DISSERTATION

EXTRINSIC INCUBATION TEMPERATURE IMPACTS ON ZIKA VIRUS

EVOLUTION AND VECTOR COMPETENCE DURING SYSTEMIC *AEDES* INFECTION

Submitted by

Reyes David A. Murrieta

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2020

Doctoral Committee:

    Advisor: Gregory Ebel

    Kenneth Olson
    Mark Stenglein
    Dan Sloan

ABSTRACT

EXTRINSIC INCUBATION TEMPERATURE IMPACTS ON ZIKA VIRUS

EVOLUTION AND VECTOR COMPETENCE DURING SYSTEMIC AEDES INFECTION

Arthropod-borne viruses (arboviruses) are distinctive in that they are required to constantly replicate in different hosts and in a wide range of temperatures for their perpetuation in nature. Vertebrate hosts tend to maintain temperatures of approximately 37°C - 40°C, but arthropods hosts are poikilotherms and subject to ambient temperatures which can have a daily temperature fluctuation of > 10°C.

Invertebrate host genus, species, and strain in combination with arbovirus strain and preparation methods are known to have large impacts on vector competence and vectorial capacity. Seemingly small differences in host geographic isolation, virus strain, and preparation methods can have significant impacts on vector competence studies. The role of temperature on the ability of an arthropod vector to acquire, maintain, and transmit a pathogen has been investigated for numerous arboviruses. Changing the extrinsic incubation temperature between distinct constant temperatures has been shown to alter arbovirus vector competence, extrinsic incubation period, and mosquito survival, in which moderate temperatures of 28°C-32°C are optimal and temperatures higher and lower have deleterious effects. The mean and range of daily temperature fluctuations (diurnal temperature) have likewise been shown to influence arbovirus perpetuation and vector competence, in which large daily temperature fluctuations negatively affect mosquito development, survival, and vector competence. However, little is known as to how temperature

alters arbovirus genetic diversity during systemic mosquito infection or how differences in arbovirus hosts and viral strains impact arbovirus genetic diversity in relationship to temperature.

Therefore in the study completed in chapter two, we characterized the impact that constant temperatures of 25°C, 28°C, 32°C, and 35°C, and the diurnal fluctuation from 25°C to 35°C during extrinsic incubation periods have on the Puerto Rican isolate of Zika virus (ZIKV) vector competence and population dynamics within *Aedes aegypti* (Poza Rica) and *Aedes albopictus* (Florida) mosquitoes. To characterize the impact that temperature has on ZIKV population diversity in different host species and viral isolates, in the study completed in chapter three, we used a Tapachula, Mexico *Aedes aegypti* line and a Chiapas, Mexico ZIKV isolate to assess ZIKV population dynamics during 20°C, 24°C, 28°C, 32°C, 34°C, and 36°C constant extrinsic incubation temperatures.

We found that vector competence varied in a unimodal manner for constant temperatures peaking between 28°C and 32°C for both *Aedes* species, while transmission peaked at 10 days post-infection for *Aedes aegypti* and 14 days post-infection in *Aedes albopictus*. The diurnal temperature group is not predicted by the constant temperature distribution. Instead, when using the mean daily temperature of the diurnal group as a predicter, its VC lies between the moderate (28°C and 32°C) and extreme (25°C and 35°C) temperature group VCs. Using RNA-seq to characterize ZIKV population structure, we identified that temperature alters the ZIKV selective environment during infection. During mosquito infection, constant temperatures more often elicited positive selection whereas diurnal temperatures led to strong purifying selection in both *Aedes* species.

These findings demonstrate that temperature has multiple impacts on ZIKV biology within mosquitoes and has distinct effects on the selective environment within mosquitoes. Additionally,

the selective pressures induced by temperature are consistent across host species and viral strain and have similar impacts on shaping the viral population structure. However, input viral populations are still a driving factor of diversity and expansion during systemic mosquito infection. While our findings and those of others suggest that vector competence is impacted unimodally regardless of temperature, this is only applicable for constant temperatures. Future work assessing daily temperature fluctuation range and mean are needed to have a clear understanding of the role extrinsic incubation temperature plays on vector competence.

ACKNOWLEDGEMENTS

Thank you for the great conversations and amazing book recommendations. Heidi Runge and Kate Sherrill, thank you both for keeping me on track in the MIP program and GAUSSI, without the two of you I am sure I would have missed far too many deadlines and events.

The Ebel lab, I am extremely thankful for the friendship of Dr. Nathan Grubaugh who always had the next great idea and was always willing to share a glass of whiskey and talk science. The mentorship from Dr. Claudia Rückert, Dr. James Weger-Lucarelli, Dr. Joseph Fauver, Dr. Nicole Sexton, and Dr. Emily Gallichotte - each of you provided unique advice and perspectives in advancing my science. Thank you for harsh critiques, the novel insights into method development, the expertise from years working in the field, and the motivation and tips to push projects to completion. Additionally, I want to thank all of my lab mates that I have worked with, Michael Young, you have been a great office mate and thanks for humoring all the times I played devil's advocate. Dr. Alex Byas, thank you reminding me that in pathology, blue is bad, or was it red? Dr. Dalit Talmi-Frank, thank you for all the great conversations in lab and geeking out with me about single-cell sequencing. Bekah McMinn, thanks for hitting the climbing gym/crag with me and keeping me in the right head space. Dr. Kendra Quicke, thanks for job advice: "Bear Down". And thanks to all the great undergraduates that I have had the pleasure of working with and mentoring, you all are going to do great things. Additionally, I want to thank my friends, peers, and anyone else I may have missed. I have too many to name and if I tried to, I would undoubtedly leave someone out unknowingly anyway. So [insert your name here], thank you for everything that you do, from walking around strange towns in Brazil (Hannah), to BJJ rolling on the mats (Megan), and breaking down on trails in Moab (Darcy & Erik), each of you have played a role in the completion of this work.

Last, I thank all of those who cannot be here to celebrate and enjoy this work. My best friend Daniel Wilcox III, I think of you often and live a better life for it. Kanamei, who brought many years of joy into our lives. My aunt Shannon, I wish you could enjoy this work with me, but I know that you would be proud, regardless of all the damage I may have caused to your home growing up. My Tia Rita, you brought so much joy to our family, I hope that I too can be a light for those who need it. My Tia Juanita, there is not a holiday meal that goes by, without me recalling the days I tried to sneak food before it was ready. I do not get caught sneaking food nearly as often anymore.

TABLE OF CONTENTS

Chapter 1: Overview of Literature

1.1 Historical perspective of arboviruses

Mosquitoes are the world's deadliest animal due to the devastating pathogens they transmit to hosts. Diseases spread by mosquitoes include malaria, dengue, West Nile, yellow fever, and the recent explosive epidemic of Zika.

For millennia, mosquitoes were overlooked as a critical component of disease spread to the host. Some of the earliest reports of arthropod-borne viruses (arboviruses) date back to 265-992 AD in China, and included clinical descriptions of dengue infection; identified then as a "water poison" [1]. Similar descriptions of illness were reported in the French West Indies and Panama in the late 1600s [1]. By the 1700-1800s dengue had spread globally, which was most likely due to commercial sailing [2]. However, it was not until the seminal work of a British Surgeon General, Sir Ronald Ross, who demonstrated that the malarial parasite *Plasmodium* was vectored by mosquitoes in 1897 [3], that the world seriously started considering mosquitoes as vectors of disease. In comparison to Sir Ronald Ross, who had an advantage of working with the known agent of malaria disease in his studies, Walter Reed had been dispatched to Cuba as a part of the Yellow Fever Commission to investigate the origins and transmission of yellow fever [4]. Walter Reed was well versed in Ross' research [3], and thus understood that mosquitoes were able to vector pathogens. Additionally he was well aware of the work of Carlos Finlay, a Cuban physician who proposed that mosquitoes were responsible for the yellow fever pestilence in 1881 [5]. Taking into account the work of Ross and Finlay, Walter Reed and Jesse William Lazear were able to characterize the spread of yellow fever by performing transmission experiments using mosquitoes obtained from Finlay, in which they had mosquitoes that had fed on a patient dying of yellow fever

also feed on William E. Dean and James Carrol, both individuals having subsequently contracted yellow fever [6]. Dean would successfully recover from yellow fever, but due to the severe attack of yellow fever Carrol would suffer of acute dilation of his heart, ultimately leading to his death in 1907 [7]. Unfortunately, Lazear would not be so fortunate, while visiting a local hospital to collect blood from a yellow fever patient, he was bitten by a mosquito presumed to be infected from feeding of yellow fever patients, twelve days later he would succumb to fatal yellow fever [6]. Shortly after these findings, in 1906 [8], the second arbovirus vector was identified, *Aedes aegypti*, as the vector of dengue.

Jumping forward to present day, yellow fever virus (YFV) and dengue virus (DENV) continue to impact public health regardless of the presence of a vaccine or not [9-12]. In addition to these two arboviruses, numerous other pathogenic arboviruses have emerged around the world, including West Nile virus (WNV), chikungunya virus (CHIKV), and Zika virus (ZIKV) [13-15]. In fact, to date we have identified over 500 different arboviruses [16, 17], and novel arboviruses continue to be identified [18]. The ongoing emergence and reemergence of arboviruses indicate that they are here to stay, and the large diversity of them are indeed what make the mosquito the world's deadliest animal.

1.2 Diversity of Flaviviruses

Flaviviruses (Family *Flaviviridae*; genus *Flavivirus*) include some of the most medically important arboviruses. With the emergence and reemergence of flaviviruses such as DENV, YFV, WNV, Powassan virus (POWV) and ZIKV it is easy to see why they are considered to be of medical importance [9, 12, 13, 19, 20]. DENV alone is known to infect an estimated 400 million humans each year, and over a quarter of the world's population live in DENV endemic areas [21].

2

Flaviviruses have a single stranded, positive sense, RNA genome that contain a single open reading frame [22]. The genus *Flavivirus* can be split into four groups of viruses based on their mode of transmission: those viruses that have no known vector (NKV), those that infect mosquitoes but not vertebrates (insect-specific flaviviruses, ISF), those that transmit to vertebrates by ticks (tick-borne flaviviruses, TBF), and by mosquitoes (mosquito-borne flaviviruses, MBF) [23, 24]. ISF are unique flaviviruses because they can only replicate in insects, an example being Nhumirim virus (NHUV) which can only replicate in mosquito cells [25]. While ISFs are not considered arboviruses, in the case of NHUV, they have been shown to play a role in moderating arbovirus acquisition and transmission within the mosquito [25-27]. With the recent characterization of ISV modulation of arboviral transmission, novel ISFs are being searched for and identified rapidly [28, 29]. There are 14 known species of NKV flavivirus that have been recognized, eight bat-associated and six rodent-associated species [30]. Of the NKV flaviviruses, the best characterized are the Rio Bravo and Modoc viruses which can only replicate in bats [31] and rodents [32], respectively. TBFs are subdivided by their host vector pairing. The mammalian tick-borne virus sub-group (M-TBF) includes some of the most pathogenic viruses: Tick-borne encephalitis virus, Powassan virus which can cause encephalitis, and Kyasanur Forest disease viruses which can cause hemorrhagic fever [33, 34]. The second sub-group of TBFs are the seabird tick-borne virus group (S-TBF) [33] and includes three species: Tyuleniy virus, Meaban virus, and Saumarez Reef virus. It is interesting to note that the Kadam virus was once grouped with both the M-TBF and S-TBF groups, but genetic distance analysis now suggests that Kadam virus makes up its own putative third group within S-TBF [33, 34]. The last flavivirus group is the MBF group which is complex and arguably the most relevant group regarding human health. Like the TBF group, the MBF group can be further sub-divided into two groups, those vectored by *Culex spp.* mosquitoes, which includes

WNV and Japanese encephalitis virus, and those that are vectored by *Aedes spp.* mosquitoes, which include DENV, YFV, and ZIKV.

Even within the above groups, there are further levels of diversity within flavivirus groups. For example, DENV has 4 serotypes (DENV-1, DENV-2, DENV-3, and DENV-4) that are globally circulating, cause similar disease, but are genetically distinct [35]. Another example of within group diversity was observed during the recent ZIKV outbreak in the Americas from 2015-2016 [15]. From previous studies, we have seen that ZIKV lineage diversity and spread across the Americas was complex and had multiple introductions into the Americas [36-40]. During this outbreak, two distinct genetic lineages of ZIKV were identified: the African and Asian lineages [41, 42]. The Asian lineage circulated during the recent ZIKV outbreak [43], however there is some evidence that there may have been an African lineage circulating in Americas as well [44]. In addition to the two principal African and Asian lineages, others have proposed that there is a third lineage of ZIKV, with conflicting views as to which lineage (Asian [20] or African [43, 44]) divided to create it. In either case, it is evident that there is a high level of diversity within ZIKV.

1.3 Global emergence of Zika virus

Over the course of history, viruses endemic to Africa have emerged as global pathogens. ZIKV is one such pathogen that was once endemic to just a small region of Uganda and was able to emerge as a global pathogen, leading to severe disease burden across Africa, Asia, and the Americas [45]. ZIKV was first discovered in the Zika Forrest of Uganda in 1947. A group of researchers including Alexander Haddow and George Dick were performing routine YFV surveillance on sentinel platforms when they collected the first known ZIKV sample from a rhesus macaque (Rhesus 766) [46]. Haddow and Dick then isolated ZIKV by infecting mice

intracerebrally and collecting a filterable transmissible agent from the cerebral tissue of the sick mice [46]. A second ZIKV isolate was collected in 1948 at the same location from *Aedes africanus* mosquitoes [46].

It wasn't until six years later that the first human cases were reported in three patients in Nigeria: a female age 10, and two males age 24 and 30 [47]. In 1964, the first reported case of *Aedes aegypti* mosquitoes vectoring ZIKV occurred [48] implicating a well known urban mosquito vector increasing the risk of Zika disease to human populations. In 1966, the first reported case of ZIKV was identified in Asia, a pool of adult *Aedes aegypti* collected from a small shop house in Bentong Malaysia tested positive for ZIKV using a hemagglutination-inhibition test [48]. This ZIKV isolate was shown to be a new lineage of ZIKV known as the Asian lineage [49]. Even though ZIKV was now known to be in Asia, it would be another 11 years and ~2,300 km further South before the first reported case of a human infected with ZIKV occurred in Asia, specifically from central Java in Indonesia [50]. In 1977 two males and one female tested seropositive for ZIKV at the Tegalyoso Hospital in central Java. The following year, three females and one male also tested seropositive for ZIKV here, indicating that ZIKV was endemic in central Java [50].

For the next 30 years, ZIKV transmission would be relatively silent, as other flaviviruses such as DENV would have a much bigger impact during these times due to the reintroduction of DENVs into China, the Caribbean, and global distribution of all 4 DENV serotypes [51]. In 2007, the first significant outbreak of ZIKV occurred with the emergence of ZIKV (Asian lineage) among humans on Yap Island. Until that point in time, ZIKV was considered to have little impact on human health with only sporadic cases of human infection [47, 50, 52-54]. However, in 2007, there were 108 confirmed or probable cases of ZIKV although epidemiologists estimated there may have been over 5,000 human cases of ZIKV on Yap Island during this outbreak [55]. The

ZIKV outbreak on Yap Island was the first indication for the epidemic potential of ZIKV. It is interesting to point out that while the Yap Island outbreak brought ZIKV to the forefront of public health concern, retrospective analysis of human and mosquito samples from Gabon in equatorial Africa showed that the African lineage of ZIKV was spreading in urban environments for the first time although in this study ZIKV was transmitted by *Aedes albopictus*, commonly called the Asian tiger mosquito [56].

The next major ZIKV outbreak occurred in 2013 in French Polynesia and other islands in Oceania. The interesting outcome during Oceania outbreaks was that ZIKV transmission could also occur through human blood or other body fluids, and that ZIKV infection was associated with neurological disorders in the form of Guillain-Barre Syndrome [57, 58]. The French Polynesian outbreak was the largest to date with estimated 28,000 cases [59]. At this point, the world could see the epidemic potential of ZIKV and since French Polynesia is a major tourist destination, epidemiologists waited to see if viremic humans and international travel could efficiently spread this disease across the globe.

In May of 2015, ZIKV cases started to be reported in Brazil [60]. However, retrospective analysis show that a French Polynesian strain of ZIKV was most likely introduced to Brazil in 2013 [36, 61], around the 2013 Confederations Cup in Recife [62]. From the time of the first confirmed case of ZIKV in Brazil [60], the virus took less than a year to spread to neighboring states, regions, and other countries in South and Central America. During this outbreak, neurological symptoms of microcephaly were associated with ZIKV infection [63-65]. ZIKV continued to spread from Brazil into other countries in the Americas [38, 66, 67]. While ZIKV cases in the Americas started to wane in 2017 [68], there are still sporadic cases of ZIKV being reported [69-71], suggesting that ZIKV transmission in the Americas may be ongoing.

1.4 Temperature impacts on mosquito development and life history traits

Environmental temperature is one of the most important physical factors to impact mosquito development, behavior, and distribution. Since mosquitoes are poikilotherms, they must use behavioral strategies to modulate their body temperature to decrease the threat of thermal stress [72, 73]. Environmental temperature fluctuations expose mosquitoes to thermal stress, put them at risk of desiccation [74], impact development and reproduction [75], decrease mobility [76], and cause other temperature related impacts. However, as mosquitoes are abundant in all continents with the exception of Antarctica, clearly, they found a way to adapt and overcome the stressors of environmental temperatures [13]. Adaptation may occur through expression of heat shock proteins [77], behavioral modifications, or thermoregulation [76] but is clear that mosquitoes species can survive and thrive in varying climates.

The effect that temperature has on *Aedes* development has been well characterized [75]. *Aedes* mosquitoes have four life cycle stages (Figure 1.1), (1) eggs are laid by adult females on a damp surface near a water line. Eggs can survive for up to eight months and are essential in overwintering. When water levels rise due to rain or flooding and covers the eggs, (2) larvae emerge and feed on microorganisms in the water, molting three time until they develop into (3) pupae. Pupae continue to develop until a (4) adult mosquito emerges (Figure 1.1) [78]. Two to three days post emergence, the adult reproductive organs are developed, and males are able to fertilize the females. Females subsequently seek a blood meal, biting humans and using the nutrient rich blood meal to develop and lay eggs. Depending on the life stage, temperature impacts development in slightly different ways. Knowing the minimum and maximum temperatures necessary for efficient *Aedes* development is essential for predicting the spread of the *Aedes* mosquito geographic range. Estimating the temperature ranges that are optimal for *Aedes*

development, along with other abiotic factors such as rainfall, humidity, photoperiod [79], allows us to more accurately predict the spread of *Aedes* and the emergence of *Aedes*-borne disease [80, 81]. One study estimates that the minimum temperature of development for *Aedes aegypti* eggs is 14°C, larvae is 11.8°C, and pupae is 10.3°C [82]. Another study shows a minimum threshold for development of 16°C [83]. The maximum temperature for development is estimated to be between 34°C and 42°C [82-84] with ranges of 36°C -38°C for eggs, 36°C -42°C for larvae and 38°C - 42°C for pupae [82], with optimal temperature of development around 32°C [84]. Interestingly, it has been shown that *Aedes albopictus* can survive and develop in a wider range of temperatures, with a minimum developmental temperature of 10.4°C, optimal temperature of 29.7°C and a maximum temperature ranging from 35°C - 40°C [85]. Likewise, there are differences in thermal limits depending on geographic regions of isolation within the same mosquito species [85, 86], suggesting that there is an adaptation to the environment that modulates thermal regulation. It is important to mention that it is not just the average temperature that plays a role in development. As mosquitoes are exposed to temperature fluctuations in nature, it has been shown that large diurnal temperature ranges increase development time and lower larval survival, while areas with small diurnal temperature range may shorten the development time [87].
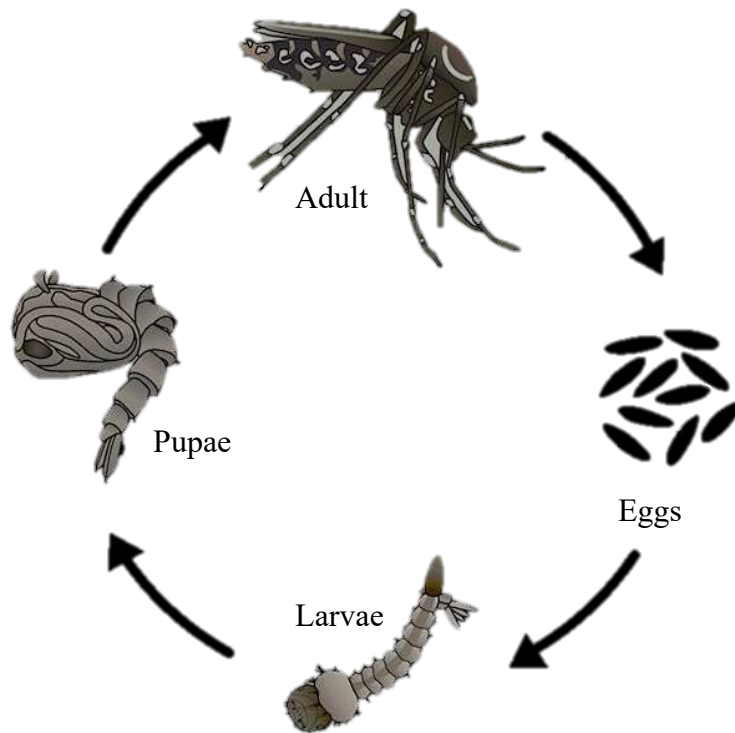
Figure 1.1 *Aedes* life cycle. Female mosquitoes lay eggs in a container that holds water. When eggs are covered with water they hatch into larvae. Larvae feed on microorganism in the water and develop into pupae which continue to develop until an adult emerges. Modified from Centers of Disease Control and Prevention [78].

Both *Aedes aegypti* and *Aedes albopictus* are anthropophilic day-biting mosquitoes that live in close proximity to humans [83, 88, 89]. The main difference between the two is that *Aedes aegypti* is endophagic and endophilic, preferring to feed on humans [83, 88]. In contrast, *Aedes albopictus* tends to be exophagic and exophilic, in that while they prefer to feed on humans, they are also opportunistic feeders, feeding from cold and warm-blooded animals when available [89, 90]. While both mosquito species have their preferred living and feeding environment, either inside or outside of dwellings, both species can be found in either environment and therefore exposed to daily and seasonal temperature fluctuations. Studies on these temperature impacts on *Aedes* behavior primary focused on flying and feeding. For female *Aedes aegypti,* it has been shown that

the optimal flight temperature is at 21°C, with a temperature range of 15°C to 32°C where flight is sustainable [76]. In the same study, it was identified that female *Aedes aegypti* can fly in an extreme temperature range of 10°C to 35°C for short periods of time [76, 83]. It is predicted that the lower optimal flight time is to aid in host-seeking during cold times of the day (early morning and late evening) when *Aedes aegypti* peak activity occurs [76, 83]. Host-seeking cues have been shown to be driven by CO_2, but odorants and heat cues also have an impact on host-seeking [91]. Early studies with *Aedes aegypti* indicated that females ceased biting at 15°C, were most aggressive at 28°C, and had an upper limit of feeding around 36°C [84]. *Aedes aegypti* are known to take multiple blood meals during a gonotrophic cycle [92]. A recent study comparing the number of bloodmeals taken by mosquitoes from Thailand (warmer) compared to Puerto Rico (cooler), observed that as environmental temperature increases so too does the number of *Aedes aegypti* meals [93]. Mosquitoes are poikilotherm, and as an pokilotherm, for hatching, pupating, flying, feeding, and other life cycle tasks, there is an optimal temperature at which the task is maximized [84]. Therefore, it is not surprising that flying, host-seeking, and feeding are optimized at different temperatures [76]. In summary, mosquito development and behavior are significantly impacted by the environmental temperature they are in, and while they may modulate their activity to survive, ideal environmental temperatures are unique to the species and the geographic regions in which they are found.

1.5 Temperature impacts Zika virus ecology

Over the past seven decades, scientists have identified numerous vertebrate and invertebrate species that are naturally infected with ZIKV. To date, there are at least seventy-nine animal species identified to be either naturally infected or experimentally susceptible to ZIKV

[94]. While primates make up the majority of animal species susceptible to ZIKV, some of the more interesting animal hosts include snakes, frogs, goats, and hippopotami, amongst others [94-96]. Of the invertebrate species collected, thirty-one wild caught mosquito species tested positive for ZIKV, of which the majority came from the *Aedes* genus (22 species) and the remainder *Culex* [97], *Eretmapodites* [98], *Anopheles*, and *Mansonia* [99]. Of the thirty-one wild caught mosquito species, twenty-five came from sylvatic settings and only six came from urban environments [94].

Since there are numerous potential hosts and vector species associated with ZIKV, it is important to highlight the ZIKV transmission cycles. The first cycle for maintaining ZIKV in nature is the sylvatic cycle, in which zoophilic mosquitoes such as *Aedes luteocephalus* and *Aedes furcifer* acquire ZIKV through feeding on a viremic animal and then following replication and dissemination, ZIKV is transmitted to susceptible wild animals, most commonly non-human primates (Figure 1.2) [94, 100]. The second cycle of ZIKV transmission is the urban cycle, in which a spillover event occurs, most likely when a ZIKV transmitting zoophilic mosquito feeds on and infects a human in close proximity to sylvatic cycles. Once ZIKV is in a human population in urban settings, an anthropophilic mosquito such as *Aedes aegypti* or *Aedes albopictus* can feed on viremic humans to acquire and transmit ZIKV to additional susceptible humans, perpetuating urban transmission (Figure 1.2). There are three other mechanisms of transmission, sexual transmission from human to human and vertical transmission, where a ZIKV positive female oviposit eggs infected with ZIKV, and venereal transmission from an infected female to a mosquito male and vice versa [94].
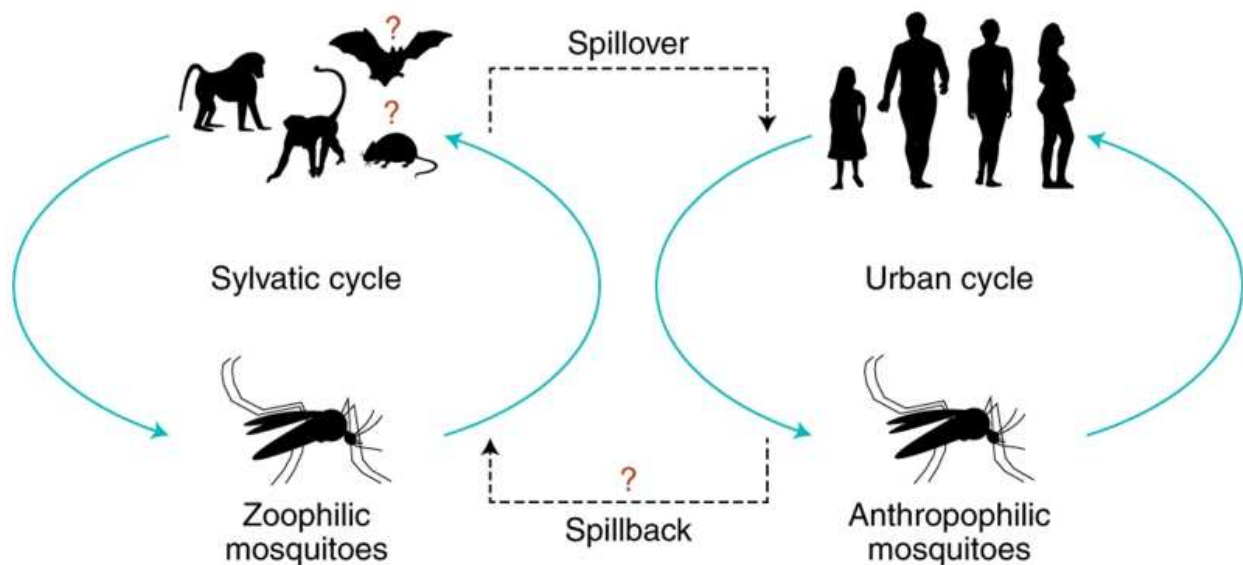
Figure 1.2 ZIKV vector-borne transmission. (a) Horizontal transmission of ZIKV in sylvatic cycle between animal vertebrate host and zoophilic mosquitoes, and urban cycle where ZIKV is transmitted between humans and anthropophilic mosquitoes. Modified from Gutierrez-Bugallo et al. (2019) [94].

The sylvatic and urban transmission cycles seem relatively straightforward understand: a competent mosquito feeds on a viremic vertebrate, becomes infected, feeds on another susceptible vertebrate and should the host become viremic, the next biting mosquito can acquire the virus thus completing the transmission cycle. In reality, the factors associated with the vector transmission cycles are quite complex. These factors include the number of mosquitoes feeding, the likelihood the vector survives the dangerous act of blood feeding and virus acquisition, how long the vector survives before it is able to transmit or the extrinsic incubation period (EIP), the number of bites a vector takes in a day, and are variables that must be taken into account. Fortunately, all of the above variables have been considered and included in Macdonald's Equation of Vectorial Capacity (V) [101]:

$$V = \frac{ma^2bp^n}{-\ln p}$$

where (*m*) is equal to the vector density in relation to the host, (*a*) is the probability of vector feeding per day, (*b*) is the vector competence (VC) or the ability of the mosquito to acquire and transmit a pathogen, (*p*) is daily survival, (*n*) is the EIP, and (-ln *p*) is the number that of days a mosquito survives after EIP. It was previously stated that mosquitoes are constantly exposed to daily and seasonal variations of temperature, and that this variation can play a large role in development and life history traits. Temperature can greatly impact vectoral capacity and thus the impact of environmental temperature on each variable with respect to *Aedes aegypti* or *Aedes albopictus* and ZIKV will be important to discuss and study.

The vector density (*m*) is directly impacted by temperature, as temperature is known to alter mosquito development and therefore host density. To recap, it has been shown that optimal temperature of development for *Aedes aegypti* is approximately 32°C and *Aedes albopictus* is 29.7°C, with increases or decreases of temperatures slowing or completely halting development [82-84]. The likelihood of a mosquito to feed per day (*a*) is also impacted by temperature: females mosquitoes are most aggressive in biting at 28°C but feeding decreases at extreme lows (15°C) and highs (36°C) [84]. Daily survival (*p*) is impacted by temperature and is dependent on the mosquito species and the geographic region of isolation. In a recent study comparing survival of *Aedes aegypti* mosquitoes collected from Mexico under eight different constant temperature (16°C, 20°C, 24°C, 28°C, 32°C, 34°C, 36°C C, 38°C) , peak survival out to 21 days was at 24°C, with population die off at 3 days at 38°C, 17 days at 36°C and ~ 50% survival at 34°C at 21 days. Interestingly, the lower temperatures of 16°C and 20°C survived with ~ 70% of the population at 21 days [102]. Also, diurnal temperature ranges are known to alter mosquito survival, in which large daily temperature fluctuations of 20°C significantly decrease *Aedes aegypti* survival [87]. EIP (*n*) for ZIKV is also impacted by temperature. In a study by Olivia et al. (2020), ZIKV EIP of

a Puerto Rico isolate in *Aedes aegypti* was held at varying temperatures of 18°C, 21°C, 26°C, and 30°C. From their findings, it was shown that EIP was shortest at 30°C with a minimum EIP of 5.1 days and a maximum of 24.2 days at 21°C [103]. A study using a Mexico isolate of ZIKV and *Aedes aegypti*, produced similar results [102]. Survival post EIP (-ln p) is similarly impacted by temperature as daily survival (p), however, it is a more complex dynamic. Increased temperatures shorten EIP while decreasing survival, and lower temperature significantly increases EIP, with a minimal impact on survival. By assessing both the work performed by Tesla et al. (2018) and Olivia et al. (2020), it is shown that optimal survival temperatures fall around 28°C, and EIP is shortest around 30°C, indicating that the temperature range of 28°C - 30°C is optimal for survival post EIP and temperatures higher or lower decrease survival and lengthen EIP [102, 103]. The temperature impacts on VC (b) is also complex. Tesla et al. (2018) show that VC is impacted by constant temperatures in a unimodal manner with peak VC being around 29°C and decreasing as it nears the extreme temperatures of 16°C or 38°C [102]. Likewise, Olivia et al. (2020) showed that as temperatures increased from 18°C to 30°C so too does VC, which is in agreement with previous findings [102, 103]. Interestingly, another study showed that when *Aedes aegypti* and *Aedes albopictus* are reared, infected, and housed under diurnal temperature fluctuations with a range of 4°C, a max of 32°C and minimum of 24°C, there is no significant difference in VC and increased temperature groups actually decreased vectorial capacity [104]. These together indicate that temperature is a major player in VC and vectorial capacity (V). However, it is important to understand that mosquito species, geographic location of isolation, and ZIKV strain all have a significant impact on VC and therefore V [105].

1.6 Zika virus replication and molecular biology

Flaviviruses are composed of a relatively conserved viral structure and replication cycle from attachment to release. Flaviviruses are small spherical enveloped viruses typically around 50 nm in diameter for mature virus particles and around 60 nm in diameter for immature virus particles. The flavivirus genome is an approximately 11 kb positive-sense single-stranded RNA (+ssRNA) genome that encodes three structural proteins and seven non-structural (NS) proteins: envelope (E), precursor membrane (pr/M), capsid (C), NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5, respectively (Figure 1.3) [106]. The +ssRNA genome contains a type I cap at the 5' end which protects viral RNA from exonucleases (e.g. Xrn1) degradation [107] and stimulates initiation of viral translation [108]. Additionally, flavivirus viral RNA can act as mRNA and be directly translated by the ribosome, but unlike mRNA they lack a poly-A tail at the 3' end [106] causing the viral RNA to be less stable than mRNA. Last, there is a 5' and 3' highly structured untranslated region flanking the protein-coding region of flavivirus viral RNA, which is essential for viral replication and translation (Figure 1.3A) [106, 108-110]. The viral structure is composed of an outer lipid envelope containing 180 copies each of E and M (a cleavage product of pr/M) proteins [111], and a nucleocapsid which is presumed to contains a linear positive strand genomic viral RNA and is comprised of several hundred copies of C [112].
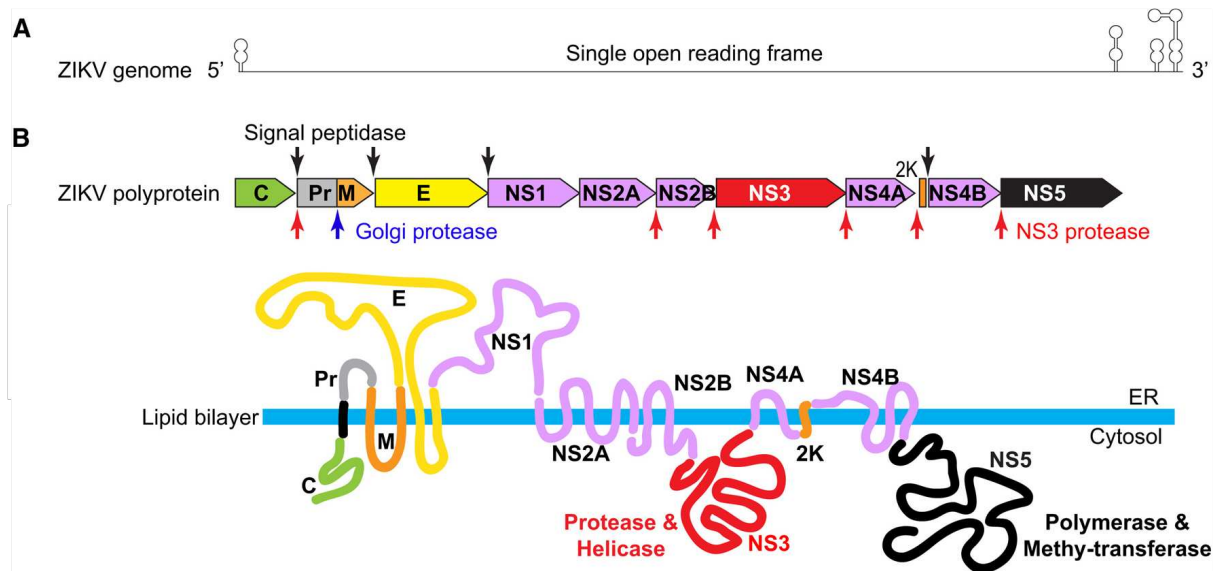
Figure 1.3 Diagram of ZIKV genome and polyprotein. (a) Flavivirus genomic RNA, ~11 kb long encoding a single open reading frame with 5' and 3' structured untranslated regions. (b) Flavivirus polyprotein with host (blue and black arrows) and viral (red arrows) cleavage sites and the topological arrangement of the polyprotein at the ER. Modified from Ming et al [113].

The flavivirus replication cycle begins when the virions attach to the cell and enter via receptor-mediated endocytosis in an E-dependent manner. While there are numerous cell surface receptors indicated in flavivirus attachment (TIM and TAM family, DC-SIGN, heat-shock proteins [114-116]), it has been show that the AXL protein may be an important receptor for ZIKV in humans [117, 118]. However, AXL is not the only receptor that can be used for successful ZIKV binding and cell entry; it has been show that AXL -/- neural progenitor cells are readily infected by ZIKV [119]. Regardless of cell type (vertebrate or invertebrate), proton pumps on the endosome are activated after cell attachment and endocytosis, resulting in a pH drop in the endosome and causing the E glycoprotein to undergo structural change from a homodimer to a homotrimer. This enables fusion of the viral and endosome membranes, and the subsequent release of the viral nucleocapsid and genomic viral RNA into the cell cytoplasm [108, 120-123]. The nucleocapsid dissociates from the 5' capped viral RNA, and the RNA is recognized by host translation

16

machinery leading to the generation of the viral polypeptide [108]. Currently, there are two perspectives as to how this translation occurs. The first perspective is that translation starts in the cytosol but stalls when the capsid transmembrane domain emerges from the ribosome. Signal recognition particles bind this domain, deliver the translation complex to the endoplasmic reticulum (ER) membrane, and the transmembrane domain binds to the ER and translation is completed [106, 108]. The second perspective stipulates that upon successful release from the nucleocapsid, the genomic viral RNA is preferentially recruited directly to ER-associated ribosomes for translation, bypassing cytoplasm associate ribosomes [124, 125]. Irrespective to how the viral RNA is carried to the ER, the flavivirus genome is translated at the ER [126] into a single polyprotein. The polyprotein is subsequently activated via co- and post-translationally cleavage by host viral and host proteases, resulting in the ten mature viral proteins described above (Figure 1.3B) [106].

After initial translation of the genomic viral RNA, the transmembrane domains of NS proteins, such as NS4A, integrate into and remodel the ER membrane to generate viral replication complexes (RC). After cleavage of the polyprotein, the C, NS3, and NS5 proteins are located on the cytoplasmic side of the ER, whereas the pr/M, E, and NS1 proteins are in the lumen of the ER (Figure 1.3B). The NS proteins NS2A, NS2B, NS4A, and NS4B are all located within the ER membrane alongside short regions between transmembrane domains (Figure 1.3B) [106]. The RC generated by the NS proteins transcribe the +ssRNA viral genome into a negative strand RNA that will be used as a template for replication. In early stages of replication, the +ssRNA alternates back and forth between translation to generate viral proteins and RCs, and acting as a template for negative strand generation [110]. The 3' to 5' long distance RNA-RNA interactions of the negative strand RNA work to cyclize the RNA into a panhandle formation. This formation is required for

the proper recruitment and initiation of the NS5 RNA-dependent RNA-polymerase (RdRp), which generates +ssRNA genomes used for translation [112, 127]. Following sufficient generation of viral genomes and viral structural proteins, virus assembly is initiated. The charged face of C binds to viral +ssRNA [128] and the hydrophobic face of C binds to the ER [110]. The C-viral RNA complex buds into the ER lumen in areas where E and pr/M proteins are located, leading to the assembly of immature virion approximately 60 nm in diameter [106, 129]. The immature virion is assembled with E trimers around the prM, preventing viral fusion with cellular membranes before egress [129]. Immature virion are transported through the Golgi and the trans-Golgi network and exposed to acidic environments. This leads to E and prM conformational changes, making prM accessible for cellular protease cleavage via furin. Cleavage produces the pr peptide and M, which remain bound to one another in the low pH environment [106, 130]. The virus then exits the cell and is exposed to a neutral pH environment, leading to the dissociation of the precursor peptide and complete maturation of the virion [131].

