where the *gp64* mutations were observed. A majority of these mutations were confined to the hrs; specifically, in six hrs (Figure 5.3a) out of the nine hrs in the AcMNPV C6 genome (Ayres et al., 1994). The hrs with no mutations (hr1a, hr2a, and hr4c) are 18–30 bp long,

while the ones with mutations are at least 150 bp long. Moreover, except for the gp64–160 target sample, the majority of the hrs mutations were either found exclusively in the gDNA pools extracted from the cell pellets, or distributed almost equally between the cell pellet and supernatant fractions (Figure 5.3a). It was further observed that most hrs mutations in the targeted samples were either common with the controls or carried over from the virus stock, as depicted by the drop in mutation counts between Figure 5.3a,b. Additionally, filtering out these common or preserved mutations revealed that not all targeted samples exhibited unique mutations in all six hrs, and except for the gp64–160 target sample, most of the mutations were in the cell pellet fractions (Figure 5.3b).

Mutations outside the hrs and the targeted *gp64* gene were observed in both the sgRNA target samples and the controls. However, these mutations exist within viruses that constitute a small fraction of the total number of viruses (Figure 5.4). CHOPCHOP’s uniqueness method, which looks for mismatches in 20 bp sequences upstream of protospacer adjacent motifs (PAMs) (Labun et al., 2019; Hsu et al., 2013), returned that there were no predictable off-targets for the sgRNA targets used in this study. It was further observed that for the sgRNAs to bind to an off-target site, at least four mismatches for gp64+131, five mismatches for gp64–160, five mismatches for gp64+278, four mismatches for gp64+378, five mismatches for gp64+418, and six mismatches for gp64+131/384 were required. However, they lacked the PAM sequence, without which CRISPR-Cas9-mediated binding and cleavage would not occur.

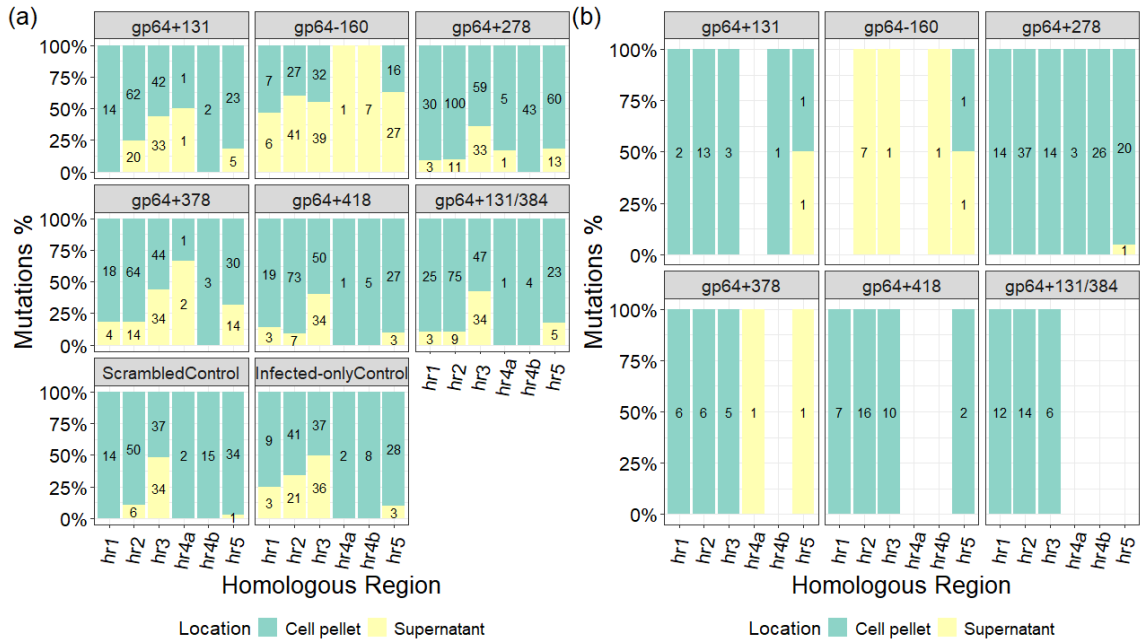


Figure 5.3: (a) Total and (b) filtered (excluding controls and conserved variants) mutation distribution in the different hrs of gDNA pools are highlighted here. The numbers within each bar represent the number of mutations observed in hrs for gDNA pools extracted from the cell pellet or supernatant of a sample. Each sample was run in duplicate and sequenced at a depth of $400\times$. Only true mutations with Fisher's exact test p -value ≤ 0.05 are considered.

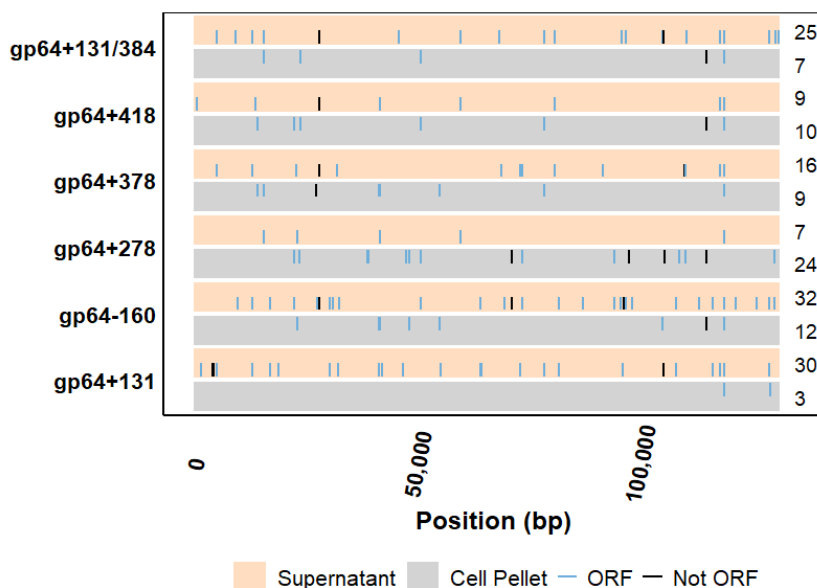


Figure 5.4: Overview of mutations in the pool of gDNAs outside the hrs and the CRISPR-Cas9 targeted *gp64* gene. The numbers on the right side of each horizontal bar represent the total number of true mutations (Fisher’s exact test p -value ≤ 0.05) observed after excluding controls and preserved variants. The blue and black vertical lines are the corresponding mutation positions within and outside ORFs, respectively, unique to the targeted samples. Each sample was run in duplicate and sequenced at a depth of $400\times$.

5.4 Discussion

Over the last two decades, the BEVS platform has been utilized for the production of ten human and five veterinary commercially approved biologics, with many other biologics currently in different phases of clinical trials (Hong et al., 2022, 2023). However, a drawback of the platform is the co-production of virions and the products of interest, like enveloped VLPs or recombinant proteins. Deletion or inactivation of genes essential for budded virus release in the final production stage can reduce baculovirus contamination (Bruder,

[Aucoin, 2023a](#)). The AcMNPV *gp64* gene codes for the major envelope fusion protein, GP64, which is essential for cell-to-cell transmission of the virus and systemic infection of the host ([Rohrmann, 2019a](#); [Monsma et al., 1996](#)). A previous study reported that CRISPR-Cas9-mediated disruption of *gp64* reduced IVT by 99% while maintaining Gag VLP production, thus demonstrating the use of CRISPR-Cas9 to address BV co-production in cell culture supernatant ([Bruder, Aucoin, 2023a](#)). While previous works on CRISPR-Cas9-mediated *gp64* disruption analyzed only the phenotypic changes ([Bruder, Aucoin, 2023a,b](#); [Hausjell et al., 2023](#)), we investigated both phenotypic and genotypic changes with a focus on the genomic changes that occurred upon CRISPR-Cas9 genome editing using *gp64* as an example. In this study, we used a CRISPR-Cas9-based T-I assay (screening assay) ([Bruder, Aucoin, 2023b](#)) to disrupt the *gp64* gene at different locations (five single sgRNA targets and one dual sgRNA target) and evaluated its impact on GFP and progeny virus production. The percentage of GFP-expressing cells (Figure 5.1) upon *gp64* gene disruption was similar to the control for all sgRNA targets while reducing the IVT and total viral particles (Figure 5.2), which is consistent with previous reports ([Bruder, Aucoin, 2023b,a](#); [Chakraborty et al., 2025](#)). Our data also suggests that a synchronous infection at a multiplicity of infection (MOI) of 3 pfu/cell does not reduce foreign protein production while transiently targeting *gp64*. This demonstrates the need for a different approach to low MOI strategies, one that may require using a subset of non-Cas9 expressing cells at the beginning of production. No matter what target site was chosen, similar outcomes were observed. Thus, the CRISPR-Cas9-mediated targeting of *gp64* is a promising strategy to reduce rBEV contamination without affecting foreign protein production when using the p6.9 promoter to drive foreign protein expression.

5.4.1 CRISPR-Cas9-Mediated Targeted Disruption of *gp64* During Virus Propagation

The CRISPR-Cas9 system uses a short RNA to guide the Cas9 endonuclease to a specific genomic location, where Cas9 then cleaves double-stranded DNA. Once the DNA is cleaved, cellular repair mechanisms will try and repair the break, which can result in indel mutations. If the cell is supplied with complementary DNA, the repair can result in “designed/premeditated” changes; otherwise, the changes should be random. As a result, the CRISPR-Cas9 system can be used for targeted gene editing (Jinek et al., 2012; Hsu et al., 2014). The indel mutations that occur are located 3–4 bp upstream of the PAM and can range between 1 and 5 bp and lead to targeted gene disruption (Sari-Ak et al., 2023; Guo et al., 2018; Hussain et al., 2021). Although we observed phenotypic effects (fluorescence and virus titer), we wanted to confirm that our observations were due to the disruption of our intended target. A secondary goal was to investigate if the mutations were confined to unpackaged (intracellular) or budded (extracellular) baculovirus genomes, or if they could be observed in both fractions.

It was postulated that our targeted rBEV genomes might have constituted a minority population. Thus, to analyze mutations upon gene disruption, we used a variant calling pipeline based on the computational tool iVar (intra-host variant analysis of replicates), originally used for measuring intra-host Zika and West Nile virus diversity (Grubaugh et al., 2019). In this study, different thresholds were used for the variant calling pipeline parameters to ensure the detection of high-quality, low-frequency variants. For instance, the *samtools view* `-q` and `-F` parameters filtered out alignments with a low mapping quality

score below 10 and excluded unmapped reads, as well as secondary (suboptimal) and supplementary (segmented) alignments from downstream analysis using the 2308 flag (Danecek et al., 2021). Additionally, *samtools ampliconclip* ensured that primers were trimmed from both ends of mapped reads, while the `-filter-len` parameter excluded short reads below 200, which are often less reliable (Danecek et al., 2021). In the *iVar variants* step, the `-q` parameter was used to set the minimum base quality score to 20 (default) for a position to be considered in variant calling, and the `-t` parameter defined the variant calling frequency threshold, enabling the detection of low-frequency variants (present in at least 1%) (Grubaugh et al., 2019).

A drawback of the plasmid-based delivery of sgRNA is that transfection is not 100% efficient, which can result in some cells not receiving the sgRNA. Additionally, the AcMNPV *gp64* gene was targeted while the virus was undergoing its infection cycle in Sf9-Cas9 cells, where the viral genome will be undergoing replication. Thus, it is possible that the double-stranded breaks (DSBs) can undergo homology-directed repair (HDR) using other copies of the genome as a DNA template. Using the variant calling pipeline with specified thresholds, we were able to identify mutations within the expected regions of the *gp64* gene for most of the sgRNA targets, and these mutations were observed in the cell pellet as well as the supernatant fractions (Table 5.1). However, even though expected phenotypic changes were observed, no *gp64* mutations were detected for the dual sgRNA target gp64+131/384 and the single sgRNA target gp64+278, as well as for the supernatant fraction of the single sgRNA target gp64+131. It is possible that targeting the *gp64* gene at two locations within close proximity induced DSBs that were beyond repair by non-homologous end joining (NHEJ) or HDR. Consequently, while the intracellular and extracellular viral genomes

from the rBEV infecting untransfected Sf9-Cas9 cells (no sgRNA) were sequenced, the targeted genomes were probably routed for destruction by nucleases. Moreover, for the gp64+278 target and the supernatant fraction of the gp64+131 target, perhaps only properly repaired or untargeted genomes were detected, and the targeted genomes with indels were present in an even lower frequency (less than 1%). Applying less stringent thresholds, such as lowering *samtools ampliconclip* `-filter-len` to 50 and *iVar variants* `-q` to 10 or 15, did not reveal mutations within *gp64* for the targets where mutations were not initially detected.

5.4.2 Are CRISPR-Cas9 Off-Targets Observed in Our System?

The CRISPR-Cas9 genome-editing technology is often associated with off-target effects, where the Cas9 endonuclease cleaves unwanted genetic locations that are similar to the target and have the required PAM. In this study, sgRNAs were designed using the online tool CHOPCHOP with advanced parameters, and only those sgRNAs that returned no off-targets were selected (Labun et al., 2019; Chakraborty et al., 2024; Hsu et al., 2013). A previous study demonstrated that while Cas9 tolerated 2–3 interspaced or concatenated mismatches in the PAM distal region, even 2 mismatches, whether interspaced or concatenated, in the PAM seed region considerably reduced or abolished Cas9 activity (Hsu et al., 2013). For Cas9 to cleave an off-target site, the appropriate PAM must be present (Hsu et al., 2013). To further investigate whether our sgRNA targets would have off-target effects, we scanned our reference genome (on Benchling [Biology Software] (2025), retrieved from <https://benchling.com>) for each sgRNA target by increasing the number of mismatches between the 20 bp spacer sequence and the reference genome. We noticed that a

minimum of 4 mismatches were required for our selected sgRNA targets to guide the Cas9 to unintended genomic locations. Further probing these locations revealed missing PAM sites, thus leading to no Cas9 binding and cleavage.

5.4.3 Variant Conservation or Random Mutations Upon Virus Propagation in Cell Culture

As we are analyzing variants in this study, we wanted to investigate if any of the variants from the T-I assay were present in our p6.9GFP rBEV stock. Two different sequencing methodologies, shotgun and tiled-amplicon sequencing, were used for the virus stock and rBEV genomes recovered from the T-I assay, respectively. Applying the variant calling pipeline to both sequencing reads revealed that 68.8% of the virus stock variants were carried over to the T-I assay. While these variants were mainly confined to the hrs (hr1, hr2, and hr3), ~12.6% of them were present in other regions of the genome. These mutations do not form the consensus sequence of the virus stock or the genomes from the T-I assay but are rather present as variants. Additionally, since they are conserved as variants, it suggests that they do not negatively impact virus fitness.

Baculovirus genomes are susceptible to mutations originating from replication errors during propagation in insect cell cultures (Miele et al., 2011; Technologies, 2016). These random mutations are often observed in the hrs as well as other regions of the genome (Pijlman et al., 2001). The AT-rich hrs are repetitive in nature and considered mutation hotspots due to their susceptibility to error-prone processes during viral replication (Pijlman et al., 2001; Chateigner et al., 2015; Boezen et al., 2022). Our data suggested that

the longer hrs (≥ 150 bp) were more prone to mutations compared to the shorter ones (18–30 bp). It is important to note that the bioinformatics pipeline used in this study was designed to identify variants (minor species) within a heterogeneous population; thus, these mutations do not represent the consensus sequences of the viral gDNAs and are not necessarily present in the same genome where *gp64* mutations exist. Variants detected outside the targeted *gp64* gene could result from baculoviruses' susceptibility to mutations during amplification in insect cell cultures (Pijlman et al., 2001; Miele et al., 2011). Moreover, most of the variants (outside *gp64*) in the targeted samples were also observed in the controls or were conserved over passages.

A previous study demonstrated that continuous propagation of an AcMNPV E2 strain-based recombinant baculovirus in Sf21 cells, at a high MOI of 20 pfu/cell, led to a drop in titer with complete loss observed from P31; and this was associated with the formation of defective interfering particles (DIPs) that can infect cells but are not capable of completing the infection cycle (Pijlman et al., 2001). Although DIP generation in high-passage viral stocks has been reported upon serial passages at high MOI in insect cell cultures, the questions remain whether the viruses need to be amplified at high MOI during the seed train and whether such a high-passage stock is required for production in bioreactors. Typically, infecting at a low MOI of 0.1 pfu/cell until the production stage would suffice to obtain a virus stock for production. To put that into perspective, 5 mL of a P1 virus stock with a titer of 1×10^8 pfu/mL can infect 2.5 L of culture seeded at a density of 2×10^6 cells/mL at an MOI of 0.1 pfu/cell. Following that, 50 L culture seeded at 2×10^6 cells/mL can be infected with 2.5 L of the P2 stock ($\sim 4 \times 10^8$ pfu/mL) at an MOI of 10 pfu/cell to achieve synchronous infection. Thus, realistically, one can say that a high-passage virus

stock is not necessary to generate enough stock to infect cells in a bioreactor. Since, in our work, we infected the Sf9-Cas9 cells with a low passage (P2) p6.9GFP rBEV stock at a moderate MOI of 3 pfu/cell, we believe that the mutations observed outside the targeted region were not due to the formation of DIPs. These deletion mutations outside *gp64* were, in fact, a maximum of 20 bp (within *egt*) or 19 bp (within *hr2*) long compared to large genomic deletions of ~43% in DIPs (Pijlman et al., 2001; Kool et al., 1991). Additionally, it was previously demonstrated that the 20 bp deletion in *egt* originates from a small shotgun sequencing reference genome assembly error, where this 20 bp region should not be present in the reference genome (Chakraborty et al., 2025).

5.5 Materials and Methods

5.5.1 Cell Line and Maintenance

Cas9 expressing Sf9 (Sf9-Cas9) cells (Bruder et al., 2021) and parental Sf9 cells were maintained in suspension in Sf-900TM III serum-free media (SFM) (Gibco, Carlsbad, CA, USA) at 27 °C and 130 rpm. To keep the cells in the exponential growth phase, they were passaged every 3–4 days until the viable cell density reached between 3 and 5×10^6 cells/mL. Since the Sf9-Cas9 cells carried a Cas9-2A-puromycinR gene cassette, puromycin (Sigma-Aldrich, Oakville, ON, Canada) was routinely added to the Sf9-Cas9 cell cultures at a concentration of 5 µg/mL to ensure *cas9* expression.

5.5.2 Plasmid Design and Construction

All primers used in this work were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA) and listed elsewhere (Chakraborty et al., 2024). sgRNA plasmids were constructed using the NEBuilder 2× HiFi DNA assembly master mix (New England Biolabs, Whitby, ON, Canada) as previously described (Chakraborty et al., 2024). Briefly, to obtain the SfU6-sgRNA insert, a fusion PCR was performed with a PCR-amplified SfU6 promoter (Mabashi-Asazuma, Jarvis, 2017) gBlock (synthesized dsDNA) gene fragment (IDT) and a gRNA scaffold with a scrambled spacer sequence (Table 5.3) at the 5′ end and a transcriptional terminator at the 3′ end (Addgene # 49411) (Port et al., 2014) as templates. Separately, the ampicillin resistance gene (ampR) and the pBR322 origin of replication (ori) from the pBR322-TIMER plasmid (Addgene # 103056) (Claudi et al., 2014) were PCR-amplified to obtain the backbone fragment. The pSfU6-sgRNA scrambled control plasmid was constructed in a two-fragment Gibson assembly reaction utilizing the SfU6-sgRNA insert and the backbone fragment.

