Characterization of emerging and potentially emerging RNA viruses from sylvatic, rural, and urban environments in Colombia.

Katherine Dayanna Laiton Donato

Universidad de Antioquia
Corporación de Ciencias Básicas Biomédicas
Facultad de Medicina
Medellín, Colombia
2023

Characterization of emerging and potentially emerging RNA viruses from sylvatic, rural and urban environments in Colombia.

Katherine Dayanna Laiton Donato

Dissertation presented of the requirements for the degree of Doctor in basic biomedical sciences with an emphasis in virology

Advisor:
José Aldemar Usme Ciro. *BSc., MSc., Ph.D.*
Full Professor. CIST Centro de Investigación en Salud para el Trópico.
School of Medicine, Universidad Cooperativa de Colombia, Santa Marta, Colombia
Co-Advisor:
María Cristina Navas Navas. *Bac., MSc., Ph.D.*
Full Professor. Coordinator. Gastrohepatology Group. School of Medicine, University of Antioquia MedellIin, Colombia.

Tutorial committee:
Gabriel Parra Henao*, MSc., Ph.D.* Full Professor. Centro de Investigación en Salud para el Trópico. School of Medicine, Universidad Cooperativa de Colombia, Santa Marta, Colombia.
Yesid Cuesta Astroz*, MSc., PhD.* Assistant Professor, Instituto Colombiano de Medicina Tropical. Universidad CES. Medellin, Colombia.
Sandra Junglen*,* Prof. Dr. rer. nat. Charité – Universitätsmedizin Berlin, Berlin, Germany.

Universidad de Antioquia
Corporación de Ciencias Básicas Biomédicas
Facultad de Medicina
Medellín, Colombia
2023

A mi Mamá por su amor, confianza y motivación.

AKNOWLEDGMENTS

To Dr. José Usme, who has given me his absolute support in the academic and personal training process. I am truly proud to be his student during these nine years and he continues to be my best professional model.

To Dr. María Cristina Navas, who welcomed me in her research group and for allowing me the autonomy during the development of the PhD.

To the tutorial committee for guiding me in my academic process.

To Dr. Marcela Mercado, the leader who has empowered and trusted in my professional capabilities.

To the Emerging Microorganisms Genomics Group of the Instituto Nacional de Salud, for allowing me to consolidate my professional passion: Virology.

To the other valuable people, for their friendship and advice

FINANCING

This work was funded by the Ministry of Science, Technology, and Innovation in Health (Minciencias), Colombia, project 210477757671, CEMIN-06-2017 and project CEMIN-4-2020 of Instituto Nacional de Salud.

TABLE OF CONTENTS

TABLE OF FIGURES

CHAPTER 2

TABLE INDEX

BACKGROUND

CHAPTER 2

ABBREVATIONS

Adenocarcinoma human alveolar basal epithelial (A549)
Adenocarcinoma human cervix epithelial (HeLa)
Adenocarcinoma human alveolar basal epithelial (A549)
Adenocarcinoma human cervix epithelial (HeLa)
Arthropod-borne viruses (arboviruses)
Centers for Disease Control and Prevention (CDC)
Chikungunya virus (CHIKV)
Coronavirus disease (COVID-19)
Cytopathic effect (CPE)
Dengue virus (DENV)
Dulbecco's phosphate buffered saline (DPBS)
Eagle's minimal essential medium (MEM)
Eastern Equine Encephalitis (EEE),
Ebola virus (EBOV)
Eliminate Yellow Fever Epidemics (EYE)
FEL (Fixed Effects Likelihood)
Fetal bovine serum (FBS)
Genome copy equivalents (GCE)
GISAID (Global Initiative on Sharing All Influenza Data)
Guachaca virus (GUAV)
Human embryonic kidney (HEK293)
Immunoglobulin (Ig)
Insect-associated *Tymoviridae* (IAT)
Insertion–deletion mutations (INDELs)
Instituto Nacional de Salud (INS)
Internal Ribosomal Entry Site (IRES)
Japanese encephalitis virus (JEV)
Limit of detection (LoD)
Mayaro virus (MAYV)
MEME (Mixed Effects Model of Evolution)
MERS coronavirus (MERS-CoV)
Metagenomic NGS (mNSG)
Middle East respiratory syndrome (MERS)
Multiple sequence alignments (MSA)
Next Generation Sequencing (NGS)
Non-human primates (NHPs)
No Template Control (NTC)

N-terminal domain (NTD)
Nucleotide variants (NVs)
Oropuche virus (OROV)
Pan American Health Organization (PAHO)
Pro monocyte from histiocytic lymphoma (U937)
Public Health Emergency of International Concern (PHEIC)
Receptor-binding Domain (RBD)
Restriction Fragment Length Polymorphism (RFLP)
Rift Valley fever virus (RVFV)
Saint Louis encephalitis virus (SLEV)
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
SH-like approximate likelihood ratio test (SH-aLRT)
Sierra Nevada de Santa Marta (SNSM)
Transmission electron microscopy (TEM)
Untranslated region (UTR)
Usutu virus (USUV)
Variants of concern (VOC)
Variant of High Consequence (VOHC)
Variants of interest (VOI)
Venezuelan Equine Encephalitis (EEV)
Metagenomics through Next Generation Sequencing (mNSG)
West Nile virus (WNV)
World Health Organization's (WHO)
Yellow fever virus (YFV)
Zika virus (ZIKV)

2.ABSTRACT

The thesis objective allowed identify and characterize emerging and potentially emerging RNA viruses in sylvatic, rural and urban ecosystems in Colombia.

Chapter 1 presents the regional and temporal distribution of arboviruses of public health concern. The biological traits of arboviruses and their inherent capacity to infect various host and vector species, and the metagenomic next generation sequencing (mNGS) approach have proven to be reliable for discovery and monitoring of arboviruses in a variety of biological samples and different ecological scenarios.

Chapter 2 describes the characterization of a wide variety of viruses in mosquitoes from the Sierra Nevada de Santa Marta in Colombia and along a sylvatic-to-rural-to-urban gradient, applying a mNGS approach, as well as other underlying evolutionary mechanisms that may partially explain the virus dynamics in this environment.

Chapter 3 describes a new virus putatively named *Guachaca virus* (GUAV) that was identified from a *Culex pipiens/Culex quinquefasciatus* pool of mosquitoes collected in the rural area of Santa Marta, Colombia. The molecular characterization comprised viral isolation, NGS, 5'/3' RACE PCR, transmission electron microscopy, and phenotypic characterization in insect and vertebrate cells.

Chapter 4 describes the origins of SARS-CoV-2 in Colombia and the routes of introduction into the country. The lineage diversity was confirmed by phylogenetic analysis and epidemiological transmission chains.

Chapter 5 describes a novel, highly divergent lineage of SARS-CoV-2 with 21 exclusive mutations (10 nonsynonymous, 8 synonymous, and 3 located in non-coding regions). The amino acid alterations L249S and E484K found at the CTD and RBD of the Spike protein were shown to be of particular relevance.

Chapter 6 describes the first report of the B.1.621 lineage of SARS-CoV-2, subsequently ratified by WHO as the Mu variant of interest with multiple amino acid changes in the Spike protein and wide dispersion in Colombia and several other countries.

Chapter 7 presents the design and evaluation of a control RNA for real-time RT-PCR YFV detection and an approximation for YFV molecular test validation on clinical, entomological, and epizootic samples.

Finally, the chapter 8 presents the general discussion, conclusions, perspectives and originality of the thesis.

2. RESUMEN

El objetivo de la tesis permitió identificar y caracterizar virus ARN emergentes y con potencial emergente en ecosistemas sélvaticos, rurales y urbanos en Colombia.

El capítulo 1 presenta la distribución regional y temporal de los arbovirus de interés para la salud pública. Los rasgos biológicos de los arbovirus y su capacidad inherente para infectar a diversas especies de hospedadores y vectores, así como el enfoque de secuenciación metagenómica de nueva generación (mNGS), han demostrado ser fiables para el descubrimiento y seguimiento de arbovirus en una variedad de muestras biológicas y diferentes escenarios ecológicos.

El capítulo 2 describe la caracterización de una amplia variedad de virus en mosquitos de la Sierra Nevada de Santa Marta, en Colombia y a lo largo de un gradiente selvático-rural-urbano, aplicando un enfoque mNGS, así como otros mecanismos evolutivos subyacentes que pueden explicar parcialmente la dinámica de los virus en este entorno.

El capítulo 3 describe un nuevo virus, denominado *Guachaca virus* (GUAV), identificado en mosquitos *Culex pipiens/Culex quinquefasciatus* de la zona rural de Santa Marta (Colombia). La caracterización molecular comprendió aislamiento viral, NGS, PCR 5'/3' RACE, microscopía electrónica de transmisión y caracterización fenotípica en células de insectos y vertebrados.

En el capítulo 4 se describen los orígenes del SARS-CoV-2 en Colombia y las rutas de introducción en el país. La diversidad de linajes se confirmó mediante análisis filogenético y de cadenas de transmisión epidemiológicas.

El capítulo 5 describe un nuevo linaje altamente divergente de SARS-CoV-2 con 21 mutaciones exclusivas (10 no sinónimas, 8 sinónimas y 3 localizadas en regiones no codificantes). Las sustituciones aminoacídicas L249S y E484K encontradas en el CTD y el RBD de la proteína Spike demostraron ser de especial relevancia.

El capítulo 6 describe el primer informe del linaje B.1.621 del SARS-CoV-2, posteriormente denominado por la OMS como la variante *Mu* de interés con múltiples sustituciones aminoacídicas en la proteína Spike y amplia dispersión en Colombia y varios países.

El capítulo 7 presenta el diseño y evaluación de un control ARN para la detección del YFV por RT-PCR en tiempo real y la validación de la prueba molecular del YFV en muestras clínicas, entomológicas y epizoóticas.

Por último, el capítulo 8 presenta la discusión general, conclusiones, perspectivas y grado de originalidad de la tesis.

3. INTRODUCTION

3.1 Virodiversity

The can be defined as the expanding number of virus species circulating in different species of hosts and vectors and through numerous ecological interactions that resemble an ecosystem (1), it is estimated that there are at least more than 10^{31} viruses on earth; however, virodiversity is still unknown in most ecosystems because they are unexplored (2).It has been shown that there are no exclusive viral populations among the different ecosystems, but instead are in constant change (3),revealing the adaptation of viruses among species.

Viruses can contribute significantly to the emergence and development of outbreaks, epidemics, and pandemics. An outbreak is the unexpected development of disease cases in a specific population with restricted time and area. An epidemic occurs when there is an increase in the number of cases of a disease in a larger population than expected. A pandemic is a worldwide epidemic that spreads across many geographic areas (4).

Highly infectious and efficiently transmitted, such as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) the etiologic agent of Coronavirus disease (COVID-19), have a high potential for causing epidemics or pandemics. Indeed, in late 2019, this virus emerged in Wuhan, China, and soon spread to other nations, resulting in a pandemic infecting more than 600 millions of people globally (5). Viruses are also the etiologic agents of outbreaks and epidemics of vector-borne illnesses transmitted by insects such as mosquitoes or ticks. Viruses, such as *West Nile* and *virus Zika virus*, are transmitted by the bite of an infected vector and can create epidemics as the described in the Americas (6).

Some of the major viruses that are considered a public health threat include:

1. *MERS-CoV*: It is a coronavirus that causes Middle East respiratory syndrome (MERS) and is transmitted mainly through close contact with infected animals and people (7).
2. *SARS-CoV-2*: It is the coronavirus responsible for COVID-19 pandemic that began in 2019 and continues to affect the entire world. It is transmitted mainly through close person-to-person contact and aerosols (8).
3. *Ebola virus* (EBOV): A highly contagious virus that causes hemorrhagic fever and can spread rapidly through contact with infected body fluids. Ebola outbreaks have occurred in West and Central Africa, causing thousands of deaths (9).
4. *Zika virus* (ZIKV): It is a virus transmitted primarily by mosquitoes and by vertical transmission and can cause serious complications during pregnancy, including microcephaly in newborns (10).
5. *West Nile virus* (WNV): It is a virus transmitted by mosquitoes and can cause West Nile fever, which can be fatal in severe cases (11).
6. *Dengue virus* (DENV): It is a virus transmitted by mosquitoes and can cause dengue fever, which can be severe and potentially fatal.

Humans have faced different epidemics throughout their history, among the most important of which are those caused by highly pathogenic viruses (1). In this regard, the emergence of viruses with pandemic potential is unpredictable and additionally more than 60% of the infectious diseases identified have been of zoonotic origin. Thus, among the emerging pathogens discovered during 1980-2005, viruses comprised 67% and almost 85% of them correspond to RNA viruses (12).

Since 1960, the development of methodologies, including viral isolation and serology, as well as the implementation of viral ecology studies in different unexplored geographical areas, allowed the characterization of new arthropod-borne viruses. The discovery of these potentially emerging viral agents with capacity to infect and develop disease in humans or animals led to the creation of the Centers for Disease Control and Prevention (CDC) International Catalog of Arboviruses

https://wwwn.cdc.gov/Arbocat/Default.aspx. Since the 1980s, the number of arboviruses has increased markedly, coinciding with the emergence of the comprehensive program supported by The Rockefeller Foundation to isolate viruses from humans, animals, and arthropods in different continents and subcontinents (13), 44.2% of which were identified in the Americas (14).

Currently, the number of discoveries has notoriously increased because there are different approaches to discover and characterize circulating viruses in different ecosystems at genetic level, through metagenomics, protein modeling, structural and functional biology, and genetic modification including the ecological approach, pathogenesis and epidemiological dynamics (15). Although it is still evident from sequence databases that virome in arthropods relative to mammals has been little explored. Another significant finding is that between 48% and 80% of viral sequences characterized from arthropods do not show homology with sequences deposited in GenBank databases (16) suggesting the possibility of a broad outlook for the search for new viruses with emerging potential.

The characterization of an ecosystems virodiversity is a challenge due to the imperative inclusion of autochthonous species and the limitation of temporal and spatial variations of the sampling and subsequently perform viral identification including viral isolation, sequencing, and computational analysis.

3.2 Emerging diseases of viral origin

Emerging infectious diseases (EIDs), and emerging viruses, are a key threat to global public health, livestock health, wildlife conservation and ecosystem functioning. Some EIDs threaten public health through pandemics with large-scale mortality (e.g., HIV/AIDS). Others cause smaller outbreaks with high fatality rates or lack effective therapies and vaccines (e.g., Ebola virus, rabies, multidrug resistant TB). As a group, EIDs and re-emerging diseases cause millions of deaths each year, and some single outbreak events (e.g., SARS) have cost the global economy tens

of billions of dollars. The World Economic Forum considers EIDs as ‘major’ risks, comprising significant likelihood of occurrence and significant economic threat over the next 10 years, comparable in scale to unsustainable population growth.
There are two types of viral infections in vertebrates, the first one when an arthropod vector is required for transmission and the second one, arthropod-independent infections in which viruses circulate exclusively in vertebrate hosts (34). Due to the high virodiversity of arthropod and vertebrate viral infections, it is very likely that viruses can cross the species barrier and cause emerging diseases in humans and animals (35,36).

When we talk about emerging diseases of viral origin, we have to think about deforestation of tropical forests that is caused by increasing urban infrastructure, crops, agriculture, livestock, mining and other human-induced activities that represent a rise in the contact between human and natural land use and land cover. types, where the infected vectors of arbovirus inhabit. The emergence of novel viral diseases is driven by socioeconomic, demographic, and environmental changes. These include land use changes such as deforestation, agricultural expansion, and habitat degradation. However, the links between land use change and disease emergence are poorly understood and probably complex.

Consequently, the emergence of viruses in human populations is unpredictable and therefore interdisciplinary approaches are needed to identify and characterize viruses with emerging potential. The best approach is exploring different ecosystems for discovery viruses that are close to human populations because viruses do not remain confined for prolonged periods of time as they have transmission cycles that are susceptible to host change (37). As well as implementing approaches to understand evolutionary processes, viral origin, viral ecology, pathogenesis, and surveillance aimed at predicting outbreaks (38) and describing in real time the viral emergence towards the human population.

Flaviviruses and alphaviruses of public health importance include *Dengue virus* (DENV), *Zika virus* (ZIKV), *Chikungunya virus* (CHIKV), *West Nile virus* (WNV), *Venezuelan Equine Encephalitis Virus* (VEE) (39). Recently the Americas faced two major epidemics, the first one by CHIKV caused about 1.1 million cases in one year and was associated with chronic joint disease (40); and them in less than two years the epidemic by ZIKV was established in the region causing only in Brazil 4000 suspected cases of microcephaly including fetuses and newborns (41). Both epidemics had invaluable repercussions at the economic and social level in the region.

In December 2019 a new coronavirus emerged, causing an unprecedented pandemic in humans. Apparently, the first cases of severe acute respiratory illness were identified in Wuhan, China, in people attending an animal market. One month later, on January 30, 2020, the situation was declared an international health emergency by the World Health Organization's (WHO) International Health Regulations Emergency Committee (2005). At that date there had been 7711 confirmed cases in Republic of China, but also reported in about 81 countries on five continents; therefore, on March 11, 2020, the pandemic was declared (42).

This recent evidence confirms the pandemic potential of RNA viruses and that is why at this moment other viruses that could be of epidemic character worry the health entities, only among the arboviruses are: *Mayaro virus* (MAYV), *Oropuche virus* (OROV) and *Usutu virus* (USUV) (43,44). Therefore, it is necessary to carry out research on viruses of jungle circulation, whose changes in transmission dynamics could directly affect human health and to consider that there is a significant number of viruses belonging to other families also transmitted by arthropods and whose genetic and phenotypic characteristics have not been described so far and whose pandemic potential is unpredictable (45).

The establishment of emerging and re-emerging viruses in new areas increases the risk of disease in Colombia. To date, it is crucial to perform routine surveillance of

circulation of OROV, ZIKV, DENV, YFV, CHIKV (46,47) and to continue with the real-time genomic characterization of SARS-CoV-2 whose epidemiological dynamics are still uncertain (48).

3.3 Arthropod-borne RNA viruses

Arthropod-borne viruses (arboviruses) include RNA viruses, which have played an important role in evolution and biological efficacy because their high mutation rates as a product of replication mediated by RNA-dependent RNA polymerase lacking proofreading activity, large population sizes and short generation times (17). Additionally, RNA viruses infecting arthropods and vertebrate hosts are subject to the process of viral population adaptation in each host, which represents a major challenge and trade-offs in viral evolution (18).

RNA viruses represent 70% of viruses that cause diseases in humans, which has been remarkable during the last three decades (19); among the factors that contribute of their role as pathogenic agents of high impact on human health are: viral adaptation to new vectors, intervention in ecosystems that involve the interaction of reservoirs and vectors with domestic animals and humans, as well as international trade allowing vectors transportation over long distances (20). Considering that mosquitoes (Diptera: Culicidae) feed on the blood of various hosts including mammals and birds that are the main viral reservoirs (21),could constitute the best alternative to access virodiversity. In this case, sequencing viral nucleic acids purified directly from arthropods is the simplest way to understand the virodiversity of an ecosystem (22).

3.4 Ecology of arbovirus transmission

There are major factors that contribute to viruses emergence among them when the arboviruses find a favorable environment infecting a susceptible host completely asymptomatic and the virus maintains a prolonged infection (23). In general, high densities of jungle mosquitoes with ability to fly long distances or the mosquitoes

inhabiting jungle areas with competence to invade urban and peri-urban regions and therefore to infect human populations with no history of previous challenges by the virus or other viruses closely related serocomplexes, allowing the transmission cycle as is the case of Aedes albopictus (24).

There is a huge panorama of arboviruses that have caused enzootics and epizootics in different regions of the Americas; however, the knowledge of their ecological habitat is partial, and they have the potential of adaptation and condition necessary for emergence. Among these different Arboviruses, Venezuelan Equine Encephalitis (EEV), *Eastern Equine Encephalitis* (EEE), *West Nile virus* (WNV), *Mayaro virus* (MAYV), *Usutu virus* (USUV), *Oropuche virus* (OROV) among others (25,26).
The emergence of arboviruses is of great concern due to their wide distribution of vector in urban, peri-urban, and rural areas; moreover, with an exclusively urban cycle which may cross the barrier and enter the jungle zone. For an arbovirus to establish a sylvatic transmission cycle, it must have the ability to infect sylvatic vectors with the capacity to transmit to vertebrate hosts and generate high viral loads in these hosts because they are viral amplifiers (27), or cryptic transmission between them vertebrate and birds but this hypothesis is still controversial (28).

