CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Tropical AgriSciences

Genetic Background of Adaptation of Crimean-Congo Haemorrhagic Fever Virus to the Different Tick Hosts

MASTER'S THESIS

Prague 2021

Author: Seyma Celina

Chief supervisor: RNDr. Jiří Černý, Ph.D.

Declaration

I hereby declare that I have done this thesis entitled “Genetic Background of Adaptation of Crimean-Congo Haemorrhagic Fever Virus to the Different Tick Hosts” independently, all texts in this thesis are original, and all the sources have been quoted and acknowledged by means of complete references and according to Citation rules of the FTA.

In Prague, 23rd of April, 2021

.....

Seyma Celina

Acknowledgements

First of all, I would like to thank my supervisor Dr. Jiří Černý, Ph.D. for giving me the opportunity of working on this topic. His support, encouragement, patience, and criticism have been invaluable in the production of this thesis and I am very fortunate to have learned many things from him.

I would also like to show my gratitude to my colleagues MVDr. Mentor Alishani and MVDr. Avni Robaj (University of Prishtina, Kosovo) for making all arrangements and helping me to collect ticks in Kosovo during COVID-19 pandemic.

Special thanks belong to Dr. Ing. Karolína Brandlová, Ph.D. (WMC Research Group, FTZ) for helping me to get in contact with ornithologists from CES project and got permission to join their bird ringing activities and collect ticks.

I would also like to acknowledge the assistance of Dr. Saima Naz (Dep.of Veterinary Sciences, FAFNR, CZU) for identifying all tick species.

My appreciation also belongs to epidemiologist Asst. Prof. Abdallah M. Samy (University of Kansas, USA; Ain Shams University, Egypt) for his cooperation, professional support, guidance, and help in the modeling study of this thesis as well as in answering various theory-related questions and commenting on my results. I am very grateful for his assistance, support and kindness.

Thanks to my friends Beatriz, Pauline, Valentina and Corrie for their supports. We shared many things together during our master study. They have all been doing researches on addax, tigers and reptiles, but they have always been interested in ticks.

Lastly, I would like to thank my husband Berk for his encouragement, moral support, patience and love, and also to my parents and siblings for their endless support.

This thesis was made possible through the generous Czech Government Scholarship that supported me during the two years of my studies at Czech University of Life Sciences.

Last but not least, I would like to express my gratitude to all those who read this thesis, to which I devoted most of my time but also learned and enjoyed. I hope you will learn and enjoy as much as I did.

Abstract

Crimean-Congo haemorrhagic fever virus (CCHFV) is a negative-sense, single-stranded RNA virus and causative agent of Crimean-Congo Haemorrhagic Fever (CCHF). It has become one of the most widely distributed tick-borne viral infection in the world. CCHF is endemic in Africa, Asia, the Middle East, and the Balkan region of Europe. However, the disease is spreading northwards following expanded distribution of the main vector, *Hyalomma marginatum*. Tick species of *Hyalomma* genus are mainly responsible for the transmission of CCHFV but other ticks such as representatives of genera *Rhipicephalus* and *Dermacentor* are also involved in virus transmission.

The framework of this thesis consists of three main parts. Firstly, it was focused to understand the evolutionary adaptation of CCHFV to *Hyalomma* and *Rhipicephalus* tick species. Understanding coevolution between CCHFV and its main vectors and whether there is a difference in evolutionary adaptation of CCHFV to the different ticks were the main focus of the first part of this thesis.

Study of tick collection from livestock in Kosovo and from birds in Czechia and investigate possible presence of CCHFV was the second focus of the thesis. The Balkan region is considered to be highly endemic region for CCHF. On the other hand, risk areas for tick presence occur in all regions of Czechia. Moreover, migratory birds play an important role in distribution of CCHFV vectors into Czechia. Tick collection, thus, were carried out in these countries to collect *Hyalomma* species and screen for CCHFV presence.

Furthermore, apart from migratory birds also other factors such as climate change and international animal trade might influence the expansion of *H. marginatum*, principal vector of CCHFV, in Europe. Up to the present, there are sporadic occurrences of this tick species in all countries of Central Europe except Czechia. In this regard, final part of this thesis was focused on ecological niche modeling analysis to map the environmental suitability of *H. marginatum* in Europe, with a focus on Central Europe.

In order to better understand the evolutionary characteristics of the virus and general codon usage pattern in CCHFV strains isolated from different tick hosts,

effective number of codons (ENC) and the GC-content at the 3rd codon position (GC3) of synonymous codons were calculated. Analysis of ENC and GC3 shows that *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains use different codons. The same result is apparent from correspondence analysis (COA) and supports the results of ENC and GC3 analysis. Furthermore, CAI analysis was performed which is a widely used technique for analyzing codon usage bias in order to measure the adaptation of the virus and predict the virus' behaviour on different tick hosts. CAI analysis shows that *Rhipicephalus*-isolated strains are more adapted for using codons which are being preferred by *Hyalomma* species and the same for *Hyalomma*-isolated strains. Results of the study indicate the CCHFV adaptation to the different tick hosts supports the effect of evolutionary processes on codon usage patterns of the virus.

In relation to tick collection, in total 171 ticks were collected from livestock in Kosovo in August 2020 and from birds in Czechia in September-October 2020. *H. marginatum* dominated in Kosovo, while *Ixodes ricinus* was the majority of collection in Czechia. No viral RNA was detected in collected ticks, suggesting seasonality of the CCHF might be related with tick abundance and activity.

Ecological niche modeling analysis and the predictions of *H. marginatum* in our model depict medium suitability for the occurrence of *H. marginatum* in many regions of southern and northern parts of Czechia.

The results of this thesis gave us new information about evolutionary adaptation of CCHFV to different ticks. Moreover, the results of this work were highly important for observing tick species in Kosovo and Czechia, and screening for possible CCHFV presence. Moreover, predicting the main vector of CCHFV's habitat suitability and making inferences regarding the virus transmission, and anticipate disease risk for the future gave us valuable information.

Key words: Crimean-Congo haemorrhagic fever virus, molecular evolution, codon usage bias, ecological niche modeling, *Hyalomma*, *Rhipicephalus*

Contents

List of tables

List of figures

List of the abbreviations used in the thesis

a.s.l: Above sea level

AIC: Akaike Information Criteria

AUC: Area Under the Curve

BSL: Biosafety Level

CAI: Codon Adaptation Index

CCHF: Crimean-Congo haemorrhagic fever

CCHFV: Crimean-Congo haemorrhagic fever virus

CHF: Crimean haemorrhagic fever

CHFV: Crimean haemorrhagic fever virus

COA: Correspondence Analysis

CSV: Comma-separated value

c_q: Quantification Cycle

DNA: Deoxyribonucleic acid

e-CAI: Expected Codon Adaptation Index

ELISA: Enzyme-linked immunosorbent assay

ENC: Effective Number of Codons

ENM: Ecological Niche Modeling

GC: Guanine-Cytosine

GC3: GC-content at the 3rd codon position

HCl: Hydrochloric acid

kDa: Kilodalton

L: Large segment

LAI: Leaf Area Index

LST: Land Surface Temperature

M: Medium segment

MoE: Ministry of the Environment of the Czech Republic

mRNA: Messenger ribonucleic acid

N: Nucleocapsid

NCBI: National Center for Biotechnology Information

NDVI: Normalized Difference Vegetation Index

OS: Omission rates

PCA: Principal Component Analysis

PCs: Principal components

PPT: Precipitation

pROC: Partial Receiver Operating Characteristic

qPCR: Real-time polymerase chain reaction

RdRp: RNA-dependent polymerases

RNA: Ribonucleic acid

RSCU: Relative Synonymous Codon Usage

S: Small segment

SDM: Species Distribution Modeling

tRNA: Transfer RNA

WHO: World Health Organization

1. Introduction

1.1. Ticks

1.1.1. Global importance of ticks

Ticks have long been known as ectoparasites of vertebrates for their associated medical importance, with tick fever described in an Egyptian papyrus scroll as early as 1550BC (Varma 1993). They occupy a wide variety of habitats, have close associations with a wide range of vertebrate hosts and can transmit various disease agents to humans, livestock, and wildlife animals (Magnarelli 2009). These highly specialized ectoparasites have variety of hosts including mammals, birds, reptiles and amphibians (Leech 2015). Ticks transmit various pathogens in comparison to any other arthropod organisms (Jongejan & Uilenberg 2004) and can induce tick paralysis and toxicosis themselves (Hall-Mendelin et al. 2011). Ticks are the most important vectors of livestock pathogens worldwide and human pathogens in Europe.

Ticks transmit deadly pathogens within livestock such as *Theileria*, a parasite which infects cattle costing the livestock industry hundreds of millions of dollars (Bishop et al. 2004), *Babesia canis*, a protozoan haemoparasite infecting domestic dogs (Irwin 2009), the Lyme disease spirochete which causes a debilitating human disease, tick-borne encephalitis virus (TBEV), a flavivirus causing serious disease in humans and occasionally its large vertebrate hosts, and various zoonotic haemorrhagic fever viruses such as Crimean Congo haemorrhagic fever virus, Omsk haemorrhagic fever virus, Kyasanur forest disease virus are important tick-borne pathogens for humans (Whitehouse 2004; Zivcec et al. 2013).

1.1.2. Ticks as important vector organism and Arboviruses

Ticks have several characteristics that make them ideal for maintenance and transmission of pathogens, one of which is their longevity and life cycle. Ticks undergo four life stages; eggs, larvae, nymphs and adults with the entire life cycle measured in years (Labuda & Nuttall 2004). Pathogens, such as tick-borne viruses, are shown to be

maintained throughout the tick life cycle making them important reservoirs of disease (Davies et al. 1986).

Different tick species produce different numbers of generations in one year, mainly affected by climatic factors (Gray et al. 2009). For instance, *Rhipicephalus sanguineus* ticks may produce up to two generations per year being active throughout winter in warm climates of Europe (Dantas-Torres 2010). *Ixodes ricinus* have two peaks of activity, one in spring and autumn in cooler northern climates. This species is known for feeding once in a year, completing one life cycle in three years (Hillyard 1996) and can overwinter in the absence of its suitable host in order to prolong the life cycle further (Kahl et al. 2002). *Dermacentor albopictus*, the winter tick, from North America is commonly associated with cervid species, particularly with moose (*Alces alces*) is a single-host tick with a life cycle of about one year. *Dermacentor reticulatus* ticks, on the other hand, have an intermediate strategy with the adults overwintering but still having one generation per year. These strategies on their life history make efficient reservoirs of diseases and vector, carrying viruses for long periods.

In order to complete their life cycle, ticks must get their blood meal from a suitable host in each life stage. Ticks can be categorized based on the number of host species they parasitise during their life cycle. Tick species in genera such as *Ixodes*, *Rhipicephalus*, and *Dermacentor* are predominantly three-host species that drop off their host each molt, while *Hyalomma* ticks (e.g., *H. marginatum* complex of species, *Hyalomma detritum*, *Hyalomma anatolicum*) being two-host species whereby the larvae and nymphs feed on the same host such as rodents, hares, ground-feeding birds or other small animals while adults feed on sheep, cattle and other large mammals (Sonenshine & Roe 2014). One-host ticks such as *D. albopictus*, *Rhipicephalus annulatus* and *Rhipicephalus microplus* remain on one host, typically a large mammal, during all life stages.

The feeding behaviour of ticks thus affects the vertebrate hosts that are in contact with pathogens. The feeding behaviour of ticks is also important in the transmission of pathogens, as it enables pathogen transmission through saliva secreted during feeding (Kaufman & Nuttall 2003; Nuttall & Labuda 2004).

Ticks are pool feeders, engorging all of the fluids that are exuded into the haemorrhagic pool of the host's skin generated by the bite (Leech 2015). Various

bioactive compounds are secreted by ticks in order to get successfully their blood meal (Kazimírová & Štibrániová 2013). These compounds include potent anticoagulants and anti-inflammatory proteins that inhibit platelet aggregation and facilitate blood flow, inhibit the coagulation pathway and suppress wound healing and inflammation (Kazimírová & Štibrániová 2013).

Arboviruses (arthropod-borne viruses) are a diverse group of viruses that are transmitted to vertebrates through an arthropod host with more than 530 identified species (Labuda & Nuttall 2004). The major arthropod vectors of arboviruses are mosquitoes, ticks, sandflies, and biting midges (Artsob & Lindsay 2008). Ticks are involved in transmission of over 70 arbovirus species from two orders, nine families, and at least 12 genera (Shi et al. 2018, Table 1). Tick-borne viruses are unique due to their ability to replicate in both vertebrate and non-vertebrate cells with less effect to the vector.

Arboviruses are RNA viruses, with the exception of African swine fever which is a DNA virus. Approximately, 10% of the 900 of recognized tick species are defined as vector for viral species either indicating a highly specialized and specific relationship or alternatively may be due to the lack of vector competence studies using tick-borne viruses.

Table 1. Classification of tick-borne viruses (Shi et al. 2018).

Family	**Genus**	**Species**
Bunyavirales		
Nairoviridae	*Orthonairovirus*	Crimean-Congo haemorrhagic fever virus
		Dugbe virus
		Nairobi sheep disease virus/Ganjam virus
		Farallon virus
		Hughes virus
		Punta Salinas virus
		Soldado virus
		Zirqa virus
Peribunyaviridae	*Orthobunyavirus*	*Tete orthobunyavirus serogroup*
		Bahig virus
		Matruh virus
Phenuiviridae	*Phlebovirus*	*Uukunimi group*
		Uukuniemi virus

Families unassigned to any order	**Genus**	**Species**
Asfarviridae	*Asfivirus*	African swine fever virus
Flaviviridae	*Flavivirus*	*Mammalian tick-borne flavivirus group*
		Kyasanur Forest disease virus
		Alkhumra haemorrhagic fever virus
		Louping ill virus
		Omsk haemorrhagic fever
		Powassan virus
		Deer tick virus
		Tick-borne encephalitis virus
		Gadgets Gully virus
		Karshi virus
		Langat virus
		Royal Farm virus
		Seabird tick-borne flavivirus

		SFTS/Heartland group
		Heartland virus
		Hunter island virus
		Severe fever with thrombocytopenia syndrome virus
		Bhanja group
		Bhanja virus
		Lone Star virus
		Palma virus
		Kaisodi group
		Kaisodi virus
		Khasan virus
		Lanjan virus
		Silverwater virus
Mononegavirales		
Nyamiviridae	*Nyavirus*	Midway virus
		Nyamanini virus
		Sierra Nevada virus
Rhabdoviridae	*Ledantevirus*	Barur virus
		Kolente virus
		Yongjia tick virus 2
	Vesiculovirus	Isfahan virus
	Unassigned rhabdoviruses	Long Island tick rhabdovirus
		Zahedan rhabdovirus
		Sawgrass virus group
		Connecticut virus
		New Minto virus
		Sawgrass virus

		group
		Meaban virus
		Saumarez Reef virus
		Tyuleniy virus
		Putative third group
		Kadam virus
Orthomyxoviridae	*Quarjavirus*	Johnston Atoll virus
		Quaranfil virus
	Thogotovirus	Dhori virus
		Jos virus
		Thogoto virus
Reoviridae	*Coltivirus* (*Spinareovirinae*)	Colorado tick fever virus
		Eyach virus
	Orbivirus (Sedoreovirinae)	*Chenuda virus species*
		Baku virus
		Chenuda virus
		Essaouira virus
		Huacho virus
		Kala Iris virus
		Mono Lake virus
		Sixgun city virus
		Chobar Gorge virus species
		Chobar Gorge virus
		Great Island virus species
		Great Island virus
		Kemerovo virus
		Lipovnik virus
		Tribec virus
		St Croix River virus
		Wad Medani virus species
		Seletar virus
		Wad Medani virus

1.2. Crimean-Congo haemorrhagic fever virus

1.2.1. Historical background

Crimean-Congo haemorrhagic fever (CCHF) is the most widely distributed tick-borne viral disease in humans, ranging from western China through South Asia and the Middle East to southeastern Europe and Africa (Bente et al. 2013).

At the beginning of the 12th century, humans have presumably had the disease now known as the Crimean-Congo haemorrhagic fever in the area what is now Tajikistan (Hoogstraal 1979).

The disease was first described in the literature in 1945 (Chumakov 1945) when the first documented CCHF outbreak occurred in Crimean region of the former Soviet Union in 1944 (Bente et al. 2013), where 200 military personnel were infected and developed an acute febrile illness with haemorrhage and shock, and nearly 10% died. In order to investigate the situation, Moscow sent a team led by Mikhail Chumakov, and the researchers found out that the cases were related to tick exposure. According to Grashchenkov's report in 1945, more than 3000 blood-sucking arthropods were collected, but ticks were the most abundant among them (Grashchenkov 1945), and *Hyalomma marginatum*, which is now recognized to be the principal vector of Crimean-Congo haemorrhagic fever virus (CCHFV) in Russia and vicinity, were collected the most. Their observation was that the population of hosts of *Hyalomma* ticks such as hares and other wild animals had increased as vast areas of cultivated lands had been left during German occupation, and soldiers and agricultural workers were suffering a great number of tick bites. Shortly after, Chumakov and his team proved that "Crimean haemorrhagic fever" (CHF) was a tick-borne viral disease (Chumakov 1965; Chumakov 1974). However, laboratory studies were limited due to the difficulties on cultivating of CHF causative agent for more than two decades after CHF was identified.

In 1967, Chumakov's research group discovered that new-born mice, injected with CCHF samples from patients or infected ticks developed a lethal illness. The technique of virus culture in suckling mice led to an explosion in research on Crimean haemorrhagic fever virus (CHFV).

In the meantime, a virus, given the name Congo virus was isolated from a 13-year-old child with fever, headache, nausea, and vomiting in the former Belgian Congo or now Democratic Republic of the Congo in 1956.

In 1969, it was discovered that the two isolated viruses were identical. As a consequence, the disease and its causative agent were called "Crimean-Congo haemorrhagic fever (CCHF)" and "Crimean-Congo haemorrhagic fever virus (CCHFV)", respectively since the late 1970s.

CCHFV is classified as the biosafety level 4 (BSL-4) agent due to its high potential for human-to-human transmission and high risks for laboratory-acquired infections, and the lack of a specific vaccine. CCHFV is also classified as a Category C bioterrorism agent by National Institute of Allergy and Infectious Disease of the National Institutes of Health in the USA.

1.2.2. Classification and Structure

The CCHFV is a tick-borne, negative-sense, single-stranded RNA virus and a member of the genus *Orthonairovirus*. *Orthonairovirus* belongs to the family Nairoviridae of the order Bunyavirales, and is comprised of at least 43 viruses assigned to 15 groups/species (Abudurexiti et al. 2019). Nairoviridae family is classified into three genera (*Orthonairovirus*, *Shaspivirus*, and *Striwavirus*). These viruses are predominantly transmitted by ticks among mammals or birds.

The CCHFV has spherical shape with a diameter of approximately 80-100 nm (Figure 1). It is a lipid-enveloped virus with an outer surface covered with glycoproteins Gn and Gc, which bind cell receptor recognition sites (Bente et al. 2013). CCHFV has three-segmented genome denoted by their size with different functions (Walter & Barr 2011): S (small) segment encodes nucleocapsid (N), M (medium) segment the glycoproteins (Gn and Gc), and L (large) segment RNA-dependent polymerases (RdRp) (Bente et al. 2013). M segment of the genus *Orthonairovirus* is larger than other members of Nairoviridae family and it is translated into polyprotein precursors PreGn (140 kDa) and PreGc (85 kDa) which result in the structural glycoproteins Gn (37 kDa) and Gc (75 kDa) (Papa et al. 2002; Sanchez et al. 2002; Altamura et al. 2007; Bergeron et al. 2007).

In the CCHFV genome, the genetic diversity of M segment increases the variation frequency of glycoproteins (Gn and Gc) in the viral life cycle. Glycoproteins are responsible of binding the virus to cellular receptors and facilitate viral infection in different vertebrate hosts.

In general, Orthonairoviruses differ from other members of Bunyavirales family with their large (L) segment which is nearly twice the size in CCHFV and other members in the genus. L segment encodes RdRp and is responsible for the initiation of transcription and genome replication in the host cell (Ergonul 2006).

Generally, a great genetic diversity with complex evolutionary patterns characterizes CCHFV (Anagnostou & Papa 2009).

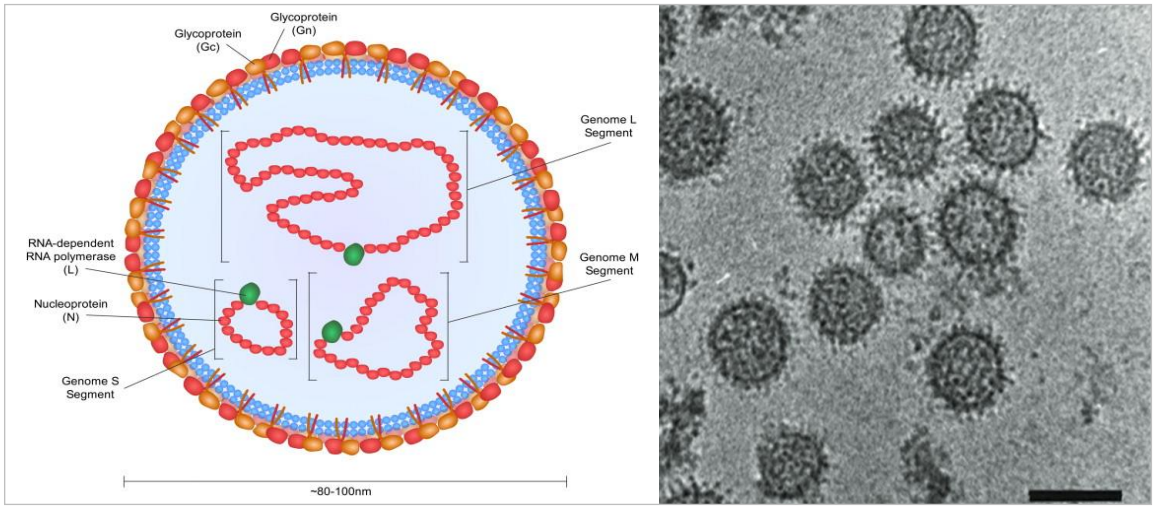


Figure 1. CCHFV virion structure. Diagrammatic representation of CCHFV virion in cross-section. The surface spikes comprise two glycoproteins termed Gn and Gc. CCHFV genome is composed of three RNA segments (L, large; M, medium; S, small) encapsidated by N protein and associated with RNA-dependent polymerases (left) (Bente et al. 2013). Transmission electron micrograph (TEM) of CCHFV (right) (International Committee on Taxonomy of Viruses 2011).

1.2.3. Epidemiology and Geographic distribution

CCHF has Old World distribution (Figure 2). The disease has been reported in Africa (Burkina Faso, Democratic Republic of the Congo, Kenya, Mauritania, Republic of the Congo, Senegal, South Africa, Sudan, Tanzania, Uganda), Asia (Afghanistan,

Armenia, China, Georgia, India, Iran, Iraq, Kazakhstan, Kuwait, Kyrgyzstan, Oman, Pakistan, Russia, Saudi Arabia, Tajikistan, Turkey, Turkmenistan, United Arab Emirates, Uzbekistan), and Europe (Albania, Bulgaria, Greece, Hungary, Kosovo, North Macedonia, Russia, Serbia, Spain, Turkey, Ukraine) (Hoogstraal 1979; Ergonul & Whitehouse 2007; Avšič 2009; Pigott et al. 2017; Al-Abri 2017).

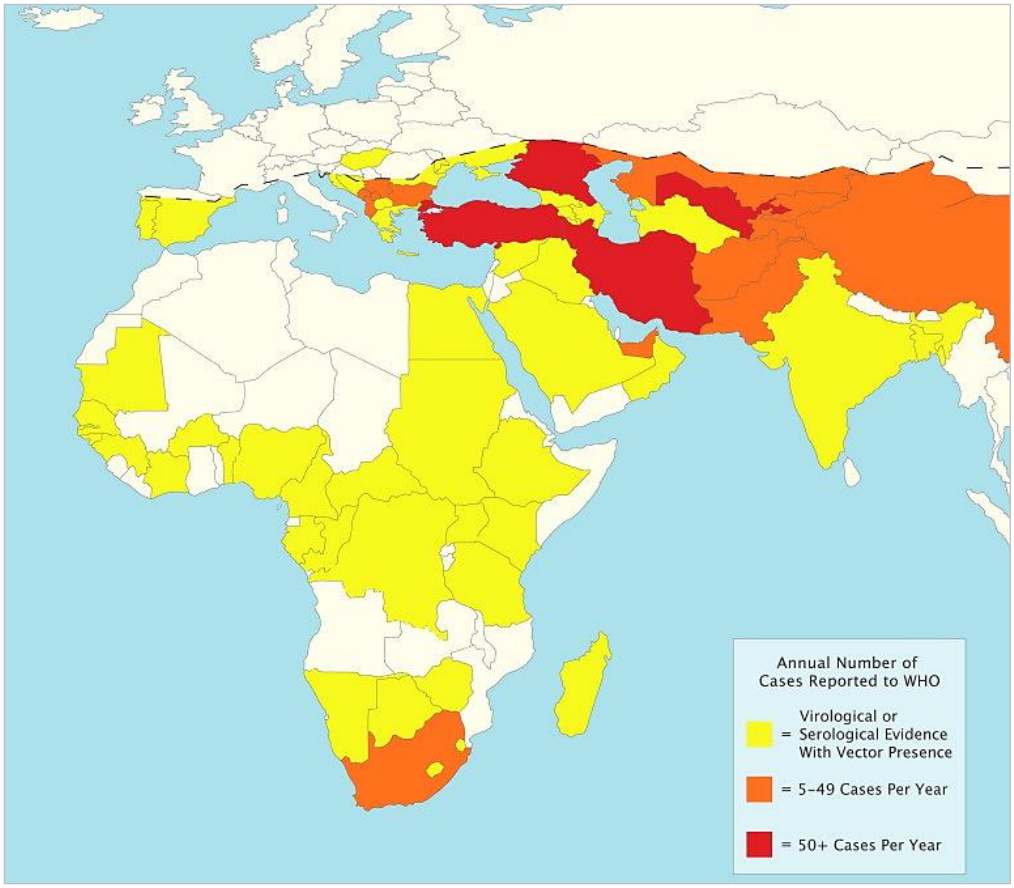


Figure 2. Geographic distribution of CCHF and annual number of reported cases of CCHF by country. Countries in red report more than 50 human cases annually to the WHO. Countries in orange report fewer than 50 cases. Countries in yellow have not reported human cases but there is serological or virological evidence of CCHFV with its vector presence (Bente et al. 2013).

However, due to lack of efficient and active surveillance systems and/or the lack of diagnostic assays, the exact annual CCHF case numbers and its frequency of occurrence remain uncertain, and the virus could be circulating unnoticed. (Ergonul & Whitehouse 2007; Avšič 2009; Al-Abri 2017; Pigott et al. 2017).

Since 2000, the virus has been reported from Turkey, Iran, India, Pakistan, Senegal, Kenya, Mauritania, Albania, Kosovo, Bulgaria, Greece, Spain leading to the concern that CCHF is expanding in its current geographical distribution (ECDC 2019).

Even though the earliest known case of CCHF in Kosovo was reported in 1954, since 1989, outbreaks have occurred every 4-5 years (Vesenjak et al. 1991, Ahmeti et al. 1996). A large increase of cases has also occurred in Turkey since the first human case was observed in 2002, over 6300 cases were diagnosed in the following 10 years. In Iran, the first human case was recognized in 1999 and the disease has increased in prevalence since 2000 (Chinikar et al. 2010). Recently, the discovery of autochthonous cases of CCHF in Spain reflects an alteration in geographical distribution of the disease. These cases in Spain, a historically unaffected region, indicate CCHFV must be viewed as an emerging pathogen (Negredo et al. 2017).

The actual number of CCHF cases is estimated to be higher than that recorded, as the disease typically occurs in remote areas (Saijo et al. 2010). In endemic countries in the Northern hemisphere, CCHFV is usually identified in the spring and early summer. Seasonality of the outbreak might be related with tick abundance and activity, as well as increasing viral load in animals in the areas (Saijo et al. 2010). CCHF outbreaks have widely occurred among agricultural workers, slaughterers, and shepherds, all of whom handle livestock. If the areas are abundant with small mammals such as hares and hedgehogs and large mammals such as sheep and cattle, the virus may circulate unnoticeably, with human cases occurring only occasionally, when farmers handling livestock are bitten by infected ticks. Contrary, in regions where density of small mammals is high, but less large mammals, humans in the region are more likely to be bitten by ticks, resulting in a high number of infections. This situation occurred in 1944 in Crimea, following the World War II, where wild hares continuously reproduced on abandoned farms during German occupation, whereas the number of livestock drastically decreased. In this situation, adult *Hyalomma* ticks were required urgently to get their blood meal from a large animal source. Since the region was less abundant with numbers of large mammals, there was an explosion in cases among agricultural workers and soldiers. The appearance of hot spots of CCHFV in other geographical locations, for instance in Turkey, may be explained by the similar situation (Bente et al. 2013).

1.2.4. Phylogeny

Earlier studies on phylogenetic analyses have indicated high diversity in the genomes of the sequenced CCHFV strains from different countries (Papa et al. 2002; Hewson et al. 2004; Deyde et al. 2006; Sun et al. 2009; Zhou et al. 2013).

Studies on phylogeny of CCHFV are based on S genome segment sequences. S segment is the most conserved segment across the CCHFV genome (Bente et al. 2013). The phylogenetic analyses based on S segment (Figure 3) showed that CCHFV strains cluster in seven distinct genetic lineages in association with the geographic distributions, West-Africa in clade I (Africa 1), Central Africa in clade II (Africa 2), South Africa and West Africa in clade III (Africa 3), Middle East and Asia in clade IV (Asia 1 and Asia 2), Europe and Balkans in clade V (Europe 1) and divergent strains isolated from ticks in Greece and Turkey in clade VI (Europe 2).