As the ten mature flavivirus proteins are known to have multiple impacts on viral replication and assembly, therefore it is important to note key interactions of the structural and nonstructural proteins. The C protein has the least genetic conservation among the flavivirus proteins, yet the structure and charge distribution is highly conserved [132]. The primary role of C is encapsulation of viral RNA to form the nucleocapsid for viral assembly. Interesting, ZIKV C has been found in the nucleoli and associated with an increase in ribosomal stress and apoptosis, or programmed cell death [133]. It is likely in part responsible for modulation of the host transcriptome [132]. Additionally, as C is indiscriminate in binding to nucleic acids, it has been suggested to interfere with dicer activity in mosquitoes [134] and it is likely that there are other capsid-nucleic acid interactions yet to be described [135]. As another structural protein, the prM

protein is essential in E protein folding, viral assembly, and egress, and thus allowing for the formation of an immature virus particle that will not bind to a cellular membrane internally. After removal of the pr peptide post virion egress, M is an essential piece of the viral envelope structure [106]. While not entirely understood, it has been suggested that amino acid changes of prM in pre-epidemic strains of ZIKV led to the adaption of an urban-based transmission cycle or an increased neuroinvasiveness [136, 137]. The final structural protein, the E protein, is made up of three envelope domains (ED): I, II, and III, a fusion loop, stem, and two transmembrane domains. EDI aids in stabilizing the E protein [138] and is involved in to viral assembly [139]. EDII undergoes conformational change during infection and contributes to endosomal membrane fusion [140], and EDIII participates in host receptor recognition [141]. As a nonstructural protein, NS1 has been found to have multiple functions: it acts as a cofactor with other viral proteins to facilitate viral replication [142], it has been associated with membrane remodeling leading to RC formation [143, 144], it is essential in viral assembly, release, and immune evasion [145-148], and NS1 has been associated with temperature-sensitive mutations. A YFV study showed that a single amino acid change from alanine to arginine at amino acid position 299 led to significantly less replication at higher temperature of 39°C, indicating that temperature can impact NS1 and lead to loss of replication [149]. NS2A is a transmembrane protein with no known enzymatic activity. However, it has been shown to recruit the viral genome to the prM, E, and C complex for virion assembly [150] and to aid in viral RNA replication [151]. It has been suggested that NS2A and NS2B may play roles in ER budding and viral egress through individual oligomerization and pore formation [152, 153]. NS2B is an essential cofactor for NS3, which in turn is essential for protease activity and processing of the viral polyprotein [144, 154]. NS2B has also been associated with immunomodulation, counteracting the functions of cGAS and STING therefore inhibiting Type I

interferon (IFN) [155, 156]. In addition to the protease activity described previously, NS3 has a helicase domain that assists in viral RNA synthesis by unwinding double stranded RNA for synthesis by NS5 [157, 158]. NS3 also has NTPase and triphosphatase activity, both of which are required for efficient viral RNA synthesis and 5' capping [144]. NS4A and NS4B are transmembrane proteins that act as RC components and play important roles during replication. [159, 160]. As an example, it has been shown that NS4A induces membrane rearrangement, leading to the formation of convoluted membranes [161, 162]. NS4B has been shown to block IFN α/β signaling [163], as well as cause membrane rearrangements by itself [164]. NS5 is the largest viral protein generated during translation, and it includes a methyltransferase which aids in capping newly synthesized viral RNA, guanylyltransferase which is likewise involved in viral RNA capping, and a RdRp domain for viral replication [165-168]. Additionally NS5 interacts with NS3 for efficient viral RNA synthesis, indicating that the helicase activity of NS3 is important for RdRp progression [169]. Interesting, RdRp can accumulate temperature -sensitive mutations which may significantly decrease the temperature of initiation [170, 171].

1.7 Characteristics of arbovirus evolution

Arboviruses such as ZIKV are RNA viruses that exists in nature as a genetically complex mixture of competing viral genomes. RNA viruses use RNA-dependent RNA polymerases (RdRps), which lack a 3' to 5' exonuclease proofreading mechanism which drive mutation accumulation [172]. Low fidelity replication, large populations, and fast replication ultimately lead to a large population of mixed viral genomes. This complex population structure is comprised of a large number of variant genomes which is termed the mutant spectra, mutant swarms, or mutant clouds [173]. As ZIKV replicates in the mosquito or host the mutant spectra of viral genomes

undergo constant mutation, inter-variant competition, and selection for the fittest set of variants in a given landscape [174, 175]. In order to help characterize the mutant spectra associated with the evolutionary dynamics of RNA viruses, the quasispecies theory is used to explain self-organization and adaptability of RNA-like molecules [176, 177], and has recently been used to describe the dynamics associated with RNA virus evolution [174, 178-181]. Additionally, quasispecies is important for defining the consensus sequence of RNA viruses and the minority variants, which are associated with fitness altering viral phenotypes [174, 182, 183]. Thus, using quasispecies theories, we have a framework that can be used to understand the mechanism of genetic diversity and fitness during arbovirus infection. In summary, arboviruses like ZIKV are found in nature as a genetically diverse population of competing viral genome. They have high mutation rates and large heterogenous populations which allow the virus to adapt to selective pressures. Likewise, different genetic components or genomes of the arbovirus quasispecies can better adapt to different host ranges, which may be essential for viral emergence or reemergence [173]. The bigger the effective population size, and higher the replication rate, the more likely the virus will persist [174]. Therefore, the combination of replication rate, viral load, genetic diversity, and replicative fitness (ability to create infectious particles) can significantly impact disease progression and expansion [173]. These factors undoubtedly contributed to the ZIKV epidemic of 2015-2016. It is predicted that adaption for transmission by urban vectors [184], such as what occurred with the CHIKV adapting to *Aedes albopictus* [185], or recently by Liu et al. (2017), which showed that an alanine to valine amino acid substitution at residue 188 in NS1, led to increase in ZIKV infectivity and transmission in *Aedes aegypti* mosquitoes may be a key factor to emergence [184]. Conversely, it has be predicted that the adaption to vertebrate hosts for higher replication in humans by changing the codon usage bias to match that of human hosts instead of invertebrate

host is hypothesized to be a key in ZIKV emergence [94, 186]. In either case the adaptive plasticity offered by RNA virus quasispecies is most likely the driving factor for vector and host expansion.

In addition to the molecular dynamics associated with RNA viruses, arboviruses are unique in that they must sufficiently infect and replicate in multiple hosts, invertebrates and vertebrates, both of which pose unique bottlenecks of infection (a severe reduction in arbovirus population size during various stages of infection) [187] and selective environments [188]. In addition to host differences, arboviruses must also replicate and disseminate through multiple tissue in invertebrates that can act as bottlenecks [189]. Bottlenecks between hosts and within hosts can drastically reduce the population size the quasispecies, which will have a significant impact on shaping the evolutionary lineage of the virus [187, 190, 191]. Additionally, it has been shown that the mutant spectrum must be at a certain threshold for the virus to overcome encountered bottlenecks and have successful infection or dissemination [192, 193]. Other studies characterized the impact of mosquito bottlenecks and their role in altering mutant spectrum. Patterson et al. (2018) showed that the midgut escape bottleneck of *Culex taeniopus* mosquitoes reduced the mutant spectrum of Venezuelan equine encephalitis virus (VEEV) and led to the accumulation of novel mutations, likely due to the regeneration of the mutant spectrum after stochastic population reduction [192]. Similar results were observed for the WNV mutant spectrum when exposed to various *Culex spp*. and *Aedes aegypti* mosquito intrinsic bottlenecks, in which stochastic reductions and expansions of the mutant spectrum were observed when encountering mosquito bottlenecks [187]. That the reduction and expansion of the mutant spectrum is likely arbovirus - vector specific, as it has been shown that depending on this combination, the mutant spectrum during systemic infection is significantly different [187, 194]. While mosquito bottlenecks play a

large role in shaping the mutant spectra, it is know that other factors such as the innate antiviral response (RNAi) also drive diversifying selection [195, 196].

As arbovirus perpetuation is dependent on being able to successfully replicate in two hosts, when the virus is exposed to its vertebrate host, it is exposed to an entirely new fitness landscape which impacts the mutant spectrum. For example, a recent study assessing ZIKV mutant spectrum in infected immunocompromised mice (Ifnar$^{-/-}$) showed that as the virus infects specific organs, the mutant spectrum is likewise unique, indicating an organ and tissue-specific bottleneck associated with mutant spectrum diversity [197]. These unique tissue-specific bottlenecks as well as host specific bottleneck have been observed in other studies, indicating that the impact of these different fitness landscapes on the mutant spectrum are far from being predictive [187, 198-201]. Another example of changes from vertebrate landscape can be observed in WNV infected birds. As we would predict, birds provide drastically different fitness landscape for WNV. For example, WNV infection in mosquitoes is under diversifying selection, increasing the diversity of the mutant spectrum, but birds exhibit strong purifying selection, decreasing the diversity of the mutant spectrum and selecting for the primary sequence in the population [201]. This strong purifying selection is believed to be in part due to the primary innate antiviral response, type I interferon pathway, and the initial bottleneck of infection [202].

In summary, arbovirus evolution is characterized by the general features that they share with RNA viruses. Arboviruses have a diverse mutant spectrum that undergoes reductions and expansions in a tissue/species-specific manner. Stochastic reduction and regeneration of the mutant spectrum is essential for the perpetuation of arboviruses in nature. Therefore, further characterization of the mutant spectrum with a focus on minority population generation in these unique fitness landscapes is needed for the future characterization and prediction of factors

23

associated with viral emergence. In addition to the intrinsic factors impacting the fitness landscape, it is important to advance the understanding of how extrinsic factors, such as temperature, alter the intrinsic factors that shape these environments.

Chapter 2: Extrinsic incubation temperatures lead to specific selective environments in Aedes mosquitoes during Zika virus infection

2.1 Introduction

Arthropod-borne viruses (arboviruses) such as Zika virus (ZIKV) are mainly RNA viruses that are transmitted by an arthropod vector to vertebrate hosts [203]. Arboviruses are required to alternate replication between hosts with drastically different body temperatures. This extreme variation in temperature may pose unique evolutionary pressures that could impact arbovirus transmission dynamics, replication rates, and population structure. While replication in vertebrates generally occurs within 2-3 degrees of 38°C [204], infection in mosquitoes may occur at a much wider range of temperatures: Mosquito vectors are distributed throughout tropical and temperate climates and their geographical range is increasing [81]. Climate variations such as heat waves, cold snaps, or daily temperature fluctuations change the host environment within which arboviruses replicate and are transmitted. Fluctuations in the temperature of the host environment are central to arbovirus biology [189] and virus-host interaction [205-207].

The impact of temperature on vector competence (VC), i.e. the ability of a mosquito to acquire, maintain, and transmit a pathogen, is well described. The extrinsic incubation temperature (EIT) influences viral replication and dissemination within vectors [208-213], altering the extrinsic incubation period (EIP), i.e. the number of days between acquisition of an infection and infectiousness to a new host [205, 214]. Most studies examining the effects of temperature on VC use single temperatures representative of average conditions [215-218]. However, diurnal temperature fluctuations more accurately model environmental conditions [87,

219-221]. EIT impacts ZIKV VC in *Aedes aegypti* mosquitoes one study shows that as

temperature increase from 20°C to 30°C so too does VC, another study goes one step further and

shows that EIT actually impacts VC in a unimodal manner, with extreme low (16°C) and high

(38°C) temperatures having low VC while median temperatures (28°C -32°C) have higher VC

[102, 103]. Temperature also exerts a strong selective pressure on RNA viruses [222, 223],

however, little is known about how temperature may influence the emergence of arbovirus

genotype in the context of infection in mosquitoes. Temperature may therefore have multiple

impacts on arbovirus replication and transmission.

RNA viruses like ZIKV and West Nile virus (WNV) have the capacity to evolve rapidly

in response to changing environments. This is in part due to short generation times and error

prone replication driven by the viral RNA-dependent RNA polymerases, which lack error-

checking and mismatch-repair mechanisms [179, 224]. In addition, mosquito RNA interference

[225] and stochastic population reductions caused by bottlenecks during mosquito infection

[191] impact arbovirus population structure. As a result, arboviruses, including WNV and ZIKV,

exist within hosts as large populations of mixed haplotypes, which is critical to their perpetuation

in nature [191, 226-229]. While there have been numerous studies assessing ZIKV VC and viral

ecology and some efforts focusing on the use of environmental data to predict virus spread, there

is limited knowledge as to how environmental factors such as temperature impact the selective

environments and mutational diversity of arboviruses within mosquitoes.

Accordingly, we sought to determine whether ZIKV mutational diversity is altered by the

EIT of *Aedes aegypti* and *Aedes albopictus*. We exposed *Aedes aegypti* and *Aedes albopictus,* the

two dominant vectors of ZIKV, to a Puerto Rican isolate of ZIKV. We held both *Aedes* species

at constant temperatures of 25°C, 28°C, 32°C, 35°C, and a diurnal fluctuation from 25°C to

35°C. Temperatures were selected to mimic environmental temperatures ranges found in Rio de Janeiro, Brazil during the 2015-2016 ZIKV outbreak. These experiments allowed us to assess EIT impacts on ZIKV vector competence and virus evolution within the mosquito host (Figure 2.1). Our results suggest that the selective environment within mosquitoes is dependent on temperature, and that daily fluctuating temperatures impose strong purifying selection.
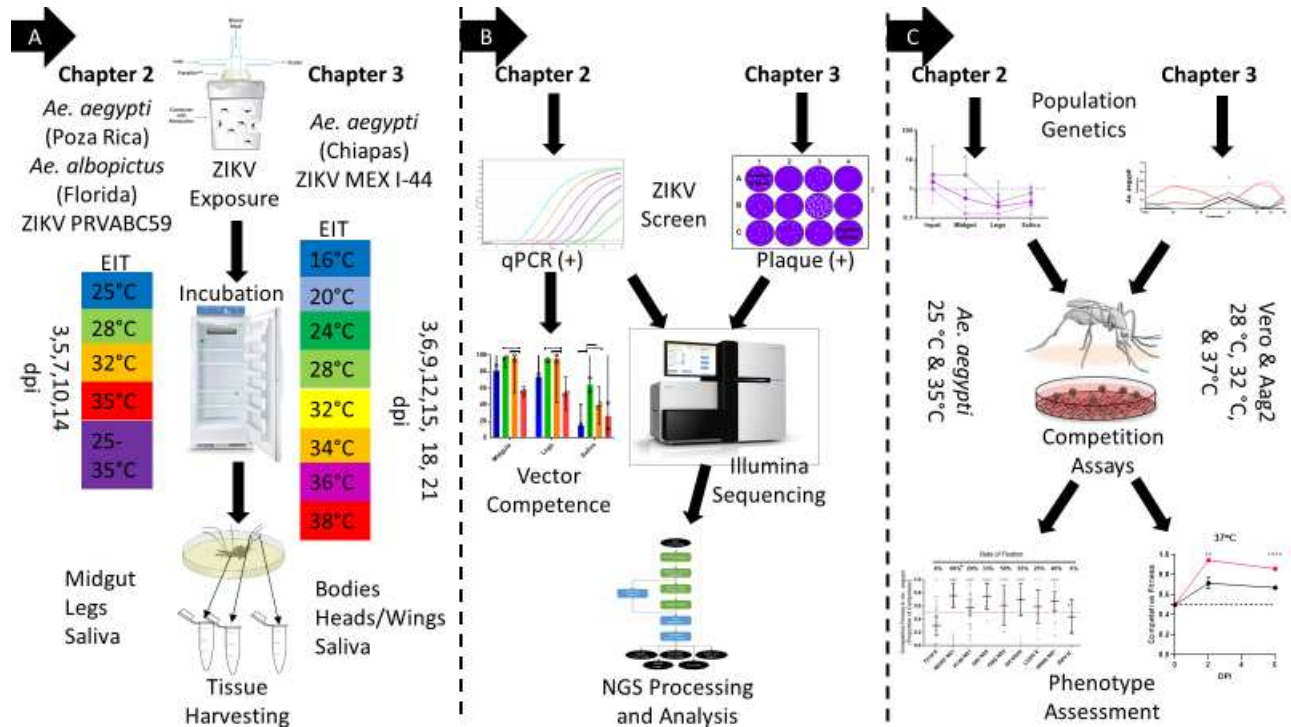
Figure 2.1. An overview of methods. Methods for Chapter 2 (left) and Chapter 3 (right) are summarized in columns A, B, and C. (A) *Aedes* mosquitoes were exposed to ZIKV blood meals, engorged females were sorted and housed at specific EITs for 3-14 dpi (Chapter 2) or 3-21 dpi (Chapter 3). Mosquito tissues were collected and (B) screened for ZIKV positive tissues via qPCR (Chapter 2) or plaque assays (Chapter 3). VC analysis was performed and biological triplicates of matching tissues were sequenced. The resulting Next-generation sequencing (NGS) data was processed and analyzed using population genetics analysis methods (C). Any variants of interest were cloned using an infectious clone (Chapter 2) or by purifying biological clones (Chapter 3) for further phenotype assessment.

2.2 Materials and Methods

2.2a Cells and Virus

African Green Monkey kidney cells (Vero; ATCC CCL-81) were maintained at 37⁰C and 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Pen-Strep). ZIKV strain PRVABC59 (ZIKV-PRVABC59; GenBank # KU501215) obtained from the Center for Disease Control and Prevention branch in Fort Collins, CO was originally isolated from the sera of a patient returning from travel to Puerto Rico in December 2015. The virus was isolated from Vero cells and a 4th passage frozen at -80°C was used for all *in vivo* and *in vitro* experiments. ZIKV-PRVABC59 infectious clone (ZIKV-PR-IC) served as a backbone for the reverse genetic platform developed by our lab [230] to introduce all point mutations. ZIKV-REF was designed using the aforementioned reverse genetic platform. ZIKV-REF incorporates 5 synonymous mutations into amino acid 108-arganie and 109-serine of the prM protein coding sequence. The ZIKV-PR-IC sequence nucleotides were changed from ZIKV-PR-IC 795-CGG TCG-800 to ZIKV-REF 795-AGA AGT-200.

2.2b Mosquitoes

Aedes aegypti colonies for this study were established from individuals collected in Poza Rica, Mexico [231] and used at F13-F18 generation. A lab adapted colony (greater than 50 generations) of *Aedes albopictus* were established from individuals collected in Florida, USA; the colony was provided by the Centers for Diseases Control and Prevention (CDC-Fort

Collins, CO, USA) in 2010. Mosquitoes were reared and maintained at 27-28°C and 70-80% relative humidity with a 12:12 L:D photoperiod. Water and sucrose were provided ad libitum.

2.2c Infection of Aedes mosquitoes and sample collection

Adult mosquitoes used for experiments were 3-7 days post-eclosion. Mosquitoes were provided a bloodmeal using a water jacketed glass membrane feeder. The bloodmeal contained calf blood mixed 1:1 with ZIKV-PRVABC59 (1E7 PFU/mL). Engorged female mosquitoes were sorted into cartons and housed at 25°C, 28°C, 32°C, 35°C at constant temperature or alternating between 25°C-35°C to simulate diurnal conditions, with 70-80% relative humidity and 12:12 L:D photoperiod. Mosquitoes were cold anesthetized in preparation for dissection. Mosquito midguts, legs/wings, and saliva from the first batch of mosquitoes were collected after 7- and 14-days post-feed for NGS processing. The mosquito carcass, legs/wings and saliva from the second batch of mosquitoes were collected at 3, 5, 7, 10, and 14 days post-feed for assessing systemic infecting dynamics. Tissues represent infection (midgut), dissemination (legs), and transmission (saliva). Tissues were removed using forceps cleaned with 70% ethanol between samples and were homogenized in 200 µl of mosquito diluent with a stainless-steel ball bearing using a Retsch Mixer Mill 400 at 24 Hz for 45 seconds, as previously described [232]. Saliva was collected by inserting mosquito mouthparts into capillary tubes containing mineral oil for 40 minutes. Saliva in oil was removed from the capillary tube by centrifugation into 100 µl of mosquito diluent for 5 minutes at >20,000 x g. All samples were stored at -80°C until manipulation.

2.2d Plaque assay

ZIKV stocks and infectious bloodmeal were quantified by plaque assay on Vero cell cultures seeded in 12-well plates. Duplicate wells were infected with 0.2 ml aliquots from serial 10-fold dilutions of virus stocks and infectious blood meals in media (DMEM supplemented with 1% FBS and 1% penicillin/streptomycin), and virus was adsorbed for one hour by incubating at 37 °C in 5% CO_2. Following incubation, the inoculum was removed, and monolayers were overlaid with tragacanth-EMEM overlay containing 1x EMEM, 5x L-glutamine, sodium bicarbonate 3.75%, tragacanth 1.2%, gentamicin (25mg/ml), and Amphotericin B 40mL/L. Cells were incubated at 37 °C in 5% CO_2 for four days for plaque development. Cell monolayers then were stained with 1 mL of overlay containing a 20% ethanol and 0.1% crystal violet. Cells were incubated at room temperature for 30-60 minutes and then gently washed and plaques were counted. Plaque assays for 3, 5, 7, 10 and 14 days post infection (dpi) mosquitoes were performed similar to above with the following changes, 50 ul of homogenized midgut and leg tissues or 30 ul of saliva samples were added to Vero cultures in 24-well plates (final volume of 200 ul), and plaques were observed post processing.

2.2e Viral RNA isolation

Viral RNA was extracted from 50 µl of either cell culture supernatant, homogenized mosquito tissues, or saliva-containing solution using the Mag-Bind® Viral DNA/RNA 96 kit (Omega Bio-Tek) on the KingFisher Flex Magnetic Particle processor (Thermo Fisher Scientific). Nucleic acid extraction was performed as directed by the manufacturer and eluted in 50 µl nuclease-free water. Viral RNA was then quantified by qRT-PCR using the iTaq™ Universal Probes One-Step Kit (BIO-RAD) according to manufacture protocol. The qRT-PCR

primers and probe were forward primer (5'- CCGCTGCCCAACACAAG-3'), reverse primer (5'- CCACTAACGTTCTTTTGCAGACAT-3'), and FAM probe (5'- AGCCTACCTTGACAAGCAGTCAGACACTCAA-3') sequences [233].

2.2f Generation of ZIKV mutant clones

An infectious clone for ZIKV-PRVABC59 was used to generate mutants [230]. To engineer the point mutations (Supplemental Table 2.1) into the ZIKV genome, the corresponding single nucleic acid substitution was introduced into the ZIKV-PR-IC using *in vivo* assembly cloning methods [234]. The infectious clone plasmids were linearized by restriction endonuclease digestion, PCR purified, and ligated with T4 DNA ligase. From the assembled fragments, capped T7 RNA transcripts were generated, and the resulting RNA was electroporated into Vero cells. Infectious virus was harvested when 50-75% cytopathic effects were observed (5-8 days post transfection). Viral supernatant was clarified by centrifugation and supplemented to a final concentration of 20% fetal bovine serum and 10 mM HEPES prior to freezing and storage as single use aliquots. Titer was measured by plaque assay on Vero cells. All stocks (both wildtype and infectious clone-derived viruses) were sequenced via sanger sequencing to verify complete genome sequence.

2.2g Competition Study

Competitive fitness was determined largely as described in previous studies [201, 235]. Competitions were conducted with orally infected *Aedes aegypti* (Poza Rica) mosquitoes. Three to seven day old mosquitoes were offered a bloodmeal containing the 1:1 mixture of viruses (ZIKV-REF and ZIKV-clone of interest) at a concentration of 1 million PFU/mL and bodies

were collected 14 days post blood feed. RNA was extracted as above, and amplicons were generated via qRT-PCR using iTaq™ Universal SYBR® Green One-Step Kit (BIO-RAD) according to manufacture protocol. A locked nucleic acid (LNA) forward primer was used to ensure amplicon specificity. The forward LNA primer 5'-A+CTTGGGTTGTGTACGG-3' and reverse primer 5'- GTTCCAAGACAACATCAACCCA-3' were used to generate amplicons for Quantitative Sanger sequencing. Genotype fitness was analyzed using polySNP software to measure the proportion of the five synonymous variants present in the ZIKV-REF sequence allowing us to compare the proportion of ZIKV-REF virus to competitor virus.

2.2h Library preparation for next-generation sequencing

Positive controls were generated in triplicate, each generated with 1 million genome equivalents of a 100% ZIKV PRVABC59 viral stock, a mixture of 90% ZIKV PRVABC59 and 10% ZIKV PA259359 (GenBank: KX156774.2), and a mixture of 99% ZIKV PRVABC59 and 1% ZIKV PA259359. The negative control was water (no template control, or NTC). Controls and sample RNA (10ul) was prepared for NGS using the Trio RNA-Seq Library Preparation Kit (NUGEN) per manufacturer standard protocol. Final libraries were pooled by tissue type and analyzed for size distribution using the Agilent High Sensitivity D1000 Screen Tape on the Agilent Tapestation 2200, final quantification was performed using the NEBNext® Library Quant Kit for Illumina® (NEB) according to manufacture's protocol. 150 nt paired-end reads were generated using the Illumina HiSeq4000 at Genewiz.

2.2i NGS processing and data analysis

NGS data were analyzed using a workflow termed "RPG (RNA virus Population Genetics) Workflow"; this workflow was generated using Snakemake [236] a detailed description can be found in Appendix A1 and workflow and related documentation can be found at https://bitbucket.org/murrieta/snakemake/src. Briefly, Read 1 and Read 2 .fastq files from paired-end Illumina HiSeq 4000 data were trimmed for Illumina adaptors and quality trimming of phred scores < 30 from the 3' and 5' read ends using cutadapt [237]. The reads were then mapped to the ZIKV-PRVABC59 reference sequence (Genbank: KU501215) using MOSAIK [238], similar to that previously described [239]. Picard [240], Genome Analysis Toolkit (GATK) [241], and SAMtools [242] were used for variant calling preprocessing. Single nucleotide variants (SNV's) and inserts and deletions (INDELS) were called using LoFreq [243] with the --call-indels command; otherwise, all settings were default. Consensus sequences were generated using the .vcf files generated above and VCFtools [244]. NTC had less than 0.02% of reads mapping to ZIKV and an average of < 8x coverage across the genome indicating little to no contamination, sequencing bleed through, or index hopping (Supplemental Table 2.2). Therefore only variants in the coding sequence (nt position 108-10379), with >100x coverage and a cut off of 0.01 frequency were used for analysis to account for low coverage (reads per genome position) in the 3' and 5' untranslated regions (Supplemental Table 2.2 Table, Supplemental Figure 2.5).

Data analysis was performed using custom Python and R code integrated into the RPG Workflow. Using .vcf files generated by LoFreq and .depth files generated by GATK DepthOfCoverage command, the workflow generates .csv files that provides sequencing coverage across the coding sequence (CDS) and results for measures of genetic diversity

(defined in section 2.2j) including complexity, richness, nucleotide diversity, selection, and divergence of a specified locus. Additionally, amino acid changes, synonymous (S) and non-synonymous (NS) changes, and Shannon entropy are reported by variant positions. The same scripts are called manually outside of the RPG Workflow to perform the above analysis on specific protein coding regions or to compare divergence of populations other than the input.

2.2j Genetic diversity

All genetic diversity calculations were incorporated into Python and R code located at https://bitbucket.org/murrieta/snakemake/src/master/scripts/. In short, richness was calculated by the sum of the intrahost SNV (iSNV) sites detected in the CDS in each population. Diversity was calculated by the sum of the iSNV and amino acid substitutions frequencies per CDS. Complexity was calculated using Shannon entropy (S) which was calculated for each intrahost population (i) using the iSNV frequency (p) at each nucleotide position (s):

$$(1) \qquad S_{i.s} = -(p_s(log_2 p_s) + (1 - p_s)log_2(1 - p_s))$$

The mean S from all sites s is used to estimate the mutant spectra complexity. Divergence was calculated using F_{ST}, or the fixation index, to estimate genetic divergence between two viral populations as described previously [239].

2.2k Selection

Intrahost selection was estimated by the ratio of nonsynonymous (d_N) to synonymous (d_S) SNVs per site (d_N/d_S) using the Jukes-Cantor formula as previously described [239], and incorporated into custom Python code found at

34

. DnaSP software [245] was used to determine the number of nonsynonymous (7822.83) and synonymous (2446.17) sites from the ancestral input ZIKV consensus sequence. When no synonymous SNVs sites were present in replicates, d_N/d_S was set to 1, and no nonsynonymous SNV's d_N/d_S was set to 0.

2.2l Statistical analysis

All analyses were performed using GraphPad Prism (version 7.04) and R. Fisher's exact test were used to determine significant difference in virus titers and viral loads. All other tests were done using Kruskal-Wallis with Dunn's correction unless otherwise noted.

To evaluate the relationship between external factors and the infection dynamics of ZIKV, we examined the data with generalized linear models. The predictors we used include days post infection (days), temperature (scaled), species, and tissue type. We evaluated the impact of these variables on consensus changes, vector competence, complexity, nucleotide diversity and richness. We assumed that consensus changes and richness follow a quasi-poisson distribution, complexity and nucleotide diversity follow a linear distribution, and assumed that dissemination efficiency and vector competence follow a binomial distribution. Our original models follow the base structure:

$$response \sim \exp\left[\beta_1(days) * \beta_2(temp)^2 * \beta_3(species) * \beta_4(tissue)\right]$$

Each model was reduced to a best fit structure using AIC values and/or a chi-square goodness of fit test. The polynomial on temperature allows us to differentiate between the linear and quadratic effect of temperature. Vector competence was evaluated with the following base structure for each tissue response.

$$response \sim \exp\left[\beta_1(days) * \beta_2(temp)^2 * \beta_3(species)\right]$$

2.3 Results

2.3a Vector competence

To assess how extrinsic incubation temperature affects vector competence for ZIKV we exposed *Aedes aegypti* and *Aedes albopictus* to ZIKV (n=72-108), held them at 25°C, 28°C, 32°C, 35°C and alternating between 25°C-35°C for 7 and 14 days. Infection rates were high in all mosquitoes except those held at 35°C (Figure 2.2). We observed that in both *Aedes* species there was a unimodal distribution across constant temperature groups in mosquitoes that disseminated and transmitted ZIKV at 7- and 14-days post infection (dpi) (Figure 2.2). Peak dissemination and transmission occurred at a median temperature of 28°C or 32°C, while extreme temperatures (25°C and 35°C) decreased dissemination and transmission. Interestingly, at earlier time points, temperature had a greater impact on dissemination and transmission. In both *Aedes* species moderate temperatures (28°C and 32°C) significantly (p < 0.05, Two-tailed Fisher's exact test) increase dissemination and transmission at 7 day post-infection (Figure 2.2A & 2.2C) with *Aedes albopictus* being the most impacted (dissemination of ~30% at 28°C increasing to ~80% at 32°C) (Figure 2.2A). Our diurnal temperature group, when graphed by mean daily temperature (30°C), did not fit with the expected unimodal distribution. Instead, this temperature group fit best between the 25°C and 28°C temperatures or 32°C and 35°C temperatures. Diurnal temperatures had significantly lower dissemination (p < 0.05, Two-tailed Fisher's exact test) and transmission when compared to the common temperature of 28°C used for most VC studies.
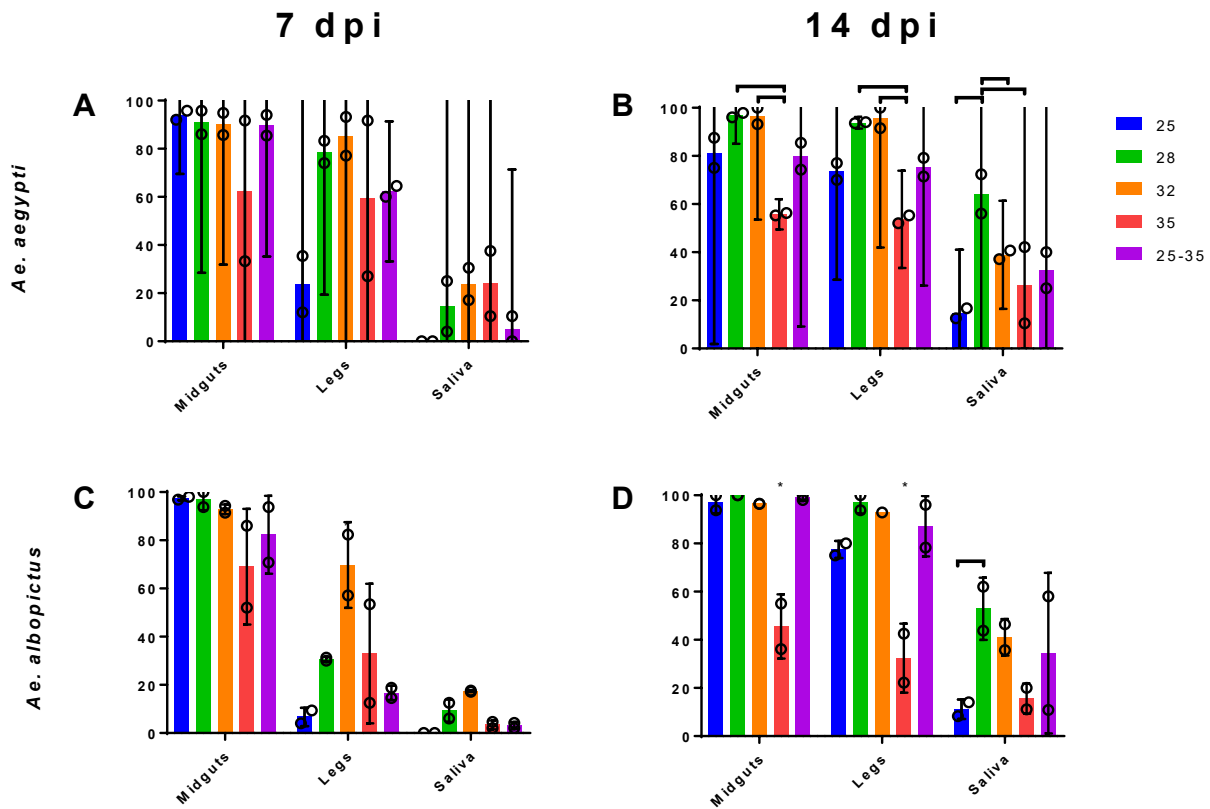
Figure 2.2 Constant extrinsic incubation temperature impacts of *Aedes* vector competence in a unimodal distribution. EIT effects on percent infected (midgut), dissemination (legs), and transmission (saliva) of *Aedes aegypti* (A & B) and *Aedes albopictus* (C & D) for 7 (A & C) and 14 (B & D) dpi groups. Black circles represent the mean of each experimental replicate, mean with 95% CI (p < 0.05, Two-tailed Fisher's exact test).

We next assessed how temperature affects systemic mosquito infection at a finer timescale, sampling exposed mosquitoes at 3, 5, 7, 10, and 14 days post exposure (Figure 2.3). The proportion of mosquitoes with detectable ZIKV infection, dissemination, and transmission increased with time. Detectable infection in *Aedes aegypti* took 5 dpi to establish at diurnal temperatures, whereas all other EIT groups took 3 dpi to establish (Figure2.3A). ZIKV infection took longer to establish itself to a detectable level at 25°C (5 dpi) than in any other EIT group (3 dpi) in *Aedes albopictus* (Figure 2.3D). *Aedes aegypti* reached peak dissemination at 7 dpi, and peak transmission at 10 dpi (Figure 2.3B & 2.3C), compared to 10 and 14 dpi in *Aedes albopictus* (Figure 2.3E & 2.3F).
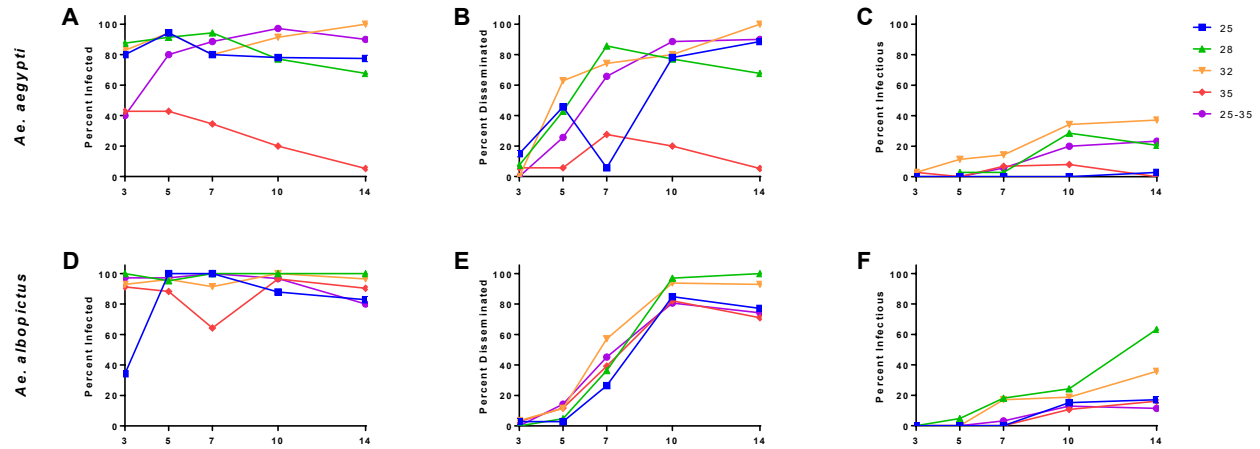
Figure 2.3 *Aedes aegypti* **mosquitoes reach peak transmission 4 days faster than** *Aedes albopictus.* Temperature effects on systemic infection dynamics in *Aedes aegypti* (A-C) and *Aedes albopictus* (D-F) over time. The relationship between days post infection (3, 5, 7, 10 & 14) and the percent of mosquitoes infected (A & D, ZIKV positive midgut), disseminated (B & E, ZIKV positive Legs), and transmitting (C & F, ZIKV positive saliva) out of the total mosquitoes exposed to ZIKV at 25°C, 28°C, 32°C, 35°C, and 25°C-35°C.

2.3b Between host genetic diversity

We assessed the effect of temperature on ZIKV genomic sequences by conducting next-generation sequencing (NGS) on 3 biological replicates of matching tissues (midgut, legs, and saliva) from *Aedes aegypti* and *Aedes albopictus* harvested at 14 days post exposure. First, we identified that there were low frequency SNVs in the input ZIKV stock population. On average we observed 19 SNVs found in the input population with a frequency ranging from 0.010 to 0.344; the majority of these SNVs were in the envelope (E), non-structural (NS) proteins, NS1, and NS5 coding sequences (Figure 2.4A & 2.4B, Supplemental Table 2.1). Next, we identified the SNV distribution in *Aedes* mosquitoes using the mean SNV. In *Aedes* aegypti held at constant temperatures, for each constant EIT group, the largest proportion of total SNVs was found within the NS1 coding sequence (20%-37%), followed by NS5 (18-34%) and E (9-21%) (Figure 2.4C & Supplemental Figure 2.1). In contrast, in the diurnal EIT group, the majority of variants were in NS5 (34%) followed by NS1 (27%) and finally E (15%) (Figure 2.4C &

Supplemental Figure 2.1). In *Aedes albopictus,* all EIT groups except 32°C had the majority of

SNVs occurring in NS1 (27-43%) followed by NS5 (22-27%) and E (13-16%) whereas 32°C

accumulated the majority of SNVs in NS5 (26%) followed by NS1 (22%) and E (15%)

(Figure 2.3C & Supplemental Figure 2.2). For both *Aedes* species, the 32°C EIT group

accumulated the most SNVs with 26% of all SNVs identified in *Aedes aegypti* and 27% in

Aedes albopictus (Figure 2.4C, Supplemental Figure 2.1 & Supplemental Figure 2.2).
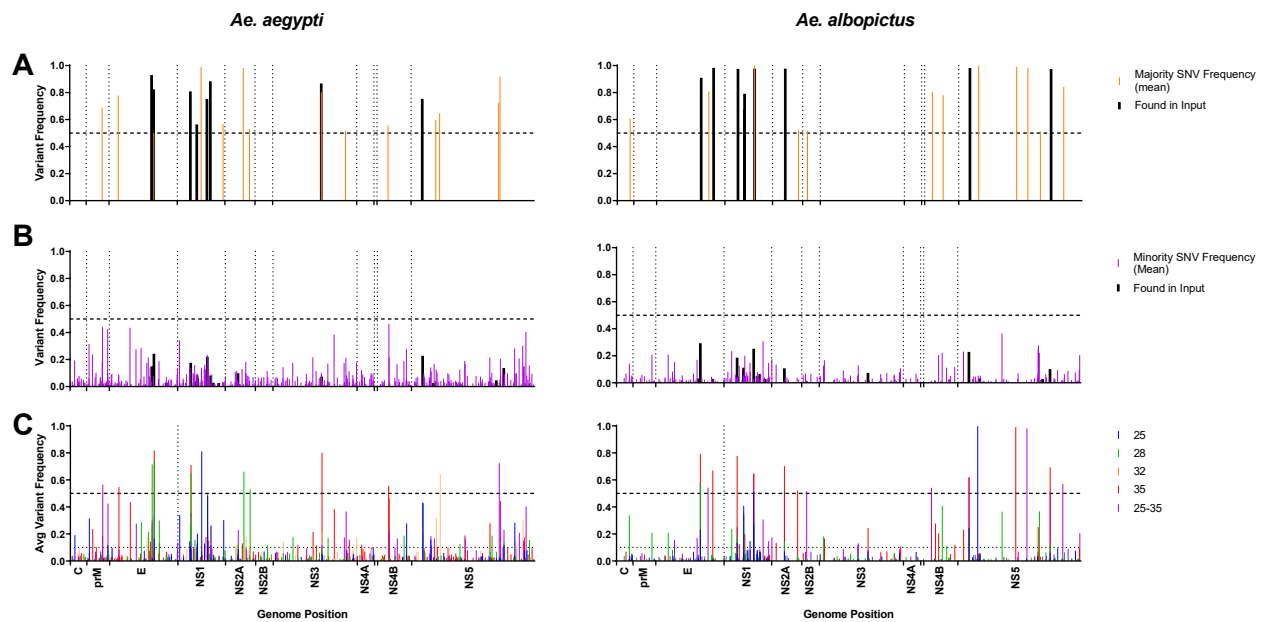
Figure 2.4 SNVs accumulate most often in the NS1 for constant temperatures and NS5 for diurnal groups.
Interhost variant analysis, majority SNV frequency (mean of all tissues per replicate) (A), minority SNV frequency
(mean of all tissues per replicate) (B), across the ZIKV CDS over all 14 dpi samples (midgut, legs, saliva) for *Aedes
aegypti* and *Aedes albopictus*. Temperature effects on SNV frequency (mean of all tissues per replicate) across the
CDS for *Aedes aegypti* and *Aedes albopictus* at 25°C, 28°C, 32°C, 35°C and alternating between 25°C-35°C (C).

