DEVELOPMENT OF A NOVEL SARS-COV-2 VIRAL VECTOR

by

Huan Liu

(Under the Direction of Dexi Liu)

ABSTRACT

At the end of 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in the city of Wuhan, which subsequently caused an unprecedented pandemic around the world. As of March 2023, the coronavirus disease 2019 (COVID-19) pandemic has led to more than 670 million people infected and over 6.8 million deaths globally. The COVID-19 pandemic severity has highlighted the urgent need of understanding SARS-CoV-2 virology and identifying novel antiviral mechanisms. The current paradigm of SARS-CoV-2 research mostly employs biosafety level 2-compatible systems such as SARS-CoV-2 spike-pseudoviruses or SARS-CoV-2 virus-like particles, which fail to resemble the complete viral structure or deliver exogeneous genes upon infection. This dissertation project is designed to develop a novel SARS-CoV-2 viral vector composed of all four SARS-CoV-2 structural proteins, the packaging signal sequence of SARS-CoV-2, a reporter gene, and an RNA amplification component of Venezuelan equine encephalitis virus (VEEV). This VEE-SARS-CoV-2 viral vector transduces target cells in an ACE2-dependent manner, and the incorporation of the VEEV self-amplification mechanism generates a significantly higher and extended gene expression profile in transduced cells. This dissertation consists of four chapters. **Chapter 1** is a literature review of the characteristics of SARS-CoV-2 and COVID-19 and a summary of current model systems of SARS-CoV-2 research. **Chapter 2** describes the

details regarding the materials, instruments and methods used to develop the VEE-SARS-CoV-2 vector system. **Chapter 3** shows the data of design, optimization, and characterization of the vector system and demonstrates its applications in neutralizing antibody quantification and antiviral drug testing. **Chapter 4** discusses the advantages, applications, limitations and future perspectives of utilizing this novel vector system for studying SARS-CoV-2 and developing antiviral strategies against COVID-19. The results presented provide direct evidence in support of successful establishment of SARS-CoV-2 viral vectors and clear demonstration of their broad applications in studying virology of SARS-CoV-2, drug and vaccine development against COVID-19, and probing the underlying mechanisms of various symptoms associated with COVID-19.

INDEX WORDS: SARS-CoV-2 viral vectors; Self-amplifying RNA viral vectors; COVID-19 drug discovery; COVID-19 vaccine

DEVELOPMENT OF A NOVEL SARS-COV-2 VIRAL VECTOR

by

HUAN LIU

B.S., College of Pharmacy, East China University of Science & Technology, Shanghai, China, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2023

DEVELOPMENT OF A NOVEL SARS-COV-2 VIRAL VECTOR

by

HUAN LIU

Major Professor: Dexi Liu
Committee: Arthur Roberts
Houjian Cai
Natalia Ivanova

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
May 2023

DEDICATION

I would like to dedicate this dissertation to my parents: my mother, Xinying Yao, and my dad, Depeng Liu. Thanks for their support and all the efforts they made for me. I also would like to dedicate this dissertation to my aunt Zoe, my uncle Charles, my cousin Jerry and his wife Nan, and my cousin Ryan. I appreciate their help while I am pursuing my Ph.D. degree in this country. This great family always encourages me and pushes me forward. Meanwhile, I would like to thank Yueze for her help and company during this journey.

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Dexi Liu, for his mentorship. I learned how to be a scientist from the numerous talks, discussions, and meetings. I always like to share my thoughts with him, and he always provides me the insightful feedback and support. I would also like to thank Guisheng Zhang for her training of molecular cloning, which laid a foundation for me to develop this novel viral vector. I want to thank Dr. Houjian Cai for providing me HEK293T cell line. I thank Dr. Marco Archetti for Vero-E6 cells and his helpful discussions. I am grateful to Dr. Jennifer Doudna and Dr. Abdullah Muhammad Syed for providing me with plasmids containing structural gene and PS9 packaging signal sequence of SARS-CoV-2. I also thank my committee members, Dr. Arthur Roberts and Dr. Natalia Ivanova for their input on my dissertation research.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

LITERATURE REVIEW OF SARS-COV-2 AND COVID-19

1.1 Emergence and spread of COVID-19

Patients with pneumonia symptoms of unknown causes were reported in late December 2019 by local hospitals in Wuhan, Hubei Province, China. On December 31, 2019, Wuhan Municipal Health Commission conducted an epidemiologic and etiologic investigation and identified a novel coronavirus from patient samples (**Fig1.1**) [1]. Subsequently, World Health Organization (WHO) set up an Incident Management Support Team for dealing with this emergent outbreak. The novel coronavirus genome sequence was promptly analyzed by independent Chinese teams by metagenomic RNA sequencing, and the results suggested that it is closely related (~90% genomic similarity) to other coronaviruses found in bats [2,3]. Its genome sequence was publicly shared by several Chinese groups and the sequence information were available on GISAID website on January 11, 2020. Two days later, a traveler from Wuhan to Thailand was confirmed as the first case out of China. Given that those early cases were linked to a seafood market in Wuhan and that multiple familial infections were reported [4, 5], WHO mission to China stated that there was evidence confirming that the novel coronavirus can be transmitted human-to-human on January 21, 2020. Due to the massive domestic travel for lunar new year, the virus spread to all 34 provinces and infected thousands of people within 1 month after the outbreak. To contain the spread of the virulent virus, Chinese government enforced lockdown policies in Wuhan and other major cities in China at the end of January 2020. At the same time, WHO announced

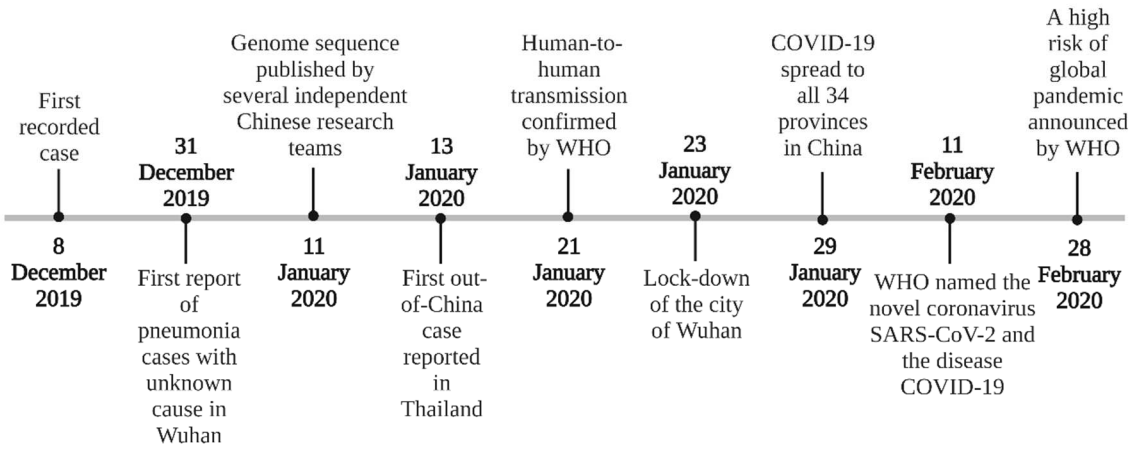


Figure 1.1 Timeline of the COVID-19 outbreak. The first pneumonia case caused by SARS-CoV-2 was recorded on 8 December 2019. On 31 December 2019, Wuhan Municipal Health Commission reported a serials of pneumonia cases with unknown causes in the city of Wuhan, China. The disease-causing virus was isolated and sequenced by a few Chinese research teams. The virus was confirmed to be transmitted human-to-human and spread to all 34 provinces in China within 1 month. The novel coronavirus was named SARS-CoV-2 and the disease was named COVID-19 by WHO. A global pandemic was notified by WHO in the end of February 2020.

that the outbreak constituted a Public Health Emergency of International Concern and gave a risk assessment of very high for China, and high at the global level. The novel coronavirus was named "severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)" by the International Committee on Taxonomy of Viruses, and the disease was named "coronavirus disease 19 (COVID-19)" on February 11-12, 2020 [6].

The spread of SARS-CoV-2 peaked in February 2020 at more than 3000 cases per day nationwide. Fortunately, the rapid spread was effectively controlled by the enforced policies including strict lockdown and social distancing later. Since then, China has conducted a Zero-COVID policy to contain the virus spread, and COVID-19 cases have been maintained at a very low level. Despite the lockdown and social distancing policies enforced similarly, COVID-19 cases spiked globally and caused one of the major catastrophes in human history. The virus quickly caused more than 118,000 cases in 114 countries and 4,291 deaths in total, therefore WHO

declared the outbreak of COVID-19 as a pandemic. As of March 6, 2023, the COVID-19 pandemic has caused more than 670 million cases around the world, and more than 6.8 million patients died [7].

1.2 Characteristics of SARS-CoV-2

1.2.1 Phylogeny of SARS-CoV-2

SARS-CoV-2 belongs to the family *Coronaviridae*, a family of large, enveloped, and mostly positive-stranded RNA viruses. *Coronaviridae* family can be genetically categorized into four genera: *alphacoronavirus*, *betacoronavirus*, *gammacoronavirus*, and *deltacoronavirus* [8]. Among these family members, there are 7 coronaviruses known to infect humans including HCoV-229E, HCoVNL63, HCoV-OC43, HCoV-HKU1, SARS-CoV-1, Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2. With the exception of two alphacoronaviruses (i.e., HCoV-229E and HCoVNL63), other 5 human coronaviruses all belong to the genus of *betacoronavirus*. *Betacoronaviruses* can be further divided into subgenera: HCoV-OC43 and HCoV-HKU1 are embecoviruses [9]; MERS-CoV belongs to merbecoviruses [10]; both SARS-CoV-1 and SARS-CoV-2 are sarbecoviruses [11]. Metagenomic RNA sequencing data suggest that SARS-CoV-2 shares highly similar (~79%) genomic information with SARS-CoV-1 [12, 13]. It is also shown that the genome of a bat coronavirus was found to be extremely similar to that of SARS-CoV-2 (>96%), indicating that the origin of SARS-CoV-2 could be from bat [14].

1.2.2 Genomic structure of SARS-CoV-2

As a member of *Coronaviridae*, SARS-CoV-2 has a positive-stranded RNA genome. Its genome size is around 29.9 kb, including a 265-nucleotide 5' untranslated region (5' UTR), a large

open-reading frame (ORF1ab) spanning ~2/3 of the genome, the 3' 1/3 of the genome encoding 4 structural proteins and other accessory proteins, and a 3' UTR with a poly(A) tail (**Fig 1.2**). The SARS-CoV-2 genome comprises of 15 (12 functional) ORFs, which encodes 31 proteins in total. Its genomic structure shares ~90% similarity with other coronaviruses such as SARS-CoV-1 and MERS [15, 16].

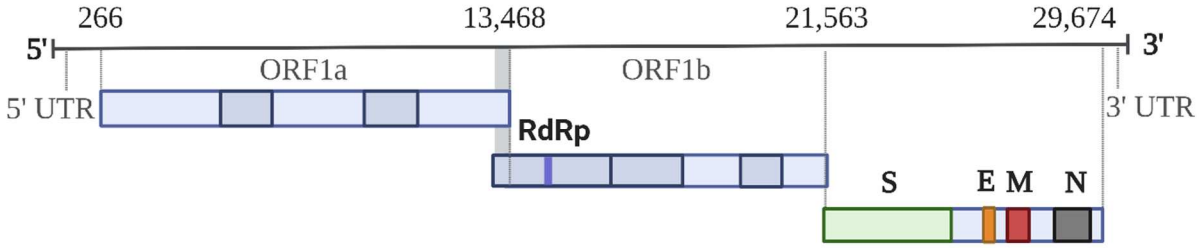


Figure 1.2 SARS-CoV-2 genome organization. SARS-CoV-2 viral genome is a ~29.9 kb positive-stranded RNA genome including a 265-nucleotide 5' untranslated region (5' UTR), a large open-reading frame (ORF1ab) spanning ~2/3 of the genome, the 3' 1/3 of the genome encoding 4 structural proteins and other accessory proteins, and a 3' UTR.

1.2.3 Non-structural proteins of SARS-CoV-2

The majority of SARS-CoV-2 genome encodes the ORF1a and ORF1ab, which ORF1ab is a product of ribosomal frameshift mechanism. The ORF1ab includes 16 non-structural protein (nsp) genes, and their individual functions are listed in **Table 1.1**. Because polymerase (encoded by nsp7, nsp8, and nsp11) and proteinase (encoded by nsp5) are encoded in the frame of ORF1ab, the translated polyprotein 1ab can then self-process and replicate its own genome. Therefore, ORF1ab polyprotein plays critical roles in the genome replication and early transcription initiation [17]. This ribosomal frameshift mechanism and genetic arrangement are also conserved in other coronaviruses [18, 19].

Table 1.1 Summary of 16 SARS-CoV-2 non-structural proteins (NSP)[a]

Protein	Genomic location (bp)	Description	Protein function
NSP1	266–805	The N-terminal of replicase	Mediate RNA replication
NSP2	806–2719	Viral replicase for proofreading	Suppress host immunity
NSP3	2720–8554	Papain-like protease	Polyprotein cleavage
NSP4	8555–10054	Transmembrane protein	Anchor replication-transcription complex
NSP5	10055–10972	Main protease	Polyprotein cleavage
NSP6	10973–11842	Putative transmembrane domain	Induce inflammatory cell death in the host
NSP7	11843–12091	RdRP cofactor	Responsible for replication
NSP8	12092–12685	RdRP cofactor	Responsible for replication
NSP9	12686–13024	RNA-binding protein	Participate in replication as an RNA-binding protein
NSP10	13025–13441	Growth-factor-like protein	Participate in genome methylation
NSP11	13442–13480	Product of translation termination at ORF1a	Unknown
NSP12	13442–16236	RdRP	Responsible for replication
NSP13	16237–18039	RNA 5’ triphosphatase	Participate in genome RNA synthesis
NSP14	18040–19620	Exoribonuclease domain for proofreading	Exoribonuclease activity
NSP15	19621–20658	EndoRNAse	Endonuclease activity
NSP16	20659–21552	Methyltransferase	RNA cap methylation

[a] Modified from reference [9]

1.2.4 Virion composition and structural proteins of SARS-CoV-2

To the 3’ end of the SARS-CoV-2 genome, four structural proteins are encoded by 4 independent ORFs. These proteins are called structural proteins as they are present on the virion structure of SARS-CoV-2 and compose the virus with its positive-strand RNA genome. The SARS-CoV-2 viral particle diameter is 70–110 nm, obtained from the observation under electron

microscopy [20]. The SARS-CoV-2 virion comprises of spike (S) protein, envelop (E) protein, membrane (M) protein, and nucleocapsid (N) protein (**Fig 1.3**).

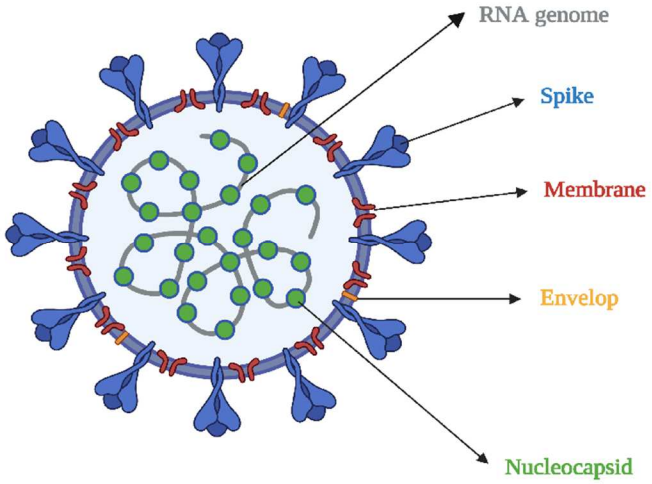


Figure 1.3 SARS-CoV-2 virion structure. SARS-CoV-2 virion is composed of 4 structural proteins including spike, membrane, envelop, and nucleocapsid protein binding to the SARS-CoV-2 RNA genome.

The S protein is a ~200 kDa large protein that consists of a 13-amino acid signal peptide at the N-terminal, the S1 subunit, the S2 subunit, and two other regions responsible for receptor binding and membrane fusion. Notably, a receptor binding domain (RBD) is in the S1 subunit, which directly binds to the angiotensin converting enzyme 2 (ACE2) receptor on the surface of host cells [21]. On the virion surface, spike protein forms trimers and is coated with polysaccharide molecules for immune escape upon entry into the host. Since the spike trimers sticks out of the viral particle and look like a crown under the electron microscopy, this family of viruses are named coronaviruses ("corona" means "crown"). The spike protein is the most studied protein of SARS-CoV-2 due to the function of receptor binding. Like SARS-CoV-1, SARS-CoV-2 enters the host cell via the biding of S protein to ACE2 receptor, which requires an acidic pH [22–26]. Upon

binding to ACE2 receptor, the host protease cleaves the S1 and S2 subunits to enable the membrane fusion of the S2 subunit. Given that the S protein dictates the viral entry of SARS-CoV-2 into host cells, the S protein is the major focus of the current vaccine and drug development paradigm.

The E protein is the smallest SARS-CoV-2 structural protein (~10 kDa) and highly conserved in the coronavirus family [27]. It is a transmembrane protein spanning the virion surface. Previous studies have shown that E protein along with M protein plays an important role of maintaining viral structure, viral assembly, and viral release [28, 29]. Recently, SARS-CoV-2 E protein was found to form viroporin pores on the host cell surface to regulate membrane permeability, membrane remodeling, and ion and protein trafficking, which may ultimately influence the SARS-CoV-2 viral replication [30]. It is reported that deletion of E protein leads to significant decrease of SARS-CoV-2 infection pathogenicity, indicating that this tiny protein could also serve as a target for drug development against SARS-CoV-2 [31].

The M protein is a ~30 kDa protein that is abundantly present on the SARS-CoV-2 virion surface. M protein is also a highly conserved protein in coronavirus family. It is known that SARS-CoV-2 M protein shares 90% similarity to SARS-CoV-1 M protein [32]. Structural biology data suggest that M proteins form dimers by binding to transmembrane domains [33]. By interacting with S, E, and N structural protein, coronavirus M protein exerts its function of maintaining virion membrane curvature and virion shape [34, 35]. In addition, Zheng et al. demonstrated that SARS-CoV-2 M protein suppresses type I and III interferon-mediated antiviral mechanisms by activating RIG-I/MDA-5–MAVS signaling pathway [36]. M protein also enables SARS-CoV-2 to evade host cellular immunity by inhibiting mitochondria antiviral signaling protein (MAVS), an adaptor protein in type I interferon signaling pathway [37]. Taken together, M protein is not only important in maintaining the viral structure of SARS-CoV-2, but also serves as an innate immunity antagonist.

The N protein is a ~48 kDa phosphoprotein, which exclusively locates inside the viral particle of SARS-CoV-2. The N protein has 5 domains: an N-terminal domain (NTD), an RNA-binding domain (RBD), a central linker domain, a dimerization domain, and a C-terminal domain (CTD). The primary function of the N protein is to anchor the viral RNA genome within the SARS-CoV-2 viral particle. On one hand, the SARS-CoV-2 N protein binds to other structural proteins such as M protein to assist the viral structure formation [38]. On the other hand, the N protein specifically binds to the viral genome and forms a helical nucleocapsid structure to be packaged in the viral particle shell [39,40]. The mechanism by which the N protein binds to the SARS-CoV-2 RNA genome represents a key event in the viral infection process and has been intensively studied. Data from independent groups suggest that the SARS-CoV-2 N protein binds to the RNA genome by liquid-liquid phase separation, thereby genomic RNA can be exclusively packaged into the virion [41–43]. Besides genome packaging, the N protein also plays important roles in immune suppression and genomic RNA synthesis. Like other viral proteins, the N protein can also interact with host immunity by altering the phosphorylation and nuclear translocation of IRF3, STAT1, and STAT2 [44]. During the viral RNA synthesis, the N protein has been shown to enhance the replicase-transcriptase complex formed by nsp7, nsp8, and nsp12 and boost viral RNA synthesis in the host [43].

1.2.5 Accessory proteins of SARS-CoV-2

Besides structural proteins and 16 non-structural proteins encoded by ORF1ab, there are 11 accessory proteins: ORF3a, ORF3b, ORF3c, ORF3d, ORF6, ORF7a, ORF7b, ORF8, ORF9b, ORF9c, and ORF10 [45]. These proteins play important roles in viral infection since they are involved in pathogenesis mechanisms and related to immune escape of SARS-CoV-2 [46]. For example, ORF3, ORF6, ORF7a, and ORF8 inhibit cellular immunity by suppressing type I

interferon (IFN), a major intracellular antiviral mechanism [47–49]. In addition, cellular events including autophagy and apoptosis can also be affected by accessory proteins. ORF3a activates NF-kB signaling and induces cell apoptosis in multiple cell lines [50]. ORF9b induces ATG5-mediated autophagy in host cells [51]. ORF9b is also associated with inflammasome activation to evade host immune responses [52]. Although current antiviral drug development paradigm neglects these accessory proteins, they could also be further exploited and serve as novel targets for pharmacotherapy.

1.2.6 SARS-CoV-2 life cycle

The life cycle of SARS-CoV-2 is initiated by the spike protein binding to ACE2 receptor on the host cell surface (**Fig 1.4**). Once receptor binding occurs, host proteases such as furin cleave at the S1 and S2 subunit boundary to expose the S2' site of the S2 subunit for further cleavage. There are two distinct mechanisms by which the virus can enter the host cell: 1) the S2 subunit will be activated by transmembrane protease serine 2 (TMPRSS2) to fuse with the host membrane; 2) when the host does not express sufficient TMPRSS2, the S2 subunit-ACE2 receptor complex is internalized by clathrin-mediated endocytosis [53]. In the host cell, SARS-CoV-2 uncoats its viral particle and releases its RNA genome. The positive-stranded RNA genome can be readily translated into polyprotein 1a and 1ab by ribosomes because they are encoded by the first ORF at the 5' end of the genome. The polyprotein 1ab is produced by a programmed -1 ribosomal frameshift at the overlapping region of ORF1a and ORF1b in which the ribosome will skip the stop codon at the end of ORF1a and continues to translate proteins encoded by ORF1b until encounters the next stop codon located at the 3' end of ORF1b [54]. ORF1ab translation generates 16 non-structural proteins including papain-like protease (encoded by nsp3) and

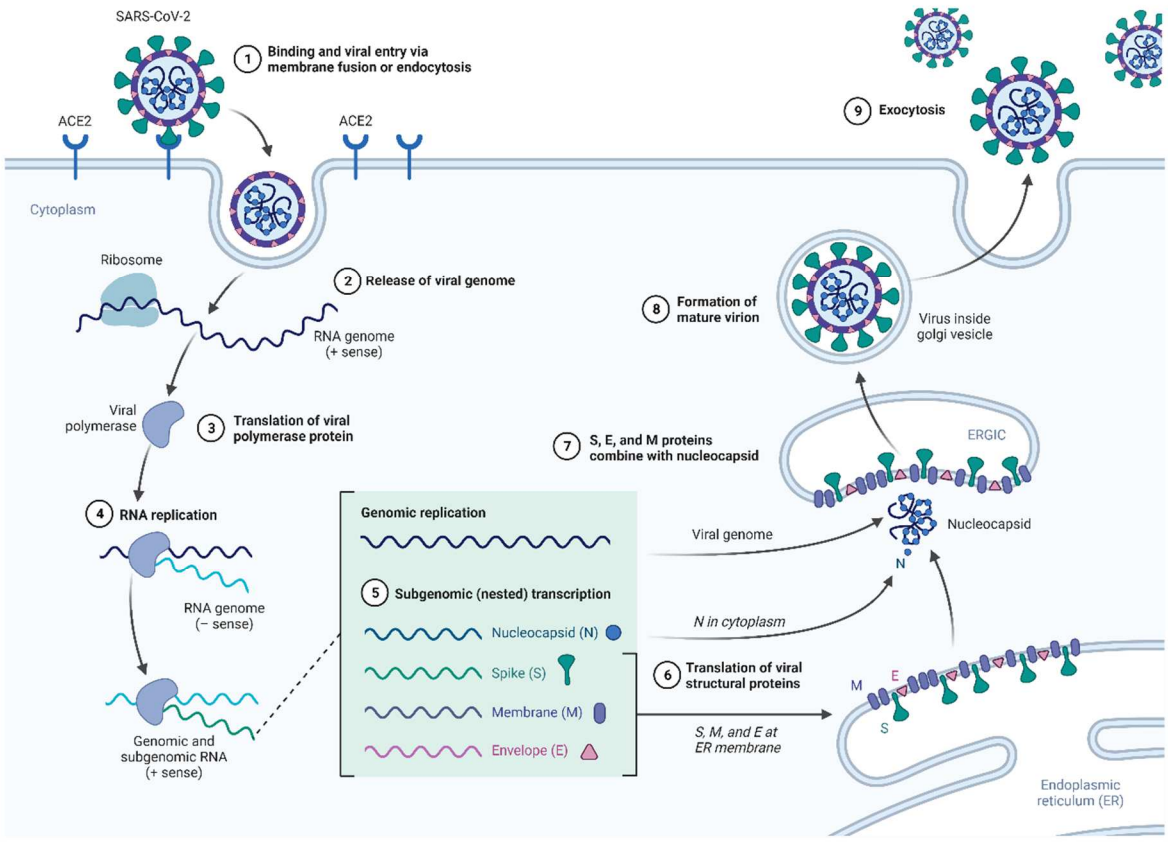


Figure 1.4 SARS-CoV-2 life cycle. SARS-CoV-2 binds to ACE2 receptor on the cell surface and enters the host cell via membrane fusion or endocytosis. Once enters the host, it releases the positive-stranded RNA genome which is translated into polyprotein 1ab by ribosomes. The polyprotein 1ab contains proteases responsible for cleaving the polyprotein into small proteins including RNA-dependent RNA polymerases. The replication-transcription complex forms to synthesize the negative-strand RNA genome using the positive strand as the template. The negative strand of RNA genome then serves as the template for 1) the synthesis of positive-strand RNA genome; 2) the synthesis of positive-stranded subgenomic RNAs encoding structural proteins and accessory proteins. Structural proteins are then translated in the endoplasmic reticulum and translocated into ER-to-Golgi intermediate compartment (ERGIC). In ERGIC, the newly synthesized positive-strand full length viral genome is packaged into viral particles composed of 4 structural proteins. The mature assembled virions are formed in ERGIC vesicles and released out of the host cell via exocytosis.

chymotrypsin-like protease (encoded by nsp5). These proteases can then cleave this bulky polyprotein 1ab and process it into individual non-structural proteins with distinct functions. The replication-transcription complex (RTC) composed of nsp12 RNA-dependent RNA polymerase (RdRP) and other two cofactors (nsp7 and nsp8) is responsible for genomic replication and

transcription. At the same time, the positive-stranded RNA genome serves as the template for synthesis of the negative-strand full length genome, which in turn is replicated into the full length positive-strand genome. The newly synthesized full length positive-stranded genomic RNAs are then packaged into viral particles. To circumvent the limitation that only the first ORF on the 5' end of a RNA molecule is translated, coronaviruses evolved a mechanism called discontinuous transcription to generate a series of nested subgenomic RNAs [55]. Briefly, when the positive-stranded RNA genome serves as the target, RTC synthesizes negative-stranded RNAs starting from the 3' end of the positive strand. The synthesis process could possibly terminate once encounters the transcriptional regulatory sequences (TRS) upstream to most ORFs in the 3' 1/3 of the genome. In that case, the RNA synthesis immediately stops and RTC jumps into the TRS to restart the RNA synthesis at 5' UTR of the viral genome, thereby a shorter subgenomic negative-stranded RNA will be produced. And this negative-stranded subgenomic RNA is then converted into the positive strand for protein translation. If the skipping does not occur at this TRS, the RTC will continue until the next TRS [56]. The discontinuously transcribed subgenomic RNAs are translated into structural proteins. The translated proteins translocate to the ER-to-Golgi intermediate compartment (ERGIC) and form viral particles in which N protein binds to the newly synthesized positive-stranded RNA genome. Finally, the mature viruses are released out of the host via exocytosis.

1.3 Pathology of COVID-19

1.3.1 Pulmonary pathology of COVID-19

Because SARS-CoV-2 infects host cells by binding to ACE2 receptor which is abundantly expressed on the surface of alveolar epithelial type 2 cells (the major cell type constituting the lung

[57]) and nasal epithelial cells [58], COVID-19 primarily involves the respiratory system including the upper and lower respiratory tract. Pneumonia is one of the major severe symptoms induced by COVID-19. The COVID-induced pneumonia is initiated by the viral replication-mediated tissue damage and followed by the infection-induced immune responses involving peripheral immune cells such as T lymphocytes, monocytes, and neutrophils. These immune cells are recruited to release the inflammatory cytokines (e.g., tumor necrosis factor-α (TNFα), interleukin-1 (IL-1), and interleukin-6 (IL-6)), which is manifested by an event called "cytokine storm" resulting in systemic inflammation [59, 60]. In more severe cases, the infection-induced tissue damage causes pulmonary edema by increasing the vascular permeability, and ultimately the patient could have difficulty breathing [61].

1.3.2 Extrapulmonary pathology of COVID-19

Since ACE2+ cells are not only located in the respiratory system, SARS-CoV-2 can also infect and impair other tissues or organs such as cardiovascular system, central nervous system (CNS), gastrointestinal (GI) tract etc. In the cardiovascular system, cardiac myocytes and endothelial cells express high level of ACE2 protein. It is known that SARS-CoV-2 infection may cause myocarditis due to its cytotoxicity in cardiac myocytes [62]. ACE2+ endothelial cells can also be infected by SARS-CoV-2, evidenced by the accumulation of viral components in endothelial cells [63]. The viral infection induces apoptosis of endothelial cells and inflammatory cells, which ultimately leads to endotheliitis. Additionally, SARS-CoV-2 is neuroinvasive. Severe COVID-19 patients experience headache, dizziness, anorexia [64]. However, the exact mechanism of neurotoxicity of COVID-19 remains to be understood. It is speculated that the infection-induced inflammation and cytokine storm could alter the integrity of the blood-brain barrier. In the GI

tract, SARS-CoV-2 infects intestinal glandular cells, rectal epithelial cells, and glandular enterocytes. It is reported that viral RNA can be detected in patients' stool in over 50% of COVID-19 cases [65].

1.4 COVID-19 vaccines

Inducing adaptive immunity against SARS-CoV-2 by vaccination is considered as the most important approach to protecting people from infection and containing COVID-19 spread. Thanks to timely genome sequence sharing, quick adaption of state-of-the-art technologies, and sufficient financial and policy support, various COVID-19 vaccines have been developed at an unprecedented speed. As of January 2023, there are 11 vaccines granted for global emergency use against COVID-19 by WHO. These 11 vaccines can be categorized into 4 major classes by their mechanism of action: protein subunit vaccine, mRNA vaccine, non-replicating viral vector vaccine, and inactivated vaccine (**Table 1.2**) [66].

1.4.1 Protein subunit vaccines

Protein subunit vaccine comprises of a viral protein or a part of a viral protein as the antigen to induce antiviral immune responses. Covovax and Nuvaxoid developed by Novavax are nanoparticle vaccines employing SARS-CoV-2 spike protein trimer with matrix adjuvant [67]. A phase 3 clinical trial demonstrated an 89.7% vaccine efficacy of Covovax in the UK [68]. This protein vaccine product received the emergency use authorization from FDA for COVID-19 prevention in July 2022.

Table 1.2 Global COVID-19 vaccines granted emergency use by WHO

Vaccine platform	Vaccine	Developer	Clinical trial efficacy	Route of administration	Approval date and country□
Protein subunit	Nuvaxovid	Novavax	89.7%	Intramuscular	July 2022 in the US
	Covovax	Serum Institute of India	95%	Intramuscular	December 2021 in India
RNA	Spikevax	Moderna	94.1%	Intramuscular	December 2020 in the US
	Comirnaty	Pfizer/BioNTech	95%	Intramuscular	December 2020 in the US
Non replicating viral vector	Convidecia	CanSino	57.5%	Intranasal	September 2022 in China
	Jcovden	Johnson& Johnson	66.9%	Intramuscular	February 2021 in the US
	Vaxzevria	AstraZeneca	74%	Intramuscular	January 2021 in Europe
	Covishield	Serum Institute of India	66.7%	Intramuscular	January 2021 in India
Inactivated virus	Covaxin	Bharat Biotech	77.8%	Intramuscular	January 2021 in India
	Covilo	Sinopharm	79%	Intramuscular	December 2020 in China
	CoronaVac	Sinovac	83.5%	Intramuscular	February 2021 in China

□ Approval is defined by the authorization for EUA.

1.4.2 mRNA vaccine

mRNA vaccine utilizes lipid nanoparticle (LNP)-encapsulated, nucleoside-modified mRNA encoding the antigen gene to induce the desired immunity. In the case of COVID-19

mRNA vaccine, mRNA is designed to express spike protein, the structural protein responsible for ACE2 receptor binding and viral entry into the host. Given that they are developed at an unprecedented speed and highly effective in prevention, COVID-19 mRNA vaccine is one of the great victories in the history of drug development. As the first approved COVID-19 vaccine in the US, BioNTech and Pfizer developed BNT162b2 which showed excellent safety and efficacy profile in clinical trials. It is reported that a two-dose regimen of BNT162b2 demonstrated 95% protection rate among 43448 subjects [69]. Likewise, Moderna's mRNA-1273 vaccine also showed 94.1% protection efficacy in a phase 3 randomized clinical trial [70, 71]. These two mRNA vaccines are widely used for COVID-19 vaccination in the US and cover more than 70% of population in the country [72].

1.4.3 Viral vector vaccine

Non-replicating viral vector induces immunity by delivering the antigen gene without further replication of vector genome. Currently approved non-replicating viral vector vaccines utilize non-replicating adenoviral vectors to deliver spike protein gene. Chinese biopharmaceutical company CanSino Biologics developed Convidecia employing a single dose of adenovirus type 5 as the vector to express SARS-CoV-2 spike protein gene. Their data showed a 57.5% of protection rate in over 20000 participants [73]. Johnson & Johnson's Jcovden uses a single dose of adenovirus type 26 as the vector to deliver full-length perfusion-stabilized spike protein gene to elicit immunity. The global phase 3 clinical trial demonstrated a 66.9% of protection rate against COVID-19 [74]. AstraZeneca and University of Oxford developed a chimpanzee adenovirus-based vaccine platform to express spike protein gene. The overall efficacy of this vaccine was 74% and its safety profile was also proven in the clinical trial [75]. Although viral vector-based vaccines

failed to achieve the equal prevention efficacy as the mRNA vaccine, they are still important constituents of COVID-19 vaccine landscape.

1.4.4 Inactivated vaccine

Inactivated vaccines are live viruses killed by different means such as chemicals, radiation, or heat. Since they are killed viruses, they lose the replication competence but can still elicit the antibody production. CoronaVac, a whole-virion SARS-CoV-2 inactivated by β-propiolactone yielded 83.5% efficacy in over 10000 volunteers [76]. Similarly, Sinopharm developed Covilo and first tested its efficacy in rhesus macaques [77]. Its safety and immunogenicity were further proven by a phase 1/2 clinical trial [78]. A recent clinical trial showed that Covilo achieved a 79% efficacy in prevention of COVID-19, which granted the approval by Chinese FDA [79]. Covaxin, developed by Indian company Bharat Biotech, showed an overall efficacy of 77.8% in 25798 participants and became a major COVID-19 vaccine in India [80].