Furthermore, arthropods are naturally infected by a wide range of viruses, a situation that allows us to inquire about viral diversity and evolution of arboviruses (29). For the mentioned reasons, there is a growing interest in investigating RNA viruses whose capacity for emergence is very high and which could affect human populations at any time. Furthermore, this type of study indicate the real impact of arboviruses on public health, especially in tropical countries where the dynamics of transmission from jungle cycles to urban cycles or vice versa can change abruptly (30).

The Sierra Nevada de Santa Marta is one of the most visited tourist destinations in Colombia; it is recognized as the largest coastal mountain in the world. Much of its extension is represented by tropical rainforests, with high biodiversity. Different

attractions such as the presence of ancestral indigenous communities, high biodiversity, and beaches, are the reasons for tourists visiting this area of the country to the relatively close jungle, rural and urban enviroments. In this sense, the ecology of the ecosystem directly influences the health of the indigenous communities, the inhabitants of the different zones, and the travelers if the context of the region and the increasing discovery of new viruses in this type of environment are taken into account (31,32). Additionally, the north of the country has proven to be an area of intense viral circulation. In the department of La Guajira and in the area of the Sierra Nevada de Santa Marta (Guachaca) in 1995, there was an outbreak of encephalitis in which 14,156 cases of people with symptoms compatible with Venezuelan equine encephalitis that had historically circulated in this region were reported, of which 1258 required hospitalization and 26 died (33).

By this reason, we need more studies aimed at investigating the virodiversity and ecology of arboviruses that allows to determine the risk of transmission of the diseases caused by these viruses and provide recommendations to health authorities for prevention, surveillance, and possible control of arbovirus in the area.

3.5 Genetic, variability phylogenetic and evolutionary analysis of SARS-CoV-2

Coronaviruses are a large family of RNA viruses and most circulate among animals such as pigs, camels, bats and cats. However, in the last two decades, three new animal coronaviruses have emerged that have caused severe and widespread disease and death in the human population; they are classified into multiple genera, including Alpha, Beta, Gamma, and Delta coronaviruses. Notably, CoVs appear to be able to adapt to new hosts and changing environments; this may be related to their ability of mutation and recombination, contributing to new viruses with variable human pathogenicity. About 40 CoVs have been identified in nature, of which 7 (including the novel coronavirus) cause disease in humans. Four of these CoVs (HCoV-229E, HCoV-OC43, HCoV-NL63 and HCoV-HKLU1) are globally distributed and generally cause acute respiratory disease (49–51).

MERS-CoV and SARS-CoV cause much more severe respiratory infections in humans than other coronaviruses. In late 2002, SARS-CoV was identified as the cause of an outbreak of severe acute respiratory syndrome (SARS) in China, documenting that 9% of patients develop fatal disease. In 2012, Middle East Respiratory Syndrome (MERS) caused by the MERS coronavirus (MERS-CoV) emerged. Transmitted from camels, the animal reservoir and still continues to cause outbreaks in the region (52,53).

SARS-CoV-2 is an RNA virus with an estimated substitution rate of 0.8-1.1 × 10-3 substitutions/site/year, which means that it evolves rapidly as it is transmitted. Through the genome sequence analysis, it is possible to explore the divergence between variants and natural selection pressures with evolutionary advantage and, therefore, the subsequent emerging variants of SARS-CoV-2 in different hosts to better understand the impact of the viral population on evolution, physiopathology, and transmission (54). Several international health authorities have classified new emerging variants as variants of concern (VOC) and variants of interest (VOI), 1. A VOI is characterized by a set of substitutions of the Spike protein related to higher transmissibility, evasion of neutralizing antibodies after vaccination or natural infection, and more severe clinical disease. A VOC also relates disease severity to negative impacts on diagnostics, treatments or vaccines developed (55).

Colombia has suffered the negative impact of the COVID-19 pandemic with an enormous contribution to global epidemiology represented in at least 6.370.000 cases and 142.000 deaths during first three years (56). During the pandemic, national health officials early recognized the importance of the genomic surveillance and basic research of SARS-CoV-2 using phylogenomic and evolutionary forces that result in the emergence of SARS-COV-2 variants in the country, as well as the lineage replacement and dynamics after new introductions (57–59).

The analyses performed have allowed characterizing the dynamics of SARS-CoV-2 variants during the last three years, highlighting the circulation of The P.1 (Gamma) variant circulating in Colombia since January 2020 from the entry through Tabatinga, Brazil, and dispersing throughout the country (60).The B.1.621 lineage named *Mu* variant was detected in January 2021 and was classified as a variant of interest (VOI), was dominant in Colombia and dispersed to 43 countries worldwide (59,61); subsequently Delta was established as the predominant variant today in July 2021 and was displaced by the arrival of omicron in December 2021. During 2022 the BA.1, BA.2, BA.4, BA.5 and BQ.1 sub lineages were dominant in Colombia. At the beginning of 2023, the establishment of the sublineage BQ. 1.1 sublineage and the recombinant XBB lineage (62).

3.6 Molecular detection and metagenomics methods

There is a limitation during the implementation and development of virological surveillance because there are restrictions during the characterization of viral agents that are generally emerging and of which the information of their morphological, genetic and pathogenesis characteristics is unknown. In this case, genetic monitoring is a strategy of impact in the prediction or generation of early warnings of possible outbreaks or epidemics according to epidemiological dynamics.

To implement the monitoring of viral agents of unknown circulation, the first step is to carry out viral identification and characterization, which consists of implementing routine monitoring of viruses through the design and development of RT-PCR and qRT-PCR techniques in real time that allow determining the incidence of viral diseases of public health interest during the acute phase of infection, demonstrating that molecular detection methods are useful to quickly discriminate acute infections caused by different viral families and could be very useful, especially in countries located in tropical areas (63).

At a general level, multiple RT-PCR is a technique with high sensitivity in the search for diverse viral agents because it allows the identification of families, genera, or species, it is also a technique with less possibility of contamination with respect to a conventional PCR technique, its sensitivity also increases with the use of specific probes and automated equipment increasingly more affordable at a clinical level (64).

However, it is essential to control the reverse transcription step and for this purpose, the *in vitro* runoff transcription technique is applied, the transcription is carried out from a DNA plasmid that has a bicatenary base promoter of 19-23 nucleotides which is located downstream of the region corresponding to the viral sequence of interest. The plasmid is mixed with RNA polymerase, rNTPs and transcription buffer and the RNA polymerase binds to the double-stranded DNA promoter and separates the two DNAs and uses the 3'-5' sense strand as a template to synthesize the complementary 5'-3' strand of DNA (65). Thus, this technique has been designed during the design of RT-PCR assays because the use of Internal control is paramount in the detection of viral agents of public health concern (66,67). Although nucleic acid detection techniques are fast and sensitive in viral identification, their design requires genomic sequences as a requirement for the design of oligonucleotides and evaluation of genetic variability.

During the last two decades the exponential emergence of massive nucleic acid sequencing methods called Next Generation Sequencing (NGS) has made it possible to obtain reads from millions of fragments of one or several genomes, which is why NGS and metagenomics (mNGS) techniques for the characterization of new viruses are fundamental. Metagenomics consists of determining the taxonomic composition of a community of microorganisms using NSG, which allows obtaining representative reads of nucleic acid sequences and through subsequent processing with bioinformatics tools allows determining the identity of each species (68).

The main advantage of metagenomic analysis is that it can be performed on any type of biological sample (16) to identify new viruses. Furthermore, through metagenomic studies it is possible to obtain precise information on the virome of a given ecosystem, allowing the simultaneous implementation of studies of viral evolution, viral ecology and transmission dynamics of potential viral agents of interest in public health (16).

The use of robust bioinformatics tools allows the assembly of individual reads in order to achieve the assembly of consensus sequences or genomes present in the mixture of nucleic acids being analyzed. These new methods has improved over the years for obtain reads and even whole genome reads, they also allow a higher depth that refers to the number of nucleotide fragment reads corresponding to the same region, giving the technology the best resolution during species identification (69). Additionally, the development of computational analysis strategies has allowed establishment protocols for the identification of new species (70) tools that allow facing the challenges that metagenomic analysis implies and at the same time, data storage alternatives in the cloud and permanent updating of databases are being implemented (71,72).

NGS is an innovative technology available to many research groups worldwide to discover and determine the emergence of circulating viruses in any ecosystem. In the lessons learned after the SARS-CoV-2 pandemic, a number of emerging disease prevention programs worldwide are booming NGS methodologies (73) and their approach is multidisciplinary and includes mathematical predictions of epidemic spread, vector control and health prevention. Therefore, monitoring and predicting the ecology, biological characteristics, and emerging potential of viruses through NGS tools would allow the establishment of prevention measures to mitigate the appearance of enzootic diseases up to human outbreaks, being a powerful approach in emerging virus programs and public health contingency measures.

4. BACKGROUND

Different studies have been developed using a metagenomic approach in the search for new viral agents. Most of them are carried out by sampling in ecosystems and natural geographical areas. In Colombia, studies are scarce and the main representative ecosystems of the country have not been characterized and virus diversity or virodiversity is unknown.

Among the main studies to characterize virodiversity in sylvatic, rural, and urban areas, those that capture arthropod vectors stand out worldwide because these species feed and interact with different host species. In the last decade virodiversity studies are described in Table 1.

Table 1. Discovered specific insect viruses and arboviruses in the last decade worldwide.

FAMILY OR GENUS	VIRUS	ARTHROPOD	GEOGRAPHIC LOCATION	RESEARCH
Flaviviridae	*Ilomantsi virus* *Lammi virus*	*Aedes sp.*	Finland	(74)
Flaviviridae	*Novel Flavivirus*	*Culex sp.*	Iquitos, Peru	(75)
Flaviviridae	*Ecuador Paraiso Escondido Virus*	*Lutzomyia abonnenci*	Ecuador	(76)
Flaviviridae	*La Tina virus* *Long Pine Key virus*	*Aedes scapularis* *Anopheles crucians*	Lima, Peru Florida	(77)

	Kampung Karu virus	*Anopheles tesselatus*	Sarawak, Malaysia	
Flaviviridae	*Pamunjeom virus*	*Aedes vexans* and *Aedes esoensis*	Republic of Korea	(78)
Flaviviridae	*Culiseta flavivirus*	*Culiseta melanura*	United States	(79)
Flaviviridae	*Nhumirim virus*	Culex chidesteri	Brazil	(80)
Flaviviridae	*Palm creek virus*	*Coquillettidia xanthogaster*	Australia	(81)
Flaviviridae	*Spanish Ochlerotatus virus*	*Ochlerotatus caspius*	Spain	(82)
Flaviviridae	*Guapiaçu virus*	*Aedes* and *Aedes scapularis*	Brazil	(83)
Flaviviridae, Rhabdoviridae	*Cuacua virus*	*Culex sp. Mansonia sp.*	Mozambique	(84)
Rhabdoviridae	*Moussa virus*	*Culex decens*	Cote d'Ivoire	(85)
Rhabdoviridae	*Coot Bay virus* *Rio Chico virus* *Balsa Virus*	*Anopheles quadrimaculatus* *Anopheles triannulatus* *Culex erraticus*	Panama Florida Colombia	(86)

Rhabdoviridae	*Lobeira virus*	*Stegomyia albopicta*	Brazil	(87)
Rhabdoviridae	*Menghai rhabdovirus*	*Aedes albopictus*	China	(88)
Rhabdoviridae	*Merida-like Turkey virus*	*Culex pipiens*	Turkey	(89)
Rhabdoviridae	*Riverside virus 1*	*Ochlerotatus*	Hungary	(90)
Rhabdoviridae	*Beaumont virus* *North creek virus*	*Anopheles annulipes* *Culex sitiens*	Australia	(91)
Bunyaviridae	*Gouleako virus* *Herbert virus* *Taï virus* *Kibale virus*	*Culex sp.* *Anopheles sp.*	Côte d'Ivoire, Ghana y Uganda	(92)
Bunyaviridae	*Tucunduba virus* *Laco virus*	*Wyeomyia sp.*	Americas	(94)
Bunyanviridae	*Yongsan bunyavirus 1* *Yongsan picorna-like virus 3* *Yongsan sobemo-like virus 1*	*Aedes vexans nipponii*	Republic of Korea	(95)

Bunyanviridae	*Massilia phlebovirus*	*Phlebotomus perniciosus*	Portugal	(95)
Peribunyanviridae	*Baakal virus Lakamha virus*	*Culex nigripalpus Wyeomyia complosa*	Mexico	(96)
Peribunyanviridae	*Tacaiuma virus*	*Anopheles sp.*	Brazil	(97)
Togaviridae	*Agua salud virus*	*Culex declarator*	Panama	(98)
Reoviridae	*Palyam virus Corriparta virus*	*Culicoides sp Culex sp*	Namibia	(99)
Reoviridae	*Skunk River virus*	*Aedes trivittatus*	United States	(100)
Orthomyxoviridae	*Sinu virus*	Hematophagous mosquitoes pools	Colombia	(101)
Mesenviridae	*Cavally virus Nam Dinh virus*	*Culex sp., Aedes sp., Anopheles Uranotaenia*	Côte d'Ivoire and Vietnam	(102)
Negevirus	*Negev virus Ngewotan virus Piura virus Loreto virus*	*Culex sp. Anopheles albimanus Lutzomyia sp. Aedes aegypti*	Brazil, Peru, United States, Israel, Côte	(23)

The health of people in the 21st century has been significantly influenced by SARS-CoV-2 and other developing urban infections. The first case was report in Wuhan, China, in December 2019 (8). The genetic component of evolution and adaptive ability in human and animal populations have been the subject of several research.

The National Institute of Health in collaboration with universities and research centers in the country has led the SARS-CoV-2 genomic studies in Colombia, which have made it possible to track changes in the genetic diversity of the virus, identify variants or lineages that are circulating in the country's human population, and better understand COVID-19 dynamics. Other studies also discuss the genetic component of the nation's regional genetic heterogeneity (Table 2).

Table 2. Genomic, phylogenetic, and evolutionary analyses of SARS-CoV-2 in Colombia.

DESCRIPTION	RESEARCH
Description of the genetic variability of the Colombian SARS-CoV-2 genomes in the oligonucleotide hybridization regions of the main methods described worldwide for molecular detection.	(103)
Identification of the frequency of substitution of SARS-CoV-2 S and N proteins in South America. The substitutions D614G in S and R203K / G204R in N were the most frequent in South America, observed in 83 and 34 % of the sequences, respectively.	(104)
Report of the genome sequence of a SARS-CoV-2 viral isolate from a patient with no history of travel and ambulatory in Cali, Colombia.	(105)

Report the first SARS-CoV-2 genomes in the Colombian-Venezuelan border region. The SARS-CoV-2 genomes from Venezuela were classified into lineage B1, lineage B.1.13.	(106)
Identification of the emergence and import routes of COVID-19 into Colombia using epidemiological observations, air travel history and phylogenetic analysis. Provided evidence of multiple evidence of multiple introductions from 12 lineages.	(57)
Description of the first case of reinfection in Colombia, exhibiting different SARS-CoV-2 lineage classifications between samples (B.1 and B.1.1.269).	(107)
Identification of a highly divergent lineage of SARS-CoV-2 containing 21 distinctive mutations and emerging in the north region of Colombia.	(58)
Identification of the emergence of the B.1.621 lineage, considered a variant of interest (VOI) with the accumulation of several substitutions affecting the populations in Colombia during the three epidemiological peak.	(59)
Description of the molecular characterization of SARS-CoV-2 in military personnel and subsequent introduction of B.1.1.7 and C.36 lineages to Colombia	(108)
Description of the genomic epidemiology of SARS-CoV-2 in one of the regions of Colombia with the largest indigenous populations.	(109)

References

1. Anthony SJ, Epstein JH, Murray KA, Navarrete-Macias I, Zambrana-Torrelio CM, Solovyov A, et al. A Strategy To Estimate Unknown Viral Diversity in Mammals. Moscona A, editor. MBio [Internet]. 2013 Nov;4(5). Available from: https://journals.asm.org/doi/10.1128/mBio.00598-13
2. Breitbart M, Rohwer F. Here a virus, there a virus, everywhere the same virus? Trends Microbiol [Internet]. 2005 Jun;13(6):278–84. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0966842X05001083
3. BREITBART M. Global distribution of nearly identical phage-encoded DNA sequences. FEMS Microbiol Lett [Internet]. 2004 Jul;236(2):249–56. Available from: http://doi.wiley.com/10.1016/j.femsle.2004.05.042
4. Piret J, Boivin G. Pandemics Throughout History. Front Microbiol [Internet]. 2021 Jan 15;11. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2020.631736/full
5. Van Damme W, Dahake R, Delamou A, Ingelbeen B, Wouters E, Vanham G, et al. The COVID-19 pandemic: diverse contexts; different epidemics—how and why? BMJ Glob Heal [Internet]. 2020 Jul 27;5(7):e003098. Available from: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2020-003098
6. Kading RC, Brault AC, Beckham JD. Global Perspectives on Arbovirus Outbreaks: A 2020 Snapshot. Trop Med Infect Dis [Internet]. 2020 Sep 7;5(3):142. Available from: https://www.mdpi.com/2414-6366/5/3/142
7. Memish ZA, Perlman S, Van Kerkhove MD, Zumla A. Middle East respiratory syndrome. Lancet [Internet]. 2020 Mar;395(10229):1063–77. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673619332210
8. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature [Internet]. 2020 Mar 12;579(7798):270–3. Available from: https://www.nature.com/articles/s41586-020-2012-7
9. Kourtis AP, Appelgren K, Chevalier MS, McElroy A. Ebola Virus Disease. Pediatr Infect Dis J [Internet]. 2015 Aug;34(8):893–7. Available from: https://journals.lww.com/00006454-201508000-00022

10. Weaver SC, Costa F, Garcia-Blanco MA, Ko AI, Ribeiro GS, Saade G, et al. Zika virus: History, emergence, biology, and prospects for control. Antiviral Res [Internet]. 2016 Jun;130:69–80. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166354216301206
11. Ronca SE, Ruff JC, Murray KO. A 20-year historical review of West Nile virus since its initial emergence in North America: Has West Nile virus become a neglected tropical disease? Vasconcelos PFC, editor. PLoS Negl Trop Dis [Internet]. 2021 May 6;15(5):e0009190. Available from: https://dx.plos.org/10.1371/journal.pntd.0009190
12. Woolhouse M, Gaunt E. Ecological Origins of Novel Human Pathogens. Crit Rev Microbiol [Internet]. 2007 Jan 11;33(4):231–42. Available from: http://www.tandfonline.com/doi/full/10.1080/10408410701647560
13. Rosenberg R, Johansson MA, Powers AM, Miller BR. Search strategy has influenced the discovery rate of human viruses. Proc Natl Acad Sci [Internet]. 2013 Aug 20;110(34):13961–4. Available from: https://pnas.org/doi/full/10.1073/pnas.1307243110
14. Institute of Medicine (US) Committee on Emerging Microbial Threats to Health, Lederberg J, Shope RE, Stanley C, Oaks J. Emerging Infections: Microbial Threats to Health in the United States [Internet]. Washington, D.C.: National Academies Press; 1992. Available from: http://www.nap.edu/catalog/2008
15. Morse SS, Mazet JA, Woolhouse M, Parrish CR, Carroll D, Karesh WB, et al. Prediction and prevention of the next pandemic zoonosis. Lancet [Internet]. 2012 Dec;380(9857):1956–65. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673612616845
16. Terry Fei Fan N. Discovery of Novel Viruses From Animals, Plants, and Insect Vectors Using Viral Metagenomics [Internet]. University of South Florida; 2010. Available from: https://digitalcommons.usf.edu/etd/3506%0A%0A
17. Smith EC, Sexton NR, Denison MR. Thinking Outside the Triangle: Replication Fidelity of the Largest RNA Viruses. Annu Rev Virol [Internet]. 2014 Nov 3;1(1):111–32. Available from: https://www.annualreviews.org/doi/10.1146/annurev-virology-031413-085507

