الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

N°Ref: .....

Centre Universitaire Abdelhafid BOUSSOUF de Mila

Institut des Sciences et de la Technologie **Département des Sciences de la Nature et de la Vie**

Mémoire préparé en vue de l'obtention du diplôme de

Master

Filière : Sciences Biologiques Spécialité : Biochimie Appliquée

Thème :

In silico characterization of Zika virus envelope protein structure (2019)

Présentée par :

- BOULDJEDJE Sofia
- AMOURA Hamida
- LEBCIR Rima

Devant le jury composé de :

Mr. BOUTALAA Saber	Maitre de Conférences B	Président
Mr. BOUNAMOUS Azzedine	Maitre de Conférences A	Examinateur
Mr. BOUBENDIR Abdelhafid	Maitre de Conférences A	Promoteur

Année Universitaire: 2019/2020

Acknowledgment

First and foremost, praises and thanks to Allah the almighty, for his showers of blessings throughout our research work to complete the research successfully And we do not forget the virtue, his mercy, his Blessings are many on us, we do not count them and we do not count praise on him and we cannot appreciate and he is glorified as he praised himself and Glory be to him the worlds.

We would first like to thank my supervisor **Mr. BOUBENDIR Abdelhafid** for its teaching and its supervision Thanks to him, we learned this new approach to working in silico and we started the path of scientific research in Biology. Throughout this memory, his relevant advice with listening, kindness and patience allowed our work to succeed and see the light of day.

We offer our respects and thanks to the members of the jury who did us the honor of evaluating this work **Mr. BOUTALAA Saber** president of the jury and **Mr. BOUNAMOUS Azzedine** examiner, who showed interest and availability for the examination of our memory.

Many thanks to everyone who contributed to the success of this work from far and near, whether material or moral.

Dedications

I dedicate this memory to the most loving mother in the world NADIA, little darling mom, every day you prove your love to me. My turn to give you the proof that I love you with all my heart. I promise to try to be more successful. I know that the greatest gift I can give you is to make an effort and to love you very much, may Allah keep you for me. And to my darling father Mahfoud May God have mercy on you and make you dwell in Paradises.

To my grandfather Ibrahim I may not have been born in the richest of families, but I grew up with the sweetest of grandfathers. Thank you, grandfather, for this beautiful childhood that you gave me and for all this love that is priceless. Thank you for working hard with me for 21 years to study me and get me out of the pangs to the light. All I got today was thanks to you. And to my grandmother Yamina may god have mercy on you and make you dwell in paradises.

I dedicate this work to my husband Bilal No words can express you my deep attachment and my gratitude for love, tenderness and kindness which you always surrounded me.Dear husband, I would like you to find in this work the expression of my feelings of gratitude the most sincere because thanks to your help and your patience with me that this work has seen the light of day.

To all my brothers who have supported me throughout my life Djamal, Aymen, Fateh, Fares, Samir, and Mouloud, and to all my sisters Fatima, Rima,and Lamis. And to the wives of my brothers Hayat, Yasmina, Fadila, and Tawas. To my loved cousins hassina, saida, adel, noureddine, ahmed, and bachir. To my all aunts and uncles, to my little chicks Adem, Tasnim, and Iyad.

Special dedications for my dear sisters and friends at the same time nadia, nawal, and nourelhouda. To my girlfriends from primary to university Hadjer, Rima, Malika, Ismahane, and Hadjer. And the great dedication with the great respect for my primary teachers Khalfa nawar and kouicem Abdelhak who taught me the real value of science and educated me on the love of learning.

Sofia

I dedicate this humble work....

✨ *To my very dear mother* ***Fatima zohra*** ✨

who worked for my success, with her love, her support, all the sacrifices made and her precious advice, for all her assistance and her presence in my life. May Almighty God preserve you and grant you health, long life and happiness.

✨ *To my very dear father* ***Abdelkrim*** ✨

who can be proud and find here the result of long years of sacrifice and hardship to help me move forward in life. Thank you for the noble values, the education and the permanent support that came from you. Almighty God preserves you and gives you health, happiness and long life.

✨ *To my dear brother* ***Aissa*** ✨

No dedication can express the love, esteem, dedication and respect that I have always had for you. I wish you a future full of success and serenity.

✨ *To my sisters* ***Aicha*** *and* ***Lamis*** ✨

Thank you, lovely sisters, for showing so much kindness and helpfulness towards me. May Allah, the Highest, grant you a happy life and a prosperous future.

✨ *To all my family: uncles, aunts, cousins and cousins* ✨

To all the members of my family, young and old Please find in this modest work the expression of my Affection.

✨ *To my dear friends Houda and Dounia* ✨

who have helped me throughout my student life.

✨ *To my girlfriends at this work* ***Rima*** *and* ***Sofia*** ✨

To those who have helped me from near or far.

To all the others that I have not mentioned but who I am also thinking of.

Hamida

I dedicate this work to:

**God Almighty*

my creator, my strong pillar for giving me this opportunity to do my Masters degree which would not have been possible otherwise.

**To my source of inspiration and understanding; my loving parents whose affection, love, encouragement ,prays and Douaa of day and night make me able to get such success and honor.*

My great beloved wonderful mother,who never stop giving of herself in countless ways, who spread the happiness tenderness and warm. in my heart she is my teacher in life

My first and my forever hero, my precious father who leads me through the valley of darkness with light of hope and support who stand by me when things look bleak.

He who always says : "I am here for you , I am living for you "

**To my dearest sisters brothers Soumia Souria Bouchra Walid Badis also brother's wife Ilham and to your pure soul broth "OUSSAMA"*

**To the sweet heart lovely princes" Nourssine" my niece.*

**To the woman of Kindness and tenderness the iron woman my aunt Naima*

**To the beautiful uant Noura.*

And i would like to take this opportunity to say warm thanks to my sweet cousins Nassiha Djenat Radia Ibtissam Ahlam Hayat Zahra Sana Mouna Nada Kamer Youssra Nedjela Safa Maroua Abir Zainab and Meriem and special feeling of gratitude to Assaad a cousin with a brother flavor i wish to you all the best.

Profound gratitude to my teachers Mr.Zouaghi Mohammed ; Mrs Harrieche Ouahiba and Mrs.Boucherit Hanane .They left a glowing imprint through my university carrer.

A deep feeling to all my friends far or near : Imane .Ikhlas never forget her help and magnanimity.Maram.Micha Bouchra and her Her valuable advice. Aya wish to you all the best Amel , Anfal and Khawla Saadi

**To thesis mate, Hamida and Sophia Thank you from the heart for the wonderful treatment in the difficult circumstances I have been through.*

**To the person who I can't classify my relation with her: friend , dream mate or sister Souad(Amina) Distance could never destroy the forever bond we have That you were never not here, you were always here to talk to share to advice to hear me to understand and believe in my dreams wish to you all the best God bless you.*

**Thanks To the lost dream ,you opened new horizons for me.*

Last but not least, deepest thanks go to all people who took part in making this thesis real.

Rim

Table of Contents

Part II: Experimental search

List of abbreviations

A: *Aedes*

AXL: anexelekto

C: capsid

CF: complement fixation

Cryo-EM: cryo-electron microscopy

DB: dumbbell

DC-SIGN: Dendritic cell specific icam grabbing non-integrin

DENV: Dengue virus

E: envelope

ER: endoplasmic reticulum

GBS : Guillain Barré syndrom

GPI: glycosylphosphatidylinositol

HI: Hemagglutination inhibition

ICTV: International Committee on Taxonomy of Viruses

ID: identifiant

IFN: interferon

IgG: Immunoglobulin G

IgM: Immunoglobulin M

Mo-DCs: Monocyte and dendritic cells

NCR: non coding region

NS: non structural

NT: neutralization tests

ORF: open reading frame

prM: premembrane

RMSD: Root-Mean-Square Deviation

RT-PCR: Reverse transcription polymerase chain reaction

sHP-3′ SL: small hairpin 3′-stem-loop

SL: Stem loop

SPOV: spondweni virus

TAM: tyrosine 3 AXL and MER

TIM: T cell/transmembrane, immunoglobulin, and mucin

TLR3: Toll-like receptor 3

UTR: untraslated region

WNV: West Nile Virus

YF: yellow fever

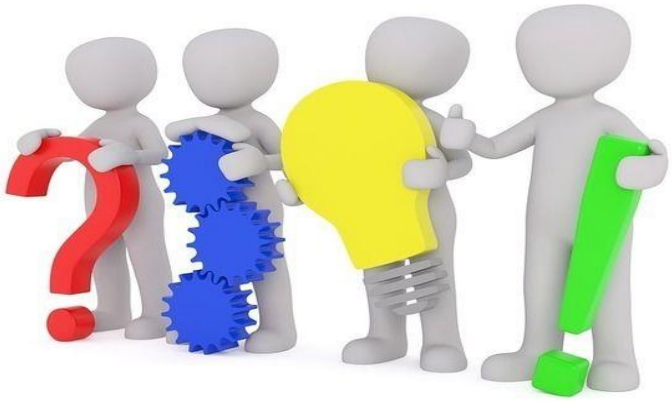
YFV: yellow fever virus

ZIKV: Zika virus

List of figures

List of tables

Introduction



Introduction

Zika virus (ZIKV) is an emerging virus that has been defined by the World Health Organization as a serious global biological-threat (Ramharack et al., 2016), it is a positive single-stranded RNA virus that is transmitted by mosquito bites (Ashfaq et al., 2016). A large number of serological studies in the half century since the discovery of ZIKV have revealed a broad but confined geographic distribution of human infection with the virus, across a relatively narrow equatorial belt running from Africa to Asia (Song et al., 2017).

The first isolation of zika virus was in Uganda in 1947(Faye et al., 2014) from Rhesus macaque and the first isolation from human was in Nigerian, Africa in 1954; it is originally transmitted in Africa through a sylvatic cycle involving mainly Aedes vectors and nonhuman primates, with humans being occasional hosts (Baronti et al., 2014).

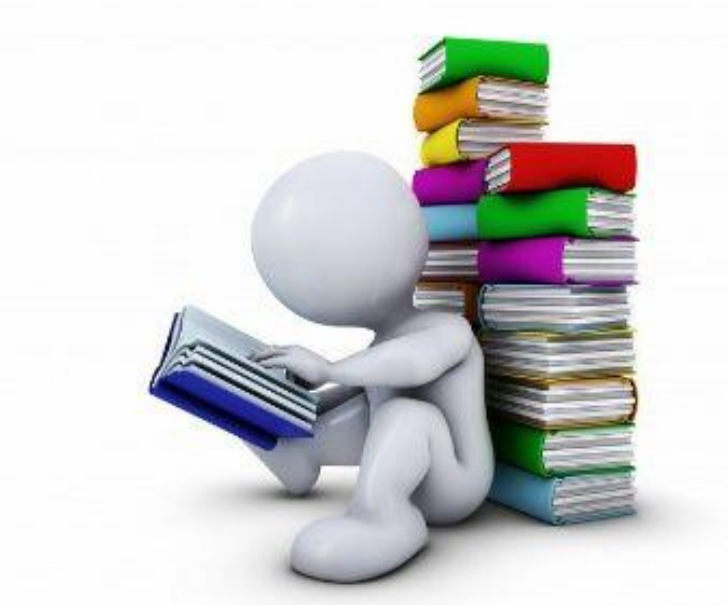
This virus causes the neurodevelopmental congenital Zika syndrome and that has been linked to the neuroinflammatory Guillain–Barré syndrome. The absence of a vaccine or a clinically approved drug to treat the disease combined with the likelihood that another outbreak will occur in the future defines an unmet medical need (Bernatchez et al., 2019).

ZIKV Envelop protein is antigenic and is involved in fusion and entry of viral particles into the cell (Ashfaq et al., 2016). For this, the structural knowledge of the ZIKV proteins (exactly ZIKV Envelop protein) may allow us to understand exposed epitopes which will facilitate the development of specific diagnostic reagents that differentiate it from dengue and other flaviviruses. Furthermore, open sharing of the three-dimensional arrangement of viral surface proteins could allow the mapping of potential neutralizing epitopes, guiding efforts to rationally design effective vaccines (Ekins et al., 2016).

The objectives of the present study are:

- Gathering protein sequences in FASTA format from UniProt.
- Alignment of amino acids sequences using Clustal Omega on UniProt.
- Phylogenetic analysis by MEGA6 software.
- Prediction of Zika virus 3D protein envelope structure using SWISS-MODEL online tool.

Part I
Bibliographic Search



Chapter I

Generalities

Chapter I: Generalities

1. History of the Zika virus discovery

ZIKV was discovered in the Zika Forest of Uganda during research supported by the Rockefeller Foundation to study the enzootic or sylvatic cycle of yellow fever (YF) virus and to identify additional arboviruses (Weaver et al., 2016).

In brief, ZIKV was first isolated in 1947 from a febrile sentinel rhesus monkey (no. 766) held in a cage on a platform in the canopy of the Zika Forest in Uganda during studies to identify the vector of sylvatic YF. A blood sample from this monkey was collected on day 3 of fever and was inoculated intracerebrally into Swiss mice and into another rhesus monkey (no. 771). All mice showed signs of illness on day 10 post inoculation, and a filterable transmissible agent was isolated from the brains of these sick mice. Monkey 766 showed no other clinical signs or symptoms, and monkey 771 remained asymptomatic. The convalescent serum from both monkeys (766 and 771) neutralized the virus isolated from monkey 766 in mice, which was designated ZIKV 766. Preinfection sera from these monkeys did not neutralize ZIKV 766 (Gubler et al., 2017).

In January 1948, mosquitoes were collected in the Zika Forest in an attempt to isolate YFV. Eighty-six *Aedes africanus* mosquitoes were collected, and mice were inoculated with the Seitz filtrate of pools of these mosquitoes. One mouse died on day 6 after inoculation, and one appeared sick on day 14. The virus isolated from *Ae. africanus* was designated ZIKV(E/1strain). The remaining portion of the Seitz filtrate was inoculated subcutaneously into rhesus monkey no.758. This monkey remained asymptomatic, but two mice inoculated intracerebrally with blood taken from this monkey died and another became sick; ZIKV (758 strain) was isolated from its serum. Rhesus monkey no. 758 developed neutralizing antibodies to the agent isolated from its serum, to the strain of virus isolated from *Ae. africanus* (ZIKVE/1strain), and to the strain isolated from rhesus monkey no. 766 (the ZIKV 766 strain). Cross neutralization tests (NT) showed that ZIKV was different from YFV, DENV, and Theiler's encephalomyelitis virus; NT with ZIKV and the antisera from other neurotropic viruses showed no relationship. Cross-reactions performed by complement fixation (CF) confirmed that ZIKV was a distinct virus (Musso et Gubler., 2016).

There is some controversy surrounding the first ZIKV isolate from humans. The first report was from serum of a 10 year old Nigerian female in 1954. The patient was clearly jaundiced, but interpretation of the clinical presentation was complicated by coinfection with malaria. In cross-neutralization tests with convalescent sera from monkeys infected with Bunyamwera, Bwamba, Mengo, Ntaya, Semliki Forest, Uganda S, West Nile, YF, and Zika viruses, only the serum from the monkey infected with ZIKV neutralized the virus isolated from the patient, strongly suggesting the girl was infected with ZIKV (Gubler et al., 2017).

2. Classification

The virus is a member of the family *Flaviviridae* (Table 1), and is transmitted to humans by *Aedes* species mosquitoes (Wikan et Smith., 2016). It is a *flavivirus* and is related to other arboviruses such as YFV, Japanese Encephalitis virus, Dengue virus, and West Nile virus (Figure 1) (Shapshak et al., 2015).

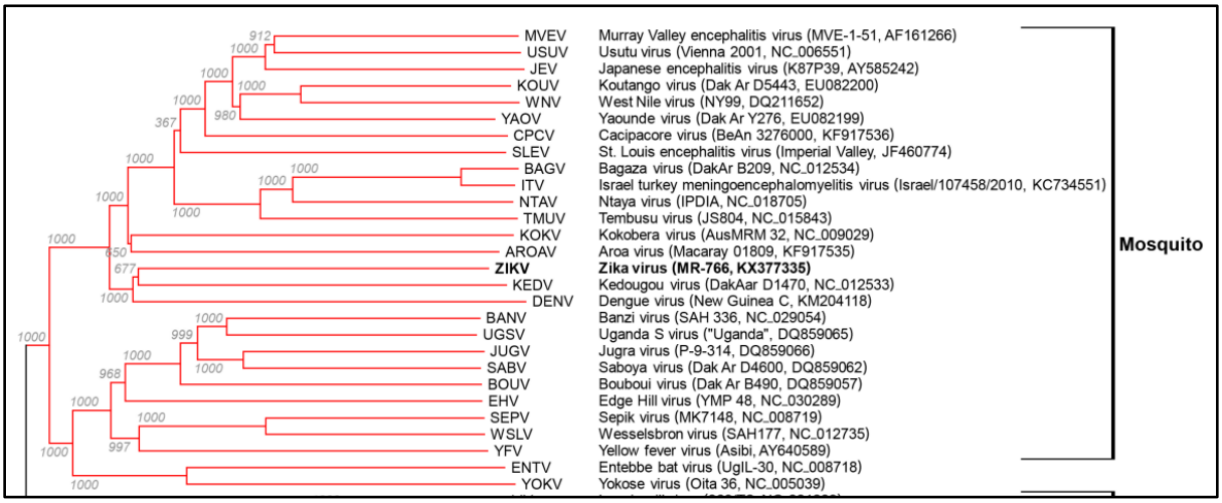


Figure 1. A rooted phylogenetic tree based on the nucleotide sequence of complete or near-complete genomes of all 46 available flaviviruses (Song et al., 2017).

Table 1. Taxonomy of Zika virus ((https://www.uniprot.org/taxonomy/64320), 2020).

Taxonomy	Zika virus
Kingdom	Virus
Realm	*Riboviria*
Family	*Flaviviridae*
Genus	*Flavivirus*

2.1. Riboviria

In February 2019 the International Committee on Taxonomy of Viruses (ICTV) has approved, by an absolute majority, the creation of the realm *Riboviria*, a likely monophyletic group encompassing all viruses with positive-strand, negative-strand and double-strand genomic RNA that use cognate RNA-directed RNA polymerases for replication (Walker et al., 2019).

2.2. Flaviviridae

The *Flaviviridae* is a family of small enveloped viruses with RNA genomes of 9000-13000 bases. Most infect mammals and birds. Many *Flaviviruses* are host-specific and pathogenic, such as hepatitis C virus in the genus Hepacivirus. The majority of known members in the genus *Flavivirus* are arthropod borne, and many are important human and veterinary pathogens (Simmonds et al., 2017).

2.3. Flavivirus

This genus consists primarily of more than 50 species of arthropod borne viruses, with distinct groups infecting mosquitoes or ticks. Mammals and birds are the usual primary hosts, in which infections range from asymptomatic to severe or fatal hemorrhagic fever or neurological disease. Important human pathogens include yellow fever virus, dengue virus, Japanese encephalitis virus, West Nile virus and tick-borne encephalitis virus. Other members cause economically important diseases in domestic or wild animals. Additional viruses

infecting only arthropods or only mammals have been described recently (Simmonds et al., 2017).

3. Mechanism of infection

Host-virus interactions that shape ZIKV infection remain poorly characterized. The human dermal fibroblasts, epidermal keratinocytes, and immature dendritic cells all were permissive to a ZIKV isolate from French Polynesia. TLR3 (Toll-like receptor 3) was identified as the initial immune receptor involved in the sensing of ZIKV infection in human fibroblasts leading to type I and type II interferon (IFN) responses. ZIKV also interacted with DC-SIGN (Dendritic cell specific icam grabbing non-integrin) to initiate infection of immature Mo-DCs (Monocyte and dendritic cells), whereas members of the TIM (T cell/transmembrane, immunoglobulin, and mucin) and TAM (tyrosine 3 AXL and MER) family of phosphatidylserine receptors possibly serve as receptors or attachment factors for other cells; in cutaneous fibroblasts and epidermal keratinocytes lacking expression of DC-SIGN, the TAM receptor AXL (anexelekto) facilitated ZIKV entry . The scientists examined the receptor repertoire of human radial glia cells in the fetal brain involved in ZIKV attachment and entry during neurogenesis. Distinct *flavivirus* entry receptor genes, including AXL receptor, were enriched in radial glia cells, astrocytes, endothelial cells, and microglia, suggesting that these cell populations may be particularly vulnerable to ZIKV infection in the developing brain (Olagnier et al., 2016).

4. Target cells

4.1. Nerve cells

Since the emergence of ZIKV virus, reports of microcephaly have increased considerably in the world. Studies of the effects of ZIKV infection in human neural stem cells growing as neurospheres and brain organoids, using immunocytochemistry and electron microscopy, showed that ZIKV targets human brain cells, reducing their viability, and neural stem cells can be infected and become apoptotic following infection. These results suggest that ZIKV abrogates neurogenesis during human brain development (Garcez et al., 2016).

4.2. Skin cells

Human skin cells are permissive for ZIKV infection and replication. Given the capacity of mosquitoes to inoculate ZIKV into the human skin during the blood-feeding process, the

potential target cells for infection with this virus are likely to be localized to the epidermis and dermis, which constitute the first line of defense. The skin fibroblasts, which have been recognized as a permissive target for various arboviruses, they were infected in vitro with ZIKV. The viral envelope protein was detected in several cells, 100% of the infected cells expressed ZIKV, the results show a gradual increase in the production of viral particles over time. Next, given the observation that the epidermal layer is comprised mainly of keratinocytes, these cells could also be a target for ZIKV. Dendritic cells have been reported to be permissive for ZIKV infection, and as such, they are recognized as an important target for propagation of this virus in the human skin, the capacity of ZIKV to replicate ex vivo in human skin cells was also studied (Hamel et al., 2015).

4.3. Placental cells

ZIKV infects primary human placental cells from mid and late gestation and chorionic villus explants from early gestation, these cells, along with a primary human umbilical vein endothelial cell line from umbilical cord, were infected with prototype ZIKV strain MR766 and were immunostained with monoclonal antibodies to E glycoprotein and nonstructural protein NS3, the results show that different types of primary cells from mid- and late-gestation placentas are permissive to infection with prototype and contemporary ZIKV strains. ZIKV virus infection during pregnancy is linked to severe birth defects, but mother to fetus transmission routes are unknown (Tabata et al., 2016).

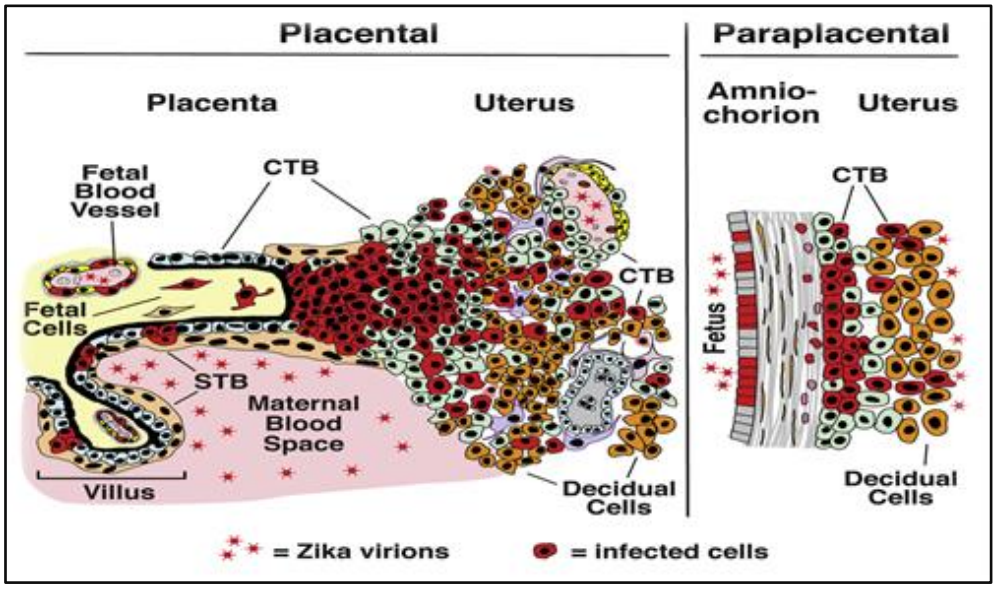


Figure 2. Infected placental cells by ZIKV presented with red color (Tabata et al., 2016).

4.4. Eye cells

Several clinical studies have observed eye malformations and pathology in neonates born to mothers infected with Zika during pregnancy. Manifestations of eye disease in newborns with zika include chorioretinal atrophy, intraretinal hemorrhages, and blindness. Viral infection in the eye can cause inflammation of uveal tissues (retina, choroid, iris, and ciliary body), also termed uveitis, which can lead to permenant vision loss if untreated. Zika causes conjunctivitis in most of infected adults. Fluid sampled from the anterior chamber of eye contained viral RNA, suggesting that Zika can replicate within the eye at some stage of clinical syndrome (Miner et al., 2016).

4.5. Blood cells

Some studies in macaques have allowed detected the presence of RNA ZIKV viral in whole blood and plasma, the animals experienced no clinical disease but developed short lived plasma viremias that cleared as neutralizing antibody developed. Despite no major histopathologic changes, many adults tissues contained RNA viral with highest levels in hemolymphatic tissues, these observations warrant further studies to investigate Zika persistence and its potential clinical implications for transmission via blood products or tissue and organ transplants (Coffey et al., 2017).

5. Clinical presentation

On 1 February 2016, the World Health Organization (WHO) declared that the recent cluster of microcephaly cases and other neurological disorders reported in the Americas, where an outbreak of Zika virus infection is ongoing, constitutes a public health emergency of international concern (Charrel et al., 2016).

ZIKV can produce a wide variety of clinical symptoms in humans. A growing body of evidence suggests that in some severe cases, ZIKV causes neurological diseases, such as Guillain-Barré syndrome in ZIKV-infected adults and microcephaly in infants born to ZIKV-infected women (Song et al., 2017).

ZIKV infection is not fatal. However, the first fatal case of ZIKV-associated encephalitis was reported in 2016 in a 47 year old non-pregnant woman, soon followed by the report of

three additional ZIKV-related fatalities with one of the patients being severely immunocompromised) (Gorshcov et al., 2019).

5.1. Common signs and symptoms

The clinical manifestations of Zika infection are very similar to those of other arboviruses such as dengue and chikungunya (Table 2) (Amorim L., 2019). ZIKV infections are symptomatic in only ~20-25% of the infected individuals who develop a mild and self-limited illness, with an incubation period of 4-10 days in symptomatic cases, the common symptoms are nonspecific and resemble those of a flu-like syndrome, including transient low-grade fever, itchy maculopapular rash (Figure 3 B), arthritis or arthralgia, and non-purulent conjunctivitis (Figure 3 A); at a lesser frequency, retro-orbital pain, headache, myalgia, edema, and vomiting are seen. Other clinical manifestations observed with acute ZIKV infection include hematospermia, hearing difficulties, thrombocytopenia, and subcutaneous bleeding. The symptoms generally appear along with the viremia and disappear spontaneously within a week, but arthralgia may persist for up to a month (Song et al., 2017).

Symptoms	Dengue	Chikungunya	Zika
Fever	++++	+++	+++
Myalgia/arthralgia	+++	++++	++
Edema of extremities	0	0	++
Maculopapular rash	++	++	+++
Retro-orbital pain	++	+	++
Conjunctivitis	0	+	+++
Lymphadenopathies	++	++	+
Hepatomegaly	0	+++	0
Leukopenia/thrombopenia	+++	+++	0
Hemorrhage	+	0	0

0: absence of symptoms. (+) (++) (+++) (++++): density of symptoms

Table 2. Comparison among dengue, chikungunya, and Zika symptoms (Amorim L., 2019).

An additional risk of ZIKV in adults is damage to the testis. A study in mice reported the persistence of ZIKV in the testis and epididymis leading to extensive tissue damage. Male mice were reported to exhibit oligospermia, diminished testosterone. A more recent study reinforced these observations and revealed that peritubular spermatogonium cells are vulnerable to ZIKV infection. Furthermore, even an acute, uncomplicated, symptomatic ZIKV infection may result in microhematospermia in the absence of hematuria (Gorshcov et al., 2019).

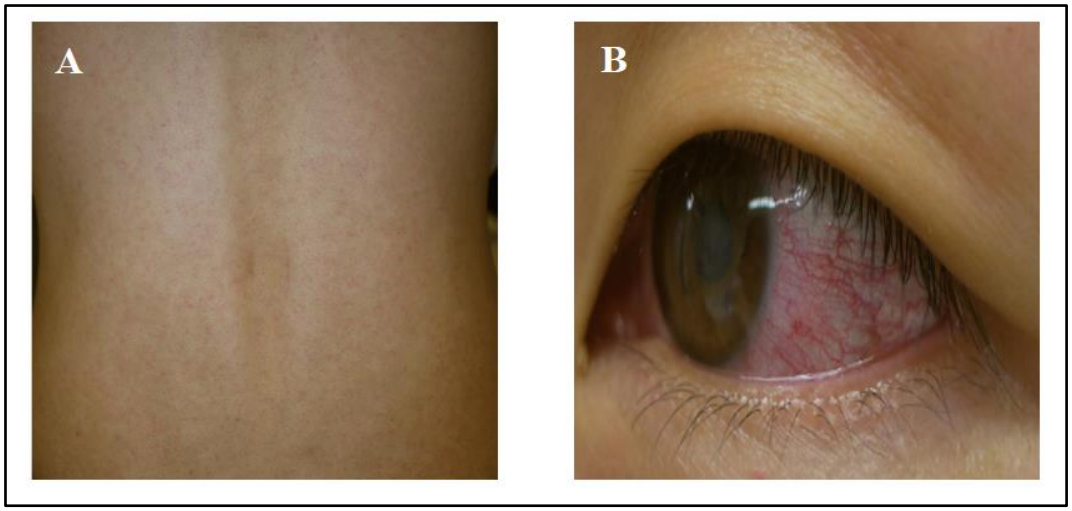


Figure 3. A- japan Rash case, B- Japan conjunctivitis case (Kutsuna et al., 2014).

5.2. Guillain-Barré syndrome

Guillain-Barré syndrome is an autoimmune disease in which the immune system attacks part of the peripheral nervous system, causing tingling, muscle weakness, paralysis, and even death. Previously, this neuromuscular complication had been associated with infection by other arboviruses, such as DENV and Chikungunya virus. The temporal and geographic association of ZIKV with Guillain-Barré syndrome was initially observed during the 2013-2014 outbreak reported in French Polynesia, and subsequently during the 2015-2016 outbreak that is still ongoing in the Americas, during the previous French Polynesia outbreak, the incidence of Guillain Barré syndrome was estimated to be ~20-fold higher than its basal incidence of 1-2 cases per 100,000 population per year more definitively, a recent case-control study revealed that anti-ZIKV IgM or IgG was detected in 41 (98%) of 42 patients with Guillain-Barré syndrome, and all had neutralizing antibodies against ZIKV, as compared to 54 (55%) of 98 controls: age-, sex- and residence-matched patients with a non-febrile

illness. The same study also showed that patients with Guillain-Barré syndrome had electrophysiological characteristics consistent with the acute motor axonal neuropathy type of the disease. Thus far, ZIKV-induced Guillain Barré syndrome has been transient, and most patients have recovered fully. Currently, the mechanism by which ZIKV infection leads to Guillain-Barré syndrome is unknown but is under active investigation (Song et al., 2017). 18 countries and territories have reported an increased incidence of GBS and/or laboratory confirmation of a Zika virus infection among GBS cases (Table 3) (WHO, 2016).

Table 3. Countries and territories reporting Guillain-Barré syndrome (GBS) potentially associated with Zika virus infection (WHO, 2016).

Classification	Country/territory
Reported increase in incidence of GBS cases, with at least one GBS case with confirmed Zika virus infection	Brazil, Colombia, Dominican Republic, El Salvador, French Guiana, French Polynesia, Honduras, Jamaica, Martinique, Suriname, Venezuela.
No increase in GBS incidence reported, but at least one GBS case with confirmed Zika virus infection	Costa Rica, Grenada, Guadeloupe, Guatemala, Haiti, Panama, Puerto Rico.

5.3. Microcephaly

Microcephaly is a neurological condition in which the brain of a baby does not develop properly, causing the head to be smaller than normal. It is divided into two types: primary or congenital microcephaly, which is present in utero or at birth; and secondary or postnatal microcephaly, which develops after birth While primary microcephaly is likely caused by a decrease in the number of neurons produced during neurogenesis, secondary microcephaly is presumably caused by a reduction in the number of dendritic processes and synaptic connections, microcephaly can be caused by a variety of genetic mutations, peri- and post-natal brain injuries, teratogenic agents, and congenital infections (Figure 4). For ZIKV, a causal relationship between prenatal infection and microcephaly emerged in Brazil, as the number of newborns with microcephaly began to rise in September 2015 (thereafter, 8,301 cases of microcephaly were recorded from November 2015 to July 2016 in that country Subsequently, the potential risk of microcephaly associated with ZIKV infection has been suggested by two retrospective studies from French Polynesia and one prospective study from

Brazil. In addition to this spatiotemporal association, ZIKV or its gene expression has been detected in the amniotic fluid and in various tissues of fetuses with microcephaly and those who died after birth or following abortion . Moreover, ZIKV-specific IgM has been identified in the cerebrospinal fluid and serum of neonates with microcephaly (Song et al., 2017).

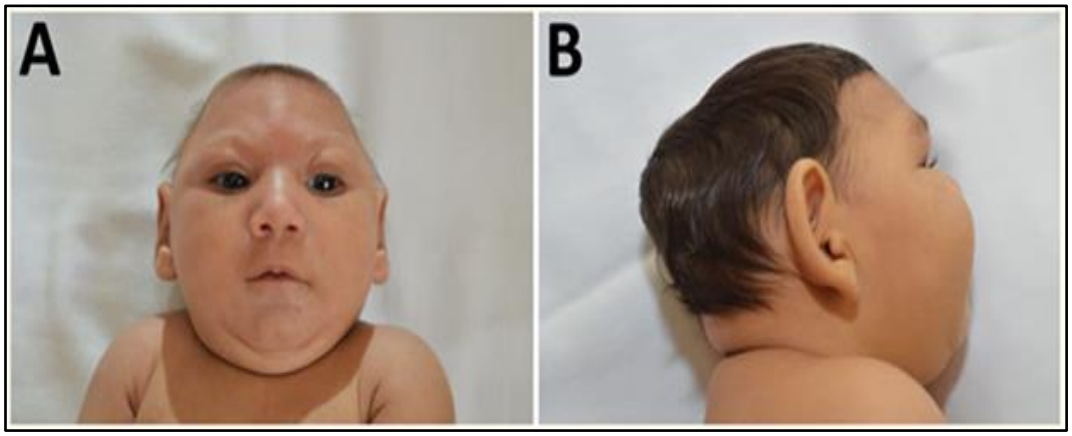


Figure 4. Characteristic phenotype of fetal brain disruption sequence in infants with probable congenital ZIKV syndrome: A) craniofacial disproportion and biparietal depression, B) prominent occiput (Moura da Saliva et al., 2016).

5.4. Other neurological complications

Case reports suggest that ZIKV infection may be associated with other neurological complications, including the case of an 81-year-old man who developed meningoencephalitis and the case of a 15-year-old girl who was diagnosed with acute myelitis research progresses, it is likely that other neurological and non-neurological complications caused by ZIKV infection will be identified in the future. Case control or cohort studies, together with well-characterized case reports, will be required in order for us to better understand the potential role of ZIKV infection in adults and newborns (Song et al., 2017).

6. Diagnosis

The diagnosis of infection by ZIKV is based on clinical, epidemiological and laboratorial criteria. Because the symptoms of ZIKV disease are nonspecific and can easily be confused with those of other arbovirus-induced diseases, such as Dengue and Chikungunya, in regions where those viruses co-circulate. Laboratorial diagnosis of ZIKV can be realized by the detection of virus, viral nucleic acid, viral antigen, or antibody or by a combination of these techniques. The choice of method depends on the purpose for which the test is performed

(clinical, epidemiological study, or vaccine development), the type of laboratory facilities and expertise available, and the sample collection time. When the sample is collected in the first few days after the onset of symptoms, a test detecting virus or viral nucleic acid may be performed. The virus detection is based on isolation from cell culture (using mosquito or mammalian cell lines), directly from mosquitoes, or intracerebrally from newborn mice. The low levels of viremia may explain this difficulty with the isolations. For the detection of viral RNA, molecular techniques, such as conventional or real-time RT-PCR, have been developed . These molecular techniques are the most widely used methods for ZIKV diagnosis, particularly because of the extensive antigenic cross-reactivity between *flaviviruses* (Zanulca et Santos., 2016). Two sets of primers that target the nonstructural 5 (NS5) and envelope (E) genes of Zika virus (Table 4) and developed RT-PCR assays for the detection of Zika virus RNA (Bhatnagar et al., 2017).