Re-targeting sgRNA plasmids were constructed using the pSfU6-sgRNA scrambled control plasmid as the template for inverse PCR (Bruder, Aucoin, 2023b). Primers with an altered spacer sequence (for different locations within the *gp64* gene), appended to their 5′ ends and annealing either to the 5′ end of the gRNA scaffold or the 3′ end of the SfU6 promoter, were used to amplify the entire sgRNA plasmid as a linear fragment. Following DpnI digestion to remove the template DNA, the gel-extracted re-targeting sgRNA fragment was re-circularized. The new spacer sequence at both ends acted as the homologous regions for Gibson assembly. The online tool “CHOPCHOP” (Labun et al., 2019) was used to design AcMNPV *gp64* targeting spacer sequences utilizing the previously outlined

spacer sequence selection criteria (Chakraborty et al., 2024). The gp64+131 sgRNA plasmid was obtained from a previous study (Bruder, Aucoin, 2023b). All the spacer sequences along with their PAM sequences used in this study are listed in Table 5.3.

Table 5.3: Spacer sequences used for control and AcMNPV *gp64* gene.

Gene	Location	Spacer Sequence (5'-3')	PAM	Strand
Scrambled control	N/A	CACCTTGAAGCGCATGAACT	N/A	N/A
<i>gp64</i>	+131	GGAAACGCTGCAAAAGGACG	TGG	Antisense
<i>gp64</i>	-160	GTTGTAGTCCGTCTCCACGA	TGG	Sense
<i>gp64</i>	+278	AACGCTGAATGTGGGCAAAG	AGG	Antisense
<i>gp64</i>	+378	GACTGTTTTTCGCGACAACGA	GGG	Antisense
<i>gp64</i>	+418	AAGGCAAAGAGTTGGTGAAG	CGG	Antisense
<i>gp64</i>	+131/ 384	GGAAACGCTGCAAAAGGACG/ TTTCGCGACAACGAGGGCCG	TGG/ CGG	Antisense

5.5.3 Baculovirus Amplification and Quantification

A previously constructed rBEV carrying the GFP monomeric Azami green under the AcMNPV p6.9 promoter (herein referred to as p6.9GFP rBEV) was used in this study (Bruder, Aucoin, 2023b). A working virus stock was obtained by amplifying a P1 stock of this rBEV in Sf9 cells at a low MOI of ~ 0.1 pfu/cell until the viable cell density dropped between 80 and 90%.

Using an end-point dilution assay (EPDA), IVTs were quantified as previously described (Reed, Muench, 1938; O'Reilly et al., 1992). Briefly, 96-well tissue-culture-treated plates (VWR International, Mississauga, ON, Canada) were seeded with 100 μ L of Sf9 cells

diluted to 2×10^5 cells/mL in Sf-900TM III SFM. The cells were allowed to adhere for around 1 h at 27 °C. In the meantime, dilutions of the virus stocks were prepared by serial dilution from 10^{-2} to 10^{-8} in Sf-900TM III SFM. Following the incubation period, cells were inoculated with 10 μ L of a virus dilution, resulting in 12 replicates per dilution per row, with an uninfected row receiving no virus. Finally, the plates were observed under a fluorescence microscope after being incubated at 27 °C for 6–7 days to determine the fluorescence. The virus titer in plaque-forming units per mL (pfu/mL) was then obtained by taking the reciprocal of TCID₅₀ expressed in mL of virus added and multiplying it by 0.68 based on the Poisson distribution.

5.5.4 Transfection-Infection Assay (T-I Assay)

The AcMNPV *gp64* gene was targeted via a previously developed CRISPR-Cas9-based T-I assay (Bruder, Aucoin, 2023b). Briefly, four 12-well tissue-culture-treated plates (VWR International, Mississauga, ON, Canada) were seeded with Sf9-Cas9 cells at a density of 0.9×10^6 cells/well and allowed to adhere for ~ 1 h at 27 °C. The cells were then transfected with either the scrambled control plasmid or *gp64* targeting sgRNA plasmids ($n = 6$) using FuGENE HD (Promega, Madison, WI, USA) transfection reagent according to the manufacturer’s protocol. An additional three wells seeded with the Sf9-Cas9 cells were left untransfected. At 16–24 h post-transfection (hpt), the media from each well was aspirated, and fresh Sf-900TM III SFM containing the p6.9GFP rBEV was added to all the wells to achieve an infection at an MOI of 3 pfu/cell. This resulted in six replicates of the transfected+infected samples for each *gp64* targeting sgRNA plasmid, six of the scrambled control, and three of the infected-only control. Approximately 48 h

post-infection (hpi), the cells were harvested by centrifugation at $800\times g$ for 15 min. Cell cultures from three wells (replicates) for each control and targeted sample were pooled for sequencing. The pooled cell pellets and supernatants were stored at $-80\text{ }^{\circ}\text{C}$ and $4\text{ }^{\circ}\text{C}$, respectively, to confirm mutations in both intracellular and extracellular p6.9GFP rBEV genomes via next-generation sequencing (NGS). Finally, the supernatants from the other replicates ($n = 3$) of the scrambled control and targeted samples were analyzed by EPDA and flow cytometry for infectious and total baculovirus titers, respectively, and the cell pellets via flow cytometry for GFP production.

5.5.5 Flow Cytometry Analysis of GFP upon *gp64* Gene Disruption

The cell pellets from the T-I assay were prepared for flow cytometry by resuspending them in 2% paraformaldehyde (Sigma-Aldrich, Oakville, ON, Canada) in phosphate-buffered saline (PBS) (Wisent Inc., Saint-Jean-Baptiste, QC, Canada), followed by incubation at $4\text{ }^{\circ}\text{C}$ for ~ 30 min. Fixed samples were diluted in $1\times$ PBS before being analyzed using a BD AccuriTM C6 Plus flow cytometer (BD Biosciences, San Jose, CA, USA) equipped with a blue laser with an excitation frequency of 488 nm. Samples were run at a low flow rate of $14\text{ }\mu\text{L}/\text{min}$, with 10,000 events collected per sample. Licensed software, FlowJoTM V10 (Tree Star, Ashland, OR, USA), was used to analyze the acquired flow data. After applying a gate to remove cell debris, a histogram of the gated population was used to bin the fluorescence detected by the FL1-H detector into high ($\text{FL1-H} \geq 10^6\text{ au}$) and low ($\text{FL1-H} < 10^6\text{ au}$) fluorescence bins, and these gates were applied to all the samples. The

R programming language was used to process the data and visualize the results further.

5.5.6 Total Baculovirus Quantification via Flow Cytometry

Total baculovirus particles were quantified by a previously described SYBR Green staining-based flow cytometry assay (Shen et al., 2002). Briefly, 10^{-2} diluted virus samples in $1\times$ PBS were fixed in 2% paraformaldehyde at 4 °C for 1 h. This was followed by a freeze-thaw cycle during which samples were subjected to freezing at -80 °C for 30 min and thawing at 27 °C for 10 min. The samples were then treated with 10% Triton X-100 (Sigma-Aldrich, Oakville, ON, Canada) for 5 min to permeabilize the membrane. SYBR Green I Nucleic Acid Gel Stain $10,000\times$ (Thermo Fisher, Mississauga, ON, Canada) diluted to 5×10^{-3} was added to the samples to stain dsDNA. The stained samples were incubated in the dark at 80 °C for 10 min, followed by a cooling step on ice for 10 min before being analyzed by flow cytometry.

The prepared samples were analyzed by a BD AccuriTM C6 Plus flow cytometer (BD Biosciences, San Jose, CA, USA) equipped with a blue laser with an excitation frequency of 488 nm. Each sample was run at a medium flow rate of 35 μ L/min for 30 s. A detailed flow cytometry analysis protocol for total baculovirus quantification can be found elsewhere (Chakraborty et al., 2024). Briefly, acquired data were processed using the FlowJoTM V10 software (Tree Star, Ashland, OR, USA) and the R programming language. For calibration purposes, 3 μ m FlowSet fluorospheres (Beckman Coulter, Mississauga, ON, Canada) with a concentration of 1×10^6 fluorospheres/mL were used.

5.5.7 Tiled-Amplicon Sequencing Assay for rBEV Genomes

The sequencing assay was adapted from a previous study (Quick et al., 2017). All custom index primers (Supplementary Table C.1) and tiled-amplicon primers (Chakraborty et al., 2025) used in this study were synthesized by IDT. The adaptation of this tiled-amplicon sequencing assay to targeted rBEV genomes has been described elsewhere (Chakraborty et al., 2025). Briefly, using PrimalScheme (Quick et al., 2017), tiled-amplicon primers were designed to amplify the rBEV genomes recovered from the T-I assay. These gDNAs were extracted using the Wizard Genomic DNA Purification kit (Promega, Madison, WI, USA) following the manufacturer’s protocol. Odd-numbered and even-numbered primer pairs were pooled into Primer Pools 1 and 2, respectively, to generate alternate amplicons via multiplex PCR. The size of the amplicons was confirmed by running the PCR products on 1% agarose gel. Subsequently, amplicons from each PCR for a sample were pooled and purified using AMPure XP magnetic beads (Beckman Coulter, Mississauga, ON, Canada) according to the manufacturer’s directions. Purified amplicons were quantified using the Qubit dsDNA high-sensitivity assay kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer’s instructions, and then prepared for DNA library construction using the Illumina DNA Prep, (M) Tagmentation kit (Illumina, San Diego, CA, USA) according to the manufacturer’s protocol. The indexed amplicons were then quantified using the Qubit dsDNA high-sensitivity assay kit (Thermo Fisher Scientific, Waltham, MA, USA) as described by the manufacturer and visualized on 1% agarose gel to estimate the median sizes of the DNA libraries. Finally, the DNA and PhiX control libraries were denatured and diluted as previously described (Chakraborty et al., 2025) and loaded onto the MiSeq reagent kit v3 (Illumina, San Diego, CA, USA) cartridge for sequencing.

5.5.8 Bioinformatics Pipeline for Minor Species

The variant calling pipeline (included in the Supplementary C), with a reference genome and two technical replicates for each sample, was run from the Ubuntu terminal command prompt (22.04.5 LTS). Briefly, for running the pipeline with a reference, the ‘bwa index reference.fasta’ (Li, 2013) command was used to prepare the FASTA file for alignment by creating index files of the reference genome with extensions .amb, .ann, .bwt, .pac, and .sa. Following this, *bwa mem* (Li, 2013) was utilized to map a sample’s forward and reverse sequencing reads to the indexed reference, creating a sam file. Using *samtools view* (Danecek et al., 2021), with `-q` and `-F` parameter thresholds 10 and 2308, respectively, the sam file was converted to a raw bam file. This raw bam file was then realigned to the reference genome by *lofreq viterbi* (Wilm et al., 2012) to correct mapping errors, especially around indels. The realigned bam file was sorted by *samtools sort* (Danecek et al., 2021) based on its position in the reference genome determined through alignment. For fast random access of data, the sorted bam file was index coordinated using *samtools index* (Danecek et al., 2021). *Samtools ampliconclip* (Danecek et al., 2021) was then utilized to clip primers from reads in the sorted bam file with the `-filter-len` parameter set to 200, and the output was a trimmed bam file. Subsequently, *samtools mpileup* (Danecek et al., 2021) generated a pileup file from the trimmed bam file, which summarized the alignment of reads to the reference genome. The *samtools mpileup* output was piped into *iVar variants* (Grubaugh et al., 2019) to detect variants (single-nucleotide variants and indels) with the `-q` parameter set to 20 and `-t` set to 0.01. To correctly call variants, the reference genome was passed using the `-r` flag, and after processing the pileup file, a tab-delimited file was generated with variant details. Finally, *iVar filtervariants* (Grubaugh et al., 2019) was

used to filter variants across the .tsv files of the technical replicates of a sample to create a single filtered variant file for the sample (Supplementary Table C.2 describes the .tsv file column headers). All the .tsv files were then analyzed using the R programming language.

5.6 Conclusions

All targeted disruption of the AcMNPV *gp64* gene at different locations reduced BV co-production in cell culture while enhancing GFP production. Our results suggest improved disruption when multiple targeting sites are used. Our sequencing pipeline successfully determined the resulting CRISPR-Cas9-mediated mutations within *gp64* for the majority of the targets (gp64+131, gp64-160, gp64+378, and gp64+418). Additionally, our analysis suggested that there were no off-target effects and confirmed that variants could be conserved over virus amplification in cell culture.

Chapter 6

Conclusions and Recommendations

The BEVS platform was first established in the 1980s ([Smith et al., 1983b](#)). Since then, ten human and five veterinary use biologics produced in this system have received regulatory approvals, with many in different phases of clinical trials ([Hong et al., 2022, 2023](#); [Anguela, High, 2024](#); [Symington et al., 2024](#)). While some progress has been made on the genetic engineering front to improve foreign protein yield and quality, only a few changes have been made to the commercially available baculovirus genomes ([Kaba et al., 2004](#); [Hitchman et al., 2010b,a](#)). Additionally, the functions of many AcMNPV genes are not yet experimentally confirmed, and the co-production of rBEVs in cell culture supernatant complicates downstream purification processes ([Bruder, 2021](#); [Marek et al., 2011](#); [Bruder, Aucoin, 2023a](#)).

The BEVS is a production platform that uses recombinant baculovirus vectors to produce products of interest in insect cell culture, thereby distinguishing it from other cell-based (mammalian or bacterial) production systems and wild-type viruses. It is a process

where the survival of the cell or virus is not of importance, but rather the survival of the production process, which dictates the essentiality of endogenous baculovirus genes in cell culture. In this thesis, we employed a previously developed CRISPR-Cas9-based T-I assay ([Bruder, Aucoin, 2023b](#)) to probe late and very late AcMNPV genes for BV production and late exogenous protein production. We categorized 19 AcMNPV genes as essential, as their disruption impaired BV and foreign protein production from the late p6.9 promoter. These genes should not be removed from the baculovirus genome to maintain production levels. Another 19 AcMNPV genes were classified as of special interest, as their disruption significantly reduced GFP production from the late p6.9 promoter while maintaining or enhancing BV levels. Complete removal or disruption of these genes can potentially improve BV production in insect cell culture, and this reduced genome viral vector can be used as a BacMam with a mammalian promoter for protein production in mammalian cells. While previous studies focused only on BV titers or foreign protein production from very late promoters (upon fragment deletion) ([Chen et al., 2021](#); [Yu et al., 2023](#); [Zhang et al., 2023](#)), we demonstrated the effect of individual gene disruptions on both progeny virus and late exogenous protein production, which is important for moving towards minimal genome rBEVs.

To establish high confidence in our results not only did we use 2–3 sgRNA targets for each AcMNPV gene and ran them in triplicates, but also implemented a conservative statistical approach (ANOVA with Dunnett’s test) that specifically compared changes in multiple treatment groups to a single control while controlling the family-wise error rate by adjusting the significance threshold. These steps ensured that the observed data were a true representation of the effects of gene disruptions. This high-throughput T-I assay

can be further enhanced by adapting it to a 24- or 48-well format, thereby increasing the number of genes that can be assessed in parallel. Additionally, as a next step, the genes identified as of special interest in our study can be probed further by completely removing them from the rBEV genome via restriction digestion and/or Gibson assembly, and investigating if the resulting rBEVs are still capable of generating progeny viruses that can be applied as BacMam vectors for gene delivery into mammalian cells. While we successfully used this T-I assay to probe the essentiality of genes for the production of virions and late intracellular exogenous reporter protein, it can be extended to investigate their effect on the production of late extracellular complex proteins, such as VLPs. The effect of AcMNPV gene disruption might differ depending on the type of foreign gene (intracellular or extracellular) being expressed, and it is recommended to investigate these endogenous genes in the next phase. Moreover, while a gene can be targeted at multiple locations using multiplexing sgRNAs to increase the chances of knockout, from a screening point of view, it might complicate the interpretation of phenotypes. Finally, this screening assay can be carried over to the High-FiveTM cells; however, to ensure high sgRNA expression, the sgRNAs should be driven by the High-FiveTM cell-specific U6 promoter.

Although the phenotypic changes could be effectively assessed using the T-I assay, we sought to investigate the genomic changes resulting from CRISPR-Cas9-mediated disruptions. Previous studies have reported on the targeted CRISPR-Cas9 mutations ([Liu et al., 2014](#); [Xiang et al., 2011](#); [Dong et al., 2016, 2018](#)); however, no reports on whole-genome sequences upon targeted gene disruptions were available. This thesis outlines different pipelines for whole-genome sequencing of untargeted and targeted rBEV genomes. The first step was to obtain the consensus sequence of the p6.9GFP rBEV used in CRISPR-

Cas9 studies to use it as a reference genome for future work. We used shotgun sequencing on the p6.9GFP rBEV stock (supernatant) amplified in Sf9 cells and generated a consensus sequence that acted as the reference genome for control and targeted rBEV genomes recovered from the T-I assay. To the best of our knowledge, it is the first whole-genome rBEV sequence that has been reported.

The downside of shotgun sequencing is that sequencing intracellular baculovirus genomes would require additional processing steps, as it cannot differentiate between the virus and cell genomes. Additionally, it is believed that disrupting an AcMNPV gene that is required for BV assembly or release would result in the targeted virions not budding out of the cells and thus complicate the sequencing process. To overcome the limitations of shotgun sequencing, we adapted a high-throughput tiled-amplicon sequencing assay to the controls and applied it to *gp64* targeted rBEV genomes when the sgRNA was delivered via a plasmid or rBEV. The tiled-amplicon sequencing assay is more precise, with tiled-amplicon primers designed to specifically amplify the rBEV genomes from both cell pellet and supernatant fractions. Upon the rBEV-based delivery of *gp64* sgRNA, we observed that the IVT dropped by ~ 3 logs, which made it difficult for us to isolate rBEV genomes from the supernatant fractions, thereby emphasizing the relevance of the tiled-amplicon sequencing approach for our system. That being said, for both plasmid- and rBEV-based delivery of *gp64* sgRNA, we were able to detect mutations within the targeted regions when the rBEV genomes were isolated from the cell pellets. These indel mutations suggested that there is indeed a repair mechanism by the cells at play that can salvage the virus if it is “targeted”. Our findings also suggested that CRISPR-Cas9-mediated *gp64* gene disruption can be used as an effective tool to reduce budded baculovirus contamination in

insect cell cultures. While the shotgun sequencing pipeline can be applied to viruses for which only best-guess sequences are known, the tiled-amplicon sequencing assay pipeline can be applied to viruses for which complete genome sequences are available, although genome-specific tiled-amplicon primers need to be designed for them.

As *gp64* disruption has been shown to lower IVT levels, we wanted to probe it further using the T-I assay and sequence the resulting intracellular and extracellular viral genomes using the more precise tiled-amplicon sequencing assay. We targeted the *gp64* gene at six different locations with single or dual sgRNA targets carrying one or two spacer sequences, respectively. In this thesis, we have demonstrated reduced BV production while enhancing or maintaining GFP production from the late p6.9 promoter, thus addressing the drawback of rBEV co-production of the BEVS platform. We used a variant calling pipeline based on the computational tool iVar, specifically designed to look into wild-type virus diversity ([Grubaugh et al., 2019](#)); however, it has not been used to look at viral subpopulations in cell culture. We adapted this tool to a process system to detect CRISPR-Cas9-mediated mutations while the baculovirus undergoes its infection cycle in cell culture. Our sequencing pipeline enabled low-frequency variant detection while successfully determining CRISPR-Cas9-mediated *gp64* mutations and identified variant conservation over virus propagation in insect cell cultures. The tiled-amplicon sequencing assay, combined with the variant calling pipeline outlined here, can be applied to other systems to detect genetic mutations.

