VIRAL DETECTION AND CHARACTERIZATION IN FREE-RANGING AQUATIC TURTLES IN THE UNITED STATES

by

WEERAPONG LAOVECHPRASIT

(Under the Direction of James B. Stanton)

ABSTRACT

Aquatic turtles, many of which are threatened or endangered, face numerous threats, including anthropogenic interaction, ocean pollution, climate change, and infectious agents. Viruses have been implicated in severe disease outbreaks in other endangered wildlife, such as outbreaks of canine distemper virus in black-footed ferrets and ranavirus-induced mass mortality events in multiple amphibian species. However, the threat of viruses on freshwater and marine turtle populations is unknown, largely due to the limited knowledge regarding viruses found in aquatic turtles. Notably, the aquatic environment, both freshwater and marine, is rich in viral diversity, most of which remains to be discovered and characterized. Thus, there is a need to fill this knowledge gap.

Metagenomic sequencing has revolutionized viral detection and characterization. Metagenomic sequencing has been used as a surveillance tool for unveiling novel viral taxa in diverse ecosystems and as an agnostic diagnostic test for the identification of pathologic viral infections in a wide array of species. Third-generation sequencing makes metagenomic sequencing more accessible due to reduced costs, methodologic flexibility, and simplified

bioinformatic analysis. This allows for metagenomic sequencing to be applied to research fields that are historically less funded, but are ecologically important, such as endangered reptiles.

This dissertation has two aims. First to use efficient third-generation sequencing technology to survey for known and unknown viruses in aquatic turtles (Chapter 1). This data can then be used in the future for determining their impact on the health of endangered aquatic turtles and improve the understanding of viral evolution, genetics, and ecology. In the second aim, molecular Koch's postulates are applied to aquatic turtle meningoencephalitis to discover the likely cause of that disease (Chapter 2). This work collectively identified and characterized twelve novel viral species in this study (eleven in the gastrointestinal tract and one in the central nervous system with spatial localization of mRNA within the affected tissues).

The foundational insight of viral discovery and diversity presented in this dissertation will lay a groundwork for future hypothesis-driven research and underline the imperative need to comprehend viral diversity in understudied species such as endangered aquatic turtles.

INDEX WORDS: Aquatic turtles, Virus, Metagenomics, MinION sequencing, Gastrointestinal system, Central nervous system, Meningoencephalomyelitis

VIRAL DETECTION AND CHARACTERIZATION IN FREE-RANGING AQUATIC TURTLES IN THE UNITED STATES

by

WEERAPONG LAOVECHPRASIT

DVM, Kasetsart University, Thailand, 2015

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2024

VIRAL DETECTION AND CHARACTERIZATION IN FREE-RANGING AQUATIC TURTLES IN THE UNITED STATES

by

WEERAPONG LAOVECHPRASIT

Major Professor:	James Stanton
Committee:	Alvin Camus
	Elizabeth Howerth
	Kathleen Colegrove

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2024

ACKNOWLEDGEMENTS

First and foremost, I would like to sincerely thank my powerful mother, a greengrocer in a local fresh market in Thailand and the most important person of my life, who has inspired me, since the day I was adopted, to be passionate in the nature and aquatic world. She proved that no matter how many life resources you were born with, you can still be successful if you stay passionate, determined and consistently hard-working throughout the impossible-looking journey. I would also like to thank the Fulbright Thailand who selected me to be a grantee for the Thai Graduate Scholarship and trusted this ordinary, local marine-life veterinarian with the life-changing opportunity. Without all of you, I would not have been able to have come this far.

For this dissertation, I would like to thank Kelsey Young, my very first and official laboratory mentor for giving me a perfect benchwork and bioinformatic training; my research collaborators (Drs. Brian Stacy, Heather Harris, Allison Tuttle, Adam Kennedy, Terry Norton, Robet Ossiboff, Kuttichantran Subramaniam and Mr/Mrs Kennedy, Sirpensky, and Zirkelbach) for providing these valuable research specimens and constant helps me throughout my PhD adventure; my research committee (Drs. Jeremiah Saliki, Alvin Camus, Kathleen Colegrove, and Elizabeth Howerth) for continuously providing me feedback and shaping me to be a qualified researcher. I would like to assure you all that I will maintain my passion and work towards veterinary virology, anatomic pathology, and wildlife conservation as much as possible.

Most importantly, I would like to thank Dr. James Brantly Stanton, my PhD mentor, for tirelessly training me with wisdom in science, molecular biology, philosophy, and life. He is a true example of how to be a good role model for a graduate trainer, an ethical researcher, and a

decent colleague. Without him, I would not have been able to walk on this path with good judgment, become a metagenomic scientist, and produce meaningful work toward marine life conservation.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Aquatic testudines, a large group of aquatic and marine reptiles encompassing freshwater turtles (e.g., family *Chelydridae*) and sea turtles (i.e., families *Dermochelyidae* and *Cheloniidae*) (Wyneken et al., 2013), inhabit a wide range of ecosystems including rivers, lakes, estuaries, coastal shorelines, seagrass beds, coral reefs, and pelagic zones, where they interact with a wide array of invertebrates, fishes, amphibians, reptiles, and marine mammals (Byles, 1988). The global distribution of aquatic turtles spans across seven biogeographical regions (i.e., North America [Nearctic], Central America [the Caribbean and northern Neotropical], South America [Southern Neotropical], Mediterranean [western Palearctic], Sub-Saharan Africa [Afrotropical], Asia [Oriental and eastern Palearctic], and Australia [Australasia]) (Buhlmann et al., 2009). This global distribution means that people across the world are aware of and interact with these animals. Additionally, their unique evolutionary features such as a primitive skull, a keratinized protective shell (carapace and plastron), and their status as "living fossils" (i.e., lack of significant morphologic changes since the time of dinosaurs) drive public interest (Joyce, 2015; Lyson et al., 2016). Apart from their unique biology and public interest, aquatic testudines play many important roles in the aquatic ecosystems, such as improving water quality by consuming fish carcasses, maintaining seagrass health by grazing on seagrass overgrowth, and interchanging the diversity of aquatic sessile species by carrying epibiota during the migration (Boyle et al., 2009).

Notably, more than half of aquatic turtle species are classified as vulnerable to critically endangered on the IUCN red list due to natural predation, harmful algal blooms, habitat degradation, environmental pollution, anthropogenic interference, and infectious diseases (Crawford et al., 2014). Of those threats, several infectious pathogens are commonly reported as a cause of disease in multiple species of turtles. These include fungi, such as *Emydomyces testavorans* causing epithelial inclusion cysts in freshwater turtles (Woodburn et al., 2021), and *Pseudallescheria ellipsoidea* or *Fusarium solani* causing hatching failure in sea turtles (Chai et al., 2023). Larger organisms, such as spirorchiid trematodes (*Neospirorchis*, *Hapalotrema*, *Learedius*, and *Carettacola* spp.) can cause multi-organ granulomatous vasculitis in green sea turtles [*Chelonia mydas*], and loggerhead sea turtles [*Caretta caretta*] (Chapman et al., 2017).

However, out of the many infectious agents that affect aquatic testudines, only a few viruses are known to have pathologic associations. Herpesvirids (*Alphaherpesvirinae*, *Herpesviridae*) are the most reported with disease association in both freshwater and sea turtles. In freshwater turtles, cheloniid herpesvirus 2, 3, and 4 have been associated with hepatic necrosis in Pacific pond turtles (Frye et al., 1977), hepatic necrosis in painted and map turtles (Cox & Barker, 1980), and necrotizing stomatitis in Argentinian tortoises (Jacobson et al., 1985), respectively. In marine turtles cheloniid herpesvirus 1, 5, and 6 have been associated with grey patch disease in cultured sea turtles, contagious fibropapillomatosis in free-ranging sea turtles (Whitmore et al., 2021), and lung-eye-tracheal disease in cultured sea turtles (Manire et al., 2017). Two less characterized herpesviruses have also been identified in sea turtles: loggerhead genital-respiratory herpesvirus associated with cloacal-genital ulceration in loggerhead turtles (Stacy et al., 2008), and loggerhead orocutaneous herpesvirus associated with tongue ulceration in loggerhead turtles (Stacy et al., 2008). Moreover, adenovirids (*Adenoviridae*) from three

genera have been reported with disease association. An atadenovirus was detected through PCR in a spur-thighed tortoise with stomatitis and esophagitis (Rivera et al., 2009). Siadenoviruses have been detected through PCR in Sulawesi tortoises with ulcerative stomatitis (Rivera et al., 2009) and in a separate study, they were also detected in impressed tortoises and a Burmese star tortoise with multisystem inflammation (duodenitis, cholangiohepatitis, myocarditis, and encephalitis) (Schumacher et al., 2012). The third adenovirid was a testadenovirus, which was detected through PCR in an ornate box turtle with hepatic degeneration and lymphocytic infiltration in hepatic sinusoids (Farkas & Gál, 2009). RNA viruses have also been associated with disease in turtles, most notably a picornavirus (*Tochivirus, Picornaviridae*, torchivirus A1), was recently associated with kidney disease and shell weakness syndrome in European tortoises (Paries et al., 2019).

Conventionally, detection of viruses was predominantly dependent on gross findings and histopathologic interpretation (severity and distribution of pathologic changes, mononuclear cell inflammation, and inclusion bodies), which suggested to pathologists that more in-depth ancillary tests were needed (Galindo-Cardiel et al., 2013). These additional tests could be broad, screening tests (e.g. electron microscopy [morphology-based microscopic detection] (Cocumelli et al., 2018) and viral isolation (Perlejewski et al., 2020), while others are agent-specific tests. Specific tests include antibody-based tests such as immunohistochemistry (IHC) (Di Guardo et al., 2013) and nuclei-acid based tests such as in-situ hybridization (ISH) (Leigh Perkins et al., 2001) and polymerase chain reaction (PCR) (Watzinger et al., 2006). Both broad and pathogen-specific approaches are widely used and accurately diagnose human and veterinary medicine. However, EM lacks genetic characterization capability, and the more targeted methods are mostly restricted to pre-existing viruses because they require prior knowledge, e.g. viral

nucleotide sequences (ISH and PCR), availability of permissible cell lines, or production of viral antibodies (IHC and serology). The ability to detect unknown viruses has greatly expanded in the era of metagenomics, with the ability to employ random sequencing. For example, viruses within the order Jingchuvirales were only just recently discovered, even though there is strong evidence that these viruses have existed in vertebrates for thousands of years (Dezordi et al., 2020). However, the great genetic distance of Jingchuvirales to more common viral orders such as Mononegavirales (monopartite negative-sense RNA viruses) and Picornavirales (monopartite and bipartite positive-sense RNA viruses), made Jingchuvirales undiscovered until 2015 (Li et al., 2015). In addition to significant genetic distance of jingchuvirals, there is no evidence that jingchuviral infection results in virus-specific pathologic effects (e.g., Guǎngdōng red-banded snake chuvirus-like in a red-banded snake and Wēnlǐng fish chu-like virus in a long-sniped fish), the ability of a virus-specific test, like PCR, to accidentally detect a jingchuviral as an off-target result is very limited. Thus, these were only detected through random metagenomic sequencing. While this is a powerful method, to more fully apply this to wildlife and other areas of veterinarian medicine, methods to improve cost efficiency are required.

Generally, sequencing platforms have been categorized into three different generations, first-sequencing (Sanger sequencing), second-generation sequencing (high throughput, short-read sequencing [e.g., Illumina, Roch 454, Ion torrent]), and third-generation sequencing (high-throughput long-read sequencing [e.g. Oxford Nanopore Technology, Pacific Bioscience]). Sanger sequencing, invented by Federick Sanger, combines the amplification of fluorescence-tagged dideoxynucleotides with gel electrophoresis to determine the order of nucleotide sequences. This assay provides high accuracy but is low throughput (i.e., small number of short DNA reads per sequencing run) (Carpente et al., 2019). Second-generation or next-generation

sequencing was invented to overcome the throughput limitation of the Sanger sequencing by massive amplification of short DNA sequences. Even though this platform provides highly accurate and massive output, the amplification of short DNA fragments results in complicated post-sequencing analysis and massive computational storage (Datta, 2015). MinION sequencing is a portable, real-time, third-generation sequencing device that utilizes membrane-embedded, pore-forming nanoproteins to sequence single, intact RNA or DNA molecules (Lu et al., 2016). As the nucleic acid molecules pass through the nanopores, the nucleotides alter the transmembrane electrical charge, and this change in the electrical signal can be converted (i.e., basecalled) into each base (A, T, C, and G) (Quick et al., 2017). This platform allows for the sequencing of long nucleic acid molecules, which simplifies bioinformatic assembly. Additionally, basecalling can be performed in real-time with sequencing, subsequently enhancing time- and cost-effectiveness (Butt et al., 2021). Due to the real-time sequencing and low initial cost-of-investment, MinION sequencing has been used to perform long-read sequencing in vastly different laboratory settings (Buysse et al., 2024; Castro-Wallace et al., 2017). Due to these efficiencies, MinION sequencing is a strategy that fits well in areas in which research funding and diagnostic laboratories are limited, particularly in wildlife and aquatic animal fields (Quick et al., 2017). Moreover, MinION sequencing can provide a semi-deep sequencing result, which again helps to minimize costs, minimize computer storage, and eases bioinformatic analysis (of benefit to researchers with limited bioinformatic experience as compared to other sequencing platforms).

As compared to the conventional PCR coupled with Sanger sequencing, high-throughput sequencing (both second- and third-generation sequencing) allows for the creation of randomly primed sequencing libraries (no specific forward and/or reverse primers, a.k.a., random or

unbiased sequencing) (Brown et al., 2018; Y. Z. Zhang et al., 2019). This technology provides a massive shift in the capabilities for the diagnosis of viral diseases, particularly diseases that were initially considered idiopathic (Xing et al., 2020). Recently, several viruses have been discovered through the application of metagenomic sequencing, e.g., SARS-CoV2 (*Sarbecovirus, Coronaviridae*), the coronaviral cause of the pandemic outbreak of severe acute respiratory syndrome in primates, captive and wild felids, and cervids (Tyson et al., 2020). Metagenomic sequencing has revealed additional SARS-CoV2 viral variants and mutations through the routine monitoring using similar platforms (Capraru et al., 2022). Additionally, the recently discovered turtle fraser viruse 1 (*Fraservirus, Tosoviridae*) was metagenomically identified in freshwater turtles (Florida softshell turtles [*Apalone ferox*], Florida red-bellied cooters (*Pseudemys nelsoni*), and a peninsula cooter [*Pseudemys peninsularis*]) with heterophilic/histiocytic meningoencephalitis and multi-organ vasculitis (Waltzek et al., 2022).

To apply MinION sequencing (third-generation sequencing platform) for phylogenetic characterization of viral agents, several molecular approaches can be used. Fully targeted sequencing, akin to a typical conventional PCR relying on 2 specific primers (forward and reverse primers), can be used (i.e., amplicon sequencing, a.k.a., ampseq). Disadvantages for this procedure require prior knowledge of the targeted sequence, it requires that the nucleic acid integrity must be sufficient to allow for amplification of that target size (i.e., if the product size is 1000 bases, then the RNA must be sufficiently intact to retain 1000-base fragments) (Figure 1), its characterization capability is limited to only the targeted region, and it requires that the two required primer-binding sites be conserved enough to prevent target:primer mismatches over a total of 40 bases (Yang & Rothman, 2020). Whereas random sequencing utilizes random hexamers as primers, which mitigates the issue of specific-region amplification of targeted

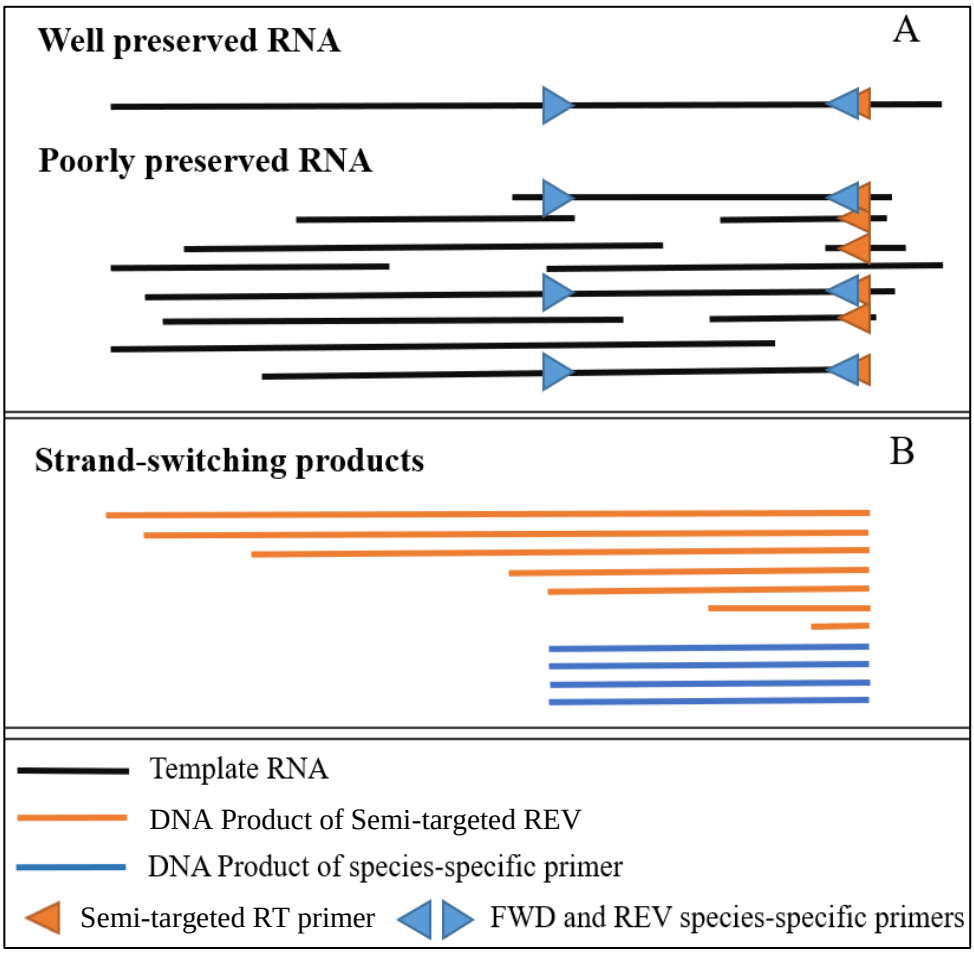
sequencing. However, this approach provides the amplification of not only viral nucleic acids, but also all nucleic acid in the sample of interest (mostly host and bacterial ribosomal RNA); thus, there is no viral enrichment. While this random sequencing has been employed for virus characterization, the random nature of the reaction requires deeper sequencing to obtain more viral characterization data, and thus the cost is increased, as compared to targeted sequencing. To alleviate this disadvantage of random sequencing, viruses detected through random sequencing can be more fully characterized after a semi-targeted sequencing approach. The semi-targeted MinION sequencing approach uses a single, virus-targeted reverse primer binding site along with the previously published strand-switching reaction (Young et al., 2021) to create a defined forward primer binding site (akin to rapid amplification of cDNA ends [RACE]), can then extend into previously uncharacterized genomic regions. This semi-targeted research strategy mitigates the issues of both fully targeted and random sequencing resulting in longer phylogenetic information. Because a single specific primer will provide lower target specificity, semi-deep sequencing can also identify target versus off target reads (Figure 1).

With a large knowledge gap in research concerning viral diversity and infection in aquatic testudines, the main goal of dissertation is to apply MinION sequencing for the detection and characterization of viruses on free-ranging aquatic turtles. It is hypothesized that free-ranging aquatic turtles carry or are affected by either known and/or undiscovered viruses. This dissertation has two major objectives to fill the existing gap. The first objective is to develop random and semi-targeted MinION sequencing methods to detect and characterize undiscovered viruses in the gastrointestinal system of free-ranging sea turtles. The second objective is to determine the cause of idiopathic mononuclear meningoencephalomyelitis in five free-ranging aquatic turtles using MinION sequencing and RNAscope in-situ hybridization. These studies

demonstrate that MinION sequencing can detect and characterize a variety of both known and undiscovered viruses in free-ranging aquatic turtles with cost-, time-, and bioinformatic skill-effectiveness.

Figure and Table

Figure 1: The benefit of single primer strand-switching



A: A representation of the impact of RNA integrity strand-switching (orange arrowhead) and routine, species-specific PCR (blue arrowhead).

B: The template for the strand-switching reaction is more numerous than templates that have the forward and reverse primer binding sites that are required for routine PCR. Additionally, if the RNA is sufficiently intact, then longer sequences can be obtained with strand-switching, whereas PCR has a defined product length based entirely on the primer binding sites.

CHAPTER 2

SURVEILLANCE OF GASTROINTESTINAL VIRUSES IN FREE-RANGING TURTLES IN THE UNITED STATES

Laovechprasit W, Stacy BA, Harris HS, Tuttle AD, Sirpenski G, Kennedy AE, Innis CJ, Norton TM, Zirkelbach B, and Stanton JB. To be submitted to a peer-reviewed journal (Virus Evolution).

Abstract

Knowledge of the virosphere is rapidly expanding due to technological advances, but understudied areas exist, including viruses to be found in sea turtles. Sea turtles, aquatic reptiles with six out of seven species classified as vulnerable, endangered, or critically endangered on the IUCN Red List, have been impacted by various threats including infectious diseases. To pave the way for understanding the effects of viruses on sea turtle health, it is crucial to establish baseline knowledge of the viruses present within their gastrointestinal tracts, and doing so will also provide important information regarding viral diversity. Viruses are extensively found in the gastrointestinal tract and can be identified as a cause of gastrointestinal disease in many species, but little is known about enteric viruses in these endangered sea turtles; thus, this study metagenomically sequenced 78 gastrointestinal specimens from six species of sea turtles with various health conditions from different locations across the United States (West Coast [CA, OR, and WA], Northeastern Coast [MA], and Southeastern Coast [GA and FL]. Viral sequences were assessed at two levels: 1) detection and taxon assignment, which was defined as at least 5 reads aligning via BLAST to the same viral taxon and 2) phylogenetic characterization, which was defined as at least 10× coverage of the virus-appropriate genomic region for species or genus demarcation. Nucleic acid sequences of 48 viruses from at least eight known viral families and three potential novel families were detected in 19 individuals across several turtle species, geographical regions, and health conditions. For a subset of detected viruses, viral reads and contigs from random sequencing were confirmed and sequencing metrics were improved through subsequent semi-targeted strand-switching MinION sequencing, allowing for phylogenetic

characterization. As a result, 19 viruses across at least seven known viral families were phylogenetically characterized. These include 11 novel viral species and three putative novel genera. Sixteen RNA viruses were phylogenetically characterized, which included four double-stranded RNA viruses (*Partitiviridae, Orthototiviridae,* and *Picobirnaviridae*), 11 positive-sense single-stranded RNA viruses (*Caliciviridae, Dicistroviridae,* and unclassified Hepelivirales*),* and one negative-sense bisegmented RNA virus (*Chuviridae*). Although RNA viruses were the target in this investigation, three DNA viruses were also characterized (*Parvoviridae, Circoviridae,* and unclassified Cressdnaviricota). Viruses detected in sea turtle gastrointestinal tracts often had closest relatives that were previously detected from invertebrates and fish. This included salmon calicivirus (host: salmonids), sea star-associated densovirus (host: echinoderms), halhan virus (host: bivalves), and Wenling crustacean virus 9 (host: crustaceans). Although these viruses could be passing through the gastrointestinal tract of turtles without any clinical significance (e.g., ingesta-associated), this study expands the knowledge of highly abundant viruses present in the gastrointestinal tracts of sea turtles, characterizes several putatively novel species, and serves as a basis for future hypothesis-driven research to understand their relevance to the health of sea turtles and their ecosystems.

Introduction

Recent advances in nucleic acid detection have resulted in exponential virus discovery. For example, in 2002, there were 1619 formally recognized viral species; in 2018 there were 5560; and in 2023 there were 14,690 recognized species (https://ictv.global/taxonomy/history, accessed June 27, 2024). The marine environment (particularly estuaries, shorelines, and open oceans) has been a recent target of viral discovery, resulting in numerous newly discovered viruses, but much remains undetermined about the virosphere of this vast environment and the animals that inhabit it (Coutinho et al., 2017; Edgar, 2023; Wolf et al., 2020). Some of the novel viruses (e.g., viruses under *Marnaviridae* and *Aoguangviridae*) detected through virosphere surveillance, including from the marine environment (Lang et al., 2021; Zhou et al., 2023), were genetically distinct from known viral taxa (Lang et al., 2021; Zhou et al., 2023), resulting the creation of new viral orders, which have increased from fourteen in 2018 to eighty-one in 2023.