To further explore the impact of temperature on ZIKV populations in *Aedes* mosquitoes,

we measured genetic selection (characterized by d_N/d_S), nucleotide diversity, richness,

complexity (Shannon entropy), and divergence (F_{ST}) (Figure 2.5). In both *Aedes* species, d_N/d_S

was significantly lower in ZIKV within diurnal-exposed mosquitoes compared to those held at

constant temperatures, and the diurnally held mosquitoes were the only group with a d_N/d_S value

much lower than 1 (Figure 2.5A). Nucleotide diversity in ZIKV was significantly increased (p-value less than 0.05) compared to the input ZIKV population in diurnal groups for both *Aedes* species, and at 28°C in *Aedes aegypti* and 35°C in *Aedes albopictus* (Figure 2.5B). Similar levels of richness, complexity, and divergence were observed in all mosquitoes and EIT groups (Figure 2.5C – 2.5E, p-value greater than 0.05).
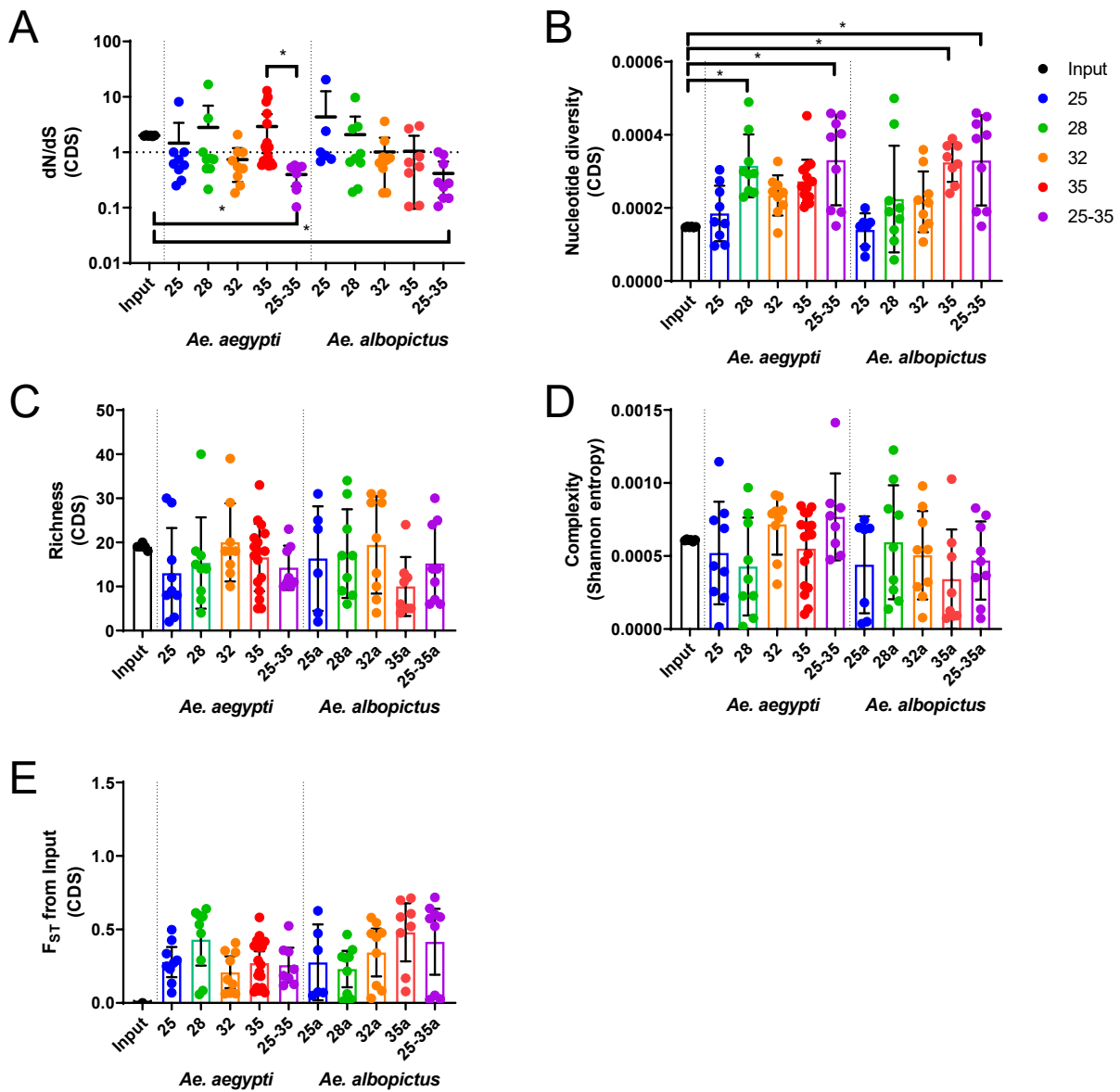
Figure 2.5 Diurnal EIT drives purifying selection across the ZIKV coding sequence. Characterization of ZIKV population diversity using d_N/d_S (A), nucleotide diversity (B), richness (C), complexity (D), and divergence from input

population (E) as markers of population diversity. Kruskal-Wallis test with Dunn's correction ($*$ $p < 0.05$), Mean and 95% CI graphed.

ZIKV variants generated during systemic mosquito infection were distributed evenly across the CDS (Figure 2.4). To assess the possibility that coding region-specific population genetic profiles may exist, we assessed complexity, nucleotide diversity, and selection in each viral protein coding region for all tissues combined (Figure 2.6). Compared to the input ZIKV population, the complexity of the NS1 sequences were reduced in both *Aedes* species (Figure 2.6A & 2.6D). By assessing nucleotide diversity, we saw that diurnal EIT increased nucleotide diversity in NS5, and 32°C EIT increased diversity in NS3 for both *Aedes* species (Figure 2.6B & 2.6E). In *Aedes aegypti,* 28°C and 35°C EIT increase diversity in E and NS1 (Figure 2.6B). Using d_N/d_S to characterize genetic selection, we saw that E and NS1 coding regions were under positive selection for both *Aedes* species as indicated by a d_N/d_S ratio greater than 1 (Figure 2.6C & 2.6F). Generally, in *Aedes aegypti,* moderate temperature of 28°C had the highest indication of positive selection within E and NS1 (Figure 2.6C) and in *Aedes albopictus* the low temperature of 25°C had the highest selective pressure in these regions (Figure 2.6F). Interestingly, we observed that NS5 underwent weak purifying selection for all constant temperature groups, while diurnal temperature groups exhibited strong purifying selection (Figure 2.6C & 2.6F). Each EIT group had a d_N/d_S near 1 for C, prM NS2A, NS2B, NS4A, NS4B (data not shown).
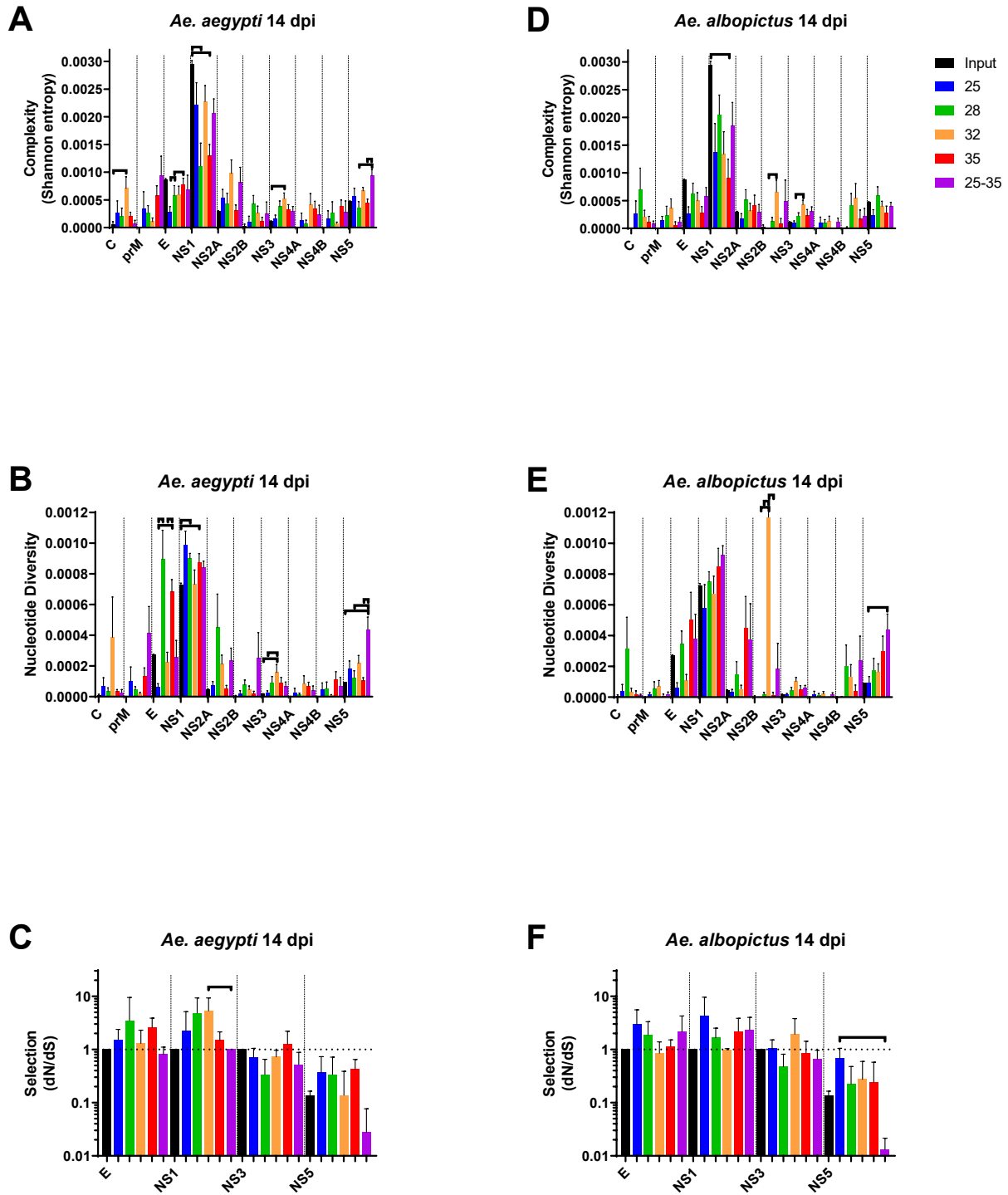
Figure 2.6 Between host ZIKV CDS region specific genetic diversity Characterization of ZIKV population diversity in 14 dpi *Aedes aegypti* and *Aedes albopictus* for each EIT group at C, prM, E, NS1, NS2A, NS3, NS4A, NS4B, and NS5 protein coding regions. Complexity (A), nucleotide diversity (B), and selection (C) as markers of population diversity. Mean and 95% CI graphed, Kruskal-Wallis test with Dunn's multiple comparison test ($p < 0.05$)

Next, we characterized EIT impacts on consensus-altering mutations (Figure 2.4A &

Figure 2.7A). We observed that 35°C resulted in the largest accumulation of non-synonymous

consensus changes across the genome in both *Aedes* species, while 25°C had the lowest

accumulation of non-synonymous and synonymous consensus changes in both *Aedes* species

(Supplemental Figure 2.3). Interestingly, diurnal EIT groups had the greatest accumulation of

synonymous mutations in both *Aedes* species compared to all constant/other EIT groups

(Supplemental Figure 2.3). The majority of non-synonymous consensus changes occurred in E

(43 total combined) and NS1 (58 total combined), while the majority of synonymous consensus

changes accumulated in NS5 (36 total combined, Figure 2.7A). Because structural and non-

structural proteins play different roles in ZIKV infection, we sought to determine how consensus

changes differed between the two genome regions (Figure 2.7B). At 35°C, *Aedes aegypti* has the

most consensus changes in both structural (n=20) and non-structural regions (n=18, Figure

2.7B). In *Aedes albopictus,* while 35°C has the most consensus changes in the structural region

(n=9), the diurnal group had more consensus changes occurring in the non-structural region

(n=21, Figure 2.7B). To determine whether these consensus changes tended to occur in one

tissue type, we assessed consensus change according to their tissue of origin (Figure 2.7C). In

general, 35°C EIT groups accumulated the most majority variants (variant frequency of 0.5 or

greater, also known as consensus change) in the midgut, while moderate to high temperatures

(28°C - 35°C) have the most consensus changes and at a similar amount in legs and saliva.

Interestingly, we saw that there were significantly more majority variants in *Aedes albopictus* for

diurnal temperatures than any constant temperature group (Figure 2.7C).
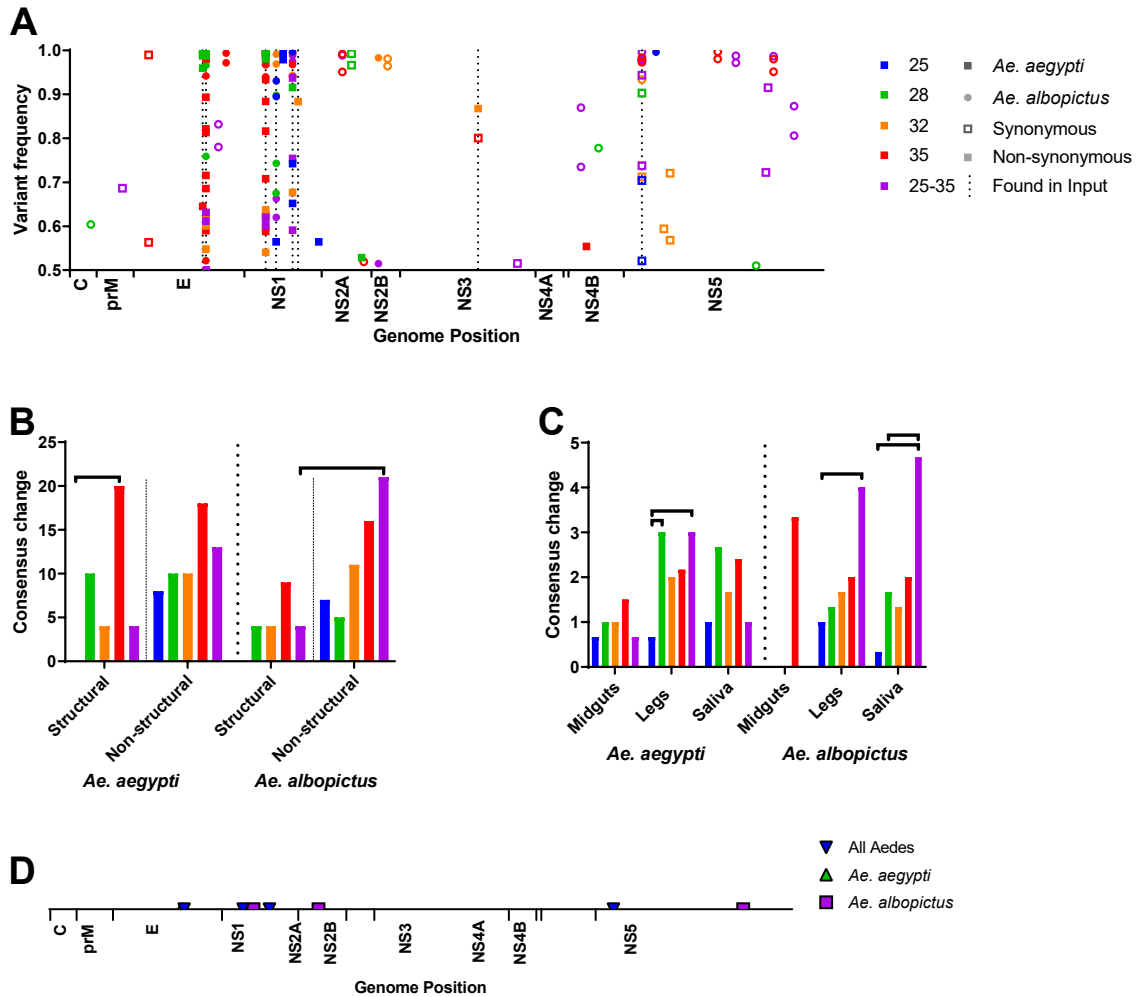
Figure 2.7 Majority variants are impacted by temperature, species, and coding sequence location. 14-dpi majority variants across the ZIKV CDS for *Aedes aegypti* (square) and *Aedes albopictus* (circle) for 4 constant EIT and diurnal EIT (A), the distribution of majority variants across structural (C, prM, E), and non-structural (NS1-NS5) protein coding regions (B), and intrahost accumulation of majority variants by tissue type (C). Reoccurring majority variants from multiple biological replicates and multiple EITs in either *Aedes aegypti*, *Aedes albopictus*, or both (D).

Further characterizing consensus changes, we found 8 consensus changes that occurred in multiple biological replicates at different temperatures (Figure 2.7D, Supplemental Table 2.1). Of these 8 SNVs, 3 nonsynonymous and 1 synonymous SNVs were found in both *Aedes aegypti* and *Aedes albopictus* (L330V E, W98G NS1, M220T NS1, and G83 NS5) samples. The remaining 4 SNVs were comprised of 1 non-synonymous mutation (T315I E) unique to *Aedes aegypti* and 3 mutations unique to *Aedes albopictus*: 1 non-synonymous mutation and 2

synonymous mutations (K146E NS1, I94 NS2A and F682 NS5). These eight variants were aligned to 150 complete ZIKV sequences from nature and assessed for sequence identity. Of these, L330V E was found 100% in this alignment indicating the 330L E in our stock virus is most likely a adaptation after isolation, K146E NS1 2%, and I94 NS2A 0.7%. W98G NS1, M220T NS1 G83 NS5, T315I E, and F682 NS5 were all novel mutations that were found uniquely in our ZIKV PRVABC59 population (Supplemental Table 2.1). Additionally, all 8 of these variants were found as minority variants (variant frequency less than 0.5) in the input virus population with mean frequencies from ~0.01 – 0.35 (Supplemental Table 2.1).

We used competition assays to determine the fitness effects of the eight consensus-changing mutations that arose during systemic infection (Figure 2.7D, Supplemental Table 2.1) in mosquitoes under low (25°C) and high (35°C) EITs to determine the likelihood of a given mutant virus rising in frequency. In comparing mosquitoes held at an EIT of 25°C (Figure 2.8A) and 35°C (Figure 2.8B), engineered viruses significantly increases (p-value less than 0.001, Two-tailed Fisher's exact test) the overall rate of fixation (32% vs 11%) in 35°C compared with 25°C. Likewise, engineered generally outcompete the reference virus at a significantly higher rate (p-value less than 0.002, Two-tailed Fisher's exact test), reaching 90% frequency or higher 37% of the time (Figure 2.8B). In orally exposed *Aedes aegypti* bodies, the NS1 M220T mutant clone had significant (p-value less than 0.05, Kruskal-Wallis and Dunn's, Figure 2.8B) competitive fitness advantage 14 days after blood feeding over the wildtype ZIKV-PR-IC . Conversely, the envelope T315I mutant had significantly decreased (p-value less than 0.05, Kruskal-Wallis and Dunn's, Figure 2.8A) competitive fitness effects compared to ZIKV-PR-IC at 25°C.
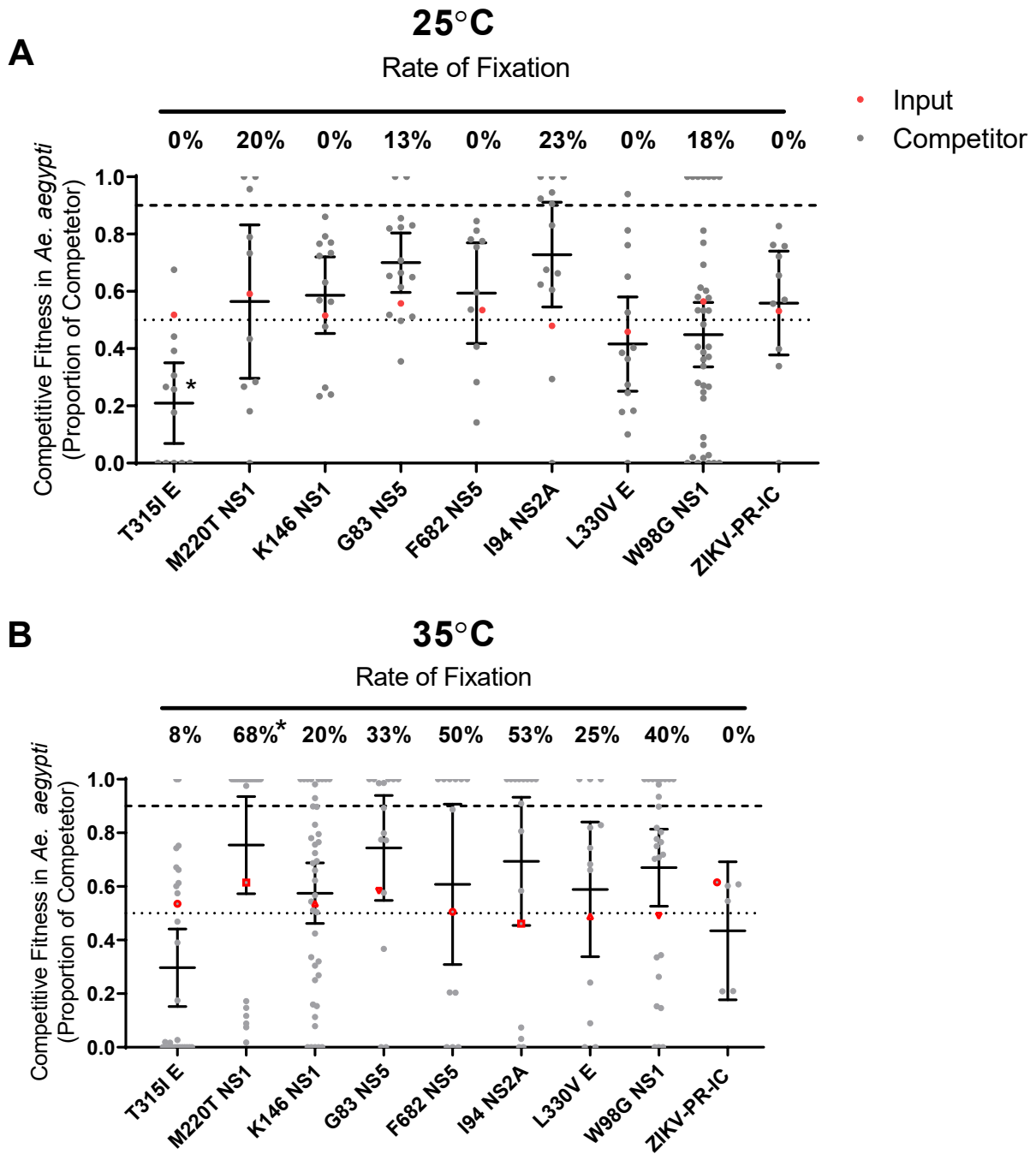
Figure 2.8 EITs of 35°C increase the overall rate of fixation of engineered ZIKV clones. Each mutation was engineered into a ZIKV-PR-IC and mixed with a ZIKV-REF virus. The proportion of each competitor (grey, mean with 95% CI, *p-value less than 0.05 compared with ZIKV-PR-IC , Kruskal-Wallis and Dunn's) and rate of fixation (*p-value less than 0.05 compared with ZIKV-PR-IC , Two-tailed Fisher's exact test) was determined from mosquito bodies at 14 dpi for *Aedes aegypti* mosquitoes held at constant EIT's 25°C (A) & 35°C (B). Fixation indicate that 100% of the sequenced nucleotides were from the competitor virus, initial viral inoculum in red symbols.

Of the eight input minority variants that arose to majority frequency during systemic mosquito infection, 4 were found in both *Aedes* species (Supplemental Table 2.1). Therefore, we used these 4 mutations to assess EIT effects on variant frequency during systemic mosquito infection (Figure 2.9). We identified two mutations L330V E and W98G NS1 that appear to have temperature-driven effects, which increased variant frequency at 28°C, 35°C (L330V), and diurnal groups (L330V and W98G, Figure 9A-9B, Figure 9E-9F) in *Aedes aegypti*. Surprisingly, in *Aedes albopictus*, we see the same temperature (28°C and 35°C) impacting variant frequency for L330V, however these effects are only observed in the legs and saliva, whereas the mutation does not show any fitness advantage in the midgut tissue of *Aedes albopictus* (Figure 2.9E). W98G is impacted differently in *Aedes albopictus* and only shows signs of fitness adaptation in 35°C EIT groups, suggesting that host species greatly impacts variant frequency (Figure 2.9F). The M220T and G83 variant frequencies are minimally impacted by constant EIT groups (Figure 2.9C-2.9D, Figure 2.9G-2.9H). However, diurnal temperature groups appear to drive fitness advantage in the legs of *Aedes aegypti* (Figure 2.9C-2.9D) as well as legs and saliva in *Aedes albopictus* (Figure 2.9G-2.9H), suggesting that some variants have fitness advantages at constant temperatures, while others only have a fitness advantage at diurnal temperatures that may better mimic environmental conditions. Thus, we have determined that there are both temperature specific effects which appear to provide a fitness advantage, and species-specific effects on these variants.

Ae. aegypti

Ae. albopictus

A L330V (E)

E L330V (E)

B W98G (NS1)

F W98G (NS1)

C M220T (NS1)

G M220T (NS1)

D G83 (NS5)

H G83 (NS5)

Legend: 25, 28, 32, 35, 25-35
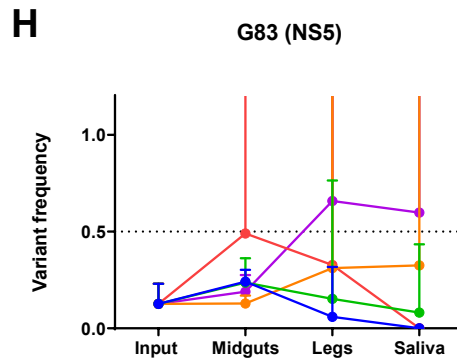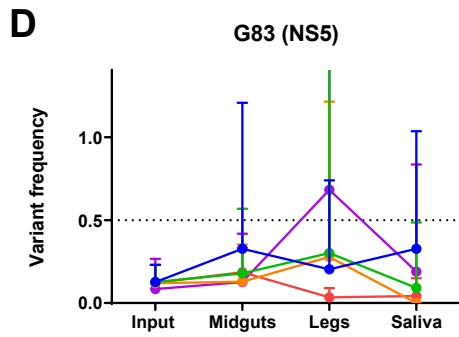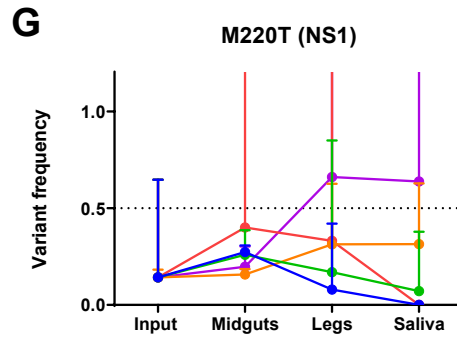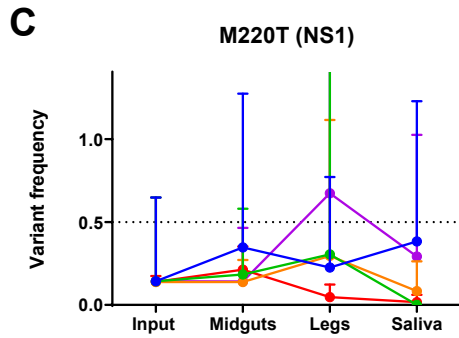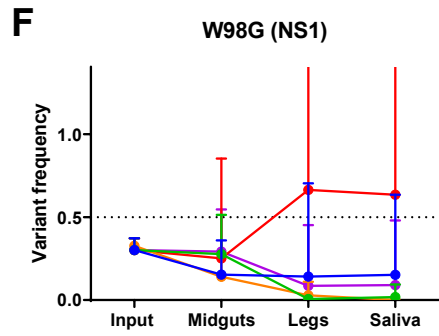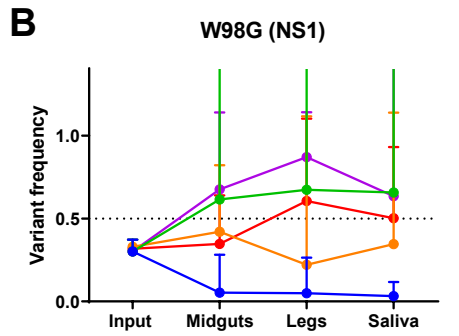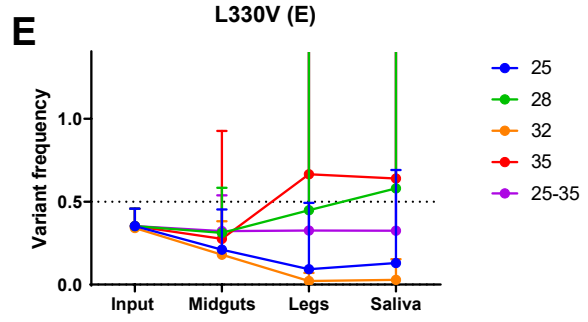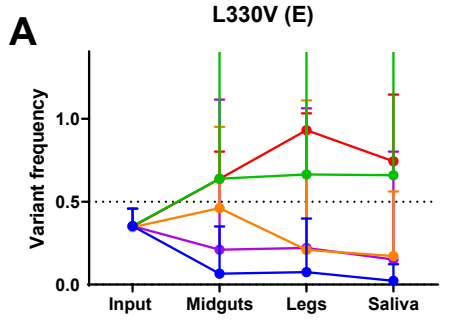
Figure 2.9 EIT and species, impact ZIKV variant frequency during systemic *Aedes* infection. Assessment of extrinsic incubation temperature for 3 non-synonymous mutations, L330V E (A&E), W98G NS1 (B&F), and M220T NS1 (C&G), and 1 synonymous mutation G83 NS5 (D&H) introduced by the input ZIKV population in *Aedes aegypti* (A-D) and *Aedes albopictus* (E-H). Mean frequency is graphed for three biological replicates. Mean and 95% CI graphed.

2.3c Intrahost genetic diversity

Intrahost viral population diversity in the midguts, legs, and saliva of mosquitoes held at varying temperatures was characterized to assess the possibility that altered temperature disproportionately influenced virus population structure in a particular tissue (Figure 2.10). In general, population diversity appears to be unimodally impacted by constant temperatures in the midgut, with 28°C and 32°C having the highest levels of richness, complexity, and divergence (p-value greater than 0.05) (Figure 2.10A-2.10C). After dissemination from the midgut, the unimodal distribution is not as clear and other factors associated with dissemination (bottlenecks, founder effect) may be overpowering the EIT effect. Interestingly, the ZIKV populations in mosquitoes subject to diurnal temperature tended to have lower diversity during midgut infection but increased genetic diversity during systemic infection that resulted in some of the highest levels of richness, complexity, and nucleotide diversity in saliva-associated virus (Figure 2.10A-2.10C). Next, we used SNV carry-through (the proportion of SNVs from the input population that are passed to the next tissue) as a proxy to identify EIT impacts on bottlenecks (Figure 2.10D). In *Aedes aegypti,* we saw that SNV carry-through in the midgut was most efficient at 32°C and 35°C whereas *in Aedes albopictus* we saw that midgut carry-through was most efficient at 25°C and 28°C (Figure 2.10D), suggesting that the midgut infection barrier was impacted by EIT in a species-specific manner.
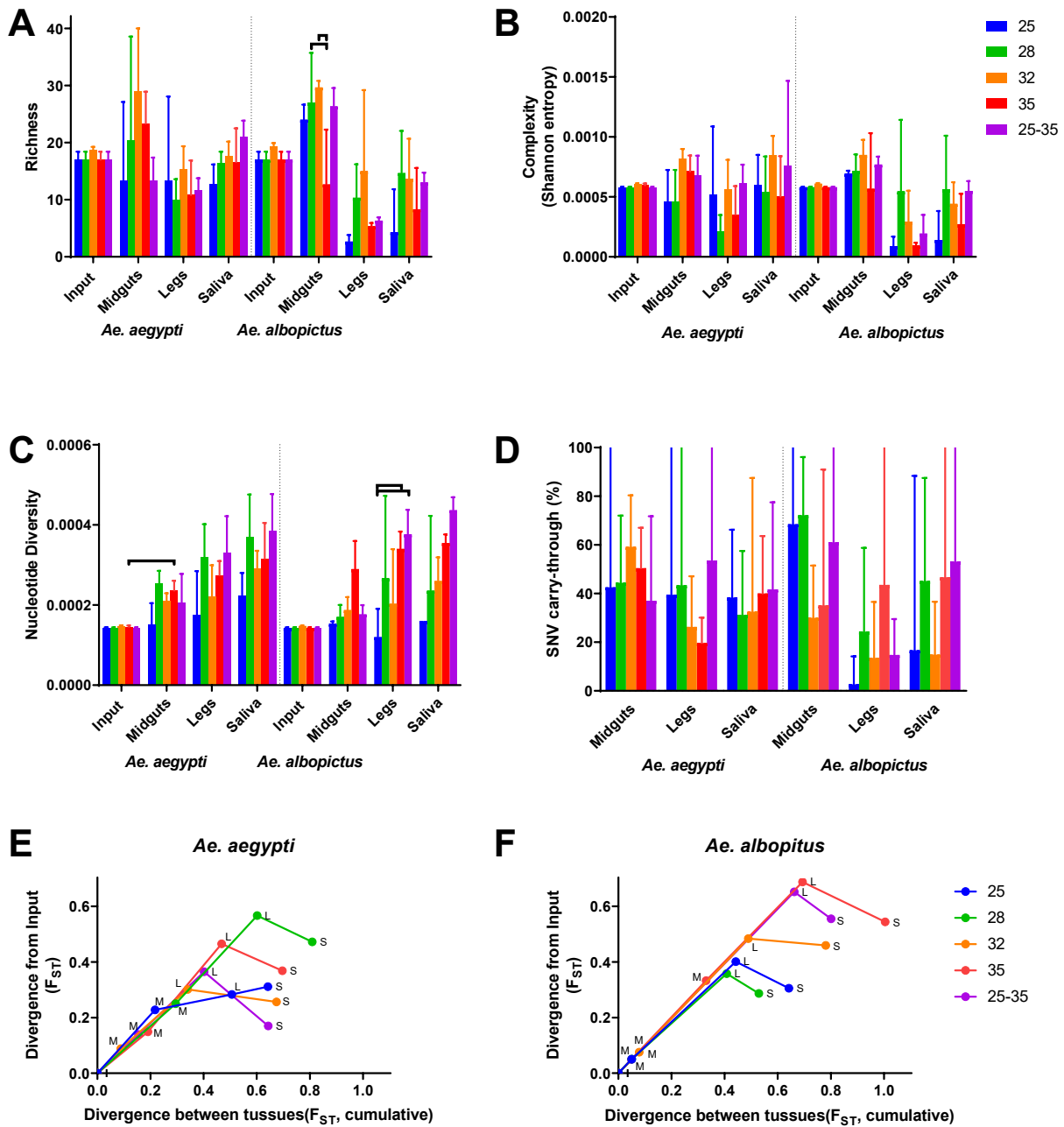
Figure 2.10 EIT and species, impact ZIKV variant frequency during systemic *Aedes* infection. Intrahost genetic diversity was characterized by measuring richness (A), complexity (B), nucleotide diversity(C). Bottlenecks were assessed by analyzing SNV carry-through (D), and divergence from input population (y-axis) and cumulative divergence between tissues (x-axis) (E-F) as markers of population diversity. (A-D) 2-way ANOVA with Tukey's (p-value less than 0.05), Mean and 95% CI graphed. Midguts (M), legs (L), Saliva (S).

We assessed the extent to which the ZIKV populations diverge from the input ZIKV

population during systemic mosquito infection. To do this, we looked at divergence from the

input in each tissue, and the cumulative divergence between tissues during infection (Figure 2.10 E-F). EIT did not appear to impact divergence in a consistent way in the main vector, *Aedes aegypti*, but in the less efficient vector, *Aedes albopictus*, exposure to higher temperatures tended to promote divergence once the infection was established. Generally, ZIKV diverged more in the midguts of *Aedes aegypti* than *Aedes albopictus*. The 28°C EIT group diverged more than any other EIT group in *Aedes aegypti* (Figure 2.10E). In *Aedes albopictus*, 35°C had similar levels of divergence in the midgut compared to *Aedes aegypti*, however the other EIT groups minimally diverged during initial midgut infection (Figure 2.10F). In both species, divergence was greatest when the population disseminated from the midgut to the legs and divergence was reduced in the saliva. These data provide evidence that divergence from the founding population was increased in the midgut and legs of both species and that there was a decrease in reduction going from legs to saliva, possibly driven by bottleneck and purifying selection removing novel minority variants.

2.3d Intrahost selective pressures

Last, we assessed EIT impacts on natural selection by estimating d_N/d_S for each EIT group, for the entire CDS, and for the structural and nonstructural regions (Figure 2.11). Our input population had a d_N/d_S ratio of 1.75 for the CDS, 3.11 for the structural region, and 0.95 for the non-structural regions. This indicates that the structural regions of our input population were under positive selection during its preparation, whereas the non-structural regions were not under positive selection (Figure 2.11). Interestingly, when ZIKV was exposed to diurnal fluctuating temperatures, it was under strong purifying selection (d_N/d_S less than 1) in both *Aedes* species (Figure 2.11E & 2.11J), whereas all constant temperatures caused ZIKV d_N/d_S generally near or

above 1 (Figure 2.11). While selective environments are complex (Figure 2.11), we see that in the saliva, 25°C and 32°C EIT groups had a d_N/d_S that neared 1, decreasing from the input in both *Aedes* species (Figure 2.11A, 2.11C, 2.11F, 2.11H). Conversely, 28°C and 35°C EIT groups maintain or increased d_N/d_S when compared to input (Figure 2.11B, 2.11D, 2.11G, 2.11I). This suggests that these temperatures may be important in the selective environment that shapes the ZIKV population.
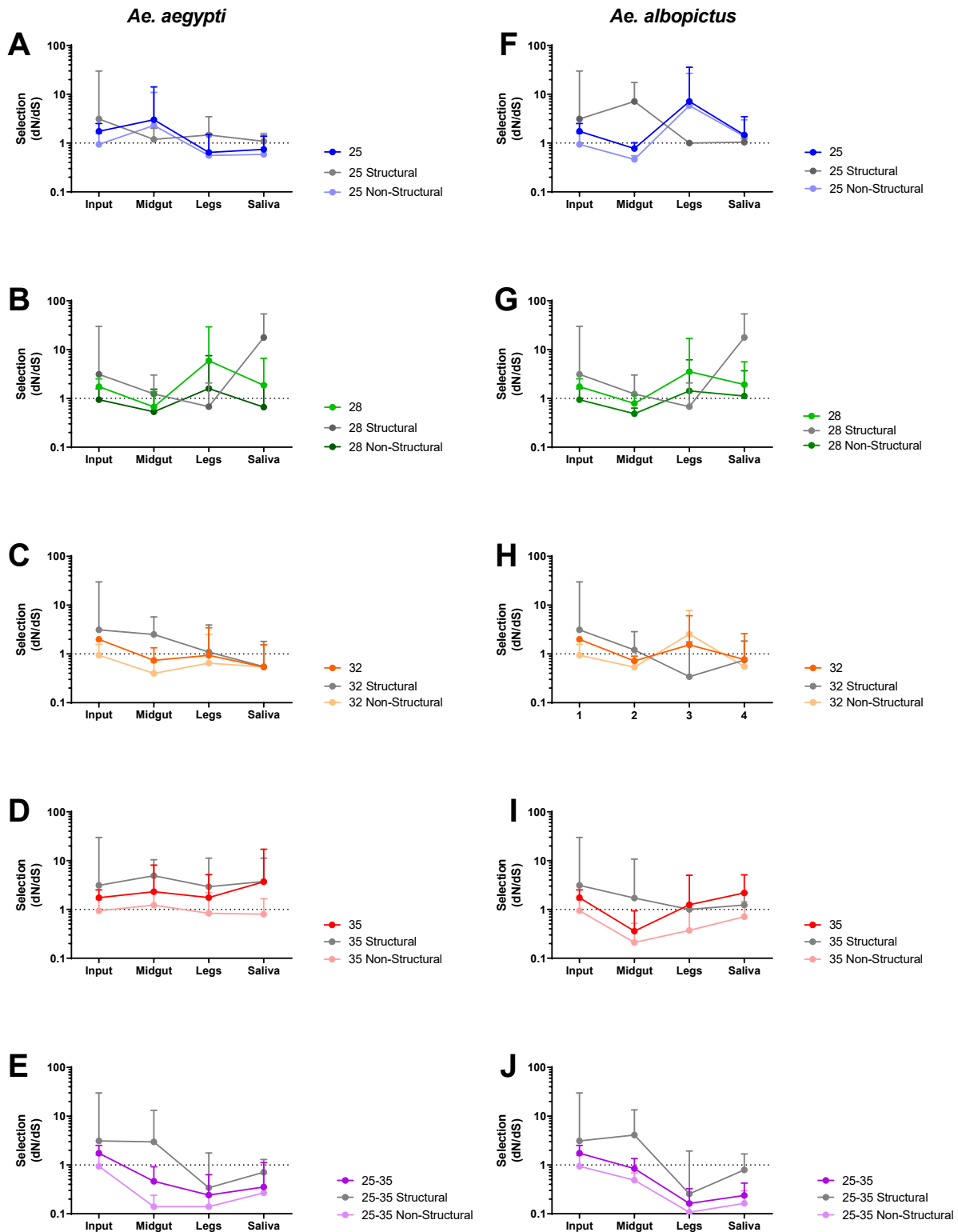
Figure 2.11 Intrahost selection under strong purifying selection during diurnal EIT. Strength of host and EIT selection on virus population (CDS, structural, and non-structural regions), estimated by d_N/d_S (mean with 95% CI)

for *Aedes aegypti* (A-E) and *Aedes albopictus* (F-J), during systemic infection for each of the 5 EIT groups (25°C, 28°C, 32°C, 35°C and 25°C -35°C).