18. Weaver SC, Brault AC, Kang W, Holland JJ. Genetic and Fitness Changes Accompanying Adaptation of an Arbovirus to Vertebrate and Invertebrate Cells. J Virol [Internet]. 1999 May;73(5):4316–26. Available from: https://journals.asm.org/doi/10.1128/JVI.73.5.4316-4326.1999
19. Bichaud L, de Lamballerie X, Alkan C, Izri A, Gould EA, Charrel RN. Arthropods as a source of new RNA viruses. Microb Pathog [Internet]. 2014 Dec;77:136–41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0882401014001260
20. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. Nature [Internet]. 2004 Jul 8;430(6996):242–9. Available from: http://www.nature.com/articles/nature02759
21. Molaei G, Andreadis TG, Armstrong PM, Anderson JF, Vossbrinck CR. Host Feeding Patterns of Culex Mosquitoes and West Nile Virus Transmission, Northeastern United States. Emerg Infect Dis [Internet]. 2006 Mar;12(3):468–74. Available from: http://wwwnc.cdc.gov/eid/article/12/3/05-1004_article.htm
22. Scott TW, Githeko AK, Fleisher A, Harrington LC, Yan G. DNA profiling of human blood in anophelines from lowland and highland sites in western Kenya. Am J Trop Med Hyg. 2006;75(2):231–7.
23. Vasilakis N, Forrester NL, Palacios G, Nasar F, Savji N, Rossi SL, et al. Negevirus: a Proposed New Taxon of Insect-Specific Viruses with Wide Geographic Distribution. J Virol [Internet]. 2013 Mar;87(5):2475–88. Available from: https://journals.asm.org/doi/10.1128/JVI.00776-12
24. Pereira dos Santos T, Roiz D, Santos de Abreu FV, Luz SLB, Santalucia M, Jiolle D, et al. Potential of Aedes albopictus as a bridge vector for enzootic pathogens at the urban-forest interface in Brazil. Emerg Microbes Infect [Internet]. 2018 Dec 1;7(1):1–8. Available from: https://www.tandfonline.com/doi/full/10.1038/s41426-018-0194-y
25. Esposito DLA, Fonseca BAL da. Will Mayaro virus be responsible for the next outbreak of an arthropod-borne virus in Brazil? Brazilian J Infect Dis [Internet]. 2017 Sep;21(5):540–4. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1413867017301630

26. Gould E, Pettersson J, Higgs S, Charrel R, de Lamballerie X. Emerging arboviruses: Why today? One Heal [Internet]. 2017 Dec;4:1–13. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2352771417300137
27. Vorou R. Zika virus, vectors, reservoirs, amplifying hosts, and their potential to spread worldwide: what we know and what we should investigate urgently. Int J Infect Dis [Internet]. 2016 Jul;48:85–90. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1201971216310578
28. Kuno G, Mackenzie J, Junglen S, Hubálek Z, Plyusnin A, Gubler D. Vertebrate Reservoirs of Arboviruses: Myth, Synonym of Amplifier, or Reality? Viruses [Internet]. 2017 Jul 13;9(7):185. Available from: https://www.mdpi.com/1999-4915/9/7/185
29. Bolling B, Weaver S, Tesh R, Vasilakis N. Insect-Specific Virus Discovery: Significance for the Arbovirus Community. Viruses [Internet]. 2015 Sep 10;7(9):4911–28. Available from: http://www.mdpi.com/1999-4915/7/9/2851
30. Huang Y-J, Higgs S, Horne K, Vanlandingham D. Flavivirus-Mosquito Interactions. Viruses [Internet]. 2014 Nov 24;6(11):4703–30. Available from: http://www.mdpi.com/1999-4915/6/11/4703
31. Romero-Alvarez D, Escobar LE. Vegetation loss and the 2016 Oropouche fever outbreak in Peru. Mem Inst Oswaldo Cruz [Internet]. 2017 Apr;112(4):292–8. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762017000400292&lng=en&tlng=en
32. Fernandes J, de Oliveira RC, Guterres A, de Carvalho Serra F, Bonvicino CR, D'Andrea PS, et al. Co-circulation of Clade C New World Arenaviruses: New geographic distribution and host species. Infect Genet Evol [Internet]. 2015 Jul;33:242–5. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134815001720
33. Coffey LL, Forrester N, Tsetsarkin K, Vasilakis N, Weaver SC. Factors shaping the adaptive landscape for arboviruses: implications for the emergence of disease. Future Microbiol [Internet]. 2013 Feb;8(2):155–76. Available from: https://www.futuremedicine.com/doi/10.2217/fmb.12.139

34. Zhang Y-Z, Zhou D-J, Xiong Y, Chen X-P, He Y-W, Sun Q, et al. Hemorrhagic fever caused by a novel tick-borne Bunyavirus in Huaiyangshan, China. Zhonghua Liu Xing Bing Xue Za Zhi [Internet]. 2011 Mar;32(3):209–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21457654
35. Li C-X, Shi M, Tian J-H, Lin X-D, Kang Y-J, Chen L-J, et al. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife [Internet]. 2015 Jan 29;4. Available from: https://elifesciences.org/articles/05378
36. Shi M, Lin X-D, Chen X, Tian J-H, Chen L-J, Li K, et al. The evolutionary history of vertebrate RNA viruses. Nature [Internet]. 2018 Apr 4;556(7700):197–202. Available from: http://www.nature.com/articles/s41586-018-0012-7
37. Wolfe ND, Dunavan CP, Diamond J. Origins of major human infectious diseases. Nature [Internet]. 2007 May;447(7142):279–83. Available from: http://www.nature.com/articles/nature05775
38. Smits SL, Osterhaus AD. Virus discovery: one step beyond. Curr Opin Virol [Internet]. 2013 Apr;3(2):e1–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S187962571300031X
39. Hwang J, Jurado KA, Fikrig E. Genetics of War and Truce between Mosquitos and Emerging Viruses. Cell Host Microbe [Internet]. 2016 May;19(5):583–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1931312816301457
40. Yactayo S, Staples JE, Millot V, Cibrelus L, Ramon-Pardo P. Epidemiology of Chikungunya in the Americas. J Infect Dis [Internet]. 2016 Dec 15;214(suppl 5):S441–5. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiw390
41. Yakob L, Walker T. Zika virus outbreak in the Americas: the need for novel mosquito control methods. Lancet Glob Heal [Internet]. 2016 Mar;4(3):e148–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2214109X16000486
42. World Health Organization. Coronavirus disease (COVID-19) pandemic [Internet]. World Health Organization. 2023. Available from:

https://www.who.int/europe/emergencies/situations/covid-19

43. Vasconcelos PFC, Calisher CH. Emergence of Human Arboviral Diseases in the Americas, 2000–2016. Vector-Borne Zoonotic Dis [Internet]. 2016 May;16(5):295–301. Available from: https://www.liebertpub.com/doi/10.1089/vbz.2016.1952

44. Ashraf U, Ye J, Ruan X, Wan S, Zhu B, Cao S. Usutu Virus: An Emerging Flavivirus in Europe. Viruses [Internet]. 2015 Jan 19;7(1):219–38. Available from: http://www.mdpi.com/1999-4915/7/1/219

45. Shi M, Lin X-D, Vasilakis N, Tian J-H, Li C-X, Chen L-J, et al. Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. Ou J-HJ, editor. J Virol [Internet]. 2016 Jan 15;90(2):659–69. Available from: https://journals.asm.org/doi/10.1128/JVI.02036-15

46. Ciuoderis KA, Berg MG, Perez LJ, Hadji A, Perez-Restrepo LS, Aristizabal LC, et al. Oropouche virus as an emerging cause of acute febrile illness in Colombia. Emerg Microbes Infect [Internet]. 2022 Dec 31;11(1):2645–57. Available from: https://www.tandfonline.com/doi/full/10.1080/22221751.2022.2136536

47. Organización Panamericana de la Salud. Actualización epidemiológica semanal para dengue, chikunguña y zika en 2022 [Internet]. PLISA Plataforma de Información en Salud para las Américas. 2022. Available from: https://www3.paho.org/data/index.php/es/temas/indicadores-dengue/boletin-anual-arbovirosis-2022.html

48. Ribeiro Dias MF, Andriolo BV, Silvestre DH, Cascabulho PL, Leal da Silva M. Genomic surveillance and sequencing of SARS-CoV-2 across South Americ. Rev Panam Salud Pública [Internet]. 2023 Jan 18;47:1. Available from: https://iris.paho.org/bitstream/handle/10665.2/56995/v47e212023.pdf?sequence=1&isAllowed=y

49. Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. J Med Virol [Internet]. 2020 Apr 7;92(4):418–23. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jmv.25681

50. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. Trends Microbiol [Internet]. 2016 Jun;24(6):490–502. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0966842X16000718
51. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol [Internet]. 2019 Mar 10;17(3):181–92. Available from: http://www.nature.com/articles/s41579-018-0118-9
52. Chan JFW, Lau SKP, To KKW, Cheng VCC, Woo PCY, Yuen K-Y. Middle East Respiratory Syndrome Coronavirus: Another Zoonotic Betacoronavirus Causing SARS-Like Disease. Clin Microbiol Rev [Internet]. 2015 Apr;28(2):465–522. Available from: https://journals.asm.org/doi/10.1128/CMR.00102-14
53. Cheng VCC, Lau SKP, Woo PCY, Yuen KY. Severe Acute Respiratory Syndrome Coronavirus as an Agent of Emerging and Reemerging Infection. Clin Microbiol Rev [Internet]. 2007 Oct;20(4):660–94. Available from: https://journals.asm.org/doi/10.1128/CMR.00023-07
54. Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, et al. The biological and clinical significance of emerging SARS-CoV-2 variants. Nat Rev Genet [Internet]. 2021 Dec 17;22(12):757–73. Available from: https://www.nature.com/articles/s41576-021-00408-x
55. European Centre for Disease Prevention and Control. SARS-CoV-2 variants of concern as of 9 February 2023 [Internet]. European Union. 2023. Available from: https://www.ecdc.europa.eu/en/covid-19/variants-concern
56. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). COVID-19 Dashboard [Internet]. 2023. Available from: https://coronavirus.jhu.edu/map.html
57. Laiton-Donato K, Villabona-Arenas CJ, Usme-Ciro JA, Franco-Muñoz C, Álvarez-Díaz DA, Villabona-Arenas LS, et al. Genomic Epidemiology of Severe Acute Respiratory Syndrome Coronavirus 2, Colombia. Emerg Infect Dis [Internet]. 2020 Dec;26(12):2854–62. Available from: http://wwwnc.cdc.gov/eid/article/26/12/20-2969_article.htm

58. Laiton-Donato K, Usme-Ciro JA, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Reales-González J, et al. Novel Highly Divergent SARS-CoV-2 Lineage With the Spike Substitutions L249S and E484K. Front Med [Internet]. 2021 Jun 28;8. Available from: https://www.frontiersin.org/articles/10.3389/fmed.2021.697605/full
59. Laiton-Donato K, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Usme-Ciro JA, Prada DA, et al. Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2. Infect Genet Evol [Internet]. 2021 Nov;95:105038. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134821003361
60. Instituto Nacional de Salud. INS Detecta Variante Brasileña en Ciudadana de ese país, atendida en Leticia. [Internet]. Instituto Nacional de Salud. 2021. Available from: https://www.ins.gov.co/Noticias/Paginas/INS-Detecta-Variante-Brasileña-en-Ciudadana-de-ese-país,-atendida-en-Leticia.aspx
61. Broteinfo. Mu Variant Report [Internet]. Brote.info. 2022. Available from: https://outbreak.info/situation-reports/mu?xmin=2022-08-23&xmax=2023-02-23&loc=COL&loc=GBR&loc=Worldwide&loc=Worldwide&selected=Worldwide
62. Instituto Nacional de Salud. COVID-19 en Colombia [Internet]. Instituto Nacional de Salud. 2023. Available from: https://www.ins.gov.co/Noticias/Paginas/coronavirus-genoma.aspx
63. Smith DR. Waiting in the wings: The potential of mosquito transmitted flaviviruses to emerge. Crit Rev Microbiol [Internet]. 2017 Jul 4;43(4):405–22. Available from: https://www.tandfonline.com/doi/full/10.1080/1040841X.2016.1230974
64. Lipkin WI, Anthony SJ. Virus hunting. Virology [Internet]. 2015 May;479–480:194–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0042682215000495
65. Gallo S, Furler M, Sigel RKO. In vitro Transcription and Purification of RNAs of Different Size. Chimia (Aarau) [Internet]. 2005 Nov 1;59(11):812. Available from: https://chimia.ch/chimia/article/view/4076

66. Zambenedetti MR, Pavoni DP, Dallabona AC, Dominguez AC, Poersch C de O, Fragoso SP, et al. Internal control for real-time polymerase chain reaction based on MS2 bacteriophage for RNA viruses diagnostics. Mem Inst Oswaldo Cruz [Internet]. 2017 Apr 6;112(5):339–47. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762017000500339&lng=en&tlng=en

67. Armas Cayarga A, Perea Hernández Y, González González YJ, Dueñas Carrera S, González Pérez I, Robaina Álvarez R. Generation of HIV-1 and Internal Control Transcripts as Standards for an In-House Quantitative Competitive RT-PCR Assay to Determine HIV-1 Viral Load. Biotechnol Res Int [Internet]. 2011;2011:1–10. Available from: https://www.hindawi.com/archive/2011/964831/

68. Rampelli S, Soverini M, Turroni S, Quercia S, Biagi E, Brigidi P, et al. ViromeScan: a new tool for metagenomic viral community profiling. BMC Genomics [Internet]. 2016 Dec 1;17(1):165. Available from: http://www.biomedcentral.com/1471-2164/17/165

69. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet [Internet]. 2016 Jun 17;17(6):333–51. Available from: http://www.nature.com/articles/nrg.2016.49

70. Bahassi EM, Stambrook PJ. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis [Internet]. 2014 Sep 1;29(5):303–10. Available from: https://academic.oup.com/mutage/article/29/5/303/1039175

71. Rose R, Constantinides B, Tapinos A, Robertson DL, Prosperi M. Challenges in the analysis of viral metagenomes. Virus Evol [Internet]. 2016 Jul 3;2(2):vew022. Available from: https://academic.oup.com/ve/article-lookup/doi/10.1093/ve/vew022

72. Li Y, Wang H, Nie K, Zhang C, Zhang Y, Wang J, et al. VIP: an integrated pipeline for metagenomics of virus identification and discovery. Sci Rep [Internet]. 2016 Mar 30;6(1):23774. Available from: https://www.nature.com/articles/srep23774

73. World Health Organization. Global genomic surveillance strategy for pathogens with pandemic and epidemic potential, 2022–2032 [Internet]. World Health Organization. 2022. Available from: https://www.who.int/publications/i/item/9789240046979
74. Huhtamo E, Cook S, Moureau G, Uzcátegui NY, Sironen T, Kuivanen S, et al. Novel flaviviruses from mosquitoes: Mosquito-specific evolutionary lineages within the phylogenetic group of mosquito-borne flaviviruses. Virology. 2014;
75. J E, C C, C G, H A, C C, TJ K, et al. Characterization of a novel flavivirus isolated from Culex (Melanoconion) ocossa mosquitoes from Iquitos, Peru. J Gen Virol. 2013 Jun;94(Pt 6):1266–72.
76. Alkan C, Zapata S, Bichaud L, Moureau G, Lemey P, Firth AE, et al. Ecuador Paraiso Escondido Virus, a New Flavivirus Isolated from New World Sand Flies in Ecuador, Is the First Representative of a Novel Clade in the Genus Flavivirus. J Virol. 2015 Dec;89(23):11773–85.
77. Guzman H, Contreras-Gutierrez MA, Travassos da Rosa APA, Nunes MRT, Cardoso JF, Popov VL, et al. Characterization of Three New Insect-Specific Flaviviruses: Their Relationship to the Mosquito-Borne Flavivirus Pathogens. Am J Trop Med Hyg [Internet]. 2018 Feb 7;98(2):410–9. Available from: https://www.ajtmh.org/view/journals/tpmd/98/2/article-p410.xml
78. Korkusol A, Takhampunya R, Hang J, Jarman RG, Tippayachai B, Kim H-C, et al. A novel flavivirus detected in two Aedes spp. collected near the demilitarized zone of the Republic of Korea. J Gen Virol [Internet]. 2017 May 1;98(5):1122–31. Available from: https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.000746
79. Misencik MJ, Grubaugh ND, Andreadis TG, Ebel GD, Armstrong PM. Isolation of a Novel Insect-Specific Flavivirus from Culiseta melanura in the Northeastern United States. Vector-Borne Zoonotic Dis [Internet]. 2016 Mar;16(3):181–90. Available from: https://www.liebertpub.com/doi/10.1089/vbz.2015.1889
80. A P-C, O S, D C-L, J K, N S-F, A B, et al. Nhumirim virus, a novel flavivirus

isolated from mosquitoes from the Pantanal, Brazil. Arch Virol. 2015 Jan;160(1):21–7.

81. Hobson-Peters J, Yam AWY, Lu JWF, Setoh YX, May FJ, Kurucz N, et al. A New Insect-Specific Flavivirus from Northern Australia Suppresses Replication of West Nile Virus and Murray Valley Encephalitis Virus in Co-infected Mosquito Cells. PLoS One. 2013 Feb;8(2):e56534.
82. Vázquez A, Sánchez-Seco M-P, Palacios G, Molero F, Reyes N, Ruiz S, et al. Novel Flaviviruses Detected in Different Species of Mosquitoes in Spain. Vector-Borne Zoonotic Dis [Internet]. 2012 Mar;12(3):223–9. Available from: https://www.liebertpub.com/doi/10.1089/vbz.2011.0687
83. de Oliveira Ribeiro G, da Costa AC, Gill DE, Ribeiro ESD, Rego MO da S, Monteiro FJC, et al. Guapiaçu virus, a new insect-specific flavivirus isolated from two species of Aedes mosquitoes from Brazil. Sci Rep [Internet]. 2021 Feb 25;11(1):4674. Available from: https://www.nature.com/articles/s41598-021-83879-6
84. H C, J H, AP A, FC M, J V-C, KI F, et al. Discovery of Novel Viruses in Mosquitoes from the Zambezi Valley of Mozambique. PLoS One. 2016 Sep;11(9).
85. Quan P-L, Junglen S, Tashmukhamedova A, Conlan S, Hutchison SK, Kurth A, et al. Moussa virus: A new member of the Rhabdoviridae family isolated from Culex decens mosquitoes in Côte d'Ivoire. Virus Res [Internet]. 2010 Jan;147(1):17–24. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168170209003475
86. Contreras MA, Eastwood G, Guzman H, Popov V, Savit C, Uribe S, et al. Almendravirus : A Proposed New Genus of Rhabdoviruses Isolated from Mosquitoes in Tropical Regions of the Americas. Am J Trop Med Hyg [Internet]. 2017 Jan 11;96(1):100–9. Available from: https://ajtmh.org/doi/10.4269/ajtmh.16-0403
87. AZ LP, M S de C, FL de M, ALM R, B MR, R DS. Novel viruses in salivary glands of mosquitoes from sylvatic Cerrado, Midwestern Brazil. PLoS One. 2017 Nov;12(11).

88. Q S, Q Z, X A, X G, S Z, X Z, et al. Complete genome sequence of Menghai rhabdovirus, a novel mosquito-borne rhabdovirus from China. Arch Virol. 2017 Apr;162(4):1103–6.
89. Ergünay K, Brinkmann A, Litzba N, Günay F, Kar S, Öter K, et al. A novel rhabdovirus, related to Merida virus, in field-collected mosquitoes from Anatolia and Thrace. Arch Virol [Internet]. 2017 Jul 10;162(7):1903–11. Available from: http://link.springer.com/10.1007/s00705-017-3314-4
90. Reuter G, Boros Á, Pál J, Kapusinszky B, Delwart E, Pankovics P. Detection and genome analysis of a novel (dima)rhabdovirus (Riverside virus) from Ochlerotatus sp. mosquitoes in Central Europe. Infect Genet Evol. 2016;39:336–41.
91. LL C, BL P, AL G, BL H, RC R, SL D, et al. Enhanced arbovirus surveillance with deep sequencing: Identification of novel rhabdoviruses and bunyaviruses in Australian mosquitoes. Virology. 2014 Jan;448:146–58.
92. M M, S H, W G, A K, A L, C D, et al. Gouleako virus isolated from West African mosquitoes constitutes a proposed novel genus in the family Bunyaviridae. J Virol. 2011 Sep;85(17):9227–34.
93. Marklewitz M, Zirkel F, Rwego IB, Heidemann H, Trippner P, Kurth A, et al. Discovery of a Unique Novel Clade of Mosquito-Associated Bunyaviruses. J Virol. 2013 Dec;87(23):12850–65.
94. Chowdhary R, Street C, Travassos da Rosa A, Nunes MRT, Tee KK, Hutchison SK, et al. Genetic characterization of the Wyeomyia group of orthobunyaviruses and their phylogenetic relationships. J Gen Virol [Internet]. 2012 May 1;93(5):1023–34. Available from: https://www.microbiologyresearch.org/content/journal/jgv/10.1099/vir.0.039479-0
95. Amaro F, Zé-Zé L, Lourenço J, Giovanetti M, Becker SC, Alves MJ. Phylogenetic Analysis of Massilia phlebovirus in Portugal. Viruses [Internet]. 2021 Jul 20;13(7):1412. Available from: https://www.mdpi.com/1999-4915/13/7/1412
96. Kopp A, Hübner A, Zirkel F, Hobelsberger D, Estrada A, Jordan I, et al.