1.5 COVID-19 therapeutics

More than 2 years into the COVID-19 pandemic, several classes of therapeutic agents (e.g., anti-inflammatory agents, antivirals, and anti-SARS-CoV-2 neutralizing antibodies) have been developed to combat this new coronavirus (**Table 1.3**). The administration of therapeutics selection is based on the clinical settings and drug mechanism of action [81].

1.5.1 Anti-inflammatory agents

Tocilizumab is a humanized IL-6 receptor monoclonal antibody. Because IL-6 is a major pro-inflammatory cytokine, inhibition of the receptor can alleviate infection-induced inflammation

Table 1.3 COVID-19 therapeutics approved or authorized for EUA[a]

Drug category	Drug	Setting	Patient condition	Dose & Route of administration	Date of approval or EUA authorization in the US
Anti-inflammatory agents	Tocilizumab	Inpatient	Severe	8mg/kg, i.v.	June 2021, EUA
	Baricitinib	Inpatient	Severe	4 mg, oral	November 2020, EUA May 2022, approval
Antiviral agents	Remdesivir	Inpatient/ outpatient	Mild/moderate	200 mg, i.v.	May 2020, EUA October 2020, approval
	Paxlovid	Outpatient	Mild/moderate	300/100 mg, oral	December 2021, EUA
	Molnupiravir	Outpatient	Mild/moderate	800 mg oral	December 2021, EUA
Antibody-based therapeutics	Convalescent plasma	Inpatient and outpatient	Severe	~200 ml, i.v.	August 2020, EUA
	Bamlanivimab/ etesevimaba	Outpatient	Mild/moderate	700/1400 mg, i.v.	February 2021, EUA
	Casirivimab/ imdevimab	Outpatient	Mild/moderate	600 mg, s.c.	November 2020, EUA
	Sotrovimab	Outpatient	Mild/moderate	500 mg, i.v.	May 2021, EUA
	Bebtelovimab	Outpatient	Mild/moderate	175 mg, i.v.	February 2022, EUA
	Tixagevimab/ cilgavimab (Evusheld)	Outpatient	Moderate to severe	300 mg, i.m.	December 2021, EUA

i.v., intravenous; s.c., subcutaneous; i.m., intramuscular; EUA, emergency use authorization.

[a] Modified from reference [75]

or cytokine storm. Clinical data obtained from a Randomized, Embedded, Multifactorial Adaptive Platform Trial for Community-Acquired Pneumonia (REMAP-CAP) showed that tocilizumab reduced patient mortality (28% vs 36%) [82]. A larger randomized clinical trial performed in the

UK showed the similar results. The mortality rate of the group receiving tocilizumab was 31%, whereas the placebo group mortality rate was 35% [83]. Considering the promising data collected, tocilizumab has been approved for the emergency use in the US and Europe.

Baricitinib is a small molecule inhibitor targeting Janus kinase (JAK). JAK is also a critical regulator of inflammation, therefore JAK inhibitor is believed to benefit COVID-19 patients. A randomized clinical trial shows that baricitinib reduced the mortality rate from 13% to 8% [84]. Another double-blinded, randomized trial recruited 1033 patients receiving either remdesivir or a combination of remdesivir and baricitinib for COVID-19 treatment. The results showed that the combinatory administration reduced the mortality rate and severe adverse effect simultaneously [85]. In May 2022, FDA approved baricitinib for the treatment of certain hospitalized patients with COVID-19.

1.5.2 Antiviral therapies

Originally designed for Ebola virus, remdesivir is a nucleoside analogue prodrug inhibiting viral RNA replication by targeting RdRP [86]. Later, remdesivir was further proven to be effective against coronaviruses such as SARS-CoV-1 and MERS *in vitro* [87]. At the early stage of the COVID-19 pandemic, remdesivir was considered as one the most promising repurposed antiviral drugs. Several independent clinical trials were conducted to evaluate its efficacy against SARS-CoV-2. A PINETREE study reported that a 3-day administration of remdesivir was safe and effective to reduce hospitalization rate by 87% to prevent severe disease progression [88]. Likewise, another double-blinded, randomized, placebo-controlled trial revealed that intravenous administration of remdesivir shortened a median recovery time from 15 days to 10 days and significantly decreased mortality rate (6.7% vs 11.9% at day 15; 11.4% vs 15.2% at day 29) [89].

The largest randomized clinical trial of remdesivir led by WHO recruited 8275 COVID-19 patients and showed that remdesivir slightly reduced hospitalization mortality (11.9% vs 13.5%) in the non-ventilated patients, however no clinical benefits were found in the ventilated patients [90]. On October 22, 2020, remdesivir became the first drug receiving FDA approval for the use against COVID-19. It is also approved by European Medicines Agency on July 3, 2020.

Paxlovid is an antiviral drug product combining nirmatrelvir and ritonavir. Nirmatrelvir is an oral inhibitor of the SARS-CoV-2 3C-like protease, which is responsible for cleavage of polyprotein 1ab [91]. Its antiviral activity was first examined in *in vitro* models and mouse-adapted SARS-CoV-2 (SARS-CoV-2 MA10) infection model. Since cytochrome P450 3A4 (CYP3A4) is the major enzyme for nirmatrelvir metabolism, a low dose (100 mg) of ritonavir is co-administrated with nirmatrelvir as a CYP3A4 inhibitor for pharmacokinetic improvement [92]. A 2246 patient clinical trial showed that nirmatrelvir combined with ritonavir reduced the rate of hospitalization and mortality in unvaccinated and non-hospitalized patients by 89% than the placebo control group [93]. Paxlovid received the EUA for COVID-19 from FDA on December 22, 2021, and the conditional marketing authorization in Europe on January 28, 2022.

Molnupiravir is an oral broad spectrum antiviral compound. As a cytidine analogue, it can disrupt viral RNA synthesis by inhibiting RdRP of SARS-CoV-2 in the host cells. Molnupiravir showed antiviral activity against SARS-CoV-1, MERS, and SARS-CoV-2 in human lung epithelial Calu-3 cells, primary human airway epithelial cell cultures, and transgenic murine models [94]. The following clinical trial recruited 1433 non-hospitalized patients and showed that molnupiravir significantly decreased the hospitalization rate and death risk [95]. On December 23, 2021, molnupiravir received the FDA approval for emergency use on COVID-19.

1.5.3 Anti-SARS-CoV-2 neutralizing antibodies

Convalescent plasma was approved by FDA for emergency use against COVID-19 at the very early stage of the COVID-19 pandemic outbreak (August 2020). However, several independent randomized clinical trials revealed that convalescent plasma failed to significantly reduce the mortality rate in hospitalized patients [96–98]. Moreover, the emergence of several new variants of SARS-CoV-2 further dampened the efficacy of convalescent plasma against COVID-19. Therefore, the convalescent plasma collected before the outbreak of omicron variants is suspended for treatment due to the low efficacy.

Recombinant monoclonal antibodies targeting the receptor binding domain of SARS-CoV-2 spike protein are developed to block viral entry into ACE2+ cells. So far, FDA has authorized 5 neutralizing antibodies under the emergency use authorization including bamlanivimab plus etesevimab, casirivimab plus imdevimab, sotrovimab, bebtelovimab, and tixagevimab plus cilgavimab. A phase 3 trial showed that the COVID-19-related hospitalization rate and mortality rate was 2.1% in the group received bamlanivimab plus etesevimab infusion, compared to 7.0% of the placebo group [99]. In a randomized, controlled, open-label, platform trial, casirivimab plus imdevimab demonstrated the effectiveness of reducing hospitalized patient mortality rate (24% vs 30%) [100]. Among these recombinant antibodies, bebtelovimab is the only one recommended for non-hospitalized patients. Evusheld, a combination of tixagevimab plus cilgavimab, was authorized for pre-exposure and prevention in vulnerable people on December 1, 2022. A clinical trial showed that an intramuscular injection of tixagevimab and cilgavimab was effective to prevent COVID-19 (0.2% vs 1%) and safe to use [101].

1.6 Model systems developed to study SARS-CoV-2

Establishing model systems for infection is imperative for both understanding basic virology of SARS-CoV-2 and developing antiviral agents such as neutralizing antibodies, antiviral drugs, and vaccines. Model systems include cell lines susceptible for SARS-CoV-2 infection, animal models prone to be infected by SARS-CoV-2, and other systems representing SARS-CoV-2 with the attenuated virulence. Their development enables researchers to study this novel virus in a safer and more efficient manner, since the authentic virus raises safety concern and requires cautious handling.

1.6.1 Cell line models for SARS-CoV-2

Mammalian cell lines are the most used model systems to study SARS-CoV-2 and evaluate antiviral agent activity because of the ease of use and scalable applications. For SARS-CoV-2 study, appropriate target cell lines express ACE2, the cellular receptor of SARS-CoV-2 infection. In addition, other proteases expressed on the host cell surface including TMPRSS2 or furin are also ideal for the enhanced viral entry of SARS-CoV-2. Another important characteristic of the ideal target cell line is the lack or decreased activity of cellular antiviral immunity such as interferon signaling mechanisms, because the functional cellular immunity may inhibit or shut down the replication of SARS-CoV-2 in the host to decease the infection. In this case, the failure of the infection will not cause cytopathic effects, the major characteristic of viral replication [102]. Thus, a SARS-CoV-2-specific cellular tropism and permissibility are both important to a proper cell line model for studying SARS-CoV-2 [103]. Commonly used cell lines for SARS-CoV-2 research are summarized in **Table 1.4**.

Table 1.4 Cell line models for SARS-CoV-2 studies

Cell line	Species	Tissue source	ACE2 level	TMPRSS2 level	CPE
Vero E6	African green monkey	Kidney	High	Negative	Present
Calu-3	Human	Lung cancer	High	Low	Absent
Caco-2	Human	Colon cancer	Moderate	Moderate	Absent
HuH-7	Human	Liver cancer	Negative	High	Absent
ACE2-TMPRSS2-293T	Human	Kidney	High	High	Present
ACE2-TMPRSS2-BHK21	Hamster	Kidney	High	High	Present

Verda reno E6 cells (Vero E6), a kidney epithelial cell line isolated from African green monkey (C. sabaeus) by Japanese scientists from Chiba University, are the most used mammalian cell line in COVID-19 research [104, 105]. The lack of interferon-mediated antiviral immunity makes it the optimal cell line for virology studies. Vero E6 cells have been used to study various viruses including simian polyomavirus (SV-40), rubella virus, arboviruses, adenoviruses, H5N1 influenza virus, Ebola hemorrhagic fever virus 19, etc [106–108]. In addition, Vero E6 cell line is also particularly useful for coronavirus studies. In late 2002, SARS-CoV-1 emerged and caused life threats in China. SARS-CoV-1 enters the host cells via the binding of its spike protein to ACE2 receptor, and Vero E6 cells express high level of ACE2 on the cell surface [109]. Therefore, Vero E6 cells were inoculated with a throat-swab specimen to evaluate and investigate the cytopathic effects induced by the viral infection [110–112]. Ten years later, another betacoronavirus named MERS emerged and caused deaths in Saudi Arabia [113]. Likewise, Vero E6 cells were used again for cytopathology studies and drug development experiments [114]. Once COVID-19 pandemic strikes, Vero E6 cell line was promptly used for isolation and propagation of the authentic SARS-CoV-2 viruses. Its high endogenous expression of ACE2 and deficiency of interferon signaling are

unique advantages over other cell lines, although the endogenous expression of TMPRSS2 in Vero E6 cells is very low. In order to improve virus isolation efficiency, Vero E6 cells were further engineered for overexpressing TMPRSS2 [115]. One of the most important Vero cell applications is the development of inactivated virus vaccine by Sinopharm. Vero E6 cells were used to propagate a large scale of SARS-CoV-2, which was then soaked in beta-propiolactone for inactivation [77]. In the field of SARS-CoV-2 research, Vero E6 is also commonly used for high-throughput drug screening [116]. It is important to note that Vero E6 is derived from African green monkey, a non-human primate species. Therefore, the experimental data generate from Vero E6 cells may not completely recapitulate viral infection in human cases. Some additional human cell line data are necessary for supplementation.

Cultured human airway epithelial cells (Calu-3) is a human lung cancer cell line that represents the human lung physiology [117]. It was previously shown that Calu-3 is permissive to both SARS-CoV-1 and SARS-CoV-2 [118]. The SARS-CoV-2 propagation in Calu-3 cells, nevertheless, is slower than in Vero E6 cells and fails to induce cytopathic effects, possibly due to the interferon-mediated immunity and the lower level of ACE2 expression [119]. This human lung epithelial cell line is commonly used for antiviral drug screening in addition to Vero E6 cells [120].

Cancer coli 2 cells (Caco-2) are isolated from a colorectal adenocarcinoma. It is reported that SARS-CoV-2 can also infect the digestive system and be found in some patients' stool. Therefore, Caco-2 cells are used to represent the small intestine enterocytes in the gastrointestinal tract upon SARS-CoV-2 infection. A recent study used Caco-2 cells to obtain the unbiased proteomic profile of SARS-CoV-2 infected cells to identify potential therapeutic targets [121]. Compared to Vero E6 and Calu-3 cells, Caco-2 expresses lower level of ACE2 but higher level of

TMPRSS2 on its surface, which leads to more difficulties for infection and a lower level of infection-induced cytopathic effects [122].

Human hepatoma 7 cells (HuH-7) are derived from human liver cancer [123]. It can be infected with human coronaviruses such as CoV 229E, CoV OC43, MERS-CoV, and SARS-CoV [124, 125]. Although HuH-7 is reported to be permissive to SARS-CoV-2, HuH-7 barely expresses ACE2 endogenously [126]. Meanwhile, the replication of SARS-CoV-2 in HuH-7 cells is also moderate, which is consistent with the clinical observation that a small fraction of COVID-19 patients also develops liver injury [127].

Other than these cell lines discussed above, ACE2 and TMPRSS2 are co-overexpressed in some non-permissive cell lines to maximize the viral entry of SARS-CoV-2. Human embryonic kidney 293T cells (HEK 293T) are engineered to overexpress both ACE2 and TMPRSS2 genes because HEK 293T cells are easy to transfect or transduce. This transgenic cell line is widely used for evaluation of neutralizing antibody and antiviral drugs against COVID-19 [128]. A hamster-derived cell line, BHK21 cells were also transfected with plasmids encoding ACE2 and TMPRSS to identify the protease responsible for viral entry of SARS-CoV-2 [129]. The limitation of these engineered cells is the need of quality control to ensure the constantly high gene expression level of ACE2 and TMPRSS2.

1.6.2 Animal models for SARS-CoV-2

In vivo animal models are indispensable for comprehensive understanding of SARS-CoV-2 pathogenesis and developing antiviral agents or vaccines. To meet the need for basic research and preclinical studies, various animal models have been established. The major models are summarized in **Table 1.5**.

Table 1.5 Animal models for SARS-CoV-2 studies[a]

Animal model	Infection features	Disease severity	Transmission	Ease of use	Cost
Hamsters	Weight loss, viral infection in lung, trachea, kidney, etc.	Mild to moderate	Efficient	Easy	Low
Transgenic mice	Weight loss, viral load in lung, intestine, kidney, etc.	Moderate to lethal	Not efficient	Easy	Low
Ferrets	Elevated body temperature, viral load in lung, trachea, intestine, etc.	Mild to moderate	Efficient	Moderate	Moderate
Nonhuman primates	Weight loss, reduced appetite, viral load in lung, trachea, nose, throat, etc.	Moderate	Efficient	Difficult	High

[a] Modified from reference [142]

Hamster models are widely used for respiratory disease studies and susceptible for SARS-CoV-2 infection. Chan et al. reported the use of golden Syrian hamsters for implications of COVID-19 pathogenesis and transmissibility. The hamster model of infection demonstrated a 10% body weight loss within one week of infection, lethargy, ruffled fur, hunched back, and rapid breath symptoms. Histological analysis showed submucosal infiltration and blood vessel congestion at 2 days post-infection. The dwelling trachea and local inflammation were also observed on day 2 post-infection. The viral infection was not lethal, and all animals recovered 14 days post-infection [130]. Likewise, Sia et al. intranasally inoculated golden Syrian hamsters with 8×10^4 $TCID_{50}$ of SARS-CoV-2 to investigate pathogenesis and transmissibility of SARS-CoV-2. The infection caused lung consolidation as early as 2 days after infection. Meanwhile, SARS-

CoV-2 N protein was detected by immunohistochemistry assay in nasal mucosa, and bronchial epithelial cells. Viral transmission from inoculated animals to naïve animals was observed. Consistent with data from Chan et al., animals gained their body weight back within 14 days [131]. In addition, hamster models were used to investigate the impact of variant mutations on virus fitness and pathogenicity. Zhou et al. used hamster models to demonstrate that SARS-CoV-2 spike D614G mutation increases transmissibility and replication [132]. Liu et al. used 4- to 6-week-old golden Syrian hamsters to show the N501Y spike substitution enhances viral infection and transmission [133]. Taken together, the susceptibility to SARS-CoV-2 and the ease of handling make hamsters an ideal model to mimic the moderate but not lethal pathological features of COVID-19.

Murine models are the most widely used for biological studies because they are cost-effective, easy to handle, and rapid breeding. However, wild-type mice are not permissive to the authentic SARS-CoV-2 infection due to the difference between human ACE2 and mouse ACE2 protein. Therefore, transgenic mouse models were generated to express human ACE2 protein for SARS-CoV-2 spike binding by various approaches including transgenic, CRISPR/Cas9 knock-in technology, adenoviral vector, and Venezuelan equine encephalitis replicon particles [134–138]. Due to the high availability, K18 hACE2 transgenic mice have become the most used animal model for SARS-CoV-2 study. This transgenic animal model was originally developed for studying SARS-CoV-1 [139]. Since SARS-CoV-2 also uses ACE2 receptor to enter the host, K18 hACE2 model was readily adopted to investigate SARS-CoV-2 [140]. Compared to the moderate pathologic features in hamsters, SARS-CoV-2 infection induced cytokine storm and is lethal to k18 hACE2 mice. Therefore, the lethality of SARS-CoV-2 infection makes it a suitable model for identification of effective vaccines or antiviral drugs. Instead of genetically modifying the animals,

SARS-CoV-2 virus can also be modified to infect mice. Gu et al. adapted a clinical isolate of SARS-CoV-2 in aged BALB/C mice for multiple passages, which resulted in a virus mutant that causes infection-associated pneumonia and inflammation [141].

Ferrets (Mustela putorius furo) are susceptible for SARS-CoV-2 infection. It is reported that the ferret model of infection is akin to human COVID-19 cases. Once infected, the ferrets showed higher body temperature and viral RNA was found in multiple tissues including nasal turbinate, lung, trachea, kidney, and intestine [142]. The infection was found not lethal, but the viral transmission to naïve animals by direct contact and acute bronchiolitis were observed. The ferret models were also used for antiviral drug evaluation. Cox et al. tested the effect of molnupiravir in mitigating infection and blocking transmission of SARS-CoV-2 in ferrets [143]. Blanco-Melo et al. demonstrated an imbalanced innate immunity and high inflammatory responses in ferret models [144].

These small animal models are widely used for SARS-CoV-2 research because of the ease of handling. Nevertheless, small animals may not represent human infections due to the genomic differences. Nonhuman primates including rhesus macaques (Macaca mulatta), cynomolgus macaques (Macaca fascicularis) are of significance for recapitulating clinical symptoms and serve as the gold standard of preclinical trials. For example, Pfizer and BioNTech used rhesus macaques to examine the BNT162b vaccine protection effects against SARS-CoV-2 before the product enters clinical trials [145]. Likewise, Moderna's mRNA-1273 vaccine was evaluated in rhesus macaques and showed robust activity of generating neutralizing antibody and protecting animals from severe symptoms [146]. In addition to COVID-19 vaccines, nonhuman primates are also used for antiviral drug evaluation. A few months since the COVID-19 emerged, the nucleoside analogue remdesivir was tested in rhesus macaques. The data suggest that the early intervention of

remdesivir reduced viral load and lung damage caused by infection [147]. The limitations of nonhuman primate models are that the infection is mostly not lethal and that the cost is significantly higher than other small animal models.

1.6.3 BSL-2-compatible SARS-CoV-2 substitute systems

One major obstacle of conducting SARS-CoV-2 research is that the authentic virus needs to be handled in a BSL-3 laboratory, which is not available to most researchers. Therefore, various substitute model systems were developed to enable conducting SARS-CoV-2 research under BSL-2 conditions (**Table 1.6**). These systems include SARS-CoV-2 spike-pseudotyped viruses, SARS-CoV-2 replicons, SARS-CoV-2 virus-like particles (VLPs), etc. They either partially or completely represent the morphological and infectious feature of SARS-CoV-2; however, their application could be limited by the lack of certain essential components found in the wild type virus.

Table 1.6 BSL-2-compatible SARS-CoV-2 substitute systems

Substitute systems	**SARS-CoV-2 structural proteins**	**Applications**	**Limitations**
Spike-pseudoviruses	Spike only	Viral entry drug screening, neutralization assay	Absence of other structural proteins
Replicons	Mostly no spike	Drug screening	Mostly no spike, no complete viral structure
VLPs	All 4 present	Viral entry modeling, drug screening, neutralization assay	Mostly can't express reporter gene

Pseudotyped-virus is a common strategy to study a highly virulent virus. A good example is the development of lentiviral vector system in which the Human immunodeficiency virus type

I (HIV-1) glycoprotein is replaced by Vesicular stomatitis virus (VSV) glycoprotein due to cytotoxicity [148]. In the case of SARS-CoV-2, multiple pre-existing viruses have been pseudotyped with spike protein of SARS-CoV-2 for conferring cellular tropism. And a reporter gene, mostly luciferase or GFP gene, is carried by the pseudovirus for the purpose of quantification. For example, lentiviruses were pseudotyped with spike protein by co-transfection of the plasmid encoding spike protein gene with other plasmids encoding HIV genes required for lentiviral vectors. Because the system is easy to use, pseudotyped lentiviruses are the most widely employed for neutralizing antibody quantification and anti-S drug screening [149–151]. Furthermore, the transduction efficiency of pseudotyped-lentiviruses was improved by pseudotyping with the delta variant spike protein containing a key mutation of D614G [152–154]. VSV is also a good model for pseudotyping because of its simple structure and ability to grow. In this case, the VSV glycoprotein gene is replaced by the SARS-CoV-2 spike while producing the viruses. Like pseudotyped lentiviruses, spike-VSV systems are also suitable for evaluating antiviral agents targeting viral entry [155, 156]. Murine leukemia virus (MLV) is another virus employed for pseudotyping SARS-CoV-2. MLV belongs to the family of retrovirus. The robust transduction-induced reporter gene expression level makes MLV an ideal model virus. MLV is reported to be pseudotyped with coronavirus's spike proteins including MERS, SARS-CoV-1, and SARS-CoV-2 [157, 158]. Although these spike-pseudotyped viruses share the same cellular tropism with SARS-CoV-2 and transduce ACE2+ cell in an efficient manner, their viral particles only contain SARS-CoV-2 spike protein. They may not be able to completely represent the viral structure of SARS-CoV-2 due to the absence of other 3 structural proteins of SARS-CoV-2.

SARS-CoV-2 replicon systems are developed to investigate the replication mechanism of SARS-CoV-2 and develop drugs targeting this critical process. The key feature of replicon systems

is that the structural gene needs to be deleted or expressed separately from the entire DNA/RNA replicon sequence, such that there is no replication-competent virion produced. Ricardo-Lax et al. constructed a SARS-CoV-2 replicon using a yeast-based system. SARS-CoV-2 spike protein gene was deleted and replaced with reporter genes to indicate viral replication [159]. Nguyen et al. developed a noninfectious SARS-CoV-2 replicon in a bacterial artificial chromosomal (BAC) vector by reverse genetics. In this system, spike protein gene was replaced by a luciferase reporter. Once transfected into Vero E6 cells, the replicon induces expression of nano luciferase, which can be sensitively detected for antiviral drug testing and gene function studies [160]. Tanaka et al. utilized the similar strategy to develop a stable cell line expressing the reporter replicon. The replicon system was then used for high throughput screening to identify potent COVID-19 therapeutics such as remdesivir, molnupiravir, and interferon-β [161]. Most SARS-CoV-2 replicon systems exclude spike protein to avoid the synthesis of live virus. Li et al., however, utilized a VEEV-derived RNA replicon system to build a self-propagating hybrid replicon. This replicon system is composed of a VEEV self-amplifying RNA encoding the SARS-CoV-2 spike protein gene. Unlike other replicon systems lacking the spike protein, it can be used for screening viral entry inhibitors ad studying cell-to-cell transmitted syncytia formation [162].

SARS-CoV-2 virus-like particles (VLP) consist of 4 structural proteins to represent the complete viral structure of the authentic virus. VLPs are typically produced by co-transfection of 4 independent plasmids encoding 4 structural proteins of SARS-CoV-2. Since the COVID-19 pandemic strikes, SARS-CoV-2 VLPs have been constructed in various systems including mammalian cells, plant cells, and insect cells, etc. Xu et al. demonstrated that SARS-CoV-2 can be assembled in both HEK293T and Vero E6 cells. And their data suggest that VLPs produced from Vero E6 cells tend to have more stable corona-like structures compared to the counterpart of

HEK293T cells [163]. Swann et al. showed that N protein is not essential for the assembly and formation of SARS-CoV-2 VLPs in mammalian cells [164]. Furthermore, Plescia et al. developed a BSL-2-compatible VLP system in which the spike protein is tagged with GFP. This fluorescently labeled VLP system can be utilized to model viral entry [165]. In addition to mammalian cells as the VLP producer, SARS-CoV-2 was also produced in plant and insect cells [166–170]. In these producer systems, the yield of VLP is significantly higher and the cost of production is lower than using mammalian cells. Since the VLPs consist of all 4 structural proteins, they not only serve as a modeling system but can also serve as an effective vaccine to elicit the immunity against SARS-CoV-2. For example, a Canadian company Medicago developed a plant-based SARS-CoV-2 VLP vaccine named COVIFENZ and received approval by Health Canada in February 2022. The adjuvanted VLP vaccine showed 69.5% vaccine efficacy against symptomatic infection and 78.8% vaccine efficacy against moderate-to-severe disease [171]. The VLP systems represent an empty viral protein shell without any RNA packaged. Therefore, VLP systems failed to serve as a quantitative model for screening antiviral drugs or evaluating neutralizing antibodies. Syed et al. established a new type of SARS-CoV-2 virus-like particles (SC2-VLPs) capable of packaging and delivering exogeneous transgene into ACE2+ cells [172]. The key component of the SC2-VLP system is the putative packaging signal sequence (PS9 sequence) identified in the SARS-CoV-2 genome, which direct the N protein selectively binding to the viral genome in the host cell cytoplasm. The SC2-VLP system was then used to examine the effect of structural gene mutations on viral fitness and infectivity of evolved variants [173]. Although the SC2-VLP system has unique advantages over other SARS-CoV-2 VLP systems, its gene delivery efficiency is still limited. There is still great room for improving the delivery efficiency to enable wider applications of this system.

1.7 Challenges of drug development against COVID-19

The emergence and rapid spread of SARS-CoV-2 posed unparalleled threat to the global public health. Although these anti-COVID policies were effective to control the spread of COVID-19, its influence on economy and inconvenience caused inevitable frustration of public. Intervention approaches against COVID-19 including prevention and treatment were in urgent need to save human lives. Unfortunately, the pandemic caused multifactorial challenges for scientific community to perform basic virology and drug development research [174].

New drug development is well-known to be time-consuming and highly expensive. A new drug typically takes more than 10 years from the early stage of target identification to the final approval by regulatory agencies, which apparently is too slow for handling this rapidly evolving situation. Therefore, repurposing existing antiviral drugs to combat the current pandemic served as a reasonable strategy [175-179]. However, since SARS-CoV-2 was a novel virus that could not be targeted by pre-existing antiviral agents, those repurposed pharmacotherapies demonstrated moderate or even no efficacy in hospitalized patients [180-182]. The disappointing results called the urgent need for comprehensive understanding of the basic SARS-CoV-2 virology and novel technologies that can be quickly adapted to combat the pandemic spread.

Due to its airborne transmission and life-threatening nature, SARS-CoV-2 is classified as a Risk Group 3 biological agent. And it requires to be handled in a Biosafety level 3 (BSL-3) laboratory which is designed for working with potentially lethal pathogens that can cause diseases via inhalation [183]. The limited access of BSL-3 laboratories to the scientific community and the high cost of running experiments under BSL-3 conditions became a major obstacle for advancing COVID-19 drug development [184]. To circumvent this obstacle, significant efforts have been made to develop BSL-2 compatible model systems mimicking SARS-CoV-2. Although model

systems (e.g., spike-pseudotyped viruses, replicons, and virus-like particles) showed some similarities to the authentic SARS-CoV-2, there is still a need for developing better systems to recapitulate both viral structure and behavior, which ultimately enables the high-throughput screening of drugs targeting SARS-CoV-2.

1.8 Objectives of Dissertation Research

This dissertation aims to develop a novel SARS-CoV-2 viral vector system (VEE-SARS-CoV-2 vectors) to deliver genetic material efficiently. This viral vector is designed to comprise of the 4 SARS-CoV-2 structural proteins (spike, envelop, membrane, nucleocapsid) and carry a self-amplifying Venezuelan equine encephalitis virus (VEEV) replicon, a putative SARS-CoV-2 packaging signal, and the gene of interest. This dissertation entails the construction, optimization, and characterization of the VEE-SARS-CoV-2 vector system. In addition, its applications for antiviral drug testing and neutralizing antibody quantitation are demonstrated. The developed VEE-SARS-CoV-2 vector system is an important and versatile tool to meet the need for investigating SARS-CoV-2 molecular virology and developing antiviral agents targeting receptor binding, RNA genome packaging, or essential but not well studied structural proteins of SARS-CoV-2.

References

1. N. Zhu *et al*., A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* **382**, 727-733 (2020).

2. F. Wu *et al*., A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).

3. P. Zhou *et al*., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).

4. J. F. Chan *et al*., A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **395**, 514-523 (2020).

5. N. Chen *et al*., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507-513 (2020).

6. V. Coronaviridae Study group of the international committee on taxonomy of the species of severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544 (2020).

7. John Hopkins Coronavirus Resource Center. Johns Hoskins University of medicine. Retrieved March 6, 2023, from https://coronavirus.jhu.edu/map.html.

8. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**, 181-192 (2019).

9. D. Forni *et al*., Adaptation of the endemic coronaviruses HCoV-OC43 and HCoV-229E to the human host. *Virus Evol* **7**, veab061 (2021).

10. V. M. Corman *et al.*, Rooting the phylogenetic tree of middle east respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *J Virol* **88**, 11297-11303 (2014).

11. M. F. Boni *et al.*, Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat Microbiol* **5**, 1408-1417 (2020).

12. F. Wu *et al.*, A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).

13. P. Zhou *et al*., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).

14. H. Zhou *et al.*, Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell* **184**, 4380-4391 (2021).

15. A. A. T. Naqvi *et al*., Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: structural genomics approach. *Biochim Biophys Acta Mol Basis Dis* **1866**, 165878 (2020).

16. R. Lu *et al*., Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574 (2020).

17. F. K. Yoshimoto, The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *Protein J* **39**, 198-216 (2020).

18. R. Lu *et al*., Complete genome sequence of middle east respiratory syndrome coronavirus (MERS-CoV) from the first imported MERS-CoV case in China. *Genome Announc* **3**, e00818-15 (2015).

19. P. A. Rota *et al*., Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**, 1394-1399 (2003).

20. C. S. Goldsmith, S. E. Miller, R. B. Martines, H. A. Bullock, S. R. Zaki, Electron microscopy of SARS-CoV-2: a challenging task. *Lancet* **395**, e99 (2020).

21. S. Xia *et al*., Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell Mol Immunol* **17**, 765-767 (2020).

22. J. Shang *et al*., Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* **117**, 11727-11734 (2020).

23. A. J. B. Kreutzberger *et al*., SARS-CoV-2 requires acidic pH to infect cells. *Proc Natl Acad Sci U S A* **119**, e2209514119 (2022).

24. C. B. Jackson, M. Farzan, B. Chen, H. Choe, Mechanisms of SARS-CoV-2 entry into cells. *Nat Rev Mol Cell Biol* **23**, 3-20 (2022).

25. R. Yan *et al*., Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444-1448 (2020).

26. M. Hoffmann *et al*., SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271-280 (2020).

27. A. A. Rabaan *et al*., SARS-CoV-2, SARS-CoV, and MERS-COV: a comparative overview. *Infez Med* **28**, 174-184 (2020).

28. K. El Omari *et al*., The structure of a prokaryotic viral envelope protein expands the landscape of membrane fusion proteins. *Nat Commun* **10**, 846 (2019).

29. E. A. J. Alsaadi, B. W. Neuman, I. M. Jones, Identification of a membrane binding peptide in the envelope protein of mhv coronavirus. *Viruses* **12**, 1054 (2020).

30. D. Cabrera-Garcia, R. Bekdash, G. W. Abbott, M. Yazawa, N. L. Harrison, The envelope protein of SARS-CoV-2 increases intra-Golgi pH and forms a cation channel that is regulated by pH. *J Physiol* **599**, 2851-2868 (2021).

31. Y. Cao *et al*., Characterization of the SARS-CoV-2 E protein: sequence, structure, viroporin, and inhibitors. *Protein Sci* **30**, 1114-1130 (2021).

32. S. Thomas, The structure of the membrane protein of SARS-CoV-2 resembles the sugar transporter semiSWEET. *Pathog Immun* **5**, 342-363 (2020).

33. Y. Cao *et al*., Probing the formation, structure and free energy relationships of M protein dimers of SARS-CoV-2. *Comput Struct Biotechnol J* **20**, 573-582 (2022).

34. B. Nal *et al*., Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *J Gen Virol* **86**, 1423-1434 (2005).

35. B. W. Neuman *et al*., A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol* **174**, 11-22 (2011).

36. Y. Zheng *et al*., Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) membrane (M) protein inhibits type I and III interferon production by targeting RIG-I/MDA-5 signaling. *Signal Transduct Target Ther* **5**, 299 (2020).

37. Y. Z. Fu *et al*., SARS-CoV-2 membrane glycoprotein M antagonizes the MAVS-mediated innate antiviral response. *Cell Mol Immunol* **18**, 613-620 (2021).

38. S. Lu *et al*., The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat Commun* **12**, 502 (2021).

39. P. S. Masters, L. S. Sturman, Functions of the coronavirus nucleocapsid protein. *Adv Exp Med Biol* **276**, 235-238 (1990).

40. R. McBride, M. van Zyl, B. C. Fielding, The coronavirus nucleocapsid is a multifunctional protein. *Viruses* **6**, 2991-3018 (2014).

41. C. Iserman *et al*., Genomic RNA elements drive phase separation of the SARS-CoV-2 Nucleocapsid. *Mol Cell* **80**, 1078-1091 (2020).

42. J. Cubuk *et al*., The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat Commun* **12**, 1936 (2021).

43. A. Savastano, A. Ibanez de Opakua, M. Rankovic, M. Zweckstetter, Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nat Commun* **11**, 6041 (2020).

44. Y. Zhao *et al*., A dual-role of SARS-CoV-2 nucleocapsid protein in regulating innate immune response. *Signal Transduct Target Ther* **6**, 331 (2021).