2.3e Statistical Modeling

The results of our generalized linear model were consistent with those of the study. We evaluated both VC and measures of genetic diversity. As described previously, our original models follow the base structure:

$$response \sim \exp\left[\beta_1(days) * \beta_2(temp)^2 * \beta_3(species) * \beta_4(tissue)\right]$$

With all other predictors held constant, successful infection (midgut) has a significantly negative effect on consensus changes when compared to successful dissemination (legs) ($p < 0.05$). Dissemination efficiency (proportion of infected that were disseminated and disseminated that were transmitted) is positively impacted by dpi and temperature (linearly) and is reduced when comparing *Aedes albopictus* species against *Aedes aegypti*; each with all other variables held constant. Richness was positively affected by both successful infection (midgut) and transmission (saliva) when compared to dissemination (leg tissue) and had a slight significant concave (non-linear) effect of temperature. Complexity is positively impacted by an interaction between dpi and temperature, and negatively impacted by an interaction between dpi and species. Temperature has a positive linear effect on nucleotide diversity as well as a significant negative effect when nucleotide diversity is measured in midgut versus legs.

As previously mentioned, VC was evaluated with the following base structure for each tissue response.

$$response \sim \exp\left[\beta_1(days) * \beta_2(temp)^2 * \beta_3(species)\right]$$

The results are consistent with those seen in Figure 2.2, with the interactions between all factors significantly affecting mosquito infection (midgut). Dissemination (leg) has similar

results, with significance in almost all interactions. Transmission (saliva) had the most variation in significance. Higher temperatures and dpi were the only significant factor, negatively impacting VC in the saliva. These evaluations demonstrate the complexity of interactions and effects that external factors (days post infection, temperature, and species) have on the infection dynamics of ZIKV. The results of this evaluation are supported by the associated results of the study.

2.4 Discussion

2.4a Extrinsic incubation temperature driven unimodal distribution of vector competence

Vector competence is largely determined by barriers to infection and escape from mosquito midgut and salivary glands [189, 246]. Our results show that EIT impacted this infection and escape mechanism in a unimodal distribution, where extreme low 25°C and high 35°C had the lowest VC, while moderate temperatures of 28°C and 32°C had peak VC over time (Figure 2.2). Similarly, when assessing rates of infection and dissemination, we observed the constant moderate temperature groups (28°C and 32°C) peaked first in dissemination and transmission over time (Figure 2.3). These constant EIT results are expected as they agree with previous studies [102] and mechanistic models predicting the ZIKV thermal optimal limit of 29°C for *Aedes aegypti* [102, 247]. However, when assessing the VC results for the diurnal group of 25°C-35°C, alternating for a mean daily temperature of 30°C, our VC results were consistently lower than the moderate temperatures of 28°C and 32°C. This suggests that while constant EIT groups appear to fit the unimodal mechanistic model previously described [102], diurnal temperatures do not fit the model. Of note in this study is the use of 25°C-35°C

as the diurnal temperature group range. We chose the mean low and high daily temperature during the 2014 ZIKV outbreak in Brazil, and while this mimics the daily max fluctuation it does not account for microclimates in which mosquito vectors may be found. This highlights the importance of accounting for environmental diurnal temperature [221, 248, 249] and microclimate [250, 251] changes in future mechanistic models. From our data we identified that moderate temperatures of 28°C and 32°C are good representations for capturing peak ZIKV VC during extrinsic incubation (Figure 2.2 & Figure 2.3). However, these temperatures may be an overestimate of what is observed in nature where diurnal temperature fluctuations and microclimates are found. These data combined show that moderate temperatures positively impact vector competence. The rate of transmission, and diurnal EIT, which may better mimic efficiency of dissemination observed in nature, are consistently lower than the moderate constant EIT groups. *Aedes albopictus,* which is generally a less efficient vector at 7 dpi, is as good a vector as *Aedes aegypti* when it is exposed to an EIT of 32°C.

2.4b Species dependent impacts and extrinsic incubation temperature impacts on viral genetic diversity

In NGS analysis, richness is used as a marker for variant expansion and reduction, while complexity is used to identify diversifying selection by using Shannon entropy to assess genetic complexity of an allele at a specific locus. During infection, richness and complexity were not significantly different among species or between EITs (Figure 2.5), instead, founder variants from infection increase in frequency during replication, which appears to be driven by selection (Figure 2.5A, Figure 2.6C & 2.6F). As temperature increases in both *Aedes* species we see an increase in nucleotide diversity (variant frequencies across the CDS), and that the constant EITs

are near or above a d_N/d_S of 1, while diurnal EIT is significantly lower (Figure 2.5A & 2.5B). As there is evidence that suggests arbovirus population diversity is directly impacted by specific host species [187], we aimed to account for this effect by using both *Aedes aegypti* and *Aedes albopictus* as vectors to determine EIT impacts of ZIKV population dynamics. Interestingly, *Aedes aegypti* appears to be under strong positive selection at 35°C (Figure 2.11D), whereas *Aedes albopictus* is not (Figure 2.11). This suggest that differences in mosquito species [215, 217, 218] and environmental response not only impact vector competence [102], but directly impact genomic diversity. These data indicate that genomic diversity in mosquitoes is driven by selection, and that the viral population, under constant EITs must adapt through positive selection to new host environments, while diurnal temperatures, which mimic a natural environment, are under purifying selection (Figure 2.11).

We observed stochastic reductions and expansion of genetic diversity across the protein coding sequence regions. We observed high levels of complexity in our input virus at NS1, but during replication there were stochastic reductions in complexity at NS1 (Figure 2.6A & 2.6D). All other protein coding regions had low complexity in viral input population (less than 0.001), but during infection complexity increased for all protein coding regions with the exception of E. The stochastic reduction of complexity in NS1 and the increase of complexity across the remaining CDS suggests that there may be a maximum threshold for complexity that can be maintained across the genome during systemic infection. This is likely due to stochastic reduction caused by bottlenecks of infection suppressing complexity of the ZIKV population [189, 246]. Nucleotide diversity, which has stochastic increases across the genome, had a significant increase in diversity in E and NS1 at 28°C and 35°C for *Aedes aegypti* and in NS3 at 32°C and NS5 at diurnal EITs for both *Aedes* species. This indicates that there are stochastic

increases in diversity unique to both *Aedes* species, and a diurnally driven diversity in NS5. However, it's important to note that the input ZIKV population had increased complexity (Figure 2.6A & 2.6D) and diversity (Figure 2.6B & 2.6E) in the envelope and NS1 coding regions, which may lead to the increased complexity and diversity in both *Aedes* species. It is likely that these increases in complexity and diversity are driven by the input population diversity and overshadow any temperature impacts that may be present.

While there may be some stochastic increase in diversity, we can see that these changes are primarily driven by positive and purifying selection (Figure 2.6B & 2.6E, Figure 2.11). The majority of the protein coding regions have a d_N/d_S ~1 which is suggestive of neutral evolution. However, this is likely due to some neutral changes along with a mix of positive an purifying selection offsetting one another. Interestingly, we see that E and NS1 appear to be under positive selection (d_N/d_S greater than 1), whereas NS5 is under weak purifying selection (d_N/d_S less than 1). These regions being impacted by the selective environment are interesting for the role that they play in viral infection and replication. E and NS1 are essential for viral infection [110, 142, 252] and by undergoing positive selection in these regions they may be adapting to host selective pressures. Conversely, NS5 is essential for viral replication [110] and therefore mutations are even more deleterious - being under purifying selection decreases the chance of deleterious mutation arising in the methyltransferase and RNA-dependent RNA polymerase. Our findings suggest that E and NS1 are adapting to the selective environments that the EITs are creating, acting as a proxy for environmental host condition, while NS5 is highly conserved and maintaining strict replication functionality.

Consensus changes are of interest as these may indicate haplotypes that have a fitness advantage over the rest of the competing viral genome population [253]. We observed that most

non-synonymous changes occurred in NS1 and E, while most synonymous variants accumulated in NS5. In *Aedes aegypti,* 35°C EIT led to the highest accumulation of consensus changes in both structural and non-structural region, however in *Aedes albopictus,* the diurnal group had more consensus changes occurring in the non-structural region. This suggests that 35°C is a strong selective pressure, leading to the accumulation of consensus changes in the structural and non-structural gene regions.

2.4c Increased incubation temperatures drive viral variant fixation in mosquitoes

Further analysis of the consensus changes led us to detect 8 consensus ZIKV mutations (5 non-synonymous and 3 synonymous) in multiple mosquitoes that were found in the input population at low frequencies (0.01- 0.35). Of the consensus changes found in both *Aedes* species, L330V E (Supplemental Figure 2.4A) is imbedded in domain III of the envelope protein, which is known to play a role in host cell receptor binding for viral entry [254]. W98G NS1 (Supplemental Figure 2.4B) is a surface exposed aromatic to aliphatic amino acid change on the wing section of NS1. The wing structure is believed to contribute to cellular membrane association [255]. M220T NS1 (Supplemental Figure 2.4C) replaces a sulfur containing side group with a hydroxylic side group and is located on the loop surface the NS1 172-352 homodimer [256]. G83 NS5 (Supplemental Figure 2.4D) is a synonymous mutation found in the middle the NS5 methyltransferase domain active site. We believe that E-L330V may be a reversion to wildtype sequence, being selected during systemic mosquito infection. The lack of reversion during *in vivo* competition studies (Figure 2.8A & 2.8B) suggests that this mutation may be a cell culture adaptation, which is in agreement with previously identified results [257]. Of the remaining 7 mutations, only NS1-M220T had significant fitness advantage during in vivo

competitive fitness experiments (Figure 2.8B). However, NS1-M220T variant frequency during

Aedes aegypti systemic infection (Figure 2.9C) does not appear to have fitness advantages at

35°C, as in the competitive fitness experiment. This may be explained by NS1-M220T being

paired with other variants, which we suspect may decrease the overall genotype fitness. In

short, while we did not discover temperature specific adaptation from our phenotype analysis,

we did identify that increased temperature (35°C) increases the overall rate of variant fixation

(Figure 2.8A & 2.8B) in mosquitoes, and may play an important role in novel virus genotypes

emerging in nature.

2.4d Intrahost genetic diversity is impacted by repeated bottlenecks and selective pressures

During systemic mosquito infection, ZIKV must infect and leave mosquito midgut and

salivary glands, with the efficiency in escaping these barriers characterizing vector competence

[189, 246]. These barriers impose genetic bottlenecks on arboviruses and impact genetic

population structure during infection. Our data point to the ZIKV population during repeated

bottlenecks and show that midgut infection allows for expansion of population diversity by

increasing richness, complexity and nucleotide diversity (Figure 2.10A-2.10C). However,

compared to previous data in WNV where continuous expansion of richness and diversity is

observed during dissemination out of the midgut [187], we observe a stochastic reduction in

variants along with a decrease in complexity and an increase in nucleotide diversity during

dissemination out of the midgut (Figure 2.10A-2.10C). Collectively, this suggests repeated

stochastic reduction of variants, lead to a loss if richness and complexity at any given locus,

while increasing the now dominant allele at said locus increases nucleotide diversity.

Additionally, SNV carry-through shows a trend where variants from the donor midgut

population are lost during escape into the mosquito hemolymph, suggesting a bottleneck is impacting SNV carry-through. This work is similar to that which has been previously described in WNV [187], however, we observe the largest genetic bottleneck in escaping the midgut, and not the salivary glands or in our case virus dissemination to saliva. Interestingly, we see that temperature impacts these bottlenecks severely in *Aedes albopictus,* where we see that 25°C, 35°C and diurnal EIT groups severely decrease richness and complexity after disseminating from the midgut. As a byproduct of this reduction of diversity, variants that successfully make it past this midgut bottleneck are able to establish themselves through founder's effects in the new environment suggesting that *Aedes albopictus* may be better able to drive low frequency variants to fixation.

2.4e Temperature impacts intrahost selection

Selection is a critical factor in RNA virus evolution [258]. Our data suggests that fluctuating diurnal temperatures increase the strength of purifying selection on both non-structural and structural ZIKV coding sequences after successful escape from the midgut (Figure 2.11E & 2.11J). This is extremely informative, as fluctuating temperatures are a better representation of what we observe in nature, and therefore the strong purifying selection is more likely to maintain the genotypes that are adapted to these temperature environments. In contrast, the constant temperatures which are more artificial, all exhibit neutral to positive selection, suggesting that the constant EIT groups are under positive selection, potentially adapting to temperature-dependent host environments. Additionally, we show that extrinsic incubation temperatures and species differences can provide sufficient selective pressures to force extinction or near fixation on individual variants (Figure 2.9). In the spectrum of variants assessed, we see

that there are not only variants or genotypes with temperature specific fitness advantages, but also species-specific advantages, indicating a complicated dynamic between temperature and host response during systemic infection. This highlights the importance of accounting for proper host and environmental factors in future genomic studies.

This study demonstrates that EIT is a driving factor in VC and RNA virus evolution for ZIKV. Temperature specific effects provide a fitness advantage, and that advantage varies between temperature and species. In addition to these effects, it is important to evaluate other factors that affect virus transmission and evolution. Chapter 3 evaluates changes in the ZIKV genetic population while varying the mosquito population within species and virus strains, both exposed to a wide range of EIT.

Chapter 3: Adaptive mutations arise during systemic *Aedes* infection under extreme extrinsic incubation temperatures

3.1 Introduction

Zika virus (ZIKV) was originally identified in Kampala, Uganda in 1947 [46], and became a global health concern in 2015-2016 during the ZIKV outbreak throughout the Americas [38, 60], resulting in an estimated > 700,000 suspected cases and over 800,000 cases worldwide to date [259]. Global travel, urbanization of the tropics, and climate change have allowed such arboviruses to explosively emerge in new geographical areas outside of previously endemic areas. ZIKV is rare in that it is an arthropod-borne virus (arbovirus) which must successfully infect and replicate in two very different hosts: vertebrates and invertebrates [188, 203]. While these hosts share some innate immune systems [260, 261], they do have distinct immune response mechanisms [262, 263], mechanical barriers of infections [189], and environmental temperatures [204, 264], all of which may cause very different selective environments. Previously, we identified that temperature plays a key role in shaping the selective environment in mosquitoes. We have seen that constant extreme temperatures increase positive selection, which may cause ZIKV to adapt to the new environment in *Aedes* mosquitoes. We also know that temperature plays a large role during mosquito transmission; the impact of environmental temperature has been modeled and described extensively. In recent years, researchers have addressed temperature impacts on vector competence in ZIKV with numerous different *Aedes* mosquitoes and ZIKV isolates [102, 103, 265, 266]. However, there is limited knowledge about how different mosquito species, strains and ZIKV isolates are impacted by the selective environment that is driven by a range of extrinsic incubation temperatures.

Differences between mosquito species [105, 216, 217, 267, 268], virus strains [105, 216, 269-271], virus preparation [216, 271], and virus passaging [272, 273] have been shown to impact arbovirus vector competence [105]. For ZIKV, it has been shown that *Aedes aegypti* is the more efficient vector when compared to *Aedes albopictus* [105]. We have observed that different African and Asian lineage ZIKV have significantly different vector competence in the same *Aedes aegypti* mosquitoes [105]. There are well documented reports of *Aedes aegypti* isolated from geographically distinct regions having different ZIKV vector competence to the same input virus [215]. When assessing difference in arbovirus population structure, we have seen that West Nile virus (WNV), another arbovirus, significantly alters the viral population during systemic infection when infected in different *Culex* mosquito species [187]. We have shown when ZIKV infects *Aedes aegypti* and *Aedes albopictus*, the latter species has strong bottlenecks in the midgut with increased divergence from the input population during systemic infection. However, we know little as to how differences within mosquito species and virus strains impact the genetic population during the selective environments induced under a range of extrinsic incubation temperatures.

To evaluate the impacts of variable species and strains, we expanded our previous studies in Chapter 2 to Tapachula, Chiapas, Mexico isolates of *Aedes aegypti* and ZIKV, and we expanded the extrinsic incubation temperature (EIT) groups to a broader temperature range. ZIKV exposed engorged female mosquitoes were sorted and housed at extrinsic incubation temperature of 16°C, 20°C, 24°C, 28°C, 32°C, 34°C, 36°C, and 38°C for 15 days at which time we collected mosquito bodies for downstream Next-generation library prep and sequencing (Figure 2.1). These samples allow us to understand how host species and virus strain leads to differences in genetic population dynamics and whether an extrinsic incubation temperature driven selective environment is conserved between strains as well as species. Our results, in comparison to our previous work,

suggest that temperature driven selective environments are consistent between mosquito and virus strains. Additionally, we identified that ZIKV input population diversity has a linear relationship with infected mosquito population diversity regardless of temperature, and last we identified an adaptive ZIKV genotype which facilitates fitness advantage in extreme temperature conditions within mosquitoes and adapting to novel cell culture environments.

3.2 Materials and Methods

3.2a Viruses

Virus was prepared and isolated as described by Tesla et al. (2018) [102]. Briefly, ZIKV MEX I-44 isolate (GenBank: KY648934) was passaged in Vero cells four times at UTMB and was passaged an additional six times in Vero cells at the University of Georgia.

3.2b Mosquitoes

Mosquitoes were collected and reared as described by Tesla et al. (2018) [102]. *Aedes aegypti* from Tapachula, Chiapas, Mexico were collected from Ovitraps in the Spring of 2016. Mosquito eggs were hatched in ddH2O under reduced pressure in a vacuum desiccator and L1 larvae dispersed into rearing trays. Larvae were fed fish food pellets (Hikari Cichlid Gold Large Pellets). Adult mosquitoes were maintained in rearing cages and provided with 10% sucrose *ad libitum*. Colonies were maintained on whole human blood (Interstate Blood Bank). Larvae and adults were maintained under standard conditions at 27°C +- 0.5°C, 80% +- 10% relative humidity, and 12:12 LD photoperiod.

3.2c Infection of Aedes mosquitoes and sample collection

Field derived *Aedes aegypti* mosquitoes were infected with ZIKV as described by Tesla et al. (2018) [102]. Females (3-5 day old, F_4 generation, *Aedes aegypti*, Tapachula, MEX) were exposed to an infectious blood meal (human blood washed in Roswell Park Memorial Institute medium and resuspended in 20% FBS, 1% sucrose, 5mmol/L ATP, DMEM) containing 1 million plaque forming units (PFU/mL) of MEX I-44. Mosquitoes were blood-fed through a water-jacketed glass membrane feeder for 30 minutes, engorged female mosquitoes were sorted and housed at each temperature treatment (Percival Scientific): 16°C , 20°C , 24°C , 28°C , 32°C , 34°C , 36°C , and 38°C , with ~80% relative humidity and a 12:12 LD photoperiod. Mosquitoes were maintained on a 10% sucrose solution for remainder of the study. 15 days post exposure, 20 ZIKV exposed mosquitoes per temperature group were cold anesthetized. Mosquito saliva, heads, legs, and bodies were collected into separate tubes (700 ul of DMEM with 1x antibiotic/antimycotic) and homogenized in a QUAGEN TissueLyzer at 30 cycles/s for 30 seconds.

3.2d Plaque assay

To identify ZIKV positive mosquito tissues, plaque assays were performed as previously described [102]. Each tissue homogenate was tested for presence or absence of ZIKV performing plaque assays on Vero cells were infected with mosquito homogenate in two full biological replicates and allowed to incubate for 1-2 hours. After incubation a 1.5% agarose DMEM (UltraPure LMP Agarose, Fisher Scientific) overlay was added and cells were incubated at 37°C with 5% CO2. Cells were formalin fixed (4% formalin) and stained with crystal violet.

3.2e Viral RNA isolation and Quantitative RT PCR

Viral RNA was extracted and qRT-PCR quantified from 50 µl of homogenized mosquito bodies, heads/legs, and saliva in biological triplicate as previously described [102]. RNA was extracted using the Mag-Bind® Viral DNA/RNA 96 kit (Omega Bio-Tek) on the KingFisher Flex Magnetic Particle processor (Thermo Fisher Scientific) and stored at -80°C. Viral RNA was quantified using qRT-PCR via the EXPRESS One-Step SuperScript qRT-PCR kit (Invitrogen) according to manufacture protocol using a forward primer (5'- CCGCTGCCCAACACAAG-3'), reverse primer (5'- CCACTAACGTTCTTTTGCAGACAT-3'), and probe (5'- AGCCTACCTTGACAAGCAGTCAGACACTCAA-3') sequences [233].

3.2f Library preparation for Next-generation sequencing

Water was used as a no template control (NTC), ZIKV PRVABC59 mixed with 10% and 1 % of ZIKV MEX I-44 genome equivalents (Supplemental Figure 3.2), and sample RNA (10ul) were prepared using the Trio RNA-Seq Library Preparation Kit (NUGEN) per manufacturer standard protocol. NTCs were used to identify contamination and sequencing bleed though and 10% and 1% spike ins were used to assess variant calling control characterization (Supplemental Table 3.2). Viral genome equivalents (GE/mL) were quantified following RNA extraction, and sequencing libraries were prepared as described above. Final libraries were pooled by tissue type (bodies, heads/legs, saliva, input virus, mixed virus, NTC) and analyzed for size distribution using the Agilent High Sensitivity D1000 Screen Tape on the Agilent Tapestation 2200, final quantification was performed using the NEBNext® Library Quant Kit for Illumina® (NEB) according to manufacturer's protocol. 150 nt pair-end reads were generated using the Illumina HiSeq4000 at Genewiz.

3.2g NGS processing and data analysis

Next-generation sequencing data were processed and analyzed using the RPG (RNA virus Population Genetics) workflow as described in Chapter 2 and Appendix A1. Workflow and specific code can be found at https://bitbucket.org/murrieta/snakemake/src. The RPG workflow was run on the ZIKV input blood meals samples using ZIKV-PRVABC59 as the reference sequence (GenBank: KU501215). Resulting consensus .fasta sequence generated from the input blood meals where then used as the reference sequence for reprocessing and final analysis. Of the mosquito body samples, only single nucleotide variants of 1% or greater in the coding sequence (nt position 91-10363), with >100x coverage were used for analysis to account for low coverage (reads per genome position, Supplemental Figure 3.1) in the 3' and 5' untranslated regions and degraded RNA/low input GE of heads/legs, and saliva samples.

It is important to point out that the NTC had 1,184 reads map to ZIKV (Supplemental table 3.2), the NTC consensus was 100% identical to the MEX I-44 input, and of the 4 low frequency variants identified only 1 was found in the input population (C-3878-T). However, the average ZIKV coverage of the NTC was equal to 15 (Supplemental table 3.2), and therefore would have been filtered out of analysis (Coverage less than 100), suggesting this contamination would have little impact on sample analysis. The ZIKV PRVABC59 and MEX I-44 consensus sequence has 45 unique differences in the ZIKV coding sequence (CDS). By spiking in ~10% and ~1% MEX I-44 into PRVABC59 (by approximate genome equivalents) and sequencing we can identify the consistency of variant calling (Supplemental Figure 3.2). ZIKV PRVABC59 and MEX I-44 have 45 different nucleotide sites, we these sites to identify variant frequency calling, the 10% spike had an average spike in of 22.04% MEX I-44. The 1% spike had an average of 5.02% spike-in of MEX I-44 (Supplemental Figure 3.2). We believe the 20% spike in was due to

dilution miscalculation and the 5% spike in was due to pipet error and differences fixation of these 45 nucleotide sites. Therefore, we selected at 1% variant frequency threshold for analysis based on inherent error associated with illumina sequencing.

Multiple alignment of MEX I-44-Extreme (MEX I-44 with NS2B-S45T and E-T470M consensus changes) and MEX I-44-Moderate (MEX I-44 with NS3-K117R, NS2A-A117V, NS1-R103T, C-G73R, and E-L491S consensus changes) consensus sequences were performed (Supplemental Table 3.1) using Geneious 10.1.3 geneious alignment function by aligning to 283 ZIKV genome sequences (https://github.com/andersen-lab/paper_2018_cuba-travel-zika/blob/master/phylogenetics/2018.10.09_alignment.fa) that have been previously described [67].

3.2h Purification of biological clones

Plaque purification were performed similar to that previously performed by our group. The MEX I-44-Extreme and MEX I-44-Moderate biological clones were purified by inoculating Vero cell monolayers in 6-well plates at 70-90% confluence in triplicate with10-fold serial dilutions (-1,-2,-3,-4) of the mosquito body homogenate (50ul). Mosquito homogenates were selected for by assessing the frequency of the extreme and moderate genotype bioinformatically and inoculating Vero cells with the homogenates that had the highest frequency of each genotype. After 1 hour of incubation at 37ºC with 5% CO_2, rocking every 10 minutes, the inoculum was removed and a 0.6% agar in EMEM overlay was added. Three days post infection a secondary agar overlay was added with 0.015% Neutral Red solution (Sigma-Aldrich). Twenty-four hours after the secondary agar overlay was added, individual plaques were picked by pipet from wells containing 1-20 plaques. Plaque plugs were pipetted into 7ml of 10% FBS

DMEM media and added to 70-90% confluent monolayer of Vero cells in a T-25 flask and incubated at 37C with 5% CO2 until ~ 50% CPE was observed and supernatant collected, and frozen at -80⁰C. RNA was extracted from frozen stocks using the Omega Mag-Bind Viral DNA/RNA kit as described above. cDNA was generated using the QIAGEN OneStep RT-PCR kit according to manufacturer's protocol. All stocks were screened by Quantitative Sanger sequencing using amplicons for the MEX I-44-Extreme and MEX I-44-Moderate mutations. Amplicons were generated using primers previously described elsewhere [274].

3.2i Cells

Vero cells (African Green Monkey kidney; ATCC CCL-81) were maintained at 37⁰C and 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Pen-Strep). Aag2 cells (*Aedes aegypti*) were maintained at 28⁰C and 5% CO_2 in Schneider's Insect Medium (Sigma-Aldrich supplemented with 8% FBS and 1% Pen-Strep in a sealed cap flask.

3.2j Competition study

Competitive fitness was determined largely as described previously [201]. Competitions were conducted in Vero and Aag2 cells. The reference vs. competitors and WT vs. competitors where mixed at a 1:1 ratio and cells where inoculated at a MOI of 0.01. Supernatants were collected at 0, 48, and 144 hours post infection. RNA was extracted as above, and cDNA was generated using EXPRESS One-Step SYBR GreenER Kit (Invitrogen) according to manufacture protocol. For all reference vs. competitors the forward LNA primer 5'-A+CTTGGGTTGTGTACGG-3' and reverse primer 5'- GTTCCAAGACAACATCAACCCA-3'

were used to generate amplicons for Quantitative Sanger sequencing. Genotype fitness was analyzed using polySNP [275] to analyze the area under the curve of the .ab1 output files.

3.2k Genetic diversity

Markers of genetic diversity including richness, nucleotide diversity, Shannon's entropy, F_{ST}, and Selection (d_N/d_S) were calculated as previously described [239]. DnaSP6 [276] was used to determine MEX I-44 synonymous (2268.33) and non-synonymous (8003.67) sites which is used to calculate d_N/d_S.

3.2l Statistical analysis

All statistical analyses were performed using GraphPad Prism (version 8.42). All tests were performed as describe in results.

3.3 Results

3.3a Input population standing variation seeds found populations

To assess how extrinsic incubation temperatures of 16°C, 20°C, 24°C, 28°C, 32°C, 34°C, 36°C, and 38°C impact our ZIKV population structure within the *Aedes aegypti* vector, we first characterized the population diversity of the input virus stock used for mosquito infections by performing Next-generation sequencing (NGS) of 3 biological replicate infectious blood meals used to orally expose *Aedes aegypti* to ZIKV. We were able to identify that the input ZIKV population had standing variation which is comprised of an average of 32 single nucleotide variants (SNVs) ranging in frequency from greater than 1%to less than 30% (Figure 3.1A). The

input population had low frequency variants (less than 0.5) variants in each gene coding region of ZIKV except for the non-structural (NS) proteins coding regions for NS4A and NS4B, (Figure 3.1A). The *Aedes aegypti* founder population was characterized for *Aedes aegypti* bodies that had a minimum of 100 times coverage across the ZIKV CDS, which included 20°C, 24°C, 28°C, 32°C, 34°C, 36°C, and 38°C extrinsic incubation groups (Figure 3.1B). 16°C *Aedes aegypti* bodies had less than 100 times coverage across the CDS and were excluded from analysis (Figure 3.1B). The heads/legs and saliva samples were excluded from analysis as there were tissues that had low input genome equivalents (16°C heads/legs & saliva, 20°C, 24°C, 36°C saliva samples) and either libraries could not be generated or sequencing resulted in coverage less than 100 times across the CDS (20°C and 24°C heads/legs, 28°C saliva) (Supplemental Figure 3.1).

Next, we characterized how SNVs are distributed upon successful mosquito infection. We identified 74 total SNV sites across the ZIKV CDS, which is comprised of 50 novel SNV sites, and 24 SNV sites found in the input population (Figure 3.1C, Table 3.1). Of the total SNVs observed we saw that the majority of SNVs accumulated in the envelope (E, 27%), followed by NS1 (16%), and then NS2B (15.5%) gene regions (Figure 3.1C). When assessing total variants by temperature across the CDS, we observed that 28°C (23%) followed by 32°C and 34°C (20% each), make up the majority of the SNV accumulation (Figure 3.1C). Unique novel SNV sites accumulated most efficiently at higher temperatures of 34°C, followed by 28°C, and 32°C (Table 3.1). We observed the fewest unique novel SNVs of 2 and 7 at our two lowest temperatures 20°C and 24°C respectively (Table 3.1). Looking at input SNV sites that successfully make it past the initial barrier of infection and seed our founder population, we observed that 75% (24/32) of the input ZIKV variant population are found in the infected *Aedes aegypti* bodies. The moderate temperature of 32°C accumulates the most input SNV sites through initial infection of *Aedes*

aegypti bodies, accumulating 67% of the input virus seeded population variants (Table 3.1). Surprisingly 24°C, 28°C, and 34°C all have the same percent of input SNV site carry-through (the proportion of SNVs from the input population that are passed to the next tissue) of 50% (Table 3.1), and 20°C and 36°C have the lowest input SNV site carry-through of 33% and 29%, respectively (Table 3.1). The 28°C EIT group has the most input SNVs per replicate (mean of 9.7) and decreasing as EIT decreases and increases from 28°C (Table 3.1)
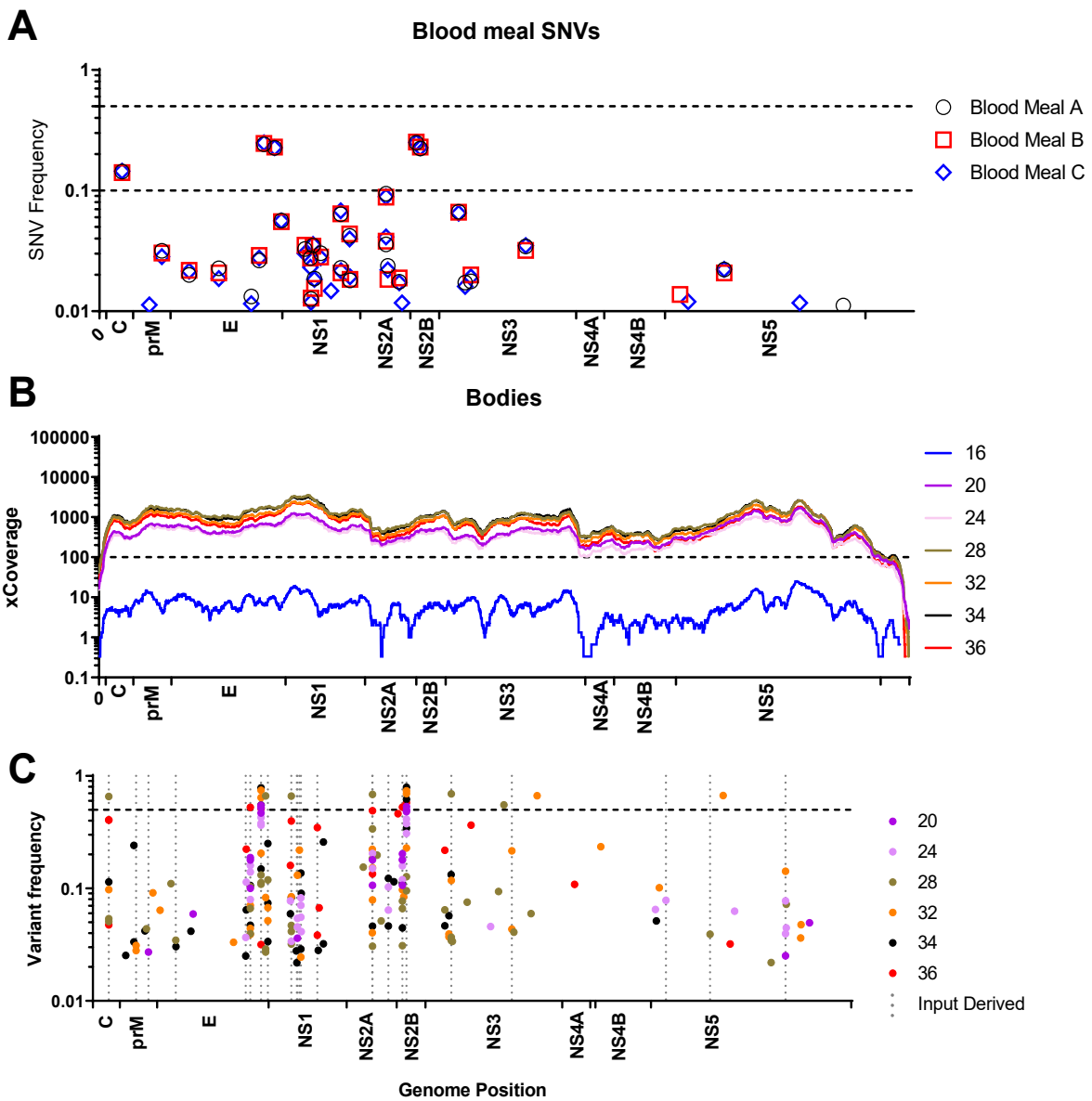
Figure 3.1 Next-generation sequencing characterization of ZIKV. Characterization of ZIKV input population (A), sequencing coverage of ZIKV positive bodies across the ZIKV genome (B), and the distribution of SNVs across the ZIKV CDS for each extrinsic incubation temperature group 20°C, 24°C, 28°C, 32°C, 34°C, & 36°C (C). Those SNVs found in the input population are indicated by a dashed line (C).

Table 3.1 1 Input virus and founder population SNV characterization. Characterization of single nucleotide variants found in 20°C, 24°C, 28°C, 32°C, 34°C, & 36°C extrinsic incubation temperature groups.

	Temperature					
	20	24	28	32	34	36
Total Novel SNV sites %	4% (2/50)	14% (7/50)	30% (15/50)	24% (12/50)	32% (16/50)	16% (8/50)
Total Input SNV sites %	33% (8/24)	50% (12/24)	50% (12/24)	67% (16/24)	50% (12/24)	29% (7/24)
Total SNV sites %	14% (10/74)	26% (19/74)	36% (27/74)	38% (28/74)	38% (28/74)	20% (15/74)
Mean Novel SNVs	0.7	2.3	5.7	4.3	6.0	2.7
Mean Input SNVs	5.7	8.7	9.7	9.0	7.3	4.3
Mean SNVs	6.3	11.0	15.3	13.3	13.3	7.0

3.3b Extreme cold and hot extrinsic incubation temperature select for adaptive mutations

To identify the role that temperature has on ZIKV variant frequency within the mosquito host, we identified 7 non-synonymous variants that were present in the input population and tracked their variant frequency during mosquito infection under our 6 extrinsic incubation temperature groups. From this, we identified 2 SNVs (NS2B-S45T and E-T470M) that have increased variant frequencies when at extreme temperature conditions of 20°C, 24°C, 32°C, and 34°C (Figure 3.2A). These two SNVs were slightly deleterious in that their variant frequency was lower than that of it input frequency at 28°C. Conversely, we identified 5 SNVs (NS3-K117R, NS2A-A117V, NS1-R103T, C-G73R, and E-L491S) that when at a moderate temperature of 28°C have a fitness increase and increase in variant frequency (Figure 3.2A). And when these 5 SNVs are at extreme temperatures of 20°C, 24°C, 32°C, and 34°C they appear to have no fitness effect or are deleterious. In both the extreme and moderate temperature adapted group, we observe the inverse effect on variant frequency when at 36°C.

74
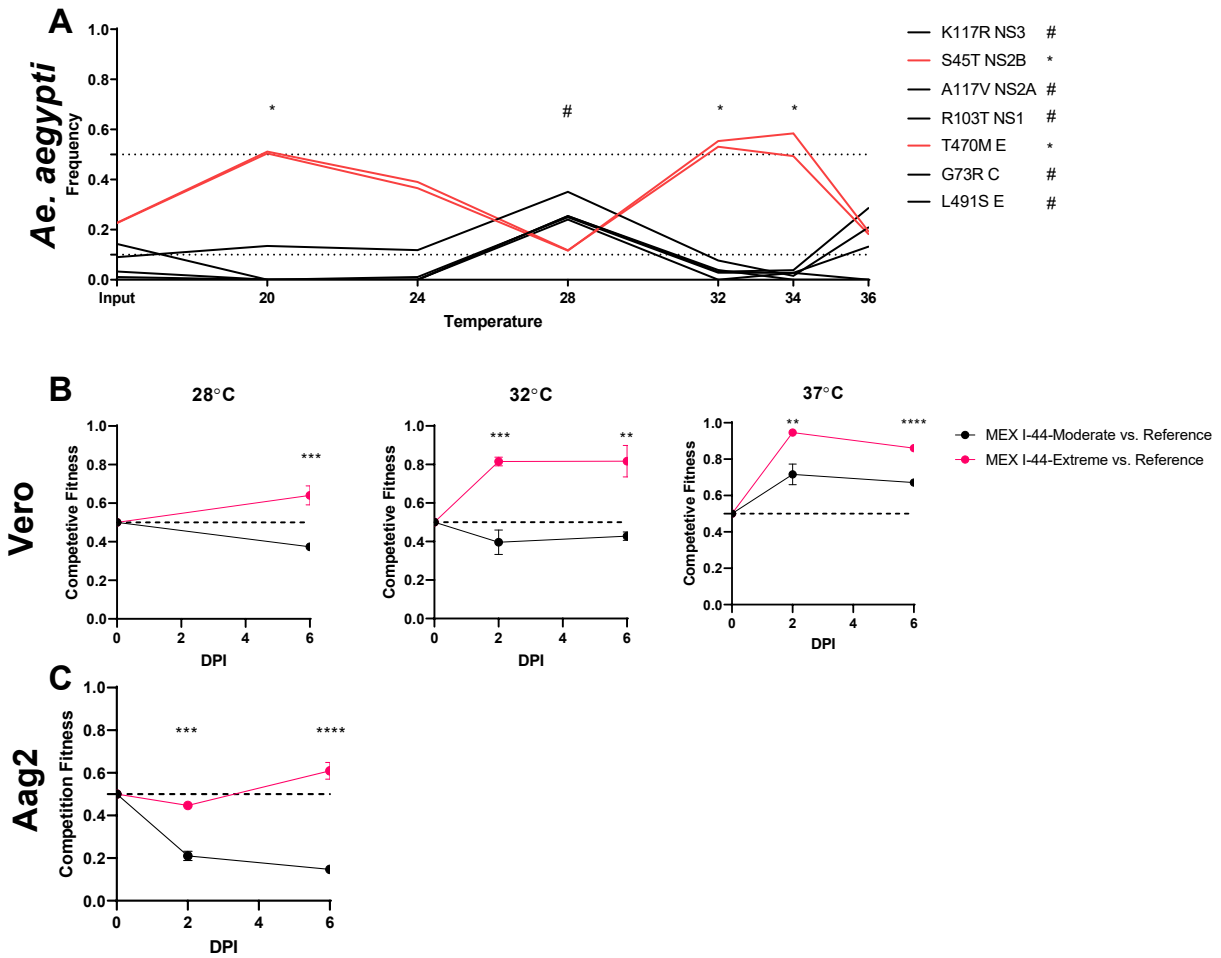
Figure 3.2 Extreme extrinsic incubation temperatures lead to the adaptive mutations arising. Assessment of extrinsic incubation temperature on variant frequency for 7 non-synonymous variants found in the input population (A). Competitive fitness of the MEX I-44-Extreme and MEX I-44-Moderate genotype in Vero cells at 28°C, 32°C, and 37°C (B) and Aag2 cells at 28°C(C). Unpaired two-tailed, T-test, p-value **** <0.0001, ***0.0001, **0.001.