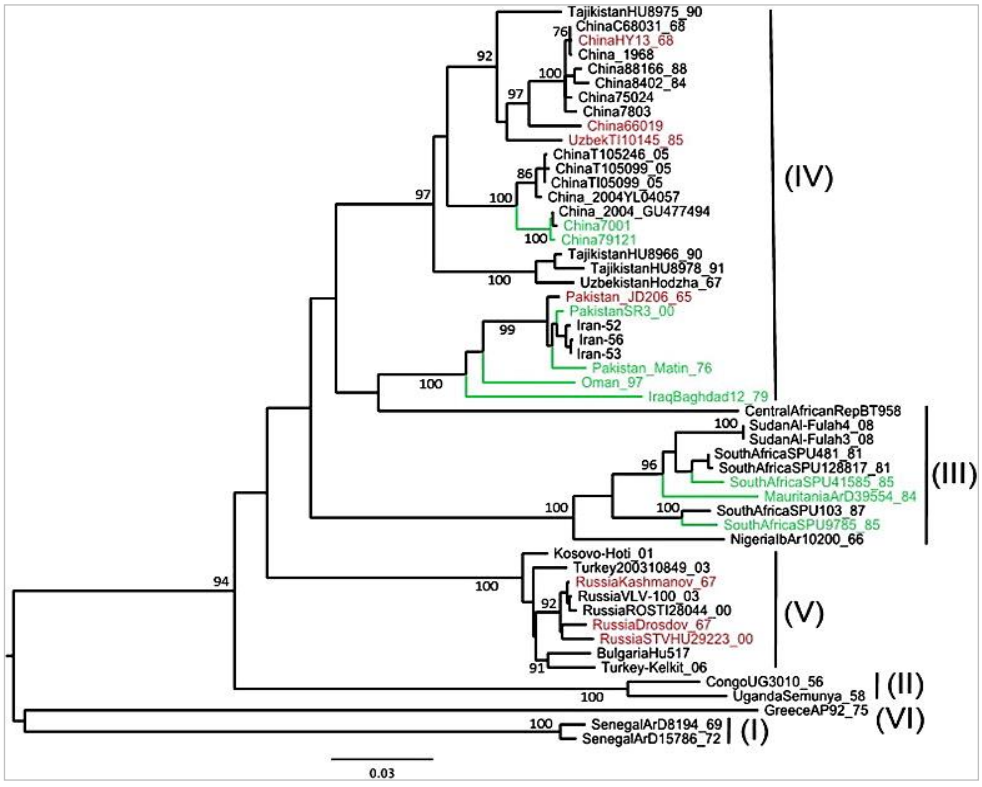


Figure 3. Phylogenetic tree of CCHFV based on S segment. It shows the genetic relationships of the small (S) segment of CCHFV. Strains marked as green show evidence of recombination and strains marked as red show evidence of reassortment (Bente et al. 2013).

Generally, the earliest lineages are found in Africa, with the exception of strain Ap92, which was isolated in Greece.

Phylogenetic analyses were also performed based on M and L segment sequences (Figure 4, Figure 5). However, due to reassortment events, many inconsistencies were identified among the clades, suggesting that the M and L segments of CCHFV are highly diverse (Anagnostou & Papa 2009; Zhou et al. 2013).

CCHFV's distinct clades are sometimes characterized by different potentials in terms of pathogenicity. For instance, very few and mostly mild cases in humans have been reported in Greece which the isolates belong to clade VI (Papa et al. 2018). Furthermore, studies showed that similar genotypes are found in distant geographical locations, supporting the idea that virus or infected ticks may be carried over long distances during bird migration (Gonzalez-Scarano & Nathanson 1996).

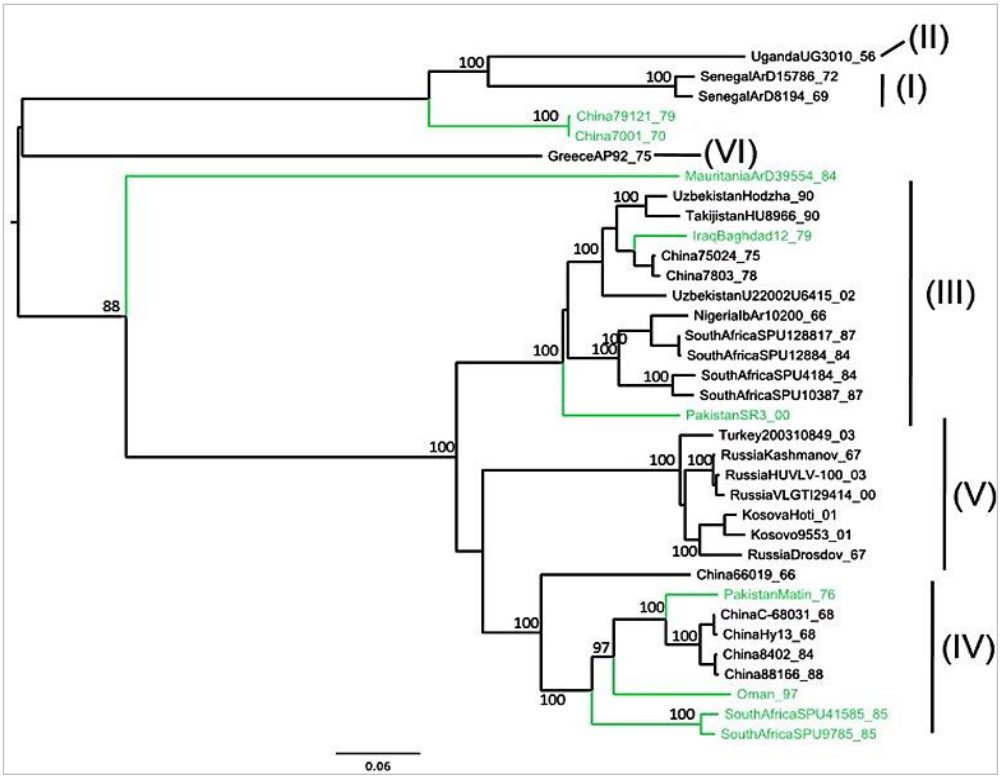


Figure 4. Phylogenetic tree of CCHFV based on M segment (Bente et al. 2013).

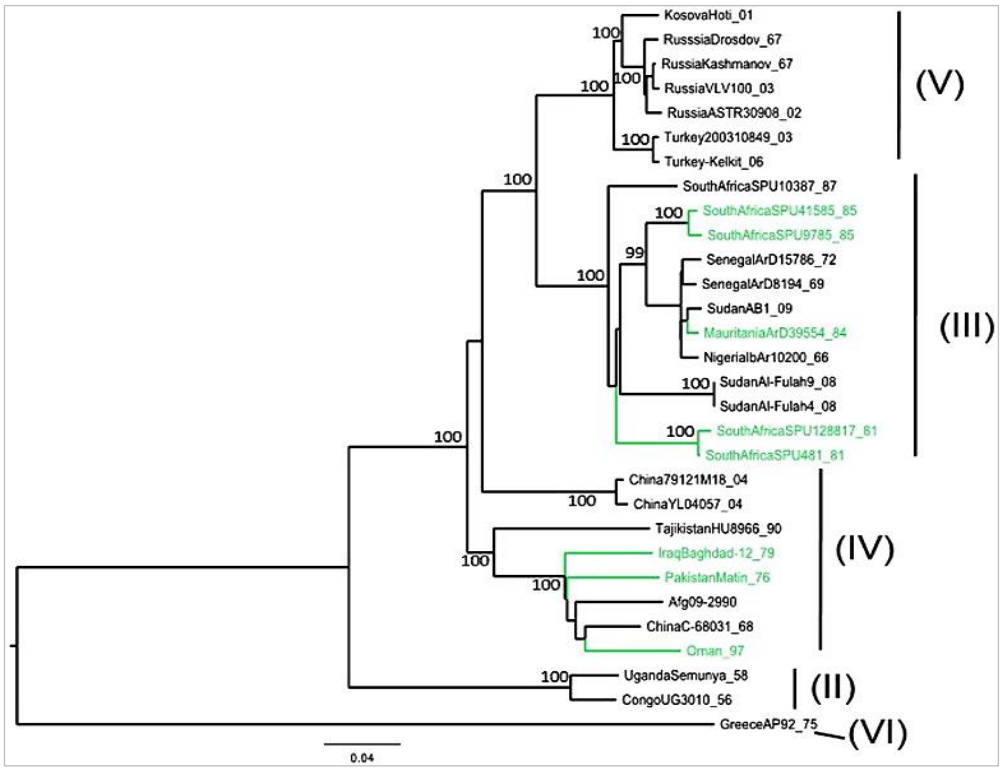


Figure 5. Phylogenetic tree of CCHFV based on L segment (Bente et al. 2013).

1.2.5. Transmission

The virus is transmitted through the bite of its main vector, tick species in the genus *Hyalomma* (Ixodidae). *Hyalomma* ticks take their blood meal from various domestic ruminants (sheep, goats, and cattle) and wild animals (hares, hedgehogs, certain rodents, and ostriches) (Figure 6).

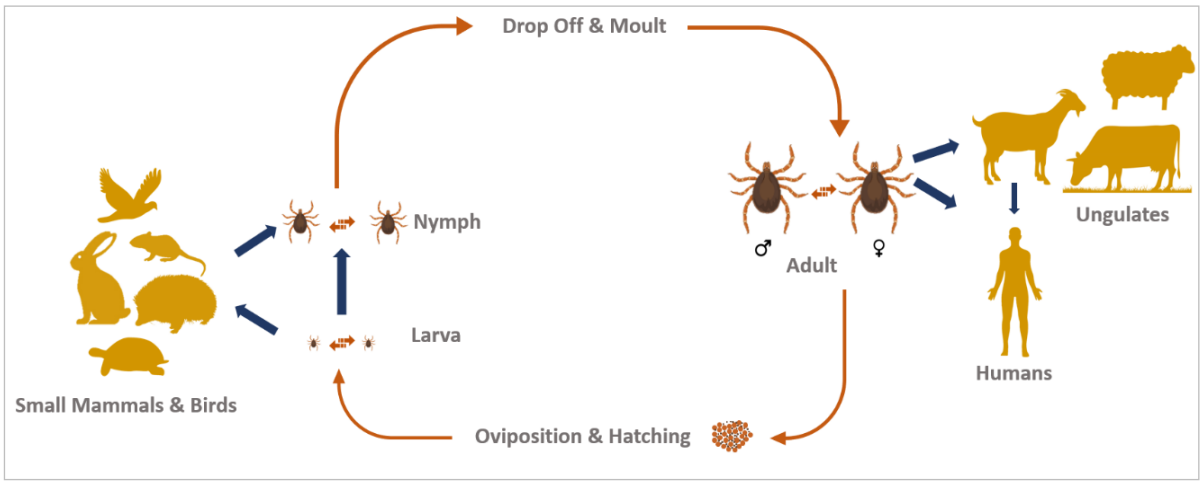


Figure 6. Life cycle of *Hyalomma* spp. ticks as two-host ticks.

Animals, in general, show no clinical manifestations, and occurrence of the disease in human populations is used to understand whether the virus circulates in a given geographic location (Bente et al. 2013).

Ticks transmit CCHFV to various mammals by taking the blood meals for their maturation and egg production, however, the occurring viremia in mammals does not last long. On the other hand, since ticks are true natural reservoirs of CCHFV, they remain infected for several years while humans are only dead-end hosts. Humans acquire infection from tick bites, and from direct contact with the blood or tissues of infected livestock or human patients. Other routes of transmission for CCHFV are through infected mother to offspring, sexual contact, transfusion of blood, and get in contact with different body fluids of infected patients and animals (Bente et al. 2013; Pshenichnaya et al. 2017).

Hyalomma spp. are the main vectors of CCHFV. Ticks of the genus *Hyalomma* actively seek their hosts (hunter strategy), which is in contrast to the *Ixodes* ticks which passively wait in the vegetation for a vertebrate host coming accidentally into contact with them (ambush strategy) (Sonenshine & Roe 2014).

H. marginatum Koch, 1844 is the major vector of CCHFV transmission from Kosovo to Pakistan (Bente et al. 2013). The wide dispersion of *Hyalomma* ticks reflects their tolerance of diverse environments, including savannah, steppe, and lightly wooded areas, and the ability of their aggressively questing larvae and nymphs to feed on a variety of hosts, including hedgehogs, hares, and ground-feeding birds, while the adults actively seek out sheep, cattle and other large mammals (Hoogstraal 1979).

Apart from *H. marginatum*, other species of *Hyalomma* were also reported to carry the CCHFV such as *Hyalomma excavatum* Koch, 1844, *Hyalomma lusitanicum* Koch, 1844, *Hyalomma rufipes* Koch, 1844, and *Hyalomma truncatum* Koch, 1844 (Causey et al. 1970; Gonzalez et al. 1991; Estrada-Peña et al. 2012; Akuffo et al. 2016). Besides, some other tick species of genera *Rhipicephalus*, such as *Rhipicephalus bursa* Koch, 1844, *Rhipicephalus turanicus* Pomerantsev, 1936 detected in Albania, Turkey, and Greece, and species of genera *Dermacentor* such as *Dermacentor marginatus* Sulzer, 1776 was tested positive in Turkey (Papadopoulos & Koptopoulos 1978; Gargili et al. 2011; Papa et al. 2011; Yesilbag et al. 2013).

1.2.6. Clinical manifestations, Diagnosis and Prevention

There are four phases that are involved in the infection of the CCHFV: Incubation period (non-symptomatic phase), pre-haemorrhagic, haemorrhagic and convalescent (symptomatic phases) (Figure 7). The incubation period lasts from 3-7 days of infection. The disease starts suddenly with the pre-haemorrhagic phase for 4-5 days. The main symptoms are headache, high fever, abdominal pain, myalgia, hypotension and flushed face (Hoogstraal 1979). Severe symptoms start appearing as the disease progresses such as petechiae (red spots on skin), ecchymosis (extravasation of blood), epistaxis (nose bleeding), gum bleeding and emesis (Ergonul et al. 2004; Bakir et al. 2005). Additionally, nausea, diarrhoea, neuropsychiatric and cardiovascular changes can occur (Whitehouse 2004). When the disease is not treated, patients may succumb due to multiorgan failure.

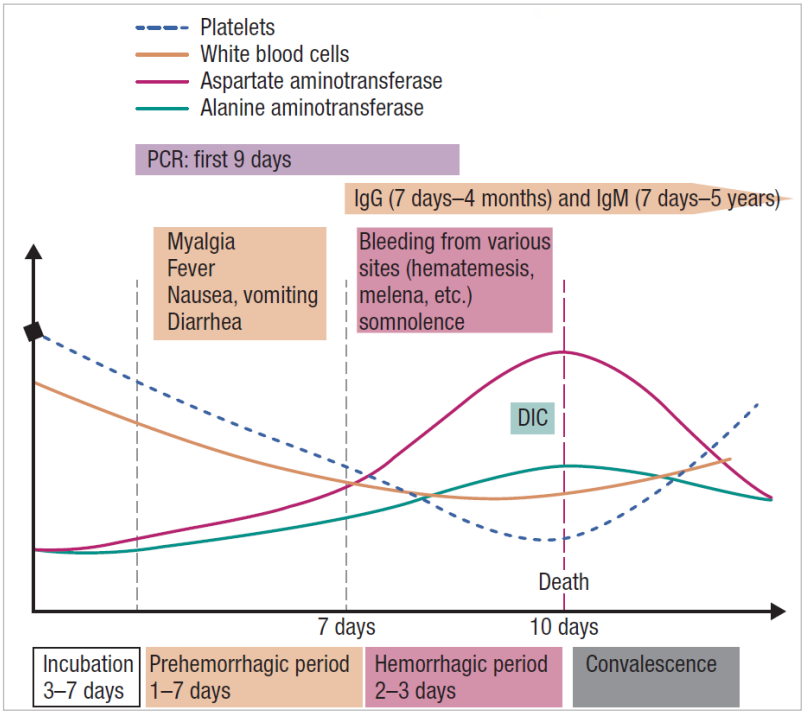


Figure 7. The clinical course of CCHF. The starting point is the entrance of the CCHFV to the human through tick bite or a contact with infected material such as body fluids (Endy 2020).

Due to classification of CCHFV as a World Health Organization (WHO) Risk Group 4 pathogen, the laboratory diagnosis is complicated. There are different techniques for the diagnosis, including enzyme-linked immunosorbent assay (ELISA), real-time polymerase chain reaction (qPCR), antigen detection, serum neutralization and isolation of the virus by cell culture (Vanhomwegen et al. 2012; WHO 2015). However, PCR diagnosis is still not available in many laboratories of rural countries where CCHFV can be considered as endemic.

The most effective measure to prevent CCHFV infection is the avoidance of tick habitats. The risk of acquiring CCHFV can be minimized by using repellents, proper clothing, proper equipment in professional settings such as hospitals, butcheries, and slaughterhouses. In order to raise awareness regarding the mode of CCHFV transmission and the measures to prevent infection, local and global educational campaigns should be undertaken.

1.3. Role of wild animals and livestock in the maintenance and transmission of CCHFV

Ticks are both natural reservoir and vector for the virus, carrying it from wild animals to domestic animals and humans. Vertebrates are crucial part of CCHF epidemiology, as they provide blood meals to allow reproduction of tick populations. They play a key role in transportation of ticks across wide geographic areas, and transmission of CCHFV to ticks and humans during the period of viremia.

Small mammals, in particular hares and hedgehogs, act as amplifying hosts for the immature stages of the ticks, whereas domestic large mammals such as cattle, goats, and sheep are the usual hosts for the adult ticks. In general, CCHFV replicates and produces a short viral replication period in small mammals (2-15 days), followed by development of anti-CCHFV antibodies. Thus, it appears that small mammals do not act efficiently as long-term CCHFV reservoirs. They, however, have an essential role in CCHFV ecology, and population increases (e.g., in hares) have been associated with disease outbreaks.

Large domestic mammals play an important role in supporting a high tick load and bringing ticks into close proximity to agricultural workers. Additionally, these animals can directly expose humans to CCHFV through infected blood or crushing of engorged ticks on the animals during slaughtering (Chinikar et al. 2010; Mustafa et al. 2011; Sharifi-Mood et al. 2014; Kadanali et al. 2009; Sargianou et al, 2013).

Apart from mammals, birds also serve as a blood meal source for ticks. Multiple host ticks such as larvae and nymphs of *Hyalomma* species frequently get the blood meal from ground-dwelling birds. However, there is no evidence reporting birds to become viremic when CCHFV-infected ticks feed on them. The only exception is ostriches, developing notable viremia and associating with human infection (Hoogstraal 1979; Zeller et al. 1994). Although birds do not develop viremia, they may have a potential role in the introduction of CCHFV-infected ticks into new geographic regions by carrying feeding immature ticks.

CCHFV has been isolated from many domesticated and wild animals from various countries including cattle, sheep, goats, water buffalo, hares (European hare, *Lepus europaeus*), hedgehogs (Four-toed hedgehog, *Erinaceus albiventris*; Long-eared

hedgehog, *Hemiechinus auritus),* and several small wild rodents (e.g., white-tailed rat, *Mystromys albicaudatus*; red rock rat, *Aethomys chrysophilus*; bushveld gerbil, *Gerbilliscus leucogaster*; four-striped grass mouse, *Rhabdomys pumilio*; highveld gerbil, *Gerbilliscus brantsii*; southern multimammate mouse, *Mastomys coucha*; natal multimammate mice, *Mastomys natalensis*; and Cape short-eared gerbil, *Desmodillus auricularis*) (Spickler 2019).

Serological evidence of exposure has been reported in many additional species, such as horses, donkeys, pigs, camels, dogs, red foxes (*Vulpes vulpes*), wild dogs (*Lycaon pictus*), Pallas cats (*Felis manul*), genets (*Genetta genetta*), a number of African ungulates (e.g., white rhinoceros, *Ceratotherium simum*; black rhinoceros, *Diceros bicornis*; giraffe, *Giraffa camelopardalis*; African buffalo, *Syncerus caffer*; many antelopes such as *Aepyceros melampus*, *Connochaetes gnou*, *Connochaetes taurinus*, *Hippotragus equinus*, *Hippotragus niger*, *Oreotragus oreotragus*, *Oryx gazella*, *Taurotragus oryx*, *Tragelaphus angasii*, *Tragelaphus scriptus*, *Tragelaphus strepsiceros*), various rodents and bats (Spengler et al. 2016; Spickler 2019).

Among seropositive species, susceptibility has been confirmed by experimental infection in equids (horses, donkeys), European hares, scrub hares (*Lepus saxatilis*), and some wild rodents (Spengler et al. 2016). However, some mammals may be resistant to this virus. For instance, one group was able to recover CCHFV from the blood of experimentally infected long-eared hedgehogs (*Hemiechinus auritus*) but not European hedgehogs (*Erinaceus europaeus*) (Spickler 2019).

In general, these studies found low level of viremia in horses, donkeys, sheep, cattle, hares, ostriches and other animals in endemic areas, without obvious signs of disease, and in some cases transmitted the virus to feeding ticks. However, all studies investigating the role of domestic and wild animals in transmission of the virus and the researches on animal susceptibility were performed in the 1970s (Spengler et al. 2016). Despite the fact that CCHFV does not harm economy on livestock animal production, it is crucial to conduct the serological screening of animal serum samples of CCHFV-specific antibodies.

1.4. Molecular evolution of CCHFV

1.4.1. Evolution of CCHFV viral genome

When Charles Darwin's "On the Origin of Species" was published in 1859, viruses had yet to be discovered. However, since that time a lot of knowledge on evolutionary trends was acquired. 40 years after publication, the concept of viruses was proposed and viruses, in particular RNA viruses, are known to evolve rapidly as a result of their error-prone replication (Anagnostou & Papa 2009).

RNA viruses generally show extremely high mutation rates that allow fast evolution than others with short generation time (Drake & Holland 1999). Arthropod-borne RNA viruses (arboviruses), on the other hand, show remarkably low levels of evolution rate comparing with other RNA viruses. This is explained by double-filter concept that the evolution of arbovirus is highly restricted by maintaining high fitness in both arthropod and vertebrate host environments (Drake & Holland 1999). Interestingly, Deyde et al. (2006) detected high levels of diversity in the complete genome sequences of CCHFV isolates. Such variation may contradict the requirement for an evolutionary compromise between host and vector. Alternatively, these high levels of diversity may be occurred due to high-efficiency vertical transmission in the tick that avoids the need for maintaining high fitness vertebrates and the double-filter concept may not apply on CCHFV (Deyde et al. 2006).

CCHFV is the most genetically diverse among arboviruses, with 20, 22, and 31% nucleotide sequence differences among S, L, and M segments, respectively (Bente et al. 2013). Discrete lineages of the viruses can be found within the same geographic locations (the specific environment of the endemic geographic regions such as ecological niches, climatic factors that might affect tick population and virus dissemination), whereas closely related viruses have been isolated in distant locations (Bente et al. 2013), suggesting that widespread expansion of CCHFV has occurred in the past, possibly by ticks carried on migratory birds or through the international animal trade. Moreover, the occurrence of genetic reassortment among genome segments and switch of different genomic regions between different strains by genetic recombination contribute to complex evolutionary history of CCHFV (Bente et al. 2013).

M segment of CCHFV genome has the highest genetic diversity compared to S and L segments, which is expected as it encodes glycoproteins that allow the virus to bind to host cells. L and S segments present coevolution patterns and although they are quite similar, it has been shown that there are some differences indicating reassortment events (Deyde et al. 2006). It is apparent that both M and L segments have higher evolution rate than the S segment, which is the most conserved part of the genome as it encodes the nucleoprotein.

An earlier study on the evolution of CCHFV indicated that current CCHF viruses have arisen from an ancestral agent that existed around 3000-3500 years ago, and suggested that the human migration and increase of number of livestock farming and trade may have contributed to its geographic dispersion (Carroll et al. 2010).

1.4.2. Standard genetic code

All organisms encode information in the form of either deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) that is transcribed to messenger ribonucleic acid (mRNA), which is then decoded by ribosomes during translation process. The genetic code defines how encoded information in genetic material (DNA or RNA sequences) is translated into amino acid sequence of proteins, and it is defined as a set of three-letter combinations of nucleotides called codons, each of which corresponds to a specific amino acid or stop signal.

In general, out of 64 codons, 61 encode 20 different amino acids and the remaining three represent stop codons (Figure 8).

Due to the degeneracy of genetic code, most amino acids are encoded by more than one codon. The various codons that correspond to the same amino acid are termed as synonymous codons. Every species have unique preferences of codons encoding the same amino acids, and therefore the codons used in their genes have unequal frequencies (Plotkin et al. 2006; Plotkin & Kudla 2011).

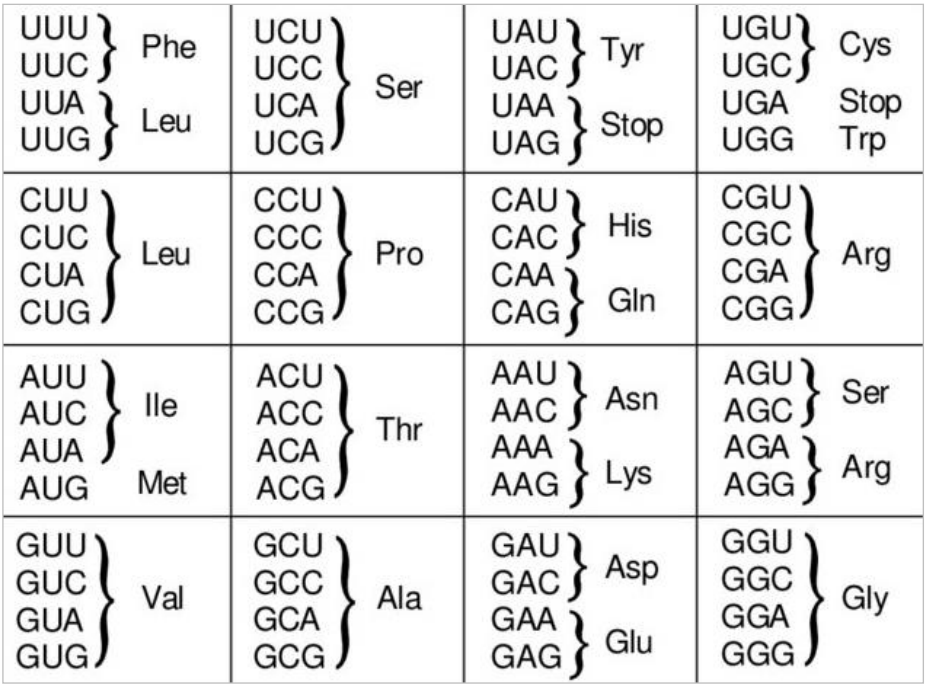


Figure 8. The codon table. The genetic code is composed of four different letters U, C, A, and G, generating in total 64 different three-letter combinations. The codon table summarizes all 64 codons and clarifies which codon specifies which amino acid.

1.4.3. Codon usage bias

Viruses do not have their own proteosyntetic apparatus and therefore have to "kidnap" the host translation machinery and use ribosomes and tRNAs of their hosts. Due to this fact, all viruses face the strong evolutionary pressure to use the same codons as their hosts.

Some synonymous codons are preferred to be used more widely than others and this concept is known as codon usage bias that refers to differential usage of synonymous codons. Codon usage bias has been studied in a wide range of organisms, from prokaryotes to eukaryotes and viruses (Sharp et al. 1993; Akashi H 1997; Butt et al. 2016). However, how such biases arise is a much-debated area of molecular evolution.

Different factors have been suggested to be related to codon usage bias, including gene expression level (reflecting selection for optimizing the translation process by tRNA abundance), guanine-cytosine content (GC content, reflecting horizontal gene transfer or mutational bias), guanine-cytosine skew (GC skew, reflecting strand-specific mutational bias), amino acid conservation, protein hydropathy,

and transcriptional selection (Ermolaeva 2001; Knight et al. 2001; Chen et al. 2014; Quax et al. 2015).

The codon usage bias is mainly affected by compositional constraints under mutational pressure and natural selection. Natural selection and mutational pressure are the two major evolutionary forces responsible for codon usage variation among genomes (Shields et al. 1988; Stenico et al. 1994).

Many studies on codon usage bias of various viruses have showed that the main driver on shaping codon usage patterns is mutational pressure than natural selection (Jenkins & Holmes 2003; Sharp et al. 2010; Wong et al. 2010; Wang et al. 2011). However, for many DNA and RNA viruses, mutational pressure is not the only responsible factor on establishing codon usage patterns (Butt et al. 2014; Rahman et al. 2018). Comparing with prokaryotic and eukaryotic genomes, viral genome has some certain features such as dependence on their hosts for replication, synthesizing protein, and transmission of proteins. The interaction between virus and host is considered to affect survival, adaptation, and evolution of virus, and escape from host's immune system (Burns et al. 2006; Mueller et al. 2006; Costafreda et al. 2014).

1.5. *Hyalomma marginatum* and modeling its environmental suitability to Central Europe

Ticks and tick-borne diseases are significant emerging threats to animals and humans. Diseases that are transmitted by ticks require interactions between vertebrate hosts, arthropod hosts, and pathogens, constrained by a set of environmental variables. Environmental conditions and host migration strongly influence tick reproduction, dispersion and survival.

Hyalomma genus is a small genus, with 27 species mainly found in the Afrotropical Region and parts of the Palaearctic Region (Guglielmone et al. 2014). *H. marginatum* Koch, 1844 is the type-species of the *H. marginatum* complex (Capek et al. 2014), and it is widely distributed through southern Europe and North Africa to Ukraine and southern Russia and the Middle East (Apanaskevic & Horak 2008). *H. marginatum*, as an important arthropod disease carrier, is a hematophagous ectoparasite of wild and domesticated animals and humans. This species is not only the vector of CCHFV but also of other pathogens such as West Nile, Thogoto, Dhori and other viruses (Kaiser & Hoogstraal 1974), as well as *Rickettsia aeschlimannii* (Sentausa et al. 2014), *Babesia caballi* and *Theileria annulata* (De Kok et al. 1993; Walker et al. 2003).