45. C. Bai, Q. Zhong, G. F. Gao, Overview of SARS-CoV-2 genome-encoded proteins. *Sci China Life Sci* **65**, 280-294 (2022).

46. N. Redondo, S. Zaldivar-Lopez, J. J. Garrido, M. Montoya, SARS-CoV-2 accessory proteins in viral pathogenesis: knowns and unknowns. *Front Immunol* **12**, 708264 (2021).

47. Y. Konno *et al*., SARS-CoV-2 ORF3b is a potent interferon antagonist whose activity is increased by a naturally occurring elongation variant. *Cell Rep* **32**, 108185 (2020).

48. J. Y. Li *et al*., The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. *Virus Res* **286**, 198074 (2020).

49. Y. Lu, H. A. Michel, P. H. Wang, G. L. Smith, Manipulation of innate immune signaling pathways by SARS-CoV-2 non-structural proteins. *Front Microbiol* **13**, 1027015 (2022).

50. Y. Ren *et al*., The ORF3a protein of SARS-CoV-2 induces apoptosis in cells. *Cell Mol Immunol* **17**, 881-883 (2020).

51. C. S. Shi *et al*., SARS-coronavirus open reading frame-9b suppresses innate immunity by targeting mitochondria and the MAVS/TRAF3/TRAF6 signalosome. *J Immunol* **193**, 3080-3089 (2014).

52. K. K. Singh, G. Chaubey, J. Y. Chen, P. Suravajhala, Decoding SARS-CoV-2 hijacking of host mitochondria in COVID-19 pathogenesis. *Am J Physiol Cell Physiol* **319**, C258-C267 (2020).

53. C. B. Jackson, M. Farzan, B. Chen, H. Choe, Mechanisms of SARS-CoV-2 entry into cells. *Nat Rev Mol Cell Biol* **23**, 3-20 (2022).

54. Y. Finkel *et al*., The coding capacity of SARS-CoV-2. *Nature* **589**, 125-130 (2021).

55. S. G. Sawicki, D. L. Sawicki, Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv Exp Med Biol* **380**, 499-506 (1995).

56. I. Sola, F. Almazan, S. Zuniga, L. Enjuanes, Continuous and discontinuous RNA synthesis in coronaviruses. *Annu Rev Virol* **2**, 265-288 (2015).

57. C. C. Macklin, The pulmonary alveolar mucoid film and the pneumonocytes. *Lancet* **266**, 1099-1104 (1954).

58. W. Sungnak *et al*., SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med* **26**, 681-687 (2020).

59. N. Chen *et al*., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507-513 (2020).

60. C. Huang *et al*., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506 (2020).

61. X. Cui *et al*., Pulmonary edema in COVID-19 patients: mechanisms and treatment potential. *Front Pharmacol* **12**, 664349 (2021).

62. A. Gupta *et al*., Extrapulmonary manifestations of COVID-19. *Nat Med* **26**, 1017-1032 (2020).

63. Z. Varga *et al*., Endothelial cell infection and endotheliitis in COVID-19. *Lancet* **395**, 1417-1418 (2020).

64. C. Huang *et al*., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506 (2020).

65. R. Mao *et al*., Manifestations and prognosis of gastrointestinal and liver involvement in patients with COVID-19: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* **5**, 667-678 (2020).

66. Y. Li *et al*., A Comprehensive review of the global efforts on COVID-19 vaccine development. *ACS Cent Sci* **7**, 512-533 (2021).

67. C. Keech *et al*., Phase 1-2 trial of a SARS-CoV-2 recombinant spike protein nanoparticle vaccine. *N Engl J Med* **383**, 2320-2332 (2020).

68. P. T. Heath *et al*., Safety and efficacy of NVX-CoV2373 Covid-19 vaccine. *N Engl J Med* **385**, 1172-1183 (2021).

69. F. P. Polack *et al*., Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med* **383**, 2603-2615 (2020).

70. L. R. Baden *et al*., Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med* **384**, 403-416 (2021).

71. L. A. Jackson *et al*., An mRNA vaccine against SARS-CoV-2 - preliminary report. *N Engl J Med* **383**, 1920-1931 (2020).

72. D. H. Barouch, Covid-19 vaccines - immunity, variants, boosters. *N Engl J Med* **387**, 1011-1020 (2022).

73. S. A. Halperin *et al*., Final efficacy analysis, interim safety analysis, and immunogenicity of a single dose of recombinant novel coronavirus vaccine (adenovirus type 5 vector) in adults 18 years and older: an international, multicentre, randomised, double-blinded, placebo-controlled phase 3 trial. *Lancet* **399**, 237-248 (2022).

74. J. Sadoff *et al*., Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. *N Engl J Med* **384**, 2187-2201 (2021).

75. A. R. Falsey *et al*., Phase 3 safety and efficacy of AZD1222 (ChAdOx1 nCoV-19) Covid-19 vaccine. *N Engl J Med* **385**, 2348-2360 (2021).

76. M. D. Tanriover *et al*., Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomized, placebo-controlled, phase 3 trial in Turkey. *Lancet* **398**, 213-222 (2021).

77. H. Wang *et al*., Development of an inactivated vaccine candidate, BBIBP-CorV, with potent protection against SARS-CoV-2. *Cell* **182**, 713-721 e19 (2020).

78. S. Xia *et al*., Safety and immunogenicity of an inactivated COVID-19 vaccine, BBIBP-CorV, in people younger than 18 years: a randomized, double-blind, controlled, phase 1/2 trial. *Lancet Infect Dis* **22**, 196-208 (2022).

79. Y. Li *et al*., A Comprehensive review of the global efforts on COVID-19 vaccine development. *ACS Cent Sci* **7**, 512-533 (2021).

80. R. Ella *et al*., Efficacy, safety, and lot-to-lot immunogenicity of an inactivated SARS-CoV-2 vaccine (BBV152): interim results of a randomized, double-blind, controlled, phase 3 trial. *Lancet* **398**, 2173-2184 (2021).

81. N. Murakami *et al*., Therapeutic advances in COVID-19. *Nat Rev Nephrol* **19**, 38-52 (2023).

82. R.-C. Investigators *et al*., Interleukin-6 receptor antagonists in critically ill patients with Covid-19. *N Engl J Med* **384**, 1491-1502 (2021).

83. R. C. Group, Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomized, controlled, open-label, platform trial. *Lancet* **397**, 1637-1645 (2021).

84. V. C. Marconi *et al*., Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): a randomized, double-blind, parallel-group, placebo-controlled phase 3 trial. *Lancet Respir Med* **9**, 1407-1418 (2021).

85. A. C. Kalil *et al*., Baricitinib plus Remdesivir for hospitalized adults with Covid-19. *N Engl J Med* **384**, 795-807 (2021).

86. T. K. Warren *et al*., Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature* **531**, 381-385 (2016).

87. T. P. Sheahan *et al.*, Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. *Sci Transl Med* **9**, eaal3653 (2017).

88. R. L. Gottlieb *et al.*, Early Remdesivir to prevent progression to severe Covid-19 in outpatients. *N Engl J Med* **386**, 305-315 (2022).

89. J. H. Beigel *et al.*, Remdesivir for the treatment of Covid-19 - final report. *N Engl J Med* **383**, 1813-1826 (2020).

90. W. H. O. S. T. Consortium, Remdesivir and three other drugs for hospitalized patients with COVID-19: final results of the WHO Solidarity randomized trial and updated meta-analyses. *Lancet* **399**, 1941-1953 (2022).

91. D. R. Owen *et al*., An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586-1593 (2021)

92. I. F. Sevrioukova, T. L. Poulos, Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir. *Proc Natl Acad Sci U S A* **107**, 18422-18427 (2010).

93. J. Hammond *et al*., Oral Nirmatrelvir for high-risk, non-hospitalized adults with Covid-19. *N Engl J Med* **386**, 1397-1408 (2022).

94. T. P. Sheahan *et al*., An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci Transl Med* **12** eabb5883 (2020).

95. A. Jayk Bernal *et al*., Molnupiravir for oral treatment of Covid-19 in non-hospitalized patients. *N Engl J Med* **386**, 509-520 (2022)

96. A. Jorda *et al*., Convalescent plasma treatment in patients with covid-19: a systematic review and meta-analysis. *Front Immunol* **13**, 817829 (2022).

97. P. Begin *et al*., Convalescent plasma for hospitalized patients with COVID-19: an open-label, randomized controlled trial. *Nat Med* **27**, 2012-2024 (2021).

98. P. Millat-Martinez *et al*., Prospective individual patient data meta-analysis of two randomized trials on convalescent plasma for COVID-19 outpatients. *Nat Commun* **13**, 2583 (2022).

99. M. Dougan *et al*., Bamlanivimab plus Etesevimab in mild or moderate Covid-19. *N Engl J Med* **385**, 1382-1392 (2021).

100. R. C. Group, Casirivimab and imdevimab in patients admitted to hospital with COVID-19 (RECOVERY): a randomized, controlled, open-label, platform trial. *Lancet* **399**, 665-676 (2022).

101. M. J. Levin *et al*., Intramuscular AZD7442 (Tixagevimab-Cilgavimab) for prevention of Covid-19. *N Engl J Med* **386**, 2188-2200 (2022).

102. Steffen Faisst, Propagation of viruses | animal. *Encyclopedia of Virology* 1408–1413 (1999).

103. G. A. Pires De Souza *et al*., Choosing a cellular model to study SARS-CoV-2. *Front Cell Infect Microbiol* **12**, 1003608 (2022).

104. N. C. Ammerman, M. Beier-Sexton, A. F. Azad, Growth and maintenance of Vero cell lines. *Curr Protoc Microbiol* Appendix 4, Appendix 4E (2008).

105. N. Osada *et al*., The genome landscape of the African green monkey kidney-derived Vero cell line. *DNA Res* **21**, 673-683 (2014).

106. D. S. Ellis *et al*., Ebola and Marburg viruses: II, their development within Vero cells and the extra-cellular formation of branched and torus forms. *J Med Virol* **4**, 213-225 (1979).

107. T. Horimoto, Y. Kawaoka, Strategies for developing vaccines against H5N1 influenza A viruses. *Trends Mol Med* **12**, 506-514 (2006).

108. S. Kiesslich, A. A. Kamen, Vero cell upstream bioprocess development for the production of viral vectors and vaccines. *Biotechnol Adv* **44**, 107608 (2020).

109. W. Li *et al*., Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).

110. T. G. Ksiazek *et al*., A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953-1966 (2003).

111. T. Mizutani, S. Fukushi, M. Saijo, I. Kurane, S. Morikawa, Phosphorylation of p38 MAPK and its downstream targets in SARS coronavirus-infected cells. *Biochem Biophys Res Commun* **319**, 1228-1234 (2004).

112. T. Mizutani *et al*., Inhibition of cell proliferation by SARS-CoV infection in Vero E6 cells. *FEMS Immunol Med Microbiol* **46**, 236-243 (2006).

113. A. M. Zaki, S. van Boheemen, T. M. Bestebroer, A. D. Osterhaus, R. A. Fouchier, Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **367**, 1814-1820 (2012).

114. A. H. de Wilde *et al*., MERS-coronavirus replication induces severe in vitro cytopathology and is strongly inhibited by cyclosporin A or interferon-alpha treatment. *J Gen Virol* **94**, 1749-1760 (2013).

115. S. Matsuyama *et al*., Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A* **117**, 7001-7003 (2020).

116. K. Yan, D. J. Rawle, T. T. Le, A. Suhrbier, Simple rapid in vitro screening method for SARS-CoV-2 anti-virals that identifies potential cytomorbidity-associated false positives. *Virol J* **18**, 123 (2021).

117. J. Fogh, J. M. Fogh, T. Orfeo, One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. *J Natl Cancer Inst* **59**, 221-226 (1977).

118. H. Chu *et al*., Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. *Lancet Microbe* **1**, e14-e23 (2020).

119. B. K. Park *et al*., Differential signaling and virus production in calu-3 cells and Vero cells upon SARS-CoV-2 infection. *Biomol Ther (Seoul)* **29**, 273-281 (2021).

120. S. Tao *et al*., Comparison of anti-SARS-CoV-2 activity and intracellular metabolism of remdesivir and its parent nucleoside. *Curr Res Pharmacol Drug Discov* **2**, 100045 (2021).

121. D. Bojkova *et al*., Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469-472 (2020).

122. H. Shuai *et al*., Differential immune activation profile of SARS-CoV-2 and SARS-CoV infection in human lung and intestinal cells: implications for treatment with IFN-beta and IFN inducer. *J Infect* **81**, e1-e10 (2020)

123. H. Nakabayashi, K. Taketa, K. Miyano, T. Yamane, J. Sato, Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res* **42**, 3858-3863 (1982).

124. F. Freymuth *et al*., Replication of respiratory viruses, particularly influenza virus, rhinovirus, and coronavirus in HuH7 hepatocarcinoma cell line. *J Med Virol* **77**, 295-301 (2005).

125. A. H. de Wilde *et al*., MERS-coronavirus replication induces severe in vitro cytopathology and is strongly inhibited by cyclosporin A or interferon-alpha treatment. *J Gen Virol* **94**, 1749-1760 (2013).

126. E. Saccon *et al*., Cell-type-resolved quantitative proteomics map of interferon response against SARS-CoV-2. *iScience* **24**, 102420 (2021).

127. F. Zhou, J. Xia, H. X. Yuan, Y. Sun, Y. Zhang, Liver injury in COVID-19: Known and unknown. *World J Clin Cases* **9**, 4980-4989 (2021).

128. J. A. Duty *et al*., Discovery and intranasal administration of a SARS-CoV-2 broadly acting neutralizing antibody with activity against multiple Omicron subvariants. *Med (N Y)* **3**, 705-721 e11 (2022).

129. M. Hoffmann *et al*., Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* **65**, 103255 (2021)

130. J. F. Chan *et al*., Simulation of the clinical and pathological manifestations of coronavirus disease 2019 (covid-19) in a golden Syrian hamster model: implications for disease pathogenesis and transmissibility. *Clin Infect Dis* **71**, 2428-2446 (2020).

131. S. F. Sia *et al*., Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **583**, 834-838 (2020).

132. B. Zhou *et al*., SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature* **592**, 122-127 (2021).

133. Y. Liu *et al*., The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature* **602**, 294-299 (2022).

134. L. Bao *et al*., The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. *Nature* **583**, 830-833 (2020).

135. S. H. Sun *et al*., A mouse model of SARS-CoV 2 infection and pathogenesis. *Cell Host Microbe* **28**, 124-133 e4 (2020).

136. A. O. Hassan *et al*., A SARS-CoV-2 infection model in mice demonstrates protection by neutralizing antibodies. *Cell* **182**, 744-753 e4 (2020).

137. E. S. Winkler *et al*., SARS-CoV-2 infection of human ACE2-transgenic mice causes severe lung inflammation and impaired function. *Nat Immunol* **21**, 1327-1335 (2020).

138. Y. N. Zhang *et al*., A mouse model for SARS-CoV-2 infection by exogenous delivery of hACE2 using alphavirus replicon particles. *Cell Res* **30**, 1046-1048 (2020).

139. P. B. McCray, Jr. *et al*., Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J Virol* **81**, 813-821 (2007).

140. F. S. Oladunni *et al*., Lethality of SARS-CoV-2 infection in K18 human angiotensin-converting enzyme 2 transgenic mice. *Nat Commun* **11**, 6122 (2020).

141. H. Gu *et al*., Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science* **369**, 1603-1607 (2020).

142. Y. I. Kim *et al*., Infection and rapid transmission of SARS-CoV-2 in ferrets. *Cell Host Microbe* **27**, 704-709 e2 (2020).

143. R. M. Cox, J. D. Wolf, R. K. Plemper, Therapeutically administered ribonucleoside analogue MK-4482/EIDD-2801 blocks SARS-CoV-2 transmission in ferrets. *Nat Microbiol* **6**, 11-18 (2021).

144. D. Blanco-Melo *et al*., Imbalanced host response to SARS-CoV-2 drives development of Covid-19. *Cell* **181**, 1036-1045 e1039 (2020).

145. A. B. Vogel *et al*., BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592**, 283-289 (2021).

146. K. S. Corbett *et al*., Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *N Engl J Med* **383**, 1544-1555 (2020).

147. B. N. Williamson *et al*., Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. *Nature* **585**, 273-276 (2020).

148. H. Chu, J. F. Chan, K. Y. Yuen, Animal models in SARS-CoV-2 research. *Nat Methods* **19**, 392-394 (2022).

149. L. Naldini *et al*., In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263-267 (1996).

150. K. H. D. Crawford *et al*., Protocol and reagents for pseudotyping lentiviral particles with SARS-CoV-2 spike protein for neutralization assays. *Viruses* **12**, 513 (2020).

151. C. Lei *et al*., Neutralization of SARS-CoV-2 spike pseudotyped virus by recombinant ACE2-Ig. *Nat Commun* **11**, 2070 (2020).

152. Y. Pisil, H. Shida, T. Miura, A neutralization assay based on pseudo-typed lentivirus with SARS Cov-2 spike protein in ace2-expressing CRFK cells. *Pathogens* **10**, 153 (2021).

153. X. Fu, L. Tao, X. Zhang, Comprehensive and systemic optimization for improving the yield of SARS-CoV-2 spike pseudotyped virus. *Mol Ther Methods Clin Dev* **20**, 350-356 (2021).

154. M. C. Johnson *et al*., Optimized pseudotyping conditions for the SARS-CoV-2 spike glycoprotein. *J Virol* **94**, e01062-20 (2020).

155. Z. Daniloski *et al*., The spike D614G mutation increases SARS-CoV-2 infection of multiple human cell types. *Elife* **10**, e65365 (2021).

156. J. M. Condor Capcha *et al*., Generation of SARS-CoV-2 spike pseudotyped virus for viral entry and neutralization assays: a 1-week protocol. *Front Cardiovasc Med* **7**, 618651 (2020).

157. M. E. Dieterle *et al*., A replication-competent vesicular stomatitis virus for studies of SAES-Cov-2 spike-mediated cell entry and its inhibition. *Cell Host Microbe* **28**, 486-496 e486 (2020).

158. Y. Zheng *et al*., Neutralization assay with SARS-CoV-1 and SARS-CoV-2 spike pseudotyped murine leukemia virions. *Virol J* **18**, 1 (2021).

159. J. K. Millet, G. R. Whittaker, Murine leukemia virus (mlv)-based coronavirus spike-pseudotyped particle production and infection. *Bio Protoc* **6**, e2035 (2016).

160. I. Ricardo-Lax *et al*., Replication and single-cycle delivery of SARS-CoV-2 replicons. *Science* **374**, 1099-1106 (2021).

161. H. T. Nguyen, D. Falzarano, V. Gerdts, Q. Liu, Construction of a noninfectious SARS-CoV-2 replicon for antiviral-drug testing and gene function studies. *J Virol* **95**, e0068721 (2021).

162. T. Tanaka *et al*., Establishment of a stable SARS-CoV-2 replicon system for application in high-throughput screening. *Antiviral Res* **199**, 105268 (2022).

163. N. Li *et al*., A new screening system for entry inhibitors based on cell-to-cell transmitted syncytia formation mediated by self-propagating hybrid VEEV-SARS-CoV-2 replicon. *Emerg Microbes Infect* **11**, 465-476 (2022).

164. R. Xu, M. Shi, J. Li, P. Song, N. Li, Construction of SARS-CoV-2 virus-like particles by mammalian expression system. *Front Bioeng Biotechnol* **8**, 862 (2020).

165. H. Swann *et al*., Minimal system for assembly of SARS-CoV-2 virus like particles. *Sci Rep* **10**, 21877 (2020).

166. C. B. Plescia *et al*., SARS-CoV-2 viral budding and entry can be modeled using BSL-2 level virus-like particles. *J Biol Chem* **296**, 100103 (2021).

167. K. B. Moon *et al*., Construction of SARS-CoV-2 virus-like particles in plant. *Sci Rep* **12**, 1005 (2022).

168. F. Hemmati *et al*., Plant-derived VLP: a worthy platform to produce vaccine against SARS-CoV-2. *Biotechnol Lett* **44**, 45-57 (2022).

169. Y. Mi *et al*., Production of SARS-CoV-2 virus-like particles in insect cells. *Vaccines (Basel)* **9**, 554 (2021).

170. M. Jaron *et al*., Baculovirus-free SARS Covid-2 virus-like particle production in insect cells for rapid neutralization assessment. *Viruses* **14**, 2087 (2022).

171. A. Naskalska *et al*., Functional severe acute respiratory syndrome coronavirus 2 virus-like particles from insect cells. *Front Microbiol* **12**, 732998 (2021).

172. K. J. Hager *et al*., Efficacy and safety of a recombinant plant-based adjuvanted Covid-19 vaccine. *N Engl J Med* **386**, 2084-2096 (2022).

173. A. M. Syed *et al*., Rapid assessment of SARS-CoV-2-evolved variants using virus-like particles. *Science* **374**, 1626-1632 (2021).

174. A. M. Syed *et al*., Omicron mutations enhance infectivity and reduce antibody neutralization of SARS-CoV-2 virus-like particles. *Proc Natl Acad Sci U S A* **119**, e2200592119 (2022).

175. J. Shi *et al*., Challenges of drug development during the COVID-19 pandemic: Key considerations for clinical trial designs. *Br J Clin Pharmacol* **87**, 2170-2185 (2021).

176. M. P. Lythgoe, P. Middleton, Ongoing clinical trials for the management of the COVID-19 pandemic. *Trends Pharmacol Sci* **41**, 363-382 (2020).

177. T. U. Singh *et al*., Drug repurposing approach to fight COVID-19. *Pharmacol Rep* **72**, 1479-1508 (2020).

178. B. Shah, P. Modi, S. R. Sagar, In silico studies on therapeutic agents for COVID-19: Drug repurposing approach. *Life Sci* **252**, 117652 (2020).

179. L. Rodrigues, R. Bento Cunha, T. Vassilevskaia, M. Viveiros, C. Cunha, Drug repurposing for Covid-19: a review and a novel strategy to identify new targets and potential drug candidates. *Molecules* **27**, 2723 (2022).

180. Y. L. Ng, C. K. Salim, J. J. H. Chu, Drug repurposing for COVID-19: approaches, challenges and promising candidates. *Pharmacol Ther* **228**, 107930 (2021).

181. W. H. Self *et al*., Effect of hydroxychloroquine on clinical status at 14 days in hospitalized patients with COVID-19: a randomized clinical trial. *JAMA* **324**, 2165-2176 (2020).

182. J. H. Beigel *et al*., Remdesivir for the treatment of COVID-19 - final report. *N Engl J Med* **383**, 1813-1826 (2020).

183. B. Cao *et al*., A trial of lopinavir-ritonavir in adults hospitalized with severe COVID-19. *N Engl J Med* **382**, 1787-1799 (2020).

184. A. M. Kaufer, T. Theis, K. A. Lau, J. L. Gray, W. D. Rawlinson, Laboratory biosafety measures involving SARS-CoV-2 and the classification as a Risk Group 3 biological agent. *Pathology* **52**, 790-795 (2020).

185. K. B. Yeh *et al*., Significance of high-containment biological laboratories performing work during the COVID-19 pandemic: biosafety level-3 and -4 labs. *Front Bioeng Biotechnol* **9**, 720315 (2021).

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

Information of plasmids used in this dissertation project is summarized in **Table 2.1**. Reagent and material information is provided in **Table 2.2**.

Table 2.1 Reagents and materials used in the dissertation project

Plasmids	**Source**	**Identifier**
CoV2-Spike-D614G	Dr. Jennifer Doudna	Addgene #177960
CoV2-M-IRES-E	Dr. Jennifer Doudna	Addgene #177938
CoV2-N-P199L	Dr. Jennifer Doudna	Addgene #177949
Luc-PS9	Dr. Jennifer Doudna	Addgene #177942
T7-VEE-GFP	Dr. Steven Dowdy	Addgene # 58977
pEGFP-N1	Clontech	6085-1
pcDNA3.1(+)-VEE-GFP-PS9	In this study	N/A
pcDNA3.1(+)-VEE-GFP-PS9-ΔNdeI	In this study	N/A
pcDNA3.1(+)-VEE-Luc-PS9-ΔndeI	In this study	N/A
pcDNA3.1(+)-VEE-GFP-ΔPS9	In this study	N/A
pcDNA3.1(+)-VEE-Luc-ΔPS9	In this study	N/A

Table 2.2 Reagents and materials used in the dissertation project

Cell lines	**Source**	**Identifier**
HEK293T	Dr. Houjian Cai	N/A
Vero E6	Dr. Marco Archetti	N/A
A549	ATCC	CCL-185
ACE2-TMPRSS2-HEK293T	InvivoGen	hkb-hace2tpsa
Reagents and Materials	**Source**	**Identifier**
SacI	New England Biolabs	R3156S
EcoRI	New England Biolabs	R3101S
NotI	New England Biolabs	R3189S
XhoI	New England Biolabs	R0146S
XbaI	New England Biolabs	R0145S
PmeI	New England Biolabs	R0560S

NdeI	New England Biolabs	R0111S
T4 ligase	New England Biolabs	M0202S
Phusion Plus DNA polymerase	Thermo Scientific	F630XL
E.Z.N.A. Plasmid DNA Mini Kit	Omega Bio-tek	D6942
E.Z.N.A. Plasmid DNA Maxi Kit	Omega Bio-tek	D6922
E.Z.N.A. Gel Extraction Kit	Omega Bio-tek	D2500
Mix & Go E. coli Transformation Kit & Buffer Set	Zymo Research	T3001
DH10B Competent Cells	Thermo Scientific	EC0113
Ethidium bromide	Sigma Aldrich	E1510
Agarose	Bio-Rad	161-3102
GeneRuler	Thermo Scientific	SM0331
50xTAE buffer	Amresco	K915
LB medium pouches	MP Biomedicals	3002-075
Bacto-agar	Difco Laboratories	0140-01
Ampicillin	Fisher Bioreagents	BP1760-25
pcDNA3.1(+) vector	Invitrogen	V79020
DMEM	Corning	10-013-CV
Trypsin/EDTA	Corning	25-053-CI
Phosphate buffered saline	Gibco	10010023
Fetal bovine serum	Gibco	26140079
Penicillin Streptomycin solution	Corning	30-002-Cl
Puromycin	Invivogen	ant-pr-5
Zeocin	Invivogen	ant-zn-05
Hygromycin	Invivogen	ant-hg-5
Opti-MEM I reduced serum medium	Gibco	31985062
6-well Clear TC-treated Plate	Corning	3516
12-well Clear TC-treated Plate	Corning	3512
24-well Clear TC-treated Plate	Corning	3527
96 well Tissue Culture Plate	CELLTREAT	229196
60 mm Culture Dish	Corning	3295
100 mm TC-treated Culture Dish	Corning	430167
Polyethylenimine (PEI)	Sigma Aldrich	408727
Sucrose	Sigma Aldrich	84097
Trizol Reagent	Invitrogen	15596026
Chloroform	JT Baker	9175-02
2-Propanol	Fisher Chemical	A416-4
Ethanol	Decon	14G10C
PES syringe filter	CellTreat	229749
10 mL Syringe	BD	309604
Amicon Ultra-15	Millipore	UFC901008
Konical Tubes	Beckman Coulter	358126
DNA-free kit	Invitrogen	AM1906
Maxima first strand cDNA synthesis kit	Thermo Scientific	K1642

PerfeCTa SYBR Green reagent	Quantabio	95073
10x Fluc-Lysis buffer	Genecopoeia	LF009
Luciferase assay system	Promega	E1501
Pierce LDH Cytotoxicity reaction mixture	Thermo Scientific	88954
SARS-CoV-2 spike neutralizing antibody	SinoBiological	40591-MM43
Cepharanthine	Targetmol	T0131
Berbamine dihydrochloride	Targetmol	T2920
Fangchinoline	Targetmol	T3122
Dimethyl sulfoxide	Sigma Aldrich	D2650
Nuclease-free water	Promega	P1195
96-well assay plate	Costar	3912
Bright-Glo luciferase assay system	Promega	E2610

2.2 Instruments

Information of instruments used in this dissertation project is summarized in **Table 2.3**.

Table 2.3 Instruments used in this dissertation project

Instruments	**Source**
AE163 analytical balance	Mettler-Toledo
SL400 analytical balances	Scientech Inc
Precision™ Reciprocal Shaking Bath	Thermo Fisher Scientific
Direct-Q® 5UV-R Water purification system	Millipore
510 benchtop pH meter	OAKTON Instruments
Revolution automated microscope	Echo
NAPCO Series 8000 Water-Jacketed CO2 Incubator	Thermo Fisher Scientific
Laboratory micro burner MF-2001	Porta-Lab
Excella® E25 Incubator shaker	New Brunswick Scientific
1300 Series A2 biological safety cabinet	Thermo Fisher Scientific
2720 Thermal Cycler	Applied Biosystems
OWL EasyCast™ horizontal gel electrophoresis chamber connected with EC105 power supply	Thermo Electron Corporation
Gel Doc™ EZ Imager	Bio-Rad Laboratories
NANODROP LITE Spectrophotometer	Thermo Fisher Scientific
Optima™ L-100 XP Ultracentrifuge	Beckman Coulter
SW 32 Ti Swinging-Bucket Rotor	Beckman Coulter
Allegra® X-15R centrifuge	Beckman Coulter
Avanti® J-30I	Beckman Coulter
Eppendorf™ Centrifuge 5424	Eppendorf
Vortex-Genie 2 Variable Speed Vortex	Scientific Industries

StepOnePlus™ Real Time PCR System	Applied Biosystems
CLARIOstar plate reader	BMG Labtech

2.3 Methods

2.3.1 Plasmid construction and preparation

2.3.1.1 Construction of pcDNA3.1(+)-VEE-GFP-PS9 plasmid

As shown in **Fig 2.1**, the transfer plasmid pcDNA3.1(+)-VEE-GFP-PS9 was constructed by sequential subcloning of VEE virus replicon from T7-VEE-GFP (a gift from Dr. Steven Dowdy, Addgene plasmid # 58977) and PS9 sequence from Luc-PS9 (a gift from Dr. Jennifer Doudna, Addgene plasmid #177942) into the pcDNA3.1(+) backbone. Firstly, the first 2141-bp sequence of T7-VEE-GFP was PCR-amplified with Phusion Plus by a primer pair (5' to 3'): ACGAGCTCTACGACTCACTATAGATGGGCG and CCGGAATTCATGGAAGGGAGGATCC. The PCR condition was: (1) initial denaturation at 98°C for 30 s; (2) denaturation at 98°C for 10 s, annealing at 60°C for 30 s, extension at 72°C for 30 s, 30 cycles; (3) final extension at 72°C for 5 min and hold at 4°C. PCR fragment was digested with *SacI* and *EcoRI* for 1 hr, gel electrophoresis purified, and ligated by T4 ligase into the pcDNA3.1(+) backbone to get the pcDNA3.1(+)-VEE-5'UTR-nsp1 construct. Secondly, the T7-VEE-GFP (2136-8331 bp) was digested with *EcoRI* and *NotI* restriction enzymes and ligated into pcDNA3.1(+)-VEE-5'UTR-nsp1-4. Thirdly, the VEE 3'UTR sequence (9561-9759 bp) in T7-VEE-GFP was PCR-amplified by primer pair (5' to 3'): CGGCTCGAGGAATTGGCAAGCTGCTTACATAG and TGCTCTAGACACATTTCCCCGAAAAGTG, digested with *XhoI* and *XbaI*, and ligated into Luc-PS9 to get the construct Luc-PS9-VEE3'UTR. Lastly, the fragment containing PS9-WPRE-

VEE 3'UTR from the Luc-PS9-VEE3'UTR construct was subcloned into the pcDNA3.1(+)-VEE-5'UTR-nsp1-4 generated in the second step using *NotI* and *PmeI*.

