

UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Aylin Sultanova

**Stable expression of Barmah Forest virus
replicons in cultured mammalian cells**

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor:
MSc, Aliar Omler

Tartu 2023

Stable expression of Barmah Forest virus replicons in cultured mammalian cells

Abstract:

Viruses are infectious agents that can infect various living organisms, including humans, animals, plants, and microorganisms. Understanding their molecular biology is crucial for developing effective prevention and treatment strategies. Alphaviruses are a group of mosquito-borne viruses that cause diseases such as rash, arthritis, and encephalitis. Barmah Forest virus (BFV) is an alphavirus that poses a public health threat in Australia. However, research on BFV is relatively limited compared to other alphaviruses. In this study, we focus on the expression of BFV replicons in cultured mammalian cells. Replicons are valuable tools for studying viral replication and developing antiviral agents. We review the successful generation of replicons for alphaviruses and Hepatitis C virus and propose a workflow for constructing BFV replicons in cultured mammalian cells.

Keywords: alphavirus, Barmah Forest virus, alphavirus replicon, Barmah Forest virus replicon, stable cell lines

CERCS:

B230 Microbiology, bacteriology, virology, mycology

Barmah Foresti viiruse replikoni stabiilne ekspressioon imetajarakkudes

Lühikokkuvõte:

Viirused on võimelised nakatama kõiki eluvorme ning nende molekulaarbioloogia tundmine on võtmetähtsusega väljatöötamiseks efektiivseid ravimeid ning vaktsiine nende poolt põhjustatud tõvede ära hoidmisel. Alfaviiruste perekonna liikmed on suuresti sääseliste poolt levitatavad viirused, mis inimesi nakatades võivad põhjustada ulatuslikku löövet, lihasvalusid, artriiti ning ka entsefaliiti. Barmah Foresti viirus (BFV) on meditsiiniliselt oluline viirus Austraalias, kuid võrreldes mõnede teiste alfaviirustega on seda oluliselt vähem uuritud. Käesolevas töös on tähelepanu all BFV replikonide ekspressioon imetajarakkude kultuurides. Viiruste replikonid on olulised tööriistad uurimaks viiruste replikatsiooni ning viirusvastaste strateegiate avastamisel. Siinses töös kirjeldatakse varasemaid edulugusid viiruste replikonide ning neid ekspresseerivate stabiilsete rakuliinide välja töötamisel, kus luubi all on positiivse polaarsusega RNA genoomsed alfaviirused ning C-hepatiidi viirus. Varasemate saavutuste põhjal on töö raames esitletud tööplaani BFV replikonide ning nendel põhinevate stabiilsete rakuliinide loomine.

Võtmesõnad: alfaviirus, Barmah Foresti viirus, alfaviiruse replikon, Barmah Foresti viiruse replikon, stabiilsed rakuliinid

CERCS:

B230 Mikrobioloogia, bakterioloogia, viroloogia, mükoloogia

TABLE OF CONTENTS

TERMS, ABBREVIATIONS AND NOTATIONS	6
INTRODUCTION	9
1. LITERATURE REVIEW	10
1.1. Overview of alphaviruses	10
1.2. Alphavirus virion	11
1.3. Alphavirus genome	12
1.4. Non-structural proteins	13
1.5. Alphavirus infection cycle	14
1.6. Review of the Barmah Forest Virus	16
1.7. Replicons of RNA viruses as tools for in laboratory studies and biotechnological industry: benefits and limitations	16
1.8. Stable cell lines for replicons	19
1.9. Examples for uses of stable replicon cell lines	25
1.10. Barmah Forest virus (BFV) cell line development workflow and challenges'	27
2. THE AIMS OF THE THESIS	28
3. THE SUMMARY	29
REFERENCES	31
NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC	36

TERMS, ABBREVIATIONS AND NOTATIONS

C: Capsid protein

E1: Envelope protein 1

E2: Envelope protein 2

6K: Protein 6K

TF: Transframe protein

ORF: Open Reading Frame

nsP1: Non-structural protein 1

nsP2: Non-structural protein 2

nsP3: Non-structural protein 3

nsP4: Non-structural protein 4

UTR: Untranslated Region

sgRNA: Subgenomic RNA

Non-Structural Proteins:

P1234: Polyprotein precursor containing nsP1, nsP2, nsP3, and nsP4

P123: Polyprotein precursor containing nsP1, nsP2, and nsP3

nsP1: Non-structural protein 1

nsP2: Non-structural protein 2

nsP3: Non-structural protein 3

nsP4: Non-structural protein 4

AUD: Alphavirus Unique Domain

HVD: Hypervariable Domain

RdRp: RNA-dependent RNA polymerase

RepRNA - Replicon RNA

CSFV - Classical Swine Fever Virus

VRPs - Virus Replicon Particles

BSL2 - Biosafety Level 2

SARS-CoV-2 - Severe Acute Respiratory Syndrome Coronavirus 2

GFP - Green Fluorescent Protein

AUD - Alphavirus Unique Domain

CHIKV - Chikungunya Virus

Pac - Puromycin N-acetyl-transferase

Pac gene - Puromycin N-acetyl-transferase gene

HCV - Hepatitis C Virus

RdRp - RNA-dependent RNA Polymerase

HIV - Human Immunodeficiency Virus

HCV: Hepatitis C virus

C: Core structural protein

E1: Envelope glycoprotein 1

E2: Envelope glycoprotein 2

NS: Non-structural

NS2: Non-structural protein 2

NS3: Non-structural protein 3

NS5B: Non-structural protein 5B

SFV: Semliki Forest virus

SINV: Sindbis virus

EMCV: Encephalomyocarditis virus

BHK: Baby hamster kidney

neo: Neomycin phosphotransferase gene

IRES: Internal ribosome-entry site

CHIKV: Chikungunya virus

BSL-3: Biosafety level 3

EGFP: Enhanced green fluorescent protein

RT-PCR: Reverse transcription polymerase chain reaction

Rluc: Renilla luciferase

NTR: Non-translated region

BHK-CHIKV-NCT: BHK cell line containing CHIKV replicon

PCR: Polymerase chain reaction

RdRp: RNA-dependent RNA polymerase

NHC: β -D-N4-hydroxycytidine

Huh-7: Human hepatocellular carcinoma cell line

BFV: Barmah Forest virus

icDNA: Infectious clone DNA

ORF: Open reading frame

Pac: Puromycin acetyltransferase

INTRODUCTION

Viruses are submicroscopic infectious agents which consist of genetic material (DNA or RNA) enclosed in a protein or lipid envelope. Viruses have been described to infect all forms of life and their reproduction can take place by hijacking the machinery of their host cell and using it to make copies of themselves, meaning that they are obligatory endoparasites. This often leads to cellular pathologies, and ultimately to disease. Viruses thus significantly impact public health and for that reason, viruses are crucial research subjects for scientists. To better understand these infectious agents and contribute to developing more effective methods for preventing or treating viral infections, the research on the molecular biology of viruses is an important frontier in developing new and effective vaccines and antivirals.

Alphaviruses are positive-sense RNA genomic viruses from the *Togaviridae* family. The *Alphavirus* genus currently contains over 30 members spread worldwide in various hosts and vectors. Alphaviruses are arboviruses and are mostly transmitted by a broad spectrum of different species of mosquitoes to vertebrate hosts. Depending on the species, alphaviruses cause rash, arthritis, muscle pain, febrile illness, and encephalitis in humans and animals of veterinary importance, resulting in a high disease burden worldwide. Specific antiviral medications or vaccines are not currently available for medical treatment of these infections. One such alphavirus of clinical significance is Barmah Forest virus (BFV). It is primarily caused outbreaks in Australia and threatens public health in areas where suitable mosquito vectors are common. Although alphavirus research is vigorous, BFV remains a relatively modestly studied subject and a better understanding in its molecular biology will contribute to our knowledge on the genus overall and help develop antiviral prevention and treatment strategies.

Viral replicon cell lines have proven to be valuable tools in general molecular virology research and in the search for antiviral agents. In this work, a review on the successful generation of replicons and replicon cell lines for several alphaviruses and Hepatitis C virus is presented and as a separate aim, we describe the proposed workflow for constructing Barmah Forest virus replicons for the generation of replicon expressing stable cell lines in cultured mammalian cells based on these prior studies described here.

1. LITERATURE REVIEW

1.1. Overview of alphaviruses

Alphaviruses are a class of viral agents categorized within the *Togaviridae* family and the *Riboviria* realm. Characteristically, these viruses possess a genomic structure consisting of a single-stranded RNA molecule that is oriented in a positive-sense configuration (Chen et al., 2018). Depending on the species, alphaviruses cause illness characterized by rash, arthritis, polyarthritis, muscle pain, febrile illness, and/or encephalitis in humans. While most Alphavirus genus members are mosquito-borne, there are some notable exceptions. The vector of Southern elephant seal virus (SESV) is most likely a seal louse (La Linn et al., 2001); Salmonid virus (SAV, also known as Salmon pancreas disease virus) has no known vector and can be transmitted horizontally between fish, yet SAV can replicate in mosquito cells. Finally, Eilat virus (EILV) and Tai Forest virus (TAFV) are insect-specific viruses unable to replicate in vertebrate cells (Laureti, Paradkar, Fazakerley, & Rodriguez-Andres, 2020; Nowee et al., 2021).