H.marginatum is the main vector species for CCHFV in Europe. The geographic distribution of the virus coincides with spreading of its vector species. In Central Europe, *H. marginatum* was first detected in southern Germany in 2007 (Kampen et al. 2007). In the globalized world, the vector can be faced anywhere as an introduced species. Last years, *H. marginatum* records in other countries of Central Europe are evidenced in 2011 in Hungary (Földvári et al. 2011), in 2014 in Slovakia (Capek et al. 2014), and in 2018 in Austria (Duscher et al. 2018). Generally, permanent populations of *H. marginatum* are limited to the warmer areas of the Mediterranean basin in Europe and its permanent populations have not been recognized so far in Central Europe, probably due to current climatic conditions. However, its possible northern spread and establishment of permanent populations are highly important due to passive transportation of immature stages of *H. marginatum* are regularly occurred by migratory birds flying to temperate Europe (Hoogstraal et al. 1961; Hillyard 1996).

Ecological niche modeling (ENM), also known as species distribution modeling (SDM), is a widely used tool that requires occurrence data of interested species and environmental data to predict habitat suitability in a particular area (Guisan & Zimmermann 2000; Peterson et al. 2015). Different types of niche concepts were recognized to define the species niche such as Grinnellian, Eltonian, and Hutchinsonian niche concepts (Chase & Leibold 2003). The Grinnellian niche concept, however, was selected to be used as it is the most appropriate approach in relation to the species and the purpose of the work. Grinnellian niche concept uses the range of values of environmental factors that are necessary and sufficient to allow a species to carry out its life and produce a correlative model of those environmental factors that meet a species' ecological requirements and predict the relative habitat suitability (Chase & Leibold 2003).

2. Aims of the Thesis

CCHF has become one of the most geographically widely distributed tick-borne viral diseases in the world. Its causative agent, CCHFV, is among the deadliest human pathogens in Africa and Eurasia. CCHFV is transmitted mainly by tick species in *Hyalomma* genus, particularly by *H. marginatum* in Europe, but other tick species such as representatives of genera *Dermacentor* and *Rhipicephalus* are also involved in virus transmission.

The first aim of this work was to deepen our understanding of evolutionary adaptation of CCHFV to the different tick species (mainly representatives of genera *Hyalomma* and *Rhipicephalus*). This adaptation is measured by so called "codon adaptation index (CAI)". The first aim of this study was to determine if viruses collected from the same (or closely related) tick species phylogenetically cluster to each other and if they show higher values of CAI for this particular tick species and to understand how CCHFV adapts on its different vector species.

The second aim was to carry out field works in Kosovo and Czechia in order to collect ticks, search for *Hyalomma* species, and investigate possible presence of CCHFV in the collected ticks. The Balkan region, in particular Kosovo, is a well-known highly endemic region for CCHF. Therefore, tick collection from livestock was conducted in hyper-endemic municipalities of Kosovo and examined for viral detection later. Czechia, on the other hand, is well-known for a high occurrence of tick-borne diseases, such as tick-borne encephalitis (TBE), anaplasmosis, and Lyme disease, and Central Bohemia, in particular, is a high-risk zone for some *Ixodid* tick species. Migratory birds play a crucial role as transportation sources of the ticks from countries to countries. In relation to this objective, ticks were collected from migratory birds which are important in the spread of ticks. All the collected ticks from Kosovo and Czechia were screened for CCHFV presence together.

The third and final aim was to perform ecological niche modeling to map the environmental suitability of *H. marginatum* in Central Europe, focusing in Czechia. *H. marginatum* is the main vector of CCHFV in Europe. Geographic expansions of CCHF have been reported due to climate change impacts, transportation of immature

ticks by migratory birds and international animal trades. These factors may be combined to modify the distributional potential of tick species and allow emergence of CCHF into new geographic regions. In Central Europe, *H. marginatum* has been reported in Hungary, Austria, Slovakia, and Germany to date, and these ticks are introduced by migratory birds to Czechia. I aimed to analyse the environmental suitability *of H. marginatum* in Czechia.

The results of this work gave us new information regarding the adaptation of CCHFV to different tick vector species and gain insight into transmission dynamics of the virus. Moreover, the results of this work were highly important for possible establishment of *Hyalomma* species in Czechia and making inferences regarding the virus transmission for the future.

3. Methods

3.1. Molecular evolution

3.1.1. Data collection

Sequences of all the three segments (S, M, and L) for CCHFV were retrieved from the National Center for Biotechnology Information (NCBI) in FASTA format, accessed on 17 December 2019. The accession numbers and other detailed information of the selected CCHFV sequences, such as isolation date, isolation place, host, and genome size were also retrieved and processed in Microsoft Excel (Microsoft, 2019). For further analyses, the data were undergone a selection process. CCHFV strains which were isolated from tick species (all retrieved sequences were extracted from ticks in *Hyalomma* and *Rhipicephalus* genera) were selected and other strains were excluded for analyses. Multiple sequence alignment were conducted in Clustal format of MAFFT v.7.427 (Katoh & Standley 2013) for sequences of all selected strains.

3.1.2. Nucleotide composition analysis

Nucleotide compositional properties of CCHFV coding sequences were calculated using CAIcal server (http://genomes.urv.es/CAIcal/) (Puigbò et al. 2008a). The overall frequency of occurrence of nucleotides (A%, C%, U% and G%), frequency of each nucleotide at the third site of codons (A3%, C3%, U3%, and G3%), frequencies of occurrence of nucleotides GC at the first (GC1), second (GC2) and third codon positions (GC3), the mean frequencies of nucleotide GC at the first and the second position (GC12), overall GC and AU contents, and AU and GC contents at the third codon positions (AU3, GC3) were calculated. Frequencies of occurrence of nucleotides GC at the third site of synonymous codons (GC3s) were calculated using CodonW (http://sourceforge.net/projects/codonw) (Peden 1999).

AUG and UGG that are only the codons for Met and Trp (no synonymous codons) were excluded from the analyses along with the termination codons (UAG, UAA, and UGA) which do not encode any amino acids and not expected to show any codon usage bias.

3.1.3. Analysis of the Effective Number of Codons (ENC)

Most amino acids (except Met and Trp) in a coding sequence are encoded by two or more codons that are termed as synonymous codons. Effective number of codons (ENC) is a concept that helps to demonstrate codon usage bias. ENC is the index that varies between 20 and 61 (Novembre 2002). A lower value indicates extreme bias in codon usage and the use of only one of the possible synonymous codons for the given amino acid while the highest value indicates no bias in the codon usage and indicates that all the available codons are used equally (Wright 1990; Comeron & Aguade 1998). As a result, the greater the spectrum of codon preference in genome, the lower the ENC value. In general, the genome is known to have highly biased codon usage if the observed ENC value is less than 35. ENC values were calculated using the program CodonW (http://sourceforge.net/projects/codonw) (Peden 1999). To determine correlation between the expected ENC and the GC3s values, the expected ENC values were calculated for different GC3 using the method of Singh et al. (2016):

$ENC\exp = 2 + s + \frac{29}{(s^2+(1-s)^2)}$ where "s" indicates GC contents at the 3rd synonymous codon position (GC3s%).

3.1.4. Relative Synonymous Codon Usage (RSCU) Analysis

RSCU analysis was performed to determine the synonymous codon usage pattern in *Hyalomma*- and *Rhipicephalus*-isolated CCHFV coding sequences. The RSCU value was described by Sharp & Li (1986) and is the ratio of the observed value to the expected value for a given amino acid. Synonymous codons with RSCU values >1.0 show positive codon usage bias and are described as "abundant" codons. RSCU values <1.0 show negative codon usage bias and are described as "less-abundant" codons. When the RSCU value is 1.0, there is no codon usage bias for that amino acid and the codons are preferred equally or randomly (Sharp & Li 1986).

Overrepresented codons possess RSCU values >1.6, while underrepresented codons have values <0.6. Codons with RSCU values ranging between 0.6 and 1.6 are considered unbiased or randomly used. All CCHFV coding sequences isolated from *Hyalomma* and isolated from *Rhipicephalus* were compared with codon usage values of their natural hosts obtained as a table in Codon and Codon Pair Usage Tables

(CoCoPUTs) (https://hive.biochemistry.gwu.edu/review/codon2) (Alexaki et al. 2019), accessed in June 2020, and analysed using CAIcal.

3.1.5. Correspondence (COA) Analysis

Multivariate statistical analyses can be used to analyse the relationships between variables and samples. Due to the multivariate nature of codon usage, it is important to analyse the data using multivariate statistical techniques such as Correspondence Analysis (COA) (Greenacre 1984). COA is a method of multivariate analysis that provides a geometric representation of row and column sets in a dataset. COA was performed based on the RSCU values using the program CodonW. In order to examine the codon usage indices, CodonW (v1.4.4) software was used (Peden 1999).

3.1.6. Codon Adaptation Index (CAI) Analysis

Codon adaptation index (CAI) was used to analyse codon usage preferences in CCHFV in relation to codon usage of two different tick hosts and to assess whether adaptation of CCHFV in terms of preferences of codon usage does not differ to two different hosts (Sharp & Li 1987). CAI is a quantitative value that shows the frequency of use of preferred codons among highly expressed genes. The most frequent codons signify the highest relative adaptation to the host, and sequences having higher CAI are known to be favoured over sequences having lower CAI. The CAI values vary from 0 to 1, and higher values indicate higher levels of codon usage bias towards the codons used in highly expressed genes (Butt et al. 2014).

CAI analysis was performed using CAIcal (Puigbo et al. 2008a) (available at: http://genomes.urv.es/CAIcal) for CCHFV strains isolated either from *Hyalomma* or *Rhipicephalus* ticks towards the codons specific for *Hyalomma* (*H. asiaticum*) and *Rhipicephalus* (*R. sanguineus*) species, as these two species displayed the highest CAI values. This method allows us to compare a given codon usage (in our case, CCHFV strains isolated from different hosts) to a predefined reference set (species of *Hyalomma* and *Rhipicephalus* genera).

The synonymous codon usage patterns of *Hyalomma* species (*H. anatolicum*, *H. asiaticum*, *H. dromedarii*, *H. excavatum*, *H. lusitanicum*, *H. marginatum*, *H. rufipes* and

H. truncatum) and *Rhipicephalu*s species (*R. bursa* and *R. sanguineus*) were used as references. The reference data sets obtained from the the CoCoPUTs database.

CAI values are influenced by nucleotide composition of analysed sequences. To minimize this effect, we used CAIcal to calculate the expected CAI (eCAI) values (Puigbo et al. 2008b) using the average CAI of 500 randomly created sequences of the same length as the experimental sequences (S, M, and L segments of CCHFV). Further, normalized CAI (nCAI) values were calculated as: nCAI = CAI/eCAI to minimalize the effect of differences in sequence composition on CAI and to make the data comparable between different host species.

E-CAI (Puigbo et al. 2008b) is also used to calculate the expected value of CAI (eCAI) at the 95% confidence interval to decide whether the statistically significant differences in the CAI values emerge from codon preferences. A Kolmogorov-Smirnov test was later performed for the expected CAI values (Puigbo et al. 2008b).

3.1.7. Statistical analysis

For the statistical analysis of the data, GraphPad Prism version 9.1 for Windows, GraphPad Software, San Diego, California USA (www.graphpad.com) was used. Graphs were created by using GraphPad Prism software version 9.1 and Excel 2019. The normal distribution of the data was checked by using Kolmogorov-Smirnov's test, allowing the selection of parametric or non-parametric methods when comparing the sample sets. In order to test significance of the difference between two means among variables with normal distribution Student's t-test was used, and the Mann-Whitney *U* test in the negative case. For all these tests, the significance level was set at as 0.05. Student's t-test and Mann-Whitney *U* test were used to determine whether nucleotide composition, the ENC, the expected ENC, CAI, eCAI and nCAI values for strains isolated from *Hyalomma* and *Rhipicephalus* were significantly different.

3.1.8. Phylogenetic analysis of CCHFV strains

The phylogenetic tree was constructed using MEGA v10.2.4 and the maximum likelihood method with GTR+G model. The tree was designed using the online tool the Interactive Tree Of Life version 5.7 (http://itol.embl.de/) (Letunic & Bork 2011). A total of 70 strains were used in this study.

3.2. Tick collection and screening for CCHFV

3.2.1. Tick collection in Kosovo

The studies for tick collection took place in Kosovo during summer, 2020. Kosovo is located in the central part of Balkan Peninsula with a total area of 10.908 km², bordered by Albania to the southwest, the Republic of North Macedonia to the southeast, Montenegro to the west, and Serbia to the north and east. Kosovo has approximately 53% agricultural land. 60% of Kosovo is covered by mountains with elevations between 276 and 2.656 meters above sea level (a.s.l.). It is characterized by a Mid-Continental climate, with a dominant influence of Adriatic-Mediterranean climate (Report of the Kosovo Environmental Protection Agency 2015).

According to Jameson et al. (2012), the risk areas for CCHF in Kosovo occupy about 50% of municipalities with hyperendemicity occurring in the municipalities of Skënderaj, Klinë, Malishevë, Rahovec, and Suharekë. The collection of ticks were conducted during August 2020 in six villages of three municipalities, namely CCHF-hyperendemic municipalities Malishevë (four villages), Rahovec (one village), and Suharekë (one village) (Figure 9).

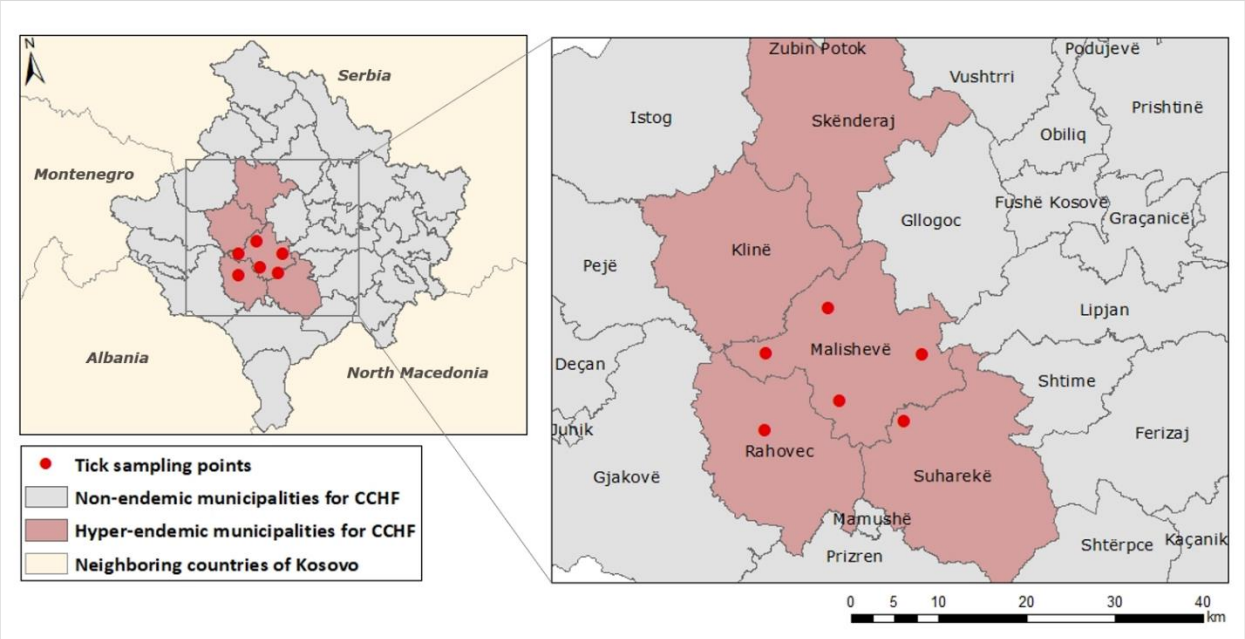


Figure 9. Area of study in Kosovo.

Malishevë, located in Prizren district, is 526 m a.s.l., with a warm and temperate climate. In Malishevë, the average annual temperature is 10.9°C and precipitation is about 609 mm. Another town that fieldwork was conducted, Rahovec, is located in Prizren district. The average temperature is 12.2°C and precipitation is 583 mm annually. The final location for collection of ticks was Suharekë which is 389 m a.s.l., a municipality located in the Prizren district of central-southern Kosovo. The average annual temperature in Suharekë is 10.7°C with a significant rainfall about 820 mm.

All the ticks from Kosovo were sampled from directly livestock (cattle) in the farms. All participants used adequate personal protective equipment during collection processes. Ticks were collected during feeding process on their host by the use of fine forceps. Collected ticks were placed into the 2 ml eppendorf tube, precisely marked, and stored at +4°C for tick identification process. All the data about daily temperature, humidity, and elevation, as well as coordinates of the station of the villages were immediately recorded after collection.

3.2.2. Tick collection in Czechia

The studies for tick collection took place in Czechia during autumn. Collection of ticks in Czechia was included in this study as introduction of larva or nymph stages of Mediterranean tick species could be transported via migratory birds during spring migration from the south to the north. Fieldwork activities were conducted during September-October 2020 with collaboration of Bird-Ringing Station of Natural History Museum of Czech Republic near Choteč (Figure 10), a village in Prague-West District in the Central Bohemian Region of Czechia.

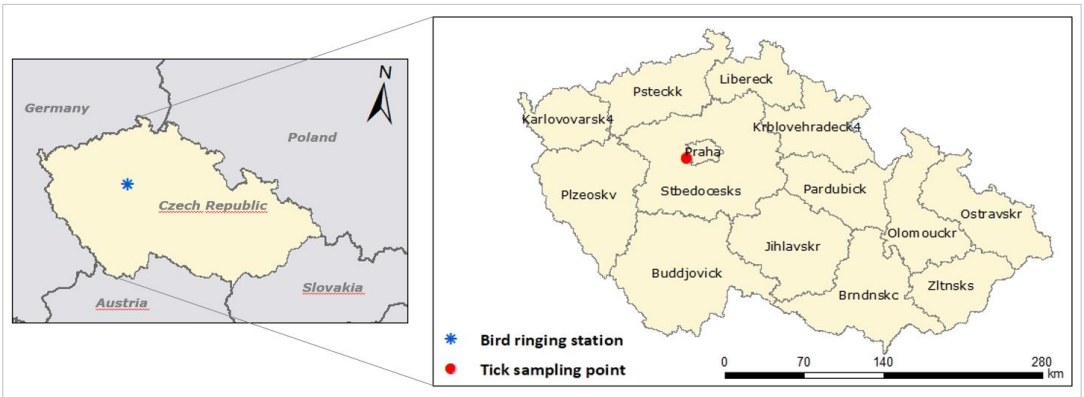


Figure 10. Area of study in Czechia.

All the ticks from Czechia were sampled from directly migratory birds. The collection of ticks was performed during the post-breeding period of birds, which coincides with the beginning of autumn migration from the north to the south. The tick collection from birds in the autumn season was a preliminary experiment to learn tick collection on birds.

Birds were captured by professional ornithologists during regular ringing activities. Each bird was identified by species and was ringed. During bird ringing activities, ticks were collected by opportunistic sampling on birds. The data about the species of birds, their age, and sex were recorded in the data sheet. All collected ticks were placed into the 2 ml eppendorf tubes and stored at +4°C for tick identification.

3.2.3. Tick identification

Before performing the molecular analyses, all collected ticks were analysed morphologically. The species, the development stage, and sex of the ticks were determined according to keys (Estrada-Peña et al. 2004) by using stereomicroscope. Before RNA extraction, the ticks were stored individually at -80°C in eppendorf tubes.

3.2.4. RNA extraction

Ticks placed in one tube from each bird or livestock were kept at -80°C until RNA extraction. Ticks were individually homogenized under sterile conditions by using plastic sticks, which were sterilized with HCl and then washed properly before the use to remove any possible contaminants. The viral RNA extraction was performed using the QiaAmp Viral RNA Minikit, Qiagen, Hilden, Germany according to the manufacturer's instructions. For more information, manufacturing protocol was added in Appendix 1. RNA extracted from all samples was eluted in 60 µl by using Buffer AVE (Elution Buffer). Resulting RNA products were marked properly, stored in the freezer at -80°C, and later used as templates for PCR amplification.

3.2.5. Quantitative Reverse Transcription Polymerase Chain Reaction (RT-qPCR) for detection of CCHFV

The initial RT-qPCR was designed in the volume of 20 µl reaction mixture containing 5 µl 4×Reliance One-Step Multiplex Supermix (Bio-Rad), 2 µl of the primer

and probe mix (10 μM each primer, 5 μM probe), 11 μl of the nuclease-free water and 2 μl of the extracted RNA.

Two primer pairs and probe that were designed for detection of CCHFV by Kurtesh et al. (2014) were used for this study (Table 2). The probe was designed with a 5'6 carboxifluorescein (FAM) fluorophore and a 3' Black Hole Quencher 1 (BHQ1). For primer mix, 10 μl of each 100 μM primer (forward and reverse), 5 μl of 100 μM probe and 75 μl nuclease-free water were mixed together in a final volume of 100 μl.

Table 2. Sequences of primer pairs and probe used in this study.

Primer/Probe	Sequence 5' to 3'	Bases	Type of Probe
Forward Primer: CCHF S1-F	AATAAATCATAATCTCAAAGAAACACGTGCC	31	
Reverse Primer: CCHF S122-P	AATAAATCATAACCTTTTTGAACTCTTCAAACC	33	
Probe: CCHF SP	ACTCAAGGKAACACTGTGGGCGTAAG	26	FAM-BHQ1

PCRs were run and analysed on a CFX connect real-time PCR detection system (Bio-Rad, Hercules, CA, USA). The cycling parameters were as follows: 10 min at 50°C, 10 min at 95°C, 50× 10 sec at 95°C and 30 sec at 60°C. Samples with c_q values lower than 40 were considered as positive.

3.3. Mapping Environmental Suitability

3.3.1. Occurrence records

H. marginatum occurrence records in Europe were assembled from the previously peer-reviewed literature (Estrada-Peña et al. 2016). Since the MaxEnt software was used for modeling later, the data was prepared in accordance MaxEnt requires that the occurrence data is in a comma-separated value (CSV) text file containing three fields, that identify the "species", "X-coordinates (longitude)" and "Y-coordinates (latitude)". Since the occurrence data were few only for Central Europe, occurrence records across Europe were used in order to make a better prediction for Central Europe, particularly for Czechia which is the main focus of this part of the thesis.

A total of 658 occurrences were included in the original dataset. The final dataset included 288 occurrence points, after removing duplicated occurrences and redundant occurrence records. Redundant occurrence records occurring in a distance ≤2.5′ (≈ 5 km) were excluded to remove spatial bias in estimating ENMs by using SDMtoolbox 2.4 (Brown et al. 2017) in ArcGIS 10.7.1 (Environmental Systems Research Institute (ESRI), Redlands, CA). The occurrence data set was divided into two equal portions: 50% for model calibration, and 50% for evaluation of model predictions using Hawth's Tools (Beyer 2004) available in ArcGIS 10.7.1.

3.3.2. Covariate variables

Data summarizing current global climates from the WorldClim version 2 archive (www.worldclim.org) were used. WorldClim data includes 19 bioclimatic variables originally derived from monthly temperature and rainfall values collected from weather stations in 1970-2000 (Fick & Hijmans 2017). 2.5 min spatial resolution (≈5 km) was selected for the data. Bioclimatic variables 8-9 and 18-19 were excluded from the analysis to avoid problems deriving from odd spatial artifacts (Samy et al. 2016). Satellite data summarizing the Normalized Difference Vegetation Index (NDVI), land surface temperature (LST), leaf area index (LAI), and precipitation were also included. NDVI data was included due to its significant role in shaping the ecological niches of tick vectors (Estrada-Pena et al. 2016). NDVI is also considered an important factor in reflecting soil moisture's availability for larvae and nymphs. (Guerra et al. 2002; Randolph 2000).

NDVI, LST, and LAI data were downloaded from the Copernicus Global Land Service (CGLS) archive available via the link https://land.copernicus.vgt.vito.be/ for 2012-2018 periods in 1 km spatial resolution. Precipitation data were downloaded from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data archive (https://data.chc.ucsb.edu/products/CHIRPS-2.0/) for the same time period. All bioclimatic and satellite raster data were resampled to the spatial resolution of 5 km in the context of climatic variables and extracted to the interested study area by using SDMtoolbox 2.4 (Brown et al. 2017). We calculated a composite of these data representing the maximum, minimum, range, median, and mean of each of these data.

All analyses of rasters were performed under the WGS 1984 geographic coordinate system in ArcGIS 10.7.1 (Environmental Systems Research Institute (ESRI), Redlands, CA). All environmental variables (15 bioclim and 20 satellite data) were subjected to principal component analysis (PCA) in ArcGIS 10.7.1 to reduce the dimensionality and multicollinearity between these variables (Samy et al. 2016). The first 6 principal components (PCs) were used to estimate the ENM as they summarized about 99% of the overall variance in the environmental data. Rasters of each of the environmental variables (bioclimatic and satellite) were converted to ASCII grid format with all rasters sharing the same exact cell size and the same exact geographic extent as required by MaxEnt.

Variables were selected considering their contribution to models and their collinearity. Final sets of variables are as follows: Set 1 (15 bioclimatic variables from WorldClim); Set 2 (20 satellite data from CGLS and CHIRPS); Set 3 (the first 6 PCs of bioclimatic variables); and Set 4 (the first 6 PCs of satellite data) (Table 3).

Table 3. List of environmental variables used for creation of models. Environmental variables were prepared based on kuenm R package instructions. Set 1 includes all bioclimatic variables, while Set 2 has only variables from satellite data. After performing PCA analysis to see collinearity between variables, only the first 6 PCs were selected. The first 6 PCs of bioclimatic variables and satellite data were included in Set 3 and Set 4, respectively.

Variables from WorldClim	Description of Variables	Set 1	Set 3
Bio1	Annual mean temperature	x	x
Bio2	Mean diurnal range	x	x
Bio3	Isothermality	x	x
Bio4	Temperature Seasonality	x	x
Bio5	Maximum temperature of warmest month	x	x
Bio6	Minimum temperature of coldest month	x	x
Bio7	Temperature annual range	x	
Bio10	Mean temperature of the warmest quarter	x	
Bio11	Mean temperature of the coldest quarter	x	
Bio12	Annual precipitation	x	
Bio13	Precipitation of the wettest month	x	
Bio14	Precipitation of the driest month	x	
Bio15	Precipitation seasonality	x	
Bio16	Precipitation of the wettest quarter	x	
Bio17	Precipitation of driest quarter	x	

Variables from Satellite Data	Description of Variables	Set 2	Set 4
max LST	Land surface temperature (maximum)	x	x
min LST	Land surface temperature (minimum)	x	x
med LST	Land surface temperature (median)	x	x
mean LST	Land surface temperature (mean)	x	x
ran LST	Land surface temperature (range)	x	x
max LAI	Leaf area index (maximum)	x	x
min LAI	Leaf area index (minimum)	x	
med LAI	Leaf area index (median)	x	
mean LAI	Leaf area index (mean)	x	
ran LAI	Leaf area index (range)	x	
max NDVI	Normalized Difference Vegetation Index (maximum)	x	
min NDVI	Normalized Difference Vegetation Index (minimum)	x	
med NDVI	Normalized Difference Vegetation Index (median)	x	
mean NDVI	Normalized Difference Vegetation Index (mean)	x	
ran NDVI	Normalized Difference Vegetation Index (range)	x	
max PPT	Precipitation (maximum)	x	
min PPT	Precipitation (minimum)	x	
med PPT	Precipitation (median)	x	
mean PPT	Precipitation (mean)	x	
ran PPT	Precipitation (range)	x	

3.3.3. Ecological niche modeling

Grinnellian ecological niche model was used to map the environmental suitability of *H. marginatum* with the purpose of predicting its possible distribution in Czechia and other countries of Europe to better understand the risk of *Hyalomma*-borne diseases. This modeling method was estimated for *H. marginatum* based on the maximum entropy algorithm implemented in MaxEnt v3.4.1 (Phillips et al. 2006) via the kuenm R package (Cobos et al. 2019). The MaxEnt program developed by Phillips, Dudík, and Schapire (2004) is a machine learning process that uses multiple iterations to train the model into creating an acceptable model. MaxEnt requires a point dataset that contains species occurrence data and environmental raster datasets. There are different computational features that are used to constrain the MaxEnt output based on the environmental variable values. Therefore, before the model was run in MaxEnt, it was crucial to evaluate the model and decide the parameters which should be used in MaxEnt to predict the best model.

In order to calibrate, evaluate, and select the best models, kuenm R package was used in R Studio v1.4.1106. Kuenm is an R package designed to make the process of model calibration and final model creation easier and more robust (Cobos et al. 2019). The package requires MaxEnt but runs all the functions outside MaxEnt.

Before running kuenm R package to evaluate and select the models, all data were prepared according to kuenm requirements (Cobos et al. 2019). Three different CSV files were prepared which included all occurrence records of *H. marginatum*, occurrence records after calibration process and occurrence records after evaluation which were obtained by using Hawth's Tools.

The accessible area "M" is an important element in the biotic, abiotic, and movement (BAM) diagram and defines the key parameters in constructing ecological niche model for the species (Barve et al. 2011). Accessible area "M" indicates the areas that the species explored and had access to in over relevant periods of species' history (Barve et al. 2011). Therefore, a polygon was created around occurrence points to extend the limits of the entire calibration region. After creating accessible area "M", all four sets of environmental variables were extracted to this area and used in kuenm R package for creating candidate models and evaluating and selecting best models.

Candidate models were created by combining four distinct sets of environmental variables, 17 values of regularization multipliers (0.1-1 with intervals of 0.1, 2-6 with intervals of 1, and 8 and 10), and all 31 possible combinations of 5 feature classes (linear = l, quadratic = q, product = p, threshold = t, and hinge = h) (Cobos et al. 2019).

The candidate models and best models were selected according to the following criteria proposed by Cobos et al. (2019): 1) significance, 2) performance, and 3) the Akaike information criteria (AIC): AICCc, Delta AICCc, and Weight AICCc. Final models were created by using bootstrap replicate type with 10 replicates and logistic outputs. The final models were then transferred from the accessible area "M" to the projection area "G". Projection area "G" is another important component of kuenm R package that projects and transfers the ENMs into other areas. The best predicted final model of suitable areas for *H. marginatum* was created based on the model evaluation in kuenm by using bootstrap replicate type with 10 replicates and selecting linear, quadratic, and product features outputs in MaxEnt. Final model output was classified into three classes representing lower, medium, and higher predictions in ArcGIS 10.7.1.