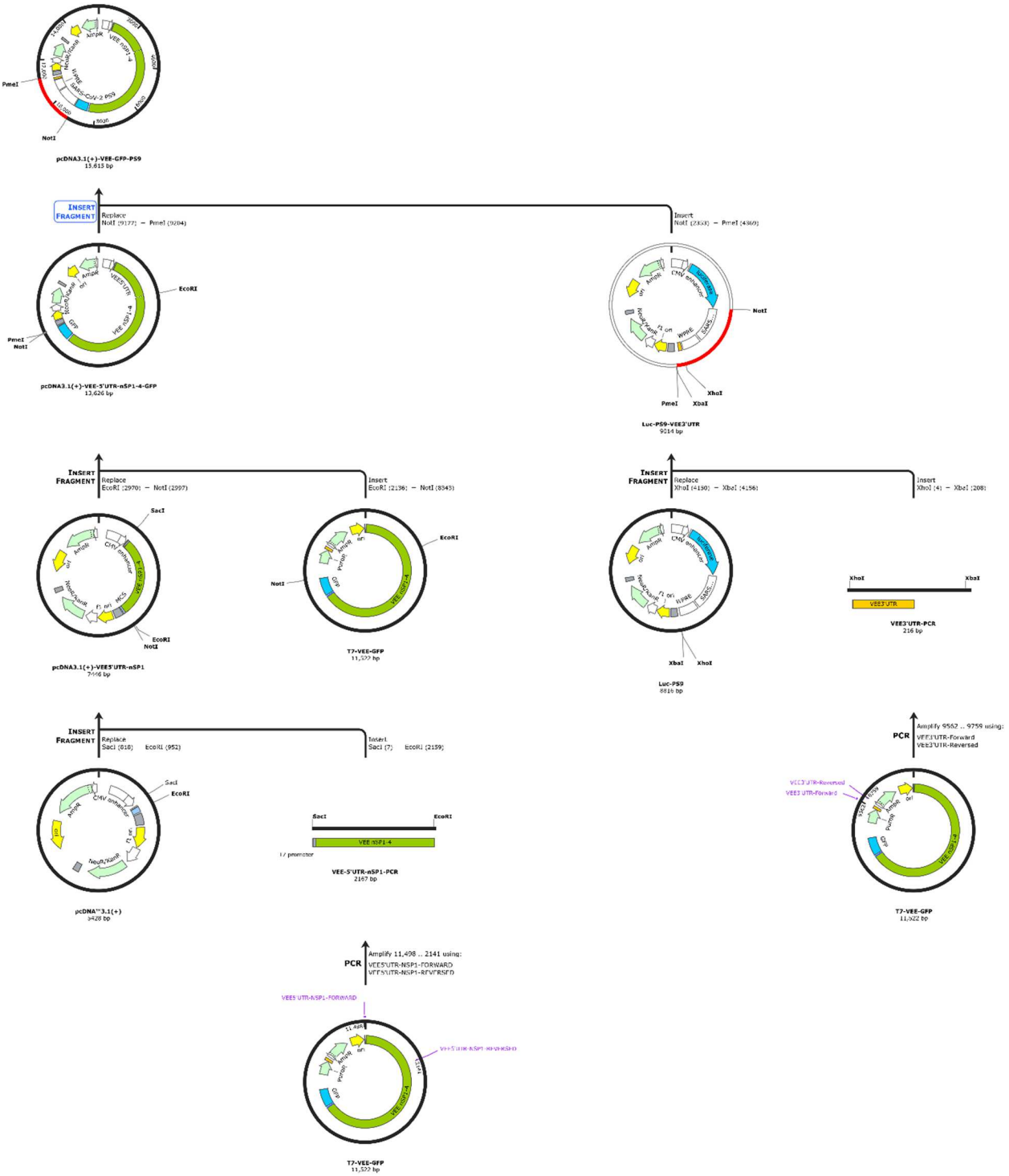


Figure 2.1 Construction procedures of pcDNA3.1(+)-VEE-GFP-PS9 plasmid

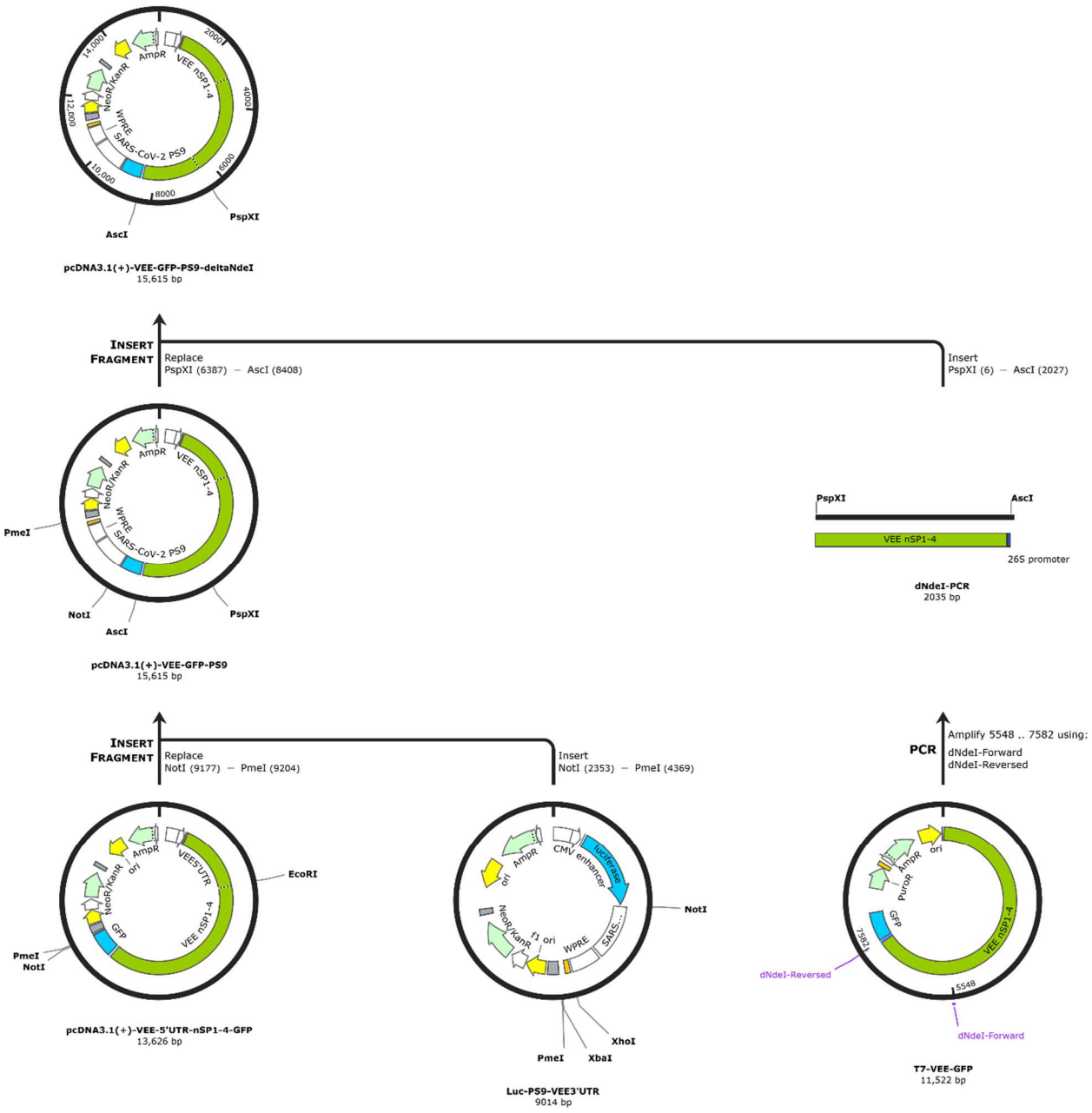


Figure 2.2 Construction procedures of pcDNA3.1(+)-VEE-GFP-PS9-ΔNdeI plasmid

2.3.1.2 Construction of pcDNA3.1(+)-VEE-GFP-PS9-ΔNdeI plasmid

To further improve protein translation efficiency, the extra start codon (ATG) in the *NdeI* site of T7-VEE-GFP template was mutated by PCR (**Fig 2.2**). A fragment (Δ*NdeI*) mutating CATAT**G** to CATAT**A** was generated by PCR using primer pairs (5' to 3'): AGGAGCTCGAGGCGCTTA

and AGGGCGCGCCTATATGCTA. The PCR condition was: (1) initial denaturation at 98°C for 30 s; (2) denaturation at 98°C for 10 s, annealing at 60°C for 30 s, extension at 72°C for 30 s, 30 cycles; (3) final extension at 72°C for 5 min and hold at 4°C. The fragment generated by PCR and the pcDNA3.1(+)-VEE-GFP-PS9 plasmid were then digested with *PspXI* and *AscI* for 1 hr, gel electrophoresis purified, and ligated into the pcDNA3.1(+)-VEE-GFP-PS9-ΔndeI by T4 ligase. Fragment generated by PCR and the pcDNA3.1(+)-VEE-GFP-PS9 plasmid were then digested with *PspXI* and *AscI* for 1 hr, gel electrophoresis purified, and ligated into the pcDNA3.1(+)-VEE-GFP-PS9-ΔndeI by T4 ligase.

2.3.1.3 Construction of pcDNA3.1(+)-VEE-Luc-PS9-ΔNdeI plasmid

cDNA encoding firefly luciferase was PCR amplified from the Luc-PS9 plasmid using primer pair (5' to 3'): AGGCGCGCCACCATGGAAGATG and TTGGCGCGCCTATCTTTATAGCCACGGAAC. The PCR condition was: (1) initial denaturation at 98°C for 30 s; (2) denaturation at 98°C for 10 s, annealing at 60°C for 30 s, extension at 72°C for 30 s, 30 cycles; (3) final extension at 72°C for 5 min and hold at 4°C. The PCR fragment encoding firefly luciferase and the pcDNA3.1(+)-VEE-GFP-PS9-ΔndeI were digested with *AscI* and *NotI* for 1 hr, gel electrophoresis purified, then ligated into pcDNA3.1(+)-VEE-Luc-PS9-ΔndeI by T4 ligase (**Fig. 2.3**).

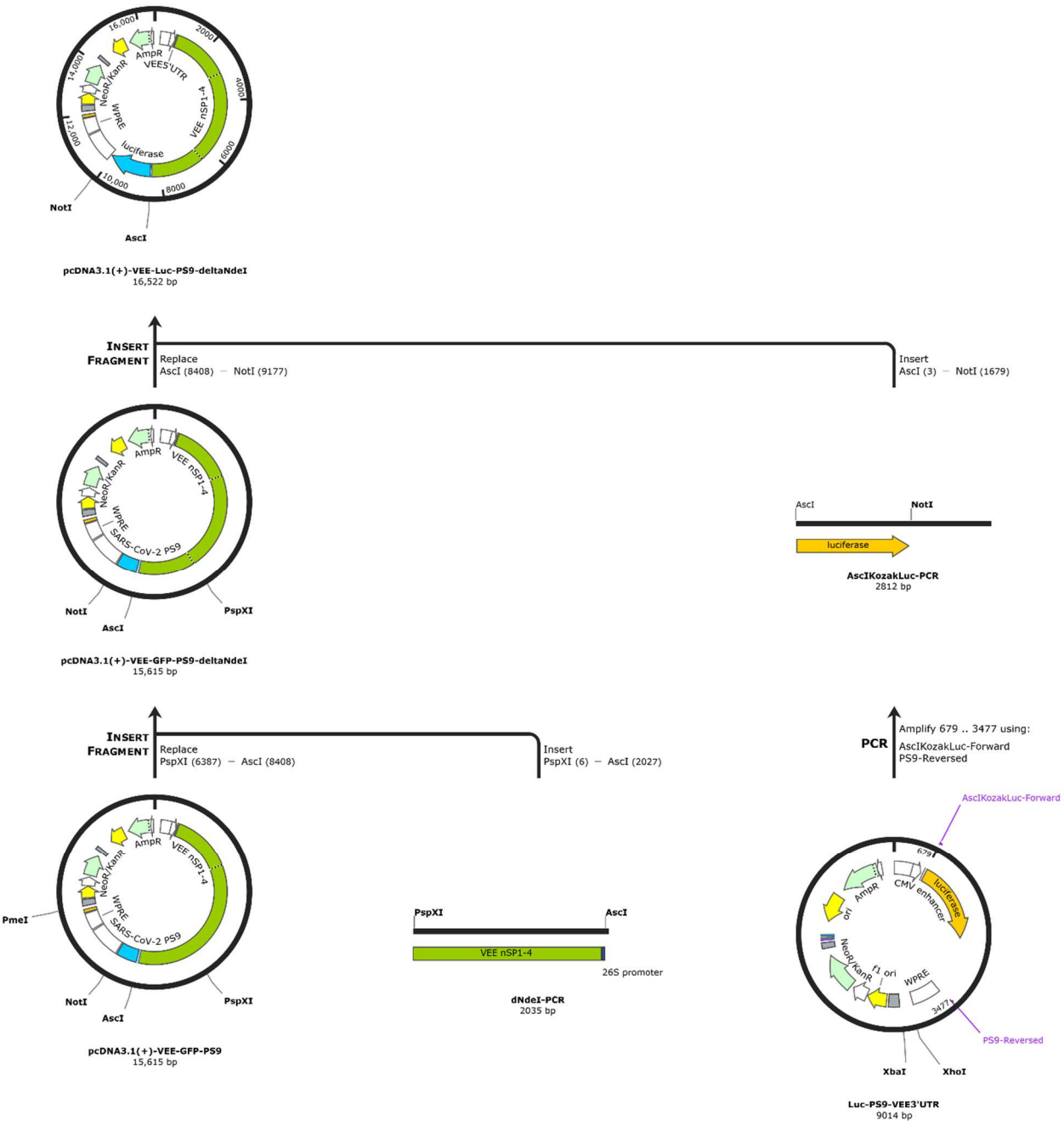


Figure 2.3 Construction procedures of pcDNA3.1(+)-VEE-Luc-PS9-ΔNdeI plasmid

2.3.1.4 Construction of pcDNA3.1(+)-VEE-Luc-ΔPS9 plasmid

As shown in **Fig. 2.4**, the PS9 sequence from the pcDNA3.1-Luc-PS9-VEE3'UTR construct obtained above was removed by digestion using *NotI* and *BamHI* and replaced by a short,

scrambled linker sequence generated by annealing of 2 oligonucleotides GGCCGCTTGTAGAAGAAACACATGGCGCGCCTGAG and GATCCTCAGGCGCGCCATGTGTTTCTTCTACAAGC. The annealing reaction condition was: (1) at 37°C for 30 min; (2) at 95°C for 5 min; (3) decreased to 25°C at the rate of 0.1°C/s. The annealed linker containing blunt overhangs complimentary with *NotI* and *BamHI* cutting sites was ligated into the digested pcDNA3.1-Luc-PS9-VEE3'UTR backbone to get the pcDNA3.1-Luc-linker-VEE3'UTR construct. Then, the Luc-linker-VEE3'UTR fragment was generated by digestion with *NotI* and *PmeI*, gel electrophoresis purified, and ligated into the digested backbone of pcDNA3.1(+)-VEE-Luc-PS9-ΔndeI by T4 ligation.

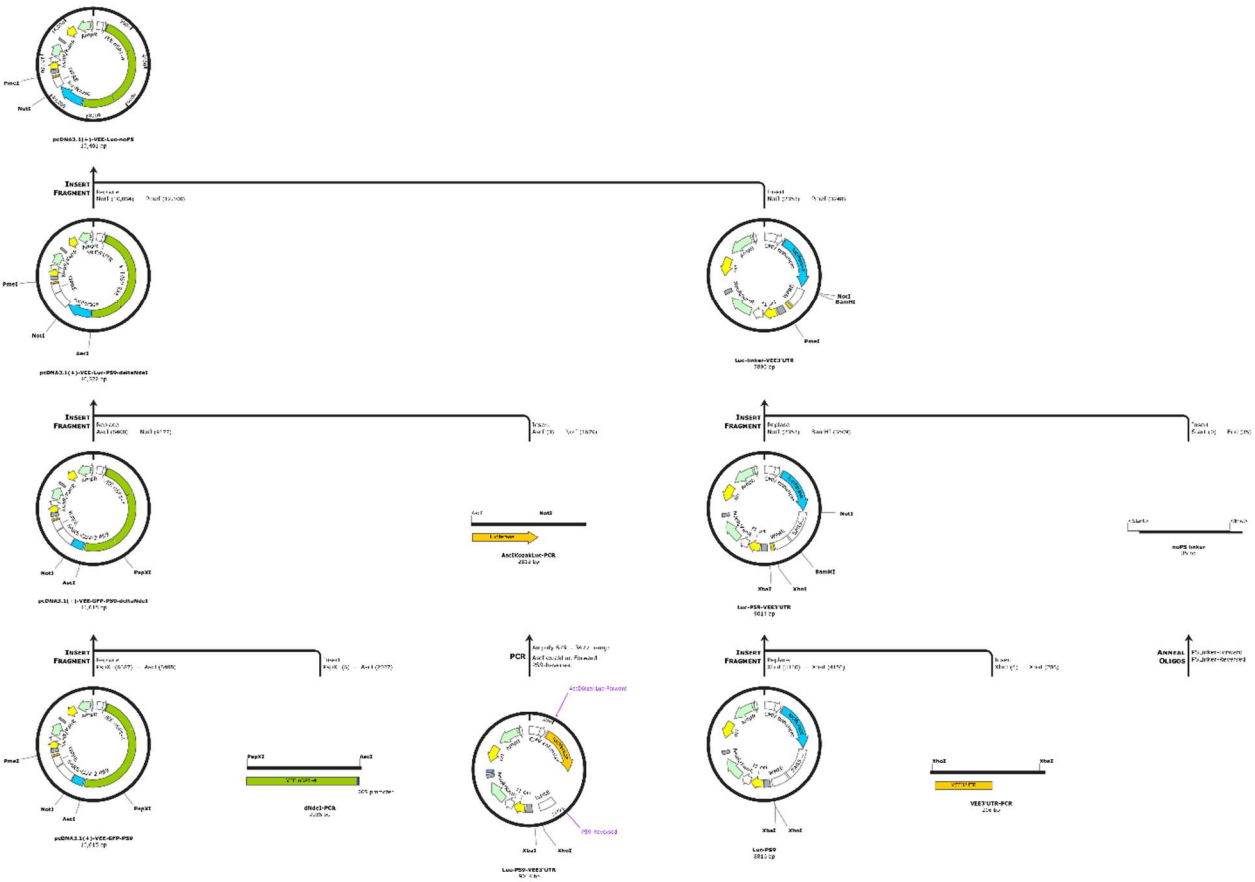


Figure 2.4 Construction procedures of pcDNA3.1(+)-VEE-Luc-ΔPS9 plasmid

2.3.1.5 Construction of pcDNA3.1(+)-VEE-GFP-ΔPS9 plasmid

The pcDNA3.1(+)-VEE-GFP-ΔPS9 plasmid was generated by digesting the Luc-linker-VEE3'UTR plasmid and the pcDNA3.1(+)-VEE-GFP-PS9-ΔndeI plasmid with *NotI* and *PmeI*. The linker-VEE3'UTR fragment and the pcDNA3.1(+)-VEE-GFP-PS9-ΔndeI backbone were gel electrophoresis purified and ligated into the pcDNA3.1(+)-VEE-GFP-ΔPS9 plasmid (**Fig 2.5**).

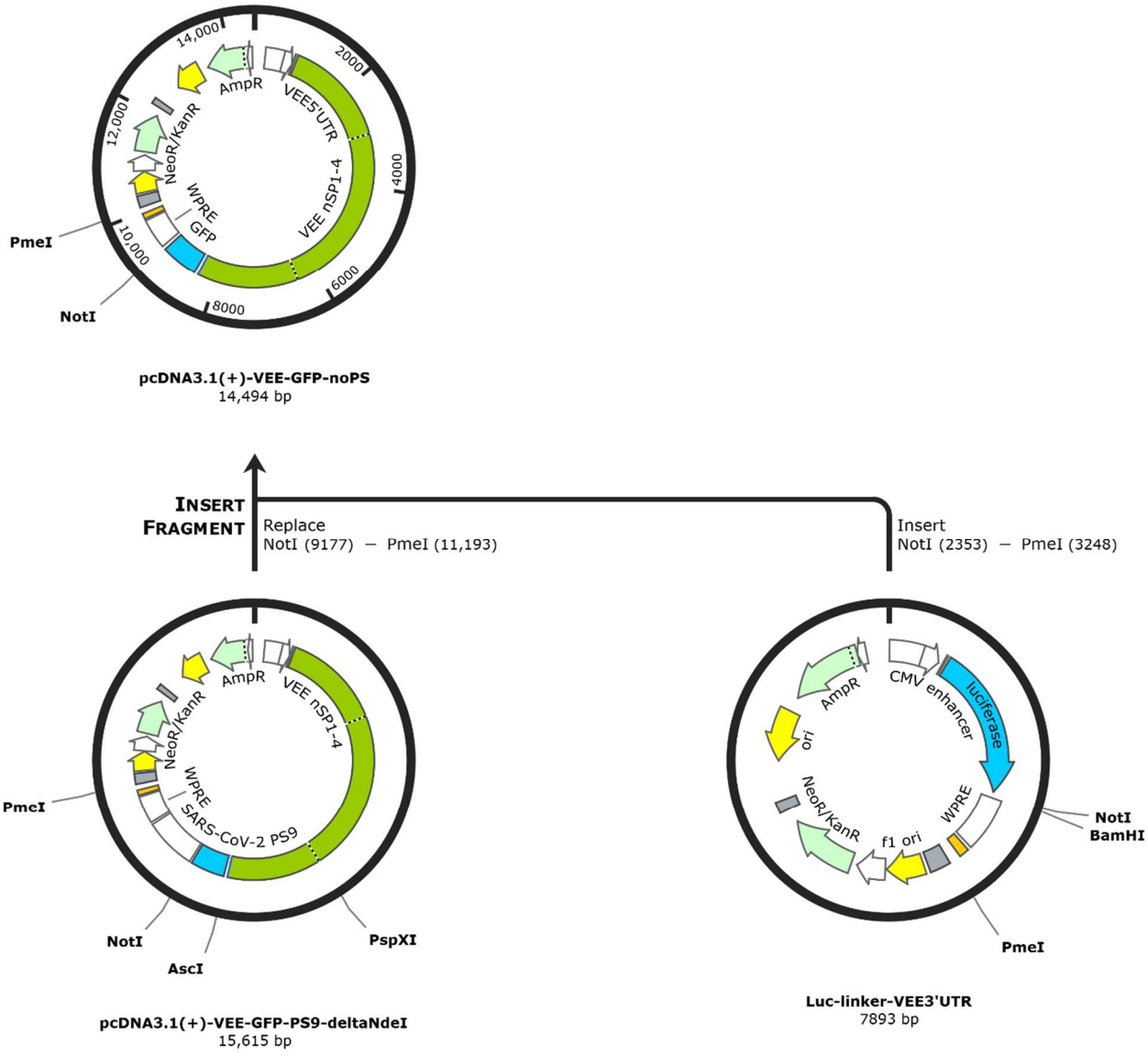


Figure 2.5 Construction procedures of pcDNA3.1(+)-VEE-GFP-ΔPS9 plasmid

The DNA sequences of key fragments constructed were confirmed by sanger sequencing and restriction digestion followed by DNA gel electrophoresis. Full sequences of constructed plasmids are listed in **Appendices**.

2.3.2 Cell culture

HEK293T, Vero E6, and A549 cells were cultured in DMEM complete medium (4.5 g/L glucose, L-glutamine, sodium pyruvate) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin and streptomycin. ACE2-TMPRSS2-HEK293T cells were purchased from InvivoGen and cultured in DMEM complete medium (4.5 g/L glucose, L-glutamine, sodium pyruvate) with 10% FBS, 1% penicillin and streptomycin, 0.5 μg/ml Puromycin, 200 μg/ml hygromycin, and 100 μg/ml zeocin. Cells are kept in 10 cm petri dish under 5% CO_2 in a humidified incubator at 37°C. All the cell lines were passed once in every 3 days in the following procedure: firstly, cells were washed by 2 mL of pre-warmed 1x phosphate buffered saline (PBS); 1x PBS wash solution was aspirated; cells were trypsinized by 2 mL of pre-warmed trypsin/EDTA solution at 37°C for 3 min; 4 mL of pre-warmed DMEM complete medium containing 10% FBS was then added to wash cells; cell suspension was centrifuged at 200 g for 5 min at room temperature and resuspended in 2 mL pre-warmed DMEM complete medium. All cell lines were passaged at 1:10 ratio.

2.3.3 Optimization of transfection procedure in 293T cells

Two transfection procedures were performed in 6-well cell culture plates to establish a standard procedure. HEK293T cells were cultured in 1 mL of DMEM complete medium with 10% FBS, 1% penicillin and streptomycin. At 70~80% confluency, each well of 293T cells was transfected with pcDNA3.1-GFP and pcDNA3.1-VEE-Luc-PS9-ΔndeI plasmids (3.6 μg each).

The plasmid DNA solution is dissolved in Opti-MEM reduced serum medium with a total volume of 150 μL. PEI stock solution was prepared in deionized water at the concentration of 1 μg/ μL. For each well of transfection, 21.6 μL of PEI solution (DNA: PEI=1:3) was dissolved in Opti-MEM reduced serum medium with a total volume of 150 μL. The 150 μL of plasmid DNA solution was mixed dropwise with 150 μL of PEI, followed by vortex-mixing for 5 sec. The DNA-PEI mixture was then incubated for 15 min at room temperature before transfection. For transfection procedure 1, 300 μL of DNA-PEI mixture was directly added into the 293T cells cultured in 1 mL of DMEM complete culture medium. After 16-hour incubation, the culture medium containing DNA-PEI mixture was removed and replaced with 1 mL fresh DMEM complete medium with 10% FBS and 1% antibiotics. For transfection procedure 2, the DMEM complete medium in each well was removed, and cells were washed with pre-warmed 1x PBS. Then the wash solution was removed, and 300 μL of DNA-PEI mixture was added into empty wells seeded with the 293T cells. The DNA-PEI mixture was incubated with 293T cells for 6 hr and replaced with 1 mL fresh DMEM complete medium with 10% FBS and 1% antibiotics. Transfection efficiency of 2 procedures was evaluated by fluorescence microscopy and luciferase assay at the indicated time points. Luciferase activity was normalized to total protein concentration determined by BCA assay. Lysis buffer is composed of 0.1M Tris-HCl, 0.05% Triton X-100, 2mM EDTA, pH 7.8.

2.3.4 Determining transgene expression in HEK293T producer cells

At 48 hr post-transfection, the culture medium was removed, and transfected producer cells were lysed with 500 μL Trizol reagent. The total RNA was extracted according to manufacturer's instructions and treated with Dnase I. One μg of treated RNA was subjected to reverse transcription using Maximal first strand cDNA synthesis kit. 10 ng of cDNA was used as PCR template to

validate transfected gene expression in producer cells. Primer sequences can be found in **Table 2.4**. Phusion plus was used for PCR and reaction condition was: (1) initial denaturation at 98°C for 30 s; (2) denaturation at 98°C for 10 s, annealing at 60°C for 30 s, extension at 72°C for 30 s, 30 cycles; (3) final extension at 72°C for 5 min and hold at 4°C. PCR product was analyzed by gel electrophoresis.

Table 2.4 Primer sequences used for validating gene transfection in producer cells

Gene	Forward (5' to 3')	Reverse (5' to 3')
Spike Protein	AGTTGCCGCCCACTTTGC	TCATGTATAGTGCAGTTTGACGC
Envelop Protein	AGCTTCGTATCAGAAGAAACCG	TAAAGAACGAGGAGATCCGGC
Membrane Protein	ATGGCCGACTCAAATGGGAC	TTACTGGACGAGCAAAGCAATG
Nucleocapsid Protein	AGCGATAACGGCCCCCAA	TTACGCCTGAGTAGAATCGGCTG
Luciferase	CGGGATCCACTAGCGCCGCCATGGAAG	CCGCATAAGCGAATTCTTACACG

2.3.5 Preparation of VEE-SARS-CoV-2 viral vectors

HEK293T cells ($8x10^6$ cells/plate) were seeded in a 10 cm petri dish 24 hr before transfection. For transfection, a total of 19 μg of plasmid DNA containing the desirable plasmids at appropriate plasmid ratio was dissolved in Opti-MEM reduced serum medium with a total volume of 500 μL. The plasmid DNA solution was mixed dropwise with 500 μL of PEI prepared by mixing 57 μL of PEI (1 mg/mL) with 443 μL of Opti-MEM. The plasmid/PEI mixture was vortexed for 5 sec and incubated at room temperature for 15 min before being added dropwise into each of the 10 cm petri dishes of HEK293T cells containing 9 mL of DMEM medium with 10% FBS and 1% penicillin/streptomycin. The transfection mixture was kept with cells overnight and the medium was replaced with fresh DMEM medium supplemented with 10% FBS and 1%

penicillin/streptomycin. At 48 hr post-transfection, the culture medium containing viral particles was harvested and centrifuged at 4,000 g for 10 min using Allegra X-15R Centrifuge. The supernatant was harvested and filtered with a 0.45 μm syringe filter and the filtrate was stored at -80°C until use.

2.3.6 Preparation of SC2-Viral Like Particles (VLPs)

Each 10-cm plate of HEK293T cells was transfected with 19 μg of plasmid DNA (6.3 μg of CoV2-N-P199L, 3.1 μg of CoV2-M-IRES-E, 15 ng of CoV2-Spike-D614G, 9.5 μg of Luc-PS9). The plasmid DNA was dissolved in Opti-MEM reduced serum medium with a total volume of 500 μL and mixed dropwise with 500 μL of PEI prepared by mixing 57 μL of PEI (1 mg/mL) with 443 μL of Opti-MEM. The plasmid/PEI mixture was vortexed for 5 sec and incubated at room temperature for 15 min before being added dropwise into each of the 10 cm petri dishes of HEK293T cells containing 9 mL of DMEM medium with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. The transfection mixture was kept with cells overnight and the medium was replaced with fresh DMEM medium supplemented with 10% FBS and 1% penicillin/streptomycin. At 48 hr post-transfection, the culture medium containing viral particles was harvested and centrifuged at 4,000 g for 10 min using Allegra X-15R Centrifuge. The supernatant was harvested and filtered with a 0.45 μm syringe filter and the filtrate was stored at -80°C until use.

2.3.7 Purification of VEE-SARS-CoV-2 viral vectors

The method of ultracentrifugation was used for purification of viral vectors. The frozen filtrate was thawed on ice and 24 mL was loaded onto 2 mL of 20% sucrose cushion (20% w/v,

1x PBS as solvent) in a conical bottom tube. The SW32Ti rotor was used for centrifugation at 28,000 g for 2 hr at 4°C. The supernatant was discarded, and pellet of viral particles was resuspended in 150 μL of ice-cold 1x PBS solution. The viral vectors were kept on ice and usually used within 1 day after preparation.

For ultrafiltration method, the filtered supernatant was concentrated by 10 kDa molecular weight cutoff Amicon filters at 2000 g for 30 min. Viral vector concentrates were diluted in ice-cold 1x PBS solution to desired volumes. The viral vectors were kept on ice and usually used within 1 day after preparation.

2.3.8 Viral vector RNA extraction

To extract RNA from viral vectors, 25 μL of the purified viral vector solution was mixed with 500 μL of Trizol solution and incubated for 5 min. Then 100 μL chloroform was added and mixed by shaking. The sample was incubated for 3 min at room temperature and centrifuged at 12000 g at 4°C for 15 min, and 250 μL of the aqueous phase containing RNA was transferred to a new microcentrifuge tube. The aqueous phase was washed by 250 μL isopropanol and centrifuged for 10 min at 12000 g at 4°C. The supernatant was discarded, and the RNA pellet was resuspended in 500 μL 75% ethanol, vortexed, and then centrifuged for 5 min at 7500 g at 4°C. The RNA pellet was air-dried for 5 min and resuspended in 31 μL of nuclease-free water.

2.3.9 Real Time PCR-based viral vector titration

Thirty μL of the extracted viral RNA was treated with 1 μL of Dnase I (2 units/μL) at 37°C for 30 min in a 50-μL reaction, and the treatment was inactivated by adding 5 μL resuspended Dnase inactivation reagent and incubating for 2 min. The sample was then centrifuged at 10000 g

for 90 sec at room temperature, and the RNA supernatant was transferred into a new tube. Fifteen µL of treated RNA was reverse transcribed into cDNA using Maxima first strand cDNA synthesis kit in a 20-µL reaction system. The reaction condition was: (1) at 25°C for 10 min; (2) at 50°C for 30 min; (3) termination at 85°C for 5 min.

A series of standard plasmid samples were prepared to generate a titration standard curve. The copy numbers/µL for each reporter plasmid was calculated as below:

	VEE-Luc-PS9	VEE-GFP-PS9	Luc-PS9	GFP-PS9
Size (bp)	16522	15615	8816	7933
DNA concentration (µg/µL)	0.531	0.0805	0.3071	0.671
Molecular Weight	16522 bp x 650 Daltons/bp (g/mole) = 10.739 x 10^6 grams/mole	15615 bp x 650 Daltons/bp (g/mole) = 10.150 x 10^6 grams/mole	8816 bp x 650 Daltons/bp (g/mole) = 5.73 x 10^6 grams/mole	7933 bp x 650 Daltons/bp (g/mole) = 5.156 x 10^6 grams/mole
Moles/µL	0.531 µg/µL x 1g/10^6 µg x1 / 10.739x10^6 g/mole = 4.94 x10^{-14} moles/ µL	0.0805 µg/µL x 1g/10^6 µg x1 / 10.150x10^6 g/mole = 7.93 x10^{-15} moles/ µL	5.36 x10^{-14} moles/ µL	1.30 x10^{-13} moles/ µL
Molecules/ µL	4.94 x10^{-14} moles/ µL x 6.022145 x 10^{23} molecules/mole = 2.97x 10^{10} molecules/µL	7.93 x10-15 moles/ µL x 6.022145 x 10^{23} molecules/mole = 4.78x 10^9 molecules/µL	3.23 x 10^{10} molecules /µL	7.83 x 10^{10} molecules /µL

To obtain a 100-µL solution at 10^9 molecules/µL template DNA: VEE-Luc-PS9: 10^{11}/2.97 x 10^{10} = 3.37 µL plasmid solution into 96.63 µL nuclease-free water; VEE-GFP-PS9: 10^{11}/4.78 x 10^9 = 20.9 µL plasmid solution into 79.1 µL nuclease-free water; Luc-PS9: 10^{11}/3.23 x 10^{10} = 3.10 µL plasmid solution into 96.9 µL nuclease-free water; GFP-PS9: 10^{11}/7.83 x 10^{10} = 1.28 µL plasmid preparation into 98.72 µL nuclease-free water. 25 µL of 10^9 molecules/L template DNA

solution was then diluted into 75 µL nuclease-free water to obtain standard samples at $2.5x10^8$ molecules/µL. A 10-fold serial dilution was performed as such till $2.5x10^2$ molecules/µL.

Real Time PCR-based titration was conducted in a 10-µL reaction for each well: 6 µL premix of SYBR green (5 µL) and primer pair (1 µL, final concentration at 500 nM) with 4 µL of diluted DNA standard or extracted viral cDNA. The primer pair (5' to 3'): GTGGTGTGCAGCGAGAATAG and CGCTCGTTGTAGATGTCGTTAG targeting luciferase reporter sequence was used. PCR condition was: (1) 95°C for 10 min; (2) (95°C for 15 sec; 60°C:1 min) x 40 cycles; (3) 95°C:15 sec, 60°C:1 min, 95°C:15 sec. Real time qPCR was performed in a StepOne Plus Real Time PCR system.

2.3.10 Gene expression analysis by Real Time PCR

To determine the RNA level of luciferase reporter in transduced cells in a 12-well plate, the culture medium was removed at the desirable times after transduction and the cells were washed once with PBS. The cells were lysed with 500 µL Trizol reagent, and the total RNA was extracted according to manufacturer's instruction. The RNA concentration was determined, and 1 µg of RNA was subjected to reverse transcription using Maximal first strand cDNA synthesis kit. The cDNA was then amplified by qPCR, and GAPDH gene served as the internal control. Luciferase gene primers were the same as above. GAPDH primer sequences were (5' to 3'): GAAGGTGAAGGTCGGAGTCAAC and CAGAGTTAAAAGCAGCCCTGGT. Real time qPCR was performed in a StepOne Plus Real Time PCR system using PerfeCTa® SYBR® Green reagent under the same condition as described above.

2.3.11 Luciferase assay

A total of 10^5 ACE2-TMPRSS2-293T cells were seeded in each well of a 12-well plate. After 48 hr, the attached cells were transduced with VEE-SARS-CoV-2 viral vectors or SC2-VLPs when needed. The transduction proceeded until the desirable time for analysis of gene expression. Cell culture medium was aspirated, and cells were rinsed with 1x PBS. The cells were lysed in 100 µL 1x Fluc-Lysis buffer with 5-min shaking at room temperature. Cell lysates (20 µL) were transferred into a 96-well microplate and mixed with 100 µL of reconstituted luciferase reagent. Luciferase activity was immediately measured using CLARIOstar plate reader.

2.3.12 Cytotoxicity assay

ACE2-TMPRSS2-293T cells (2.5 x10^4 cells/well) in 100 µL DMEM complete medium were seeded in a 96-well plate for 24 hr and transduced with VEE-SARS-CoV-2 vectors at desired multiplicity of infection (MOI). Forty-eight hr post-transduction, 50 µL of cell culture medium of transduced cells was transferred into a new 96-well plate and mixed with 50 µL of Pierce LDH Cytotoxicity reaction mixture at room temperature for 30 min. The reaction was terminated by addition of 50 µL of stop solution. The absorbance was measured at 490 nm and 680 nm, respectively. For positive control, the culture medium of the untreated cells was removed and replaced with the same volume of lysis buffer. Cells were incubated for 5 min to complete cell lysis and 50 µL of cell lysis was used for spectrometric measurement.

2.3.13 Antibody neutralization assay

A total of 2.5 x10^4 ACE2-TMPRSS2-293T cells in 100 µL of DMEM complete medium were seeded in each well of 96-well plate for 24 hr. VEE-SARS-CoV-2-Luc viral vectors were

incubated with SARS-CoV-2 spike neutralizing antibody (40591-MM43, SinoBiological) at 37°C for 1 hr. The vector-antibody mixture was added to pre-seeded cells and incubated for 48 hr. The level of reporter gene expression was determined using the luciferase assay described above.

2.3.14 Inhibition of VEE-SARS-CoV-2-Luc vector-mediated transduction by anti-viral agents

Cepharanthine, Berbamine dihydrochloride, and Fangchinoline were purchased from TargetMol and dissolved in dimethyl sulfoxide with final concentration of 10 mM. ACE2-TMPRSS2-293T cells ($2.5x10^4$ cells/well) in 100 μL DMEM complete medium were seeded in a 96-well plate. Twenty-four hr later, VEE-SARS-CoV-2-Luc viral vectors pre-incubated for 1 hr at 37°C with each compound at desirable concentrations were added to the pre-seeded cells. The cells were cultured for 3 hr, then the mixture was removed and replaced with fresh DMEM complete medium. Luciferase assay was performed 48 hr post-transduction.