The mosquito-borne alphaviruses have historically been divided into two groups based on the symptoms they cause: arthritogenic and encephalitic alphaviruses (Chen et al., 2018; Powers & Roehrig, 2011). However, the genus can also be categorized based on their geographical origin, namely Old World (OW) and New World (NW) alphaviruses in which OW alphaviruses tend to be associated with arthritic symptoms and NW alphaviruses with encephalitic symptoms and which may lead to lethal outcomes. The Old-World alphaviruses Sindbis (SINV), chikungunya (CHIKV), o'nyong-nyong (ONNV), and Ross River viruses (RRV) are found predominantly or originated in Africa, Europe, Australia and Asia. They are often characterized as causing febrile illnesses, with rash and debilitating arthralgia as common symptoms. Conversely, the New World alphaviruses (Venezuelan, Eastern and Western equine encephalitis (V/E/W/EEV) and related viruses, Una (UNAV) and Mayaro viruses (MAYV) are mainly found in the Americas. They are largely characterized as causing encephalitic disease. Interestingly, UNAV infections are associated with arthralgia as opposed to encephalitic disease, much like Old World alphaviruses (Chen et al., 2018; Nowee et al., 2021; Powers & Roehrig, 2011).

The latest report of the International Committee on the Taxonomy of Viruses (ICTV) recognizes Venezuelan equine encephalitis complex, Eastern equine encephalitis complex, Western equine encephalitis complex (including also SINV), Barmah Forest complex with Barmah Forest virus (BFV) as the sole member, Middelburg complex, Ndumu complex and Semliki Forest complex that includes, among others, SFV, CHIKV, o'nyong'nyong virus (ONNV), Mayaro virus (MAYV), and Ross River virus (RRV) (Chen et al., 2018). But as (alpha)viruses become more widely dispersed, it is no longer advisable to divide them according to their place of origin. A more comprehensive classification approach for alphaviruses is based on their serocomplex, which takes into account their genetic similarity and serological reactivity. This method offers a deeper understanding of the categorization of alphaviruses. The taxonomic structure of the alphavirus genus was originally organized according to antigenic relationships determined in serologic assays

(Chen et al., 2018). Several arthritogenic alphaviruses have (re-)emerged and posed a severe risk to the public health. There are no approved antiviral medications or vaccines to treat or prevent these viral infections. Alphaviruses are arthropod-borne viruses and, as it was mentioned before, are predominantly transmitted via mosquito vectors. When a mosquito carrying the virus bites a person, the virus can enter the bloodstream, resulting in viremia. Climate change may also contribute to the emergence and re-emergence of Alphaviruses. This is made possible by fluctuations in abiotic factors such as temperature, rainfall, and humidity, which may shift the vectors' dynamics (Chen et al., 2018; Lu et al., 2020). Chikungunya virus (CHIKV) is an emerging alphavirus that has caused recent outbreaks in Asia, the Indian Ocean, and Europe and can potentially spread to the Continental US and other parts of the Americas. Alphaviruses of greatest immediate importance in the US as naturally emerging pathogens and/or potential biological weapons include eastern (EEEV), western (WEEV), and Venezuelan equine encephalitis (VEEV) viruses (Greene et al., 2005; Powers & Roehrig, 2011; Tsetsarkin, Chen, Sherman, & Weaver, 2011).

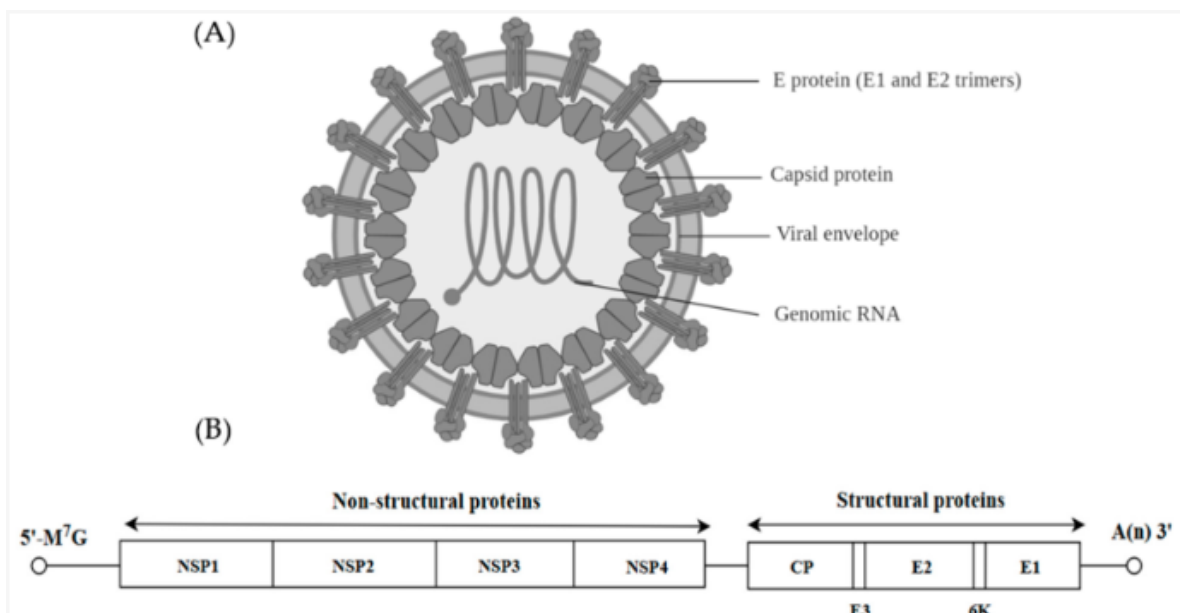


Figure 1 . A. The arrangement of the E protein (E1 and E2 trimers), capsid protein, and genomic RNA is visually illustrated in the alphavirus structure. B. The genetic structure of the alphavirus is displayed, indicating the 5' and 3' untranslated regions. Non-structural proteins (NSP1, NSP2, NSP3, and NSP4) and structural proteins (CP, E3, E2, 6K, and E1) are represented by open boxes (Mandary et al., 2019)

1.2. Alphavirus virion

The nucleocapsid, consisting of the viral genome and approximately 240 copies of the C protein, forms the core of the virion. Surrounding the nucleocapsid is a lipid bilayer derived from the host cell, which contains approximately 240 copies of glycoproteins E1/E2 heterodimers. The E2 glycoprotein interacts with the nucleocapsid C proteins in a 1:1 ratio. On the surface of the virion, the 240 heterodimers of E1/E2 glycoproteins form trimers, resulting in the formation of characteristic spike structures known as "80 spikes."

The outer domains of the E2 glycoprotein in these spikes serve as viral anti receptors, facilitating the binding of the virion to specific cellular receptors. In addition to the C protein and E1/E2 glycoproteins, other proteins such as 6K and TF have been observed to be packaged into the virions, although to a lesser extent compared to the C protein and glycoproteins. Overall, the composition of alphavirus virions consists of a nucleocapsid core comprising the viral genome and C proteins, surrounded by a lipid bilayer derived from the host cell. The lipid bilayer contains E1/E2 glycoproteins that form trimers on the virion surface, leading to the formation of spike structures Figure 1 (Mandary, Masomian, & Poh, 2019). These glycoproteins play a crucial role in viral binding to receptors. Additionally, other proteins, such as 6K and TF, are present in the virions but in smaller quantities.(Cheng et al., 1995; Kääriäinen & Söderlund, 1978; Mandary et al., 2019)

1.3. Alphavirus genome

The alphavirus genome is a positive-sense, RNA genome approximately 12 kb in length with a 5'cap and a 3'-PolyA tail, with the alphavirus genome mimicking that of a cellular mRNA. The genome contains two open reading frames (ORFs) where the first ORF encodes for the non-structural proteins (nsP1, nsP2, nsP3, and nsP4) that constitute the viral replicase. These proteins are initially synthesized as polyprotein precursors P1234 or P123 straight from the genomic RNA, where the type of precursor translated is regulated by the presence of an opal stop codon UGA at the nsP3/nsP4 junction. The second ORF encodes the structural proteins (capsid, E1, E2, and 6K/TF) responsible for the assembly of the viral particle. These proteins are translated from a shorter subgenomic RNA (sgRNA) that is produced during later infection from negative-strand RNA templates and the sgRNA promoter is inactive in the context of the viral genome. The alphavirus genome also contains conserved regions, such as the 5' untranslated region (UTR) and the 3' UTR, which play important roles in initiating viral genome replication, transcription, protein translation, and packaging(Kääriäinen & Söderlund, 1978; Mandary et al., 2019; Tsetsarkin et al., 2011)

1.4. Non-structural proteins

As mentioned above, the alphavirus non-structural proteins are translated as two precursor polyproteins: P1234 and/or P123 due to the UGA codon at the nsP3/nsP4 junction. The P1234 precursor is thus produced more seldom but readthrough of the stop codon leads to the production of P1234 and thus the essential component nsP4. All alphavirus nsP-s are essential to be included in a functional replicase and they all serve specific enzymatic properties needed for viral RNA propagation(Jose, Snyder, & Kuhn, 2009; E. G. Strauss, Levinson, Rice, Dalrymple, & Strauss, 1988; J. H. Strauss & Strauss, 1994).

nsP1 is involved in the capping of viral RNAs, facilitating translation and immune evasion. nsP1 serves as a viral methyltransferase and guanylyltransferase and these properties are essential in the synthesis of a cap structure at the 5' end of viral RNA, crucial for stabilizing the RNA and evading host immune responses (Zhang et al., 2021). A crucial characteristic is also its membrane-binding capability as alphavirus replication complexes are active and stable when membrane bound. And more recent studies have described the formation of a characteristic ring structure that is formed by 12 subunits of nsP1 that will sit at the neck area of modified membrane invaginations, or spherules, that are apparent upon alphavirus infection. Such structures recruit other nsP-s to the replication complex together with template RNAs (Jones et al., 2023; Zhang et al., 2021).

nsP2, like alphavirus nsPs overall, is multifunctional and plays essential roles in viral replication and pathogenesis. N-terminal regions of the protein contain RNA helicase, NTPase and RNA triphosphatase activities. As an RNA helicase, unwinding RNA structures to facilitate viral RNA replication and the NTPase activity is supposedly the energetic driver of the former activity. The RNA triphosphatase activity is important in preparing the substrates for nsP1 capping activity, thus nsP2 is also essential in providing the proper capping activity of nsP1. C-terminally, nsP2 possesses protease activity that plays a crucial role in regulating the RNA synthetic properties of the alphaviral replicase complex in a sequential manner. nsP2 cleaves the viral polyprotein into cleavage intermediates and individual nonstructural proteins (Hardy & Strauss, 1989). It also modulates cellular processes and suppresses host immune responses to create a favorable environment for viral replication. In the case of Old-World alphaviruses, nsP2 serves as a mediator for the shutdown of transcription and translation processes in host cells where an important role is in the nuclear localization of nsP2 for which a nuclear localization signal has been described (K. Tamm, A. Merits, & I. Sarand, 2008). It plays a crucial role as an antagonist to effectively suppress the host's antiviral responses, thereby facilitating viral replication and evasion of host immune defenses (Breakwell et al., 2007; Rikkinen, Peränen, & Kääriäinen, 1994).