Detection of Two Highly Diverse Peribunyaviruses in Mosquitoes from Palenque, Mexico. Viruses. 2019 Sep;11(9).

97. de Melo AB, de Souza WM, Acrani GO, Carvalho VL, Romeiro MF, Tolardo AL, et al. Genomic characterization and evolution of Tacaiuma orthobunyavirus (Peribunyaviridae family) isolated in Brazil. Infect Genet Evol [Internet]. 2018 Jun;60:71–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134818300716
98. K H, M M, F Z, GJ O, RA P, JR L, et al. Agua Salud alphavirus defines a novel lineage of insect-specific alphaviruses discovered in the New World. J Gen Virol. 2020;101(1):96–104.
99. Guggemos HD, Fendt M, Hermanns K, Hieke C, Heyde V, Mfune JKE, et al. Orbiviruses in biting midges and mosquitoes from the Zambezi region, Namibia. J Gen Virol [Internet]. 2021 Sep 23;102(9). Available from: https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.001662
100. CS T, J C, SL H, BM D, RC S, LC B, et al. Skunk River virus, a novel orbivirus isolated from Aedes trivittatus in the United States. J Gen Virol. 2019;100(2):295–300.
101. MA C-G, MRT N, H G, S U, JD SV, JF C, et al. Sinu virus, a novel and divergent orthomyxovirus related to members of the genus Thogotovirus isolated from mosquitoes in Colombia. Virology. 2017 Jan;501:166–75.
102. Vasilakis N, Guzman H, Firth C, Forrester NL, Widen SG, Wood TG, et al. Mesoniviruses are mosquito-specific viruses with extensive geographic distribution and host range. Virol J 2014 111. 2014 May;11(1):1–12.
103. Álvarez-Díaz DA, Franco-Muñoz C, Laiton-Donato K, Usme-Ciro JA, Franco-Sierra ND, Flórez-Sánchez AC, et al. Molecular analysis of several in-house rRT-PCR protocols for SARS-CoV-2 detection in the context of genetic variability of the virus in Colombia. Infect Genet Evol [Internet]. 2020 Oct;84:104390. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134820302215
104. Franco-Muñoz C, Álvarez-Díaz DA, Laiton-Donato K, Wiesner M, Escandón

P, Usme-Ciro JA, et al. Substitutions in Spike and Nucleocapsid proteins of SARS-CoV-2 circulating in South America. Infect Genet Evol [Internet]. 2020 Nov;85:104557. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134820303889

105. Lopez-Alvarez D, Parra B, Cuellar WJ. Genome Sequence of SARS-CoV-2 Isolate Cali-01, from Colombia, Obtained Using Oxford Nanopore MinION Sequencing. Dennehy JJ, editor. Microbiol Resour Announc [Internet]. 2020 Jun 25;9(26). Available from: https://journals.asm.org/doi/10.1128/MRA.00573-20
106. Paniz-Mondolfi A, Muñoz M, Florez C, Gomez S, Rico A, Pardo L, et al. SARS-CoV-2 spread across the Colombian-Venezuelan border. Infect Genet Evol [Internet]. 2020 Dec;86:104616. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134820304470
107. Ramírez JD, Muñoz M, Ballesteros N, Patiño LH, Castañeda S, Rincón CA, et al. Phylogenomic Evidence of Reinfection and Persistence of SARS-CoV-2: First Report from Colombia. Vaccines [Internet]. 2021 Mar 19;9(3):282. Available from: https://www.mdpi.com/2076-393X/9/3/282
108. Ramírez JD, Muñoz M, Patiño LH, Ballesteros N, Paniz-Mondolfi A. Will the emergent SARS-CoV2 B.1.1.7 lineage affect molecular diagnosis of COVID-19? J Med Virol [Internet]. 2021 May 9;93(5):2566–8. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jmv.26823
109. Serrano-Coll H, Miller H, Rodríguez-Van Der Hamen C, Gastelbondo B, Novoa W, Oviedo M, et al. High Prevalence of SARS-CoV-2 in an Indigenous Community of the Colombian Amazon Region. Trop Med Infect Dis [Internet]. 2021 Oct 27;6(4):191. Available from: https://www.mdpi.com/2414-6366/6/4/191

5. PROBLEM RESEARCH

In recent years we have faced the emergence and epidemics of viruses such as MERS-CoV, *Ebola virus* (EBOV), *Yellow fever virus* (YFV), *Chikungunya virus* (CHIKV) and *Zika virus* (ZIKV) that until a few decades ago were confined to sylvatic circulation (1). and since the end of 2019 we are facing the disastrous SARS-CoV-2 pandemic (2). Successful viral emergence events are established by changes at the ecological level allowing interaction between different host species and the emergence of new selection pressures and adaptation of viral populations to new environments (3).

Considering that viruses with RNA genomes exist as highly diverse viral populations (quasispecies) of closely related genetic variants (4) because of their high mutation rates, it is possible to understand their enormous adaptive potential that on multiple occasions has led to the successful emergence of these viruses into new host populations (5).

Although the strengthening of the surveillance of events of public health interest has been significant in recent years, with the availability of highly sensitive and specific laboratory assays for the confirmation of cases (6), The response to disease-causing agents of unknown etiologies shows a gap in the knowledge of our "virodiversity" and a lack of specific methods for the detection of viral agents based on local genetic variability as well as the absence of protocols for preparedness and response to the introduction of agents of interest and global concern.

Historically, Colombia has reported outbreaks of febrile, respiratory, hemorrhagic, and encephalitis-causing diseases, among others, some of them fatal, whose etiological agents cannot be identified. There is an extensive list of viruses that have been described in countries of the region, which have had some impact on public health (7,8). These viruses could be contributing to the epidemiology of febrile illnesses at the national level; however, the diagnostic battery available is not sufficiently broad and some pathologies remain undiagnosed. Among the viruses

that have been identified in the country and that can be attributed to cases of unknown etiology are: *Guaroa virus (9), Ilheus virus (10),San Luis encephalitis virus (11), Leticia virus, Buenaventura virus, Necocli virus (12), Mayaro virus, Wyeomya virus, Manzanilla virus and Bussuquara virus (13) and Oropuche virus* (14) in which it is necessary to characterize the actual distribution in the country.

The Next-Generation Sequencing or NGS methods, which allow the identification and characterization of nucleic acids present in a sample for discovery of pathogens. In recent years, thousands of new viruses have been described through these methodologies (15–17), some of them representing new genera and even new families and new viruses such as SARS-CoV-2 (18–20).

On January 30, 2020 a Public Health Emergency of International Concern by COVID-19 was declared (21) and in Colombia, genomic surveillance of SARS-CoV-2 was established at the beginning of the pandemic in March 2020. The high genetic variability allowed the emergence of SARS-CoV-2 variants since the last quarter of 2020 worldwide several convergent substitutions have been evidenced and are explained by a high rate of genetic variability by selection pressure scenarios such as monoclonal antibody therapies (22,23) and vaccination (24). Substitutions in the spike protein are common, although substitutions have been relevant features, for example, the presence of E484K has been associated with lower convalescent plasma neutralizing activity (23). The presence of deletion 69/70 together with E484K and N501Y substitutions decreases the ability to neutralize antibodies (25).

The wide genetic diversity and the appearance of new variants of interest (VOI) and of concern (VOC) (26), due to its RNA genome, is unstable and prone to present a series of substitutions, insertions, and deletions, resulting in viral variants that worldwide have been associated with biological implications that are of concern, not only because of the risk of increased transmissibility, failures in molecular detection for diagnosis but also due to antigenic changes that may affect therapies with monoclonal antibodies and vaccination (24,25)**.**

Obtaining information about circulating viruses at the sylvatic level is necessary to determine their potential emergence and would allow their inclusion in the differential diagnosis and their possible emergence in public health, additionally, the identification and genetic characterization of SARS-CoV-2 would allow us to adjust molecular detection methods and contribute to the refinement of detection tests, vaccines and antivirals.

The present study proposes the active search for circulating viruses in populations of hematophagous mosquitoes in the tropical rainforest of the northwestern slope of the Sierra Nevada de Santa Marta, and evaluate the phylogenetic relationships and evolutionary mechanisms that shape the genetic variability of SARS-CoV-2 during the pandemic in Colombia the with the main objective of recognizing the virodiversity present in this ecosystem and characterizing at the molecular level SARS-CoV-2 in the human population.

This implies new questions and opportunities for the characterization of RNA viruses in rural and urban rainforest environments and their possible present and future public health implications. What viruses circulate in the vector populations present in the tropical rainforest of the Sierra Nevada de Santa Marta? What mechanisms have shaped the evolution and genetic diversity of SARS-CoV-2 in Colombia?

References

1. Tsetsarkin KA, Weaver SC. Sequential Adaptive Mutations Enhance Efficient Vector Switching by Chikungunya Virus and Its Epidemic Emergence. Heise M, editor. PLoS Pathog [Internet]. 2011 Dec 8;7(12):e1002412. Available from: https://dx.plos.org/10.1371/journal.ppat.1002412

2. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature [Internet]. 2020 Mar 12;579(7798):265–9. Available from: http://www.nature.com/articles/s41586-020-2008-3

3. Coffey LL, Forrester N, Tsetsarkin K, Vasilakis N, Weaver SC. Factors shaping

the adaptive landscape for arboviruses: implications for the emergence of disease. Future Microbiol [Internet]. 2013 Feb;8(2):155–76. Available from: https://www.futuremedicine.com/doi/10.2217/fmb.12.139

4. Domingo E, Sheldon J, Perales C. Viral Quasispecies Evolution. Microbiol Mol Biol Rev [Internet]. 2012 Jun;76(2):159–216. Available from: https://journals.asm.org/doi/10.1128/MMBR.05023-11

5. Mahy BWJ. The Evolution and Emergence of RNA Viruses. Emerg Infect Dis [Internet]. 2010 May;16(5):899–899. Available from: http://wwwnc.cdc.gov/eid/article/16/5/10-0164_article.htm

6. Burd EM. Validation of Laboratory-Developed Molecular Assays for Infectious Diseases. Clin Microbiol Rev [Internet]. 2010 Jul;23(3):550–76. Available from: https://journals.asm.org/doi/10.1128/CMR.00074-09

7. SC W, WK R. Present and future arboviral threats. Antiviral Res. 85:328–345.

8. Howard CR, Fletcher NF. Emerging virus diseases: can we ever expect the unexpected? Emerg Microbes Infect [Internet]. 2012 Jul 25;1(1):1–9. Available from: https://www.tandfonline.com/doi/full/10.1038/emi.2012.47

9. Barreto-Reyes P, Bernal C, Oya A, Groot H. Guaroa Virus, a New Agent Isolated in Colombia, South America. Am J Trop Med Hyg [Internet]. 1959 Sep 1;8(5):604–9. Available from: https://ajtmh.org/doi/10.4269/ajtmh.1959.8.604

10. Downs WG, Anderson CR, Aitken THG. The Isolation of Ilhéus Virus from Wild Caught Forest Mosquitoes in Trinidad 1. Am J Trop Med Hyg [Internet]. 1956 Jul 1;5(4):621–5. Available from: https://ajtmh.org/doi/10.4269/ajtmh.1956.5.621

11. Hoyos-López R, Soto SU, Rúa-Uribe G, Gallego-Gómez JC. Molecular identification of Saint Louis encephalitis virus genotype IV in Colombia. Mem Inst Oswaldo Cruz [Internet]. 2015 Aug 21;110(6):719–25. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-

02762015000600719&lng=en&tlng=en

12. Montoya-Ruiz C, Cajimat MNB, Milazzo ML, Diaz FJ, Rodas JD, Valbuena G, et al. Phylogenetic Relationship of Necoclí Virus to Other South American Hantaviruses (Bunyaviridae: Hantavirus). Vector-Borne Zoonotic Dis [Internet]. 2015 Jul;15(7):438–45. Available from: https://www.liebertpub.com/doi/10.1089/vbz.2014.1739

13. Groot Liévano H. Estudios sobre virus transmitidos por artrópodos en Colombia. Rev la Acad Colomb Ciencias Exactas, Físicas y Nat [Internet]. 2017 Dec 26;41(Suplemento):12. Available from: https://www.raccefyn.co/index.php/raccefyn/article/view/565

14. Ciuoderis KA, Berg MG, Perez LJ, Hadji A, Perez-Restrepo LS, Aristizabal LC, et al. Oropouche virus as an emerging cause of acute febrile illness in Colombia. Emerg Microbes Infect [Internet]. 2022 Dec 31;11(1):2645–57. Available from: https://www.tandfonline.com/doi/full/10.1080/22221751.2022.2136536

15. Smits SL, Osterhaus AD. Virus discovery: one step beyond. Curr Opin Virol [Internet]. 2013 Apr;3(2):e1–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S187962571300031X

16. Edwards RA, Rohwer F. Viral metagenomics. Nat Rev Microbiol [Internet]. 2005 Jun 10;3(6):504–10. Available from: http://www.nature.com/articles/nrmicro1163

17. Junglen S, Drosten C. Virus discovery and recent insights into virus diversity in arthropods. Curr Opin Microbiol [Internet]. 2013 Aug;16(4):507–13. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1369527413000799

18. Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M. Human viruses: discovery and emergence. Philos Trans R Soc B Biol Sci [Internet]. 2012 Oct 19;367(1604):2864–71. Available from:

https://royalsocietypublishing.org/doi/10.1098/rstb.2011.0354

19. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host Microbe [Internet]. 2020 Mar;27(3):325–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S193131282030072X

20. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, et al. Redefining the invertebrate RNA virosphere. Nature. 2016;

21. World Health Organization. Coronavirus disease (COVID-19) pandemic [Internet]. World Health Organization. 2023. Available from: https://www.who.int/europe/emergencies/situations/covid-19

22. Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med [Internet]. 2021 Apr 2;27(4):622–5. Available from: http://www.nature.com/articles/s41591-021-01285-x

23. Liu Z, VanBlargan LA, Bloyet L-M, Rothlauf PW, Chen RE, Stumpf S, et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. Cell Host Microbe [Internet]. 2021 Mar;29(3):477-488.e4. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1931312821000445

24. Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature [Internet]. 2021 Apr 22;592(7855):616–22. Available from: http://www.nature.com/articles/s41586-021-03324-6

25. Xie X, Liu Y, Liu J, Zhang X, Zou J, Fontes-Garfias CR, et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. Nat Med [Internet]. 2021 Apr 8;27(4):620–1. Available from: http://www.nature.com/articles/s41591-021-01270-4

26. European Centre for Disease Prevention and Control. SARS-CoV-2 variants of concern as of 9 February 2023 [Internet]. European Union. European Union; 2023. Available from: https://www.ecdc.europa.eu/en/covid-19/variants-concern

6. OBJECTIVES

6.1 Main objective

Identify and characterize emerging and potentially emerging RNA viruses in sylvatic, rural and urban environments in Colombia.

6.2 Specific objectives

To establish phylogenetic relationships of RNA viruses identified in urban, rural and sylvatic environments in the Sierra Nevada de Santa Marta.

To evaluate through *in vitro* assays the emergence potential of RNA viruses identified in the environments of the Sierra Nevada de Santa Marta.

To analyze the phylogenetic relationships and evolutionary mechanisms that shape the genetic variability of SARS-CoV-2 in Colombia.

To establish methodological strategies for genomic surveillance of emerging viruses with public health impact.

7. HYPOTHESIS

It is possible to identify and characterize emerging and potentially emerging RNA viruses in sylvatic, rural, and urban ecosystems in Colombia using virological, bioinformatic, and epidemiological approximations.

The emergence of RNA viruses will be determinate through i*n vitro* virological assays virological as well as the mechanisms involved in the genetic variability of the different RNA viruses characterized in Colombia will be described through phylogenetic and evolutionary analysis.

It is possible to propose and develop alternative methodological strategies for the detection and characterization of RNA viruses of public health relevance.

8. CHAPTER 1: EVOLUTION AND EMERGENCE OF MOSQUITO-BORNE VIRUSES OF MEDICAL IMPORTANCE: TOWARD A ROUTINE METAGENOMIC SURVEILLANCE APPROACH.

Laiton-Donato, K., Guzmán-Cardozo, C., Peláez-Carvajal, D., Ajami, N., Navas, M., Parra-Henao, G., & Usme-Ciro, J. (2023). Journal of Tropical Ecology, 39, E13. doi:10.1017/S0266467423000019

Link:https://www.cambridge.org/core/journals/journal-of-tropical-ecology/article/abs/evolution-and-emergence-of-mosquitoborne-viruses-of-medical-importance-towards-a-routine-metagenomic-surveillance-approach/15B9D7F53E33D41C819EB6F8408D6C1D

Annex 1

9. CHAPTER 2: VIRAL METAGENOMICS OF MOSQUITOES (DIPTERA: CULICIDAE) COLLECTED IN DIFFERENT ECOLOGICAL ENVIROMENT IN THE SIERRA NEVADA DE SANTA MARTA, COLOMBIA.

In preparation

9.1 Introduction

Arthropods of the order Diptera and the family Culicidae are a very diverse group of insects, known as mosquitoes, which have evolved the hematophagy strategy as part of their life history for a successful reproduction, being critical for the egg's development (1). Male mosquitoes exclusively feed on plants to obtain the sugars and nutrients they require for survival, while females commonly feed on vertebrate species (2). While sucking on the feeding sources, they are exposed to the source microbiota, including plant and animal viruses (3). An evolutionarily stable strategy has evolved for several viruses, which have adapted to exploit the mosquito cells (insect-specific viruses) exclusively and persistently (4), or to alternately infect the vertebrate and mosquito cells (arboviruses) establishing transmission cycles, with mosquitoes and vertebrates acting as vectors and hosts, respectively (5). The alternance between hosts has been found to be a successful strategy for arbovirus transmission, some of which have caused an enormous impact in public health, resulting of their emergence into the human population (6). Several methods based on RT-PCR and direct sequencing have been used historically to detect and characterize the virus diversity in mosquitoes of medical importance, mainly those based on the previous knowledge of the infectious agent or closely related species (7–9)). During the last decade, a drastic change in the approaches to the study of the virus diversity from complex DNA/RNA extracts of different biological sources

has been mediated by the next-generation sequencing (NGS) technologies and has led to the foundation of the viral metagenomics field (10). Through this strategy, an enormous amount of data is generated from the nucleic acid sample by an unbiased way, without previous knowledge of the virus species being present. This strategy has allowed the identification of previously described arboviruses (11), as well as novel putative virus species and major taxa (12,13).

Due to the intense and sometimes promiscuous feeding behavior of female mosquitoes, they are expected to store a good representation of the virus diversity in ecosystems where they interact (14). Therefore, the viral metagenomics approach could help us to understand the virus diversity in natural ecosystems. Here, we used a metagenomic approach to the unbiased characterization of mosquitoes from a natural ecosystem in Colombia and through the gradient from sylvatic-to-rural-to-urban environments. Our results revealed virus signatures that can directly demonstrate the presence of a high diversity of viruses and other underlying evolutionary processes that can in part explain the current arbovirus dynamics.

9.2 Materials and methods

9.2.1 Sampling area and mosquito collection

Two different linear 2-km transects corresponding to the sylvatic and rural settings, as well as several random sampling points in the urban area were selected for the present study. The study area corresponded to the northwest slope of the Sierra Nevada de Santa Marta, Magdalena, Colombia, a unique mountain range near the Caribbean Sea, which combine several thermal floors with an enormous biodiversity. The routine of the entomological fieldwork performed two days per month, during

June 2018 to March 2019 was the following: 7:00-12:00 h and 14:00-16:00 h for the manual capture with entomological nets and aspirators, 18:00-overnight for the CO_2-baited CDC light trap, and 19:00-23:00 h for active collection with the Shannon trap. Living mosquitoes were transported in lateral window bottles to the field station or the Entomology Lab, where they were separated according to the morphotype on a cool surface. Mosquito specimens were pooled (up to 14 individuals) and stored in liquid nitrogen for molecular and virologic studies, in dry for specimen mounting and classic taxonomic identification, and finally, at least one specimen of each morphotype was preserved in absolute ethanol for subsequent mosquito DNA barcoding.