To determine if these variants have arisen previously, we performed a multiple-alignment against 283 naturally occurring ZIKV isolates (Supplemental Table 3.1). We found that only 2 of the 7 variants have been identified before. Both variants had a fitness advantage at moderate temperatures. E-L491S had a percent similarity (percent of identical sequences in the 283 ZIKV sequences used for alignment) of 0.35% and was previously characterized in a ZIKV isolate from human plasma in Brazil (GenBank: KY785429). NS2A-A117V had a percent similarity of 1.06% and was found in the original MEX I-44 isolate, a second Chiapas, Mexico isolate from

mosquitoes (GenBank: KX446951), and from human blood from Brazil (GenBank: KX520666, Supplemental Table 3.1).

To determine the phenotypic effect that the extreme (NS2B-S45T and E-T470M) and moderate (NS3-K117R, NS2A-A117V, NS1-R103T, C-G73R, and E-L491S) genotypes have, we purified biological clones of each genotype which we term MEX I-44-Extreme and MEX I-44-Moderate, in which these are purified viral stocks that have the respective mutations present. We then used these biological clones in a competition assay in which we determined the fitness of each clone under different temperature conditions (28°C, 32°C, and 37°C) in a vertebrate cell line (Figure 3.2B) and at 28°C in an invertebrate cell line (Figure 3.2C) . We identified that MEX I-44-Extreme consistently had significantly (p-value < 0.05, unpaired two-tailed T test) higher fitness when compared to MEX I-44-Moderate clones, regardless of temperature and cell type (Figure 3.2B & 3.2C). Surprisingly, MEX I-44-Moderate had no fitness advantage when compared to the reference ZIKV strain or it was slightly deleterious at the moderate temperature of 28°C and 32°C in vertebrate cells (Figure 3.2B) and at 28°C in invertebrate cells (Figure 3.2C). In fact, only at a temperature of 37°C do we see any fitness increase in MEX I-44-Moderate when compared to the reference ZIKV clone.

3.3c Extrinsic incubation temperature impacts within host population diversity in a unimodal manner

Population diversity within *Aedes aegypti* bodies was characterized to determine how varying temperatures impact richness (the number of variants present per sample), complexity (the uncertainty with sampling a certain allele), nucleotide diversity (sum of variant frequencies across a specific locus), divergence (fixation index, F_{ST} compared to input population), and

selection (d_N/d_S) across the ZIKV CDS, structural protein coding regions, and non-structural coding sequence regions. In general, we see that population diversity is impacted by constant temperatures in a unimodal distribution, in which moderate temperatures of 28°C and 32°C have the highest richness, complexity, nucleotide diversity, and divergence across the ZIKV CDS (Figure 3.3A-3.3D). Selection is inversely impacted and has the lowest d_N/d_S at 28°C and 32°C (Figure 3.3E).

When assessing richness in *Aedes aegypti* bodies we see that there is a 4.8-fold (20°C) to 2.1-fold (28°C) reduction in richness across the ZIKV CDS. Extreme temperatures of 20°C and 36°C have the highest fold change reduction from the input virus richness across the CDS of 4.8-fold and 4.5-fold respectively. On the other hand, we see the smallest fold change across the CDS (2.1-fold reduction) occurring at 28°C. Similar observations hold true for the structural and non-structural gene regions, and we observe that generally there are 1.4-3.1-fold more variants in the non-structural regions compared to the structural counterpart group (Figure 3.3A). Complexity decreases from the input population across all temperature groups, whereas nucleotide diversity is increased from the input population at 28°C and 32°C across the CDS (Figure 3.3B & 3.3C). Divergence is inversely related to selection, and we see that divergence from the input population in highest at 28°C and 32°C across the CDS while selection is lowest at 28°C and 32°C across the CDS (Figure 3.3D & 3.3E). We see that constant temperatures excluding 28°C are under constant positive selection, whereas 28ºC groups are under neutral evolution or weak purifying selection (Figure 3.3E).
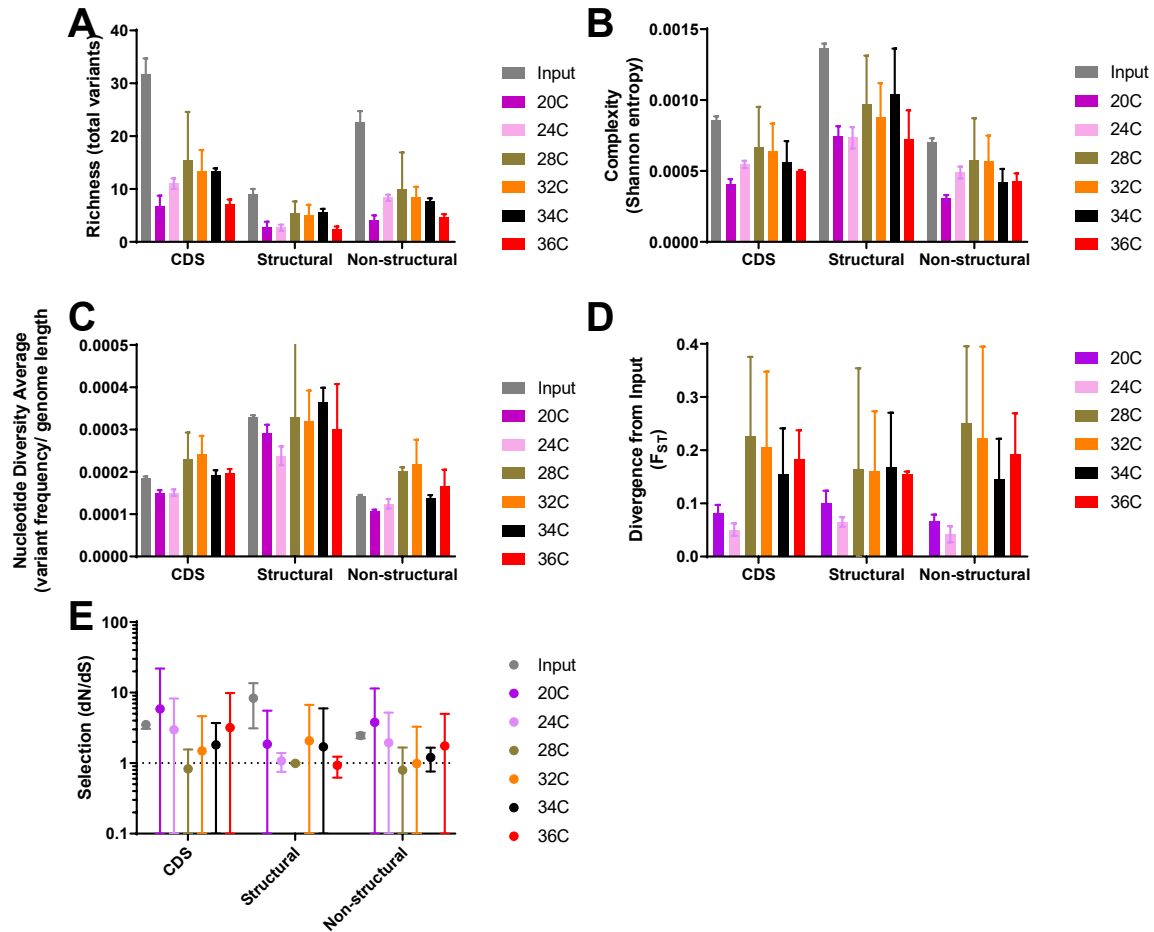
Figure 3.3 Extrinsic incubation temperature impacts on ZIKV genetic diversification in the CDS, structural, and non-structural regions. Genetic diversity was measured by assessing richness (A), complexity (B), nucleotide diversity (C), divergence (D), and selection (E) at each extrinsic incubation temperature group. Mean and 95% CI graphed.

3.3d Intrahost viral gene region diversity is driven by viral input population diversity

Last we assessed how specific ZIKV protein coding region (capsid, pre-membrane, envelope, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) population diversity is impacted during mosquito infection under varying temperature ranges (Figure 3.4A & 3.4B). We used nucleotide diversity and complexity as markers of diversity and characterized varying constant temperature impacts across the gene regions (Figure 3.4A & 3.4B). We observed inherently high

levels of nucleotide diversity and complexity in the NS2B gene region for all temperatures with

28°C having the lowest diversity (Figure 3.4A & 3.4B). The input population diversity and

complexity were highest in NS2B. Generally, we observed that temperatures of 28°C, 32°C, and

36°C had random increases in nucleotide diversity and complexity across the gene specific

regions, but temperature specific impacts do not appear to be present. Therefore, we sought to

identify the relationship of input level nucleotide diversity and complexity to that observed

across treatment groups (Figure 3.4C & 3.4D). When graphing the mean input diversity in

relationship to the mean diversity of all treatment groups pooled by gene region, we observed a

clear linear relationship in nucleotide diversity (Equation Y=1.058*X-1.002e-005, R^2=0.9715)

and complexity (Equation Y=0.6960*X-5.006e-005, R^2=0.9205). This indicated that as input

population diversity was increased in a coding region-specific manner, so too did the coding

region-specific diversity in infected *Aedes aegypti* bodies.
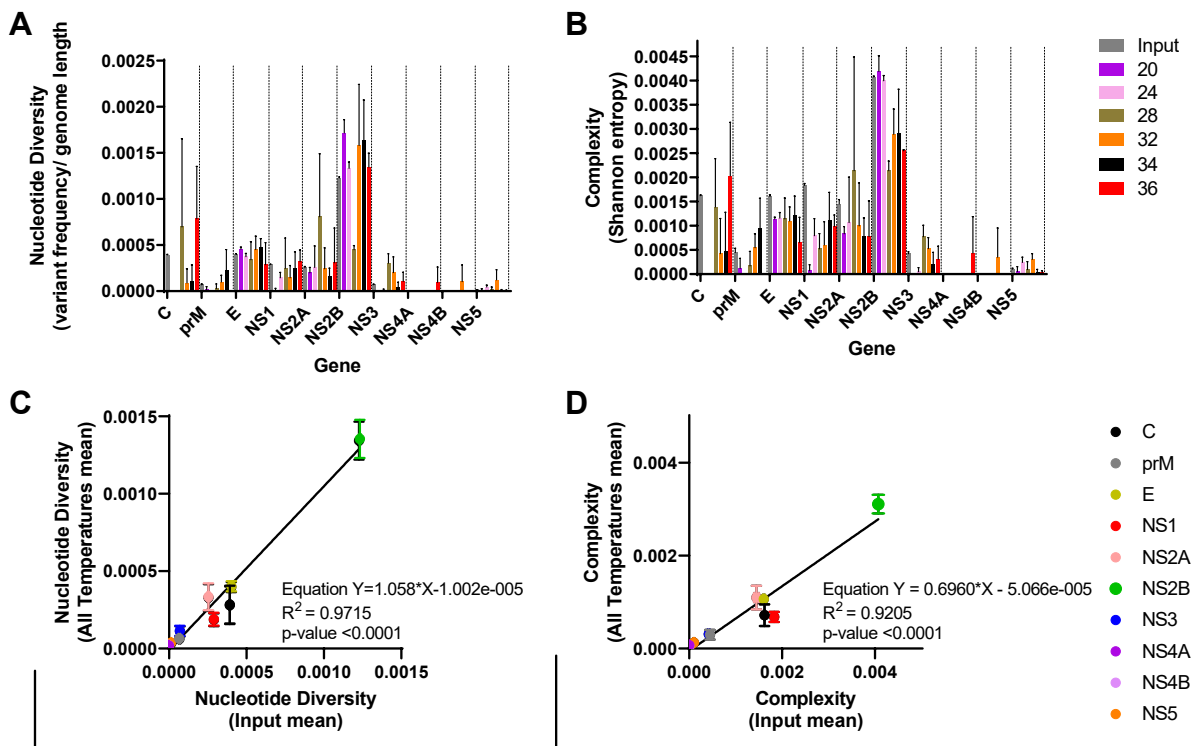
Figure 3.4 Impacts of extrinsic incubation temperature on ZIKV gene regions. Nucleotide diversity (A), and complexity (B) were characterized for each EIT at each protein coding region; structural (C, prM, E) and non-structural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). Linear regression of input (mean) against all temperature groups combined (mean) for each gene region was assessed for nucleotide diversity (C) and complexity(D). Mean and 95% CI graphed.

3.4 Discussion

3.4a Input population structure as a predictor of founder populations and gene region diversity

Through NGS analysis, we characterized the input blood meal population and founder virus population when infecting *Aedes aegypti* (Tapachula) with ZIKV (MEX I-44). In this study, we were unable to generate quality sequencing data from heads/legs and saliva samples. Therefore, we were unable to analyze within host population diversity. However, as we were able to successfully sequence the mosquito bodies, this provided a picture of the viral population structure as a whole. During infection we identified that 75% of the input virus population SNV sites make up 32% of the total (all observed SNV sites) founder population SNV sites during successful mosquito infection (Figure 3.1A & 3.1C). Moderate temperatures of 28°C and 32°C had the lowest impact of perceived infection bottlenecks by accounting for the highest input SNV site accumulation (28°C =12 & 32°C =16 sites), and mean input SNV sites per replicate (28°C = 10 & 32°C =9) (Table 3.1). When temperatures were extremes (low and high), we observed that bottleneck strength is higher (larger reduction of viral population) which led to the lowest number of input SNV sites successfully carried-through infection at 20°C (8) and 36°C (7), with mean input SNVs of 6 (20°C and 4 (36°C) per replicate (Table 3.1). This suggested that constant temperature impacts founder population richness in a unimodal manner with moderate temperature of 28°C and 32°C having the highest richness.

The input ZIKV population structure had the greatest impact on coding region-specific diversity. We observed that NS2B had high levels of nucleotide diversity and complexity in the input population; upon successful infection in *Aedes aegypti* the observed NS2B gene region nucleotide diversity and complexity was high among all temperature groups (Figure 3.4A & 3.4B). During successful infection, we saw stochastic reduction and promotion of diversity across the gene coding regions, with a general trend of moderate to high temperatures increasing in diversity (Figure 3.4A & 3.4B). However, because there was no clear relationship to extrinsic incubation temperature and gene region diversity, we combined all temperature groups to assess the role that input population region-specific diversity has on founder population region-specific diversity (Figure 3.4C & 3.4D). From this, we found protein coding region nucleotide diversity and complexity had an ~1:1 relationship with input population diversity, regardless of extrinsic incubation temperature groups (Figure 3.4C & 3.4D). This is an interesting observation as we have previously shown that in the Poza Rica, Mexico *Aedes aegypti* line infected with a Puerto Rican isolated ZIKV had inherently high gene region diversity in NS1 in the input population, we also saw high diversity in that region during systemic mosquito infection. This highlights that gene region diversity, which we expected to be greatly impacted by temperature, is highly susceptible to the input virus population region-specific diversity, and differences in input population structure are a direct reflection of gene region population diversity. These results elucidate the role that input ZIKV diversity and extrinsic incubation temperature have on founder population structure. This allows us to better predict how environmental temperature impacts founder populations across the CDS by imposing stronger bottlenecks at extreme temperatures (Figure 3.3A-3.3D), whereas at the region-specific level, input population region-specific diversity is the driving factor for founder population region-specific diversity (Figure 3.4).

3.4b Extreme extrinsic incubation temperatures lead to the accumulation of adaptive mutations

When characterizing the role that constant extrinsic incubation temperature had on ZIKV (MEX I-44) population structure during infection of *Aedes aegypti*, we identified 7 non-synonymous variants that appeared to be on one of two haplotypes or at least had a similar response to our temperature treatments. Of these 7 variants, we broke them into two haplotypes. MEX I-44-Moderate was comprised of 5 variants which have a fitness advantage (based on variant frequency increase) at moderate temperature of 28°C and to a lesser extent 36°C, but deleterious (based on variant frequency decrease) at other temperatures (Figure 3.2A). The second haplotype MEX I-44-Extreme was comprised of 2 variants which had fitness advantage at all temperatures except 28°C and 36°C in which those temperatures were deleterious (Figure 3.2A). Of the variants found in these two haplotypes, only two were found occurring in nature (E-L491S and NS2A-A117V) and at very low frequencies (Supplemental Table 3.1). Considering that these two variants were isolated in two very distinct geographic regions, this suggested that they may be undergoing convergent evolution. Of the remaining 5 variants, none have been found in nature, which suggested that each were undergoing independent evolution. With respect to the variants found on MEX I-44-Extreme haplotype, these variants appeared to be selected for by the hosts response caused by extreme temperatures.

Contrary to our original assumption that the two haplotypes MEX I-44-Extreme and MEX I-44-Moderate would have fitness advantages based solely on temperate (moderate vs extreme), we identified that MEX I-44-Extreme was a generally adaptive haplotype consistently outcompeting our reference clone and having significantly higher fitness that the MEX I-44-Moderate haplotype in each cell type and temperature that it was exposed (Figure 3.2B & 3.2C). This suggested that the two "extreme" non-synonymous variants (NS2B-S45T and E-T470M)

incorporated into the MEX I-44 genome infer a level of adaptive plasticity that may be essential for adapting to host response, host switching, or some combination of environmental stressors. The envelope substitution of T470M is an amino acid with a polar hydrophilic side chain (threonine) which changes to an amino acid with a non-polar, hydrophobic side chain (methionine) found as a surface exposed residue in the C-terminal transmembrane anchor [277]. The NS2B substitution of S45T is a polar hydrophilic amino acid (serine) to a similar polar hydrophilic amino acid (threonine) just four residues before the n-terminal residues of the NS2B-NS3 protease [278]. There are numerous examples of seemingly trivial amino acid changes occurring in arboviruses in nature, which lead to significant fitness impacts [279, 280]. Whether the MEX I-44-Extreme haplotype has the potential to cause similar phenotypic effects is unclear, but further investigation is warranted to identify the mechanism and limit that this haplotype can infer adaptive plasticity on ZIKV infection. While we predict that the E-T470M substitution may be a driving factor for adaptation, as the amino acid dissimilarities are greatest and the envelope protein is essential for successful infection [281, 282] , substitutions in this protein may lead to altered host affinity [280]. We cannot rule out the potential synergistic effects of NS2B and S45T, and therefore further studies should be concluded to elucidate the roles of these two amino acid substitutions for ZIKV infection.

3.4c Extrinsic incubation temperatures impact CDS population diversity in a predictable manner regardless of mosquito species and virus strain

Through NGS analysis we used genetic markers to characterize genetic diversity (Figure 3.3). Richness was used to assess mutational expansion and reduction upon successful mosquito infection. Complexity was used to determine the likelihood of sampling a specific allele at a

particular locus. Nucleotide diversity was used to determine how extrinsic incubation

temperature impacts variant frequencies across the ZIKV CDS. The fixation index (F_{ST}) was

used to determine divergence from the input population. In general, when assessing richness and

complexity across the ZIKV CDS, we saw a decrease in both genetic markers compared to the

input population (Figure 3.3A & 3.3B). Additionally, we saw a unimodal distribution of both

markers (richness and complexity) with 28°C and 32°C having the highest levels of diversity,

and the extreme temperatures of 20°C and 36°C having the lowest levels (Figure 3.3A & 3.3B).

When we compared these finding to previous work in a different *Aedes aegypti* and *Aedes*

albopictus ZIKV infection experiment (Chapter 2), we saw numerous similarities in the

distribution of diversity in our current data and previous work. Previously, the midgut was used

to represent initial infection and characterize CDS level diversity. In midgut tissues we showed

that mean richness is highest at 32°C and decreased unimodally with 25°C having the lowest

richness in *Aedes aegypti* and 35°C the lowest in *Aedes albopictus*. Here, we show that there is a

clear unimodal distribution of richness with a peak of 28°C and a minimum at 20°C followed by

36°C (Figure 3.3A). Our previous findings also show a unimodal distribution with peak

complexity of 32°C in both *Aedes aegypti* and *Aedes albopictus*, but the 25°C and 28°C groups

are much closer in complexity in the current study (Figure 3.3B). Last, we saw that 32°C and

28°C had the highest levels of nucleotide diversity with the unimodal distribution with low

temperatures of 20°C and 24°C being impacted the most. Previously we showed that in *Aedes*

aegypti nucleotide diversity is highest at 28°C; however, the 32°C and 35°C groups do not

follow unimodal distribution as seen in the previous two genetic markers. These differences may

be explained by the narrower range of extrinsic incubation temperatures of 25°C -35°C

performed previously (Chapter 2), compared to the current studies 20°C -36°C. The combination

of our current findings and those previously performed indicate that extrinsic incubation temperature impacts ZIKV genetic richness, complexity, and nucleotide diversity in a constant and predictable manner regardless of strain or species.

Divergence between the input population and ZIKV positive body population had a unimodal distribution across the CDS, but only for 24°C - 34°C temperature range. The 28°C group had peak divergence from the input population with 24°C on the low end and 34°C on the high end both having the lowest divergence from the input population (Figure 3.3D). In our case, low divergence can most likely be explained in two ways. For the 24°C group, decreased temperature reduces replicative capacity of the virus [208] and novel variants are less likely to arise, maintaining a close similarity to the founder population. At a high temperature of 34°C, increased temperature increases replicative capacity [208], which may increase the proportion of founder variants compared to novel variants and ultimately reduce the perceived bottleneck influence. For example, as founder variants now make up a high proportion of the population, they also have high likelihood of infection and maintenance. Surprisingly, 20°C and 36°C had higher divergence than their next closest temperature group (24°C and 34°C respectively) which indicates there may be some biological response induced by extreme temperatures that increases divergence. Again, our current results are in agreement with previous findings, in which we show that divergence across the CDS is unimodal in *Aedes aegypti* when all tissues are combined. However, we previously showed that *Aedes albopictus* increases in divergence up to 35°C which suggest that divergence is impacted by temperature but also by host species and host temperature response.

When we looked at d_N/d_S as an indicator of selection during ZIKV infection, the input virus was under positive selection with a d_N/d_S greater than 2 (Figure 3.3E). Selection based on

EIT groups showed that there was a unimodal distribution of selection across the CDS (Figure 3.3E). The 28°C temperature groups suggested neutral selection, whereas when temperature was raised further to the extremes (20°C and 36°C), d_N/d_S increased in a unimodal manner. This indicated that extreme environments generated by temperature led to an environment where positive selection was the driving force for population structure. Which is in agreement with our previous finding when we assessed selection in the midgut tissues, we found that 28°C and 32°C extrinsic incubation temperature groups were near a d_N/d_S of 1, but as temperature diverged to the extremes, the selective environment changed to positive selection.

The combined results of the studies completed in Chapters 2 and 3 highlight the importance of considering multiple factors when evaluating vector-borne disease transmission. Not only does temperature impact the virus at a genetic level, but its interactions vary by virus strains, the host, and with other environmental factors. Chapter 4 summarizes the impacts that extrinsic incubation temperature has on ZIKA virus evolution and vector competence and provides future considerations for research.

Chapter 4: Summary and Future Considerations

4.1 Summary

Vector-borne disease persistence and transmission is dependent on the dynamic interactions between the pathogen, its hosts, and the environment [283, 284]. We characterized the impact of extrinsic incubation temperature (EIT) on *Aedes* vector competence (VC) and Zika virus (ZIKV) evolution during systemic mosquito infection. Overall, we found that EIT impacts VC for constant temperatures in a unimodal manner and its effect increases over time. These findings are consistent with previously described work on ZIKV in *Aedes aegypti* [102, 103]. In addition to our examination of the impact of constant temperatures, we assessed fluctuating (diurnal) temperatures. Our diurnal temperatures were designed to have mean daily temperature of 30°C, which is close to the optimal VC temperature range (between 28°C and 32°C, Figure 2.1). However, mosquitoes that experienced diurnal temperatures were less efficient ZIKV vectors than constant EIT groups with similar mean temperatures. While assessing EIT impacts on interhost evolution of the virus, we see that at the coding sequence (CDS) level there is little impact on viral genetic variation, but when assessing gene region diversity in genetic selection, we see that envelope (E) and NS1 coding regions appear to be under positive selection with species and temperature specific differences. Conversely, NS5 is under purifying selection for all EIT groups, suggesting NS5 conservation is essential for thermal stability and replication. Additionally, we identified and characterized 8 consensus change mutations that allow varying fitness advantages during systemic ZIKV infection.

Through competitive fitness experiments, we found that increased extrinsic incubation temperatures (35°C) significantly increase the rate of variant fixation. Additionally, we were

able to characterize interhost evolution and found that moderate temperatures have slightly more diversity per CDS, but when assessing selection during systemic infection, we saw that 28°C and 35°C increase selective pressure for transmitted viruses, and diurnal EIT groups exhibit strong purifying selection on transmitted viruses which may better represent natural environments.

Following the above finding, we sought to unravel the interactions of EIT groups on distinct hosts and ZIKV isolates to determine if our previous findings were consistent regardless of host species isolation and ZIKV strain. Thus, we characterized the population structure of a Mexican ZIKV strain exposed to six different constant EITs during systemic infection in a Mexico *Aedes aegypti* line. As expected, we found that the input virus population variants directly seed the founder population diversity with 75% of the input variants sites making up 32% of the infected *Aedes aegypti* mosquito variant sites. We saw that EIT has a direct impact on the founder population. When temperatures were near moderate levels of 28°C, we saw the highest accumulation of variants, conversely, when ZIKV was exposed to extreme temperatures of 20°C and 36°C, we saw the least number of variants. This suggests extreme temperatures greatly impact bottlenecks of infection with extreme temperature increasing the strength of bottlenecks. Of the variants that were characterized in our study, we identified two variants of interest which appeared to have fitness advantages at extreme temperatures, or those temperatures outside of 28°C. The first variant was found in the envelope protein and the other in the NS2B protein (Supplemental Table 3.1). By purifying biological clones, we were able to successfully create a biological clone containing both variants. In characterizing phenotypic effects of this clone, we found that they were not strictly driven to adapt to temperature variation, but to different environments, such as cell type and temperature, suggesting that these mutations aid in the adaption to novel environments.

We next characterized the population structure of ZIKV during these EIT groups. As we have previously shown, there was a unimodal distribution of genetic markers, but it is interesting that we saw the same unimodal distribution of diversity with moderate temperatures having peak richness, complexity, and nucleotide diversity but extreme temperatures negatively impacting population diversity (Figure 3.3) when using a different virus isolate and mosquito strain. Last, we identified the temperature that has minimal impact on gene region nucleotide diversity and complexity (Figure 3.4A & 3.4B). However, we saw that the input population diversity had a near one to one relationship with observed population diversity in infected mosquitoes regardless of temperature.

Our data combined help to elucidate the dynamic interplay between host, pathogen, and environment including the finding that temperature directly impacts host VC and ZIKV selective environment. Genetic complexity continues to be multifaceted and is dependent on many more factors than just EIT. Input diversity, host species, and host metabolic and immunological response all play a role in mutational spectrum. Using our finding we can better predict what conditions lead to high population diversity, and therefore robustness or population decline. Additionally, we have identified two novel variants that appear to provide a mechanism for adaptation in novel environments, and further assessment of mode of action of these variants may provide insight into the requirements needed for successful host adaption.

4.2 Future Considerations

Collectively this body of work allowed us to make great progress in unraveling the interaction of environmental temperature and arbovirus population diversity. However, a caveat of this work is that the majority of it focuses on constant temperatures, which does not model

what mosquito vectors would be exposed to in nature. Therefore, future direction should include detailed assessment of extrinsic incubation temperatures using a range of diurnal temperature groups to identify how fluctuating temperatures impact arbovirus population structure during systemic infection. Additionally, it is important to take these finding back into a field setting in order to perform an assessment of the unimodal temperature distribution of population genetic markers and their relationships between mosquito-borne disease and temperature in the field.

References

1. Gubler, D.J., *Dengue and dengue hemorrhagic fever.* Clin Microbiol Rev, 1998. **11**(3): p. 480-96.
2. Weaver, S.C. and N. Vasilakis, *Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease.* Infect Genet Evol, 2009. **9**(4): p. 523-40.
3. Ross, R., *On some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood.* Br Med J, 1897. **2**(1929): p. 1786-8.
4. Holbrook, M.R., *Historical Perspectives on Flavivirus Research.* Viruses, 2017. **9**(5).
5. Chiong, M.A., *Dr. Carlos Finlay and yellow fever.* CMAJ, 1989. **141**(11): p. 1126.
6. Reed, W., et al., *The Etiology of Yellow Fever-A Preliminary Note.* Public Health Pap Rep, 1900. **26**: p. 37-53.
7. del Regato, J.A., *James Carroll: a biography.* Ann Diagn Pathol, 1998. **2**(5): p. 335-49.
8. Henchal, E.A. and J.R. Putnak, *The dengue viruses.* Clin Microbiol Rev, 1990. **3**(4): p. 376-96.
9. Gubler, D.J., *Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century.* Trends Microbiol, 2002. **10**(2): p. 100-3.
10. Halstead, S.B., *Dengue.* Lancet, 2007. **370**(9599): p. 1644-52.
11. Robertson, S.E., et al., *Yellow fever: a decade of reemergence.* JAMA, 1996. **276**(14): p. 1157-62.
12. Tomashek, K.M., et al., *Disease Resurgence, Production Capability Issues and Safety Concerns in the Context of an Aging Population: Is There a Need for a New Yellow Fever Vaccine?* Vaccines (Basel), 2019. **7**(4).
13. Kramer, L.D., L.M. Styer, and G.D. Ebel, *A global perspective on the epidemiology of West Nile virus.* Annu Rev Entomol, 2008. **53**: p. 61-81.
14. Weaver, S.C. and M. Lecuit, *Chikungunya virus and the global spread of a mosquito-borne disease.* N Engl J Med, 2015. **372**(13): p. 1231-9.
15. Pierson, T.C. and M.S. Diamond, *The emergence of Zika virus and its new clinical syndromes.* Nature, 2018. **560**(7720): p. 573-581.
16. Karabatsos, N., *International Catalogue of Arboviruses*. 3rd ed. Am Soc Trop Med Hyg. 1985, San Antonio, TX.
17. Prevention, C.f.D.C.a. *Arbovirus Catalog*. [cited 2020 6/16/2020]; Available from: https://wwwn.cdc.gov/Arbocat/Default.aspx.
18. Junglen, S. and C. Drosten, *Virus discovery and recent insights into virus diversity in arthropods.* Curr Opin Microbiol, 2013. **16**(4): p. 507-13.
19. Hermance, M.E. and S. Thangamani, *Powassan Virus: An Emerging Arbovirus of Public Health Concern in North America.* Vector Borne Zoonotic Dis, 2017. **17**(7): p. 453-462.
20. Gubler, D.J., N. Vasilakis, and D. Musso, *History and Emergence of Zika Virus.* J Infect Dis, 2017. **216**(suppl_10): p. S860-S867.
21. Bhatt, S., et al., *The global distribution and burden of dengue.* Nature, 2013. **496**(7446): p. 504-7.
22. *International Committee on Taxonomy of Viruses ICTV*. March 2020 [cited 2020 6/16/2020]; Available from: https://talk.ictvonline.org/.

23.   Gould, E.A., et al., *Origins, evolution, and vector/host coadaptations within the genus Flavivirus.* Adv Virus Res, 2003. **59**: p. 277-314.

24.   Kuno, G., et al., *Phylogeny of the genus Flavivirus.* J Virol, 1998. **72**(1): p. 73-83.

25.   Kenney, J.L., et al., *Characterization of a novel insect-specific flavivirus from Brazil: potential for inhibition of infection of arthropod cells with medically important flaviviruses.* J Gen Virol, 2014. **95**(Pt 12): p. 2796-2808.

26.   Romo, H., et al., *Restriction of Zika virus infection and transmission in Aedes aegypti mediated by an insect-specific flavivirus.* Emerg Microbes Infect, 2018. **7**(1): p. 181.

27.   Goenaga, S., et al., *Potential for Co-Infection of a Mosquito-Specific Flavivirus, Nhumirim Virus, to Block West Nile Virus Transmission in Mosquitoes.* Viruses, 2015. **7**(11): p. 5801-12.

28.   Nouri, S., et al., *Insect-specific viruses: from discovery to potential translational applications.* Curr Opin Virol, 2018. **33**: p. 33-41.

29.   Guzman, H., et al., *Characterization of Three New Insect-Specific Flaviviruses: Their Relationship to the Mosquito-Borne Flavivirus Pathogens.* Am J Trop Med Hyg, 2018. **98**(2): p. 410-419.

30.   Blitvich, B.J. and A.E. Firth, *A Review of Flaviviruses that Have No Known Arthropod Vector.* Viruses, 2017. **9**(6).

31.   Sulkin, S.E., R.A. Sims, and R. Allen, *Isolation of St. Louis encephalitis virus from bats (Tadarida b. mexicana) in Texas.* Science, 1966. **152**(3719): p. 223-5.

32.   Davis, J.W. and J.L. Hardy, *Characterization of persistent Modoc viral infections in Syrian hamsters.* Infect Immun, 1974. **10**(2): p. 328-34.

33.   Grard, G., et al., *Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy.* Virology, 2007. **361**(1): p. 80-92.

34.   Shi, J., et al., *Tick-Borne Viruses.* Virol Sin, 2018. **33**(1): p. 21-43.

35.   Guzman, M.G., et al., *Dengue infection.* Nat Rev Dis Primers, 2016. **2**: p. 16055.

36.   Faria, N.R., et al., *Zika virus in the Americas: Early epidemiological and genetic findings.* Science, 2016. **352**(6283): p. 345-349.

37.   Faria, N.R., et al., *Establishment and cryptic transmission of Zika virus in Brazil and the Americas.* Nature, 2017. **546**(7658): p. 406-410.

38.   Grubaugh, N.D., et al., *Genomic epidemiology reveals multiple introductions of Zika virus into the United States.* Nature, 2017. **546**(7658): p. 401-405.

39.   Metsky, H.C., et al., *Zika virus evolution and spread in the Americas.* Nature, 2017. **546**(7658): p. 411-415.

40.   Worobey, M., *Epidemiology: Molecular mapping of Zika spread.* Nature, 2017. **546**(7658): p. 355-357.

41.   Pettersson, J.H., et al., *How Did Zika Virus Emerge in the Pacific Islands and Latin America?* mBio, 2016. **7**(5).

42.   Zhang, Y., et al., *Highly diversified Zika viruses imported to China, 2016.* Protein Cell, 2016. **7**(6): p. 461-4.

43.   Beaver, J.T., et al., *Evolution of Two Major Zika Virus Lineages: Implications for Pathology, Immune Response, and Vaccine Development.* Front Immunol, 2018. **9**: p. 1640.

44.   de Almeida, P.R., et al., *Detection of a novel African-lineage-like Zika virus naturally infecting free-living neotropical primates in Southern Brazil.* bioRxiv, 2019: p. 828871.

45.   Weaver, S.C., *Emergence of Epidemic Zika Virus Transmission and Congenital Zika Syndrome: Are Recently Evolved Traits to Blame?* mBio, 2017. **8**(1).

46.   Dick, G.W., S.F. Kitchen, and A.J. Haddow, *Zika virus. I. Isolations and serological specificity.* Trans R Soc Trop Med Hyg, 1952. **46**(5): p. 509-20.

47.   Macnamara, F.N., *Zika virus: a report on three cases of human infection during an epidemic of jaundice in Nigeria.* Trans R Soc Trop Med Hyg, 1954. **48**(2): p. 139-45.

48.   Marchette, N.J., R. Garcia, and A. Rudnick, *Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia.* Am J Trop Med Hyg, 1969. **18**(3): p. 411-5.

49.   Wang, L., et al., *From Mosquitos to Humans: Genetic Evolution of Zika Virus.* Cell Host Microbe, 2016. **19**(5): p. 561-5.

50.   Olson, J.G., et al., *Zika virus, a cause of fever in Central Java, Indonesia.* Trans R Soc Trop Med Hyg, 1981. **75**(3): p. 389-93.

51.   Vasilakis, N. and S.C. Weaver, *The history and evolution of human dengue emergence.* Adv Virus Res, 2008. **72**: p. 1-76.

52.   Simpson, D.I., *Zika Virus Infection in Man.* Trans R Soc Trop Med Hyg, 1964. **58**: p. 335-8.

53.   Fagbami, A.H., *Zika virus infections in Nigeria: virological and seroepidemiological investigations in Oyo State.* J Hyg (Lond), 1979. **83**(2): p. 213-9.

54.   Filipe, A.R., C.M. Martins, and H. Rocha, *Laboratory infection with Zika virus after vaccination against yellow fever.* Arch Gesamte Virusforsch, 1973. **43**(4): p. 315-9.

55.   Duffy, M.R., et al., *Zika virus outbreak on Yap Island, Federated States of Micronesia.* N Engl J Med, 2009. **360**(24): p. 2536-43.

56.   Grard, G., et al., *Zika virus in Gabon (Central Africa)--2007: a new threat from Aedes albopictus?* PLoS Negl Trop Dis, 2014. **8**(2): p. e2681.

57.   Oehler, E., et al., *Zika virus infection complicated by Guillain-Barre syndrome--case report, French Polynesia, December 2013.* Euro Surveill, 2014. **19**(9).

58.   Buathong, R., et al., *Detection of Zika Virus Infection in Thailand, 2012-2014.* Am J Trop Med Hyg, 2015. **93**(2): p. 380-383.

59.   Musso, D., E.J. Nilles, and V.M. Cao-Lormeau, *Rapid spread of emerging Zika virus in the Pacific area.* Clin Microbiol Infect, 2014. **20**(10): p. O595-6.

60.   Hennessey, M., M. Fischer, and J.E. Staples, *Zika Virus Spreads to New Areas - Region of the Americas, May 2015-January 2016.* MMWR Morb Mortal Wkly Rep, 2016. **65**(3): p. 55-8.

61.   Passos, S.R.L., et al., *Detection of Zika Virus in April 2013 Patient Samples, Rio de Janeiro, Brazil.* Emerg Infect Dis, 2017. **23**(12): p. 2120-2121.

62.   Aliota, M.T., et al., *Zika in the Americas, year 2: What have we learned? What gaps remain? A report from the Global Virus Network.* Antiviral Res, 2017. **144**: p. 223-246.

63.   Cardoso, C.W., et al., *Outbreak of Exanthematous Illness Associated with Zika, Chikungunya, and Dengue Viruses, Salvador, Brazil.* Emerg Infect Dis, 2015. **21**(12): p. 2274-6.

64.   Pacheco, O., et al., *Zika Virus Disease in Colombia - Preliminary Report.* N Engl J Med, 2016.

65.   Lozier, M., et al., *Incidence of Zika Virus Disease by Age and Sex - Puerto Rico, November 1, 2015-October 20, 2016.* MMWR Morb Mortal Wkly Rep, 2016. **65**(44): p. 1219-1223.

66. Baud, D., et al., *An update on Zika virus infection.* Lancet, 2017. **390**(10107): p. 2099-2109.

67. Grubaugh, N.D., et al., *Travel Surveillance and Genomics Uncover a Hidden Zika Outbreak during the Waning Epidemic.* Cell, 2019. **178**(5): p. 1057-1071 e11.

68. O'Reilly, K.M., et al., *Projecting the end of the Zika virus epidemic in Latin America: a modelling analysis.* BMC Med, 2018. **16**(1): p. 180.

69. Phumee, A., et al., *Molecular Epidemiology and Genetic Diversity of Zika Virus from Field-Caught Mosquitoes in Various Regions of Thailand.* Pathogens, 2019. **8**(1).

70. Yadav, P.D., et al., *Zika virus outbreak in Rajasthan, India in 2018 was caused by a virus endemic to Asia.* Infect Genet Evol, 2019. **69**: p. 199-202.

71. Ruchusatsawat, K., et al., *Long-term circulation of Zika virus in Thailand: an observational study.* Lancet Infect Dis, 2019. **19**(4): p. 439-446.

72. Abram, P.K., et al., *Behavioural effects of temperature on ectothermic animals: unifying thermal physiology and behavioural plasticity.* Biol Rev Camb Philos Soc, 2017. **92**(4): p. 1859-1876.

73. HUEY, R.B. and R.D. STEVENSON, *Integrating Thermal Physiology and Ecology of Ectotherms: A Discussion of Approaches.* American Zoologist, 2015. **19**(1): p. 357-366.

74. Juliano, S.A., et al., *Desiccation and thermal tolerance of eggs and the coexistence of competing mosquitoes.* Oecologia, 2002. **130**(3): p. 458-469.

75. Couret, J. and M.Q. Benedict, *A meta-analysis of the factors influencing development rate variation in Aedes aegypti (Diptera: Culicidae).* BMC Ecol, 2014. **14**: p. 3.