Three domains of alphavirus nsP3 are recognized: the N-terminal macro domain, the intermediate alphavirus unique domain (AUD) and the C-terminal hypervariable domain (HVD) (Ahola, McInerney, & Merits, 2021). The macro domain includes the only enzymatic activity described for nsP3 as it is able to hydrolyse ADP-ribose from mono-ADP-ribosylated host proteins. CHIKV mutants defective in this enzymatic property have been described to be lethal for the virus, suggesting that this activity is indeed essential for alphavirus infection (Shin et al., 2012). AUD has also been called the zinc-binding-domain of nsP3 as a zinc ion is described bound to this domain within replicase complexes and this has been characterized as essential in SINV replication (Shin et al., 2012). HVD, as the name suggests, is most heterogenous in structure within the alphavirus genus and it can also significantly vary in length. While disordered in structure, it serves as an important hub for interactions with host factors. An example of such is the attribute of some alphavirus nsP3 binding the host cell factor G3BP, which in turn inhibits the formation of stress granules in response to the viral infection. Such structures down-regulate translation and it is in the virus' interest to keep this activity intact (Panas, Ahola, & McInerney, 2014).

nsP4 functions as the RNA-dependent RNA polymerase (RdRp) and provides the viral RNA synthetic capabilities in the host cell cytoplasm. This activity lies within the C-terminal domain of nsP4 and shows high similarity to other viral RdRp-s (Rubach et al., 2009). It has also been characterized as a terminal adenylyltransferase, adding a polyA tail to the 3' end of the viral mRNAs (Tomar, Hardy, Smith, & Kuhn, 2006). The N-terminal domain is described as alphavirus specific and is conserved within the genus. This domain contains residues essential for nsP4 to interact with the rest of the replicase complex. Furthermore, there are elements here, like the opal stop codon, and a destabilizing terminal tyrosine residue, that serve as regulators of nsP4 amounts upon infection (de Groot, Rümenapf, Kuhn, Strauss, & Strauss, 1991; Li & Rice, 1989).

1.5. Alphavirus infection cycle

For alphaviruses to initiate infection, they must attach to target cells and engage an entry receptor. Over the years, alphavirus interactions with several ubiquitously expressed cell surface molecules, termed attachment factors, have been described (Hoornweg, T.E., et al., 2015). Alphaviruses utilize several cell surface molecules, including HS (heparan sulfate), DC-SIGN/L-SIGN, and PS (phosphatidylserine receptors) receptors to carry out an initial attachment to target cells (Holmes et al., 2020). Matrix remodeling associated protein 8 or Mxra8 has been described as an important receptor mediating the attachment and internalization of several alphaviruses, including CHIKV, RRV and BFV. Yet, this is not a universally used strategy within the genus, as SFV was shown not to utilize the receptor (Zhang et al., 2018). After successfully binding to a specific receptor, the virus is internalized into the host cell via clathrin-mediated endocytosis. When internalization is completed, fusion between the virion envelope protein E1 and the endosomal membrane at the acidic pH, causes releasing of the nucleocapsid into the cytoplasm. (Leung et al., 2011)

Uncoating of nucleocapsid occurs immediately after their penetration into the cytoplasm and genome RNA is released (Singh et al., 1992). After uncoating, the 5' two-thirds of the genome RNA, or the first ORF, is translated into two RNA replicase polyproteins P123 and P1234, and the nature of the opal stop codon has been described above. A series of proteolytic events mediated by nsP2 then cleave the polyprotein in a specific sequence (Strauss J.H., E.G.Strauss, 1994). Firstly, P1234 is cleaved at the 3/4 site to yield replication complex P123+nsP4. This complex takes part in the synthesis of negative-strand RNAs and is termed the early alphavirus replicase (Merits et al., 2001). Next, cleavage of the P123 at the 1/2 site yields in the intermediate replicase complex nsP1+P23+nsP4 which is shown to capable of both negative- and positive-strand RNA synthesis. The final cleavage event at the 2/3 site gives rise to the nsP1+nsP2+nsP3+nsP4 containing complex or the late replication complex, capable of synthesizing positive strand genomic and sgRNA utilizing the previously produced negative-strand RNAs as templates (Strauss J.H., E.G.Strauss, 1994).

Concurrently, alphaviruses undertake active modifications of diverse cellular membranes to establish favorable platforms and a safeguarded milieu to facilitate efficient replication.

The replication of RNA occurs in close association with these cellular membranes, and alphaviruses generate diminutive invaginations termed spherules, typically with a diameter of approximately 60 nm. These spherules serve as specialized compartments where viral RNA replication occurs (Elmasri et al., 2021). A dodecameric ring of nsP1 subunits has been recently shown to position at the neck of the spherules and directs the localization of other replicase proteins and template RNAs to the site (Tan et al., 2022).

As mentioned above, the translation of the structural polyprotein is initiated from sgRNA, a product of the late replicase activity. The polyprotein includes the individual proteins in the following order: C-PE2(E3+E2)-6K/TF-E1 (Strauss and Strauss, 1994). C protein is cleaved autoproteolytically, and the rest of the polyprotein is transported to the endoplasmic reticulum (ER). In the endoplasmic reticulum (ER), the enzymatic activity of cellular proteases facilitates the generation of PE2 and 6K-E1 subunits. Subsequently, another cleavage event liberates the E1 subunit, enabling the formation of PE2-E1 heterodimers. These heterodimers are then transported via the Golgi complex, and the furin protease facilitates the cleavage of E3 from the dimers, resulting in the formation of E2-E1 heterodimers. These dimers are transported through the Golgi complex and furin cleaves E3 from the dimers leading to formation of E2-E1 heterodimers (Ivanova, L.; Schlesinger, M.J., 1993). As the C protein recognizes genomic RNA, it is then seen interacting with the E2 proteins in the heterodimers which have now been transported to the plasma membrane and the aforementioned interaction gives rise to the initiation of virion budding. Successful budding also results in the virions obtaining their characteristic bilipid membrane (Garoff et al., 2004)

1.6. Review of the Barmah Forest virus

Barmah Forest virus (BFV) is a member of the *Togaviridae* family, specifically an alphavirus, and holds substantial medical importance as a mosquito-borne virus within Australia. Infection with BFV leads to notable morbidity, commonly characterized by symptoms including maculopapular rash, fever, fatigue, and debilitating arthralgia or arthritis. Importantly, the clinical presentation of BFV disease closely resembles that of Ross River virus (RRV), another medically significant alphavirus prevalent in Australia (Cashman et al., 2008; Jacups et al., 2008). This study presents a comprehensive analysis of the evolutionary properties, relationships, and dynamics of Barmah Forest virus (BFV). By compiling a near-whole-genome dataset of 34 taxa, the researchers were able to characterize three distinct genotypes of BFV (G1-G3) circulating in both Australia and Papua New Guinea (PNG). The analysis reveals evidence of temporal and geographical co-circulation of the G2 and G3 genotypes in Australia, with G3 being the contemporary lineage. The study also compares the evolutionary dynamics of BFV with Ross River virus (RRV), highlighting some differences in their emergence patterns and genetic diversity (Michie et al., 2020).

1.7. Replicons of RNA viruses as tools for in laboratory studies and biotechnological industry: benefits and limitations

A viral replicon is a modified virus genome that is capable of autonomous replication within host cells. It typically consists of a segment of viral RNA or DNA that contains the necessary elements for replication, including the cis-sequence essential for replication and coding regions for virus-encoded replicase proteins (Démoulins et al., 2021; Hannemann, 2020). Viral replicons are engineered in the laboratory by removing certain non-essential viral genes (typically encoding for structural proteins) or components from the viral genome while retaining the elements required for replication. This modification allows researchers to use such genomes to study specific aspects of viral replication and gene expression in a controlled manner.

By introducing the viral replicon into host cells, researchers can observe and manipulate the replication and expression of the viral genetic material (Frolov et al., 1999; He et al., 2019). Replicons have many advantages over full-length virus genomes that allow to express various proteins of interest in controlled and enhanced way. These advantages include reduced complexity of the viral genome which makes it a convenient and easy way to be used as an expression vector. In addition to that, replicon can be engineered to prioritize protein expression which in turn produces proteins in high yield and efficiency. Moreover, controlled expression and simple handling made these replicons an excellent way as expression vectors for various studies. In one study, researchers developed self-amplifying replicon RNA (RepRNA) derived from classical swine fever virus (CSFV, genus Pestivirus, family Flaviviridae) for delivering influenza virus antigens. The researchers engineered RepRNA constructs and compared different delivery methods in pigs. They found that RepRNA delivered using virus replicon particles (VRPs) and a specific synthetic formulation showed the most effective induction of immune defenses. In contrast, “naked” RepRNA combined with an adjuvant had limited effectiveness. This study highlighted the potential of RepRNA-based vaccines as viral replicon controlled approach for influenza virus antigens expression (Démoulins et al., 2021). It should be mentioned, however, that used of advanced RNA delivery systems currently used in mRNA vaccines can have potential to change the field – delivery of replicon RNA in form of liposomes (or other type of lipid containing particle) is hugely more efficient than delivery of naked RNA and may even exceed the efficiency of delivery in form of VRPs. Even if the latter is not the case, the production of lipid-RNA particles is a well-established technology which makes it less expensive, safer and better controlled than packaging of replicons into VRPs. For safety considerations, replicons can be engineered so that risky viral infectious particles won't be formed and ability to produce diseases is minimized. This will lower the risk of infecting laboratory workers. In a particular investigation, researchers introduced a novel system based on single-cycle infectious SARS-CoV-2 VRPs. This system was designed to include a dual reporter system comprising luciferase and green fluorescent protein (GFP). It allowed for the safe handling of SARS-CoV-2 in biosafety level 2 (BSL2) laboratories (Malicoat et al., 2022).