Table 4. Oligonucleotide primers used for RT-PCR (Bhatnagar et al., 2017).

Primers	**Sequence 5'3'**	**Gene target**	**Product size pb**	**Annealing temperature**
Forward Reverse	AAGTACACATACCAAAACAAAGTGGT TGTTAAGAGCGTAAGTGACAAC	NS5	127	56
Forward Reverse	TGCCCAACACAAGGTGAAGC ACTGACAGCATTATCCGGTACTC	E	209	58

7. Phylogeny

ZIKV belongs to the Spondweni serocomplex, and phylogenetic analyses revealed the existence of two main virus lineages (African and Asian). The results suggest that a different ZIKV subtype of the West African circulated in the *Aedes* species in Central Africa. Molecular evolution studies indicated that ZIKV might have undergone several natural recombination events, which is an unusual feature among members of genus *Flavivirus*. A specific adaptive genetic change, the recurrent loss and gain of the N-linked glycosylation site in the E protein, was observed, and it has been suggested that this genetic alteration could be related to mosquito-cell infectivity. During the current epidemics in the Americas, a growing number of ZIKV genome sequences are being determined their phylogenetic relationship with other members of the *Flavivirus* genus revisited (Zanulca et Santos., 2016).

Several studies developed the phylogeny of ZIKV over the world, in this part examples of studies are provided with their corresponding phylogenetic trees. Haddow et al. (2012), reported that ZIKV is distributed throughout much of Africa and Asia. To elucidate the genetic relationships of geographically distinct ZIKV strains and the origin of the strains responsible for the 2007 outbreak on Yap Island and a 2010 Cambodian pediatric case of ZIKV infection, the nucleotide sequences of the open reading frame of five isolates from Cambodia, Malaysia, Nigeria, Uganda, and Senegal collected between 1947 and 2010 were determined. Phylogenetic analyses of these and previously published ZIKV sequences revealed the existence of two main virus lineages African and Asian (Figure 5) and that the strain responsible for the Yap epidemic and the Cambodian case most likely originated in Southeast Asia. Examination of the nucleotide and amino acid sequence alignments revealed the loss of a potential glycosylation site in some of the virus strains, which may correlate with the passage history of the virus.

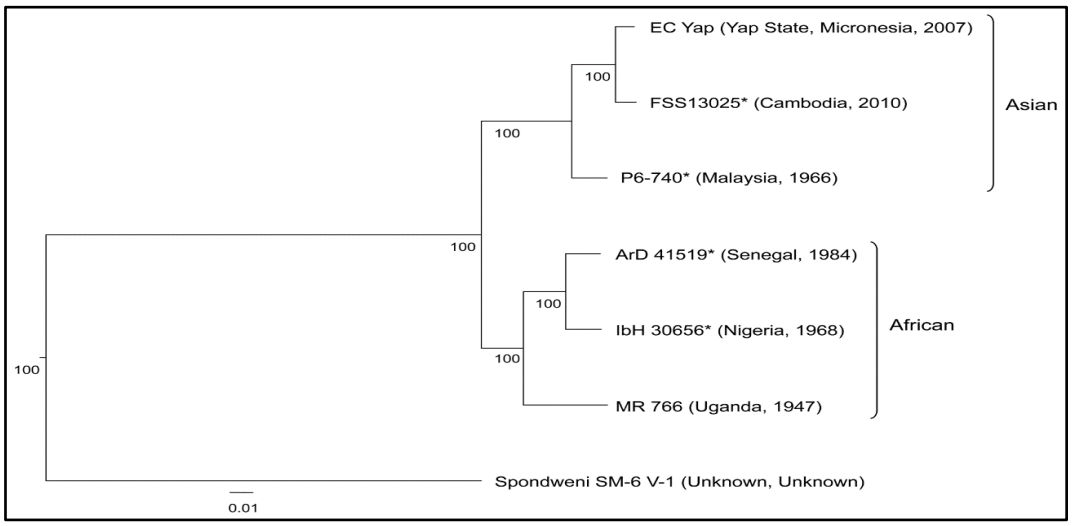


Figure 5. ZIKV nucleotide and amino acid alignments. Neighbor-joining phylogeny generated from open reading frame nucleotide sequences of Zika virus strains. The tree was rooted with Spondweni virus (GenBank accession number DQ859064). The scale at the bottom of the tree represents genetic distance in nucleotide substitutions per site. Numbers at the nodes represent percent bootstrap support values based on 1,000 replicates. Isolates are represented according to strain name, country of origin, and year of isolation. The lineage of each virus is indicated to the right of the tree (Haddow et al., 2012).

According to Faye et al. (2014), the trees for E, NS5 and the two concatenated genes (Figure 6) reinforced that ZIKV strains could be classified in three major clusters.

Accordingly, the African strains were arranged into two groups: the MR766 prototype strain cluster (yellow sector on Figure 6) and the Nigerian cluster (green sector on Figure 6); and the Micronesian and Malaysian strains clustered together forming the Asian clade (Figure 6), For West Africa, the strains from Ivory Coast and Senegal were found in both African clusters, suggesting that at least two distinct lineages of ZIKV circulated in these countries. Interestingly, we found that the position of the Senegalese cluster, comprising viruses isolated from 1998 to 2001 associated with *Aedes dalzieli*, branching as a sister group of HD78788 isolated in Senegal in 1991, was not simply explained by recombination or poor rooting of the tree, since it did not depend on the inclusion (Figure 6) or exclusion of the Spondweni, which is a bonafide outgroup.

Figure 6. Maximum likelihood phylogenetic tree inferred for concatenated of sequences from Envelope and NS5 genes of Zika virus Consensus tree summarized after 1000 non-parametric

bootstrap replicates, with support values greater than 60% shown in the nodes. The cluster the Ugandan MR766 prototype strain was highlighted by the yellow sector and the Nigerian cluster was highlighted by the green sector. The strains from Senegal and Côte d'Ivoire are shown in green and orange, respectively. The tree has been rooted with the Spondweni lineage isolated in South Africa was used as outgroup to root the tree (Faye et al., 2014).

According to Enfissi et al. (2016), on October 1, 2015, a 52-year-old man was hospitalised with exanthema and conjunctivitis at the Academic Hospital in Paramaribo, Suriname. During the next few days, four patients were admitted with mild symptoms including exanthema. Sera from these patients were negative for dengue and chikungunya viruses but positive for ZIKV by specific real-time reverse transcription PCR. Soon after, the first evidence was found of the emergence of ZIKV in the Americas, in northeast Brazil in May, 2015. 4 autochthonous circulation of ZIKV in other countries started on Oct 16, 2015, in Colombia, followed by Suriname on Nov 12, 2015.The first five autochthonous cases detected in Suriname were confirmed by the French National Reference Centre for arboviruses, located at the Pasteur Institute in French Guiana. Viral sequencing was done directly from the sera of four of these viraemic patients. Complete coding of the ZIKV sequence was obtained for one patient and envelope protein coding sequences for the three others. Few complete genomes are available for ZIKV and, until this analysis, none for ZIKV circulating in the Americas. Phylogenetic analyses were conducted for the NS5 protein coding region, the envelope protein coding region, and the complete coding region, against the sequences available in databases: all the phylogenetic trees showed the same topology. The Suriname strains belong to the Asian genotype and seem to be most closely related to the strain that was circulating in French Polynesia in 2013, with which they share more than 99.7% and 99.9% of nucleotide and amino-acid identity, respectively (Figure 7).

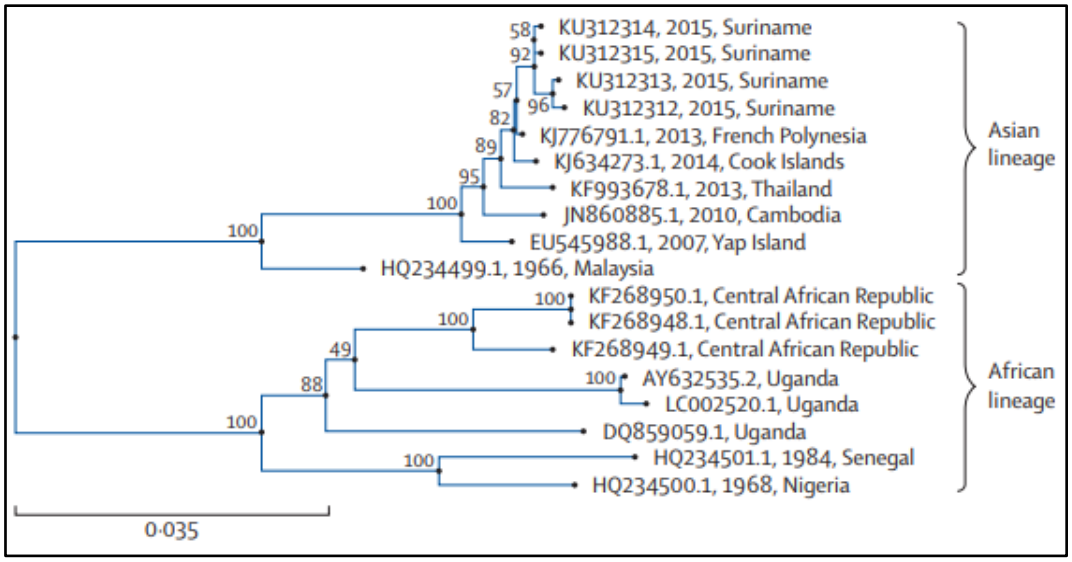


Figure 7. Phylogenetic relations between the envelope gene sequences of Suriname ZIKV and other ZIKV (Enfissi et al., 2016).

In another study conducted by Charrel et al. (2016), the phylogenetic analysis reveals the existence of two lineages the African lineage, which has shown no propensity to disseminate outside of Africa, and the Asian lineage, which continues to seed in previously unaffected regions of the world. All recently disseminated strains belong to the Asian lineage. ZIKV genomes from patients infected in Brazil and Suriname in 2015 are closely related to the strain that circulated in French Polynesia in 2013 (Figure 8), with more than 99.7% and 99.9% level of nucleotide and amino acid identities, respectively.

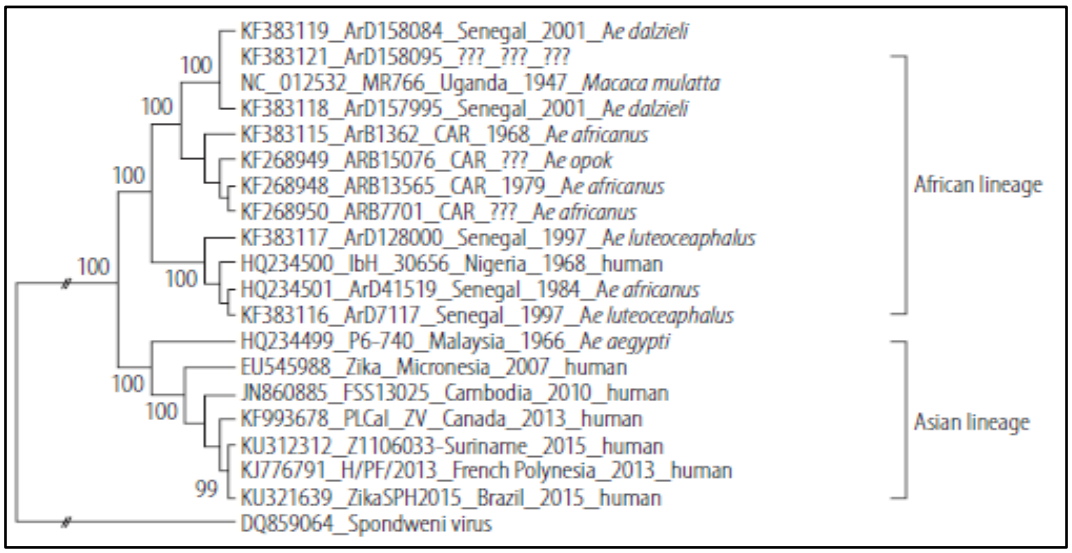


Figure 8. Phylogenetic relationships among selected Zika virus strains belonging to the African and Asian lineages based on complete genomic sequence (Maximum Likelihood analysis) (Charrel et al., 2016).

Chapter II

Molecular Biology

Chapter II: Molecular Biology

1. The structure of the virion

The Zika virions are 40–60 nm in diameter, spherical in shape and contain lipid envelope (Gurumayum et al., 2018). The structure of Zika virus is similar to other known flavivirus structures, except for the ~10 amino acids that surround the Asn154 glycosylation site (Sirohi et al., 2016).

Flaviviruses are enveloped viruses complexed with multiple copies of the capsid protein, surrounded by an icosahedral shell consisting of 180 copies each of the envelope (E) glycoprotein and the membrane (M) protein or precursor membrane(prM) protein, all anchored in a lipid membrane (Figure 9). During their life cycle, flavivirus virions exist in three major states— immature, mature, and fusogenic—which are non-infectious, infectious, and host membrane–binding states, respectively. The mature ZIKV structure is similar to mature DENV and WNV structures (Sirohi et al., 2016).

ZIKV particles may expand into smooth surfaced particles when incubated at higher temperatures, making the lipid envelope more fluid, and allowing the structure to revert to its normal state (Kostyuchenko et al., 2016).

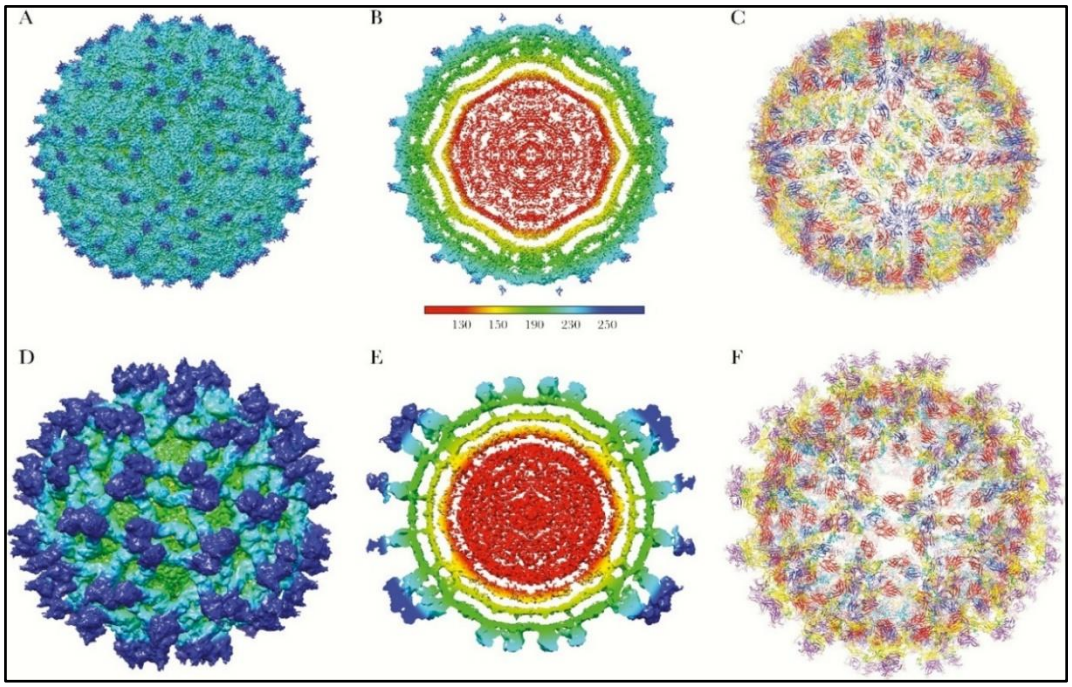


Figure 9. Structure of Zika virus (ZIKV). A- C depict the mature form of ZIKV (3.8 Å resolution), whereas D-F represent the immature form (9.1 Å). C displays the icosahedral

arrangement and Cα-backbone of the E and M proteins derived from the 3.8 Å density map of mature ZIKV (Protein Data Bank [PDB] 5IRE). Two asymmetric units related by 180° define the raft subunit of the virus consisting of 3 pairs of E and M homodimers. F shows the Cα-backbone of the DENV2 prM-E heterodimer (PDB 3C6E) and transmembrane domains of ZIKV E and M proteins (PDB 5IRE) fitted into the immature ZIKV map (PDB 5U4W). E protein is colored as follows: domain I (red), II (yellow) with fusion loop in green, III (blue) and stem-transmembrane helices (pink). pr peptide is shown in purple. The soluble region of M protein is displayed in magenta, and the stem-transmembrane helices are represented in cyan. The glycans projecting from the surface on prM and E proteins are highlighted (Sirohi et Kuhn., 2017).

2. The viral genome

ZIKV contains a non-segmented, linear, single-stranded positive-sense ribonucleic acid (RNA) genome, typically 10,807 nucleotides in length (Figure 10.A), although some differences in length have been reported among different isolates and even among the prototype MR-766 strains with different passage histories. The genomic RNA has a type I cap structure (m7GpppAmG) at its 5' end, followed by a 5' non-coding region (NCR) of 106-107 nucleotides, a single open reading frame (ORF) of 10,272 nucleotides, and a 3' NCR of 428-429 nucleotides, with no poly-A tail at the 3' end (Figure 10.A) (Yun et Lee., 2017).

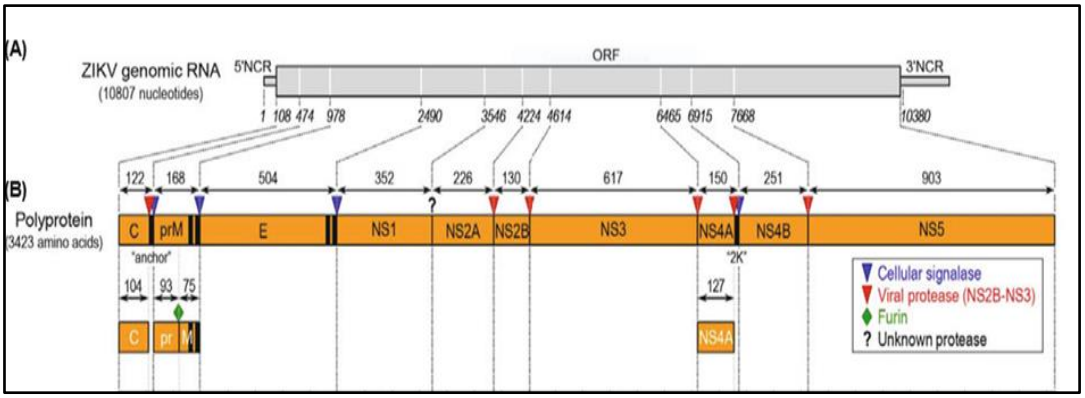


Figure 10. Genome structure, polyprotein processing of ZIKV. (A) Genome structure. The positive-sense genomic RNA of ZIKV is composed of a 5'NCR, a single long ORF, and a 3'NCR. (B) Polyprotein processing. The viral ORF encodes a 3,423-amino-acid polyprotein, which is co- or post-translationally processed by host- and virus-encoded proteases (as indicated) into three structural (C, prM, and E) and at least seven nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) proteins. Vertical black bars represent one or two

transmembrane domains located at the junctions of C/prM (designated “anchor”), prM/E, E/NS1, and NS4A/NS4B (designated “2K”). Also indicated are the putative cleavage sites conserved among flaviviruses and the lengths of the cleavage products (yun et Lee., 2017).

The 3′ untranslated regions (UTR) is further divided into three domains, including the highly variable proximal domain 1 that directly follows the stop codon, the moderately conserved domain 2 that contains the stem-loop (SL) and dumbbell (DB) structures, and the highly conserved domain 3 that contains the complementary cyclization elements and the conserved sHP-3′ SL structure (Figure 11). Deletion of the SL sequences in the 5′- or 3′-UTR is lethal for flavivirus infectious clones.

A Y-shape stem-loop A (SLA) structure is found at the 5′-end of the ZIKV genome. At the 3′-end of the viral genome, a small hairpin 3′-stem-loop (sHP-3′ SL) structure, three additional SL structures, and a dumbbell (DB) structure are found (Figure 11). Notably, the external loop of the SLI in domain 1 of the 3′-UTR just distal to the stop codon of the NS5 in the 1947 prototype pre-epidemic strain is replaced by a large bulge of nine nucleotide bases (UAG UCA GCC) in the representative epidemic ZIKV strain. Short conserved sequences within the 3′ terminal SL structure include the terminal 5′-CU-3′ and a 5′-ACAG-3′ in the top loop of the sHP-3′ SL structure. There are three pairs of inverted complementary sequences (GAU CUG UG-CAC AGA UC, UGG AUU U-AAA UCC A and GAG UUU CUG GUC-GAC CAG AGA CUC and GAG UUU CUG GUC-GAC CAG AGA CUC that may mediate genome cyclization (Zhu et al., 2016).

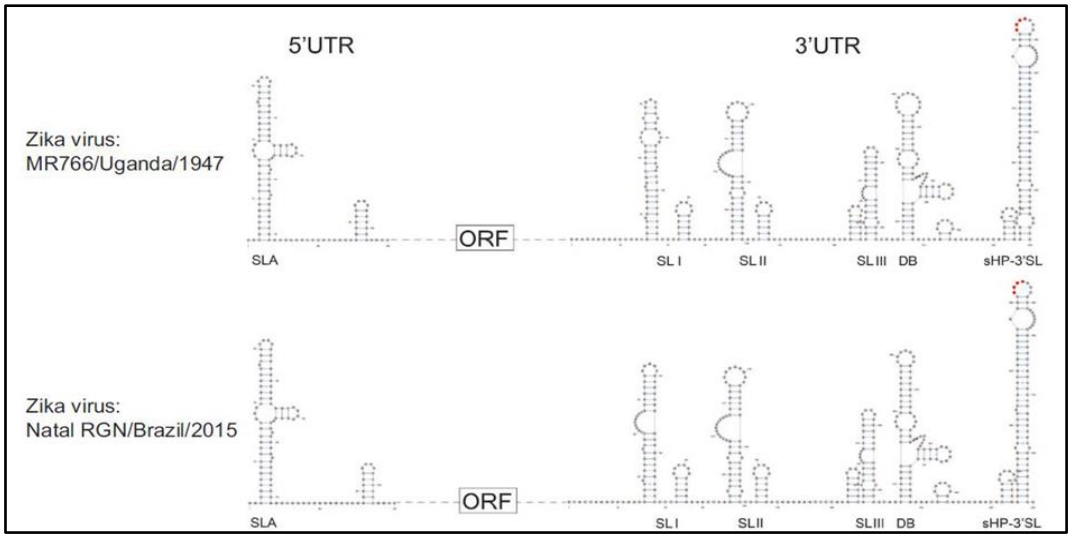


Figure 11. Schematic representations of the Zika virus genome RNA secondary structures. The short conserved 5′-ACAG-3′ sequences in the top loop of the sHP-3′ SL structure are indicated in red (Zhu et al., 2016).

3. The viral proteins

The RNA is translated into a single polyprotein (3423 amino acids in length) (Figure 10.B). This polyprotein is processed by host and viral proteases into three structural and seven non-structural proteins (Chambers et al., 2018). These structural proteins consist of capsid (C), pre-membrane (prM), and envelope (Env) proteins, which are predominantly involved in viral pathogenesis and virion structure. The seven non-structural proteins, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 proteins, largely contribute towards the purposes of viral pathogenesis, replication, and immune evasion (Mohd Ropidi et al., 2020). The coding region orders and NS protein motifs of ZIKV are arranged in the order of 5′-Capsid (C)- pre-Membrane (prM)-Envelope (E)-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-3′ (Figure 12.b). The complete polyprotein sequences of ZIKV have low similarity with those of other human-pathogenic flaviviruses (DENV-2, 58.1% to 58.9%; SPOV, 68.3% to 69.0% nucleotide similarity) (Zhu et al., 2016).

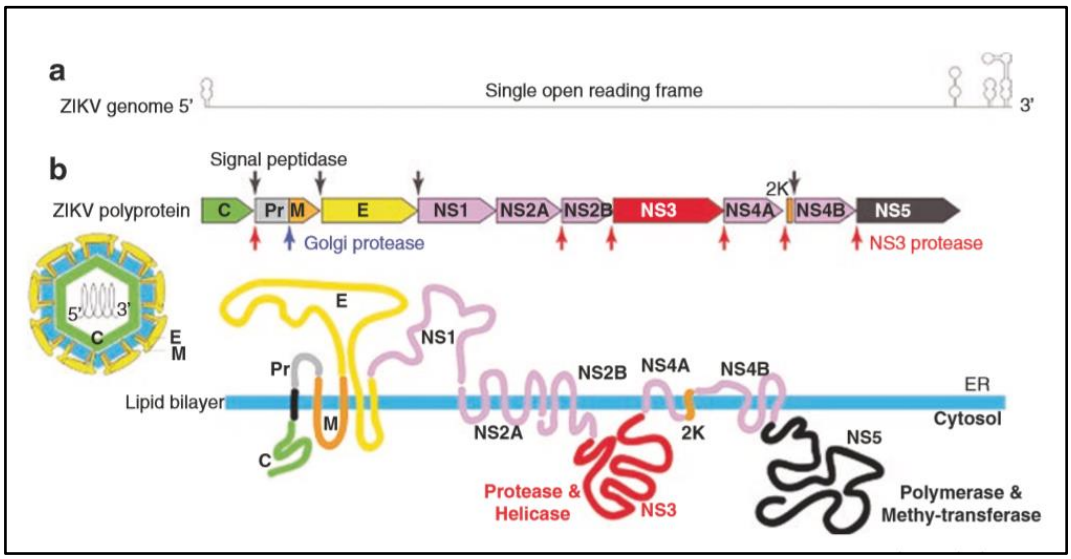


Figure 12. Zika genome. (a) A diagram of ZIKV genomic RNA. (b) Processing strategy and protein products. The polyprotein is processed at various sites by host (red arrow heads) and viral (down black arrows) proteases (Amorim L., 2019).

3.1. Structural proteins

3.1.1. Capsid (C) protein

Capsid (C) protein is made up of 122aa (Table 5). It is present in the cytoplasm of the infected cells (Figure 12.b, 13) while form a nucleocapsid complex with RNA in viral

particle. All capsid proteins of the virion have positive net charge and are similar in size but with minor sequence conservation. This protein is released into the cytosol and assembles homodimers after polyprotein cleavage. One side domain of the capsid protein contains the basic residues that bind with the RNA genome while other side domain contains hydrophobic residues that cooperate with the lipid envelope of the virus. After endosomal membrane fusion of the virus, viral genome entering remains related with the C dimers to evade RNA sensors and nucleases from the host. In addition to role in the synthesis of viral nucleocapsid, C protein functions as RNA chaperone. Thus, resulting nucleocapsid buds formation in the lumen of endoplasmic reticulum to make viral particles with E and prM proteins (Javed et al., 2017).

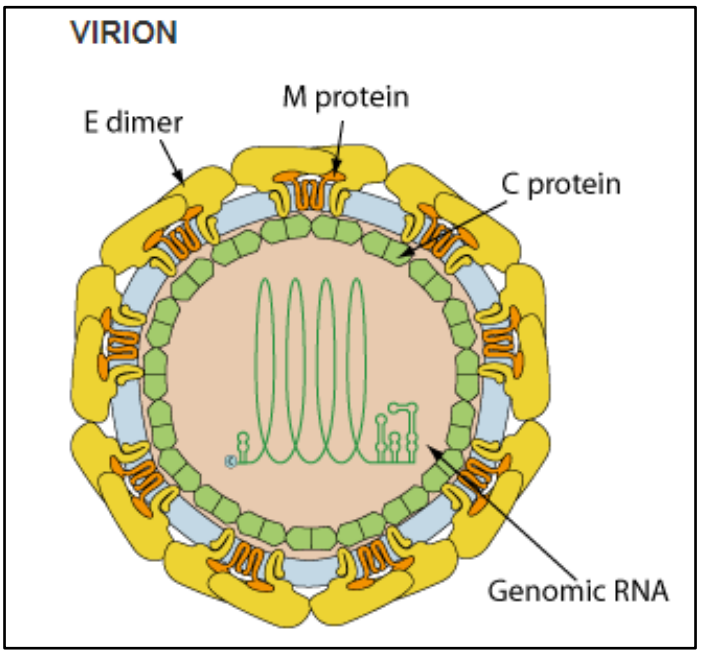


Figure 13. Structural proteins of Zika virus (Lin et al., 2018).

3.1.2. Pre membrane (PrM) protein

Pre membrane (PrM) protein is made up of 178aa (Table 5), which is buried under the layer of E-protein. The M and E proteins are arranged in icosahedral symmetry consisting of repeating 60 units, and each of asymmetrical unit comprise of three individual E-proteins (Javed et al., 2017). In the Golgi apparatus, PrM is cleaved by furin-like protease to produce mature M protein and Pr protein product (Figure 12.b). After maturation, PrM and E are released and 90 E protein homodimers rearrange in a herringbone-like array forming mature Zika virus (Almansour et al., 2019).

3.1.3. Envelope (E) protein

The E protein is the major component involved in receptor binding, membrane fusion, and host immune recognition (Shi et Gao., 2017). ZIKV E proteins have a characteristic "herringbone" structure with a single glycosylation site at residue Asp154 (Lin et al., 2018). The flavivirus E proteins belong to the class-II fusion proteins, which lie flat on the virus surface in the form of antiparallel homodimers (Shi et al., 2018). Each E protein monomer consists of three domains: DI, DII, and DIII13 (Figure 14), which undergo major rearrangements during the virus maturation cycle (Almansour et al., 2019).

DI is a central beta-barrel domain; DII is a finger- like dimerization domain; and DIII is an immunoglobulin-like domain. DI, which connects DII to DIII, is essential for the conformational changes required for viral entry into cells (Almansour et al., 2019). The central domain I contains around 130 residues in three segments, residues 1–51, 132–192, and 280–295 (Figure 14.A). The central domain I is folded into an eight-stranded β-barrel with an additional N-terminal A0 strand (Figure 14.B). The 150-loop (residues 147–161, between strands E_0 and F_0), which contains the potential N154 glycosylation site, likely represents a highly flexible loop in this domain due to a lack of electron density in this region (Dai et al., 2016).

DII contains a fusion loop (FL) that interacts with the endosomal membrane, whereas DIII contains the receptor-binding site and is thus essential for attachment of virus particles to the host cell (Almansour et al., 2019). The finger-like domain II is formed by two segments, residues 52–131 and residues 193–279 (Figure 14) (Dai et al., 2016).

DIII also plays an essential role in mediating the fusion of virus particles with the endosomal membrane after endocytosis (Almansour et al., 2019). The C-terminal domain III (residues 296–403) displays an IgG-like fold where the AA'BE sheet and disordered D strand are contacted by the cd loop of the adjacent E monomer (Dai et al., 2016).

Domain II is responsible for the dimerization of E monomer, leading to an extended, but interrupted, dimer interface, and thus there are ''holes'' in the dimer at either side of domain II (Figure 14.B). The central dimer interface is mainly constituted by the aB helix and j strand elements of each sE monomer, whereas the distal dimer interface is mainly created by the hydrophobic interaction between the cd loop and the crevice formed by domains I and III of

the adjacents E monomer (Figure 14.B).The hydrophobic cd loop represents the fusion loop (residues 98–109), which is responsible for the membrane fusion between host cell and virus membranes during virus entry, and is highly conserved in flaviviruses (Dai et al., 2016).

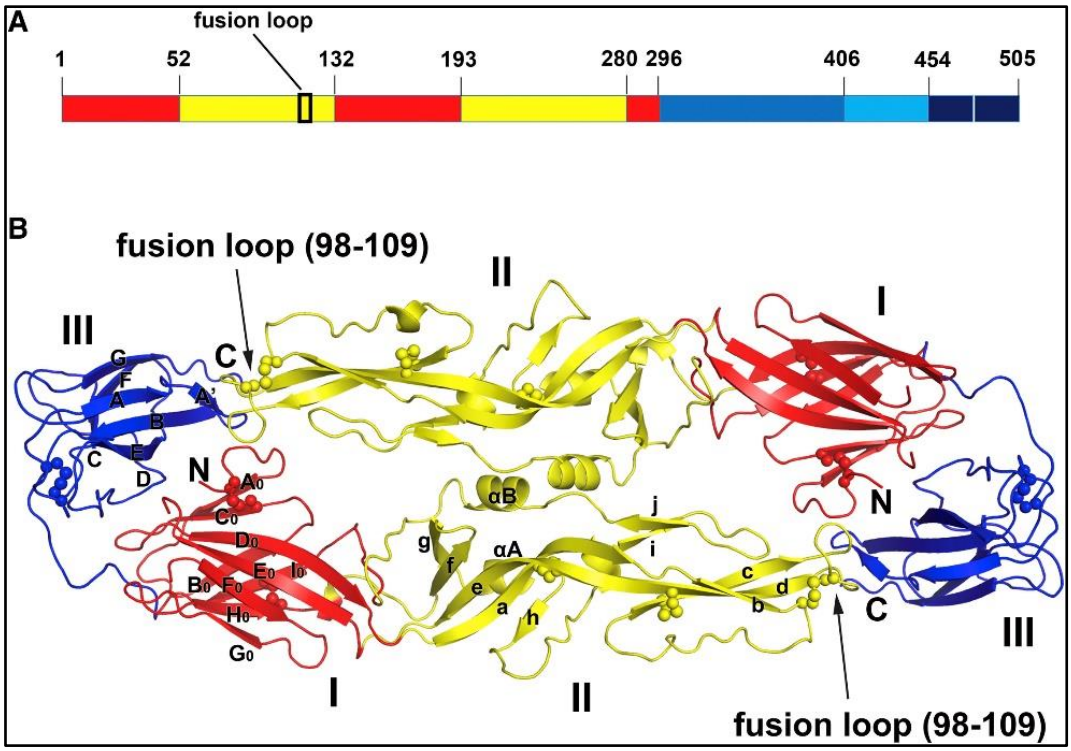


Figure 14. Overall Structure of the ZIKV-E Protein. A: Schematic diagram of domain organization for ZIKV-E. Domain I (red), domain II (yellow), and domain III (blue). A 48-residue stem segment links the stably folded ZIKV-E ectodomain with the C-terminal transmembrane anchor. B: Dimer structure of ZIKV-E. ZIKV-E has three distinct domains: b-barrel-shaped domain I, finger-like domain II, and immunoglobulin-like domain III. Domain II is responsible for the dimerization of ZIKV-E. The fusion loop is buried by the domains I and II of the other ZIKV-E monomer (Dai et al., 2016).

3.2. Non-structural proteins

3.2.1. NS1

The non-structural protein NS1 is a 46 KDa glycoprotein containing 2–3 glycosylation sites and 12 conserved cysteine residues that can form disulphide bonds. Mutations in glycosylation sites Asn130 and Asn207 drastically affect virus replication and virus production. Although the NS1protein has no hydrophobic transmembrane domain, it associates with the membrane through a glycosylphosphatidylinositol (GPI) anchor. It is

located inside the cell. Upon proteolytic separation from the envelope protein, it is secreted out of the cell. Cell surface expression of NS1 could elicit a strong humoral immune response, which further aids in antibody-mediated killing of virus infected cells. The NS1 and NS4 non-structural proteins interact with each other and co-localize in the viral replicase complex to help the viral genome replication (Routhu et Byrareddy., 2017).

The ZIKV NS1protein crystallized as a rod-like homodimer with a length of ~9 nm (Figure 15.a). Sedimentation velocity analytical ultracentrifugation analyses confirmed that the ZIKV NS1 protein exists as a homodimer in solution (Supplementary Figure 15).the ZIKV NS1 homodimer structure has a continuous β-sheet on one surface, with 20 β-strands arranged like the rungs of a ladder (Figure 15.a), in which each monomer contributes ten rungs to the antiparallel β-ladder. On the opposite side of the homodimer, an irregular surface is formed by a complex arrangement of loop structures (Figure 15.a). Most of those interstrand loops are short, except for a long 'spaghetti loop' between β4 and β5 (Figure 15.a, b). A potential N-linked glycosylation site that is highly conserved in the Flaviviridae family is located in the β3–β4 loop (Song et al., 2016).