4. Results

4.1. Codon usage bias of CCHFV to two tick hosts

4.1.1. Nucleotide composition analysis of CCHFV in relation to two tick hosts

Codon usage patterns are considered to be strongly influenced by the nucleotide composition (Jenkins & Holmes, 2003; Wong et al. 2010; Nasrullah et al. 2015). Therefore, we first analysed the nucleotide composition of *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains belong to three segments separately in order to evaluate the potential influence of the nucleotide constraints on codon usage patterns and determine whether CCHFV strains isolated from two hosts differ from each other.

At the beginning of data collection, nearly 3000 data were collected. All data were subjected to a selection process and only sequences isolated from species of *Hyalomma* and *Rhipicephalus* genera were selected. In total, 70 CCHFV coding sequences, namely 24 complete sequences of S segment, 21 complete and 2 partial sequences of M segment, and 17 complete and 6 partial sequences of L segment were obtained and analyzed in relation to this aim (Appendix 2). The mean contents of almost all nucleotides were significantly different between *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains in M and L segments but not in S segment (Tables 4A, 4B, 4C; Figure 11).

The nucleotide contents of the third position of codons (A3, U3, G3, C3) and of GC1, GC1,2, GC3 and AU3 also play an important role on influencing overall codon usage preferences. The results showed that there is a statistically significant difference between the frequency of A3 and G3 in all segments, while C3 is significantly different between S and L segments of *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains. Similarly, GC content differs between *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains on almost all codon positions including GC3 (except in M segment).

Taken together, nucleotide composition analysis showed that there are substantial differences on frequencies of occurrence of nucleotides between *Hyalomma*- and *Rhipicephalus*-isolated CCHFV variants.

Table 4 (A, B, C). Nucleotide composition analysis of S, M, and L segments of CCHFV isolated from *Hyalomma* and *Rhipicephalus* species (%). Significant values on the 95% confidence limit between strains isolated from two tick hosts are marked in bold. ($p<0.05$). ENC represents the effective number of codons. GC12 represents the GC content at the first and second positions of codons. GC3 represents the GC content at the third positions of codons. AU3 represents the AU content at the third positions of codons.

Table 4A. Nucleotide composition analysis of S genome segment sequences of CCHFV isolated from *Hyalomma* and *Rhipicephalus* species (%).

SEQUENCES	HOST	COUNTRY		A	C	U	G	A3	U3	G3	C3	AU	GC	GC1	GC2	AU3	GC3	GC12
DQ076415.1	*Hyalomma* spp.	Uganda	*S segment*	30.23	22.29	22.36	25.12	20.50	25.26	26.29	27.95	52.59	47.41	50.10	37.89	45.76	54.24	44.00
KF793333.1	*Hyalomma* spp.	Mali		30.71	22.77	22.22	24.29	22.15	25.05	23.60	29.19	52.93	47.07	50.31	38.10	47.20	52.80	44.20
KU707899.1	*H.anatolicum*	Iran		30.99	21.60	23.40	24.02	22.57	27.54	23.40	26.50	54.38	45.62	49.28	37.68	50.10	49.90	43.48
KY484036.1	*H.excavatum*	Nigeria		30.50	22.43	22.57	24.50	21.33	25.88	24.43	28.36	53.07	46.93	49.90	38.10	47.20	52.80	44.00
KY484037.1	*H.anatolicum*	Pakistan		30.23	22.22	22.98	24.57	20.70	25.88	25.05	28.36	53.21	46.79	49.28	37.68	46.58	53.42	43.48
AF481799.1	*H.asiaticum*	Uzbekistan		30.78	22.57	22.02	24.64	22.36	23.40	24.84	29.40	52.80	47.20	49.69	37.68	45.76	54.24	43.69
MG659724.1	*H.asiaticum*	China		30.16	22.15	22.98	24.71	20.50	26.09	25.26	28.16	53.14	46.86	49.28	37.89	46.58	53.42	43.58
MG659727.1	*H.asiaticum*	China		30.02	22.50	22.64	24.84	20.08	25.05	25.67	29.19	52.66	47.34	49.28	37.89	45.13	54.87	43.58
MH688497.1	*H.asiaticum*	China		30.64	22.15	22.64	24.57	22.15	25.26	24.64	27.95	53.28	46.72	49.48	38.10	47.41	52.59	43.79
KU707898.1	*H.dromedarii*	Iran		30.43	22.22	22.91	24.43	21.33	25.88	24.43	28.36	53.35	46.65	49.28	37.89	47.20	52.80	43.58
KU707900.1	*H.dromedarii*	Iran		30.57	22.22	22.98	24.22	21.95	25.88	23.81	28.36	53.55	46.45	49.28	37.89	47.83	52.17	43.58
MF547415.1	*H. lusitanicum*	Spain		30.57	22.77	22.08	24.57	21.74	24.64	24.43	29.19	52.66	47.34	50.31	38.10	46.38	53.62	44.20
AY277672.1	*H.marginatum*	Russia		30.37	22.71	22.36	24.57	20.91	24.84	25.26	28.99	52.73	47.27	49.90	37.68	45.76	54.24	43.79
KR814833.1	*H.marginatum*	Russia		30.09	23.81	21.12	24.98	20.50	21.12	26.50	31.88	51.21	48.79	50.31	37.68	41.61	58.39	44.00
KR814834.1	*H.marginatum*	Russia		30.09	23.81	21.12	24.98	20.50	21.12	26.50	31.88	51.21	48.79	50.31	37.68	41.61	58.39	44.00
KY484031.1	*H.marginatum*	China		31.19	22.71	21.74	24.36	23.60	22.77	24.02	29.61	52.93	47.07	50.10	37.47	46.38	53.62	43.79
KY484044.1	*H.marginatum*	Uganda		30.23	22.29	22.29	25.19	20.50	25.26	26.29	27.95	52.52	47.48	50.10	38.10	45.76	54.24	44.10
DQ211641.1	*H.rufipes*	Mauritania		30.71	22.64	22.36	24.29	21.95	25.47	24.02	28.57	53.07	46.93	50.31	37.89	47.41	52.59	44.10
MF511219.1	*H.rufipes*	S. Africa		30.23	22.50	22.02	25.26	20.50	24.43	26.50	28.57	52.24	47.76	50.10	38.10	44.93	55.07	44.10
DQ211639.1	*H.truncatum*	Senegal		31.06	23.05	21.81	24.09	22.36	24.02	23.60	30.02	52.86	47.14	50.31	37.47	46.38	53.62	43.89
KY484027.1	*H.truncatum*	Senegal		31.06	23.05	21.81	24.09	22.36	24.02	23.60	30.02	52.86	47.14	50.31	37.47	46.38	53.62	43.89
	Mean ± STD			30.52 ± 0.35	22.59 ± 0.51	22.30 ± 0.58	24.59 ± 0.36	**21.45 ± 0.93**	24.71 ± 1.53	**24.86 ± 1.05**	**28.98 ± 1.23**	**52.82 ± 0.68**	**47.18 ± 0.68**	**49.87 ± 0.43**	37.83 ± 0.22	**46.16 ± 1.82**	**53.84 ± 1.82**	**43.85 ± 0.23**
DQ211638.1	*R.bursa*	Greece		30.16	22.98	21.88	24.98	18.63	23.60	27.74	30.02	52.04	47.96	48.24	37.89	42.24	57.76	43.06
MG516211.1	*R.bursa*	Greece		29.95	23.26	21.95	24.84	19.25	22.98	26.50	31.26	51.90	48.10	48.24	38.30	42.24	57.76	43.27
U04958.1	*R.bursa*	Greece		30.23	22.98	21.81	24.98	18.63	23.60	27.74	30.02	52.04	47.96	48.24	37.89	42.24	57.76	43.06
	Mean ± STD			30.11 ± 0.12	23.07 ± 0.13	21.88 ± 0.06	24.94 ± 0.07	**18.84 ± 0.29**	23.40 ± 0.29	**27.33 ± 0.58**	**30.43 ± 0.58**	**51.99 ± 0.07**	**48.01 ± 0.07**	**48.24 ± 0.00**	38.03 ± 0.19	**42.24 ± 0.00**	**57.76 ± 0.00**	**43.13 ± 0.10**

Table 4B. Nucleotide composition analysis of M genome segment sequences of CCHFV isolated from *Hyalomma* and *Rhipicephalus* species (%).

SEQUENCES	HOST	COUNTRY		A	C	U	G	A3	U3	G3	C3	AU	GC	GC1	GC2	AU3	GC3	GC12
DQ157174.1	*Hyalomma* spp.	S. Africa		31.39	21.70	24.97	21.94	30.98	26.41	17.92	24.69	56.36	43.64	44.51	43.8	57.39	42.61	44.16
KY484038.1	*H.anatolicum*	Pakistan		32.06	22.44	24.15	21.36	31.92	24.48	18.23	25.37	56.20	43.8	43.54	44.25	56.4	43.6	43.90
MG659726.1	*H.asiaticum*	China		30.83	22.29	24.62	22.27	28.88	26.09	19.23	25.80	55.44	44.56	44.5	44.14	54.97	45.03	44.32
MG659723.1	*H.asiaticum*	China		30.85	22.27	24.67	22.21	28.93	25.98	19.11	25.98	55.52	44.48	44.5	43.85	54.91	45.09	44.18
MH688498.1	*H.asiaticum*	China		31.31	22.01	24.5	22.18	30.33	25.85	19.26	24.56	55.81	44.19	44.99	43.76	56.18	43.82	44.38
NC_005300.2	*H.excavatum*	Nigeria		31.55	22.02	24.65	21.78	31.39	25.82	17.63	25.16	56.20	43.8	44.57	44.04	57.21	42.79	44.31
KY484035.1	*H.excavatum*	Nigeria		31.53	22.06	24.67	21.74	31.34	25.82	17.63	25.22	56.20	43.8	44.51	44.04	57.15	42.85	44.28
AF467768.2	*H.excavatum*	Nigeria		31.55	22.02	24.65	21.78	31.39	25.82	17.63	25.16	56.20	43.8	44.57	44.04	57.21	42.79	44.31
MF547416.1	*H. lusitanicum*	Spain		31.41	22.08	24.63	21.88	30.98	26.29	17.92	24.81	56.04	43.96	44.99	44.15	57.27	42.73	44.57
KY484045.1	*H.marginatum*	Uganda		31.39	21.70	24.97	21.94	30.98	26.41	17.92	24.69	56.36	43.64	44.51	43.80	57.39	42.61	44.16
KY484032.1	*H.marginatum*	China		31.28	21.50	25.38	21.83	30.47	26.09	18.58	24.85	56.67	43.33	43.31	43.25	56.57	43.43	43.28
AY900145.1	*H.marginatum*	China		31.28	21.62	25.29	21.81	30.53	25.98	18.52	24.97	56.57	43.43	43.37	43.43	56.51	43.49	43.4
AY900141.1	*H.marginatum*	S. Africa		31.37	21.68	24.97	21.98	31.04	26.41	17.92	24.63	56.34	43.66	44.51	43.92	57.45	42.55	44.22
AY179961.1	*H.marginatum*	Russia		31.34	22.68	24.02	21.97	30.20	24.51	19.60	25.7	55.36	44.64	44.17	44.46	54.71	45.29	44.32
DQ211628.1	*H.rufipes*	Mauritania		31.15	23.11	23.84	21.89	28.95	24.71	19.34	27.0	54.99	45.01	44.22	44.46	53.66	46.34	44.34
MF511236.1	*H.rufipes*	S. Africa		31.61	21.66	25.10	21.62	31.16	26.82	17.45	24.57	56.72	43.28	44.33	43.50	57.98	42.02	43.92
KF793334.1	*Hyalomma* spp.	Mali		30.68	19.65	26.93	22.74	27.15	28.81	18.87	25.17	57.62	42.38	43.38	39.74	55.96	44.04	41.56
KY484026.1	*H.truncatum*	Senegal		31.92	21.58	24.77	21.73	31.04	25.81	18.87	24.28	56.69	43.31	43.21	43.56	56.85	43.15	43.39
DQ211626.1	*H.truncatum*	Senegal		31.92	21.58	24.77	21.73	31.04	25.81	18.87	24.28	56.69	43.31	43.21	43.56	56.85	43.15	43.39
	Mean ± STD		***M segment***	**31.39 ± 0.35**	**21.88 ± 0.67**	**24.82 ± 0.63**	**21.91 ± 0.28**	**30.46 ± 1.15**	26.00 ± 0.90	**18.45 ± 0.67**	25.10 ± 0.65	56.21 ± 0.59	43.79 ± 0.59	44.15 ± 0.59	43.67 ± 0.98	**56.45 ± 1.11**	**43.55 ± 1.11**	**43.91 ± 0.67**
MG516212.1	*R.bursa*	Greece		30.37	23.45	23.60	22.58	27.42	23.76	20.75	28.07	53.97	46.03	44.81	44.46	51.18	48.82	44.64
DQ211625.1	*R.bursa*	Greece		30.29	23.17	23.82	22.72	27.18	24.17	21.34	27.30	54.11	45.89	45.22	43.81	51.36	48.64	44.52
EF189752.1	*R.bursa*	Turkey		31.40	14.73	30.19	23.67	25.36	30.43	21.74	22.46	61.59	38.41	36.23	34.78	55.80	44.20	35.51
EF189751.1	*R.bursa*	Turkey		31.16	14.73	30.43	23.67	25.36	30.43	21.74	22.46	61.59	38.41	36.23	34.78	55.80	44.20	35.51
	Mean ± STD			**30.81 ± 0.48**	**19.02 ± 4.29**	**27.01 ± 3.30**	**23.16 ± 0.51**	**26.33 ± 0.97**	27.20 ± 3.24	**21.39 ± 0.41**	25.07 ± 2.63	57.82 ± 3.78	42.19 ± 3.78	40.62 ± 4.39	39.46 ± 4.68	**53.54 ± 2.27**	**46.47 ± 2.27**	**40.04 ± 4.54**

Table 4C. Nucleotide composition analysis of L genome segment sequences of CCHFV isolated from *Hyalomma* and *Rhipicephalus* species (%).

SEQUENCES	HOST	COUNTRY		A	C	U	G	A3	U3	G3	C3	AU	GC	GC1	GC2	AU3	GC3	GC12
KY484039.1	*H.anatolicum*	Pakistan		32.65	19.04	26.38	21.93	29.24	28.48	21.41	20.86	59.03	40.97	45.21	35.43	57.73	42.27	40.32
MG659725.1	*H.asiaticum*	China		32.56	19.18	26.27	21.99	29.02	28.10	21.62	21.26	58.84	41.16	45.24	35.38	57.12	42.88	40.31
MG659722.1	*H.asiaticum*	China		32.50	19.14	26.34	22.02	28.92	28.31	21.64	21.14	58.84	41.16	45.34	35.38	57.22	42.78	40.36
MH688499.1	*H.asiaticum*	China		32.51	19.56	25.9	22.02	28.56	27.47	22.12	21.84	58.41	41.59	45.64	35.15	56.03	43.97	40.40
KY484034.1	*H.excavatum*	Nigeria		32.77	19.23	26.09	21.90	29.73	27.39	21.36	21.52	58.86	41.14	44.86	35.68	57.12	42.88	40.27
AY389508.2	*H.excavatum*	Nigeria		32.76	19.24	26.09	21.91	29.73	27.37	21.36	21.54	58.84	41.16	44.88	35.68	57.10	42.90	40.28
NC_005301.3	*H.excavatum*	Nigeria		32.73	19.27	26.09	21.92	29.70	27.39	21.36	21.54	58.81	41.19	44.93	35.73	57.10	42.90	40.33
AY422209.2	*H.excavatum*	Nigeria		32.77	19.25	26.08	21.90	29.73	27.37	21.36	21.54	58.84	41.16	44.86	35.71	57.10	42.90	40.29
AY389361.2	*H.excavatum*	Nigeria		32.73	19.27	26.09	21.92	29.70	27.39	21.36	21.54	58.81	41.19	44.93	35.73	57.10	42.90	40.33
AY947891.1	*H.excavatum*	Nigeria		32.77	19.23	26.09	21.9	29.73	27.39	21.36	21.52	58.86	41.14	44.86	35.68	57.12	42.88	40.27
MF547417.1	*H. lusitanicum*	Spain		32.94	19.29	26.04	21.74	30.06	27.22	21.03	21.69	58.98	41.02	44.70	35.63	57.27	42.73	40.17
KY484043.1	*H.marginatum*	Uganda		32.88	19.33	25.92	21.88	29.80	27.27	21.44	21.49	58.79	41.21	45.03	35.66	57.07	42.93	40.35
KY484033.1	*H.marginatum*	China		32.74	18.99	26.42	21.84	29.55	28.64	21.21	20.60	59.17	40.83	45.18	35.5	58.19	41.81	40.34
DQ211615.1	*H.rufipes*	Mauritania		32.77	19.23	26.12	21.88	29.65	27.55	21.36	21.44	58.89	41.11	44.98	35.55	57.20	42.80	40.27
MF511202.1	*H.rufipes*	S. Africa		32.95	19.29	25.94	21.81	30.06	27.32	21.24	21.39	58.90	41.10	45.01	35.68	57.37	42.63	40.35
KY484025.1	*H.truncatum*	Senegal		32.85	19.24	26.07	21.84	29.83	27.34	21.24	21.59	58.92	41.08	44.80	35.61	57.17	42.83	40.21
DQ211613.1	*H.truncatum*	Senegal		32.85	19.24	26.07	21.84	29.83	27.34	21.24	21.59	58.92	41.08	44.8	35.61	57.17	42.83	40.21
	Mean ± STD			**32.75 ± 0.13**	**19.24 ± 0.12**	**26.12 ± 0.15**	**21.90 ± 0.07**	**29.58 ± 0.40**	27.61 ± 0.45	**21.39 ± 0.23**	**21.42 ± 0.30**	**58.87 ± 0.15**	**41.13 ± 0.15**	**45.01 ± 0.23**	**35.58 ± 0.15**	**57.19 ± 0.40**	**42.81 ± 0.40**	**40.30 ± 0.06**
MG516213.1	*R.bursa*	Greece		32.25	20.11	24.96	22.67	27.95	25.49	23.47	23.09	57.21	42.79	46.05	35.76	53.45	46.55	40.91
DQ211612.1	*R.bursa*	Greece		32.42	20.16	24.88	22.54	28.51	25.49	22.99	23.01	57.30	42.70	46.30	35.81	54.0	46.0	41.06
KY963542.1	*R.bursa*	Turkey		30.72	20.04	24.4	24.84	22.88	28.1	22.88	26.14	55.12	44.88	46.41	39.22	50.98	49.02	42.82
KY963541.1	*R.bursa*	Turkey		30.28	20.04	24.4	25.27	21.57	28.76	24.18	25.49	54.68	45.32	47.06	39.22	50.33	49.67	43.14
KY963540.1	*R.bursa*	Turkey		30.20	20.36	23.71	25.73	22.15	27.52	24.83	25.50	53.91	46.09	47.65	40.27	49.66	50.34	43.96
KY963543.1	*R. sanguineus*	Turkey		30.95	19.26	24.89	24.89	23.38	31.17	22.73	22.73	55.84	44.16	47.40	39.61	54.55	45.45	43.51
	Mean ± STD			**31.14 ± 0.89**	**20.00 ± 0.35**	**24.54 ± 0.44**	**24.32 ± 1.25**	**24.41 ± 2.77**	27.76 ± 1.96	**23.51 ± 0.76**	**24.33 ± 1.40**	**55.68 ± 1.25**	**44.32 ± 1.25**	**46.81 ± 0.59**	**38.32 ± 1.82**	**52.16 ± 1.90**	**47.83 ± 1.90**	**42.56 ± 1.17**

L segment

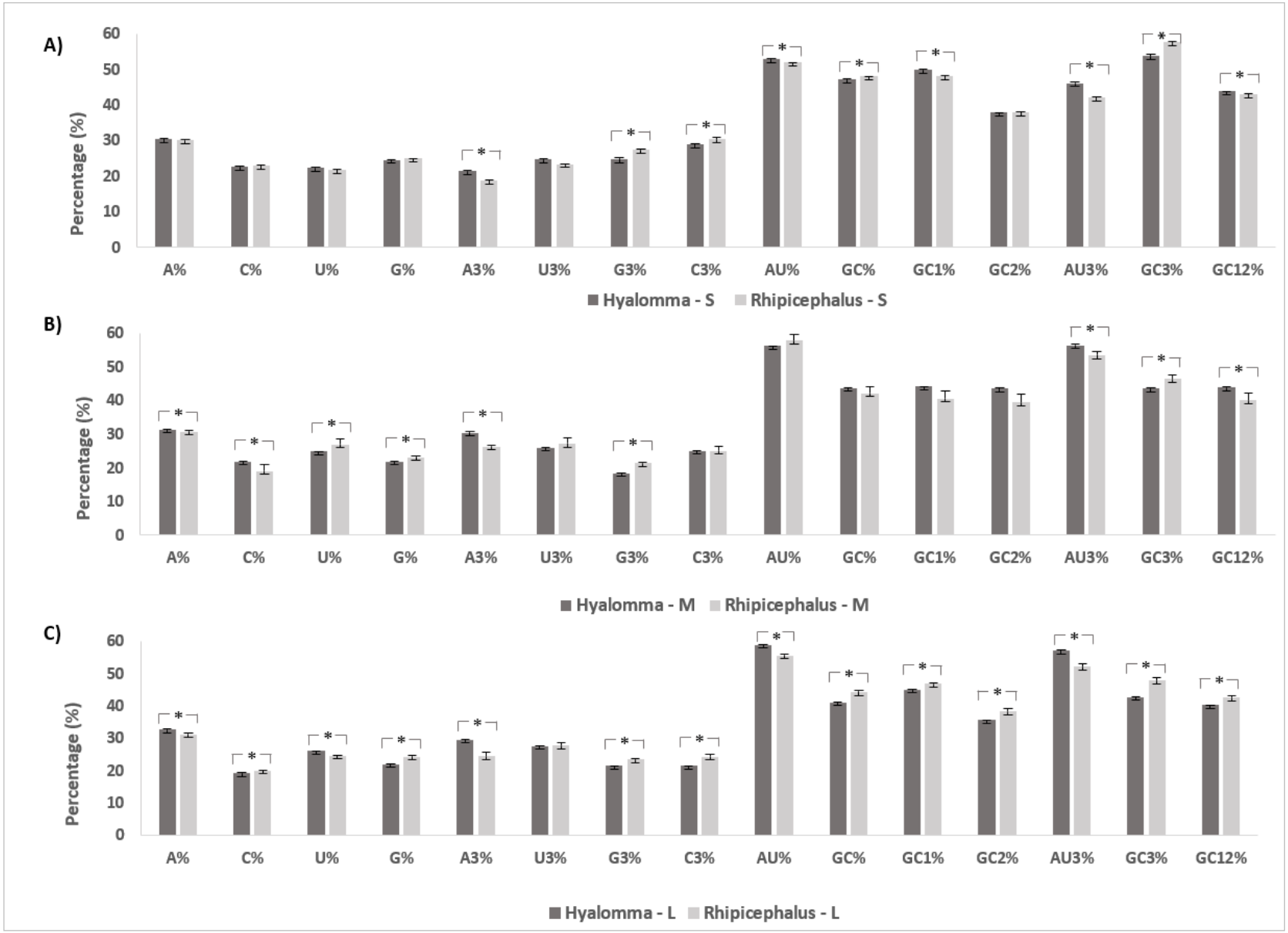


Figure 11. Nucleotide content distribution and composition in *Hyalomma*- and *Rhipicephalus*-isolated S, M, and L segments (11A, 11B, 11C), respectively. Standard deviation was marked in the plot. Asterisk (*) show a significant difference between variables ($p<0.05$).

4.1.2. General codon usage pattern among CCHFV strains isolated from *Hyalomma* and *Rhipicephalus*

To analyse if the codon usage bias differs between CCHFV isolated from *Hyalomma* and *Rhipicephalus* tick vectors ENC's values were calculated.

The ENC-GC3s plot was performed to determine whether the codon usage of given strains is solely due to mutational pressure or selectional pressure. Genetic evolution is influenced by mutational pressure where there is no natural selection, and codon usage bias can be affected only by the nucleotide composition of the genome. Each data point, therefore, falls onto the expected curve or near the expected curve. Contrarily, if the points fall below the expected curve, the codon usage is subject to natural selection. As Figure 12 shows, all the points lie below the expected curve, which suggests that in addition to the mutation pressure, translation selection also influences the codon usage bias of CCHFV.

Calculated ENC values were as follows: ENC values of S segment isolated from *Hyalomma* ranged from 53.14 to 55.27 (mean = 53.33 ± 1.33) and isolated from *Rhipicephalus* ranged from 52.28 to 53.20 (mean = 52.85 ± 0.41). For M segment, ENC values of *Hyalomma*-isolated strains ranged from 49.96 to 51.84 (mean = 50.89 ± 0.45) and ENC values of *Rhipicephalus*-isolated strains ranged from 46.93 to 51.73 (mean = 49.50 ± 1.91). ENC values of L segment isolated from *Hyalomma* ranged from 51.13 to 52.65 (mean = 51.92 ± 0.32) and isolated from *Rhipicephalus* ranged from 45.27 to 51.57 (mean = 47.64 ± 2.74). ENC values showed significant differences only in L segment isolated from two hosts ($p < 0.05$).

Hyalomma- and *Rhipicephalus*-isolated M and L segments share similar patterns in relation to GC3 values, with the mean GC3 values of *Hyalomma*-isolated M and L segments are 0.418 and 0.411, respectively. GC3 values for *Rhipicephalus*-isolated M and L segments are 0.448 and 0.458, respectively. However, GC3 values of *Hyalomma*- and *Rhipicephalus*-isolated S segment are considerably higher compared with the others 0.517 and 0.558, respectively.

In the ENC-GC3 plot, the CCHFV strains of all segments isolated from both *Hyalomma* and *Rhipicephalus* clustered together below the expected ENC curve, indicating that GC compositional constraints might influence the codon usage of CCHFV in relation to *Hyalomma* and *Rhipicephalus* hosts (Figure 12B, 12C, 12D). To

clarify the relationship between the GC3 and ENC values, the expected ENC values for different GC3 were calculated. All CCHFV segments isolated from two hosts showed significant differences in expected ENC values ($p < 0.05$) (Figure 13).

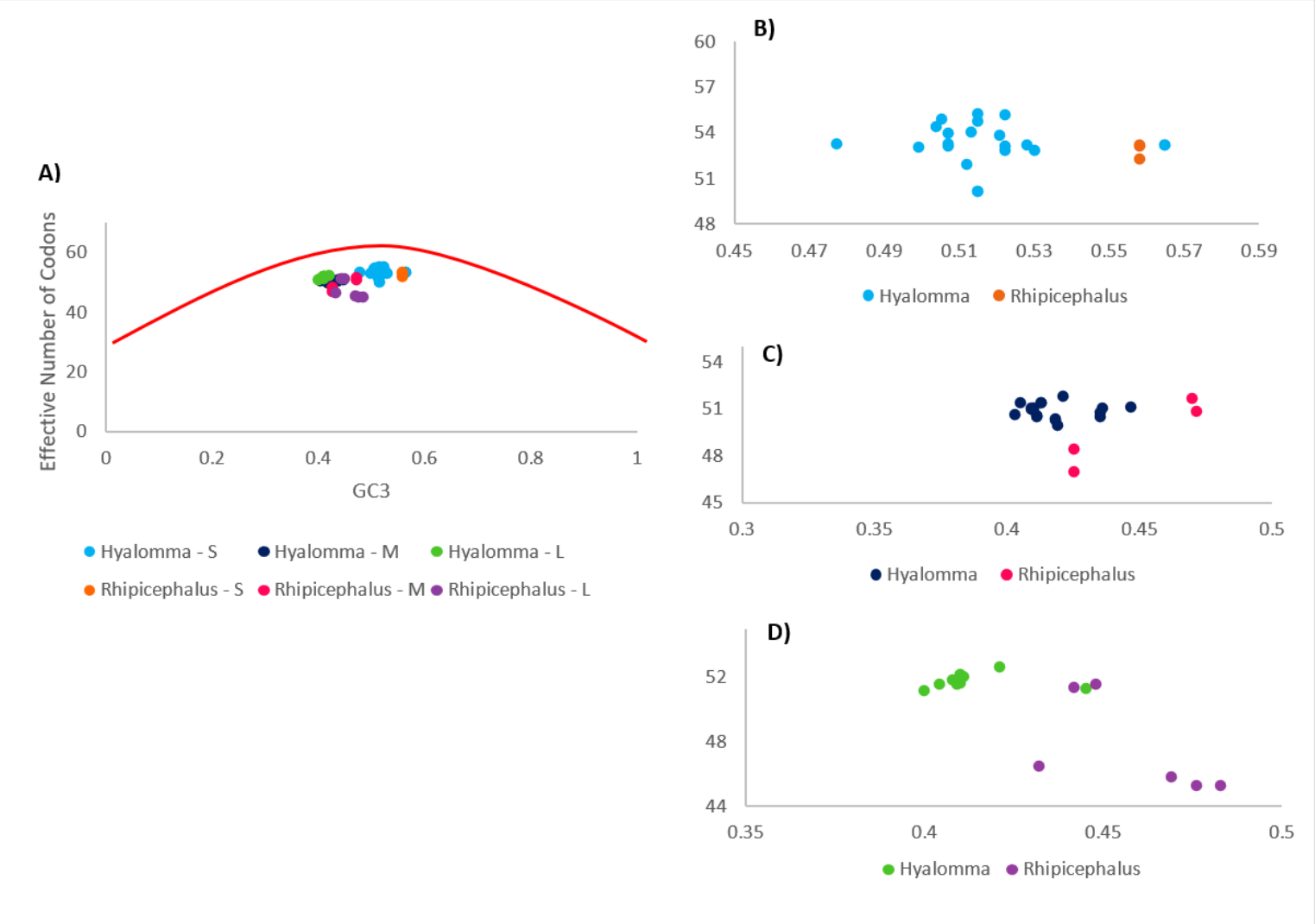


Figure 12. ENC-GC3 plots of CCHFV genomes. (A) The effective number of codons (ENC values, Y-axis) was plotted against the GC content at the third synonymous codon positions (GC3 values, X-axis). The curve (red line) indicates the expected codon usage if GC compositional constraints alone account for the codon usage bias. (B), (C), (D) ENC-GC3 plots were separately produced for *Hyalomma*- and *Rhipicephalus*-isolated S, M, and L segments, respectively.