References

- Airenne Kari J., Hu Yu Chen, Kost Thomas A., Smith Richard H., Kotin Robert M., Ono Chikako, Matsuura Yoshiharu, Wang Shu, Ylä-Herttuala Seppo.* Baculovirus: An insect-derived vector for diverse gene transfer applications // *Molecular Therapy*. 2013. 21, 4. 739–749.
- Anderson M. A. E., Purcell J., Verkuyl 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*. 2020. 9, 3. 678–681.
- Anguela Xavier M., High Katherine A.* Hemophilia B and gene therapy: a new chapter with etranacogene dezaparvovec // *Blood Advances*. 2024. 8, 7. 1796–1803.
- Arif B., Pavlik L.* Insect cell culture : Virus replication and applications in biotechnology // *Journal of Invertebrate Pathology*. 2013. 112. S138–S141.
- Arunkarthick S., Asokan R., Aravintharaj R., Niveditha M., Krishna N. K.* A Review of Insect Cell Culture : Establishment , Maintenance and Applications in Entomological Research 1 // *Journal of Entomological Science*. 2017. 52(3). 261–273.

- Aucoin M. G.* Characterization and optimization of the production of adeno-associated viral vectors using a baculovirus expression vector/insect cell system // PhD Thesis. 2007. 1. 1–318.
- Aucoin M. G., Mena J. A., Kamen A. A.* Bioprocessing of Baculovirus Vectors: A Review // Current Gene Therapy. 2010. 10, 3. 174–186.
- 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. 2006. 95, 6. 1081–1092.
- 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. 1994. 202, 2. 586–605.
- Barrett P. N., Darner F., Mundt W., Reiter M.* Development of a mammalian cell (Vero) derived candidate influenza virus vaccine // Vaccines. 1998. 16. 960–968.
- Bassett A. R., Tibbit C., Ponting C. P., Liu J. L.* Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9 // Biology Open. 2014. 3, 1. 42–49.
- Bauer M. M. T., Schnapp G.* Protein Production for Three-Dimensional Structural Analysis // Comprehensive Medicinal Chemistry II. 2007. 3. 411–432.
- Beames Burton, Summers Max D.* Location and nucleotide sequence of the 25K protein missing from baculovirus free polyhedra (FP) mutants // Virology. 1989. 168, 2. 344–353.
- Berger I., Fitzgerald D. J., Richmond T. J.* Baculovirus expression system for heterologous multiprotein complexes // Nature Biotechnology. 2004. 22, 12. 1583–1587.

- Berretta M. F., Ferrelli M. L., Salvador R., Sciocco A., Romanowski V.* Baculovirus Gene Expression // InTech. 2013. 1. 13.
- Birnbaum M. J., Clem R. J., Miller L. K.* An apoptosis-inhibiting gene from a nuclear polyhedrosis virus encoding a polypeptide with Cys/His sequence motifs // Journal of Virology. 1994. 68. 2521–2528.
- Black B. C., Brennan L. A., Dierks P. M., Gard I. E.* Commercialization of Baculoviral Insecticides // The Baculoviruses. 1997. 1. 341–387.
- Blissard G. W., Wenz J. R.* Baculovirus gp64 envelope glycoprotein is sufficient to mediate pH-dependent membrane fusion // Journal of Virology. 1992. 66. 6829–6835.
- Boezen Dieke, Ali Ghulam, Wang Manli, Wang Xi, Werf Wopke van der, Vlak Just M., Zwart Mark P.* Empirical estimates of the mutation rate for an alphabaculovirus // PLoS Genetics. 2022. 18, 6. 1–22.
- Boldogkői Zsolt, Moldován Norbert, Szűcs Attila, Tombácz Dóra.* Data descriptor: Transcriptome-wide analysis of a baculovirus using nanopore sequencing // Scientific Data. 2018. 5, 1. 1–10.
- Bolger Anthony M., Lohse Marc, Usadel Bjoern.* Trimmomatic: A flexible trimmer for Illumina sequence data // Bioinformatics. 2014. 30, 15. 2114–2120.
- Bolotin A., Quinquis B., Sorokin A., Dusko Ehrlich S.* Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin // Microbiology. 2005. 151, 8. 2551–2561.

- Bossert Maïke, Carstens Eric B.* Sequential deletion of AcMNPV homologous regions leads to reductions in budded virus production and late protein expression // *Virus Research*. 2018. 256, 1. 125–133.
- Braunagel S. C., Parr R., Belyavskiy M., Summers M. D.* *Autographa californica* nucleopolyhedrovirus infection results in Sf9 cell cycle arrest at G2/M phase // *Virology*. 1998. 244. 195–211.
- Braunagel S. C., Russell W. K., Rosas-Acosta G., Russell D. H., Summers M. D.* Determination of the protein composition of the occlusion-derived virus of *Autographa californica* nucleopolyhedrovirus // *Proceedings of the National Academy of Sciences of the United States of America*. 2003. 100, 17. 9797–9802.
- Bruder M. R.* Toward Optimization of the Baculovirus Expression Vector System- Development of Genetic Tools to Improve Biologics Production // PhD Thesis. 2021. 1. 1–283.
- Bruder Mark R., Aucoin Marc G.* Evaluation of Virus-Free Manufacture of Recombinant Proteins Using CRISPR-Mediated Gene Disruption in Baculovirus-Infected Insect Cells // *Vaccines*. 2023a. 11, 2. 225.
- Bruder Mark R., Aucoin Marc G.* A sensitive assay for scrutiny of *Autographa californica* multiple nucleopolyhedrovirus genes using CRISPR-Cas9 // *Applied Microbiology and Biotechnology*. 2023b. 107, 13. 4323–4335.

- Bruder Mark R., Walji Sadru-Dean, Aucoin Marc G.* Comparison of CRISPR-Cas9 Tools for Transcriptional Repression and Gene Disruption in the BEVS // *Viruses-Basel*. 2021. 13, 10. 1925.
- Bruder Mark R. R., Aucoin Marc G.* Utility of Alternative Promoters for Foreign Gene Expression Using the Baculovirus Expression Vector System // *Viruses-Basel*. 2022. 14, 12. 2670.
- Bushnell Brian, Rood Jonathan, Singer Esther.* BBMerge – Accurate paired shotgun read merging via overlap // *PLoS ONE*. 2017. 12, 10. 1–15.
- Butler M.* Animal cell cultures: Recent achievements and perspectives in the production of biopharmaceuticals // *Applied Microbiology and Biotechnology*. 2005. 68, 3. 283–291.
- Cai Yi, Long Zhao, Qiu Jianxiang, Yuan Meijin, Li Guanghong, Yang Kai.* An ac34 Deletion Mutant of *Autographa californica* Nucleopolyhedrovirus Exhibits Delayed Late Gene Expression and a Lack of Virulence In Vivo // *Journal of Virology*. 2012. 86, 19. 10432–10443.
- Carstens Eric B., Wu Yuntao.* No single homologous repeat region is essential for DNA replication of the baculovirus *Autographa californica* multiple nucleopolyhedrovirus // *The Journal of General Virology*. 2007. 88, 1. 114–122.
- Chakraborty M., Nielsen L., Nash D., Nissimov J. I., Charles T. C., Aucoin M. G.* Adapting next-generation sequencing to in process CRISPR-Cas9 genome editing of recombinant AcMNPV vectors: From shotgun to tiled-amplicon sequencing // *Viruses*. 2025. 17, 3. 437.

- Chakraborty Madhuja, Powichrowski Jacqueline, Bruder Mark R., Nielsen Lisa, Sung Christopher, Boegel Scott J., Aucoin Marc G.* Probing Baculovirus Vector Gene Essentiality for Foreign Gene Expression Using a CRISPR-Cas9 System // *Methods in Molecular Biology* (Clifton, N.J.). 2024. 2829, 1. 127–156.
- Chan Leslie C. L., Reid Steven.* Development of serum-free media for lepidopteran insect cell lines // *Methods in Molecular Biology* (Clifton, N.J.). 2016. 1350. 161–196.
- 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*. 2020. 30, 5. 757–767.
- Chateigner Aurélien, Bézier Annie, Labrousse Carole, Jiolle Davy, Barbe Valérie, Herniou Elisabeth A.* Ultra deep sequencing of a baculovirus population reveals widespread genomic variations // *Viruses*. 2015. 7, 7. 3625–3646.
- Chaves L. C. S., Ribeiro B. M., Blissard G. W.* Production of GP64-free virus-like particles from baculovirus-infected insect cells // *Journal of General Virology*. 2018. 99, 2. 265–274.
- Chen Tong, Duan Xiaoyan, Hu Hengrui, Shang Yu, Hu Yangbo, Deng Fei, Wang Hualin, Wang Manli, Hu Zhihong.* Systematic Analysis of 42 *Autographa Californica* Multiple Nucleopolyhedrovirus Genes Identifies An Additional Six Genes Involved in the Production of Infectious Budded Virus // *Virologica Sinica*. 2021. 36, 4. 762–773.

- Chen Y., Zhong S., Fei Z., Hashimoto Y., Xiang J. Z., Zhang S., Blissard W.* The Transcriptome of the Baculovirus *Autographa californica* Multiple Nucleopolyhedrovirus in *Trichoplusia ni* Cells // *Journal of Virology*. 2013. 87, 11. 6391–6405.
- Chisholm G. E., Henner D. J.* Multiple Early Transcripts and Splicing of the *Autographa californica* Nuclear Polyhedrosis Virus IE-1 Gene // *Journal of Virology*. 1988. 62, 9. 3193–3200.
- Claudi B., Spröte P., Chirkova A., Personnic N., Zankl J., Schürmann N., Schmidt A., Bumann D.* Phenotypic variation of salmonella in host tissues delays eradication by antimicrobial chemotherapy // *Cell*. 2014. 158, 4. 722–733.
- Clem R. J., Miller L. K.* Apoptosis reduces both the in vitro replication and the in vivo infectivity of a baculovirus // *Journal of Virology*. 1993. 67. 3730–3738.
- Clement Kendell, Rees Holly, Canver Matthew C., Gehrke Jason M., Farouni Rick, Hsu Jonathan Y., Cole Mitchel A., Liu David R., Joung J. Keith, Bauer Daniel E., Pinello Luca.* CRISPResso2 provides accurate and rapid genome editing sequence analysis // *Nature Biotechnology*. 2019. 37, 3. 220–224.
- Cochran Mark A., Faulkner Peter.* Location of Homologous DNA Sequences Interspersed at Five Regions in the Baculovirus AcMNPV Genome // *Journal of Virology*. 1983. 45, 3. 961–970.
- Cohen David P.A., Marek Martin, Davies Bryn G., Vlak Just M., van Oers Monique M.* Encyclopedia of *Autographa californica* nucleopolyhedrovirus genes // *Virologica Sinica*. 2009. 24, 5. 359–414.

- Condreay J. P., Witherspoon S. M., Clay W. C., Kost T. A.* Transient and stable gene expression in mammalian cells transduced with a recombinant baculovirus vector // Proceedings of the National Academy of Sciences of the USA. 1999. 96. 127–132.
- Cory J.S., Clarke E.E., Brown M.L., Hails R.S., O'Reilly D.R.* Microparasite manipulation of an insect: the influence of the egt gene on the interaction between a baculovirus and its lepidopteran host // Functional Ecology. 2004. 18, NA. 443–450.
- Cox M. J., Hashimoto Y.* A fast track influenza virus vaccine produced in insect cells // Journal of Invertebrate Pathology. 2011. 107. S31–S41.
- Cox M. J., Hollister J. R.* FluBlok, a next generation influenza vaccine manufactured in insect cells // Biologicals. 2009. 37. 182–189.
- Danecek Petr, Bonfield James K., Liddle Jennifer, Marshall John, Ohan Valeriu, Pollard Martin O., Whitwham Andrew, Keane Thomas, McCarthy Shane A., Davies Robert M.* Twelve years of SAMtools and BCFtools // GigaScience. 2021. 10, 2. 1–4.
- 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. 2011. 471, 7340. 602–607.
- Demain A. L., Vaishnav P.* Production of recombinant proteins by microbes and higher organisms // Biotechnology Advances. 2009. 27, 3. 297–306.
- Detvisitsakun Chanitchote, Berretta Marcelo F., Lehiy Christopher, Passarelli A. Lorena.* Stimulation of cell motility by a viral fibroblast growth factor homolog: Proposal for a role in viral pathogenesis // Virology. 2005. 336, 2. 308–317.

- Detvisitsakun Chanitchote, Hutfless Erica L., Berretta Marcelo F., Passarelli A. Lorena.*
 Analysis of a baculovirus lacking a functional viral fibroblast growth factor homolog // *Virology*. 2006. 346, 2. 258–265.
- Dolgin E.* Early clinical data raise the bar for hemophilia gene therapies // *Nature Biotechnology*. 2016. 34. 999–1001.
- Dong Z., Chen T., Zhang J., Hu N., Cao M., Dong F., Jiang Y., Chen P., Lu C., Pan M.*
 Establishment of a highly efficient virus-inducible CRISPR/Cas9 system in insect cells // *Antiviral Research*. 2016. 130. 50–57.
- Dong Z., Huang L., Dong F., Hu Z., Qin Q., Long J., Cao M., Chen P., Lu C., Pan M.*
 Establishment of a baculovirus-inducible CRISPR/Cas9 system for antiviral research in transgenic silkworms // *Applied Microbiology and Biotechnology*. 2018. 102. 9255–9265.
- Doroshenko A., Halperin S. A.* Trivalent MDCK cell culture-derived influenza vaccine Optaflu . sup .[R] (Novartis Vaccines) // *Expert Review Vaccines*. 2009. 8.
- Doudna J. A., Charpentier E.* The new frontier of genome engineering with CRISPR-Cas9 // *Science*. 2014. 346, 6213.
- Eren A. Murat, Kiefl Evan, Shaiber Alon, Veseli Iva, Miller Samuel E., Schechter Matthew S., Fink Isaac, Pan Jessica N., Yousef Mahmoud, Fogarty Emily C., Trigodet Florian, Watson Andrea R., Esen Özcan C., Moore Ryan M., Clayssen Quentin, Lee Michael D., Kivenson Veronika, Graham Elaina D., Merrill Bryan D., Karkman Antti, Blankenberg Daniel, Eppley John M., Sjödin Andreas, Scott Jarrod J., Vázquez-Campos Xabier, McKay Luke J., McDaniel Elizabeth A., Stevens Sarah L.R., Anderson Rika E.,*

- Fuessel Jessika, Fernandez-Guerra Antonio, Maignien Lois, Delmont Tom O., Willis Amy D.* Community-led, integrated, reproducible multi-omics with anvi'o // *Nature Microbiology*. 2021. 6, 1. 3–6.
- European Medicine Agency EMA.* Questions and Answers on the Review of Preflucel and Associated Names (Influenza Vaccine, Purified Antigen) // Available Online. 2012. 1. 1.
- Evans J T, Leisy D J, Rohrmann G F.* Characterization of the interaction between the baculovirus replication factors LEF-1 and LEF-2 // *Journal of Virology*. 1997. 71, 4. 3114–3119.
- Fan Xiaomin, McLachlin Jeanne R., Weaver Robert F.* Identification and characterization of a protein kinase-interacting protein encoded by the *Autographa californica* nuclear polyhedrosis virus // *Virology*. 1998. 240, 2. 175–182.
- Felberbaum R. S.* The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors // *Biotechnology Journal*. 2015. 10. 702–714.
- 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*. 2013. 12, 2. 225–236.
- Funk C. J., Braunagel S. C., Rohrmann G. F.* Baculovirus Structure // *The Baculoviruses*. 1997. 1. 7–32.
- Funk C. J., Consigli R. A.* Phosphate Cycling on the Basic Protein of *Plodia interpunctella* Granulosis Virus // *Virology*. 1993. 193, 1. 396–402.

- Fuxa J. R.* Insect control with baculoviruses // *Biotechnology Advances*. 1991. 9. 425–442.
- Gaj T., Gersbach C. A., Barbas C. F.* ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering // *Trends in Biotechnology*. 2013. 31, 7. 397–405.
- Galibert L., Jacob A., Savy A., Dickx Y., Bonnin D., Lecomte C., Rivollet L., Sanatine P., Fontaine M. B., Bec C. L., Merten O. W.* Monobac system—a single baculovirus for the production of rAAV // *Microorganisms*. 2021. 9.
- Gandhi Kamal M., Ohkawa Taro, Welch Matthew D., Volkman Loy E.* Nuclear localization of actin requires AC102 in *Autographa californica* multiple nucleopolyhedrovirus-infected cells // *Journal of General Virology*. 2012. 93, 8. 1795–1803.
- Garavaglia Matías Javier, Miele Solange Ana Belén, Iserte Javier Alonso, Belaich Mariano Nicolás, Ghiringhelli Pablo Daniel.* The ac53 , ac78 , ac101 , and ac103 Genes Are Newly Discovered Core Genes in the Family Baculoviridae // *Journal of Virology*. 2012. 86, 22. 12069–12079.
- Gauthier David, Thirunavukkarasu Kannan, Faris Brian L., Russell Darcy L., Weaver Robert 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*. 2012. 93, 2. 364–373.
- Ge Jing, Wei Zong, Huang Yishu, Yin Juan, Zhou Ziqian, Zhong Jiang.* AcMNPV ORF38 protein has the activity of ADP-ribose pyrophosphatase and is important for virus replication // *Virology*. 2007. 361, 1. 204–211.

- George S.* Use and Control of Co-expression in the Baculovirus-Insect Cell System for the Production of Multiple Proteins and Complex Biologics // PhD Thesis. 2016. 1. 1–165.
- George S., Jauhar A. M., Mackenzie J., Kie 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. 2015. 112. 1822–1831.
- Gibson D. G., Young L., Chuang R. Y., Venter J. C., Hutchison C. A., Smith H. O.* Enzymatic assembly of DNA molecules up to several hundred kilobases // Nature Methods. 2009. 6, 5. 343–345.
- 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. 1997. 22. 730–735.
- Grabherr R., Ernst W., Oker-Blom C., Jones I.* Developments in the use of baculoviruses for the surface display of complex eukaryotic proteins // Trends in Biotechnology. 2001. 19. 231–236.
- Grace T. D. C.* Establishment of four strains of cells from insect tissues grown *in vitro* // Nature. 1962. 195. 788–789.
- Grissa I., Vergnaud G., Pourcel C.* CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats // Nucleic Acids Research. 2007. 35. 52–57.

Grubaugh Nathan D., Gangavarapu Karthik, Quick Joshua, Matteson Nathaniel L., De Jesus Jaqueline Goes, Main Bradley J., Tan Amanda L., Paul Lauren M., Brackney Doug E., Grewal Saran, Gurfield Nikos, Van Rompay Koen K.A., Isern Sharon, Michael Scott F., Coffey Lark L., Loman Nicholas J., Andersen Kristian G. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar // *Genome Biology*. 2019. 20, 1. 1–19.