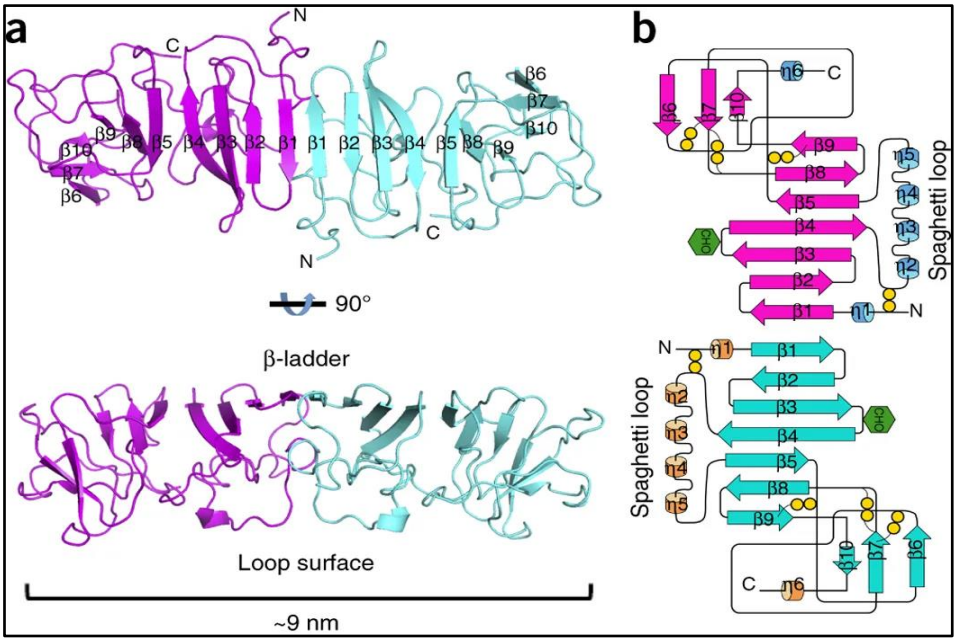


Figure 15. Overall structure of protein NS1. (a) ZIKV NS1 forms a head-to-head dimer with one extended β-ladder platform and one loop surface on the opposite side. (b) Topology diagram for NS1. Glycosylation sites are indicated with green hexagons, and disulfide bonds are indicated with yellow circles. η represents the 3_{10} helix, and β represents the β-sheet (Song et al., 2016).

3.2.2. NS2A

The NS2A is a 22 kDa hydrophobic protein and plays an important role in infectious viral particle production. Cleavage of NS2A upstream of the C-terminus by viral serine protease produces a smaller 20 kDa protein, which is essential in viral replication. The NS2A non-structural protein interacts with the NS3, NS5, and co-localizes with the replication complex. These interactions are essential in managing the shift between RNA packaging and RNA replication in virus particle production (Routhu et Byrareddy., 2017).

3.2.3. NS2B

The NS2B is a 14 kDa membrane-associated protein that mediates membrane interaction by its conserved central hydrophobic region. These interactions increase the host cell membrane permeability (Routhu et Byrareddy., 2017). This protein is made up of l30aa (Table 5). Two or more membrane spanning regions are present in this protein and play central role in the association and complexes of virus replication on the endoplasmic reticular membrane. NS2B collaborates with the C-terminal protease domain of NS3 to make the complex of serine protease that is involved in the viral polyprotein cleavage (Javed et al., 2017).

3.2.4. NS3

NS3 is a multifunctional 70 kDa protein that has multiple enzymatic activities, such as trypsin-like serine protease activity, Mg2+ inhibited and poly(A)-stimulated NTPase activity. In addition, it also has Mg2+ stimulated, poly(A)-inhibited RNA triphosphatase (RTPase) activity and ATP-driven RNA duplex unwinding activity at its C-terminus. Therefore, mutations in the putative NTP-binding site (Lys199) destroy the NTPase and RTPase activities. The NS3 localizes to the membranes by interacting with the membrane-localizing transmembrane viral protein NS2B and helps in viral polyprotein processing and viral replication. The trypsin-like serine protease activity has been mapped to His53, Asp77, and Ser138 (catalytic triad) (Routhu et Byrareddy., 2017).

3.2.5. NS4A

NS4A is a 16 kDa multifunctional, hydrophobic protein. It regulates membrane proliferation, localizes the replication complex to the membrane, aids in polyprotein

processing, and is crucial for immune evasion. The NS4A C-terminus interacts and translocates NS4B to the ER lumen (Routhu et Byrareddy., 2017).

3.2.6. NS4B

The NS4B is a 27 KDa transmembrane protein (Figure 12.b) (Routhu et Byrareddy., 2017) about 252aa (Table 5) long and poorly conserved protein. Multiple potential membrane spanning hydrophobic regions is present in NS4B. Membrane constituents of the viral replication complex are formed by NS4B and is involved in localization in membranes of NS3 protein (Javed et al., 2017).

3.2.7. NS5

NS5 is a highly conserved 103 kDa protein and plays a crucial role in viral replication. It has N-terminal RNA capprocessing activity (homology with the S-adenosylmethionine (SAM)-dependent methyltransferases), GTP-binding activity. The interaction of NS5 protein stimulates NS3 NTPase activity. As a result, this helps in nuclear localization (Routhu et Byrareddy., 2017).

Table 5. Structural and non-structural proteins of ZIKA virus (Javed et al., 2017).

Proteins name	Symbol	Size (aa)	Location in cell	Function in cell
Structural				
Capsid	C	122	Cytoplasm	Viral nucleocapsid formation
Pre membrane	prM	178	Cytoplasm	Viral capsid formation, host cell fusion, and stabilization
Envelope	E	500	Cytoplasm	Host receptor binding, fusion, and entry
Non-Structural				
NS1 protein	NS1	384	Cytoplasm	Emission, virulence, and replication
NS2 protein	NS2A	226	Cytoplasm	Viral transcription and assembly
NS2 protein	NS2B	130	Cytoplasm	NS3 cofactor for serine protease function, polyprotein cleavage

NS3 protein	NS3	617	Cytoplasm	Unwinding of structured protein template region and processing of viral polyprotein via serine protease, helicase and triphosphatase activity
NS4 protein	NS4A	127	Cytoplasm	Viral replication
NS4 protein	NS4B	252	Cytoplasm	Viral replication complex
NS5 protein	NS5	902	Cytoplasm	RNA replication via RNA dependent RNA polymerase and RNA capping

4. Replication cycle

Viral structural glycoproteins, in particular envelope E glycoprotein, mediate binding to cellular receptors, thereby triggering endocytotic pathways. These interactions between cellular receptors and glycoproteins allow ZIKV to infect specific cellular types. The uptake of viral particles occurs primarily through clathrin-dependent endocytosis. Surface glycoproteins of internalized viral particles undergo conformational changes due to the endosomal lumen's acidic environment, which promote viral envelope fusion with the endosomal membrane. This completes the entry process, which implies the delivery of viral RNA into the cytoplasm of the host cell. The positive-sense RNA is translated into a polyprotein, which is subsequently cleaved to release structural and NS proteins (Esteves et al., 2017). Cellular compartments such as the endoplasmic reticulum (ER) and the Golgi apparatus, seem to be crucial for viral replication and propagation. First, ER membranes give rise to the vesicles involved in autophagic flux, a cellular mechanism exploited and manipulated by *Flaviviruses* in order to enhance their own replication and initiate infection (Carneiro et al., 2016). Second, immature viral particles are assembled within the ER and virions traffic through the Golgi network for particle maturation prior to the release from the infected cell. Mature particles are then delivered into the extracellular environment where they are ready to commence a new infectious lifecycle (Figure 16) (Esteves et al., 2017).

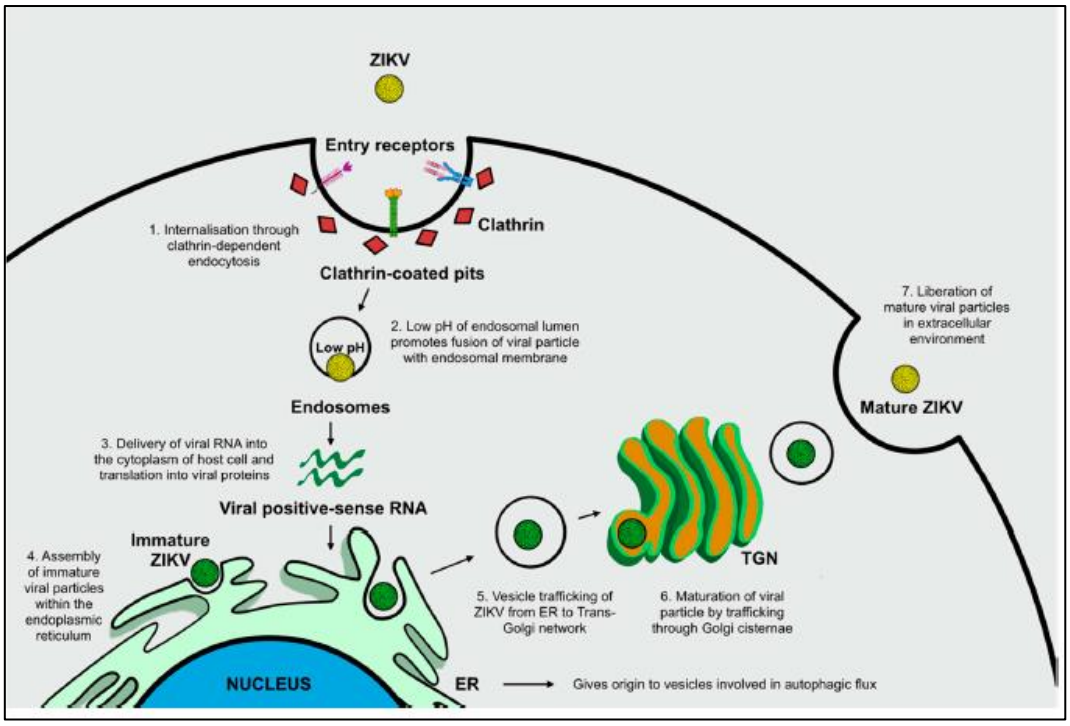


Figure 16. Replication cycle of ZIKV on the host cell (Gratton et al., 2019).

5. Zika virus 3D protein envelope structure

5.1. 3D structure by microscopy

Barba-Spaith et al. (2016), have three-dimensional cryo-electron microscopy (cryo-EM) structures of the mature ZIKV particles have recently been reported to near atomic resolution (3.8 Å), showing that the virus has essentially the same organization as the other flaviviruses of known structure, such as dengue virus and West Nile virus. The E protein is about 500 amino acids long, with the 400 N-terminal residues forming the ectodomain essentially folded as β-sheets with three domains, named I, II and III, aligned in a row with domain I at the centre (Figure 17). The conserved fusion loop is at the distal end of the rod in domain II, buried at the E dimer interface. At the C terminus, the E ectodomain is followed by the stem, featuring two α-helices lying flat on the viral membrane (the stem helices), which link to two C-terminal transmembrane α-helices. The main distinguishing feature of the ZIKV virion is an insertion within a glycosylated loop of E (the '150' loop), which protrudes from the mature virion surface.

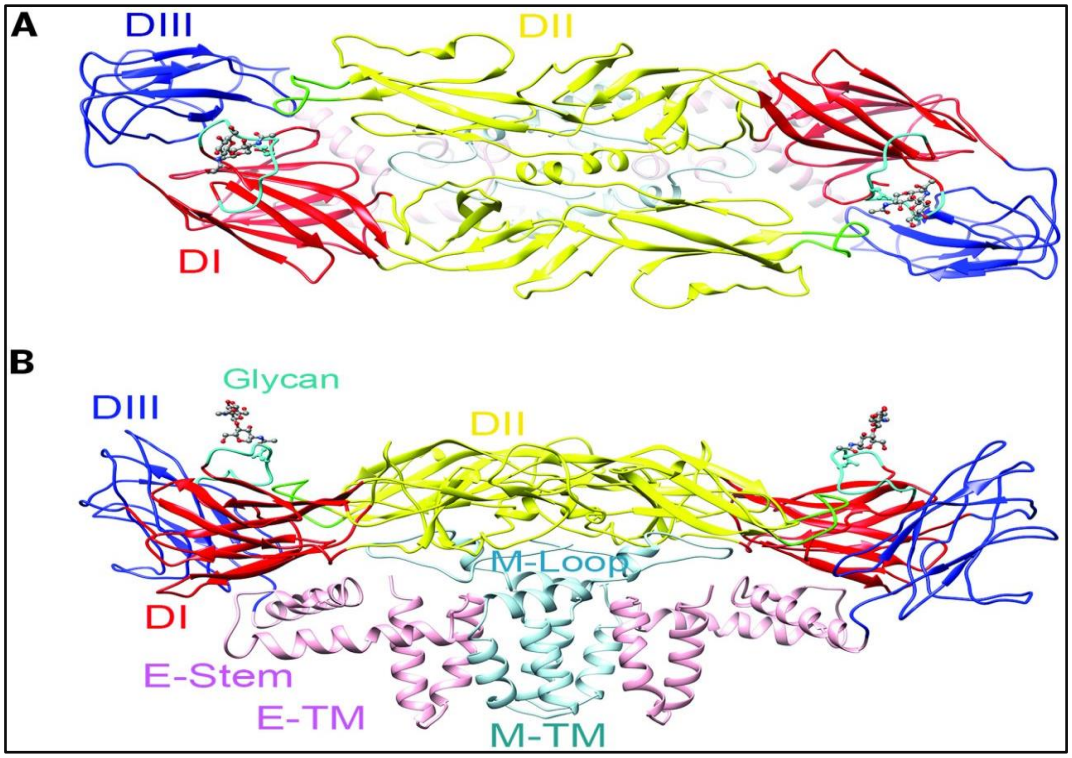


Figure 17. The structures of the ZIKV E and M proteins. (A) The E protein dimer is shown in ribbon form, viewed down the twofold axis. The color code follows the standard designation of E protein domains I (red), II (yellow) and III (blue). The underlying stem and transmembrane residues are shown (light pink). The fusion loop (green), the Asn^{154} glycan (ball-and-stick representation), and the variable loop surrounding the Asn^{154} glycan (cyan; residues 145 to 160) are shown. (B) Side view of the (E-M), showing the three E ectodomains, as well as the E stem-transmembrane domains (pink) and the M loop and stem-transmembrane domains (light blue; TM, transmembrane) (Sirohi et al., 2016).

In other study Fontes-Garfias et al. (2017), the flavivirus E protein is a major surface glycoprotein involved in modulating the viral infection cycle and eliciting antibody response. The E protein of most flaviviruses is posttranslationally modified by N-linked glycosylation at amino acid 153/154 within a highly conserved glycosylation motif of N-X-T/S at positions154-156 (Figure 18), indicating the biological importance of this modification; however, some flaviviruses isolates lack E glycosylation, suggesting that the function of E can be achieved without the N-linked glycan.

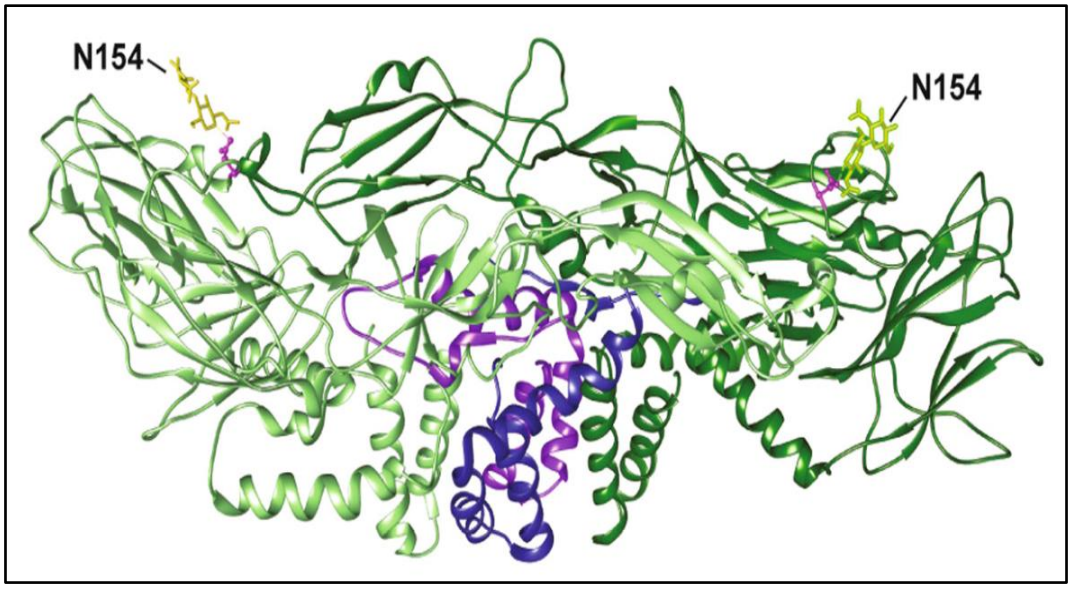


Figure 18. Three-dimensional (3D) dimer structure of ZIKV envelope protein. E protein sdimer is shown in ribbon form; E protein monomers are colored in light and dark green, and the transmembrane regions are colored in blue and purple. The N154 glycans on each monomer are labeled and shown projecting on the E protein surface (PDB: 5IRE) (Fontes Garfias et al., 2017).

5.2. 3D structure by modeling

Ekins et al. (2016) conducted a qualitative analysis of the Zika virion (which was constructed based on the dengue virion) can be compared to the dengue cryo-EM virion (Figure 19) and indicates that Zika appears to have slightly more raised ‘pimples’ on the surface. The glycoprotein E dimer in ZIKV also has a narrow ‘letter-box’ groove while the dengue virion has a bigger ‘pore‘ between the intersection of 5 dimers (5 fold axis). These differences are considerably more apparent in the animation. It is important to note that the differences may also be artifacts of the homology modeling approach and template used for modeling ZIKV glycoprotein E.

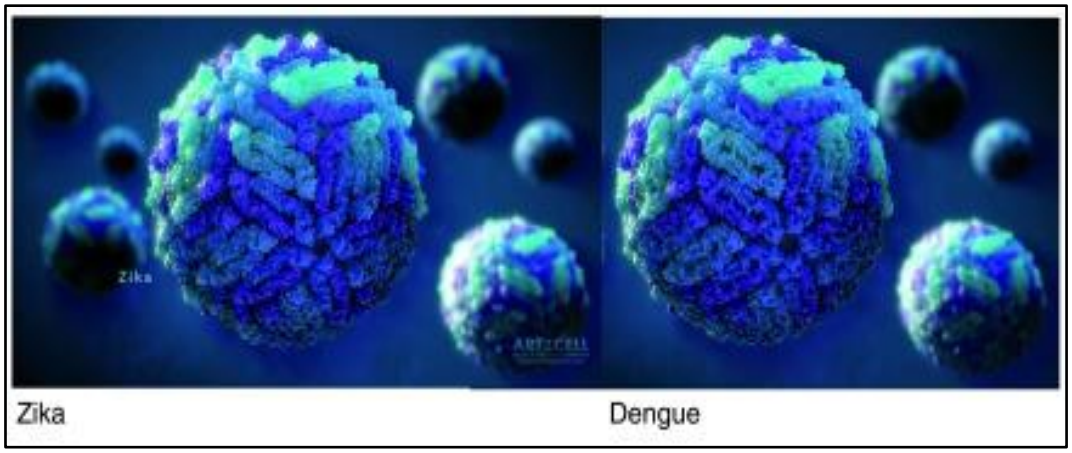


Figure 19. Comparison of Zika and Dengue virion illustrations (Ekins et al., 2016).

The homology models developed using two different templates namely the immature protein which was based on the dengue crystal structure 4gsx as a template and the mature protein which was based on PDB ID: 3P54 from Japanese encephalitis virus showed a large difference (RMSD 13.47Å) (Figure 20). These proteins also demonstrate differences around the pocket used centered on the residues 270-277.

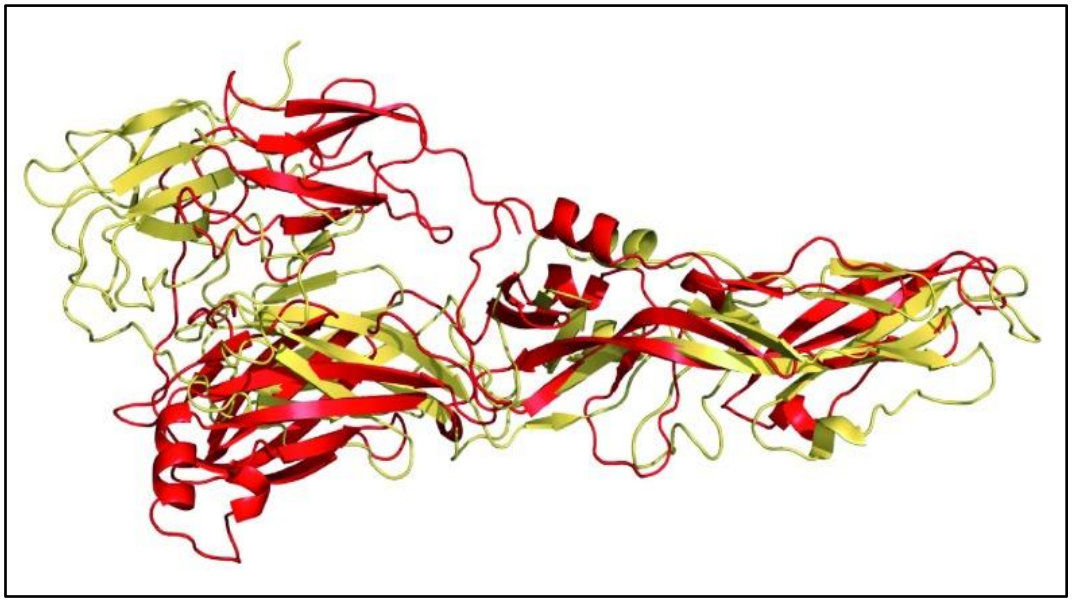


Figure 20. Overlap of ZIKV homology models for glycoprotein E, Yellow = mature conformation (this study) compared with the immature conformation (red) (Ekins et al., 2016).

A model of the Zika virion was constructed as an illustration using the homology model of the glycoprotein E dimer (Figure 20). While the combined protein sequence of glycoprotein E

and the immunoglobulin like domain is closest to dengue virus 1 (57 percent identity) the closest template was for the crystal structure of the Japanese encephalitis virus envelope protein (53.12 percent identity). This would suggest the virion should more closely resemble that of dengue virus 1, while producing a homology model based on a more distant virus might not be ideal. The homology model of glycoprotein E developed for the mature conformation in this study is significantly different from that developed previously for the immature conformation (Figure 18). The proposed binding site centered around residues 270-277 appears shallower in the mature conformation and this would certainly affect the kinds of molecules that it could interact with. It might also point to the need to interfere with the immature conformation as preferable versus the mature conformation (Ekins et al., 2016).

In other study reported by Alam et al. (2016) the secondary structure of protein E from its amino acid sequence describes the α-helix, β-sheets and random coil. Our ZIKV envelope glycoprotein is 504 residues long, of which 180 residues (35,7%) form sheet, 61 residues form turn and 305 residues (60,5%) form helix regions of the protein (Figure 21). In the secondary structure, 35,7% region of the target protein remains as β-sheet. In several experiments, it was shown that the antigenic part of the protein was more likely to belong to the β-sheet region. For the docking analysis, the 3D structures of the selected peptides were designed using the PEP-FOLD Peptide Structure Prediction server that searches for known 3D protein structures from PDB that are homologous to the epitope source sequence. Keeping this in mind we carried out homology modelling of full ZIKV glycoprotein instead of peptide. The 3D structure built by the I-TASSER used the top three templates 3J65, 3J27 and 4CCT of 10 PDB templates. The average confidence score of the predicted model was 1,50; the *z*-scores were 2,96, 4,11 and 3,55 for the top three templates, respectively; and the average RMSD of our predicted structure was 0,481 from the top three templates. The Ramachandran plot for our model showing residues in the allowed region was more 85% (Figure 22). Verified 3D predicted that 87,50% of the residues had an average 3D score of 0,2 for the best predicted model and at least 80% of the amino acids should have scored ≥ 0,2 in the 3D/1D profile. We superimposed predicted peptides with the our modelled 3D structure and found that mostly epitopes were in the accessible area of proteins, which means that interactions occur easily.

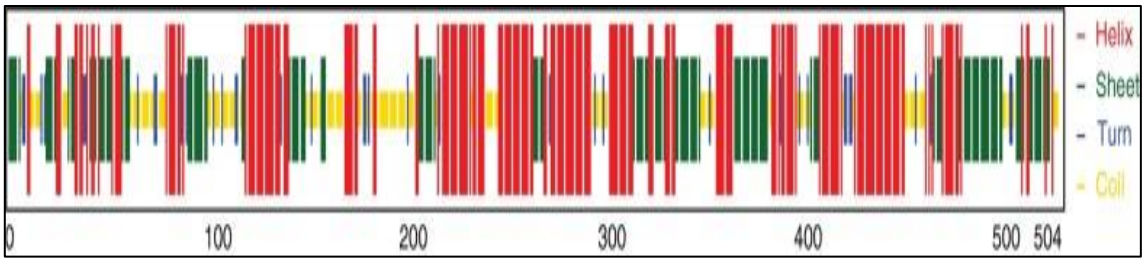


Figure 21. Represents the composition of secondary structure from amino acid residues of ZIKV envelope glycoprotein. Only 35,7% residues form sheet, 60,5% form helices and 3,8% residues form the turn region using an online server cfssp (Chou & Fasman Secondary Structure Prediction) (Alam et al., 2016).

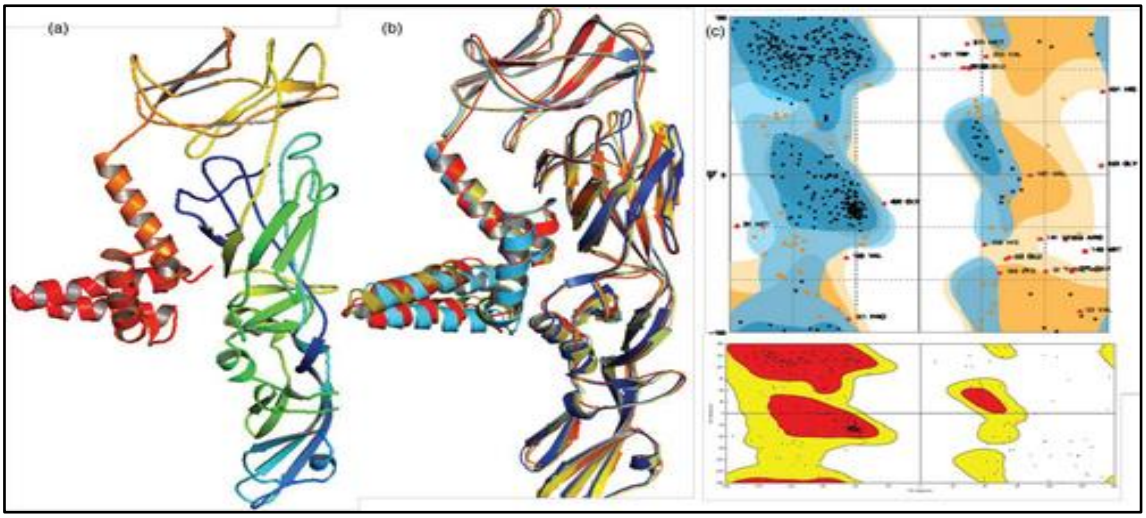


Figure 22. (a) Modelled three-dimenasional structure of ZIKV envelope glycoprotein; (b) Top three templates superimposed with our modelled protein with an average of RMSD = 0,481; (c) Ramachandran plot of our modelled protein showing the residues in allowed the region (Alam et al., 2016).

Chapter III

Epidemiology

Chapter III: Epidemiology

1. Transmission of ZIKV

1.1. Vectoriel transmission

In Africa, ZIKV exists in a sylvatic transmission cycle involving nonhuman primates and forest-dwelling species of *Aedes* mosquitoes (Figure 23). In Asia, a sylvatic transmission cycle has not yet been identified. Several mosquito species, primarily belonging to the stegomyia and diceromyia subgenera of Aedes, including *A. africanus*, *A. luteocephalus*, *A. furcifer*, and *A. taylori*, are likely enzootic vectors in Africa and Asia. In urban and suburban environments, ZIKV is transmitted in a human–mosquito–human transmission cycle (Figure 23). Two species in the stegomyia subgenus of *Aedes A. aegypti* and, to a lesser extent, *A. albopictus* have been linked with nearly all known ZIKV outbreaks, although two other species, *A. hensilli* and *A. polynesiensis*, were thought to be vectors in the Yap and French Polynesia outbreaks, respectively. *A. aegypti* and *A. albopictus* are the only known aedes (stegomyia) species in the Americas. Despite the association of *A. aegypti* and *A. albopictus* with outbreaks, both were found to have unexpectedly low but similar vector competence for the Asian genotype ZIKV strain, as determined by a low proportion of infected mosquitoes with infectious saliva after ingestion of an infected blood meal. However, *A. aegypti* is thought to have high vectorial capacity because it feeds primarily on humans, often bites multiple humans in a single blood meal, has an almost imperceptible bite, and lives in close association with human habitation. Both *A. aegypti* and *A. albopictus* bite primarily during the daytime and are widely distributed throughout the tropical and subtropical world. *A. albopictus* can exist in more temperate areas, than *A. aegypti*, thus extending the potential range where outbreaks may occur (Petersen et al., 2016).

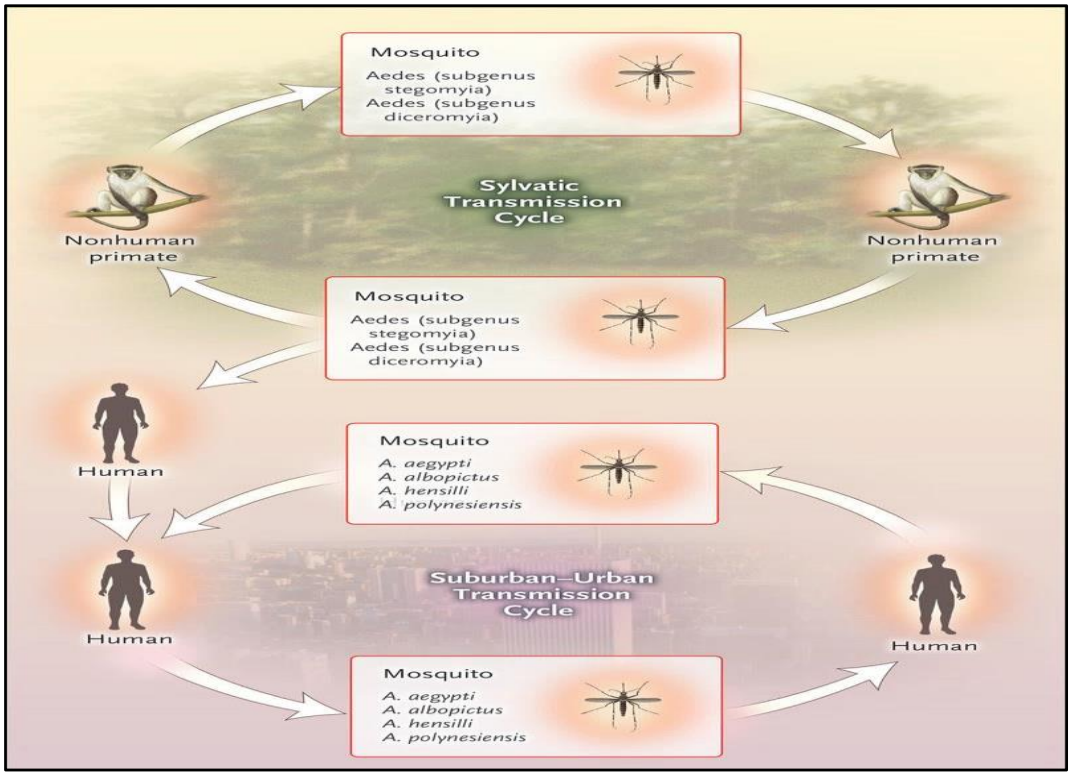


Figure 23. Vectoriel transmission (Petersen et al., 2016).

1.2. Non vectoriel transmission

ZIKV is transmitted to humans primarily through the bite of infected mosquitoes, but it can also be passed from mother to child during pregnancy or spread through sexual contact, breastfeeding, or blood transfusion. The multiple modes of ZIKV transmission make it difficult to develop control strategies against the pathogen **(**Song et al., 2017).

1.2.1. Sexual transmission

ZIKV can be sexually transmitted from an infected person to his or her partners, as shown by the detection of the virus in patient's semen, often in high titer, and transmission of the virus between both sexes, with the male-to-female transmission occurring more frequently than female-to-male or male-to male transmission (Song et al., 2017). It was realized that more than 30 cases of the sexual transmission were from males to females, one case from male to male, and also, one case from female to male has been reported (Pielnaa et al., 2020). In addition to semen, it is important to note that ZIKV has also been detected in urine, saliva, and nasopharyngeal swabs, potentially facilitating the non-vector-borne transmission of ZIKV **(**Song et al., 2017).

1.2.2. Vertical transmission

ZIKV can be passed from an infected mother to her fetus during pregnancy, as evidenced not only by the detection of viral RNA in the amniotic fluid, urine, or serum of mothers whose fetuses had brain abnormalities, but also by probing of viral RNA/proteins/particles in the brain, placenta, or serum of newborns with microcephaly and those aborted **(**Song et al., 2017**)**.

1.2.3. Blood transfusion

ZIKV is likely to be transmitted through transfusions of blood from donors who have been infected with the virus (Song et al., 2017).The first case of transmission of Zika by blood transfusions would be reported in Brazil in 2015 and published in 2016 (Amorim L., 2019), this complicates the viral transmission as most ZIKV- infected patients are asymptomatic. Asymptomatic blood donates can easily transmit ZIKV to blood recipients. About 3 % of blood donates were tested positive for the virus. The situation is further aggravated by the fact that ZIKV can be preserved in whole blood of infected persons close to two months (Pielnaa et al., 2020).

1.2.4. Transmission by breast milk

Transmission through breast milk was suspected when ZIKV RNA was detected in breast milk from two mothers, during the French Polynesian epidemic. However, when viruses were inoculated in vero cells, in vitro, there was no viral replication, which made the transmission via this route very unlikely. Both infants had ZIKV in their plasma, even though the viremia had started 24 h before breastfeeding begins. The first analyzed report was the one from French Polynesia, abovementioned. More recently, there was a case report from a Venezuelan patient, who was exclusively breast feeding her 5-month-old child. She had acquired Zika infection; the virus was isolated in breast milk as well as in the child's urine. The virus was successfully cultured in milk. There was a 99% homology between the mother's and child's virus. The mother had symptoms of mild disease, whereas the child remained asymptomatic. So far, there is no conclusive evidence that the virus is transmitted through breast milk, but it is also not possible to rule out this hypothesis. Therefore, many doubts remain about the recommendation to be made to lactating women who acquire Zika, whether or not to interrupt breastfeeding (Amorim L., 2019).

2. Seroprevalence in Africa and Asia prior to the year 2000

A large number of serological studies in the half century since the discovery of ZIKV have revealed a broad but confined geographic distribution of human infection with the virus, across a relatively narrow equatorial belt running from Africa to Asia (Figure 24) (Song et al., 2017).