9.2.2 Cell lines and virus isolation

The C6/36 cells, derived from whole larvae of the Asian tiger mosquito, *Aedes albopictus,* was cultured in Eagle's minimal essential medium (MEM) supplemented with 10% or 2% fetal bovine serum (FBS) for growing and maintenance, respectively, and incubated at 28°C in a 95% relative humidity and 5% CO_2 atmosphere. Vero cells, derived from kidney of the African green monkey, *Cercopithecus aethiops*, were cultured in MEM supplemented with 8% or 2% FBS for growing and maintenance, respectively, and incubated at 37°C in a 95% relative humidity and 5% CO_2 atmosphere.

Mosquito pools were mechanically homogenized for 25 sec and 1500 rpm in the BeadBug instrument (BenchMark Inc.), with the use of ceramic Magna lyser green beads (Roche, Mannheim, Germany) in 1.3 ml of PBS, supplemented with 10% FBS, and 1% penicillin/streptomycin. Homogenizes were subsequently cleared in a

microcentrifuge at 14,000 rpm for 15 minutes and passed through 0.4 mm filter membranes, aliquoted and stored at -70°C.

A 100-µL of every homogenized was inoculated simultaneously in fresh cultures of C6/36 and Vero cells, growing in 24-well plates. After the adhesion phase for 1 h at the corresponding temperature, 800 µL of maintenance medium was added and incubated for a week, with daily inspection of cytopathic effect (CPE). Each plate contained a positive control (DENV-2) and a Mock infection. Supernatants were collected when CPE was observed or 7 days after inoculation. Subsequently, a second passing was carried out by following the same protocol. Finally, each supernatant was collected, cleared, aliquoted and stored at -70°C.

9.2.3 Total RNA extraction

Two types of samples, I) mosquito pools homogenized as described above or II) cell culture supernatants from virus isolation attempts displaying cytopathic effect, were used for viral RNA extraction from 140-µL aliquots, through the QIAamp Viral RNA minikit (QIAGEN Inc, Germany), by following the manufacturer's instructions. Pools of 4 RNA extracts were made for a total of 48 extraction extracts. RNA analysis and generic RT-PCR for Alphavirus Flavivirus and Bunyavirus.

9.2.4 Whole transcriptome amplification

For the metagenomic characterization of mosquito samples, the whole transcriptome amplification 2 kit (WTA2, Sigma-Aldrich) was used as adapted by NetoVIR for virome research. A volume of 2.82 µL of the total RNA extract (50 ng/µL) was added to 0.5 µL of Library Synthesis Solution in RNase-free PCR tube, and incubated at 95

°C for 2 min and then cooled down at 18 °C. The resulting volume was immediately mixed with the following mix, prepared in another tube on ice: 0.5 µL of library synthesis buffer, 0.78 µL of RNase-free water, and 0.4 µL of library synthesis enzyme. Subsequently, the mix was incubated, as follows: 18°C for 10 min, 25 °C for 10 min, 37 °C for 30 min, 42 °C for 10 min, 70 °C for 20 min, and 4 °C. A new mix prepared on ice (7.5 µL of amplification mix, 60.2 µL of Rnase-free water, 1.58 µL of WTA dNTP mix, 0.75 µL of amplification enzyme) was added to 5 µL of the DNA library. Finally, the mix was amplified by PCR, at 94 °C for 2 min, followed by 17 cycles of 94 °C for 30 s, and 70 °C for 5 min, and a final cooling down to 4°C.

9.2.5 Library preparation and mNGS

DNA libraries were prepared by following the Nextera XT protocol (Illumina XXX). The first step consisted in the DNA Tagmentation, for which 2,5 µL of the WTA2 PCR product at a concentration of 1.2 ng/µL was mixed with 5 µL of Tagment DNA (TD) Buffer and 2.5 µL of Amplicon tagment mix (ATM) in a PCR tube. The mix was centrifuged at 280 × g for 1 min and incubated in a thermal cycler for 4 min at 55 °C, followed by cooling to 10 °C to ensure tagmentation. Immediately, 2.5 µL of Neutralize Tagment (NT) Buffer were added and incubated at room temperature for 5 min to stop the tagmentation process. In PCR tube strings, the following components were added to the tagmented DNA: 2.5 µL of primer 1 (i7, N70X), 2.5 µL of primer 2 (i5, S50X), and 7.5 µL of Nextera PCR master (NPM) mix, and submitted to the following thermal profile: 72°C for 3 min; 95°C for 30 s; 15 cycles (95°C for 10 s, 55°C for 30 s, 72°C for 45 s); and holding at 4°C. Libraries were cleaned up by using Agencourt AMPure magnetic beads, recovered in 25 µL of

Resuspension Buffer (RSB), and quantified through the Qubit dsDNA HS Assay kit. Library nanomolar concentration was calculated by the following formula: (ng/µL x 10^6) / (660 g/mol x 600 bp). Subsequently, libraries were manually normalized, denaturated, and loaded into the MiSeq cell flow for cluster generation.

9.2.6 Viral metagenomics pipeline

The NGS raw reads were quality-filtered, assembled, and used for taxonomic assignment by following the VirMAP pipeline. This bioinformatic protocol allows to merge nucleotide and amino acid information in both, mapping assembly to generate contigs through the algorithms of BBMap and DIAMOND, or de novo assembly and MEGAHIT or Tadpole; followed by filtering of reads outside of the superkingdom and taxonomic determination through a specific scoring system. Viral signatures corresponding to the mock infection of virus isolations were removed from the VirMAP output results.

9.3 Results

9.3.1 A high diversity of mosquito species identified in the different ecological settings

A total of 1900 mosquito and sand fly specimens were collected during the fieldwork, from which 559, 646, and 695 were obtained from the urban, rural, and sylvatic settings, respectively (Table 1). Although the major proportion of specimens remained undetermined, the study allowed to recognize an enormous diversity of mosquitoes (16 morphotypes identified at the species level, 13 identified at the genus level, and several undetermined morphotypes from the family *Culicidae*)

(Figure 1), mainly in sylvatic and rural settings and the need for further studies on the *Culicidae* taxonomy in this important ecosystem and potential speciation and endemism center. While the major abundance of sylvatic species mosquitoes was registered in the sylvatic setting, the presence of that species also in the rural and urban settings is of major importance (Figure 2).

Table 1. Mosquitoes and Sand flies specimens collected in the different ecological settings.

Setting	Collection Method	Specimens in Liquid Nitrogen	Specimens in Dry	Specimens in Ethanol	Total
Urban	Manual	430	42	11	483
	CDC Trap	16	34	0	50
	Shannon Trap	22	4	0	26
Rural	Manual	239	87	31	357
	CDC Trap	214	26	4	244
	Shannon Trap	36	9	0	45
Sylvatic	Manual	383	94	81	558
	CDC Trap	8	129	0	137
Total specimens		**1348**	**425**	**127**	**1900**

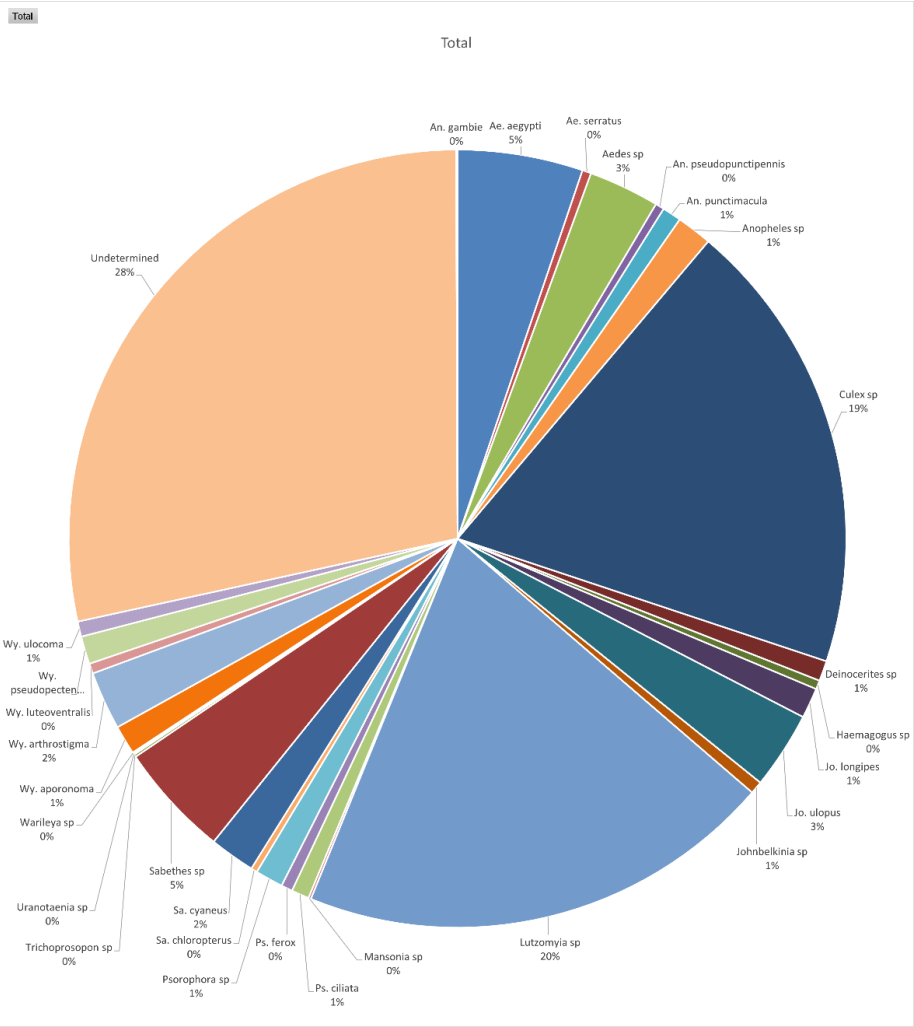

Figure 1. Mosquitoes and Sand flies species identified in the present study at different taxonomic levels.

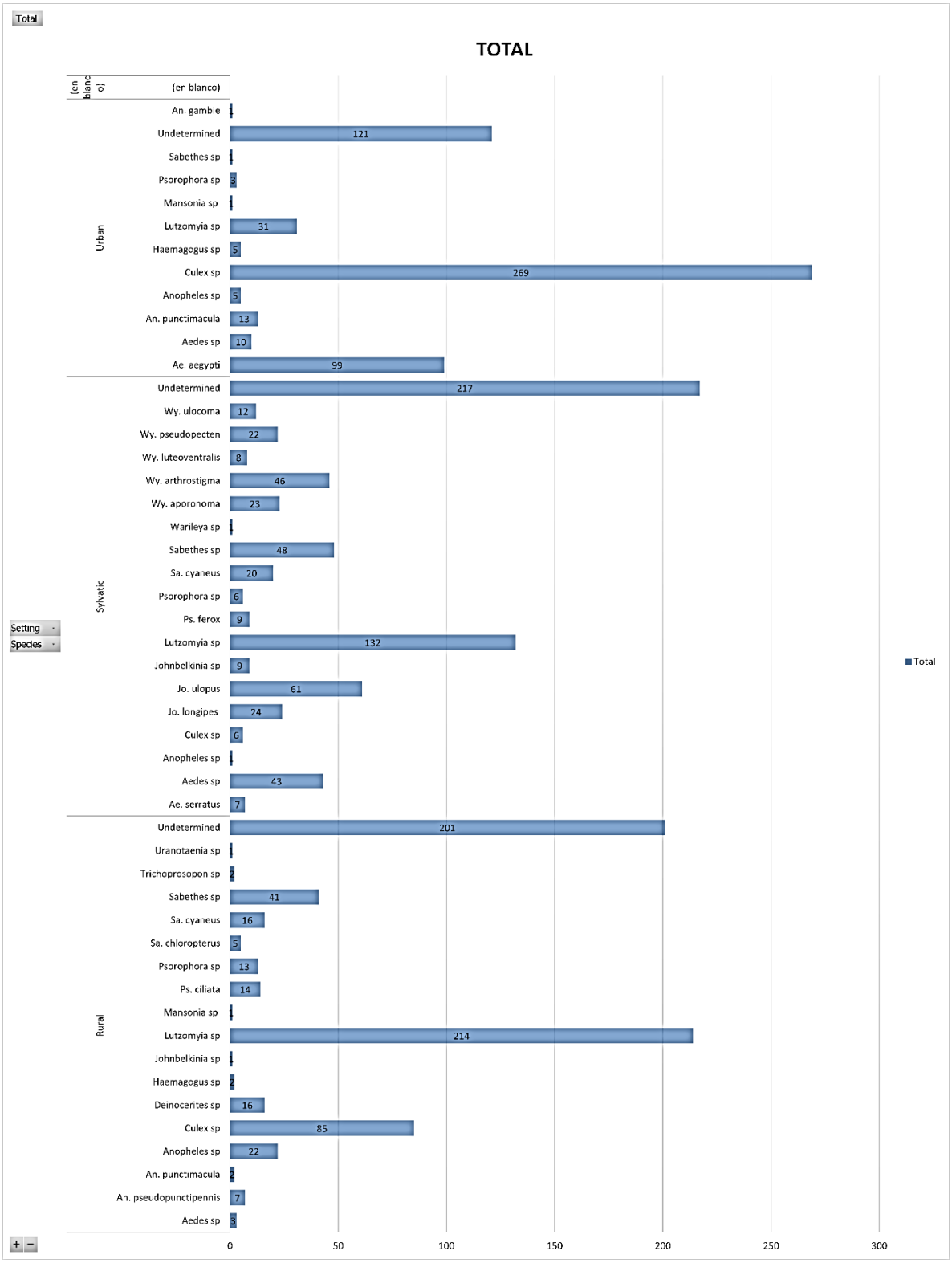

Figure 2. Mosquitoes and sand flies abundance in the different ecological settings in the SNSM, Colombia.

9.3.2 Several viral signatures identified in the mosquitoes pools

Four mNGS runs were performed, and the raw data generated were trimmed and filtered for subsequent VirMAP pipeline. The processed NGS data amount is presented in Supp. Table 1. A total of 41 viral species were identified at different taxonomic levels, from which the *Bovine viral diarrhea virus 1* (taxId=11099), *Bovine viral diarrhea virus 2* (taxId=54315), *Culex pipiens pallens densovirus* (taxId=465914), and the unclassified *Parvoviridae* (taxId=535600) were found in the mock infection and therefore removed, because they are commonly found as contaminants of the fetal bovine serum or persistently infecting the C6/36 cells. Also, *Brochothrix phage A9* (taxId=857312), *Cronobacter phage CR5* (taxId=1195085), *Human papillomavirus 107* (taxId=427343) were removed, as these viral signatures may constitute environmental contaminants from the assays. The final list of 34 viruses more closely related to the virus signatures identified in hematophagous mosquitoes from the SNSM is shown in Table 2. From the 12 identified virus families, *Flaviviridae*, *Orthomyxoviridae*, *Peribunyaviridae*, *Phenuiviridae*, *Rhabdoviridae*, *parvoviridae*, and *picornaviridae* contain species of medical and veterinary importance.

The mosquito species were partially identified through classic and molecular taxonomy through the DNA barcoding approach by Cytochrome oxidase I gene amplification, and by the NGS data obtained from sequenced mosquito pools homogenizes. The distribution of viral diversity according to the mosquito species (Figure 3) demonstrates distinct abundances among different species of mosquitoes

Table 2. Viral signatures identified in mosquitoes (Diptera: Culicidae) from the SNSM, Colombia.

Order	Family	Genus	More closely related virus species
Amarillovirales	***Flaviviridae***	*Flavivirus*	*Culex flavivirus*;taxId=390844
		Pestivirus	Pestivirus A;taxId=2170080
Articulavirales	***Orthomyxoviridae***	*Quaranjavirus*	*Wuhan Mosquito Virus 6*;taxId=1608131
Bunyavirales	***Peribunyaviridae***	Unclassified	Bunyaviridae environmental sample;taxId=1628184
			Zhee Mosquito virus;taxId=1608147
	Phenuiviridae	*Phasivirus*	*Phasi Charoen-like phasivirus*;taxId=1521189
		Tenuivirus	*European wheat striate mosaic virus*;taxId=2661631
			Fitzroy Crossing tenui-like virus 1;taxId=2755159
			Rice hoja blanca tenuivirus;taxId=12332
		Unclassified	*Atrato Gouko-like virus 1*;taxId=2689369
	unclassified	Unclassified	Bunyavirales;taxId=1980410
Mononegavirales	***Rhabdoviridae***	*Almendravirus*	*Menghai rhabdovirus*;taxId=1919071
			Puerto Almendras virus;taxId=1479613
	Xinmoviridae	*Anphevirus*	*Aedes anphevirus*;taxId=2230910
Piccovirales	***Parvoviridae***	Unclassified	*Culex densovirus*;taxId=2304518
Picornavirales	***Picornaviridae***	Unclassified	Picornaviridae sp.;taxId=1530251
Sobelivirales	***Solemoviridae***	*Sobemovirus*	*Bat sobemovirus*;taxId=1340805
		Unclassified	*Atrato Sobemo-like virus 4*;taxId=2689350
Tolivirales	***Luteoviridae***	Unclassified	*Bat luteovirus*;taxId=1340807
			Culex-associated Luteo-like virus;taxId=2304522
			unclassified Luteoviridae;taxId=94699
Tymovirales	***Tymoviridae***	Unclassified	*Ek Balam virus*;taxId=2488704
Unclassified	***Partitiviridae***	Unclassified	*Atrato Partiti-like virus 2*;taxId=2689327
	Unclassified	*Negevirus*	*Rinkaby virus*;taxId=2651953
		Unclassified	*Culex luteo-like virus*;taxId=2010270
			Guato virus;taxId=1795437
			Hubei noda-like virus 5;taxId=1922985
			Hubei virga-like virus 2;taxId=1923335
			Mayapan virus;taxId=2488589
			Salarivirus Mos8CM0;taxId=1925501
			superkingdom=Viruses;Riboviria;taxId=2559587
			unclassified RNA viruses ShiM-2016;taxId=1922348
			uncultured virus;taxId=340016
			Wenzhou sobemo-like virus 3;taxId=1923659

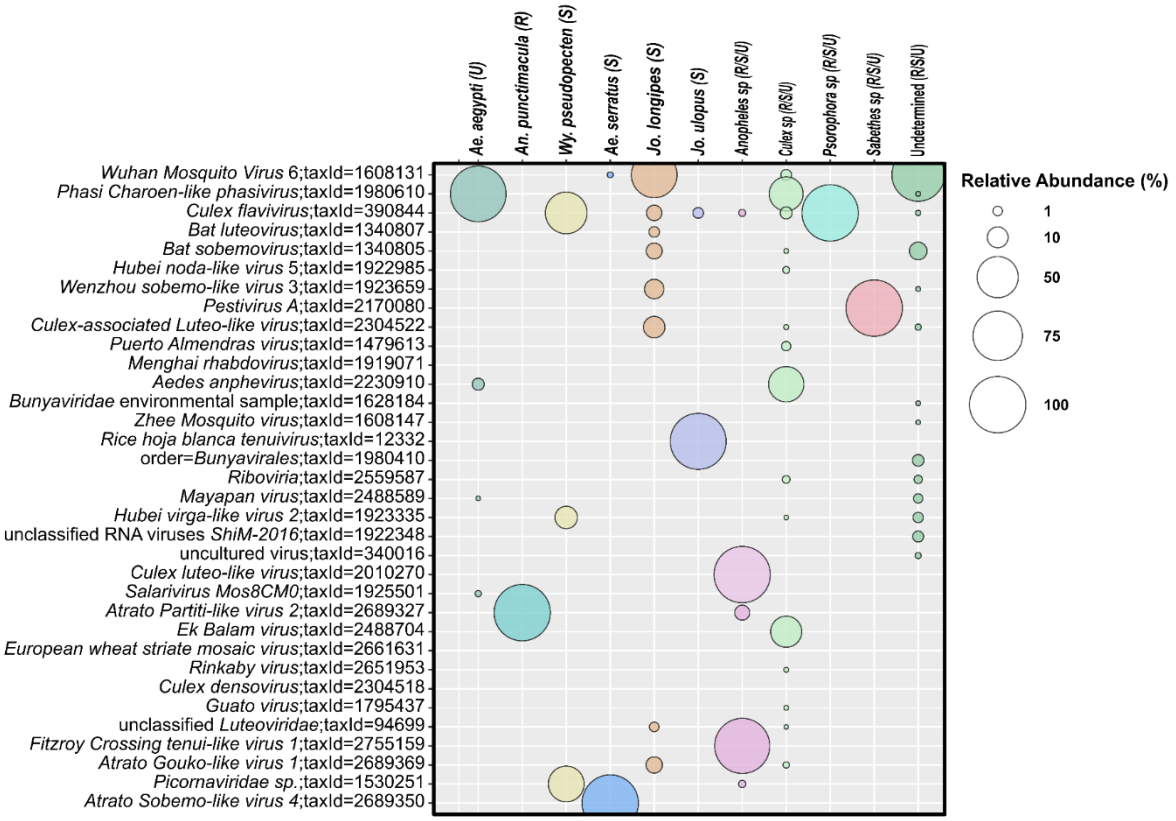

Figure 3. The distribution of viral diversity according to the mosquito species collected in the SNSM, Colombia.