Guarino Linda A., Mistretta Toni-Ann, Dong Wen. Baculovirus lef-12 Is Not Required for Viral Replication // *Journal of Virology*. 2002a. 76, 23. 12032–12043.

Guarino Linda A., Mistretta Toni Ann, Dong Wen. DNA binding activity of the baculovirus late expression factor PP31 // *Virus Research*. 2002b. 90, 1-2. 187–195.

Guarino Linda A., Smith Gayle, Dong Wen. Ubiquitin is attached to membranes of baculovirus particles by a novel type of phospholipid anchor // *Cell*. 1995. 80, 2. 301–309.

Guarino Linda A., Xu Bin, Jin Jianping, Dong Wen. A Virus-Encoded RNA Polymerase Purified from Baculovirus-Infected Cells // *Journal of Virology*. 1998. 72, 10. 7985–7991.

Guo Jinliang, Li Shangrong, Bai Lisha, Zhao Huimin, Shang Wenyu, Zhong Zhaojun, Maimaiti Tuerxunjiang, Gao Xueyan, Ji Ning, Chao Yanjie, Li Zhaoifei, Du Dijun. Structural transition of GP64 triggered by a pH-sensitive multi-histidine switch // *Nature Communications*. 2024. 15, 1. 1–14.

Guo Tao, Feng Yi Li, Xiao Jing Jing, Liu Qian, Sun Xiu Na, Xiang Ji Feng, Kong Na, Liu Si Cheng, Chen Guo Qiao, Wang Yue, Dong Meng Meng, Cai Zhen, Lin Hui, Cai Xiu Jun, Xie An Yong. Harnessing accurate non-homologous end joining for efficient

- precise deletion in CRISPR/Cas9-mediated genome editing // *Genome Biology*. 2018. 19, 1. 1–20.
- Haase S., Sciocco-Cap A., Romanowski V.* Baculovirus insecticides in Latin America: historical overview, current status and future perspectives // *Viruses*. 2015. 7, 5. 2230–2267.
- 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*. 2000. 74.
- Haddley K.* Alipogene tiparvovec for the treatment of lipoprotein lipase deficiency // *Drugs of today (Barcelona, Spain : 1998)*. 2013. 49. 161–70.
- Haft D. H., Selengut J., Mongodin E. F., Nelson K. E.* A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes // *PLoS Computational Biology*. 2005. 1, 6. 0474–0483.
- Harper D. M.* Impact of vaccination with Cervarix™ on subsequent HPV-16/18 infection and cervical disease in women 15-25 years of age // *Gynecologic Oncology*. 2008. 110. 11–17.
- Harrison Robert L., Bonning Bryony C.* Comparative analysis of the genomes of *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedroviruses // *Journal of General Virology*. 2003. 84, 7. 1827–1842.
- Harrison Robert L., Jarvis Donald L., Summers Max D.* The role of the AcMNPV 25K gene, 'FP25,' in baculovirus polh and p10 expression // *Virology*. 1996. 226, 1. 34–46.

- Hausjell Christina Sophie, Klausberger Miriam, Ernst Wolfgang, Grabherr Reingard.* Evaluation of an inducible knockout system in insect cells based on co-infection and CRISPR/Cas9 // PLoS ONE. 2023. 18, 7. 1–13.
- 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. 1997. 238, 2. 243–253.
- Héricourt F., Blanc S., Redeker V., Jupin I.* Evidence for phosphorylation and ubiquitinylation of the turnip yellow mosaic virus RNA-dependent RNA polymerase domain expressed in a baculovirus–insect cell system // Biochemical Journal. 2015. 349. 417–425.
- Hernandez G. J., 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. 2006. 35. 21–34.
- 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. 2010a. 26, 1. 57–68.
- 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. 2010b. 56, 3. 85–93.

- 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*. 2004. 318. 833–838.
- Hodder A. N., Crewther P. E., Matthewll M. L. S. M., Reid G. E., Moritz R. L., Simpson R. J., Anders R. F.* The disulfide bond structure of Plasmodium apical membrane antigen-1 // *Journal of Biological Chemistry*. 1996. 271. 29446–29452.
- Hofmann C., Sandig V., Jennings G., Rudolph M., Schlag P., Strauss M.* Efficient gene transfer into human hepatocytes by baculovirus vectors // *Proceedings of the National Academy of Sciences*. 1995. 92. 10099–10103.
- Hoggan M. D., Blacklow N. R., Rowe W. P.* Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. // *Proceedings of the National Academy of Sciences of the USA*. 1966. 55. 1467–1474.
- Hong G. P., Park J. H., Lee H. H., Jang K. O., Chung D. K., Kim W., Chung I. S.* Production of influenza virus-like particles from stably transfected *Trichoplusia ni* BT1 TN-5B1-4 cells // *Biotechnology and Bioprocess Engineering*. 2015. 20. 506–514.
- Hong Mingqing, Li Tingting, Xue Wenhui, Zhang Sibao, Cui Lingyan, Wang Hong, Zhang Yuyun, Zhou Lizhi, Gu Ying, Xia Ningshao, Li Shaowei.* Genetic engineering of baculovirus-insect cell system to improve protein production // *Frontiers in Bioengineering and Biotechnology*. 2022. 10, September. 1–15.
- Hong Qiaonan, Liu Jian, Wei Yuquan, Wei Xiawei.* Application of Baculovirus Expression Vector System (BEVS) in Vaccine Development // *Vaccines*. 2023. 11, 7. 1–15.

Hsu P. D., Lander E. S., Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering // *Cell*. 2014. 157, 6. 1262–1278.

Hsu Patrick D., Scott David A., Weinstein Joshua A., Ran F. Ann, Konermann Silvana, Agarwala Vineeta, Li Yingqing, Fine Eli J., Wu Xuebing, Shalem Ophir, Cradick Thomas J., Marraffini Luciano A., Bao Gang, Zhang Feng. DNA targeting specificity of RNA-guided Cas9 nucleases // *Nature Biotechnology*. 2013. 31, 9. 827–832.

Hussain Suleman S., Majumdar Rahul, Moore Grace M., Narang Himanshi, Buechelmaier Erika S., Bazil Maximilian J., Ravindran Pavithran T., Leeman Jonathan E., Li Yi, Jalan Manisha, Anderson Kyrie S., Farina Andrea, Soni Rekha, Mohibullah Neeman, Hamzic Edin, Rong-Mullins Xiaoqing, Sifuentes Christopher, Damerla Rama R., Viale Agnes, Powell Simon N., Higginson Daniel S. Measuring nonhomologous end-joining, homologous recombination and alternative end-joining simultaneously at an endogenous locus in any transfectable human cell // *Nucleic Acids Research*. 2021. 49, 13. 1–14.

James D. C., Goldman M. H., Hoare M., Jenkins N., Oliver R. W., Green B. N., Freedman R. B. Posttranslational processing of recombinant human interferon-gamma in animal expression systems. // *Protein science : a publication of the Protein Society*. 1996. 5. 331–340.

James R. M. An Alternative to the Scale-up and Distribution of Pandemic Influenza Vaccine // *Vaccines Manufacturing*. 2009. 1. 12–20.

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*. 2006. 151, 7. 1257–1266.

- 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*. 2012. 337, August. 816–822.
- Joshi P. R. H., Venereo-Sanchez A., Chahal P. S., Kamen A. A.* Advancements in molecular design and bioprocessing of recombinant adeno-associated virus gene delivery vectors using the insect-cell baculovirus expression platform // *Biotechnology Journal*. 2021. 16.
- 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*. 2004. 122, 1. 113–118.
- Kang S. M., Song J. M., Quan F. S., Compans R. W.* Influenza vaccines based on virus-like particles // *Virus Research*. 2009. 143. 140–146.
- Katoh Kazutaka, Rozewicki John, Yamada Kazunori D.* MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization // *Briefings in Bioinformatics*. 2019. 20, 4. 1160–1166.
- Keech C., Albert G., Cho I., Robertson A., Reed P., Neal S., Pledsted J. S., Zhu M., Cloney-Clark S., Zhou H., Smith G. E., Patel N., Frieman M. B., Haupt R. E., Logue J., McGrath M., Weston S., Piedra P. A., Desai C., Callahan K., Lewis M., Price-Abbott P., Formica N., Shinde V., Fries L., Lickliter J. D., Griffin P., Wilkinson B., Glenn G. M.* Phase 1–2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine // *New England Journal of Medicine*. 2020. 383. 2320–2332.

- Kelly B. J., King L. A., Possee R. D.* Introduction to Baculovirus Molecular Biology // Methods in Molecular Biology. 2007. 388. 25–53.
- Kirnbauer R., Booy F., Cheng N., Lowy D. R., Schiller J. T.* Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic // Proceedings of the National Academy of Sciences of the USA. 1992. 89. 12180–12184.
- Kitts P. A., Ayres M. D., Possee R. D.* Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors // Nucleic Acids Research. 1990. 18, 19. 5667–5672.
- Kitts P. A., Possee R. D.* A method for producing recombinant baculovirus expression vectors at high frequency // Biotechniques. 1993. 14. 810–817.
- 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. 2017. 25. 2661–2675.
- Kool M., Berg P. M. M. M. van den, Tramper J., Goldbach R. W., Vlak J. M.* Location of two putative origins of DNA replication of *Autographa californica* nuclear polyhedrosis virus // Virology. 1993. 192. 94–101.
- Kool M., Voncken J. W., Van Lier F. L. J., Tramper J., Vlak J. M.* Detection and analysis of *Autographa californica* nuclear polyhedrosis virus mutants with defective interfering properties // Virology. 1991. 183. 739–746.

- Kool Marcel, Ahrens Christian H., Goldbach Rob W., Rohrmann George F., Vlak Just M.* Identification of genes involved in DNA replication of the *Autographa californica* baculovirus // Proceedings of the National Academy of Sciences of the United States of America. 1994. 91, 23. 11212–11216.
- Kost T. A., Condreay J. P., Jarvis D. L.* Baculovirus as versatile vectors for protein expression in insect and mammalian cells // Nature Biotechnology. 2005. 23. 567–575.
- Kumar S., Miller L. K.* Effects of serial passage of *Autographa californica* nuclear polyhedrosis virus in cell culture // Virus research. 1987. 7, 4. 335–349.
- 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. 2019. 47, W1. W171–W174.
- Lai Qingying, Xu Lixia, Wang Yanling, Luo Wangtai, Zhu Leyuan, Yuan Meijin, Wu Wenbi, Yang Kai.* AcMNPV PKIP is required for hyperexpression of very late genes and involved in the hyperphosphorylation of the viral basic protein P6.9 // Virus Research. 2020. 279. 197889.
- Le A., Jacob D., Transfiguracion J., Ansoerge S., Henry O., Kamen A. A.* Scalable production of influenza virus in HEK-293 cells for efficient vaccine manufacturing // Vaccine. 2010. 28. 3661–3671.
- 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. 2015. 199. 68–76.

- 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., Airene K. J.* Production and purification of lentiviral vectors generated in 293T suspension cells with baculoviral vectors // *Gene Therapy*. 2011. 18. 531–538.
- Lesch H. P., Turpeinen S., Niskanen E. A., Mähönen A. J., Airene K. J., Ylä-Herttuala S.* Generation of lentivirus vectors using recombinant baculoviruses // *Gene Therapy*. 2008. 15. 1280–1286.
- Li Heng.* Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM // *arXiv:1303.3997v2*. 2013. N/A, N/A. 1–3.
- Li Lulin, Harwood Steve H., Rohrmann G. F.* Identification of additional genes that influence baculovirus late gene expression // *Virology*. 1999. 255, 1. 9–19.
- Liang Changyong, Li Min, Dai Xuejuan, Zhao Shuling, Hou Yanling, Zhang Yongli, Lan Dandan, Wang Yun, Chen Xinwen.* *Autographa californica* multiple nucleopolyhedrovirus PK-1 is essential for nucleocapsid assembly // *Virology*. 2013. 443, 2. 349–357.
- Liang Changyong, Su Xia, Xu Guodong, Dai Xuejuan, Zhao Shuling.* *Autographa californica* multiple nucleopolyhedrovirus PK1 is a factor that regulates high-level expression of very late genes in viral infection // *Virology*. 2017. 512, 12. 56–65.
- Lin Guangyun, Blissard Gary W.* Analysis of an *Autographa californica* Multicapsid Nucleopolyhedrovirus lef-6 -Null Virus: LEF-6 Is Not Essential for Viral Replication but Appears To Accelerate Late Gene Transcription // *Journal of Virology*. 2002a. 76, 11. 5503–5514.

- Lin Guangyun, Blissard Gary W.* Analysis of an *Autographa californica* nucleopolyhedrovirus lef-11 knockout: LEF-11 is essential for viral DNA replication // *Journal of Virology*. 2002b. 76, 6. 2770–2779.
- Liu C., Zhou Q., Li Y., Garner L. V., Watkins S. P., Carter L. J., Smoot J., Gregg A. C., Daniels A. D., Jerve S., Albaiu D.* Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases // *ACS Central Science*. 2020. 6. 315–331.
- Liu Chao, Li Zhaoferi, Wu Wenbi, Li Lingling, Yuan Meijin, Pan Lijing, Yang Kai, Pang Yi.* *Autographa californica* multiple nucleopolyhedrovirus ac53 plays a role in nucleocapsid assembly // *Virology*. 2008. 382, 1. 59–68.
- 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*. 2013. 90. 104–116.
- Liu Y., Chen D., Zhang X., Chen S., Yang D., Tang L., Yang X., Wang Y., Luo X., Wang M., Hu Z., Huang Y.* Construction of baculovirus-inducible CRISPR/Cas9 antiviral system targeting BmNPV in *Bombyx mori* // *Viruses*. 2022. 14, 1.
- 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*. 2014. 49, 1. 35–42.

- Lu A, Miller L K.* Identification of three late expression factor genes within the 33.8- to 43.4-map-unit region of Autographa californica nuclear polyhedrosis virus // Journal of Virology. 1994. 68, 10. 6710–6718.
- Lu A, Miller L K.* The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication // Journal of Virology. 1995. 69, 2. 975–982.
- Lu A., Miller L. K.* Regulation of Baculovirus Late and Very Late Gene Expression // The Baculoviruses. 1997. 1. 193–216.
- 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. 1993. 67, 8. 4566–4579.
- Lung Oliver Y., Cruz-Alvarez Marilyn, Blissard Gary W.* Ac23, an Envelope Fusion Protein Homolog in the Baculovirus Autographa californica Multicapsid Nucleopolyhedrovirus, Is a Viral Pathogenicity Factor // Journal of Virology. 2003. 77, 1. 328–339.
- 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. 2010. 406. 336–341.
- 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. 2011. 421. 34–41.

- 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*. 2014. 4. 1–6.
- 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 USA*. 2017. 114, 34. 9068–9073.
- Maghodia Ajay B., Jarvis Donald L., Geisler Christoph.* Complete genome sequence of the *Autographa californica* multiple nucleopolyhedrovirus strain E2 // *Genome Announcements*. 2014. 2, 6. 6–7.
- 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*. 2002. 30, 2. 482–496.
- 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*. 2006. 1. 1–26.
- 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., Van Der Oost J., Koonin E. V.* Evolution and classification of the CRISPR-Cas systems // *Nature Reviews Microbiology*. 2011. 9, 6. 467–477.

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., Van Der Oost J., Backofen R., Koonin E. V. An updated evolutionary classification of CRISPR-Cas systems // *Nature Reviews Microbiology*. 2015. 13, 11. 722–736.

Makela Anna R., Oker-Blom Christian. The Baculovirus Display Technology - An Evolving Instrument for Molecular Screening and Drug Delivery // *Combinatorial Chemistry and High Throughput Screening*. 2008. 11. 86–98.

Maranga L., Brazao T. F., Carrondo M. J. T. Virus-like particle production at low multiplicities of infection with the baculovirus insect cell system // *Biotechnology and Bioengineering*. 2003. 84. 245–253.

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*. 2011. 108, 5. 1056–1067.

Marek Martin, Romier Christophe, Galibert Lionel, Merten Otto-Wilhelm, van Oers Monique M. Baculovirus VP1054 Is an Acquired Cellular PUR α , a Nucleic Acid-Binding Protein Specific for GGN Repeats // *Journal of Virology*. 2013. 87, 15. 8465–8480.

Martínez-solís María, Herrero Salvador, Targovnik Alexandra Marisa. Engineering of the baculovirus expression system for optimized protein production // *Applied Microbiology and Biotechnology*. 2019. 103. 113–123.

- McLachlin Jeanne R., Escobar Julian C., Harrelson Jennifer A., Clem Rollie J., Miller Lois 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*. 2001. 81, 1-2. 77–91.
- McLachlin Jeanne R., Yang Song, Miller Lois K.* A baculovirus mutant defective in PKIP, a protein which interacts with a virus-encoded protein kinase // *Virology*. 1998. 246, 2. 379–391.
- Meghrouis 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*. 2005. 21. 154–160.
- Mena J. A., Aucoin M. G., Montes J., Chahal P. S., Kamen A. A.* Improving adeno-associated vector yield in high density insect cell cultures // *The Journal of Gene Medicine*. 2010. 12. 157–167.
- Mena J. A., Kamen A. A.* Insect cell technology is a versatile and robust vaccine manufacturing platform // *Expert Review of Vaccines*. 2011. 10, 7. 1063–1081.
- 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. 2011. 1–15.
- Mikhailov Victor S., Rohrmann George F.* Baculovirus Replication Factor LEF-1 Is a DNA Primase // *Journal of Virology*. 2002. 76, 5. 2287–2297.

- Milián E., Kamen A. A.* Current and Emerging Cell Culture Manufacturing Technologies for Influenza Vaccines // BioMed Research International. 2015. 2015. 1–11.
- Mishra Gourav, Chadha Pooja, Das Rakha H.* Serine/threonine kinase (pk-1) is a component of Autographa californica multiple nucleopolyhedrovirus (AcMNPV) very late gene transcription complex and it phosphorylates a 102 kDa polypeptide of the complex // Virus Research. 2008. 137, 1. 147–149.
- Mojica F. J. M., Díez-Villaseñor C.r, García-Martínez J., Soria E.* Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements // Journal of Molecular Evolution. 2005. 60, 2. 174–182.
- Moldován Norbert, Tombácz Dóra, Szucs Attila, Csabai Zsolt, Balázs Zsolt, Kis Emese, Molnár Judit, Boldogkoi Zsolt.* Third-generation Sequencing Reveals Extensive Polycistronism and Transcriptional Overlapping in a Baculovirus // Scientific Reports. 2018. 8, 1. 1–11.
- 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. 1996. 70, 7. 4607–4616.
- 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. 2012. 3. 1–16.