At 48 hr post-treatment, cytotoxicity assay was performed using Pierce LDH Cytotoxicity reaction mixture as described above to determine the cytotoxicity of anti-viral compounds.

2.3.15 Fluorescence microscopy

Time-dependent gene expression induced by VEE-SARS-CoV-2 vectors and SC2-VLPs were documented by monitoring GFP expression using Revolution automated microscope.

2.3.16 Statistical analysis

Transduction and luciferase assays were performed in triplicates in all experiments and RLU value was presented as mean ± standard deviation. Statistical analysis was performed using

one -way ANOVA with Tukey Post Hoc tests by GraphPad Prism 8. GraphPad Prism 8 was also used to plot all the figures and calculate IC_{50} values for assessing neutralizing antibody and antiviral drugs.

CHAPTER 3

RESULTS

3.1 Optimizing PEI transfection procedures in 293T cells

Two different transfection procedures were performed to introduce pcDNA3.1-GFP and pcDNA3.1-VEE-Luc-PS9-ΔndeI plasmids into 293T cells. The major difference between two procedures is the length of PEI-DNA mixture incubation with cells. The transfection efficiency of two procedures was compared by GFP expression and luciferase activity. The fluorescence microscopy results revealed that an overnight incubation significantly increased transfection efficiency compared to 6-hr incubation procedure at both 24 hr and 48 hr post-transfection (**Fig. 3.1A**). Consistently, the luciferase assay data show that luciferase activity was more than 3-fold higher when cells were transfected overnight (**Fig. 3.1B**). Therefore, a transfection procedure with 16-hr incubation was adopted for producing SARS-CoV-2 viral vectors in 293T cells.

3.2 Production of VEE-SARS-CoV-2 viral vectors in producing cells and mechanism of transgene amplification in target cells

A functional packaging signal regulating the packaging of RNA into the viral particle is the critical element of building a SARS-CoV-2 viral vector. Thanks to Syed et al., a putative packaging signal named PS9 sequence (nt 20080-21171) was identified [1]. Although harboring the functional PS9 packaging sequence, the SC2-VLP system fails to generate a high transgene expression level, which could attribute from the rapid degradation of RNA molecules and the lack

of replication or amplification mechanism. Therefore, certain modifications are needed to build a more efficient SARS-CoV-2 vector system.

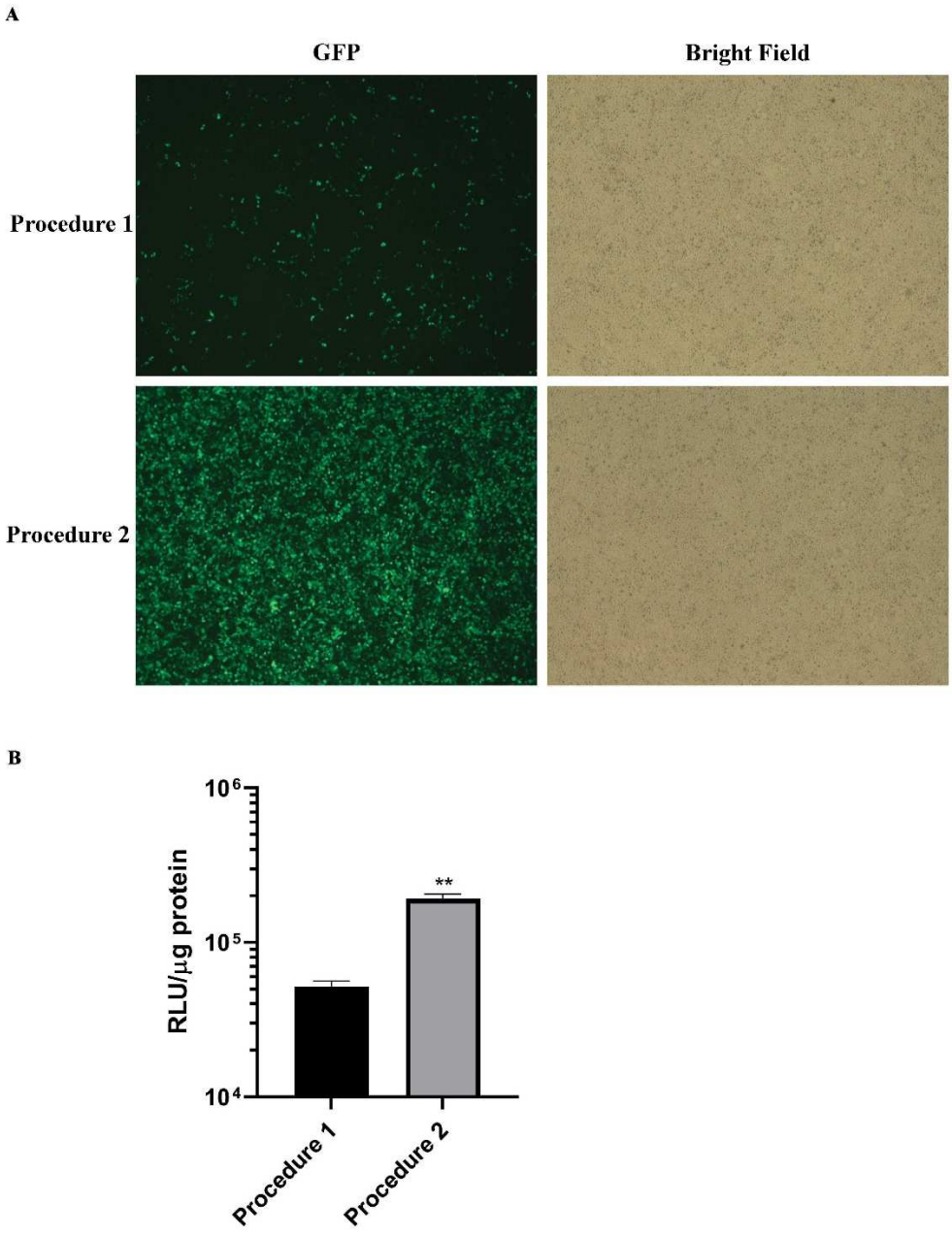


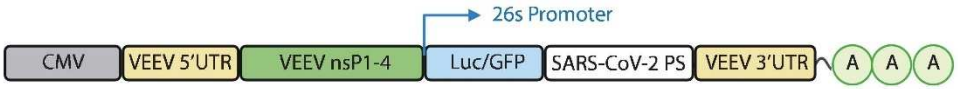
Figure 3.1 Optimizing PEI transfection procedures in 293T cells. (A) Comparison of GFP expression at 48 hr post-transfection and **(B)** Comparison of luciferase activity normalized to total protein at 48 hr post-transfection.

To overcome the limitation of low gene delivery efficiency, a self-amplifying RNA mechanism derived from VEEV is incorporated to build the novel SARS-CoV-2 vector system. Like SARS-CoV-2, VEEV is also a single-stranded positive-sense RNA virus that replicates in the cytoplasm and does not enter the host nucleus. Once entering the host, the VEEV RNA genome can be readily translated into non-structural proteins 1–4 to form RNA-dependent RNA polymerase (RdRP) complex, which in turn synthesizes the negative-stranded full length genomic RNA [2]. The negative-stranded genomic RNA can serve as the template for the synthesis of the positive-stranded full length RNA genome or the positive-stranded subgenomic RNAs under the control of 26S promoter. The VEEV-derived self-amplification mechanism was employed because it is a single positive-stranded RNA without activity of genome-integrating [3]; it can generate high copy number of transgene and extended gene expression [4, 5]; and its safety profile has been demonstrated by ongoing clinical trials for vaccine development [6, 7].

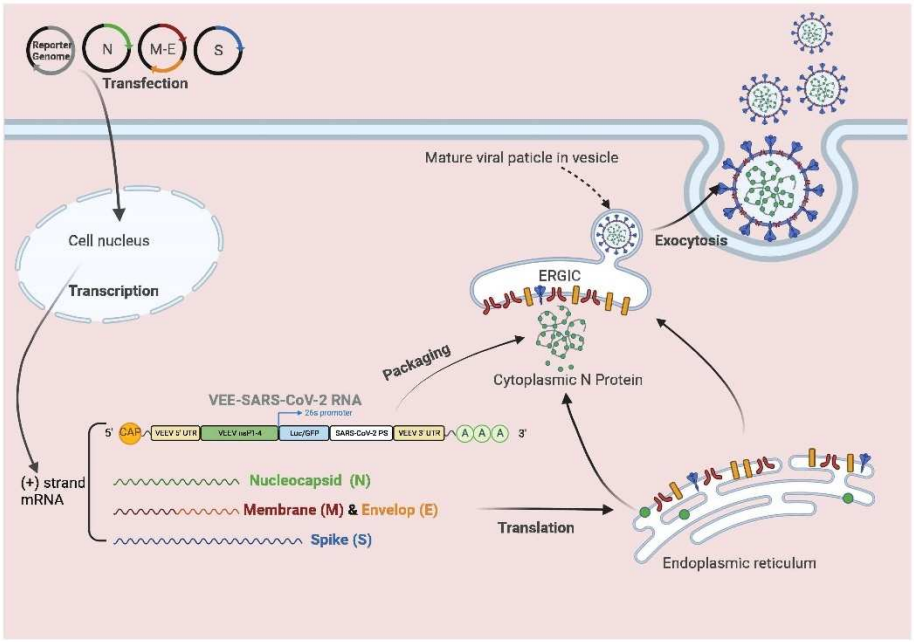
pcDNA3.1 plasmid vector was employed for construction of the reporter-carrying transfer plasmid (named VEE-Luc/GFP-PS9). As shown in **Fig. 3.2A**, the transfer plasmid has a cytomegalovirus (CMV) promoter, the VEEV 5' untranslated region (UTR), an open reading frame (ORF) of VEEV non-structural proteins (nsP) 1-4 encoding for VEEV RNA replication machinery, a 26 subgenomic promoter to drive downstream gene expression, the SARS-CoV-2 packaging signal sequence (PS9) [1], and the VEEV 3' UTR with a poly A tail for RNA stability. Other plasmids carrying SARS-CoV-2 structural protein genes were obtained from Addgene.

Fig 3.2B is a proposed mechanism for generating VEE-SARS-CoV-2 viral vectors in producer cells. The transfer plasmid is co-transfected into HEK293T cells with other 3 plasmids carrying each of the 4 structural protein genes of SARS-CoV-2 (S, M-E, N), respectively. After transfection, all SARS-CoV-2 structural proteins are produced and translocated into endoplasmic-

A


26s Promoter
CMV
VEEV 5'UTR
VEEV nsP1-4
Luc/GFP
SARS-CoV-2 PS
VEEV 3'UTR
A
A
A

B


Reporter Genome
N
M-E
S
Transfection
Mature viral paticle in vesicle
Cell nucleus
Transcription
Exocytosis
ERGIC
Packaging
Cytoplasmic N Protein
VEE-SARS-CoV-2 RNA
26s promoter
5'
CAP
VEEV 5' UTR
VEEV nsP1-4
Luc/GFP
SARS-CoV-2 PS
VEEV 3' UTR
A
A
A
3'
(+) strand mRNA
Nucleocapsid (N)
Membrane (M) & Envelop (E)
Spike (S)
Translation
Endoplasmic reticulum

C

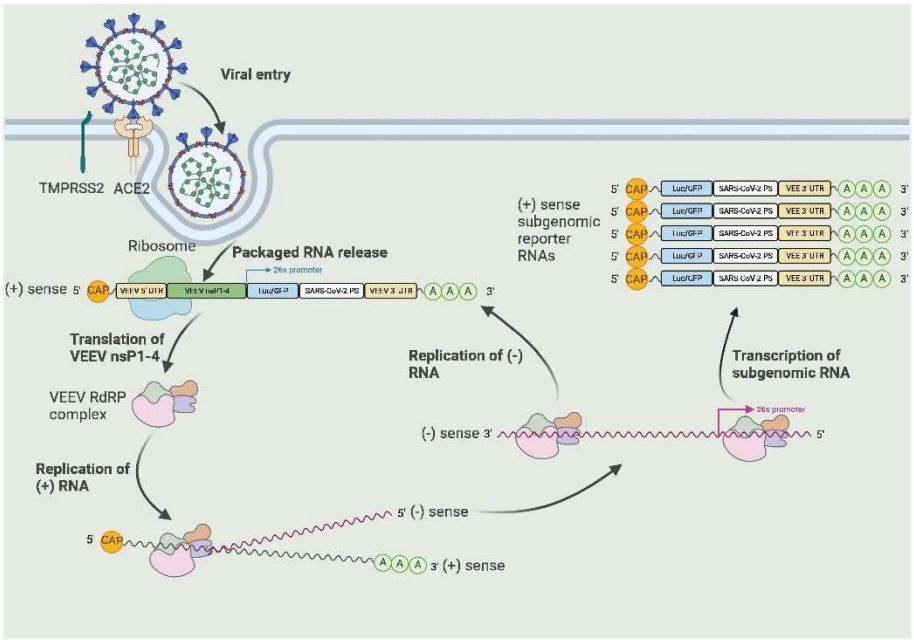

Viral entry
TMPRSS2
ACE2
Ribosome
Packaged RNA release
(+) sense 5'
CAP
26s promoter
3'
(+) sense subgenomic reporter RNAs
Translation of VEEV nsP1-4
VEEV RdRP complex
Replication of (+) RNA
5' (-) sense
(+) sense
Replication of (-) RNA
(-) sense 3'
Transcription of subgenomic RNA

Figure 3.2 Production of VEE-SARS-CoV-2 viral vectors in producing cells and mechanism of transgene amplification in target cells. (A) Schematic diagram of the core elements in VEE-SARS-CoV-2 plasmid including a CMV promoter, 5' UTR and nsP1-4 gene of VEEV, 26 subgenomic VEEV promoter (26S), reporter genes (Luc or GFP), the SARS-CoV-2 packaging signal (PS9 sequence), 3' UTR of VEEV genome, and a poly(A) signal. **(B)** Mechanistic illustration of production of VEE-SARS-CoV-2 viral vectors. Components of viral vectors were produced in 293T cells by co-transfection with plasmids carrying structural protein genes (S, M-E, N) and reporter gene with regulatory elements for amplification and expression. The viral vectors are assembled and released from 293T cells. **(C)** Schematic presentation of transduction of viral vectors. RNA amplification and reporter gene expression are accomplished in target cells via ACE2-mediated internalization, translation of VEEV nsP1-4 genes, formation of VEEV nsP1-4 protein complex, and synthesis of the negative-strand RNA. The newly synthesized negative-strand RNA serves as the template for synthesis of positive strand RNA and production of sub-genomic RNA encoding reporter gene.

Reticulum-Golgi intermediate compartment (ERGIC) to be assembled into viral particles. Meanwhile, the positive-stranded artificial RNA genome transcribed from the transfer plasmid is selectively packaged into the viral vectors with the help of N protein and PS9 package signal sequence in the RNA transcript. The mature viral particles are then released into the cell culture media via exocytosis.

Fig 3.2C is a proposed mechanism for VEE-SARS-CoV-2 viral vector-mediated gene delivery and gene expression in target cells. The entry of VEE-SARS-CoV-2 vectors into target cells is initiated by the biding of S protein of VEE-SARS-CoV-2 vectors to ACE2 receptor, primed by TMPRSS2, followed by the release of VEE-SARS-CoV-2 RNA into target cells. The ORF encoding VEEV nsP1-4 on the 5' end of the positive-sense RNA can be readily translated into proteins to form the VEEV RdRP complex. Subsequently, the RdRP complex synthesizes the negative strand of the introduced RNA. The newly synthesized negative-strand RNA serves as the template for two events: synthesis of the positive-strand RNA, and transcription of the subgenomic RNA encoding the reporter gene, which is driven by the 26 subgenomic (26S) promoter on the negative strand of VEE-SARS-CoV-2 RNA. The transcription process is highly efficient and the

copy number of reporter-coding subgenomic RNAs is at least 10-fold higher than that of the full-length VEE-SARS-CoV-2 RNA [8]. Notably, the SARS-CoV-2 structural protein genes are encoded by separate plasmids not by the VEE-SARS-CoV-2 RNA. Therefore, this self-amplifying RNA vector transduces target cells in a single-round fashion without synthesizing new progeny of virions.

3.3 Validation of gene transfection in producer cells

Reverse transcription PCR (RT-PCR) was performed for mRNA extracted from 293T cells transfected with 4 plasmids (S, M-E, N, pcDNA3.1-VEE-Luc-PS9) to validate the plasmid construction and transfected gene expression. PCR product size was analyzed by DNA gel electrophoresis. As shown in **Fig. 3.3**, all 4 structural protein genes were amplified by RT-PCR. The DNA band size of E, M, and N genes was consistent with the gene size. Since spike protein gene length is more than 3000 bp which is too long to be amplified by PCR, a primer pair was designed to amplify a fragment of the spike protein gene. In addition, luciferase reporter gene was also detected by gel electrophoresis at correct size. cDNA obtained from naïve HEK293T cells was used as negative control, no band was observed on the DNA gel.

3.4 Titration of VEE-SARS-CoV-2 viral vectors

Real Time PCR-based viral vector titration procedure was established to quantitively determine the titer of viral vectors generated from each condition or preparation. To test the reproducibility of the titration procedure, three equal aliquots of the VEE-SARS-CoV-2-Luc viral vectors (25 μL/each) were titrated independently. The results show that the standard curve (ranging from 10^4 to 10^8 copies) generated by a 10x serial dilution of standard plasmid is linear (R^2=0.9952) (**Fig.**

3.4). The titer of three independent aliquots was close, and the average titer was $4.27x10^6$ copies/ 25 μL. These data suggest that this Real Time PCR-based titration procedure is robust and generated the same titration results for the same batch of viral vector preparations.

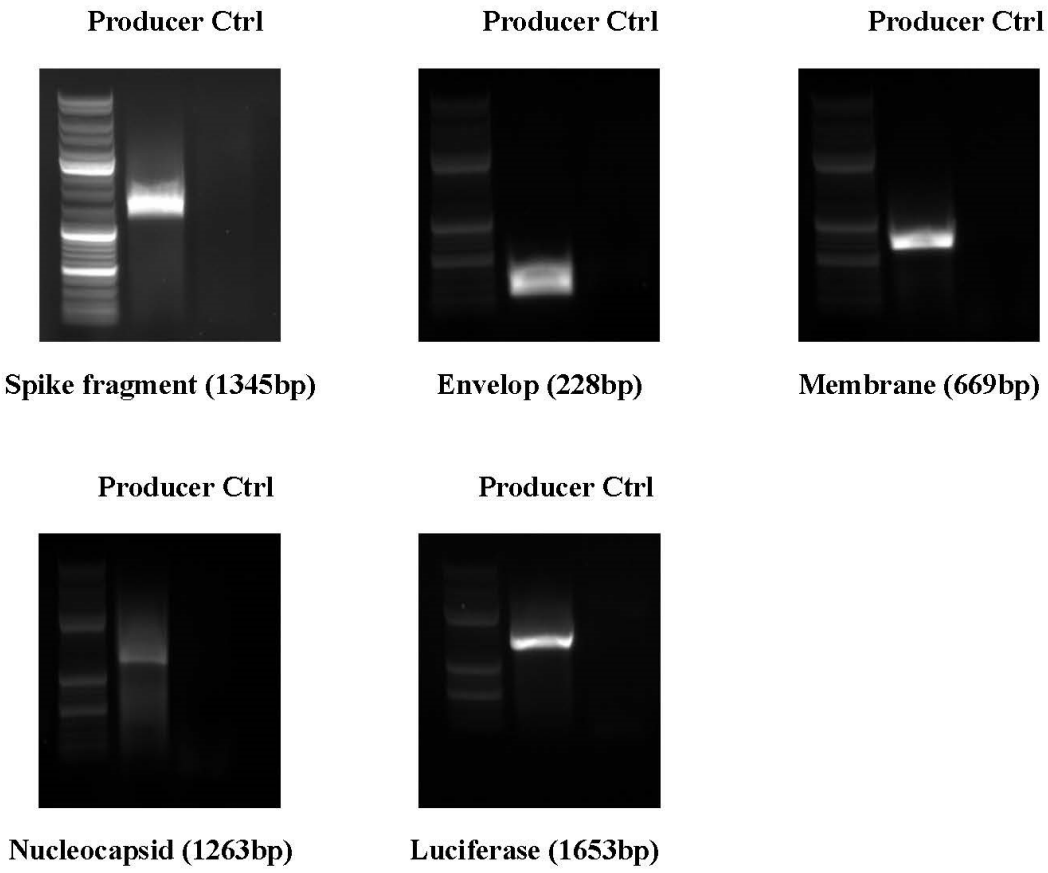
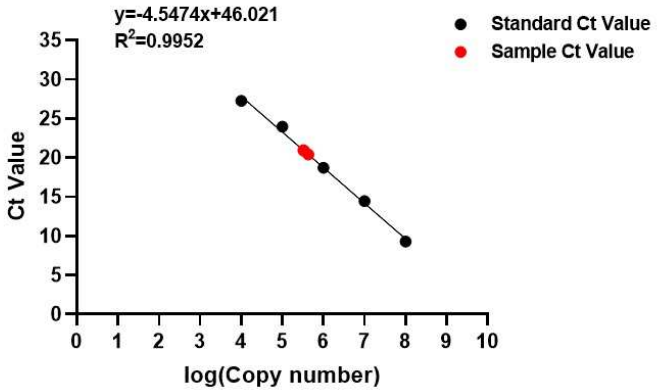


Figure 3.3 Validation of gene transfection in producer cells by RT-PCR. Five transfected genes (Spike fragment, Envelop, Membrane, Nucleocapsid, Luciferase) were PCR amplified from cDNA of HEK293T cells transfected with 4 plasmids (S, M-E, N, VEE-Luc-PS9). Naïve HEK293T cells serve as the negative control.

Standard Curve	
Log copy number	Ct Value
8	**9.290065765**
7	**14.45486927**
6	**18.6999054**
5	**23.97456551**
4	**27.26705551**
3	30.25265884
NTC	30.97834206



Sample	Ct Value	log(copy number)	Titer of 25 µL solution (copy number)	Average Titer of 25 µL solution (copy number)
VEE-SARS-CoV2-1-1	20.9469	5.513941742	3.81E+06	
VEE-SARS-CoV2-1-2	20.4038	5.633373963	5.02E+06	
VEE-SARS-CoV2-1-3	20.85771	5.533555464	3.99E+06	4.27E+06

Figure 3.4 Real Time qPCR-based titration of VEE-SARS-CoV-2 viral vectors. A standard curve was generated by a 10x serial dilution of the plasmid template. The titration procedure was verified by determining the copies of three aliquots from the same batch of viral vector preparation.

3.5 Optimization of VEE-SARS-CoV-2 viral vector production and transduction

To optimize the yield and transduction activity of VEE-SARS-CoV-2 vectors, we first focused on the ratio between the plasmid carrying spike protein gene and the total of plasmids used in transfection. The ratio impact on the yield of viral vectors was quantified by qPCR-based titration of purified viral vectors and luciferase reporter assay in transduced cells. Over a range of 3.8 ng to 3.8 µg of a total 19 µg plasmids used for transfection of 293T cells, 380 ng (2% of total plasmids) of plasmid with spike protein gene generated the highest luciferase activity in ACE2-TMPRSS2-293T cells. The qPCR-based titration assay also showed the highest level of viral vectors at 380 ng of S plasmid in a total of 19 ug plasmids (2%) (**Fig. 3.5A**). Next, we fixed the amount of plasmid with spike protein gene at 380 ng and optimized the ratio of other 3 plasmids (M-E, N, and VEE-Luc-PS9). The results showed that the plasmid mass ratio (S: M-E: N: VEE-

Luc-PS9) at 0.12:2:2:2 was optimal (**Fig. 3.5B**). These data suggest that the SARS-CoV-2 viral vectors require much lower amount of spike protein with respect to other structural proteins, which is consistent with the finding of a previous study [1].

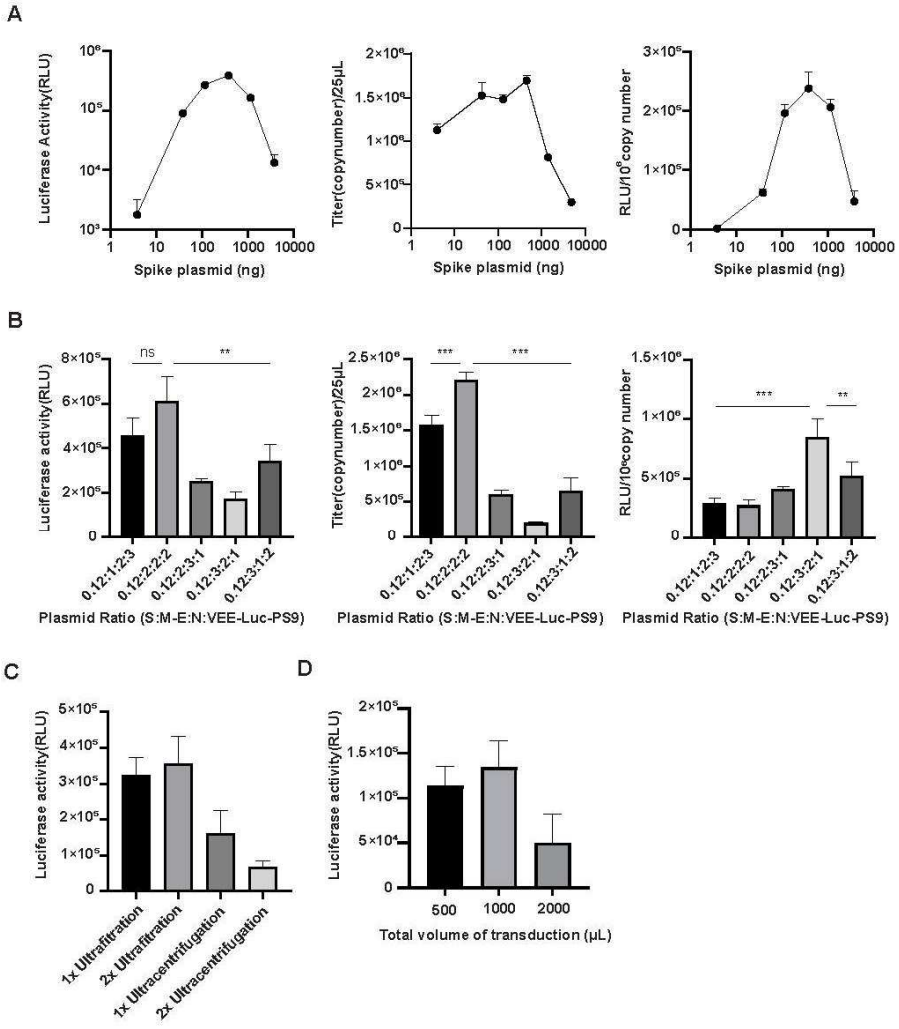


Figure 3.5 Optimization of VEE-SARS-CoV-2 viral vector production and transduction. (A) Optimization of the ratio between the plasmid carrying spike protein gene and the rest of plasmids including those carrying M-E, N, and the transfer plasmid for transduction activity and viral vector titer. **(B)** Optimization of transduction activity and titer of viral vectors by varying the mass ratio of each of the plasmids at fixed amount of plasmids carrying spike protein gene. **(C)** Relative yield of VEE-SARS-CoV-2 viral vectors by ultracentrifugation and ultrafiltration of the same amount

of culture medium of transfected 293T cells. Transduction activity of purified VEE-SARS-CoV-2-Luc vectors was assessed in ACE2-TMPRSS2-293T cells. **(D)** Effect of volume in transduction of ACE2-TMPRSS2-293T cells by the same amount of VEE-SARS-CoV-2-Luc viral vectors. Luciferase activity was measured at 24 hr post-transduction. Data are presented as the mean ± SD of 3 independent experiments. Statistical significance was determined by one-way ANOVA with Tukey Post Hoc tests. ** $p < 0.01$, *** $p < 0.001$.

Two methods were used to compare the purification yield of viral vectors. Significantly higher yield was seen with the method of ultrafiltration comparing to a ~50% purification yield for the method of ultracentrifugation (**Fig. 3.5C**). Further studies revealed high level of serum protein contamination in purified viral vectors by ultrafiltration (data not shown) and the method was not adopted for the following experiments.

The volume Impact on transduction efficiency of purified VEE-SARS-CoV-2 vectors was assessed to optimize the transduction efficiency of target cells. Results in **Fig. 3.5D** show the optimal transduction when 1000 μL of volume was used per well of cells in a 12-well plate comparing to the volume of 500 and 2000 μL. Taken together, these experiments established an optimal procedure for producing VEE-SARS-CoV-2 vectors and their transduction into target cells.

3.6 Characterization of VEE-SARS-CoV-2 viral vectors

The role of each functional components in VEE-SARS-CoV-2 vector system was examined using a series of viral vectors generated using minus one approach. HEK293T cells were transfected with plasmids withdrawing one of the 4 plasmids carrying structural protein genes or the PS9 packaging signal sequence. Viral vectors were purified separately and transduced into ACE2-TMPRSS2-293T cells. The luciferase assay data show that the viral vectors completely lost their transduction activity when missing any plasmids carrying the structural protein gene (**Fig. 3.6A**). Notably, ΔPS9 viral vectors showed no transducing activity and extremely low genome copy number revealed by Real Time PCR-based titration assay, indicating the packaging activity

of VEE-SARS-CoV-2 vectors is dependent upon the PS9 packaging signal sequence. In addition, although the moderate copy number of VEE-SARS-CoV-2 genome was detected in viral vectors lacking structural protein genes, lack of luciferase activity suggests that all structural protein genes

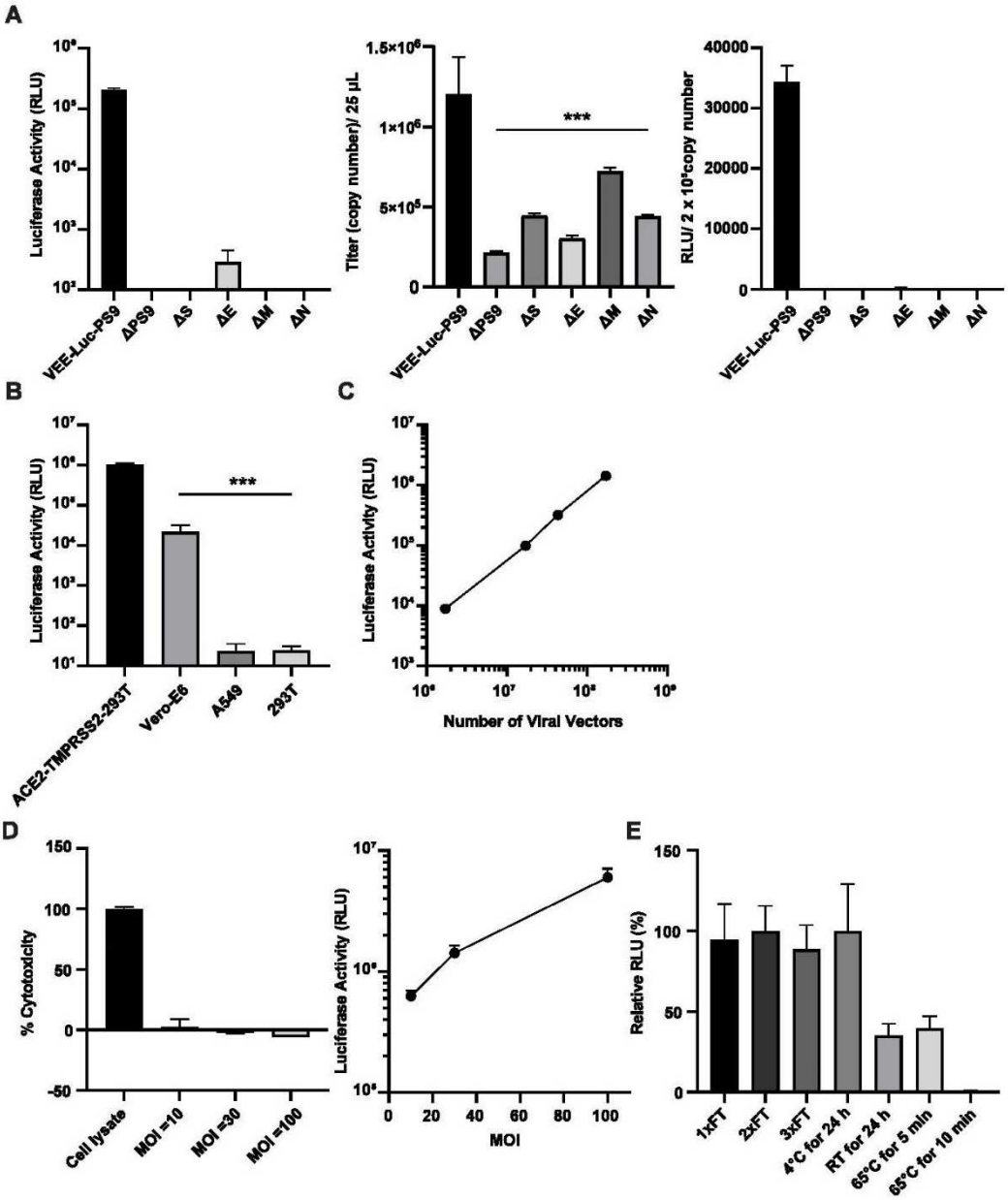


Figure 3.6 Characterization of VEE-SARS-CoV-2 viral vectors. (A) The effects of structural gene and package signal sequence on transduction activity of VEE-SARS-CoV-2 vectors prepared by co-transfection of 293T cells with all but lacking one structural gene or package signal sequence.

The viral vectors were prepared and titrated by qPCR and examined for transduction activity in ACE2-TMPRSS2-293T cells. **(B)** Cell type-dependent transduction activity of VEE-SARS-CoV-2 viral vector. Same number of cells (2 x10^5 /well) were transduced with equal number of VEE-SARS-CoV-2-Luc vectors. The transduction activity was determined by luciferase assay at 48 hr post-transduction. **(C)** Dose-dependent transduction of VEE-SARS-CoV-2 viral vectors. Same number of ACE2-TMPRSS2-293T cells (2 x10^5 /well) were transduced with increasing amount of VEE-SARS-CoV-2-Luc vectors and luciferase activity was determined. **(D)** Cytotoxicity assessment of VEE-SARS-CoV-2-Luc vectors in ACE2-TMPRSS2-293T cells. Lactate dehydrogenase (LDH) activity in the supernatant of transduced cells was determined 48 hr post-transduction. Cell lysate was used as the positive control for total release of LDH. Luciferase activity was measured to confirm the transduction activity of the viral vectors. **I** VEE-SARS-CoV-2-Luc vector stability under different conditions. All luciferase activity in transduced cells was measured 24 hr post-transduction except for **(B)** and **(D)**. Data are presented as the mean ± SD of 3 independent experiments. Statistical significance was determined by one-way ANOVA with Tukey Post Hoc tests. *** $p < 0.001$.

are required for the vector transduction activity. Overall, these results confirm that all functional components included in the vector design are essential for producing functional VEE-SARS-CoV-2 viral vectors.