9.3.3 Nearly complete viral genome of segmented virus assembled from viral metagenomics data

Accounting for the sample source, viral enrichment was expected for samples extracted from cell culture supernatants, while mayor diversity of viral signatures was expected from mosquitos' pools homogenizes, as their whole body was processed, and some specimens had previously feed on vertebrate sources. It does not imply vector competence but allows the identification of viruses actively circulating in every ecological scenario. The number of reads and deep of coverage in that samples was generally low, but it was enough for taxonomic determination from NGS data. On the

other hand, the use of supernatants of virus isolations constitutes a powerful mean to recover nearly complete genomes from infecting viruses.

9.3.4 High coverage of delimitated viral regions suggests their presence as subviral genomes.

An increasing number of endogenous viral elements derived from non-retroviral genomes have been identified in mosquitoes, and their functional role in the mosquito innate immunity is being unveiled. While analyzing the viral signatures identified through the VirMAP pipeline, a notorious pattern was found, characterized by the presence of covered sequences at very high depth, flanked by uncovered regions where there was no evidence of genetic information from the sample.

9.3.5 Low identity of some viral signatures with more closely related species of viruses

Some viral signatures identified in the present study showed a very high identity to the more closely related sequences available in the GenBank. However, a high number of the viral signatures were distantly related to the candidate species.

9.4 Discussion

Tropical environments are recognized as biodiversity hotspots. In terms of insect species, the diversity is as enormous as poorly studied. Mosquitoes and sandflies are present in high abundance in sylvatic settings where several species interact and suck of several vertebrate feeding sources. The identification of several viral signatures in the mNGS data of the present study performed on homogenizes of

whole mosquito pools can be the result of active replication in the biological vector, or simply the result of a recent feeding on an infected source.

The genome coverage was found to be variable, with some segmented viruses being identified at high coverage with all segments represented in the genome assembly and other viruses being only partially covered. Although NGS sequencing techniques allow the detection of viral reads, the present study was successful and allowed the detection of some viruses with complete genes and even genomes in some viral species with excellent coverage and depth, demonstrating the methodological approach as ideal for determining viral presence in a vector, host, or specific environment. In addition, molecular detection is the input to implement *in vitro* analysis studies focused on characterizing each of the new viruses discovered at the phenotype level.

Studies in the region highlight the discovery of new viral agents in mosquito species and this study highlighted the enormous virodiversity through the detection of approximately 31 new viruses in circulating mosquito species in sylvatic, urban, and rural areas of SMSM, demonstrating viral establishment in different areas. and environments whose interaction between vectors and hosts allow for favorable scenarios for emergence towards human populations. In addition, Colombia has different types of ecosystems still without viral characterization, which is essential to establish routine virological surveillance strategies focused on viruses that may represent a threat to public health.

9.5 Conclusion

The SNSM has a high diversity of viral agents present in blood-sucking mosquitoes are of special importance due to its ecological interaction with vertebrate species such as NHP and due to the risk of contact with humans, causing jumping events in the species barrier. Faced with the eventual emergence of some unknown viral agent that could have an impact on human health, current methods of diagnoses are limited, and it is almost impossible to determine the etiologic agent of a given syndrome. This is how the study of virodiversity in sylvatic, urban and rural environments through mNGS is essential to recognize the virus species present in different environments, determine their cycles of transmission and its potential impact on human health.

9.6 References

1. Shaw WR, Catteruccia F. Vector biology meets disease control: using basic research to fight vector-borne diseases. Nat Microbiol [Internet]. 2018 Aug 27;4(1):20–34. Available from: https://www.nature.com/articles/s41564-018-0214-7
2. Takken W, Verhulst NO. Host Preferences of Blood-Feeding Mosquitoes. Annu Rev Entomol [Internet]. 2013 Jan 7;58(1):433–53. Available from: https://www.annualreviews.org/doi/10.1146/annurev-ento-120811-153618
3. Brinkmann A, A N, Kohl C. Viral Metagenomics on Blood-Feeding Arthropods as a Tool for Human Disease Surveillance. Int J Mol Sci [Internet]. 17: 1743. Available from: https://doi.org/10.3390/ijms17101743
4. Roundy CM, Azar SR, Rossi SL, Weaver SC, Vasilakis N. Chapter Four - Insect-Specific Viruses: A Historical Overview and Recent Developments. In:

Advances in Virus Research [Internet]. 2017. p. 119–46. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065352716300641

5. Ebel GD. Promiscuous viruses—how do viruses survive multiple unrelated hosts? Curr Opin Virol [Internet]. 2017 Apr;23:125–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1879625717300202

6. Weaver SC, Charlier C, Vasilakis N, Lecuit M. Zika, Chikungunya, and Other Emerging Vector-Borne Viral Diseases. Annu Rev Med [Internet]. 2018 Jan 29;69(1):395–408. Available from: https://www.annualreviews.org/doi/10.1146/annurev-med-050715-105122

7. Sánchez-Seco MP, Rosario D, Domingo C, Hernández L, Valdés K, Guzmán MG, et al. Generic RT-nested-PCR for detection of flaviviruses using degenerated primers and internal control followed by sequencing for specific identification. J Virol Methods [Internet]. 2005 Jun;126(1–2):101–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S016609340500042X

8. Rico-Hesse R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. Virology [Internet]. 1990 Feb;174(2):479–93. Available from: https://linkinghub.elsevier.com/retrieve/pii/004268229090102W

9. Lanciotti RS, Roehrig JT, Deubel V, Smith J, Parker M, Steele K, et al. Origin of the West Nile Virus Responsible for an Outbreak of Encephalitis in the Northeastern United States. Science (80-) [Internet]. 1999 Dec 17;286(5448):2333–7. Available from: https://www.science.org/doi/10.1126/science.286.5448.2333

10. Zhang Y-Z, Shi M, Holmes EC. Using Metagenomics to Characterize an Expanding Virosphere. Cell. 2018 Mar;172(6):1168–72.

11. Russell JA, Campos B, Stone J, Blosser EM, N B-C, Jacobs JL. Unbiased Strain-Typing of Arbovirus Directly from Mosquitoes Using Nanopore Sequencing: A Field-forward Biosurveillance Protocol. Sci Rep. 8:5417.

12. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, et al. Redefining the invertebrate RNA virosphere. Nature. 2016;

13. Xiao P, Li C, Zhang Y, Han J, Guo X, Xie L, et al. Metagenomic Sequencing From Mosquitoes in China Reveals a Variety of Insect and Human Viruses. Front Cell Infect Microbiol [Internet]. 2018 Oct 19;8. Available from: https://www.frontiersin.org/article/10.3389/fcimb.2018.00364/full

14. Batson J, Dudas G, Haas-Stapleton E, Kistler AL, Li LM, Logan P, et al. Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. Elife [Internet]. 2021 Apr 27;10. Available from: https://elifesciences.org/articles/68353

10. CHAPTER 3: NOVEL PUTATIVE *Tymoviridae*-like virus ISOLATED FROM *Culex* MOSQUITOES IN COLOMBIA.

Laiton-Donato, K.; Guzmán, C.; Perdomo-Balaguera, E.; Sarmiento, L.; Torres-Fernandez, O.; Ruiz, H.A.; Rosales-Munar, A.; Peláez-Carvajal, D.; Navas, M.-C.; Wong, M.C.; Junglen, S.; Ajami, N.J.; Parra-Henao, G.; Usme-Ciro, J.A. Novel Putative Tymoviridae-like Virus Isolated from Culex Mosquitoes in Colombia. Viruses 2023, 15, 953. doi.org/10.3390/v15040953

Link: https://www.mdpi.com/1999-4915/15/4/953

Annex 2

11. CHAPTER 4: GENOMIC EPIDEMIOLOGY OF SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS 2, COLOMBIA

Laiton-Donato K, Villabona-Arenas CJ, Usme-Ciro JA, Franco-Muñoz C, Álvarez-Díaz DA, Villabona-Arenas LS, Echeverría-Londoño S, Cucunubá ZM, Franco-Sierra ND, Flórez AC, Ferro C, Ajami NJ, Walteros DM, Prieto F, Durán CA, Ospina-Martínez ML, Mercado-Reyes M.Emerging Infectious Diseases. 2020 Dec;26(12):2854-2862. doi: 10.3201/eid2612.202969

Link: https://pubmed.ncbi.nlm.nih.gov/33219646/

Annex 3

12. CHAPTER 5 NOVEL HIGHLY DIVERGENT SARS-CoV-2 LINEAGE WITH THE SPIKE SUBSTITUTIONS L249S AND E484K

Laiton-Donato K, Usme-Ciro JA, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Reales-González J, Prada DA, Corchuelo S, Herrera-Sepúlveda MT, Naizaque J, Santamaría G, Wiesner M, Walteros DM, Ospina Martínez ML, Mercado-Reyes M.Frontiers in Medicine . 2021 Jun 28;8:697605. doi:10.3389/fmed.2021.697605

Link: https://pubmed.ncbi.nlm.nih.gov/34262921/

Annex 4

13. CHAPTER 6: CHARACTERIZATION OF THE EMERGING B.1.621 VARIANT OF INTEREST OF SARS-CoV-2

Laiton-Donato K, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Usme-Ciro JA, Prada DA, Reales-González J, Corchuelo S, Herrera-Sepúlveda MT, Naizaque J, Santamaría G, Rivera J, Rojas P, Ortiz JH, Cardona A, Malo D, Prieto-Alvarado F, Gómez FR, Wiesner M, Martínez MLO, Mercado-Reyes M.. Infection Genetic and Evolution. 2021 Nov;95:105038. doi: 10.1016/j.meegid.2021.105038

Link: https://pubmed.ncbi.nlm.nih.gov/34403832/

Annex 5

14. CHAPTER 7: USEFULNESS OF AN *IN VITRO*-TRANSCRIBED RNA CONTROL FOR THE DETECTION AND QUANTIFICATION OF *Yellow fever virus* THROUGH REAL-TIME REVERSE TRANSCRIPTION-POLYMERASE CHAIN REACTION.

Laiton-Donato K, Quintero-Cortés P, Franco-Salazar JP, Peláez-Carvajal D, Navas MC, Junglen S, Parra-Henao G, Usme-Ciro JA. Infection Diseases Now. 2023 Jan 26;53(3):104654. doi: 10.1016/j.idnow.2023.104654.

Link: https://pubmed.ncbi.nlm.nih.gov/36709865/

Annex 6

15. GENERAL DISCUSSION

Viral emergency in sylvatic, rural and urban areas.

Over the past twenty-five years, the emergence and the reemergence of numerous infectious diseases around the world have coincided with unprecedented rates of change in the structure and diversity of the environment and human social and economic systems. The coincidence of broad scale environmental changes, the expansion of human social and economic networks, and the emergence of infectious diseases may point to underlying predictable ecological relationships. The emergence of an infectious disease caused by a virus and its rapid spreading in a human population can have serious consequences for public health (1). In general, viral emergencies are events that require a rapid and coordinated response by public health authorities, including measures to control and prevent the spread of the virus, as well as the development of antivirals and vaccines (2).

Most of the emerging infectious diseases affecting humans have zoonotic and vectorial origins. The zoonotic process occurs through direct contact of human with infected animals, consumption of contaminated meat, exposure to infected body fluids, or the bite of an infected vector. Vector cycles involve mainly mosquitoes and ticks that feed on infected hosts (3). Viral emergence is a complex phenomenon that can be caused by several factors and determinants. Environmental change due to deforestation, urbanization, expansion of agriculture and livestock including international trade in animals and animal products, climate change, and human mobility, including tourism, international travel, and migration, can increase human-animal-vector contact (4). In addition, the limitations to access to health care, including access to medicines and vaccines, and failure to maintain sustainability in international collaboration can increase the risk of the spread of viral diseases.

Therefore, it is imperative to prevent and control the transmission of infectious diseases from animals to humans through monitoring in hosts and vectors to detect

viruses with emerging potential and to assess the prevalence of diseases in animal and human populations. This could be done by wildlife surveillance, collection of samples of blood, saliva, tissues, or other body fluids from wildlife and domestic animals, and through the capture of circulating vectors in wild, rural, and urban areas and its viral characterization. The information gathered through viral monitoring can be used to develop and complement cross-sectoral strategies with collaboration of various disciplines and sectors for the implementation of safer animal production and handling practices, education in zoonotic disease promotion and prevention, and the development of public policies for the development of effective treatments and vaccines.

Viruses with emerging and emerging potential in Colombia

Colombia is one of the most biodiverse countries, with several ecosystems that resemble the tropical world. These ecosystems are home to multiple vertebrate and invertebrate species that ecologically interact in an ancestral and stable relationship. Mosquitoes (Diptera: Culicidae) is a significant diverse taxonomic group in the country, including several species of medical importance as disease vectors. Viruses circulating in natural environments, such as those recently described in the SNSM are considered Insect-specific virusesHowever, due to the vector feeding behavior, these viruses are unavoidably exposed to the vertebrate cells, a critical step for adaptation, species jumping and virus emergence. There is an extensive list of emerging viruses that have been described in countries of the Americas, which have had some impact on public health (5,6)

The interaction between sylvatic, rural, and peri-urban vector species and people has been mainly determined by the human activity that leads to alteration in the natural dynamics of the ecosystems. The resulting interactions could allow enzootic arboviruses to potentially adapt to new vectors and hosts, hence leading to their emergence in epizootic or epidemic transmission cycles with significant impact in animal and human health. Special attention has been given to arboviruses circulating in the urban cycle, mainly DENV, whose sylvatic origin has been evidenced by the

isolation of strains from mosquitoes and primates in Asian and African forests displaying a basal (ancestral) position at the phylogenetic level, with respect to the strains circulating endemically/epidemically in human population (7). Another virus of interest is YFV, which was introduced to the Americas from Africa through slavery of people (8), This virus was successfully maintained in the urban cycle due to advantages of mosquito Aedes aegypti to invade during decades until aggressive vector control and immunization campaigns allowed to interrupt its transmission (9,10). However, yellow fever epizootics and outbreaks continued occurring in Latin American countries where the virus successfully infected and transmitted between sylvatic vectors and non-human primates. Similarly, the origin of viruses such as CHIKV and ZIKV has been traced to the sylvatic area in Africa during the 1950's (11,12). Interestingly, the reemergence of these viruses several decades after its emergence supposes the successful virus maintenance, for which several mechanisms have been hypothesized, including trans ovarian transmission in vector, the existence of a natural vertebrate reservoir or a stable transmission cycle between vectors and natural vertebrate hosts.

SARS-CoV-2 is an emerging RNA virus becoming the best example of viral evolution, adaptation to a new host and emergence with public health consequences (13). This coronavirus rapidly spread to other countries, causing a pandemic that affected millions of people worldwide and forced to respond to the challenges imposed by this infectious disease with high public health impact. In Colombia, the acquired research skills and laboratory capacity derived from next generation sequencing and metagenomics was determinant to respond to the SARS-CoV-2 emergence and pandemic explosion, and to align the country efforts to the global challenge of genomic surveillance as a response to future events with epidemic and pandemic potential proposed by WHO (14).

Viral metagenomics through Next Generation Sequencing (mNGS)

The availability and robustness of metagenomics by Next Generation Sequencing (mNGS) applied to the detection of viral species in a biological sample has made it possible to devise robust strategies and protocols for the identification of novel viral species, describe virodiversity in each environment (15), and incriminate some viruses with the potential to emerge and cause a new public health problem (16).

Metagenomics is not only a valuable technique for viral detection as the basis for subsequent approaches to determine the viral fitness, strain virulence, transmission ability, and epidemic potential (17). When viral sequences are enriched, this approximation also provides data for other research methodologies, such as genotype surveillance supporting epidemiological changes, phylogeographic analysis and host receptor interaction (18,19). The present study is the first one in Colombia to contribute to design and implementation of a metagenomic approach using mNGS next generation sequencing for virus discovery and characterization of viruses with emerging potential, and to contribute to the description of the virodiversity of natural ecosystems in the country, becoming the reference for local and regional research interested in implementing similar or broader studies.

Discovery of new insect-associated viruses in mosquitoes from the Sierra Nevada de Santa Marta.

Different studies have identified an enormous virodiversity in several species of mosquitoes contributing to discover novels viral species for subsequent taxonomic characterization (15,20–24) and to the scientific community in determining those viruses with the biological characteristics to establish themselves in phases of zoonotic or anthropophilic transmission (25).

Historically, the emergence of arboviruses of public health interest has been demonstrated from sylvatic cycles to rural and urban transmission, being essential to identify in natural ecosystems such as the Sierra Nevada de Santa Marta the

reservoirs and vectors that interact with the discovered viruses; although spillover events are infrequent, there is the possibility of adaptation in an ecosystem that maintains favorable conditions so that the set of determinants of vector species, viruses and reservoirs are adequate to establish and maintain constant transmission cycles (26) and with the emerging potential towards the human population in conditions where human populations intervene in the ecosystems (19). CHAPTER 1.

The present study demonstrated a broad panorama of new or unreported viruses in the Sierra Nevada de Santa Marta and provides a precedent for evaluation of other ecosystems and highlights the role of hematophagous mosquitoes analysis in the routine practical in transects defined of sylvatic and rural areas to determine the temporal, seasonal and spatial variation of the populations of mosquitoes and their virodiversity.

The metagenomic strategy implemented in the study revealed a high diversity of virus species in the mosquito populations analyzed, and a total of 13 genera were identified. The virus species identified were mainly: *Wuhan Mosquito Virus, Phasi Charoen-like phasivirus, Culex flavivirus, Bat luteovirus, Bat sobemovirus, Hubei noda-like virus 5, Wenzhou sobemo-like virus, Pestivirus A, Culex-associated Luteo-like virus, Puerto Almendras virus, Menghai rhabdovirus, Aedes anphevirus, Zhee Mosquito virus, Rice white leaf tenuivirus, Riboviria, Mayapan virus, Hubei virga-like virus 2, unclassified RNA viruses ShiM-2016, Culex luteo-like virus, Salarivirus, Atrato Partiti-like virus 2, Guachaca virus, European wheat striate mosaic virus, Rinkaby virus, Culex densovirus, Guato virus, unclassified Luteoviridae, Fitzroy Crossing tenui-like virus 1, Atrato Gouko-like virus 1, Picornaviridae sp, Atrato Sobemo-like virus 4.* None of the species are classified as arboviruses, although they have been evidenced in different types of phylogenetically distant hosts including invertebrates, vertebrates, and plants. CHAPTER 2 and CHAPTER 3

The wide virodiversity found is consistent with the biological diversity in Colombia (27) and with a successful methodological design of collection in transects in rural and urban sylvatic regions. However, most of the readings of viral sequences

present low identity at nucleotide and amino acid level with respect to previously described viral species; although, probably several of the described viral species are new and belong to the identified genera and families, this limitation is due to a lower percentage of coverage of the viral genome and because the taxonomic assignment is performed by means of the identity with sequences deposited in viral databases (28) CHAPTER 2. Therefore, it is essential to complement metagenomics with NGS sequencing techniques including the use of random primers, targeted probes or amplicon sequencing to obtain sequences with a higher percentage of coverage and depth and even the characterization of complete viral genomes of interest (29–31)

Virus-Vector interaction

Several viruses have developed an evolutionary strategy to be either insect-specific viruses or alternately infect vertebrate and mosquito cells (arboviruses) to establish transmission cycles, with mosquitoes and vertebrates as the respective vectors and hosts (32). Colombia ranks second in biodiversity worldwide, and its different ecosystems harbor a high diversity of viral agents (27). Of this wide diversity, virus species in hematophagous mosquitoes are of special importance due to their ecological interaction with vertebrate species that may have contact with humans in sylvatic and rural areas, leading to species barrier jumping events (33).The emergence of an unknown viral agent can have a negative impact on human health, being fundamental the study of virodiversity in rural and sylvatic ecosystems to recognize the circulating virus species, their transmission cycles and their potential impact on human health (34).