The use of viral replicons in scientific research and, especially in industrial applications has, however, certain limitations. First, it can be costly due to various factors. These include the initial development and engineering of replicon systems, the need for specific cell cultures and their maintenance, quality control, and characterization processes, licensing and intellectual property considerations, as well as the scale-up and manufacturing for larger-scale applications (Chang, Sosnovtsev, Belliot, King, & Green, 2006; Fernandes et al., 2020; Yang, Zhao, Xun, Ma, & Wang, 2021). The costs associated with viral replicons can vary depending on research requirements and intended applications. In general, it is less problematic for basic and applied studies (compared to industrial use), in part because most of replicons were originally developed for these purposes. Furthermore, the expenses involved, viral replicons offer unique advantages and capabilities in research and biotechnology fields (Ehteshami et al., 2017; Tamm, Merits, & Sarand, 2008; Utt et al., 2015). In industrial use, the problems are more evident as replicon studies are often conducted in cell culture systems, which may not be suitable for scale up required, for example, for large-scale protein production. As a result, culturing and maintaining cells in sufficient quantities for protein expression experiments can be time-consuming and costly (Frolov et al., 1999). Additionally, replicons may not be able to produce high yields of protein compared to other expression systems like recombinant protein expression in bacteria or yeast (Baghban et al., 2019; Weber et al., 2020) or mammalian cell lines developed for production of recombinant proteins. There is also an issue with possible contamination of product by replicon or replicon encoded replicase proteins, hence more sophisticated (and more costly) purification procedures may be required.

However, viral replicon systems provide several advantages over other expression systems in molecular biology and genetic engineering. They efficiently infect different types of mammalian cells, allowing for studies in a biologically relevant context or, in industry, expression of protein of interest in the right type of cells. Moreover, viral replicons exhibit high-level gene expression in mammalian cells, enabling robust analysis of proteins and pathways and in industry this translates into lower costs of the product. In addition to that, the replicons of viruses with RNA genomes can be delivered as naked RNA, liposomes or VRPs i.e. in forms where no DNA is present. Obviously, this eliminates risks associated with genomic integration of genetic material or replicon. The downside of use of RNA replicons is also obvious – it is very hard to achieve precise control of gene expression as vast majority of available regulatory elements (promoters, enhancers etc) do not work in context of RNA genome.

Replicon systems are adaptable to different cell types, providing flexibility in target selection. Although it may be costlier, such a replicon system provides valuable benefits in research and biotechnology applications (Casales et al., 2008). The utilization of replicon systems offers a significant advantage in terms of genetic stability. The stability of replicons is particularly noteworthy when compared to the use of mutant viruses in certain studies. Mutant viruses, due to their RNA genomic nature, are prone to rapid reversions and acquisition of compensatory mutations during the process of virus passaging cycles. These cycles, characterized by the replication and propagation of viral genomes, create opportunities for the emergence of secondary mutations within the viral genome or the

reversion of low cytotoxicity or non-cytotoxic viruses (Hannemann, 2020). The sheer abundance of viral genomes produced during these passaging cycles increases the likelihood of a more "successful" virus genome emerging, either through reversion or the acquisition of secondary mutations elsewhere in the genome. This phenomenon poses challenges for studies that rely on mutant viruses as their genetic instability can complicate the interpretation of experimental results. The main problem is not, however, appearance of mutations as such. In propagating viruses (i.e. in conditions where infection spreads from one cell to another) there is always competition of different viruses both in double infected cells and for the speed of how quickly these viruses can deliver their genetic material to other cells. Competition leads to selection and, as consequence, as single mutant genome with significant growth advantage may overgrow the original virus in a few passages. Replicons also use RNA replication machinery that makes mutations. However, there is no cell to cell spread and therefore competition (and selection pressure) is much lower. Still, it represents a problem. To overcome this problem researchers have designed an approach which uses selection towards genetic stability. In most cases, this is achieved by the use of replicons in combination with antibiotic selection which provides a means to maintain the desired replicon construct. The inclusion of an antibiotic resistance gene allows for the application of selective pressure, ensuring the preservation of the replicon of interest. The high selective pressure exerted by the antibiotic enables the replicon to persist and propagate without the risk of genetic instability (Fernandes et al., 2020; Yang et al., 2021). The most important benefit for this approach is, however, the possibility to generate cell lines permanently harbouring replicating RNA that constantly expresses protein of interest.

The use of replicons is not limited to the production of recombinant proteins. Replicons are also useful for studies of molecular biology of viruses. One good example is a study performed using human hepatocarcinomatous cell line Huh-7 and CHIKV replicons. In this study, AUD of nsP3 of CHIKV was subjected to mutagenic analysis. Based on the Sindbis virus AUD structure and sequence alignments of various alphaviruses, mutations were introduced into a CHIKV sub-genomic replicon. The study demonstrated that the AUD in CHIKV plays a critical role in RNA replication, showing cell-type and species-specific phenotypes. The mutants were further examined in the context of infectious CHIKV to explore the involvement of the AUD in other stages of the virus lifecycle. It was found that the AUD was necessary for virus assembly, with one mutant (P247A/V248A) showing a significant decrease in infectious virus production. This phenotype was attributed to a block in subgenomic RNA transcription, resulting in reduced synthesis of structural proteins and subsequent decrease in virus production. The study also demonstrated that the P247A/V248A mutant nsP3 exhibited reduced binding to viral genomic RNA and decreased affinity of the mutant AUD for subgenomic promoter RNA. These findings suggest that the AUD serves as a pleiotropic protein domain with multiple functions in CHIKV RNA synthesis (Gao, Goonawardane, Ward, Tuplin, & Harris, 2019).

1.8. Stable cell lines for replicons

Stable replicon cell lines refer to cultured cells that have been genetically modified to contain a replicon that can autonomously replicate within the host cell. These cell lines serve as powerful tools in virology research for studying viral replication processes and have the potential use for developing antiviral strategies. The viral replicons used in stable cell lines are designed the same way as typical RNA replicons described in previous chapter i.e. it consists of the essential elements required for RNA replication, including the viral RNA sequence encoding for replicase proteins and the necessary cis-acting elements, such as the RNA-dependent RNA polymerase binding sites. The difference from standard replicons is the presence of selection element – typically gene encoding the resistance to antibiotic used for the selection – and, in many cases, also one or several mutations that change properties of the RNA (and/or proteins encoded by this RNA) the way that allows long time coexistence with the host cell. The latter is often the main challenge as many viruses – alphaviruses included – are cytolytic i.e. their replication tends to kill the host cell. Turning off this function does, however, aside of practical value also provide important know-how about virus-host cell interactions. By this reason construction and analysis of stable replicon cell lines offer several advantages in virology research (Casales et al., 2008; Frolov et al., 1999). Specifically, they provide a controlled and stable system to study viral replication, allowing researchers to investigate various aspects of the replication process, such as viral protein expression, RNA synthesis, and host-virus interactions. Furthermore, these cell lines enable the evaluation of antiviral compounds and the development of therapeutic strategies, as they mimic certain aspects of viral infection while eliminating the risk associated with working with live, replication-competent viruses (Chang et al., 2006; Kishine et al., 2002).

To generate stable replicon cell lines, a replicon construct is created by cloning an element allowing for selection into a viral RNA replicon vector. This selection element typically consists of puromycin N-acetyl-transferase (Pac) gene that encoded for a protein that exhibits resistance against puromycin. Puromycin is favoured because of rapid killing of non-transfected cells (few days, in comparison, for many antibiotics it takes weeks) as well as small size of Pac gene. The replicon construct is introduced into host cells through transfection or viral transduction. Following transfection, the cells are exposed to puromycin, a potent antibiotic that inhibits protein synthesis by targeting the ribosomes. The replicon construct will allow cells that have successfully established the replicon to survive in the presence of puromycin. These resistant cells are isolated, either by selecting individual colonies or using sorting techniques and expanded to establish stable replicon cell lines. Verification and characterization techniques by confirming expression level of marker proteins or by PCR and sequencing, are then employed to confirm the presence and stability of the replicon in the established cell lines (Agapov et al., 1998; Casales et al., 2008; Frolov et al., 1999).