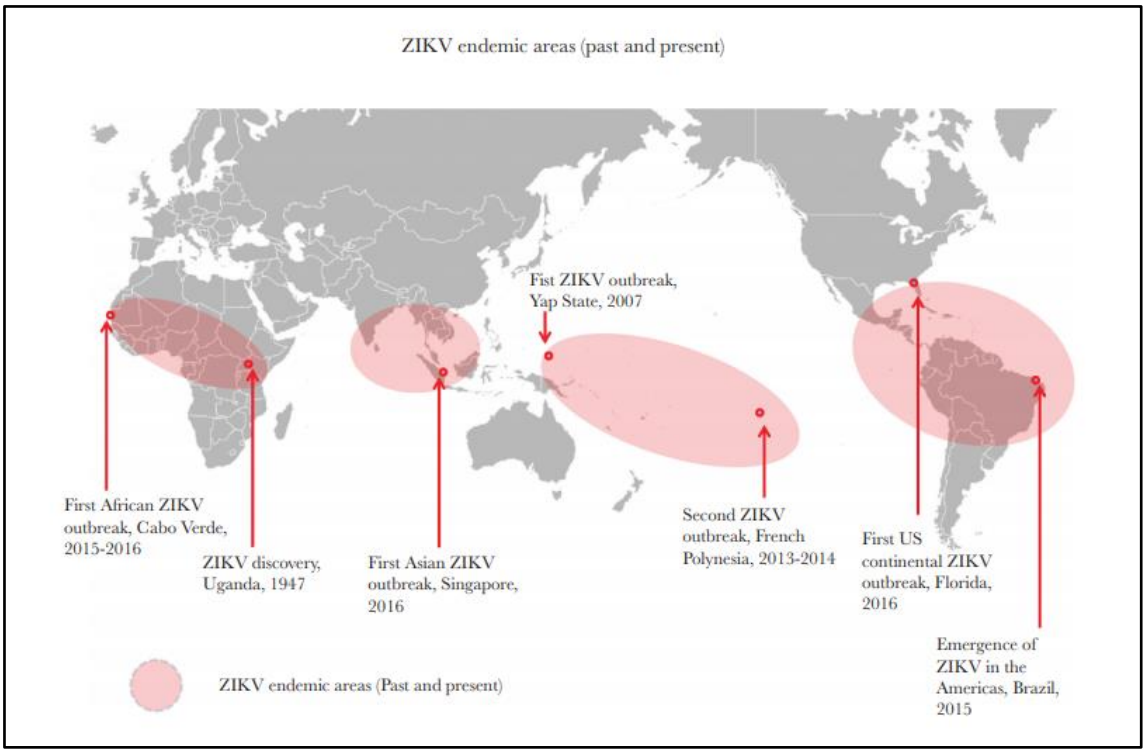


Figure 24. Geographic regions where ZIKV is caused epidemics (Gubler et al., 2017).

2.1. Senegal

A study about Arboviruses circulation of medical interest in southeastern Senegal was conducted from 1988 to 1991. A low level of yellow fever virus activity was detected both in humans and mosquitoes in 1988 to 1990. A Dengue 2 epizootic occurred in 1989-1990. Dengue 2 virus was isolated from humans and mosquitoes in 1990. A ZIKV epizootic outbreak was observed each year (Monlun et al., 1993).

2.2. Sierra-Leon

In a serological and entomological survey on yellow fever carried out in Sierra-Leone in 1972, altogether 899 sera from children 0 to 14 years were tested with 12 antigens by

hemagglutination-inhibition and complement fixation tests. Mouse neutralization test with Yellow Fever, West-Nile and Zika viruses were also performed on selected sera. Generally speaking, the incidence of antibodies for some viruses was found to vary considerably between different areas (Robin et Mouchet., 1975).

2.3. Gabon

Serological studies for arbovirus antibodies were carried out on 1279 human serum specimens collected from adults in south-eastern part of Gabon from June to September 1975 during a multipurpose epidemiological survey. More than 80% of the population surveyed has neutralizing antibodies for Yellow Fever virus as consequence of mass vaccination campaign. Chikungunya, Zika, Wesselsbron and Koutango virus showed some activity, especially in woodland Savannahs (Jan et al., 1978).

2.4. Central African Republic

A serological survey of antibodies to arboviruses was carried out in the human population of the south-east part of Central African Republic in April 1979. Four hundred and fifty-nine serum samples were tested using the hemagglutination inhibition test (HI) and fifty of them by the complement fixation test (CF). Only 11% of the population tested had no HI antibodies against ZIKV (Saluzzo et al., 1981).

2.5. Egypt

Sera from indigenous residents of ten widely scattered localities in Egypt were tested for capacity to neutralize ten different viruses, each known or believed to be arthropod borne. A few sera neutralized ZIKV, but the results do not indicate that this agent is medically important in any of the localities concerned (Smithburn et al., 1954 b).

2.6. Kenya

Arbovirus infections are of public health interest in East Africa, where a very widespread epidemic of o'nyong-nyong fever was reported in 1959 and where the threat of yellow fever, present in neighbouring areas such as Ethiopia, remains. Sera collected in a serological survey in Kenya were therefore tested for antibodies against 3 group-A arboviruses (chikungunya, o'nyong-nyong and Sindbis), 6 group-B arboviruses (Zika, yellow fever, West Nile, Banzi,

Wesselsbron and dengue 1), and Bunyamwera virus. The sera were examined mainly by the haemagglutination-inhibition test but small proportion was also subjected to virus neutralization tests (Geser et al., 1970).

2.7. Pakistan

Complement-fixation test reactions to eight viruses of the family Togaviridae were studied in 372 serum samples (157 rodents, 172 domestic animals, and 43 humans) from Pakistan. Antibodies to each tested virus were detected. The highest over-all prevalence rates were for West Niles (7.8%), Japanese encephalitis (3.2%) and Zika (2.4%) viruses, followed by Sindbis, Chikungunya, Uganda S and Royal Farm viruses (1.6 to 1.3%) (Darwish et al., 1983).

2.8. India

Blood samples obtained from indigenous residents of 38 localities in 6 states in India have been tested for neutralizing antibody against each of 15 different viruses known or believed to be arthropod borne. Antibodies against Dengue types 1 and 2 and West Nile viruses were found in sera from residents of many localities; in some localities the incidence of protection was quite high. Significant numbers of the sera neutralized ZIKV (Smithburn et al., 1954 a).

2.9. Malaysia

A strain of Zika virus (P6-740) was isolated from one of 58 pools of 1277 *Aedes aegypti* mosquitoes collected in cities and towns of peninsular Malaya. The mosquitoes in the positive pool were collected from shop houses in Bentong, a small town in West Central Malaya. No strains of Zika virus were isolated from 59 pools of 4,492 *Aedes albopictus* collected in suburban and rural areas and in rubber plantations, nor from any of 179 pools of 27,636 mosquitoes of 23 other *Aedes* species collected in rural areas, rain forests, mangrove swamps, and fresh-water swamp forests throughout Malaya. P6-740 was readily identified as a strain of ZIKV by the use of antiserum from monkeys in the standard hemagglutination-inhibition test and the plaque-reduction neutralization test in Vero cell cultures. Specific identification of P6-740 as Zika virus was also made in cross-neutralization tests in mice with hyperimmune mouse serum prepared to itself, Zika, Spondweni, and other group B arboviruses. The latter method, however, was less specific than the plaque-reduction neutralization test for comparison of Zika and Spondweni viruses (Marchette et al., 1969).

3. The Yap outbreak and sporadic cases in Southeast Asia during the late 2000s to mid-2010s

3.1. Yap Island

In 2007, however, ZIKV caused the first large outbreak outside of Africa and Asia on Yap Island, a part of the Federated States of Micronesia in the northwestern Pacific Ocean, with a relatively mild disease characterized by fever, rash, arthralgia, and conjunctivitis During this outbreak, approximately 73% of the 6892 Yap residents aged more than 3 years were estimated to be infected with ZIKV, and approximately 18% of the infected people had a clinical illness that was probably attributable to ZIKV infection. Sequence analysis suggested that ZIKV was introduced to Yap Island from Southeast Asia (Song et al., 2017).

3.2. Cambodia

In the early to mid-2010s, a handful of sporadic cases of ZIKV infection were also reported in Southeast Asian countries, such as Cambodia (Song et al., 2017). In August 2010, a blood specimen was collected from a 3-year-old boy at a health clinic in Kampong Speu Province, Cambodia. The child's reported clinical symptoms included 4 days of fever, sore throat, cough and a headache for 3 days. A maculopapular rash was not observed, and the boy was not hospitalized. The clinic staff conducted a follow-up interview and reported that the patient recovered fully. ZIKV infection was confirmed in this patient by using PCR, sequencing, and serology and through virus isolation. ELISA for Chikungunya and Dengue virus IgM and IgG antibodies on acute- and convalescent-phase serum was negative (Heang et al., 2012).

4. The French Polynesia outbreak and spread in the Pacific Islands in the early 2010s

4.1. French Polynesia

No further transmission was identified in the Pacific until October 2013, when French Polynesia (FP) reported the first cases; a subsequent explosive outbreak resulted in an estimated 28 000 cases seeking medical care (approximately 11% of the population) Phylogenetic analyses demonstrated that the FP Zika strain was closely related to Cambodia 2010 and Yap State 2007 strains, corroborating previous findings of the expansion of the ZIKV Asian lineage (Musso et al., 2014).

4.2. New Caledonia

In New Caledonia, the first cases of ZIKV infection imported from French Polynesia were confirmed at the end of November 2013, and the first autochthonous cases were reported by mid-January 2014. Early in February 2014, the New Caledonia Health Authority declared an outbreak situation. Since February 2014, a total of 1,385 ZIKV laboratory-confirmed cases have been detected, including 35 imported cases (32 from French Polynesia, 2 from Vanuatu, and 1 from the Cook Islands). Concomitant with this ZIKV outbreak, circulation of DENV-1 and DENV-3 was reported; more than 150 cases were biologically confirmed during January–September 2014. During this outbreak, patients in New Caledonia were tested for DENV, Chikungunya virus, and ZIKV within the framework of the arboviruses sentinel network, which enabled detection of co-infections. Thus, clinicians should be aware of infections with multiple pathogens in the differential diagnosis of dengue-like illness, especially in patients who returned from tropical regions. This diagnostic procedure could be improved by using multiplex RT-PCR for travelers, given the frequent co-circulation of multiple arboviruses in tropical regions (Dupont-Rouzeyrol et al., 2015).

4.3. Easter Island

In 2013, a large outbreak of Zika was reported on the archipelago of French Polynesia. In this study, the detection and molecular characterization of Zika virus for the first time in Chile from an outbreak among the inhabitants of Easter Island. A total of 89 samples from patients suspected of having ZIKV infection were collected between the period from January to May, 2014. Molecular diagnosis of the virus was performed by RT-PCR followed by the sequencing of the region containing the NS5 gene. A comparison of the viral nucleic acid sequence with those of other strains of ZIKA virus was performed using the MEGA software. Fifty-one samples were found positive for ZIKV by RT-PCR analysis. Further analysis of the NS5 gene revealed that the ZIKV strains identified in Easter Island were most closely related to those found in French Polynesia (99.8 to 99.9 % nucleotides and 100 % aminoacids sequence identity). These results strongly suggest that the transmission pathway leading to the introduction of Zika virus on Easter Island has its origin in French Polynesia (Tognarelli et al., 2016).

4.4. Australia

A female resident of Townsville, Queensland, Australia has been diagnosed with Zika virus infection following a recent trip to the Cook Islands. An initial serum sample collected in March, 2014 was positive by two separate Zika virus TaqMan real-time RT-PCRs and a pan-Flavivirus RT-PCR. Nucleotide sequencing and phylogenetics of the complete Cook Islands Zika virus envelope gene revealed 99.1% homology with a previous Cambodia 2010 sequence within the Asian lineage. In addition, IgG and IgM antibody seroconversions were detected between paired acute and convalescent phase sera using recombinant Zika virus serology assays. This is the first known imported case of Zika virus infection into northern Queensland where the potential mosquito vector *Aedes aegypti* is present and only the second such reported case diagnosed within Australia (Pyke et al., 2014).

4.5. Italy

Returning viremic travelers may ignite autochthonous infections in countries like Italy, which are infested by *Aedes albopictus,* a suitable vector for ZIKV the first two cases of laboratory confirmed Zika virus (ZIKV) infections imported into Italy from French Polynesia (Zammarchi et al., 2015).

4.6. Japan

Two cases of imported Zika fever to Japan, in travelers returning from French Polynesia, where an outbreak due to ZIKV is ongoing since week 41 of 2013 (Kutsuna et al., 2014).

5. The Brazil outbreak and spread in the Americas in 2015–2016

5.1. Brazil

A new challenge has arisen in Brazil with the emergence of ZIKV and co-circulation with others arboviruses (DENV and Chikungunya virus). ZIKV infection in Brazil associated with a recent ongoing outbreak in Camaçari, Bahia, Brazil, of an illness characterized by maculopapular rash, fever, myalgias/arthralgia, and conjunctivitis (Campos et al., 2015).

5.2. America

ZIKV is still causing an unprecedented ongoing epidemic in Latin America and threatening North America and potentially the rest of the world, the explosive pandemic of

Zika virus infection occurring throughout South America, Central America, and the Caribbean and potentially threatening the United States is the most recent of four unexpected arrivals of important arthropod-borne viral diseases in the Western Hemisphere over the past 20 years. It follows dengue, which entered this hemisphere stealthily over decades and then more aggressively in the 1990s; West Nile virus, which emerged in 1999; and Chikungunya, which emerged in 2013 (Fauci et Morens., 2016).

ZIKV was first identified in the Americas in March 2015, when an outbreak of an exanthematous illness occurred in Bahia, Brazil. Epidemiology data indicate that in Salvador, the capital of Bahia, the outbreak had begun in February and extended to June 2015. By October, the virus had spread to at least 14 Brazilian states, and in December 2015, the Brazil Ministry of Health estimated that up to 1.3 million suspected cases had occurred. In October 2015, Colombia reported the first autochthonous transmission of Zika virus outside Brazil, and by March 3, 2016, a total of 51473 suspected cases of Zika virus had been reported in that country. By March 2016, the virus had spread to at least 33 countries and territories in the Americas (Figure 25) (Petersen and al., 2016).

Figure 25. Outbreak of ZIKV in Americas, brown color represents outbreak from January to October 2015 and dark red color represents outbreak from November 2015 to March 2016 (Petersen et al., 2016).

As of November 17, 2016, 48 countries and territories in the Americas had reported the autochthonous mosquito-borne transmission of ZIKV, with an accumulated number of 171,553 confirmed cases : Anguilla; Antigua and Barbuda; Argentina; Aruba; the Bahamas; Barbados; Belize; Bolivia; Bonaire, Sint Eustatius and Saba; the British Virgin Islands; Cayman Islands; Costa Rica; Cuba; Curaçao; Dominica; the Dominican Republic; Ecuador; El Salvador; French Guiana; Grenada; Guadeloupe; Guatemala; Guyana; Haiti; Honduras; Jamaica; Martinique; Mexico; Montserrat; Nicaragua; Panama; Paraguay; Peru; Puerto Rico; Saint Barthélemy; Saint Kitts and Nevis; Saint Lucia; Saint Martin; Saint Vincent and the Grenadines; Sint Maarten; Suriname; Trinidad and Tobago; Turks and Caicos Islands; the United States of America (US); the US Virgin Islands; and Venezuela (Song et al., 2017).

Part II

Experimental Search

Material and Methods

1. Material and Methods

1.1. Gathering of protein sequences from UniProt and alignment

The sequences of ZIKV envelope protein were gathered from UniProt in FASTA format during the month of May 2020. The whole gathered sequences were 720 sequences that represent ZIKV polyproteins and envelope protein. Among those, only 23 sequences represent ZIKV envelope protein. The alignment was conducted using Clustal Omega always on UniProt.

1.1.1. Presentation of UniProt

UniProt, or the Universal Protein Resource is a comprehensive resource for protein sequence and annotation data. The UniProt Web site receives approximately 400,000 unique visitors per month and is the primary means to access UniProt. Along with various datasets that you can search, UniProt provides three main tools. These are the 'BLAST' (Basic Local Alignment Search Tool) tool for sequence similarity searching, the 'Align' tool for multiple sequence alignment, and the 'Retrieve/ID Mapping' tool for using a list of identifiers to retrieve UniProtKB proteins and to convert database identifiers from UniProt to external databases or vice versa. This unit provides three basic protocols, three alternate protocols, and two support protocols for using UniProt tools. These tools are available on their own dedicated pages on the UniProt Web site and are also accessible directly from other parts of the Web site such as the basket, search/tool results pages, and protein entry pages. Having these tools in the UniProt Web site creates an integrated hub of data and analysis tools, allowing both to leverage each other. The UniProt Web site can be accessed at http://www.uniprot.org/ (Pundir et al., 2016).

UniProt is comprised of three components, each optimized for different uses:

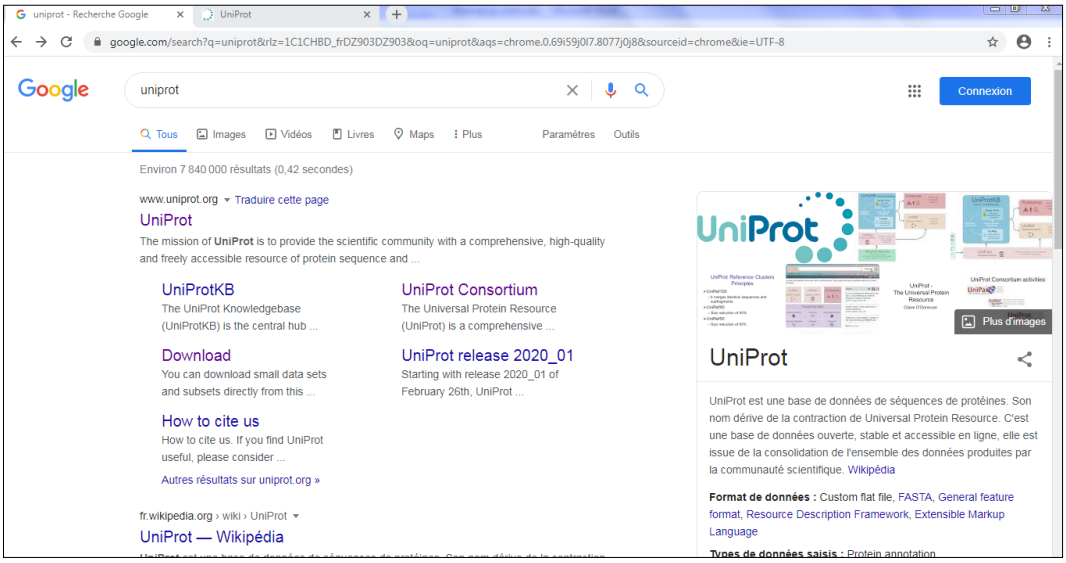
- **UniProt Knowledgebase (UniProtKB)** is the central access point for extensive curated protein information, including function, classification, and cross-reference. UniProtKB comprises two sections:

- **UniProtKB/Swiss-Prot** which is manually annotated and is reviewed.
- **UniProtKB/TrEMBL** which is automatically annotated and is not reviewed.

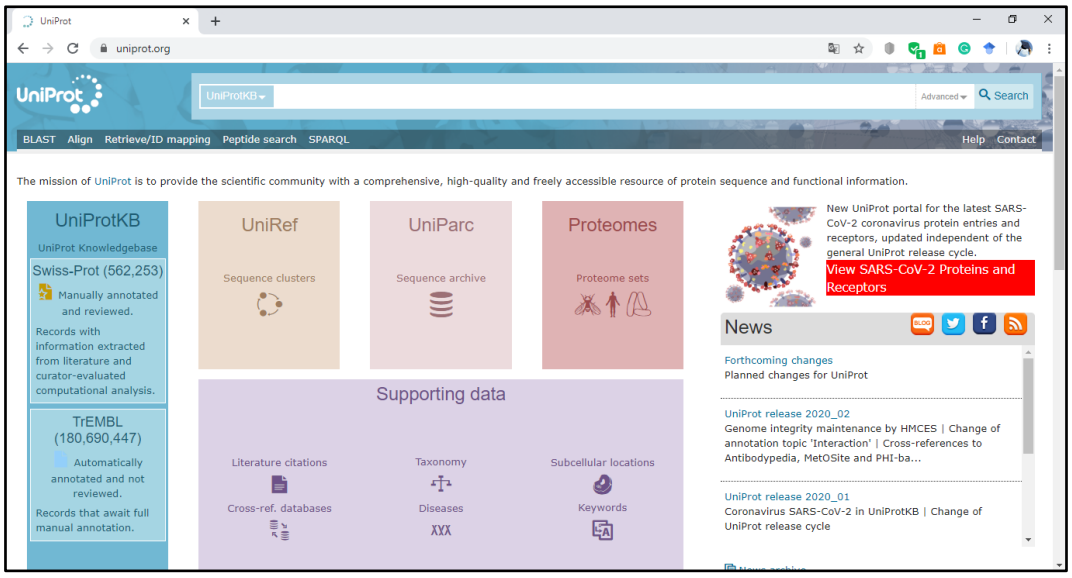
- **UniProt Reference Clusters (UniRef)** databases provide clustered sets of sequences from the UniProtKB and selected UniProt Archive records to obtain complete coverage of sequence space at several resolutions while hiding redundant sequences.
- **UniProt Archive (UniParc)** is a comprehensive repository, used to keep track of sequences and their identifiers (UniProt 2020).

1.1.2. Steps of research on UniProt

The following illustrations summarize the steps of protein sequences research on UniProt.

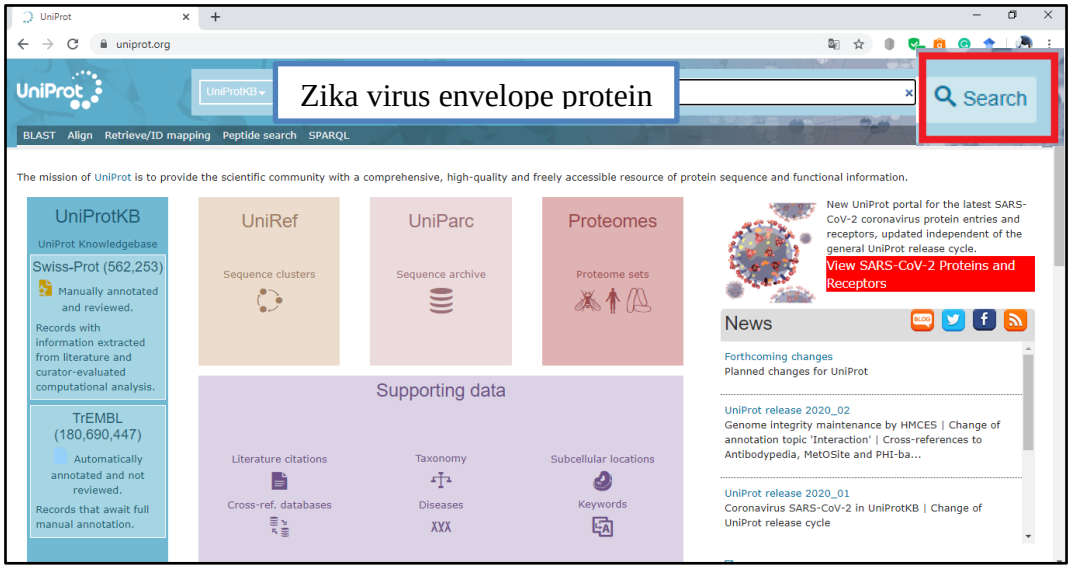


Step 1. Opening of Google browser page (https://www.uniprot.org/).



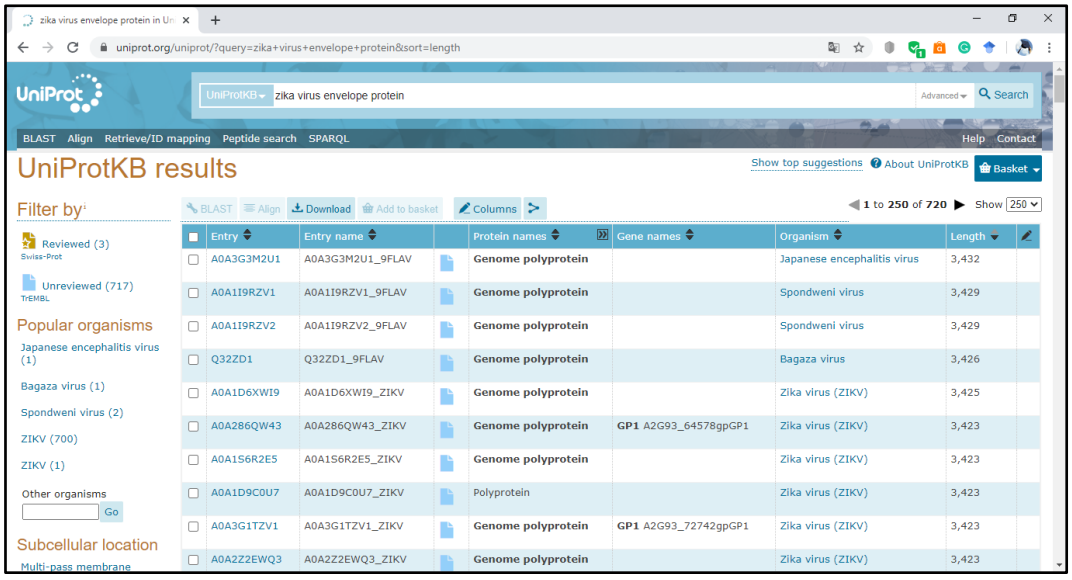
Step 2. UniProt home page.

✓ The search for protein sequences is carried by typing Zika virus envelope protein key words in the search field and clicking on Search.



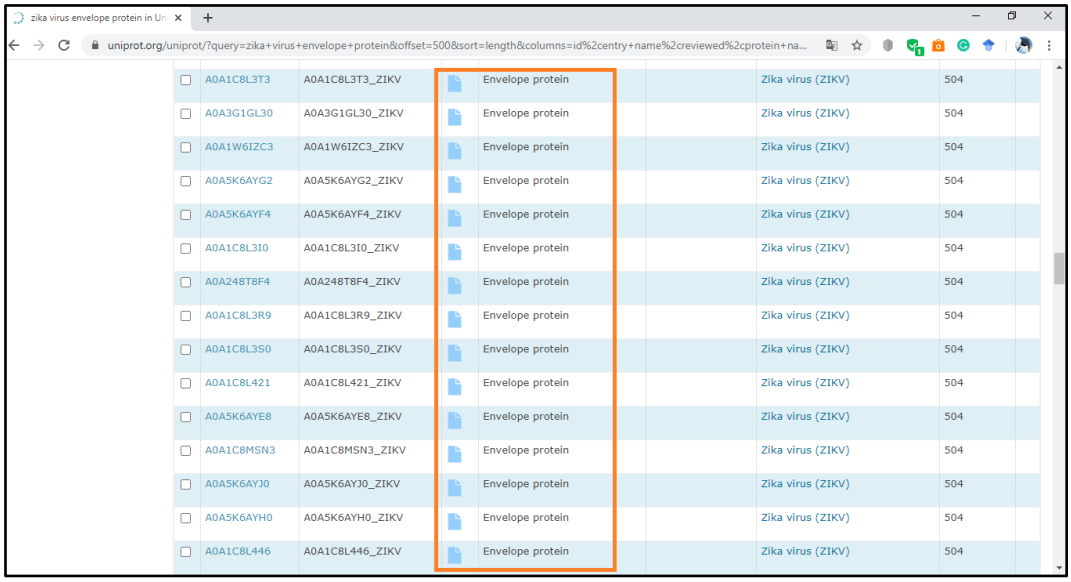
Step 3. Entry of the key words.

✓ After entering the key words, all of the envelope proteins are charged in a table containing 720 samples.



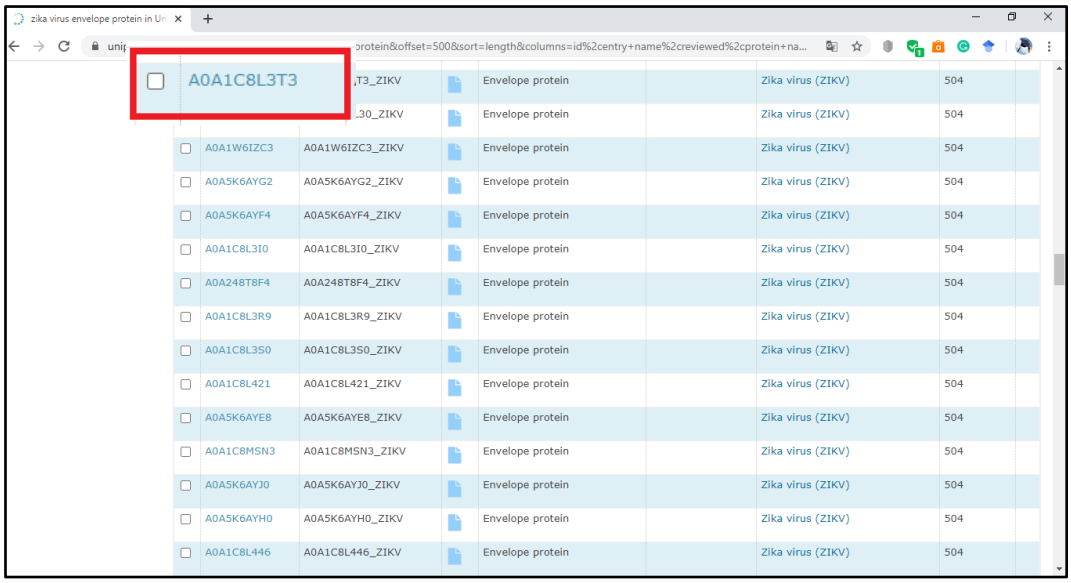
Step 4. The results of ZIKV envelope proteins available on UniProt the month of May 2020.

✓ The table contains the polyprotein and envelope proteins of ZIKV, only 23 existing the envelope proteins are selected.



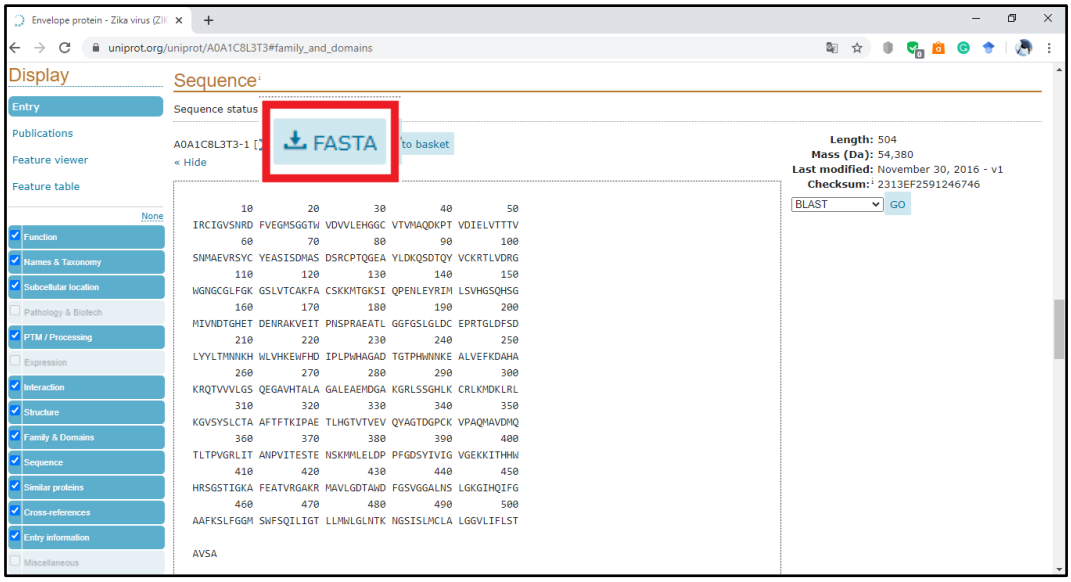
Step 5. The selection of ZIKV envelope proteins.

✓ To download the FASTA sequence go to "entry" and click on the accession number.

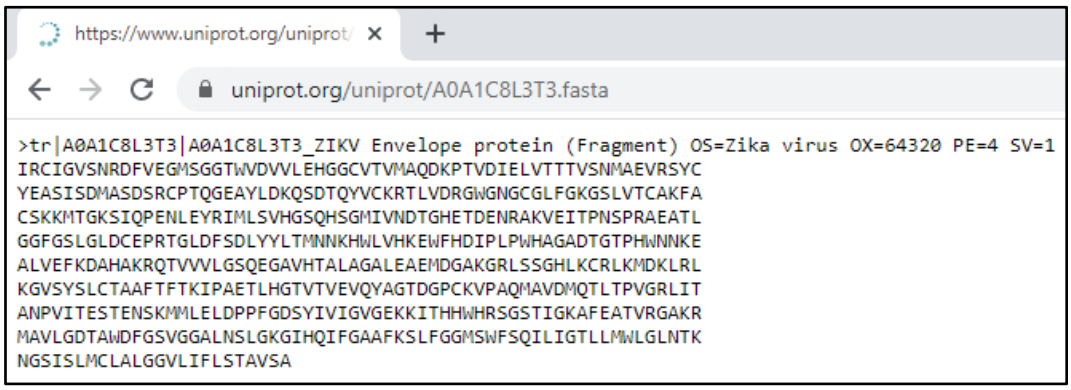


Step 6. FASTA sequences entry.

✓ Click on FASTA to save the file.

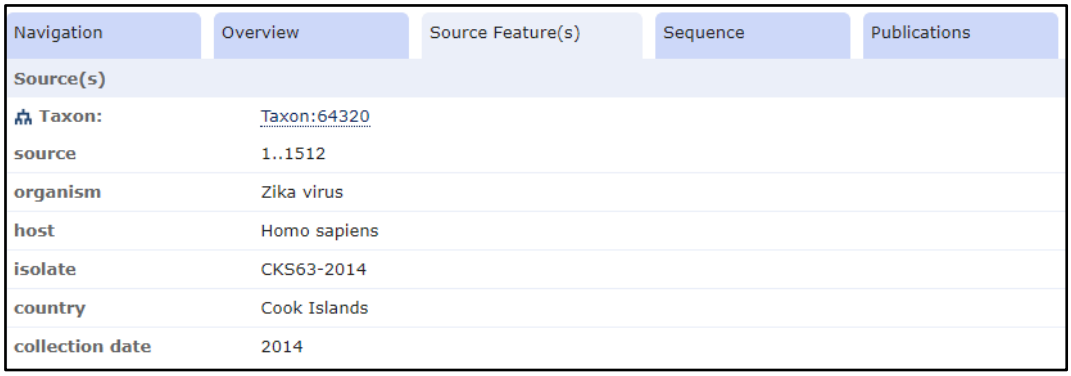


Step 7. FASTA sequences download.



Step 8. The FASTA sequence.

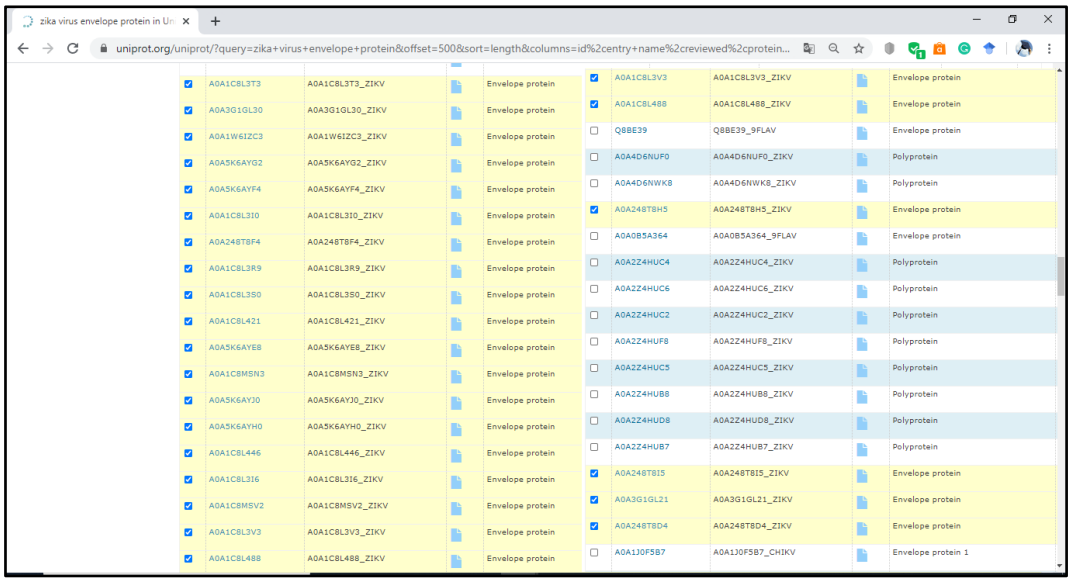
✓ Collection of the data on each sample: host, length, country, isolation source and collection date.