Detection of such genetically diverse viruses required molecular tools other than conventional PCR, which is reliant on primer-target specificity. Metagenomic/random sequencing is a genome analysis of all nucleic acids (RNA/DNA) in a sample or a bulk of samples, which may contain mixed populations of targeted, off targeted, and random genetic materials from host, microorganisms, and environment in understanding of microbiotic and/or host genetic community and diversity (Lee, 2019). Random sequencing strategy can be applied to a variety of high-throughput sequencing platforms. For example, random sequencing has been applied to aquatic invertebrates resulting in discovery of such as families of bacteriophages (*Myoviridae, Podoviridae,* and *Siphoviridae*) and DNA viruses (*Herpesviridae, Iridoviridae,* and *Poxviridae*) in reef species and jingchuvirals and picornavirals in marine annelids, mollusks, and crustaceans (Jackson et al., 2022; Laffy et al., 2018; Y. Y. Zhang et al., 2022). However, much

fewer viruses have been reported in oceanic chordates, and most were from fish species, resulting in the identification of such as picornavirids, astrovirids, arenavirids, and chuvirids in a variety of teleost species and some species of agnathans and elasmobranch (Costa et al., 2021; Geoghegan et al., 2021; Shi et al., 2018). Therefore, knowledge of viral populations in other oceanic chordates is extremely limited and requires further exploration.

Sea turtles (families: *Dermochelyidae* and *Cheloniidae*) inhabit a vast geographical range and interact with various transoceanic ecosystems for feeding, breeding, and nesting (Boyle et al., 2009). Multiple threats, such as anthropogenic influences, habitat loss, climate change, biotoxins, and infectious diseases, have resulted in a global population decline and many populations are classified as vulnerable, endangered, and critically endangered on the IUCN red list (Rhodin et al., 2021). The first sea turtle virus was detected through conventional assays such as pan-viral-family (e.g., pan-herpesvirus) PCRs (Vandevanter et al., 1996) in sea turtles with fibropapilloma formation (Quackenbush et al., 1998). More recently, novel viruses in sea turtles such as sea turtle tornovirus 1, unclassified CRESS DNA virus, and betanodavirus have been recently discovered through metagenomic sequencing (Fichi et al., 2016; Gainor et al., 2023; Ng et al., 2009). Collectively, the disease association between viruses and free-ranging sea turtle populations remains undetermined in many cases, but some viruses are thought to be associated with disease, e.g., chelonid alphaherpesvirus-5 (fibropapillomatosis) (Whitmore et al., 2021), loggerhead genital-respiratory herpesvirus, loggerhead orocutaneous herpesvirus (Stacy et al., 2008), and a novel poxvirus associated with dark discoloration of skin and eosinophilic inclusion in vacuolated keratinocytes (Sarker et al., 2021).

Interestingly, approximately 60% of stranded sea turtles have gastrointestinal lesions (Orós et al., 2004). This categorization includes a wide range of abnormalities including

ulceration to perforation, obstruction, impaction, inflammation, and intussusception at variable sites of gastrointestinal tract, which can be influenced by noninfectious and/or infectious causes (Manire et al., 2017). Multiple, non-viral infectious agents have been commonly reported as a cause of gastrointestinal disease in sea turtles, including bacteria (*Vibrio spp*., *Aeromonas spp*., *Salmonella spp*., and *Escherichia coli*), protozoa (*Caryospora cheloniae* and *Entamoeba invadens*), and endoparasites (*Rameshwarotrema spp*., *Sulcascaris spp*., *Hapalotrema spp*., *Anisakis spp*., and *Cucullanus spp*.) (Manire et al., 2017). Among these agents, infection of protozoa (e.g., *Caryospora cheloniae*) and spirorchiid trematodes (e.g., *Hapalotrema* sp.) have been commonly documented in free-ranging sea turtles and often associated with fibronecrotizing enteritis (*C. cheloniae*) and granulomatous inflammation of multiple organs, including the gastrointestinal tract (spirorchiid trematodes). Even though many infectious agents can cause serious mortalities in these species, no studies have shown the role of viruses in the enteric health of sea turtles.

In most species, including humans, gastrointestinal viral infections are common and high-impact problems due to their highly contagious nature and significant morbidity (Peery et al., 2012). In veterinary medicine, gastrointestinal viral infections are a major concern in both domestic and wildlife populations, especially those that are immunocompromised or that congregate in large groups, such as animal shelters or rehabilitation facilities, where high probability of the viral transmission exists (e.g., parvovirus enteritis in sheltered dogs and cats [Lechner et al., 2010]). However, there are a variety of viruses detected without disease association reported in several systems in humans and animals (Holtz et al., 2014; Ng et al., 2014). These viruses form communities (a.k.a., viromes) in body systems and help control the

equilibrium of co-existing normal and pathogenic microbiota (e.g., bacteria and protozoa) within hosts and ecosystems (Metzger et al., 2018; Nakatsu et al., 2018).

While less studied, a few gastrointestinal viruses have been reported in wild reptiles in association with gastrointestinal disease. Ranavirus infection in Hermann's tortoises and eastern mud turtles results in ulcerative necrotizing stomatitis, enteritis, and hepatitis (Marschang et al., 1999; Winzeler et al., 2015), and reovirus infection in Chinese vipers and ball pythons results in enteritis (Marschang, 2011). Unlike viromes, these enteric viral infections cause a variety of effects, which may include mortality, but also may result in more subtle disease leading to decreased reproductive fitness (Lechner et al., 2010)

Because so little is known about viral infections in sea turtles, it is imperative to develop methods to efficiently survey for viruses found in gastrointestinal contents prior to more targeted studies to elucidate the scope of enteric viral infections in sea turtles. To address the crucial knowledge gap regarding sea turtle enteric viruses, this study aims to detect viruses that are highly abundant within the gastrointestinal tract of free-ranging sea turtles through random MinION sequencing and then phylogenetically characterize the most relevant viruses via semi-targeted strand-switching MinION sequencing. This study creates a baseline knowledge of the highly abundant viruses that can be observed in free-ranging turtles and provides data for future studies on oceanic viral diversity and on potential disease association in sea turtles.

Material and Methods

Animal ethics and approval

This study was approved by the Florida Fish and Wildlife service (permit number: MTP-23-110), the Georgia Department of Natural Resources (permit number: 121223110), and the

National Oceanic and Atmospheric Administration (permit number: 21111-02). Samples within this study were either pre-collected or unused specimens from either routine antemortem or postmortem examinations. Therefore, a research permit was not required by the Institutional Animal Care and Use Committee at the University of Georgia.

Samples collection

To broaden the viral detection and characterization, specimens (e.g., fecal, colonic, intestinal, and gastric content) were retrieved from six sea turtle species from five states of United States with various health conditions. However, both archived and freshly collected specimens were restricted by sample availability (i.e., gastrointestinal content of necropsied turtles with various health conditions and feces of cold-stunned turtles primarily from the Northeast region). Feces were freely voided from live animals during rehabilitation or pre-collected by freely voiding from the capture-and-release project for toxicological study. The colonic, intestinal, and/or gastric contents were collected during routine necropsy from the stranding response programs. All specimens were placed into 50-mL conical tubes, stored at -80°C, and ultimately transported under 4°C to the University of Georgia for metagenomic sequencing.

Viral enrichment, RNA extraction, and random sequencing

Approximately 50–100 mg of each gastrointestinal specimen was suspended in 400 μL of 1× phosphate-buffered saline prior to the analysis. PBS-suspended contents were enriched for viral particles based on previously published protocols (Conceição-Neto et al., 2018; Vibin et al., 2018). Briefly, specimens were homogenized at 35 Hz for 2 minutes for 2 times with a sterile,

stainless steel, 0.5 mm metal bead (Qiagen). Homogenized samples were centrifuged at 17,000 × *g*, at room temperature for 3 minutes, passed through a 0.8 μm PES filter (Sartorius), and treated with a cocktail of 2.0 μL of micrococcal nuclease (2,000,000 gel units/mL, NEB Inc) and 1.0 μL Benzonase® (>250 units/μL, Millipore Sigma) in 7.0 μL of resolving enzyme buffer (Conceição-Neto et al., 2018) at 37°C. Enzyme activities were terminated with 3 μL of 500 mM EDTA (Thermo Fisher Scientific). RNA was subsequently extracted following the manufacturer's protocol using Trizol™ LS Reagent (Thermo Fisher Scientific). For detection, sequencing libraries were prepared using a random strand-switching protocol as previously described (Young et al., 2021). In short, the purified RNA was reverse transcribed using the SuperScript IV reverse transcriptase kit (Thermo Fisher Scientific) with 1 μM PCR-RH-RT primer (5'-ACTTGCCTGTCGCTCTATCTTCNNNNNN-3') as a reverse primer and 10 μM strand-switching oligo (5' -TTTCTGTTGGTGCTGATATTGCTGCCATTACGGCCmGmGmG-3'; both synthesized by Integrated DNA Technologies [IDT]). The reaction mixture was incubated at 50°C for 30 minutes, 42°C for 10 minutes, and 80°C for 10 minutes before being bead purified at a 0.7× beads:solution (volume/volume) ratio (KAPA Biosystem). cDNA then was used as a template for barcoding PCR following Oxford Nanopore Technologies (ONT)'s protocol (SQK-LSK110 with EXP-PBC096) and LongAmp Taq 2× Master Mix (NEB) using 18 cycles. The barcoded amplicons were bead purified at a 0.8× beads:solution ratio before being pooled by equal volume. Five to twelve samples were pooled per sequencing run for end repair and ligation of the sequencing adapter using suggested kits for ligation sequencing amplicons (SQK-LSK110) library preparation. Bead purification was also performed after end repair and adapter ligation at 1.0× and 0.4× beads:solution ratios, respectively. The library was washed with long fragment buffer (LFB) and eluted in elution buffer (EB) before mixing with sequencing

buffer (SBII) and loading beads II (LBII), as suggested by ONT's protocol. The final library was loaded onto a FLO-MIN106 R9.4.1 flow cell in a MinION Mk1B sequencing device (ONT). Sequencing was controlled using MinKNOW v.23.07.5 (ONT) until at least 800,000 reads generated with default settings and hourly MUX scans.

Sequence analysis, consensus sequence assembly, and annotation of open reading frames

Sequencing analysis followed the randomly primed MinION-based sequencing as previously described (Young et al., 2021). In brief, raw reads were basecalled with the GPU version of Guppy v.6.4.2 or Dorado v.0.5.3 with default settings and high accuracy setting, respectively (ONT). Porechop (http://github.com/rrwick/Porechop) was used for trimming and demultiplexing with 99% adapter threshold and a million read checked for adapter matching. Potential viral reads were initially screened for detectable viruses using all reads by pairwise aligning to non-redundant Basic Local Alignment Search Tool (BLAST+) using BLASTN database (updated on August 02, 2023) through the Georgia Advanced Computing Research Center (GACRC) using default settings with minimum expected value at 1×10^{-4}. Virus detection was defined as at least 5 viral reads that had similarity to the same genus via BLASTN. To phylogenetically characterize the viral genomic regions for species and genus demarcation, viral contigs (at least 10× coverage) were assembled through a de novo method (Flye with high read quality and metagenomic setting) (Kolmogorov et al., 2020) and/or reference-based method (Geneious Prime 2019.1.3 and 2023.2 [https://www.geneious.com] using "map-to-reference" function with the top hit result from BLASTN and BLASTX alignment based on number of hits, e-value, bit score, % pairwise identities, and % query coverage). Potential viral contigs were initially translated and searched against Pfam database (Mistry et al., 2021) to predict protein

function. Subsequently, protein-predicted viral contigs were confirmed and expand the genotyping regions by the semi-targeted strand-switching PCR prior to MinION sequencing. In several samples, multiple viral variants (either a result of the quasispecies or high mutation of a single virus) within the same sample were identified and manually sorted based on disagreement of read populations to the consensus sequence. To further distinguish these viral variants, different primers targeting each viral variant were designed for semi-targeted PCR to gain adequate sequencing depth (10× coverage) and sequencing breadth. Eventually, coding regions were predicted by using the plug-in ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder) along with the comparison to their closely related sequences on Geneious Prime. To annotate the RdRp and capsid protein domains, the Interproscan protein domain searching tool by using PfamA database was used on Geneious.

Primer design and semi-targeted strand-switching sequencing

To increase sequencing coverage of the genome of potential viral contigs, specific reverse primers targeting the genomic sense of the viral contigs were designed using a primer design tool on Geneious Prime, targeting specific areas of the genomes that remained incomplete after random sequencing (Supplemental table 2). Viral-specific reverse primers (2 μL of 10 μM) were used in reverse transcription, strand-switching reactions as described in the random sequencing protocol, substituting for the random hexamers. For some double-stranded RNA viruses, the reverse transcription, strand-switching protocol was modified with additional incubation of 25°C for 1 hour prior to the previously described incubations. Purified cDNA was used in semi-targeted strand-switching PCR using LongAmp Taq 2× Master Mix (NEB) with 10 μM viral-specific reverse primers and 10 μM strand-switching oligo as the forward primer. The

reaction mixture was incubated at 95°C for 30 seconds, followed by 25 cycles of denaturation at 95°C for 10 seconds, annealing at 3–5°C below the melting temperature of viral-specific reverse primers, extension at 65°C for 13.5 minutes, and a final extension step at 65°C for 15 minutes before being bead purified at a 0.8× beads:solution ratio. Purified first-round PCR products were used in 15 cycles of the barcoding PCR reaction, followed by end-preparation, adapter ligation, final library preparation, and MinION sequencing were conducted similarly to the random sequencing protocol.

Viral abundance and diversity analysis

Calculation of viral relative abundance was performed in reads per millions (RPM, in $\log_{10}$ scale). The reads from random sequencing that mapped to the final consensus sequences of each classified virus were used as the number of detected reads. The calculated RPM for each classified virus was plotted on a heatmap in relation to host species, stranding location, health condition, viral taxonomic classification, viral genomic structure, and viral nucleic acid types for viral diversity estimation via RStudio 2023.12.1.

Genetic comparison and phylogenetic analysis

After identification and annotation of the coding regions, viral consensus sequences were analyzed by using species or genus characterization criteria for each detected viral taxon from the International Committee on Taxonomy of Viruses (ICTV) or if unavailable, from recent publications. For the sequence alignment, multiple sequencing aligners were used (depending on the virus) to align and occasionally to later refine corresponding to the phylogenetic analysis of each viral taxon (Table 1). Amino acid sequences were used to construct viral phylogenies,

except Cressdnaviricota (complete nucleotide sequences), on Geneious, MEGA X, and MEGA 11 (Table 1). Best substitution models of aligned sequences for each viral taxon were selected based on the lowest Bayesian information criterion and Akaike scores using the best substitution model analysis for maximum-likelihood analysis and suggested models from ICTV on MEGA X and MEGA 11. Each viral phylogenetic analysis was constructed using the maximum-likelihood method, except picobirnaviruses (Neighbor joining method) at 500 to 1,000 bootstraps (Table 1).

Results

Sample collection and biographical information

Gastrointestinal specimens from 78 turtles were used in this study. The samples represented six sea turtle species: green turtles (*Chelonia mydas*; n = 23), loggerhead turtles (*Caretta caretta*; n = 20), Kemp's ridley turtles (*Lepidochelys kempii*; n = 12), olive ridley turtles (*Lepidochelys olivacea*; n = 11), hawksbill turtle (*Eretmochely imbricata*; n = 1), and leatherback turtles (*Dermochelys coriacea*; n = 11). Samples were collected from the West Coast (Washington, Oregon, and California; n = 37), Northeast Coast (Massachusetts; n = 22), and the Southeast Coast (Georgia and Florida; n = 19) of the United States (See Fig. 2.1). Conditions of animals were categorized, by reviewing records of animal history, gross necropsy, and histopathology, as presumably healthy (i.e., capture-and-released; n = 12), gastrointestinal disease (e.g., diarrhetic or histological lesions; n = 12), systemically ill (e.g., cold-stunned, septic, or papillomatosis; n = 41), and unknown origin (i.e., found dead and lacked significant postmortem changes; n = 13) (Fig. 2.1). The gastrointestinal affected group was further subdivided as inflammatory (n = 8), traumatic (n = 1), toxic (n = 1), and unclear causation (n =

2). Within the rehabilitated group, rehabilitation durations (<1 week [n = 6], 1–4 weeks [n = 4], and >1 month [n = 26]) were analyzed along with the detection of viruses.

Sequencing metrics

Barcoded libraries were prepared by pooling 5–12 specimens per library. Semi-deep random sequencing was performed, generating approximately 800,000–12,000,000 raw reads per library, to target only highly abundant viral nucleic acid. The N50 of each library was 491–3,450 bases.

Viral abundance and diversity: Summary

In general, viruses were analyzed from two depths of coverage. Initially, 48 viruses were detected through random sequencing, which was defined in this study as having at least 5 reads in which top BLASTN pairwise alignments had a similarity to viruses within the same genus or family (Fig. 2.2). Of these 48, 23 were assignable to a known viral family. Secondly, at least 38 potential viral contigs were created by combining detected viral reads and de novo-assembled contigs through reference-based consensus building. Of these potential viral contigs, 19 viruses were phylogenetically characterized (12 complete coding sequences [CDS] and 7 partial CDSs, from 11 individuals) were confirmed through semi-targeted strand-switching sequencing, which was defined in this study has having sufficient depth (at least 10×) of the regions used for phylogenetic characterization specific for each viral taxon (e.g., RdRp) (Fig. 2.2).

Viral detection: Abundance

Using the detection threshold, 24% of turtles (19 individuals) had detectable viral reads. Viruses were detected in all geographical regions: Northeast = 27%, Southeast = 37%, and West Coast =16% (Fig. 2.2A). Viruses were detected in most species, except for hawksbill turtles, which only had one sample, and olive ridley turtles. Approximately 55% of loggerhead turtles had at least one detectable virus, whereas the percentage of green, leatherback, and Kemp's ridley turtles with detectable viruses was 13%–27% (Fig. 2.2B). Viruses were detected in all health conditions 15%–33% (Fig. 2.2C). Within a group of gastrointestinal affected turtles, viruses were detected in 1/1 turtle with gastrointestinal toxicosis (crude oil in GI tract) and 2/2 turtle with gastrointestinal signs but unclear pathology. No viruses were found within cases of inflammatory gastrointestinal disease (n = 8), nor in the single traumatic gastrointestinal case. Within a group of rehabilitated turtles, viruses were detected in all rehabilitation durations: 3/6 (50%) turtles with less than a week rehabilitation, 1/4 (25%) turtles with rehabilitation time between 1 week to 4 weeks, and 5/21 (19%) turtles with more than a month rehabilitation (Fig 2.2D).

Frequency analysis was performed on the random sequencing runs. Viral abundance and diversity showed that the gastrointestinal viral community of free-ranging sea turtles is diverse across sea turtle species, geographical locations, ages, and health conditions (Fig. 2.2).

Nominal logistic regression analysis suggested that only species has a statistically significant impact on the detection of viruses; however, the only significant pairwise comparison was between loggerhead and olive ridley turtles.

Viral detection: Taxonomic assignment

Individual read assignment (defined using BLASTN) showed that of the 48 detected viruses, 23 could be assigned to the family level resulting in eight different viral families (*Chuviridae, Caliciviridae, Dicistroviridae, Partitiviridae, Picobirnaviridae, Orthototiviridae, Parvoviridae,* and *Circoviridae*). Additionally, there were several viruses that had coverage in the genome areas relevant to family assignment; however, family designations for some viral orders are incomplete and thus the viruses in this study could not be assigned to known families based on BLAST. For hepelivirals, there were eleven viruses that could not be assigned to known families, but at least two new families are possible (see "Viral phylogenetic analysis by taxa" section). Similarly, for cressdnaviricots, there was one virus that could not assigned to a known family, but at least one new family is possible (see "Viral phylogenetic analysis by taxa" section). Lastly, thirteen additional viruses lacked coverage of regions relevant to family assignment and were only assigned as bunyaviral (n = 1), picornavirals (n = 5), and RNA viruses (realm: Riboviria; n = 7).

In addition to those 48 detected viruses, several reads and contigs with similarity to bacteriophages (e.g., *Microviridae, Myoviridae, Podoviridae,* and *Siphoviridae*) and plant-infecting viruses (e.g. *Tombusviridae, Virgaviridae,* and *Alphaflexiviridae*) were detected, but were not further investigated due to the focus on potential animal-infecting viruses in this study. However, these sequences are available for investigators in these fields.

Virus phylogenetic characterization: Abundance

Viral consensus sequences were phylogenetically characterized using contigs with at least 10× depth of coverage across the required region used for phylogenetic analysis for those

particular viral taxa. In total, 11 turtles (14%), with at least one turtle of all facilities and all geographical regions, had phylogenetically characterizable viruses. Viruses were phylogenetically characterizable in all geographical locations (13%–40%) and all heath status (7.7%–17%) (Fig. 2.2). Viruses were phylogenetically characterized only from loggerhead turtles (30%), green turtles (13%), and Kemp's ridley turtles (17%) (Fig 2.3).

Nineteen viruses across eight known viral families and two unclassified orders were phylogenetically characterizable including 11 putatively novel species and three putatively novel genera. Even though only RNA was targeted, both DNA and RNA viruses were phylogenetically characterized within the RNA viruses, four double-stranded across three families (*Partitiviridae, Orthototiviridae, Picobirnaviridae*), 1 negative-sense (*Chuviridae*), and 11 positive-sense RNA viruses were characterized (*Dicistroviridae, Caliciviridae,* and unclassified Hepelivirales). Within the DNA viruses, three DNA viruses, including both linear (i.e., *Parvoviridae*) and circular genomes (i.e., *Circoviridae,* and unclassified Cressdnaviricota) were characterized. Phylogenetically characterized viruses are reported below based on taxonomy.

Viral phylogenetic analysis by taxa

RNA viruses (single-stranded RNA viruses)

Family *Chuviridae*

Two genomic segments of a negative-sense, linear, bisegmented virus, closely related to Wenling crustacean virus 14 (*Chuvivirus brunnichi,* family *Chuviridae*), were identified in a cold-stunned loggerhead turtle stranded in Massachusetts (Figs. 2.1 and 2.2). The first genomic segment had 6,717 nucleotides (nt) encoding a large protein (2,180 amino acids) and the second genomic segment had 4,371 nt encoding 3 putative open reading frames (ORFs): glycoprotein

(655 amino acids), a nucleoprotein (477 amino acids), and a viral structural protein 4 (150 amino acids). The sequencing coverage of both segments was at least 10× (except for the 117, 5'-most bases of glycoprotein coding region with at least 3×). The predicted large protein had 89.36% pairwise amino acid identity to Wenling crustacean virus 14 (GenBank: KX884451) and 31.91% pairwise amino acid identity to Wenzhou crab virus 2 (KM817601) through CLUSTAL W and MUSCLE alignments (Table S1). Phylogenetic analysis showed that this virus shared a common ancestor with these two chuviviruses (96% bootstrap value), which are the only currently fully recognized members of the *Chuvivirus* genus (Fig. 2.2.1). Based on the current classification criteria of Jingchuvirales using 91% amino acid similarity (Di Paola et al., 2022) for species demarcation and monophyletic clustering with other chuviviruses, this virus is suggested to be a novel species with the genus *Chuvivirus.* The virus was putatively named sea turtle feces associated chuvivirus 1.

Family *Caliciviridae*

Two distinct caliciviruses (positive-sense, linear, unsegmented genome) were detected in two different green turtles, both with fibropapillomatosis, stranded in Georgia (CM-C21249) and Florida (Myrtle-23228) (Figs. 2.1 and 2.2). A partial ORF1 (with at least 10× coverage, except for the 15, 3'-most bases of the 3' ORF1 with 9×) was detected in the Georgia-stranded turtle. Within the partial ORF1, a complete viral protein 1 (VP1, a.k.a., capsid protein) coding region (796 amino acids) had 81.41–84.15% pairwise amino acid identity to two different Nordland viruses (Atlantic salmon calicivirus, species *Salovirus nordlandense*, family Caliciviridae, GenBank: KJ577139-140) through CLUSTAL W alignment (Table S1). While species demarcation criteria for saloviruses have not been formally established, other well-studied genera

of calivirids (e.g., *Lagovirus*, *Norovirus*, and *Vesivirus*) use monophyletic clustering of VP1 amino acid sequences as a speciation criterion. Phylogenetic analysis showed that this turtle-associated calicivirid consistently formed a monophyletic group with the saloviruses (100% bootstrap support) (Fig. 2.4.2). Therefore, this calicivirid is presumably assigned to the *S. nordlandense* species. Additionally, reads similar to Nordland virus were also detected in a different sea turtle species from the same rehabilitation facility. However, random sequencing reads were insufficient to build a confident consensus sequence of the VP1 genomic region and semi-targeted sequencing was not performed on those samples.