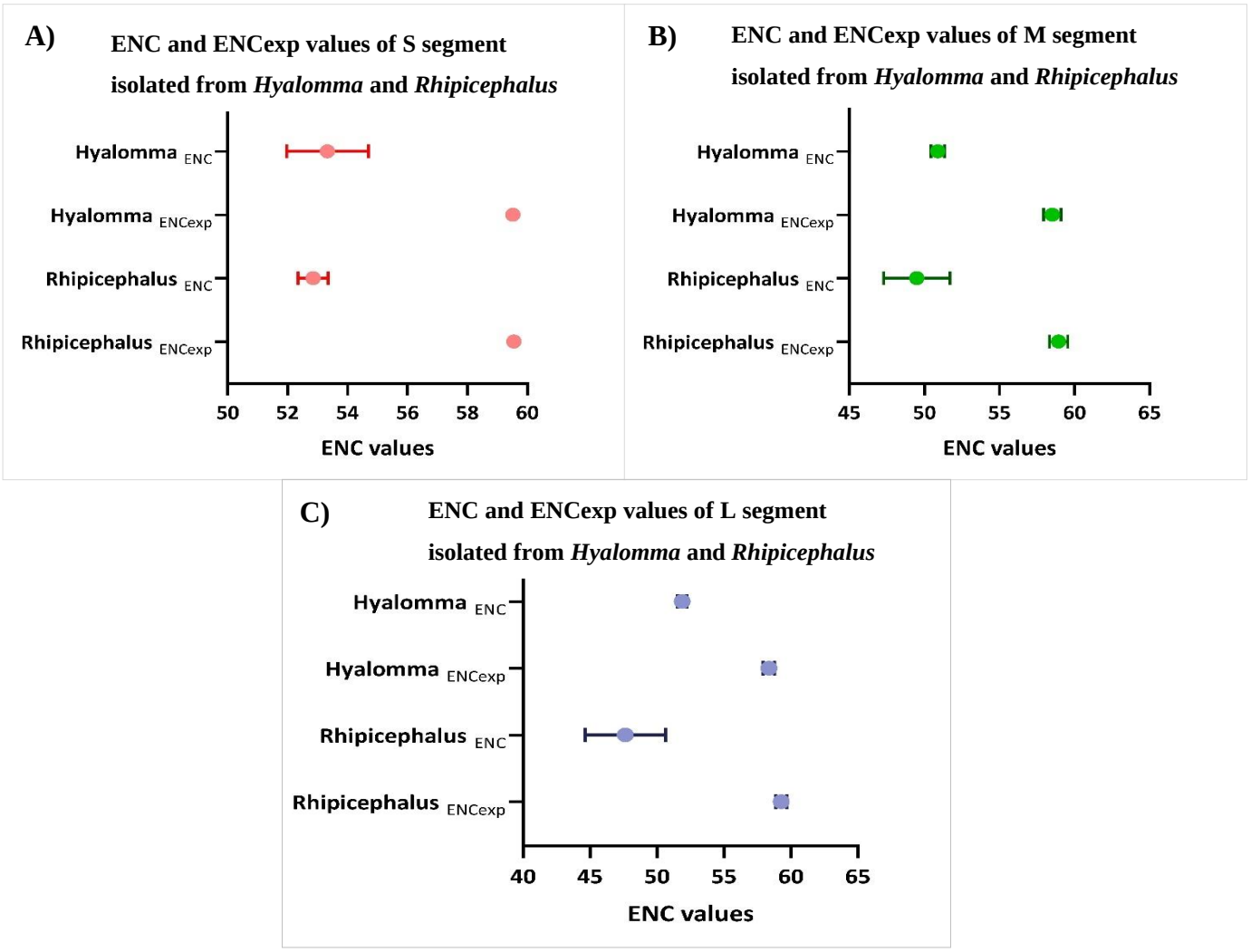


Figure 13. Bar plots representing mean and standard deviation of ENC and ENCexp values. (A), (B), (C) represent small, medium and large segments isolated two different hosts, respectively.

4.1.3. Variation in Codon Usage

Correspondence analysis (COA) was performed to detect the codon usage variation among different CCHFV strains (S, M, and L segments) isolated from *Hyalomma* and *Rhipicephalus* hosts.

The analysis is used to identify the systematic relationships among variables. Additionally, it simplifies complex data to deliver different strains or genes in multidimensional space (Butt et al. 2014; Greenacre 1984; Kumar et al. 2016). The COA was performed based on the relative synonymous codon usage (RSCU) values for

each segment strain (S, M, and L) of CCHFV, and the distribution of the strains in the plane of first two principal axes of COA was determined.

The results revealed that CCHFV strains isolated from *Hyalomma* and *Rhipicephalus* were collected into clusters (Figure 14) and support the results of ENC analysis. Scattered points in the plot describe the connection of CCHFV strains with each other depending on their host, and it clearly indicates that *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains have considerable variation in codon usage patterns.

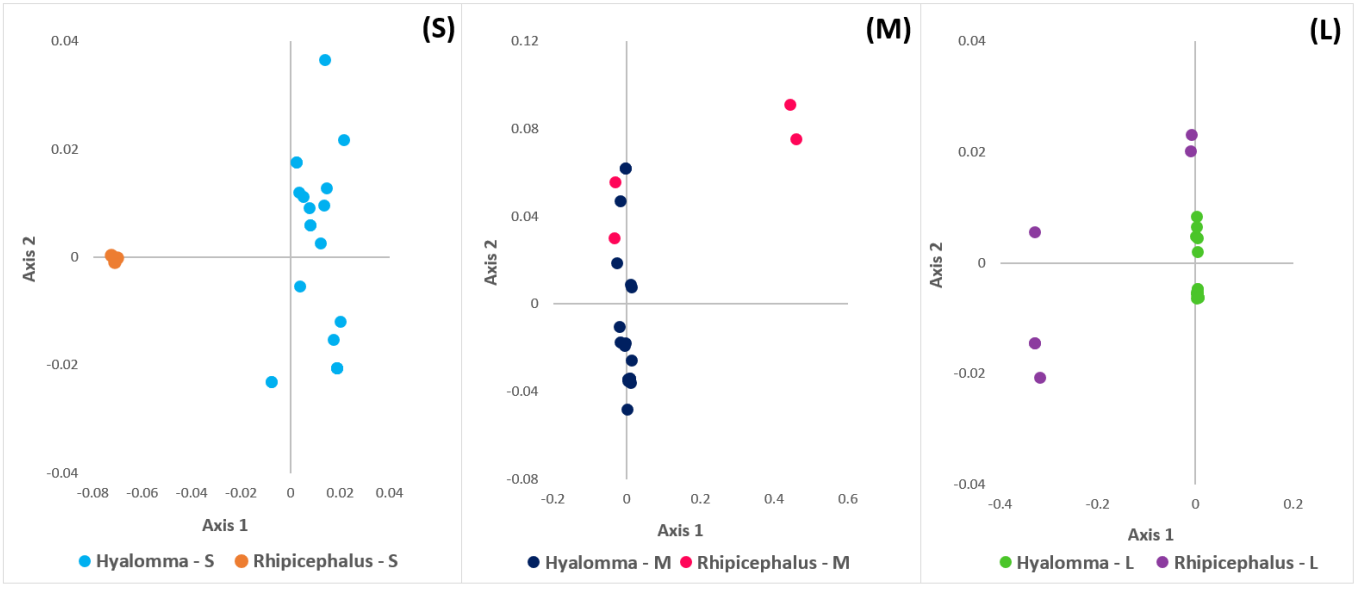
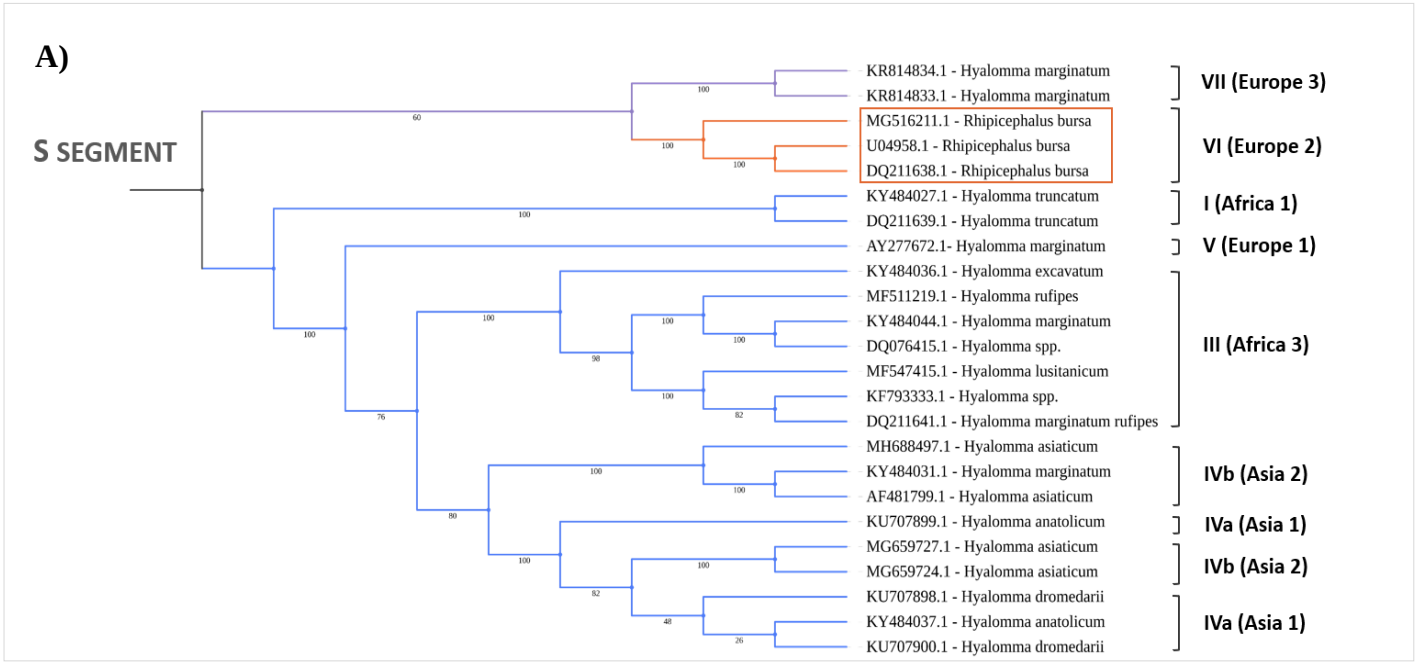
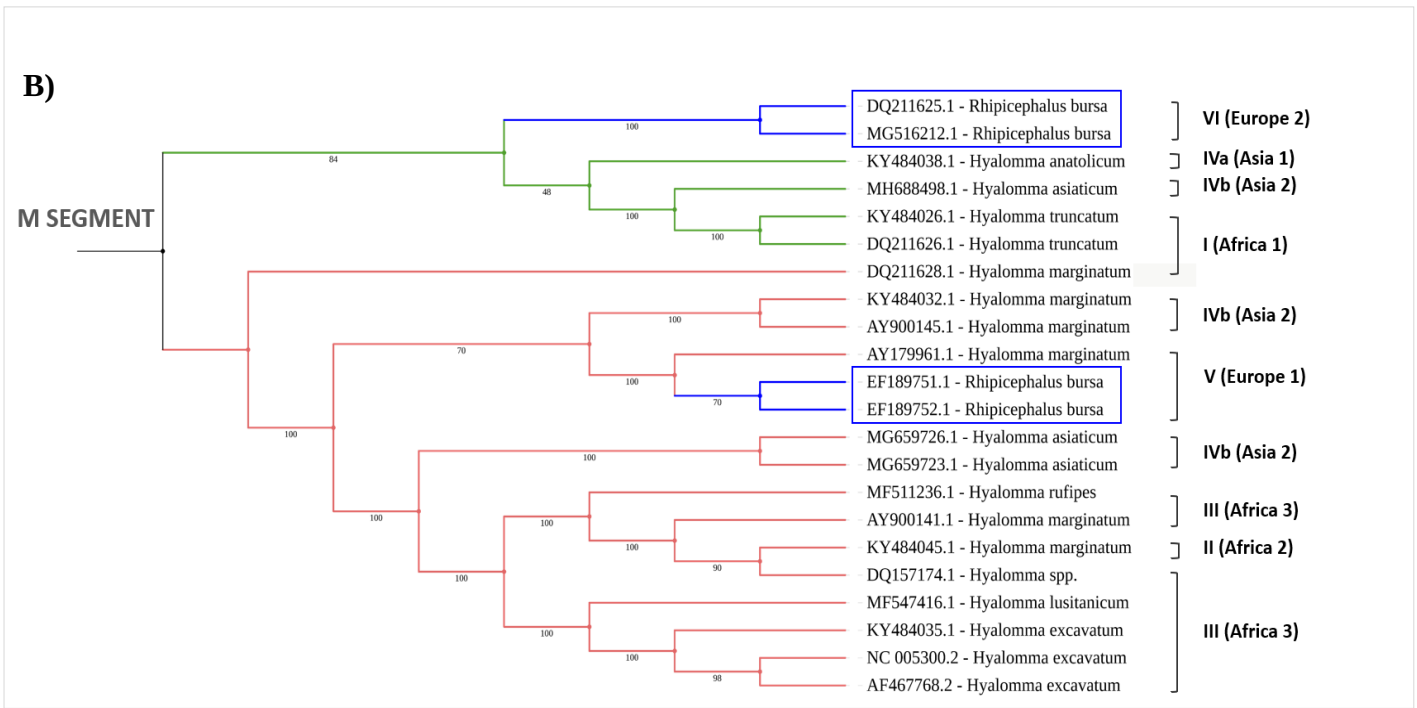


Figure 14. Correspondence analysis (COA). COA values are based on the RSCU values of three genomic parts of CCHFV isolated from *Hyalomma* and *Rhipicephalus*. All sthe strains are plotted in variance plane. Each point represents a strain, and *Hyalomma* and *Rhipicephalus*-isolated strains were marked in different colours.

Furthermore, phylogenetic analysis using the maximum likelihood method was performed in order to display how CCHFV strains isolated from *Hyalomma* and *Rhipicephalus* are assembled into clusters (Figure 15A, 15B, 15C). The analysis indicates that CCHFV strains isolated from *Rhipicephalus* are genetically distant comparing with *Hyalomma*-isolated strains. This highlights the evidence of strong selection pressure on host adaptation which is in agreement to the CAI analysis.


A)
S SEGMENT
KR814834.1 - Hyalomma marginatum
KR814833.1 - Hyalomma marginatum
MG516211.1 - Rhipicephalus bursa
U04958.1 - Rhipicephalus bursa
DQ211638.1 - Rhipicephalus bursa
KY484027.1 - Hyalomma truncatum
DQ211639.1 - Hyalomma truncatum
AY277672.1- Hyalomma marginatum
KY484036.1 - Hyalomma excavatum
MF511219.1 - Hyalomma rufipes
KY484044.1 - Hyalomma marginatum
DQ076415.1 - Hyalomma spp.
MF547415.1 - Hyalomma lusitanicum
KF793333.1 - Hyalomma spp.
DQ211641.1 - Hyalomma marginatum rufipes
MH688497.1 - Hyalomma asiaticum
KY484031.1 - Hyalomma marginatum
AF481799.1 - Hyalomma asiaticum
KU707899.1 - Hyalomma anatolicum
MG659727.1 - Hyalomma asiaticum
MG659724.1 - Hyalomma asiaticum
KU707898.1 - Hyalomma dromedarii
KY484037.1 - Hyalomma anatolicum
KU707900.1 - Hyalomma dromedarii
VII (Europe 3)
VI (Europe 2)
I (Africa 1)
V (Europe 1)
III (Africa 3)
IVb (Asia 2)
IVa (Asia 1)
IVb (Asia 2)
IVa (Asia 1)


B)
M SEGMENT
DQ211625.1 - Rhipicephalus bursa
MG516212.1 - Rhipicephalus bursa
KY484038.1 - Hyalomma anatolicum
MH688498.1 - Hyalomma asiaticum
KY484026.1 - Hyalomma truncatum
DQ211626.1 - Hyalomma truncatum
DQ211628.1 - Hyalomma marginatum
KY484032.1 - Hyalomma marginatum
AY900145.1 - Hyalomma marginatum
AY179961.1 - Hyalomma marginatum
EF189751.1 - Rhipicephalus bursa
EF189752.1 - Rhipicephalus bursa
MG659726.1 - Hyalomma asiaticum
MG659723.1 - Hyalomma asiaticum
MF511236.1 - Hyalomma rufipes
AY900141.1 - Hyalomma marginatum
KY484045.1 - Hyalomma marginatum
DQ157174.1 - Hyalomma spp.
MF547416.1 - Hyalomma lusitanicum
KY484035.1 - Hyalomma excavatum
NC 005300.2 - Hyalomma excavatum
AF467768.2 - Hyalomma excavatum
VI (Europe 2)
IVa (Asia 1)
IVb (Asia 2)
I (Africa 1)
IVb (Asia 2)
V (Europe 1)
IVb (Asia 2)
III (Africa 3)
II (Africa 2)
III (Africa 3)

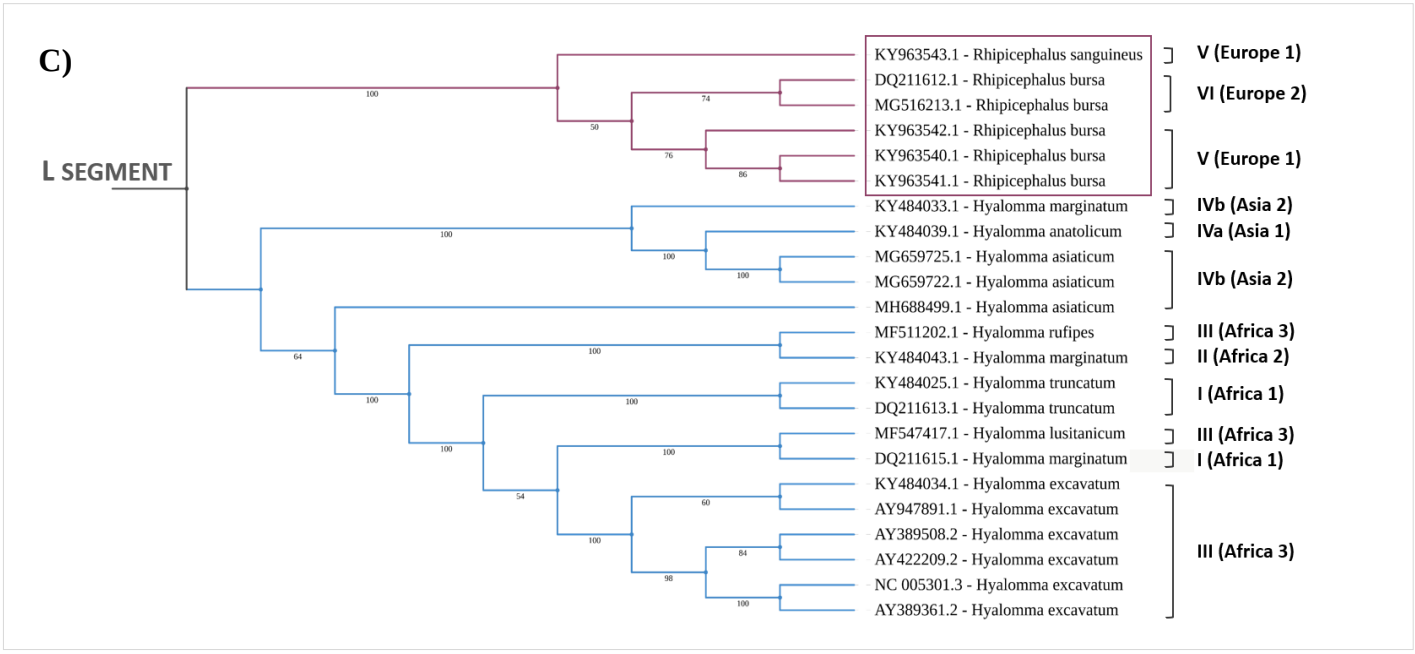


Figure 15. Phylogenetic trees based on the polyprotein-coding regions of 70 CCHFV strains isolated from *Hyalomma* and *Rhipicephalus*. The tree was generated by the maximum likelihood method using MEGA. "iTOL" was further used to design the trees. Phylogenetic trees were separately produced in relation to S (A), M (B), and L (C) segments, respectively. Phylogenetic clades which strains belong to were attached next to the strains.

4.1.4. Codon usage preferences using RSCU analysis

In order to compare the codon usage preferences of three different segments of CCHFV being isolated from *Hyalomma* and *Rhipicephalus*, the RSCU values for each synonymous codon were calculated separately for *Hyalomma*- and *Rhipicephalus*-isolated S, M, and L segments. The results of this analysis are shown in Table 5.

G or C-ended codons are preferred to be used over A or U-ended codons in S segment isolated from *Hyalomma* and *Rhipicephalus*. However, the frequencies of A/U-ended codons are different, for instance, strains of *Rhipicephalus*-isolated S segment have more codons ended with A, while *Hyalomma* ones have a higher frequency in U-ended codons.

M and L segments isolated from two hosts have higher tendency to use A/U-ended codons over G/C-ended codons. M segment isolated from *Hyalomma* show higher preference to use A-ended codons, while strains isolated from *Rhipicephalus* have a preference to highly use U-ended codons. Although M segments isolated from

both hosts use different codons, both show same frequencies preferring to use C-ended codons more, and G-ended codons less. L segments isolated from two hosts have similar preferences having more A-ended codons than U. However, *Rhipicephalus*-isolated L segments have higher frequencies in C-ended codons over G, while *Hyalomma*-isolated L segments shows equal choice for G or C-ended codons.

Codon over- and underrepresentation analysis emphasized that RSCU values of the majority preferred and non-preferred codons ranged from 0.6 to 1.6. Interestingly, we found that the nucleotide frequencies at the end of most over-represented codons (RSCU > 1.6) differ between *Hyalomma* and *Rhipicephalus*-isolated CCHFV strains (Table 5).

Ser, Pro, Thr, Ala, Arg are over-represented in M segments isolated from both hosts. However, *Hyalomma*-isolated strains have codon preference on UCA for Ser, while *Rhipicephalus* ones prefer to use AGC. In addition, they differ on codon usage preferences for Ala, where GCA is over-represented in *Hyalomma*-isolated M segments and *Rhipicephalus* ones have preferences to use GCC codon.

Similar to M segments isolated from *Hyalomma*, L segments have only two over-represented codons ending with A and one with G, however, the frequencies of ten over-represented codons of *Rhipicephalus*-isolated L segments are U: 3; C:2, A:3, G:2. In the case of Arg, *Hyalomma*-isolated strains have codon preference on AGA, while *Rhipicephalus*-isolated strains prefer to use AGG.

Furthermore, to conclude whether codon usage patterns of *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains show differences, we compared their codon usage preferences. The results showed that out of 59 synonymous codons, 6 codons were differently used, whereas 12 codons were equivalent between *Hyalomma*- and *Rhipicephalus*-isolated S and M segments (Table 5). The most similar patterns were observed in *Hyalomma*- and *Rhipicephalus*-isolated L segments where 15 codons were equally used, while only 5 codon preferences show differences.

Table 5. The relative synonymous codon usage frequency (RSCU) of CCHFV strains isolated from *Hyalomma* and *Rhipicephalus*. AA represents amino acid; the "RSCU" value represents the pattern of relative synonymous codon usage; green colours represents the optimal codons favoured by CCHFV isolated from two different hosts (RSCU > 1); over-represented (RS > 1.6) codons are marked as bold, and under-represented (RSCU < 0.6) codons are marked as underline, respectively.

		CCHFV - S SEGMENT		CCHFV - M SEGMENT		CCHFV - L SEGMENT	
AA	Codons	Hyalomma	Rhipicephalus	Hyalomma	Rhipicephalus	Hyalomma	Rhipicephalus
Phe	UUU	0.96	0.85	1.09	1.33	1.14	1.16
	UUC	1.04	1.15	0.91	0.68	0.86	0.84
Leu	UUA	0.13	0	0.98	0.83	0.83	0.28
	UUG	0.56	0.84	1.17	1.19	1.32	**1.76**
	CUU	**2.13**	**1.80**	1.08	1.49	1.10	1.15
	CUC	1.43	1.57	0.67	0.62	0.92	1.22
	CUA	0.44	0.33	1.05	0.81	0.93	0.71
	CUG	1.30	1.46	1.06	1.08	0.91	0.89
Ile	AUU	1.25	1.20	1.13	1.41	1.20	**1.64**
	AUC	0.99	1.23	0.84	0.62	0.68	0.60
	AUA	0.76	0.57	1.03	0.98	1.12	0.76
Val	GUU	1.03	0.89	1.28	1.25	1.28	1.34
	GUC	1.23	1.18	0.96	1.18	0.86	0.80
	GUA	0.27	0.49	0.72	0.29	0.62	0.58
	GUG	1.47	1.43	1.04	1.26	1.23	1.28
Ser	UCU	1.46	1.49	0.83	0.47	1.18	0.88
	UCC	1.08	1.05	0.58	0.48	0.61	0.67
	UCA	0.97	0.72	**1.74**	1.05	1.15	0.50
	UCG	0.26	0.22	0.18	0.16	0.20	0.08
	AGU	1.02	0.72	1.21	**1.61**	1.40	**1.74**
	AGC	1.21	**1.82**	1.46	**2.23**	1.47	**2.13**
Pro	CCU	1.17	1.19	1.25	1.43	1.42	0.57
	CCC	0.49	0.77	0.83	0.41	0.76	0.90
	CCA	**1.85**	**1.68**	**1.71**	**1.93**	1.52	**2.44**
	CCG	0.49	0.35	0.22	0.24	0.29	0.09
Thr	ACU	0.98	1.16	1.0	0.66	1.21	0.86
	ACC	1.58	1.16	0.91	0.96	1.14	**1.88**
	ACA	1.31	1.30	**1.78**	**2.00**	1.44	0.98
	ACG	0.13	0.37	0.30	0.39	0.21	0.29

		CCHFV - S SEGMENT		CCHFV - M SEGMENT		CCHFV - L SEGMENT	
AA	Codons	Hyalomma	Rhipicephalus	Hyalomma	Rhipicephalus	Hyalomma	Rhipicephalus
Ala	GCU	1.02	0.85	0.97	0.58	1.10	0.92
	GCC	1.46	1.39	1.13	**1.92**	0.62	0.46
	GCA	1.36	**1.63**	**1.78**	1.44	**2.02**	**2.19**
	GCG	0.15	0.13	0.12	0.06	0.26	0.42
Tyr	UAU	0.62	0.72	0.83	0.87	0.99	0.59
	UAC	1.38	1.28	1.17	1.13	1.01	1.41
His	CAU	0.71	0.56	0.93	0.53	1.23	**1.64**
	CAC	1.29	1.44	1.07	1.47	0.77	0.36
Gln	CAA	0.70	0.36	0.86	0.60	0.95	0.55
	CAG	1.30	**1.64**	1.14	1.40	1.05	1.45
Asn	AAU	0.64	0.70	0.97	1.12	1.01	1.11
	AAC	1.36	1.30	1.03	0.88	0.99	0.90
Lys	AAA	0.85	0.65	1.12	1.02	1.13	1.25
	AAG	1.15	1.35	0.88	0.99	0.87	0.75
Asp	GAU	0.93	0.67	0.97	1.12	1.13	1.47
	GAC	1.07	1.33	1.03	0.88	0.87	0.54
Glu	GAA	0.84	0.59	1.24	0.99	1.27	1.55
	GAG	1.16	1.41	0.76	1.01	0.73	0.45
Cys	UGU	1.25	1.33	0.98	0.84	1.07	0.93
	UGC	0.75	0.67	1.02	1.16	0.93	1.08
Arg	CGU	1.34	0.47	0.13	0.06	0.19	0.09
	CGC	0.23	0.59	0.16	0.04	0.29	0.53
	CGA	0.16	0.35	0.14	0.04	0.33	0.11
	CGG	0.37	0.12	0.14	0.06	0.21	0.03
	AGA	**1.71**	**1.76**	**3.49**	**3.89**	**2.66**	**2.04**
	AGG	**2.20**	**2.70**	**1.93**	**1.90**	**2.32**	**3.21**
Gly	GGU	0.77	1.14	1.00	0.73	1.26	1.38
	GGC	1.26	1.03	1.23	1.55	0.91	1.16
	GGA	1.20	1.22	1.06	1.10	1.14	0.80
	GGG	0.77	0.61	0.72	0.62	0.69	0.68

4.1.5. Codon Usage Adaptation

Codon adaptation index (CAI) analyses were performed to gain insight into the codon preferences of CCHFV in relation to its tick hosts. CAI index varies from 0 to 1, being 1 if the frequency of codon usage by CCHFV equals the frequency of usage of the reference set of tick host. Higher CAI values signify higher levels of codon usage bias. The most frequent codons show the highest relative adaptation to the host, and sequences with higher CAIs are considered to be preferred over those with lower CAIs.

CAI values of CCHFV strains isolated from *Hyalomma* and *Rhipicephalus* were shown in Table 6 in relation to *Hyalomma* and *Rhipicephalus* codon usage.

Table 6. Results of CAI analysis. CAI, expected CAI and normalized CAI values of *Hyalomma* (Hy.) and *Rhipicephalus* (Rh.) isolated S, M, and L segments in relation to *Hyalomma* (*H. asiaticum*) and *Rhipicephalus* (*R. sanguineus*) codon usage. Significant differences in CAI values in relation to codon usage of two hosts are marked as bold.