Step 9. Collection of Envelope protein data.

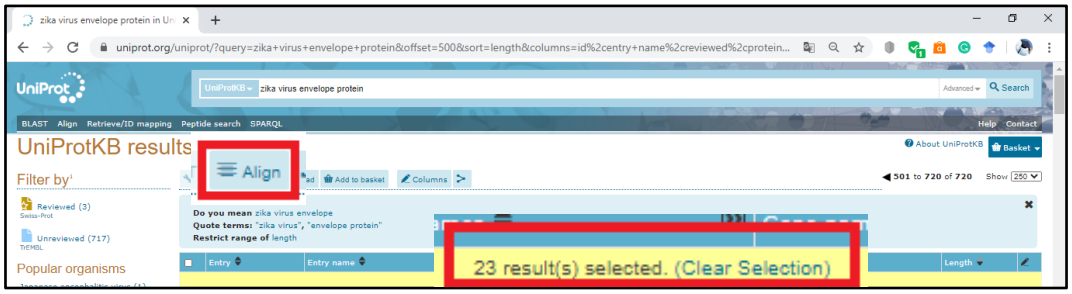
1.1.3. Steps of alignment on UniProt

✓ Click on the small square next to the entry to select only Envelope protein.

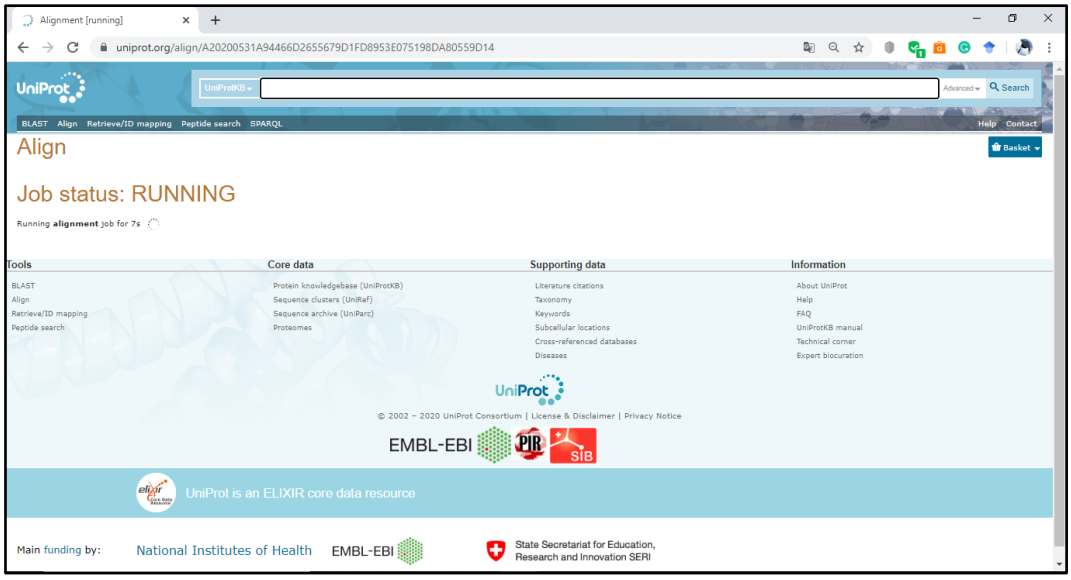


Step 1. Selection of Envelope protein.

✓ Return again to the top of page, and click on Align to make alignment.

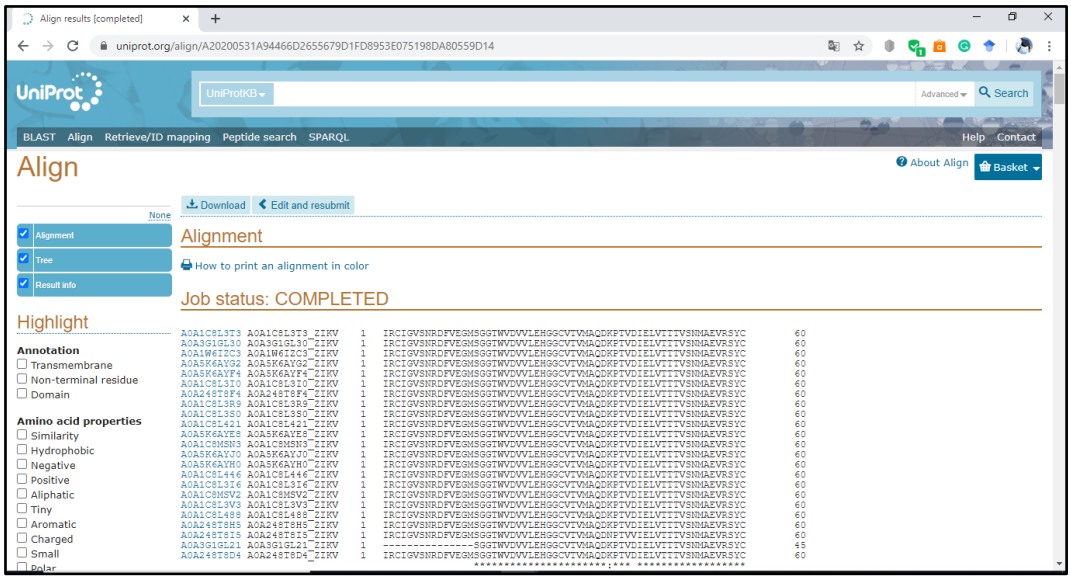


Step 2. Alignment activation.



Step 3. Alignment running.

✓ The alignment results appear in a new window.

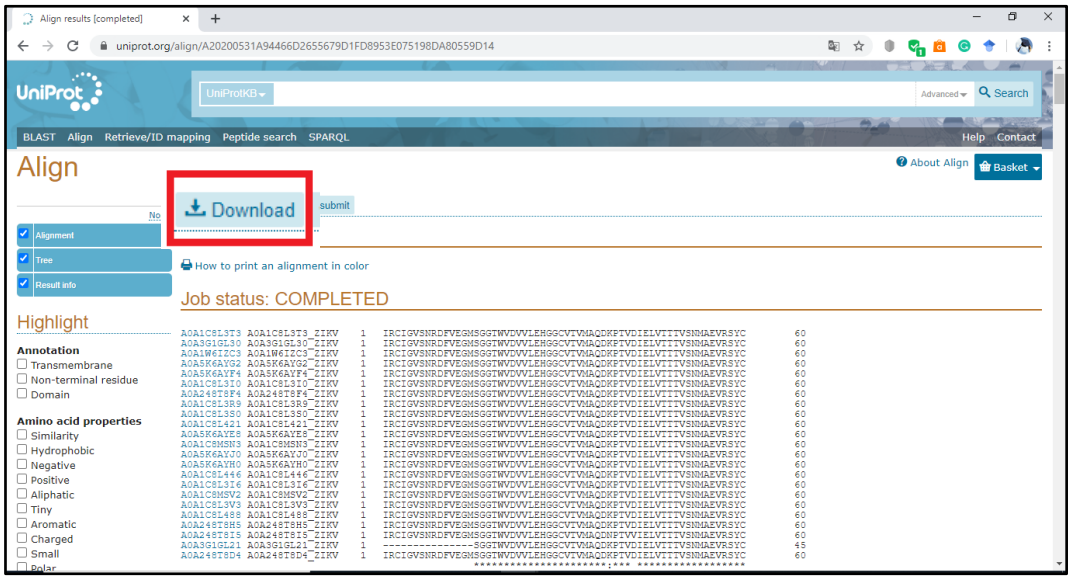


Step 4. Alignment results.

Date of job execution	2020-05-31
Job identifier	A20200531A94466D2655679D1FD8953E075198DA80559D14 (jobs are stored for 7 days)
Running time	19.6 seconds
Identical positions	405
Identity	80.357%
Similar positions	26
Program	clustalo
Default parameters	Default parameters: The default transition matrix is Gonnet, gap opening penalty is 6 bits, gap extension is 1 bit. Clustal-Omega uses the HHalign algorithm and its default settings as its core alignment engine. The algorithm is described in Söding, J. (2005) 'Protein homology detection by HMM–HMM comparison'. Bioinformatics 21, 951-960.

Step 5. Alignment results details: identical position, percentage of identity and similar position.

✓ In the end, download the results of alignment by clicking on the download button.



Step 6. Download of alignment results.

1.2. Phylogenetic analysis

1.2.1. Presentation of MEGA 6 software

MEGA 6 software is the Molecular Evolutionary Genetics Analysis software, in version 6.0, MEGA now enables the inference of timetrees, as it implements the RealTime method for estimating divergence times for all branching points in a phylogeny. A new *Timetree Wizard* in MEGA6 facilitates this timetree inference by providing a graphical user interface (GUI) to specify the phylogeny and calibration constraints step-by-step. This version also contains enhanced algorithms to search for the optimal trees under evolutionary criteria and implements more advanced memory management that can double the size of sequence data sets to which MEGA can be applied (Tamura et al., 2013).

The main window in MEGA contains a menu bar, a main toolbar (just beneath the menu bar), a secondary toolbar near the bottom of the window, and a bottom status bar.

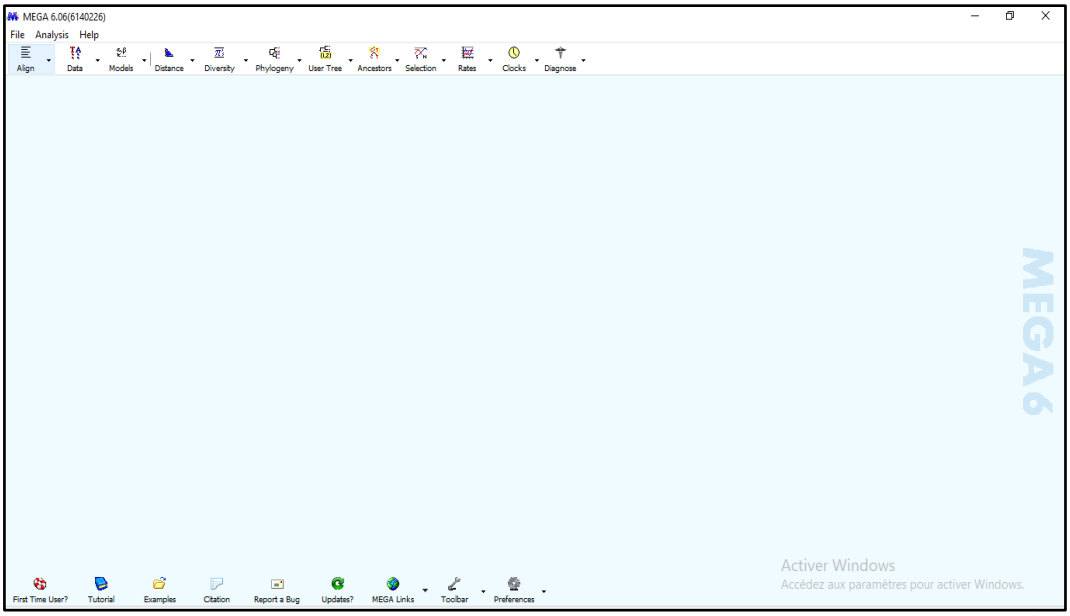
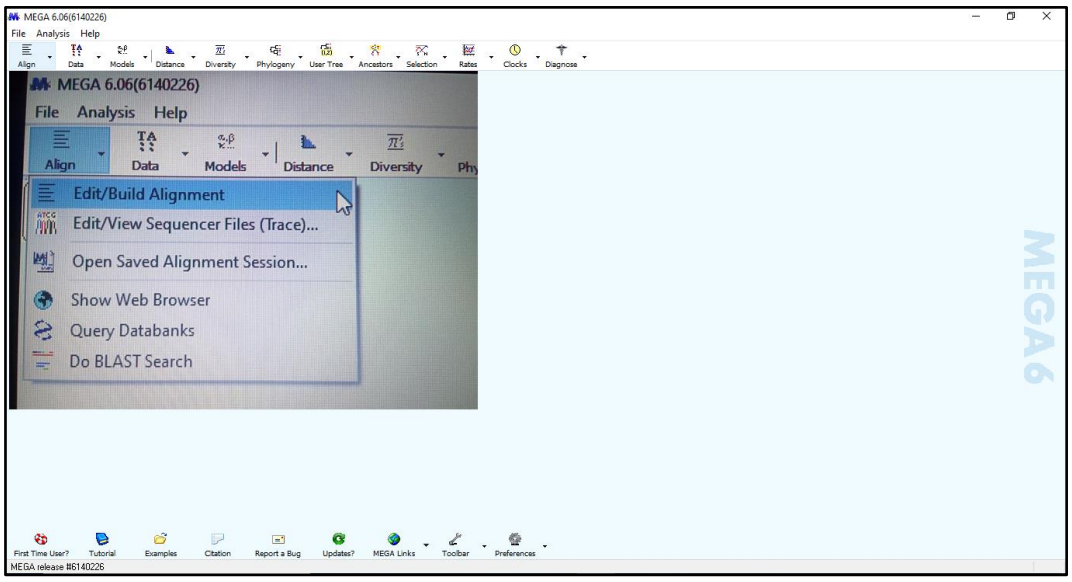


Figure 26. The Main MEGA Window.

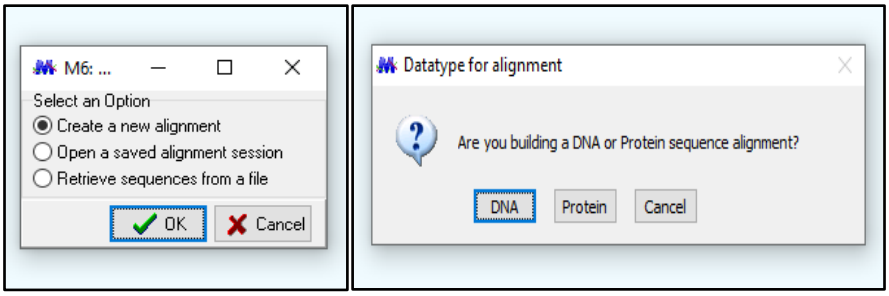
The steps and comments on the actions to be taken for handling the MEGA 6 software are described below.

1.2.2. Aligning sequences by ClustalW

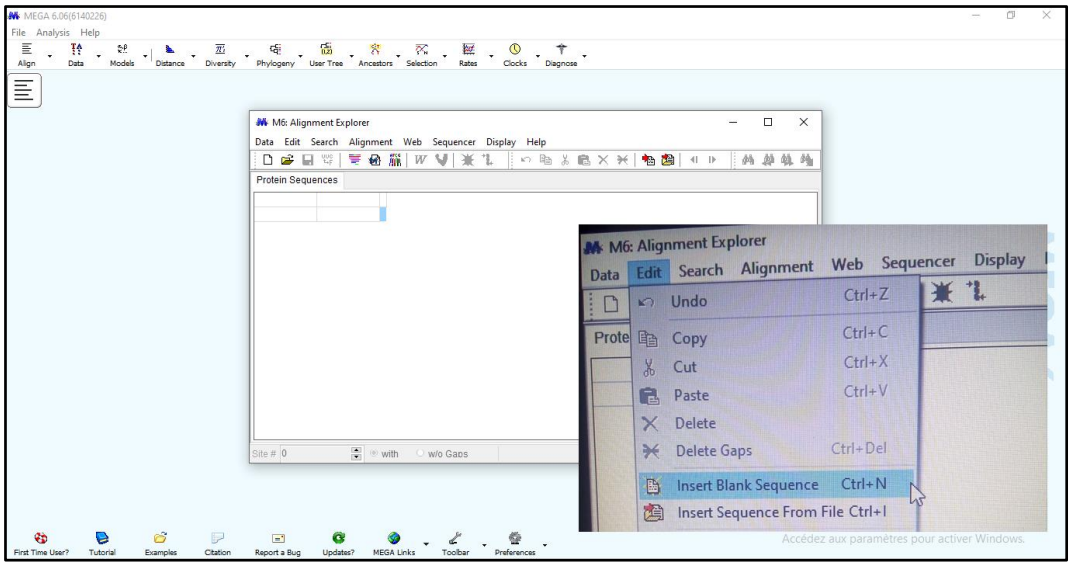
✓ Launch the Alignment Explorer by selecting the Align | Edit/Build Alignment on the launch bar of the main MEGA window.



✓ Select Create New Alignment and click Ok. A dialog will appear asking “Are you building a DNA or Protein sequence alignment?” Click the button labeled “protein”.

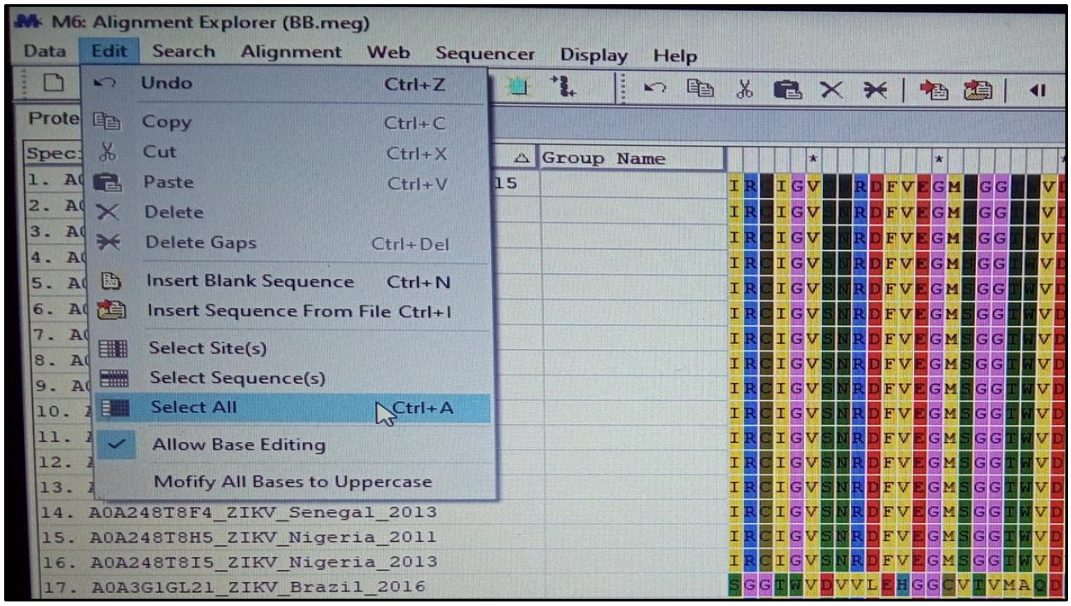


✓ The window that appears allows the introduction of the sequences of interest by clicking on Edit / Insert Blank Sequence and inserting the species name, its access number and the sequence.



✓ Insertion of the 23 samples of the Envelope protein obtained from UniProt.

✓ Select the Edit | Select All menu command to select all sites for every sequence in the data set.



✓ Select Alignment | Align by ClustalW from the main menu to align the selected sequences data using the ClustalW algorithm. Click the “Ok” button to accept the default settings for ClustalW.

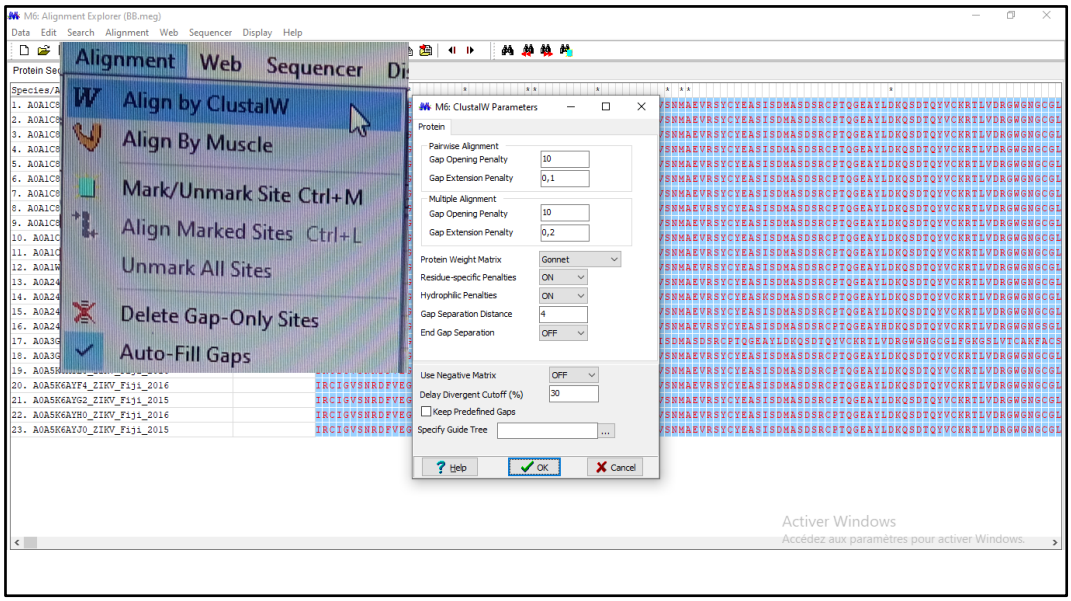


Figure 27. Alignment by ClustalW.

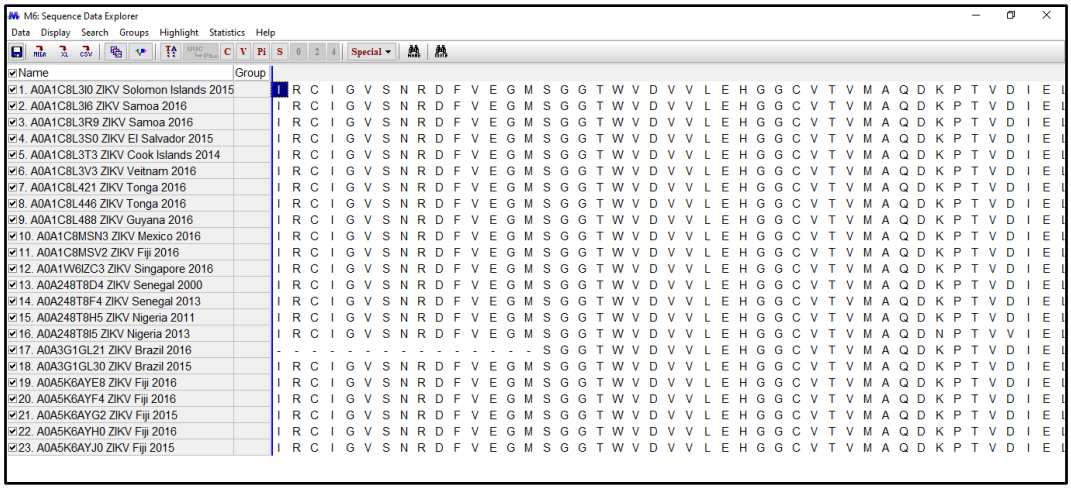


Figure 28. The results of the alignment.

✓ Once the alignment is complete, save the current alignment session by selecting Data | Save Session from the main menu. Give the file an appropriate name, such as "align". This will allow the current alignment session to be restored for future editing.

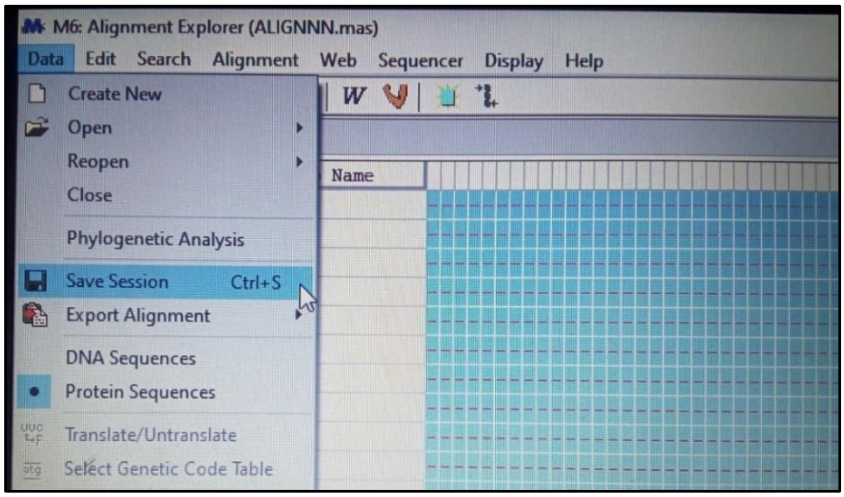
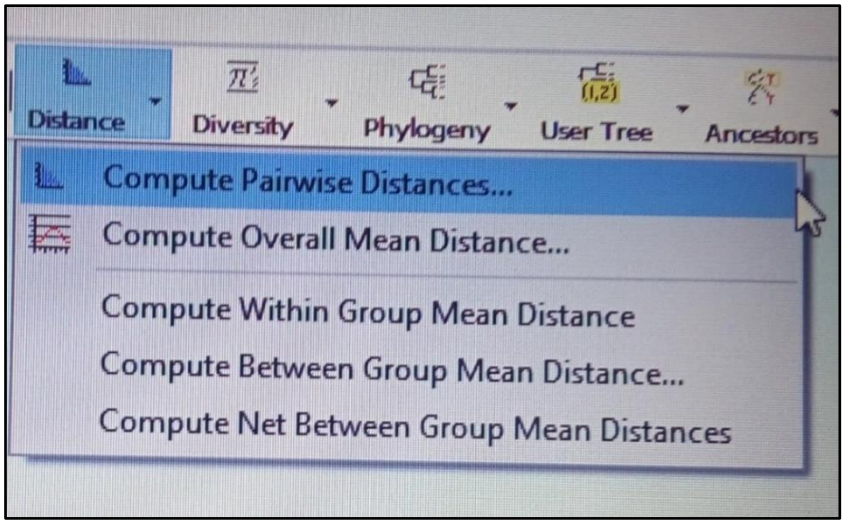


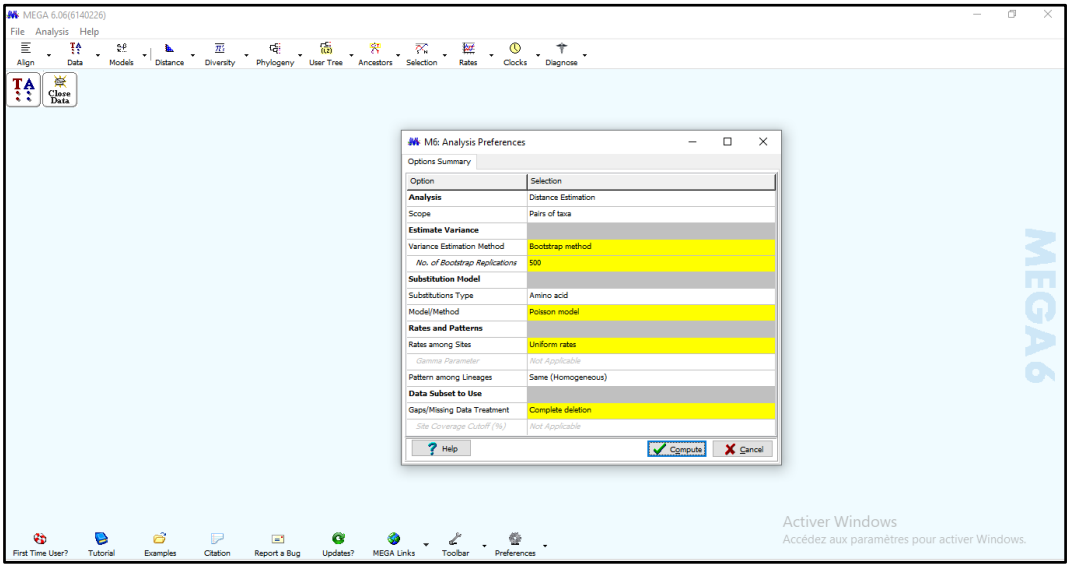
Figure 29. Saving of alignment results.

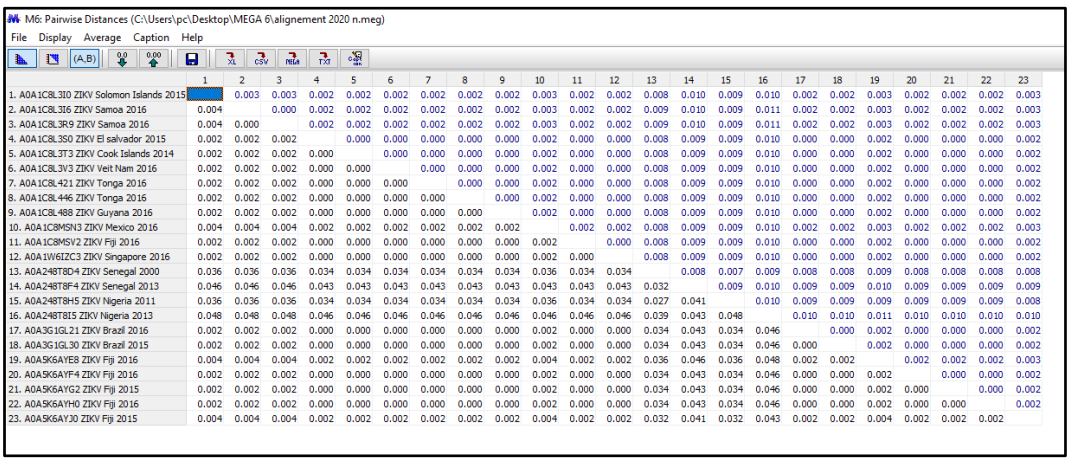
1.2.3. Estimating Evolutionary Distances Using Pairwise Distance

✓ From the main MEGA launch bar, select Distance | Compute Pairwise Distance.



✓ Click Compute to begin the computation.



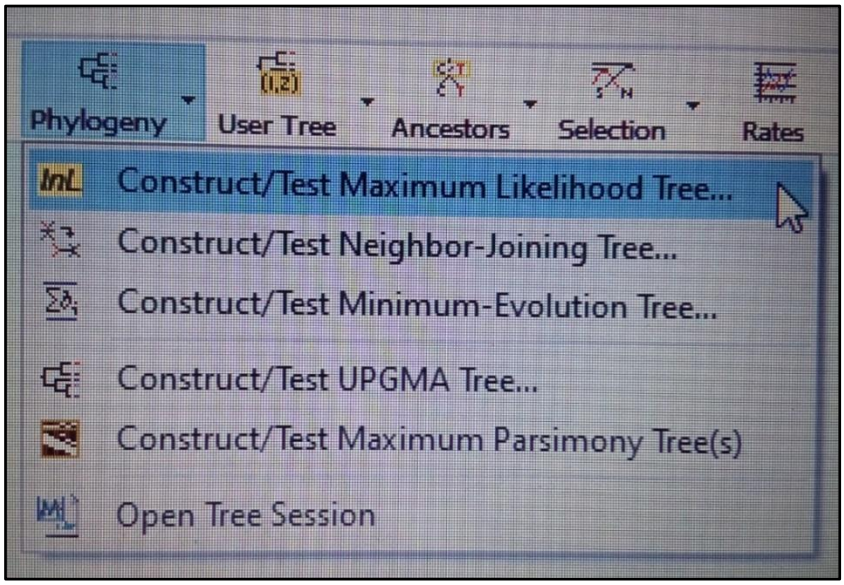


	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. A0A1C8L3I0 ZIKV Solomon Islands 2015		0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.008	0.010	0.009	0.010	0.002	0.002	0.003	0.002	0.002	0.002	0.003
2. A0A1C8L3I6 ZIKV Samoa 2016	0.004		0.000	0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.009	0.010	0.009	0.011	0.002	0.002	0.003	0.002	0.002	0.002	0.003
3. A0A1C8L3R9 ZIKV Samoa 2016	0.004	0.000		0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.009	0.010	0.009	0.011	0.002	0.002	0.003	0.002	0.002	0.002	0.003
4. A0A1C8L3S0 ZIKV El salvador 2015	0.002	0.002	0.002		0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
5. A0A1C8L3T3 ZIKV Cook Islands 2014	0.002	0.002	0.002	0.000		0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
6. A0A1C8L3V3 ZIKV Veit Nam 2016	0.002	0.002	0.002	0.000	0.000		0.000	0.000	0.000	0.002	0.000	0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
7. A0A1C8L421 ZIKV Tonga 2016	0.002	0.002	0.002	0.000	0.000	0.000		0.000	0.000	0.002	0.000	0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
8. A0A1C8L446 ZIKV Tonga 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000		0.000	0.002	0.000	0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
9. A0A1C8L488 ZIKV Guyana 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000		0.002	0.000	0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
10. A0A1C8MSN3 ZIKV Mexico 2016	0.004	0.004	0.004	0.002	0.002	0.002	0.002	0.002	0.002		0.002	0.002	0.008	0.009	0.009	0.010	0.002	0.002	0.003	0.002	0.002	0.002	0.003
11. A0A1C8MSV2 ZIKV Fiji 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002		0.000	0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
12. A0A1W6IZC3 ZIKV Singapore 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000		0.008	0.009	0.009	0.010	0.000	0.000	0.002	0.000	0.000	0.000	0.002
13. A0A248T8D4 ZIKV Senegal 2000	0.036	0.036	0.036	0.034	0.034	0.034	0.034	0.034	0.034	0.036	0.034	0.034		0.008	0.007	0.009	0.008	0.008	0.009	0.008	0.008	0.008	0.008
14. A0A248T8F4 ZIKV Senegal 2013	0.046	0.046	0.046	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.032		0.009	0.010	0.009	0.009	0.010	0.009	0.009	0.009	0.009
15. A0A248T8H5 ZIKV Nigeria 2011	0.036	0.036	0.036	0.034	0.034	0.034	0.034	0.034	0.034	0.036	0.034	0.034	0.027	0.041		0.010	0.009	0.009	0.009	0.009	0.009	0.009	0.008
16. A0A248T8I5 ZIKV Nigeria 2013	0.048	0.048	0.048	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.039	0.043	0.048		0.010	0.010	0.011	0.010	0.010	0.010	0.010
17. A0A3G1GL21 ZIKV Brazil 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.034	0.043	0.034	0.046		0.000	0.002	0.000	0.000	0.000	0.002
18. A0A3G1GL30 ZIKV Brazil 2015	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.034	0.043	0.034	0.046	0.000		0.002	0.000	0.000	0.000	0.002
19. A0A5K6AYE8 ZIKV Fiji 2016	0.004	0.004	0.004	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.002	0.002	0.036	0.046	0.036	0.048	0.002	0.002		0.002	0.002	0.002	0.003
20. A0A5K6AYF4 ZIKV Fiji 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.034	0.043	0.034	0.046	0.000	0.000	0.002		0.000	0.000	0.002
21. A0A5K6AYG2 ZIKV Fiji 2015	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.034	0.043	0.034	0.046	0.000	0.000	0.002	0.000		0.000	0.002
22. A0A5K6AYH0 ZIKV Fiji 2016	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.034	0.043	0.034	0.046	0.000	0.000	0.002	0.000	0.000		0.002
23. A0A5K6AYJ0 ZIKV Fiji 2015	0.004	0.004	0.004	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.002	0.002	0.032	0.041	0.032	0.043	0.002	0.002	0.004	0.002	0.002	0.002	

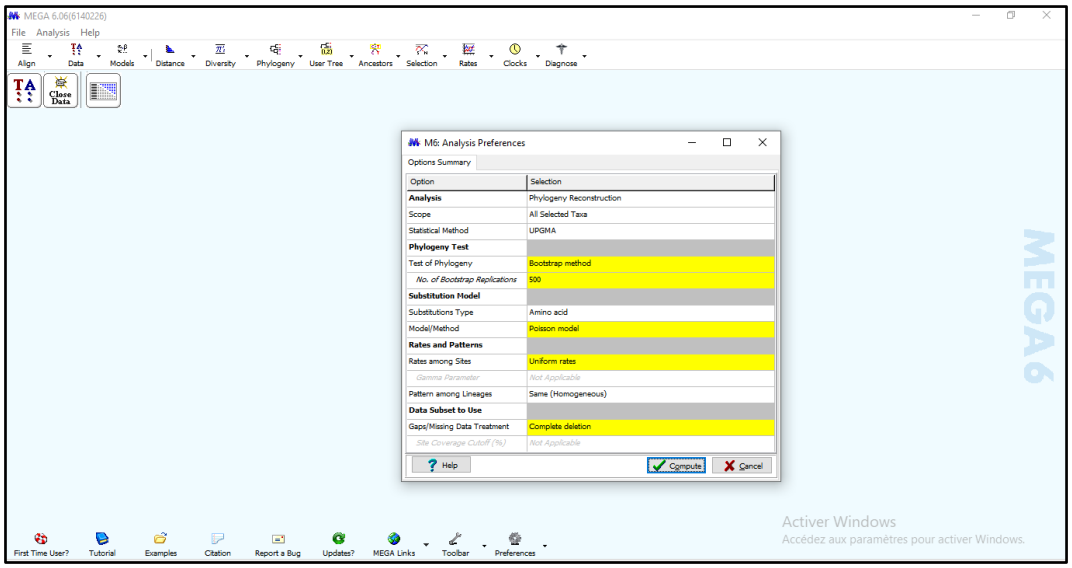
Figure 30. The result of Estimating Evolutionary Distances.