A partial genome encoding ORF1 and ORF2 (with at least 10× coverage, except for the 102, 5'-most bases of the 5' ORF1, which had 3×–8×) of another calicivirid was detected in a green turtle stranded in Florida. A complete VP1 protein (511 amino acids) shared 23.36% pairwise amino acid identity to Minovirus A (fathead minnow calicivirus, species *Minovirus pimephalis*, family Caliciviridae, GenBank: KX371097) through CLUSTAL W alignment (Table S1). Based on the current genus demarcation criterion, this calicivirid is proposed to be a novel species of a novel genus. This virus was putatively named sea turtle feces associated calicivirus 1. Phylogenetic analysis showed that sea turtle feces associated calicivirus 1 shared a common ancestor with fathead minnow calicivirus (82% bootstrap support), and that these genera of caliciviridsin turn share a common ancestor with the saloviruses (87% bootstrap support) (Fig. 2.4.2).

Order Hepelivirales: Unclassified

A complete CDS of a hepeliviral (positive-sense, linear, unsegmented genome, order Hepelivirales*)* were identified in a loggerhead turtle with gastrointestinal disease stranded in

California (SWFSC-MT-2014-05). The complete genomic CDS had 6,332 nt encoding the nonstructural polyprotein (1,210 amino acids) and the capsid protein (674 amino acids). Within the nonstructural polyprotein, the RNA-dependent RNA polymerase (RdRp) motif, 315 amino acids located towards the 3′ end of the nonstructural protein, was used for amino acid comparison to other hepelivirals, which demonstrated that the RdRp motif of this virus shared 38.91% and 38.82% pairwise amino acid identity to rana hepevirus (GenBank: MH330682) and Bastrovirus/VietNam/Porcine/17489_85 (GenBank: KX907134), respectively. The RdRp-based phylogeny of hepelivirals revealed that this sea turtle feces associated virus clustered with rana hepevirus (91% bootstrap value), and that this paired cluster shared a common ancestor (63% bootstrap value) with viruses detected in feces of animals (pigs and rats) and humans (Fig. 2.4.3), which are a group of unclassified hepelivirals termed “bastroviruses”. Due to unestablished taxonomic classification of bastroviruses, this virus is not confirmed as a novel species or genus. The capsid protein shared 38.6% pairwise amino acid identity to Fujian spotted paddle-tail newt astrovirus (GenBank: MG599913, unclassified astrovirid). These viruses were putatively named sea turtle feces associated hepeliviral 1.

Order Hepelivirales: “bastro-like”

Three sequences with high similarity to “bastro-like” viruses were identified in a loggerhead turtle (SWFCS-MT-2014-05) and a green turtle (SWFSC-MT-2017-03). Both turtles had unknown health conditions and were stranded in California. In the loggerhead turtle, two complete genomic CDSs of hepelivirals were detected. Genomes were similar in size (5,789 nt and 5,799 nt) and encoded two putative open reading frames (nonstructural polyprotein and capsid proteins). Both nonstructural polyproteins were the same size (1,429 amino acids), but the

capsid proteins were different in size due to nonsynonymous mutation (C5,676T) causing a creation of premature stop-codon (367 amino acids and 392 amino acids). The pairwise amino acid identity of nonstructural polyprotein between the two viruses was 97.9% and the pairwise amino acid identity of capsid protein was 95.9%. The pairwise amino acid identities of nonstructural polyprotein to the close related viruses were 62.43% to Hepelivirales sp. strain MR233-17E/26350 (GenBank: OP699129) and 59.67% to Pyongtaek Culex bastrovirus isolate-like virus (MT568535). The capsid protein had 38.6% pairwise amino acid identity to Barn Ness breadcrumb sponge noda-like virus 1 (GenBank: MF189996, *Nodaviridae*). These viruses were putatively named sea turtle feces associated bastro-like viruses 1 and 2, which were likely to be a detection of a novel species with two different variants in a single animal.

In the green turtle, another partial CDS of a hepeliviral was detected. This partial genome (5,431 nt) encoded two open reading frames (complete nonstructural polyprotein [1,433 amino acid] and partial capsid protein [>268 amino acids]). The pairwise amino acid identities of the nonstructural polyprotein between this hepeliviral and the two previously discussed hepelivirals from the loggerhead turtle were 41.29% and 41.79%, and the pairwise amino acid identities of capsid protein were 32.97% and 33.04%. This hepeliviral was likely a novel species and the virus was putatively named sea turtle feces associated bastro-like virus 3.

The phylogenetic analysis using hepelivirals RdRp motifs, 230 and 258 amino acids towards the 3′ end of the nonstructural protein, showed that both sea turtle feces associated bastro-like virus 1 and 2 clustered together (100% bootstrap value) and shared a common ancestor with other bastro-like viruses (Fig. 2.2.3). Sea turtle feces associated bastro-like virus 3 clustered with Hepelivirales sp. strain MR233-17E-17006 (GenBank: OP699128) (100%

bootstrap value) and shared a common ancestor with other bastro-like viruses including sea turtle feces associated bastro-like virus 1 and 2 (Fig. 2.4.3).

The capsid-based BLAST results yielded significantly different results than the RdRp-based results; therefore, the capsid protein was phylogenetically analyzed. This demonstrated that sea turtle associated bastro-like viruses 1 and 2 clustered together (100% bootstrap value) and were highly similar. These two sequences shared a common ancestor with Barn Ness breadcrumb sponge noda-like virus 1 (88% bootstrap value). Sea turtle associated bastro-like virus 3 eventually formed a monophyletic group (99% bootstrap value) with the other sea turtle bastro-like viruses and the sponge noda-like virus. This viral cluster shared a common ancestor with nodavirids (Order: Nodamuvirales), and later with bastro-like viruses (Fig 2.4.4).

Family *Dicistroviridae*

Two highly similar dicistrovirids, positive-sense, linear genome, order Picornavirales, (98.86% pairwise nonstructural (NS) nucleotide similarity) were detected from a systemically ill Kemp's stranded in Massachusetts (C21338). Complete genomes were unavailable due to a lack of sufficient depth (i.e., <10×) in the middle of the genomes (due to the primer binding location for semi-targeted sequencing); thus, two partial genomic sequences were reported for each virus (a 5' fragment and a 3' fragment). For the first dicistrovirus, the 5' fragment contained a partial NS and was 4,065 nt. The 3' fragment contained a partial NS and a complete capsid protein, and was 4,092 nt. For the second dicistrovirus, the 5' fragment contained a partial NS and was 3,945 nt. The 3' fragment contained a partial NS and a partial capsid protein and was 3,749 nt. Based on the BLASTX alignment of ORF consensus sequences, the NS CDS had 60.5% pairwise amino acid identity to Riboviria sp (GenBank: MW347392) and 41% to Wufeng shrew

dicistrovirus 9 (GenBank: OQ715898); and the capsid CDS had 62.6% pairwise amino acid identity to Riboviria sp. and 45.4% to Wufeng shrew dicistrovirus 9. Phylogenetic analysis of dicistrovirus RdRp amino acid sequences showed that these two viruses clustered together (100% bootstrap value) and shared common ancestors with other cripaviruses (Fig. 2.4.5). These viruses were putatively named sea turtle feces associated dicistrovirus 1 and 2.

Additionally, two complete CDSs (9,924 nt and 8,871 nt) and 1 partial CDS (6,807 nt) of different dicistroviruses were detected in a presumably healthy loggerhead turtle stranded in Florida (NMFS22-46). These viruses were putatively named sea turtle feces associated dicistroviruses 3, 4, and 5. The NS ORF of dicistrovirus 3, 4, and 5 were 2,018 amino acids, 1,831 amino acids, and 1,175 amino acids (partial CDS); respectively. The capsid proteins of dicistroviruses 3, 4, and 5 were 941 amino acids, 846 amino acids, and 815 amino acids (presumably complete ORF of a partial sequence); respectively. The percent pairwise amino acid identity between three NS ORFs were 37.16–47.58%. The percent pairwise amino acid identity between three capsid ORFs were 23.02–41.25%. Phylogenetic analysis of RdRp amino acid sequences of dicistroviruses showed that sea turtle feces associated dicistroviruses 3, 4, and 5 formed a monophyletic group with other dicistroviruses (76% bootstrap value) (Fig. 2.4.5). While sea turtle feces associated dicistrovirus 3 clustered within viruses under the genus *Aparavirus* (83% bootstrap value), this virus is relatively distinct from the majority of existing aparavirus with great genetic distance. While the sea turtle feces associated dicistroviruses 4 and 5 did not share a nearest common ancestor with each other, they both clustered with other highly similar aparaviruses (e.g., Wenzhou picorna-like virus 28, Strongylocentrotus intermedius associated picornavirus 1, mute swan feces associated picorna-like virus 15, and mute swan feces assocaited picorna-like virus 16) (Fig. 2.4.5).

Notably, sub-populations of these sea turtle feces associated dicistroviruses were detected in which the subpopulations were highly similar to the dominant population, but multiple single nucleotide polymorphisms (SNPs) were consistently detected within multiple long reads (i.e., multiple SNPs were consistently linked on long reads to other SNPs). Variable sites were evaluated at a 75% consensus threshold. For sea turtle feces associated dicistroviruses 3, 4, and 5; 137, 28, and 86 SNP sites were identified, respectively.

RNA viruses (double-stranded RNA viruses)

Family *Picobirnaviridae*

A complete RNA-dependent RNA-polymerase (RdRp) segment of picobirnavirus (double-stranded, linear, bisegmented genome) was detected in a gastrointestinally affected Kemp's ridley turtle stranded in Florida (NMFS22-350). This RdRp segment contained 2,312 nt encoding 740 amino acids. The RdRp of this picobirnavirus shared 63.75% pairwise amino acid identity to the atypically large genome of porcine picobirnaviruses (GenBank: MW977143-4). RdRp-based phylogeny showed that this newly identified picobirnavirus clustered with other closely related porcine picobirnaviruses (100% bootstrap value), but greatly distinct from most ICTV-classified picobirnaviruses (genogroup 1, 2 and 3) (Fig. 2.4.6). This virus is putatively named sea turtle feces associated picobirnavirus 1.

Family *Partitiviridae*

Two complete RdRp CDSs of partitivirids (double-stranded, linear, bisegmented genome) were identified in a systemically ill green turtle stranded in Georgia (CM-C21249) only through mis-priming amplification of semi-targeted strand-switching sequencing originally

targeting the sea turtle feces associated totivirus 1. Complete RdRp segments had 1,620 nt with at least 10× coverage and 1,415 nt with at least 10× coverage except for 5'-most 102 bases, which had 3×–8×. The amino acid coding regions of both partitivirids were different in size (495 amino acids and 463 amino acids) and the pairwise amino acid identity between these newly identified partitivirids was only 18.27%. However, the partitiviral phylogeny by using RdRp amino acid sequences showed that these newly identified partitiviruses clustered together, albeit with poor bootstrap support and shared a common ancestor with gammapartitiviruses, partitivirids isolated from various species of fungi (Fig. 2.4.7). Based on the current ICTV genus demarcation criteria using <24% RdRp amino acid sequence identity in pairwise comparisons, each newly identified partitivirus was novel species. These viruses were putatively named sea turtle feces associated partitivirus 1 and 2) and belonged to their own novel genera. However, the ICTV genus demarcation criteria also include host characteristics (*Alpha- and Betapartititivirus:* plant, *Gammapartititivirus:* fungi, and *Deltapartititivirus:* protozoa), protein length, e.g., RdRp (*Alphapartititivirus:* 581–621 amino acids, *Betapartititivirus:* 663–746 amino acids, *Gammapartititivirus:* 519–539 amino acids, and *Deltapartititivirus:* 472–479 amino acids), and organization of phylogenetic group (Fig. 2.4.7). Apart from the unclear host sources due to isolation of the virus from colonic content of a green sea turtle, both sea turtle feces associated partitiviridss do not belong to previously identified genera and support the generation of at least one new partitiviral genus.

Family *Orthototiviridae*

A single orthototivirus (double-stranded, linear, nonsegmented genome) from a systemically ill green turtle stranded in Georgia (CM-21249). Due to a lack of sufficient depth

(i.e., <10×) in the middle of the genomes; thus, two partial genomic sequences were reported for this single orthototivirus (a 5' fragment and a 3' fragment). The 5' fragments contained a capsid protein (2,218 nt [598 amino acids]). The 3' fragments contained only a partial RdRp sequence (2,686 nt [>864 amino acids]). The pairwise amino acid identity of RdRp to an unclassified Totiviridae sp. virus (GenBank: MZ218683) was 43.88% and pairwise amino acid identity of RdRp to Eimeria stiedai RNA viruses (GenBank: OQ821982 and MT129769) were 25.79 and 28.12%; respectively. The phylogenetic analysis using RdRp amino acid sequences of totiviruses showed that this newly identified totivirus shared a common ancestor with, but separately clustered from, the totiviruses from *E. stiedai* and the unclassified totivirus metagenomically detected from feces of a pigeon (Fig. 2.4.8). Based on the ICTV speciation demarcation criteria (isolated host and >50% amino acid identity) of family *Totiviridae*, this totivirus was likely classified as a novel species, and was putatively named sea turtle feces associated orthototivirus 1, but closely related to the unclassified totiviruses isolated from *Eimeria stiedae* and pigeon feces.

DNA viruses (single-stranded DNA virus)

Family *Parvoviridae*

A putative novel species of aquambidensovirus (single-stranded, linear DNA genome with hairpin termini, subfamily *Densovirinae*, family *Parvoviridae*) was detected in Kemp's ridley turtle with unknown health status from Massachusetts (21-0802-Lk). The complete CDS (5,868 nt) was composed of three major positive-sense, nonstructural (NS) proteins (NS1 [567 amino acids], NS2 [308 amino acids] and NS3 [360 amino acids]) and 1 negative-sense, viral structural protein (894 amino acids). This novel parvovirid was highly similar to

ambidensoviruses detected in oceanic environments such as 78.48% pairwise identity to uncultured densovirus (GenBank: MT733049) and 74.74% pairwise identity to sea star-associated densovirus (GenBank: KM052275). Phylogenetic analysis using NS1 protein 1 showed that this newly identified virus monophyletically clustered with aquambidensoviruses isolated from aquatic hosts, e.g., freshwater crayfish, marine sea stars, and oysters (99% bootstrap value) (Fig. 2.4.9). Based on the NS1 phylogeny, and the current speciation (>85% similarity) and genus (>30% similarity) demarcation criteria, this parvovirid was a novel species and was putatively named sea turtle feces associated aquambidensovirus 1, under the previously described genus *Aquambidensovirus* in subfamily *Densovirinae* and family *Parvoviridae.*

DNA viruses (circular DNA viruses)

Order Cressdnaviricota

Complete genomic sequences of a novel circovirus (single-stranded circular DNA genome, family *Circoviridae*) and an unclassified CRESS DNA virus (single-stranded circular DNA genome, phylum Cressdnaviricota) were identified in a Kemp's ridley turtle (21-0802-Lk) and a loggerhead turtle (21-0662-Cc); respectively. Both turtles were stranded in Massachusetts and were categorized as systemically ill. An 1,819-nt novel cressdnaviricot contained two putative ORF (replicase 915 nt [305 amino acids] and capsid protein 516 nt [172 amino acids]) with the nucleotide sequence "TAGTATTAC" at the origin of replication (oriented in the same direction as replicase protein) and an 11-nt complimentary stem-loop, consistent with circovirids. A 1,742-nt unclassified CRESS virus contained three putative major coding regions (replicase 1,035 nt [345 amino acids], capsid protein 624 nt [208 amino acids], and ORF3 597 nt [199 amino acids]) with the unusual nucleotide sequence "TTCTATTAC" at the origin of replication

(oriented in the same direction as replicase protein) and a 10–11 nt complimentary stem-loop. These two cressdnaviricots were only 16.1% identical based on the pairwise comparison of the RdRp amino acid sequences. The most closely related viruses to this newly identified circovirus were human stool-associated circular virus (GenBank: GQ404856) and barbel circovirus (GenBank: GU799606) with 45.53% and 45.17% pairwise amino acid identity using RdRp CDS. The most closely related viruses to the newly identified unclassified CRESS virus were Cressdnaviricota sp. isolate ctbd27 (GenBank: MH616968) and Ecklonia radiata-associated virus 3 (GenBank: NC_040823) with 98.84 and 38.83% pairwise amino acid identity using the replicase CDS; respectively. The whole genome phylogeny revealed that the newly identified circovirus shared a higher common ancestor with other circoviruses identified in European catfish, bat, and human feces (79% and 98% bootstrap value; respectively) (Fig. 2.4.10). Given the phylogenetic clustering with other circovirids, similar directional orientation of replicase protein to the origin of replicase, and speciation demarcation criteria using the whole genome nucleotide comparison, this circovirus is a novel species, and it was putatively named sea turtle feces associated circovirus 1. On the other hand, the newly unclassified CRESS DNA virus along with closely related viruses formed a monophyletic group (100% bootstrap value) and shared a common ancestor with viruses from order Mulpavirales. However, an unclassified CRESS DNA virus remained unclassified due to perplexing classification of CRESS virus and unclear phylogenetic analysis. This virus was putatively named Cressdnaviricota sp. isolate sea turtle feces associated.

Discussion

While research has exponentially increased the number of known viruses (https://www.ncbi.nlm.nih.gov/nucleotide/ accessed June 28, 2024), including viruses from marine environments, little work has been reported on viral discovery in aquatic reptiles. This is likely due to limitations of accessible sample collection, quality of collected specimens, and financial support in conducting wildlife research. While this study was purposely not designed for a full interrogation of the sea turtle gastrointestinal viromes (e.g., depth of sequencing selected for detection of more abundant viruses, phage sequences were not analyzed), it is important to unveil undiscovered viruses and catalog viral sequences, particularly ones that are highly abundant in gastrointestinal tract of sea turtles. This genetic archive can help identify genetically distinct viruses, better understand viral evolution and ecology, and shed light on their impact on these understudied aquatic species.

While there are a few viruses have been reported in association with diseases in reptiles, e.g., serpentovirus, paramyxovirus, reptarenavirus, turtle fraser virus 1 (Marschang, 2016, 2019; Waltzek et al., 2022), disease association is not always known (Waltzek et al., 2022) due to a lack of knowledge regarding the reptilian virosphere (Harding et al., 2022a). Much less known in sea turtles where only nine viral species are reported: herpesviruses (chelonid herpesvirus1, 5, and 6), papillomaviruses (chelonia mydas and caretta caretta papilloma virus 1), cheloniid poxvirus 1, sea turtle tornovirus 1, unclassified CRESS DNA virus, and betanodavirus, and the recently discovered sea turtle neural viruses (Fichi et al., 2016; Laovechprasit et al., 2024; Marschang, 2011; Ng et al., 2009; Sarker, Hannon, et al., 2021; Stacy et al., 2008; Whitmore et al., 2021). While the gastrointestinal tract is suggested to be a common site to observe pathologic changes in stranded sea turtles (Orós et al., 2004), no gastrointestinal viruses have been

identified. Before being able to determine correlation or causation with the health status of free-ranging sea turtle populations, the community and diversity of gastrointestinal viruses in these endangered marine species needs to first be established and methods for efficient detection and characterization need to be developed.

This study was designed to broaden the knowledge regarding the virosphere of sea turtles using a hybrid metagenomic approach (semi-deep random sequencing followed by semi-targeted sequencing). Because of limited knowledge regarding viral abundance and diversity in sea turtles, the samples were selected from a range of species, locations, and health conditions to maximize the likelihood of identifying viruses in sea turtles. Regardless, gastrointestinal viruses in sea turtles were diverse and were broadly found across all the variables (except for olive ridley turtles [n=11] and hawksbill turtles [n=1]). These variables (e.g., species, geographical locations, facilities, health conditions, rehabilitation durations, and specific causation of gastrointestinal diseased animals) were statistically assessed, but only species had a statistically significant impact on the detection of viruses. Pairwise analysis of the nominal linear regression indicated a statistically significant correlation between detectable viruses in loggerhead turtles (highest number of detectable and phylogenetically characterizable viruses) and olive ridley turtles (no detectable virus found). However, nominal logistic regression failed to support any significant correlation of other variables. It is noted, however, that the study design limits the ability to perform statistical comparisons between groups. Therefore, a few potential trends are noted below and may be worth further consideration. Even though statistical analysis failed to support that loggerhead turtles have an impact on the detection of viruses, 55% of loggerhead turtles had at least one detectable virus. Determining if loggerhead turtles truly have more gastrointestinal viruses as compared to olive ridley turtles or other turtles requires future, hypothesis-driven

studies. Even though statistical analysis failed to support that the correlation of health condition and the detection of viruses, it is interesting that within the gastrointestinal affected turtles which has detectable viruses, viruses were not detected in turtles with gastroenteritis (n=8) but were detected in 1/1 turtle with gastrointestinal toxicosis (100%; crude oil in GI tract) and 2/2 turtle with gastrointestinal signs but unclear pathology (see Fig 2.2D). Within the rehabilitated turtles, viruses were seemingly more often detected (50% of animal) in early rehabilitated turtles (<1 week) as compared to prolonged rehabilitation (19–25%) or no rehabilitation (24%). Future studies are required to test if animals taken into rehabilitation facilities have increased numbers of viruses and to determine if the viruses are clinically relevant and infectious.

This study shows that the gastrointestinal tract of sea turtles contains numerous different taxa of viruses in which most of detected viruses were genetically distinct, regardless of the turtle species, geographical location, or health condition. Because enteric sea turtle viruses have not been previously explored, it is not surprising that 14 out of 19 phylogenetically characterized viruses are putatively novel species with at least three putatively novel genera. This reflects that viral surveillance through metagenomic sequencing is crucial research to expand knowledge of the virosphere, particularly in species that are not commonly studied, e.g., wildlife and aquatic species. Notable, the number of detectable viruses (48 viruses in 19 individuals) are more than phylogenetically characterized viruses (19 viruses in 11 individuals), which could be impacted from the sample collection methods, transportation (wet ice vs. dry ice), and sampling of specimens prior to RNA extraction.

Not only can these data be investigated from the host perspective, but they contribute to the understanding of viral genetics and evolutions. For example, genetic recombination has been well documented in many viral taxa (RNA-RNA recombination of plant viruses [Sztuba-

Solińska et al., 2011] and DNA recombination of herpesvirus and bacteriophages [Pérez-Losada et al., 2015]). This genetic recombination suggested the history of genetic exchange, viral evolution, and explained the emergence of novel viral taxa (Kelly et al., 2016). Moreover, there is evidence for inter-family recombination of the newly emerging, but poorly characterized viruses (i.e., hepelivirals [genetically recombined with] astrovirids) (Kelly et al., 2016; Shi et al., 2018). This study found additional examples of this in the detected hepelivirals. The order Hepelivirales consists of positive-sense, monopartite, linear RNA viruses with two open reading frames [nonstructural polyprotein and capsid protein] including four formally recognized families (i.e., *Alphatetraviridae*, *Benyviridae*, *Matonaviridae*, *Hepeviridae,* and many unclassified hepelivirals [>50% of available hepeliviral sequences on GenBank]) (https://www.ncbi.nlm.nih.gov/nucleotide/ accessed June 28, 2024). Notably, these unclassified hepelivirals were reported to have genetic recombination of RdRp domain within the nonstructural polyprotein that is similar to hepevirids, but capsid protein that is more similar to astrovirids (Kelly et al., 2016; Shi et al., 2018). This also suggested the creation of a divergent clade termed “bastrovirus,” a group of unclassified hepelivirals. Due to unestablished taxonomic classification of bastroviruses, this term is not confirmed as a novel species or genus. In this study, the sea turtle feces associated hepevirals 1 was found to have genetic recombination similar to the “bastrovirus” group (see Fig. 2.4.3). However, there was only one closely related sequence (Rana hepevirus [GenBank accession no. MH330682]) available on GenBank which limited the evolutionary understanding of this divergent virus. Given that these viruses are genetically adjacent to hepevirids, which are known as hepatitis E viruses infecting human and multiple animals and “bastroviruses” detected from chordate species, it might be noteworthy to study these viruses for pathological research in the future. Additionally, other unclassified

hepeliviral group termed “bastro-like viruses” were reported to contain RdRp domains similar to other formally classified families (*Alphatetraviridae, Benyviridae,* and *Matonaviridae*). In this study, sea turtle feces associate bastro-like viruses 1, 2, and 3 were observed in two different turtles from the West Coast. These three sea turtle feces associate bastro-like viruses were found to have RdRp domains closely related to other bastro-like viruses. However, the capsid proteins of all sea turtle feces associate bastro-like viruses formed a distinct cluster with nodavirids but shared a common ancestor to other bastro-like viruses with high bootstrap support (93% bootstrap values) (see Fig. 2.4.4). This further suggests the inter-order genetic exchange of these newly emerging viruses and the importance of cataloging viral genetic information.