We then set to investigate the cellular tropism of VEE-SARS-CoV-2 vector in various cell lines. In addition to ACE2-TMPRSS2-293T cell line, VEE-SARS-CoV-2 vectors also exhibited robust transduction activity in Vero-E6 cells, an African green monkey kidney cell line endogenously expressing ACE2. Not surprisingly, our vectors failed to transduce ACE2-negative cell lines, such as A549 and HEK293T (**Fig. 3.6B**). Taken together, VEE-SARS-CoV-2 vectors share the same cellular tropism with the authentic SARS-CoV-2. Its transduction activity requires ACE2 receptor on the target cell surface and is further enhanced by TMPRSS2.

Next, the dose-dependent transduction of VEE-SARS-CoV-2 viral vectors was characterized. **Fig 3.6C** shows that the level of luciferase expression was correlated with the copy number of VEE-SARS-CoV-2 vectors used for transduction. This linear dose dependence suggests that VEE-SARS-CoV-2 vectors are suitable for quantitation of antiviral agents including neutralizing antibody and antiviral drugs. More importantly, lactate dehydrogenase assay suggests

that three ascending doses (Multiple of Infection=10, 30, 100) did not induce any significant cytotoxicity in transduced cells at 48 hr post-transduction (**Fig. 3.6D**). The low cytopathic effect of the VEE-SARS-CoV-2 vectors could be attributed to mutations in the 5' UTR and nsP2/3 [9].

Lastly, the stability of VEE-SARS-CoV-2 vectors was tested under different conditions. The results show that the transduction activity remains at the same level upon 3x freeze-thaw cycles on ice or storage at 4°C for 24 hr. However, its transduction activity decreased by 65% when kept for 24 hr at room temperature. The VEE-SARS-CoV-2 vectors were completely inactivated when kept at 65°C for 10 min (**Fig. 3.6E**). The stability of VEE-SARS-CoV-2 viral vectors is consistent with the authentic SARS-CoV-2, which can be completely inactivated under 70°C for 5 min [10]. The similar stability profile results from the structural similarity between the VEE-SARS-CoV-2 vectors and the authentic virus.

3.7 Comparison between the VEE-SARS-CoV-2 vectors and the SC2-VLP system

A side-by-side comparison was conducted between VEE-SARS-CoV-2 vectors and the SARS-CoV-2-based viral like particles (SC2-VLP) previously established [1]. SC2-VLP-transduced cells showed positive luciferase activity (RLU>10^3) as early as 2 hr post-transduction, the luciferase activity peaked at 24 hr (**Fig. 3.7A**) and decreased thereafter by half in every 24 hr. In comparison, VEE-SARS-CoV-2 vectors exhibited a delayed onset of reporter expression and plateaued at 48 hr post-transduction. The VEE-SARS-CoV-2 vectors generated ~5-fold ($4.0x10^5$ vs $8.7x10^4$ RLU) and ~65-fold ($2.6x10^6$ vs $4.0x10^4$ RLU) higher luciferase gene expression than SC2-VLPs at 24 and 48 hr, respectively. This distinct expression kinetics attributes to the VEEV self-amplification and subsequent translation of reporter RNA sequence by nsP1-4 proteins. This mechanism was further manifested by qPCR analysis measuring the copy number change of

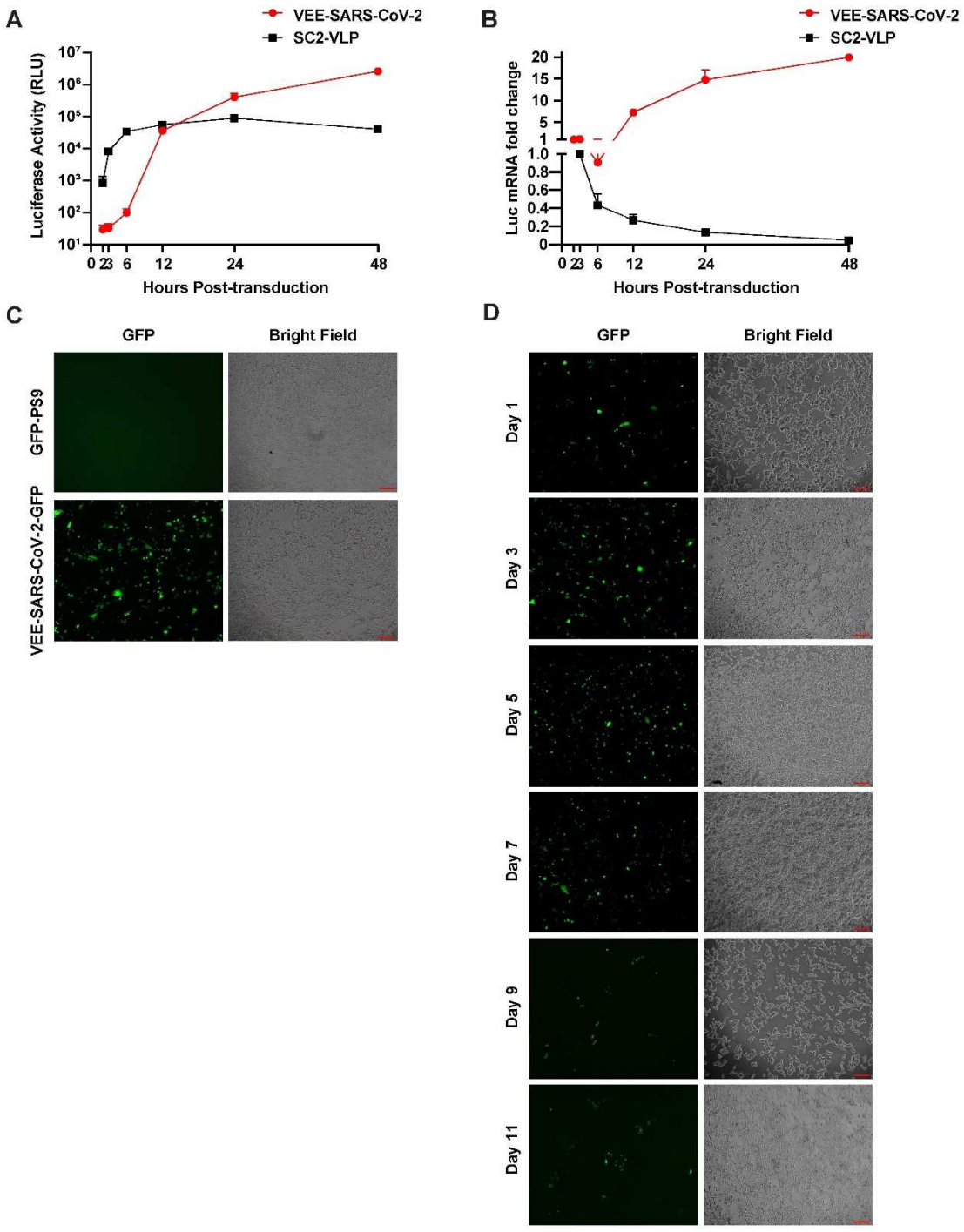


Figure 3.7 Comparison between VEE-SARS-CoV-2 vectors and SC2-VLP system. Luciferase level **(A)** and folds of luciferase mRNA change **(B)** in ACE2-TMPRSS2-293T cells after

transduction by VEE-SARS-COV-2-Luc vectors and SC2-Luc-VLP. Transduced cells were harvested at the indicated times, and luciferase assay and real-time qPCR analysis were performed. The fold change was normalized to the first time point determined (2 hr post-transduction). GAPDH was used as the internal control. Data are presented as the mean ± SD of independent triplicates. **(C)** GFP expression in ACE2-TMPRSS2-293T cells transduced with VEE-SARS-CoV-2-GFP vectors or SC2-GFP VLPs at 48 hr post-transduction. **(D)** Time-dependent GFP expression in ACE2-TMPRSS2-293T cells transduced by VEE-SARS-CoV-2-GFP vectors. Transduced cells were passaged every 4 days and GFP expression was recorded under a fluorescence microscope and cell density was photographed under regular microscope at the indicated times. Scale bar = 200 μm.

luciferase transcript in transduced cells. The results show that at 48 hr post-transduction, the reporter transcript copy number increased by 19-fold in the VEE-SARS-CoV-2-transduced cells but decreased by 24-fold in SC2-VLPs transduced cells (**Fig. 3.7B**).

To prove that VEE-SARS-CoV-2 transduction is not reporter-specific, we also produced VEE-SARS-CoV-2-GFP to deliver GFP into ACE2-TMPRSS2-293T cells. A substantially high number of GFP+ cells was observed in target cells transduced by VEE-SARS-CoV-2 vectors, whereas SC2-VLPs failed to generate any visible green cells (**Fig. 3.7C**). These data confirmed that VEE-SARS-CoV-2 vector induces significantly higher level of reporter gene expression compared to the SC2-VLPs.

The long-term transgene expression resulted from VEE-SARS-CoV-2 viral vectors was also investigated using GFP reporter. As shown in **Fig. 3.7D**, VEE-SARS-CoV-2 transduced cells exhibited an extended gene expression which persists at least 11 days after 2 passages of transduced cells. Taken together, these data demonstrate that VEE-SARS-CoV-2 vectors are highly effective in transduction and capable of generating prolonged transgene expression.

3.8 Evaluation of neutralizing antibody and antiviral compounds using the VEE-SARS-CoV-2 vectors

Proof-of-concept experiments were performed to demonstrate that VEE-SARS-CoV-2-Luc vector system can be used for assessing neutralizing antibody against SARS-CoV-2. A dose-dependent curve was obtained from a neutralization assay using VEE-SARS-CoV-2-Luc vectors pre-treated with a monoclonal antibody against SARS-CoV-2 spike protein (**Fig. 3.8A**). Notably, the IC_{50} obtained is 0.91 μg/mL, close to the IC_{50} value of 1.41 μg/mL against spike-pseudotyped lentiviruses generated by the manufacturer.

VEE-SARS-CoV-2-Luc vectors were also used to test antiviral activity of compounds such as cepharanthine, berbamine hydrochloride, and fangchinoline with reported activity in inhibiting SARS-CoV-2 infection. These three compounds were known for their activities in blocking spike protein-mediated viral entry [11-13] and were selected as model drugs to validate the drug testing assay using the VEE-SARS-CoV-2-Luc vectors. All model drugs showed antiviral activities in a dose-dependent manner (**Fig. 3.8B**) without inducing cytotoxicity to target cells in the concentration range tested (**Fig. 3.8C**). Among these compounds, cepharanthine exhibited the highest activity with an IC_{50} of 0.73 μM, which is close to IC_{50} (0.35 μM [11] and 1.90 μM [14]) determined by infection assays with authentic SARS-CoV-2 viruses. Consistent with the reported data obtained from SARS-CoV-2 [14], berbamine hydrochloride and fangchinoline showed lower activity against VEE-SARS-CoV-2 vector. Collectively, these data suggest that our novel VEE-SARS-CoV-2-Luc vector is a robust model for evaluating neutralizing antibody and antiviral drugs against SARS-CoV-2.

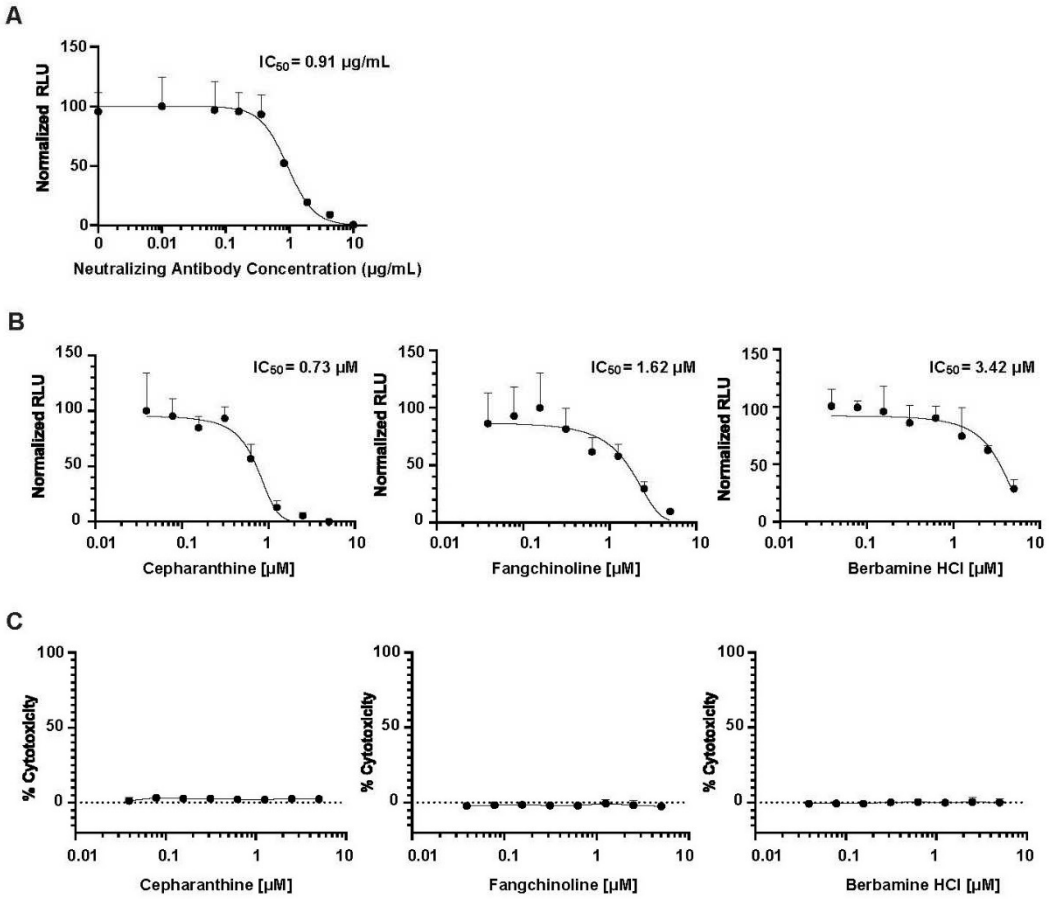


Figure 3.8 Evaluation of neutralizing antibody (A) and antiviral compounds (B) using the VEE-SARS-CoV-2 vectors. VEE-SARS-CoV-2-Luc vectors were incubated with MM43 SARS-CoV-2 spike neutralizing antibody, antiviral compound of cepharanthine, fangchiboline, or berbamine hydrochloride, for 1 hour in a volume of 100 µL at indicated concentration. The mixture was added to ACE2-TMPRSS2-293T cells (2.5 x10^4 /well) pre-seeded one day earlier. Transduction proceeded for 3 hr and the medium was replaced with fresh medium. **(C)** Cytotoxicity of compounds at different concentrations was determined at 48 hours post-treatment by LDH cytotoxicity assay. Luciferase activity was determined at 48 hr post-transduction. Data are presented as the mean ± SD of independent triplicates.

References

1. A. M. Syed *et al.*, Rapid assessment of SARS-CoV-2-evolved variants using virus-like particles. *Science* **374**, 1626-1632 (2021).

2. A. Sharma, B. Knollmann-Ritschel, Current understanding of the molecular basis of 89enezuelan equine encephalitis virus pathogenesis and vaccine development. *Viruses* **11** (2019)

3. R. M. Kinney, B. J. Johnson, J. B. Welch, K. R. Tsuchiya, D. W. Trent, The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative, strain TC-83. *Virology* **170**, 19–30 (1989).

4. H. Huysmans *et al.*, Expression kinetics and innate immune response after electroporation and LNP-mediated delivery of a self-amplifying mRNA in the skin. *Mol. Ther. Nucleic Acids* **17**, 867-878 (2019).

5. B. Leyman *et al.*, Comparison of the expression kinetics and immunostimulatory activity of replicating mRNA, nonreplicating mRNA, and pDNA after intradermal electroporation in pigs. *Mol. Pharm.* **15**, 377-384 (2018).

6. K. M. Pollock *et al.*, Safety and immunogenicity of a self-amplifying RNA vaccine against COVID-19: COVAC1, a phase I, dose-ranging trial. *EclinicalMedicine* **44**, 101262 (2022).

7. A. K. Blakney, S. Ip, A. J. Geall, An update on self-amplifying mRNA vaccine development. *Vaccines (*Basel) **9**, 97 (2021).

8. R. Kulasegaran-Shylini, V. Thiviyanathan, D. G. Gorenstein, I. Frolov, The 5'UTR-specific mutation in VEEV TC-83 genome has a strong effect on RNA replication and subgenomic RNA synthesis, but not on translation of the encoded proteins. *Virology* **387**, 211-221 (2009).

9. O. Petrakova *et al.*, Noncytopathic replication of Venezuelan equine encephalitis virus and eastern equine encephalitis virus replicons in Mammalian cells. *J. Virol.* **79**, 7597-7608 (2005).

10. A. W. H. Chin *et al.*, Stability of SARS-CoV-2 in different environmental conditions. *Lancet Microbe* **1**, e10 (2020).

11. H. Ohashi *et al*., Potential anti-COVID-19 agents, cepharanthine and nelfinavir, and their usage for combination treatment. *iScience* **24**, 102367 (2021).

12. B. A. Rodriguez-Rodriguez *et al*., Atovaquone and berberine chloride reduce SARS-CoV-2 replication in vitro. *Viruses* **13**, 2437 (2021).

13. N. Li *et al*., A new screening system for entry inhibitors based on cell-to-cell transmitted syncytia formation mediated by self-propagating hybrid VEEV-SARS-CoV-2 replicon. *Emerg. Microbes Infect* **11**, 465-476 (2022).

14. A. Hijikata *et al*., Evaluating cepharanthine analogues as natural drugs against SARS-CoV-2. *FEBS Open Bio* **12**, 285-294 (2022).

CHAPTER 4

DISCUSSION AND FUTURE PERSPECTIVES

4.1 Unique advantages of the VEE-SARS-CoV-2 vectors

The emergence of SARS-CoV-2 leads to an unparalleled global pandemic COVID-19, which has caused more than 670 million cases and more than 6.8 million deaths [7]. Understanding SARS-CoV-2 virology and identifying novel antiviral mechanisms are imperative to tackle this severe pandemic. Being a highly infectious and virulent virus, SARS-CoV-2 needs to be handled in BSL-3 laboratories. The majority of scientists has limited access to BSL-3 laboratories, which drastically hinders the SARS-CoV-2 research advancement. To address the issue, a novel SARS-CoV-2 viral vector system, dabbed VEE-SARS-CoV-2 viral vector, is developed in this dissertation project to offer an alternative approach with its unique advantages over other substitute systems.

4.1.1 Safety profile of the VEE-SARS-CoV-2 viral vectors

Firstly, the VEE-SARS-CoV-2 viral vectors are safe to use under BSL-2 conditions. Since all structural protein genes of SARS-CoV-2 are expressed *in trans*, the VEEV self-amplification machinery produces more copies of reporter gene without synthesizing new virions in transduced target cells (**Fig. 3.2B**). Meanwhile, adaptive mutations in VEEV 5' UTR, nsP2, and nsP3 significantly reduced the cytopathic effect of amplification and enabled the persistent transgene expression [1]. Its safety profile was also confirmed by LDH cytotoxicity assay. As shown in **Fig. 3.6D**, transduction of three ascending doses of VEE-SARS-CoV-2 vectors into ACE2-TMPRSS2-

293T cells generated a dose-dependent luciferase activity but did not induce any significant cytotoxicity. Because of the superior safety profile and low cytotoxicity, VEE-SARS-CoV-2 vectors are ideal to replace the disease-causing wild type SARS-CoV-2 for certain experiments.

4.1.2 Structural resemblance to SARS-CoV-2

Secondly, the VEE-SARS-CoV-2 vectors are composed of 4 SARS-CoV-2 structural (S, E, M, and N) proteins. Therefore, the VEE-SARS-CoV-2 vectors bear high structural resemblance to SARS-CoV-2 and, more importantly, can recapitulate the process of viral RNA packaging, viral release, and entry into ACE2+ cells. Compared to pseudotyped viruses bearing only one structural protein (S protein), the VEE-SARS-CoV-2 vectors comprise of all 4 structural proteins and an artificial genome encoding a reporter gene, which enables quantitation of transduced gene expression. The structural similarity could be utilized to investigate the critical steps in the SARS-CoV-2 life cycle such as viral entry, genome packaging, and budding. Additionally, the presence of all structural proteins makes the VEE-SARS-CoV-2 vector system suitable for screening antiviral drugs targeting SARS-CoV-2 structural proteins (including but not limited to spike protein).

4.1.3 Efficient and prolonged gene delivery by the VEE-SARS-CoV-2 vectors

Thirdly, the VEE-SARS-CoV-2 vectors exhibit a high gene delivery efficiency and a prolonged gene expression profile (**Fig. 3.7**). To my knowledge, the VEE-SARS-CoV-2 viral vector system represents the most efficient SARS-CoV-2 gene delivery vehicle system generating significantly more GFP+ cells compared to other systems. Chen et al. generated VLPs carrying GFP reporter to infect ACE2-293T cells. In that system, the GFP gene was linked to a putative

packaging signal PS576 sequence (19786-20361 nt of SARS-CoV-2 genome). Although the consistently positive fluorescence intensity was detected in infected ACE2+ cells, only single GFP+ cell was displayed in the data [2]. To enhance the gene delivery efficiency, Hetrick et al. employed Semliki Forest virus (SFV)-based self-amplifying RNA system to develop a hybrid SARS-CoV-2 pseudovirion system. Although an amplification mechanism is included, very limited number of GFP+ cells were generated by the system. One major limitation of that system is that the putative packaging signal was not optimized specifically for SARS-CoV-2, therefore the gene delivery efficiency is suboptimal. Also, it is suggested that other structural proteins except spike protein are not essential for its transduction activity [3]. To identify a functional packaging signal for SARS-CoV-2, Syed et al. performed extensive experiments and established a SC2-VLP system with the PS9 sequence obtained [4]. However, SC2-VLPs failed to generate GFP+ cells that can be observed by fluorescence microscopy since the copy number of RNA delivered is very limited. To this end, the VEE-SARS-CoV-2 vector system represents the optimal SARS-CoV-2-derived gene delivery system with improved transgene expression level and prolonged duration (**Fig. 3.7**). Notably, the VEEV self-amplifying mechanism generated a 65-fold increase of luciferase activity in transduced cells. The drastically higher gene delivery efficiency of the VEE-SARS-CoV-2 vector offers a wider range of quantitation for drug screening and neutralization antibody assay.

4.2 Applications of the VEE-SARS-CoV-2 vectors

4.2.1 Drug screening and neutralizing antibody evaluation

One important application of VEE-SARS-CoV-2 vectors is for drug screening and neutralizing antibody evaluation. Comparing to S-pseudotyped viruses that only contain SARS-

CoV-2 S protein instead of all 4 structural viral proteins, the VEE-SARS-CoV-2 vectors are better model viral particles for drug testing. Data shown in **Fig. 3.8** provide direct evidence in support of the utility of the VEE-SARS-CoV-2 system for evaluating neutralizing antibody and small molecules targeting S protein. In addition to S protein, other 3 structural proteins (E, M, and N proteins) are also present in the VEE-SARS-CoV-2 vectors. Therefore, this system can also be utilized to screen antiviral drugs against these 3 proteins, which are neglected in the current paradigm of COVID-19 drug development.

4.2.2 Emerging SARS-CoV-2 variant investigation

SARS-CoV-2 genomic sequence mutates while the virus spreads, leading to the emergence of new variants with distinct features. Evolving variants of SARS-CoV-2 require additional investigations to comprehensively understand the transmissibility and pathogenicity. To this end, the VEE-SARS-CoV-2 viral vector offers a convenient and reliable system for assessing the impact of structural protein mutations on the characteristics of new variants. Upon the emergence of a new SARS-CoV-2 variant, plasmids encoding the structural protein genes of the new variant can be readily constructed and used for producing evolved viral vectors. The transmissibility of the evolved vectors can be evaluated by comparing with the wild type proteins by luciferase assay. Additionally, this approach can also be utilized to evaluate the response of emerging variants to neutralizing antibodies generated by current COVID vaccines.

4.2.3 As a potential COVID vaccine platform

The VEE-SARS-CoV-2 vector is potentially a vaccine candidate against COVID-19. Since they are composed of multiple SARS-CoV-2 structural proteins, the VEE-SARS-CoV-2 virions

are believed to elicit immunity and induce antibody production in the host. This hypothesis is supported by several reports of testing SARS-CoV-2 VLPs produced through various approaches as COVID-19 vaccines in murine models. Lu et al. developed a mRNA vaccine encoding SARS-CoV-2 VLPs with immunogenicity in BALB/c mice [5]. Kumar et al. found that their VLPs conjugated with adjuvants elicited humoral and cellular immunity in murine models [6]. Furthermore, Yilmaz et al. protected K-18 ACE2 transgenic mice challenged with SARS-CoV-2 using the SARS-CoV-2 VLP vaccine produced from transiently transfected mammalian cells [7]. Compared to these VLP-based vaccines, the VEE-SARS-CoV-2 vector system not only comprises of all four structural proteins but also is able to deliver the antigen of interest, which makes it an attractive vaccine platform. In this regard, the potential of VEE-SARS-CoV-2 vectors for gene delivery and gene therapy is another important feature as a vaccine platform. Since the VEE-SARS-CoV-2 vectors infect the same target cells as the authentic SARS-CoV-2, one can foresee the possibility of delivery of siRNA [8, 9] or CRISPR-Cas13-based RNA editing machinery [10–12] to target cells of SARS-CoV-2 virus to inhibit viral replication.

4.3 Limitations and Future perspectives

One limitation of this vector system is the dependence on the high expression level of ACE2 in target cells. As shown in **Fig 3.6B**, luciferase activity was ~50 fold higher in 293T cells overexpressing hACE2 and TMPRSS2 compared to in Vero E6 cells that express lower level of endogenous ACE2. Another limitation of VEE-SARS-CoV-2 as a drug screening system is the absence of the SARS-CoV-2 replication machinery. In contrary to other SARS-CoV-2 model systems such as SARS-CoV-2 replicons [13–16], trans-complementation system [17, 18], and attenuated SARS-CoV-2 with deletions of accessory genes [19], VEE-SARS-CoV-2 fails to

recapitulate SARS-CoV-2 viral replication. Hence, this system can only be used for screening drugs targeting SARS-CoV-2 structural proteins but not non-structural proteins.

The data shown in this dissertation were obtained in cell lines, primarily ACE2-TMPRSS2-293T cells. It is essential to examine the transduction activity of VEE-SARS-CoV-2 vectors in other models such as primary human cells, human organoids, and animal models.

Compared to specific cell lines, human organoids are composed of different cell types and able to recapitulate physiological conditions of human organs, which can be utilized to model viral infections. In this case, VEE-SARS-CoV-2 vectors can replace the wild type virus for modelling infections in various tissues such as kidney [20], liver [21], and the cardiovascular system [22].

Furthermore, the transduction activity of VEE-SARS-CoV-2 vector system in animal models remains to be examined. A worthwhile endeavor would be introducing the VEE-SARS-CoV-2 vectors into murine models that are permissive to SARS-CoV-2 infection such as K18-hACE2 transgenic mice [23, 24]. As a surrogate system for the wild type virus, our system could model the viral infection and shed insights on mechanistic virology studies *in vivo*. Meanwhile, the *in vivo* data would potentially lay a foundation for exploiting VEE-SARS-CoV-2 vector system as a vaccine platform. The advantages of the VEE-SARS-CoV2 vectors in vaccine development due to its characteristics of prolonged gene expression warrantee further studies. Practically, S-protein gene could be included into VEE-SARS-CoV-2 vectors as a transgene to elevate S-protein expression and immune response against S-protein. Similarly, any of the 4 structural protein gene can be cloned and overexpressed in the presence of all structural proteins for maximal immune stimulation.

In conclusion, this dissertation demonstrates the successful development of novel VEE-SARS-CoV-2 viral vectors consisting of 4 structural proteins of SARS-CoV-2, a self-amplifying

machinery, and a transgene. The transducing activity of the vectors is dependent upon ACE2 and all structural proteins and outperforms other vector systems developed so far. The viral vectors can be used under BSL-2 conditions as a versatile model system for SARS-CoV-2 virology and antiviral agent development. The viral vectors developed could be explored for other applications such as vaccine development and gene delivery.

References

1. O. Petrakova *et al.*, Noncytopathic replication of Venezuelan equine encephalitis virus and eastern equine encephalitis virus replicons in Mammalian cells. *J Virol* **79**, 7597-7608 (2005).

2. M. Chen, C. Yan, F. Qin, L. Zheng, X. E. Zhang, The intraviral protein-protein interaction of SARS-CoV-2 reveals the key role of N protein in virus-like particle assembly. *Int J Biol Sci* **17**, 3889-3897 (2021).

3. B. Hetrick *et al.*, Development of a hybrid alphavirus-SARS-CoV-2 pseudovirion for rapid quantification of neutralization antibodies and antiviral drugs. *Cell Rep. Methods* **2**, 100181 (2022).

4. A. M. Syed *et al.*, Rapid assessment of SARS-CoV-2-evolved variants using virus-like particles. *Science* **374**, 1626-1632 (2021)

5. J. Lu *et al*., A COVID-19 mRNA vaccine encoding SARS-CoV-2 virus-like particles induces a strong antiviral-like immune response in mice. *Cell Res* **30**, 936-939 (2020).

6. C. S. Kumar *et al*., Virus-like particles of SARS-CoV-2 as virus surrogates: morphology, immunogenicity, and internalization in neuronal cells. *ACS Infect. Dis* **8**, 2119-2132 (2022).

7. I. C. Yilmaz *et al*., Development and preclinical evaluation of virus-like particle vaccine against COVID-19 infection. *Allergy* **77**, 258-270 (2022).

8. Y. C. Chang *et al.*, A siRNA targets and inhibits a broad range of SARS-CoV-2 infections including Delta variant. *EMBO Mol Med* **14**, e15298 (2022)

9. A. Idris *et al*., A SARS-CoV-2 targeted siRNA-nanoparticle therapy for COVID-19. *Mol Ther* **29**, 2219-2226 (2021).

10. T. R. Abbott *et al.*, Development of CRISPR as an antiviral strategy to combat SARS-CoV-2 and influenza. *Cell* **181**, 865-876 (2020)

11. M. Fareh *et al.*, Reprogrammed CRISPR-Cas13b suppresses SARS-CoV-2 replication and circumvents its mutational escape through mismatch tolerance. *Nat Commun* **12**, 4270 (2021)

12. E. L. Blanchard *et al.*, Treatment of influenza and SARS-CoV-2 infections via mRNA-encoded Cas13a in rodents. *Nat Biotechnol* **39**, 717-726 (2021)

13. H. Xia *et al*., Evasion of Type I Interferon by SARS-CoV-2. *Cell Rep* **33**, 108234 (2020).

14. J. Malicoat *et al*., Development of a single-cycle infectious SARS-CoV-2 virus replicon particle system for use in biosafety level 2 laboratories. *J. Virol* **96**, e01837-21 (2022).

15. X. He *et al*., Generation of SARS-CoV-2 reporter replicon for high-throughput antiviral screening and testing. *Proc. Natl. Acad. Sci. U. S. A*. **118**, e2025866118 (2021).

16. T. Tanaka *et al*., Establishment of a stable SARS-CoV-2 replicon system for application in high-throughput screening. *Antiviral Res* **199**, 105268 (2022).

17. X. Zhang *et al*., A trans-complementation system for SARS-CoV-2 recapitulates authentic viral replication without virulence. *Cell* **184**, 2229-2238 (2021).

18. C. Kurhade, X. Xie, PY Shi, Reverse genetic systems of SARS-CoV-2 for antiviral research. *Antiviral Res* **210**, 105486 (2023).

19. Y. Liu *et al*. A live-attenuated SARS-CoV-2 vaccine candidate with accessory protein deletions. *Nat Commun* **13**, 4337 (2022).

20. J. Jansen *et al.*, SARS-CoV-2 infects the human kidney and drives fibrosis in kidney organoids. *Cell Stem Cell* **29**, 217-231 (2022).

21. L. Yang *et al.*, A Human pluripotent stem cell-based platform to study SARS-CoV-2 tropism and model virus infection in human cells and organoids. *Cell Stem Cell* **27**, 125-136 (2020)

22. J. A. Perez-Bermejo *et al.*, SARS-CoV-2 infection of human iPSC-derived cardiac cells reflects cytopathic features in hearts of patients with COVID-19. *Sci Transl Med* **13**, eabf7872 (2021)

23. F. S. Oladunni *et al.*, Lethality of SARS-CoV-2 infection in K18 human angiotensin-converting enzyme 2 transgenic mice. *Nat Commun* **11**, 6122 (2020)

24. S. H. Sun *et al*., A mouse model of SARS-CoV-2 infection and pathogenesis. *Cell Host Microbe* **28**, 124-133 (2020).