Although the present study did not identify arboviruses specific to the urban cycle, nor those in which enzootic circulation has been demonstrated, a base of viruses circulating in the study area has been defined, with a predominance of viruses considered insects-specific demonstrated the usefulness of viral metagenomics applied to the characterization of mosquito virome (35) and in some cases,

demonstrating its usefulness in monitoring arbovirus activity in regions where there is close contact with the human population (36,37)

Most vector species identified in this study are exclusive to sylvatic areas; however, the proximity between peri-urban rural and sylvatic areas is becoming increasingly frequent, so it is important to consider the latent risk of interaction between arboviruses of urban cycles and vertebrate vectors or reservoirs in sylvatic areas, known as spillback (26,38). In Colombia, different arboviruses of public health importance circulate in urban and rural areas, among them DENV, CHIKV and ZIKV, whose vectors has a wide dissemination and geographical distribution (39–43), implying the availability of mosquito sampling strategies and viral detection methods to evaluate the ecology and dynamics of these anthropophilic viruses to colonize potential new vectors, as well as to continue with studies in different regions of the country to determine arboviruses in sylvatic and rural mosquito species (19). CHAPTER 1.

Finally, in the Sierra Nevada de Santa Marta, the presence and abundance of hematophagous mosquitoes and viral species are a result of ecological factors that favor contact with the human population, transmission, and emergence. Most of the species identified in this study have not been previously described in the country, and their biology and transmission cycles are unknown. CHAPTER 2. For these reasons, they serve as crucial starting points for future research to determine the emerging potential of the viruses and to evaluate the actual risk of transmission in the studied region as well as their effects on human and animal health.

Phylogenetic relationships and potential transmission scenarios of RNA viruses identified in urban, rural and sylvatic environments in the Sierra Nevada de Santa Marta.

The taxonomic classification of most of the viral species detected in the Sierra Nevada de Santa Marta was carried out by means of similarity with sequences available in genomic databases, in only some of them the genomic fragments obtained presented adequate length and representation of variable regions to

perform phylogenetic analysis. For the viral species in which it was possible to perform phylogenetic analysis, monophyletic groups with Bootstrap greater than 70% were evidenced, demonstrating the circulation of new viral species in the Sierra Nevada de Santa Marta. CHAPTER 2.

The phylogenetic analysis of a virus isolated in a pool of Culex sp mosquitoes allowed the denomination of *Guachaca virus*, forming a monophyletic group with species exclusively infecting insects, and more interestingly, the colonization of hematophagous mosquitoes and supporting the need to assign it as a new virus within a specific genus of specific insect viruses in the family *Tymoviridae*. Additionally, this new phylogeny suggests 3 new genera of insect-specific viruses and is consistent with previous studies conducted in Mexico and Brazil. The characterization will be submitted to the International Committee on Taxonomy of Viruses (44) to include *Guachaca virus* and reconsider the current taxonomic classification of the family *Tymoviridae*. CHAPTER 3.

In the Sierra Nevada de Santa Marta there is a close and almost imperceptible proximity between the rural and Sylvatic areas; in addition, the areas are frequently visited by tourists and inhabited by indigenous people and local farmers becoming the perfect scenario for the emergence of viruses towards the human population. CHAPTER 2. Perhaps the most significant finding was the discovery of the *Guachaca virus* circulation, whose family virus is related to infections in plants and with adaptive potential towards mosquito populations, implying to consider the potential emergence of circulating viruses in species of vertebrate and mosquitoes in sylvatic regions. CHAPTER 3.

Emerging potential of RNA viruses identified in the ecosystem of the Sierra Nevada de Santa Marta.

The present study included characterization of the virome by means of NGS and moreover *in vitro* analysis of mosquito pools collected in the Sierra Nevada de Santa Marta, by means of attempted viral isolation in mosquito cells and in cells of vertebrates as an initial approximation to observe the ability of viruses to replicate in

cells other than those of their presumed natural hosts. Additionally, RT-PCR was performed on all pools to search Alphavirus, Flavivirus and Bunyavirus sequences to assess the presence of anthropophilic viruses with an impact on public health.

The approach designed for *in vitro* characterization was successful considering the complete genome sequence of the putative *Guachaca virus* from a pool of *Culex sp* and subsequently the complete characterization carried out by end corroboration using RACE strategy, ultrastructural analysis, infection assays and growth curves in different cell lines, demonstrating its inability to infect vertebrate cells, including human cells (Hela, HEK293, A549 and U937). CHAPTER 3.

One methodological approach to describe the successful process of viral adaptation is by *in vitro* adaptive evolution assays observing the generation of mutations through successive viral passages in culture (45). A limitation of the study was the initial implementation of adaptive evolution assays, to analyze the permissiveness of different cell lines to viral infection and adaptation in hybrid cell culture assays. In a first approximation a hybrid cultures of C6/36 HT cells and Vero cells were established, and viability was evidenced for seven days post infection under controlled conditions of temperature, humidity, and CO_2. However, it is necessary to standardize viral infection assays and to evaluate the existence of biological efficacy by determining viral infectious particles in the different cell passages and to identify genomic mutations associated with evolutionary selection by NGS sequencing.

Vector competence is the measure of the potential of a vector to be infected by a virus and subsequently transmit it, (46) and is determined by the intrinsic factors of the vector, the virus and the environment (47). Therefore, is necessary to implement *in vivo* assays in mosquito species in which interested viral species were detected and to evaluate infection rate, spread rate, transmission rate and transmission efficiency.

Phylogenetic relationships and evolutionary mechanisms that shape the genetic variability of SARS-CoV-2 in Colombia.

Since the declaration of pandemic in February 2020, SARS-CoV-2 monitoring has been intensified at national level and WHO has recommended genome sequencing as a determinant to address the health emergency and monitor the circulation of different variants (48). The characterization of SARS-CoV-2 is essential for surveillance on real-time of dynamics of the circulation of variants in the country, viral dispersal routes, association with transmissibility, severity, and evasion of the innate and acquired immune response, as well as for the design and refinement of molecular diagnostic methods, antivirals and vaccines (49–51).

The availability of SARS-CoV-2 genomes obtained in the present study allowed the articulation of Colombia in a global initiative to provide genomic data of SARS-CoV-2, which has an immediate application in reconstruction of the history of transmission, analysis of genetic variability and viral evolution, identify the emergence of new variants and describe the dynamics of global dispersion. Although limited genomic-level data on SARS-CoV-2 were available, the first approach to the origins and introductions of SARS-CoV-2 in Colombia was made, allowing us to understand the routes of introduction and emergence, the diversity of circulating lineages, distribution, and establishment by analyzing epidemiological phylogenetic data and air travel histories. The study provided evidence of multiple introductions distributed in 12 lineages and originating mainly from Europe, the United States of America and Mexico. The phylogenetic and evolutionary findings highlighted the broad genetic diversity represented in the A and B lineages described at the beginning of the pandemic, identified multiple importation events, and demonstrated the presence of linked genetic and epidemiological chains of transmission. In general terms, the initial evolutionary history of SARS-CoV-2 in Colombia and in the region was reconstructed as one of the first studies that used and highlighted the importance of genome sequencing to complement the investigations of the COVID-

19 pandemic (52). Subsequently, other regional studies joined the initiative to perform NGS genomic sequencing of SARS-CoV-2 to analyze the introduction and emergence of new lineages together with the epidemiological dynamics, identifying at a general level the rapid dispersion and establishment of the lineage. B.1 lineage in the region (53–60) CHAPTER 4.

During 2021, SARS-CoV-2 continued to show a wide genetic diversity and in Colombia, viral microevolution and genetic variability were evidenced, allowing the identification of the emergence of autochthonous lineages with genetic determinants associated with increases in the transmission and escape of the humoral immune response. Initially, the B.1+L249S+E484K variant was described by phylogenetic analysis of a group of highly divergent sequences and 21 substitutions along the genome and of special attention the amino acid substitutions L249S and E484K located in the CTD and RBD of the Spike protein and that arose by evolutionary convergence (61) in different circulating lineages in that period of the pandemic (62). The E484K substitution was associated with decreased neutralizing antibodies (63) and increased ACE2 affinity (64) and the S249L substitution is located in the le terminal domain which frequently interacts with neutralizing antibodies (65).

Phylogenetic analysis showed a close relationship with the B.1.111 lineage and no recombination was evident. CHAPTER 5. Subsequently, micro-neutralization assays demonstrated the ability of this variant to evade neutralizing antibodies compare to ancestral variants circulating in Colombia (66). The frequency of detection of this new lineage decreased over time, without being established, whose most plausible explanation are evolutionary factors including the founder effect along with epidemiological factors including confinement measures, mandatory use of masks and social distancing established in Colombia.

Continuing genomic surveillance in the context of describing emerging variants of SARS-CoV-2, a new lineage was detected with a series of particular mutations in Spike protein, including I95I, Y144T, Y145S and the 146 N insertion in the N-terminal

domain, R346K, E484K and N501Y in the receptor binding domain (RBD) and P681H in the S1/S2 cleavage site of the spike protein which were evidenced in VOI and VOC independently. Phylogenetic and evolutionary analysis suggested that this was an emerging lineage with B.1 as parental and subsequently identified as B.1.621 in the PANGOLIN algorithm, adaptation analyses showed 9 codons under positive selection (67). B.1.621 lineage was detected in the north of the country (Caribbean Coast) since January, the spread to 14 departments, CHAPTER 6. and circulated in 58 countries around the world with a greater presence in Colombia, the Dominican Republic and Ecuador. On August 30, 2021, the B.1.621 lineage was recognized as the *Mu* variant, a new variant of interest (VOI) of SARS-CoV-2 by the World Health Organization (68). This variant dominated Colombia's third epidemiological peak, reaching about 73% of cases between July and September 2021 according to the genomic characterization of SARS-CoV-2 by probability sampling in all regions of the country. (69). Subsequently, studies associated infection by the *Mu* variant with severity, being found in 69% of fatal outcomes, and the increase in the frequency of cases was related to the dominant *Mu* variant during the third epidemiological peak in Colombia (70).

Phylogenetic and evolutionary analyses carried out in Colombia and in other countries of the region demonstrated the enormous genetic diversity and potential for adaptation of SARS-CoV-2 to selection pressures induced mainly by natural infection or vaccination (71) and its relationship with the emergence of emerging variants and catalogued as VOI and VOC (72).In Colombia, these findings were determinant to consolidate routine sampling based on epidemiological, virological and geographical criteria and a probabilistic strategy every six months of the total number of randomly selected samples based on reports from the national diagnostic database (69) and persuading national authorities to continue genomic surveillance of SARS-CoV-2 in real time to understand the dynamics of viral dispersion and evolutionary biology, as well as search for variants of interest and concern that could have implications for public health.

Methodological strategies for identification and genomic surveillance of emerging viruses with public health impact.

All molecular information obtained in the present study is decisive for the design, refinement, and validation of molecular tests used in viral detection and diagnosis of public health importance. It is also essential to consider the imminent risk of interaction between arboviruses of urban cycles and vertebrate vectors or reservoirs in sylvatic areas, which implies having robust detection methods to evaluate the ecology of these anthropophilic viruses to colonize potential new vectors. Therefore, the viral sequences identified in the NGS sequencing phase and those of viruses present in other countries of the region were raw material for the implementation of molecular tests including the design of oligonucleotides for the specific detection by RT-qPCR of *Culex flavivirus*, *Guachaca virus*, as well as the design of the control of transcription *in vitro* for the detection of YFV. CHAPTER 3 and CHAPTER 7 (73). The identification of viral agents with emergency potential was carried out by conventional RT-PCR, real-time RT-PCR, and multiplex RT-PCR tests, which have been designed in Colombia according to studies in which viral isolates of DENV, ZIKV, and CHIKV circulating in Colombia were made and which guarantee the incorporation of local and regional genetic variability (74,75).

Yellow fever virus was one of the most important urban arboviruses, through vaccination efforts and spraying of the *Ae. aegypti* mosquito (9) urban transmission was prevented; however, YFV adapted to infect sylvatic mosquitoes, such as *Sabethes spp.* and *Haemagogus spp*. establishing a sylvatic cycle in different countries of the Americas (19). Epizootics and human cases are reported annually in the endemic countries of the Americas, with the risk of Spillover. The most recent outbreak stands out in Brazil, during 2016-2020, with at least 2,240 human cases and 760 deaths (76). In 2022, three countries in the Region (Bolivia, Brazil, Peru) reported confirmed cases of yellow fever and of the 13 endemic countries in 9 countries vaccination coverage was less than 80% (77). The re-emergence of YFV in the region is imminent and it is necessary to contribute to routine surveillance for arbovirus research, diagnostic and epidemiological surveillance laboratories by

allowing access to positive controls for assay validation without relying on the availability of positive samples with variable viral loads that are depleted after successive use. Therefore, a plasmid control was developed and successfully standardized that allows *in vitro* RNA transcription and real-time PCR amplification of YFV for quantitative detection with two sets of oligonucleotides and probes used independently, determining the efficiency and analytical sensitivity of the real-time RT-PCR assay, a standard molecular amplification curve was generated from serial dilutions in base 10 of the RNA control (73). CHAPTER 7.

The metagenomic strategy through NGS sequencing, standardized and implemented in this study during 2019 from the biological collections carried out in the Sierra Nevada de Santa Marta was the technical and scientific precedent that allowed the co-executing entity of the National Institute of Health project to respond to the imperative need for genomic sequencing of SARS-CoV-2 with the objective of understanding the virus at the genetic and biological level. Within a year, the national genomic characterization program was set up to identify changes in the genetic diversity of SARS-CoV-2, monitor the epidemiological dynamics of the proportion, establishment and replacement of variants circulating in the country and contribute to the clarification of the epidemiological dynamics of COVID-19 (78)(79).

The genomic sequencing is necessity for monitoring in real time the circulation of different viruses of public health interest, generating technical guidelines, working protocols, inter-institutional collaboration agreements, transfer of sequencing reagents, strengthening the installed capacities and infrastructure of participating laboratories and training of scientific personnel in sequencing SARS-CoV-2 in countries of the Americas (80) (81) and at the national level, technical guidelines are issued on the design of protocols, diagnostic reports and genomic surveillance for public policy decision-making (69). Finally, a real time genomic surveillance in Colombia and in the region, it is a mandatory and complementary strategy to routine and sentinel virological surveillance to face outbreaks of unknown etiology and the goal of this thesis is a direct application with the World Health Organization Road Map for the control of Neglected Tropical Diseases, the road map for neglected

tropical diseases 2021–2030 and Global Genomic Surveillance Strategy for Pathogens with Pandemic and Epidemic Potential, 2022–2032.

15.6 References

1. Murphy FA, Nathanson N. The emergence of new virus diseases: an overview. Semin Virol [Internet]. 1994 Apr;5(2):87–102. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1044577384710109

2. Morse SS. The Public Health Threat of Emerging Viral Disease. J Nutr. 1997;127(5):951S–957S.

3. Perveen N, Muhammad K, Muzaffar S Bin, Zaheer T, Munawar N, Gajic B, et al. Host-pathogen interaction in arthropod vectors: Lessons from viral infections. Front Immunol [Internet]. 2023 Jan 31;14. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2023.1061899/full

4. Church DL. Major factors affecting the emergence and re-emergence of infectious diseases. Clin Lab Med [Internet]. 2004 Sep;24(3):559–86. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0272271204000642

5. Weaver SC, Reisen WK. Present and future arboviral threats. Antiviral Res [Internet]. 2010 Feb;85(2):328–45. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166354209004951

6. Howard CR, Fletcher NF. Emerging virus diseases: can we ever expect the unexpected? Emerg Microbes Infect [Internet]. 2012 Jul 25;1(1):1–9. Available from: https://www.tandfonline.com/doi/full/10.1038/emi.2012.47

7. Wang E, Ni H, Xu R, Barrett ADT, Watowich SJ, Gubler DJ, et al. Evolutionary Relationships of Endemic/Epidemic and Sylvatic Dengue Viruses. J Virol [Internet]. 2000 Apr;74(7):3227–34. Available from: https://journals.asm.org/doi/10.1128/JVI.74.7.3227-3234.2000

8. Barrett AD., Monath TP. Epidemiology and ecology of yellow fever virus. In 2003. p. 291–315. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065352703610079

9. Soper FL. The Elimination of Urban Yellow Fever in the Americas Through the Eradication of Aedes aegypti. Am J Public Heal Nations Heal [Internet]. 1963 Jan;53(1):7–16. Available from: https://ajph.aphapublications.org/doi/full/10.2105/AJPH.53.1.7

10. Furtado ND, Raphael L de M, Ribeiro IP, de Mello IS, Fernandes DR, Gómez MM, et al. Biological Characterization of Yellow Fever Viruses Isolated From Non-human Primates in Brazil With Distinct Genomic Landscapes. Front Microbiol [Internet]. 2022 Feb 14;13. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2022.757084/full

11. Kindhauser MK, Allen T, Frank V, Santhana RS, Dye C. Zika: the origin and spread of a mosquito-borne virus. Bull World Health Organ [Internet]. 2016 Sep 1;94(9):675-686C. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5034643/pdf/BLT.16.171082.pdf/

12. Weaver SC, Forrester NL. Chikungunya: Evolutionary history and recent epidemic spread. Antiviral Res [Internet]. 2015 Aug;120:32–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166354215001084

13. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature [Internet]. 2020 Mar 12;579(7798):270–3. Available from: https://www.nature.com/articles/s41586-020-2012-7

14. World Health Organization. Global genomic surveillance strategy for pathogens with pandemic and epidemic potential, 2022–2032 [Internet]. World Health Organization. 2022. Available from: https://www.who.int/publications/i/item/9789240046979

15. Batson J, Dudas G, Haas-Stapleton E, Kistler AL, Li LM, Logan P, et al. Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. Elife [Internet]. 2021 Apr 27;10. Available from: https://elifesciences.org/articles/68353

16. World Health Organization. Global genomic surveillance strategy for pathogens with pandemic and epidemic potential, 2022–2032. [Internet]. World Health Organization. 2022. Available from: https://apps.who.int/iris/handle/10665/352580

17. Roux S, Matthijnssens J, Dutilh BE. Metagenomics in Virology. In: Encyclopedia of Virology [Internet]. Elsevier; 2021. p. 133–40. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128096338209576

18. French RK, Holmes EC. An Ecosystems Perspective on Virus Evolution and Emergence. Trends Microbiol [Internet]. 2020 Mar;28(3):165–75. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0966842X19302653

19. Laiton-Donato K, Guzmán-Cardozo C, Peláez-Carvajal D, Ajami NJ, Navas M-C, Parra-Henao G, et al. Evolution and emergence of mosquito-borne viruses of medical importance: towards a routine metagenomic surveillance approach. J Trop Ecol [Internet]. 2023 Feb 9;39:e13. Available from: https://www.cambridge.org/core/product/identifier/S0266467423000019/type/journal_article

20. Charles J, Tangudu CS, Hurt SL, Tumescheit C, Firth AE, Garcia-Rejon JE, et al. Detection of novel and recognized RNA viruses in mosquitoes from the Yucatan Peninsula of Mexico using metagenomics and characterization of their in vitro host ranges. J Gen Virol [Internet]. 2018 Dec 1;99(12):1729–38. Available from: https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.001165

21. Oguzie JU, Nwangwu UC, Oluniyi PE, Olumade TJ, George UE, Kazeem A, et al. Metagenomic sequencing characterizes a wide diversity of viruses in field mosquito samples in Nigeria. Sci Rep [Internet]. 2022 May 10;12(1):7616. Available from: https://www.nature.com/articles/s41598-022-11797-2

22. Öhlund P, Hayer J, Lundén H, Hesson JC, Blomström A-L. Viromics Reveal a Number of Novel RNA Viruses in Swedish Mosquitoes. Viruses [Internet]. 2019 Nov 5;11(11):1027. Available from: https://www.mdpi.com/1999-4915/11/11/1027

23. Maia LMS, Pinto AZ de L, Carvalho MS de, Melo FL de, Ribeiro BM, Slhessarenko RD. Novel Viruses in Mosquitoes from Brazilian Pantanal. Viruses [Internet]. 2019 Oct 17;11(10):957. Available from: https://www.mdpi.com/1999-4915/11/10/957

24. Sadeghi M, Altan E, Deng X, Barker CM, Fang Y, Coffey LL, et al. Virome of > 12 thousand Culex mosquitoes from throughout California. Virology [Internet]. 2018 Oct;523:74–88. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0042682218302319

25. Wolfe ND, Dunavan CP, Diamond J. Origins of major human infectious diseases. Nature [Internet]. 2007 May;447(7142):279–83. Available from: http://www.nature.com/articles/nature05775

26. Dennehy JJ, Friedenberg NA, Holt RD, Turner PE. Viral Ecology and the Maintenance of Novel Host Use. Am Nat [Internet]. 2006 Mar;167(3):429–39. Available from: http://www.journals.uchicago.edu/doi/10.1086/499381