Protecting cell against death caused by antibiotic is straight forwarded – all it takes is expression of corresponding resistance gene at sufficient level. The other side of the coin is to protect cell against damage caused by the replicon. In some cases, it is easy enough –

some viruses are non-cytolytic and their replication does not damage cells. The most famous example of virus of this kind is hepatitis C virus (HCV) replicons. HCV is classified within the Flaviviridae family and is a single-stranded positive-sense RNA virus. Its genome spans 9.6 kb and encodes 10 proteins, comprising three structural proteins and seven non-structural proteins. Among these non-structural proteins, the NS5B RdRp is particularly appealing as a target for the development of antiviral therapies against HCV. HCV is often called as “silent killer” as symptoms of this virus develop very slowly taking many years or several decades. It is partly because of inefficient virus replication – it is intrinsically slow and has many requirements to occur. Rather frustratingly, HCV does not replicate in small animals (hence no in vivo model is available) and with rare exceptions (not known till early 2000-ds) fails to replicate in cell culture. These factors hugely complicated the studies of the virus and, from more practical point, effectively eliminated possibility to develop drugs against this virus. Indeed, compared to human immunodeficiency virus (HIV) where it took 4 years from discovery to develop first antiviral drug it took more than two decades to do the same for HCV. The breakthrough in anti-HCV drug development was in large extent possible thanks to replicon cell lines. To do this, researchers established cell culture systems for the replication and maintenance of a subgenomic replicon of the HCV over an extended period. HCV replicons are designed based on selection of mutations in non-structural proteins allowing more efficient propagation of these minigenome in specific cell line and expression of antibiotic resistance gene and specific markers for detection of replicon replication efficiency (Figure 2) (Bartenschlager, 2002). RNA replicons of HCV containing a neomycin-resistant gene were constructed using HCV genome RNA. Following transfection, Huh-7 cells resistant to the antibiotic, were selected and subsequently subcloned. The chosen cellular clone exhibited the successful synthesis of HCV proteins and the generation of replicon RNA, demonstrating the effective replication of the constructed viral RNA within the cultured cells, thus confirming its functionality as a stable replicon. (Chen et al., 2021; Kishine et al., 2002).

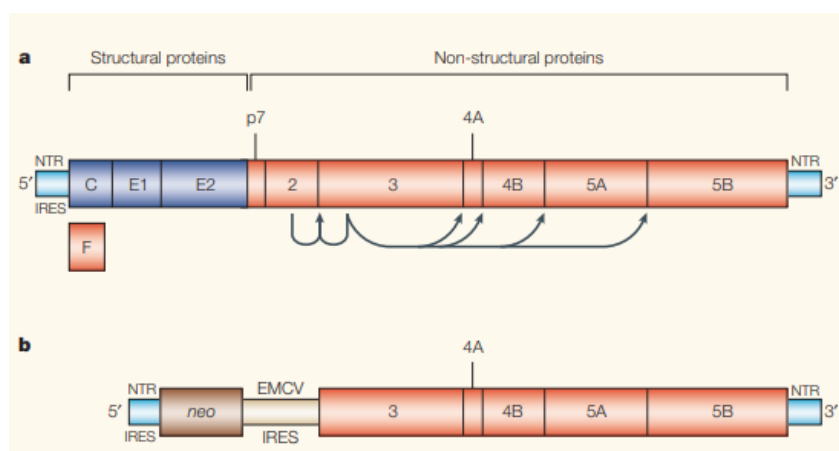


Figure 2. a. The sections of the hepatitis C virus (HCV) genome responsible for encoding the core structural protein (C) and the envelope glycoproteins (E1 and E2) are highlighted in blue, while the non-structural (NS) proteins NS2 to NS5B are depicted in red. Cleavage events facilitated by the NS2-NS3 and NS3 proteases are indicated by arrows. **b.** illustrates

the generation of a subgenomic replicon for the hepatitis C virus (HCV). The replicon was created by replacing the region responsible for encoding the structural proteins up to the NS2-encoding region with the neomycin phosphotransferase gene (neo) and the internal ribosome-entry site (IRES) derived from encephalomyocarditis virus (EMCV). The IRES from EMCV enables the expression of HCV replication proteins, while the selectable marker neo is translated under the control of the HCV IRES located in the 5' non-translated region (NTR)(Bartenschlager, 2002)

Compared to HCV, alphaviruses are, however, mostly cytotoxic, and often this effect is caused by replicase proteins of the virus i.e., cytotoxic properties of the virus are reflected in properties of replicon construct. Usually, the ways how to eliminate cytotoxicity are not obvious – profound knowledge of functions of virus proteins is required. If this knowledge does not exist (which is the most common situation) advantage is taken from error-prone replication of virus replication and selection process. In a study focused on the construction of a viral replicon system for expression in mammalian cells, researchers aimed to develop a noncytopathic SFV vector that has a poor ability to synthesize nsP2 which was suspected (and later clearly confirmed) to mediate shut-down of transcription and induction of apoptosis in infected cells. To achieve this, they introduced mutations previously reported to convert a cytopathic SINV into a noncytopathic form into an SFV vector expressing LacZ as a reporter gene. Among the mutants generated, a specific mutation (P718T) in the nsP2 subunit of SFV replicase showed replication capability only in a small fraction of BHK cells, resulting in the formation of beta-galactosidase-expressing colonies without any selection pressure. Subsequently, the LacZ gene was replaced with the Pac gene, enabling selection of a noncytopathic SFV replicon that turned out to harbour an additional mutation (R649H) in the nsP2 affecting putative nuclear localization signal of the protein. This study demonstrates the successful development of a noncytopathic SFV vector by introducing specific mutations in the nsP2 subunit of the replicase, leading to efficient expression of recombinant proteins in mammalian cells. The generated stable cell lines represent valuable tools for sustained high-level protein production, highlighting the potential utility of this vector system in various applications requiring robust protein expression (Chen et al., 2021). However, the double mutant SFV showed reduced replicase protein synthesis, leading to significantly lower levels of genomic and subgenomic RNA compared to the wild-type vector (125-fold and 66-fold lower, respectively). Despite this decrease in RNA levels, the mutant SFV exhibited similar levels of beta-galactosidase expression as the parental vector (Casales et al., 2008), most likely because of lack of replicon-induced shutdown of cellular translation. SFV is a virus with minor, or possibly no, medical importance. In contrast, CHIKV causes severe disease and infects millions of people. Therefore, studies aiming to develop stable replicon cell lines were performed.

In one study that outlines the advancement of an innovative methodology designed for screening bioactivity and conducting molecular inquiries related to CHIKV. Specifically, this involves the establishment of a durable BHK cell line that accommodates the CHIKV replicon, denoted as BHK-CHIKV-NCT. The containment of CHIKV, an infectious pathogen, is typically required in BSL-3 facilities in most countries. The primary goal of this study was to create a more user-friendly assay system for screening inhibitors of

CHIKV replication. This involved the development of a BHK cell line that could sustainably replicate CHIKV replicon RNA. To achieve this, the CHIKV-LR replicon sequence obtained from a La Réunion isolate was modified by inserting the Pac selection marker and two reporter genes (Renilla luciferase, Rluc, and EGFP) (Fig. 3A). To mitigate the cytotoxic effects associated with the wild-type CHIKV-LR replicon, a Pro718 to Gly substitution was incorporated within the protease-encoding region, as this modification had previously been demonstrated to reduce cytotoxicity in SFV and SINV vectors. It is noteworthy that the absence of this mutation invariably resulted in cell death following transfection with transcripts from such vectors. Figure 3B illustrates the selection process employed in this study. BHK cells were transfected with in vitro transcripts derived from the CHIKV-PG vector. Following transfection, cells were plated and subjected to puromycin selection at a concentration of 5 µg/ml, starting 16 hours post-transfection. The majority of the cells experienced cell death within four days, while a small fraction of cells (estimated to be approximately one in every 1×10^5 transfected cells) expanded into cell clones. These cell clones were transferred to separate plates and further expanded as cell lines, while being continuously exposed to puromycin selection.

RNA was isolated from 12 distinct cell lines, and reverse transcription polymerase chain reaction (RT-PCR) was conducted to amplify the regions corresponding to CHIKV nsP2. The resulting PCR products were then subjected to sequencing to identify the specific mutations associated with the non-cytotoxic properties of the replicon. These identified mutations were individually incorporated into the CHIKV-PG vector, and BHK-21 cells were transfected with the mutant replicons for further analysis. Cell viability assays were conducted, although the data obtained from these assays are not shown in this study. After careful examination of the experimental findings, a specific mutation was chosen, which entailed the insertion of five amino acid residues (GEEGS) between residues 647 and 648 of CHIKV nsP2. The corresponding insert sequence in the replicon RNA was identified as GGG GAG GAA GGG AGU. This insertion site coincided with the location of a nuclear localization signal previously identified in SFV nsP2. The selected mutation, along with the fusion of an Rluc marker to nsP3, was integrated into the CHIKV-PG vector, resulting in the generation of the CHIKV-NCT replicon vector (Fig. 2A). Transfection of BHK cells with this replicon vector demonstrated cell viability when subjected to continuous puromycin selection, and these cells were designated as BHK-CHIKV-NCT cells. (Pohjala et al., 2011)