Identification of multiple viral taxa is observed within single animals, for instance; 3 viruses from different taxa (*Caliciviridae, Partitiviridae,* and *Totiviridae*) were identified in a green sea turtle with fibropapillomatosis (CM-C21249). The most abundant viral taxon in this study were hepelivirals (unclassified and bastro-like viruses), and picornavirals (dicistrovirids and unclassified picornavirals). Of note, multiple viruses of these taxa are often identified within single animals with high genetic diversity (at least two genetic populations). Whether this genetic variation represents detection of highly related viral strains or high mutation of a single strain within a single animal (quasispecies) remains to be determined. Similar observations were observed in several animals from different geographical origins (MA and FL).

Interestingly, many detectable viruses in this study were genetically related to viruses infecting marine invertebrates and fish were also identified. These included the sea turtle feces associated salovirus (84.4% amino acid identity to Atlantic salmon calicivirus [host: salmonids]), sea turtle feces associated dicistroviruses (41.38% amino acid identity to Beihai picorna-like virus 71 and 46.14% to Wenzhou picorna-like virus 28 [host: crustaceans]), sea turtle feces

associated picorna-like viruses (56.88% amino acid identity to Halhan virus 1 and 55.92% bivalve RNA virus 5G [host: mollusks]), and sea turtle feces associated ambidensovirus (75.18% amino acid identity to sea star-associated densovirus [host: echinoderms]. These animals are common food sources of sea turtle species and detection of viruses related to these species suggests that the exchange of viral nucleic acids of environment and sea turtles through the digestive system. For example, Atlantic salmon calicivirus was detected in a green turtle that was stranded and rehabilitated in a Georgia facility due to fibropapillomatosis. Based on the history, frozen salmonid fish were fed to the turtles at this rehabilitation facility. This suggests that the Atlantic salmon calicivirus originated from the foodstuffs. Similarly, given the high genetic similarity of many of the detected viruses to viruses-infecting mollusks, crustaceans, and fishes, it is likely that most detectable and phylogenetically characterized viruses in sea turtles in this study were from food and environmental sources. While this study only assessed for viral nucleic acid and not infectivity, this suggests that viral genetic materials could be carried by free-ranging sea turtles and transmitted to other populations through transoceanic migration as well as to other ill animals during rehabilitation within the same facility. Future studies are required to test if the viral infectivity remains and confirm if the viruses are acquired through food, and/or if they can form active infections in sea turtles.

Even though the infectivity and pathogenicity of these newly identified viruses are still unclear, it is crucial to establish baseline information on poorly studied species such as sea turtles. This could then help reflect viral dynamics in the larger ecosystem and provide a steppingstone for further hypothesis-driven research. Ultimately, further research is required to determine the impact of any one given virus; however, that research begins with cataloging viruses in this system. Moreover, the raw sequence data are deposited in NCBI’s archive. This

also allows for future studies to retrospectively analyze this data for phage virus sequences, or even bacterial or protozoal sequence analysis.

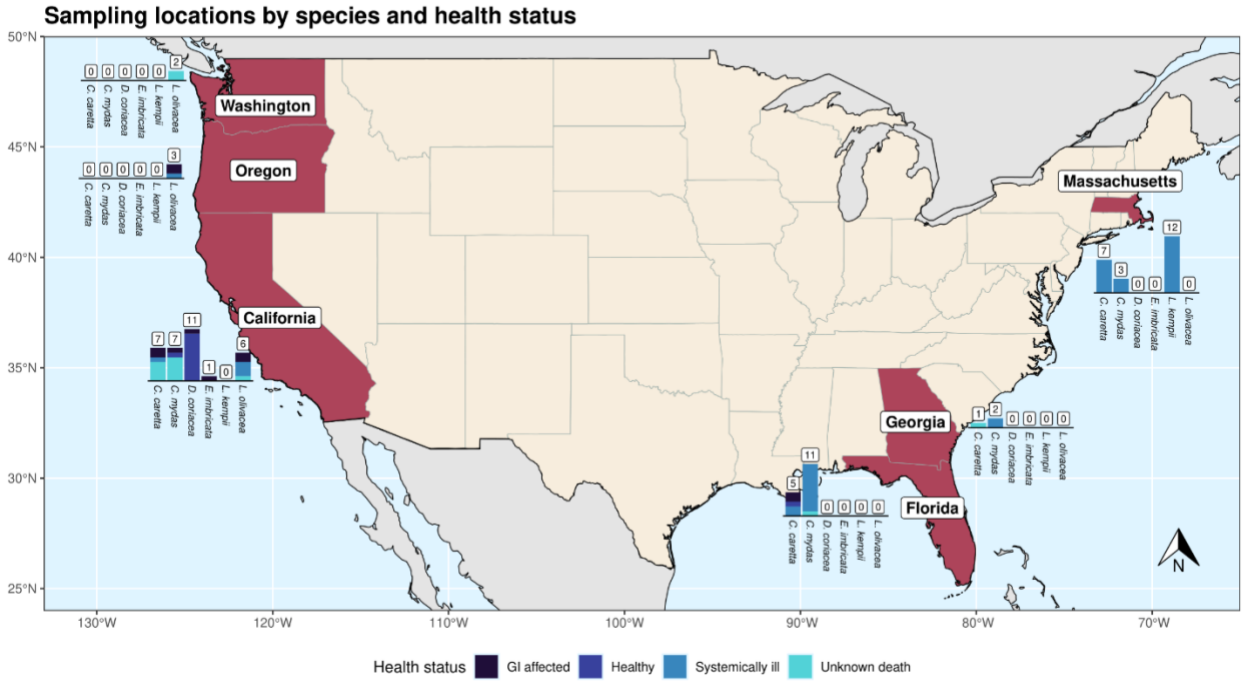
Figure and Table

Table 2.1: Sequence alignment and phylogenetic models

Viral family	**Aligned genetic regions**	**Alignment software**	**Alignment model**	**Phylogenetic software**	**Phylogenetic model**
Hepelivirales	RdRp and Capsid	Geneious	MUSCLE and CLUASTAL-Omega	MEGA X	rtREV+G+I+F (RdRp), LG+F+G (capsid)
Dicistroviridae	RdRp	Geneious	MUSCLE and CLUASTAL-Omega	MEGA X	LG+G+I+F
Caliciviridae	Viral protein 1	MEGA 11	ClustalW	MEGA 11	LG+G+I
Chuviridae	Large protein	MEGA X	ClustalW and MUSCLE	MEGA X	LG+G+I+F
Partitiviridae	Nonstructural protein	Geneious	MAFFT	MEGA X	LG+G
Orthototiviridae	Nonstructural protein	Geneious	MAFFT	MEGA 11	LG+G
Picobirnaviridae	RdRp protein	MEGA X	MUSCLE and ClustalW	MEGA X	*Neighbor-joining
Parvoviridae	Nonstructural protein 1	MEGA 11	ClustalW	MEGA 11	LG+G+F
Cressdnaviricota	Replicase	MEGA X	MUSCLE	MEGA X	rtREV+G+I+F

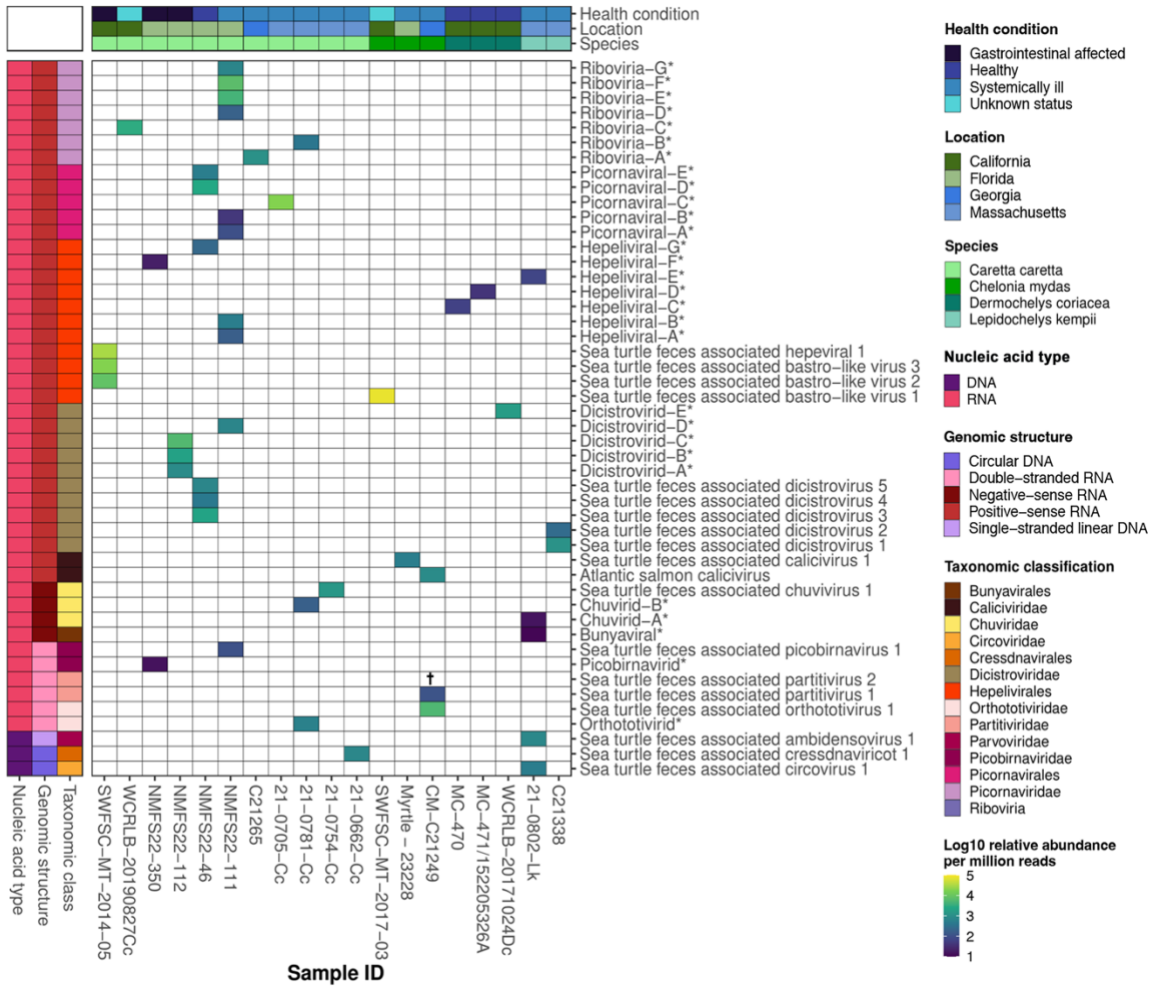
Table 2.1 The aligned genetic regions, alignment software, alignment models, phylogenetic software, and phylogenetic models used for phylogenetic analysis of each vial group. *All phylogenetic analyses were constructed under maximum-likelihood algorithm, except phylogeny of picobirnavirids (Neighbor-joining algorithm).

Figure 2.1: Sampling location by species and health status



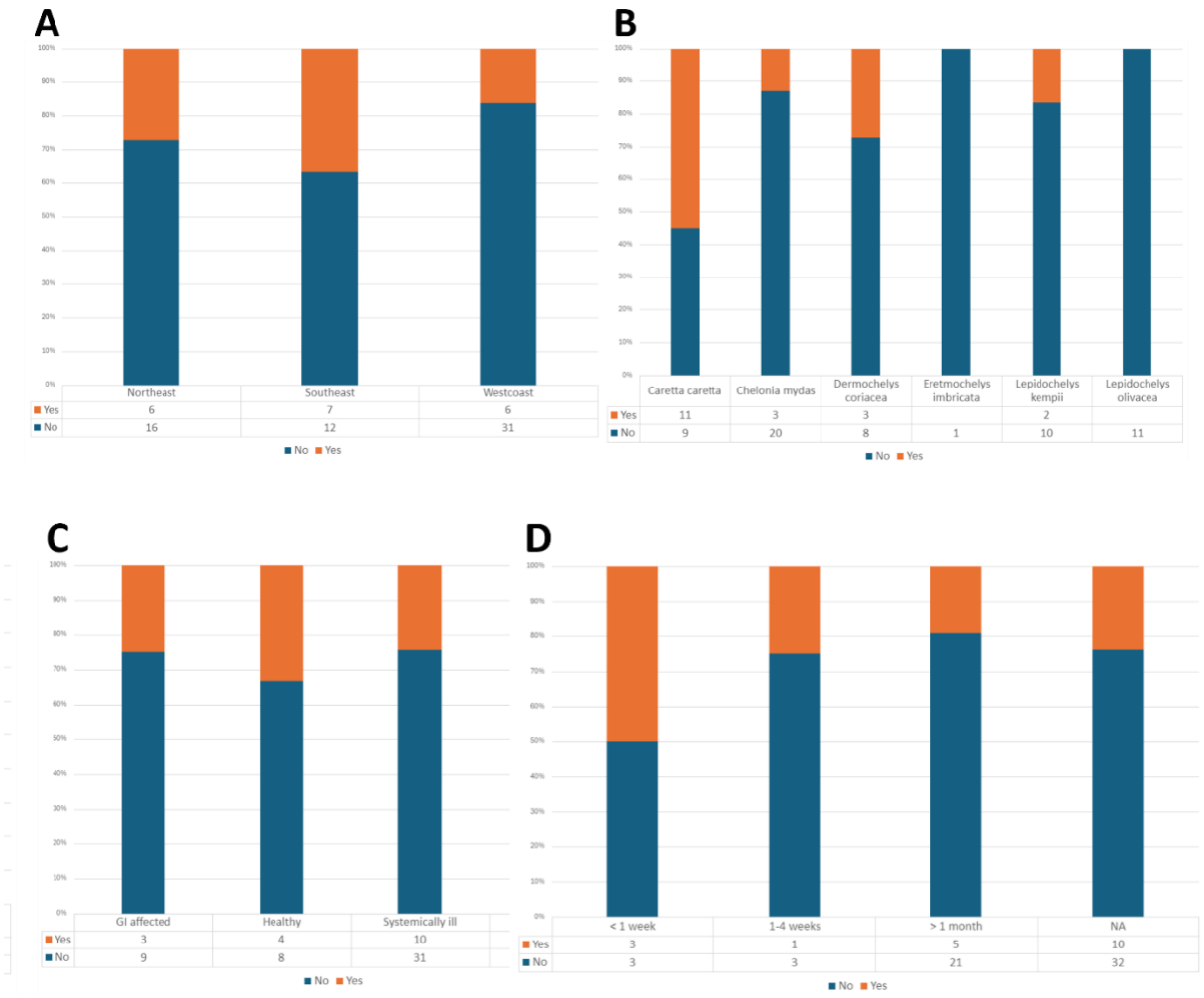
Map of the United State showing samples in this study. Burgundy highlights the states where samples were collected. The histogram adjacent to each state represents the number of turtles for each species from each state, with the health conditions being color coded.

Figure 2.2: Heatmap showing read abundance and diversity of phylogenetically characterizable gastrointestinal viruses in free-ranging sea turtles



The relative abundance of viral reads (log_{10} scale) and the diversity (viral nucleic acid types, viral genomic structures, and viral taxonomic classification) of detectable gastrointestinal viruses, based on health conditions of turtles, stranding locations, species of turtles, and age estimation of turtles. Viruses with complete names represent phylogenetically characterizable viruses. Viruses with short names represent those detected but not characterizable; letters denote different viruses from different animals. *Viral contigs detected only through random sequencing. †Viral reads only detected through semi-targeted, strand-switching sequencing.

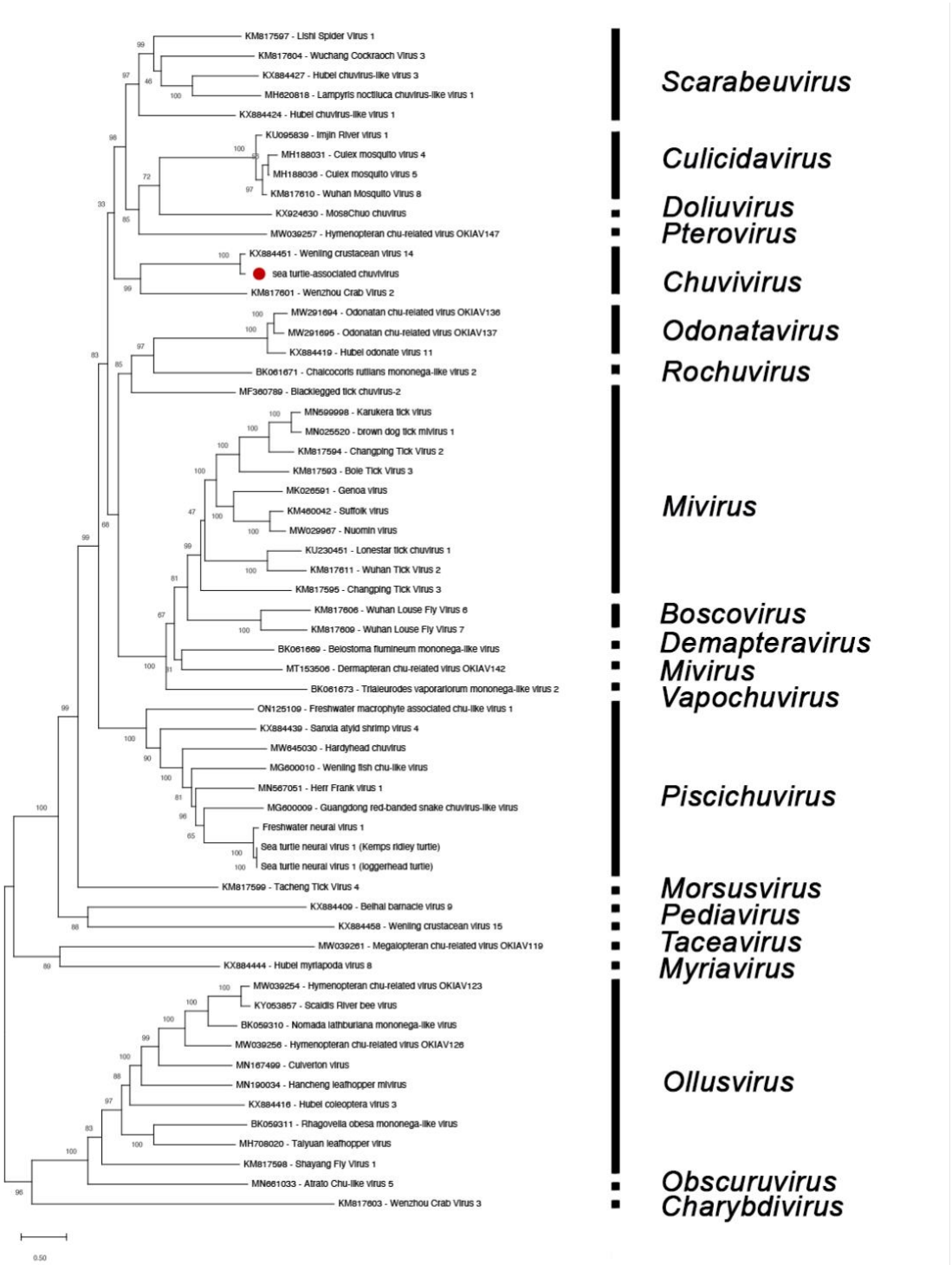
Figure 2.3 Comparison of turtles with and without detectable viruses in percentage



Turtles with and without detectable viruses (19 total) grouped by A) geographic category, B) species, C) by health condition, and D) Turtle from rehabilitation centers, sub-grouped by rehabilitation times.

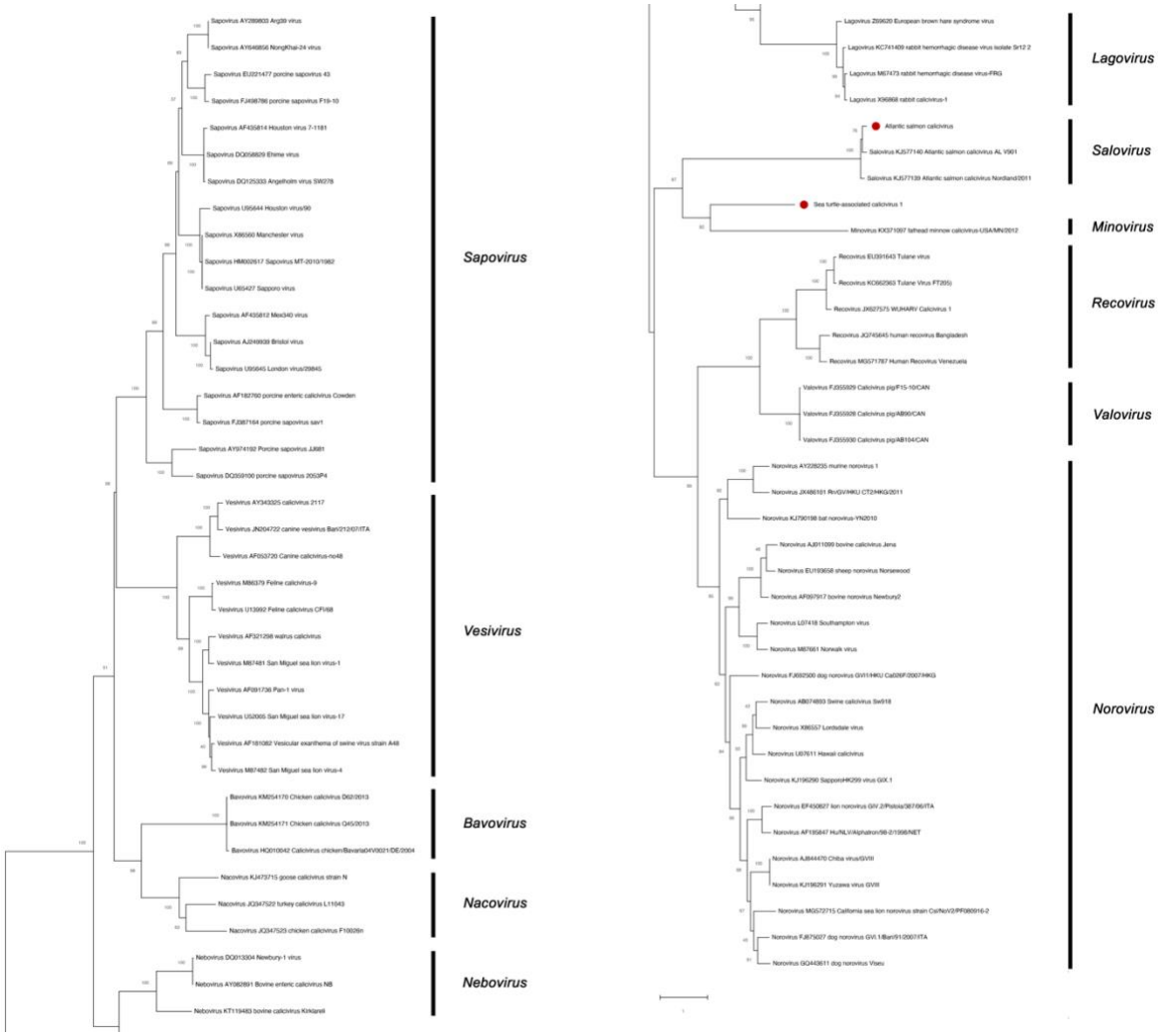
Figure 2.4: Phylogenetic analyses of each viral group. Red dots represent viral sequences identified in this study. Bars represent assigned viral groups.

Figure 2.4.1: The phylogenetic analysis of the sea turtle jingchuviral



Phylogenetic analysis of jingchuviral large protein (L) amino acid sequences from sea turtle feces associated chuvivirus 1 (red dot) and a subset of related sequences from GenBank. Complete L amino acid sequences were aligned by using ClustalW and refined by using MUSCLE with default settings. The phylogenetic analysis was performed on MEGA X by using the maximum-likelihood method and Le Gascuel matrix plus observed amino acid frequencies plus 5 discrete gamma categories distribution with parameter of 1.0879 plus invariant sites with 0.87% sites. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 60 sequences.