76. Rowley, W.A. and C.L. Graham, *The effect of temperature and relative humidity on the flight performance of female Aedes aegypti.* J Insect Physiol, 1968. **14**(9): p. 1251-7.

77. Benoit, J.B., et al., *Heat shock proteins contribute to mosquito dehydration tolerance.* J Insect Physiol, 2010. **56**(2): p. 151-6.

78. *Life Cycle of Aedes aegypti and Ae. albopictus Mosquitoes*. March 5, 2020 [cited 2020 7/4/2020]; Available from: https://www.cdc.gov/mosquitoes/about/life-cycles/aedes.html.

79. Tjaden, N.B., et al., *Mosquito-Borne Diseases: Advances in Modelling Climate-Change Impacts.* Trends Parasitol, 2018. **34**(3): p. 227-245.

80. Kraemer, M.U.G., et al., *Past and future spread of the arbovirus vectors Aedes aegypti and Aedes albopictus.* Nat Microbiol, 2019. **4**(5): p. 854-863.

81. Ryan, S.J., et al., *Global expansion and redistribution of Aedes-borne virus transmission risk with climate change.* PLoS Negl Trop Dis, 2019. **13**(3): p. e0007213.

82. Eisen, L., et al., *The impact of temperature on the bionomics of Aedes (Stegomyia) aegypti, with special reference to the cool geographic range margins.* J Med Entomol, 2014. **51**(3): p. 496-516.

83. Christophers, S.R., *Aedes aegypti (L.) the Yellow Fever Mosquito.* 1960, London, UK: Cambridge University Press.

84. Reinhold, J.M., C.R. Lazzari, and C. Lahondere, *Effects of the Environmental Temperature on Aedes aegypti and Aedes albopictus Mosquitoes: A Review.* Insects, 2018. **9**(4).

85. Delatte, H., et al., *Influence of temperature on immature development, survival, longevity, fecundity, and gonotrophic cycles of Aedes albopictus, vector of chikungunya and dengue in the Indian Ocean.* J Med Entomol, 2009. **46**(1): p. 33-41.

86. Teng, H.J. and C.S. Apperson, *Development and survival of immature Aedes albopictus and Aedes triseriatus (Diptera: Culicidae) in the laboratory: effects of density, food, and competition on response to temperature.* J Med Entomol, 2000. **37**(1): p. 40-52.

87. Carrington, L.B., et al., *Large diurnal temperature fluctuations negatively influence Aedes aegypti (Diptera: Culicidae) life-history traits.* J Med Entomol, 2013. **50**(1): p. 43-51.

88. Harrington, L.C., J.D. Edman, and T.W. Scott, *Why do female Aedes aegypti (Diptera: Culicidae) feed preferentially and frequently on human blood?* J Med Entomol, 2001. **38**(3): p. 411-22.

89. Delatte, H., et al., *Blood-feeding behavior of Aedes albopictus, a vector of Chikungunya on La Reunion.* Vector Borne Zoonotic Dis, 2010. **10**(3): p. 249-58.

90. Kamgang, B., et al., *Notes on the blood-feeding behavior of Aedes albopictus (Diptera: Culicidae) in Cameroon.* Parasit Vectors, 2012. **5**: p. 57.

91. McMeniman, C.J., et al., *Multimodal integration of carbon dioxide and other sensory cues drives mosquito attraction to humans.* Cell, 2014. **156**(5): p. 1060-71.

92. Scott, T.W., et al., *Detection of multiple blood feeding in Aedes aegypti (Diptera: Culicidae) during a single gonotrophic cycle using a histologic technique.* J Med Entomol, 1993. **30**(1): p. 94-9.

93. Scott, T.W., et al., *Longitudinal studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: blood feeding frequency.* J Med Entomol, 2000. **37**(1): p. 89-101.

94. Gutierrez-Bugallo, G., et al., *Vector-borne transmission and evolution of Zika virus.* Nat Ecol Evol, 2019. **3**(4): p. 561-569.

95. Johnson, B.K., et al., *Arbovirus isolations from, and serological studies on, wild and domestic vertebrates from Kano Plain, Kenya.* Trans R Soc Trop Med Hyg, 1977. **71**(6): p. 512-7.

96. Ragan, I.K., et al., *Investigating the Potential Role of North American Animals as Hosts for Zika Virus.* Vector Borne Zoonotic Dis, 2017. **17**(3): p. 161-164.

97. Guedes, D.R., et al., *Zika virus replication in the mosquito Culex quinquefasciatus in Brazil.* Emerg Microbes Infect, 2017. **6**(8): p. e69.

98. Haddow, A.D., et al., *Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage.* PLoS Negl Trop Dis, 2012. **6**(2): p. e1477.

99. Althouse, B.M., et al., *Impact of climate and mosquito vector abundance on sylvatic arbovirus circulation dynamics in Senegal.* Am J Trop Med Hyg, 2015. **92**(1): p. 88-97.

100. Faye, O., et al., *Quantitative real-time PCR detection of Zika virus and evaluation with field-caught mosquitoes.* Virol J, 2013. **10**: p. 311.

101. Marquardt WC, B.W., Freier JE, Hagedorn HH, Hemingway J, Higgs S, James AA, Kondratieff B, Moore CG, *Biology of Disease Vectors*. 2nd ed. 2005: Elsevier Academic Press.

102. Tesla, B., et al., *Temperature drives Zika virus transmission: evidence from empirical and mathematical models.* Proc Biol Sci, 2018. **285**(1884).

103. Winokur, O.C., et al., *Impact of temperature on the extrinsic incubation period of Zika virus in Aedes aegypti.* PLoS Negl Trop Dis, 2020. **14**(3): p. e0008047.

104. Onyango, M.G., et al., *Increased temperatures reduce the vectorial capacity of Aedes mosquitoes for Zika virus.* Emerg Microbes Infect, 2020. **9**(1): p. 67-77.

105. Azar, S.R. and S.C. Weaver, *Vector Competence: What Has Zika Virus Taught Us?* Viruses, 2019. **11**(9).

106.   *Flaviviruses*, in *Fields Virology*, P.M.H. David M. Knipe, Jeffrey Cohen, Diane Griffin, Robert lamb, Malcolm Martin, Vincent Racaniello, Bernard Roizman, Editor. 2013, Lippincott Williams & Wilkins: Philadelphia, PA. p. 747-794.

107.   Hsu, C.L. and A. Stevens, *Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure.* Mol Cell Biol, 1993. **13**(8): p. 4826-35.

108.   Barrows, N.J., et al., *Biochemistry and Molecular Biology of Flaviviruses.* Chem Rev, 2018. **118**(8): p. 4448-4482.

109.   Gebhard, L.G., C.V. Filomatori, and A.V. Gamarnik, *Functional RNA elements in the dengue virus genome.* Viruses, 2011. **3**(9): p. 1739-56.

110.   Brinton, M.A., *Replication cycle and molecular biology of the West Nile virus.* Viruses, 2013. **6**(1): p. 13-53.

111.   Sirohi, D., et al., *The 3.8 A resolution cryo-EM structure of Zika virus.* Science, 2016. **352**(6284): p. 467-70.

112.   Garcia-Blanco, M.A., et al., *Flavivirus RNA transactions from viral entry to genome replication.* Antiviral Res, 2016. **134**: p. 244-249.

113.   Ming, G.L., H. Tang, and H. Song, *Advances in Zika Virus Research: Stem Cell Models, Challenges, and Opportunities.* Cell Stem Cell, 2016. **19**(6): p. 690-702.

114.   Meertens, L., et al., *The TIM and TAM families of phosphatidylserine receptors mediate dengue virus entry.* Cell Host Microbe, 2012. **12**(4): p. 544-57.

115.   Tassaneetrithep, B., et al., *DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells.* J Exp Med, 2003. **197**(7): p. 823-9.

116.   Salas-Benito, J., et al., *Evidence that the 45-kD glycoprotein, part of a putative dengue virus receptor complex in the mosquito cell line C6/36, is a heat-shock related protein.* Am J Trop Med Hyg, 2007. **77**(2): p. 283-90.

117.   Tabata, T., et al., *Zika Virus Targets Different Primary Human Placental Cells, Suggesting Two Routes for Vertical Transmission.* Cell Host Microbe, 2016. **20**(2): p. 155-66.

118.   Hamel, R., et al., *Biology of Zika Virus Infection in Human Skin Cells.* J Virol, 2015. **89**(17): p. 8880-96.

119.   Wells, M.F., et al., *Genetic Ablation of AXL Does Not Protect Human Neural Progenitor Cells and Cerebral Organoids from Zika Virus Infection.* Cell Stem Cell, 2016. **19**(6): p. 703-708.

120.   Perera, R., M. Khaliq, and R.J. Kuhn, *Closing the door on flaviviruses: entry as a target for antiviral drug design.* Antiviral Res, 2008. **80**(1): p. 11-22.

121.   Chu, J.J. and M.L. Ng, *Infectious entry of West Nile virus occurs through a clathrin-mediated endocytic pathway.* J Virol, 2004. **78**(19): p. 10543-55.

122.   Mosso, C., et al., *Endocytic pathway followed by dengue virus to infect the mosquito cell line C6/36 HT.* Virology, 2008. **378**(1): p. 193-9.

123.   Mukhopadhyay, S., R.J. Kuhn, and M.G. Rossmann, *A structural perspective of the flavivirus life cycle.* Nat Rev Microbiol, 2005. **3**(1): p. 13-22.

124.   Reid, D.W. and C.V. Nicchitta, *Diversity and selectivity in mRNA translation on the endoplasmic reticulum.* Nat Rev Mol Cell Biol, 2015. **16**(4): p. 221-31.

125.   Jan, C.H., C.C. Williams, and J.S. Weissman, *LOCAL TRANSLATION. Response to Comment on "Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling".* Science, 2015. **348**(6240): p. 1217.

126.   Reid, D.W., et al., *Dengue Virus Selectively Annexes Endoplasmic Reticulum-Associated Translation Machinery as a Strategy for Co-opting Host Cell Protein Synthesis.* J Virol, 2018. **92**(7).

127.   Villordo, S.M. and A.V. Gamarnik, *Genome cyclization as strategy for flavivirus RNA replication.* Virus Res, 2009. **139**(2): p. 230-9.

128.   Byk, L.A. and A.V. Gamarnik, *Properties and Functions of the Dengue Virus Capsid Protein.* Annu Rev Virol, 2016. **3**(1): p. 263-281.

129.   Zhang, Y., et al., *Structures of immature flavivirus particles.* EMBO J, 2003. **22**(11): p. 2604-13.

130.   Stadler, K., et al., *Proteolytic activation of tick-borne encephalitis virus by furin.* J Virol, 1997. **71**(11): p. 8475-81.

131.   Oliveira, E.R.A., R.B. de Alencastro, and B.A.C. Horta, *New insights into flavivirus biology: the influence of pH over interactions between prM and E proteins.* J Comput Aided Mol Des, 2017. **31**(11): p. 1009-1019.

132.   Sotcheff, S. and A. Routh, *Understanding Flavivirus Capsid Protein Functions: The Tip of the Iceberg.* Pathogens, 2020. **9**(1).

133.   Slomnicki, L.P., et al., *Ribosomal stress and Tp53-mediated neuronal apoptosis in response to capsid protein of the Zika virus.* Sci Rep, 2017. **7**(1): p. 16652.

134.   Samuel, G.H., et al., *Yellow fever virus capsid protein is a potent suppressor of RNA silencing that binds double-stranded RNA.* Proc Natl Acad Sci U S A, 2016. **113**(48): p. 13863-13868.

135.   Wang, C.C., et al., *Analysis of the nucleoside triphosphatase, RNA triphosphatase, and unwinding activities of the helicase domain of dengue virus NS3 protein.* FEBS Lett, 2009. **583**(4): p. 691-6.

136.   Bos, S., et al., *The structural proteins of epidemic and historical strains of Zika virus differ in their ability to initiate viral infection in human host cells.* Virology, 2018. **516**: p. 265-273.

137.   Yuan, L., et al., *A single mutation in the prM protein of Zika virus contributes to fetal microcephaly.* Science, 2017. **358**(6365): p. 933-936.

138.   Yu, K., et al., *Structural, antigenic, and evolutionary characterizations of the envelope protein of newly emerging Duck Tembusu Virus.* PLoS One, 2013. **8**(8): p. e71319.

139.   Hanna, S.L., et al., *N-linked glycosylation of west nile virus envelope proteins influences particle assembly and infectivity.* J Virol, 2005. **79**(21): p. 13262-74.

140.   Zhang, X., et al., *Structures and Functions of the Envelope Glycoprotein in Flavivirus Infections.* Viruses, 2017. **9**(11).

141.   Zhang, S., et al., *Role of BC loop residues in structure, function and antigenicity of the West Nile virus envelope protein receptor-binding domain III.* Virology, 2010. **403**(1): p. 85-91.

142.   Glasner, D.R., et al., *The Good, the Bad, and the Shocking: The Multiple Roles of Dengue Virus Nonstructural Protein 1 in Protection and Pathogenesis.* Annu Rev Virol, 2018. **5**(1): p. 227-253.

143.   Ci, Y., et al., *Zika NS1-induced ER remodeling is essential for viral replication.* J Cell Biol, 2020. **219**(2).

144.   Mazeaud, C., W. Freppel, and L. Chatel-Chaix, *The Multiples Fates of the Flavivirus RNA Genome During Pathogenesis.* Front Genet, 2018. **9**: p. 595.

145.   Yang, Y., et al., *A cDNA Clone-Launched Platform for High-Yield Production of Inactivated Zika Vaccine.* EBioMedicine, 2017. **17**: p. 145-156.

146.   Mackenzie, J.M., M.K. Jones, and P.R. Young, *Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication.* Virology, 1996. **220**(1): p. 232-40.

147.   Muylaert, I.R., et al., *Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence.* Virology, 1996. **222**(1): p. 159-68.

148.   Muller, D.A. and P.R. Young, *The flavivirus NS1 protein: molecular and structural biology, immunology, role in pathogenesis and application as a diagnostic biomarker.* Antiviral Res, 2013. **98**(2): p. 192-208.

149.   Muylaert, I.R., R. Galler, and C.M. Rice, *Genetic analysis of the yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation.* J Virol, 1997. **71**(1): p. 291-8.

150.   Zhang, X., et al., *Zika Virus NS2A-Mediated Virion Assembly.* mBio, 2019. **10**(5).

151.   Wu, R.H., et al., *Scanning mutagenesis studies reveal a potential intramolecular interaction within the C-terminal half of dengue virus NS2A involved in viral RNA replication and virus assembly and secretion.* J Virol, 2015. **89**(8): p. 4281-95.

152.   Shrivastava, G., et al., *NS2A comprises a putative viroporin of Dengue virus 2.* Virulence, 2017. **8**(7): p. 1450-1456.

153.   Leon-Juarez, M., et al., *Recombinant Dengue virus protein NS2B alters membrane permeability in different membrane models.* Virol J, 2016. **13**: p. 1.

154.   Chambers, T.J., et al., *Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein.* Proc Natl Acad Sci U S A, 1990. **87**(22): p. 8898-902.

155.   Aguirre, S., et al., *DENV inhibits type I IFN production in infected cells by cleaving human STING.* PLoS Pathog, 2012. **8**(10): p. e1002934.

156.   Yu, C.Y., et al., *Dengue virus targets the adaptor protein MITA to subvert host innate immunity.* PLoS Pathog, 2012. **8**(6): p. e1002780.

157.   Brand, C., M. Bisaillon, and B.J. Geiss, *Organization of the Flavivirus RNA replicase complex.* Wiley Interdiscip Rev RNA, 2017. **8**(6).

158.   Takahashi, H., et al., *Establishment of a robust dengue virus NS3-NS5 binding assay for identification of protein-protein interaction inhibitors.* Antiviral Res, 2012. **96**(3): p. 305-14.

159.   Welsch, S., et al., *Composition and three-dimensional architecture of the dengue virus replication and assembly sites.* Cell Host Microbe, 2009. **5**(4): p. 365-75.

160.   Mackenzie, J.M., et al., *Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A.* Virology, 1998. **245**(2): p. 203-15.

161.   Roosendaal, J., et al., *Regulated cleavages at the West Nile virus NS4A-2K-NS4B junctions play a major role in rearranging cytoplasmic membranes and Golgi trafficking of the NS4A protein.* J Virol, 2006. **80**(9): p. 4623-32.

162.   Miller, S., et al., *The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner.* J Biol Chem, 2007. **282**(12): p. 8873-82.

163. Munoz-Jordan, J.L., et al., *Inhibition of interferon signaling by dengue virus.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14333-8.

164. Kaufusi, P.H., et al., *Induction of endoplasmic reticulum-derived replication-competent membrane structures by West Nile virus non-structural protein 4B.* PLoS One, 2014. **9**(1): p. e84040.

165. Egloff, M.P., et al., *An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization.* EMBO J, 2002. **21**(11): p. 2757-68.

166. Ray, D., et al., *West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5.* J Virol, 2006. **80**(17): p. 8362-70.

167. Tan, B.H., et al., *Recombinant dengue type 1 virus NS5 protein expressed in Escherichia coli exhibits RNA-dependent RNA polymerase activity.* Virology, 1996. **216**(2): p. 317-25.

168. Issur, M., et al., *The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure.* RNA, 2009. **15**(12): p. 2340-50.

169. Tay, M.Y., et al., *The C-terminal 50 amino acid residues of dengue NS3 protein are important for NS3-NS5 interaction and viral replication.* J Biol Chem, 2015. **290**(4): p. 2379-94.

170. Yang, H., et al., *Temperature requirements for initiation of RNA-dependent RNA polymerization.* Virology, 2003. **314**(2): p. 706-15.

171. Ackermann, M. and R. Padmanabhan, *De novo synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase.* J Biol Chem, 2001. **276**(43): p. 39926-37.

172. Steinhauer, D.A., E. Domingo, and J.J. Holland, *Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase.* Gene, 1992. **122**(2): p. 281-8.

173. Domingo, E. and C. Perales, *Viral quasispecies.* PLoS Genet, 2019. **15**(10): p. e1008271.

174. Lauring, A.S. and R. Andino, *Quasispecies theory and the behavior of RNA viruses.* PLoS Pathog, 2010. **6**(7): p. e1001005.

175. Andino, R. and E. Domingo, *Viral quasispecies.* Virology, 2015. **479-480**: p. 46-51.

176. Eigen, M., *Selforganization of matter and the evolution of biological macromolecules.* Naturwissenschaften, 1971. **58**(10): p. 465-523.

177. Eigen, M. and P. Schuster, *The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle.* Naturwissenschaften, 1977. **64**(11): p. 541-65.

178. Domingo, E., et al., *Viruses as quasispecies: biological implications.* Curr Top Microbiol Immunol, 2006. **299**: p. 51-82.

179. Domingo, E., et al., *Nucleotide sequence heterogeneity of an RNA phage population.* Cell, 1978. **13**(4): p. 735-44.

180. Domingo, E. and C. Perales, *Quasispecies and virus.* Eur Biophys J, 2018. **47**(4): p. 443-457.

181. Mandary, M.B., M. Masomian, and C.L. Poh, *Impact of RNA Virus Evolution on Quasispecies Formation and Virulence.* Int J Mol Sci, 2019. **20**(18).

182. Perales, C., *Quasispecies dynamics and clinical significance of hepatitis C virus (HCV) antiviral resistance.* Int J Antimicrob Agents, 2018: p. 105562.

183.   Garcia-Arriaza, J., et al., *Dynamics of mutation and recombination in a replicating population of complementing, defective viral genomes.* J Mol Biol, 2006. **360**(3): p. 558-72.

184.   Liu, Y., et al., *Evolutionary enhancement of Zika virus infectivity in Aedes aegypti mosquitoes.* Nature, 2017. **545**(7655): p. 482-486.

185.   Tsetsarkin, K.A., et al., *Multi-peaked adaptive landscape for chikungunya virus evolution predicts continued fitness optimization in Aedes albopictus mosquitoes.* Nat Commun, 2014. **5**: p. 4084.

186.   Wang, H., et al., *Analysis of Synonymous Codon Usage Bias of Zika Virus and Its Adaption to the Hosts.* PLoS One, 2016. **11**(11): p. e0166260.

187.   Grubaugh, N.D., et al., *Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching.* Cell Host Microbe, 2016. **19**(4): p. 481-92.

188.   Huang, Y.S., S. Higgs, and D.L. Vanlandingham, *Arbovirus-Mosquito Vector-Host Interactions and the Impact on Transmission and Disease Pathogenesis of Arboviruses.* Front Microbiol, 2019. **10**: p. 22.

189.   Hardy, J.L., et al., *Intrinsic factors affecting vector competence of mosquitoes for arboviruses.* Annu Rev Entomol, 1983. **28**: p. 229-62.

190.   Gutierrez, S., Y. Michalakis, and S. Blanc, *Virus population bottlenecks during within-host progression and host-to-host transmission.* Curr Opin Virol, 2012. **2**(5): p. 546-55.

191.   Ciota, A.T., et al., *Quantification of intrahost bottlenecks of West Nile virus in Culex pipiens mosquitoes using an artificial mutant swarm.* Infect Genet Evol, 2012. **12**(3): p. 557-64.

192.   Patterson, E.I., et al., *Mosquito bottlenecks alter viral mutant swarm in a tissue and time-dependent manner with contraction and expansion of variant positions and diversity.* Virus Evol, 2018. **4**(1): p. vey001.

193.   Manrubia, S.C., et al., *High mutation rates, bottlenecks, and robustness of RNA viral quasispecies.* Gene, 2005. **347**(2): p. 273-82.

194.   Lequime, S., et al., *Genetic Drift, Purifying Selection and Vector Genotype Shape Dengue Virus Intra-host Genetic Diversity in Mosquitoes.* PLoS Genet, 2016. **12**(6): p. e1006111.

195.   Prasad, A.N., D.E. Brackney, and G.D. Ebel, *The role of innate immunity in conditioning mosquito susceptibility to West Nile virus.* Viruses, 2013. **5**(12): p. 3142-70.

196.   Blair, C.D. and K.E. Olson, *The role of RNA interference (RNAi) in arbovirus-vector interactions.* Viruses, 2015. **7**(2): p. 820-43.

197.   Noval, M.G., et al., *Mapping the evolutionary landscape of Zika virus infection in immunocompromised mice.* bioRxiv, 2019: p. 839803.

198.   Coffey, L.L., et al., *Factors shaping the adaptive landscape for arboviruses: implications for the emergence of disease.* Future Microbiol, 2013. **8**(2): p. 155-76.

199.   Forrester, N.L., et al., *Vector-borne transmission imposes a severe bottleneck on an RNA virus population.* PLoS Pathog, 2012. **8**(9): p. e1002897.

200.   Riemersma, K.K., et al., *Chikungunya Virus Fidelity Variants Exhibit Differential Attenuation and Population Diversity in Cell Culture and Adult Mice.* J Virol, 2019. **93**(3).

201. Grubaugh, N.D., et al., *Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness.* PLoS Pathog, 2015. **11**(5): p. e1004874.

202. Samuel, M.A. and M.S. Diamond, *Pathogenesis of West Nile Virus infection: a balance between virulence, innate and adaptive immunity, and viral evasion.* J Virol, 2006. **80**(19): p. 9349-60.

203. Kuno, G., et al., *Vertebrate Reservoirs of Arboviruses: Myth, Synonym of Amplifier, or Reality?* Viruses, 2017. **9**(7).

204. Garmel, G.M., *An Introduction to Clinical Emergency Medicine.* 2nd ed. 2012.

205. Chamberlain, R.W. and W.D. Sudia, *The effects of temperature upon the extrinsic incubation of eastern equine encephalitis in mosquitoes.* Am J Hyg, 1955. **62**(3): p. 295-305.

206. Turell, M.J., C.A. Rossi, and C.L. Bailey, *Effect of extrinsic incubation temperature on the ability of Aedes taeniorhynchus and Culex pipiens to transmit Rift Valley fever virus.* Am J Trop Med Hyg, 1985. **34**(6): p. 1211-8.

207. Lane, W.C., et al., *The Efficacy of the Interferon Alpha/Beta Response versus Arboviruses Is Temperature Dependent.* mBio, 2018. **9**(2).

208. Reisen, W.K., Y. Fang, and V.M. Martinez, *Effects of temperature on the transmission of west nile virus by Culex tarsalis (Diptera: Culicidae).* J Med Entomol, 2006. **43**(2): p. 309-17.

209. Richards, S.L., et al., *Environmental and biological factors influencing Culex pipiens quinquefasciatus Say (Diptera: Culicidae) vector competence for Saint Louis encephalitis virus.* Am J Trop Med Hyg, 2009. **81**(2): p. 264-72.

210. Keyel, A.C., et al., *Seasonal temperatures and hydrological conditions improve the prediction of West Nile virus infection rates in Culex mosquitoes and human case counts in New York and Connecticut.* PLoS One, 2019. **14**(6): p. e0217854.

211. Kilpatrick, A.M., et al., *Temperature, viral genetics, and the transmission of West Nile virus by Culex pipiens mosquitoes.* PLoS Pathog, 2008. **4**(6): p. e1000092.

212. Turell, M.J., *Effect of environmental temperature on the vector competence of Aedes taeniorhynchus for Rift Valley fever and Venezuelan equine encephalitis viruses.* Am J Trop Med Hyg, 1993. **49**(6): p. 672-6.

213. Vogels, C.B., et al., *Vector competence of northern European Culex pipiens biotypes and hybrids for West Nile virus is differentially affected by temperature.* Parasit Vectors, 2016. **9**(1): p. 393.

214. Patz, J.A., et al., *Global climate change and emerging infectious diseases.* JAMA, 1996. **275**(3): p. 217-23.

215. Garcia-Luna, S.M., et al., *Variation in competence for ZIKV transmission by Aedes aegypti and Aedes albopictus in Mexico.* PLoS Negl Trop Dis, 2018. **12**(7): p. e0006599.

216. Weger-Lucarelli, J., et al., *Vector Competence of American Mosquitoes for Three Strains of Zika Virus.* PLoS Negl Trop Dis, 2016. **10**(10): p. e0005101.

217. Gendernalik, A., et al., *American Aedes vexans Mosquitoes are Competent Vectors of Zika Virus.* Am J Trop Med Hyg, 2017. **96**(6): p. 1338-1340.

218. Kenney, J.L., et al., *Transmission Incompetence of Culex quinquefasciatus and Culex pipiens pipiens from North America for Zika Virus.* Am J Trop Med Hyg, 2017. **96**(5): p. 1235-1240.

219. Paaijmans, K.P., A.F. Read, and M.B. Thomas, *Understanding the link between malaria risk and climate.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13844-9.

220. Carrington, L.B., et al., *Fluctuations at a low mean temperature accelerate dengue virus transmission by Aedes aegypti.* PLoS Negl Trop Dis, 2013. **7**(4): p. e2190.

221. Lambrechts, L., et al., *Impact of daily temperature fluctuations on dengue virus transmission by Aedes aegypti.* Proc Natl Acad Sci U S A, 2011. **108**(18): p. 7460-5.

222. Alto, B.W., et al., *Stochastic temperatures impede RNA virus adaptation.* Evolution, 2013. **67**(4): p. 969-79.

223. McGee, L.W., et al., *Payoffs, not tradeoffs, in the adaptation of a virus to ostensibly conflicting selective pressures.* PLoS Genet, 2014. **10**(10): p. e1004611.

224. Holland, J., et al., *Rapid evolution of RNA genomes.* Science, 1982. **215**(4540): p. 1577-85.

225. Brackney, D.E., J.E. Beane, and G.D. Ebel, *RNAi targeting of West Nile virus in mosquito midguts promotes virus diversification.* PLoS Pathog, 2009. **5**(7): p. e1000502.

226. Jerzak, G., et al., *Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection.* J Gen Virol, 2005. **86**(Pt 8): p. 2175-83.

227. Brackney, D.E., et al., *Homogeneity of Powassan virus populations in naturally infected Ixodes scapularis.* Virology, 2010. **402**(2): p. 366-71.

228. Ruiz-Jarabo, C.M., et al., *Memory in viral quasispecies.* J Virol, 2000. **74**(8): p. 3543-7.

229. Vignuzzi, M., et al., *Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population.* Nature, 2006. **439**(7074): p. 344-8.

230. Weger-Lucarelli, J., et al., *Development and Characterization of Recombinant Virus Generated from a New World Zika Virus Infectious Clone.* J Virol, 2016.

231. Vera-Maloof, F.Z., et al., *Coevolution of the Ile1,016 and Cys1,534 Mutations in the Voltage Gated Sodium Channel Gene of Aedes aegypti in Mexico.* PLoS Negl Trop Dis, 2015. **9**(12): p. e0004263.

232. Fauver, J.R., et al., *Temporal and Spatial Variability of Entomological Risk Indices for West Nile Virus Infection in Northern Colorado: 2006-2013.* J Med Entomol, 2016. **53**(2): p. 425-34.

233. Lanciotti, R.S., et al., *Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007.* Emerg Infect Dis, 2008. **14**(8): p. 1232-9.

234. Garcia-Nafria, J., J.F. Watson, and I.H. Greger, *IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly.* Sci Rep, 2016. **6**: p. 27459.

235. Deardorff, E.R., et al., *West Nile virus experimental evolution in vivo and the trade-off hypothesis.* PLoS Pathog, 2011. **7**(11): p. e1002335.

236. Koster, J. and S. Rahmann, *Snakemake-a scalable bioinformatics workflow engine.* Bioinformatics, 2018. **34**(20): p. 3600.

237. Martin, M., *Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads.* EMBnet.journal. **17**(1).

238. Lee, W.P., et al., *MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping.* PLoS One, 2014. **9**(3): p. e90581.

239. Grubaugh, N.D., et al., *Mosquitoes Transmit Unique West Nile Virus Populations during Each Feeding Episode.* Cell Rep, 2017. **19**(4): p. 709-718.

240. *Picard.* 6/21/2020]; Available from: https://broadinstitute.github.io/picard/.

241. Van der Auwera, G.A., et al., *From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline.* Curr Protoc Bioinformatics, 2013. **43**: p. 11 10 1-11 10 33.

242. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.

243. Wilm, A., et al., *LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets.* Nucleic Acids Res, 2012. **40**(22): p. 11189-201.

244. Danecek, P., et al., *The variant call format and VCFtools.* Bioinformatics, 2011. **27**(15): p. 2156-8.

245. Librado, P. and J. Rozas, *DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.* Bioinformatics, 2009. **25**(11): p. 1451-2.

246. Franz, A.W., et al., *Tissue Barriers to Arbovirus Infection in Mosquitoes.* Viruses, 2015. **7**(7): p. 3741-67.

247. Mordecai, E.A., et al., *Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models.* PLoS Negl Trop Dis, 2017. **11**(4): p. e0005568.

248. Paaijmans, K.P., et al., *Influence of climate on malaria transmission depends on daily temperature variation.* Proc Natl Acad Sci U S A, 2010. **107**(34): p. 15135-9.

249. Carrington, L.B., et al., *Reduction of Aedes aegypti vector competence for dengue virus under large temperature fluctuations.* Am J Trop Med Hyg, 2013. **88**(4): p. 689-97.

250. Paaijmans, K.P. and M.B. Thomas, *The influence of mosquito resting behaviour and associated microclimate for malaria risk.* Malar J, 2011. **10**: p. 183.

251. Paaijmans, K.P., et al., *Relevant microclimate for determining the development rate of malaria mosquitoes and possible implications of climate change.* Malar J, 2010. **9**: p. 196.

252. Jemielity, S., et al., *TIM-family proteins promote infection of multiple enveloped viruses through virion-associated phosphatidylserine.* PLoS Pathog, 2013. **9**(3): p. e1003232.

253. Stapleford, K.A., et al., *Emergence and transmission of arbovirus evolutionary intermediates with epidemic potential.* Cell Host Microbe, 2014. **15**(6): p. 706-16.

254. Crill, W.D. and J.T. Roehrig, *Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells.* J Virol, 2001. **75**(16): p. 7769-73.

255. Xu, X., et al., *Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS1 structure.* EMBO J, 2016. **35**(20): p. 2170-2178.

256. Song, H., et al., *Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses.* Nat Struct Mol Biol, 2016. **23**(5): p. 456-8.

257. Duggal, N.K., et al., *Mutations present in a low-passage Zika virus isolate result in attenuated pathogenesis in mice.* Virology, 2019. **530**: p. 19-26.

258. Dolan, P.T., Z.J. Whitfield, and R. Andino, *Mapping the Evolutionary Potential of RNA Viruses.* Cell Host Microbe, 2018. **23**(4): p. 435-446.

259. PAHO. *Cases of Zika Virus Disease.* 6/21/2020]; Available from: https://www.paho.org/data/index.php/en/?option=com_content&view=article&id=524:zika-weekly-en&Itemid=352.

260. Liu, T., et al., *Antiviral systems in vector mosquitoes.* Dev Comp Immunol, 2018. **83**: p. 34-43.

261. O'Neill, L.A., D. Golenbock, and A.G. Bowie, *The history of Toll-like receptors - redefining innate immunity.* Nat Rev Immunol, 2013. **13**(6): p. 453-60.

262. Halbach, R., S. Junglen, and R.P. van Rij, *Mosquito-specific and mosquito-borne viruses: evolution, infection, and host defense.* Curr Opin Insect Sci, 2017. **22**: p. 16-27.

263. Netea, M.G., et al., *Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens.* Cell Host Microbe, 2019. **25**(1): p. 13-26.

264. Marinho, R.A., et al., *Effects of temperature on the life cycle, expansion, and dispersion of Aedes aegypti (Diptera: Culicidae) in three cities in Paraiba, Brazil.* J Vector Ecol, 2016. **41**(1): p. 1-10.

265. Hugo, L.E., et al., *Vector competence of Australian Aedes aegypti and Aedes albopictus for an epidemic strain of Zika virus.* PLoS Negl Trop Dis, 2019. **13**(4): p. e0007281.

266. Tramonte, A.R. and R.C. Christofferson, *Investigating the probability of establishment of Zika virus and detection through mosquito surveillance under different temperature conditions.* PLoS One, 2019. **14**(3): p. e0214306.

267. Main, B.J., et al., *Vector competence of Aedes aegypti, Culex tarsalis, and Culex quinquefasciatus from California for Zika virus.* PLoS Negl Trop Dis, 2018. **12**(6): p. e0006524.

268. Boyer, S., et al., *An overview of mosquito vectors of Zika virus.* Microbes Infect, 2018. **20**(11-12): p. 646-660.

269. Azar, S.R., et al., *Differential Vector Competency of Aedes albopictus Populations from the Americas for Zika Virus.* Am J Trop Med Hyg, 2017. **97**(2): p. 330-339.

270. Roundy, C.M., et al., *Variation in Aedes aegypti Mosquito Competence for Zika Virus Transmission.* Emerg Infect Dis, 2017. **23**(4): p. 625-632.

271. Ciota, A.T., et al., *Effects of Zika Virus Strain and Aedes Mosquito Species on Vector Competence.* Emerg Infect Dis, 2017. **23**(7): p. 1110-1117.

272. Ciota, A.T., et al., *Cell-specific adaptation of two flaviviruses following serial passage in mosquito cell culture.* Virology, 2007. **357**(2): p. 165-74.

273. Read, J.A. and F.C. Miller, *The bupivacaine paracervical block in labor and its effect on quantitative uterine activity.* Obstet Gynecol, 1979. **53**(2): p. 166-70.

274. Weger-Lucarelli, J., et al., *Rescue and Characterization of Recombinant Virus from a New World Zika Virus Infectious Clone.* J Vis Exp, 2017(124).

275. Hall, G.S. and D.P. Little, *Relative quantitation of virus population size in mixed genotype infections using sequencing chromatograms.* J Virol Methods, 2007. **146**(1-2): p. 22-8.

276. Rozas, J., et al., *DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets.* Mol Biol Evol, 2017. **34**(12): p. 3299-3302.

277. Dai, L., et al., *Structures of the Zika Virus Envelope Protein and Its Complex with a Flavivirus Broadly Protective Antibody.* Cell Host Microbe, 2016. **19**(5): p. 696-704.

278. Phoo, W.W., et al., *Structure of the NS2B-NS3 protease from Zika virus after self-cleavage.* Nat Commun, 2016. **7**: p. 13410.

279. Moudy, R.M., et al., *A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by Culex mosquitoes.* Am J Trop Med Hyg, 2007. **77**(2): p. 365-70.

280. Tsetsarkin, K.A., et al., *A single mutation in chikungunya virus affects vector specificity and epidemic potential.* PLoS Pathog, 2007. **3**(12): p. e201.

281. Fibriansah, G., et al., *Structural changes in dengue virus when exposed to a temperature of 37 degrees C.* J Virol, 2013. **87**(13): p. 7585-92.

282. Fan, Y.C., et al., *Formalin Inactivation of Japanese Encephalitis Virus Vaccine Alters the Antigenicity and Immunogenicity of a Neutralization Epitope in Envelope Protein Domain III.* PLoS Negl Trop Dis, 2015. **9**(10): p. e0004167.

283. Parham, P.E., et al., *Climate, environmental and socio-economic change: weighing up the balance in vector-borne disease transmission.* Philos Trans R Soc Lond B Biol Sci, 2015. **370**(1665).

284. Mordecai, E.A., et al., *Thermal biology of mosquito-borne disease.* Ecol Lett, 2019. **22**(10): p. 1690-1708.

285. Weger-Lucarelli, J., et al., *Using barcoded Zika virus to assess virus population structure in vitro and in Aedes aegypti mosquitoes.* Virology, 2018. **521**: p. 138-148.

286. Stephens, Z.D., et al., *Big Data: Astronomical or Genomical?* PLoS Biol, 2015. **13**(7): p. e1002195.

287. Grabherr, M.G., et al., *Full-length transcriptome assembly from RNA-Seq data without a reference genome.* Nat Biotechnol, 2011. **29**(7): p. 644-52.

288. Bankevich, A., et al., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.* J Comput Biol, 2012. **19**(5): p. 455-77.

289. Bushnell, B., J. Rood, and E. Singer, *BBMerge - Accurate paired shotgun read merging via overlap.* PLoS One, 2017. **12**(10): p. e0185056.

290. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.

291. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.

292. Yang, X., et al., *V-Phaser 2: variant inference for viral populations.* BMC Genomics, 2013. **14**: p. 674.

293. Poplin, R., et al., *A universal SNP and small-indel variant caller using deep neural networks.* Nat Biotechnol, 2018. **36**(10): p. 983-987.

294. Di Tommaso, P., et al., *Nextflow enables reproducible computational workflows.* Nat Biotechnol, 2017. **35**(4): p. 316-319.

295. Vivian, J., et al., *Toil enables reproducible, open source, big biomedical data analyses.* Nat Biotechnol, 2017. **35**(4): p. 314-316.

296. Kotliar, M., A.V. Kartashov, and A. Barski, *CWL-Airflow: a lightweight pipeline manager supporting Common Workflow Language.* Gigascience, 2019. **8**(7).

297. Genomes Project, C., et al., *A map of human genome variation from population-scale sequencing.* Nature, 2010. **467**(7319): p. 1061-73.

298. Genomes Project, C., et al., *An integrated map of genetic variation from 1,092 human genomes.* Nature, 2012. **491**(7422): p. 56-65.

299. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.

300. Goecks, J., et al., *Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences.* Genome Biol, 2010. **11**(8): p. R86.

301. Okonechnikov, K., et al., *Unipro UGENE: a unified bioinformatics toolkit.* Bioinformatics, 2012. **28**(8): p. 1166-7.

302. Trapnell, C., et al., *Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.* Nat Biotechnol, 2010. **28**(5): p. 511-5.

303. Fumagalli, M., et al., *Quantifying population genetic differentiation from next-generation sequencing data.* Genetics, 2013. **195**(3): p. 979-92.

304. Cornish, A. and C. Guda, *A Comparison of Variant Calling Pipelines Using Genome in a Bottle as a Reference.* Biomed Res Int, 2015. **2015**: p. 456479.

305. Grubaugh, N.D., et al., *Transmission bottlenecks and RNAi collectively influence tick-borne flavivirus evolution.* Virus Evol, 2016. **2**(2): p. vew033.

306. Grubaugh, N.D., et al., *West Nile Virus Population Structure, Injury, and Interferon-Stimulated Gene Expression in the Brain From a Fatal Case of Encephalitis.* Open Forum Infect Dis, 2016. **3**(1): p. ofv182.

307. Bush, S.J., et al., *Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism-calling pipelines.* Gigascience, 2020. **9**(2).

308. Sandmann, S., et al., *Evaluating Variant Calling Tools for Non-Matched Next-Generation Sequencing Data.* Sci Rep, 2017. **7**: p. 43169.

309. McCrone, J.T. and A.S. Lauring, *Measurements of Intrahost Viral Diversity Are Extremely Sensitive to Systematic Errors in Variant Calling.* J Virol, 2016. **90**(15): p. 6884-95.

310. Said Mohammed, K., et al., *Evaluating the performance of tools used to call minority variants from whole genome short-read data.* Wellcome Open Res, 2018. **3**: p. 21.