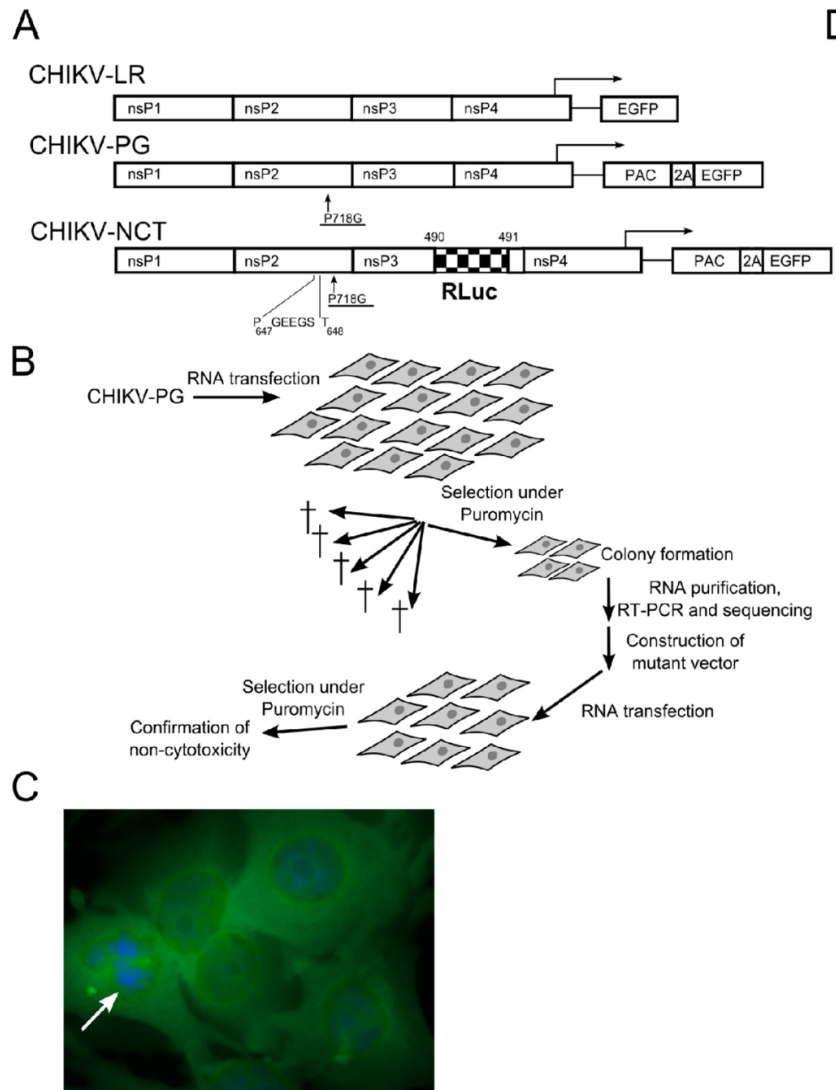


Figure3 A. A schematic diagram depicting the CHIKV replicons utilized in this study is presented. The procedure for selecting non-cytotoxic (NCT) CHIKV replicons, identifying mutations linked to the NCT phenotype, and validating their phenotypic effects was carried out in the study. C. The phenotype of BHK-CHIKV-NCT cells was examined, and the presence of green fluorescence resulting from EGFP expression was observed. An arrow is used to indicate a dividing cell in the image. (Pohjala et al., 2011)

Originally, the molecular bases of CHIKV attenuation were not well understood and in addition, the replicons were working only in BHK-21 cell but not in human cells. These issues were addressed in an additional study aimed to understand the molecular bases responsible for weakening CHIKV and to explore the potential of CHIKV replicon cell lines based on human cells for identifying effective drug candidates against this viral infection. The impact of introducing a P718G (PG) mutation and additional mutations in

CHIKV nsP2 on the replication and cytotoxicity of CHIKV replicons in BHK-21 cells was investigated by the researchers. The combination of the P718G (PG) mutation with either an E116K (EK) substitution or the insertion of a GEEGS sequence after residue T648 (5A) resulted in a notable hindrance to RNA synthesis, without affecting the nuclear translocation of nsP2. The introduction of the PG mutation into recombinant nsP2 resulted in the inhibition of proteolytic cleavage, a reduction in GTPase and RNA helicase activities, and the loss of RNA stimulation for GTPase activity. The presence of 5A and EK influenced the effects of PG, with both mutations specifically decreasing the RNA helicase activity of nsP2. These findings provided insights into the molecular mechanisms contributing to the observed cytotoxicity during the replication of alphaviruses. Furthermore, through the adaptation of the CHIKV replicon that incorporated both the 5A insertion and the PG mutation, scientists successfully identified a CHIKV replicon variant with adaptive mutations in nsP1 and nsP3. This adaptation allowed the replicon to persist in human cell lines. These CHIKV replicon cell lines serve as valuable experimental platforms for studying host factors and screening inhibitors of CHIKV replication at lower biosafety levels. This study by Utt et al. (2015) thus contributes to our understanding of the molecular basis of CHIKV attenuation and highlights the utility of CHIKV replicon cell lines for investigating host factors and identifying potential inhibitors against CHIKV replication (Utt et al., 2015).

1.9. Examples for uses of stable replicon cell lines

The global health challenge posed by HCV has stimulated the necessity for innovative approaches in treating the disease. Using HCV replicon systems, researchers can assess the activity of drug candidates against the viral replication machinery, particularly the NS5B RdRp. This enzyme is essential for HCV RNA replication and serves as a prime target for developing antiviral therapies. Scientists have recognized the NS5B enzyme of HCV as a potential target for therapeutic development. In one study, the researchers employed HCV replicon models containing NS5B to evaluate the efficacy of drugs designed to inhibit NS5B. The researcher utilized HCV replicon to evaluate the effectiveness of phosphoramidate prodrugs. The process began with the investigation of the structure-activity relationship around the phosphoramidate portion of the molecule. The synthesis of the uridine nucleoside 5, which served as the starting material, involved a two-step process. First, the benzoyl protected 20-deoxy-20-R-F- β -C-methylcytidine (7) was converted into the protected uridine 8 by heating with 80% acetic acid and then treating with methanolic ammonia. The uridine nucleoside 5 was obtained in a 78% yield. Subsequent investigations of compound that is now known as sofosbuvir, a specific blockbuster drug targeting NS5B, have exhibited promising results and later radically changed the treatment of chronic HCV infections (Sofia et al., 2010; Bhatia, Singh, Grewal, Natt, & pharmacotherapeutics, 2014; Li, Chen, Fang, Li, & Lin, 2019).

In a study focused on the development of inhibitor drugs for CHIKV utilizing a stable cell line, it was noted that there is a current absence of CHIKV vaccines and treatment options primarily rely on non-specific antiviral agents, which yield sub-optimal results. Within this context, β -D-N4-hydroxycytidine (NHC) has been identified as a novel CHIKV inhibitor. NHC exhibits properties akin to a pyrimidine ribonucleoside and demonstrates selective inhibition of CHIKV replication specifically within cell culture. The anti-CHIKV activity of small molecules was assessed through their evaluation in Huh-7-CHIKV replicon cell culture. In this experiment, 10 μ M of each compound was incubated with 5,000 cells per well in triplicate within 96-well plates for 48 hours at 37°C. The percentage inhibition of the Renilla luciferase signal was calculated by normalizing it to the signal observed in control lanes with no compound. A scatter plot was generated to illustrate the percent inhibition data for 631 compounds (Figure 4). Remarkably, only one compound, specifically NHC, demonstrated a significant inhibitory effect on the replicon, with an inhibition rate exceeding 80%. The chemical structure of the NHC compound is depicted in Figure 4B. (Ehteshami et al., 2017)

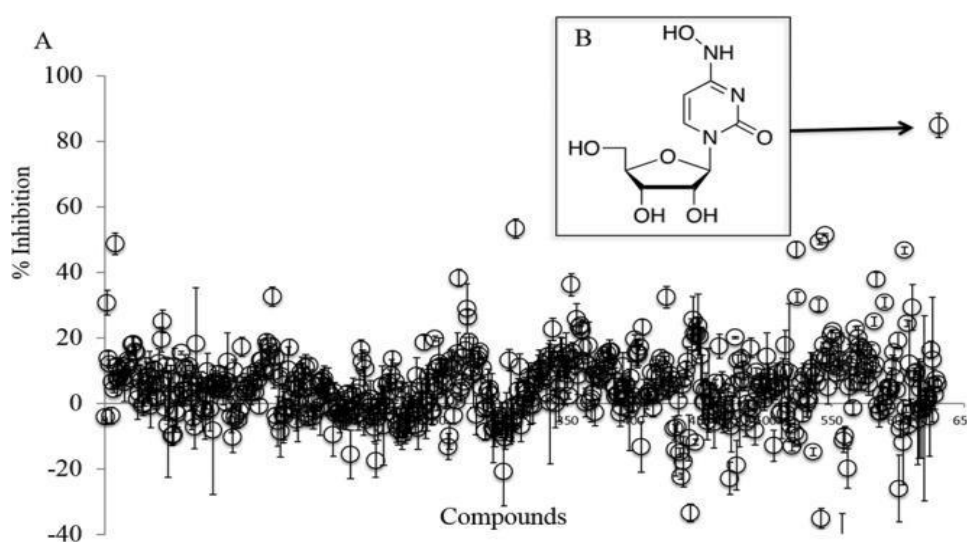


Figure 4 .A.The evaluation of 631 small molecules for their potential to hinder CHIKV replication in Huh-7-CHIKV replicon cell culture. Among the tested compounds, β -D-N4-hydroxycytidine (NHC) exhibited robust inhibitory activity, surpassing an inhibition rate of 80%. This noteworthy discovery emphasizes the promising nature of NHC as a prospective candidate for future development as an anti-CHIKV agent. B. The chemical structure of the NHC compound is illustrated in the figure..(Ehteshami et al., 2017)

1.10. Barmah Forest virus (BFV) cell line development workflow and challenges

One of the aims of the thesis at hand was to develop the design, workflow, and potential application of BFV replicon constructs and BFV replicon stable cell lines. The molecular biology of BFV remains insufficiently understood to immediately target mutations that decrease the virus's cytotoxicity, as none have been identified. Additionally, the specific protein responsible for the cytotoxic effects is not well characterized. As a result, it is uncertain whether the selection approach used for the SINV replicon, which involves starting a selection using the wild-type replicon, is applicable to BFV (Frolov et al., 1999). Alternatively, it is also unclear if the method employed for the SFV or CHIKV replicons, where a previously characterized cytotoxicity-reducing mutation in nsP2, would be more appropriate (K. Tamm et al., 2008). If the latter is the case we have additional challenge – no strong cytotoxicity reducing mutations have been described for BFV, so the start material must use what is available or, as an alternative, try to model structure of BFV replicase proteins, find similarities with replicase proteins of more intensively studied alphaviruses and attempt to make educated guesses what substitutions are most likely going to result in a desired outcome (reduction of cytotoxicity).