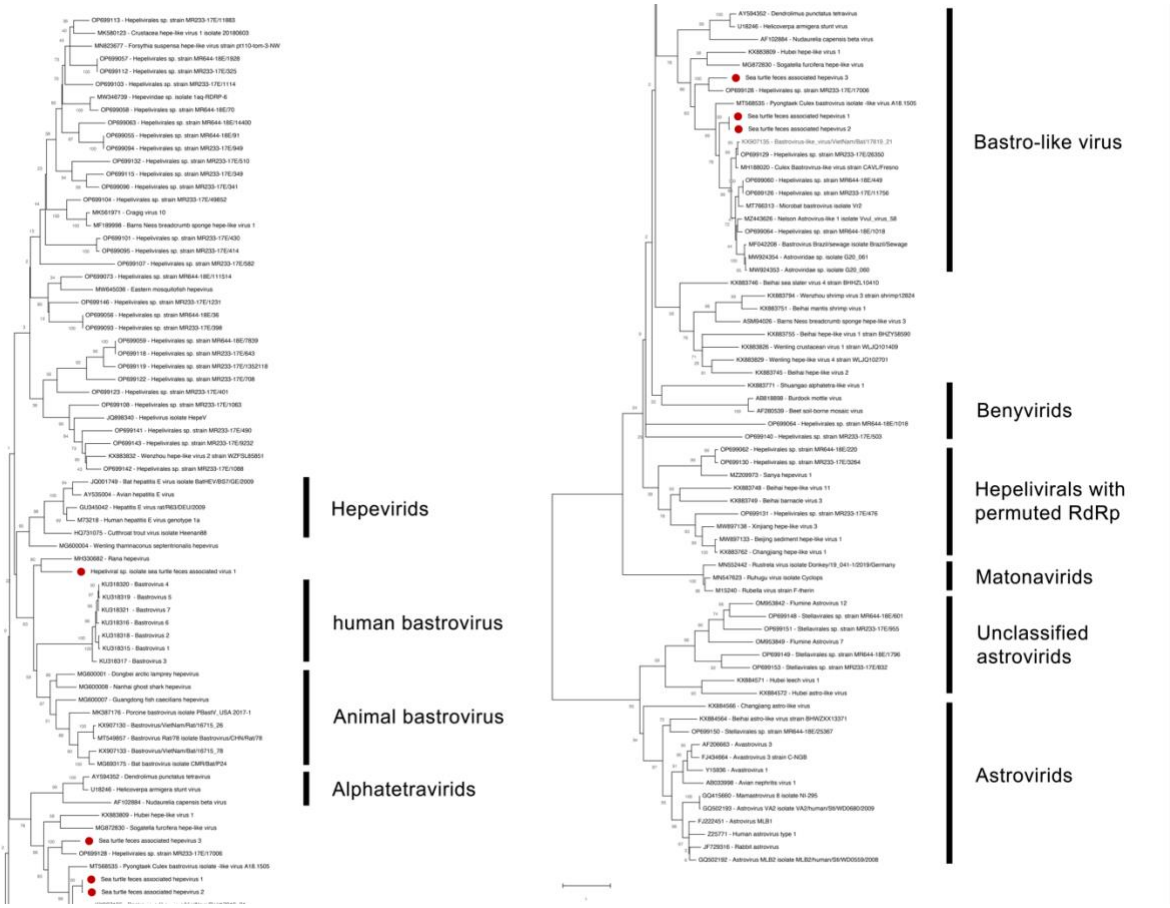
Figure 2.4.2: The phylogenetic analysis of the sea turtle calicivirids



Phylogenetic analysis of calicivirid VP1 amino acid sequences from sea turtle feces associated calicivirus 1, sea turtle feces associated salovirus 1 (red dots), and related sequences. VP1 amino acid sequences were aligned by using ClustalW with default settings. The phylogenetic analysis was performed on MEGA 11 by using the maximum-likelihood method and Le Gascuel matrix plus observed amino acid frequencies plus 5 discrete gamma categories distribution with parameter of 1.4688. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is

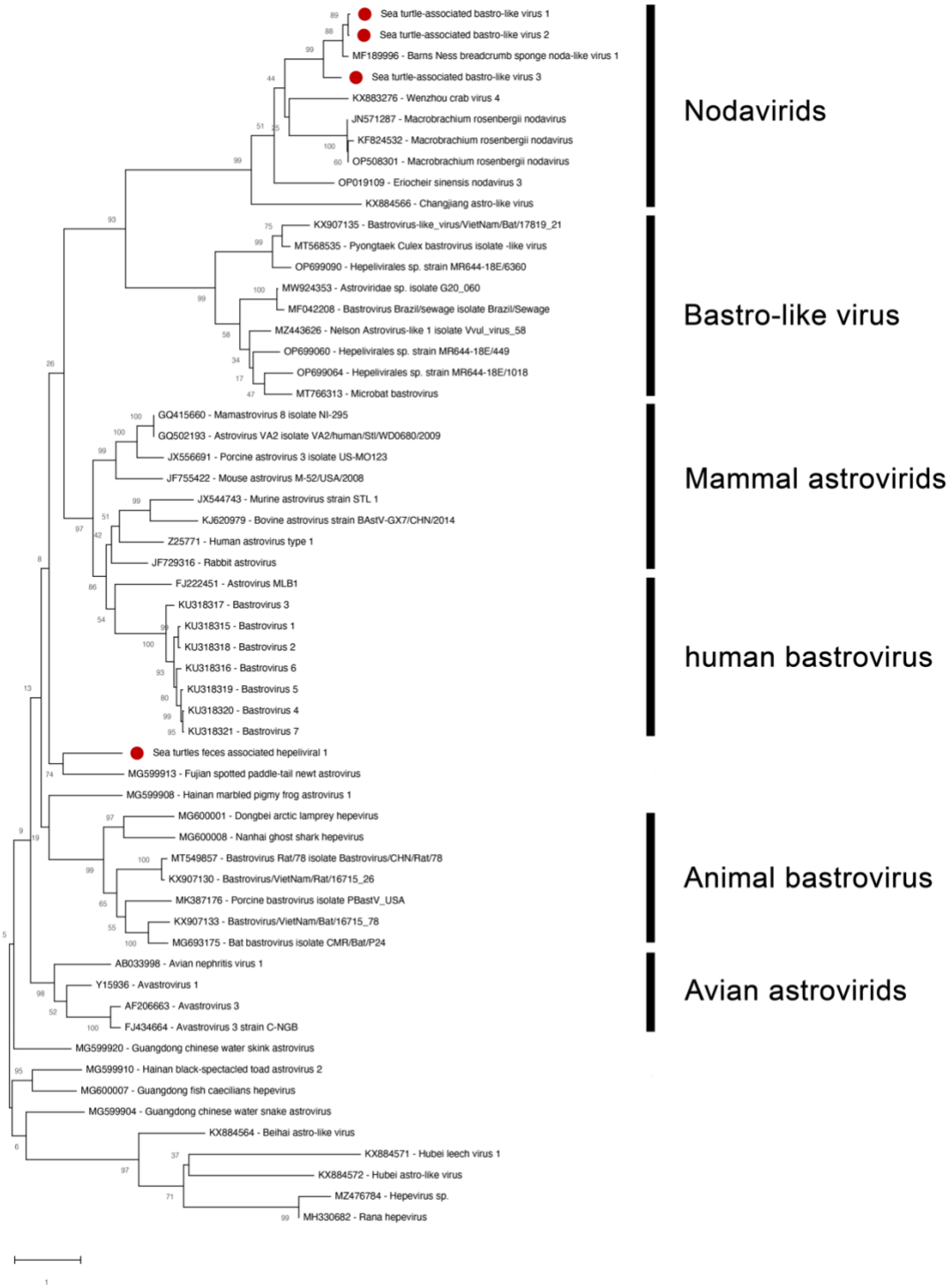
drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 75 aa sequences.

Figure 2.4.3: The RdRp-based phylogenetic analysis of the sea turtle hepelivirals



Phylogenetic analysis of hepeliviral RdRp amino acid sequences from sea turtle feces associated hepeliviral 1 and sea turtle feces associated bastro-like viruses 1, 2, and 3 (red dots), and related sequences. RdRp amino acid sequences were aligned by using MUSCLE and Clustal-Omega with default settings. The phylogenetic analysis was performed on MEGA X by using the maximum-likelihood method and rtREV plus observed amino acid frequencies plus 8 discrete gamma categories distribution with parameter of 1.1099 plus invariant sites with 1.41% sites. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 127 aa sequences.

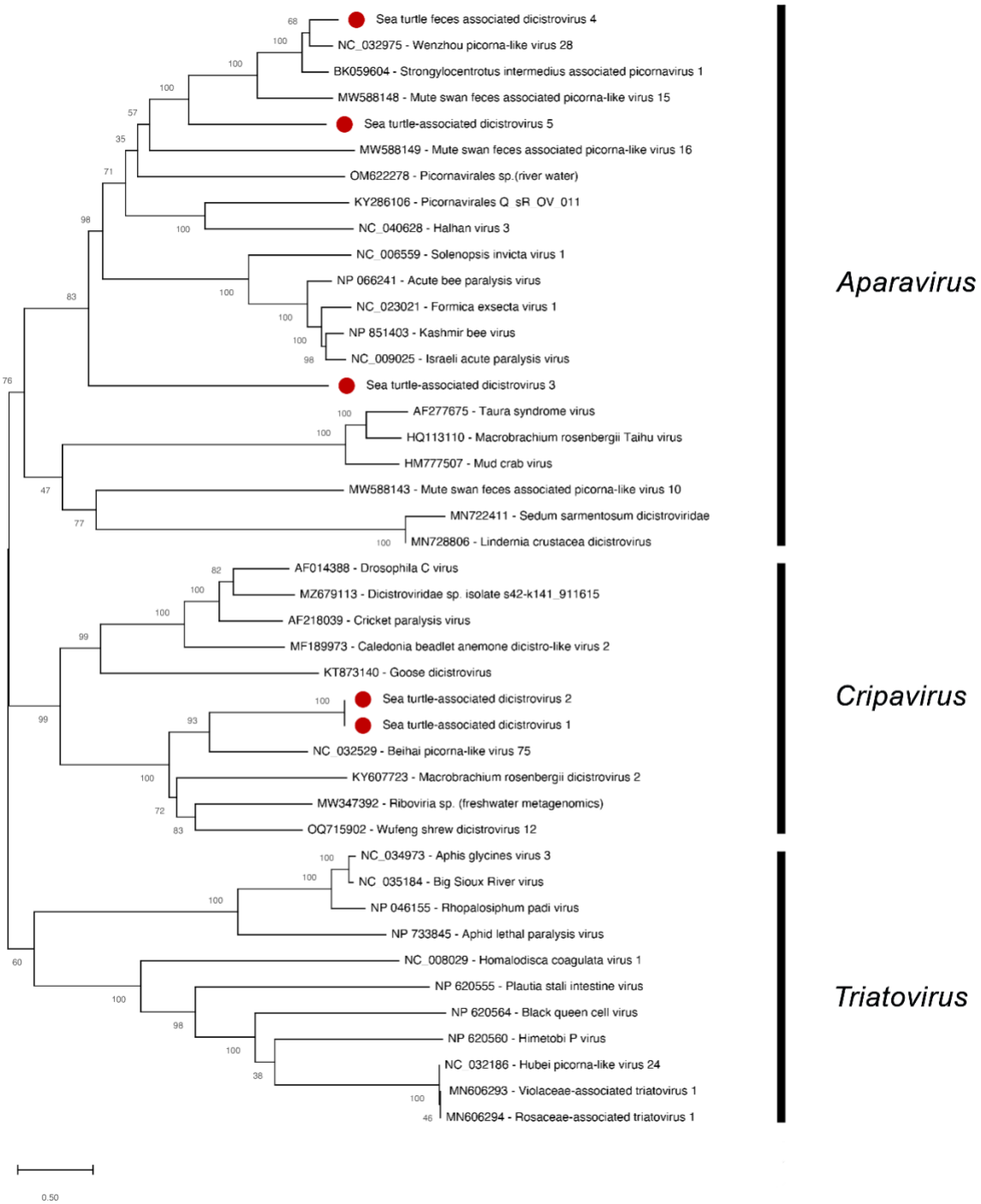
Figure 2.4.4: The capsid protein-based phylogenetic analysis of the sea turtle hepelivirals



Phylogenetic analysis of hepeliviral capsid amino acid sequences from sea turtle feces associated hepeliviral 1 and sea turtle feces associated bastro-like viruses 1, 2, and 3 (red dots), and related

sequences**.** Capsid amino acid sequences were aligned by using MUSCLE and Clustal-Omega with default settings. The phylogenetic analysis was performed on MEGA X by using the maximum-likelihood method and LG plus observed amino acid frequencies plus 8 discrete gamma categories distribution with parameter of 2.8242. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 32 aa sequences.

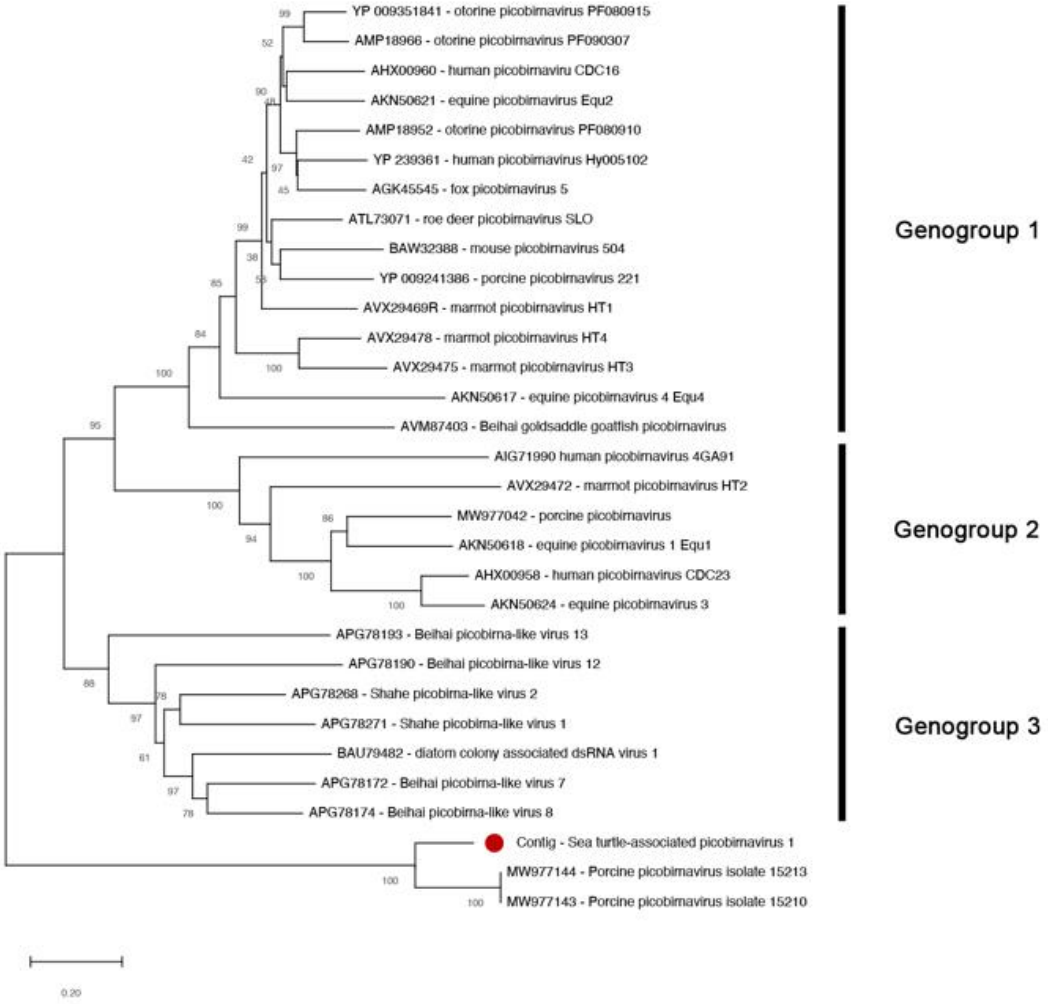
Figure 2.4.5: The phylogenetic analysis of the sea turtle dicistrovirids



Phylogenetic analysis of dicistrovirid RdRp amino acid sequences from sea turtle feces associated 1, 2, 3, 4, and 5 (red dots), and related sequences. RdRp amino acid sequences were

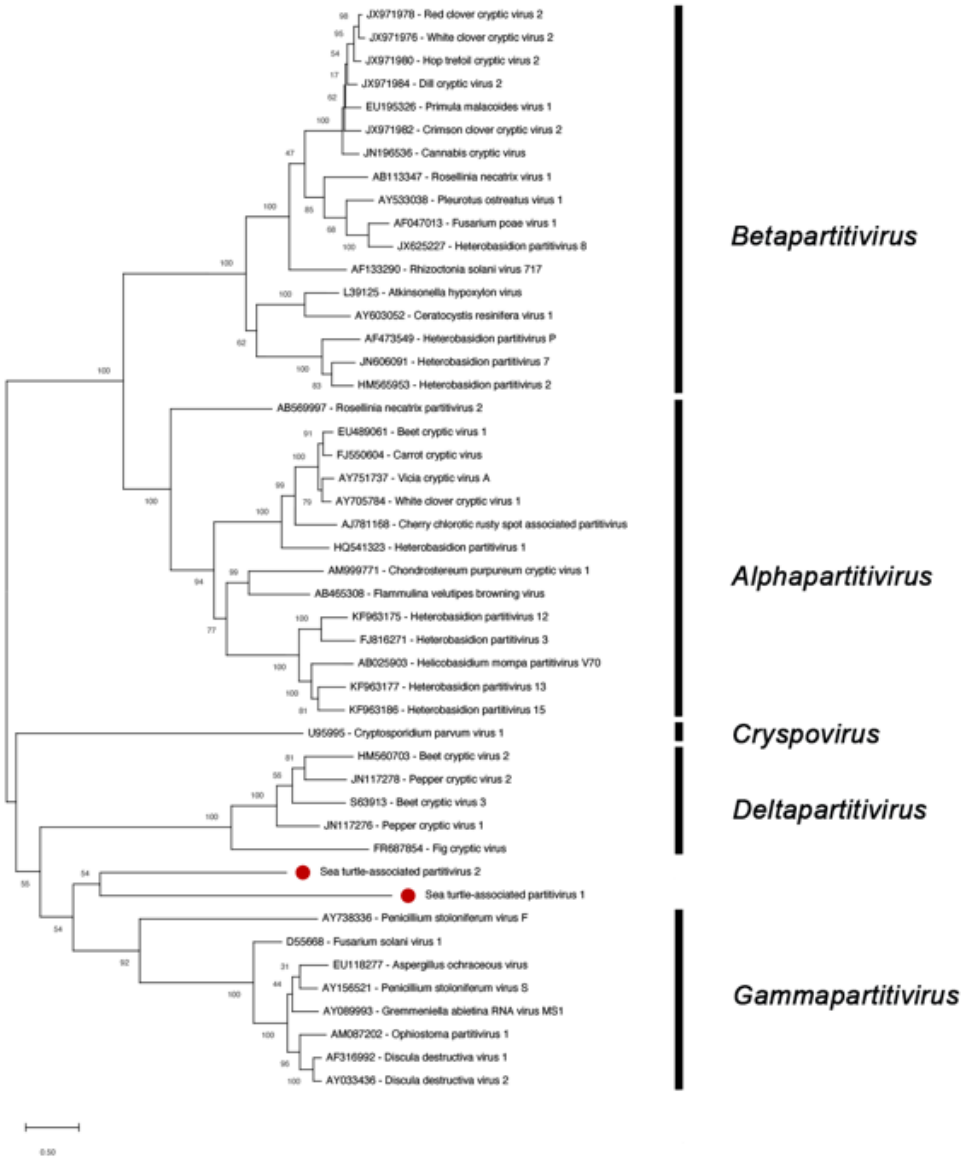
aligned by using MUSCLE and Clustal-Omega with default settings. The phylogenetic analysis was performed on MEGA X by using the maximum-likelihood method and LG plus observed amino acid frequencies plus 8 discrete gamma categories distribution with parameter of 1.5242 plus invariant sites with 1.49% sites. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 43 aa sequences.

Figure 2.4.6: The phylogenetic analysis of the sea turtle picobirnavirids



Phylogenetic analysis of picobirnavirid NS amino acid sequences from sea turtle feces associated picobirnavirus 1 (red dot) and reference sequences. NS amino acid sequences were aligned by using MUSCLE and ClustalW with default settings. The phylogenetic analysis was performed on MEGA X by using the neighbor-joining method. The substitution model was constructed with 1,000 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 31 aa sequences.

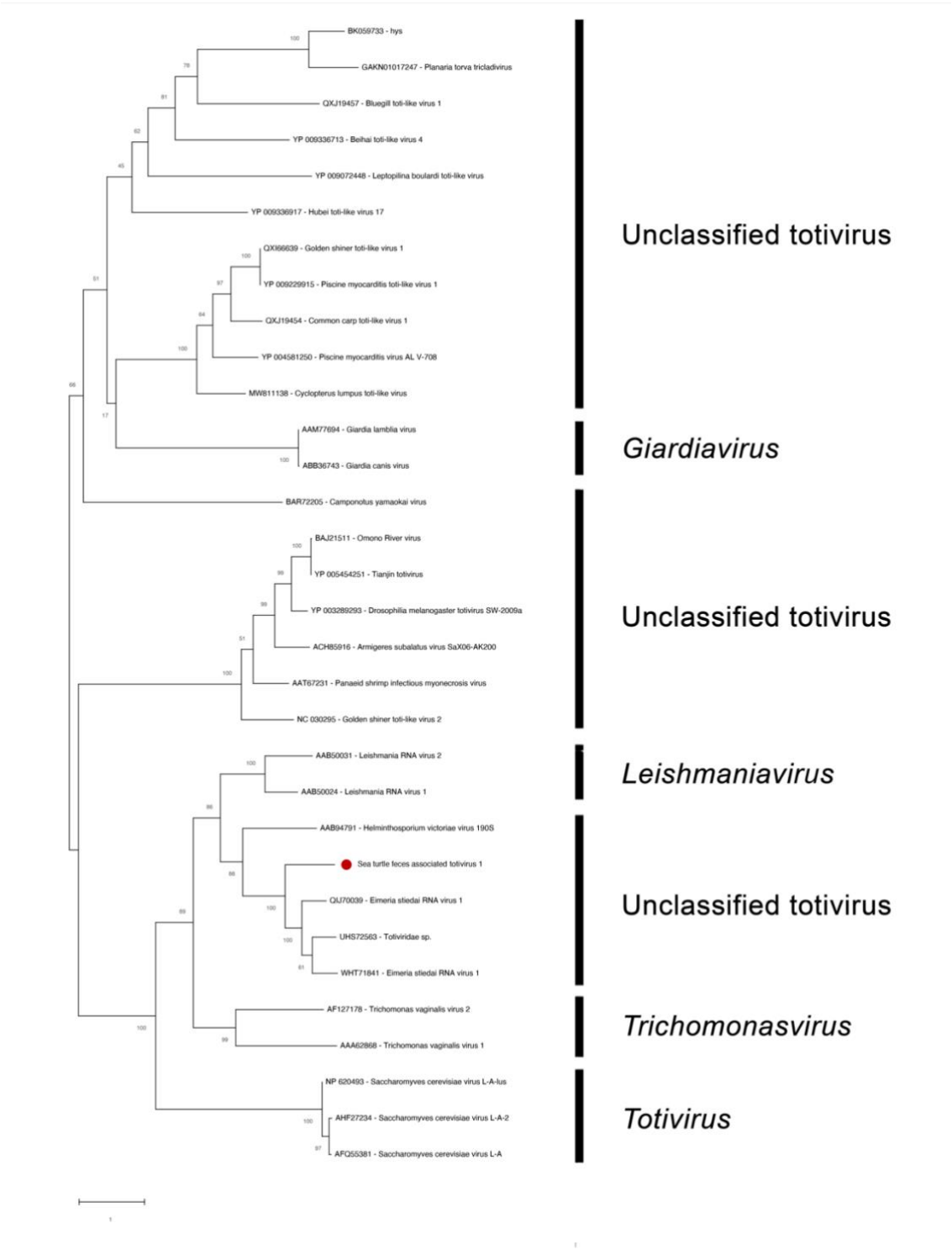
Figure 2.4.7, The phylogenetic analysis of the sea turtle partitivirids



Phylogenetic analysis of partitivirid NS amino acid sequences from sea turtle feces associated partitiviruses 1 and 2 (red dots), and related sequences. NS amino acid sequences were aligned by using MAFFT with default settings. The phylogenetic analysis was performed on MEGA 11 by using the maximum-likelihood method and Le Gascuel matrix plus 5 discrete gamma categories distribution with parameter of 1.6920. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a

percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 47 aa sequences.