- Murguía-Meca F., Plata-Muñoz J. J., Hitchman R. B., Danquah J. O., Hughes D., Friend P. J., Fuggle S. V., King L. A.* Baculovirus as delivery system for gene transfer during hypothermic organ preservation // *Transplant International*. 2011. 24. 820–828.
- 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*. 2003. 84. 3029–3039.
- Novavax .* Novavax Files for COVID-19 Vaccine Authorization with Health Canada and Completes Submission for Rolling Review to European Medicines Agency // *Press Release and Statements*. 2021. 1. 1.
- OET Oxford Expression Technologies.* baculoCOMPLETE: A Complete Laboratory Guide to the Baculovirus Expression System and Insect Cell Culture // *User Guide: Oxford Expression Technologies*. 2019. 1. 1–95.
- O'Reilly D. R., Miller L. K., Luckow V. A.* Baculovirus expression vectors: A laboratory manual // *W. H. Freeman and Company*. 1992. 1. 1–347.
- Ohkawa T., Volkman L. E., Welch M. D.* Actin-based motility drives baculovirus transit to the nucleus and cell surface // *Journal of Cell Biology*. 2010. 190, 2. 187–195.
- Ohkawa Taro, Rowe Annette R., Volkman Loy E.* Identification of Six *Autographa californica* Multicapsid Nucleopolyhedrovirus Early Genes That Mediate Nuclear Localization of G-Actin // *Journal of Virology*. 2002. 76, 23. 12281–12289.

- Ohkawa Taro, Washburn Jan O., Sitapara Ronika, Sid Eric, Volkman Loy E.* Specific Binding of Autographa californica M Nucleopolyhedrovirus Occlusion-Derived Virus to Midgut Cells of Heliothis virescens Larvae Is Mediated by Products of pif Genes Ac119 and Ac022 but Not by Ac115 // *Journal of Virology*. 2005. 79, 24. 15258–15264.
- Okano K., Vanarsdall A. L., Rohrmann G. F.* A baculovirus alkaline nuclease knockout construct produces fragmented DNA and aberrant capsids // *Virology*. 2007. 359. 46–54.
- Olszewski J, Miller L K.* Identification and characterization of a baculovirus structural protein, VP1054, required for nucleocapsid formation // *Journal of Virology*. 1997. 71, 7. 5040–5050.
- Ono C., Kamagata T., Taka H., Sahara K., Ichiro Asano S., Bando H.* Phenotypic grouping of 141 BmNPVs lacking viral gene sequences // *Virus Research*. 2012. 165, 2. 197–206.
- Oomens A. G.P., Blissard G. W.* Requirement for GP64 to drive efficient budding of Autographa californica multicapsid nucleopolyhedrovirus // *Virology*. 1999. 254, 2. 297–314.
- Palomares L. A., Realpe M., Ramírez O. T.* An Overview of Cell Culture Engineering for the Insect Cell-Baculovirus Expression Vector System (BEVS) // *Animal Cell Culture*. 2015. 9. 501–519.
- 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*. 2020. 382. 29–40.
- Passarelli A L, Miller L K*. Identification and transcriptional regulation of the baculovirus lef-6 gene // *Journal of Virology*. 1994. 68, 7. 4458–4467.
- Passarelli A L, Todd J W, Miller L K*. A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases // *Journal of Virology*. 1994. 68, 7. 4673–4678.
- Pau M. G., Ophorst C., Koldijk M. H., Schouten G., Mehtali M., Uytdehaag F*. The human cell line PER . C6 provides a new manufacturing system for the production of influenza vaccines // *Vaccines*. 2001. 19. 2716–2721.
- 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*. 2019. 116, 11. 2823–2833.
- Pearson Margot N., Rohrmann George F*. Transfer, Incorporation, and Substitution of Envelope Fusion Proteins among Members of the Baculoviridae , Orthomyxoviridae , and Metaviridae (Insect Retrovirus) Families // *Journal of Virology*. 2002. 76, 11. 5301–5304.
- Pei Jie, Liu Ruilun, Li Wei, Qian Shasha, Yang Jie, Meng Shengli, Shen Shuo, Guo Jing*. Enhancement of exogenous protein stability in AcMNPV by overexpressing lef5 gene during passaging // *Applied Microbiology and Biotechnology*. 2025. 109, 1. 1–9.

- Pijlman G. P., Schijndel J. E. V., Vlak J. M.* Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells // *Journal of General Virology*. 2003. 84, 10. 2669–2678.
- Pijlman G. P., Van Den Born E., Martens D. E., Vlak J. M.* *Autographa californica* baculoviruses with large genomic deletions are rapidly generated in infected insect cells // *Virology*. 2001. 283. 132–138.
- Pinello Luca, Canver Matthew C., Hoban Megan D., Orkin Stuart H., Kohn Donald B., Bauer Daniel E., Yuan Guo Cheng.* Analyzing CRISPR genome-editing experiments with CRISPResso // *Nature Biotechnology*. 2016. 34, 7. 695–697.
- 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 USA*. 2014. 111, 29.
- Possee R. D., Chambers A. C., Graves L. P., Aksular M., King L. A.* Recent developments in the use of baculovirus expression vectors // *Current Issues in Molecular Biology*. 2020. 34. 215–230.
- 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*. 2005. 151, 3. 653–663.
- Pushko P., Kort T., Nathan M., Pearce M. B., Smith G. E., Tumpey T. M.* Recombinant H1N1 virus-like particle vaccine elicits protective immunity in ferrets against the 2009 pandemic H1N1 influenza virus // *Vaccine*. 2010. 28. 4771–4776.

- Pushko P., Tumpey T. M., Bu F., Knell J., Robinson R., Smith G, E.* Influenza virus-like particles comprised of the HA, NA, and M1 proteins of H9N2 influenza virus induce protective immune responses in BALB/c mice // *Vaccine*. 2005. 23. 5751–5759.
- Qi Yong, Wang Shan-Shan, Li Lu-Lin.* IE1 of *Autographa californica* Multiple Nucleopolyhedrovirus Activates Low Levels of Late Gene Expression in the Absence of Virus RNA Polymerase // *Microbiology Spectrum*. 2023. 11, 1. 1–16.
- Quick Joshua, Grubaugh Nathan D., Pullan Steven T., Claro Ingra M., Smith Andrew D., Gangavarapu Karthik, Oliveira Glenn, Robles-Sikisaka Refugio, Rogers Thomas F., Beutler Nathan A., Burton Dennis R., Lewis-Ximenez Lia Laura, Jesus Jacqueline Goes de, Giovanetti Marta, Hill Sarah C., Black Allison, Bedford Trevor, Corroll Miles W., Nunes Marcio, Alcantara Jr. Luiz Carlos, Sabino Ester C., Baylis Sally A., Faria Nuno R., Loose Matthew, Simpson Jared T., Pybus Oliver G., Andersen Kristian G., Loman Nicholas J.* Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples // *Nature Protocols*. 2017. 12, 6. 1261–1276.
- Ran F. A., Hsu P. D., Wright J., Agarwala V., Scott D. A., Zhang F.* Genome engineering using the CRISPR-Cas9 system // *Nature Protocols*. 2013. 8, 11. 2281–2308.
- Rapp Jeffrey C., Wilson Joyce A., Miller Lois K.* Nineteen Baculovirus Open Reading Frames, Including LEF-12, Support Late Gene Expression // *Journal of Virology*. 1998. 72, 12. 10197–10206.
- Reed L. J., Muench H.* A simple method for estimating fifty per cent endpoints // *The American Journal of Hygiene*. 1938. 27. 493–497.

- Reid W., O'Brochta D. A.* Applications of genome editing in insects // Current Opinion in Insect Science. 2016. 13. 43–54.
- Reilly L. M., Guarino L. A.* The pk-1 gene of *Autographa californica* multinucleocapsid nuclear polyhedrosis virus encodes a protein kinase // Journal of General Virology. 1994. 75, 11. 2999–3006.
- Reilly Linda M., Guarino Linda A.* The viral ubiquitin gene of *Autographa californica* nuclear polyhedrosis virus is not essential for viral replication // Virology. 1996. 218, 1. 243–247.
- 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 USA. 2013. 110, 47. 19012–19017.
- Rimmelzwaan G. F., Baars M., Claas E. C. J., Osterhaus A. D. M. E.* Comparison of RNA hybridization , hemagglutination assay , titration of infectious virus and immunofluorescence as methods for monitoring influenza virus replication in vitro // Journal of Virological Methods. 1998. 74. 57–66.
- Road S. P.* High-level expression of five foreign genes by a single recombinant baculovirus // Gene. 1995. 156. 229–233.
- Rohrmann G. F.* Baculovirus Molecular Biology - The AcMNPV genome: Gene content, conservation, and function // National Center for Biotechnology Information. 2019a. 4. 201–275.

- Rohrmann G. F.* Baculovirus Molecular Biology- Baculovirus infection: The cell cycle and apoptosis // National Center for Biotechnology Information. 2019b. 4. 141–150.
- Rohrmann G. F.* Baculovirus Molecular Biology- Introduction to the baculoviruses, their taxonomy, and evolution // National Center for Biotechnology Information. 2019c. 4. 1–19.
- Rohrmann G. F.* Baculovirus Molecular Biology- The baculovirus replication cycle: Effects on cells and insects // National Center for Biotechnology Information. 2019d. 4. 1–298.
- 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. 2010. 9, 10. 1149–1176.
- 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. 2011. 412. 167–178.
- Sander J. D., Joung J. K.* CRISPR-Cas systems for editing, regulating and targeting genomes // Nature Biotechnology. 2014. 32, 4. 347–350.
- Sari-Ak Duygu, Alomari Omar, Shomali Raghad Al, Lim Jackwee, Thimiri Govinda Raj Deepak B.* Advances in CRISPR-Cas9 for the Baculovirus Vector System: A Systematic Review // Viruses. 2023. 15, 1. 1–17.
- 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. 2009. 83, 21. 11123–11132.

- Sergeeva A., Kolonin M. G., Molldrem J. J., Pasqualini R., Arap W.* Display technologies: Application for the discovery of drug and gene delivery agents // *Advanced Drug Delivery Reviews*. 2006. 58. 1622–1654.
- Shah S. A., Erdmann S., Mojica F. J. M., Garrett R. A.* Protospacer recognition motifs: Mixed identities and functional diversity // *RNA Biology*. 2013. 10, 5. 891–899.
- 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*. 2017. 6, 7. 1393–1402.
- Shen Chun Fang, Meghrous Jamal, Kamen Amine.* Quantitation of baculo v irus particles by flow cytometry // *Journal of Virological Methods*. 2002. 105. 321–330.
- Simonin J. A., Cuccovia Warlet F. U., Bauzá M. d. R., Plastine M. d. P., Alfonso V., Olea F. D., Cerrudo C. S., Belaich M. N.* Early to Late VSV-G Expression in AcMNPV BV Enhances Transduction in Mammalian Cells but Does Not Affect Virion Yield in Insect Cells // *Vaccines*. 2025. 13, 7. 693.
- 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*. 2006. 36, 6. 454–465.
- 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*. 1983a. 46, 2. 584–593.

- 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 USA. 1985. 82, 24. 8404–8408.
- Smith G. E., Summers M. D., Fraser M. J.* Production of Human Beta Interferon in Insect Cells Infected with a Baculovirus Expression Vector // J. Molecular and Cellular Biology. 1983b. 3. 2156–2165.
- 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. 1983c. 45. 215–225.
- Smith R. H., Levy J. R., Kotin R. M.* A simplified baculovirus-AAV expression vector system coupled with one-step affinity purification yields high-titer rAAV stocks from insect cells // Molecular Therapy. 2009. 17. 1888–1896.
- 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. 2012. 30. 766–781.
- Summers M. D.* Milestones Leading to the Genetic Engineering of Baculoviruses as Expression Vector Systems and Viral Pesticides // Advances in Virus Research. 2006. 68. 3 – 73.
- Symington E., Rangarajan S., Lester W., Madan B., Pierce G. F., Raheja P., Millar C., Osmond D., Li M., Robinson T. M.* Valoctocogene roxaparvovec gene therapy provides

durable haemostatic control for up to 7 years for haemophilia A // *Haemophilia*. 2024. 30, 5. 1138–1147.

Takagi Toshimitsu, Taylor Gregory S., Kusakabe Takahiro, Charbonneau Harry, Buratowski Stephen. A protein tyrosine phosphatase-like protein from baculovirus has RNA 5'-triphosphatase and diphosphatase activities // *Proceedings of the National Academy of Sciences of the United States of America*. 1998. 95, 17. 9808–9812.

Tao Xue Ying, Choi Jae Young, Wang Yong, Roh Jong Yul, Lee Joo Hyun, Liu Qin, Park Jong Bin, Kim Jae Su, Kim Woojin, Je Yeon Ho. Functional characterization of *Autographa californica* multiple nucleopolyhedrovirus ORF43 and phenotypic changes of ORF43-knockout mutant // *Journal of Microbiology*. 2013. 51, 4. 515–521.

Technologies Oxford Expression. Baculovirus Gene Mutations and Protein Expression // *Oxford Expression Technologies Blog*. 2016. N/A. Available online.

Tian J. H., Patel N., Haupt R., Zhou H., Weston S., Hammond H., Logue J., Portnoff A. D., Norton J., Guebre-Xabier M., Zhou B., Jacobson K., Maciejewski S., Khatoon R., Wisniewska M., Moffitt W., Kluepfel-Stahl S., Ekechukwu B., Papin J., Boddapati S., Jason Wong C., Piedra P. A., Frieman M. B., Massare M. J., Fries L., Bengtsson K. L., Stertman L., Ellingsworth L., Glenn G., Smith G. E. SARS-CoV-2 spike glycoprotein vaccine candidate NVX-CoV2373 immunogenicity in baboons and protection in mice // *Nature Communications*. 2021. 12.

Todd Jason W, Passarelli A Lorena, Miller Lois K. Eighteen Baculovirus Genes, Including lef-11, p35, 39K, and p47 , Support Late Gene Expression // *Microbiology*. 1995. 69, 2. 968–974.

- Tomalski M. D., Wu J. G., Miller L. K.* The location, sequence, transcription, and regulation of a baculovirus DNA polymerase gene // *Virology*. 1988. 167, 2. 591–600.
- Tomalski Michael D., Eldridge Russ, Miller Lois K.* A baculovirus homolog of aCu/Zn superoxide dismutase gene // *Virology*. 1991. 184, 1. 149–161.
- Toth A. M., Kuo C., Khoo K., Jarvis D. L.* A new insect cell glycoengineering approach provides baculovirus-inducible glycoprotein expression and increases human-type glycosylation efficiency // *Journal of Biotechnology*. 2014. 182-183. 19–29.
- Tsai C. H., Wei S. C., Lo H. R., Chao Y. C.* Baculovirus as versatile vectors for protein display and biotechnological applications // *Current Issues in Molecular Biology*. 2020. 34. 231–255.
- Urabe M., Ding C., Kotin R. M.* Insect cells as a factory to produce adeno-associated virus type 2 vectors // *Human Gene Therapy*. 2002. 13. 1935–1943.
- Van Der Oost J., Westra E. R., Jackson R. N., Wiedenheft B.* Unravelling the structural and mechanistic basis of CRISPR-Cas systems // *Nature Reviews Microbiology*. 2014. 12, 7. 479–492.
- Vicente T., Roldão A., Peixoto C., Carrondo M. J. T., Alves P. M.* Large-scale production and purification of VLP-based vaccines // *Journal of Invertebrate Pathology*. 2011. 107. S42–S48.
- Vijayachandran L. S., Thimiri Govinda Raj D. B., Edelweiss E., Gupta K., Maier J., Gordeliy V., Fitzgerald D. J., Berger I.* Gene gymnastics // *Bioengineered*. 2013. 4, 5. 279–287.

- Volkman L. E., Goldsmith P. A.* Mechanism of neutralization of budded *Autographa californica* nuclear polyhedrosis virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis // *Virology*. 1985. 143. 185–195.
- Volkman L. E., Oppenheimer D. I.* Evidence for rolling circle replication of *Autographa californica* M nucleopolyhedrovirus genomic DNA Brief Report // *Archives of Virology*. 1997. 142. 2107–2113.
- Walker Bruce J., Abeel Thomas, Shea Terrance, Priest Margaret, Abouelliel Amr, Sakthikumar Sharadha, Cuomo Christina A., Zeng Qiandong, Wortman Jennifer, Young Sarah K., Earl Ashlee M.* Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement // *PLoS ONE*. 2014. 9, 11. 1–14.
- Wang RanRan, Deng Fei, Hou Dianhai, Zhao Yong, Guo Lin, Wang Hualin, Hu Zhihong.* Proteomics of the *Autographa californica* Nucleopolyhedrovirus Budded Virions // *Journal of Virology*. 2010. 84, 14. 7233–7242.
- Weissmann F., Petzold G., VanderLinden R., Huis in't Veld P. J., Brown N. G., Lampert F., Westermann S., Stark H., Schulman B. A., Peters J. M.* BiGBac enables rapid gene assembly for the expression of large multisubunit protein complexes // *Proceedings of the National Academy of Sciences of the USA*. 2016. 113, 19. E2564–E2569.
- Wickham T. J., Davis T., Granados R. R., Hammer D. A., Shuler M. L., Wood H. A.* Baculovirus defective interfering particles are responsible for variations in recombinant protein production as a function of multiplicity of infection // *Biotechnology Letters*. 1991. 13. 483–488.

- Wickham T. J., Davis T., Granados R. R., Shuler M. L., Wood H. A.* Screening of Insect Cell Lines for the Production of Recombinant Proteins and Infectious Virus in the Baculovirus Expression System // *Biotechnology Progress*. 1992. 8. 391–396.
- Wilm Andreas, Aw Pauline Poh Kim, Bertrand Denis, Yeo Grace Hui Ting, Ong Swee Hoe, Wong Chang Hua, Khor Chiea Chuen, Petric Rosemary, Hibberd Martin Lloyd, Nagarajan Niranjan.* LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets // *Nucleic Acids Research*. 2012. 40, 22. 11189–11201.
- Wu Carol P., Huang Yi-Ju, Wang Jen-Yeu, Wu Yueh-Lung, Lo Huei-Ru, Wang Jui-Ching, Chao Yu-Chan.* *Autographa californica* Multiple Nucleopolyhedrovirus LEF-2 Is a Capsid Protein Required for Amplification but Not Initiation of Viral DNA Replication // *Journal of Virology*. 2010. 84, 10. 5015–5024.
- Xiang Xingwei, Chen Lin, Hu Xiaolong, Yu Shaofang, Yang Rui, Wu Xiaofeng.* *Autographa californica* multiple nucleopolyhedrovirus odv-e66 is an essential gene required for oral infectivity // *Virus Research*. 2011. 158, 1-2. 72–78.
- Xiang Zaiying, Ye Qiaoyuan, Zhao Zihan, Wang Naian, Li Jinrong, Zou Minghai, Lau Cia Hin, Zhu Haibao, Wang Shu, Ding Yuanlin.* Development of a baculoviral CRISPR/Cas9 vector system for beta-2-microglobulin knockout in human pluripotent stem cells // *Molecular Genetics and Genomics*. 2024. 299, 1. 1–13.
- Xu Xiaodong, Zhou Xinyu, Nan Hao, Zhao Yu, Bai Yu, Ou Yanmei, Chen Hongying.* Aggregation of AcMNPV LEF-10 and its impact on viral late gene expression // *PLoS ONE*. 2016. 11, 5. 1–14.