	CAI	eCAI (p<0.05)	nCAI
S SEGMENT			
Hy.-isolated strains to Hy. codon usage	**0.698 ± 0.01**	0.706	0.989
Hy.-isolated strains to Rh. codon usage	**0.681 ± 0.01**	0.710	0.959
Rh.-isolated strains to Hy. codon usage	0.717 ± 0.004	0.715	1.002
Rh.-isolated strains to Rh. codon usage	0.702 ± 0.003	0.723	0.971
M SEGMENT			
Hy.-isolated strains to Hy. codon usage	**0.656 ± 0.01**	0.666	0.985
Hy.-isolated strains to Rh. codon usage	**0.633 ± 0.01**	0.654	0.968
Rh.-isolated strains to Hy. codon usage	**0.680 ± 0.01**	0.682	0.997
Rh.-isolated strains to Rh. codon usage	**0.638 ± 0.02**	0.662	0.964
L SEGMENT			
Hy.-isolated strains to Hy. codon usage	**0.641 ± 0.002**	0.666	0.962
Hy.-isolated strains to Rh. codon usage	**0.619 ± 0.002**	0.650	0.952
Rh.-isolated strains to Hy. codon usage	**0.668 ± 0.01**	0.679	0.984
Rh.-isolated strains to Rh. codon usage	**0.641 ± 0.01**	0.667	0.961

The Student's t-test and Mann-Whitney U test were performed to observe if the differences in CAI values between strains isolated from two hosts in relation to *Hyalomma* and *Rhipicephalus* codon usage were statistically significant. Statistical analyses showed that there are statistically significant differences in CAI values between *Hyalomma*-isolated S, M, and L and *Rhipicephalus*-isolated M and L segments ($p<0.05$).

In order to ascertain whether statistically significant differences in CAI values result from codon preferences (Puigbo et al. 2008b), the expected CAI (eCAI) values were computed for CCHFV sequences isolated from two ticks in relation to S, M, and L segments. The eCAI is an algorithm (Puigboet al. 2008b) that is used to calculate the expected value of the CAI by generating 500 random sequences with the same nucleotide and amino acid composition as sequences of interest. In order to show if the generated sequences have a normal distribution, a Kolmogorov-Smirnov test for the eCAI of these random sequences was applied. The results for eCAI were displayed in Table 6 and suggested that there was a normal distribution of all the generated sequences below of critical value of 0.061.

Furthermore, normalized CAI values were calculated once CAI and eCAI values were obtained (Table 6, Figure 16). Normalized CAI value is a ratio of observed CAI to expected CAI value. These values showed that both *Hyalomma*- and *Rhipicephalus*-isolated strains have a higher tendency and adaptation to use codons that are preferred by *Hyalomma* spp.

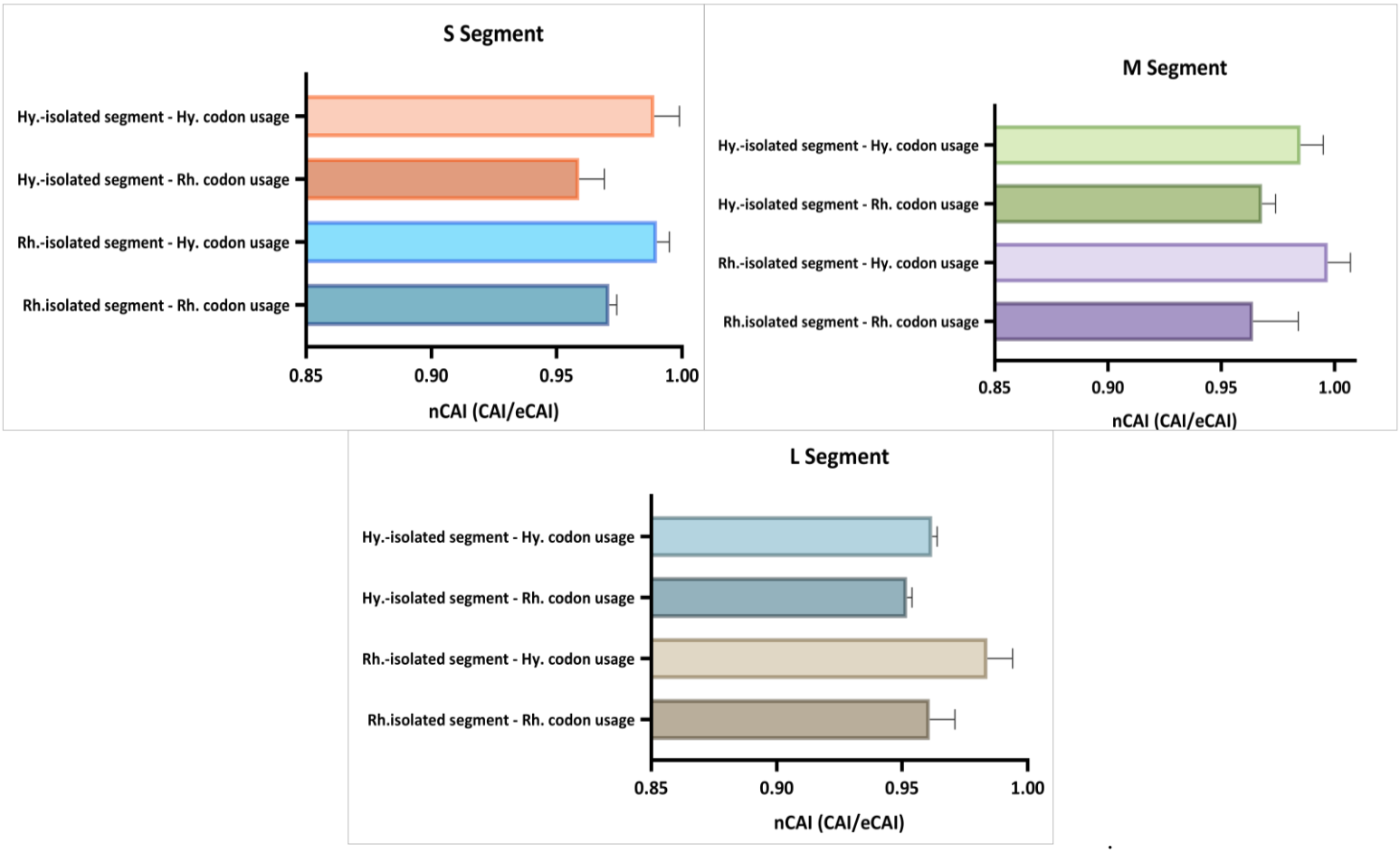


Figure 16. Comparison of nCAI values. nCAI values of *Hyalomma*-isolated strains to *Hyalomma* and *Rhipicephalus* codon usage and *Rhipicephalus*-isolated strains to *Hyalomma* and *Rhipicephalus* codon usage in relation to small (A), medium (B) and large (C) segments, respectively.

4.2. Tick collection and screening for CCHFV

A total 152 ixodid ticks (151 adults, 71 females, 80 males; 1 larva, female) were collected from cattle in three municipalities of Kosovo in August 2020 (Figure 17, Table 7). The ticks belonged to six species, namely *Hyalomma marginatum* (78, 51.3%), *Rhipicephalus annulatus* (66, 43.4%), *Hyalomma anatolicum* (5, 3.3%), *Haemaphysalis punctata* (1, 0.7%), *Rhipicephalus bursa* (1, 0.7%) and *Rhipicephalus sanguineus* (1, 0.7%).

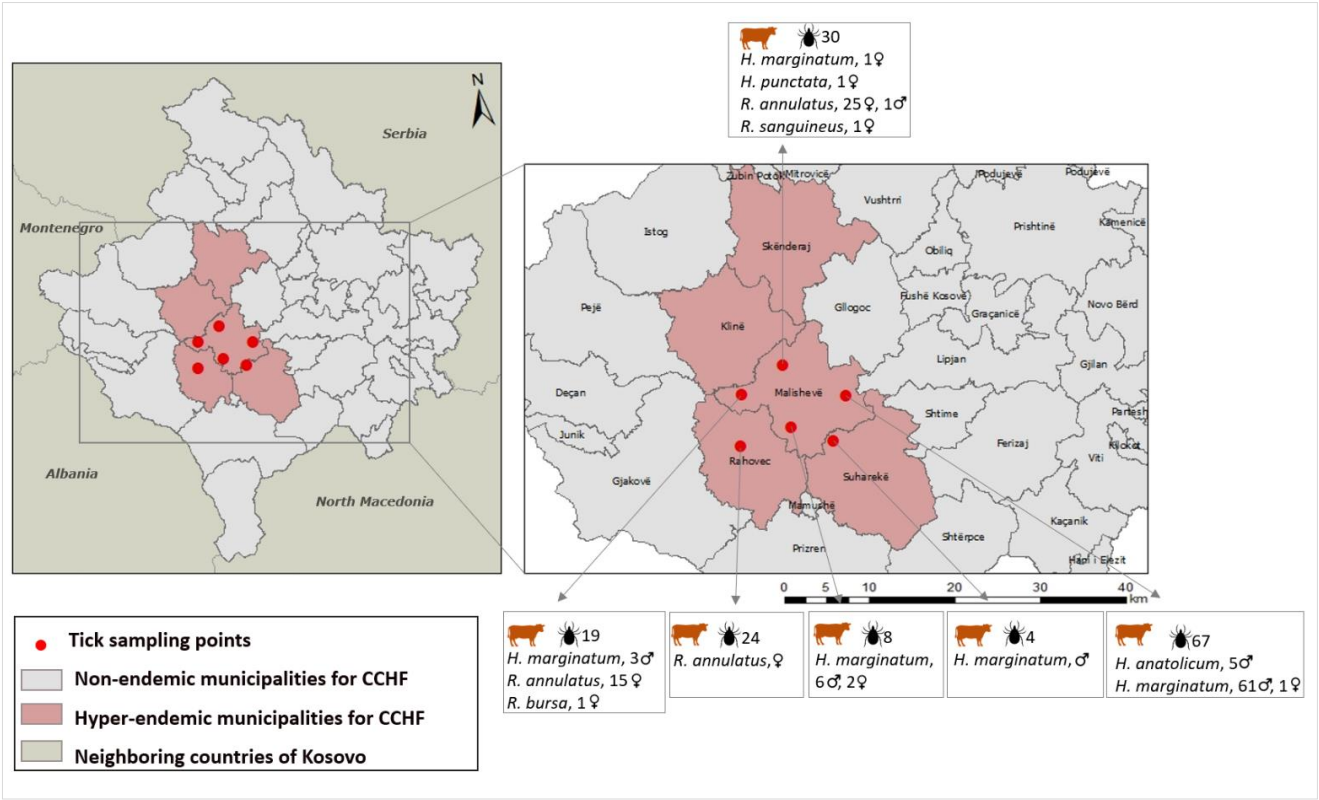


Figure 17. Study area in Kosovo and tick collection points from livestock. The figure shows total number of ticks per municipality, total number of collected ticks per species in municipality, and their sex.

Table 7. Number of tick species collected from Kosovo, 2020.

			Number of specimens collected per species no. (% of total collection)						
Collection Area	Host	Municipality	*Hyalomma marginatum*	*Rhipicephalus annulatus*	*Hyalomma anatolicum*	*Haemaphysalis punctata*	*Rhipicephalus bursa*	*Rhipicephalus sanguineus*	**Total**
Kosovo	Cattle	Malisheve	74	42	5	1	1	1	124
		Rahovec	-	24	-	-	-	-	24
		Suhareke	4	-	-	-	-	-	4
		Total	**78 (51.3%)**	**66 (43.4%)**	**5 (3.3%)**	**1 (0.7%)**	**1 (0.7%)**	**1 (0.7%)**	**152**

In Czechia, a total 19 ixodid ticks (17 adults, female; 1 larva; 1 nymph) were collected from 13 individual birds in the bird ringing station near Choteč during September-October 2020 (Figure 18, Table 8).

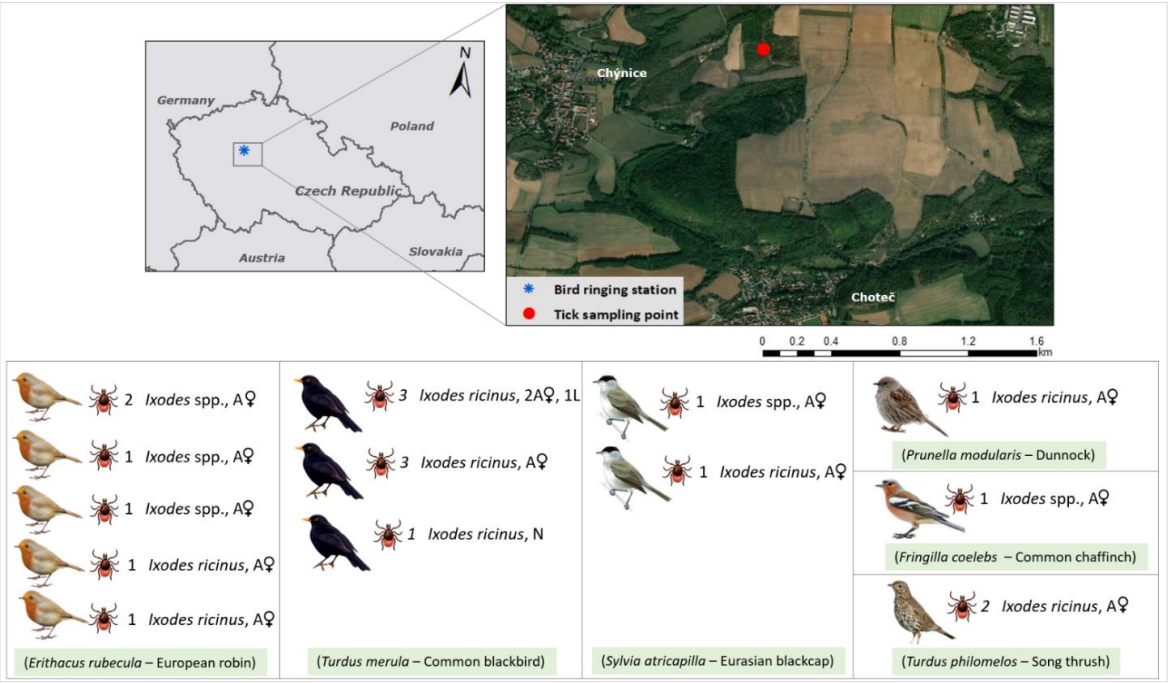


Figure 18. Tick collection point in bird ringing station in Czechia. Total number of collected ticks per individual migratory birds, their sex and life stages (L = larva, N = nymph, A = adult).

Table 8. Number of tick species collected from birds in Czechia, 2020. 1K and +1K indicate 1st calendar year and older than 1st calendar year, respectively, for birds' ages. M = male, F = female, A = adult, N = nymph, L = larva.

			Ticks (Species, Life Stage, Sex)	
Collection Area	Individual birds	Age & Sex of Birds	*Ixodes ricinus*	*Ixodes* spp.
Bird ringing station, Choteč	*Sylvia atricapilla*	1K, M		1 (A, F)
	Sylvia atricapilla	1K, M	1 (A, F)	
	Erithacus rubecula	1K, F		2 (A, F)
	Erithacus rubecula	1K, F		1 (A, F)
	Erithacus rubecula	1K, M		1 (A, F)
	Erithacus rubecula	1K, F	1 (A, F)	
	Erithacus rubecula	1K, F	1 (A, F)	
	Prunella modularis	1K, F	1 (A, F)	
	Fringilla coelebs	+1K, M		1 (A, F)
	Turdus merula	+1K, M	3 (2A, F; 1L)	
	Turdus merula	1K, F	3 (A, F)	
	Turdus merula	1K, F	1 (N)	
	Turdus philomelos	1K, F	2 (A, F)	
	Total individuals (13)		**Total ticks (19)**	

All ticks collected from livestock in Kosovo and migratory birds in Czechia tested negative for CCHFV.

4.3. Mapping environmental suitability of *H. marginatum* to Central Europe

Total of 2108 candidate models were built with four sets of variables in estimating the ENM of *H. marginatum*. Candidate models were evaluated with parameters reflecting all combinations of 17 regularization multiplier settings, 31 feature class combinations, and four distinct sets of environmental variables.

Model performance was evaluated based on statistical significance (partial ROC), omission rates (OR), and the Akaike information criterion (AICc). Partial ROC and omission rates were evaluated based on models created with training occurrences, whereas AICc values were calculated for models created with the full set of occurrences (Warren & Seifert 2011). The performance of the model was evaluated under optimal parameters using a set of environmental predictors (Set 1), statistically significant models (N = 2108), best candidate models (N = 3), regularization multiplier (N = 2.0), features classes (linear, quadratic and product), mean AUC ratio (N = 1.24), partial ROC (N = 0.00), omission rate 5% (N = 0.044), Akaike Information Criterion corrected (N = 6161.85), Delta Akaike Information Criterion corrected (N = 0.3), Akaike Information Criterion corrected weight (N = 0.28), and number parameters (N = 13) for *H. marginatum.*

All candidate models were statistically significant. At the end, only three models met the three selection criteria and were identified as the best candidate models based on their performance. Set 1 (Bioclimatic variables) were the variables selected for the construction of the ENM (Table 3). According to the estimated permutation importance, the results showed that annual mean temperature (Bio1) was the variable that contributed most to the model performance (63.6%), while the mean temperature of the warmest quarter (Bio10) and the temperature seasonality contributed with 7% and 6.8%, respectively (Table 9).

Table 9. List of Set 1 environmental variables used for creation of model and their percentage contribution in final model.

Bioclimate Variables	Description of Variables	Percent contribution	Permutation importance
Bio1	Annual mean temperature	63.6	77.6
Bio10	Mean temperature of the warmest quarter	7	0.4
Bio4	Temperature Seasonality	6.8	9.4
Bio5	Maximum temperature of warmest month	6.5	3.6
Bio11	Mean temperature of the coldest quarter	4.8	1.2
Bio2	Mean diurnal range	2.6	1.6
Bio3	Isothermality	2.3	0.9
Bio7	Temperature annual range	1.8	0.4
Bio15	Precipitation seasonality	1.1	1.2
Bio14	Precipitation of the driest month	1.1	0.3
Bio16	Precipitation of the wettest quarter	0.8	1
Bio6	Minimum temperature of coldest month	0.6	0.3
Bio17	Precipitation of driest quarter	0.5	1.6
Bio13	Precipitation of the wettest month	0.4	0.3
Bio12	Annual precipitation	0.2	0.2

Statistical results from model prediction showed that the prediction was significantly better than random expectations, with AUC ratios uniformly above the null expectations ($p<0.05$, Table 10).

Table 10. Results of partial area under the curve (AUC) ratios. It summarizes evaluations of ecological niche model of *H. marginatum*.

Descriptive statistics for AUC ratio	
Mean	1.95
Median	1.96
Minimum	1.90
Maximum	1.98
Range	0.08
Standard Deviation	0.02

No replicate had AUC ratio below 1, indicating the predicted model performed better than random ($p<0.05$).

Modeling results showed a high probability of suitable conditions for *H. marginatum* in southern Europe, especially in the Mediterranean parts of Spain,

Portugal, Italy, and Greece and along the Adriatic shore (Figure 19). Our model depicted high suitability also in the southern part of France. The model also depicted medium suitability in parts of the Netherlands and Belgium.

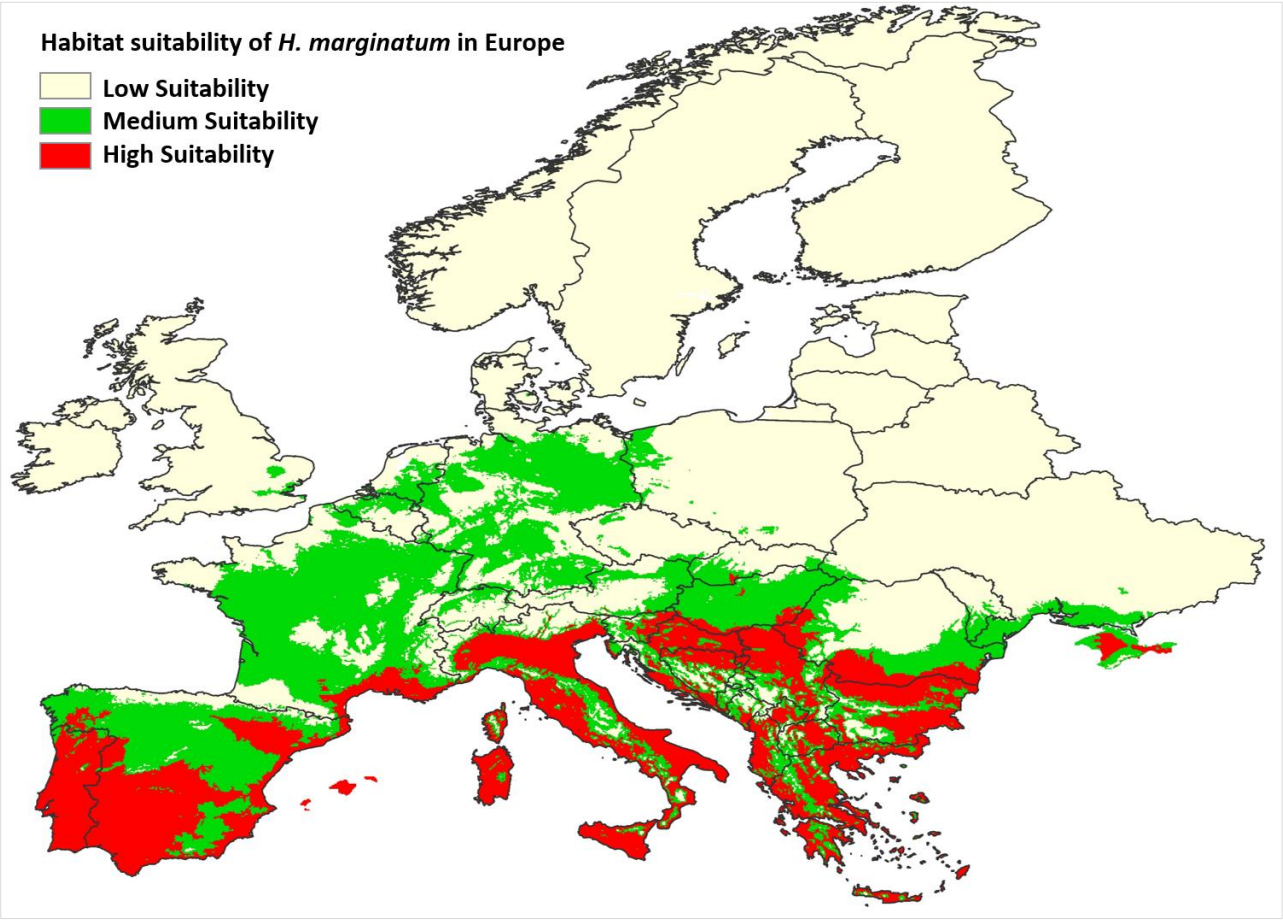


Figure 19. Environmental suitability of main vector of Crimean-Congo haemorrhagic fever virus, *Hyalomma marginatum*, in Europe.

The ENM of *H. marginatum* in Central Europe anticipated its distribution in all countries of Central Europe (Figure 19). *H. marginatum* showed broader environmental suitability in Austria, Germany, Hungary, and Slovakia, followed by Czechia and Poland (Figure 20).

Taken together, large areas of Czechia have low environmental suitability for *H. marginatum*. However, some regions in the southern and northern parts of Czechia cover a wide range of medium habitat suitability for *H. marginatum* (Figure 20).

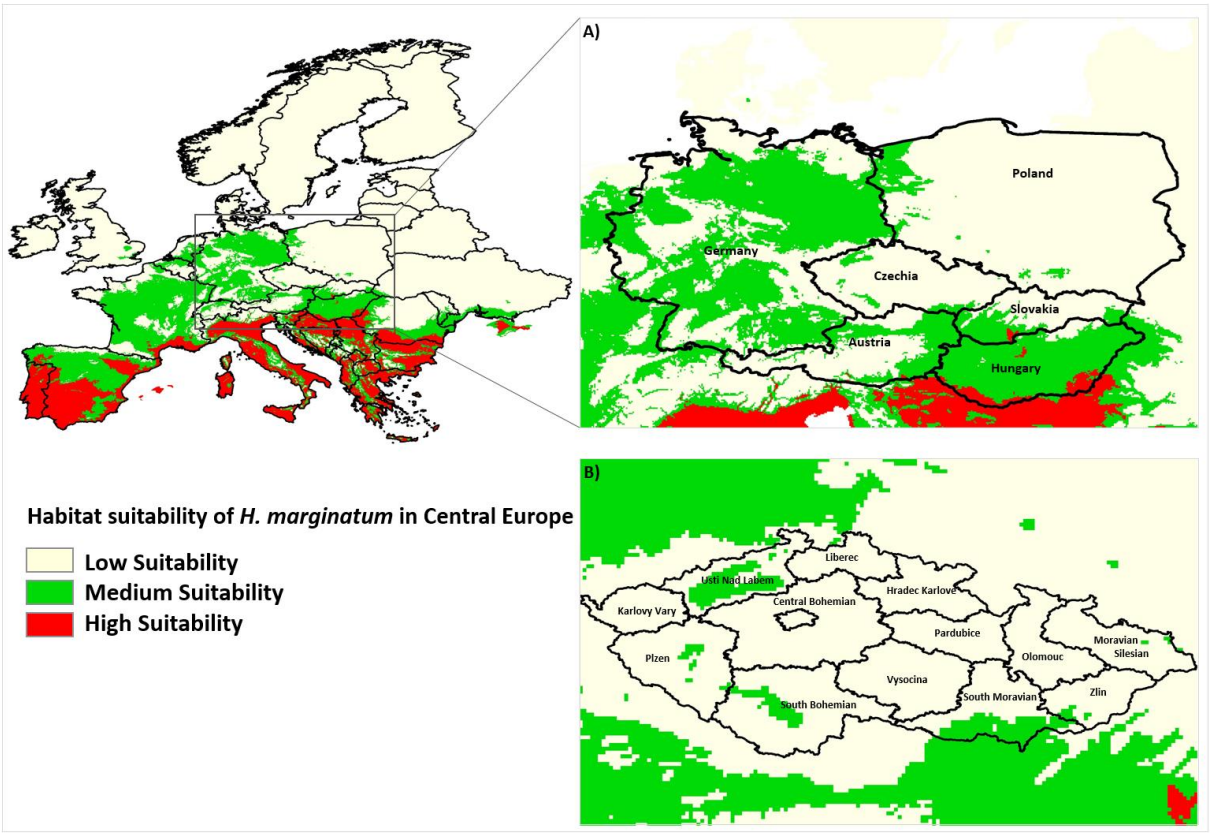


Figure 20. ENM of *Hyalomma marginatum*, in Europe. Closer view of predicted distribution of *H. marginatum* in Central Europe (A) and Czechia (B).

5. Discussion

5.1. Coevolution between CCHFV and its vectors

In the first part of this thesis, codon usage in coding sequences from *Hyalomma*- and *Rhipicephalus*-isolated CCHFV genomes was analysed to understand its molecular evolution in relation to two different tick hosts. Previously, it has been shown that codon usage bias, or preference for one type of codon over another, can be greatly influenced by overall nucleotide composition in genome (Jenkins & Holmes 2003; Wong et al. 2010; Nasrullah et al. 2015).

The analyses showed that there are substantial differences on frequencies of occurrence of nucleotides between *Hyalomma* and *Rhipicephalus*-isolated CCHFV variants (Table 4, Figure 11). Despite the differences, it is clear that the M and L segments of *Hyalomma* and *Rhipicephalus*-isolated CCHFV strains are AU rich; and A/U ending codons appear to be preferred. This is consistent with previous reports indicating substantial portion of mammalian-host isolated CCHFV strains are enriched with AU (Nasrullah et al. 2015, Rahman et al. 2018). However, S segment isolated from two ticks are GC rich and preferentially use G/C ending codons. The biological importance of this condition is uncertain and therefore it is important to investigate the factors influencing different nucleotide frequencies of CCHFV segments (van Hemert & Berkhout 2016).

Previous studies on codon usage bias have also suggested that the composition of amino acids was also the key factor in determining the nucleotide contents at the first and second codon positions of viral genomes, while the variation in proteins was forced by functional selection. However, at the third codon positions of a viral gene, 69% of the alteration at the third codon position always denoted synonymous or silent mutations, which are not restricted by functional selection of amino acids (van Hemert et al. 2016).

The RSCU analysis also showed that CCHFV strains isolated from *Hyalomma* and *Rhipicephalus* have different codon usage preferences.

Moreover, CCHFV S segment exhibit greater codon usage bias toward G- and C-ended codons, while M and L segments show contrary preferences toward A- and U-ended codons. Remarkably, M segment isolated from *Hyalomma* have higher tendency

to use A-ended codons, while strains isolated from *Rhipicephalus* have a preference to highly use U-ended codons. For L segment, although two hosts have similar preferences having more A-ended codons than U, they exhibit different codon preferences for same amino acids such as Arg, Cys, and Thr. A previous study (Rahman et al. 2018) reported that Arg (AGA, AGG) is over-represented in CCHFV and two potential hosts (*Bos taurus*, *Ovis aries*). Our findings are consistent with that study, showing both codons (AGA, AGG) for Arg are over-represented in CCHFV strains isolated from *Hyalomma* and *Rhipicephalus*.

Once it was shown that there are different codon usage patterns and codon usage bias in CCHFV strains from two hosts, the extent of this bias in CCHFV strains from two tick hosts was determined.

A number of systemic analytical approaches were performed to explore differences in codon usage patterns of CCHFV isolated from *Hyalomma* and *Rhipicephalus*. To start with, ENC concept was used, and then ENC-GC3 analysis was performed. ENC is the index that varies between 20 and 61 and helps to demonstrate codon usage bias (Novembre 2002). As it was highlighted in the results, *Hyalomma* and *Rhipicephalus*-isolated CCHFV strains of three genomic parts have remarkably different codon usage patterns as clusters of both hosts can be clearly identified on the graphs. The overall codon usage bias of CCHFV isolated from two ticks were lower. Many other RNA viruses have tendency to have low codon usage bias, for instance, Ebola virus (ENC: 57.23) (Cristina et al. 2015), chikungunya virus (ENC: 55.56) (Butt et al. 2014), hepatitis C virus (ENC: 52.62) (Hu et al. 2011), and West Nile virus (ENC: 53.81) (Moratorio et al. 2013). It has been indicated that the low codon usage bias of virus is beneficial for the efficient replication in its host cells and the reduced competition between virus and its hosts for the protein synthesis. CCHFV segments isolated from both vectors did not show very-high ENC values, suggesting that evolution of low codon bias within CCHFV coding sequences has allowed it to successfully maintain its survival cycle within both of its vectors each of which possesses distinct codon usage preferences from each other. The results also reflect that the different usage of codons between *Hyalomma* and *Rhipicephalus*-isolated CCHFV strains may increase the ability of viral genes to be involved in the translation process.