1.2.4. Constructing the phylogenetic tree

✓ Select Phylogeny | Construct/Test Maximum Likelihood Tree option from the main MEGA window launch bar.



✓ Click on the button labeled Compute.



✓ The phylogenetic tree obtained is below.

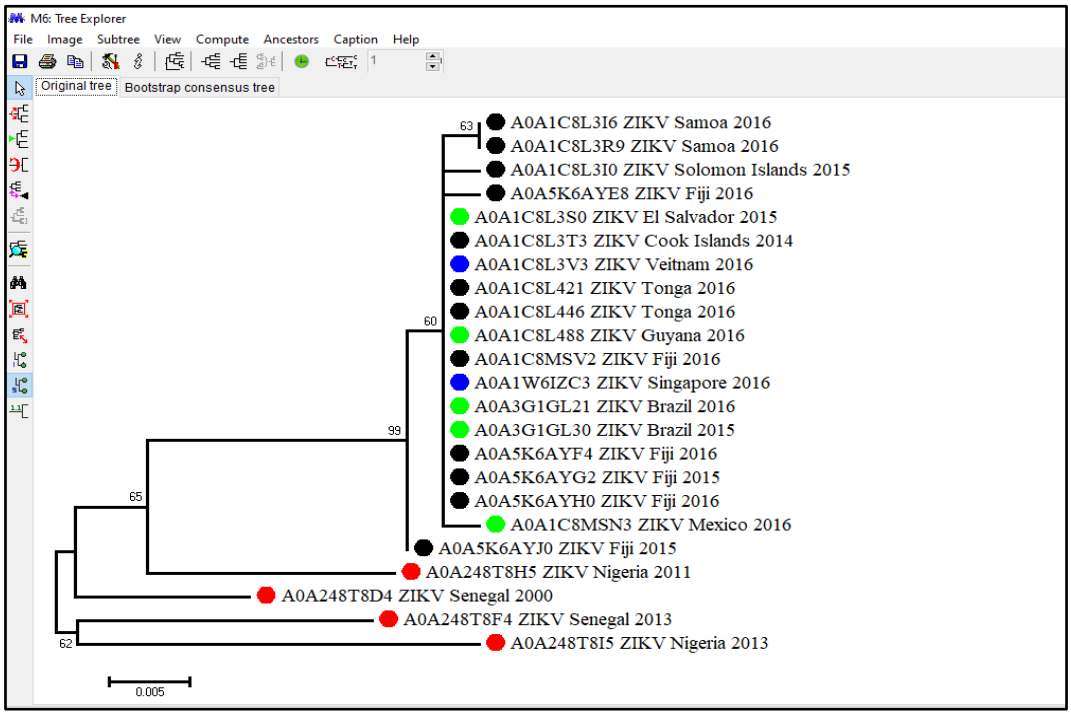


Figure 31. Result of phylogenetic analysis.

1.3. Protein structure homology-modelling

The construction of the three-dimensional structure of ZIKV envelope protein, from Senegal, carrying the access code A0A248T8F4 on UniProt submitted in 10 July 2013, is done by SWISS-MODEL server. The choice of this structure was based on the number of amino acids which is 504, because the template consists of the same number, and also the only African strain with this number.

1.3.1. Presentation of SWISS-MODEL

The aim of the SWISS-MODEL repository is to provide access to an up-to-date collection of annotated 3D protein models generated by automated homology modelling for relevant model organisms and experimental structure information for all sequences in UniProtKB. Regular updates ensure that target coverage is complete, that models are built using the most recent sequence and template structure databases, and that improvements in the underlying modelling pipeline are fully utilized. It also allows users to assess the quality of the models using the latest QMEAN results. If a sequence has not been modelled, the user can build

models interactively via the SWISS-MODEL workspace. Currently the repository contains 1,678,572 models from SWISS-MODEL for UniProtKB targets as well as 155,451 structures from PDB (Bienert et al., 2017).

1.3.2. Start a new modelling project

✓ To reach to SWISS-MODEL server, after writing it in Google bar, click on SWISS-MODEL home page.

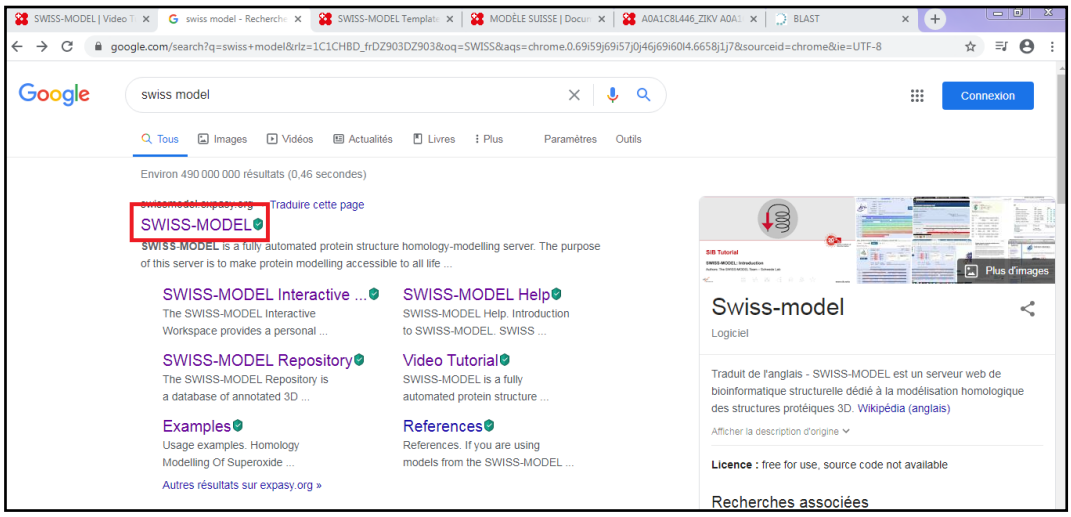


Figure 32. The access to SWISS-MODEL home page.

✓ To start a new project, click on the start modelling bar.

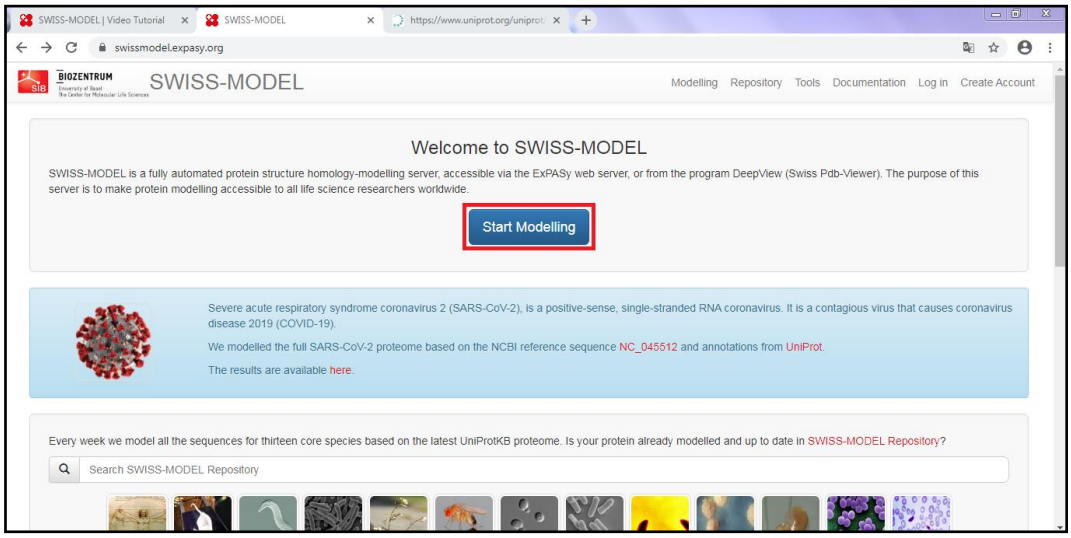


Figure 33. Start modelling.

✓ There are four different types of inputs, the most frequently used input is protein sequences, choose sequence(s) bar.

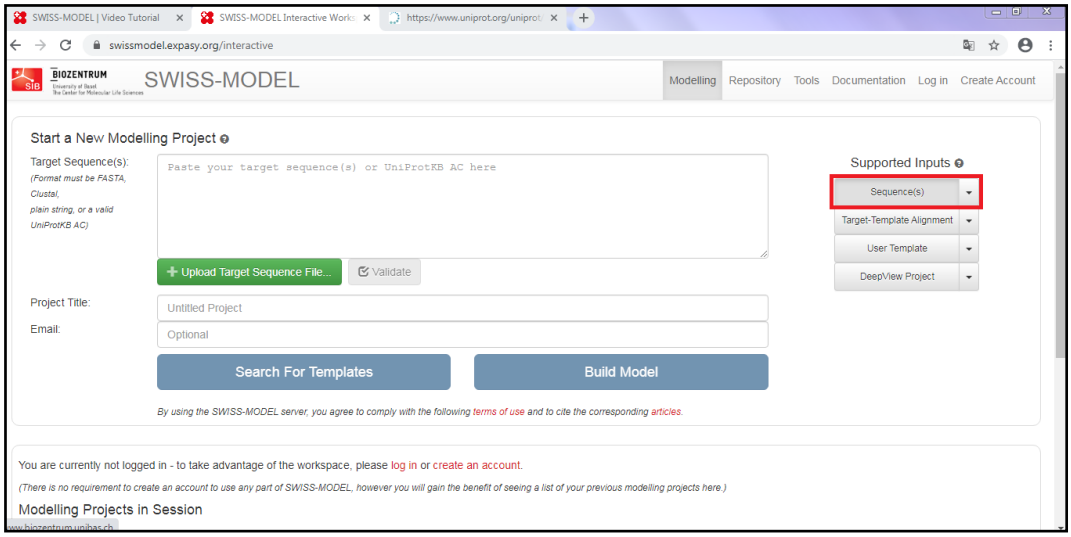


Figure 34. Sequence(s) input.

✓ Paste access code of target sequence of Senegalese strain isolated in 2013 (A0A248T8F4) from UniProt.

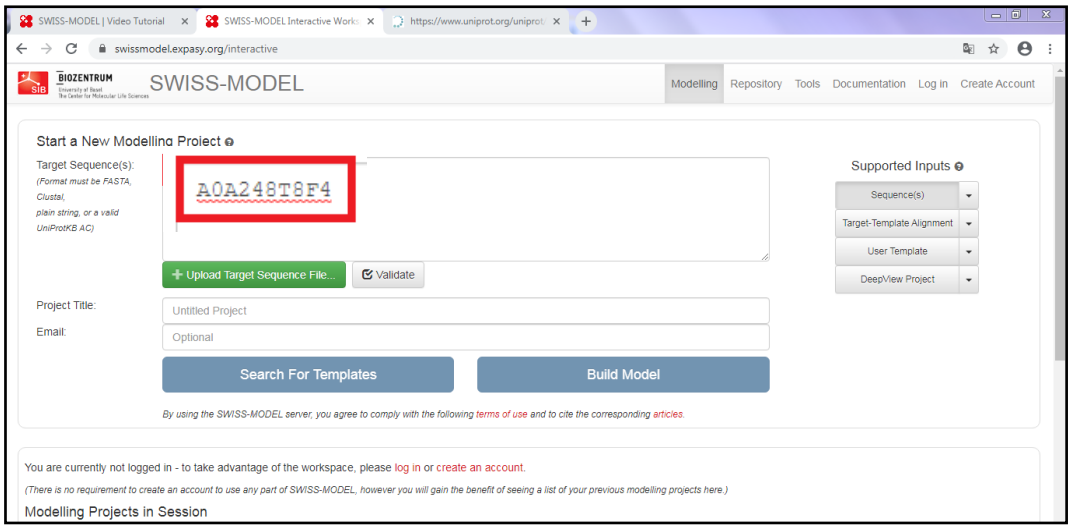


Figure 35. Access code insertion.

✓ Upload target sequence on FASTA format, make sure no PDF document, and no screenshot.

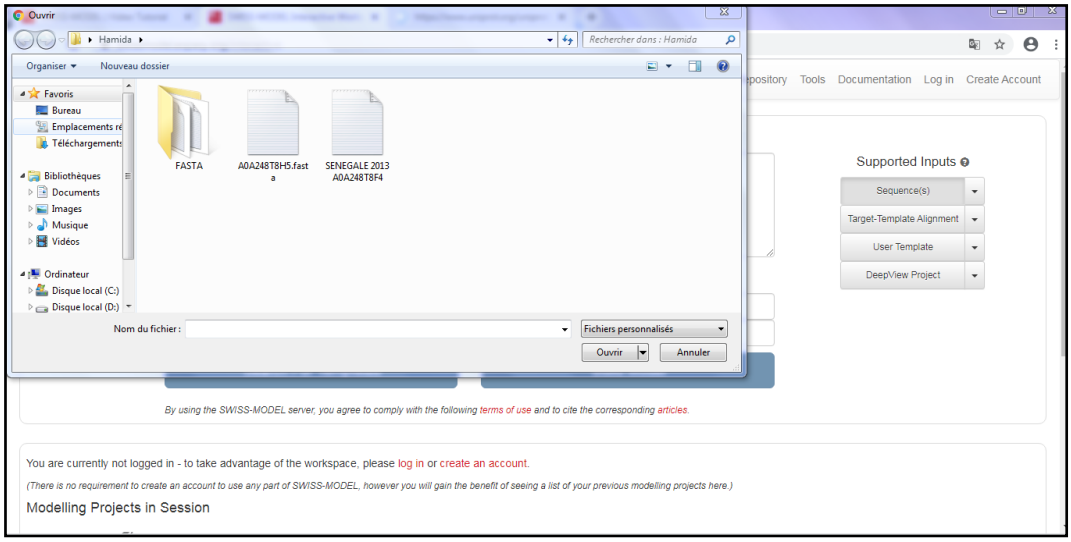


Figure 36. Upload target sequence.

1.3.3. Template search results and modelling

✓ The important point in homology modelling is the identification of a good template, click on search for templates bar.

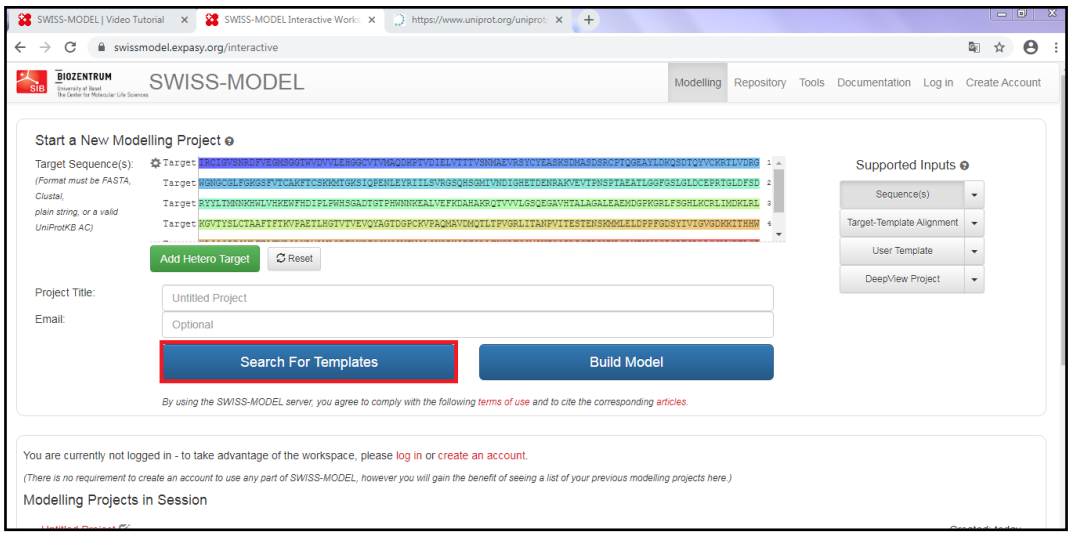


Figure 37. The search for template.

✓ The template search, took less than fifty minutes, it's the slowest process in the modeling.

Figure 38. Template search running.

✓ When the template search is finished, a list of many templates appears, for target protein sequence, the good template either up of the list, select it in the small box, and click on Build Models.

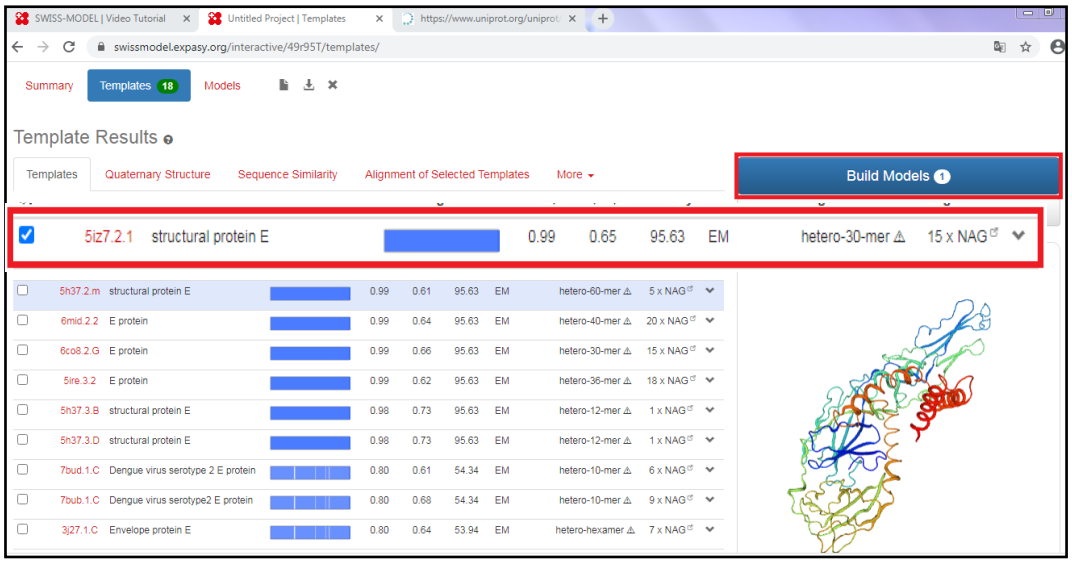


Figure 39. Template results.

✓ The model result appears after a few seconds, the three-dimensional structure appears to the right of the page, with other results.

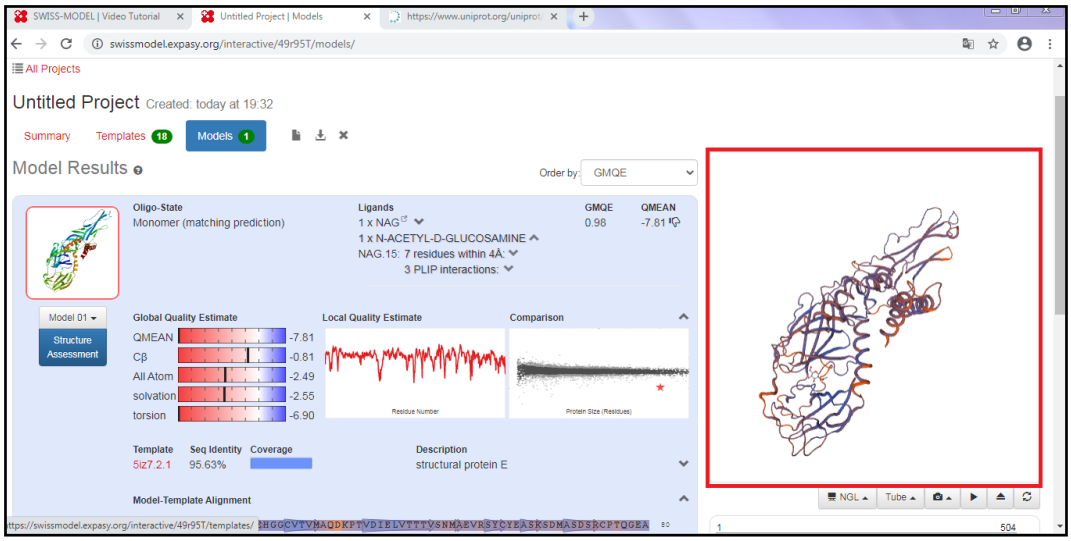


Figure 40. Model results.

Results and discussions

2. Results and discussions

2.1. Gathering of data about ZIKV envelope protein

The gathering of data about ZIKV envelope protein was done on UniProt database the month of May, 2020 (Table 6).

Table 6. Envelope proteins data gathered from UniProt.

	Accession number	Host	Country	Isolation source	Collection date	References
1	A0A5K6AYG2	Homo sapiens	Fiji	Saliva	11 Jul 2015	Kama et al. 2017
2	A0A5K6AYF4	Homo sapiens	Fiji	Serum	03 Aug 2016	Kama et al. 2017
3	A0A5K6AYH0	Homo sapiens	Fiji	Serum	22 Jul 2016	Kama et al. 2017
4	A0A5K6AYJ0	Homo sapiens	Fiji	Saliva	11 Jul 2015	Kama et al. 2017
5	A0A5K6AYE8	Homo sapiens	Fiji	Serum	26 May 2016	Kama et al. 2017
6	A0A248T8I5	Homo sapiens	Nigeria	NI	05 Jul 2013	Herrera et al. 2017
7	A0A248T8H5	Homo sapiens	Nigeria	NI	05 Apr 2011	Herrera et al. 2017
8	A0A248T8D4	Homo sapiens	Senegal	NI	11 Nov 2000	Herrera et al. 2017
9	A0A1C8MSN3	Homo sapiens	Mexico	NI	2016	Pyke et Moore. 2016
10	A0A1C8L3I0	Homo sapiens	Solomon Islands	NI	2015	Pyke et Moore. 2016
11	A0A248T8F4	Homo sapiens	Senegal	NI	10 Jul 2013	Herrera et al. 2017
12	A0A1C8L488	Homo sapiens	Guyana	NI	2016	Pyke et Moore. 2016
13	A0A1C8L3R9	Homo sapiens	Samoa	NI	2016	Pyke et Moore. 2016
14	A0A1C8L421	Homo sapiens	Tonga	NI	2016	Pyke et Moore. 2016
15	A0A1W6IZC3	Homo sapiens	Singapore	NI	2016	Hapuarachchi et al. 2017
16	A0A1C8L3V3	Homo sapiens	Vietnam	NI	2016	Pyke et Moore. 2016
17	A0A1C8L3I6	Homo sapiens	Samoa	NI	2016	Pyke et Moore. 2016
18	A0A1C8MSV2	Homo sapiens	Fiji	NI	2016	Pyke et Moore. 2016

19	A0A1C8L3T3	Homo sapiens	Cook Islands	NI	2014	Pyke et Moore. 2016
20	A0A1C8L3S0	Homo sapiens	El Salvador	NI	2015	Pyke et Moore. 2016
21	A0A1C8L446	Homo sapiens	Tonga	NI	2016	Pyke et Moore. 2016
22	A0A3G1GL21	*Aedes aegypti	Brazil	NI	Jan 2016	Ayllon et al. 2017
23	A0A3G1GL30	*Aedes aegypti	Brazil	NI	Apr 2015	Ayllon et al. 2017

NI: Not Indicated, *Mosquito

The table represents all of the 23 envelope protein of different ZIKV in the world obtained from UniProt, May 2020. The collection dates of these strains were from 2000 to 2016 distributed across African, Asian, Oceanian and American countries. Some of these viruses were collected from sera and saliva of human, and others collected from the mosquito *Aedes aegypti* which represents the main vector of ZIKV. Each protein has its own access number in the database (UniProt).

The Table 7 details the length of amino acids sequences and 3D structure availability.

Table 7. The length of amino-acids sequences gathered and 3D structure.

	Accession number	Sequence length	3D structure
1	A0A5K6AYG2	504	-
2	A0A5K6AYF4	504	-
3	A0A5K6AYH0	504	-
4	A0A5K6AYJ0	504	-
5	A0A5K6AYE8	504	-
6	A0A248T8I5	476	-
7	A0A248T8H5	494	-
8	A0A248T8D4	463	-
9	A0A1C8MSN3	504	-
10	A0A1C8L3I0	504	-
11	A0A248T8F4	504	-
12	A0A1C8L488	504	-
13	A0A1C8L3R9	504	+
14	A0A1C8L421	504	+

15	A0A1W6IZC3	504	+
16	A0A1C8L3V3	504	+
17	A0A1C8L3I6	504	+
18	A0A1C8MSV2	504	+
19	A0A1C8L3T3	504	+
20	A0A1C8L3S0	504	+
21	A0A1C8L446	504	+
22	A0A3G1GL21	463	-
23	A0A3G1GL30	504	+

(+): presence of 3D structure, (-): absence of 3D structure

The length of envelope proteins varied between 463 and 504 amino acids, where 19 out of 23 samples have the same length of 504 aa, two of them have 463 aa (Brazilian and Senegalese strains), and the remaining two possess a different length (476 and 494 aa), who were from the same country (Nigeria). Likewise, 10 out of 23 strains have its three-dimensional structure available.

2.2. Alignment of ZIKV envelope proteins

The alignment of sequences of ZIKV envelope proteins was done on UniProt database by Clustal Omega.

```
A0A1C8L3T3 A0A1C8L3T3_ZIKV    1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L446 A0A1C8L446_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1W6IZC3 A0A1W6IZC3_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L3S0 A0A1C8L3S0_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A5K6AYH0 A0A5K6AYH0_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L3R9 A0A1C8L3R9_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A5K6AYE8 A0A5K6AYE8_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A248T8F4 A0A248T8F4_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A3G1GL30 A0A3G1GL30_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L3I0 A0A1C8L3I0_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L3I6 A0A1C8L3I6_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L488 A0A1C8L488_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L3V3 A0A1C8L3V3_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A5K6AYF4 A0A5K6AYF4_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A5K6AYJ0 A0A5K6AYJ0_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8MSN3 A0A1C8MSN3_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8MSV2 A0A1C8MSV2_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A5K6AYG2 A0A5K6AYG2_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A1C8L421 A0A1C8L421_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A248T8H5 A0A248T8H5_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A248T8I5 A0A248T8I5_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDNPTVDIELVTTTVSNMAEVRSYC   60
A0A248T8D4 A0A248T8D4_ZIKV     1   IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   60
A0A3G1GL21 A0A3G1GL21_ZIKV     1   ---------------SGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC   45
                                                  **********************:*** ******************
```

Figure 41. Amino-acids sequences alignment of 23 samples of ZIKV envelope proteins from amino acid number 1 to 60. Residues with asterisk represent regions of 100% sequence homology. Dashes indicate gaps introduced to optimize the alignment. All sequences were

obtained from UniProt. The alignment displays the following symbols denoting the degree of conservation: (*), conservative mutations (:), semi-conservative mutation (.), and non-conservative mutations ( ).

After alignment of these samples, in which the number of amino-acids in this crossing is 60, all amino acids were conserved except in A0A248T8I5 (Nigeria/2013), where two mutations are observed. The first Lysine is substituted conservatively by asparagine (K38N), this mutation concerns two amino acids from different chemical families, but they belong to the polar amino acids in Venn diagram. The second non conservative mutation, where Valine took place of Aspartic acid (D42V), where the two belong to the different chemical families.

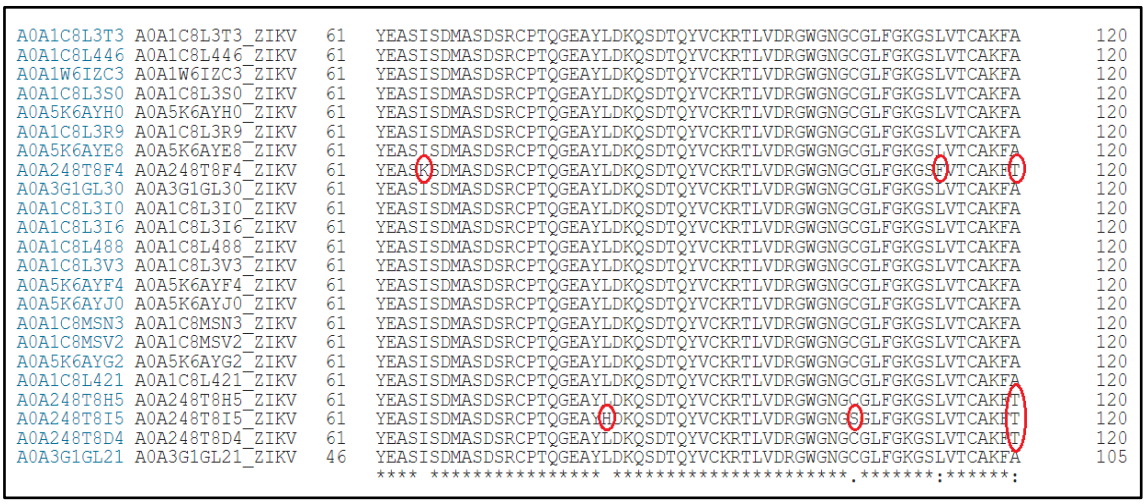


Figure 42. Amino-acids alignment from 61 to 120.

This section includes amino acids from 61 to 120, there are five mutations, the first one is non-conservative mutation in A0A248T8F4 (Senegal/2013) where Isoleucine non-polar is substituted by Lysine positively charged (S65K). The second mutation in A0A248T8I5 (Nigeria/2013), in which Leucine non-polar is replaced non-conservatively by Histidine positively charged (L82H). Semi-conservative mutation in A0A248T8I5 (Nigeria/2013), where Cysteine is substituted by Serine (C105S), these two are uncharged polar amino acids. Conservative mutation in A0A248T8F4 (Senegal/2013), where Glutamic acid (negatively charged) took a place of Leucine (aliphatic) (L113E), and they are far from each other in Venn diagram. The last also is conservative mutation for four samples A0A248T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2013), A0A248T8H5 (Nigeria/2011), and A0A248T8F4 (Senegal/2013), in which Alanine (non-polar amino acid) has been replaced by

Threonine (uncharged polar amino acid), (A120T) but they are near from each other in Venn diagram.

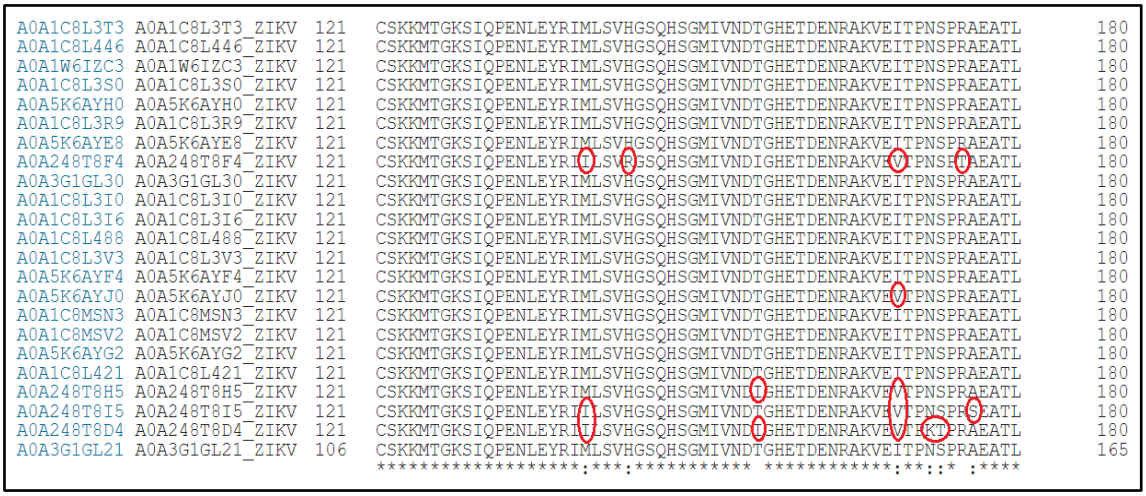


Figure 43. Amino-acids alignment from number 121 to 180.

In this section there are six conservative mutations, the first is in A0A248T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2013), and A0A248T8F4 (Senegal/2013), where Methionine is substituted by Isoleucine (M140I), knowing that they are both non polar amino acids. The second mutation is in A0A248T8F4 (Senegal/2013), where Histidine is substituted by Arginine (H144M), both of them are charged positively amino-acids. The third mutation is in A0A248T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2013), A0A248T8H5 (Nigeria/2011), A0A5K6AYJ0 (Fiji/2015), and A0A248T8F4 (Senegal/2013), in which Isoleucine is replaced by Valine (I169V), the two represent aliphatic and non-polar amino acids. The fourth mutation is in A0A248T8D4 (Senegal/2000), where Asparagine (polar) is substituted by Lysine (polar and charged positively) (N172K). The fifth mutation is in A0A248T8D4 (Senegal/2000), where Serine is replaced by Threonine (S173T), the two are polar uncharged amino acids. The last mutation is in A0A248T8I5 (Nigeria/2013), in which Alanine has been replaced by Serine (A176S). With the presence of two non-conservative mutations, the first in A0A248T8D4 (Senegal/2000), and A0A248T8H5 (Nigeria/2011), where Threonine polar uncharged is replaced by Isoleucine non-polar (T156I), the second in A0A248T8F4, where positively charged Arginine is replaced by polar uncharged Threonine (R175T).

Figure 44. Amino-acid alignment from 181 to 240.

In this section there are six mutation, semi-conservative mutation in A0A248T8H5 (Nigeria/2011), the amino acid Glycine has been replaced by Alanine (G184A), which they belong to non-polar amino acid and fall into the same group of minuscule in Venn diagram. Conservative mutation in A0A248T8I5 (Nigeria/2013), where Leucine is replaced by Valine (L204V), they are considered aliphatic amino acids. Semi-conservative mutation in A0A1C8L3I0 (Solomon/2015), in which Asparagine is replaced by Serine (N208S), the two are polar uncharged amino acids. The three rest mutations are all non-conservative, in A0A248T8F4 (Senegal/2013), where non-polar Leucine is replaced by positively charged Arginine (L205R), the last two in A0A248T8I5, in which non-polar Alanine is replaced by polar uncharged Serine (A127S), and aromatic Tryptophan is replaced by positively charged Arginine (W136R).

Figure 45. Amino-acids alignment from number 241 to 300.

Through the section above, there are seven mutations, conservative mutation in A0A248T8D4 (Senegal/2000), where Glutamine took a place of Histidine (Q288H), while the first amino acid is polar uncharged, and the second positively charged. With the presence of semi-conservative mutation in A0A248T8I5 (Nigeria/2013), in which Cysteine is replaced by Serine (C291S), the two belong to the minuscule group in Venn diagram, also polar uncharged amino acids. And the rest are all non-conservative mutations, the first in A0A248T8I5 (Nigeria/2013), where polar uncharged Threonine took a place of positively charged Arginine (R252T). The second in A0A248T8F4 (Senegal/2013), in which polar uncharged Proline took a place of non-polar Alanine (A280P). The third in A0A248T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2013), and A0A248T8F4 (Senegal/2013), where aromatic Phenylalanine took a place of polar uncharged Serine (S285F). The fourth in A0A248T8H5 (Nigeria/2011), where positively charged Arginine took a place of polar uncharged Proline (P292R). The last in A0A248T8F4 (Senegal/2013), where positively charged Lysine is replaced by non-polar Isoleucine (K294I).

Figure 46. Amino-acids alignment from number 301 to 360.