APPENDIX A

Full sequences of pcDNA3.1(+)-VEE-GFP-PS9 plasmid

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTG

ATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGT

AGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATG

AAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATAT

ACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA

GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT

GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA

GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT

GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTC

AATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT

CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT

GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC

ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA

AATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG

GAGGTCTATATAAGCAGAGCTCTACGACTCACTATAGATGGGCGGCGCATGAGAGA

AGCCCAGACCAATTACCTACCCAAAATGGAGAAAGTTCACGTTGACATCGAGGAAG

ACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTTCCCGCAGTTTGAGGTAGAAGCCA

AGCAGGTCACTGATAATGACCATGCTAATGCCAGAGCGTTTTCGCATCTGGCTTCAA

AACTGATCGAAACGGAGGTGGACCCATCCGACACGATCCTTGACATTGGAAGTGCG

CCCGCCCGCAGAATGTATTCTAAGCACAAGTATCATTGTATCTGTCCGATGAGATGT

GCGGAAGATCCGGACAGATTGTATAAGTATGCAACTAAGCTGAAGAAAAACTGTAA

GGAAATAACTGATAAGGAATTGGACAAGAAAATGAAGGAGCTGGCCGCCGTCATGA

GCGACCCTGACCTGGAAACTGAGACTATGTGCCTCCACGACGACGAGTCGTGTCGCT

ACGAAGGGCAAGTCGCTGTTTACCAGGATGTATACGCGGTTGACGGACCGACAAGT

CTCTATCACCAAGCCAATAAGGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACC

ACCCCTTTTATGTTTAAGAACTTGGCTGGAGCATATCCATCATACTCTACCAACTGG

GCCGACGAAACCGTGTTAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATG

GAGCGGTCACGTAGAGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAA

CAATGTTCTATTCTCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAG

GAGCTGGCACCTGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCG

GTGTGAGACTATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCC

AGGCCTGTATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTT

GTGCTGCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCAC

GTATGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAG

TGCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAACG

GTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGGCCC

AGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAAGGCC

ACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAGAAGGCA

CAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAAGTGAACA

GCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGGAGATCGGGC

TGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGTCACCTCTCATT

ACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGGCTAAGGAGGTGC

GTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGCTGATGTTGAGGAG

CCCACTCTGGAGGCAGACGTCGACTTGATGTTACAAGAGGCTGGGGCCGGCTCAGT

GGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATGGCGAGGACAAGATCG

GCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAGTGAAAAATTATCTTGCAT

CCACCCTCTCGCTGAACAAGTCATAGTGATAACACACTCTGGCCGAAAAGGGCGTT

ATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGCCAGAGGGACATGCAATACCC

GTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCACCATTGTGTACAACGAACGTGAG

TTCGTAAACAGGTACCTGCACCATATTGCCACACATGGAGGAGCGCTGAACACTGAT

GAAGAATATTACAAAACTGTCAAGCCCAGCGAGCACGACGGCGAATACCTGTACGA

CATCGACAGGAAACAGTGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCACAG

GCGAGCTGGTGGATCCTCCCTTCCATGAATTCGCCTACGAGAGTCTGAGAACACGAC

CAGCCGCTCCTTACCAAGTACCAACCATAGGGGTGTATGGCGTGCCAGGATCAGGC

AAGTCTGGCATCATTAAAAGCGCAGTCACCAAAAAAGATCTAGTGGTGAGCGCCAA

GAAAGAAAACTGTGCAGAAATTATAAGGGACGTCAAGAAAATGAAAGGGCTGGAC

GTCAATGCCAGAACTGTGGACTCAGTGCTCTTGAATGGATGCAAACACCCCGTAGA

GACCCTGTATATTGACGAAGCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCAT

AGCCATTATAAGACCTAAAAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTT

TTTTTAACATGATGTGCCTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCT

TCCACAAAAGCATCTCTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTT

GTTTTACGACAAAAAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTG

ACACTACCGGCAGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAG

GGTGGGTGAAGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCT

GCCTCTCAAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGA

AAATCCTCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGA

GGACCGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTG

CCAAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT

GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATAA

GGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCATAG

ACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCTCAC

TCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATCTGGAC

TCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCACTGGGATA

ACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTCAGCTCTCTC

GCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTATGACATGAAC

ACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCTGTAAACAGAAG

ACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGAGTGACTTTTCTTCA

TTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGGGGAAAAGTTGTCCGTC

CCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAGGCTACCTTCAGAGCTCGG

CTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATGACATAATATTTGTTAATGTG

AGGACCCCATATAAATACCATCACTATCAGCAGTGTGAAGACCATGCCATTAAGCTT

AGCATGTTGACCAAGAAAGCTTGTCTGCATCTGAATCCCGGCGGAACCTGTGTCAGC

ATAGGTTATGGTTACGCTGACAGGGCCAGCGAAAGCATCATTGGTGCTATAGCGCG

GCAGTTCAAGTTTTCCCGGGTATGCAAACCGAAATCCTCACTTGAAGAGACGGAAGT

TCTGTTTGTATTCATTGGGTACGATCGCAAGGCCCGTACGCACAATTCTTACAAGCTT

TCATCAACCTTGACCAACATTTATACAGGTTCCAGACTCCACGAAGCCGGATGTGCA

CCCTCATATCATGTGGTGCGAGGGGATATTGCCACGGCCACCGAAGGAGTGATTATA

AATGCTGCTAACAGCAAAGGACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAA

GAAATTCCCGGAAAGCTTCGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGG

TCAAAGGTGCAGCTAAACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTT

CGGAGGTTGAAGGTGACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATT

GTCAACGATAACAATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTT

TCCGGGAACAAAGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGAC

ACCACTGATGCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCT

CAAGGAAGCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGAC

TCTTCAGTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTG

GCTGGAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGG

GACCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCG

TTGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAGC

AGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCACCTAGCAC

GCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAAAAGC

CTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTATAGAAT

CACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGAAAGTGCC

TGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGACGAGACTCC

GGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACAACCACCACTTA

TAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCATCATCGAAGAGGAA

GAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCACCAGGTGCTGCAAGT

CGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCATCCTGGTCCATTCCTCA

TGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGACACCCTGGAGGGAGCTAG

CGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTCTTACTTCGCAAAGAGTATGG

AGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACAGTATTCAGGAACCCTCCACATC

CCGCTCCGCGCACAAGAACACCGTCACTTGCACCCAGCAGGGCCTGCTCGAGAACC

AGCCTAGTTTCCACCCCGCCAGGCGTGAATAGGGTGATCACTAGAGAGGAGCTCGA

GGCGCTTACCCCGTCACGCACTCCTAGCAGGTCGGTCTCGAGAACCAGCCTGGTCTC

CAACCCGCCAGGCGTAAATAGGGTGATTACAAGAGAGGAGTTTGAGGCGTTCGTAG

CACAACAACAATGACGGTTTGATGCGGGTGCATACATCTTTTCCTCCGACACCGGTC

AAGGGCATTTACAACAAAAATCAGTAAGGCAAACGGTGCTATCCGAAGTGGTGTTG

GAGAGGACCGAATTGGAGATTTCGTATGCCCCGCGCCTCGACCAAGAAAAAGAAGA

ATTACTACGCAAGAAATTACAGTTAAATCCCACACCTGCTAACAGAAGCAGATACC

AGTCCAGGAAGGTGGAGAACATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGC

CTAGGGCATTATTTGAAGGCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCC

TGTTCCTTTGTATTCATCTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTG

GAAGCCTGTAACGCCATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATT

ATTCCAGAGTACGATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGAC

ACTGCCAGTTTTTGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTG

GAACCCACAATACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGT

CCTGGCAGCTGCCACAAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCG

TATTGGATTCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATG

AATATTGGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAAACGTGGTA

AATTACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACAT

AATTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAG

AGACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG

GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAGA

GCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGATAT

GTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATTGTGT

TCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGCTCTGAC

CGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACGCTGATTGA

GGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTAAATTTAAATT

CGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAACACAGTCATTAA

CATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGGATCACCATGTGCAG

CATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCGGACAAATTAATGGCA

GACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATTATAGATGCTGTGGTGGG

CGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGTGTGACTCCGTGACCGGCAC

AGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTTTAAGCTTGGCAAACCTCTGGC

AGCAGACGATGAACATGATGATGACAGGAGAAGGGCATTGCATGAAGAGTCAACA

CGCTGGAACCGAGTGGGTATTCTTTCAGAGCTGTGCAAGGCAGTAGAATCAAGGTA

TGAAACCGTAGGAACTTCCATCATAGTTATGGCCATGACTACTCTAGCTAGCAGTGT

TAAATCATTCAGCTACCTGAGAGGGGCCCCTATAACTCTCTACGGCTAACCTGAATG

GACTACGACATAGTCTAGTCCGCCAAGTCTAGCATATGGGCGCGCCCTCAGCATCGA

TTCAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCC

ATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGA

GGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG

GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT

GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC

CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG

ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA

GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACT

ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTG

AACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA

CCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACC

TGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC

CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA

GTAGTCTAGAGTCGACCCGGGCGGCCGCCTCCTTTAAGAAGGAGATATAGATCTCTG

TAGGTCCCAAACAAGCTAGTCTTAATGGAGTCACATTAATTGGAGAAGCCGTAAAA

ACACAGTTCAATTATTATAAGAAAGTTGATGGTGTTGTCCAACAATTACCTGAAACT

TACTTTACTCAGAGTAGAAATTTACAAGAATTTAAACCCAGGAGTCAAATGGAAATT

GATTTCTTAGAATTAGCTATGGATGAATTCATTGAACGGTATAAATTAGAAGGCTAT

GCCTTCGAACATATCGTTTATGGAGATTTTAGTCATAGTCAGTTAGGTGGTTTACATC

TACTGATTGGACTAGCTAAACGTTTTAAGGAATCACCTTTTGAATTAGAAGATTTTA

TTCCTATGGACAGTACAGTTAAAAACTATTTCATAACAGATGCGCAAACAGGTTCAT

CTAAGTGTGTGTGTTCTGTTATTGATTTATTACTTGATGATTTTGTTGAAATAATAAA

ATCCCAAGATTTATCTGTAGTTTCTAAGGTTGTCAAAGTGACTATTGACTATACAGA

AATTTCATTTATGCTTTGGTGTAAAGATGGCCATGTAGAAACATTTTACCCAAAATT

ACAATCTAGTCAAGCGTGGCAACCGGGTGTTGCTATGCCTAATCTTTACAAAATGCA

AAGAATGCTATTAGAAAAGTGTGACCTTCAAAATTATGGTGATAGTGCAACATTACC

TAAAGGCATAATGATGAATGTCGCAAAATATACTCAACTGTGTCAATATTTAAACAC

ATTAACATTAGCTGTACCCTATAATATGAGAGTTATACATTTTGGTGCTGGTTCTGAT

AAAGGAGTTGCACCAGGTACAGCTGTTTTAAGACAGTGGTTGCCTACGGGTACGCTG

CTTGTCGATTCAGATCTTAATGACTTTGTCTCTGATGCAGATTCAACTTTGATTGGTG

ATTGTGCAACTGTACATACAGCTAATAAATGGGATCTCATTATTAGTGATATGTACG

ACCCTAAGACTAAAAATGTTACAAAAGAAAATGACTCTAAAGAGGGTTTTTTCACTT

ACATTTGTGGGTTTATACAACAAAAGCTAGCTCTTGGAGGTTCCGTGGCTATAAAGA

TAGATCCCAACTCCATAAGTGTGTCGACATATAAGGATCCACTCGACGACTCGACAA

TCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCT

CCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCC

GTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGA

GTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAAC

CCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTC

CCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACA

GGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCC

TTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCT

ACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCT

GCGGCCTCTTCCGCATCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCC

GCCTCCCCGCCTGGAATTAATTCGAGCTCGGTACCTTGATTCTCGAGGAATTGGCAA

GCTGCTTACATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTTATTT

TTTCTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTCAAAAAAAAAAAAAAAAA

AAAAAAAAACGCGTCGAGGGGAATTAATTCTTGAAGACGAAAGGGCCAGGTGGCA

CTTTTCGGGGAAATGTGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTG

TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCT

GGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTG

TCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGG

AGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCT

GAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGG

CGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA

GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGG

CTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTA

CGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCG

CCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGAC

TCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATA

AGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATT

TAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGC

TCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTG

TGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT

AGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCA

GTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGA

GGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT

AGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGA

GACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCC

GGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCT

GCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCA

AGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCG

TGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCG

GGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCA

CCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAC

GCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAG

CACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCAT

CAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGG

CGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAA

TGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCA

GGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTG

ACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTA

TCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAA

GCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGG

TTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGAT

CTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACA

AATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTA

GTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTC

TAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCG

CTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGC

CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC

GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCG

GTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT

TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG

AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG

GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA

GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA

GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC

GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC

TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG

CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG

TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT

TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT

ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT

TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT

TTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT

TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT

TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGA

AGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC

TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT

GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG

CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC

AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC

CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG

CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG

GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTG

TGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCC

GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCAT

CCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT

GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC

ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT

CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT

GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC

AAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC

TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA

TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAA

AAGTGCCACCTGACGTC

Note: CMV promoter sequence is marked in dark red; VEEV 5' and 3' UTR sequence are marked in orange; VEEV NSP 1-4 sequence is marked in green; 26S promoter sequence is marked in blue; GFP sequence is marked in red; Packaging Signal (PS9) is marked in purple; polyadenylation tail is marked in light blue; pcDNA3.1 (+) backbone sequence is marked in black.

APPENDIX B

Full sequences of pcDNA3.1(+)-VEE-GFP-PS9-ΔNdeI plasmid

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTG

ATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGT

AGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATG

AAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATAT

ACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA

GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT

GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA

GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT

GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTC

AATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT

CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT

GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC

ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA

AATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG

GAGGTCTATATAAGCAGAGCTCTACGACTCACTATAGATGGGCGGCGCATGAGAGA

AGCCCAGACCAATTACCTACCCAAAATGGAGAAAGTTCACGTTGACATCGAGGAAG

ACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTTCCCGCAGTTTGAGGTAGAAGCCA

AGCAGGTCACTGATAATGACCATGCTAATGCCAGAGCGTTTTCGCATCTGGCTTCAA

AACTGATCGAAACGGAGGTGGACCCATCCGACACGATCCTTGACATTGGAAGTGCG

CCCGCCCGCAGAATGTATTCTAAGCACAAGTATCATTGTATCTGTCCGATGAGATGT

GCGGAAGATCCGGACAGATTGTATAAGTATGCAACTAAGCTGAAGAAAAACTGTAA

GGAAATAACTGATAAGGAATTGGACAAGAAAATGAAGGAGCTGGCCGCCGTCATGA

GCGACCCTGACCTGGAAACTGAGACTATGTGCCTCCACGACGACGAGTCGTGTCGCT

ACGAAGGGCAAGTCGCTGTTTACCAGGATGTATACGCGGTTGACGGACCGACAAGT

CTCTATCACCAAGCCAATAAGGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACC

ACCCCTTTTATGTTTAAGAACTTGGCTGGAGCATATCCATCATACTCTACCAACTGG

GCCGACGAAACCGTGTTAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATG

GAGCGGTCACGTAGAGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAA

CAATGTTCTATTCTCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAG

GAGCTGGCACCTGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCG

GTGTGAGACTATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCC

AGGCCTGTATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTT

GTGCTGCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCAC

GTATGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAG

TGCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAACG

GTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGGCCC

AGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAAGGCC

ACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAGAAGGCA

CAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAAGTGAACA

GCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGGAGATCGGGC

TGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGTCACCTCTCATT

ACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGGCTAAGGAGGTGC

GTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGCTGATGTTGAGGAG

CCCACTCTGGAGGCAGACGTCGACTTGATGTTACAAGAGGCTGGGGCCGGCTCAGT

GGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATGGCGAGGACAAGATCG

GCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAGTGAAAAATTATCTTGCAT

CCACCCTCTCGCTGAACAAGTCATAGTGATAACACACTCTGGCCGAAAAGGGCGTT

ATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGCCAGAGGGACATGCAATACCC

GTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCACCATTGTGTACAACGAACGTGAG

TTCGTAAACAGGTACCTGCACCATATTGCCACACATGGAGGAGCGCTGAACACTGAT

GAAGAATATTACAAAACTGTCAAGCCCAGCGAGCACGACGGCGAATACCTGTACGA

CATCGACAGGAAACAGTGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCACAG

GCGAGCTGGTGGATCCTCCCTTCCATGAATTCGCCTACGAGAGTCTGAGAACACGAC

CAGCCGCTCCTTACCAAGTACCAACCATAGGGGTGTATGGCGTGCCAGGATCAGGC

AAGTCTGGCATCATTAAAAGCGCAGTCACCAAAAAAGATCTAGTGGTGAGCGCCAA

GAAAGAAAACTGTGCAGAAATTATAAGGGACGTCAAGAAAATGAAAGGGCTGGAC

GTCAATGCCAGAACTGTGGACTCAGTGCTCTTGAATGGATGCAAACACCCCGTAGA

GACCCTGTATATTGACGAAGCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCAT

AGCCATTATAAGACCTAAAAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTT

TTTTTAACATGATGTGCCTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCT

TCCACAAAAGCATCTCTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTT

GTTTTACGACAAAAAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTG

ACACTACCGGCAGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAG

GGTGGGTGAAGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCT

GCCTCTCAAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGA

AAATCCTCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGA

GGACCGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTG

CCAAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT

GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATAA

GGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCATAG

ACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCTCAC

TCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATCTGGAC

TCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCACTGGGATA

ACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTCAGCTCTCTC

GCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTATGACATGAAC

ACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCTGTAAACAGAAG

ACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGAGTGACTTTTCTTCA

TTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGGGGAAAAGTTGTCCGTC

CCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAGGCTACCTTCAGAGCTCGG

CTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATGACATAATATTTGTTAATGTG

AGGACCCCATATAAATACCATCACTATCAGCAGTGTGAAGACCATGCCATTAAGCTT

AGCATGTTGACCAAGAAAGCTTGTCTGCATCTGAATCCCGGCGGAACCTGTGTCAGC

ATAGGTTATGGTTACGCTGACAGGGCCAGCGAAAGCATCATTGGTGCTATAGCGCG

GCAGTTCAAGTTTTCCCGGGTATGCAAACCGAAATCCTCACTTGAAGAGACGGAAGT

TCTGTTTGTATTCATTGGGTACGATCGCAAGGCCCGTACGCACAATTCTTACAAGCTT

TCATCAACCTTGACCAACATTTATACAGGTTCCAGACTCCACGAAGCCGGATGTGCA

CCCTCATATCATGTGGTGCGAGGGGATATTGCCACGGCCACCGAAGGAGTGATTATA

AATGCTGCTAACAGCAAAGGACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAA

GAAATTCCCGGAAAGCTTCGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGG

TCAAAGGTGCAGCTAAACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTT

CGGAGGTTGAAGGTGACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATT

GTCAACGATAACAATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTT

TCCGGGAACAAAGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGAC

ACCACTGATGCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCT

CAAGGAAGCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGAC

TCTTCAGTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTG

GCTGGAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGG

GACCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCG

TTGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAGC

AGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCACCTAGCAC

GCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAAAAGC

CTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTATAGAAT

CACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGAAAGTGCC

TGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGACGAGACTCC

GGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACAACCACCACTTA

TAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCATCATCGAAGAGGAA

GAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCACCAGGTGCTGCAAGT

CGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCATCCTGGTCCATTCCTCA

TGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGACACCCTGGAGGGAGCTAG

CGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTCTTACTTCGCAAAGAGTATGG

AGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACAGTATTCAGGAACCCTCCACATC

CCGCTCCGCGCACAAGAACACCGTCACTTGCACCCAGCAGGGCCTGCTCGAGAACC

AGCCTAGTTTCCACCCCGCCAGGCGTGAATAGGGTGATCACTAGAGAGGAGCTCGA

GGCGCTTACCCCGTCACGCACTCCTAGCAGGTCGGTCTCGAGAACCAGCCTGGTCTC

CAACCCGCCAGGCGTAAATAGGGTGATTACAAGAGAGGAGTTTGAGGCGTTCGTAG

CACAACAACAATGACGGTTTGATGCGGGTGCATACATCTTTTCCTCCGACACCGGTC

AAGGGCATTTACAACAAAAATCAGTAAGGCAAACGGTGCTATCCGAAGTGGTGTTG

GAGAGGACCGAATTGGAGATTTCGTATGCCCCGCGCCTCGACCAAGAAAAAGAAGA

ATTACTACGCAAGAAATTACAGTTAAATCCCACACCTGCTAACAGAAGCAGATACC

AGTCCAGGAAGGTGGAGAACATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGC

CTAGGGCATTATTTGAAGGCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCC

TGTTCCTTTGTATTCATCTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTG

GAAGCCTGTAACGCCATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATT

ATTCCAGAGTACGATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGAC

ACTGCCAGTTTTTGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTG

GAACCCACAATACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGT

CCTGGCAGCTGCCACAAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCG

TATTGGATTCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATG

AATATTGGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAAACGTGGTA

AATTACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACAT

AATTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAG

AGACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG

GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAGA

GCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGATAT

GTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATTGTGT

TCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGCTCTGAC

CGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACGCTGATTGA

GGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTAAATTTAAATT

CGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAACACAGTCATTAA

CATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGGATCACCATGTGCAG

CATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCGGACAAATTAATGGCA

GACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATTATAGATGCTGTGGTGGG

CGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGTGTGACTCCGTGACCGGCAC

AGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTTTAAGCTTGGCAAACCTCTGGC

AGCAGACGATGAACATGATGATGACAGGAGAAGGGCATTGCATGAAGAGTCAACA

CGCTGGAACCGAGTGGGTATTCTTTCAGAGCTGTGCAAGGCAGTAGAATCAAGGTA

TGAAACCGTAGGAACTTCCATCATAGTTATGGCCATGACTACTCTAGCTAGCAGTGT

TAAATCATTCAGCTACCTGAGAGGGGCCCCTATAACTCTCTACGGCTAACCTGAATG

GACTACGACATAGTCTAGTCCGCCAAGTCTAGCATATAGGCGCGCCCTCAGCATCGA

TTCAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCC

ATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGA

GGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG

GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT

GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC

CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG

ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA

GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACT

ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTG

AACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA

CCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACC

TGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC

CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA

GTAGTCTAGAGTCGACCCGGGCGGCCGCCTCCTTTAAGAAGGAGATATAGATCTCTG

TAGGTCCCAAACAAGCTAGTCTTAATGGAGTCACATTAATTGGAGAAGCCGTAAAA

ACACAGTTCAATTATTATAAGAAAGTTGATGGTGTTGTCCAACAATTACCTGAAACT

TACTTTACTCAGAGTAGAAATTTACAAGAATTTAAACCCAGGAGTCAAATGGAAATT

GATTTCTTAGAATTAGCTATGGATGAATTCATTGAACGGTATAAATTAGAAGGCTAT

GCCTTCGAACATATCGTTTATGGAGATTTTAGTCATAGTCAGTTAGGTGGTTTACATC

TACTGATTGGACTAGCTAAACGTTTTAAGGAATCACCTTTTGAATTAGAAGATTTTA

TTCCTATGGACAGTACAGTTAAAAACTATTTCATAACAGATGCGCAAACAGGTTCAT

CTAAGTGTGTGTGTTCTGTTATTGATTTATTACTTGATGATTTTGTTGAAATAATAAA

ATCCCAAGATTTATCTGTAGTTTCTAAGGTTGTCAAAGTGACTATTGACTATACAGA

AATTTCATTTATGCTTTGGTGTAAAGATGGCCATGTAGAAACATTTTACCCAAAATT

ACAATCTAGTCAAGCGTGGCAACCGGGTGTTGCTATGCCTAATCTTTACAAAATGCA

AAGAATGCTATTAGAAAAGTGTGACCTTCAAAATTATGGTGATAGTGCAACATTACC

TAAAGGCATAATGATGAATGTCGCAAAATATACTCAACTGTGTCAATATTTAAACAC

ATTAACATTAGCTGTACCCTATAATATGAGAGTTATACATTTTGGTGCTGGTTCTGAT

AAAGGAGTTGCACCAGGTACAGCTGTTTTAAGACAGTGGTTGCCTACGGGTACGCTG

CTTGTCGATTCAGATCTTAATGACTTTGTCTCTGATGCAGATTCAACTTTGATTGGTG

ATTGTGCAACTGTACATACAGCTAATAAATGGGATCTCATTATTAGTGATATGTACG

ACCCTAAGACTAAAAATGTTACAAAAGAAAATGACTCTAAAGAGGGTTTTTTCACTT

ACATTTGTGGGTTTATACAACAAAAGCTAGCTCTTGGAGGTTCCGTGGCTATAAAGA

TAGATCCCAACTCCATAAGTGTGTCGACATATAAGGATCCACTCGACGACTCGACAA

TCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCT

CCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCC

GTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGA

GTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAAC

CCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTC

CCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACA

GGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCC

TTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCT

ACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCT

GCGGCCTCTTCCGCATCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCC

GCCTCCCCGCCTGGAATTAATTCGAGCTCGGTACCTTGATTCTCGAGGAATTGGCAA

GCTGCTTACATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTTATTT

TTTCTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTCAAAAAAAAAAAAAAAAA

AAAAAAAAACGCGTCGAGGGGAATTAATTCTTGAAGACGAAAGGGCCAGGTGGCA

CTTTTCGGGGAAATGTGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTG

TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCT

GGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTG

TCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGG

AGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCT

GAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGG

CGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA

GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGG

CTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTA

CGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCG

CCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGAC

TCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATA

AGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATT

TAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGC

TCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTG

TGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT

AGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCA

GTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGA

GGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT

AGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGA

GACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCC

GGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCT

GCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCA

AGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCG

TGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCG

GGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCA

CCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAC

GCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAG

CACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCAT

CAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGG

CGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAA

TGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCA

GGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTG

ACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTA

TCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAA

GCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGG

TTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGAT

CTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACA

AATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTA

GTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTC

TAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCG

CTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGC

CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC

GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCG

GTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT

TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG

AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG

GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA

GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA

GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC

GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC

TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG

CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG

TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT

TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT

ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT

TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT

TTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT

TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT

TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGA

AGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC

TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT

GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG

CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC

AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC

CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG

CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG

GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTG

TGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCC

GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCAT

CCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT

GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC

ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT

CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT

GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC

AAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC

TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA

TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAA

AAGTGCCACCTGACGTC

Note: CMV promoter sequence is marked in dark red; VEEV 5' and 3' UTR sequence are marked in orange; VEEV NSP 1-4 sequence is marked in green; 26S promoter sequence is marked in blue; GFP sequence is marked in red; Packaging Signal (PS9) is marked in purple; polyadenylation tail is marked in light blue; pcDNA3.1 (+) backbone sequence is marked in black.

APPENDIX C

Full sequences of pcDNA3.1(+)-VEE-Luc-PS9-ΔNdeI plasmid

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTG

ATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGT

AGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATG

AAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATAT

ACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA

GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT

GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA

GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT

GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTC

AATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT

CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT

GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC

ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA

AATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG

GAGGTCTATATAAGCAGAGCTCTACGACTCACTATAGATGGGCGGCGCATGAGAGA

AGCCCAGACCAATTACCTACCCAAAATGGAGAAAGTTCACGTTGACATCGAGGAAG

ACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTTCCCGCAGTTTGAGGTAGAAGCCA

AGCAGGTCACTGATAATGACCATGCTAATGCCAGAGCGTTTTCGCATCTGGCTTCAA

AACTGATCGAAACGGAGGTGGACCCATCCGACACGATCCTTGACATTGGAAGTGCG

CCCGCCCGCAGAATGTATTCTAAGCACAAGTATCATTGTATCTGTCCGATGAGATGT

GCGGAAGATCCGGACAGATTGTATAAGTATGCAACTAAGCTGAAGAAAAACTGTAA

GGAAATAACTGATAAGGAATTGGACAAGAAAATGAAGGAGCTGGCCGCCGTCATGA

GCGACCCTGACCTGGAAACTGAGACTATGTGCCTCCACGACGACGAGTCGTGTCGCT

ACGAAGGGCAAGTCGCTGTTTACCAGGATGTATACGCGGTTGACGGACCGACAAGT

CTCTATCACCAAGCCAATAAGGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACC

ACCCCTTTTATGTTTAAGAACTTGGCTGGAGCATATCCATCATACTCTACCAACTGG

GCCGACGAAACCGTGTTAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATG

GAGCGGTCACGTAGAGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAA

CAATGTTCTATTCTCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAG

GAGCTGGCACCTGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCG

GTGTGAGACTATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCC

AGGCCTGTATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTT

GTGCTGCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCAC

GTATGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAG

TGCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAACG

GTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGGCCC

AGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAAGGCC

ACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAGAAGGCA

CAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAAGTGAACA

GCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGGAGATCGGGC

TGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGTCACCTCTCATT

ACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGGCTAAGGAGGTGC

GTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGCTGATGTTGAGGAG

CCCACTCTGGAGGCAGACGTCGACTTGATGTTACAAGAGGCTGGGGCCGGCTCAGT

GGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATGGCGAGGACAAGATCG

GCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAGTGAAAAATTATCTTGCAT

CCACCCTCTCGCTGAACAAGTCATAGTGATAACACACTCTGGCCGAAAAGGGCGTT

ATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGCCAGAGGGACATGCAATACCC

GTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCACCATTGTGTACAACGAACGTGAG

TTCGTAAACAGGTACCTGCACCATATTGCCACACATGGAGGAGCGCTGAACACTGAT

GAAGAATATTACAAAACTGTCAAGCCCAGCGAGCACGACGGCGAATACCTGTACGA

CATCGACAGGAAACAGTGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCACAG

GCGAGCTGGTGGATCCTCCCTTCCATGAATTCGCCTACGAGAGTCTGAGAACACGAC

CAGCCGCTCCTTACCAAGTACCAACCATAGGGGTGTATGGCGTGCCAGGATCAGGC

AAGTCTGGCATCATTAAAAGCGCAGTCACCAAAAAAGATCTAGTGGTGAGCGCCAA

GAAAGAAAACTGTGCAGAAATTATAAGGGACGTCAAGAAAATGAAAGGGCTGGAC

GTCAATGCCAGAACTGTGGACTCAGTGCTCTTGAATGGATGCAAACACCCCGTAGA

GACCCTGTATATTGACGAAGCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCAT

AGCCATTATAAGACCTAAAAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTT

TTTTTAACATGATGTGCCTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCT

TCCACAAAAGCATCTCTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTT

GTTTTACGACAAAAAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTG

ACACTACCGGCAGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAG

GGTGGGTGAAGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCT

GCCTCTCAAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGA

AAATCCTCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGA

GGACCGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTG

CCAAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT

GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATAA

GGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCATAG

ACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCTCAC

TCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATCTGGAC

TCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCACTGGGATA

ACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTCAGCTCTCTC

GCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTATGACATGAAC

ACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCTGTAAACAGAAG

ACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGAGTGACTTTTCTTCA

TTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGGGGAAAAGTTGTCCGTC

CCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAGGCTACCTTCAGAGCTCGG

CTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATGACATAATATTTGTTAATGTG

AGGACCCCATATAAATACCATCACTATCAGCAGTGTGAAGACCATGCCATTAAGCTT

AGCATGTTGACCAAGAAAGCTTGTCTGCATCTGAATCCCGGCGGAACCTGTGTCAGC

ATAGGTTATGGTTACGCTGACAGGGCCAGCGAAAGCATCATTGGTGCTATAGCGCG

GCAGTTCAAGTTTTCCCGGGTATGCAAACCGAAATCCTCACTTGAAGAGACGGAAGT

TCTGTTTGTATTCATTGGGTACGATCGCAAGGCCCGTACGCACAATTCTTACAAGCTT

TCATCAACCTTGACCAACATTTATACAGGTTCCAGACTCCACGAAGCCGGATGTGCA

CCCTCATATCATGTGGTGCGAGGGGATATTGCCACGGCCACCGAAGGAGTGATTATA

AATGCTGCTAACAGCAAAGGACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAA

GAAATTCCCGGAAAGCTTCGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGG

TCAAAGGTGCAGCTAAACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTT

CGGAGGTTGAAGGTGACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATT

GTCAACGATAACAATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTT

TCCGGGAACAAAGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGAC

ACCACTGATGCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCT

CAAGGAAGCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGAC

TCTTCAGTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTG

GCTGGAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGG

GACCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCG

TTGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAGC

AGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCACCTAGCAC

GCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAAAAGC

CTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTATAGAAT

CACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGAAAGTGCC

TGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGACGAGACTCC

GGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACAACCACCACTTA

TAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCATCATCGAAGAGGAA

GAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCACCAGGTGCTGCAAGT

CGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCATCCTGGTCCATTCCTCA

TGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGACACCCTGGAGGGAGCTAG

CGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTCTTACTTCGCAAAGAGTATGG

AGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACAGTATTCAGGAACCCTCCACATC

CCGCTCCGCGCACAAGAACACCGTCACTTGCACCCAGCAGGGCCTGCTCGAGAACC

AGCCTAGTTTCCACCCCGCCAGGCGTGAATAGGGTGATCACTAGAGAGGAGCTCGA

GGCGCTTACCCCGTCACGCACTCCTAGCAGGTCGGTCTCGAGAACCAGCCTGGTCTC

CAACCCGCCAGGCGTAAATAGGGTGATTACAAGAGAGGAGTTTGAGGCGTTCGTAG

CACAACAACAATGACGGTTTGATGCGGGTGCATACATCTTTTCCTCCGACACCGGTC

AAGGGCATTTACAACAAAAATCAGTAAGGCAAACGGTGCTATCCGAAGTGGTGTTG

GAGAGGACCGAATTGGAGATTTCGTATGCCCCGCGCCTCGACCAAGAAAAAGAAGA

ATTACTACGCAAGAAATTACAGTTAAATCCCACACCTGCTAACAGAAGCAGATACC

AGTCCAGGAAGGTGGAGAACATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGC

CTAGGGCATTATTTGAAGGCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCC

TGTTCCTTTGTATTCATCTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTG

GAAGCCTGTAACGCCATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATT

ATTCCAGAGTACGATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGAC

ACTGCCAGTTTTTGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTG

GAACCCACAATACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGT

CCTGGCAGCTGCCACAAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCG

TATTGGATTCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATG

AATATTGGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAAACGTGGTA

AATTACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACAT

AATTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAG

AGACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG

GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAGA

GCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGATAT

GTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATTGTGT

TCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGCTCTGAC

CGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACGCTGATTGA

GGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTAAATTTAAATT

CGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAACACAGTCATTAA

CATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGGATCACCATGTGCAG

CATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCGGACAAATTAATGGCA

GACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATTATAGATGCTGTGGTGGG

CGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGTGTGACTCCGTGACCGGCAC

AGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTTTAAGCTTGGCAAACCTCTGGC

AGCAGACGATGAACATGATGATGACAGGAGAAGGGCATTGCATGAAGAGTCAACA

CGCTGGAACCGAGTGGGTATTCTTTCAGAGCTGTGCAAGGCAGTAGAATCAAGGTA

TGAAACCGTAGGAACTTCCATCATAGTTATGGCCATGACTACTCTAGCTAGCAGTGT

TAAATCATTCAGCTACCTGAGAGGGGCCCCTATAACTCTCTACGGCTAACCTGAATG

GACTACGACATAGTCTAGTCCGCCAAGTCTAGCATATAGGCGCGCCACCATGGAAG

ATGCCAAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGACC

GCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCAT

CGCCTTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGAT

GAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGA

TCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCCCGTGTTGGGTGCCCTGT

TCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGCGCGAGCTGCTG

AACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCGTGAGCAAGAAAGGGCTGCA

AAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAAGATCATCATCATGG

ATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATT

TGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGACCGGGACAAA

ACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGCGTAGC

CCTACCGCACCGCACCGCTTGTGTCCGATTCAGTCATGCCCGCGACCCCATCTTCGG

CAACCAGATCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCTT

CGGCATGTTCACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTA

CCGCTTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGATTCAATCTGC

CCTGCTGGTGCCCACACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTA

CGACCTAAGCAACTTGCACGAGATCGCCAGCGGCGGGGCGCCGCTCAGCAAGGAGG

TAGGTGAGGCCGTGGCCAAACGCTTCCACCTACCAGGCATCCGCCAGGGCTACGGC

CTGACAGAAACAACCAGCGCCATTCTGATCACCCCCGAAGGGGACGACAAGCCTGG

CGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTGGACTTGGACACCG

GTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGGCCCCATGATC

ATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCTCTCATCGACAAGGACGG

CTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT

GGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAAC

TGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCCGGGGTCGCCGGCCTGC

CCGACGACGATGCCGGCGAGCTGCCCGCCGCAGTCGTCGTGCTGGAACACGGTAAA

ACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAA

GAAGCTGCGCGGTGGTGTTGTGTTCGTGGACGAGGTGCCTAAAGGACTGACCGGCA

AGTTGGACGCCCGCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGCGGCAAG

ATCGCCGTGTAAGAATTCGCTTATGCGGCCGCCTCCTTTAAGAAGGAGATATAGATC

TCTGTAGGTCCCAAACAAGCTAGTCTTAATGGAGTCACATTAATTGGAGAAGCCGTA

AAAACACAGTTCAATTATTATAAGAAAGTTGATGGTGTTGTCCAACAATTACCTGAA

ACTTACTTTACTCAGAGTAGAAATTTACAAGAATTTAAACCCAGGAGTCAAATGGAA

ATTGATTTCTTAGAATTAGCTATGGATGAATTCATTGAACGGTATAAATTAGAAGGC

TATGCCTTCGAACATATCGTTTATGGAGATTTTAGTCATAGTCAGTTAGGTGGTTTAC

ATCTACTGATTGGACTAGCTAAACGTTTTAAGGAATCACCTTTTGAATTAGAAGATT

TTATTCCTATGGACAGTACAGTTAAAAACTATTTCATAACAGATGCGCAAACAGGTT

CATCTAAGTGTGTGTGTTCTGTTATTGATTTATTACTTGATGATTTTGTTGAAATAAT

AAAATCCCAAGATTTATCTGTAGTTTCTAAGGTTGTCAAAGTGACTATTGACTATAC

AGAAATTTCATTTATGCTTTGGTGTAAAGATGGCCATGTAGAAACATTTTACCCAAA

ATTACAATCTAGTCAAGCGTGGCAACCGGGTGTTGCTATGCCTAATCTTTACAAAAT

GCAAAGAATGCTATTAGAAAAGTGTGACCTTCAAAATTATGGTGATAGTGCAACATT

ACCTAAAGGCATAATGATGAATGTCGCAAAATATACTCAACTGTGTCAATATTTAAA

CACATTAACATTAGCTGTACCCTATAATATGAGAGTTATACATTTTGGTGCTGGTTCT

GATAAAGGAGTTGCACCAGGTACAGCTGTTTTAAGACAGTGGTTGCCTACGGGTAC

GCTGCTTGTCGATTCAGATCTTAATGACTTTGTCTCTGATGCAGATTCAACTTTGATT

GGTGATTGTGCAACTGTACATACAGCTAATAAATGGGATCTCATTATTAGTGATATG

TACGACCCTAAGACTAAAAATGTTACAAAAGAAAATGACTCTAAAGAGGGTTTTTTC

ACTTACATTTGTGGGTTTATACAACAAAAGCTAGCTCTTGGAGGTTCCGTGGCTATA

AAGATAGATCCCAACTCCATAAGTGTGTCGACATATAAGGATCCACTCGACGACTCG

ACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATG

TTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCT

TCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGA

GGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGC

AACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGC

TTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGG

ACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGAC

GTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTC

TGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCG

GCTCTGCGGCCTCTTCCGCATCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTT

GGGCCGCCTCCCCGCCTGGAATTAATTCGAGCTCGGTACCTTGATTCTCGAGGAATT

GGCAAGCTGCTTACATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATT

TTATTTTTTCTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTCAAAAAAAAAAAA

AAAAAAAAAAAAAACGCGTCGAGGGGAATTAATTCTTGAAGACGAAAGGGCCAGG

TGGCACTTTTCGGGGAAATGTGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTC

GACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTG

ACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCG

CATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA

GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG

CTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGT

AGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACT

TGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTC

GCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGT

GCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG

CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATA

GTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGA

TTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACA

AAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCC

CCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAAC

CAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATC

TCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTC

CGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGA

GGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGG

AGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGA

TCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAG

GTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACA

ATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTT

TTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCG

GCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCAC

TGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGT

CATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGC

TGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCG

AGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAA

GAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCC

CGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGT

GGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCG

CTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATG

GGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGC

CTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACC

GACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTAT

GAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGC

GGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATG

GTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGC

ATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTC

GACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG

TTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCT

GGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT

TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGG

AGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT

CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT

CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAA

GGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC

TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC

TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA

CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC

TCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG

CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG

TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA

CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC

CTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG

TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA

GCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAA

GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA

GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA

AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAC

CAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG

TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCC

CCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAA

TAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC

TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT

AGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTG

GTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCA

TGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT

TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT

GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGA

ATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC

GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAA

AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC

CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG

GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT

CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGC

GGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT

CCCCGAAAAGTGCCACCTGACGTC

Note: CMV promoter sequence is marked in dark red; VEEV 5’ and 3’ UTR sequence are marked in orange; VEEV NSP 1-4 sequence is marked in green; 26S promoter sequence is marked in blue;

luciferase sequence is marked in red; Packaging Signal (PS9) is marked in purple; polyadenylation tail is marked in light blue; pcDNA3.1 (+) backbone sequence is marked in black.