Although codon usage bias can be estimated using ENC values, these values cannot be used alone to indicate the driving force of codon usage bias. According to

that, ENC-GC3 plot was used and showed that both natural selection and mutational pressure have affected the codon usage patterns in CCHFV strains isolated from two tick hosts.

Apart from ENC, CAI analysis is also used to calculate codon usage bias. Higher CAI values indicate higher levels of codon bias adaptation for the host and vice versa. According to CAI analysis, the results reflect that the selection pressure from hosts may influence the codon usage pattern of CCHFV. CCHFV strains isolated from *Hyalomma* and *Rhipicephalus* in relation to three genomic segments show significant differences in CAI values, suggesting these differences are risen due to different codon preferences in their codon usage. Comparative analysis of nCAI values between CCHFV strains from two hosts and *Hyalomma* and *Rhipicephalus* codon usage showed that both *Hyalomma*- and *Rhipicephalus*-isolated strains display higher adaptation to use codons that are preferred by *Hyalomma* species. These results suggest that CCHFV genome has optimized their codon usage patterns to utilize the translational resources of *Hyalomma* species more efficiently than that of *Rhipicephalus* species. Higher selection pressure from *Hyalomma* spp. can affect the codon usage of CCHFV and that the evolution of codon usage in CCHFV has allowed it to use the translation resource of species of *Hyalomma* genus more efficiently. Higher genetic adaptation of CCHFV strains isolated from two vectors for using codons which are preferred by *Hyalomma* ticks can be explained by the fact that *Hyalomma* tick species are the main vectors of CCHFV. Contrarily, lower adaptation of CCHFV strains from two hosts to *Rhipicephalus* codon usage highlights that *Rhipicephalus* ticks are only occasional vectors or evolutionary new vectors that are used by CCHFV at areas where *Hyalomma* ticks are absent.

The previous study (Rahman et al. 2018) reported that CAI values of CCHFV polyprotein-coding regions in relation to different hosts such as *Homo sapiens*, *B. taurus* and *O. aries* were higher than to *Hyalomma* spp., suggesting the greatest adaptation of CCHFV was to *H. sapiens*, closely followed by *B. taurus*, *O. aries* and then *Hyalomma* spp. According to that study, selection pressure from the hosts (*H. sapiens*, *B. taurus*, *O. aries*) in shaping codon usage patterns of CCHFV differs from that *Hyalomma* spp. The results of our study were consistent with this prior study (Rahman et al. 2018) acquiring similar CAI values of CCHFV in relation to *Hyalomma*. Considering the previous (Rahman et al. 2018) and our study results, the replication of

CCHFV proteins could be more efficient within the potential hosts cells than within the vector cells.

5.2. Tick collection in Kosovo and Czechia and screening for CCHFV

CCHFV is considered the most geographically distributed tick-borne virus worldwide. It has Old World distribution and its presence is closely linked to *Hyalomma* spp. ticks. CCHFV is also present in Balkan Peninsula. Kosovo, in particular, is a highly endemic country for CCHF which has a significant fatality rate. The previous reports (Sherifi et al. 2014; Sherifi et al. 2018) are evidence of the presence of CCHFV in Kosovo. There are two strains which are circulating in the country. One is "Kosova Hoti" which belongs to Europe 1 (clade V) lineage and is highly pathogenic among humans. Its competent vector is considered to be *H. marginatum*. The second strain is AP92 which belongs to Europe 2 (clade VI) lineage and has mild or non-pathogenicity for humans. Its competent vector is assumed to be *R. bursa*.

Schuster et al. (2016) detected a high rate of seropositivity in livestock, being 23% in Albania and 49% in North Macedonia showing that CCHFV is circulating in ticks and animals in these countries. Circulation of CCHFV in ticks in Kosovo and Albania and the local outbreaks in humans every 3-5 years in southwestern Kosovo and North Albania are a strong indicator of CCHFV's "high pathogenic" strain exists in those regions and its relation with the abundance of *H. marginatum* ticks (Sherifi et al. 2018).

Kosovo has five hyper-endemic areas for CCHF that ticks frequently collected from those areas are tested positive for the virus. Although 152 tick samples were analysed in this study for the presence of CCHFV that were collected from cattle in hyper-endemic areas of Kosovo (Malishevë, Rahovec, and Suharekë), we could not detect viral RNA in any tick. Although relatively low amount of ticks were collected in our study, our results were surprising not to detect CCHFV, as the previous report showed an 8.6% CCHFV prevalence in ticks in Malishevë, where 243 ticks were tested and 21 were positive for CCHFV (Sherifi et al. 2014).

Ticks collected in Suharekë, another hyper-endemic region for CCHF, were also tested negative for the virus. As the amount of samples collected from Suharekë was very low, the probability of ticks being negative to CCHFV was expected in our results. Moreover, another study also showed that only one tick out of 199 *H. marginatum* ticks (0.5%) was tested positive (Sherifi et al. 2018) in Suharekë.

In these previous studies, all ticks that tested positive for CCHFV were collected during April-May-June. However, the collected and analysed ticks in our study were collected during August. The absence of CCHFV could be explained by the fact that the sampling was carried out in late summer. Furthermore, a relatively smaller sample size might have influenced the results. In the previous report, a total of 1285 ticks were collected in May and June in 2012 and only 40 of them were tested positive (Kurtesh et al. 2014). Moreover, in another study, a total of 646 ticks were collected during spring and early summer in 2014 and CCHFV could be detected in only 9 tick samples (Kurtesh et al. 2018). This highlights that CCHFV prevalence in ticks is closely related with high number of tick populations during April-May-June.

In endemic countries in the Northern hemisphere, CCHFV infection is usually reported in the spring/early summer, and seasonality of the outbreak might be related to tick abundance and activity (Saijo et al. 2010). In accordance with previous studies and our results, the prevalence of CCHFV in ticks decreases in late summer which can be influenced by decreasing tick populations and their activity. There is no reported study in which ticks are collected and tested positive for the virus during late summer in Kosovo so far. The results of this study, hence, provide valuable information with relation to CCHFV prevalence in ticks in Kosovo. Taken together, these results indicate that a higher number of ticks has to be investigated. Furthermore, collection of ticks in spring and/or early summer may increase the probability of detection of CCHFV.

Apart from the absence of CCHFV in ticks collected from Kosovo, the same result appeared for the ticks collected from Czechia. Previous studies have been reported that migratory birds play a role as long-distance transporters of ticks containing numerous pathogens (Elfving et al. 2010; Lindeborg et al. 2012). Many bird species are important tick hosts, which significantly contribute to the local maintenance of tick populations and to long-distance dispersal of several tick species. Immature stages of

ticks can attach to bird hosts during migration and if biotic and abiotic factors are favourable, they thereafter detach from birds in their breeding or stopover sites.

To date, there is no reported stable population of *Hyalomma* spp. ticks in Central Europe. However, due to the significant roles of migratory birds in transportation of immature stages of *Hyalomma* ticks into new geographical locations, *Hyalomma* tick species may be observed in the areas where the species are absent. In Central Europe, migratory birds occur between April and October and after the fall migration, the birds spend the rest of the year in the Afrotropics, where *Hyalomma* ticks breed (Capek et al. 2014). During spring migration, when migratory birds migrate from Africa and Mediterranean to southwards, the *Hyalomma* ticks are introduced into Central Europe. Many countries of Central Europe such as Germany, Hungary, Slovakia, and Austria reported *H. marginatum* occurrences and the spread of this tick species in different regions indicates a risk of spreading the virus to Central Europe.

Collection of ticks from migratory birds in Czechia coincided with their fall migration period from the north to the south. However, it is worth mentioning that the tick collection from birds in the autumn season was a preliminary experiment to learn tick collection on birds.

According to the results, it was not surprising to not detect the presence of CCHFV since all collected ticks were belong to *Ixodes* genus and they are not as a vector species for the virus. Tick collection from birds in Czechia was carried out during September-October, the period indicating their southward migration. Due the fact that ticks collected in late period of their migration routes, no tick species in *Hyalomma* genus could be observed. In order to increase the chance of collecting *Hyalomma* tick species, birds should be caught in earlier period of their migration from south to north. Furthermore, monitoring migratory birds carrying CCHFV-infected ticks might give a useful tool to disease-prevention authorities for predicting the potential emergence of new disease foci in Europe.

5.3. Habitat suitability of *H.marginatum* in Central Europe

H. marginatum is a tick species which has veterinary and public health importance as it can transmit various tick-borne pathogens aside from CCHFV such as Spotted Fever Rickettsia to humans, *Babesia caballi* and *Theileria equi* (piroplasmosis) to horses and *Theileria annulata* (tropical theileriosis) to bovines (EFSA 2010; Bakheit et al. 2012; Wallménius et al. 2014; Tirosh-Levy et al. 2021).

H. marginatum is the main vector species for CCHFV. The distribution of CCHFV is related with spreading of its vector species. Climate change impacts, transportation of larvae and nymphs of ticks through international animal trade, and migratory birds have a crucial role on geographic expansion of CCHFV.

H. marginatum is widely distributed in Africa, Asia, and southern and eastern Europe. However, some sporadic records of this tick species have been reported in Central Europe. *H. marginatum* was first reported in Germany in 2007 (Kampen et al. 2007). Although *H. marginatum* records in other countries of Central Europe, for instance in Hungary (Földvári et al. 2011), in Slovakia (Capek et al. 2014), and in Austria (Duscher et al. 2018) have been reported, these do not represent stable populations of *H. marginatum* in Central Europe.

H. marginatum ticks have an ecological plasticity that can support a wide range of temperature and humidity conditions (Bouattour et al. 1999). Stable populations of *H. marginatum* in Europe are restricted to the warm areas of the Mediterranean basin and are absent in Central Europe, likely due to environmental conditions.

A previous study anticipated significant distributional potential of *H. marginatum* and high environmental suitability across Southern, Western and Central Europe (Okely et al. 2020). In my thesis, we updated the ENM of *H. marginatum* to predict potential suitable habitats for their viability and potential risk areas for CCHFV across Europe. Our model improved the prediction and attempted to give a more reliable and detailed map of habitat suitability for *H. marginatum*.

Our model has both consistencies and inconsistencies to the prior model (Okely et al. 2020). Predictions of *H. marginatum* in our model cover wide ranges in Europe (Figure 20).

As observed in the previous ecological niche model (Okely et al. 2020), the highest environmental suitability for *H. marginatum* was predicted across Southern Europe, namely Greece, Italy, Portugal, and Spain.

In South-eastern part, the highest suitability was predicted across all Balkan countries. In contrast to the previous model (Okely et al. 2020), our model predicted broader high suitability particularly, in Albania, Croatia, Bosnia and Herzegovina, Kosovo, Montenegro, and Slovenia.

For Western Europe, our model has similar predictions to Okely's model (2020), particularly in France and the United Kingdom. One of the important findings in this study is the predictions of wide medium suitability in Belgium and the Netherlands which was underestimated by the study of Okely and colleagues (Okely et al. 2020).

In respect of *H. marginatum* distribution in Central Europe, the focus of this thesis, our model prediction differed from Okely's model (2020). The previous model prediction identified a wide range of medium habitat suitability across Central Europe but showing also extensive high suitability in all countries of Central Europe. Our model provided strong evidence for occurrence of medium suitable areas in Central Europe, which is consistent with Okely et al. (2020), but indicating very less high suitability across the region (Figure 21). Our model predicted high suitability for *H. marginatum* only in some areas in the southern part of Hungary followed by Slovakia and Austria.

Recently, the study that assessed seroprevalence of CCHF in Hungary showed seropositivity and Hungary can be considered as a potentially new geographical area in the distribution of CCHFV (Magyar et al. 2021). Moreover, some sporadic occurrence records of *H. marginatum* in other countries of Central Europe aside from Hungary shows that there can be potentially suitable habitats for breeding of *H. marginatum.*

In Czechia, many regions in the southern and northern parts of the country show medium suitability, although low suitability was predicted in many parts of the country. Our model depicts a wide range of medium probability of suitable conditions for *H. marginatum* across Usti nad Labem and South Moravian, followed by some areas in South Bohemian, Plzen, Moravian Silesian, and Zlin (Figure 21).

Even though *Hyalomma* ticks have relatively low mobility by themselves, they can be transported over great distances by their vertebrate hosts, particularly migratory birds and ungulates. Migratory birds have a key role in transportation of immature stages of *H. marginatum* ticks and can lead to the northward spread of *H. marginatum* and the establishment of permanent populations. During the spring migration of migratory birds from the south to the north, *Hyalomma* ticks are introduced to Central Europe. The spread of these ticks into new geographic locations can also cause the emergence of CCHFV. The presence of the virus, its vectors, reservoirs, and amplifying hosts are necessary for the emergence of CCHF, but suitable environmental conditions are also essential (Portillo et al. 2021).

Appropriate climatic and biotic conditions in the regions that our map depicted may provide a suitable environment for introduced *H. marginatum* ticks. Particularly, a trend towards a warmer climate may be more favourable for maintaining infected *H. marginatum* ticks. An expansion of climatically suitable habitats in Central Europe is expected in the near future for *H. marginatum*. In our modeling study, climate change scenarios have not been considered. As known from the climate scenarios developed by the MoE, even under optimistic scenarios, the temperature in Czechia is expected to continue to rise, with an increase of values between 0.9-3°C in 2050 (NPACCI in the CR). It indicates that climate change will unquestionably increase the winter temperatures, leading to the increase of the probability for overwintering of *H. marginatum*, and consequently increasing the risk of establishment of *H. marginatum* in some parts of the Czechia within 30 years. Therefore, modeling studies based on future scenarios should be implemented to better understand climate change impacts on *H. marginatum* dispersion.

The predicted map of our study has significant public health importance, and the results will provide valuable information for vector and disease surveillance. Furthermore, the model prediction should develop further to define potential changes in environmental suitability based on future climate scenarios.

6. Conclusions

The first part of this thesis suggested that analysis of codon usage bias of *Hyalomma*-isolated and *Rhipicephalus*-isolated CCHFV strains may improve the knowledge of understanding the evolution and genetic background of adaptation of CCHFV to its vector species. Our findings indicated that CCHFV strains isolated from *Hyalomma* and *Rhipicephalus* have significant differences in codon usage variations and patterns. Furthermore, this thesis highlighted that both *Hyalomma*- and *Rhipicephalus*-isolated CCHFV strains are more adapted to choose the codons that are preferred by species of *Hyalomma* genus. This research not only provided the knowledge about the variation in CCHFV codon usage patterns in relation to their two vectors but also contributed to analysing the factors that influence adaptation of virus to the hosts.

The second part of this thesis provided important information regarding CCHFV prevalence in ticks in Kosovo. The results from this part clearly showed that CCHFV prevalence is related with tick abundance and activity which is completely connected with seasonality.

The final part of this thesis indicated the potential distribution of *H. marginatum* as the main vector species of CCHFV in Europe. The map has significant public health importance and anticipated probabilities of *H. marginatum* occurrence across areas that might be at risk of CCHFV occurrence. The results will provide valuable information for vector and disease surveillance.

7. References

Abudurexiti A, Adkins S, Alioto D, Alkhovsky SV, Avsic-Zupanc T, Ballinger MJ. 2019. Taxonomy of the order Bunyavirales: update 2019. Archives of Virology **164**(7):1949-65.

Ahmeti S, Kutllovci M, Bajrami M. 1996. Crimean Congo haemorrhagic fever in Kosova during 1995. Praxis Medica **39**:11-6.

Akashi H. 1997. Codon bias evolution in Drosophila. Population genetics of mutation-selection drift. Gene **205**(1-2):269-78.

Akuffo R, Brandful J, Zayed A, Adjei A, Watany N, Fahmy N, Hughes R, Doman B, Voegborlo S, Aziati D. 2016. Crimean-Congo haemorrhagic fever virus in livestock ticks and animal handler seroprevalence at an abattoir in Ghana. BMC Infectious Diseases **16**:324.

Al-Abri SS, Abaidani IA, Fazlalipour M. 2017. Current status of Crimean-Congo haemorrhagic fever in the World Health Organization Eastern Mediterranean Region: issues, challenges, and future directions. International Journal of Infectious Diseases **58**:82-89.

Alexaki A, Kames J, Holcomb DD, Athey J, Santana-Quintero LV, Lam PVN, Hamasaki-Katagiri N, Osipova E, Simonyan V, Bar H, Komar AA, Kimchi-Sarfaty C. 2019. Codon and Codon-Pair Usage Tables (CoCoPUTs): Facilitating Genetic Variation Analyses and Recombinant Gene Design. Journal of Molecular Biology **431**(13):2434-2441.

Altamura LA, Bertolotti-Ciarlet A, Teigler J, Paragas J, Schmaljohn CS, Doms RW. 2007. Identification of a novel C-terminal cleavage of Crimean-Congo haemorrhagic fever virus PreGN that leads to generation of an NSM protein. Journal of Virology **81**:6632-6642.

Anagnostou V, Papa A. 2009. Evolution of Crimean-Congo Haemorrhagic Fever virus. Infection, Genetics and Evolution **9**(5):948-54.

Apanaskevich DA, Horak IG. 2008. The genus Hyalomma Koch, 1844: V. Reevaluation of the taxonomic rank of taxa comprising the H. (Euhyalomma)

marginatum Koch complex of species (Acari: Ixodidae) with re-description of all parasitic stages and notes on biology. International Journal of Acarology **34**:13-42.

Artsob H, Lindsay R. 2008. Arboviruses, in: Heggenhougen, H.K. (Kris) (Ed.), International Encyclopedia of Public Health. Academic Press, Oxford, pp. 219–226.

Avšič-Županc T. 2009. Epidemiology and current geographical distribution of Crimean-Congo haemorrhagic fever. Arbo-Zoonet Newsletter **2**:8-15.

Bakheit MA, Latif AA, Vatansever Z, Seitzer U, Ahmed J. 2012. The Huge Risks Due to Hyalomma Ticks. H. Melhorn, editor. Arthropods as Vectors of Emerging Diseases. Parasitology Research Monographs, vol **3**. Springer, Berlin, Heidelberg.

Bakir M, Ugurlu M, Dokuzoguz B, Bodur H, Tasyaran MA, Vahaboglu H. 2005. Turkish CCHF Study Group: Crimean-Congo haemorrhagic fever outbreak in Middle Anatolia: A multicentre study of clinical features and outcome measures. Journal of Medical Microbiology **54**:385-389.

Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modeling **222**:1810-1819.

Bente DA, Forrester NL, Watts DM, McAuley AJ, Whitehouse CA, Bray M. 2013. Crimean-Congo haemorrhagic fever: history, epidemiology, pathogenesis, clinical syndrome and genetic diversity. Antiviral Research **100**(1):159-189.

Bergeron E, Vincent MJ, Nichol ST. 2007. Crimean-Congo haemorrhagic fever virus glycoprotein processing by the endoprotease SKI-1/S1P is critical for virus infectivity. Journal of Virology **81**:13271-13276.

Beyer H. 2004. Hawth's analysis tools for ArcGIS. http://www.spatialecology.com/htools.

Bishop R, Musoke A, Morzaria S, Gardner M, Nene V. 2004. Theileria: intracellular protozoan parasites of wild and domestic ruminants transmitted by ixodid ticks. Parasitology **129**(S1):271-283.

Bouattour A, Darghouth MA, Daoud A. 1999. Distribution and ecology of ticks (Acari: Ixodidae) infesting livestock in Tunisia: an overview of eighth years field collections. Parassitologia. **41**(1):5-10.

Brown JL, Bennett JR, French CM. 2017. SDMtoolbox 2.0: the next generation python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ **5**:e4095.

Burns CC, Shaw J, Campagnoli R, Jorba J, Vincent A, Quay J, Kew O, Irol JV. 2006. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. Journal of Virology **80**:3259-3272.

Butt AM, Nasrullah I, Tong Y. 2014. Genome-wide analysis of codon usage and influencing factors in chikungunya viruses. PLoS One **9**:e90905.

Butt AM, Nasrullah I, Qamar R, Tong Y. 2016. Evolution of codon usage in Zika virus genomes is host and vector specific. Emerging Microbes & Infections **5**(10):e107.

Capek M, Literak I, Kocianova E, Sychra O, Najer T, Trnka A, Kverek P. 2014. Ticks of the Hyalomma marginatum complex transported by migratory birds into Central Europe. Ticks Tick Borne Diseases **5**(5):489-93.

Carroll SA, Bird BH, Rollin PE, Nichol ST. 2010. Ancient common ancestry of Crimean-Congo haemorrhagic fever virus. Molecular Phylogenetics and Evolution **55**:1103-1110.

Causey OR, Kemp GE, Madbouly MH, David-West TS. 1970. Congo virus from domestic livestock, African hedgehog, and arthropods in Nigeria. American Journal of Tropical Medicine and Hygiene **19**:846-850.

Chase JM, Leibold MA. 2003. Ecological Niches: Linking Classical and Contemporary Approaches. University Of Chicago Press, USA.

Chen H, Sun S, Norenburg JL, Sundberg P. 2014. Mutation and selection cause codon usage and bias in mitochondrial genomes of ribbon worms (Nemertea). PloS One **9**(1):e85631.

Chinikar S, Ghiasi SM, Moradi M, Goya MM, Shirzadi MR, Zeinali M. 2010. Geographical distribution and surveillance of Crimean-Congo haemorrhagic fever in Iran. Vector-Borne and Zoonotic Diseases **10**(7):705-708.

Chumakov MP. 1945. A new tick-borne virus disease - Crimean haemorrhagic fever [In Russian]. In: Sokolov AA, Chumakov MP, Kolachev AA (eds) Crimean Haemorrhagic Fever (Acute Infectious Capillary Toxicosis). Izd Otd Primorskoi Armii, Simferopol, pp 13-43.

Chumakov MP. 1974. On 30 years of investigation of Crimean haemorrhagic fever [Russian; English translation NAMRU3-T950]. Trudi Inst Polio Virusn Entsefalitov Akad Med Nauk **22**:5-18.

Cobos ME, Peterson AT, Barve N, Osorio-Olvera L. 2019. kuenm: an R package for detailed development of ecological niche models using Maxent. PeerJ **7**:e6281.

Comeron JM, Aguade M. 1998. An evaluation of measures of synonymous codon usage bias. Journal of Molecular Evolution **47**:268-274.

Costafreda MI, Pérez-Rodriguez FJ, D'Andrea L, Guix S, Ribes E, Bosch A, Pintó RM. 2014. Hepatitis A virus adaptation to cellular shutoff is driven by dynamic adjustments of codon usage and results in the selection of populations with altered capsids. Journal of Virology **88**:5029-5041.

Dantas-Torres F. 2010. Biology and ecology of the brown dog tick, *Rhipicephalus sanguineus*. Parasites Vectors **3**:26.

Davies CR, Jones LD, Nuttall PA. 1986. Experimental studies on the transmission cycle of Thogoto virus, a candidate orthomyxovirus, in *Rhipicephalus appendiculatus*. American Journal of Tropical Medicine and Hygiene **35**:1256-1262.

De Kok JB, d'Oliveira C, Jongejan F. 1993. Detection of the protozoan parasite *Theileria annulata* in *Hyalomma* ticks by the polymerase chain reaction. Experimental and Applied Acarology **17**:379-81.

Deyde VM, Khristova ML, Rollin PE, Ksiazek TG, Nichol ST. 2006. Crimean-Congo haemorrhagic fever virus genomics and global diversity. Journal of Virology, **80**:8834-8842.

Drake JW, Holland JJ. 1999. Mutation rates among RNA viruses. Proceedings of the National Academy of Sciences USA **96**(24):13910-3.

Duscher GG, Hodžić A, Hufnagl P, Wille-Piazzai W, Schötta AM, Markowicz MA, Estrada-Peña A, Stanek G, Allerberger F. 2018. Adult *Hyalomma marginatum* tick positive for *Rickettsia aeschlimannii* in Austria, October 2018. Eurosurveillance **23**(48):1800595.

Scientific Opinion on the Role of Tick Vectors in the Epidemiology of Crimean-Congo Haemorrhagic Fever and African Swine Fever in Eurasia. 2010. EFSA Journal, **8**(8), 1703.

Elfving K, Olsen B, Bergström S, Waldenström J, Lundkvist A, Sjöstedt A, Mejlon H, Nilsson K. 2010. Dissemination of spotted fever rickettsia agents in Europe by migrating birds. PLoS One **5**(1):e8572.

Endy TP. 2020. Chapter 37 - Viral Haemorrhagic Fevers. Pages 351-381 in Edward T. Ryan, David R. Hill, Tom Solomon, Naomi E. Aronson and Timothy P. Endy, editors. Hunter's Tropical Medicine and Emerging Infectious Diseases (Tenth Edition), Elsevier.

Ergonul O, Whitehouse CA. 2007. Crimean-Congo haemorrhagic fever: a global perspective. Amsterdam, The Netherlands, Springer.

Ergonul O, Celikbas A, Dokuzoguz B, Eren S, Baykam N, Esener H. 2004. Characteristics of patients with Crimean-Congo haemorrhagic fever in a recent outbreak in Turkey and impact of oral ribavirin therapy. Clinical Infectious Diseases **39**(2):284-7.

Ermolaeva MD. 2001. Synonymous codon usage in bacteria. Current Issues in Molecular Biology **3**(4):91-7.

Estrada-Peña A, Bouattour A, Camicas JL, Walker AR. 2004. Ticks of Domestic Animals in the Mediterranean Region: A Guide to Identification of Species. Zaragoza: University of Zaragoza.

Estrada-Peña A, Palomar AM, Santibáñez P, Sánchez N, Habela MA, Portillo A, Romero L, Oteo JA. 2012. Crimean-Congo Haemorrhagic Fever Virus in Ticks, Southwestern Europe, 2010. Emerging Infectious Diseases **18**:179-180.

Estrada-Peña A, Alexander N, Wint GW. 2016. Perspectives on modeling the distribution of ticks for large areas: so far so good? Parasites & Vectors **9**:179.

European Centre for Disease Prevention and Control. 2019. Crimean-Congo haemorrhagic fever. In: ECDC. Annual epidemiological report for 2018. Stockholm: ECDC.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. International Journal of Climatology **37**(12):4302-4315.

Földvári G, Rigó K, Jablonszky M, Biró N, Majoros G, Molnár V, Tóth M. 2011. Ticks and the city: ectoparasites of the Northern white-breasted hedgehog (*Erinaceus roumanicus*) in an urban park. Ticks and Tick Borne Diseases **2**(4):231-4.

Gargili A, Midilli K, Ergonul O, Ergin S, Alp HG, Vatansever Z, Iyisan S, Cerit C, Yilmaz G, Altas K, Estrada-Peña A. 2011. Crimean-Congo haemorrhagic fever in European part of Turkey: genetic analysis of the virus strains from ticks and a seroepidemiological study in humans. Vector Borne Zoonotic Diseases **11**:747-52.

Gonzalez JP, Cornet JP, Wilson M, Camicas JL. 1991. Crimean-Congo haemorrhagic fever virus replication in adult *Hyalomma truncatum* and *Amblyomma variegatum* ticks. Research in Virology **142**:483-488.

Gonzalez-Scarano F, Nathanson N. 1996. Bunyaviridae. Pages 1473-1504 in B. Fields, D. Knipe, and P. Howley, editors. Fields Virology. Lippincott-Raven, Philadelphia, PA.

Grashchenkov N. 1945. Investigations of Etiology, Pathogenesis, and Clinical Symptomatology of Crimean Haemorrhagic Fever. (In English: NAMRU3-T1189). Reports 1944 Sci Investig Inst Neurol Akad Med Nauk SSSR; Moskva 100-7.

Gray JS, Dautel H, Estrada-Peña A, Kahl O, Lindgren E. 2009. Effects of climate change on ticks and tick-borne diseases in europe. Interdisciplinary Perspectives on Infectious Diseases **2009**:593232.

Greenacre MJ. 1984. Theory and Applications of Correspondence Analysis. Academic Press, London.

Guerra M, Walker E, Jones C, Paskewitz S, Cortinas MR, Stancil A, Beck L, Bobo M, Kitron U. 2002. Predicting the risk of Lyme disease: habitat suitability for

Ixodes scapularis in the north central United States. Emerging Infectious Diseases **8**:289.

Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology: Ecological Modeling **135**(2-3):147-186.

Guglielmone AA, Robbins RG, Apanaskevich DA, Petney TN, Estrada-Peña A, Horak IG. 2014. The hard ticks of the world (Acari: Ixodida: Ixodidae). Berlin, Springer.

Hall-Mendelin S, Craig SB, Hall RA, O'Donoghue P, Atwell RB, Tulsiani SM, Graham GC. 2011. Tick paralysis in Australia caused by *Ixodes holocyclus* Neumann. Annals of Tropical Medicine and Parasitology **105**(2):95-106.

Hewson R, Gmyl A, Gmyl L, Smirnova SE, Karganova G, Jamil B, Hasan R, Chamberlain J, Clegg C. 2004. Evidence of segment reassortment in Crimean-Congo haemorrhagic fever virus. Journal of General Virology **85**:3059-3070.

Hillyard PD. 1996. Ticks of North-west Europe: Keys and Notes for Identification of the Species. Linnean Society of London and the Estuarine and Coastal Sciences Association.

Hoogstraal H, Kaiser MN, Traylor MA, Gaber S, Guindy E. 1961. Ticks (Ixodoidea) on birds migrating from Africa to Europe and Asia. Bulletin of the World Health Organization **24**(2):197-212.

Hoogstraal H. 1979. The epidemiology of tick-borne Crimean-Congo haemorrhagic fever in Asia, Europe, and Africa. Journal of Medical Entomology **15**(4):307-417.