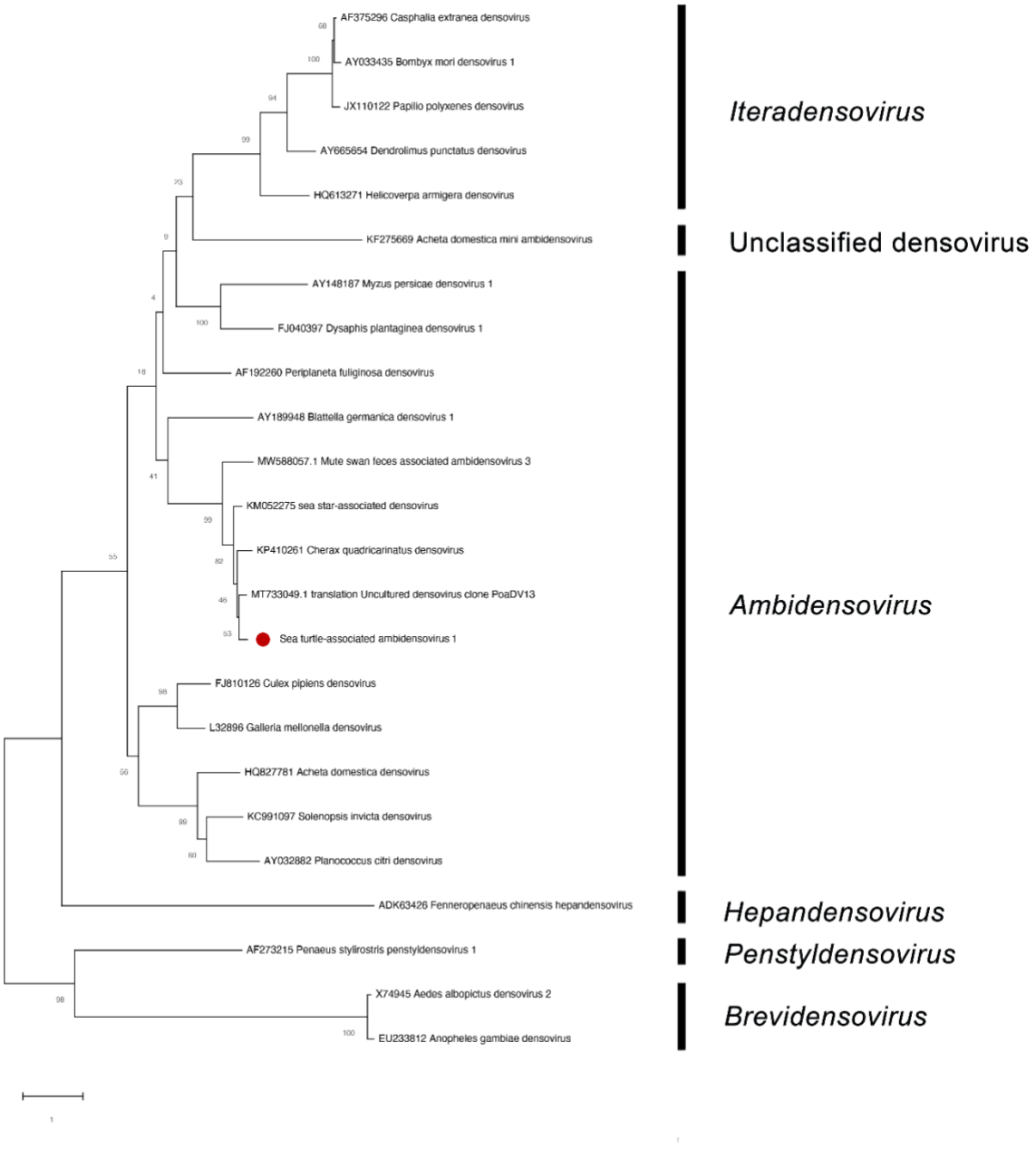
Figure 2.4.8: The phylogenetic analysis of the sea turtle orthotitivirids



Phylogenetic analysis of orthototivirid NS amino acid sequences from sea turtle feces associated orthototivirus (red dots) and related sequences. NS amino acid sequences were aligned by using

MAFFT with default settings. The phylogenetic analysis was performed on MEGA 11 by using the maximum-likelihood method and Le Gascuel matrix plus 8 discrete gamma categories distribution with parameter of 1.4566. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 32 aa sequences.

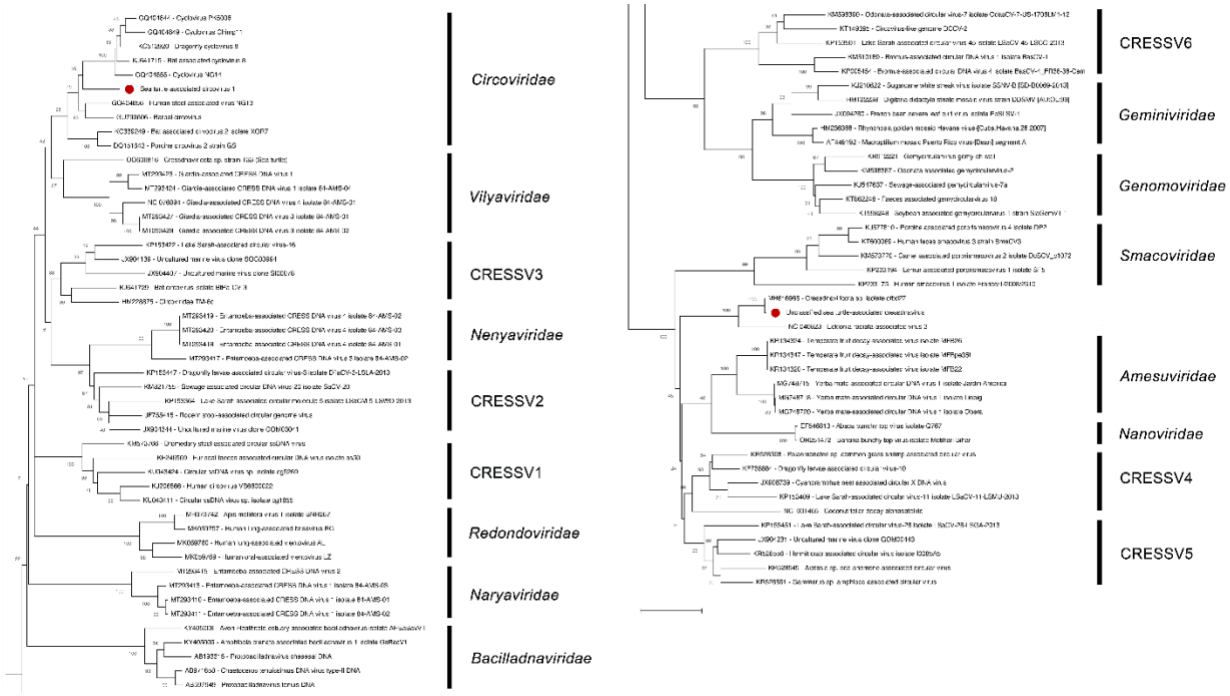
Figure 2.4.9: The phylogenetic analysis of the sea turtle densovirins



Phylogenetic analysis of densovirins NS1 amino acid sequences from sea turtle feces associated ambidensovirus 1 (red dots) and related sequences. NS1 amino acid sequences were aligned by using ClustalW with default settings. The phylogenetic analysis was performed on MEGA 11 by using the maximum-likelihood method and Le Gascuel matrix plus observed amino acid

frequencies plus 5 discrete gamma categories distribution with parameter of 1.2783. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 24 aa sequences.

Figure 2.4.10: The phylogenetic analysis of the sea turtle cressdnaviricots



Phylogenetic analysis of cressdnaviricots replicase (rep) amino acid sequences from sea turtle feces associated circovirus 1 and sea turtle feces associated cressdnaviricot 1 (red dots), and related sequences. Rep amino acid sequences were aligned by using MUSCLE with default settings. The phylogenetic analysis was performed on MEGA X by using the maximum-likelihood method and rtREV plus observed amino acid frequencies plus 8 discrete gamma categories distribution with parameter of 1.2193 plus invariant sites with 0.35% sites. The substitution model was constructed with 500 bootstrap replicates; the node numbers represent the bootstrap values (expressed as a percentage of all trees). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis included 89 aa sequences.

CHAPTER 3

PISCICHUVIRUS-ASSOCIATED SEVERE MENINGOENCEPHALOMYELITIS IN AQUATIC TURTLES, UNITED STATES, 2009-2022

Laovechprasit W, Young KT, Stacy BA, Steven TB, Ossiboff RJ, Vann JA, Subramanian K, Agnew DW, Howerth EW, Zhang J, Whitaker S, Walker A, Orgill AM, Howell LN, Shaver DJ, Donnelly K, Foley AM, Stanton JB. 2024. *Emerging Infectious Diseases*, 30(2):280-288.

Reprinted here with permission of the publisher.

Abstract

Viruses from a new species of piscichuvirus were strongly associated with severe lymphocytic meningoencephalomyelitis in several free-ranging aquatic turtles from 3 coastal US states during 2009–2021. Sequencing identified 2 variants (freshwater turtle neural virus 1 [FTuNV1] and sea turtle neural virus 1 [STuNV1]) of the new piscichuvirus species in 3 turtles of 3 species. In-situ hybridization localized viral mRNA to the inflamed region of the central nervous system in all 3 sequenced isolates and in 2 of 3 additional nonsequenced isolates. All 3 sequenced isolates phylogenetically clustered with other vertebrate chuvirids within the genus Piscichuvirus. FTuNV1 and STuNV1 shared ≈92% pairwise amino acid identity of the large protein, which narrowly places them within the same novel species. The in-situ association of the piscichuviruses in 5 of 6 turtles (representing 3 genera) with lymphocytic meningoencephalomyelitis suggests that piscichuviruses are a likely cause of lymphocytic meningoencephalomyelitis in freshwater and marine turtles.

Introduction

Wild populations of aquatic turtles are imperiled because of anthropogenic activities (e.g., consumption, collection, fisheries bycatch) (Senko et al., 2022); more than half (186/357) of the recognized species of aquatic turtles in the world are designated as critically endangered, endangered, or vulnerable (Rhodin et al., 2021). In addition to anthropogenic threats, infectious agents also negatively affect free-ranging turtles. For example, chelonid herpesvirus 5 (Scutavirus chelonidalpha5: Alphaherpesvirinae) is associated with transmissible fibropapillomatosis in sea turtles around the world (Greenblatt et al., 2005), and epizootic outbreaks of meningoencephalitis in Florida freshwater turtles have been attributed to the recently discovered turtle fraservirus 1 (Fraservirus testudinis: Tosoviridae) (Waltzek et al., 2022). However, learning about infectious agents in such turtles is complicated by their aquatic nature and cryptic lifestyles, which prevents full appreciation of the threat posed by viruses to free-ranging turtles.

Recently, numerous viruses in captive and free ranging nondomesticated animals have been identified. Among the newly discovered viruses, chuvirids (class Monjiviricetes, order Jingchuvirales, family Chuviridae) (Harding et al., 2022b) are of particular interest. First, they have a broad host range (e.g., phototrophs, a wide array of invertebrates, and vertebrates [fish and snakes]) (Hahn et al., 2020; Han et al., 2020; Harding et al., 2022b; Li et al., 2015; Shi et al., 2016, 2018). Second, the genomic structure of viruses in that family is unusual (Dezordi et al., 2020; Di Paola et al., 2022). Although other jingchuvirals have nonsegmented linear genomes, chuvirids have been reported to have circular segmented, circular nonsegmented, linear segmented, and linear nonsegmented genomes (Di Paola et al., 2022; Hahn et al., 2020; Han et al., 2020; Shi et al., 2016, 2018). Phylogenetic analysis of the large (L) protein suggests that

jingchuvirals have a unique history among viruses in the class Monjiviricetes (Di Paola et al., 2022; Li et al., 2015). Although recent discoveries of chuvirids and their varying genomic structures draw interest from a virologic standpoint, the pathogenicity of chuvirids and jingchuvirals has not been confirmed (Di Paola et al., 2022). To our knowledge, only 1 published study has demonstrated any jingchuviral from an ill animal: a piscichuvirus (Herr Frank virus 1 [HFrV1]; Chuviridae: Piscichuvirus franki) in clinically ill snakes (3 of 4 boa constrictors [Boa constrictor constrictor] with boid inclusion body disease; all 4 boas were positive for reptarenaviruses) (Argenta et al., 2020). However, because no piscichuviral in-situ studies were performed with those snakes and meningoencephalitis was not reported, clinical significance of chuvirids remains unresolved.

To identify the potential cause of lymphocytic meningoencephalomyelitis in several aquatic turtles, we randomly sequenced central nervous system (CNS) tissues from 3 affected turtles and performed in-situ hybridization (ISH) on CNS tissues of 6 turtles. We obtained complete genomes for 3 isolates, providing phylogenetic classification and in silico identification of conserved secondary structures at the genome termini and a hypothetical fourth open reading frame (ORF). The relative dissimilarity between the freshwater and sea turtle piscichuviruses raises questions regarding jingchuviral speciation criteria.

Materials and methods

Samples and Pathologic Examination

We examined isolates from 6 aquatic turtles with idiopathic CNS inflammation (meningoencephalitis to meningoencephalomyelitis): 1 freshwater (alligator snapping turtle, morphologically consistent with either Macrochelys temminckii or M. apalachicolae) and 5

marine (1 Kemp's ridley [Lepidochelys kempii] and 4 loggerhead [Caretta caretta]), by using sequencing (turtles 1–3), ISH (turtles 1–6), or both (Table 1; Figure 1). We used 6 additional retrospective cases as controls: 2 loggerhead turtles with spirorchiid-induced mononuclear meningoencephalitis, 1 Kemp's ridley turtle with bacterial encephalitis, and 3 turtles lacking meningoencephalomyelitis (1 of each species).

All turtles died spontaneously or were euthanized via pentobarbital overdose if their illness was advanced. Gross necropsy included systematic evaluation of all organ systems. We determined sexual maturity by evaluating the reproductive system. We aseptically collected fresh tissue samples, including from the brain and spinal cord (turtle 1–3), and stored them at –80°C. We preserved samples in neutral-buffered 10% formalin fixative for 24–48 hours before processing for histopathologic analysis.

Random Sequencing

We homogenized a section of cerebrum from the alligator snapping turtle (turtle 1) and a section of the brainstem from the loggerhead turtle (turtle 3) in 450 μL 1× phosphate-buffered saline by using a QIAGEN TissueLyser LT (https://www.qiagen.comExternal Link) at 35 Hz for 2 minutes with a 5-mm sterile stainless-steel bead (QIAGEN). Homogenized samples underwent viral enrichment (Conceição-Neto et al., 2018; Vibin et al., 2018). We extracted RNA by using Trizol LS Reagent (Thermo Fisher Scientific, https://www.thermofisher.comExternal Link) and prepared the cDNA library as previously described (Kelsey T. Young, Jazz Q. Stephens & David E. Stallknecht, Kiril M. Dimitrov, Salman L. Butt, 2022; Young et al., 2021), using manufacturer-suggested kits for ligation-based sequencing of amplicons (SQK-LSK110 with EXP-PBC096) and sequencing on a FLO-MIN106 R9.4.1 flow cell in a MinION Mk1B sequencing device (Oxford Nanopore Technologies, https://nanoporetech.comExternal Link).

The postsequencing analysis workflow followed the randomly primed, MinION-based sequencing as previously described (Kelsey T. Young, Jazz Q. Stephens & David E. Stallknecht, Kiril M. Dimitrov, Salman L. Butt, 2022; Young et al., 2021). We first accomplished screening for potential pathogens by pairwise alignment of reads (BLAST, https://blast.ncbi.nlm.nih.gov/Blast.cgiExternal Link]) to the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.govExternal Link) nonredundant nucleotide database (updated June 4, 2022) through the Georgia Advanced Computing Research Center (https://gacrc.uga.eduExternal Link) by using default settings. After identifying piscichuvirus-like reads, we used a long-read aligner (Centrifuge version 1.0.4) (Kim et al., 2016) with a custom index to filter out host reads by using the green sea turtle genome (GCF_015237465.1_rCheMyd.pri.v2, accessed July 2021) combined with a publicly available Centrifuge index for bacteria and archaea (https://ccb.jhu.edu/software/centrifuge/manual.shtmlExternal Link). We mapped the remaining reads >50 nt to the closest piscichuviral genome to iteratively assemble the turtle virus genomes in Geneious Prime 2019.1.3 (https://www.geneious.comExternal Link). We confirmed assembly with de novo assembly by using Flye (Kolmogorov et al., 2020).

We extracted RNA from turtle 2 by using an RNeasy Mini kit (QIAGEN) and generated a cDNA library by using a NEBNext Ultra RNA Library Prep Kit (Illumina, https://www.illumina.comExternal Link), which we sequenced on the iSeq 100 Sequencing System (Illumina). We processed raw data to remove host reads by first running Kraken version 2 (Wood et al., 2019) against a custom database created by using the green sea turtle genome (NCBI assembly GCA_000344595, accessed March 2020). We assembled the remaining paired-end reads (742,979) by using SPAdes version 3.15.3 with default parameters (Bankevich et al.,

2012). We then subjected the assembled contigs to BLASTX searches in OmicsBox version 2.0 (BioBam Bioinformatics, https://www.biobam.comExternal Link) against the NCBI nonredundant protein database.

Genome Analysis

We annotated the ORFs by using the NCBI ORFfinder (https://www.ncbi.nlm.nih.gov/orffinderExternal Link) and by manually comparing them with the annotations of other piscichuviruses. We filled the gaps in the consensus sequences by PCR for MinION (turtles 1 and 3) or Sanger (turtle 2) sequencing. To predict RNA secondary structures at genomic termini, we used the Vienna package RNAfold tool (Hofacker, 2003).

In-situ hybridization (ISH)

We used ISH to localize turtle piscichuviral mRNA in all 6 turtles with idiopathic meningoencephalomyelitis and 6 control turtles. RNAscope 2.5 HD double z-probes (Advanced Cell Diagnostics, Inc., http://acdbio.comExternal Link) were designed by using the large protein gene (L) gene (of freshwater turtle neural virus 1 [FTuNV1] probe: 1375–2368, sea turtle neural virus 1 [STuNV1] probe: 1535–2585). We used probes targeting testudine rRNA as positive probe controls and the Bacillus subtilis dihydrodipicolinate reductase (DapB) gene as negative probe controls (Advanced Cell Diagnostics, Inc.). RNAscope (Advanced Cell Diagnostics, Inc.) ISH was performed by following the manufacturer’s protocols and analyzed by light microscopy.

Phylogenetic Analyses

To infer evolutionary relationships, we phylogenetically analyzed the translated L ORFs from FTuNV1, 2 isolates of STuNV1, and 56 other complete jingchuvirals in MEGA X (Kumar et al., 2018). We performed multiple sequence alignments of each coding sequence separately by using ClustalW (https://www.clustal.orgExternal Link) and MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle
External Link) with default settings. We based selection of the best substitution model of aligned amino acid sequences on the lowest Bayesian information criterion and Akaike scores and used the best substitution model analysis for maximum-likelihood analysis. We constructed phylogenetic analyses by using the maximum-likelihood method and the Le Gascuel matrix plus observed amino acid frequencies plus 5 discrete gamma categories distribution plus invariant sites substitution model with 500 bootstrap replicates. We used subtree-pruning-regrafting level 3 (MEGA X) for maximum-likelihood tree inference and used all gaps and missing data to construct the phylogenetic tree.

Results

Animal Histories and Pathologies

Six test turtles (turtles 1–6) with histories of persistent neurologic signs (e.g., weakness, lethargy, asymmetric buoyancy, circling, head tremors, cervical ventroflexion, and unresponsiveness) were collected from the southeastern United States; 5 were piscichuvirus-positive (see Table 1; Figure 1). The CNSs of all turtles were grossly normal; however, those turtles had moderate to severe, multifocal to diffuse mononuclear meningoencephalomyelitis with severe lymphoplasmacytic cuffs. Most severely affected were the cerebrum, optic tectum, and

cerebellum (Figure 2, panels A, C). The associated neuroparenchyma was vacuolated, and some neurons exhibited central chromatolysis (Figure 2, panel C). Subsequent results from ancillary testing (i.e., Ziehl-Neelsen staining [turtles 2–4] or PCR for herpesvirus [turtle 1] [24] and turtle fraservirus 1 [turtle 4] [4]) were negative. Turtles 1–3 were used for metagenomic sequencing; turtles 1–6, along with 6 control turtles, were used for ISH.

Viral Genomes

FTuNV1/Alligator Snapping Turtle

Through reference-based alignment of all reads that passed quality filtering when BLASTN was used, we detected only a few piscichuviral-like reads, including hits to Wēnlǐng fish chu-like virus (WFClV; Piscichuvirus wenlingense, GenBank accession no. MG600011) and HFrV1 (GenBank accession nos. MN567051, MN567057, MN56703). Mapping filtered reads to WFClV in Geneious resulted in 1,491 piscichuviral reads and a draft FTuNV1 genome. After targeted sequencing to close gaps, we obtained a 10,781-nt complete FTuNV1 genome with at least 10 times coverage: (Table 2; isolate FTuNV1/Alligator_snapping_turtle/Florida/ST0994/ 2009, GenBank accession no, OQ547744).

STuNV1/Kemp's Ridley Turtle

Using BLASTX, we identified 4 de novo contigs with highest similarity to HFrV1 (GenBank accession no. MN567051) and Guǎngdōng red-banded snake chuvirus-like virus (GRSCV; Piscichuvirus lycodontis, GenBank accession no. MG600009). After performing targeted sequencing to close gaps, we obtained a 10,839-nt complete genome (GenBank accession no. OQ547745).

STuNV1/Loggerhead Turtle

After initially detecting piscichuvirus-like reads by using Centrifuge (with a custom index [16] containing FTuNV1 and STuNV1), we identified piscichuviral reads by mapping to the Kemp's ridley STuNV1 consensus sequence. That process resulted in 258 reads building a 10,839-nt complete genome with at least 10 times coverage, except for the first 9 bases of 5′ terminus, which had 6–9 times coverage (GenBank accession no. OQ547746.)

Genome Comparison of Piscichuviruses

The genomic structures of FTuNV1 and STuNV1 were linear, nonsegmented, and had the following ORF orientation: 3′-ORF4, nucleoprotein (N), glycoprotein (G), large protein (L)-5′. In addition, we identified that the genomic termini were complementary (i.e., inverted terminal repeat sequences), and in silico modeling predicted the formation of a genomic panhandle structure for each virus (see Figure 3).

To taxonomically classify FTuNV1 and STuNV1, we used the recent jingchuviral taxonomic classification, which is based on the L protein amino acid identity (*21*). The percentage pairwise amino acid identities <90%, <31%, and <21% support the differentiation of jingchuvirals as novel species, genera, and families, respectively (*12*). We determined that the L protein of the STuNV1 isolates had the same predicted length (2,145 aa; Table 2) and were 99.3% identical (Appendix); thus, they were considered to be 2 isolates of a single variant. The FTuNV1 and STuNV1 L protein sequences had the same predicted length but were ≈92% identical (Appendix), which is close to the initially proposed speciation cutoff criterion. Thus, all

3 isolates are considered to be within the same new species, but FTuNV1 and STuNV1 are proposed as variants within this species.

We identified a fourth ORF4 in FTuNV1 and STuNV1. The predicted ORF4 amino acid sequence length was the same in all 3 turtle isolates (105 aa; Table 2), and the predicted amino acid sequences were identical for the 2 isolates of STuNV1. Predicted identity between FTuNV1 and STuNV1 was ≈77.36% (Appendix). Of note, we also identified putative, but unannotated, ORF4s in previously NCBI-deposited piscichuviral sequences (Table 2). Among the ORFs, ORF4 is predicted to have the most amino acid variation across piscichuviruses (Appendix). For piscichuviruses that were previously deposited in GenBank, the putative ORF4 was 225–276 nt (74–91 aa) long and 3′ prime of the N ORF (Table 2). In addition, ORF4 lacked evidence of transmembrane domains (https://services.healthtech.dtu.dk/TMHMM-2.0External Link), signal sequences (https://services.healthtech.dtu.dk/SignalP-5.0External Link), and N-linked glycosylation sites (https://services.healthtech.dtu.dk/NetNGlyc-1.0External Link).