Appendix A: RNA Virus Population Genetics (RGP) Workflow

A1.1 Introduction

RNA virus population genetics is complex and can be performed in numerous ways. The most common methods are using markers of selection, diversity, and richness. Common analyses consider loci regions to identify genes under positive selection or nucleotides that are under selective pressures. Additionally, haplotype characterization and artificial barcoded viruses are used to characterize the population diversity, allowing us to unravel impacts on quasispecies [187, 285]. As RNA virus adapt and mutate rapidly, we can use the number of accumulated mutations or single nucleotide variants (SNVs) as an indicator of accumulated richness. SNV frequency is another marker of diversity as it enables the identification of variant frequency differences across a locus from one population to another. Other common population genetics analyses calculate Shannon's entropy or the uncertainty of sampling a specific allele which we call complexity, and the fixation index (F_{ST}) to determine how two different populations diverge. Viral demographics can be assessed by characterizing haplotypes in the population, and Tajima's D and Harpending's raggedness index can aid in identifying the impacts of viral progeny [187]. Each of these tests characterizes a unique aspect of RNA virus population genetics, and the combination of these methods enable a deeper understanding of a population through identification of population dynamics.

With the advent of "Big Data" and the exponential increase in genomic data being generated [286], we have seen the increasing need for high throughput processing and analysis methods. Numerous tools have been developed to process RNA-sequencing data in a way that enables identification of novel sequences using de novo assembly methods [287, 288], as well as

targeted processing through alignment of sequencing data to organisms of interest. Some of the most common methods of genome alignment for short-read sequence alignments generated by illumina sequencing methods include BBMap [289], Bowtie 2 [290], BWA [291], and MOSAIK [238]. Once alignment of the sequencing data is complete and contigs are generated to allow identification of genomes, it's important to determine the differences in the genetic population. Thus, a next step in bioinformatic processing is variant calling. The list of variant callers is diverse, but commonly used callers are LoFreq [243], GATK Haplotype Caller [241], V-Phaser 2 [292], SAMtools [242], and DeepVariant [293]. As there are numerous combinations of aligners and variant callers that can be used to process sequencing data, it is important to have the flexibility to choose the aligner and variant caller of choice while providing a high level of reproducibility. The time and effort often required for these processes can be greatly reduced with an automated pipeline that implements methods to process genomic data.

Accordingly, we designed a bioinformatic pipeline that was robust enough to allow the use of either Bowtie or MOSAIK as reference-based aligners and the choice of LoFreq, or Vphaser2 as the variant callers. We designed our pipeline in the Snakemake workflow environment [236] to ensure it is a reproducible and scalable framework for data processing and analysis. In addition to the processing portion of our pipeline we developed custom Python and R scripts to perform preliminary population genetics analysis on RNA virus sequencing data to facilitate high-throughput processing and analysis.

A1.2 Methods

A1.2a Workflow Manager

Workflow managers have been in use for many years, and the most common among them are Snakemake [236], Nextflow [294], Toil [295], and CWL [296]. These workflow managers were developed to streamline complex bioinformatic processing and analysis that involve numerous steps, each of which call unique software, dependencies, and environment resources. One of the key benefits of workflow managers is that they can perform processing and analysis while transparently managing individual processes and issues that arise when running a shell command. Our workflow manager, Snakemake, is a Python based language manager that we used to streamline processing and analysis (Figure A1.1.). Snakemake creates a workflow that is comprised of python scripts that make defined rules which describe how specific output files are generated from input files. Snakemake rules can be comprised of shell commands, Python code and R scripts to create output files from said input files. On top of the use of rules, Snakemake incorporates useful commands that aid in testing new workflows, debugging, generating log files, and outputting .dag files for visualization of workflows.
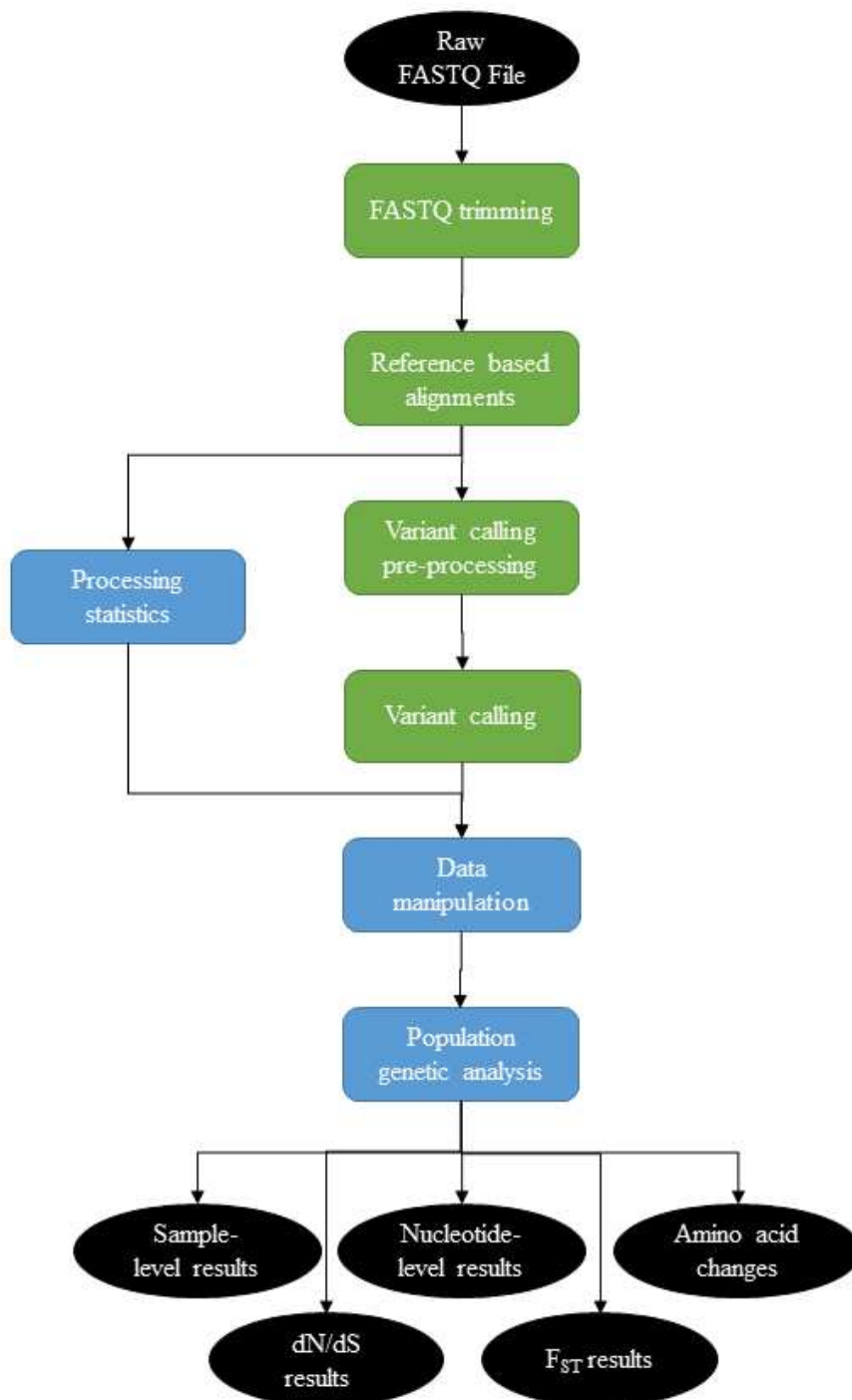
Figure A1.1. RPG Workflow Flow Chart. Generalized workflow structure for processing and analyzing one FASTQ.GZ input. Black ovals indicate input and output files, green boxes indicate bioinformatic tools, blue boxes indicate custom scripts.

A1.2b Required Inputs

The RNA-virus Population Genetics (RPG) workflow manager that we designed can be found and cloned from https://bitbucket.org/murrieta/snakemake/src/master/. The RPG workflow manager was designed to work with Illumina paired-end reads. The structure of the workflow is comprised of annotations, config, and scripts directories. The Snakefile and README.txt are contained in the working directory. Users will have to update the annotations directory to include a directory with the RNA virus species abbreviation. Chikungunya virus (CHIKV), Dengue virus 2 (DENV2), West Nile virus (WNV), and Zika virus (ZIKA) are the default directories present in our repository. The reference .fasta file to be used for alignment needs to be placed in the RNA virus species annotation directory (i.e. annotations/ZIKA/PRVABC59.fasta). A .dict and .fai file needs to be created from the reference .fasta file. Users can use Picard Tools CreateSequnceDictionary command to create the .dict file and the SamTools faidx command to generate the .fai file. A genome model .csv file needs to be imported into the annotations/species/ directory. This .csv should be comprised of three columns with the headings "Genome position", "Ref Seq", and "Virus position". Genome position is comprised of the numbered position of the genome (ie 1-10807 for Zika virus (ZIKV) PRVABC59), Ref Seq is the reference nucleotide at the corresponding genome position, and Virus position is the gene region (i.e. 5'UTR, C, prM, E, NS1, ect) at the corresponding genome position. An example model can be found here https://bitbucket.org/murrieta/snakemake/src/master/annotations/ZIKA/ZIKV_model_PRVABC 59.csv. Last, the users need to generate a Samples.txt file in the annotations/"species"/ directory that contains a list of sample base names. For example, if your .fastq file names are Sample1_S1_L001_R1_001.fastq and Sample1_S1_L001_R2_001.fastq, the Sample.txt file would be comprised of Sample1_S1_L001 to represent both reads.

After proper annotations have been added to the workflow structure, users may modify the config.yml (https://bitbucket.org/murrieta/snakemake/src/master/config/config.yml) file located in the config/ directory. In the config.yml, users update the REFERENCE, SAMPLE_LIST, and MODEL variables to contain the hard path for the above generated annotations. If using V-Phaser 2 as the variant caller, users update the VPHASER2_BASE to include the reference .fasta base name (i.e. PRVABC59.fasta would be PRBABC59). A final input is the nucleotide position that analysis should begin and stop at for the START_CODON and STOP_CODON variables, and the base name of the sample for fixation index (F_{ST}) to use a reference population for the FST_REFERENCE variable. For the F_{ST} reference we typically use the stock input virus.

The final requirement is to create a directory and a symbolic link within the directory for the raw data and workflow outputs. To do so, the mkdir command is used to create a directory in the RPG workflow working directory that is named after the virus species abbreviation (i.e. if West Nile virus, the directory would be WNV). Next, use the ln -s command to generate a symbolic link in the virus species directory that is linked to the raw data directory and named raw_data. The raw data directory linked includes Illumina paired-end sequencing FASTQ.gz files. An example of the final directory structure with essential files is depicted in figure A1.2.
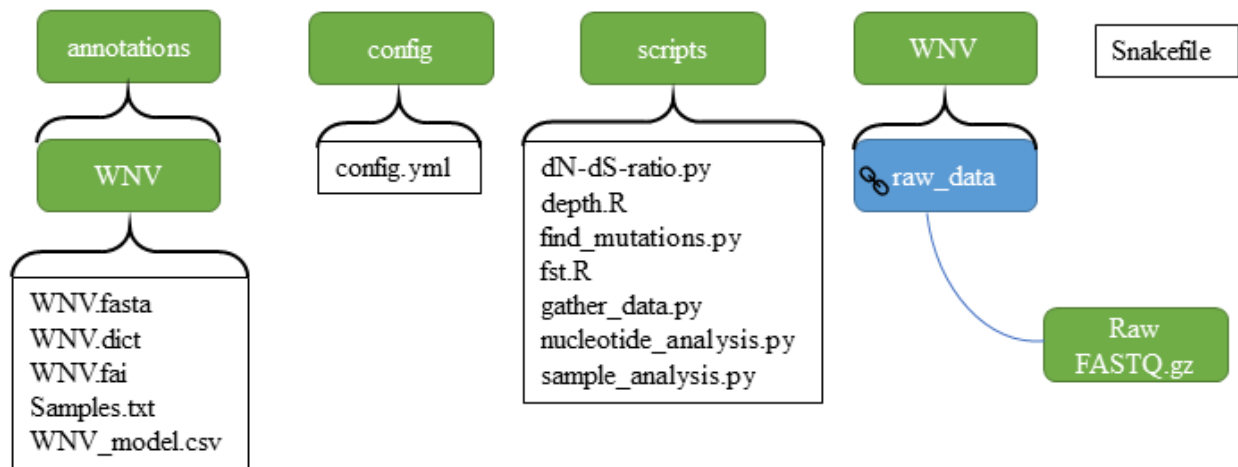
Figure A1.2. RPG Workflow Directory Structure. Example workflow directory structure for WNV processing and analysis. The top line represents the working directory, with all subdirectories and contents listed beneath. Green boxes indicate directories, blue boxes indicate symbolic links, and white boxes indicate files.

A1.2c Running the RPG Workflow

Once initial set up is complete, the RPG workflow can be executed using the following set of command: snakemake flags/"species"/"aligner"/"variant_caller"_analysis_complete, where "species", "aligner", and "variant_caller" are specified by the user. This will run the pipeline using the species directory ("species"), i.e. WNV, DENV, ZIKA, ect. The virus species directory (WNV, DENV2, ZIKA, etc.) is designated to align with the either MOSAIK 2 or Bowtie 2 ("aligner"). The variant is called using either LoFreq or V-Phaser2 ("variant_caller") and complete population genetics analysis. To perform the alignment and processing without population genetic analysis, the following command is used: snakemake flags/"species"/"aligner"/"variant_caller"_complete. If numerous sample runs are required, the processing time can be shortened by running the RPG workflow in parallel using the --cores X command, where X is equal to the number of cores available for use (i.e. snakemake --cores 22 flags/"species"/"aligner"/"variant_caller"_analysis_complete).

113

A1.2d FASTQ Trimming

The first step in the RPG workflow is to trim all Illumina adapter sequences from our paired-end reads using cutadapt [237]. To ensure a consistent quality of data for all subsequent analysis, reads are trimmed with quality score <30. The output of this step is adapter-trimmed FASTQ files with quality score ≥ 30 that are ready for alignment.

A1.2e Reference based alignment

Short-read genome sequence alignment is performed using one of two reference-based alignment tools. MOSAIK 2 [238] is a reference guided open-source alignment tool that works for a range of sequencing technologies and has been used in numerous bioinformatic pipelines including the 1000 genome project [297, 298]. When MOSAIK 2 is selected as the reference based aligner the following parameters are used: -hs 13 -mmp 0.05 -minp 0.8 -mms -9 -ms 1 -hgop 4 -gop 5 -gep 2 -m all -mfl 250 -st illumina [238]. Bowtie 2 is a fast memory-efficient short read aligner that uses a Burrows-Wheeler index to maintain a small memory footprint [299]. The index strategy that Bowtie implements allows for an ultrafast short read alignment and had been validated on data from the 1000 genomes project [299] and has been incorporated in numerous bioinformatic workflows [287, 300-302]. Bowtie indexes used for alignments are generated from the reference .fasta using the bowtie2-build command and outputs to the annotations/"species"/btindex directory. The implementation of Bowtie 2 in the RPG workflow uses the following parameters: --phred33 --rdg 5,2 -I 0 -X 700 --very-sensitive-local --score-min C,120,1 [299]. Both MOSAIK 2 and Bowtie 2 generate a .sam output that is mapped to the reference .fasta file and ready variant calling pre-processing.

A1.2f Variant Calling Pre-Processing

To prepare the reference-based alignment output for variant calling, the resulting .sam files are first sorted using Picard tools [240] SortSam command (LoFreq variant calling pipeline) or the SamTools [242] sort command (Vphaser2 variant calling pipeline) generating a sorted .bam file. Read groups are added back to the sorted .bam files using Picard AddOrReplaceReadGroups command using the following variables: AddOrReplaceReadGroups RGID=id RGLB=library RGPL=platform RGPU=machine RGSM=sample". This results in a sorted .bam file that has replaced read groups. Duplicate reads are then identified and removed using the Picard MarkDuplicates command with the following variables: MarkDuplicates CREATE_INDEX=true VALIDATION_STRINGENCY=LENIENT REMOVE_DUPLICATES=true M=output.metrics. Note that if using amplicon-based library preparation methods, the REMOVE_DUPLICATES=true can be changed to REMOVE_DUPLICATES=false to keep count of total reads mapped to the reference. The resulting sorted, read group added, deduplicated .bam file can then be used for Vphaser2 variant calling or move forward for LoFreq variant calling prep-processing.

LoFreq variant calling requires multiple preprocessing steps before single nucleotide variants can be identified. First, the sorted, read group added, deduplicated .bam file is processed using the LoFreq Viterbi realignment command to correct any mapping errors associated with the reference-based alignment. Next, the LoFreq indelqual command is performed to insert indel qualities into the realigned .bam file. Last, the .bam file is sorted by the leftmost coordinate once more using the SamTools sort command.

A1.2g Variant calling

The .bam files resulting from the variant calling pre-processing workflow are now able to be processed by either V-Phaser2 or LoFreq. V-Phaser 2 is an open-source software package that is used for SNV and length polymorphism variant (LPV) identification from read alignment .bam files. V-Phaser 2 is called in the RPG workflow using default variable and results in a tab delimited variant text file that identified SNVs and LPVs by genome position (Ref_Pos), variant nucleotide (Var), reference nucleotide (Cons), variant percent (Var_perc), strand bias significance (Strd_bias_pval), and the variant profile distribution (SNP_or_LP_Profile). LoFreq is a variant calling tool that can identify SNVs as wells as LPVs similar to V-Phaser 2. However, one of the niceties of LoFreq is that it takes into account base-call qualities and sequencing errors inherent in Next generation sequencing when making SNVs and LPVs calls. LoFreq is called using the default parameters with the addition of --call-indels command to output LPVs. The resulting output is a .vcf file that characterizes SNVs and LPVs by genome position (POS ID), reference (REF) and variant nucleotide (ALT), quality scores (QUAL), depth across samples (DP), variant frequency (AF), and strand-bias (SB).

A1.2h Processing Statistics and Data Manipulation

During the RPG workflow, processing statistics are output in the "species/statistics/"aligner"/"variant_caller" directory. One specific output used for downstream analysis is the depth of coverage output. Sequencing depth of coverage is generated using the GATK DepthOfCoverage command on the deduplicated .bam file. From the depth of coverage output we use the depth.R script (https://bitbucket.org/murrieta/snakemake/src/master/scripts/depth.R) to generate a .pdf file that

creates a xy plot that has the sequencing overage on the y-axis by the genome position on the x-axis allowing for easy visualization of alignment results. For more general statistics, such as percent reads mapped and unmapped to reference and the average coverage across the genome, the SamTools flagstat and view command are used.

As the two variant callers generate two very different outputs, a python script is incorporated into the RPG workflow so that analysis can be performed on either variant caller pipeline. To return the .VCF output in a simple table, the GATK VariantsToTable command is used with the following variables: CHROM -F POS -F FILTER -F ID -F AC -F TRANSITION -F REF -F ALT -F QUAL -GF DP -F FS -F TYPE -F AF -F AN -F SB -AMD -SMA -GF GT -GF AF. The resulting .table file or the V-Phaser 2 .var.raw.txt output is used as the input for the gather_data.py script (https://bitbucket.org/murrieta/snakemake/src/master/scripts/gather_data.py). The gather_data.py script combines the variant caller output and the .depth files generated from the GATK DepthOfCoverage command for all samples processed through the RPG workflow into one .csv file named all_samples.csv. The all_samples.csv file includes the sample name (sample), reference file name (chrom), variant position (pos), filter results (filter), variant coverage (coverage), reference nucleotide (ref), variant nucleotide (alt), quality score (qual), SNP or LPV (type), variant frequency (af), and strand-bias (SB).

A1.2i Population Genetics Analysis Scripts

Population genetics metrics are calculated using the suite of scripts found in the scripts directory (https://bitbucket.org/murrieta/snakemake/src/master/scripts/). First, the RPG workflow uses the all_samples.csv, variant caller output, and depth of coverage output as the input for the

sample_analysis.py script. Additionally, this script uses the START_CODON and STOP_CODON nucleotide position designated in the config.yml as the target locus. For each sample, sample_analysis.py calculates the average coverage across the genome (coverage(variants)), the average coverage across the target locus (coverage (CDS)), the total number of variants, single nucleotide polymorphisms (SNPs) and total length polymorphisms (LPs) across the locus (richness (lp)). Complexity is calculated across the locus for all variants (complexity (CDS)), SNPs (complexity (snp)), and LPs (complexity(lp)) using Shannon entropy (*S*) which was calculated for each intrahost population (*i*) using the iSNV (all variants, SNPs, and LPs respectively) frequency (*p*) at each nucleotide postion (*s*):

$$S_{i.s} = -(p_s(log_2 p_s) + (1 - p_s)log_2(1 - p_s)$$

The mean (S) from all sites (s) is used to determine mutant spectra complexity. Last, we estimated diversity across the locus for all variants (nucleotide_diversity (CDS)), SNPs (nucleotide_diversity (snp)), and LPs (nucleotide_diversity (lp)) by taking the sum of the SNV frequencies per locus.

Next, the nucleotide_analysis.py script functions similar to the gather_data.py script with following exceptions: it only requires the all_samples.csv as an input and uses the Shannon entropy equation above to calculate per variant complexity (entropy (S)) to include in the output. The nucleotide_analysis.csv output removes the chrom, filter, id, ac, transition, fs, and an columns that are found in the all_samples.csv output and only outputs results for the config START_CODON and STOP_CODON.

The find_mutations.py script uses the nucleotid_analysis.csv and the MODEL from the config.yml file to characterize the SNP position reference codon and amino acid in comparison to the SNP variant codon and amino. Additionally, this script summarizes if the amino acid change

118

is synonymous (S) or non-synonymous (NS) while still retaining variant position, coverage, and frequency in the mutation_analysis.csv output.

Using the mutation_analysis.csv and the config.yml START_CODON and STOP_CODON locus, the dN-dS-ratio.py script estimates the level of selection by calculating d_N, d_S, and the d_N/d_S ratio using the Jukes-Cantor formula as previously described [239]. Users should use DnaSP [245] to determine the number of non-synonymous and synonymous sites from the ancestral consensus sequence. The non-synonymous sites can be accounted for in the dN-dS-ratio.py script by using the --nSynPos variable and the synonymous sites by using the --synPos variable. The default value for nSynPos = 7822.83 and the synPos = 2446.17 based on the ZIKV PRVAB59 reference sequence. If separate runs and analysis will be performed using this script for the same reference genome, it is recommended to change the default nSynPos and synPos values on line 55 and 56 of the script to the respective reference values.

Last, the fst.R script uses the nucleotide_analysis.csv and the FST_REFERENCE variable from the config.yml file as the inputs to estimate the genetic divergence between two viral population as described previously [303]. The RPG workflow fst.R will calculate the F_{ST} for all sample populations in comparison to the FST_REFERENCE population. Coding sequence locus is hard coded for ZIKV PRVABC59 (GenBank: KU501215), WNV FtC-3699 (GenBank: KR868734), CHIKV 99659 (GenBank: KJ451624.1), and DENV 2 (GenBank: JN819407). If the locus of interest is not in the above regions update the hard-coded regions in lines 80-99 of fst.R.

A1.3 Discussion

A1.3a Selection of aligner and variant caller to use

In the RPG workflow we provide the option of using one of two aligners MOSAIK and Bowtie2. The rationale behind the options of these two aligners is that they both have slightly different strengths. In a the case of Bowtie2 it has been shown to be a very vast aligner [290] that has a high average SNV positive predictive value (PPV) of 98.69% with a SNV sensitivity of 49.19% and an average indel PPV of 45.45% [304]. In the same study, MOSAIK was shown to have slightly less SNV PPV of 98.51% and significantly lower SNV sensitivity at 35.79%. However, MOSAIK has one of the highest indel PPVs of 52.95% [304]. Further in this study, when they look at both of these aligners with the same variant caller, GATK HalotypeCaller (HC), [241] they see that when using Bowtie2 21,631 true positive (TP) SNVs and 273 false positive (FP) SNVs which is an 1.2% FP rate. MOSAIK alignment only lead to 13,528 TP and 111 FP SNVs being identified which is a 0.8% FP rate. Interesting, when Cornish et al (2015) looked at indel FP rates they saw that Bowtie2 with GATK HC they saw an indel FP rate of 58.9% in comparison to the 47.9% FP rate derived from MOSAIK with GATK HC [304]. In general, we see that Bowtie2 leads to the identification of more SNVs but at the cost of greater false positives, this however should not be an issue if you are only aiming to identify organisms present in the sequenced population at a consensus level. On the other hand, MOSAIK leads to more conservative SNV and LP calling which is beneficial when taking into account population genetics and characterizing minority variants that may be preset.

While there are numerous variant callers available to identify SNVs and LPs we provide the option of using V-Phaser 2 and LoFreq in our RPG workflow. V-Phaser2 is a variant caller

that has been used to asses population genetics in our group in the past [187, 239, 305, 306]. The updated version 2 released in 2013 lead to an increase in specificity from version 1 by increasing to 99.58% compared to the 93.84% [292]. However V-Phaser 2 has an inherently high rate of False positives. It is shown that while V-Phaser 2 has almost half of the FPs identified compared to V-Phaser, it still maintains a FP rate near 28% [292]. In addition to the high FP rate associated with V-phaser2, it appears that this tool is no longer supported as the last update was in March of 2013 (https://www.broadinstitute.org/viral-genomics/v-phaser-2-release-history). LoFreq on the other hand is actively supported with the latest version release January 2020 (https://csb5.github.io/LoFreq/blog/). LoFreq has been shown to have FP rates as low as <0.00005% in one particular data set [243]. When comparing LoFreq to other SNV calling pipelines, LoFreq is one of the better variant callers for large genome SNV calling [307]. However there have been examples that when screening for variants <10% there is an increase in FPs and a decrease in sensitivity with LoFreq [308, 309], but in general it is still considered one of the more conservative callers [310].

Taking the above into consideration, if the goal of the experiment is to characterize majority variants any combination of the aligners and variant callers should work. However, we believe that the high FP rate associated with V-Phaser 2 makes for one of the less desirable callers. In this case of characterizing consensus sequences, we believe that Bowtie2 and LoFreq make for the best combination for maximizing reads mapped and SNVs called. If we are concerned with performing minority variant analysis for population genetics work, we believe that the MOSAIK LoFreq combination which is the most conservative would be ideal.

A1.3b Robustness of Analysis Scripts

One of the major strengths of the RPG workflow is the automated analysis of processed data. The Python and R scripts are designed to not only incorporate seamlessly with the workflow, but to also be useable individually after initial processing. Likewise anyone can use the set of scripts to perform analysis on their data sets assuming they have a .vcf file for variant calls and the .depth file generated by the GATK DepthOfCoverage command. The main variables common to change for reanalysis with our scripts suit are the coding sequence start and stop position to allow for analysis of specific gene regions, variant frequency cut off to assess genetic impacts on minority and majority variants, and last is to change the F_{ST} reference population so that comparison between control groups and experimental groups can be assessed for population divergence. All of these options are available when applicable and annotation for optional commands is provided in the respective script file (https://bitbucket.org/murrieta/snakemake/src/master/scripts/).

A1.3c Conclusions

We developed a robust workflow that is works with many types of Illumina paired end data including whole genome sequencing and amplicon sequencing. Using Snakemake as a framework we have developed a reproducible pipeline that is scaleable to work with large samples sets in parallel and also to automate common population genetic analysis. The RPG workflow was developed to help increase reproducibility associated with working with big data and to facilitate faster dissemination of data generated through Next-generation sequencing.

B1.1 Terms

Term	Definition
Bottleneck of Infection	A sharp reduction in the size of a viral population when establishing infection
Complexity	Shannon's entropy or the uncertainty with sampling a certain allele
Founders Effect	The loss of genetic variation when a new population is established by a small number of variants from a larger population
Founder Population	A population that has been impacted by a bottleneck leading to a sharp decline in parental population that make it to the establishing population
F_{ST} or Divergence	Fixation index, divergence of one population compared to another population
Input	Virus stock used for experimental infection
Majority Variant	A single nucleotide variant that is found at a frequency of 0.5 or greater, also known as a consensus changing mutation
Minority Variant	A single nucleotide variant that is found at a frequency less than 0.5
Nonsynonymous Variant	An amino acid altering mutation
Nucleotide Diversity	The sum of variant frequencies across a specific locus
Positive Selection	Selective pressure that increases the frequency of single nucleotide variants that have some fitness advantage in the selective environment, ultimately leading to fixation
Purifying Selection	Selective pressure that purge single nucleotide variants that may have deleterious impacts on fitness
Richness	The number of single nucleotide variants present per sample
Selection	The ratio of nonsynonymous to synonymous SNVs per site or d_N/d_S
Selective Environment	An environment that is expressing positive or purifying selective pressures of evolution
Synonymous Variant	A silent mutation that does not alter the reference amino acid

C1.1 Supplemental Figures

C1.1a Supplemental Figure 2.1
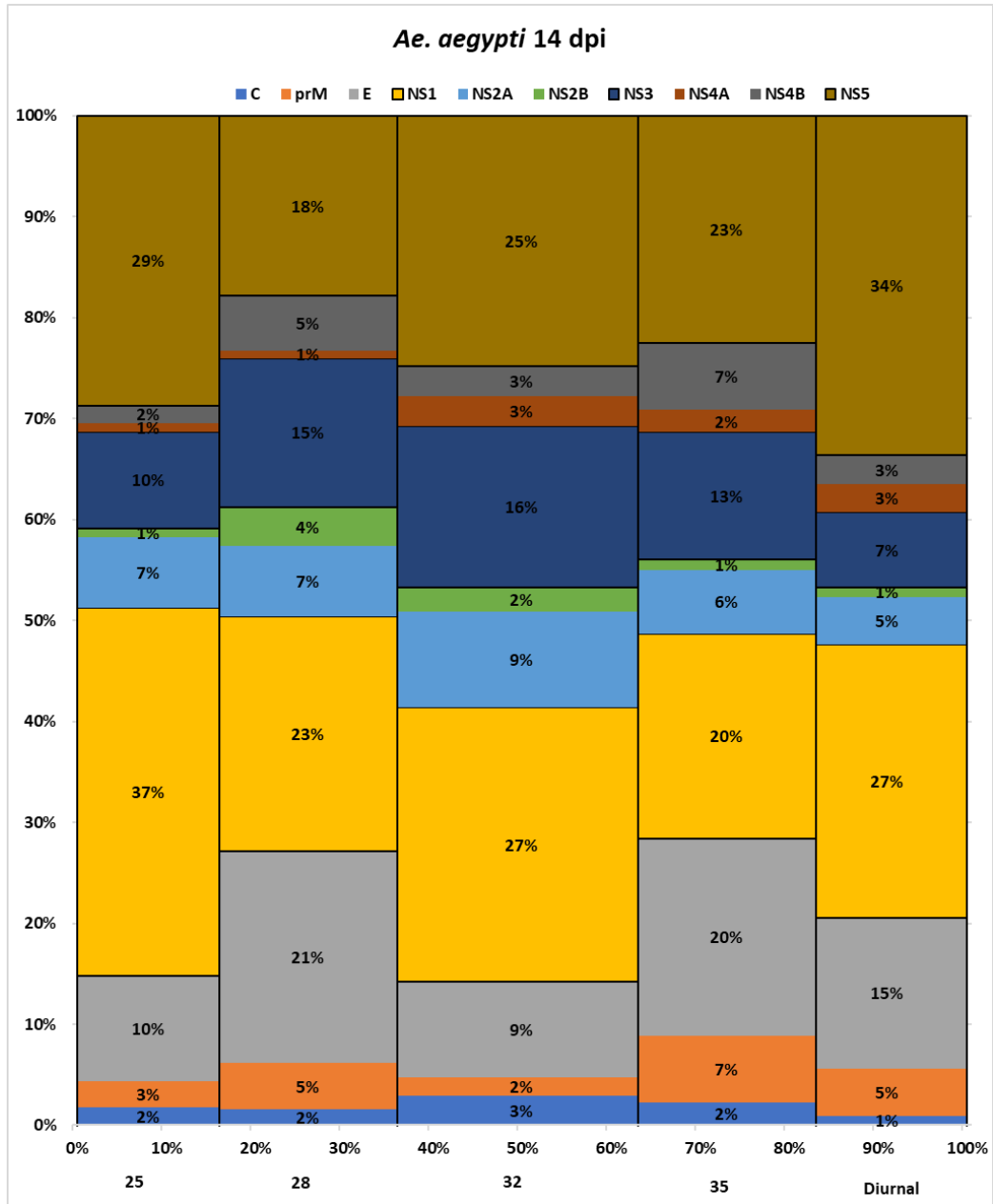
Supplemental Figure 2.1. Interhost variant summary of *Aedes aegypti* combined tissues. The percent of the total ZIKV variants (SNV's and indels) at each protein coding region (C, prM, E, NS1, NS2A, NS2B, NS3, NS4A,

NS4B, and NS5), for each EIT (25°C, 28°C, 32°C, 35°C and diurnal). The x-axis shows the percent of total SNVs for all samples and the y-axis shows the percent of SNV's within each EIT group.

C1.1b Supplemental Figure 2.2
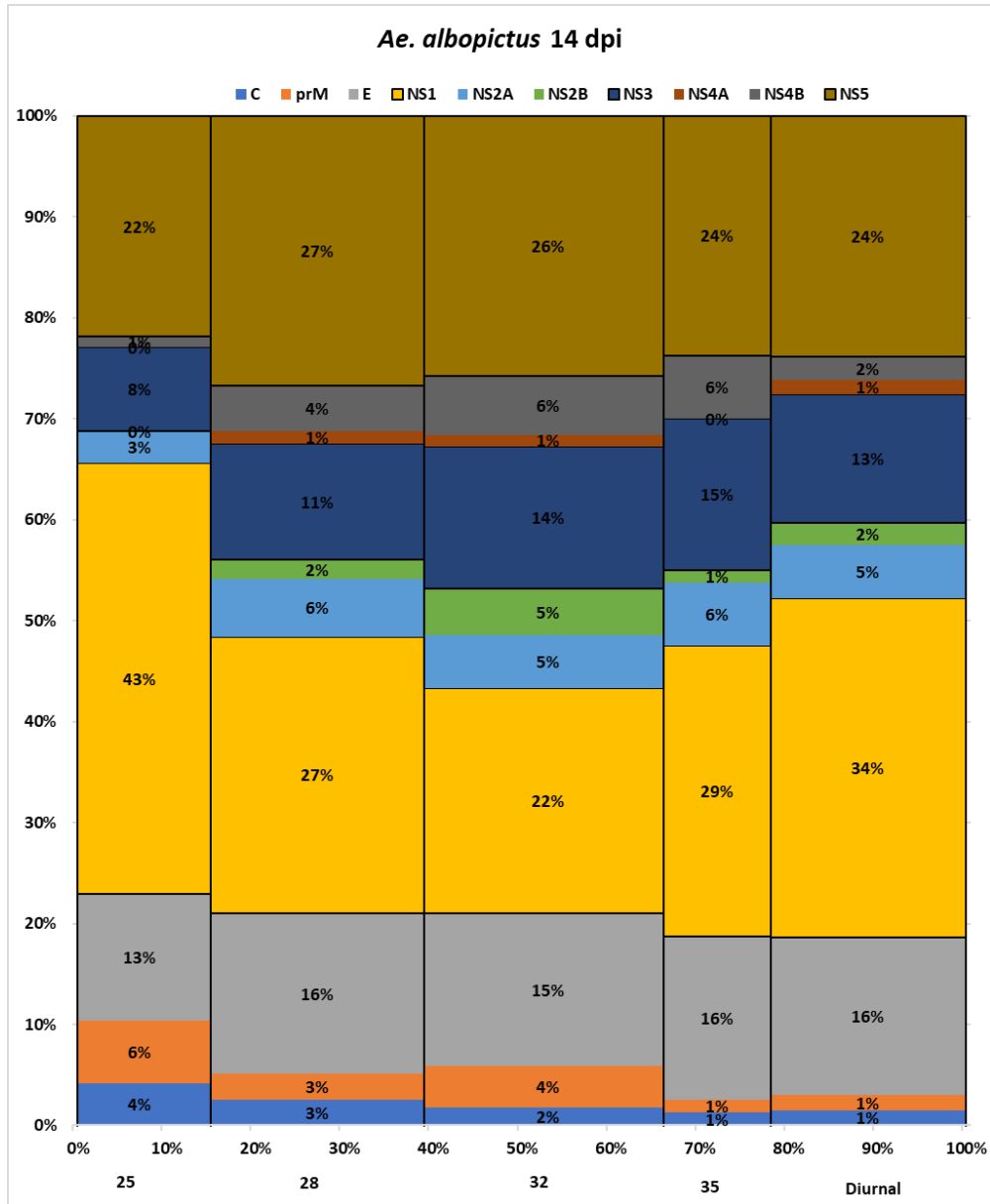
Supplemental Figure 2.2 Interhost variant summary of *Aedes albopictus* combined tissues. The percent of the total ZIKV variants (SNV's and indels) at each protein coding region (C, prM, E, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5), for each EIT (25°C, 28°C, 32°C, 35°C and diurnal). The x-axis is the percent of total SNVs for all samples and the y-axis is the percent of SNV's within each EIT group.

C1.1c Supplemental Figure 2.3
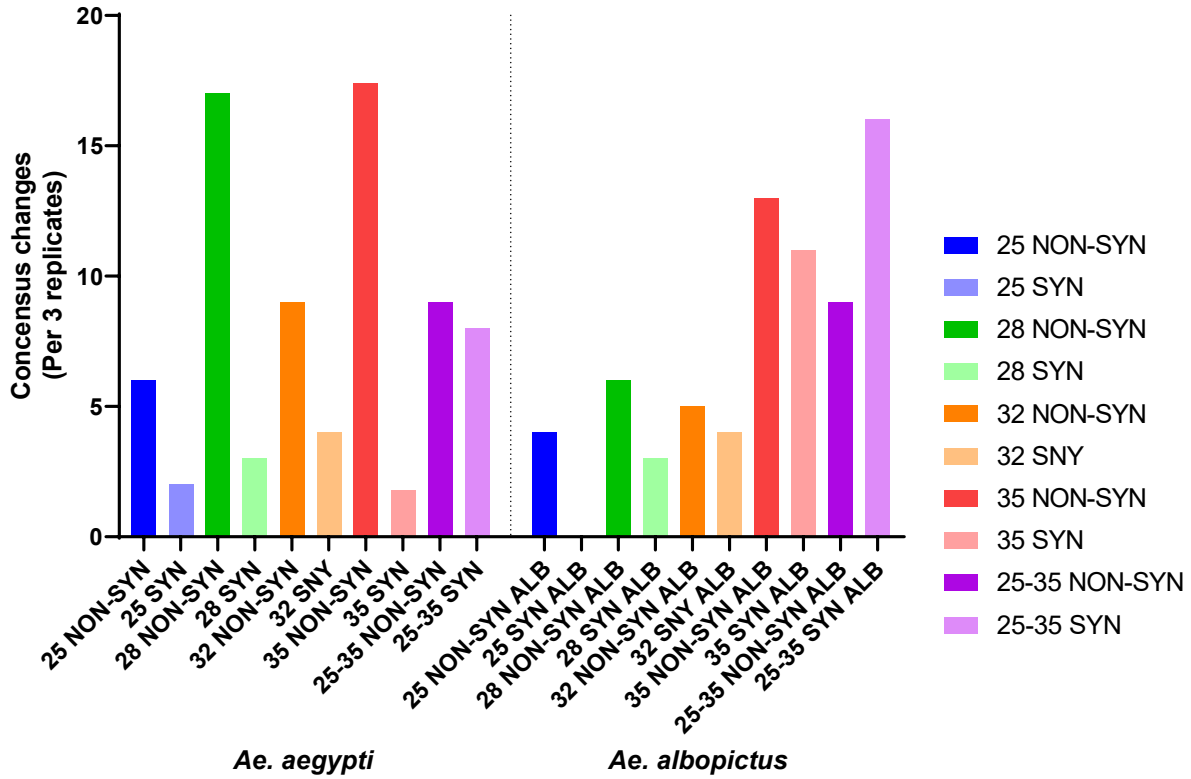
Supplemental Figure 2.3 Total consensus changes found in all tissues of three biological replicates combined.
Non-synonymous (NON-SYN) and synonymous (SYN) consensus changes are shown for each EIT (25°C, 28°C, 32°C, 35°C and diurnal) for *Aedes aegypti* and *Aedes albopictus.*

C1.1d Supplemental Figure 2.4
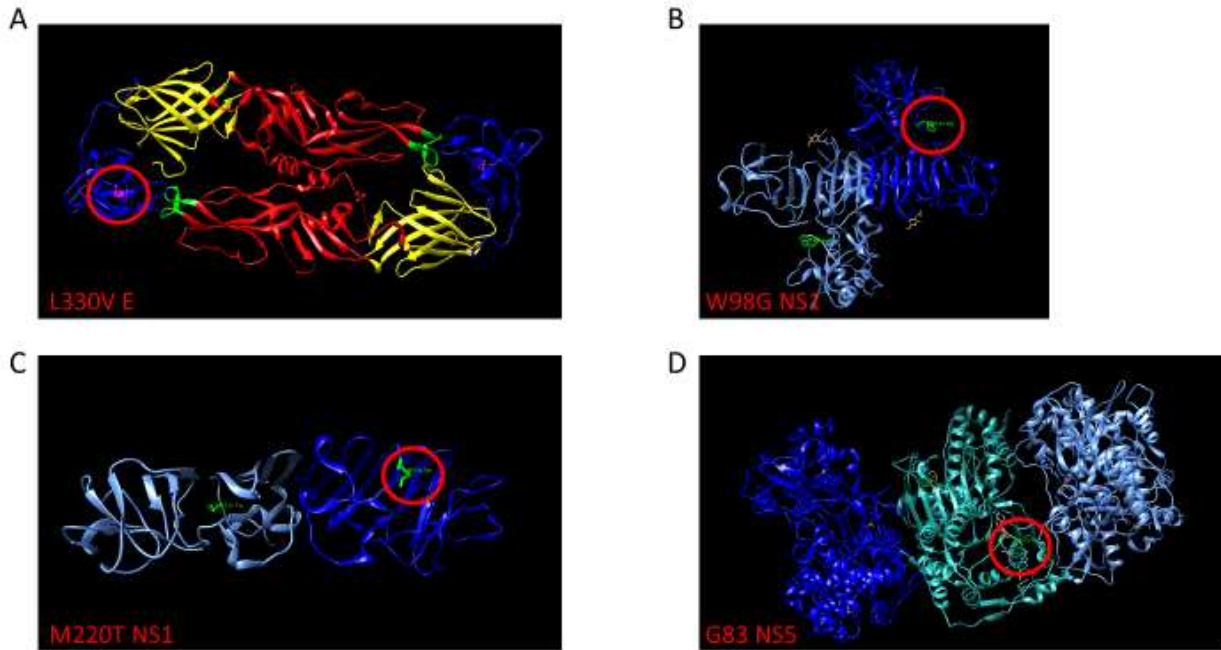
Supplemental Figure 2.4 Protein structure and location of *Aedes* consensus changes. L330V E (A), W98G NS1 (B), M220T NS1(C), and G83 NS5 (D) variants. Location of variant is circled in red.