This passage of alignment contains five mutations. Conservatively mutation in A0A248T8F4 (Senegal/2013), Serine was changed by Threonine (S304T), both of them are polar uncharged amino acids. Another conservatively mutation in A0A248T8F4 (Senegal/2013), A0A048T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2013), and A0A248T8H5 (Nigeria/2011), in which Isoleucine is replaced by Valine (I317V), the two are non-polar amino acids. Conservatively mutation in A0A248T8I5 (Nigeria/2013), in which Glutamine is substituted by Glutamic acid (Q350E), the two are polar amino acids, the second is greater than the first, as it is negatively charged. Conservative mutation, located in

A0A248T8D4 (Senegal/2000), and in A0A248T8H5 (Nigeria/2011), so that the Serine amino acid was replaced by Threonine (S353T), both of them are polar uncharged amino acids. And non-conservative mutation in A0A248T8H5 (Nigeria/2011), where aromatic Tyrosine is replaced by polar Asparagine (Y332N).

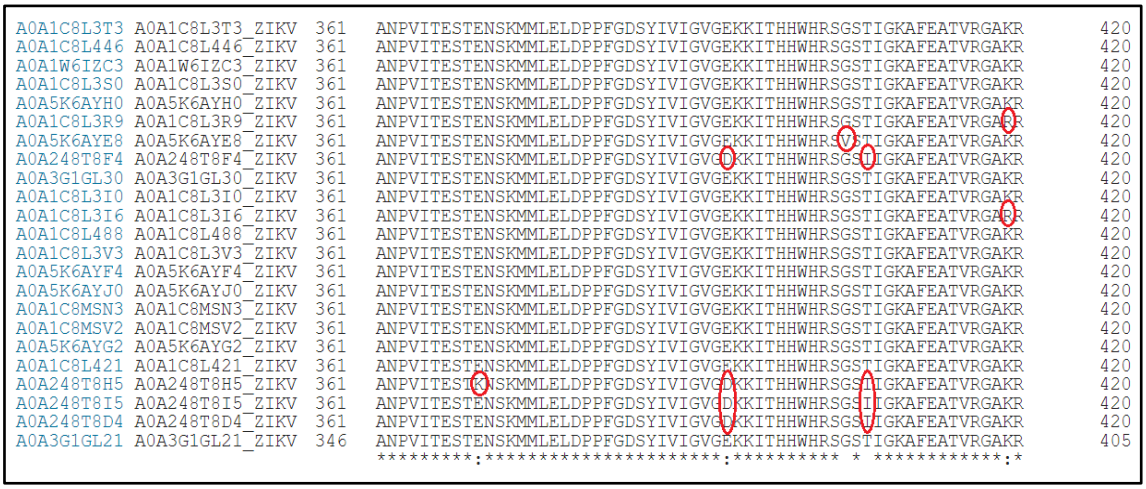


Figure 47. Amino-acids alignment from number 361 to 420.

This sequence of amino acids contains three conservative mutations, the first is in A0A248T8H5 (Nigeria/2011), in which Lysine took place of Glutamic acid (E370K), they are positevely and negatively charged respectively. The second is in A0A248T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2013), A0A248T8H5 (Nigeria/2011), and A0A248T8F4 (Senegal/2013), where was the change between polar negatively charged amino acids, Glutamic acid, and Aspartic acid (E393D). The third is in A0A1C8L3I6 (Samoa/2015), and in A0A1C8L3R9 (Samoa/2015), in which Lysine is replaced by Arginine (K419R), these amino acids are polar and positively charged. With the presence of two non-conservative mutations, the first in A0A5K6AYE8 (Fiji/2016), in which Glycine is replaced by Valine (G404V). The second in A0A248T8D4 (Senegal/2000), A0A248T8H5 (Nigeria/2011), A0A248T8I5 (Nigeria/2013), and A0A248T8F4 (Senegal/2013), where non-polar Isoleucine took a place of polar uncharged Threonine (T406I).

Figure 48. Amino-acids alignment from number 421 to 480.

The sequence of amino acids existing on the top bears five mutations. Non-conservative mutation in A0A248T8D4, where aromatic Tryptophan is replaced by positively charged Arginine (W429R). Semi-conservative mutation in A0A248T8D4 (Senegal/2000), A0A248T8I5 (Nigeria/2011), A0A248T8H5 (Nigeria/2013), and A0A248T8F4 (Senegal/2013), in which Alanine is replaced by Valine (A437V), both of them are non-polar amino acids. As usual conservative mutation in the previous four, where Phenyl alanine took a place of Leucine (L438F), the two are hydrophobic, but the first are aromatic and the second is non-polar amino acid. Another conservative mutation in A0A248T8H5 (Nigeria/2011), where Isoleucine is substituted by Valine (I445V), these amino acids are non-polar. The last one is conservative mutation, also in the same previous strain, the aromatic amino acid Tyrosine took a place of Phenyl alanine (F463Y).

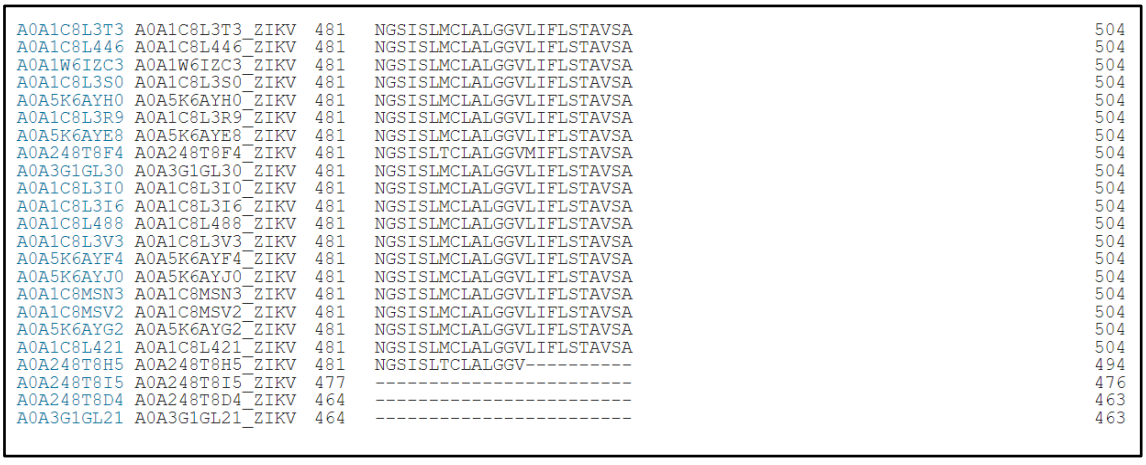


Figure 49. Amino-acids alignment from number 481 to 504.

The Table 8 summarizes the whole of mutations, observed in multiple sequences alignment. The strain accession number, mutation nature and position are mentioned.

Table 8. The whole of mutations in alignment.

Strains	Mutation nature	Amino-acid Position
A0A248T8I5	Conservative	K38N
A0A248T8I5	Non-conservative	D42V
A0A248T8F4	Non-conservative	S65K
A0A248T8I5	Non-conservative	L82H
A0A248T8I5	Semi-conservative	C105S
A0A248T8F4	Conservative	L113E
A0A248T8I5/ A0A248T8H5/ A0A248T8D4 A0A248T8F4	Conservative	A120T
A0A248T8D4/A0A248T8I5/A0A248T8F4	Conservative	M140I
A0A248T8F4	Conservative	H144M
A0A248T8I5/ A0A248T8H5/A0A248T8D4/ A0A248T8F4/ A0A5K6AYJ0	Conservative	I169V
A0A248T8D4	Conservative	N172K
A0A248T8D4	Conservative	S173T
A0A248T8I5	Conservative	A176S
A0A248T8D4/A0A248T8H5	Non-conservative	T156I
A0A248T8F4	Non-conservative	R175T
A0A248T8H5	Semi-conservative	G184A
A0A248T8I5	Conservative	L204V
A0A1C8L3I0	Semi-conservative	N208S
A0A248T8F4	Non-conservative	L205R
A0A248T8I5	Non-conservative	A127S
A0A248T8I5	Non-conservative	W136R
A0A248T8D4	Conservative	Q288H
A0A248T8I5	Semi-conservative	C291S
A0A248T8I5	Non-conservative	R252T
A0A248T8F4	Non-conservative	A280P
A0A248T8D4/ A0A248T8F4/ A0A248T8I5	Non-conservative	S285F
A0A248T8H5	Non-conservative	P292R
A0A248T8F4	Non-conservative	K294I
A0A248T8F4	Conservative	S304T
A0A248T8D4/A0A248T8F4/A0A248T8H5 /A0A248T8I5	Conservative	I317V
A0A248T8I5	Conservative	Q350E
A0A248T8D4/A0A248T8H5	Conservative	S353T
A0A248T8H5	Non-conservative	Y332N
A0A248T8H5	Semi-conservative	E370K
A0A248T8D4 /A0A248T8H5 /A0A248T8I5 A0A248T8F4	Semi-conservative	E393D
A0A1C8L3I6/A0A1C8L3R9	Semi-conservative	K419R

A0A55K6AYE8	Non-conservative	G404V
A0A248T8D4 /AA248T8H5 /A0A248T8I5 A0A248T8F4	Non-conservative	T406I
A0A248T8D4	Non-conservative	W429R
A0A248T8D4/A0A248T8H5/ A0A248T8I5 A0A248T8F4	Semi-conservative	A437V
A0A248T8D4 /A0A248T8H5 /A0A248T8I5 A0A248T8F4	Conservative	L438F
A0A248T8H5	Conservative	I445V
A0A248T8H5	Conservative	F463Y

2.3. Phylogenetic analysis of ZIKV envelope proteins

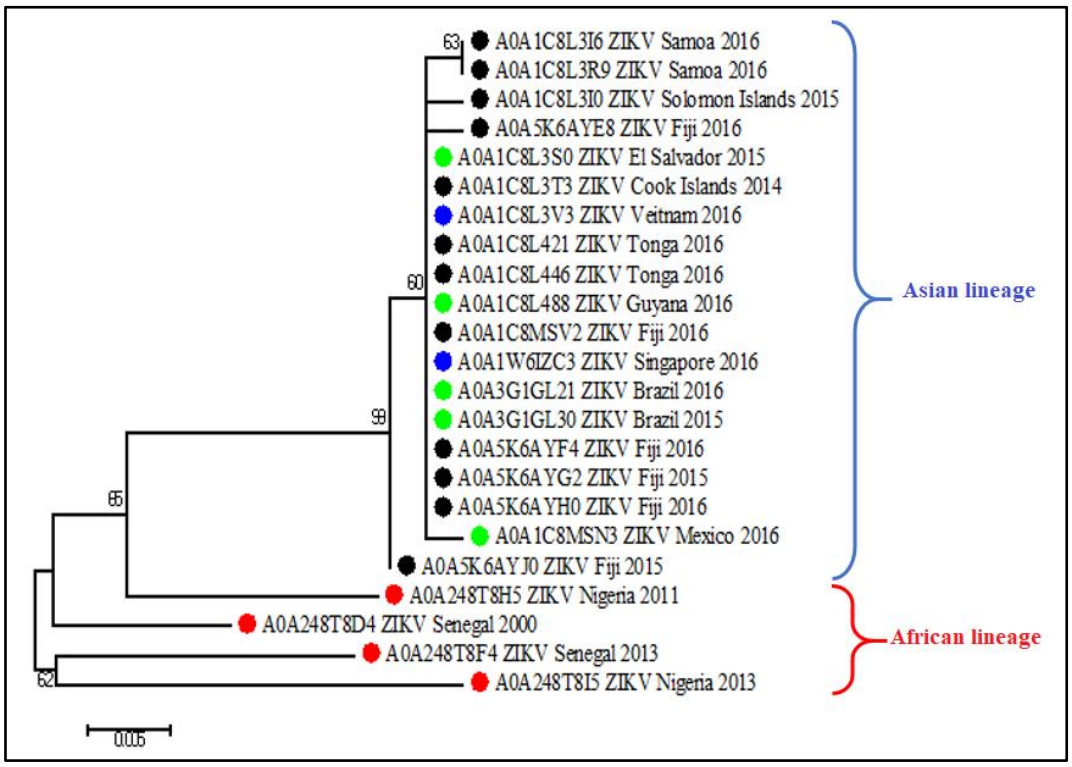


Figure 50. Phylogenetic tree of ZIKV Envelope proteins from different geographical origins. Isolates are represented with access number, country of origin, and year of isolation, in which the red color indicates the African countries, blue color for Asian countries, American countries in green color, and black color for Oceanian countries.

The evolutionary history was inferred using the Test Maximum Likelihood tree. The evolutionary distances were computed using the Pairwise Distances method. The analysis involved 23 amino acid sequences. Evolutionary analyses were conducted in MEGA6.

The phylogenetic tree revealed the presence of two major clusters (African and Asian). The first group is the African lineage includes Senegalese (A0A248T8F4) and Nigerian (A0A248T8I5) strains collected in the same year (2013), and in the same month (July), only five days separate their collection dates, from human tissue, these two strains surround from the same node in the tree, since Senegal country is located in west Africa, but Nigeria is central African country. In addition, two other African strains from Senegal (A0A248T8D4) isolated in 2000 and Nigeria (A0A248T8H5) isolated in 2011 are present above the first two in the cluster, whereas their branches don't start from the same node. This proximity can be explained by the collection of viral strains from the same countries, and for this reason it can be said that genetic similarities made them together in one group in the phylogenetic tree.

The second cluster includes countries from three continents (Asia, America, and Oceania): five islands of Oceania in the Pacific Ocean, four American countries, and two Asian countries. The two strains from Samoa island carrying A0A1C8L3I6, and A0A1C8L3R9, isolated in 2016 from the human tissue, descended from the same branch at the head of the group, because they belong to the same area. The six samples gathered from Fiji island between 2015 and 2016 are scattered on this cluster, one of them bears access number (A0A5K6AYE8), it took a place with Samoa and Solomon islands samples, another bears access number (A0A1C8MSV2) is located between Asian and American samples, three of them are situated successively which one bears access number A0A5K6AYG2 and their twin bears access number A0A5K6AYJ0 had the same collection date (11 July 2015) and isolated from saliva of human but are situated far from each other, the last one took a place at bottom of the cluster that explain genetic evolution. The rest of samples who bear access numbers respectively: A0A1C8L3S0, A0A1C8L3T3, A0A1C8L3V3, A0A1C8L421, A0A1C8L446, A0A1C8L488, A0A1C8MSV2, A0A1W6IZC3, A0A3G1GL21, A0A3G1GL30, A0A5K6AYF4, A0A5K6AYG2, and A0A5K6AYH0, descended from the same node in the second cluster, this is evidence that they are genetically similar, even though they are from different continents, this confirms what was reported by Charrel et al. (2014), the African lineage, which has shown no propensity to disseminate outside of Africa, and the Asian lineage, which continues to seed in previously unaffected regions of the world.

The two Brazilian samples (A0A3G1GL21, and A0A3G1GL30) fall between the Singaporean strain and Fijian strain, this confirms that the American ancestry is of Asian origin, as mentioned by Enfissi et al. (2016), were drawn the tree depending gene sequences of envelope between Suriname ZIKV and other ZIKV, and they found two lineages Asian and

African. The Suriname strains belong to the Asian genotype and seem to be most closely related to the strain that was circulating in French Polynesia in 2013, with which they share more than 99.7% and 99.9% of nucleotide and amino acid identity, respectively. Also Charrel et al. (2016) have drawn the tree discerns the phylogenetic relationships among selected Zika virus strains belonging to the African and Asian lineages based on complete genomic sequence Maximum Likelihood analysis Zika virus genomes from patients infected in Brazil and Suriname in 2015 are closely related to the strain that circulated in French Polynesia in 2013, with more than 99.7% and 99.9% level of nucleotide and amino acid identities, respectively.

The Mexican strains bearing the access number A0A1C8MSN3 descend from a node different from the above strains node but it falls between two Fijian strains A0A5K6AYH0 and A0A5K6AYH0, this suggests that there is a genetic relationship between the strains of ZIKV found in North America and the strains found in the island of Oceania. Also, we notice from the tree that all American strains are located near Fijian strains perhaps this convergence indicates that the movement of people from the islands of Oceania to America.

Moreover Faye et al. (2014) constructed, the trees for protein E, NS5 and the two concatenated genes by Maximum likelihood reinforced that ZIKV strains could be classified in three major clusters, African lineage is divided in two cluster MR766 prototype strain, and the Nigerian cluster. The Asian lineage alone represents the third cluster. On the other hand, Haddow et al. (2012) realized a tree by Neighbor-joining phylogeny generated from open reading frame nucleotide sequences of Zika virus strains, phylogenetic analyses of ZIKV sequences revealed the existence of two main virus lineages African and Asian.

2.4. Three-dimensional structure of envelope protein of Senegalese strain A0A248T8F4

The Senegalese strain A0A248T8F4, with non-available 3D protein E structure on UniProt, was selected to construct a predictive model of its envelope protein structure using SWISS-MODEL online tool. This strain is isolated in 10 July 2013, from human tissue (Harrera et al., 2017).

2.4.1. Alignment between the target Senegalese strain and the template.

The SWISS-MODEL proposes the nearest sequence 5iz7.2.1 as template. The template / target alignment result is given in Table 9. The alignment of amino-acids sequence of envelope protein between the two strains was done on UniProt using Clustal Omega tool.

```
5IZ7_1|Chains                 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A248T8F4|A0A248T8F4_ZIKV IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
                              ************************************************************

5IZ7_1|Chains                 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A248T8F4|A0A248T8F4_ZIKV YEASKSDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSFVTCAKFT 120
                              **** ********************************************:******:

5IZ7_1|Chains                 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A248T8F4|A0A248T8F4_ZIKV CSKKMTGKSIQPENLEYRIILSVRGSQHSGMIVNDIGHETDENRAKVEVTPNSPTAEATL 180
                              ******************:***:*********** ************:***** *****

5IZ7_1|Chains                 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A248T8F4|A0A248T8F4_ZIKV GGFGSLGLDCEPRTGLDFSDRYYLTMNNKHWLVHKEWFHDIPLPWHSGADTGTPHWNNKE 240
                              ******************** *************************:**************

5IZ7_1|Chains                 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A248T8F4|A0A248T8F4_ZIKV ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGPKGRLFSGHLKCRLIMDKLRL 300
                              ************************************* **** ******** ******

5IZ7_1|Chains                 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A248T8F4|A0A248T8F4_ZIKV KGVTYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
                              ***:************:********************************************

5IZ7_1|Chains                 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A248T8F4|A0A248T8F4_ZIKV ANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSIIGKAFEATVRGAKR 420
                              *********************************:************ ***************

5IZ7_1|Chains                 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A248T8F4|A0A248T8F4_ZIKV MAVLGDTAWDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLVWLGLNTK 480
                              ****************.::************************************:********

5IZ7_1|Chains                 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A248T8F4|A0A248T8F4_ZIKV NGSISLTCLALGGVMIFLSTAVSA 504
                              ****** ********:*********
```

Figure 51. Alignment result between the template 5iz7.2.1 and Senegalese strain A0A248T8F4.

There are 22 amino-acids mutations of envelope protein of Senegalese strain compared to the template 5iz7.2.1: 12 conservative mutations, 1 semi-conservative, and 9 non-conservatives demonstrated in the Table 9.

Table 9. Alignment between the template (5iz7.2.1) and Senegalese strain (A0A248T8F4).

Mutation nature	Amino-acids substituted	Position
Conservative	Leucine by Phenylalanine	L113F
Conservative	Alanine by Threonine	A120T
Conservative	Methionine by Isoleucine	M140I
Conservative	Histidine by Arginine	H144R
Conservative	Isoleucine by Valine	I169V
Conservative	Alanine by Serine	A127S
Conservative	Serine by Threonine	S304T
Conservative	Isoleucine by Valine	I317V
Conservative	Glutamic acid by Aspartic acid	E393D
Conservative	Leucine by Phenylalanine	L438F
Conservative	Methionine by Valine	M473V
Conservative	Leucine by Methionine	L494M
Semi-conservative	Alanine by Valine	A437V
Non-conservative	Isoleucine by Lysine	I65K
Non-conservative	Threonine by Isoleucine	T156I
Non-conservative	Arginine by Threonine	R175T
Non-conservative	Leucine by Arginine	L201R
Non-conservative	Lysine by Isoleucine	K294I
Non-conservative	Serine by Phenylalanine	S285F
Non-conservative	Alanine by Proline	A280P
Non-conservative	Threonine by Isoleucine	T406I
Non-conservative	Methionine by Threonine	M487T

2.4.2. The three-dimensional structure of envelope protein view

The name of the template in the SWISS-MODEL library is (5iz7.2.1), this one is a strain of Zika virus isolated, in French Polynesia in 2013, bearing the code H/FP/2013 during its big outbreak. The structure of ZIKV (5iz7.2.1) envelope protein is found by cryo-electron microscopy (cryoEM) at the 3.7 Å of resolution (Kostyuchenko and al., 2016).

The sequence of ZIKV envelope protein of Senegalese strain (A0A248T8F4), is completely covered by the template (5iz7.2.1), because the rectangle located below word of coverage is completely colored with bleu (Figure 52). Identity shows how many amino acids are exactly aligned here, there is 95.63%.

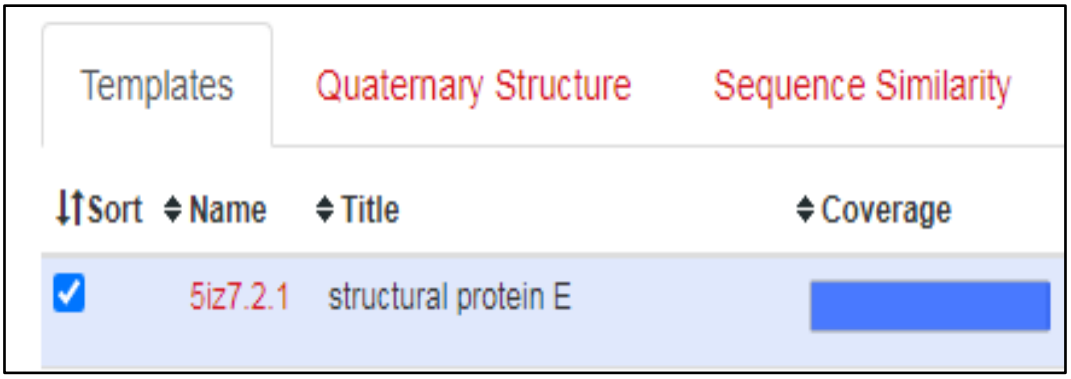


Figure 52. Coverage between the template (5iz7.2.1) and Senegalese strain (A0A248T8F4).

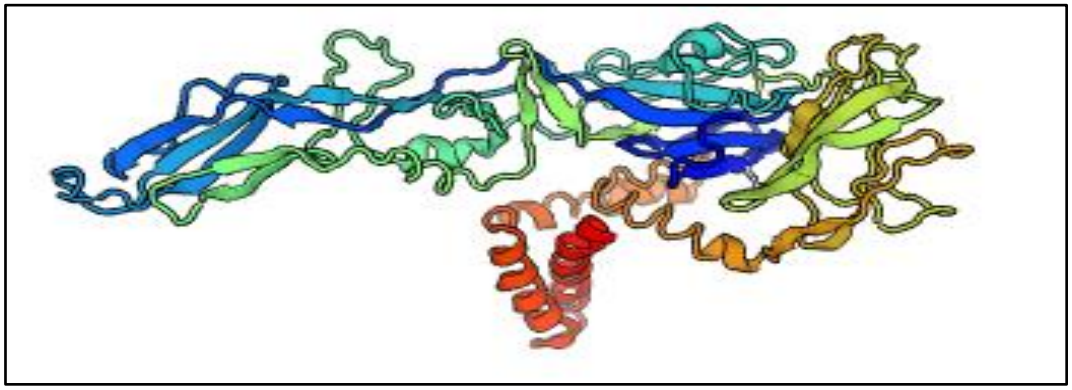


Figure 53. The three-dimensional structure of template 5iz7.2.1 E protein.

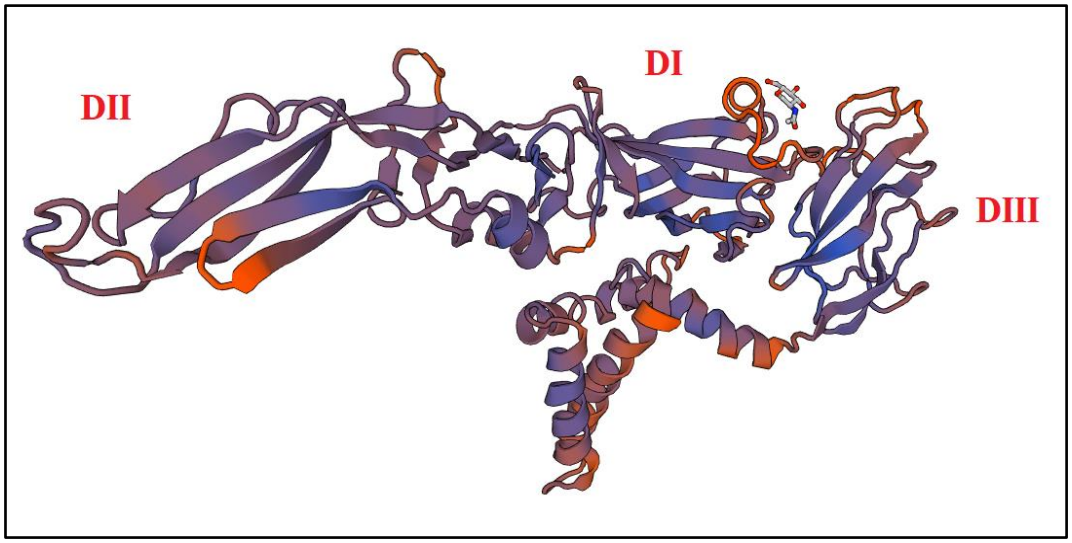


Figure 54. The predicted three-dimensional structure of E protein Senegalese strain (A0A248T8F4).

The general structure of envelope protein of A0A248T8F4 is similar to the structure of template, and almost similar to the model of E protein of ZIKV constructed by Alam et al. (2016) in the server I-TASSER , which depends to the structure of envelope protein of Dengue virus type I and type II finding by cryo-electron microscopy, the side view of the protein gives us two parts one part in ribbon form and the other part like the letter v.

This E protein is monomer consist of 504 aa owns three domains lying on straightening the domain I is located in the center between domain II and III depending on the data of the template with underlying regions form the transmembrane regions possesses two α-helices. Barba-Spaith et al. (2016) showed that the E protein is about 500 amino acids long, with three domains, named I, II and III, aligned in a row with domain I at the center, At the C terminus, the E ectodomain is followed by the stem, featuring two α-helices lying flat on the viral membrane (the stem helices). Almansour et al. (2019) each E protein monomer consists of three domains: DI, DII, and DIII, which undergo major rearrangements during the virus maturation cycle DI, which connects DII to DIII, is essential for the conformational changes required for viral entry into cells.

The residue from 98-109 form a random coils in domain II of Senegalese strain without a presence of cd loop, but Dai et al. (2016) stated that the hydrophobic cd loop represents the fusion loop (residues 98–109), which is responsible for the membrane fusion between host cell and virus membranes during virus entry, and is highly conserved in flaviviruses.

This monomer contains one site of glycosylation in Asparagine number 154 it's N-acetyl glucosamine (Figure 55) like the template, it is an indicator that Senegalese strain resemble to the other strains which allows to say that the envelope protein of A0A248T8F4 is a glycoprotein, the glycosylation sequence motif is [NDI] at positions 154-156. Lin et al. (2018) have demonstrated that protein E with a single glycosylation site at residue Asp154. Dai et al. (2016), showed that protein E contains the potential N154 glycosylation site. Fontes- Garfias et al. (2017) have demonstrated that E protein of most flaviviruses is post-translationally modified by N-linked glycosylation at amino acid 153/154 within a highly conserved glycosylation motif of N-X-T/S at positions154-156.

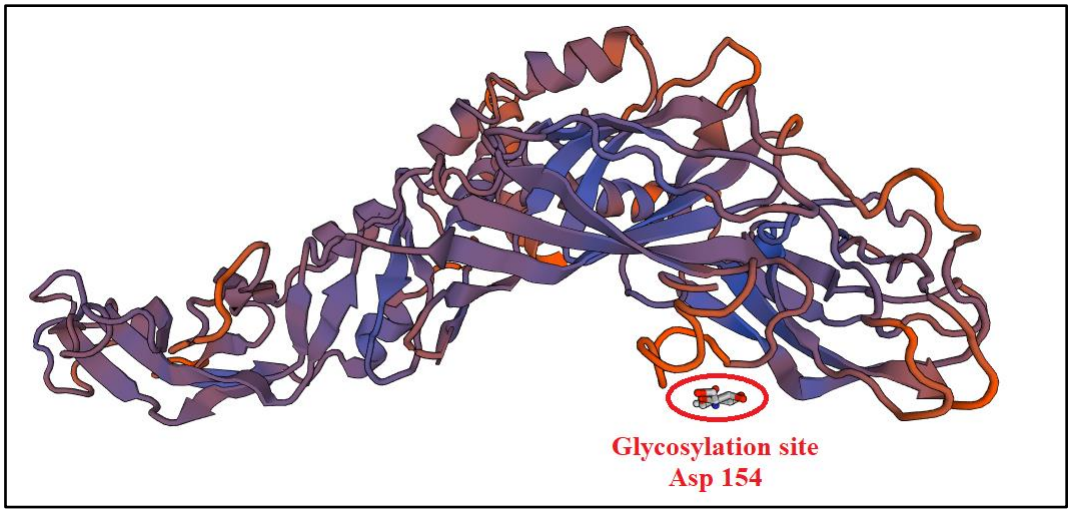


Figure 55. The site of glycosylation.

Domain I of envelope protein of Senegalese strain contains 10 β-sheets and 0 α-helices, contrariwise the domain I of envelope protein of template contains 15 β-sheets and 1 α-helices, in which the sequence is [TGHE] bearing number [156-159], the absence of this helice on Senegalese strain can be explained by the non-conservative mutation occurred between Threonine and Isoleucine in position 156. Domain II of A0A248T8F4 has 16 β-sheets and 4 α-helices, but the Polynesian template has 15 β-sheets and 1 α-helice. Domain III of envelope protein of Senegalese strain owns 11 β-sheets and 7 α-helices, when the template has 13 β-sheets and 6 α-helices. The total sum is 37 β-sheets and 11 α-helices for Senegalese strains, and the rest of the amino-acids made up of random coils. It means that the large percentage has taken by random coils, then β-sheets, and the small percentage for α-helices. The number of amino-acids cysteine is 13 is the same in both strains and also in the same positions, which explains the same number of disulfide bridges, keeping the tertiary structure. But Alam and al. (2016) ZIKV envelope glycoprotein is 504 residues long, of which 180 residues (35,7%) form sheet, 61 residues form turn and 305 residues (60,5%) form helix regions of the protein.

Conclusion

Conclusion

In the present study, we gathered the amino-acids sequences of the ZIKV envelope protein, isolated from different regions in the world using the bioinformatics resources of UniProt, to construct a phylogenetic tree by MEGA6 software. Also, we have made a three-dimensional model for envelope protein of Senegalese strain bearing the access number A0A248T8F4 using SWISS-MODEL, based on the structure of a Polynesian strain analyzed by electron microscopy.

Concerning the phylogenetic tree realized with 23 strains isolated from four continents: Asia, Africa, America, and Oceania we noticed that there are two principal lineages Asian and African, where the Asian lineage extends to the American continent and Oceania. Whereas, the African lineage remained confined in the tropical countries.

The phylogenetic tree shows the relationships of viral strains, and facilitates the determination of the epidemic focus of the virus, which creates a clear path for epidemiological studies. The envelope is the carrier for the determinants of virus-host cell recognition in enveloped viruses. These glycoproteins allow the virus to recognize the target cell. The NMR and x-ray crystallography structural biology techniques of structural biology allow building the 3D structure of the envelopes, but they are expensive and take long time. Thus, homology modeling allows building the 3D structure in less time and less efforts.

The 3D model of envelope protein of Senegalese strain built by SWISS-MODEL, showed that the 3D structure of protein E consists of 504 amino-acids forming a rubal monomer on the surface of the virus and an elongated part downwards builds the transmembrane regions, this monomer consists of three domains, domain I is located in the center with a glycosylation site at Asparagine number 154, which makes it possible to say the envelope protein is a glycoprotein. The absence of fusion loop in position 98-109, replaced by random coils is observed.

As perspectives, we think in the future to focus our study on the 3D structure of the NS5 protein, as the most important protein component of the ZIKV replication complex, it plays a key role in the life cycle and survival of the virus through its domains to develop and accelerate anti-viral drugs against ZIKV.

To complement high-throughput screening efforts, we could perform virtual screening against the proteins in ZIKV. While there are crystal structures for proteins from dengue, yellow fever, West Nile virus and other flaviviruses there are (to date) none for ZIKV. Therefore, we are limited to generating homology models, although the close evolutionary

relationships between flavivirus and their component proteins and genomes represents a valid approach.

Understanding the three-dimensional structure of antigenic ZIKV proteins may help in accelerating the development of antibodies for diagnostics and rationally designed vaccines. In addition, the comparison of the assembled surface glycoprotein of ZIKV with that of dengue virus may help understand the accessible epitopes for the development of anti-flaviviral vaccines in general.

List of references

List of references

- A -

- Alam A, Ali S, Ahamad S, Malik MZ and Ishrat R (2016) From ZIKV genome to vaccine: in silico approach for the epitope-based peptide vaccine against Zika virus envelope glycoprotein. Immunology. 149(4):386-399.
- Almansour I, Alfares R and Aljofi H (2019) Large-scale analysis of B-cell epitopes of envelope: Implications for Zika vaccine and immunotherapeutic development. F1000Research. 2019 **7**:1624
- Amorim L (2019) Zika virus, in: Hua Shan and Roger Y. Dodd (Eds) Blood Safety, Springer, Cham. 163-186. https://doi.org/10.1007/978-3-319-94436-4_9.
- Ashfaq UA and Ahmed B (2016) De Novo Structural Modeling and Conserved Epitopes Prediction of Zika Virus Envelop Protein for Vaccine Development. Viral Immunology. 29(7):436–443. doi:10.1089/vim.2016.0033.

- B -

- Barba-Spaeth G, Dejnirattisai W, Rouvinski A,Vaney MC, Medits I, Sharma A, Simon-Lorière E, Sakuntabhai A, Cao-Lormeau VM, HaouzA, England P, Stiasny K, Mongkolsapaya J, Heinz FX, Screaton GR and Rey FA (2016) Structural basis of potent Zika–dengue virus antibody cross-neutralization. Nature. 536(7614) : 48-53. doi:10.1038/nature18938
- Baronti C, Piorkowski G, Charrel RN, Boubis L, Leparc-Goffart I and de Lamballerie X (2014) Complete coding sequence of Zika virus from a French Polynesia outbreak in 2013. Genome Announcements. 2(3): e00500-14. doi:10.1128/genomeA.00500-14.
- Bernatchez JA, Tran LT, Li J, Luan Y, Siqueira-Neto JL and Li R (2019) Drugs for the Treatment of Zika Virus Infection. Journal of Medicinal Chemistry. 63(2):470–489 https://doi.org/10.1021/acs.jmedchem.9b00775.
- Bhatnagar J, Rabeneck D, Martines RB, Steiner S, Ermias Y, Estetter L, Suzuki T, Ritter J, Keating M, Hale G, J Gary, Muehlenbachs A, Lambert A, Lanciotti R, Oduyebo T, Meaney-Delman D, Bolaños F, Saad E, Shieh W and Zaki S (2017) Zika Virus RNA Replication and Persistence in Brain and Placental Tissue. Emerging infectious diseases. 23(3): 405–414.
- Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L, Schwede T (2017) The SWISS-MODEL Repository - new features and functionality Nucleic Acids Res. 45(D1): D313-D319.

- C -

- Campos GS, Bandeira AC and Sardi SI (2015) Zika virus outbreak, Bahia, Brazil Emerging. Infectious. Diseases. 21(10): 1885-1886.
- Carneiro LAM and Travassos LH (2016) Autophagy and viral diseases transmitted by Aedes aegypti and Aedes albopictus. Microbes and Infection. 18:169–171.
- Chambers MT, Schwarz MC, Sourisseau M, Gray ES, and Evans MJ (2018) Probing Zika virus neutralization determinants with glycoprotein mutants bearing linear epitope insertions. Journal of Virology. 92 (18) e00505-18; DOI: 10.1128/JVI.00505-18.
- Charrel RNa, Leparc-Goffart I, Pas S, de Lamballerie X, Koopmans M and Reusken Ch (2016) Background review for diagnostic test development for Zika virus infection. Bull World Health Organ. 94(8): 574–584.