Phylogeny of Jingchuvirals

Phylogenetic analysis of the predicted L protein amino acid sequences from 59 chuvirids demonstrated that FTuNV1 and both STuNV1 isolates clustered with other piscichuviruses; bootstrap value was 100%. All piscichuviruses detected from reptiles form a single branch with a 92% bootstrap value (Figure 4). Similarly, all piscichuviruses detected from vertebrates form a single branch with a 100% bootstrap value (Figure 4).

Localization of FTuNV1 and STuNV1 Nucleic Acid

To more definitively associate piscichuviral infection with the clinical and histopathologic findings, we conducted RNAscope ISH on all 3 isolates that were positive for piscichuvirus by sequencing (from turtles 1–3), all 3 of which demonstrated viral RNA within inflamed areas of the CNS (Figure 2, panels C, D). ISH testing of 3 additional turtles (turtles 4–6), in which no sequencing was performed indicated that 2 of the 3 were positive. Thus, 5 of 6 cases that were originally considered idiopathic were proposed to be associated with piscichuvirus (Table 1).

ISH demonstrated disseminated, strong, punctate reactivity for piscichuviral RNA in areas of inflammation throughout the CNS. Piscichuviral RNA was detected predominantly in the gray matter (Figure 2, panel B), most notably within the cytoplasm of large neurons (often chromatolytic), small neurons, glial cells, but occasionally in ependymal cells with subtle intensity (Figure 2, panel D). Testing of nonneural tissues of turtles 1 (tested for FTuNV1 nucleic acid) and 2 (tested for STuNV1 nucleic acid) did not demonstrate viral mRNA staining.

None of the control brain tissues demonstrated viral RNA staining. The host control probe reacted appropriately in all tissues that were virus negative by ISH. Neither probe detected the other variants.

Discussion

According to the current criteria of using the L protein amino acid sequence similarity for jingchuviral speciation (12), these novel turtle jingchuvirals represent a new species within the genus Piscichuvirus. However, the original speciation criteria might need to be revisited to determine if the 2 variants (i.e., FTuNV1 and STuNV1) actually represent 2 different

piscichuviral species given their relative dissimilarity (92%) and host differences. For example, although sea turtles are known to forage within tidal areas of rivers (Byles, 1988) and freshwater turtles, including alligator snapping turtles, are occasionally documented in estuarine and marine waters , those ecosystems are relatively nonoverlapping. As additional studies reveal more about the diversity within and between chuvirids, their evolutionary timeline, and their host restrictions, it is foreseeable that these 2 turtle variants might ultimately be divided into at least 2 species (e.g., freshwater [chelydroid] and marine [chelonioid] turtle).

The predicted terminal panhandle structures of the turtle neural virus genomes are similar to those of many other viruses of phylum Negarnaviricota, including bunyavirals, orthomyxovirids, paramyxovirids, and rhabdovirids (Auperin et al., 1984; Hsu et al., 1987; Obijeski et al., 1980) but have not been reported for jingchuvirals. In orthomyxovirids, those structures serve as promoters for transcription (Flick et al., 1996; Fodor et al., 1994; Neumann & Hobom, 1995), but by creating double-stranded RNA, they also induce the antiviral activity of retinoic acid-inducible gene I. Although the biological effect of this structure remains to be determined, the putative panhandle-forming untranslated regions could be used for the in-silico identification of genomic ends in chuvirids discovered in the future through metagenomics and might provide more insight into the development of genomic structure diversity within Chuviridae.

Recent viral zoonoses (e.g., severe acute respiratory syndrome, Ebola virus disease, AIDS) demonstrate that wildlife species can be reservoirs (Keatts LO, Robards M, Olson SH, Hueffer K, Insley SJ, et al, 2021); thus, it is imperative to fully document the repertoire of viruses in wildlife and their association with disease. Chuvirids are the only jingchuvirals that have been identified in vertebrates, including fish and reptiles (Argenta et al., 2020; Shi et al.,

2018). However, any associations with the disease have been weak and lacked in-situ viral localization (Argenta et al., 2020). Our study successfully localized chuvirid mRNA within the areas of inflammation in multiple individuals across 3 turtle species from 2 different ecosystems. Because sequence-based approaches have become a common platform for disease detection and characterization, modifications to Koch's postulates have been proposed to establish the causal association of a novel agent in which Koch's postulates cannot be fulfilled (e.g., infection of novel agents in endangered species and a likely irreversible condition [meningoencephalomyelitis]) (Fredricks DN, Relman DA, 1996). The 2 turtle piscichuviral variants have met 3 of the 7 proposed criteria: 1) FTuNV1 and STuNV1 nucleic acid sequences were detected in diseased tissues, 2) no nucleic acid sequence was detected in tissues without disease, and 3) infection was confirmed at the cellular level via ISH. Although further research on this disease is required to verify reproducibility and to identify similar biological properties in other hosts, those findings strongly suggest that FTuNV1 and STuNV1 are a cause of severe mononuclear meningoencephalomyelitis in aquatic turtles in multiple ecosystems throughout the southeastern United States. The identification of closely related chuvirids in other reptiles and fish suggests that chuvirids should be considered as potential emerging viruses in at least fish and reptiles, if not mammals. Further surveillance is needed to better determine the effect of chuvirids on those and other turtles.

All of the turtle species in which a chuvirid was found are considered imperiled. Affected turtles included mature adults, which are especially vital to the stability and recovery of turtle populations (Thomas et al., 2014). Of note, all 3 turtles with known body condition scores (turtles 1, 4, and 5) were in good nutritional condition at the time of death, and all infected turtles lacked any predisposing conditions that would increase susceptibility to virus infection. In

addition, 2 STuNV1-infected loggerhead turtles (turtles 4 and 5) were stranded ≈1 month apart within the same geographic region. The potential to infect and cause disease in relatively healthy individuals and the identification of multiple diseased turtles from the same areas and time indicate a serious wildlife health concern. In addition, an observation associated with 1 of the cases reported here raises the possibility of human-mediated pathogen pollution. The genus *Macrochelys* is proposed to include either 2 or 3 species (Chaffin et al., 2008; Thomas et al., 2014). The alligator snapping turtle infected by FTuNV1 was morphologically consistent with the more western member(s) of the genus, either *M. temminckii* or *M. apalachicolae*, neither of which should be located where that turtle was found. Thus, the discovery of that turtle outside of its natural range suggests that it may have been transported and released. Future studies are needed to understand the diversity and prevalence of chuvirids among turtles, the pathogenesis of chuvirid infections, and the effects of environment on disease susceptibility.

In summary, we identified 2 variants of a new piscichuviral species in 5 aquatic turtles that died of idiopathic meningoencephalomyelitis. FTuNV1 and STuNV1 most likely cause lymphocytic meningoencephalomyelitis in multiple aquatic turtle species.

Acknowledgement

We thank Paula Baker and staff and volunteers of the Amos Rehabilitation Keep and the University of Florida College of Veterinary Medicine Zoological Medicine Service for their contributions to veterinary care and husbandry; the participants in the Sea Turtle Stranding and Salvage Network for response to and documentation of stranded sea turtles; Paul Moler for his

expertise in species identification of the alligator snapping turtle; and Nicole Stacy for the cytologic analysis and interpretation of cerebrospinal fluid.

Raw sequence data are available in the NCBI Sequence Read Archive (PRJNA936591). Complete viral genome sequences are available in GenBank (OQ547744–6)

Figures and Tables

Table 3.1: Biographical information and results summary of 5 turtles positive for FTuNV1 or STuNV1, United States, 2009–2021*

Animal no., common name (taxonomic name)	Sex	SCL, cm	Life stage†	Stranding location‡		Clinical signs	Sequencing	ISH (FTuNV1/ STuNV1)
				Latitude, °N	Longitude, °W			
1, Alligator snapping turtle (*Macrochelys* sp.)	M	56.5	M	29.524879	82.300594	Weak and lethargic	FTuNV1	(+/−)
2, Kemp's ridley turtle (*Lepidochelys kempii*)	F	60.3	I	27.67338	97.16880	Persistent circling and asymmetric buoyancy	STuNV1	(–/+)
3, Loggerhead turtle (*Caretta caretta*)	M	94.0	M	30.230591	87.910237	Unresponsive	STuNV1	(ND/+)
4, Loggerhead turtle (*C. caretta*)	F	81.8	M	28.04815	80.57892	Weak	ND	(ND/+)
5, Loggerhead turtle (*C. caretta*)	F	83.7	I	28.20726	80.65725	Head tremor and cervical ventroflexion	ND	(ND/+)

*FTuNV1, freshwater turtle neural virus 1; I, immature; ISH, in-situ hybiridization; M, mature; ND, not done; SCL, straight carapace length; STuNV1, sea turtle neural virus 1. †Life stage estimation was based on the maturation of the reproductive system. ‡Latitude and longitude values correspond to the stranding locations as shown in Figure 1.

Table 3.2: Genome comparison of piscichuviruses from 5 turtles positive for FTuNV1 or StuNV1, United States, 2009–2021, and reference sequences*

Virus	GenBank accession no.	Genome, nt	3′ UTR, nt	ORF4, nt	N, nt	G, nt	L, nt	5′ UTR, nt
FTuNV1	OQ547744	10,781	89	318	1,446	2,052	6,438	91
STuNV1 (Kemp's ridley)	OQ547745	10,839	93	318	1,500	2,052	6,438	91
STuNV1 (Loggerhead)	OQ547746	10,839	93	318	1,500	2,052	6,438	91
GRSCV	MG600009	10,625	>59	276‡	1,482	1,995	6,423	>67
WFClV	MG600011	10,385	>64	225‡	1,344	1,929	6,348	>45
HFrV1	MN567051	10,718	79	255‡	1,509	1,983	6,435	204
HhCV	MW645030–2	10,858	NA	NA	1,566	1,956	6,363	NA
SxASC4	KX884439	11,270	>177	234†	1,794	2,046	6,468	>97
FMCV1	ON125109	10,991	25	396	1,395	2,178	6,615	104

*FMCV1, freshwater macrophyte associate chu-like virus 1; FTuNV1, freshwater turtle neural virus 1; G, glycoprotein; GRSCV, Guǎngdōng red-banded snake chuvirus-like; HFrV1, Herr Frank virus 1; HhCV, hardyhead chuvirus; L, large protein; N, nucleoprotein; NA, data not available; ORF, open reading frame; STuNV1, sea turtle neural virus 1; SxASC4, Sānxiá atyid shrimp virus 4; UTR, untranslated region; WFClV, Wēnlǐng fish chu-like virus. †Annotations were performed in this study because ORF4 was not originally annotated in the National Center for Biotechnology Information.

Figure 3.1: Geographic location and year found for 5 stranded piscichuvirus-infected aquatic turtles with meningoencephalomyelitis, United States, 2009–2021.

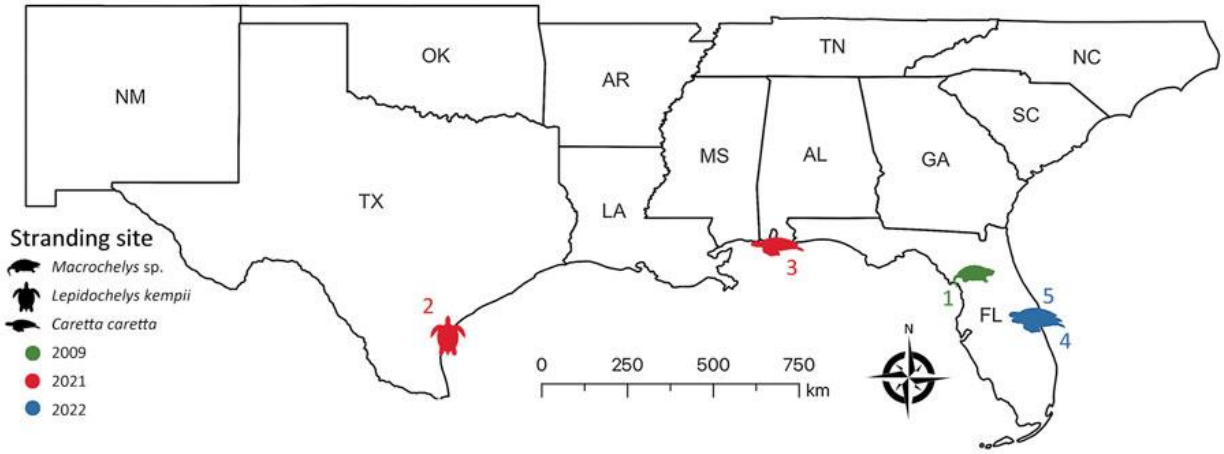
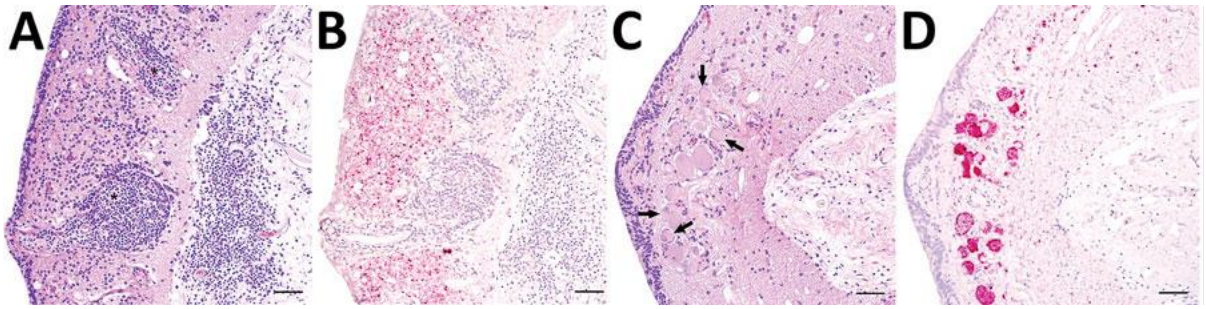
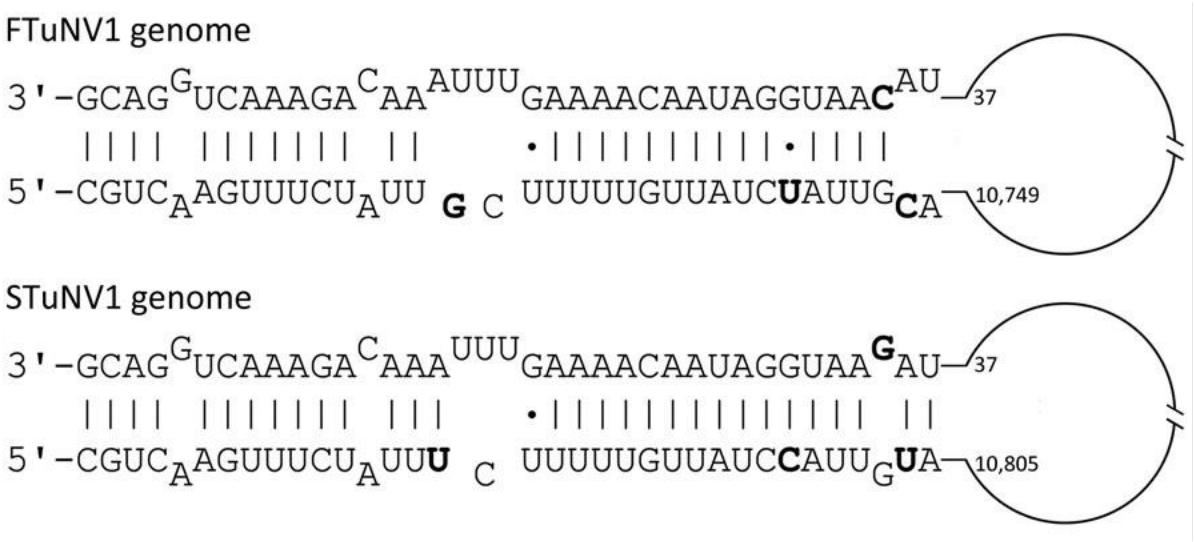


Figure 3.2: Representative tissue sections from the central nervous system of an alligator snapping turtle (*Macrochelys* sp.) with meningoencephalomyelitis, United States, 2009



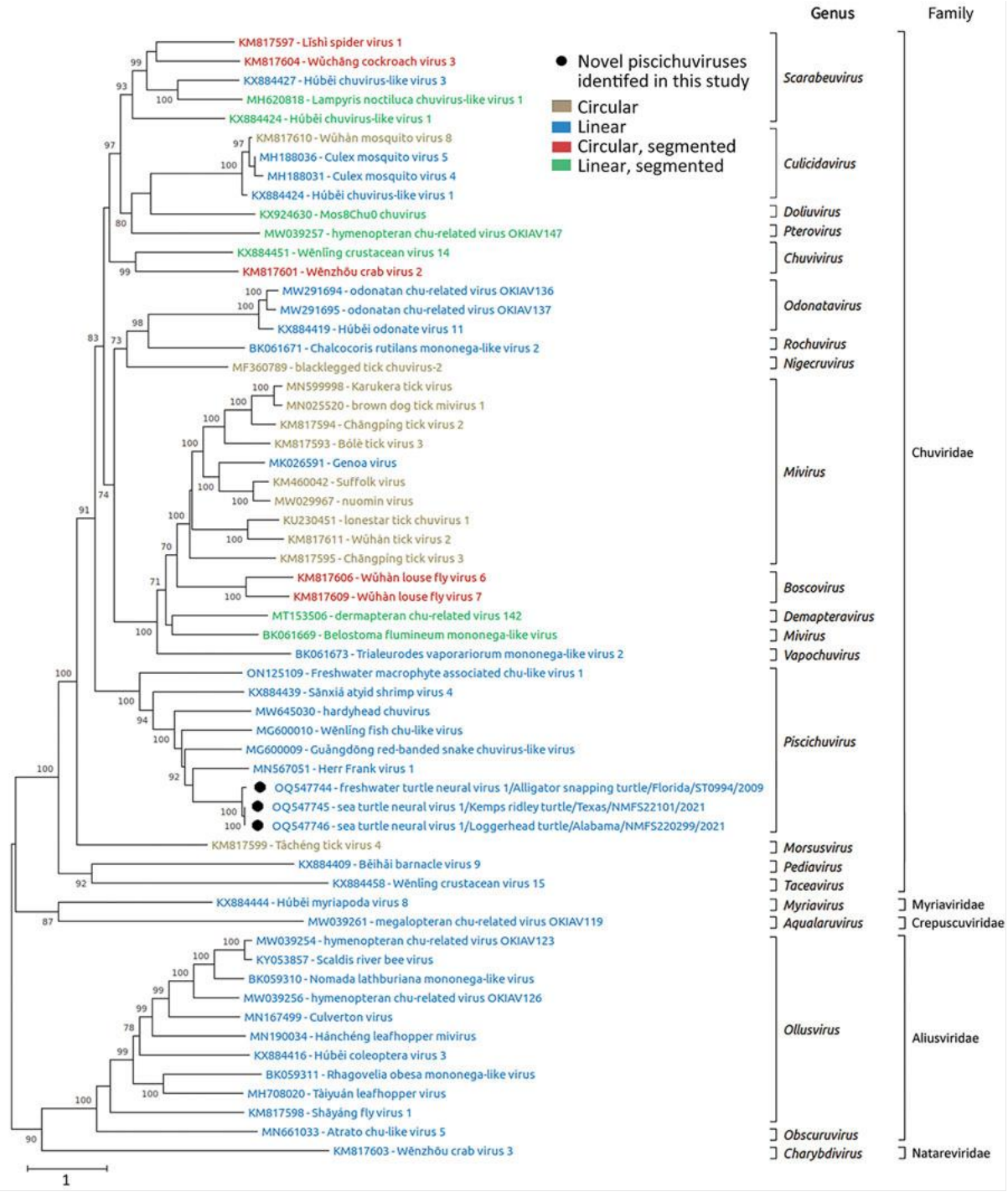
A) Cerebellum; lymphoplasmacytic perivascular cuffs (asterisk) and infiltrates are widely disseminated in the gray matter and the adjacent leptomeninges. Hematoxylin and eosin stain. B) Replicate section of the same tissue shown in panel A. There is strong in-situ hybridization signal (red) against freshwater turtle neural virus 1 (FTuNV1) in the cytoplasm of small neurons and glial cells throughout the gray matter and associated with the lymphoplasmacytic infiltrates. Hematoxylin counterstain. C) Optic tectum; several neurons have central chromatolysis (arrows). Hematoxylin and eosin stain. D) A replicate section of the tissue shown in panel C. Intense in-situ hybridization signal (red) against FTuNV1 was within the neuronal and glial cytoplasm. Scale bars indicate 50 µm.

Figure 3.3: Terminal structure showing panhandle sequences for the 3′ and 5′ termini of FTuNV1 and STuNV1 genomes from piscichuvirus-infected aquatic turtles with meningoencephalomyelitis, United States, 2009–2021.



Sequences were predicted by using the Vienna package RNAfold tool on Geneious (https://www.geneious.com). Offset bases indicate mispairing. Boldface bases indicate nucleotide differences between FTuNV1 and STuNV1. Dots indicate G/U pairings. The remaining genome is indicated by the loop structure (not to scale). FTuNV1, freshwater turtle neural virus 1; STuNV1, sea turtle neural virus 1.

Figure 3.4: Phylogenetic analysis of jingchuviral large protein (L) amino acid sequences from piscichuvirus-infected aquatic turtles with meningoencephalomyelitis, United States, 2009–2021 (black dots), and reference sequences**.**



Complete L amino acid sequences were aligned by using ClustalW (https://www.clustal.org) and refined by using MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle) with default settings. The phylogenetic analysis was performed on MEGA X (*23*) by using the maximum-likelihood method and Le Gascuel matrix plus observed amino acid frequencies plus 5 discrete gamma categories distribution with parameter of 1.0728 plus invariant sites with 0.65% sites. The substitution model was constructed with 500 bootstrap replicates. The tree is drawn to scale; bootstrap values are measured in the number of substitutions per site. This analysis included 59 aa sequences. Sequences are color coded based on their genomic structure.

CHAPTER 4

DISCUSSION

Aquatic testudines are threatened by a variety of factors including infectious diseases, which can be caused by multiple agents (e.g., bacteria, viruses, fungi, and protozoa) leading to serious morbidity and mortality in both captive and free-ranging populations. However, only a few viruses have been reported in freshwater and marine turtles with a strong disease association, e.g. cheloniid herpesvirus-5, cheloniid herpesvirus-6, and turtle fraser virus1. For other viruses, the disease association remains inconclusive, e.g. sea turtle tornovirus1 and betanodavirus. Before further investigation into the role of these viruses in the health of aquatic testudines can progress, baseline information on the viral community and diversity is required. This dissertation provides such knowledge and also describes advanced sequencing technique development to efficiently detect and characterize viruses in free-ranging aquatic testudines.

This study successfully developed random sequencing and semi-targeted sequencing for the MinION platform and used it for 1) detection and characterization of viral communities from the gastrointestinal tract of free-ranging sea turtles and for 2) identifying the cause of idiopathic meningonecephalomyelitis in free-ranging turtles. Overall, 25 viruses across 10 families were detected and phylogenetically characterized through random and semi-targeted sequencing in 14 individuals of 4 species of aquatic testudines (green sea turtles, loggerhead turtles, Kemp's ridley turtles, and an alligator snapping turtle). Out of these viruses, 15 were novel species across at least 2 novel genera.

The first major objective of this study was to develop random sequencing and semi-targeted, strand-switching MinION sequencing methods to detect and characterize undiscovered viruses in the gastrointestinal system of free-ranging sea turtles. Forty-eight viruses were successfully detected, and 19 viruses were phylogenetically characterized across eight viral families from three species of sea turtles (loggerhead turtles, green sea turtles, leatherback turtles (only detected), and Kemp's ridley turtles) from three geographical regions (West Coast, Northeast Coast, and Southeast Coast) in the United States. This demonstrates that MinION sequencing can be used to identify viral communities in specimens (i.e., feces, colonic content, digesta, and ingesta) from free-ranging sea turtles with high nucleic acid contamination from food, host, and a variety of infectious organisms. Additional studies can use these data to better determine disease correlation in controlled viral discovery assays of clinically affected turtles.

The second major objective of this study was to use MinION sequencing to determine the cause of idiopathic mononuclear meningoencephalomyelitis in five free-ranging aquatic turtles. This study successfully identified three piscichuviruses of two strains within a single novel piscichuviral species in the brain and spinal cord of freshwater and sea turtles with meningoencephalomyelitis. Viruses were putatively named freshwater turtle neural virus 1 and sea turtle neural virus 1. Moreover, double-z probes were custom designed specifically against the large protein of these novel piscichuviruses for localization of viral mRNA within tissues through RNAscope in-situ hybridization. Viral mRNA was colocalized within inflammatory areas of the brain and spinal cord tissues. This shows that MinION sequencing can be used for identifying disease causation, especially idiopathic cases with a suspicion of viral etiology. Future studies can use this data to investigate the impact of this virus across host species and geographical areas.