C1.1e Supplemental Figure 2.5
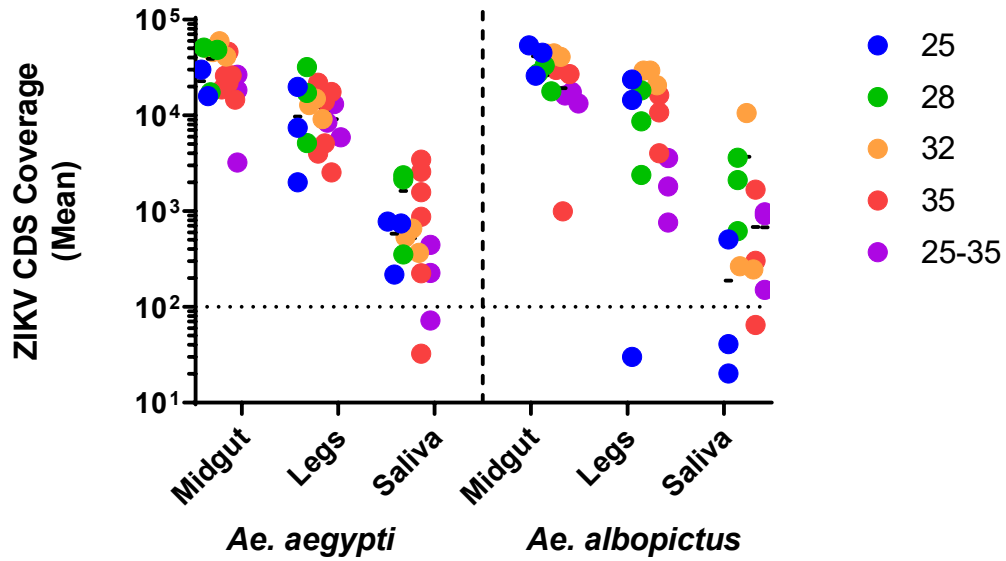
Supplemental Figure 2.5 Mean depth of coverage across the ZIKV CDS for 14 dpi *Aedes aegypti* and *Aedes albopictus*. Mean coverage of all biological samples for constant EIT (25°C, 28°C, 32°C, 35°C) and diurnal (25°C-35°C). Only samples with 100x coverage or greater were used for subsequent analysis.
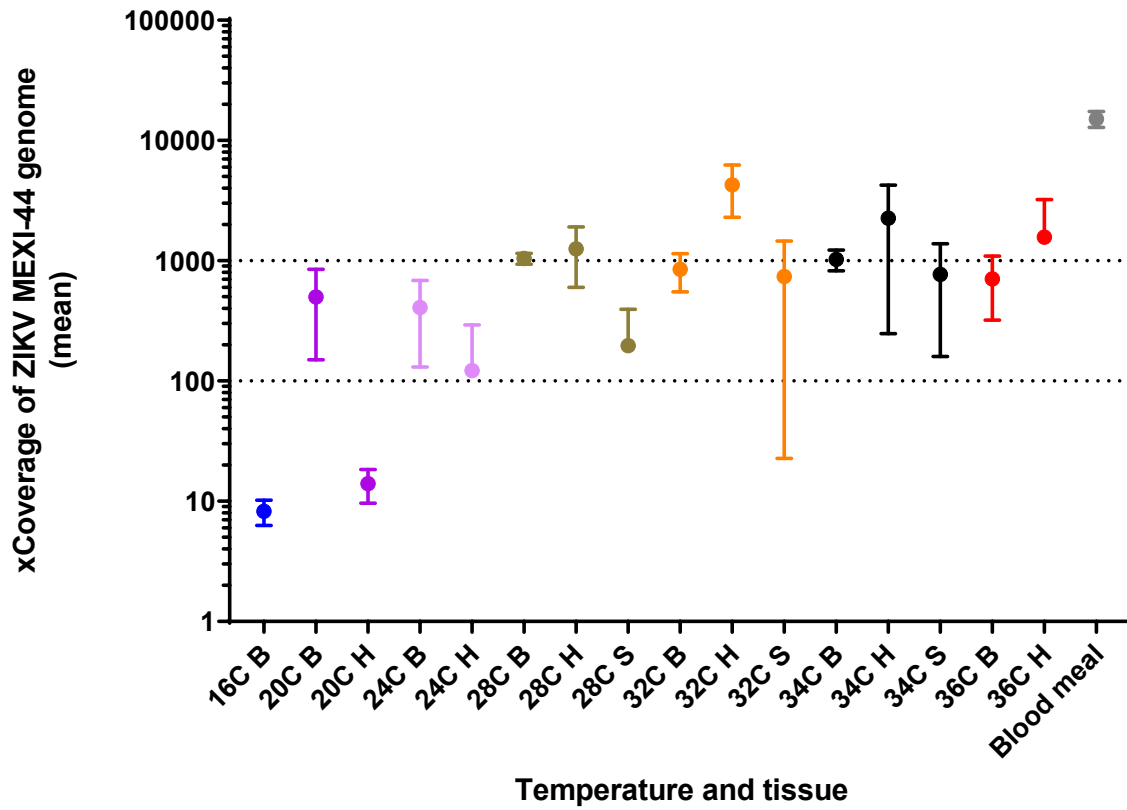
Supplemental Figure 3.1. The average depth of coverage across the ZIKV CDS for 15 day post infection
Aedes aegypti. Bodies (B), heads/legs (H), saliva (S), and input blood meal. Only samples with 100x coverage or greater across the CDS for all three replicates were used for subsequent analysis. Head/leg and saliva samples not listed had low starting GE and libraries were unable to be generated. Mean and 95% CI graphed.

PRVABC59:MEXI-44 mixtures
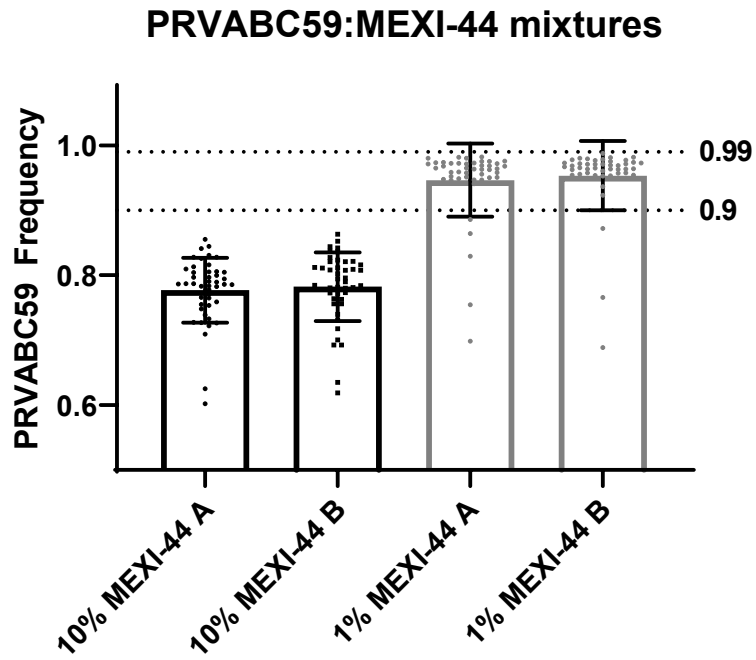
Supplemental Figure 3.2. ZIKV PRVABC59:MEX I-44 control mixtures to assess NGS variant calling. Mixed ZIKV populations were assessed in duplicate for 10% MEX I-44 mixed with 90% PRBABC59 and 1% MEX I-44 mixed with 99% PRBABC59. Mean and 95% CI graphed.

C1.2 Supplemental Tables

C1.2a Supplemental Table 2.1

Supplemental Table 2.1 ZIKV majority variants found in multiple biological samples and multiple EITs.

Species	Nt Δ	AA Δ	Gene	Virus	Titer	Input freq	%Sequence identity
Aedes spp.	T1965G	L330V	E	ZIKV-RM08	8.50E+06	0.35	100%
	T2781G	W98G	NS1	ZIKV-RM09	1.75E+06	0.30	0%
	T3148C	M220T	NS1	ZIKV-RM02	1.25E+06	0.14	0%
	C7916T	G83	NS5	ZIKV-RM05	1.50E+06	0.13	0%
Ae. aegypti	C1921T	T315I	E	ZIKV-RM01	8.50E+05	0.05	0%
Ae. albopictus	A2925G	K146E	NS1	ZIKV-RM03	1.75E+06	0.01	2%
	C3827T	I94	NS2A	ZIKV-RM07	6.50E+05	0.03	0.7%
	C9713T	F682	NS5	ZIKV-RM06	1.45E+06	0.03	0%
Viral Stock	R108-S109	pRM	ZIKV Ref	2.25E+06		NA	NA

C1.2b Supplemental Table 2.2

Supplemental Table 2.2 Sequencing summary for all biological samples. [a] 150nt paired-end reads from Illumina HiSeq 4000 (both reads represented in total) after duplicate removal. [b] The average number of nucleotides sequenced, and weather they are biological replicates [c], or technical replicates [d]. DPI, days post infection; Cov, coverage; iSNV, intrahost single nucleotide variant; nt, nucleotide; CDS, coding sequence; SNP, single nucleotide polymorphism; lp, length polymorphism.

Species	Temp	DPI	Tissue (replicate) [c]	Total no. of reads [a]	% reads ZIKV	Cov. depth [b]	iSNV sites	SNP sites	INDEL sites	nt diversity (cds) [e]	nt diversity (snp) [e]	nt diversity (lp) [e]
Ae. aegypti	25ºC	14	Leg (A)	5,474,311	2.77%	1,998.5	8	8	0	0.000125	0.000125	0.000000
			Leg (B)	5,582,946	10.10%	7,415.5	2	2	0	0.000098	0.000098	0.000000
			Leg (C)	8,224,513	18.10%	19,809.1	30	30	0	0.000304	0.000304	0.000000
			Midgut (A)	6,674,210	33.92%	30,079.3	8	8	0	0.000203	0.000203	0.000000
			Midgut (B)	7,475,366	15.94%	15,957.8	3	3	0	0.000096	0.000096	0.000000
			Midgut (C)	6,688,259	25.55%	22,495.7	29	29	0	0.000162	0.000162	0.000000
			Saliva (A)	6,414,476	0.26%	217.9	13	13	0	0.000299	0.000299	0.000000
			Saliva (B)	8,330,706	0.71%	778.8	9	8	1	0.000157	0.000145	0.000013
			Saliva (C)	8,816,495	0.65%	743.9	16	16	0	0.000244	0.000244	0.000000
		7	Leg (A)	11,264,841	0.35%	528.3	7	7	0	0.000175	0.000175	0.000000
			Leg (B)	11,343,324	0.22%	334.4	5	5	0	0.000287	0.000287	0.000000
			Leg (C)	6,192,324	0.37%	305.5	8	7	1	0.000122	0.000109	0.000013
			Midgut (A)	12,903,956	11.57%	20,007.3	45	45	0	0.000210	0.000210	0.000000
			Midgut (B)	14,776,832	8.42%	16,700.7	26	26	0	0.000185	0.000185	0.000000
			Midgut (C)	6,657,171	9.85%	8,773.7	8	8	0	0.000137	0.000137	0.000000
	28ºC	14	Leg (A)	8,631,013	4.46%	5,163.0	7	7	0	0.000300	0.000300	0.000000
			Leg (B)	12,120,334	10.62%	17,184.7	9	9	0	0.000247	0.000247	0.000000
			Leg (C)	20,419,985	11.74%	32,069.8	14	13	1	0.000415	0.000413	0.000002
			Midgut (A)	8,665,499	14.97%	17,381.7	4	4	0	0.000292	0.000292	0.000000
			Midgut (B)	18,488,371	19.57%	48,368.1	40	40	0	0.000230	0.000230	0.000000
			Midgut (C)	18,484,646	20.67%	51,201.1	17	17	0	0.000242	0.000242	0.000000
			Saliva (A)	7,761,325	2.11%	2,144.3	17	15	2	0.000363	0.000355	0.000008
			Saliva (B)	4,132,406	0.67%	353.9	14	10	4	0.000295	0.000270	0.000025
			Saliva (C)	3,279,969	5.57%	2,375.9	18	17	1	0.000490	0.000487	0.000003
		7	Leg (A)	13,916,975	1.27%	2,375.6	6	6	0	0.000264	0.000264	0.000000
			Leg (B)	21,770,118	0.74%	2,169.9	4	4	0	0.000175	0.000175	0.000000
			Leg (C)	8,703,459	2.26%	2,620.6	18	17	1	0.000204	0.000202	0.000002
			Midgut (A)	8,114,678	34.01%	36,833.6	24	24	0	0.000234	0.000234	0.000000
			Midgut (B)	11,948,388	22.63%	36,142.2	30	30	0	0.000215	0.000215	0.000000
			Midgut (C)	23,128,145	7.26%	22,468.0	40	40	0	0.000208	0.000208	0.000000
			Saliva (A)	18,365,239	0.04%	82.5	4	4	0	0.000224	0.000224	0.000000
			Saliva (B)	13,171,864	0.09%	151.4	8	7	1	0.000168	0.000159	0.000009
			Saliva (C)	6,065,250	0.43%	339.3	18	16	2	0.000296	0.000268	0.000028
	32ºC	14	Leg (A)	28,308,140	3.91%	14,922.9	13	13	0	0.000270	0.000270	0.000000
			Leg (B)	12,010,599	7.90%	12,826.5	20	20	0	0.000263	0.000263	0.000000
			Leg (C)	46,561,778	1.45%	9,184.5	13	13	0	0.000131	0.000131	0.000000
			Midgut (A)	13,733,011	24.82%	45,981.9	29	29	0	0.000230	0.000230	0.000000
			Midgut (B)	13,724,452	32.48%	59,795.7	40	39	1	0.000209	0.000208	0.000001
			Midgut (C)	17,226,801	17.75%	41,194.2	18	18	0	0.000190	0.000190	0.000000
			Saliva (A)	6,936,923	0.59%	Cov. depth b	18	15	3	0.000327	0.000315	0.000013
			Saliva (B)	15,930,283	0.32%	660.5	20	18	2	0.000304	0.000288	0.000016
			Saliva (C)	24,462,484	0.12%	367.9	15	10	5	0.000242	0.000202	0.000040
		7	Leg (A)	16,986,201	2.00%	4,636.9	13	13	0	0.000271	0.000271	0.000000
			Leg (B)	21,176,724	1.83%	5,240.5	13	13	0	0.000301	0.000301	0.000000
			Leg (C)	7,849,124	1.23%	1,321.5	10	10	0	0.000198	0.000198	0.000000
			Midgut (A)	11,795,286	21.53%	34,447.7	26	26	0	0.000197	0.000197	0.000000
			Midgut (B)	15,708,226	19.15%	40,552.5	32	32	0	0.000187	0.000187	0.000000
			Midgut (C)	9,042,542	37.33%	45,531.7	21	21	0	0.000205	0.000205	0.000000
			Saliva (A)	5,368,124	0.25%	184.0	13	13	0	0.000296	0.000296	0.000000
			Saliva (B)	25,015,532	0.02%	72.1	4	4	0	0.000093	0.000093	0.000000
			Saliva (C)	3,912,356	1.64%	861.2	18	16	2	0.000278	0.000270	0.000008

35ºC	7	Leg (A)	9,435,715	3.77%	4,735.8	8	8	0	0.000295	0.000295	0.000000
		Leg (B)	8,380,399	7.14%	7,951.9	8	8	0	0.000216	0.000216	0.000000
		Leg (C)	4,313,893	12.77%	7,330.5	5	5	0	0.000297	0.000297	0.000000
		Midgut (A)	9,846,467	15.32%	20,135.7	9	9	0	0.000221	0.000221	0.000000
		Midgut (B)	3,058,273	10.17%	4,130.0	3	3	0	0.000197	0.000197	0.000000
		Midgut (C)	5,194,522	37.25%	25,898.7	17	17	0	0.000276	0.000276	0.000000
		Saliva (A)	5,294,484	2.73%	1,887.0	27	20	7	0.000377	0.000352	0.000025
		Saliva (B)	5,372,932	0.10%	72.8	11	9	2	0.000332	0.000307	0.000025
		Saliva (C)	8,491,506	0.05%	58.1	5	5	0	0.000255	0.000255	0.000000
	14	Leg (D)	7,849,221	4.84%	5,112.3	7	7	0	0.000284	0.000284	0.000000
		Leg (E)	6,406,632	2.92%	2,535.3	6	4	2	0.000217	0.000210	0.000006
		Leg (F)	4,866,001	6.09%	3,996.4	5	5	0	0.000317	0.000317	0.000000
		Midgut (D)	5,610,653	19.13%	14,632.5	22	22	0	0.000255	0.000255	0.000000
		Midgut (E)	8,968,724	21.80%	26,162.9	18	18	0	0.000202	0.000202	0.000000
		Midgut (F)	8,725,897	15.83%	18,648.6	24	24	0	0.000230	0.000230	0.000000
		Saliva (D)	25,422,986	0.46%	1,573.8	22	19	3	0.000322	0.000313	0.000009
		Saliva (E)	5,354,453	3.61%	2,585.6	12	9	3	0.000230	0.000220	0.000010
		Saliva (F)	27,538,574	0.24%	876.2	18	16	2	0.000259	0.000255	0.000004
	14	Leg (A)	7,714,494	13.65%	14,002.2	11	11	0	0.000259	0.000259	0.000000
		Leg (B)	9,036,700	14.65%	17,643.7	16	16	0	0.000304	0.000304	0.000000
		Leg (C)	14,784,634	11.17%	21,943.3	20	20	0	0.000282	0.000282	0.000000
		Midgut (A)	9,826,665	35.37%	46,405.8	18	18	0	0.000276	0.000276	0.000000
		Midgut (B)	5,677,159	28.91%	21,841.7	25	25	0	0.000237	0.000237	0.000000
		Midgut (C)	6,805,406	28.73%	25,906.7	33	33	0	0.000218	0.000218	0.000000
		Saliva (A)	5,573,553	0.31%	224.9	9	8	1	0.000320	0.000299	0.000021
		Saliva (B)	1,949,180	0.13%	32.5	4	2	2	0.000161	0.000125	0.000036
		Saliva (C)	26,595,662	0.98%	3,452.1	22	20	2	0.000477	0.000460	0.000017
25º-35ºC	14	Leg (A)	3,580,461	12.52%	5,921.9	10	10	0	0.000228	0.000228	0.000000
		Leg (B)	11,290,631	8.73%	13,148.5	11	11	0	0.000354	0.000354	0.000000
		Leg (C)	4,181,020	15.24%	8,444.3	14	13	1	0.000418	0.000413	0.000005
		Midgut (A)	4,625,258	42.79%	26,528.7	11	11	0	0.000133	0.000133	0.000000
		Midgut (B)	6,814,360	20.23%	18,349.6	11	11	0	0.000267	0.000267	0.000000
		Midgut (C)	2,111,075	11.45%	3,223.7	18	18	0	0.000215	0.000215	0.000000
		Saliva (A)	9,255,529	0.37%	444.4	19	18	1	0.000313	0.000305	0.000007
		Saliva (B)	19,759,639	0.09%	225.7	23	21	2	0.000461	0.000447	0.000013
		Saliva (C)	7,638,697	0.07%	72.3	11	10	1	0.000409	0.000390	0.000019
	7	Leg (A)	9,380,921	0.82%	1,024.3	13	13	0	0.000244	0.000244	0.000000
		Leg (B)	7,100,825	0.84%	798.1	5	5	0	0.000216	0.000216	0.000000
		Leg (C)	10,130,930	1.08%	1,467.5	6	6	0	0.000180	0.000180	0.000000
		Midgut (A)	5,737,435	22.63%	17,399.2	11	11	0	0.000203	0.000203	0.000000
		Midgut (B)	6,886,118	24.24%	22,415.1	6	6	0	0.000252	0.000252	0.000000
		Midgut (C)	9,710,840	19.29%	24,900.6	9	9	0	0.000148	0.000148	0.000000
		Saliva (A)	2,079,544	0.23%	62.1	7	6	1	0.000210	0.000197	0.000013
		Saliva (B)	1,799,070	0.23%	55.4	4	3	1	0.000174	0.000132	0.000043
		Saliva (C)	24,535,689	0.29%	920.6	12	10	2	0.000233	0.000224	0.000009

Species	Temp	DPI	Tissue (replicate) [c]	Total no. of reads [a]	% reads ZIKV	Cov. depth [b]	iSNV sites	SNP sites	INDEL sites	nt diversity (cds) [e]	nt diversity (snp) [e]	nt diversity (lp) [e]
Ae. albopictus	25ºC	14	Leg (A)	11,894,852	9.15%	14,485.4	4	4	0	0.000199	0.000199	0.000000
			Leg (B)	15,524,135	11.52%	23,751.7	2	2	0	0.000093	0.000093	0.000000
			Leg (C)	13,888,605	0.02%	30.1	2	2	0	0.000066	0.000066	0.000000
			Midgut (A)	11,720,955	34.33%	53,757.2	31	31	0	0.000156	0.000156	0.000000
			Midgut (B)	8,497,807	39.86%	45,184.9	25	25	0	0.000161	0.000161	0.000000
			Midgut (C)	5,789,575	33.62%	26,064.8	23	23	0	0.000154	0.000154	0.000000
			Saliva (A)	6,865,073	0.05%	41.0	8	6	2	0.000312	0.000233	0.000078
			Saliva (B)	5,481,289	0.70%	505.3	13	10	3	0.000159	0.000143	0.000016
			Saliva (C)	6,609,037	0.02%	20.2	2	2	0	0.000063	0.000063	0.000000
		7	Leg (A)	150,111	33.67%	646.4	17	14	3	0.000148	0.000126	0.000022
			Leg (B)	112,456	5.98%	84.0	8	7	1	0.000196	0.000191	0.000005
			Leg (C)	528,927	4.46%	310.5	8	8	0	0.000333	0.000333	0.000000
			Midgut (A)	1,769,489	65.17%	15,445.5	27	27	0	0.000159	0.000159	0.000000
			Midgut (B)	39,344	0.20%	1.0	0	0	0	0.000000	0.000000	0.000000
			Midgut (C)	435,447	47.22%	2,749.8	17	17	0	0.000160	0.000160	0.000000
	28ºC	14	Leg (A)	6,231,408	10.47%	8,658.9	18	18	0	0.000503	0.000503	0.000000
			Leg (B)	3,431,245	5.25%	2,392.0	6	6	0	0.000113	0.000113	0.000000
			Leg (C)	9,542,620	14.56%	18,291.2	9	9	0	0.000193	0.000193	0.000000
			Midgut (A)	3,379,791	74.29%	33,146.6	32	32	0	0.000204	0.000204	0.000000
			Midgut (B)	4,869,042	42.61%	27,377.2	35	35	0	0.000172	0.000172	0.000000
			Midgut (C)	2,071,365	65.47%	17,840.4	17	17	0	0.000138	0.000138	0.000000
			Saliva (A)	6,564,404	2.47%	2,120.5	23	20	3	0.000425	0.000418	0.000007
			Saliva (B)	8,914,496	3.12%	3,625.0	10	7	3	0.000063	0.000057	0.000006
			Saliva (C)	6,149,421	0.77%	616.9	12	10	2	0.000222	0.000213	0.000009
		7	Leg (A)	16,191,698	1.33%	2,862.8	23	16	7	0.000191	0.000163	0.000028
			Leg (B)	17,084,021	0.49%	1,109.5	13	12	1	0.000318	0.000315	0.000003
			Leg (C)	20,983,358	1.94%	5,407.4	17	15	2	0.000268	0.000264	0.000004
			Midgut (A)	12,849,057	8.45%	14,438.5	24	24	0	0.000173	0.000173	0.000000
			Midgut (B)	10,874,731	14.75%	21,355.4	31	31	0	0.000173	0.000173	0.000000
			Midgut (C)	13,187,862	13.91%	24,482.0	25	25	0	0.000161	0.000161	0.000000
			Saliva (A)	3,420,869	0.36%	165.5	6	5	1	0.000098	0.000095	0.000003
			Saliva (C)	2,709,173	1.10%	397.6	11	11	0	0.000411	0.000411	0.000000
	32ºC	14	Leg (A)	9,247,726	16.92%	20,665.8	31	27	4	0.000359	0.000355	0.000004
			Leg (B)	11,993,785	18.60%	29,504.9	10	10	0	0.000142	0.000142	0.000000
			Leg (C)	12,048,169	18.27%	29,367.4	4	4	0	0.000107	0.000107	0.000000
			Midgut (A)	6,473,475	51.69%	44,445.6	29	29	0	0.000222	0.000222	0.000000
			Midgut (B)	5,940,538	52.14%	41,014.6	31	31	0	0.000183	0.000183	0.000000
			Midgut (C)	6,216,784	53.50%	44,068.5	29	29	0	0.000157	0.000157	0.000000
			Saliva (A)	2,871,904	27.71%	10,553.3	21	18	3	0.000327	0.000323	0.000004
			Saliva (B)	232,030	8.75%	266.6	7	6	1	0.000214	0.000207	0.000007
			Saliva (C)	490,773	3.84%	246.3	13	10	3	0.000238	0.000218	0.000021
		7	Leg (A)	14,621,226	7.66%	14,914.9	28	27	1	0.000335	0.000333	0.000001
			Leg (B)	6,886,174	4.29%	3,936.8	16	16	0	0.000218	0.000218	0.000000
			Leg (C)	14,675,060	6.16%	12,116.2	25	23	2	0.000210	0.000208	0.000003
			Midgut (A)	6,264,920	43.23%	36,110.7	25	25	0	0.000165	0.000165	0.000000
			Midgut (B)	6,215,531	42.14%	34,964.2	26	26	0	0.000172	0.000172	0.000000
			Midgut (C)	7,268,119	29.97%	29,220.1	22	22	0	0.000169	0.000169	0.000000
			Saliva (A)	366,460	7.00%	339.1	17	15	2	0.000417	0.000398	0.000019
			Saliva (B)	112,696,885	0.70%	10,424.0	21	20	1	0.000189	0.000187	0.000002
			Saliva (C)	509,939	4.50%	302.8	7	7	0	0.000279	0.000279	0.000000

Species	Temp	DPI	Tissue (replicate)[d]	Total no. of reads [a]	% reads ZIKV	Cov. depth [b]	iSNV sites	SNP sites	INDEL sites	nt diversity (cds)[e]	nt diversity (snp)[e]	nt diversity (lp)[e]
	35ºC	14	Leg (A)	2,397,204	12.80%	4,025.3	6	6	0	0.000394	0.000394	0.000000
			Leg (B)	4,840,193	16.76%	10,771.1	6	6	0	0.000326	0.000326	0.000000
			Leg (C)	6,400,756	19.37%	16,291.8	5	5	0	0.000306	0.000306	0.000000
			Midgut (A)	5,348,964	1.40%	998.3	4	4	0	0.000374	0.000374	0.000000
			Midgut (B)	13,060,788	17.13%	29,763.6	24	24	0	0.000263	0.000263	0.000000
			Midgut (C)	5,291,176	38.45%	27,089.1	12	12	0	0.000268	0.000268	0.000000
			Saliva (A)	4,164,661	0.12%	64.9	9	9	0	0.000380	0.000380	0.000000
			Saliva (B)	4,959,867	2.60%	1,681.6	13	10	3	0.000399	0.000388	0.000012
			Saliva (C)	3,921,628	0.60%	302.9	14	14	0	0.000384	0.000384	0.000000
		7	Leg (A)	476,718	7.09%	443.8	9	9	0	0.000583	0.000583	0.000000
			Leg (B)	379,125	2.02%	100.1	4	4	0	0.000263	0.000263	0.000000
			Leg (C)	513,028	11.56%	758.6	7	4	3	0.000261	0.000217	0.000044
			Midgut (A)	476,406	32.45%	2,051.3	13	13	0	0.000185	0.000185	0.000000
			Midgut (B)	658,975	33.73%	2,970.0	15	15	0	0.000140	0.000140	0.000000
			Midgut (C)	362,063	34.92%	1,651.1	13	9	4	0.000141	0.000118	0.000023
	25º-35ºC	14	Leg (A)	1,064,451	5.47%	763.2	6	6	0	0.000430	0.000430	0.000000
			Leg (B)	1,862,646	7.29%	1,812.8	7	7	0	0.000306	0.000306	0.000000
			Leg (C)	2,179,543	12.35%	3,579.4	6	6	0	0.000394	0.000394	0.000000
			Midgut (A)	3,027,512	40.16%	16,210.2	31	31	0	0.000193	0.000193	0.000000
			Midgut (B)	2,411,977	41.61%	13,296.7	24	24	0	0.000193	0.000193	0.000000
			Midgut (C)	2,865,841	46.21%	17,557.8	25	25	0	0.000151	0.000151	0.000000
			Saliva (A)	6,888,366	1.02%	912.0	17	12	5	0.000520	0.000450	0.000070
			Saliva (B)	6,180,800	1.21%	977.3	14	10	4	0.000403	0.000378	0.000025
			Saliva (C)	4,576,506	0.25%	151.3	11	9	2	0.000459	0.000447	0.000012
		7	Leg (A)	12,456,840	13.45%	22,196.6	11	10	1	0.000195	0.000193	0.000001
			Leg (B)	16,152,339	1.48%	3,187.4	7	6	1	0.000202	0.000200	0.000001
			Leg (C)	12,801,048	3.63%	6,199.6	25	24	1	0.000347	0.000346	0.000001
			Midgut (A)	12,771,118	6.71%	11,491.2	5	4	1	0.000290	0.000288	0.000002
			Midgut (B)	12,739,585	13.44%	22,658.2	36	36	0	0.000185	0.000185	0.000000
			Midgut (C)	11,135,540	19.66%	28,977.3	23	23	0	0.000280	0.000280	0.000000
ZIKV PRVABC	N/A	N/A	Virus Stock (A)	8,872,562	16.39%	16,324.6	18	18	0	0.000141	0.000141	0.000000
			Virus Stock (B)	14,647,924	12.62%	20,622.5	16	16	0	0.000144	0.000144	0.000000
ZIKV PRVABC	N/A	N/A	Virus Stock (C)	5,105,354	63.69%	46,669.5	18	18	0	0.000146	0.000146	0.000000
			Virus Stock (D)	4,266,471	68.95%	42,305.0	19	19	0	0.000148	0.000148	0.000000
			Virus Stock (E)	10,458,068	51.50%	75,825.3	19	19	0	0.000149	0.000149	0.000000
ZIKV PRVABC	N/A	N/A	Virus Stock (F)	10,930,085	45.95%	66,299.2	19	19	0	0.000147	0.000147	0.000000
			Virus Stock (G)	12,407,214	41.38%	67,596.0	19	19	0	0.000147	0.000147	0.000000
			Virus Stock (H)	13,260,118	38.71%	67,469.9	20	20	0	0.000149	0.000149	0.000000
NTC	N/A	N/A	NTC (A)	1,222,836	0.04%	6.9	0	0	0	N/A	N/A	N/A
			NTC (B)	1,280,640	0.02%	3.2	0	0	0	N/A	N/A	N/A
			NTC (C)	2,794,018	0.02%	7.2	0	0	0	N/A	N/A	N/A

C1.2c Supplemental Table 3.1

Supplemental Table 3.1. Multiple alignment results of MEX I-44-Extreme and MEX I-44-Moderate to 283 naturally occurring ZIKV isolates. AA, amino acid; Pos, position; Ref, reference.

Biological Clone	Protein	AA Pos	Ref AA	Alt AA	Pairwise Identity(n=283)	Accession Match
MEXI-44-Extreme	E	470	T	M	0.00%	N/A
	NS2B	45	S	T	0.00%	N/A
MEXI-44-Moderate	C	73	G	R	0.00%	N/A
	E	491	L	S	0.35%	KY785429
	NS1	103	R	T	0.00%	N/A
	NS2A	117	A	V	1.06%	KY648934, KX446951, KX520666
	NS3	117	K	R	0.00%	N/A

C1.2d Supplemental Table 3.2

Supplemental Table 3.2. Sequencing summary for all biological samples. [a] 150nt paired-end reads from Illumina HiSeq 4000 (both reads represented in total) after duplicate removal. [b] The average number of nucleotides sequenced per site, and weather they are biological replicates [c], or technical replicates [d]. DPI, days post infection; Cov, coverage; iSNV, intrahost single nucleotide variant; nt, nucleotide; CDS, coding sequence; SNP, single nucleotide polymorphism; INDEL, length polymorphism.

Species	Temp	DPI	Tissue[c]	Total no. of reads[a]	% reads ZIKV	ZIKV Cov. Depth[b]	iSNV sites	SNP sites	INDEL sites	Complexity (cds)[e]	nt diversity (snp)[f]
Ae. aegypti	16	15	Body (A)	7,207,994.00	0.01%	7.0	3	3	0	2.55E-04	8.57E-05
			Body (B)	11,466,337.00	0.00%	N/A	N/A	N/A	N/A	N/A	N/A
			Body (C)	7,164,319.00	0.01%	N/A	N/A	N/A	N/A	N/A	N/A
		18	Body (A)	9,568,657.00	0.01%	9.3	3	3	0	2.81E-04	1.18E-04
			Body (B)	10,100,306.00	0.01%	7.6	2	2	0	1.81E-04	6.79E-05
			Body (C)	12,498,962.00	0.01%	8.7	3	3	0	2.90E-04	1.50E-04
		21	Body (A)	12,188,036.00	0.01%	7.9	2	2	0	1.94E-04	9.74E-05
			Body (B)	16,299,113.00	0.01%	11.8	8	8	0	7.41E-04	3.93E-04
			Body (C)	12,889,694.00	0.01%	12.1	4	4	0	3.61E-04	1.38E-04
	20	15	Body (A)	13,548,294.00	0.51%	932.4	9	9	0	4.47E-04	1.55E-04
			Body (B)	8,504,494.00	0.33%	377.3	6	6	0	3.98E-04	1.52E-04
			Body (C)	7,690,100.00	0.24%	249.7	5	5	0	3.82E-04	1.41E-04
			Head/Legs (A)	18,933,927.00	0.01%	17.8	14	14	0	1.25E-03	5.32E-04
			Head/Legs (B)	5,516,287.00	0.02%	11.3	5	4	1	4.12E-04	1.57E-04
		24	Body (A)	11,408,571.00	0.49%	749.6	10	10	0	5.48E-04	1.56E-04
			Body (B)	7,742,944.00	0.32%	335.2	12	12	0	5.71E-04	1.55E-04
			Body (C)	5,816,889.00	0.24%	191.7	11	11	0	5.20E-04	1.42E-04
			Head/Legs (A)	12,825,250.00	0.19%	332.5	11	11	0	7.84E-04	2.40E-04
			Head/Legs (B)	6,667,439.00	0.03%	25.2	5	5	0	2.04E-04	2.53E-04
			Head/Legs (C)	6,004,809.00	0.03%	23.5	3	3	0	2.40E-04	1.39E-04
		28	Body (A)	7,938,880.00	0.89%	960.2	10	10	0	4.57E-04	1.75E-04
			Body (B)	10,829,619.00	0.80%	1165.0	10	10	0	5.45E-04	2.98E-04
			Body (C)	12,449,849.00	0.69%	1144.7	26	26	0	9.92E-04	2.20E-04
			Head/Legs (A)	9,000,412.00	1.64%	1985.9	2	2	0	2.72E-05	1.91E-04
			Head/Legs (B)	13,147,476.00	0.75%	1326.4	4	4	0	1.86E-05	3.87E-04
			Head/Legs (C)	3,566,796.00	1.31%	616.7	11	11	0	6.05E-04	2.00E-04
			Saliva (A)	2,576,422.00	0.11%	37.7	5	5	0	3.52E-04	2.84E-04
			Saliva (B)	4,513,263.00	0.24%	140.8	9	8	1	4.24E-04	3.21E-04
			Saliva (C)	5,854,277.00	0.57%	432.9	13	13	0	7.94E-04	2.68E-04
		32	Body (A)	7,322,003.00	1.24%	1212.1	14	14	0	5.61E-04	2.19E-04
			Body (B)	5,854,342.00	1.08%	840.0	9	9	0	4.99E-04	2.92E-04
			Body (C)	9,333,256.00	0.48%	598.6	17	17	0	8.62E-04	2.13E-04
			Head/Legs (A)	4,462,710.00	6.12%	3600.7	10	9	1	3.40E-04	2.68E-04
			Head/Legs (B)	17,241,287.00	2.96%	6790.1	13	12	1	3.53E-04	3.80E-04
			Head/Legs (C)	10,602,964.00	2.09%	2937.3	18	18	0	6.75E-04	3.90E-04
			Saliva (A)	2,454,750.00	0.87%	279.9	11	9	2	5.58E-04	3.08E-04
			Saliva (B)	28,874,443.00	0.44%	1632.6	20	16	4	8.54E-04	2.67E-04
			Saliva (C)	11,078,302.00	0.28%	402.9	7	6	1	3.96E-04	3.06E-04
		34	Body (A)	5,829,983.00	1.41%	1095.1	13	13	0	5.36E-04	1.80E-04
			Body (B)	10,961,609.00	0.58%	849.2	13	13	0	4.30E-04	2.05E-04
			Body (C)	12,921,216.00	0.74%	1264.6	14	14	0	7.22E-04	1.90E-04
			Head/Legs (A)	4,108,891.00	3.22%	1750.2	15	14	1	6.09E-04	2.47E-04
			Head/Legs (B)	13,446,087.00	2.60%	4670.2	10	10	0	2.10E-04	2.23E-04
			Head/Legs (C)	3,758,537.00	1.25%	621.9	12	12	0	6.29E-04	1.96E-04
			Saliva (A)	13,161,790.00	0.44%	760.9	15	12	3	7.04E-04	2.55E-04
			Saliva (B)	6,529,930.00	1.71%	1458.3	24	19	5	9.29E-04	3.60E-04
			Saliva (C)	2,920,405.00	0.49%	188.4	13	11	2	7.96E-04	2.88E-04
		36	Body (A)	3,633,290.00	1.00%	481.2	6	6	0	4.92E-04	2.05E-04
			Body (B)	7,571,414.00	0.52%	528.1	8	8	0	5.08E-04	1.87E-04
			Body (C)	14,286,146.00	0.63%	1201.9	7	7	0	4.86E-04	2.01E-04
			Head/Legs (A)	232,838.00	4.95%	152.2	5	5	0	3.50E-04	1.89E-04
			Head/Legs (B)	12,252,805.00	2.16%	3517.3	9	9	0	5.03E-04	2.41E-04
			Head/Legs (C)	15,258,448.00	0.61%	1229.3	9	8	1	3.11E-04	2.06E-04
Water	N/A	N/A	A	911,236.00	0.13%	15.0	4	4	0	3.08E-04	1.68E-04
PRVABC59 w/ 10% MEXI-44[d]			A	5,562,862.00	14.79%	10537.0	56	56	0	3.60E-03	3.46E-03
			B	16,612,942.00	12.39%	26241.6	58	58	0	3.55E-03	3.48E-03
PRVABC59 w/ 1% MEXI-44[d]			A	2,562,422.00	22.38%	7258.4	53	53	0	1.39E-03	4.18E-03
			B	15,168,448.00	17.18%	33130.3	54	54	0	1.30E-03	4.22E-03
Input Blood Meal[c]			A	16,472,192.00	6.99%	15246.9	31	31	0	8.55E-04	1.85E-04
			B	4,988,493.00	20.57%	13576.4	29	29	0	8.31E-04	1.82E-04
			C	13,175,876.00	10.41%	18223.8	35	35	0	8.86E-04	1.90E-04