- Coffey LL, Pesavento PA, Keesler RI, Singapuri A, Watanabe J et al (2017) Zika Virus Tissue and Blood Compartmentalization in Acute Infection of rhesus macaques. ploS one. 12(1). e0171148.

- 𝒟 -

- Dai L, Song J, Lu X, Deng YQ, Musyoki AM, Cheng H, Xiao H, and al. (2016) Structures of the Zika Virus Envelope Protein and Its Complex with a Flavivirus Broadly Protective Antibody. Cell Host and Microbe. 19(5): 696–704. doi:10.1016/j.chom.2016.04.013.
- Darwish MA, Hoogstraal H, Roberts TJ, Ahmed IP and Omar F (1983) A sero-epidemiological survey for certain arboviruses (Togaviridae) in Pakistan. Transactions of The Royal Society of Tropical Medicine and Hygiene. 77(4) 442-445.
- Dupont-Rouzeyrol M, O'Connor O, Calvez E, Daures M, John M, Grangeon JP and Gourinat AC (2015) Co-infection with Zika and dengue viruses in 2 patients, New Caledonia, 2014. Emerging Infectious Diseases. 21 (2)381-382.

- ℰ -

- Ekins S, Liebler J, Neves BJ, Lewis WG, Coffee M, Bienstock R, Southan Ch and Andrade CH (2016) Illustrating and homology modeling the proteins of the Zika virus. F1000Research. 5:275.
- Enfissi A, Codrington J, Roosblad J, Kazanji M and Rousset D (2016). Zika virus genome from the Americas. The lancet. 387(10015): 227-228.
- Esteves E, Rosa N, Correia MJ, Arrais JP and Barros M (2017) New Targets for Zika Virus Determined by Human-Viral Interactomic: A Bioinformatics Approach. Biomed Research International. 2017:1-15.

- ℱ -

- Fauci AS & Morens DM (2016) Zika Virus in the Americas — Yet Another Arbovirus Threat. New England Journal of Medicine. 374(7): 601–604. doi:10.1056/nejmp1600297
- Faye O, Freire CCM, Iamarino A, Faye O, de Oliveira JVC, Diallo M, Zanotto PMA and Sall AA (2014) Molecular Evolution of Zika Virus during Its Emergence in the 20th Century. PLoS Negl Trop Dis 8(1): e2636. doi:10.1371/journal.pntd.0002636.
- Fontes-Garfias CR, Shan C, Luo H, Muruato AE, Medeiros DBA, Mays E and al (2017) Functional Analysis of Glycosylation of Zika Virus Envelope Protein. Cell Reports. 21(5):1180–1190.

- 𝒢 -

- Garcez PP, Loiola EC, Madeiro da Costa RM, Higa LM, and Trindade P et al (2016) Zika virus impairs growth in human neurospheres and brain organoids. Science. 352(6287): 816-818.
- Geser A, Henderson BE and Christensen S (1970) A multipurpose serological survey in Kenya:2. Results of arbovirus serological tests. Bulletin of the World Health Organization. 43 (4):539-552.
- Gorshkov K, Shiryaev SA, Fertel S, Lin YW , Huang CT , Pinto A, Farhy Ch, Strongin AY Zheng W and Terskikh AV (2019) Zika Virus: Origins, Pathological Action, and Treatment Strategies. Frontiers in Microbiology. 9(3252).
- Gratton, R., Agrelli, A., Tricarico PM, Brandão L and Crovella, S (2019) Autophagy in Zika Virus Infection: A Possible Therapeutic Target to Counteract Viral Replication. International Journal of Molecular Sciences. 20(5):1048. doi:10.3390/ijms20051048.

- Gubler DJ, Vasilakis N and Musso D (2017) History and Emergence of Zika Virus. The Journal of Infectious Diseases. 216: 860–867. doi:10.1093/infdis/jix451.

- H -

- Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V, Huy R, Guzman H, Tesh RB, and Weaver S (2012) Genetic Characterization of Zika Virus Strains: Geographic Expansion of the Asian Lineage. Plos neglected tropical diseases. 6(2) e1477. https://doi.org/10.1371/journal.pntd.0001477.
- Hamel R., Dejarnac O, Wichit S, Ekchariyawat P, Neyret A, Luplertlop N, Perera-Lecoin M, Surasombatpattana P, Talignani L, Thomas F, Cao-Lormeau VM, Choumet V, Briant L, Desprès Ph, Amara A, Yssel H, Misséa D (2015) Biology of Zika Virus Infection in Human Skin Cells. Journal of Virology. 89:8880-8896.
- Heang V, Yasuda CY, Sovann L, Haddow AD, Travassos da Rosa AP, Tesh RB and KasperMR (2012) Zika virus infection, Cambodia, 2010. Emerging Infectious Diseases. 18 (2): 349-351.https://doi.org/10.1038/nature18938.
- https://www.ebi.ac.uk/uniprot/index), 2020.
- https://www.uniprot.org/taxonomy/64320.
- https://www.who.int/

- J -

- Jan C, Languillat G, Renaudet J and Robin Y (1978) A serological survey of arboviruses in Gabon. Bulletin de la Société de Pathologie Exotique et de ses filiales. 71(2) : 140-146.
- Javed F, Manzoor KN, Ali M, Haq IU, Khan AA, Zaib A and Manzoor S (2017) Zika virus: what we need to know? Journal of Basic Microbiology. 58(1):3–16. doi:10.1002/jobm.201700398.

- K -

- Kostyuchenko VA, Lim EXY, Zhang S, Fibriansah G, Ng T-S, Ooi JSG and al (2016) Structure of the thermally stable Zika virus. Nature. 533(7603): 425-428. doi:10.1038/nature17994.
- KutsunaS, Kato Y, Takasaki T , Moi ML, Kotaki A, Uemura H, Matono T, Fujiya Y, Mawatari M , Takeshita N , Hayakawa K , Kanagawa S, Ohmagari N (2014) Two cases of Zika fever imported from French Polynesia to Japan, December 2013 to January 2014 .Euro surveillance. 19(4) 4):pii2068.

- L -

- Lin H-H, Yip B-S, Huang L-M, and Wu S-C (2018) Zika virus structural biology and progress in vaccine development. Biotechnology Advances. 36(1): 47–53. doi:10.1016/j.biotechadv.2017.09.004.

- M -

- Marchette NJ, Garcia R and Rudnick A (1969) Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia. The American Journal of Tropical Medicine and Hygiene .18 (3):411-415.
- Miner JJ, Sene A, Richner JM, Smith AM, Santeford A, Ban N, Weger-Lucarelli J, Manzella F, Rückert C, Govero J, Noguchi KK, Ebel GD, Diamond MS and Apte RS

(2016) zika virus infection in mice causes panuveitis with shedding of virus in tears. Cell reports. 16 (12): 3208-3218.

- Mohd Ropidi MI, Khazali AS, Nor Rashid N and Yusof R (2020) Endoplasmic reticulum: a focal point of Zika virus infection. Journal of Biomedical Science. 27(1). doi:10.1186/s12929-020-0618-6.
- Monlun E, Zeller H, Le Guenno B, Traore-Lamizana M Hervy JP, Adam F Ferrara L, Fontenille D, Sylla R and Mondo M (1993) Surveillance of the circulation of arbovirus of medical interest in the region of eastern Senegal. Bulletin de la Societe de Pathologie Exotique (1990). 86(1):21-28.
- Moura da Silva AA, Ganz JSS, Da Silva Sousa PD, Doriqui MJ, Ribeiro MR, Branco MD, Queiroz RC, Pacheco MJ, Vieira da Costa FR, Silva FS, Simões VM, Pacheco MA, Lamy-Filho F, Lamy ZC and Alves M (2016) Early Growth and Neurologic Outcomes of Infants with Probable Congenital Zika Virus Syndrome. Emerging infectious diseases. 22(11): 1953–1956. doi:10.3201/eid2211.160956.
- Musso D and Gubler DJ (2016) Zika Virus. Clinical Microbiology Reviews. 29(3): 487–524.
- Musso D, Nilles EJ, Cao-Lormeau VM (2014) Rapid spread of emerging Zika virus in the Pacific area. Clinical Microbiology and Infection. 20 (10): 595-596.

- O -

- Olagnier D, Muscolini M, Coyne CB, Diamond MS and Hiscott J (2016) Mechanisms of Zika Virus Infection and Neuropathogenesis. DNA and cell Biology.35(8):367-372.

- P -

- Petersen LR, Jamieson DJ, Powers AM and Honein MA (2016) Zika Virus. the new England journal of medicine.374(16):1552-1563.
- Pielnaa P, Al-Saadawe M, Saro A, Dama MF, Zhou M, Huang Y, Jufang Huang Xia Z (2020) Zika virus-spread, epidemiology, genome, transmission cycle, clinical manifestation, associated challenges, vaccine and antiviral drug development. Virology. doi:10.1016/j.virol.2020.01.015
- Pundir S, Martin MJ, & O'Donovan C and The UniProt Consortium (2016) UniProt Tools. Current Protocols in Bioinformatics. 53 :1. 29.1-1.29.15.
- Pyke AT, Daly MT, Cameron JN, Moore PR, Taylor CT, Hewitson GR, Humphreys JL and Gair R (2014) Imported Zika virus infection from the Cook Islands into Australia, 2014. PLos Currents. 6.

- R -

- Ramharacka P and Soliman MES (2016) Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap. RSC Advances .6(73).
- Robin Y and Mouchet J (1975) Serological and entomological study on yellow fever in Sierra Leone. Bulletin de la Societe de Pathologie Exotique et de ses filiales. 68(3) : 249-258.
- Routhu NK and Byrareddy SN (2017) Host-Virus Interaction of ZIKA Virus in Modulating Disease Pathogenesis. Journal of Neuroimmune Pharmacology. 12(2): 219–232. doi:10.1007/s11481-017-9736-7.

- S -

- Saluzzo JF, Gonzalez JP, Herve JP and Georges AJ (1981) Serological survey for the prevalence of certain arboviruses in the human population of the south-east area of Central African Republic (author's transl). Bulletin de la Societe de Pathologie Exotique et de ses filiales. 74(5) : 490-499.
- Shapshak P, Somboonwit C, Foley BT, Alrabaa SF, Wills T & Sinnott JT (2015) Zika Virus. Global Virology I - Identifying and Investigating Viral Diseases. 477–500. doi:10.1007/978-1-4939-2410-3_18
- Shi Y and Gao G F (2017) Structural Biology of the Zika Virus. Trends in Biochemical Sciences. 42(6): 443–456. doi:10.1016/j.tibs.2017.02.009.
- Shi Y, Dai L, Song H and Gao G F (2018) Structures of Zika Virus E and NS1: Relations with Virus Infection and Host Immune Responses Dengue and Zika. Control and Antiviral Treatment Strategies.77–87. doi:10.1007/978-981-10-87271_6.
- Simmonds P, Becher P, Bukh J, Gould E A, Meyers G, Monath T, Muerhoff S, Pletnev A, Rico-Hesse R, Smith D B,Stapleton JT and ICTV Report Consortium (2017) ICTV virus taxonomy profile: Flaviviridae. Journal of General Virology. 98(1): 2–3.
- Sirohi D and Kuhn R J (2017) Zika Virus Structure, Maturation, and Receptors. The Journal of Infectious Diseases. 216(suppl_10):935944. doi:10.1093/infdis/jix515.
- Sirohi D, Chen Z, Sun L, Klose T, Pierson TC, Rossmann MG, Kuhn RJ (2016) The 3.8 Å resolution cryo-EM structure of Zika virus. Science. 352: 467–470.
- Smithburn KC, Kerr JA and Gatne PB (1954 a) Neutralizing antibodies against certain viruses in the sera of residents of India. The Journal of Immunology. 72 (4):248-257.
- Smithburn KC, Taylor RM, Rizk F, Kader A (1954 b) Immunity to certain arthropod-borne viruses among indigenous residents of Egypt. The American Journal of Tropical Medicine and Hygiene. 3(1): 9-18.
- Song BH, Yun SI, Woolley M and Lee YM (2017) Zika virus: History, epidemiology, transmission, and clinical presentation. Journal of Neuroimmunology. 308:50–64.
- Song H, Qi J, Haywood J, Shi Y and Gao GF (2016) Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. Nature Structural & Molecular Biology. 23(5):456–458. doi:10.1038/nsmb.3213.

- T -

- Tabata T, Petitt M, Puerta-Guardo H, Michlmayr D, Wang Ch, Fang-Hoover J, Harris E and Pereira L (2016) Zika virus targets different primary human placental cells, Suggesting Two Routes for Vertical Transmission. Cell host and microbes. 20(2): 155-166.
- Tamura K, Stecher G, Peterson D, Filipski A & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution. 30(12), 2725–2729. doi:10.1093/molbev/mst197.
- Tognarelli J, Ulloa S, Villagra E, Lagos J, Aguayo C, Fasce R, Parra B, Mora Becerra JN , Lagos N, Vera L, Olivares B, Vilches M and Fernandez J (2016) A report on the outbreak of Zika virus on Easter Island, South Pacific, 2014.Archives of Virology.161 (3): 665-668.

- W -

- Walker PJ, Siddell, SG, Lefkowitz E, Mushegian AR, Dempsey DM, Dutilh BE, Harrach B, Harrison RL, Hendrickson RC, Junglen S, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, Nibert M, Rubino L, Sabanadzovic S, Simmonds P, Varsani A, Zerbini FM and

Davison AJ (2019) Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses. Archives of Virology. 164(9): 2417–2429. https://doi.org/10.1007/s00705-019-04306-w

- Weaver SC, Costa F, Garcia-Blanco MA, Ko AI, Ribeiro GS, Saade G, Shi PY and Vasilakis N (2016) Zika Virus: History, Emergence, Biology, and Prospects for Control. Antiviral Research. 130: 69-80.
- Wikan N & Smith D (2016) Zika virus: history of a newly emerging arbovirus. The Lancet infectious diseases. 16(7): e119-e126. doi:10.1016/s1473-3099(16)30010-x.

- Y -

- Yun SI and Lee YM (2017) Zika virus: An emerging flavivirus. Journal of Microbiology. 55:204–219. doi:10.1007/s12275-017-7063-6

- Z -

- Zammarchi L, Stella G, Mantella A, Bartolozzi D, Tappe D, Gunther S, Oestereich L, Cadar D, Munoz-Fontela C, Bartoloni A and Schmidt-Chanasit J (2015) Zika virus infections imported to Italy: clinical, immunological and virological findings, and public health implications. Journal of Clinical Virology. 63: 32-35.
- Zanluca C and Santos CND (2016). Zika virus an overview. Microbes and Infection. 18 (5):295-301.
- Zhu Z, Chan J F-W, Tee K-M, Choi G K-Y, Lau S K-P, Woo P C-Y and al (2016) Comparative genomic analysis of pre-epidemic and epidemic Zika virus strains for virological factors potentially associated with the rapidly expanding epidemic. Emerging Microbes & Infections. 5(3): e22–e22. doi:10.1038/emi.2016.48.

Annexes

Annex 1. FASTA sequences

>tr|A0A1C8L3I0|A0A1C8L3I0_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNSKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L3I6|A0A1C8L3I6_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGARRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L3R9|A0A1C8L3R9_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGARRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L3S0|A0A1C8L3S0_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L3T3|A0A1C8L3T3_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L3V3|A0A1C8L3V3_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L421|A0A1C8L421_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L446|A0A1C8L446_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8L488|A0A1C8L488_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLIW
LGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8MSN3|A0A1C8MSN3_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHVGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1C8MSV2|A0A1C8MSV2_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A1W6IZC3|A0A1W6IZC3_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A3G1GL21|A0A3G1GL21_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

SGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDMASDSRCPTQGEAYL
DKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENLEYRIMLSVHGSQHS
GMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKE
WFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAK
GRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVD
MQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGA
KRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLN

>tr|A0A3G1GL30|A0A3G1GL30_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A5K6AYE8|A0A5K6AYE8_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSVS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A5K6AYF4|A0A5K6AYF4_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL

EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A5K6AYG2|A0A5K6AYG2_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A5K6AYH0|A0A5K6AYH0_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A5K6AYJ0|A0A5K6AYJ0_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFACSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEVTPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTD
GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGS
TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLM
WLGLNTKNGSISLMCLALGGVLIFLSTAVSA

>tr|A0A248T8D4|A0A248T8D4_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFTCSKKMTGKSIQPENL
EYRIILSVHGSQHSGMIVNDIGHETDENRAKVEVTPKTPRAEATLGGFGSLGLDCEPRTGLDFSDLYYLT
MNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTALA
GALEAEMDGAKGRLFSGQLKCRLKMDKLRLKGVSYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTDG
PCKVPAQMAVDMQTLSPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSII
GKAFEATVRGAKRMAVLGDTARDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWF

>tr|A0A248T8F4|A0A248T8F4_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASKSD
MASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSFVTCAKFTCSKKMTGKSIQPEN
LEYRIILSVRGSQHSGMIVNDIGHETDENRAKVEVTPNSPTAEATLGGFGSLGLDCEPRTGLDFSDRYYL
TMNNKHWLVHKEWFHDIPLPWHSGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGPKGRLFSGHLKCRLIMDKLRLKGVTYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTD

GPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGS
IIGKAFEATVRGAKRMAVLGDTAWDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLVW
LGLNTKNGSISLTCLALGGVMIFLSTAVSA

>tr|A0A248T8H5|A0A248T8H5_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFTCSKKMTGKSIQPENL
EYRIMLSVHGSQHSGMIVNDIGHETDENRAKVEVTPNSPRAEATLGGFASLGLDCEPRTGLDFSDLYYL
TMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTAL
AGALEAEMDGAKGRLSSGHLKCPLKMDKLRLKGVSYSLCTAAFTFTKVPAETLHGTVTVEVQNAGTD
GPCKVPAQMAVDMQTLSPVGRLITANPVITESTKNSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGS
IIGKAFEATVRGAKRMAVLGDTAWDFGSVGGVFNSLGKGVHQIFGAAFKSLFGGMSWYSQILIGTLLV
WLGLNTKNGSISLTCLALGGV

>tr|A0A248T8I5|A0A248T8I5_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDNPTVVIELVTTTVSNMAEVRSYCYEASISDM
ASDSRCPTQGEAYHDKQSDTQYVCKRTLVDRGWGNGSGLFGKGSLVTCAKFTCSKKMTGKSIQPENL
EYRIILSVHGSQHSGMIVNDTGHETDENRAKVEVTPNSPRSEATLGGFGSLGLDCEPRTGLDFSDLYYVT
MNNKHWLVHKEWFHDIPLPWHSGADTGTPHRNNKEALVEFKDAHAKTQTVVVLGSQEGAVHTALAG
ALEAEMDGAKGRLFSGHLKSRLKMDKLRLKGVSYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTDGP
CKVPAQMAVDMETLTPVGRLITANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSIIG
KAFEATVRGAKRMAVLGDTAWDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLVWLG

>tr|A0A2Z2CKG2|A0A2Z2CKG2_ZIKV Envelope protein (Fragment) OS=Zika virus OX=64320 PE=4 SV=1

GTCHHKKGEARRSRRAVTLPSHSTRKLQTRSQTWLESREYTKHLIRVENWIFRNPGFALAAAAIAWLL
GSSTSQKVIYLVMILLIAPAYSIRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVT
TTVSNMAEVRSYCYEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLV
TCAKFACSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATLGGF
GSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFKDAH
AKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRLKGVSYSLCAAAFTFT
KIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPPFG
DSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGA
AFKSLFGGMSWFSQILIGTLLVWLGLNTKNGSISLACLALGGVLIFLSTAVSADVGCSVDFSKKETRCGT
GVFVYNDVEA

Annex 2.

Amino acids, one and three letter codes

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine or aspartic acid	Asx	B
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glutamine or glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Annex 3. Multiple alignment of ZIKV envelope protein by Clustal Omega on Uniprot, May 2020.

```
TR|A0A1C8L3T3_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A3G1GL30_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1W6IZC3_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A5K6AYG2_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A5K6AYF4_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L3I0_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A248T8F4_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L3R9_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L3S0_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L421_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A5K6AYE8_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8MSN3_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A5K6AYJ0_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A5K6AYH0_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L446_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L3I6_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8MSV2_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L3V3_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A1C8L488_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A248T8H5_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
TR|A0A248T8I5_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDNPTVVIELVTTTVSNMAEVRSYC 60
TR|A0A3G1GL21_ZIKV 1 ---------------SGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 45
TR|A0A248T8D4_ZIKV 1 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVRSYC 60
                                  **********************:*** ******************

TR|A0A1C8L3T3_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A3G1GL30_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1W6IZC3_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A5K6AYG2_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A5K6AYF4_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L3I0_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A248T8F4_ZIKV 61 YEASKSDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSFVTCAKFT 120
TR|A0A1C8L3R9_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L3S0_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L421_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A5K6AYE8_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8MSN3_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A5K6AYJ0_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A5K6AYH0_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L446_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L3I6_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8MSV2_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L3V3_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A1C8L488_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 120
TR|A0A248T8H5_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFT 120
TR|A0A248T8I5_ZIKV 61 YEASISDMASDSRCPTQGEAYHDKQSDTQYVCKRTLVDRGWGNGSGLFGKGSLVTCAKFT 120
TR|A0A3G1GL21_ZIKV 46 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFA 105
TR|A0A248T8D4_ZIKV 61 YEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFT 120
                     **** ***************** **********************.*******:******:

TR|A0A1C8L3T3_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A3G1GL30_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1W6IZC3_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A5K6AYG2_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A5K6AYF4_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L3I0_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A248T8F4_ZIKV 121 CSKKMTGKSIQPENLEYRIILSVRGSQHSGMIVNDIGHETDENRAKVEVTPNSPTAEATL 180
TR|A0A1C8L3R9_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L3S0_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L421_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A5K6AYE8_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8MSN3_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A5K6AYJ0_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEVTPNSPRAEATL 180
TR|A0A5K6AYH0_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L446_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L3I6_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8MSV2_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L3V3_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A1C8L488_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 180
TR|A0A248T8H5_ZIKV 121 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDIGHETDENRAKVEVTPNSPRAEATL 180
TR|A0A248T8I5_ZIKV 121 CSKKMTGKSIQPENLEYRIILSVHGSQHSGMIVNDTGHETDENRAKVEVTPNSPRSEATL 180
TR|A0A3G1GL21_ZIKV 106 CSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIVNDTGHETDENRAKVEITPNSPRAEATL 165
```

```
TR|A0A248T8D4_ZIKV 121 CSKKMTGKSIQPENLEYRIILSVHGSQHSGMIVNDIGHETDENRAKVEVTPKTPRAEATL 180
                       ********************:***:*********** ************:**::* :****

TR|A0A1C8L3T3_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A3G1GL30_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1W6IZC3_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A5K6AYG2_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A5K6AYF4_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L3I0_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNSKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A248T8F4_ZIKV 181 GGFGSLGLDCEPRTGLDFSDRYYLTMNNKHWLVHKEWFHDIPLPWHSGADTGTPHWNNKE 240
TR|A0A1C8L3R9_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L3S0_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L421_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A5K6AYE8_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8MSN3_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHVGADTGTPHWNNKE 240
TR|A0A5K6AYJ0_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A5K6AYH0_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L446_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L3I6_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8MSV2_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L3V3_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A1C8L488_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A248T8H5_ZIKV 181 GGFASLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
TR|A0A248T8I5_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYVTMNNKHWLVHKEWFHDIPLPWHSGADTGTPHRNNKE 240
TR|A0A3G1GL21_ZIKV 166 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 225
TR|A0A248T8D4_ZIKV 181 GGFGSLGLDCEPRTGLDFSDLYYLTMNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKE 240
                       ***.**************** **:***.****************** ******** ****

TR|A0A1C8L3T3_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A3G1GL30_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1W6IZC3_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A5K6AYG2_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A5K6AYF4_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L3I0_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A248T8F4_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGPKGRLFSGHLKCRLIMDKLRL 300
TR|A0A1C8L3R9_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L3S0_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L421_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A5K6AYE8_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8MSN3_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A5K6AYJ0_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A5K6AYH0_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L446_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L3I6_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8MSV2_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L3V3_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A1C8L488_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 300
TR|A0A248T8H5_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCPLKMDKLRL 300
TR|A0A248T8I5_ZIKV 241 ALVEFKDAHAKTQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLFSGHLKSRLKMDKLRL 300
TR|A0A3G1GL21_ZIKV 226 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLSSGHLKCRLKMDKLRL 285
TR|A0A248T8D4_ZIKV 241 ALVEFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLFSGQLKCRLKMDKLRL 300
                       *********** ************************** **** **:**. * ******

TR|A0A1C8L3T3_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A3G1GL30_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1W6IZC3_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A5K6AYG2_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A5K6AYF4_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L3I0_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A248T8F4_ZIKV 301 KGVTYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L3R9_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L3S0_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L421_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A5K6AYE8_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8MSN3_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A5K6AYJ0_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A5K6AYH0_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L446_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L3I6_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8MSV2_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L3V3_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A1C8L488_ZIKV 301 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 360
TR|A0A248T8H5_ZIKV 301 KGVSYSLCTAAFTFTKVPAETLHGTVTVEVQNAGTDGPCKVPAQMAVDMQTLSPVGRLIT 360
TR|A0A248T8I5_ZIKV 301 KGVSYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMETLTPVGRLIT 360
TR|A0A3G1GL21_ZIKV 286 KGVSYSLCTAAFTFTKIPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRLIT 345
TR|A0A248T8D4_ZIKV 301 KGVSYSLCTAAFTFTKVPAETLHGTVTVEVQYAGTDGPCKVPAQMAVDMQTLSPVGRLIT 360
                       ***:************:************** *****************:**:******* 
```

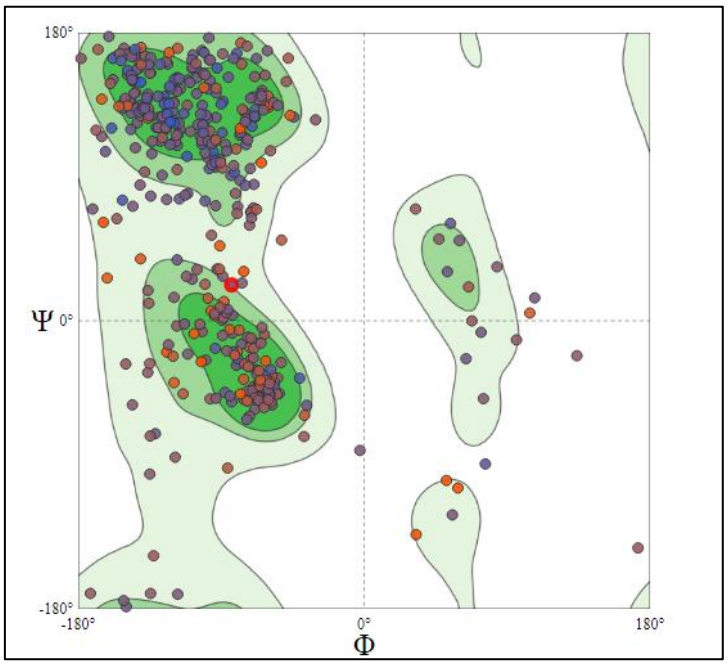
```
TR|A0A1C8L3T3_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A3G1GL30_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1W6IZC3_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A5K6AYG2_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A5K6AYF4_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1C8L3I0_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A248T8F4_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSIIGKAFEATVRGAKR 420
TR|A0A1C8L3R9_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGARR 420
TR|A0A1C8L3S0_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1C8L421_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A5K6AYE8_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSVSTIGKAFEATVRGAKR 420
TR|A0A1C8MSN3_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A5K6AYJ0_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A5K6AYH0_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1C8L446_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1C8L3I6_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGARR 420
TR|A0A1C8MSV2_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1C8L3V3_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A1C8L488_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 420
TR|A0A248T8H5_ZIKV 361 ANPVITESTKNSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSIIGKAFEATVRGAKR 420
TR|A0A248T8I5_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSIIGKAFEATVRGAKR 420
TR|A0A3G1GL21_ZIKV 346 ANPVITESTENSKMMLELDPPFGDSYIVIGVGEKKITHHWHRSGSTIGKAFEATVRGAKR 405
TR|A0A248T8D4_ZIKV 361 ANPVITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSIIGKAFEATVRGAKR 420
                        *********:**********************:********** * ************:*

TR|A0A1C8L3T3_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A3G1GL30_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1W6IZC3_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A5K6AYG2_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A5K6AYF4_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L3I0_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A248T8F4_ZIKV 421 MAVLGDTAWDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLVWLGLNTK 480
TR|A0A1C8L3R9_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L3S0_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L421_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A5K6AYE8_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8MSN3_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A5K6AYJ0_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A5K6AYH0_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L446_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L3I6_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8MSV2_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L3V3_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLNTK 480
TR|A0A1C8L488_ZIKV 421 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLIWLGLNTK 480
TR|A0A248T8H5_ZIKV 421 MAVLGDTAWDFGSVGGVFNSLGKGVHQIFGAAFKSLFGGMSWYSQILIGTLLVWLGLNTK 480
TR|A0A248T8I5_ZIKV 421 MAVLGDTAWDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLVWLG---- 476
TR|A0A3G1GL21_ZIKV 406 MAVLGDTAWDFGSVGGALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIGTLLMWLGLN-- 463
TR|A0A248T8D4_ZIKV 421 MAVLGDTARDFGSVGGVFNSLGKGIHQIFGAAFKSLFGGMSWF----------------- 463
                        ******** *******.:******:******************:

TR|A0A1C8L3T3_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A3G1GL30_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1W6IZC3_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A5K6AYG2_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A5K6AYF4_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L3I0_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A248T8F4_ZIKV 481 NGSISLTCLALGGVMIFLSTAVSA 504
TR|A0A1C8L3R9_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L3S0_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L421_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A5K6AYE8_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8MSN3_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A5K6AYJ0_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A5K6AYH0_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L446_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L3I6_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8MSV2_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L3V3_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A1C8L488_ZIKV 481 NGSISLMCLALGGVLIFLSTAVSA 504
TR|A0A248T8H5_ZIKV 481 NGSISLTCLALGGV---------- 494
TR|A0A248T8I5_ZIKV ------------------------
TR|A0A3G1GL21_ZIKV ------------------------
TR|A0A248T8D4_ZIKV ------------------------
```

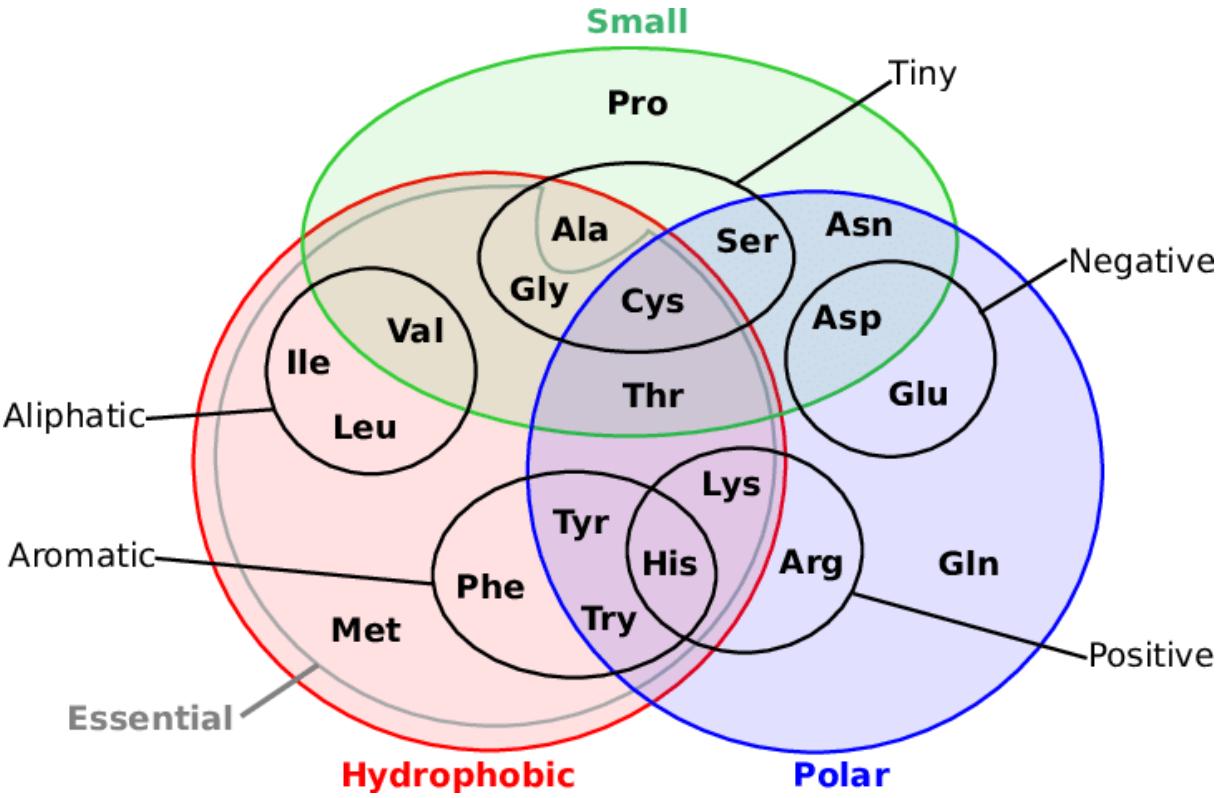
Annex 4.

The Ramachandran Plot



Annex 5.

Venn diagram of amino acids the properties.



Abstract

Zika virus (ZIKV) is a positive single-stranded RNA virus that is transmitted by mosquito bites. ZIKV envelope protein is antigenic and is involved in fusion and entry of viral particles into the cell. In this study, we have conducted a multiple sequence alignment of envelope protein structures collected from UniProt, using Clustal Omega tool on UniProt. Also, a phylogenetic analysis of Zika strains, isolated from several regions of Africa, Asia, America and Oceania was performed by MEGA 6.0 software. In addition, the three-dimensional structure of ZIKV envelope protein was predicted using SWISS-MODEL.

Key words: Zika virus, envelope protein, phylogenetic analysis, 3D protein structure prediction

Résumé

Le virus Zika (ZIKV) est un virus à ARN positif simple brin transmis par les piqûres de moustiques. La protéine d'enveloppe ZIKV est un antigène et est impliquée dans la fusion et l'entrée de particules virales dans la cellule. Dans cette étude, nous avons effectué plusieurs alignements de séquences de structures de protéines d'enveloppes collectées à partir d'UniProt, à l'aide de l'outil Clustal Omega sur UniProt. De plus, l'analyse phylogénétique des souches de Zika, isolées dans plusieurs régions d'Afrique, d'Asie, d'Amérique et d'Océanie, a été réalisée par le logiciel MEGA 6.0. De plus, la structure 3D de la protéine d'enveloppe ZIKV a été prédite en utilisant SWISS-MODEL.

Mots clés : virus Zika, protéine d'enveloppe, analyse phylogénétique, prédictions de la structure des protéines 3D

ملخص

فيروس زيكا (ZIKV) هو فيروس ايجابي RNA وحيد الخيط ينتقل عن طريق لدغات البعوض. بروتين غلاف ZIKV هو مستضد و يشارك في اندماج و دخول الجزيئات الفيروسية إلى الخلية. في هذه الدراسة، أجرينا محاذاة تسلسلية متعددة لهياكل بروتين الغلاف التي تم جمعها من UniProt ، باستخدام أداة Clustal Omega على UniProt . أيضا، تم إجراء تحليل phylogenetic لسلالات زيكا، المعزولة من عدة مناطق في إفريقيا و اسيا و أمريكا و أوقيانوسيا بواسطة برنامج MEGA.6.0 . بالإضافة إلى ذلك ، تم التنبؤ بالبنية ثلاثية الأبعاد لبروتين غلاف باستخدام SWISS-MODEL

الكلمات المفتاحية: فيروس زيكا ، بروتين الغلاف ، تحليل phylogenetic ، تنبؤ ببنية البروتين الثلاثية الأبعاد