In conclusion, these MinION sequencing methods can be used to efficiently interrogate viruses in sea turtles. Constraints inherent to PCR diagnostic assays, such as inadequate phylogenetic information from short amplicons (100-1,000 bases), binary results (detected and not detected), and requirement for specific primers can be alleviated by random and semi-targeted MinION sequencing. This dissertation provides a methodology for other researchers willing to pursue metagenomic studies of viruses in wildlife and aquatic animals, particularly gastrointestinal viromes and identification of viral agents causing idiopathic diseases.

REFERENCES

1. Argenta, F. F., Hepojoki, J., Smura, T., Szirovicza, L., Hammerschmitt, M. E., Driemeier, D., Kipar, A., & Hetzel, U. (2020). Identification of reptarenaviruses, hartmaniviruses and a novel chuvirus in captive brazilian native boa constrictors with boid inclusion body disease. *Virology*, *January*, 1–19. https://doi.org/10.1101/2020.01.02.893420
2. Auperin, D. D., Romanowski, V., Galinski, M., & Bishopt, D. H. L. (1984). Sequencing Studies of Pichinde Arenavirus S RNA Indicate a Novel Coding Strategy, an Ambisense Viral S RNA. In *JOURNAL OF VIROLOGY*. https://journals.asm.org/journal/jvi
3. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021
4. Boyle, M. C., FitzSimmons, N. N., Limpus, C. J., Kelez, S., Velez-Zuazo, X., & Waycott, M. (2009). Evidence for transoceanic migrations by loggerhead sea turtles in the southern Pacific Ocean. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1664), 1993–1999. https://doi.org/10.1098/rspb.2008.1931
5. Brown, J. R., Bharucha, T., & Breuer, J. (2018). Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. *Journal of Infection*, *76*(3), 225–240. https://doi.org/10.1016/j.jinf.2017.12.014

6. Butt, S. L., Erwood, E. C., Zhang, J., Sellers, H. S., Young, K., Lahmers, K. K., & Stanton, J. B. (2021). Real-time, MinION-based, amplicon sequencing for lineage typing of infectious bronchitis virus from upper respiratory samples. *Journal of Veterinary Diagnostic Investigation*, *33*(2), 179–190. https://doi.org/10.1177/1040638720910107

7. Buysse, M., Koual, R., Binetruy, F., de Thoisy, B., Baudrimont, X., Garnier, S., Douine, M., Chevillon, C., Delsuc, F., Catzeflis, F., Bouchon, D., & Duron, O. (2024). Detection of Anaplasma and Ehrlichia bacteria in humans, wildlife, and ticks in the Amazon rainforest. *Nature Communications*, *15*(1), 3988. https://doi.org/10.1038/s41467-024-48459-y

8. Byles, R. A. (1988). Behavior and ecology of sea turtles from Chesapeake Bay, Behavior and ecology of sea turtles from Chesapeake Bay, Virginia Virginia. *The College of William and Mary ProQuest Dissertations Publishing*, *8907936*. https://doi.org/10.25773/v5-h9nv-c205

9. Carpente, L., Cerdeira-Pena, A., Lorenzo-Freire, S., & Places, Á. S. (2019). Optimization in Sanger sequencing. *Computers and Operations Research*, *109*, 250–262. https://doi.org/10.1016/j.cor.2019.05.011

10. Castro-Wallace, S. L., Chiu, C. Y., John, K. K., Stahl, S. E., Rubins, K. H., McIntyre, A. B. R., Dworkin, J. P., Lupisella, M. L., Smith, D. J., Botkin, D. J., Stephenson, T. A., Juul, S., Turner, D. J., Izquierdo, F., Federman, S., Stryke, D., Somasekar, S., Alexander, N., Yu, G., … Burton, A. S. (2017). Nanopore DNA Sequencing and Genome Assembly on the International Space Station. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-18364-0

11. Chaffin, K., Norton, T. M., Gilardi, K., Poppenga, R., Jensen, J. B., Moler, P., Cray, C., Dierenfeld, E. S., Chen, T., Oliva, M., Origgi, F. C., Gibbs, S., Mazzaro, L., & Mazet, J. (2008). Health assessment of free-ranging alligator snapping turtles (Macrochel ys temminckii) in Georgia and Florida. *Journal of Wildlife Diseases*, *44*(3), 670–686. https://doi.org/10.7589/0090-3558-44.3.670

12. Chai, C. X., Samat, A., Mohd-Taib, F. S., Bharudin, I., & Tim, J. (2023). Fungal infection of sea turtle eggs in the sea turtle hatcheries in Peninsular Malaysia. *Fungal Ecology*, *63*. https://doi.org/10.1016/j.funeco.2023.101243

13. Chapman, P. A., Owen, H., Flint, M., Soares Magalhães, R. J., Traub, R. J., Cribb, T. H., Kyaw-Tanner, M. T., & Mills, P. C. (2017). Molecular epidemiology and pathology of spirorchiid infection in green sea turtles (Chelonia mydas). *International Journal for Parasitology: Parasites and Wildlife*, *6*(1), 39–47. https://doi.org/10.1016/j.ijppaw.2017.03.001

14. Cocumelli, C., Fichi, G., Marsili, L., Senese, M., Cardeti, G., Cersini, A., Ricci, E., Garibaldi, F., Scholl, F., Di Guardo, G., & Terracciano, G. (2018). Cetacean Poxvirus in two striped dolphins (Stenella coeruleoalba) stranded on the Tyrrhenian coast of Italy: Histopathological, ultrastructural, biomolecular, and ecotoxicological findings. *Frontiers in Veterinary Science*, *5*(SEP). https://doi.org/10.3389/fvets.2018.00219

15. Conceição-Neto, N., Yinda, K. C., Van Ranst, M., & Matthijnssens, J. (2018). NetoVIR: Modular approach to customize sample preparation procedures for viral metagenomics. In *Methods in Molecular Biology* (Vol. 1838, pp. 85–95). Humana Press Inc. https://doi.org/10.1007/978-1-4939-8682-8_7

16. Crawford, B. A., Maerz, J. C., Nibbelink, N. P., Buhlmann, K. A., & Norton, T. M. (2014). Estimating the consequences of multiple threats and management strategies for semi-aquatic turtles. *Journal of Applied Ecology*, *51*(2), 359–366. https://doi.org/10.1111/1365-2664.12194

17. Datta, S. (2015). Next-generation sequencing in clinical virology: Discovery of new viruses. *World Journal of Virology*, *4*(3), 265. https://doi.org/10.5501/wjv.v4.i3.265

18. Dezordi, F. Z., Vasconcelos, C. R. dos S., Rezende, A. M., & Wallau, G. L. (2020). In and Outs of Chuviridae Endogenous Viral Elements: Origin of a Potentially New Retrovirus and Signature of Ancient and Ongoing Arms Race in Mosquito Genomes. *Frontiers in Genetics*, *11*(October). https://doi.org/10.3389/fgene.2020.542437

19. Di Guardo, G., Di Francesco, C. E., Eleni, C., Cocumelli, C., Scholl, F., Casalone, C., Peletto, S., Mignone, W., Tittarelli, C., Di Nocera, F., Leonardi, L., Fernández, A., Marcer, F., & Mazzariol, S. (2013). Morbillivirus infection in cetaceans stranded along the Italian coastline: Pathological, immunohistochemical and biomolecular findings. *Research in Veterinary Science*, *94*(1), 132–137. https://doi.org/10.1016/j.rvsc.2012.07.030

20. Di Paola, N., Dheilly, N. M., Junglen, S., Paraskevopoulou, S., Postler, T. S., Shi, M., & Kuhn, J. H. (2022). Jingchuvirales: a New Taxonomical Framework for a Rapidly Expanding Order of Unusual Monjiviricete Viruses Broadly Distributed among Arthropod Subphyla. *Applied and Environmental Microbiology*, *88*(6). https://doi.org/10.1128/aem.01954-21

21. Flick, R., Neumann, G., Hoffmann, E., Neumeier, E., & Hobom, G. (1996). Promoter Elements in the Influenza vRNA Terminal Structure. *RNA*, *2*(10), 1046–1057.

22. Fodor, E., Pritlove, D. C., & Brownlee, G. G. (1994). The Influenza Virus Panhandle Is Involved in the Initiation of Transcription. In *JOURNAL OF VIROLOGY* (Vol. 68, Issue 6).

23. Galindo-Cardiel, I., Ballester, M., Solanes, D., Nofrarías, M., López-Soria, S., Argilaguet, J. M., Lacasta, A., Accensi, F., Rodríguez, F., & Segalés, J. (2013). Standardization of pathological investigations in the framework of experimental ASFV infections. In *Virus Research* (Vol. 173, Issue 1, pp. 180–190). https://doi.org/10.1016/j.virusres.2012.12.018

24. Greenblatt, R. J., Work, T. M., Dutton, P., Sutton, C. A., Spraker, T. R., Casey, R. N., Diez, C. E., Parker, D., St. Leger, J., Balazs, G. H., & Casey, J. W. (2005). Geographic variation in marine turtle fibropapillomatosis. *Journal of Zoo and Wildlife Medicine*, *36*(3), 527–530. https://doi.org/10.1638/04-051.1

25. Hahn, M. A., Rosario, K., Lucas, P., & Dheilly, N. M. (2020). Characterization of viruses in a tapeworm: phylogenetic position, vertical transmission, and transmission to the parasitized host. *ISME Journal*, *14*(7), 1755–1767. https://doi.org/10.1038/s41396-020-0642-2

26. Han, X., Wang, H., Wu, N., Liu, W., Cao, M., & Wang, X. (2020). Leafhopper Psammotettix alienus hosts chuviruses with different genomic structures. *Virus Research*, *285*(April), 197992. https://doi.org/10.1016/j.virusres.2020.197992

27. Harding, E. F., Russo, A. G., Yan, G. J. H., Mercer, L. K., & White, P. A. (2022). Revealing the uncharacterised diversity of amphibian and reptile viruses. *ISME Communications*, *2*(1). https://doi.org/10.1038/s43705-022-00180-x

28. Hewson, I., Button, J. B., Gudenkauf, B. M., Miner, B., Newton, A. L., Gaydos, J. K., Wynne, J., Groves, C. L., Hendler, G., Murray, M., Fradkin, S., Breitbart, M., Fahsbender,

E., Lafferty, K. D., Kilpatrick, A. M., Miner, C. M., Raimondi, P., Lahner, L., Friedman, C. S., … Harvell, C. D. (2014). Densovirus associated with sea-star wasting disease and mass mortality. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(48), 17278–17283. https://doi.org/10.1073/pnas.1416625111

29. Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Research*, *31*(13), 3429–3431. https://doi.org/10.1093/nar/gkg599

30. Hsu, M.-T., Parvin, J. D., Gupta, S., Krystal, M., & Palese, P. (1987). Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. In *Proc. Nati. Acad. Sci. USA* (Vol. 84). https://www.pnas.org

31. Joyce, W. G. (2015). The origin of turtles: A paleontological perspective. In *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* (Vol. 324, Issue 3, pp. 181–193). John Wiley and Sons Inc. https://doi.org/10.1002/jez.b.22609

32. Kelsey T. Young, Jazz Q. Stephens, a R. L. P., & David E. Stallknecht, Kiril M. Dimitrov, Salman L. Butt, J. B. S. (2022). Putative Novel Avian Paramyxovirus (AMPV) and Reidentification of APMV-2 and APMV-6 to the Species Level Based on Wild Bird Surveillance (United States, 2016-2018). *Applied and Environmental Microbiology*, 2016–2018.

33. Kim, D., Song, L., Breitwieser, F. P., & Salzberg, S. L. (2016). Centrifuge: rapid and accurate classificaton of metagenomic sequences, version 1.0.4_beta. *BioRxiv*, *26*(12), 054965. https://doi.org/10.1101/gr.210641.116.Freely

34. Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T. P. L., & Pevzner, P. A. (2020). metaFlye:

scalable long-read metagenome assembly using repeat graphs. *Nature Methods*, *17*(11), 1103–1110. https://doi.org/10.1038/s41592-020-00971-x

35. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

36. Lechner, E. S., Cynda Crawford, P., Levy, J. K., Edinboro, C. H., Dubovi, E. J., & Caligiuri, R. (2010). Prevalence of protective antibody titers for canine distemper virus and canine parvovirus in dogs entering a Florida animal shelter. *Journal of the American Veterinary Medical Association*, *236*(12), 1317–1321. https://doi.org/10.2460/javma.236.12.1317

37. Leigh Perkins, L. E., Campagnoli, R. P., Harmon, B. G., Gregory, C. R., Steffens, W. L., Latimer, K., Clubb, S., & Crane, M. (2001). Detection and confirmation of reptilian adenovirus infection by in situ hybridization. *Journal of Veterinary Diagnostic Investigation*, *13*(4), 365–368. https://doi.org/10.1177/104063870101300418

38. Li, C. X., Shi, M., Tian, J. H., Lin, X. D., Kang, Y. J., Chen, L. J., Qin, X. C., Xu, J., Holmes, E. C., & Zhang, Y. Z. (2015). Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *ELife*, *2015*(4), 1–26. https://doi.org/10.7554/eLife.05378

39. Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. In *Genomics, Proteomics and Bioinformatics* (Vol. 14, Issue 5, pp. 265–279). Beijing Genomics Institute. https://doi.org/10.1016/j.gpb.2016.05.004

40. Lyson, T. R., Rubidge, B. S., Scheyer, T. M., de Queiroz, K., Schachner, E. R., Smith, R. M. H., Botha-Brink, J., & Bever, G. S. (2016). Fossorial Origin of the Turtle Shell. *Current Biology*, *26*(14), 1887–1894. https://doi.org/10.1016/j.cub.2016.05.020

41. Marschang, R. E. (2011). Viruses infecting reptiles. *Viruses*, *3*(11), 2087–2126. https://doi.org/10.3390/v3112087

42. Marschang, R. E., Becher, P., Posthaus, H., Wild, P., Thiel, H. J., Müller-Doblies, U., Kaleta, E. F., & Bacciarini, L. N. (1999). Isolation and characterization of an iridovirus from Hermann's tortoises (Testudo hermanni). *Archives of Virology*, *144*(10), 1909–1922. https://doi.org/10.1007/s007050050714

43. Mikalsen, A. B., Nilsen, P., Frøystad-Saugen, M., Lindmo, K., Eliassen, T. M., Rode, M. R., & Evensen, Ø. (2014). Characterization of a novel calicivirus causing systemic infection in Atlantic salmon (Salmo salarL.): Proposal for a new genus of caliciviridae. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0107132

44. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419. https://doi.org/10.1093/nar/gkaa913

45. Neumann, G., & Hobom, G. (1995). Mutational analysis of influenza virus promoter elements in vivo. In *Journal of General Virology* (Vol. 76).

46. Ng, T. F. F., Manire, C., Borrowman, K., Langer, T., Ehrhart, L., & Breitbart, M. (2009). Discovery of a Novel Single-Stranded DNA Virus from a Sea Turtle Fibropapilloma by Using Viral Metagenomics. *Journal of Virology*, *83*(6), 2500–2509. https://doi.org/10.1128/jvi.01946-08

47. Obijeski, J. F., Mccauley, J., & Skehel, J. J. (1980). *Nucleotide sequences at the termini of La Crosse virus RNAs* (Vol. 8). https://doi.org/10.1093/nar/8.11.2431

48. Orós, J., Calabuig, P., & Déniz, S. (2004). Digestive pathology of sea turtles stranded in the Canary Islands between 1993 and 2001. *Veterinary Record*, *155*(6), 169–174. https://doi.org/10.1136/vr.155.6.169

49. Paries, S., Funcke, S., Kershaw, O., Failing, K., & Lierz, M. (2019). The role of Virus "X" (Tortoise Picornavirus) in kidney disease and shell weakness syndrome in European tortoise species determined by experimental infection. *PLoS ONE*, *14*(2). https://doi.org/10.1371/journal.pone.0210790

50. Peery, A. F., Dellon, E. S., Lund, J., Crockett, S. D., McGowan, C. E., Bulsiewicz, W. J., Gangarosa, L. M., Thiny, M. T., Stizenberg, K., Morgan, D. R., Ringel, Y., Kim, H. P., Dibonaventura, M. D., Carroll, C. F., Allen, J. K., Cook, S. F., Sandler, R. S., Kappelman, M. D., & Shaheen, N. J. (2012). Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology*, *143*(5), 1179-1187.e3. https://doi.org/10.1053/j.gastro.2012.08.002

51. Perlejewski, K., Bukowska-Ośko, I., Rydzanicz, M., Pawełczyk, A., Caraballo Cortès, K., Osuch, S., Paciorek, M., Dzieciątkowski, T., Radkowski, M., & Laskus, T. (2020). Next-generation sequencing in the diagnosis of viral encephalitis: sensitivity and clinical limitations. *Scientific Reports*, *10*(1), 1–7. https://doi.org/10.1038/s41598-020-73156-3

52. Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T. F., Beutler, N. A., Burton, D. R., Lewis-Ximenez, L. L., De Jesus, J. G., Giovanetti, M., Hill, S. C., Black, A., Bedford, T., Carroll, M. W., Nunes, M., … Loman, N. J. (2017). Multiplex PCR method for MinION

and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols*, *12*(6), 1261–1266. https://doi.org/10.1038/nprot.2017.066

53. Rhodin, A. G. J., Iverson, J. B., Bour, R., Fritz, U., Georges, A., Shaffer, H. B., & Van Dijk, P. P. (2021). Turtles of the World: Annotated Checklist and Atlas of Taxonomy, Synonymy, Distribution, and Conservation Status (9th Ed.). *Chalonian Research Foundation and Turtle Conservancy*, *8*(9).

54. Sarker, S., Hannon, C., Athukorala, A., & Bielefeldt-Ohmann, H. (2021). *Emergence of a Novel Pathogenic Poxvirus Infection in the*. https:// doi.org/10.3390/v13020219

55. Senko, J. F., Burgher, K. M., del Mar Mancha-Cisneros, M., Godley, B. J., Kinan-Kelly, I., Fox, T., Humber, F., Koch, V., Smith, A. T., & Wallace, B. P. (2022). Global patterns of illegal marine turtle exploitation. *Global Change Biology*, *28*(22), 6509–6523. https://doi.org/10.1111/gcb.16378

56. Shi, M., Lin, X. D., Chen, X., Tian, J. H., Chen, L. J., Li, K., Wang, W., Eden, J. S., Shen, J. J., Liu, L., Holmes, E. C., & Zhang, Y. Z. (2018). The evolutionary history of vertebrate RNA viruses. *Nature*, *556*(7700), 197–202. https://doi.org/10.1038/s41586-018-0012-7

57. Shi, M., Lin, X. D., Tian, J. H., Chen, L. J., Chen, X., Li, C. X., Qin, X. C., Li, J., Cao, J. P., Eden, J. S., Buchmann, J., Wang, W., Xu, J., Holmes, E. C., & Zhang, Y. Z. (2016). Redefining the invertebrate RNA virosphere. *Nature*, *540*(7634), 539–543. https://doi.org/10.1038/nature20167

58. Stacy, B. A., Wellehan, J. F. X., Foley, A. M., Coberley, S. S., Herbst, L. H., Manire, C. A., Garner, M. M., Brookins, M. D., Childress, A. L., & Jacobson, E. R. (2008). Two herpesviruses associated with disease in wild Atlantic loggerhead sea turtles (Caretta

caretta). *Veterinary Microbiology*, *126*(1–3), 63–73. https://doi.org/10.1016/j.vetmic.2007.07.002

59. Thomas, T. M., Granatosky, M. C., Bourque, J. R., Krysko, K. L., Moler, P. E., Gamble, T., Suarez, E., Leone, E., Enge, K. M., & Roman, J. (2014). Taxonomic assessment of Alligator Snapping Turtles (Chelydridae: Macrochelys), with the description of two new species from the southeastern United States. *Zootaxa*, *3786*(2), 141–165. https://doi.org/10.11646/zootaxa.3786.2.4

60. Tyson, J. R., James, P., Stoddart, D., Sparks, N., Wickenhagen, A., Hall, G., Choi, J. H., Lapointe, H., Kamelian, K., Smith, A. D., Prystajecky, N., Goodfellow, I., Wilson, S. J., Harrigan, R., Snutch, T. P., Loman, N. J., & Quick, J. (2020). Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *BioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2020.09.04.283077

61. Vibin, J., Chamings, A., Collier, F., Klaassen, M., Nelson, T. M., & Alexandersen, S. (2018). Metagenomics detection and characterisation of viruses in faecal samples from Australian wild birds. *Scientific Reports*, *8*(1), 1–23. https://doi.org/10.1038/s41598-018-26851-1

62. Waltzek, T. B., Stacy, B. A., Ossiboff, R. J., Stacy, N. I., Fraser, W. A., Yan, A., Mohan, S., Koonin, E. V., I. Wolf, Y., Rodrigues, T. C. S., Viadanna, P. H. O., Subramaniam, K., Popov, V. L., Guzman-Vargas, V., & Shender, L. A. (2022). A novel group of negative-sense RNA viruses associated with epizootics in managed and free-ranging freshwater turtles in Florida, USA. *PLoS Pathogens*, *18*(3), 1–28. https://doi.org/10.1371/journal.ppat.1010258

63. Watzinger, F., Ebner, K., & Lion, T. (2006). Detection and monitoring of virus infections by real-time PCR. *Molecular Aspects of Medicine*, *27*(2–3), 254–298. https://doi.org/10.1016/j.mam.2005.12.001

64. Whitmore, L., Yetsko, K., Farrell, J. A., Page-karjian, A., Daniel, W., Donna, J., Frandsen, H. R., Walker, J. S., Crowder, W., Bovery, C., Rollinson, D., Burkhalter, B., Ryan, E., & Duffy, D. J. (2021). *Evolutionary genomic comparisons of chelonid herpesvirus 5 (ChHV5) from fibropapillomatosis-afflicted green (Chelonia mydas), olive ridley (Lepidochelys olivacea), and Kemp's Ridley (Lepidochelys kempii) sea turtles. Animals, 11(9)*. https://doi.org/10.3390/ani11092489

65. Winzeler, M. E., Hamilton, M. T., Tuberville, T. D., & Lance, S. L. (2015). First case of ranavirus and associated morbidity and mortality in an eastern mud turtle Kinosternon subrubrum in South Carolina. *Diseases of Aquatic Organisms*, *114*(1), 77–81. https://doi.org/10.3354/dao02849

66. Woodburn, D. B., Kinsel, M. J., Poll, C. P., Langan, J. N., Haman, K., Gamble, K. C., Maddox, C., Jeon, A. B., Wellehan, J. F. X., Ossiboff, R. J., Allender, M. C., & Terio, K. A. (2021). Shell Lesions Associated With Emydomyces testavorans Infection in Freshwater Aquatic Turtles. *Veterinary Pathology*, *58*(3), 578–586. https://doi.org/10.1177/0300985820985217

67. Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, *20*(1). https://doi.org/10.1186/s13059-019-1891-0

68. Wyneken, J., Lohmann, K. J., & Musick, J. A. (2013). *The biology of sea turtles. Volume III*.

69. Xing, X. W., Zhang, J. T., Ma, Y. B., He, M. W., Yao, G. E., Wang, W., Qi, X. K., Chen, X. Y., Wu, L., Wang, X. L., Huang, Y. H., Du, J., Wang, H. F., Wang, R. F., Yang, F., & Yu, S. Y. (2020). Metagenomic Next-Generation Sequencing for Diagnosis of Infectious Encephalitis and Meningitis: A Large, Prospective Case Series of 213 Patients. *Frontiers in Cellular and Infection Microbiology*, *10*(March). https://doi.org/10.3389/fcimb.2020.00088

70. Young, K. T., Lahmers, K. K., Sellers, H. S., Stallknecht, D. E., Poulson, R. L., Saliki, J. T., Tompkins, S. M., Padykula, I., Siepker, C., Howerth, E. W., Todd, M., & Stanton, J. B. (2021). Randomly primed, strand-switching, MinION-based sequencing for the detection and characterization of cultured RNA viruses. *Journal of Veterinary Diagnostic Investigation*, *33*(2), 202–215. https://doi.org/10.1177/1040638720981019

71. Zhang, Y. Z., Chen, Y. M., Wang, W., Qin, X. C., & Holmes, E. C. (2019). Expanding the RNA Virosphere by Unbiased Metagenomics. *Annual Review of Virology*, *6*, 119–139. https://doi.org/10.1146/annurev-virology-092818-015851