- Yamagishi Junya, Burnett Erik D., Harwood Steven H., Blissard Gary W.* The AcMNPV pp31 gene is not essential for productive AcMNPV replication or late gene transcription but appears to increase levels of most viral transcripts // *Virology*. 2007. 365, 1. 34–47.
- 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 and Engineering Chemistry Research*. 2018. 57, 31. 10061–10070.
- Yu Yue, Zhang Tong, Lu Dongbo, Wang Jing, Xu Zhenhe, Zhang Yuanxing, Liu Qin.* Genome-wide nonessential gene identification of *Autographa californica* multiple nucleopolyhedrovirus // *Gene*. 2023. 863. 147239.
- Zhang J., Kalogerakis N., Behie L. A., Iatrou K.* Investigation of Reduced Serum and Serum-Free Media for the Cultivation of Insect Cells (Bm5) and the Production of Baculovirus (BmNPV) // *Biotechnology and Bioengineering*. 1992. 40, 10. 1165–1172.
- Zhang Xiaoyue, He Aiping, Zong Yuyu, Tian Houlu, Zhang Zhihui, Zhao Kaixia, Xu Xiaodong, Chen Hongying.* Improvement of protein production in baculovirus expression vector system by removing a total of 10 kb of nonessential fragments from *Autographa californica* multiple nucleopolyhedrovirus genome // *Frontiers in Microbiology*. 2023. 14, April. 1–13.
- Zhu J.* Mammalian cell protein expression for biopharmaceutical production // *Biotechnology Advances*. 2012. 30, 5. 1158–1170.

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*. 2015. 5. 1–13.

Zu Putlitz J., Kubasek W. L., Duchene M., Marget M., Specht B. U. von, Domdey H. Antibody production in baculovirus-infected insect cells. // *Biotechnology* (Nature Publishing Company). 1990. 8. 651–654.

van Oers M. M. Opportunities and challenges for the baculovirus expression system // *Journal of Invertebrate Pathology*. 2011. 107, SUPPL. S3–S15.

van Oers M. M., Pijlman G. P., Vlak J. M. Thirty years of baculovirus-insect cell protein expression: From dark horse to mainstream technology // *Journal of General Virology*. 2015. 96. 6–23.

Appendix A

Chapter 3 supplementary

Table A.1: Spacer sequences used in this study for the control and AcMNPV genes.

Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
Scrambled control	N/A	N/A	CACCTTGAAGCGCATGAACT	N/A	N/A	N/A
<i>ptp</i>	+87	1	GTCACGTACGCAAACAACCTC	GGG	Antisense	60.07
<i>ptp</i>	-206	2	TGGTGTGCATTTTTTTGCGGG	CGG	Sense	54.87
<i>AcOrf-4</i>	-78	1	CGAGTCAAGTGATCAAAGTG	TGG	Sense	64.63
<i>AcOrf-4</i>	+116	2	CGACGGCATGATTTAAAAGCG	AGG	Antisense	61.82
<i>lef2</i>	+30	1	AAACATGACGCTCTAATGAG	CGG	Antisense	71.37
<i>lef2</i>	-101	2	ATTGACCCTAACTCCATACA	CGG	Sense	66.79
<i>lef2</i>	-164	3	GCGATTGTACATGCTGTAA	CGG	Sense	44.21
<i>pk-1</i>	-75	1	AAAATAATCAACGGGCGCTT	TGG	Sense	36.36
<i>pk-1</i>	+194	2	TTGGGTGGTCGCTCATTAAC	TGG	Antisense	34.76
<i>pk-1</i>	-251	3	CAACAACCAAGTGATCGTGA	TGG	Sense	62.85
<i>AcOrf-13</i>	-61	1	TGCATTGAACCGATCTTCGG	CGG	Sense	64.85
<i>AcOrf-13</i>	+329	2	CAAACGCGATCTCAACTCGT	TGG	Antisense	54.37

Table A.1 continued from previous page

Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
<i>lef1</i>	-68	1	CCGTCATGAACGCGTACTTG	CGG	Sense	59.50
<i>lef1</i>	+269	2	GAAAATTTACATTGGCGCCA	CGG	Antisense	59.33
<i>AcOrf-19</i>	+106	1	CAATGTGCTTACAAACGCGA	TGG	Antisense	59.82
<i>AcOrf-19</i>	+157	2	CGAGTTGATCAAATTGCACG	CGG	Antisense	75.24
<i>AcOrf-22</i>	+142	1	GTAGTGCCTTCTACGCATTA	CGG	Antisense	50.82
<i>AcOrf-22</i>	-197	2	TGATACTAATCTCGGCACGT	TGG	Sense	67.89
<i>AcOrf-22</i>	-296	3	GATATGCGATAACCCGTCTG	CGG	Sense	75.41
<i>env-prot</i>	-87	1	TCGACGACGAGTACAGTTGT	CGG	Sense	69.33
<i>env-prot</i>	+558	2	AACATTACTCTGTCTTTGGG	CGG	Antisense	57.06
<i>pkip</i>	+116	1	CGACAAAATGTTGTGCATAG	CGG	Antisense	64.03
<i>pkip</i>	+137	2	GGCCGATATCAAAGGCCAAG	TGG	Antisense	73.76
<i>AcOrf-26</i>	+75	1	GCGCATTTGGAATGACTAAA	AGG	Antisense	44.66
<i>AcOrf-26</i>	-123	2	GTTTTGAAGCGGTTGAGCAA	CGG	Sense	57.93

Table A.1 continued from previous page

Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
<i>iap1</i>	+61	1	CATGTCGAATACGTGTTTCGG	CGG	Antisense	73.90
<i>iap1</i>	-333	2	AACGCTGTACTAGTGAAAGA	AGG	Sense	55.75
<i>lef6</i>	-79	1	ATTTGAAAAACAGCGTCGAC	TGG	Sense	52.51
<i>lef6</i>	-163	2	ACTGCAACGGCAGATACTAC	TGG	Sense	56.43
<i>sod</i>	-99	1	AATTTGCCTCGAGGTTTGCA	CGG	Sense	59.81
<i>sod</i>	+169	2	GGTGGGATTAAAGTGCTCAC	CGG	Antisense	62.24
<i>fgf</i>	-89	1	ACTGCCGGTTAATAAACAGC	TGG	Sense	60.16
<i>fgf</i>	-190	2	GACGATGCGATTTCTGTCAA	CGG	Sense	54.94
<i>AcOrf-34</i>	-75	1	CGTTGAAGGGAAATAATTCG	TGG	Sense	58.09
<i>AcOrf-34</i>	+196	2	GGGAGGTGATCGATACGATG	GGG	Antisense	68.76
<i>v-ubi</i>	-29	1	CAAACCATTACCGCCGAAA	CGG	Sense	49.98
<i>v-ubi</i>	+91	2	TACGGGCACACCTTCTTTAT	CGG	Antisense	20.48
<i>39k/pp31</i>	-107	1	AATTTACAATAACGGCCAGC	TGG	Sense	54.18

Table A.1 continued from previous page

Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
<i>39k/pp31</i>	+259	2	CTAAGATCAAGCAGCCCGAG	TGG	Antisense	66.56
<i>lef11</i>	-4	1	TAAGTGCGTGCAATTTTTGG	GGG	Sense	54.31
<i>lef11</i>	+59	2	CGAAATACAAGCGCTGTTCA	GGG	Antisense	57.73
<i>AcOrf-38</i>	+23	1	TATGATAATCGAGCCGGACA	AGG	Antisense	60.98
<i>AcOrf-38</i>	-94	2	AAAAGTGTCGTTTCATGTCCG	CGG	Sense	56.46
<i>AcOrf-38</i>	-209	3	TGAACGCGCTGTCAAAAAAC	CGG	Sense	49.25
<i>p47</i>	-114	1	ACTTCGCGCTTTACAAAACG	AGG	Sense	66.53
<i>p47</i>	+295	2	TACAGCTTGTGTGCCGAGAT	CGG	Antisense	51.60
<i>AcOrf-41</i>	-78	1	TTGAACGTTTTACGACAGCA	GGG	Sense	65.79
<i>AcOrf-41</i>	-128	2	TTTGTGCGTGTCAGACGACA	CGG	Sense	67.55
<i>AcOrf-41</i>	+178	3	GCGGAACGATACAAAATTGC	AGG	Antisense	54.68
<i>gta</i>	+376	1	TGCCAATAAAACGTCGTACG	TGG	Antisense	68.43
<i>gta</i>	-430	2	CAAGTCTGTTTTCAACCCGC	TGG	Sense	56.09

Table A.1 continued from previous page

Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
<i>gta</i>	-523	3	CTTTGACCGCAACAAACCGA	TGG	Sense	63.82
<i>AcOrf-43</i>	-62	1	GTTTAGTAACATGTCCCCTT	CGG	Sense	53.26
<i>AcOrf-43</i>	-105	2	CGCATGGCCATTGTTAAAAA	CGG	Sense	38.78
<i>AcOrf-45</i>	-118	1	AAGTTAACGAACGCATCAGG	CGG	Sense	68.12
<i>AcOrf-45</i>	-323	2	TATGAAGCAATTGCTGCGCG	AGG	Sense	59.70
<i>odv-e66</i>	+157	1	GTTGGTAGTGGTGGTAGCAA	CGG	Antisense	59.74
<i>odv-e66</i>	-551	2	TTTCACAATCACAATGCCCG	AGG	Sense	71.50
<i>odv-e66</i>	-732	3	ACGTACAGTCAAATCTTGCG	CGG	Sense	69.40
<i>pcna</i>	-69	1	TCACAATCAAAAGTAGCGTG	CGG	Sense	65.73
<i>pcna</i>	-198	2	CTGTTTATGGACACGTTGAG	CGG	Sense	70.96
<i>pcna</i>	+254	3	CAGCTCTGTACTGATGAAAG	CGG	Antisense	76.61
<i>lef8</i>	+476	1	CGCTAGCGATTACGTAGTGA	CGG	Antisense	63.69
<i>lef8</i>	-951	2	CCGACGCGATATTTTAACAC	GGG	Sense	59.05

Table A.1 continued from previous page

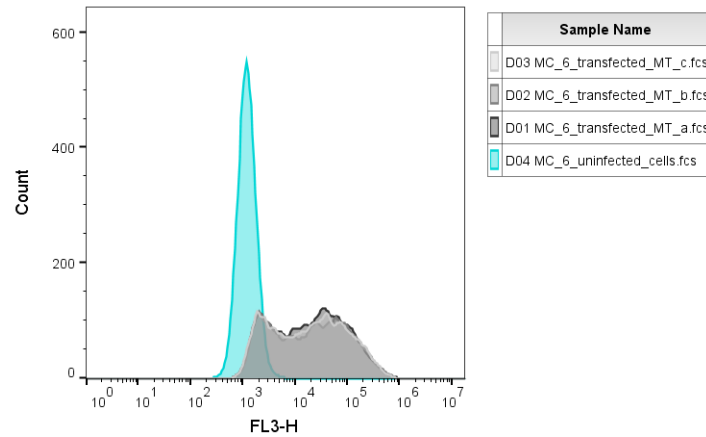
Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
<i>AcOrf-51</i>	+205	1	AACTTGTTTCGTCACCACGT	TGG	Antisense	61.58
<i>AcOrf-51</i>	-248	2	CGAAAAGCACCATTACAACA	CGG	Sense	71.33
<i>AcOrf-52</i>	+107	1	AAAGTGGACCCAAGACGTCA	TGG	Antisense	58.89
<i>AcOrf-52</i>	-182	2	TCACAACCGACAAGCACATT	TGG	Sense	41.56
<i>AcOrf-53</i>	-99	1	GAGAAAATTGACGATAACGG	GGG	Sense	73.06
<i>AcOrf-53</i>	+136	2	CAAGTTTAACATGCCAGTGT	CGG	Antisense	51.34
<i>AcOrf-53</i>	-172	3	ACGAACAATGTATTCAGCGT	TGG	Sense	52.53
<i>lef10</i>	-2	1	GACGAACGTATGGTTCGCGA	CGG	Sense	66.31
<i>lef10</i>	+112	2	GCACAGAGGTCTAACTTGAT	CGG	Antisense	50.57
<i>AcOrf-54</i>	+40	1	GAAGGGCGTTAATTTACCG	AGG	Antisense	69.79
<i>AcOrf-54</i>	+105	2	ACTTTGCAATTCGCTCGTCG	AGG	Antisense	59.74
<i>AcOrf-54</i>	-332	3	ACATTTTAGAAGCGTCGACG	AGG	Sense	69.60
<i>AcOrf-58/59</i>	-39	1	AATAACTTTTTGGCGCGAGG	AGG	Sense	60.96

Table A.1 continued from previous page

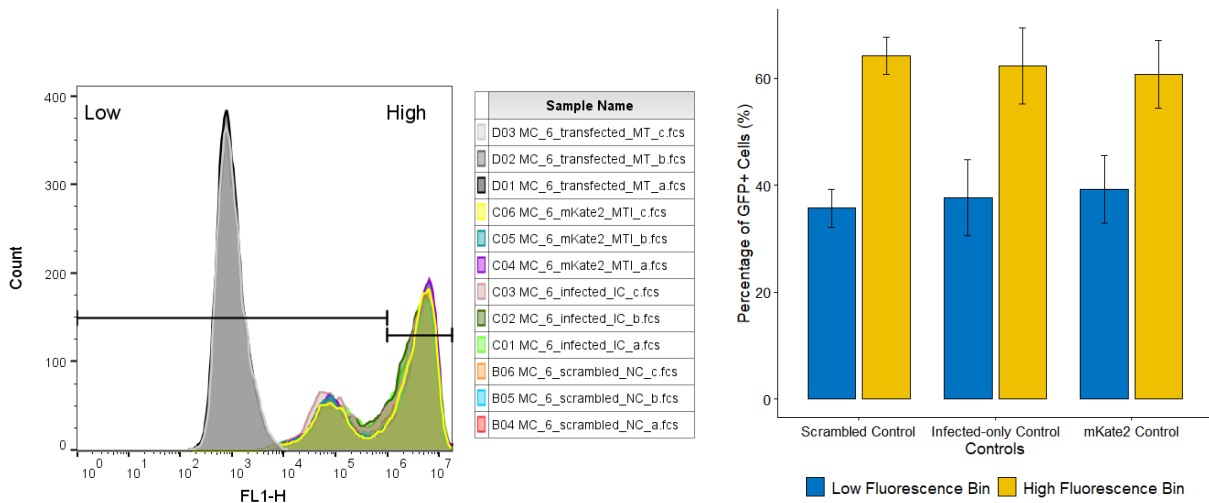
Gene	Location [†]	Target#	Spacer sequence (5'-3')	PAM	Strand	Efficiency [‡]
<i>AcOrf-58/59</i>	+400	2	AAAATTTTCCTATCGGCGAAG	TGG	Antisense	60.48
<i>AcOrf-60</i>	-23	1	TCCTCTTACCGTTATAGGGA	AGG	Sense	49.72
<i>AcOrf-60</i>	+185	2	CGACACTACTACTAGTACAG	AGG	Antisense	59.61
<i>fp</i>	+56	1	GCAAATCGACGAAAATGTGT	CGG	Antisense	69.90
<i>fp</i>	+148	2	TATACGGTATTCACGACAGC	AGG	Antisense	57.23
<i>lef9</i>	+116	1	ACTTGTGAGGGTCTAATATG	AGG	Antisense	60.67
<i>lef9</i>	-590	2	GTACGCGACGTTTCTCAACA	CGG	Sense	66.23
<i>lef9</i>	-649	3	TCAACGAAATTATGCCGCCG	CGG	Sense	66.82
<i>AcOrf-63</i>	-122	1	TTGTGAAACTCGTCTGAAAC	CGG	Sense	45.69
<i>AcOrf-63</i>	-149	2	CAGTTGCAACACTTTAACTG	CGG	Sense	71.35

[†] Distance of spacer sequence from the 5' end of a gene/ORF;

[‡] Targeting efficiency calculated by the sgRNA target designing web tool CHOPCHOP.



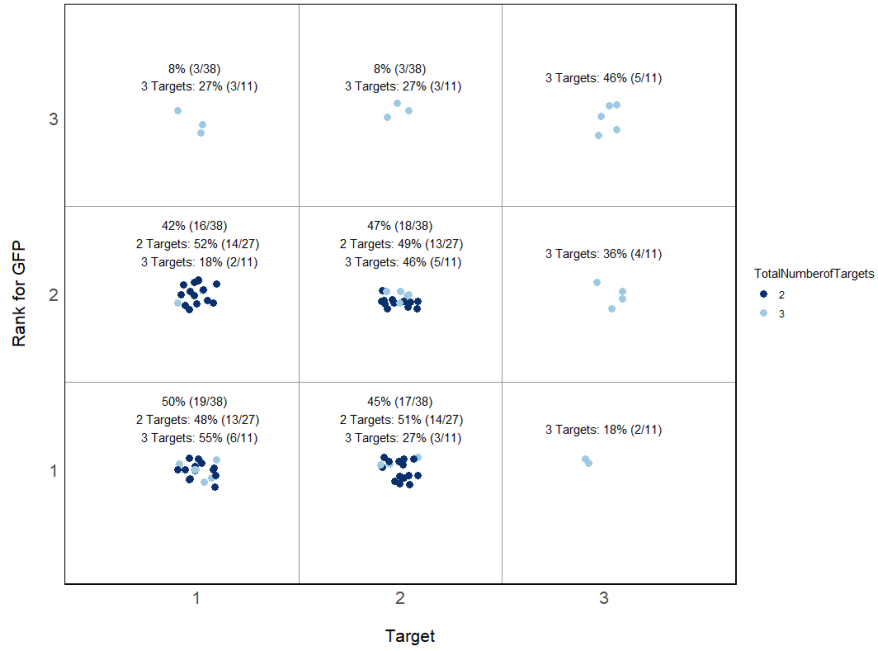
(a) Fluorescence distribution of transfected-only control and only cells