APPENDIX D

Full sequences of pcDNA3.1(+)-VEE-GFP-ΔPS9 plasmid

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTG

ATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGT

AGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATG

AAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATAT

ACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA

GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT

GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA

GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT

GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTC

AATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT

CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT

GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC

ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA

AATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG

GAGGTCTATATAAGCAGAGCTCTACGACTCACTATAGATGGGCGGCGCATGAGAGA

AGCCCAGACCAATTACCTACCCAAAATGGAGAAAGTTCACGTTGACATCGAGGAAG

ACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTTCCCGCAGTTTGAGGTAGAAGCCA

AGCAGGTCACTGATAATGACCATGCTAATGCCAGAGCGTTTTCGCATCTGGCTTCAA

AACTGATCGAAACGGAGGTGGACCCATCCGACACGATCCTTGACATTGGAAGTGCG

CCCGCCCGCAGAATGTATTCTAAGCACAAGTATCATTGTATCTGTCCGATGAGATGT

GCGGAAGATCCGGACAGATTGTATAAGTATGCAACTAAGCTGAAGAAAAACTGTAA

GGAAATAACTGATAAGGAATTGGACAAGAAAATGAAGGAGCTGGCCGCCGTCATGA

GCGACCCTGACCTGGAAACTGAGACTATGTGCCTCCACGACGACGAGTCGTGTCGCT

ACGAAGGGCAAGTCGCTGTTTACCAGGATGTATACGCGGTTGACGGACCGACAAGT

CTCTATCACCAAGCCAATAAGGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACC

ACCCCTTTTATGTTTAAGAACTTGGCTGGAGCATATCCATCATACTCTACCAACTGG

GCCGACGAAACCGTGTTAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATG

GAGCGGTCACGTAGAGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAA

CAATGTTCTATTCTCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAG

GAGCTGGCACCTGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCG

GTGTGAGACTATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCC

AGGCCTGTATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTT

GTGCTGCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCAC

GTATGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAG

TGCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAACG

GTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGGCCC

AGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAAGGCC

ACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAGAAGGCA

CAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAAGTGAACA

GCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGGAGATCGGGC

TGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGTCACCTCTCATT

ACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGGCTAAGGAGGTGC

GTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGCTGATGTTGAGGAG

CCCACTCTGGAGGCAGACGTCGACTTGATGTTACAAGAGGCTGGGGCCGGCTCAGT

GGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATGGCGAGGACAAGATCG

GCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAGTGAAAAATTATCTTGCAT

CCACCCTCTCGCTGAACAAGTCATAGTGATAACACACTCTGGCCGAAAAGGGCGTT

ATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGCCAGAGGGACATGCAATACCC

GTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCACCATTGTGTACAACGAACGTGAG

TTCGTAAACAGGTACCTGCACCATATTGCCACACATGGAGGAGCGCTGAACACTGAT

GAAGAATATTACAAAACTGTCAAGCCCAGCGAGCACGACGGCGAATACCTGTACGA

CATCGACAGGAAACAGTGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCACAG

GCGAGCTGGTGGATCCTCCCTTCCATGAATTCGCCTACGAGAGTCTGAGAACACGAC

CAGCCGCTCCTTACCAAGTACCAACCATAGGGGTGTATGGCGTGCCAGGATCAGGC

AAGTCTGGCATCATTAAAAGCGCAGTCACCAAAAAAGATCTAGTGGTGAGCGCCAA

GAAAGAAAACTGTGCAGAAATTATAAGGGACGTCAAGAAAATGAAAGGGCTGGAC

GTCAATGCCAGAACTGTGGACTCAGTGCTCTTGAATGGATGCAAACACCCCGTAGA

GACCCTGTATATTGACGAAGCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCAT

AGCCATTATAAGACCTAAAAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTT

TTTTTAACATGATGTGCCTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCT

TCCACAAAAGCATCTCTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTT

GTTTTACGACAAAAAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTG

ACACTACCGGCAGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAG

GGTGGGTGAAGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCT

GCCTCTCAAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGA

AAATCCTCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGA

GGACCGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTG

CCAAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT

GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATAA

GGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCATAG

ACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCTCAC

TCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATCTGGAC

TCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCACTGGGATA

ACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTCAGCTCTCTC

GCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTATGACATGAAC

ACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCTGTAAACAGAAG

ACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGAGTGACTTTTCTTCA

TTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGGGGAAAAGTTGTCCGTC

CCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAGGCTACCTTCAGAGCTCGG

CTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATGACATAATATTTGTTAATGTG

AGGACCCCATATAAATACCATCACTATCAGCAGTGTGAAGACCATGCCATTAAGCTT

AGCATGTTGACCAAGAAAGCTTGTCTGCATCTGAATCCCGGCGGAACCTGTGTCAGC

ATAGGTTATGGTTACGCTGACAGGGCCAGCGAAAGCATCATTGGTGCTATAGCGCG

GCAGTTCAAGTTTTCCCGGGTATGCAAACCGAAATCCTCACTTGAAGAGACGGAAGT

TCTGTTTGTATTCATTGGGTACGATCGCAAGGCCCGTACGCACAATTCTTACAAGCTT

TCATCAACCTTGACCAACATTTATACAGGTTCCAGACTCCACGAAGCCGGATGTGCA

CCCTCATATCATGTGGTGCGAGGGGATATTGCCACGGCCACCGAAGGAGTGATTATA

AATGCTGCTAACAGCAAAGGACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAA

GAAATTCCCGGAAAGCTTCGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGG

TCAAAGGTGCAGCTAAACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTT

CGGAGGTTGAAGGTGACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATT

GTCAACGATAACAATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTT

TCCGGGAACAAAGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGAC

ACCACTGATGCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCT

CAAGGAAGCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGAC

TCTTCAGTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTG

GCTGGAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGG

GACCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCG

TTGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAGC

AGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCACCTAGCAC

GCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAAAAGC

CTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTATAGAAT

CACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGAAAGTGCC

TGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGACGAGACTCC

GGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACAACCACCACTTA

TAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCATCATCGAAGAGGAA

GAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCACCAGGTGCTGCAAGT

CGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCATCCTGGTCCATTCCTCA

TGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGACACCCTGGAGGGAGCTAG

CGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTCTTACTTCGCAAAGAGTATGG

AGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACAGTATTCAGGAACCCTCCACATC

CCGCTCCGCGCACAAGAACACCGTCACTTGCACCCAGCAGGGCCTGCTCGAGAACC

AGCCTAGTTTCCACCCCGCCAGGCGTGAATAGGGTGATCACTAGAGAGGAGCTCGA

GGCGCTTACCCCGTCACGCACTCCTAGCAGGTCGGTCTCGAGAACCAGCCTGGTCTC

CAACCCGCCAGGCGTAAATAGGGTGATTACAAGAGAGGAGTTTGAGGCGTTCGTAG

CACAACAACAATGACGGTTTGATGCGGGTGCATACATCTTTTCCTCCGACACCGGTC

AAGGGCATTTACAACAAAAATCAGTAAGGCAAACGGTGCTATCCGAAGTGGTGTTG

GAGAGGACCGAATTGGAGATTTCGTATGCCCCGCGCCTCGACCAAGAAAAAGAAGA

ATTACTACGCAAGAAATTACAGTTAAATCCCACACCTGCTAACAGAAGCAGATACC

AGTCCAGGAAGGTGGAGAACATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGC

CTAGGGCATTATTTGAAGGCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCC

TGTTCCTTTGTATTCATCTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTG

GAAGCCTGTAACGCCATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATT

ATTCCAGAGTACGATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGAC

ACTGCCAGTTTTTGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTG

GAACCCACAATACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGT

CCTGGCAGCTGCCACAAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCG

TATTGGATTCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATG

AATATTGGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAAACGTGGTA

AATTACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACAT

AATTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAG

AGACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG

GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAGA

GCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGATAT

GTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATTGTGT

TCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGCTCTGAC

CGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACGCTGATTGA

GGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTAAATTTAAATT

CGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAACACAGTCATTAA

CATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGGATCACCATGTGCAG

CATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCGGACAAATTAATGGCA

GACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATTATAGATGCTGTGGTGGG

CGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGTGTGACTCCGTGACCGGCAC

AGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTTTAAGCTTGGCAAACCTCTGGC

AGCAGACGATGAACATGATGATGACAGGAGAAGGGCATTGCATGAAGAGTCAACA

CGCTGGAACCGAGTGGGTATTCTTTCAGAGCTGTGCAAGGCAGTAGAATCAAGGTA

TGAAACCGTAGGAACTTCCATCATAGTTATGGCCATGACTACTCTAGCTAGCAGTGT

TAAATCATTCAGCTACCTGAGAGGGGCCCCTATAACTCTCTACGGCTAACCTGAATG

GACTACGACATAGTCTAGTCCGCCAAGTCTAGCATATAGGCGCGCCCTCAGCATCGA

TTCAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCC

ATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGA

GGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG

GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT

GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC

CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG

ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA

GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACT

ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTG

AACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA

CCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACC

TGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC

CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA

GTAGTCTAGAGTCGACCCGGGCGGCCGCTTGTAGAAGAAACACATGGCGCGCCTGA

GGATCCACTCGACGACTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTG

ACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGC

CTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCC

TGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTG

TGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAG

CTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCG

CCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG

TGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGAT

TCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCT

TCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCATCTTCGCCTTCGCCCTCAGA

CGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGGAATTAATTCGAGCTCGGTAC

CTTGATTCTCGAGGAATTGGCAAGCTGCTTACATAGAACTCGCGGCGATTGGCATGC

CGCCTTAAAATTTTTATTTTATTTTTTCTTTTCTTTTCCGAATCGGATTTTGTTTTTAAT

ATTTCAAAAAAAAAAAAAAAAAAAAAAAACGCGTCGAGGGGAATTAATTCTTG

AAGACGAAAGGGCCAGGTGGCACTTTTCGGGGAAATGTGTCTAGAGGGCCCGTTTA

AACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCC

TCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA

ATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGG

TGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA

TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGT

ATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGC

AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT

CCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTT

AGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGA

TGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA

GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTAT

CTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAA

AATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGT

TAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAT

CTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAG

TATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCC

CATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATT

TTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGT

GAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATAT

CCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAG

ATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACT

GGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAG

GGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAG

GACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGT

GCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGG

GGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTG

ATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAG

CGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAG

GATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCT

CAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCT

TGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGC

TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAG

AGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCG

ATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTG

GGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTC

CACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTG

GATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTT

TATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAA

AGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAT

CATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTG

TTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGC

ATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTG

CGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC

GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTC

ACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG

GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC

AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC

ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG

CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT

GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG

GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC

GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCC

TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG

GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA

GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTG

CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA

ACAAACCACCGCTGGTAGCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA

AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA

CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTA

GATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC

TTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCT

ATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA

GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGC

TCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTC

CTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAA

GTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG

TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC

GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA

TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGC

ATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC

AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC

AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAA

AACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGA

TGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC

TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA

CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGG

GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG

GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Note: CMV promoter sequence is marked in dark red; VEEV 5' and 3' UTR sequence are marked in orange; VEEV NSP 1-4 sequence is marked in green; 26S promoter sequence is marked in blue; GFP sequence is marked in red; polyadenylation tail is marked in light blue; pcDNA3.1 (+) backbone sequence is marked in black.

APPENDIX E

Full sequences of pcDNA3.1(+)-VEE-Luc-ΔPS9 plasmid

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTG

ATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGT

AGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATG

AAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATAT

ACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA

GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT

GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA

GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT

GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTC

AATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT

CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT

GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC

ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA

AATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG

GAGGTCTATATAAGCAGAGCTCTACGACTCACTATAGATGGGCGGCGCATGAGAGA

AGCCCAGACCAATTACCTACCCAAAATGGAGAAAGTTCACGTTGACATCGAGGAAG

ACAGCCCATTCCTCAGAGCTTTGCAGCGGAGCTTCCCGCAGTTTGAGGTAGAAGCCA

AGCAGGTCACTGATAATGACCATGCTAATGCCAGAGCGTTTTCGCATCTGGCTTCAA

AACTGATCGAAACGGAGGTGGACCCATCCGACACGATCCTTGACATTGGAAGTGCG

CCCGCCCGCAGAATGTATTCTAAGCACAAGTATCATTGTATCTGTCCGATGAGATGT

GCGGAAGATCCGGACAGATTGTATAAGTATGCAACTAAGCTGAAGAAAAACTGTAA

GGAAATAACTGATAAGGAATTGGACAAGAAAATGAAGGAGCTGGCCGCCGTCATGA

GCGACCCTGACCTGGAAACTGAGACTATGTGCCTCCACGACGACGAGTCGTGTCGCT

ACGAAGGGCAAGTCGCTGTTTACCAGGATGTATACGCGGTTGACGGACCGACAAGT

CTCTATCACCAAGCCAATAAGGGAGTTAGAGTCGCCTACTGGATAGGCTTTGACACC

ACCCCTTTTATGTTTAAGAACTTGGCTGGAGCATATCCATCATACTCTACCAACTGG

GCCGACGAAACCGTGTTAACGGCTCGTAACATAGGCCTATGCAGCTCTGACGTTATG

GAGCGGTCACGTAGAGGGATGTCCATTCTTAGAAAGAAGTATTTGAAACCATCCAA

CAATGTTCTATTCTCTGTTGGCTCGACCATCTACCACGAGAAGAGGGACTTACTGAG

GAGCTGGCACCTGCCGTCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCG

GTGTGAGACTATAGTTAGTTGCGACGGGTACGTCGTTAAAAGAATAGCTATCAGTCC

AGGCCTGTATGGGAAGCCTTCAGGCTATGCTGCTACGATGCACCGCGAGGGATTCTT

GTGCTGCAAAGTGACAGACACATTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGCAC

GTATGTGCCAGCTACATTGTGTGACCAAATGACTGGCATACTGGCAACAGATGTCAG

TGCGGACGACGCGCAAAAACTGCTGGTTGGGCTCAACCAGCGTATAGTCGTCAACG

GTCGCACCCAGAGAAACACCAATACCATGAAAAATTACCTTTTGCCCGTAGTGGCCC

AGGCATTTGCTAGGTGGGCAAAGGAATATAAGGAAGATCAAGAAGATGAAAGGCC

ACTAGGACTACGAGATAGACAGTTAGTCATGGGGTGTTGTTGGGCTTTTAGAAGGCA

CAAGATAACATCTATTTATAAGCGCCCGGATACCCAAACCATCATCAAAGTGAACA

GCGATTTCCACTCATTCGTGCTGCCCAGGATAGGCAGTAACACATTGGAGATCGGGC

TGAGAACAAGAATCAGGAAAATGTTAGAGGAGCACAAGGAGCCGTCACCTCTCATT

ACCGCCGAGGACGTACAAGAAGCTAAGTGCGCAGCCGATGAGGCTAAGGAGGTGC

GTGAAGCCGAGGAGTTGCGCGCAGCTCTACCACCTTTGGCAGCTGATGTTGAGGAG

CCCACTCTGGAGGCAGACGTCGACTTGATGTTACAAGAGGCTGGGGCCGGCTCAGT

GGAGACACCTCGTGGCTTGATAAAGGTTACCAGCTACGATGGCGAGGACAAGATCG

GCTCTTACGCTGTGCTTTCTCCGCAGGCTGTACTCAAGAGTGAAAAATTATCTTGCAT

CCACCCTCTCGCTGAACAAGTCATAGTGATAACACACTCTGGCCGAAAAGGGCGTT

ATGCCGTGGAACCATACCATGGTAAAGTAGTGGTGCCAGAGGGACATGCAATACCC

GTCCAGGACTTTCAAGCTCTGAGTGAAAGTGCCACCATTGTGTACAACGAACGTGAG

TTCGTAAACAGGTACCTGCACCATATTGCCACACATGGAGGAGCGCTGAACACTGAT

GAAGAATATTACAAAACTGTCAAGCCCAGCGAGCACGACGGCGAATACCTGTACGA

CATCGACAGGAAACAGTGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCACAG

GCGAGCTGGTGGATCCTCCCTTCCATGAATTCGCCTACGAGAGTCTGAGAACACGAC

CAGCCGCTCCTTACCAAGTACCAACCATAGGGGTGTATGGCGTGCCAGGATCAGGC

AAGTCTGGCATCATTAAAAGCGCAGTCACCAAAAAAGATCTAGTGGTGAGCGCCAA

GAAAGAAAACTGTGCAGAAATTATAAGGGACGTCAAGAAAATGAAAGGGCTGGAC

GTCAATGCCAGAACTGTGGACTCAGTGCTCTTGAATGGATGCAAACACCCCGTAGA

GACCCTGTATATTGACGAAGCTTTTGCTTGTCATGCAGGTACTCTCAGAGCGCTCAT

AGCCATTATAAGACCTAAAAAGGCAGTGCTCTGCGGGGATCCCAAACAGTGCGGTT

TTTTTAACATGATGTGCCTGAAAGTGCATTTTAACCACGAGATTTGCACACAAGTCT

TCCACAAAAGCATCTCTCGCCGTTGCACTAAATCTGTGACTTCGGTCGTCTCAACCTT

GTTTTACGACAAAAAAATGAGAACGACGAATCCGAAAGAGACTAAGATTGTGATTG

ACACTACCGGCAGTACCAAACCTAAGCAGGACGATCTCATTCTCACTTGTTTCAGAG

GGTGGGTGAAGCAGTTGCAAATAGATTACAAAGGCAACGAAATAATGACGGCAGCT

GCCTCTCAAGGGCTGACCCGTAAAGGTGTGTATGCCGTTCGGTACAAGGTGAATGA

AAATCCTCTGTACGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCACGGA

GGACCGCATCGTGTGGAAAACACTAGCCGGCGACCCATGGATAAAAACACTGACTG

CCAAGTACCCTGGGAATTTCACTGCCACGATAGAGGAGTGGCAAGCAGAGCATGAT

GCCATCATGAGGCACATCTTGGAGAGACCGGACCCTACCGACGTCTTCCAGAATAA

GGCAAACGTGTGTTGGGCCAAGGCTTTAGTGCCGGTGCTGAAGACCGCTGGCATAG

ACATGACCACTGAACAATGGAACACTGTGGATTATTTTGAAACGGACAAAGCTCAC

TCAGCAGAGATAGTATTGAACCAACTATGCGTGAGGTTCTTTGGACTCGATCTGGAC

TCCGGTCTATTTTCTGCACCCACTGTTCCGTTATCCATTAGGAATAATCACTGGGATA

ACTCCCCGTCGCCTAACATGTACGGGCTGAATAAAGAAGTGGTCCGTCAGCTCTCTC

GCAGGTACCCACAACTGCCTCGGGCAGTTGCCACTGGAAGAGTCTATGACATGAAC

ACTGGTACACTGCGCAATTATGATCCGCGCATAAACCTAGTACCTGTAAACAGAAG

ACTGCCTCATGCTTTAGTCCTCCACCATAATGAACACCCACAGAGTGACTTTTCTTCA

TTCGTCAGCAAATTGAAGGGCAGAACTGTCCTGGTGGTCGGGGAAAAGTTGTCCGTC

CCAGGCAAAATGGTTGACTGGTTGTCAGACCGGCCTGAGGCTACCTTCAGAGCTCGG

CTGGATTTAGGCATCCCAGGTGATGTGCCCAAATATGACATAATATTTGTTAATGTG

AGGACCCCATATAAATACCATCACTATCAGCAGTGTGAAGACCATGCCATTAAGCTT

AGCATGTTGACCAAGAAAGCTTGTCTGCATCTGAATCCCGGCGGAACCTGTGTCAGC

ATAGGTTATGGTTACGCTGACAGGGCCAGCGAAAGCATCATTGGTGCTATAGCGCG

GCAGTTCAAGTTTTCCCGGGTATGCAAACCGAAATCCTCACTTGAAGAGACGGAAGT

TCTGTTTGTATTCATTGGGTACGATCGCAAGGCCCGTACGCACAATTCTTACAAGCTT

TCATCAACCTTGACCAACATTTATACAGGTTCCAGACTCCACGAAGCCGGATGTGCA

CCCTCATATCATGTGGTGCGAGGGGATATTGCCACGGCCACCGAAGGAGTGATTATA

AATGCTGCTAACAGCAAAGGACAACCTGGCGGAGGGGTGTGCGGAGCGCTGTATAA

GAAATTCCCGGAAAGCTTCGATTTACAGCCGATCGAAGTAGGAAAAGCGCGACTGG

TCAAAGGTGCAGCTAAACATATCATTCATGCCGTAGGACCAAACTTCAACAAAGTTT

CGGAGGTTGAAGGTGACAAACAGTTGGCAGAGGCTTATGAGTCCATCGCTAAGATT

GTCAACGATAACAATTACAAGTCAGTAGCGATTCCACTGTTGTCCACCGGCATCTTT

TCCGGGAACAAAGATCGACTAACCCAATCATTGAACCATTTGCTGACAGCTTTAGAC

ACCACTGATGCAGATGTAGCCATATACTGCAGGGACAAGAAATGGGAAATGACTCT

CAAGGAAGCAGTGGCTAGGAGAGAAGCAGTGGAGGAGATATGCATATCCGACGAC

TCTTCAGTGACAGAACCTGATGCAGAGCTGGTGAGGGTGCATCCGAAGAGTTCTTTG

GCTGGAAGGAAGGGCTACAGCACAAGCGATGGCAAAACTTTCTCATATTTGGAAGG

GACCAAGTTTCACCAGGCGGCCAAGGATATAGCAGAAATTAATGCCATGTGGCCCG

TTGCAACGGAGGCCAATGAGCAGGTATGCATGTATATCCTCGGAGAAAGCATGAGC

AGTATTAGGTCGAAATGCCCCGTCGAAGAGTCGGAAGCCTCCACACCACCTAGCAC

GCTGCCTTGCTTGTGCATCCATGCCATGACTCCAGAAAGAGTACAGCGCCTAAAAGC

CTCACGTCCAGAACAAATTACTGTGTGCTCATCCTTTCCATTGCCGAAGTATAGAAT

CACTGGTGTGCAGAAGATCCAATGCTCCCAGCCTATATTGTTCTCACCGAAAGTGCC

TGCGTATATTCATCCAAGGAAGTATCTCGTGGAAACACCACCGGTAGACGAGACTCC

GGAGCCATCGGCAGAGAACCAATCCACAGAGGGGACACCTGAACAACCACCACTTA

TAACCGAGGATGAGACCAGGACTAGAACGCCTGAGCCGATCATCATCGAAGAGGAA

GAAGAGGATAGCATAAGTTTGCTGTCAGATGGCCCGACCCACCAGGTGCTGCAAGT

CGAGGCAGACATTCACGGGCCGCCCTCTGTATCTAGCTCATCCTGGTCCATTCCTCA

TGCATCCGACTTTGATGTGGACAGTTTATCCATACTTGACACCCTGGAGGGAGCTAG

CGTGACCAGCGGGGCAACGTCAGCCGAGACTAACTCTTACTTCGCAAAGAGTATGG

AGTTTCTGGCGCGACCGGTGCCTGCGCCTCGAACAGTATTCAGGAACCCTCCACATC

CCGCTCCGCGCACAAGAACACCGTCACTTGCACCCAGCAGGGCCTGCTCGAGAACC

AGCCTAGTTTCCACCCCGCCAGGCGTGAATAGGGTGATCACTAGAGAGGAGCTCGA

GGCGCTTACCCCGTCACGCACTCCTAGCAGGTCGGTCTCGAGAACCAGCCTGGTCTC

CAACCCGCCAGGCGTAAATAGGGTGATTACAAGAGAGGAGTTTGAGGCGTTCGTAG

CACAACAACAATGACGGTTTGATGCGGGTGCATACATCTTTTCCTCCGACACCGGTC

AAGGGCATTTACAACAAAAATCAGTAAGGCAAACGGTGCTATCCGAAGTGGTGTTG

GAGAGGACCGAATTGGAGATTTCGTATGCCCCGCGCCTCGACCAAGAAAAAGAAGA

ATTACTACGCAAGAAATTACAGTTAAATCCCACACCTGCTAACAGAAGCAGATACC

AGTCCAGGAAGGTGGAGAACATGAAAGCCATAACAGCTAGACGTATTCTGCAAGGC

CTAGGGCATTATTTGAAGGCAGAAGGAAAAGTGGAGTGCTACCGAACCCTGCATCC

TGTTCCTTTGTATTCATCTAGTGTGAACCGTGCCTTTTCAAGCCCCAAGGTCGCAGTG

GAAGCCTGTAACGCCATGTTGAAAGAGAACTTTCCGACTGTGGCTTCTTACTGTATT

ATTCCAGAGTACGATGCCTATTTGGACATGGTTGACGGAGCTTCATGCTGCTTAGAC

ACTGCCAGTTTTTGCCCTGCAAAGCTGCGCAGCTTTCCAAAGAAACACTCCTATTTG

GAACCCACAATACGATCGGCAGTGCCTTCAGCGATCCAGAACACGCTCCAGAACGT

CCTGGCAGCTGCCACAAAAAGAAATTGCAATGTCACGCAAATGAGAGAATTGCCCG

TATTGGATTCGGCGGCCTTTAATGTGGAATGCTTCAAGAAATATGCGTGTAATAATG

AATATTGGGAAACGTTTAAAGAAAACCCCATCAGGCTTACTGAAGAAAACGTGGTA

AATTACATTACCAAATTAAAAGGACCAAAAGCTGCTGCTCTTTTTGCGAAGACACAT

AATTTGAATATGTTGCAGGACATACCAATGGACAGGTTTGTAATGGACTTAAAGAG

AGACGTGAAAGTGACTCCAGGAACAAAACATACTGAAGAACGGCCCAAGGTACAG

GTGATCCAGGCTGCCGATCCGCTAGCAACAGCGTATCTGTGCGGAATCCACCGAGA

GCTGGTTAGGAGATTAAATGCGGTCCTGCTTCCGAACATTCATACACTGTTTGATAT

GTCGGCTGAAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCTGGGGATTGTGT

TCTGGAAACTGACATCGCGTCGTTTGATAAAAGTGAGGACGACGCCATGGCTCTGAC

CGCGTTAATGATTCTGGAAGACTTAGGTGTGGACGCAGAGCTGTTGACGCTGATTGA

GGCGGCTTTCGGCGAAATTTCATCAATACATTTGCCCACTAAAACTAAATTTAAATT

CGGAGCCATGATGAAATCTGGAATGTTCCTCACACTGTTTGTGAACACAGTCATTAA

CATTGTAATCGCAAGCAGAGTGTTGAGAGAACGGCTAACCGGATCACCATGTGCAG

CATTCATTGGAGATGACAATATCGTGAAAGGAGTCAAATCGGACAAATTAATGGCA

GACAGGTGCGCCACCTGGTTGAATATGGAAGTCAAGATTATAGATGCTGTGGTGGG

CGAGAAAGCGCCTTATTTCTGTGGAGGGTTTATTTTGTGTGACTCCGTGACCGGCAC

AGCGTGCCGTGTGGCAGACCCCCTAAAAAGGCTGTTTAAGCTTGGCAAACCTCTGGC

AGCAGACGATGAACATGATGATGACAGGAGAAGGGCATTGCATGAAGAGTCAACA

CGCTGGAACCGAGTGGGTATTCTTTCAGAGCTGTGCAAGGCAGTAGAATCAAGGTA

TGAAACCGTAGGAACTTCCATCATAGTTATGGCCATGACTACTCTAGCTAGCAGTGT

TAAATCATTCAGCTACCTGAGAGGGGCCCCTATAACTCTCTACGGCTAACCTGAATG

GACTACGACATAGTCTAGTCCGCCAAGTCTAGCATATAGGCGCGCCACCATGGAAG

ATGCCAAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGACC

GCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCAT

CGCCTTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGAT

GAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGA

TCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCCCGTGTTGGGTGCCCTGT

TCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGCGCGAGCTGCTG

AACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCGTGAGCAAGAAAGGGCTGCA

AAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAAGATCATCATCATGG

ATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATT

TGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGACCGGGACAAA

ACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGCGTAGC

CCTACCGCACCGCACCGCTTGTGTCCGATTCAGTCATGCCCGCGACCCCATCTTCGG

CAACCAGATCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCTT

CGGCATGTTCACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTA

CCGCTTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGATTCAATCTGC

CCTGCTGGTGCCCACACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTA

CGACCTAAGCAACTTGCACGAGATCGCCAGCGGCGGGGCGCCGCTCAGCAAGGAGG

TAGGTGAGGCCGTGGCCAAACGCTTCCACCTACCAGGCATCCGCCAGGGCTACGGC

CTGACAGAAACAACCAGCGCCATTCTGATCACCCCCGAAGGGGACGACAAGCCTGG

CGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTGGACTTGGACACCG

GTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGGCCCCATGATC

ATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCTCTCATCGACAAGGACGG

CTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT

GGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAAC

TGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGACGCCGGGGTCGCCGGCCTGC

CCGACGACGATGCCGGCGAGCTGCCCGCCGCAGTCGTCGTGCTGGAACACGGTAAA

ACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAA

GAAGCTGCGCGGTGGTGTTGTGTTCGTGGACGAGGTGCCTAAAGGACTGACCGGCA

AGTTGGACGCCCGCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGCGGCAAG

ATCGCCGTGTAAGAATTCGCTTATGCGGCCGCTTGTAGAAGAAACACATGGCGCGCC

TGAGGATCCACTCGACGACTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGA

TTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAA

TGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAA

TCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTG

GTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGT

CAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCG

CCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCG

TGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTG

GATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCT

TCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCATCTTCGCCTTCGCCCT

CAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGGAATTAATTCGAGCTCG

GTACCTTGATTCTCGAGGAATTGGCAAGCTGCTTACATAGAACTCGCGGCGATTGGC

ATGCCGCCTTAAAATTTTTATTTTATTTTTTCTTTTCTTTTCCGAATCGGATTTTGTTTT

TAATATTTCAAAAAAAAAAAAAAAAAAAAAAAAACGCGTCGAGGGGAATTAATT

CTTGAAGACGAAAGGGCCAGGTGGCACTTTTCGGGGAAATGTGTCTAGAGGGCCCG

TTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTG

CCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAA

TAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGT

GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTG

GGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGG

GGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTAC

GCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTC

CCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCC

CTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGG

GTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTT

GGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCC

TATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTA

AAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTC

AGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATG

CATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAG

AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCC

GCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT

AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAG

TAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGT

ATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAA

CAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTAT

GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCG

CAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTG

CAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGC

TGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGC

CGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGG

CTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACC

AAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGAT

CAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAG

GCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCT

GCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCC

GGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTG

AAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTC

CCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACT

CTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGA

TTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGG

CTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTT

GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAA

TAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT

TATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAG

CTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGA

AGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC

GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG

AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTC

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC

AAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT

GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT

TTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG

GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC

TCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC

TTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA

GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG

CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC

ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA

CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA

TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG

GCAAACAAACCACCGCTGGTAGCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA

GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT

GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC

ACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG

TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATC

TGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC

GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCA

CCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAG

TGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGA

GTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC

GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA

GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTC

CGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC

TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA

CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGC

GTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGG

AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC

GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTT

TCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGA

CACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCA

GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT

AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Note: CMV promoter sequence is marked in dark red; VEEV 5' and 3' UTR sequence are marked in orange; VEEV NSP 1-4 sequence is marked in green; 26S promoter sequence is marked in blue; luciferase sequence is marked in red; polyadenylation tail is marked in light blue; pcDNA3.1 (+) backbone sequence is marked in black.