27. Rodríguez-Zapata MA, Ruiz-Agudelo CA. Environmental liabilities in Colombia: A critical review of current status and challenges for a megadiverse country. Environ Challenges [Internet]. 2021 Dec;5:100377. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2667010021003516

28. Bahassi EM, Stambrook PJ. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis [Internet]. 2014 Sep 1;29(5):303–10. Available from: https://academic.oup.com/mutage/article/29/5/303/1039175

29. Brejová B, Boršová K, Hodorová V, Čabanová V, Gafurov A, Fričová D, et al. Nanopore sequencing of SARS-CoV-2: Comparison of short and long PCR-tiling amplicon protocols. Abd El-Aty AM, editor. PLoS One [Internet]. 2021 Oct 29;16(10):e0259277. Available from: https://dx.plos.org/10.1371/journal.pone.0259277

30. Li K, Shrivastava S, Brownley A, Katzel D, Bera J, Nguyen AT, et al. Automated degenerate PCR primer design for high-throughput sequencing improves efficiency of viral sequencing. Virol J [Internet]. 2012 Dec 6;9(1):261. Available from: https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-9-261

31. Koehler JW, Hall AT, Rolfe PA, Honko AN, Palacios GF, Fair JN, et al. Development and Evaluation of a Panel of Filovirus Sequence Capture Probes for Pathogen Detection by Next-Generation Sequencing. Kuhn JH, editor. PLoS One [Internet]. 2014 Sep 10;9(9):e107007. Available from: https://dx.plos.org/10.1371/journal.pone.0107007

32. Ebel GD. Promiscuous viruses—how do viruses survive multiple unrelated hosts? Curr Opin Virol [Internet]. 2017 Apr;23:125–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1879625717300202

33. Weaver SC, Barrett ADT. Transmission cycles, host range, evolution and emergence of arboviral disease. Nat Rev Microbiol [Internet]. 2004 Oct;2(10):789–801. Available from: http://www.nature.com/articles/nrmicro1006

34. Vasilakis N, Tesh RB, Popov VL, Widen SG, Wood TG, Forrester NL, et al. Exploiting the Legacy of the Arbovirus Hunters. Viruses [Internet]. 2019 May 23;11(5):471. Available from: https://www.mdpi.com/1999-4915/11/5/471

35. Öhlund P, Lundén H, Blomström A-L. Insect-specific virus evolution and potential effects on vector competence. Virus Genes [Internet]. 2019 Apr 10;55(2):127–37. Available from: http://link.springer.com/10.1007/s11262-018-01629-9

36. Hameed M, Khan S, Xu J, Zhang J, Wang X, Di D, et al. Detection of Japanese encephalitis virus in mosquitoes from Xinjiang during next-generation sequencing arboviral surveillance. Transbound Emerg Dis [Internet]. 2021 Mar 15;68(2):467–76. Available from: https://onlinelibrary.wiley.com/doi/10.1111/tbed.13697

37. Xiao P, Li C, Zhang Y, Han J, Guo X, Xie L, et al. Metagenomic Sequencing From Mosquitoes in China Reveals a Variety of Insect and Human Viruses. Front Cell Infect Microbiol [Internet]. 2018 Oct 19;8. Available from: https://www.frontiersin.org/article/10.3389/fcimb.2018.00364/full

38. Guth S, Hanley KA, Althouse BM, Boots M. Ecological processes underlying the emergence of novel enzootic cycles: Arboviruses in the neotropics as a case study. Hol F, editor. PLoS Negl Trop Dis [Internet]. 2020 Aug 13;14(8):e0008338. Available from: https://dx.plos.org/10.1371/journal.pntd.0008338

39. Laiton-Donato K, Usme-Ciro JA, Rico A, Pardo L, Martínez C, Salas D, et al. Análisis filogenético del virus del chikungunya en Colombia: evidencia de selección purificadora en el gen E1. Biomédica [Internet]. 2015 Oct 23;36:25. Available from: http://www.revistabiomedica.org/index.php/biomedica/article/view/2990

40. Laiton-Donato K, Alvarez DA, Peláez-Carvajal D, Mercado M, Ajami NJ, Bosch I, et al. Molecular characterization of dengue virus reveals regional diversification of serotype 2 in Colombia. Virol J [Internet]. 2019 Dec 8;16(1):62. Available from: https://virologyj.biomedcentral.com/articles/10.1186/s12985-019-1170-4

41. Mendez JA, Usme-Ciro JA, Domingo C, Rey GJ, Sanchez JA, Tenorio A, et al. Phylogenetic history demonstrates two different lineages of dengue type 1 virus in Colombia. Virol J [Internet]. 2010 Dec 14;7(1):226. Available from: https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-7-226

42. Usme-Ciro JA, Mendez JA, Tenorio A, Rey GJ, Domingo C, Gallego-Gomez JC. Simultaneous circulation of genotypes I and III of dengue virus 3 in Colombia. Virol J [Internet]. 2008 Dec 2;5(1):101. Available from: https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-5-101

43. Black A, Moncla LH, Laiton-Donato K, Potter B, Pardo L, Rico A, et al. Genomic epidemiology supports multiple introductions and cryptic transmission of Zika virus in Colombia. BMC Infect Dis [Internet]. 2019 Dec 12;19(1):963. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-019-4566-2

44. Relich RF, Loeffelholz MJ. Taxonomic Changes for Human Viruses, 2020 to 2022. Humphries RM, editor. J Clin Microbiol [Internet]. 2023 Jan 26;61(1). Available from: https://journals.asm.org/doi/10.1128/jcm.00337-22

45. Cuevas JM, Moya A, Elena SF. Evolution of RNA virus in spatially structured heterogeneous environments. J Evol Biol [Internet]. 2003 May 17;16(3):456–66. Available from: https://onlinelibrary.wiley.com/doi/10.1046/j.1420-9101.2003.00547.x

46. Beerntsen BT, James AA, Christensen BM. Genetics of Mosquito Vector Competence. Microbiol Mol Biol Rev [Internet]. 2000 Mar;64(1):115–37. Available from: https://journals.asm.org/doi/10.1128/MMBR.64.1.115-137.2000

47. Hardy JL, Houk EJ, Kramer LD, Reeves WC. Intrinsic Factors Affecting Vector Competence of Mosquitoes for Arboviruses. Annu Rev Entomol [Internet]. 1983 Jan;28(1):229–62. Available from: https://www.annualreviews.org/doi/10.1146/annurev.en.28.010183.001305

48. World Health Organization. End-to-end integration of SARS-CoV-2 and influenza sentinel surveillance: revised interim guidance [Internet]. WHO. 2022. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-Integrated_sentinel_surveillance-2022.1

49. John G, Sahajpal NS, Mondal AK, Ananth S, Williams C, Chaubey A, et al. Next-Generation Sequencing (NGS) in COVID-19: A Tool for SARS-CoV-2 Diagnosis, Monitoring New Strains and Phylodynamic Modeling in Molecular Epidemiology. Curr Issues Mol Biol [Internet]. 2021 Jul 30;43(2):845–67. Available from: https://www.mdpi.com/1467-3045/43/2/61

50. European Centre for Disease Prevention and Control. Testing strategies for SARS-CoV-2 [Internet]. European Union. 2022. Available from: https://www.ecdc.europa.eu/en/covid-19/surveillance/testing-strategies

51. Mercer TR, Salit M. Testing at scale during the COVID-19 pandemic. Nat Rev Genet [Internet]. 2021 Jul 4;22(7):415–26. Available from: https://www.nature.com/articles/s41576-021-00360-w

52. Laiton-Donato K, Villabona-Arenas CJ, Usme-Ciro JA, Franco-Muñoz C, Álvarez-Díaz DA, Villabona-Arenas LS, et al. Genomic Epidemiology of Severe Acute Respiratory Syndrome Coronavirus 2, Colombia. Emerg Infect Dis [Internet]. 2020 Dec;26(12):2854–62. Available from: http://wwwnc.cdc.gov/eid/article/26/12/20-2969_article.htm

53. Castillo AE, Parra B, Tapia P, Acevedo A, Lagos J, Andrade W, et al. Phylogenetic analysis of the first four SARS-CoV-2 cases in Chile. J Med Virol [Internet]. 2020 Sep 8;92(9):1562–6. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jmv.25797

54. Márquez S, Prado-Vivar B, Guadalupe JJ, Gutierrez B, Becerra-Wong M, Jibaja M, et al. Metagenome of a Bronchoalveolar Lavage Fluid Sample from a Confirmed COVID-19 Case in Quito, Ecuador, Obtained Using Oxford Nanopore MinION Technology. Roux S, editor. Microbiol Resour Announc [Internet]. 2020 Oct 8;9(41). Available from: https://journals.asm.org/doi/10.1128/MRA.00996-20

55. Garcés-Ayala F, Araiza-Rodríguez A, Mendieta-Condado E, Rodríguez-Maldonado AP, Wong-Arámbula C, Landa-Flores M, et al. Full genome sequence of the first SARS-CoV-2 detected in Mexico. Arch Virol [Internet]. 2020 Sep 18;165(9):2095–8. Available from: https://link.springer.com/10.1007/s00705-020-04695-3

56. Franco D, Gonzalez C, Abrego LE, Carrera J-P, Diaz Y, Caicedo Y, et al. Early Transmission Dynamics, Spread, and Genomic Characterization of SARS-CoV-2 in Panama. Emerg Infect Dis [Internet]. 2021 Feb;27(2):612–5. Available from: https://wwwnc.cdc.gov/eid/article/27/2/20-3767_article.htm

57. Padilla-Rojas C, Lope-Pari P, Vega-Chozo K, Balbuena-Torres J, Caceres-Rey O, Bailon-Calderon H, et al. Near-Complete Genome Sequence of a 2019 Novel Coronavirus (SARS-CoV-2) Strain Causing a COVID-19 Case in Peru. Roux S, editor. Microbiol Resour Announc [Internet]. 2020 May 7;9(19). Available from: https://journals.asm.org/doi/10.1128/MRA.00303-20

58. Padilla-Rojas C, Jimenez-Vasquez V, Hurtado V, Mestanza O, Molina IS, Barcena L, et al. Genomic analysis reveals a rapid spread and predominance of lambda (C.37) SARS-COV-2 lineage in Peru despite circulation of variants of concern. J Med Virol [Internet]. 2021 Dec 20;93(12):6845–9. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jmv.27261

59. Juscamayta-López E, Carhuaricra D, Tarazona D, Valdivia F, Rojas N, Maturrano L, et al. Phylogenomics reveals multiple introductions and early spread of SARS-CoV-2 into Peru. J Med Virol [Internet]. 2021 Oct 16;93(10):5961–8. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jmv.27167

60. Elizondo V, Harkins GW, Mabvakure B, Smidt S, Zappile P, Marier C, et al. SARS-CoV-2 genomic characterization and clinical manifestation of the COVID-19 outbreak in Uruguay. Emerg Microbes Infect [Internet]. 2021 Jan 1;10(1):51–65. Available from: https://www.tandfonline.com/doi/full/10.1080/22221751.2020.1863747

61. Ferrareze PAG, Franceschi VB, Mayer A de M, Caldana GD, Zimerman RA, Thompson CE. E484K as an innovative phylogenetic event for viral evolution: Genomic analysis of the E484K spike mutation in SARS-CoV-2 lineages from Brazil. Infect Genet Evol [Internet]. 2021 Sep;93:104941. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134821002380

62. Laiton-Donato K, Usme-Ciro JA, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Reales-González J, et al. Novel Highly Divergent SARS-CoV-2 Lineage

With the Spike Substitutions L249S and E484K. Front Med [Internet]. 2021 Jun 28;8. Available from: https://www.frontiersin.org/articles/10.3389/fmed.2021.697605/full

63. Jangra S, Ye C, Rathnasinghe R, Stadlbauer D, Krammer F, Simon V, et al. SARS-CoV-2 spike E484K mutation reduces antibody neutralisation. The Lancet Microbe [Internet]. 2021 Jul;2(7):e283–4. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2666524721000689

64. Laffeber C, de Koning K, Kanaar R, Lebbink JHG. Experimental Evidence for Enhanced Receptor Binding by Rapidly Spreading SARS-CoV-2 Variants. J Mol Biol [Internet]. 2021 Jul;433(15):167058. Available from: https://linkinghub.elsevier.com/retrieve/pii/S002228362100276X

65. Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med [Internet]. 2021 Apr 2;27(4):622–5. Available from: http://www.nature.com/articles/s41591-021-01285-x

66. Álvarez-Díaz DA, Laiton-Donato K, Torres-García OA, Ruiz-Moreno HA, Franco-Muñoz C, Beltran MA, et al. Reduced levels of convalescent neutralizing antibodies against SARS-CoV-2 B.1+L249S+E484K lineage. Virus Res [Internet]. 2022 Jan;308:198629. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168170221003361

67. Laiton-Donato K, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Usme-Ciro JA, Prada DA, et al. Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2. Infect Genet Evol [Internet]. 2021 Nov;95:105038. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134821003361

68. World Health Organization. Tracking SARS-CoV-2 variants [Internet]. World Health Organization. 2023. Available from: https://www.who.int/activities/tracking-SARS-CoV-2-variants

69. Instito Nacional de Salud. Documentos técnicos genómica [Internet]. Instituto Nacional de Salud. 2022. Available from: https://www.ins.gov.co/Paginas/Documentos-tecnicos-genomica.aspx

70. Álvarez-Díaz DA, Ruiz-Moreno HA, Zapata-Bedoya S, Franco-Muñoz C, Laiton-Donato K, Ferro C, et al. Clinical outcomes associated with Mu variant infection during the third epidemic peak of COVID-19 in Colombia. Int J Infect Dis [Internet]. 2022 Dec;125:149–52. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1201971222005665

71. López-Cortés GI, Palacios-Pérez M, Veledíaz HF, Hernández-Aguilar M, López-Hernández GR, Zamudio GS, et al. The Spike Protein of SARS-CoV-2 Is Adapting Because of Selective Pressures. Vaccines [Internet]. 2022 May 28;10(6):864. Available from: https://www.mdpi.com/2076-393X/10/6/864

72. Centers for Disease Control and Prevention CDC. SARS-CoV-2 Variant Classifications and Definitions [Internet]. CDC. 2023. Available from: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html

73. Laiton-Donato K, Quintero-Cortés P, Franco-Salazar JP, Peláez-Carvajal D, Navas M-C, Junglen S, et al. Usefulness of an in vitro-transcribed RNA control for the detection and quantification of Yellow fever virus through real-time reverse transcription-polymerase chain reaction. Infect Dis Now [Internet]. 2023 Apr;53(3):104654. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2666991923000167

74. Álvarez-Díaz DA, Quintero PA, Peláez-Carvajal D, Ajami NJ, Usme-Ciro JA. Novel pan-serotype control RNA for dengue virus typing through real-time reverse transcription-polymerase chain reaction. J Virol Methods [Internet]. 2019 Sep;271:113677. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166093419300527

75. Álvarez-Díaz DA, Valencia-Álvarez E, Rivera JA, Rengifo AC, Usme-Ciro JA, Peláez-Carvajal D, et al. An updated RT-qPCR assay for the simultaneous detection and quantification of chikungunya, dengue and zika viruses. Infect Genet Evol [Internet]. 2021 Sep;93:104967. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134821002641

76. Sacchetto L, Silva NIO, Rezende IM de, Arruda MS, Costa TA, de Mello ÉM, et al. Neighbor danger: Yellow fever virus epizootics in urban and urban-rural transition areas of Minas Gerais state, during 2017-2018 yellow fever outbreaks in Brazil. Carvalho MS, editor. PLoS Negl Trop Dis [Internet]. 2020 Oct 5;14(10):e0008658. Available from: https://dx.plos.org/10.1371/journal.pntd.0008658

77. Organización Panamericana de la Salud, Organización Mundial de la Salud. Alerta Epidemiológica: Fiebre amarilla - 31 de agosto de 2022 [Internet]. OPS/OMS. 2022. Available from: https://www.paho.org/es/documentos/alerta-epidemiologica-fiebre-amarilla-31-agosto-2022

78. Instito Nacional de Salud. COVID-19 en Colombia [Internet]. Instituto Nacional de Salud. 2023. Available from: https://www.ins.gov.co/Noticias/Paginas/coronavirus-genoma.aspx

79. Laiton-Donato K, Álvarez-Díaz DA, Franco-Muñoz C, Ruiz-Moreno HA, Rojas-Estévez P, Prada A, et al. Monkeypox virus genome sequence from an imported human case in Colombia. Biomédica [Internet]. 2022 Sep 2;42(3):541–5. Available from: https://revistabiomedica.org/index.php/biomedica/article/view/6647

80. Álvarez-Díaz DA, Laiton-Donato K, Franco-Muñoz C, Mercado-Reyes M. Secuenciación del SARS-CoV-2: la iniciativa tecnológica para fortalecer los sistemas de alerta temprana ante emergencias de salud pública en Latinoamérica y el Caribe. Biomédica [Internet]. 2020 Oct 30;40(Supl. 2):188–97. Available from: https://revistabiomedica.org/index.php/biomedica/article/view/5841

81. Yek C, Pacheco AR, Vanaerschot M, Bohl JA, Fahsbender E, Aranda-Díaz A, et al. Metagenomic pathogen sequencing in resource-scarce settings: Lessons learned and the road ahead. Front Epidemiol [Internet]. 2022 Aug 15;2. Available from: https://www.frontiersin.org/articles/10.3389/fepid.2022.926695/full

16. CONCLUSIONS

RNA viruses still represent a challenge in research and the new viral discoveries through mNGS are relevant in the description of the different ecological and evolutionary scenarios to face the emergence towards the human population.

The mNGS characterization allowed us to define the general panorama of circulating viruses in the sylvatic, rural, and urban of the Sierra Nevada de Santa Marta and was a significant contribution to demonstrate a high diversity of viruses and understand the evolutionary processes.

The *in vitro* virological assays were complementary to the genomic characterization and constitute the basis for evaluating the adaptation potential of the viruses found in hematophagous mosquito species.

The phylogenetic and evolutionary mechanisms that shaped the genetic variability of SARS-CoV-2, the distribution, establishment and replacement of lineages and variants were described and contributed to the generation of real-time genomic data in Colombia.

Genomic data was essential in the design, and validation of molecular tests, positive controls, and implementation of methodological strategies for genomic surveillance of emerging viruses.

17. PERSPECTIVES

Continuing mNGS studies in mosquitoes, vertebrate hosts, and human that inhabit different Colombian ecosystems is essential to contribute to the ecological and biological knowledge of viruses with emerging potential and their future implications.

Genomic characterization of different viral species must be accompanied by *in silico*, *in vitro*, and *in vivo* assays to biologically support the findings and understand the factors that contribute to viral emergence.

Studies evaluating the evolutionary mechanisms and genetic diversity of SARS-CoV-2 will continue to be essential to assess transmissibility, pathogenesis, vaccine design and refinement, and molecular tests.

It is necessary to define a comprehensive strategy that includes technical-scientific training, financing, and adherence to genomic characterization programs in Colombia and the Americas region through collaboration networks.

It is essential to implement NGS detection methods in the diagnosis of unknown etiologies, in the search for new viruses and the characterization during public health alarms.

18. LEVEL OF ORIGINALITY

The thesis is an original research project in which, for the first time, was performed the sampling of several mosquito species and the identification of the virodiversity of a tropical rainforest ecosystem in Colombia, including sylvatic, rural and urban transects that represent the general panorama of the emerging viral potential are highlighted. In addition, new viruses were discovered among them *Guachaca virus* including the genetic and *in vitro* characterization to establish the emerging potential.

The methodological approaches at the level of NGS sequencing established in the thesis allowed sequencing the first cases of SARS-CoV-2 in Colombia and subsequently establishing the routine genomic characterization of an emerging RNA virus, providing original genomic data and highlighting the genetic characterization of emerging variants among them *Mu* with substitutions of biological relevance and of national and international epidemiological scope.

For Colombia, the development of the thesis allowed the consolidation of new NGS sequencing methodologies, phylogenetic and evolutionary analysis, allowing the dissemination of the findings in real time to decision makers in public health.

Molecular tests were designed and implemented for the detection and characterization of specific insect and arbovirus viruses that can be freely used in research or diagnostic laboratories contributing to routine virological surveillance of yellow fever.

Overall, the thesis work describes a strategy of methodological development and implementation, and effective scientific divulgation for the emerging and re-emerging potential of RNA viruses of interest at sylvatic, rural, and urban levels and with public health implications.