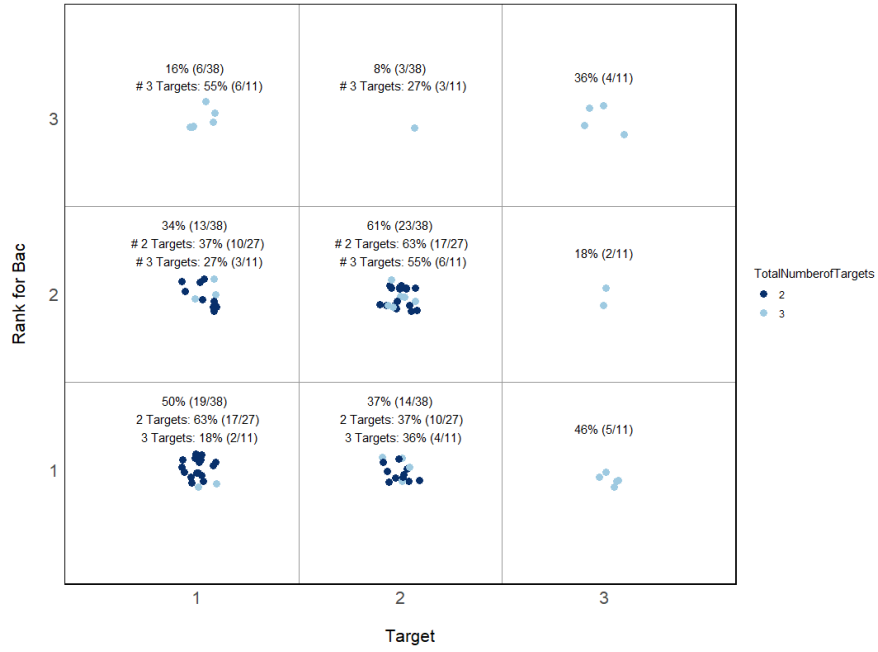
(b) Fluorescence intensity distribution

(c) GFP production

Figure A.1: Flow cytometry profiles of controls for T-I assay

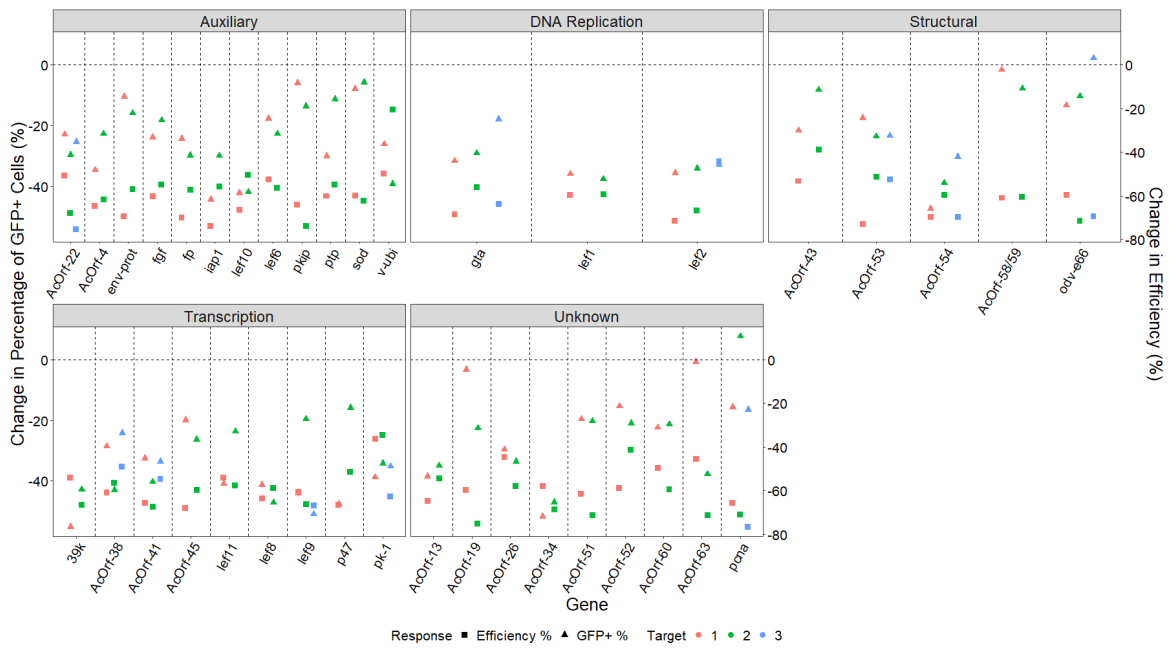


(a) Ranked Effect on GFP

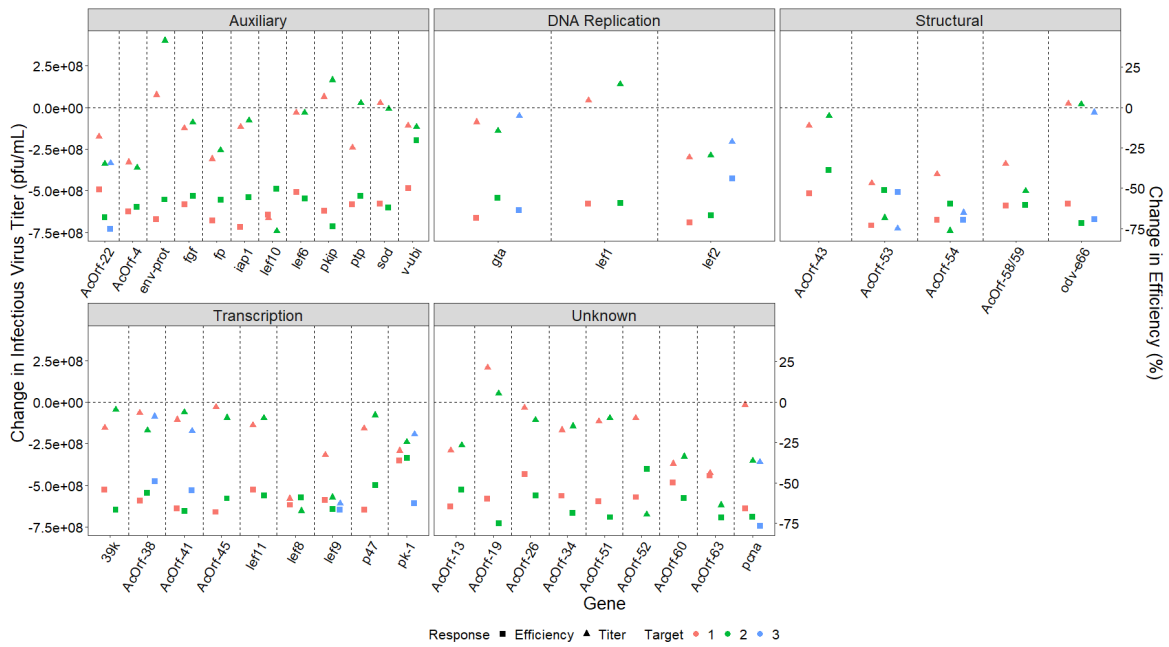


(b) Ranked Effect on Baculovirus Titer

Figure A.2: Ranked effect



(a) Change in GFP production and corresponding targeting efficiency



(b) Change in virus titer and corresponding targeting efficiency

Figure A.3: Change in GFP or BV with respect to targeting efficiency

Appendix B

Chapter 4 supplementary

Table B.1: Primers used to construct the scrambled control plasmid.

Template	Primer Sequence (5'-3')	Amplicon
SfU6 gBlock gene fragment	atctgcttagggttaggcgt cggtggtcgagcacgaattt	SfU6 promoter fragment ¹
pCFD4-U6:1_U6:3tandemgRNAs (Port et al., 2014)	tcgaccaccgcaccttgaagcgcataactgttttagagctagaaatagc cacaccacaaatatactggt	gRNA scaffold with scrambled spacer sequence and SfU6 homologous sequence ²
Amplicon 1 + Amplicon 2	atctgcttagggttaggcgt cacaccacaaatatactggt	SfU6-sgRNA DNA fragment
SfU6-sgRNA DNA fragment	tagaaacgcaatctgcttagggttaggc [†] aagcggagacacaccacaaatatactgttgc [†]	SfU6-sgRNA insert
pBR322-TIMER (Claudi et al., 2014)	ttgtggtgtgtcttccgcttctcgtc [†] ctaagcagattgcgtttctacaaactcttttg [†]	Backbone fragment containing ampR and ori

[†] Gibson primers with overhangs for constructing pSfU6-sgRNA scrambled control plasmid by Gibson assembly; ¹ Amplicon 1;

² Amplicon 2.

Table B.2: Positions of the tiled-amplicon primers in the p6.9GFP genome (shotgun-sequenced rBEV) as generated by Primal Scheme (Quick et al., 2017) and used in this study. Forward primers are denoted by ‘+’ or LEFT and reverse primers by ‘-’ or RIGHT.

Genome	Start Position (bp)	End Position (bp)	Primer Name	Pool	Forward/ Reverse Primer	Sequence (5'-3')
p6.9GFP	215	237	p6.9GFP_NGS_1_LEFT	1	+	TGACATCATCCACTGATCGTGC
p6.9GFP	5168	5190	p6.9GFP_NGS_1_RIGHT	1	-	ATGTTCGTAGAAGAAGCAGTCGC
p6.9GFP	4861	4883	p6.9GFP_NGS_2_LEFT	2	+	AAGATCAAGCTGTGCATGAGGG
p6.9GFP	9662	9684	p6.9GFP_NGS_2_RIGHT	2	-	GCTCCCCTCCTTGCGAAATTAT
p6.9GFP	9286	9308	p6.9GFP_NGS_3_LEFT	1	+	CGATTTGCTCCAACACTTCACG
p6.9GFP	14277	14299	p6.9GFP_NGS_3_RIGHT	1	-	GTGCTTGGCAAAACACGTACTC
p6.9GFP	13904	13926	p6.9GFP_NGS_4_LEFT	2	+	TTTGACACACACCATCACCTCC
p6.9GFP	18910	18932	p6.9GFP_NGS_4_RIGHT	2	-	ATGCCGTACTTTTATCCAGGCC
p6.9GFP	18464	18486	p6.9GFP_NGS_5_LEFT	1	+	GTCTCAATCAGTGGACGTGCAT
p6.9GFP	23598	23620	p6.9GFP_NGS_5_RIGHT	1	-	TCGGCGATCTTGTTAGCATACG
p6.9GFP	23081	23103	p6.9GFP_NGS_6_LEFT	2	+	ACAAGTATGTGGACGTGTGCTC
p6.9GFP	28095	28117	p6.9GFP_NGS_6_RIGHT	2	-	TATCGCTTGCTGGCACTTGTAG
p6.9GFP	27734	27756	p6.9GFP_NGS_7_LEFT	1	+	TCGTTCTTCAGTGCCACATACG

Table B.2 continued from previous page

Genome	Start Position (bp)	End Position (bp)	Primer Name	Pool	Forward/ Reverse Primer	Sequence (5'-3')
p6.9GFP	32781	32803	p6.9GFP_NGS_7_RIGHT	1	–	AACATGGATTACATACGCGGCA
p6.9GFP	32393	32415	p6.9GFP_NGS_8_LEFT	2	+	ATTGATTGATCTCGAGCCACCG
p6.9GFP	37376	37398	p6.9GFP_NGS_8_RIGHT	2	–	AGTAGCGGTTGTAGCATTGAGC
p6.9GFP	36843	36865	p6.9GFP_NGS_9_LEFT	1	+	CAACGACGACGACCCCTATTTT
p6.9GFP	41987	42009	p6.9GFP_NGS_9_RIGHT	1	–	ACTCTGTGGCGGTAAACAAGTC
p6.9GFP	41476	41498	p6.9GFP_NGS_10_LEFT	2	+	GCTGCGGTAAACACACCTTTC
p6.9GFP	46513	46535	p6.9GFP_NGS_10_RIGHT	2	–	CAGCTGCCCAAATTCCTGACTA
p6.9GFP	46014	46036	p6.9GFP_NGS_11_LEFT	1	+	GCCGAGGCGAGAAAAACAATTC
p6.9GFP	51107	51129	p6.9GFP_NGS_11_RIGHT	1	–	ATGATATATGTCGGCGCCACAC
p6.9GFP	50809	50831	p6.9GFP_NGS_12_LEFT	2	+	GGCTAGAGATGTTGTTGCGTGA
p6.9GFP	55567	55589	p6.9GFP_NGS_12_RIGHT	2	–	CATATTTCCGTCTTGCCGCAAC
p6.9GFP	55190	55212	p6.9GFP_NGS_13_LEFT	1	+	GATGCTTTCGACATGTTGTGGC
p6.9GFP	60024	60046	p6.9GFP_NGS_13_RIGHT	1	–	GCGGGTTCAACGACATGGTATA
p6.9GFP	59646	59668	p6.9GFP_NGS_14_LEFT	2	+	ACAACGAGAGAATAAGAGCGGC
p6.9GFP	64499	64521	p6.9GFP_NGS_14_RIGHT	2	–	TAATGAACCACGAATCCTCCGC

Table B.2 continued from previous page

Genome	Start Position (bp)	End Position (bp)	Primer Name	Pool	Forward/ Reverse Primer	Sequence (5'-3')
p6.9GFP	64211	64233	p6.9GFP_NGS_15_LEFT	1	+	GCCGGGCTGACGATAATAAACA
p6.9GFP	69071	69093	p6.9GFP_NGS_15_RIGHT	1	-	TCGTTAAC TTTACTACTGGCCGG
p6.9GFP	68778	68800	p6.9GFP_NGS_16_LEFT	2	+	CGAGTTTATTTTGAGCGGCGAC
p6.9GFP	73714	73736	p6.9GFP_NGS_16_RIGHT	2	-	CACAAAAATCAGAGCCGTGCTG
p6.9GFP	73308	73330	p6.9GFP_NGS_17_LEFT	1	+	CCCGAAATGCCTATCAACACCA
p6.9GFP	78317	78339	p6.9GFP_NGS_17_RIGHT	1	-	TGACTTTAACAGCTCGGACTGC
p6.9GFP	78016	78038	p6.9GFP_NGS_18_LEFT	2	+	CACCGCTACAGCAGAGCAATTA
p6.9GFP	82847	82869	p6.9GFP_NGS_18_RIGHT	2	-	GCCGACGAGCTCAGCATTTATA
p6.9GFP	82526	82548	p6.9GFP_NGS_19_LEFT	1	+	CCATCAAGCAGACTTTTAGCGC
p6.9GFP	87495	87517	p6.9GFP_NGS_19_RIGHT	1	-	GCGGTACCGAAATTCCGTTTTG
p6.9GFP	87140	87162	p6.9GFP_NGS_20_LEFT	2	+	ATGTGGTACCGGTTGAAGAACG
p6.9GFP	92156	92178	p6.9GFP_NGS_20_RIGHT	2	-	AAGCAACTGTGACGCCATAGAC
p6.9GFP	91767	91789	p6.9GFP_NGS_21_LEFT	1	+	CCAACAAACAGCCCAACATGAG
p6.9GFP	96662	96684	p6.9GFP_NGS_21_RIGHT	1	-	GCCTGTGCTGCACTATGGATAA
p6.9GFP	96138	96160	p6.9GFP_NGS_22_LEFT	2	+	GGTTCCACCAAATTGTGAGGGA

Table B.2 continued from previous page

Genome	Start Position (bp)	End Position (bp)	Primer Name	Pool	Forward/ Reverse Primer	Sequence (5'-3')
p6.9GFP	101172	101194	p6.9GFP_NGS_22_RIGHT	2	–	CGTGTGCATTGCCTTCGATTAC
p6.9GFP	100792	100814	p6.9GFP_NGS_23_LEFT	1	+	TGCCATTTGTCCGCAATTGTTT
p6.9GFP	105698	105720	p6.9GFP_NGS_23_RIGHT	1	–	AGGCGAGACTTGAACTCACAAC
p6.9GFP	105199	105221	p6.9GFP_NGS_24_LEFT	2	+	GTAACGGCCAATTCAACGTGAC
p6.9GFP	110209	110231	p6.9GFP_NGS_24_RIGHT	2	–	GGTCAACGACGCAAACATGATG
p6.9GFP	109698	109720	p6.9GFP_NGS_25_LEFT	1	+	GGGAGAGTGCCGTTTTTCAAGA
p6.9GFP	114802	114824	p6.9GFP_NGS_25_RIGHT	1	–	CCGAATTTTTTGAACGACGACGG
p6.9GFP	114292	114314	p6.9GFP_NGS_26_LEFT	2	+	TCGGGTCCTATACGAAGCGTTA
p6.9GFP	119276	119298	p6.9GFP_NGS_26_RIGHT	2	–	ATATAGTGTTGCAGCGCTACCG
p6.9GFP	118924	118949	p6.9GFP_NGS_27_LEFT	1	+	TTCGTTGTGCATTTCAAAGCTTTTG
p6.9GFP	123722	123744	p6.9GFP_NGS_27_RIGHT	1	–	TGGTAAAAGCGAATGGTCCGT
p6.9GFP	123430	123452	p6.9GFP_NGS_28_LEFT	2	+	CGTTAAAATGCTAAGCCGCGAG
p6.9GFP	128206	128228	p6.9GFP_NGS_28_RIGHT	2	–	CAACTGAACCCGTCGTCTGATT
p6.9GFP	126523	126545	p6.9GFP_NGS_29_LEFT	1	+	TTCCGTTGTCCGACGCTATAAC
p6.9GFP	131340	131363	p6.9GFP_NGS_29_RIGHT	1	–	TTGGGGTCAACATCGATAGTGTC

Table B.2 continued from previous page

Genome	Start Position (bp)	End Position (bp)	Primer Name	Pool	Forward/ Reverse Primer	Sequence (5'-3')
p6.9GFP	129190	129212	p6.9GFP_NGS_30_LEFT	2	+	CAAACAAATCTGGCGACTGTGG
p6.9GFP	2409	2431	p6.9GFP_NGS_30_RIGHT	2	-	CGGTGTACAGATACTTGTGCGT

Table B.3: Tiled-amplicon primers generated by Primal Scheme (Quick et al., 2017) and used in this study.

Primer Name	Pool	Sequence (5'-3')	Size (bp)	%GC	Tm (use 65)
p6.9GFP_NGS_1_LEFT	1	TGACATCATCCACTGATCGTGC	22	50	60.92
p6.9GFP_NGS_1_RIGHT	1	ATGTCGTAGAAGAAGCAGTCGC	22	50	60.91
p6.9GFP_NGS_2_LEFT	2	AAGATCAAGCTGTGCATGAGGG	22	50	61.13
p6.9GFP_NGS_2_RIGHT	2	GCTCCCCTCCTTGCGAAATTAT	22	50	60.93
p6.9GFP_NGS_3_LEFT	1	CGATTTGCTCCAACACTTCACG	22	50	60.84
p6.9GFP_NGS_3_RIGHT	1	GTGCTTGGCAAAACACGTACTC	22	50	61.03
p6.9GFP_NGS_4_LEFT	2	TTTGACACACACCATCACCTCC	22	50	60.93
p6.9GFP_NGS_4_RIGHT	2	ATGCCGTACTTTTATCCAGGCC	22	50	60.93
p6.9GFP_NGS_5_LEFT	1	GTCTCAATCAGTGGACGTGCAT	22	50	61.11
p6.9GFP_NGS_5_RIGHT	1	TCGGCGATCTTGTTAGCATACG	22	50	61.03
p6.9GFP_NGS_6_LEFT	2	ACAAGTATGTGGACGTGTGCTC	22	50	61.05
p6.9GFP_NGS_6_RIGHT	2	TATCGCTTGCTGGCACTTGTAG	22	50	61.17
p6.9GFP_NGS_7_LEFT	1	TCGTTCTTCAGTGCCACATACG	22	50	61.1
p6.9GFP_NGS_7_RIGHT	1	AACATGGATTACATACGCGGCA	22	45.45	60.92

Table B.3 continued from previous page

Primer Name	Pool	Sequence (5'-3')	Size (bp)	%GC	Tm (use 65)
p6.9GFP_NGS_8_LEFT	2	ATTGATTGATCTCGAGCCACCG	22	50	60.98
p6.9GFP_NGS_8_RIGHT	2	AGTAGCGGTTGTAGCATTGAGC	22	50	61.17
p6.9GFP_NGS_9_LEFT	1	CAACGACGACGACCCCTATTTT	22	50	61.11
p6.9GFP_NGS_9_RIGHT	1	ACTCTGTGGCGGTAAACAAGTC	22	50	60.99
p6.9GFP_NGS_10_LEFT	2	GCTGCGGTTAAACACACCTTTC	22	50	61.03
p6.9GFP_NGS_10_RIGHT	2	CAGCTGCCCAAATTCCTGACTA	22	50	60.8
p6.9GFP_NGS_11_LEFT	1	GCCGAGGCGAGAAAAACAATTC	22	50	61.16
p6.9GFP_NGS_11_RIGHT	1	ATGATATATGTCGGCGCCACAC	22	50	61.11
p6.9GFP_NGS_12_LEFT	2	GGCTAGAGATGTTGTTGCGTGA	22	50	61.11
p6.9GFP_NGS_12_RIGHT	2	CATATTTCCGTCTTGCCGCAAC	22	50	60.97
p6.9GFP_NGS_13_LEFT	1	GATGCTTTCGACATGTTGTGGC	22	50	61.16
p6.9GFP_NGS_13_RIGHT	1	GCGGGTTCAACGACATGGTATA	22	50	60.92
p6.9GFP_NGS_14_LEFT	2	ACAACGAGAGAATAAGAGCGGC	22	50	60.91
p6.9GFP_NGS_14_RIGHT	2	TAATGAACCACGAATCCTCCGC	22	50	60.92
p6.9GFP_NGS_15_LEFT	1	GCCGGGCTGACGATAATAAACA	22	50	61.24