The workflow for developing BFV replicon cell lines will encompass several key steps. As a basis, we will use our BFV infectious clone plasmid construct (icDNA) that has been previously developed in our research group. The initial cloning step would be to swap out the structural ORF region in the icDNA with that of a marker cassette used in similar studies before: a Pac selection marker, 2A autoprotease signal sequence and a fluorescent protein marker sequence (Pohjala et al., 2011; K. Tamm, A. Merits, & I. J. T. J. o. g. v. Sarand, 2008; Utt et al., 2015). Subsequently, based upon the resulting WT BFV replicon, a mutated nsP2 containing replicon can then be generated by site-directed mutagenesis to include a similar mutation that has been previously introduced into SFV and CHIKV replicons to generate stable cell lines (Frolov et al., 1999; Pohjala et al., 2011; Kristi Tamm et al., 2008). However, the problem is that BFV is only distantly related to these well studied viruses and chances that selected mutation have predicted impact on the replicon of this virus are not high. These replicon constructs will then be transfected into BHK21 cells, a widely employed cell line in alphavirus research. Selective pressure will be applied by introducing an antibiotic, enabling only those cells with successful replicon uptake and expression to survive and propagate. Following the selection process, individual antibiotic-resistant cell lines exhibiting sustained replication of the modified replicons will be identified and isolated. RNA will be extracted from these selected cell lines, and the replicon RNA will be sequenced to determine the presence and characteristics of the introduced mutations or possible adaptations to selective pressure in the case of WT BFV replicase transfection. If apparent, mutations within the replicon will be described, including the precise nucleotide or amino acid changes, their locations within the replicon genome, and potential implications for viral replication or protein function. (Agapov et al., 1998; Chang et al., 2006; Gao et al., 2019)

As stated above no mutation with strong attenuating effect is known for BFV. There is also reasonable chance that mechanisms of cytotoxicity of BFV virus may be substantially different from these of other alphaviruses and, if this is the case, the approach described above (rational selection of attenuating mutations) has high chance to fail. Cytotoxic functions of nsP2 of SINV, SFV and CHIKV are due the ability of this protein to enter the nucleus. It is not clear if the nsP2 of BFV has similar subcellular localization patterns and a similar biological effect. Furthermore, a specific characteristic of BFV infection has recently been shown that the nsP3 of BFV that also localizes in the vertebrate host cell nucleus; the significance of such a property is yet to be described (Jacups et al., 2008). Previous work in our lab has shown that disrupting the NLS in the BFV nsP3 AUD resulted in viable virus yet it was attenuated in type I IFN competent murine embryonic fibroblasts. Therefore, it is reasonable to include such a mutation in a BFV replicon that will be used for selection of stable cell lines. Theoretically it is possible that this mutation would allow for stable expression of BFV replicons with defective localization of nsP3. Of course, as this approach has never been used for attenuation of alphavirus replicons there is a high chance that the result will not be what we expect. However, on the other hand, if obtaining stable replicon cell lines is not facilitated, we will acquire more knowledge about biological significance of nuclear localization of nsP3 of BFV. For example, it would be interesting to observe whether the nuclear localization of nsP3 influences the cell cycle progression of the vertebrate host cell. Therefore, such replicons will serve as a valuable tool in future studies.

To validate the impact of the identified mutations, reverse genetics experiments will be conducted. This will involve introducing the identified mutations into a wild-type BFV genome and evaluating their effects on replication, protein expression, or other pertinent phenotypic attributes (He et al., 2019). By executing this comprehensive workflow, researchers will be able to gain valuable insights into the behaviour and functionality of BFV replicons, contributing to a better understanding of the virus and its associated molecular biology. BFV replicon cell lines hold potential for various applications in scientific research. One prospective application is to pursue purely scientific goals by elucidating mutations that result in the attenuation of BFV. These identified mutations can subsequently be employed in the design of attenuated viruses, serving as promising candidates for vaccine development. Another envisioned utilization of BFV replicon cell lines is in the screening of antiviral compound as these cell lines can be instrumental in evaluating the effectiveness of various compounds as potential inhibitors of BFV replication, contributing to the development of novel antiviral therapies.

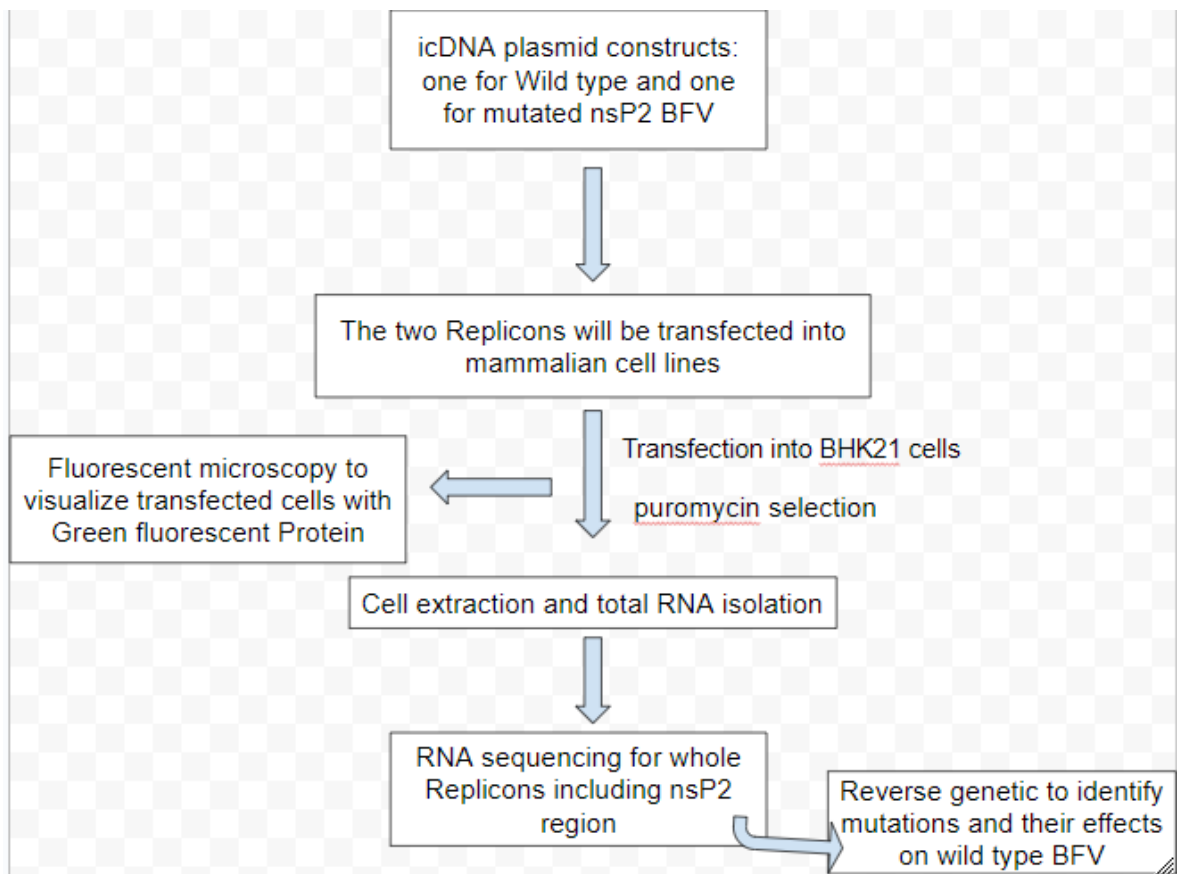


Figure 5. The figure represents key aspects of BFV replicon design for wt and mutated nsP2 BFV genomes, including the introduction and validation of nsP2 mutations and the use of reverse genetics to study their effects in wild-type BFV.

Figure 5 represents key aspects of BFV replicon design, including the introduction and validation of nsP2 mutations and the use of reverse genetics to study their effects in wild-type BFV. This comprehensive workflow enables researchers to gain insights into the behavior and functionality of BFV replicons, contributing to a better understanding of the virus and its molecular biology. BFV replicon cell lines have potential applications in scientific research, such as identifying attenuating mutations for vaccine development and screening antiviral compounds as potential therapies (Figure 5)

2. THE AIMS OF THE THESIS

As the title suggests, it is about construction of BFV replicons and developing mammalian cell line that allows for stable expression of BFV replicon.

3. The Summary

Alphaviruses are a collection of viruses transmitted by mosquitoes that induce various illnesses like skin rash, joint inflammation, and brain inflammation. Among these alphaviruses, BFV is a significant public health concern in Australia. However, the research conducted on BFV is comparatively limited in comparison to other alphaviruses. The main objective of this study was to propose a workflow for the construction of BFV replicons and for their stable expression in cultivated mammalian cells as viral replicons serve as valuable tools for examining viral replication mechanisms and developing antiviral treatments. This work provides a review of the known successful generation of replicons and such stable expression cell lines for both alphaviruses and another positive-sense RNA genomic virus, HCV. The review brings out the many difficulties that have needed to be overcome and strategies that have resulted in successful generation of CHIKV, SFV, SINV and HCV replicons and their stable expression in mammalian cell cultures. On the basis of these prior (successful) attempts, we propose here a systematic procedure for constructing BFV replicons and a mammalian cell line allowing for its stable expression. Our approach involves the design of a replicon specific to BFV harbouring a Pac gene for antibiotic selection, which will then be introduced into a mammalian cell line to identify and analyse the impact of critical mutations on the virus's life cycle. As BFV is relatively distinct from CHIKV, SFV and SINV, then there is reason to doubt the exact same mutations will result in a BFV replicon stable cell line and thus, when the selection process yields in replicon positive colonies, it will be crucial to sequence the replicon genome. Such approaches have previously resulted in describing attenuating mutations in alphavirus studies and provide valuable information on the molecular virology of these viruses and producing vaccine candidates. This research is essential for the development of antiviral therapies targeting BFV and have potential to reveal antiviral strategies within the genus.