International Committee on Taxonomy of Viruses. 2011. ICTV 9th Report, Negative Sense RNA Viruses (2011), Bunyaviridae.

Irwin PJ. 2009. Canine babesiosis: from molecular taxonomy to control. Parasitology Vectors **2**(S4).

Jameson LJ, Ramadani N, Medlock JM. 2012. Possible drivers of Crimean-Congo haemorrhagic fever virus transmission in Kosova. Vector Borne Zoonotic Diseases **12**(9):753-7.

Jenkins GM, Holmes EC. 2003. The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Research **92**:1-7.

Jongejan F, Uilenberg G. 2004. The global importance of ticks. Parasitology **129**(S1):S3-S14.

Kadanali A, Erol S, Ozkurt Z, Ozden K. 2009. Epidemiological risk factors for Crimean-Congo haemorrhagic fever patients. Turkish Journal of Medical Sciences **39**(13):829-832.

Kahl O, Gern L, Eisen L, Lane RS. 2002. Ecological research on *Borrelia burgdorferi* sensu lato: terminology and some methodological pitfalls. Pages 29-46 in Gray O. Kahl RS. Lane, and G. Stanek, editors. Lyme Borreliosis: Biology, Epidemiology and Control. Wallingford, CABI.

Kaiser MN, Hoogstraal H. 1974. Ticks (Ixodoidea) on migrating birds in Cyprus, fall 1967 and spring 1968, and epidemiological considerations. Bulletin of Entomological Research **64**:97-110.

Kamal M, Kenawy MA, Rady MH, Khaled AS, Samy AM. 2018. Mapping the global potential distributions of two arboviral vectors *Aedes aegypti* and *Ae. albopictus* under changing climate. PLoS One **13**:e0210122.

Kampen H, Poltz W, Hartelt K, Wölfel R, Faulde M. 2007. Detection of a questing *Hyalomma marginatum marginatum* adult female (Acari, Ixodidae) in southern Germany. Experimental and Applied Acarology **43**(3):227-31.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution **30**(4):772-780.

Kaufman WR, Nuttall PA. 2003. *Rhipicephalus appendiculatus* (Acari: Ixodidae): dynamics of Thogoto virus infection in female ticks during feeding on guinea pigs. Experimental Parasitology **104**:20-25.

Kazimírová M, Štibrániová I. 2013. Tick salivary compounds: their role in modulation of host defences and pathogen transmission. Front Cellular Infection Microbiology **3**:43.

Knight RD, Freeland SJ, Landweber LF. 2001. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. Genome Biology **2**(4):RESEARCH0010.

Kosovo Institute for Nature Portection. 2015. State of Nature Report 2010-2014. Prishtine, Kosovo.

Ministry of the Environment of Czech Republic. 2017. National Programme to Abate the Climate Change Impacts in the Czech Republic. Prague, Czech Republic.

Kumar N, Bera BC, Greenbaum BD, Bhatia S, Sood R, Selvaraj P, Anand T, Tripathi BN, Virmani N. 2016. Revelation of influencing factors in overall codon usage bias of equine influenza viruses. PLoS One **11**:e0154376.

Leech SL. 2015. Investigation into the vector competence of Ixodes ricinus ticks to Hazara virus and Crimean-Congo Haemorrhagic Fever virus. PhD thesis, London School of Hygiene & Tropical Medicine.

Letunic I, Bork P. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Research **39**(Web Server issue):W475-8.

Lindeborg M, Barboutis C, Ehrenborg C, Fransson T, Jaenson TG, Lindgren PE, Lundkvist A, Nyström F, Salaneck E, Waldenström J, Olsen B. 2012. Migratory birds, ticks, and crimean-congo haemorrhagic fever virus. Emerging Infectious Diseases **18**(12):2095-7.

Magnarelli LA. 2009. Global Importance of Ticks and Associated Infectious Disease Agents. Clinical Microbiology Newsletter **31**(5):33-37.

Magyar N, Kis Z, Barabás É, Nagy A, Henczkó J, Damjanova I, Takács M, Pályi B. 2021. New geographical area on the map of Crimean-Congo haemorrhagic fever virus: First serological evidence in the Hungarian population. Ticks Tick Borne Diseases **12**(1):101555.

Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E. 2006. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. Journal of Virology **80**:9687-9696.

Mustafa ML, Ayazi E, Mohareb E, Yingst S, Zayed A, Rossi C. 2011. Crimean-Congo haemorrhagic fever, Afghanistan, 2009. Emerging Infectious Diseases **17**(10):1940-1941.

Nasrullah I, Butt AM, Tahir S, Idrees M, Tong Y. 2015. Genomic analysis of codon usage shows influence of mutation pressure, natural selection, and host features on Marburg virus evolution. BMC Ecology and Evolution **15**:174.

Negredo A, de la Calle-Prieto F, Palencia-Herrejón E, Mora-Rillo M, Astray-Mochales J, Sánchez-Seco MP, Bermejo Lopez E, Menárguez J, Fernández-Cruz A, Sánchez-Artola B, Keough-Delgado E, Ramírez de Arellano E, Lasala F, Milla J, Fraile JL, Ordobás Gavín M, Martinez de la Gándara A, López Perez L, Diaz-Diaz D, López-García MA, Delgado-Jimenez P, Martín-Quirós A, Trigo E, Figueira JC, Manzanares J, Rodriguez-Baena E, Garcia-Comas L, Rodríguez-Fraga O, García-Arenzana N, Fernández-Díaz MV, Cornejo VM, Emmerich P, Schmidt-Chanasit J, Arribas JR; Crimean Congo Haemorrhagic Fever@Madrid Working Group. 2017. Autochthonous Crimean-Congo Haemorrhagic Fever in Spain. The New England Journal of Medicine **377**(2):154-161.

Novembre JA. 2002. Accounting for background nucleotide composition when measuring codon usage bias. Molecular Biology and Evolution **19**:1390-1394.

Nuttall PA, Labuda M. 2004. Tick-host interactions: saliva-activated transmission. Parasitology **129**:S177S189.

Okely M, Anan R, Gad-Allah S, Samy AM. 2020. Mapping the environmental suitability of etiological agent and tick vectors of Crimean-Congo haemorrhagic fever. Acta Tropica **203**:105319.

Papa A, Ma B, Kouidou S, Tang Q, Hang C, Antoniadis A. 2002. Genetic characterization of the M RNA segment of Crimean Congo haemorrhagic fever virus strains, China. Emerging Infectious Diseases **8**:50-53.

Papa A, Tzala E, Maltezou HC. 2011. Crimean-Congo haemorrhagic fever virus, northeastern Greece. Emerging Infectious Diseases **17**:141-143.

Papa A, Papadopoulou E, Tsioka K, Kontana A, Pappa S, Melidou A, Giadinis ND. 2018. Isolation and whole-genome sequencing of a Crimean-Congo haemorrhagic fever virus strain, Greece. Ticks and Tick-Borne Diseases **9**:788-791.

Papadopoulos O, Koptopoulos G. 1978. Isolation of Crimean-Congo haemorrhagic fever (C-CHF) virus from *Rhipicephalus bursa* ticks in Greece. Acta Microbiologica Hellenica **23**:20-28.

Peden JF. 1999. Analysis of Codon Usage. Doctoral dissertation. Department of Genetics, University of Nottingham.

Peterson AT, Papeş M, Soberón J. 2015. Mechanistic and Correlative Models of Ecological Niches. European Journal of Ecology **1**(2).

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modeling **190**:231-259.

Pigott DM, Deshpande A, Letourneau I. 2017. Local, national, and regional viral haemorrhagic fever pandemic potential in Africa: a multistage analysis. Lancet **390**(10113):2662-72.

Plotkin JB, Dushoff J, Desai MM, Fraser HB. 2006. Codon usage and selection on proteins. Journal of Molecular Evolution **63**(5):635-53.

Plotkin JB, Kudla G. 2011. Synonymous but not the same: the causes and consequences of codon bias. Nature Reviews Genetics **12**:32-42.

Portillo A, Palomar AM, Santibáñez P, Oteo JA. 2021. Epidemiological Aspects of Crimean-Congo Haemorrhagic Fever in Western Europe: What about the Future? Microorganisms **9**(3):649.

Pshenichnaya NY, Leblebicioglu H, Bozkurt I, Sannikova IV, Abuova GN, Zhuravlev AS, Barut S, Shermetova MB, Fletcher TE. 2017. Crimean-Congo haemorrhagic fever in pregnancy: A systematic review and case series from Russia, Kazakhstan and Turkey. International Journal of Infectious Diseases **58**:58-64.

Puigbò P, Bravo IG, Garcia-Vallvé S. 2008a. CAIcal: a combined set of tools to assess codon usage adaptation. Biology Direct **3**:38.

Puigbò P, Bravo IG, Garcia-Vallvé S. 2008b. E-CAI: a novel server to estimate an expected value of codon adaptation index (eCAI). BMC Bioinformatics **9**:65.

Quax TE, Claassens NJ, Söll D, van der Oost J. 2015. Codon Bias as a Means to Fine-Tune Gene Expression. Molecular Cell **59**(2):149-61.

Rahman SU, Yao X, Li X, Chen D, Tao S. 2018. Analysis of codon usage bias of Crimean-Congo haemorrhagic fever virus and its adaptation to hosts. Infection, Genetics and Evolution **58**:1-16.

Randolph SE. 2000. Ticks and tick-borne disease systems in space and from space. Advances in Parasitology **47**:217-243.

Saijo M, Morikawa S, Kurane I. 2010. Recent progress in the treatment of Crimean-Congo haemorrhagic fever and future perspectives. Future Virology **5**(6):801-809.Sanchez AJ, Vincent MJ, Nichol ST. 2002. Characterization of the glycoproteins of Crimean-Congo haemorrhagic fever virus. Journal of Virology **76**:7263-7275.

Sargianou M, Panos G, Tsatsaris A, Gogos C, Papa A. 2013. Crimean-Congo haemorrhagic fever: seroprevalence and risk factors among humans in Achaia, western Greece. International Journal of Infectious Diseases **17**(12):e1160ee1165.

Schuster I, Mertens M, Mrenoshki S, Staubach C, Mertens C, Brüning F, Wernike K, Hechinger S, Berxholi K, Mitrov D, Groschup MH. 2016. Sheep and goats as indicator animals for the circulation of CCHFV in the environment. Experimental and Applied Acarology **68**(3):337-46.

Sentausa E, El Karkouri K, Michelle C, Raoult D, Fournier PE. 2014. Draft genome sequence of *Rickettsia aeschlimannii*, associated with *Hyalomma marginatum* ticks. Genome Announc. **2**:e00666-14.

Sharifi-Mood B, Metanat M, Alavi-Naini R. 2014. Prevalence of Crimean-Congo haemorrhagic fever among high risk human groups. International Journal of High Risk Behaviors and Addiction **3**(1):e11520.

Sharp PM, Li WH. 1986. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for “rare” codons. Nucleic Acids Research **14**:7737-7749.

Sharp PM, Li WH. 1987. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Research **15**:1281-95.

Sharp PM, Stenico M, Peden JF, Lloyd AT. 1993. Codon usage: mutational bias, translational selection, or both? Biochemical Society Transactions **21**:835-841.

Sharp PM, Emery LR, Zeng K. 2010. Forces that Influence the Evolution of Codon Bias. Philos Trans R Soc Lond B Biol Sci **365**(1544):1203-12.

Sherifi K, Cadar D, Muji S, Robaj A, Ahmeti S, Jakupi X, Emmerich P, Krüger A. 2014. Crimean-Congo haemorrhagic fever virus clades V and VI (Europe 1 and 2) in ticks in Kosovo, 2012. PLOS Neglected Tropical Diseases **8**(9):e3168.

Sherifi K, Rexhepi A, Berxholi K, Mehmedi B, Gecaj RM, Hoxha Z, Joachim A, Duscher GG. 2018. Crimean-Congo Haemorrhagic Fever Virus and *Borrelia burgdorferi* sensu lato in Ticks from Kosovo and Albania. Frontiers in Veterinary Science **5**:38.

Shi J, Hu Z, Deng F, Shen S. 2018. Tick-Borne Viruses. Virologica Sinica **33**(1):21-43.

Shields DC, Sharp PM, Higgins DG, Wright F. 1988. "Silent" sites in Drosophila genes are not neutral: evidence of selection among synonymous codons. Molecular Biology and Evolution **5**:704-16.

Singh NK, Tyagi A, Kaur R, Verma R, Gupta PK. 2016. Characterization of codon usage pattern and influencing factors in Japanese encephalitis virus. Virus Research **221**:58-65.

Sonenshine DE, Roe RM. 2014. Biology of Ticks. Oxford University Press.

Spengler JR, Estrada-Peña A, Garrison AR, Schmaljohn C, Spiropoulou CF, Bergeron É, Bente DA. 2016. A chronological review of experimental infection studies of the role of wild animals and livestock in the maintenance and transmission of Crimean-Congo haemorrhagic fever virus. Antiviral Research **135**:31-47.

Spickler AR. 2019. Crimean-Congo Haemorrhagic Fever. Retrieved from https://www.cfsph.iastate.edu/Factsheets/pdfs/crimean_congo_haemorrhagic_fever.pdf.

Stenico M, Lloyd AT, Sharp PM. 1994. Codon usage in Caenorhabditis elegans: delineation of translational selection and mutational biases. Nucleic Acids Research **22**:2437-46.

Phillips SJ, Dudík M, Schapire RE. Maxent software for modeling species niches and distributions (Version 3.4.1). Available from

http://biodiversityinformatics.amnh.org/open_source/maxent/. (Accessed on 15 February, 2020).

Samy AM, Elaagip AH, Kenawy MA, Ayres CF, Peterson AT, Soliman DE. 2016. Climate change influences on the global potential distribution of the mosquito *Culex quinquefasciatus*, vector of West Nile virus and lymphatic filariasis. PLoS One **11**:e0163863.

Samy AM, Peterson AT. 2016. Climate change influences on the global potential distribution of bluetongue virus. PLoS One **11**:e0150489.

Sumrandee C, Hirunkanokpun S, Doornbos K, Kitthawee S, Baimai V, Grubhoffer L, Trinachartvanit W, Ahantarig A. 2014. Molecular detection of Rickettsia species in Amblyomma ticks collected from snakes in Thailand. Ticks Tick-Borne Diseases **5**(6):632-40.

Sun S, Dai X, Aishan M, Wang X, Meng W, Feng C, Zhang F, Hang C, Hu Z, Zhang Y. 2009. Epidemiology and phylogenetic analysis of Crimean-Congo haemorrhagic fever viruses in Xeinjiang, China. J Clin Microbiol, **47**:2536-2543.

Smith RL, Schnack JA, Schaefe EF, Kehr AI. 2008. Ticks, *Amblyomma rotundatum* (Acari: ixodidae), on toads, *Chaunus schneideri* and *Chaunus granulosus* (Anura: bufonidae), in northern Argentina. Journal of Parasitology **94**(2):560-562.

Tirosh-Levy S, Mazuz ML, Savitsky I, Pinkas D, Gottlieb Y, Steinman A. 2021. Serological and Molecular Prevalence of Babesia caballi in Apparently Healthy Horses in Israel. Pathogens. **10**(4):445.

Vanhomwegen J, Alves MJ, Zupanc TA, Bino S, Chinikar S, Karlberg H, Korukluoglu G, Korva M, Mardani M, Mirazimi A. 2012. Diagnostic assays for Crimean-Congo haemorrhagic fever. Emerging Infectious Diseases **18**:1958-1965.

van Hemert F, Berkhout B. 2016. Nucleotide composition of the Zika virus RNA genome and its codon usage. Virology Journal **13**:95.

van Hemert F, van der Kuyl AC, Berkhout B. 2016. Impact of the biased nucleotide composition of viral RNA genomes on RNA structure and codon usage. Journal of General Virology **97**(10):2608-2619.

Varma MRG. 1993. Ticks and mites (Acari). Medical Insects and Arachnids. Springer, Dordrecht 97-658.

Vesenjak-Hirjan J, Punda-Polić V, Dobe M. 1991. Geographical distribution of arboviruses in Yugoslavia. J Hyg Epidemiol Microbiol Immunology **35**:129-40.

Walker AR, Bouattour A, Camicas JL, Estrada-Peña A, Horak IG, Latif AA, Pegram P, Preston P. 2003. Ticks of domestic animals in Africa: guide to identification of species. Edinburgh, Bioscience Reports.

Wallménius K, Barboutis C, Fransson T, Jaenson TGT, Lindgren PE, Nyström F, Olsen B, Salaneck E, Nilsson K. 2014. Spotted fever Rickettsia species in Hyalomma and Ixodes ticks infesting migratory birds in the European Mediterranean area. Parasites Vectors **7**:318.

Walter CT, Barr JN. 2011. Recent advances in the molecular and cellular biology of bunyaviruses. Journal of General Virology **92**:2467-2484.

Whitehouse CA. 2004. Crimean-Congo haemorrhagic fever. Antiviral Research **64**:145-160.

WHO: Crimean Congo haemorrhagic fever: Fact Sheet No. 208 January, 2013. Available from https://www.who.int/news-room/fact-sheets/detail/crimean-congo-haemorrhagic-fever. Accessed March, 2021.

Wang M, Liu YS, Zhou JH, Chen HT, Ma LN, Ding YZ, Liu WQ, Gu YX, Zhang J. 2011. Analysis of codon usage in Newcastle disease virus. Virus Genes **42**:245-253.

Warren DL, Seifert SN. 2011. Ecological niche modeling in Maxent: The importance of model complexity and the performance of model election criteria. Ecological Applications **21**: 335-342.

Wong EHM, Smith DK, Rabadan R, Peiris M, Poon LLM. 2010. Codon usage bias and the evolution of influenza A viruses. Codon usage biases of influenza virus. BMC Ecology and Evolution**10**:253.

Wright F. 1990. The “effective number of codons” used in a gene. Gene **87**:23-29.

Yesilbag K, Aydin L, Dincer E, Alpay G, Girisgin AO, Tuncer P, Ozkul A. 2013. Tick survey and detection of Crimean-Congo haemorrhagic fever virus in tick species from a non-endemic area, South Marmara region, Turkey. Experimental and Applied Acarology **60**:253-61.

Zeller HG, Cornet JP, Camicas JL. 1994. Crimean-Congo haemorrhagic fever virus infection in birds: field investigations in Senegal. Research in Virology **145**(2):105e109.

Zhou Z, Deng F, Han N, Wang H, Sun S, Zhang Y, Hu Z, Rayner S. 2013. Reassortment and migration analysis of Crimean-Congo haemorrhagic fever virus. Journal of General Virology **94**:2536-2548.

Zivcec M, Safronetz D, Feldmann H. 2013. Animal models of tick-borne haemorrhagic Fever viruses. Pathogens **2**(2):402-421.

Appendices

List of the Appendices:

Appendix 1: Viral RNA extraction protocol.

Protocol: Purification of Viral RNA (Spin Protocol)

This protocol is for purification of viral RNA from 140 µl plasma, serum, urine, cell-culture media or cell-free body fluids using a microcentrifuge. For automated purification of viral RNA using the QIAamp Viral RNA Mini Kit on QIAcube Connect, refer to the *QIAcube Connect User Manual* and the relevant protocol sheet.

Larger starting volumes, up to 560 µl (in multiples of 140 µl), can be processed by increasing the initial volumes proportionally and loading the QIAamp Mini column multiple times, as described below in the protocol. Some samples with very low viral titers should be concentrated before the purification procedure; see "Protocol: Sample Concentration" (page 35).

Alternatively, larger sample volumes can be processed using one of the following kits, which provide simultaneous purification of viral DNA and RNA:

- QIAamp MinElute® Spin Kit* – 200 µl
- QIAamp MinElute Vacuum Kit – 500 µl
- QIAamp UltraSens® Virus Kit – 1000 µl

Important points before starting

- Read "Important Notes" (page 16) before starting the protocol.
- All centrifugation steps are carried out at room temperature.

Things to do before starting

- Equilibrate samples to room temperature.
- Equilibrate Buffer AVE to room temperature for elution in step 11.
- Check that Buffer AW1 and Buffer AW2 have been prepared according to the instructions on page 19.
- Add carrier RNA reconstituted in Buffer AVE to Buffer AVL according to instructions on page 17.

Procedure

1. Pipet 560 µl prepared Buffer AVL containing carrier RNA into a 1.5 ml microcentrifuge tube.

   Note: If the sample volume is larger than 140 µl, increase the amount of Buffer AVL–carrier RNA proportionally (e.g., a 280 µl sample will require 1120 µl Buffer AVL–carrier RNA) and use a larger tube.

2. Add 140 µl plasma, serum, urine, cell-culture supernatant or cell-free body fluid to the Buffer AVL–carrier RNA in the microcentrifuge tube. Mix by pulse-vortexing for 15 s.

   Note: To ensure efficient lysis, it is essential that the sample is mixed thoroughly with Buffer AVL to yield a homogeneous solution. Frozen samples that have only been thawed once can also be used.

3. Incubate at room temperature for 10 min.

   Note: Viral particle lysis is complete after lysis for 10 min at room temperature. Longer incubation times have no effect on the yield or quality of the purified RNA.

4. Briefly centrifuge the tube to remove drops from the inside of the lid.
5. Add 560 µl ethanol (96–100%) to the sample, and mix by pulse-vortexing for 15 s. After mixing, briefly centrifuge the tube to remove drops from inside the lid.

   Note: Use only ethanol, since other alcohols may result in reduced RNA yield and purity. Do not use denatured alcohol, which contains other substances, such as methanol or methylethylketone. If the sample volume is greater than 140 µl, increase the amount of ethanol proportionally (e.g., a 280 µl sample will require 1120 µl ethanol). To ensure efficient binding, it is essential that the sample is mixed thoroughly with the ethanol to yield a homogeneous solution.

6. Carefully apply 630 µl of the solution from step 5 to the QIAamp Mini column (in a 2 ml collection tube) without wetting the rim. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini column into a clean 2 ml collection tube, and discard the tube containing the filtrate.

   Note: Close each spin column to avoid cross-contamination during centrifugation.

   Note: Centrifugation is performed at 6000 x *g* (8000 rpm) to limit microcentrifuge noise. Centrifugation at full speed will not affect the yield or purity of the viral RNA. If the solution has not completely passed through the membrane, centrifuge again at a higher speed until all of the solution has passed through.

7. Carefully open the QIAamp Mini column, and repeat step 6. If the sample volume was greater than 140 µl, repeat this step until all of the lysate has been loaded onto the spin column.
8. Carefully open the QIAamp Mini column, and add 500 µl Buffer AW1. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini column in a clean 2 ml collection tube (provided), and discard the tube containing the filtrate.

   Note: It is not necessary to increase the volume of Buffer AW1 even if the original sample volume was larger than 140 µl.

9. Carefully open the QIAamp Mini column, and add 500 µl Buffer AW2. Close the cap and centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 3 min. Continue directly with step 11; or to eliminate possible Buffer AW2 carryover, perform step 10 and then continue with step 11.

   Note: Residual Buffer AW2 in the eluate may cause problems in downstream applications. Some centrifuge rotors may vibrate upon deceleration, resulting in flow-through, containing Buffer AW2, contacting the QIAamp Mini column.

   Removing the QIAamp Mini column and collection tube from the rotor may also cause flow-through to come into contact with the QIAamp Mini column. In these cases, the optional step 10 should be performed.

10. **Recommended**: Place the QIAamp Mini column in a new 2 ml collection tube (not provided), and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 min.
11. Place the QIAamp Mini column in a clean 1.5 ml microcentrifuge tube (not provided). Discard the old collection tube containing the filtrate. Carefully open the QIAamp Mini column and add 60 µl Buffer AVE equilibrated to room temperature. Close the cap, and incubate at room temperature for 1 min.
12. Centrifuge at 6000 x *g* (8000 rpm) for 1 min.

    A single elution with 60 µl Buffer AVE is sufficient to elute at least 90% of the viral RNA from the QIAamp Mini column. Performing a double elution using 2 x 40 µl Buffer AVE will increase yield by up to 10%. Elution with volumes of less than 30 µl will lead to reduced yields and will not increase the final concentration of RNA in the eluate.

    Viral RNA is stable for up to 1 year when stored at –30 to –15°C or at –90 to –65°C.

Appendix 2: The strain name, accession number, origin, isolation host and collection date of polyprotein-coding region of each CCHFV isolates used in this study.

	STRAIN	STRAIN NAME	LOCALITY	HOST	DATE OF COLLECTION
S SEGMENT	DQ076415.1	SPU128/81/7	Uganda	*Hyalomma* spp.	1981
	KF793333.1	Daral 2012	Mali	*Hyalomma* spp.	2012
	KU707899.1	CJH	Iran	*H.anatolicum*	2015
	KY484036.1	IbAr10200	Nigeria	*H.excavatum*	1996
	KY484037.1	JD-206	Pakistan	*H.anatolicum*	1965
	AF481799.1	Uzbek/TI10145	Uzbekistan	*H.asiaticum*	1985
	MG659724.1	WJQ16206	China	*H.asiaticum*	2017
	MG659727.1	FK16116	China	*H.asiaticum*	2017
	MH688497.1	YL16204	China	*H.asiaticum*	2016
	KU707898.1	CHH	Iran	*H.dromedarii*	2015
	KU707900.1	GRH	Iran	*H.dromedarii*	2015
	MF547415.1	Caceres 2014	Spain	*H. lusitanicum*	2014
	AY277672.1	ROS/TI28044	Russia	*H.marginatum*	2000
	KR814833.1	59-TK-2012	Russia	*H.marginatum*	2012
	KR814834.1	128-TK-2012	Russia	*H.marginatum*	2012
	KY484031.1	HY-13	China	*H.marginatum*	1968
	KY484044.1	SPU 128/81	Uganda	*H.marginatum*	1981
	DQ211641.1	ArD39554	Mauritania	*H.rufipes*	1984
	MF511219.1	SPUD8_81_7_813051_S	South Africa	*H.rufipes*	1981
	DQ211639.1	ArD8194	Senegal	*H.truncatum*	1969
	KY484027.1	DAK8194	Senegal	*H.truncatum*	1969
	DQ211638.1	AP92	Greece	*R.bursa*	1975
	MG516211.1	Pentalofos-Greece-2015	Greece	*R.bursa*	2015
	U04958.1	AP92	Greece	*R.bursa*	1994

	STRAIN	STRAIN NAME	LOCALITY	HOST	DATE OF COLLECTION
M SEGMENT	DQ157174.1	SPU128/81/7	South Africa	*Hyalomma* spp.	1981
	KY484038.1	JD-206	Pakistan	*H.anatolicum*	1965
	MG659726.1	FK16116	China	*H.asiaticum*	2017
	MG659723.1	WJQ16206	China	*H.asiaticum*	2017
	MH688498.1	YL16204	China	*H.asiaticum*	2016
	NC_005300.2	IbAr10201	Nigeria	*H.excavatum*	1966
	KY484035.1	IbAr10200	Nigeria	*H.excavatum*	1996
	AF467768.2	IbAr10200	Nigeria	*H.excavatum*	2002
	MF547416.1	Caceres 2014	Spain	*H. lusitanicum*	2014
	KY484045.1	SPU 128/81	Uganda	*H.marginatum*	1981
	KY484032.1	HY-13	China	*H.marginatum*	1968
	AY900145.1	Hy13	China	*H.marginatum*	2005
	AY900141.1	SPU128/84	South Africa	*H.marginatum*	2005
	AY179961.1	VLG/TI29414	Russia	*H.marginatum*	2000
	DQ211628.1	ArD39554	Mauritania	*H.rufipes*	1984
	MF511236.1	SPUD8_81_7_813051_M	South Africa	*H.rufipes*	1981
	KF793334.1	Daral 2012	Mali	*Hyalomma* spp.	2012
	KY484026.1	DAK8194	Senegal	*H.truncatum*	1969
	DQ211626.1	ArD8194	Senegal	*H.truncatum*	1969
	MG516212.1	Pentalofos-Greece-2015	Greece	*R.bursa*	2015
	DQ211625.1	AP92	Greece	*R.bursa*	1975
	EF189752.1	Kelkit/Turkey-RB2/2005	Turkey	*R.bursa*	2005
	EF189751.1	Kelkit/Turkey-RB1/2005	Turkey	*R.bursa*	2005

	STRAIN	STRAIN NAME	LOCALITY	HOST	DATE OF COLLECTION
L SEGMENT	KY484039.1	JD-206	Pakistan	*H.anatolicum*	1965
	MG659725.1	FK16116	China	*H.asiaticum*	2017
	MG659722.1	WJQ16206	China	*H.asiaticum*	2017
	MH688499.1	YL16204	China	*H.asiaticum*	2016
	KY484034.1	IbAr10200	Nigeria	*H.excavatum*	1996
	AY389508.2	IbAr10200	Nigeria	*H.excavatum*	1966
	NC_005301.3	IbAr10200	Nigeria	*H.excavatum*	1966
	AY422209.2	IbAr10200	Nigeria	*H.excavatum*	1966
	AY389361.2	IbAr10200	Nigeria	*H.excavatum*	1966
	AY947891.1	IbAr10201	Nigeria	*H.excavatum*	1966
	MF547417.1	Caceres 2014	Spain	*H. lusitanicum*	2014
	KY484043.1	SPU 128/81	Uganda	*H.marginatum*	1981
	KY484033.1	HY-13	China	*H.marginatum*	1968
	DQ211615.1	ArD39554	Mauritania	*H.rufipes*	1984
	MF511202.1	SPUD8_81_7_813051_L	South Africa	*H.rufipes*	1981
	KY484025.1	DAK8194	Senegal	*H.truncatum*	1969
	DQ211613.1	ArD8194	Senegal	*H.truncatum*	1969
	MG516213.1	Pentalofos-Greece-2015	Greece	*R.bursa*	2015
	DQ211612.1	AP92	Greece	*R.bursa*	1975
	KY963542.1	ET35	Turkey	*R.bursa*	2016
	KY963541.1	ET37	Turkey	*R.bursa*	2016
	KY963540.1	ET36	Turkey	*R.bursa*	2016
	KY963543.1	KM6	Turkey	*R. sanguineus*	2015