Table B.3 continued from previous page

Primer Name	Pool	Sequence (5'-3')	Size (bp)	%GC	Tm (use 65)
p6.9GFP_NGS_15_RIGHT	1	TCGTTAACTTTACTACTGGCCGG	22	50	61.05
p6.9GFP_NGS_16_LEFT	2	CGAGTTTATTTTGAGCGGCGAC	22	50	60.96
p6.9GFP_NGS_16_RIGHT	2	CACAAAAATCAGAGCCGTGCTG	22	50	61.1
p6.9GFP_NGS_17_LEFT	1	CCCGAAATGCCTATCAACACCA	22	50	61.13
p6.9GFP_NGS_17_RIGHT	1	TGACTTTAACAGCTCGGACTGC	22	50	61.05
p6.9GFP_NGS_18_LEFT	2	CACCGCTACAGCAGAGCAATTA	22	50	61.17
p6.9GFP_NGS_18_RIGHT	2	GCCGACGAGCTCAGCATTTATA	22	50	61.04
p6.9GFP_NGS_19_LEFT	1	CCATCAAGCAGACTTTTAGCGC	22	50	60.65
p6.9GFP_NGS_19_RIGHT	1	GCGGTACCGAAATTCCGTTTTG	22	50	61.15
p6.9GFP_NGS_20_LEFT	2	ATGTGGTACCGGTTGAAGAACG	22	50	61.05
p6.9GFP_NGS_20_RIGHT	2	AAGCAACTGTGACGCCATAGAC	22	50	61.37
p6.9GFP_NGS_21_LEFT	1	CCAACAAACAGCCCAACATGAG	22	50	60.73
p6.9GFP_NGS_21_RIGHT	1	GCCTGTGCTGCACTATGGATAA	22	50	60.93
p6.9GFP_NGS_22_LEFT	2	GGTTCACCAAATTGTGAGGGA	22	50	60.94
p6.9GFP_NGS_22_RIGHT	2	CGTGTGCATTGCCTTCGATTAC	22	50	60.96

Table B.3 continued from previous page

Primer Name	Pool	Sequence (5'-3')	Size (bp)	%GC	T _m (use 65)
p6.9GFP_NGS_23_LEFT	1	TGCCATTTGTCCGCAATTGTTT	22	40.91	60.67
p6.9GFP_NGS_23_RIGHT	1	AGGCGAGACTTGAACTCACAAC	22	50	60.98
p6.9GFP_NGS_24_LEFT	2	GTAACGGCCAATTCAACGTGAC	22	50	60.84
p6.9GFP_NGS_24_RIGHT	2	GGTCAACGACGCAAACATGATG	22	50	61.15
p6.9GFP_NGS_25_LEFT	1	GGGAGAGTGCCGTTTTTCAAGA	22	50	61.25
p6.9GFP_NGS_25_RIGHT	1	CCGAATTTTTTGAACGACGACGG	22	50	61.14
p6.9GFP_NGS_26_LEFT	2	TCGGGTCCCTATACGAAGCGTTA	22	50	60.93
p6.9GFP_NGS_26_RIGHT	2	ATATAGTGTTCAGCGCTACCG	22	50	61.04
p6.9GFP_NGS_27_LEFT	1	TTCGTTGTGCATTTCAAAGCTTTTG	25	36	60.82
p6.9GFP_NGS_27_RIGHT	1	TGGTAAAAAGCGAATGGTCCGT	22	45.45	60.99
p6.9GFP_NGS_28_LEFT	2	CGTTAAAATGCTAAGCCGCGAG	22	50	61.02
p6.9GFP_NGS_28_RIGHT	2	CAACTGAACCCGTCTGCTGATT	22	50	61.05
p6.9GFP_NGS_29_LEFT	1	TTCCGTTGTCCGACGCTATAAC	22	50	60.91
p6.9GFP_NGS_29_RIGHT	1	TTGGGGTCAACATCGATAGTGTC	23	47.83	60.62
p6.9GFP_NGS_30_LEFT	2	CAAACAAATCTGGCGACTGTGG	22	50	60.78

Table B.3 continued from previous page

Primer Name	Pool	Sequence (5'-3')	Size (bp)	%GC	Tm (use 65)
p6.9GFP_NGS_30_RIGHT	2	CGGTGTACAGATACTTGTGCGT	22	50	60.85

Table B.4: Custom primers with unique i7 and i5 index pairs for each sample, used to construct MiSeq DNA libraries.

Sample	i7 Primer ID	i7 Index (5'-3') [†]	i5 Primer ID	i5 Index (5'-3') [‡]
p6.9GFP_sgRNA_gp64+131 rBEV	N724	CGCTCAGT	S515	TTCTAGCT
p6.9GFP rBEV infected-only control	N721	CCTCTCTG	S517	AGAGTAGA
p6.9GFP rBEV gp64+131 plasmid	N714	TCATGAGC	S515	TTCTAGCT
p6.9GFP rBEV scrambled control	N707	GTAGAGGA	S517	AGAGTAGA

[†] All i7 indexes have the sequence CAAGCAGAAGACGGCATACGAGAT appended to their 5' end and GTCTCGTGGGCTCGG to their 3' end;

[‡] All i5 indexes have the sequence AATGATACGGCGACCACCGAGATCTACAC appended to their 5' end and TCGTCCGACGCGTC to their 3' end.

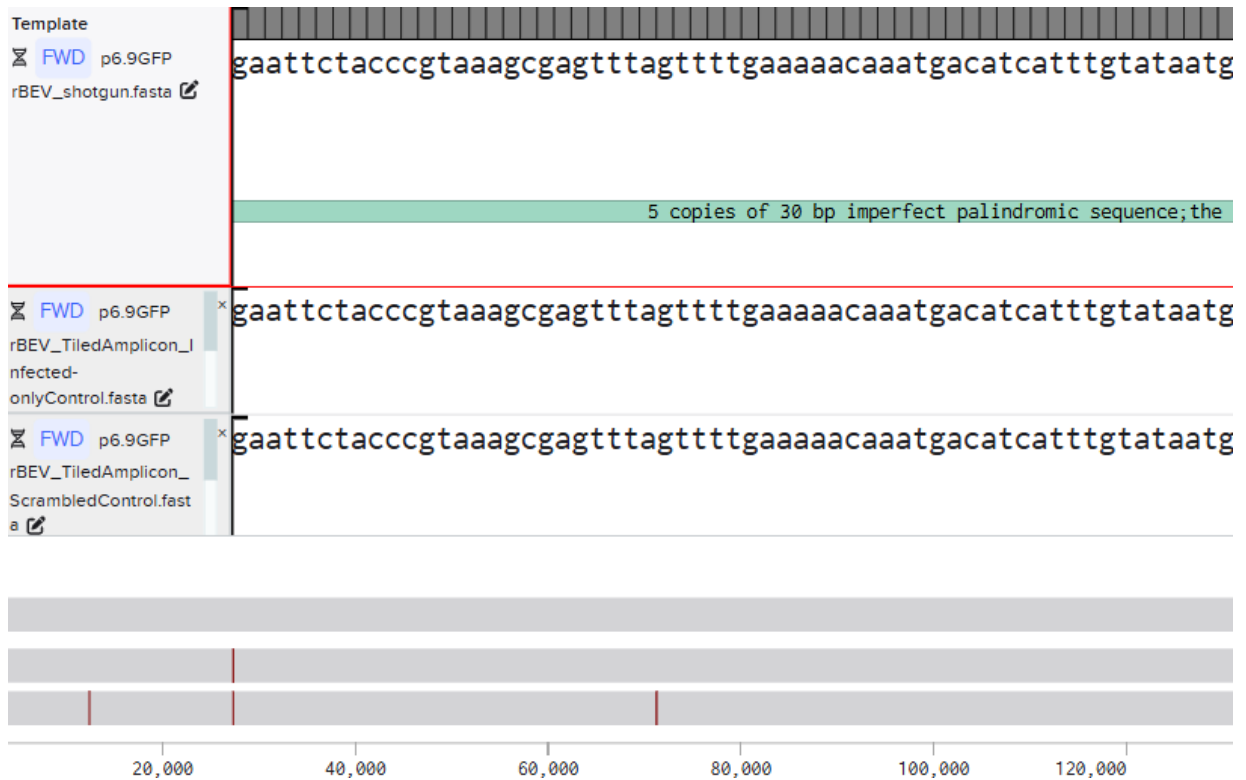


Figure B.1: Visual representation of sequence alignments (Benchling). The first panel is the shotgun-sequenced reference genome (p6.9GFP rBEV_shotgun.fasta), the second panel is the p6.9GFP rBEV genome from the infected-only control (p6.9GFP rBEV_TiledAmplicon_Infected-onlyControl.fasta), and the third panel is the p6.9GFP rBEV genome from the scrambled control (p6.9GFP rBEV_TiledAmplicon_ScrambledControl.fasta). The red vertical lines in the grey bars represent the mutations compared to the reference genome.

Table B.5: Mutations within the *gp64* gene from tiled-amplicon sequencing of the rBEVs p6.9GFP_sgRNA_gp64+131 and p6.9GFP upon *gp64* gene disruption. Table 4.6 has been extended to include position on the p6.9GFP reference genome.

Read	Mutation Region	Mutation Type	Mutation Length (bp)	Reference/Targeted <i>gp64</i> Sequence	Position on Targeted gDNA	Position on p6.9GFP Reference
Read 1	<i>gp64</i>	Deletion	12	gtccttttcag/-	108,669	107,220
Read 2	<i>gp64</i>	Deletion	2	CC/-	108,671	107,222
Read 3	<i>gp64</i>	Deletion	1	C/-	108,671	107,222
Read 4	<i>gp64</i>	Deletion	2	GT/-	108,669	107,220

Appendix C

Chapter 5 supplementary

Table C.1: Custom primers (unique indexes with sequences appended to their 5' and 3' ends) used in this study to construct MiSeq DNA libraries. These unique i7 and i5 index pairs for each sample are read by the machine to distinguish reads from different samples.

Target	Replicate	i7 Primer ID	i7 Index (5'-3') [†]	i5 Primer ID	i5 Index (5'-3') [‡]
gp64+131_C	A	N721	GCAGCGTA	S515	TTCTAGCT
gp64+131_C	B	N722	CTGCGCAT	S515	TTCTAGCT
gp64+131_S	A	N721	GCAGCGTA	S516	CCTAGAGT
gp64+131_S	B	N722	CTGCGCAT	S516	CCTAGAGT
gp64-160_C	A	N723	GAGCGCTA	S515	TTCTAGCT
gp64-160_C	B	N724	CGCTCAGT	S515	TTCTAGCT
gp64-160_S	A	N723	GAGCGCTA	S516	CCTAGAGT
gp64-160_S	B	N724	CGCTCAGT	S516	CCTAGAGT
gp64+278_C	A	N714	TCATGAGC	S515	TTCTAGCT
gp64+278_C	B	N715	CCTGAGAT	S515	TTCTAGCT
gp64+278_S	A	N714	TCATGAGC	S516	CCTAGAGT
gp64+278_S	B	N715	CCTGAGAT	S516	CCTAGAGT

Table C.1 continued from previous page

Target	Replicate	i7 Primer ID	i7 Index (5'-3') [†]	i5 Primer ID	i5 Index (5'-3') [‡]
gp64+378_C	A	N716	TAGCGAGT	S515	TTCTAGCT
gp64+378_C	B	N718	GTAGCTCC	S515	TTCTAGCT
gp64+378_S	A	N716	TAGCGAGT	S516	CCTAGAGT
gp64+378_S	B	N718	GTAGCTCC	S516	CCTAGAGT
gp64+418_C	A	N719	TACTACGC	S515	TTCTAGCT
gp64+418_C	B	N720	AGGCTCCG	S515	TTCTAGCT
gp64+418_S	A	N719	TACTACGC	S516	CCTAGAGT
gp64+418_S	B	N720	AGGCTCCG	S516	CCTAGAGT
gp64+131/384_C	A	N727	ACTGATCG	S515	TTCTAGCT
gp64+131/384_C	B	N729	GACGTCGA	S515	TTCTAGCT
gp64+131/384_S	A	N727	ACTGATCG	S516	CCTAGAGT
gp64+131/384_S	B	N729	GACGTCGA	S516	CCTAGAGT
ScrambledControl_C	A	N716	TAGCGAGT	S517	GCGTAAGA

Table C.1 continued from previous page

Target	Replicate	i7 Primer ID	i7 Index (5'-3') [†]	i5 Primer ID	i5 Index (5'-3') [‡]
ScrambledControl_C	B	N718	GTAGCTCC	S517	GCGTAAGA
ScrambledControl_S	A	N714	TCATGAGC	S517	GCGTAAGA
ScrambledControl_S	B	N715	CCTGAGAT	S517	GCGTAAGA
Infected-onlyControl_C	A	N721	GCAGCGTA	S517	GCGTAAGA
Infected-onlyControl_C	B	N722	CTGCGCAT	S517	GCGTAAGA
Infected-onlyControl_S	A	N719	TACTACGC	S517	GCGTAAGA
Infected-onlyControl_S	B	N720	AGGCTCCG	S517	GCGTAAGA

[†] All i7 indexes have the sequence CAAGCAGAAGACGGCATAACGAGAT appended to their 5' end and GTCTCGTGGGCTCGG to their 3' end;

[‡] All i5 indexes have the sequence AATGATACGGCGACCACCGAGATCTACAC appended to their 5' end and TCGTCCGCAGCGTC to their 3' end.

Variant calling pipeline with a reference genome and two technical replicates:

```
bwa index reference.fasta
```

```
for sample in $(cat samples)
```

```
do
```

```
echo "On Sample: $sample"
```

```
bwa mem -t 20 reference.fasta $sample_a_R1.fq.gz $sample_a_R2.fq.gz >
```

```
$sample_A.sam
```

```
samtools view -b -S -q 10 -F 2308 -o $sample_A-RAW.bam $sample_A.sam
```

```
./pkgs/lofreq_star-2.1.2/bin/lofreq viterbi -f reference.fasta
```

```
-o $sample_A_realigned.bam $sample_A-RAW.bam
```

```
samtools sort -o $sample_A.sort.bam $sample_A_realigned.bam &&
```

```
samtools index $sample_A.sort.bam
```

```
samtools ampliconclip --both-ends --filter-len 200 -b scheme.primer.bed
```

```
$sample_A.sort.bam -o $sample_A.trim.bam
```

```
rm $sample_A.sam $sample_A-RAW.bam $sample_A_realigned.bam
```

```
$sample_A.sort.bam $sample_A.sort.bam.bai
```

```
samtools mpileup -aa -A -d 0 -B -Q 0 -f reference.fasta $sample_A.trim.bam —
```

```
ivar variants -p $sample_A_variants -q 20 -t 0.01 -m 0 -r reference.fasta
```

```
bwa mem -t 20 reference.fasta $sample_b_R1.fq.gz $sample_b_R2.fq.gz >
```

```
$sample_B.sam
```

```
samtools view -b -S -q 10 -F 2308 -o $sample_B-RAW.bam $sample_B.sam
```

```
./pkgs/lofreq_star-2.1.2/bin/lofreq viterbi -f reference.fasta
```

```
-o $sample_B_realigned.bam $sample_B-RAW.bam
```

```
samtools sort -o $sample_B.sort.bam $sample_B_realigned.bam &&
```

```
samtools index $sample_B.sort.bam
```

```
samtools ampliconclip --both-ends --filter-len 200 -b scheme.primer.bed
```

```
$sample_B.sort.bam -o $sample_B.trim.bam
```

```
rm $sample_B.sam $sample_B-RAW.bam $sample_B_realigned.bam
```

```
$sample_B.sort.bam $sample_B.sort.bam.bai
```

```
samtools mpileup -aa -A -d 0 -B -Q 0 -f reference.fasta $sample_B.trim.bam —
```

```
ivar variants -p $sample_B_variants -q 20 -t 0.01 -m 0 -r reference.fasta
```

```
ivar filtervariants -p $sample_variants $sample_A_variants.tsv $sample_B_variants.tsv
```

```
done
```

Table C.2: Description of the column headers appearing in each .tsv output file obtained from running the variant calling pipeline with a reference genome and two technical replicates. Here, 1_A_variants.tsv and 1_B_variants.tsv represent the .tsv file names for sample 1 with replicates A and B.

Column header	Description
REGION	Common region across all replicate variant .tsv output files
POS	Common position across all variant .tsv output files
REF	Common reference nucleotide across all variant .tsv output files
ALT	Common alternate nucleotide across all variant .tsv output files
GFF_FEATURE	A feature used for translation
REF_CODON	Codon using the reference nucleotide
REF_AA	Translated amino acid from reference codon
ALT_CODON	Codon using the alternate nucleotide
ALT_AA	Translated amino acid from alternate codon
REF_DP_1_A_variants.tsv	Reference nucleotide depth in replicate A
REF_RV_1_A_variants.tsv	Reverse reads reference nucleotide depth in replicate A
REF_QUAL_1_A_variants.tsv	Average quality of reference nucleotide in replicate A

Table C.2 continued from previous page

Column header	Description
ALT_DP_1_A_variants.tsv	Alternate nucleotide depth in replicate A
ALT_RV_1_A_variants.tsv	Reverse reads alternate nucleotide depth in replicate A
ALT_QUAL_1_A_variants.tsv	Average quality of alternate nucleotide in replicate A
ALT_FREQ_1_A_variants.tsv	Frequency of alternate nucleotide in replicate A
TOTAL_DP_1_A_variants.tsv	Total depth at the position in replicate A
PVAL_1_A_variants.tsv	p -value of fisher's exact test in replicate A
PASS_1_A_variants.tsv	p -value (≤ 0.05) result in replicate A
REF_DP_1_B_variants.tsv	Reference nucleotide depth in replicate B
REF_RV_1_B_variants.tsv	Reverse reads reference nucleotide depth in replicate B
REF_QUAL_1_B_variants.tsv	Average quality of reference nucleotide in replicate B
ALT_DP_1_B_variants.tsv	Alternate nucleotide depth in replicate B
ALT_RV_1_B_variants.tsv	Reverse reads alternate nucleotide depth in replicate B
ALT_QUAL_1_B_variants.tsv	Average quality of alternate nucleotide in replicate B

Table C.2 continued from previous page

Column header	Description
ALT_FREQ_1_B_variants.tsv	Frequency of alternate nucleotide in replicate B
TOTAL_DP_1_B_variants.tsv	Total depth at the position in replicate B
PVAL_1_B_variants.tsv	<i>p</i> -value of fisher's exact test in replicate B
PASS_1_B_variants.tsv	<i>p</i> -value (≤ 0.05) result in replicate B