REFERENCES

- Ahola, T., McInerney, G., & Merits, A. (2021). Alphavirus RNA replication in vertebrate cells. *Adv Virus Res*, 111, 111-156. doi:10.1016/bs.aivir.2021.07.003
- Bartenschlager, R. J. N. R. D. D. (2002). Hepatitis C virus replicons: potential role for drug development. 1(11), 911-916.
- Breakwell, L., Dosenovic, P., Karlsson Hedestam, G. B., D'Amato, M., Liljeström, P., Fazakerley, J., & McInerney, G. M. (2007). Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J Virol*, 81(16), 8677-8684. doi:10.1128/jvi.02411-06
- Cashman, P., Hueston, L., Durrheim, D., Massey, P., Doggett, S., & Russell, R. C. J. C. D. I. (2008). Barmah Forest virus serology; implications for diagnosis and public health action. 32(2), 263-266.
- Chen, R., Mukhopadhyay, S., Merits, A., Bolling, B., Nasar, F., Coffey, L. L., . . . ICTV Report Consortium, n. J. J. o. G. V. (2018). ICTV virus taxonomy profile: Togaviridae. 99(6), 761-762.
- Cheng, R. H., Kuhn, R. J., Olson, N. H., Rossmann, M. G., Choi, H. K., Smith, T. J., & Baker, T. S. (1995). Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell*, 80(4), 621-630. doi:10.1016/0092-8674(95)90516-2
- de Groot, R. J., Rümenapf, T., Kuhn, R. J., Strauss, E. G., & Strauss, J. H. (1991). Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc Natl Acad Sci U S A*, 88(20), 8967-8971. doi:10.1073/pnas.88.20.8967
- Ehteshami, M., Tao, S., Zandi, K., Hsiao, H.-M., Jiang, Y., Hammond, E., . . . chemotherapy. (2017). Characterization of β -d-N 4-hydroxycytidine as a novel inhibitor of chikungunya virus. 61(4), e02395-02316.
- Frolov, I., Agapov, E., Hoffman Jr, T. A., Prágai, B. I. M., Lippa, M., Schlesinger, S., & Rice, C. M. J. J. o. v. (1999). Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. 73(5), 3854-3865.
- Greene, I. P., Paessler, S., Anishchenko, M., Smith, D. R., Brault, A. C., Frolov, I., & Weaver, S. C. (2005). Venezuelan equine encephalitis virus in the guinea pig model: evidence for epizootic virulence determinants outside the E2 envelope glycoprotein gene. *Am J Trop Med Hyg*, 72(3), 330-338.

- Hardy, W. R., & Strauss, J. H. (1989). Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J Virol*, 63(11), 4653-4664. doi:10.1128/jvi.63.11.4653-4664.1989
- He, Y., Liu, P., Wang, T., Wu, Y., Lin, X., Wang, M., . . . Zhao, X. J. V. (2019). Genetically stable reporter virus, subgenomic replicon and packaging system of duck Tembusu virus based on a reverse genetics system. 533, 86-92.
- Jacups, S. P., Whelan, P. I., Currie, B. J. J. V.-B., & Diseases, Z. (2008). Ross River virus and Barmah Forest virus infections: a review of history, ecology, and predictive models, with implications for tropical northern Australia. 8(2), 283-298.
- Jones, R., Hons, M., Rabah, N., Zamarreño, N., Arranz, R., & Reguera, J. (2023). Structural basis and dynamics of Chikungunya alphavirus RNA capping by nsP1 capping pores. *Proc Natl Acad Sci U S A*, 120(12), e2213934120. doi:10.1073/pnas.2213934120
- Jose, J., Snyder, J. E., & Kuhn, R. J. (2009). A structural and functional perspective of alphavirus replication and assembly. *Future Microbiol*, 4(7), 837-856. doi:10.2217/fmb.09.59
- Kääriäinen, L., & Söderlund, H. (1978). Structure and replication of alpha-viruses. *Curr Top Microbiol Immunol*, 82, 15-69. doi:10.1007/978-3-642-46388-4_2
- La Linn, M., Gardner, J., Warrilow, D., Darnell, G. A., McMahon, C. R., Field, I., . . . Suhrbier, A. (2001). Arbovirus of marine mammals: a new alphavirus isolated from the elephant seal louse, *Lepidophthirus macrorhini*. *J Virol*, 75(9), 4103-4109. doi:10.1128/jvi.75.9.4103-4109.2001
- Laureti, M., Paradkar, P. N., Fazakerley, J. K., & Rodriguez-Andres, J. J. V. (2020). Superinfection exclusion in mosquitoes and its potential as an arbovirus control strategy. 12(11), 1259.
- Lello, L. S., Bartholomeeusen, K., Wang, S., Coppens, S., Fragkoudis, R., Alpey, L., . . . Utt, A. (2021). nsP4 Is a Major Determinant of Alphavirus Replicase Activity and Template Selectivity. *J Virol*, 95(20), e0035521. doi:10.1128/jvi.00355-21
- Li, G. P., & Rice, C. M. (1989). Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational readthrough and its effect on virus replication. *J Virol*, 63(3), 1326-1337. doi:10.1128/jvi.63.3.1326-1337.1989
- Lu, G., Chen, R., Shao, R., Dong, N., Liu, W., & Li, S. (2020). Getah virus: An increasing threat in China. *J Infect*, 80(3), 350-371. doi:10.1016/j.jinf.2019.11.016

- Lulla, V., Karo-Astover, L., Rausalu, K., Saul, S., Merits, A., & Lulla, A. (2018). Timeliness of Proteolytic Events Is Prerequisite for Efficient Functioning of the Alphaviral Replicase. *J Virol*, 92(14). doi:10.1128/jvi.00151-18
- Mandary, M. B., Masomian, M., & Poh, C. L. J. I. j. o. m. s. (2019). Impact of RNA virus evolution on quasispecies formation and virulence. 20(18), 4657.
- Michie, A., Ernst, T., Chua, I.-L. J., Lindsay, M. D., Neville, P. J., Nicholson, J., . . . Imrie, A. J. V. (2020). Phylogenetic and timescale analysis of Barmah forest virus as inferred from genome sequence analysis. 12(7), 732.
- Nowee, G., Bakker, J. W., Geertsema, C., Ros, V. I., Göertz, G. P., Fros, J. J., . . . Biology, D. (2021). A tale of 20 alphaviruses; inter-species diversity and conserved interactions between viral non-structural protein 3 and stress granule proteins. 9, 625711.
- Panas, M. D., Ahola, T., & McInerney, G. M. J. J. o. v. (2014). The C-terminal repeat domains of nsP3 from the Old World alphaviruses bind directly to G3BP. 88(10), 5888-5893.
- Pohjala, L., Utt, A., Varjak, M., Lulla, A., Merits, A., Ahola, T., & Tammela, P. J. P. o. (2011). Inhibitors of alphavirus entry and replication identified with a stable Chikungunya replicon cell line and virus-based assays. 6(12), e28923.
- Powers, A. M., & Roehrig, J. T. (2011). Alphaviruses. *Methods Mol Biol*, 665, 17-38. doi:10.1007/978-1-60761-817-1_2
- Rikkonen, M., Peränen, J., & Kääriäinen, L. (1994). ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J Virol*, 68(9), 5804-5810. doi:10.1128/jvi.68.9.5804-5810.1994
- Rubach, J. K., Wasik, B. R., Rupp, J. C., Kuhn, R. J., Hardy, R. W., & Smith, J. L. (2009). Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology*, 384(1), 201-208. doi:10.1016/j.virol.2008.10.030
- Shin, G., Yost, S. A., Miller, M. T., Elrod, E. J., Grakoui, A., & Marcotrigiano, J. J. P. o. t. N. A. o. S. (2012). Structural and functional insights into alphavirus polyprotein processing and pathogenesis. 109(41), 16534-16539.
- Strauss, E. G., Levinson, R., Rice, C. M., Dalrymple, J., & Strauss, J. H. (1988). Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: sequence and comparison with those of other alphaviruses. *Virology*, 164(1), 265-274. doi:10.1016/0042-6822(88)90644-7

- Strauss, J. H., & Strauss, E. G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev*, 58(3), 491-562. doi:10.1128/mr.58.3.491-562.1994
- Tamm, K., Merits, A., & Sarand, I. (2008). Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. *J Gen Virol*, 89(Pt 3), 676-686. doi:10.1099/vir.0.83320-0
- Tamm, K., Merits, A., & Sarand, I. J. T. J. o. g. v. (2008). Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. 89(Pt 3), 676.
- Tan, Y. B., Lello, L. S., Liu, X., Law, Y. S., Kang, C., Lescar, J., . . . Luo, D. (2022). Crystal structures of alphavirus nonstructural protein 4 (nsP4) reveal an intrinsically dynamic RNA-dependent RNA polymerase fold. *Nucleic Acids Res*, 50(2), 1000-1016. doi:10.1093/nar/gkab1302
- Tomar, S., Hardy, R. W., Smith, J. L., & Kuhn, R. J. (2006). Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J Virol*, 80(20), 9962-9969. doi:10.1128/jvi.01067-06
- Tsetsarkin, K. A., Chen, R., Sherman, M. B., & Weaver, S. C. (2011). Chikungunya virus: evolution and genetic determinants of emergence. *Curr Opin Virol*, 1(4), 310-317. doi:10.1016/j.coviro.2011.07.004
- Utt, A., Das, P. K., Varjak, M., Lulla, V., Lulla, A., & Merits, A. J. J. o. v. (2015). Mutations conferring a noncytotoxic phenotype on chikungunya virus replicons compromise enzymatic properties of nonstructural protein 2. 89(6), 3145-3162.
- Zhang, K., Law, Y. S., Law, M. C. Y., Tan, Y. B., Wirawan, M., & Luo, D. (2021). Structural insights into viral RNA capping and plasma membrane targeting by Chikungunya virus nonstructural protein 1. *Cell Host Microbe*, 29(5), 757-764.e753. doi:10.1016/j.chom.2021.02.018

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC

I, Aylin Sultanova

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, Stable expression of Barmah Forest virus replicons in cultured mammalian cells

supervised by Ailar Omler

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Aylin Sultanova

24/05/2024