

T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

**IDENTIFICATION OF VIROME CONTENT IN
FECAL SAMPLES OF BATS IN NORTHWEST
AND CENTRAL ANATOLIA**

ÖYKÜ DURAK

MOLECULAR BIOLOGY AND GENETICS
MASTER OF SCIENCE THESIS

IZMIR-2022

TEZ KODU: DEU.IBG.MSc-2019850020

T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

IDENTIFICATION OF VIROME CONTENT IN FECAL SAMPLES OF BATS IN NORTHWEST AND CENTRAL ANATOLIA

MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

ÖYKÜ DURAK

Supervisor: Assistant Prof. Zeynep Ahsen KOÇER

This study was funded by Dokuz Eylul University, Department of Scientific Research
Project (BAP Project) #2021.KB.FEN.031

Thesis ID: DEU.IBG.MSc-2019850020

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü
Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı,
Moleküler Biyoloji ve Genetik Yüksek Lisans Programı öğrencisi Öykü
DURAK '**IDENTIFICATION OF VIROME CONTENT IN FECAL
SAMPLES OF BATS IN NORTHWEST AND CENTRAL ANATOLIA**'
konulu Yüksek Lisans tezini 25.08.2022 tarihinde başarılı olarak
tamamlamıştır.

BAŞKAN

Dr. Öğr. Üyesi Zeynep Ahsen KOÇER

Dokuz Eylül Üniversitesi, İzmir Uluslararası Biyotıp ve Genom
Enstitüsü, Biyotıp ve Sağlık Teknolojileri ABD

ÜYE

Doç. Dr. Efe SEZGİN

İzmir Yüksek Teknoloji Enstitüsü,
Mühendislik Fakültesi, Gıda
Mühendisliği Bölümü

ÜYE

Prof. Dr. İbrahim Mehmet Ali ÖKTEM

Dokuz Eylül Üniversitesi, Tıp Fakültesi,
Temel Tıp Bilimleri Bölümü, Tıbbi
Mikrobiyoloji ABD

YEDEK ÜYE

Doç. Dr. Barış Emre DAYANÇ

Ekonomi Üniversitesi, Tıp
Fakültesi, Temel Tıp Bilimleri,
Tıbbi Mikrobiyoloji ABD

YEDEK ÜYE

Prof. Dr. Ayça Arzu SAYINER

Dokuz Eylül Üniversitesi, Tıp Fakültesi,
Temel Tıp Bilimleri Bölümü, Tıbbi
Mikrobiyoloji ABD

Dokuz Eylul University Izmir International Biomedicine and Genome
Institute

Department of Genomics and Molecular Biotechnology,
Molecular Biology and Genetics Graduate Program, Master of Science
student Öykü DURAK has successfully completed her Master of Science
thesis titled

**“IDENTIFICATION OF VIROME CONTENT IN FECAL SAMPLES OF
BATS IN NORTHWEST AND CENTRAL ANATOLIA’** on 25.08.2022

CHAIR

Assistant Prof. Zeynep Ahsen KOÇER

Dokuz Eylul University, Izmir International Biomedicine and Genome Institute,
Department of Biomedicine and Health Technologies

MEMBER

Associate Prof. Efe SEZGİN

Izmir Institute of Technology, Faculty of
Engineering, Department of Food
Engineering

MEMBER

Prof. Dr. İbrahim Mehmet Ali ÖKTEM

Dokuz Eylul University, Faculty of Medicine,
Department of Basic Medical Sciences,
Department of Medical Sciences, Medical
Microbiology

SUBSTITUTE MEMBER

Associate Prof. Barış Emre DAYANÇ

Izmir University of Economics
Faculty of Medicine
Department of Basic Medical, Medical
Microbiology

SUBSTITUTE MEMBER

Prof. Dr. Ayça Arzu SAYINER

Dokuz Eylul University, Faculty of Medicine,
Department of Basic Medical Sciences,
Department of Medical Sciences, Medical
Microbiology

Table of Content

TABLE OF CONTENT	I
LIST OF TABLES.....	III
LIST OF FIGURES.....	IV
LIST OF ABBREVIATIONS.....	V
ACKNOWLEDGEMENT	VII
ABSTRACT	1
ÖZET	2
1. INTRODUCTION.....	3
1.1. Statement and Importance of the Problem.....	3
1.2. Aim of the Study.....	4
1.3. Hypothesis of the study.....	5
2. REVIEW OF THE LITERATURE.....	5
2.1. Bats	5
2.1.1. Bat Species in Türkiye	6
2.1.1.1. <i>Miniopterus schreibersii</i>	7
2.1.1.2. <i>Miniopterus pallidus</i>	8
2.1.1.3. <i>Myotis capaccinii</i>	8
2.1.1.4. <i>Myotis myotis</i>	8
2.1.1.5. <i>Rhinolophus euryale</i>	9
2.1.1.6. <i>Rhinolophus ferrumequinum</i>	9
2.1.2. Ecological Niche	10
2.1.3. The Mechanism Against Pathogens	12
2.2. Bat-Borne Viruses	13
2.2.1. Viruses in Bat Blood.....	13
2.2.2. Viruses in Bat Saliva.....	14
2.2.3. Viruses in Bat Urine.....	14
2.2.4. Viruses in Bat Feces.....	15
2.2.4.1. Plant Viruses	16
2.2.4.2. Insect Viruses.....	17
2.2.4.3. Fungal Viruses	18
2.2.4.4. Vertebrate Viruses	19
2.2.4.5. Bacteriophages.....	21
2.2.5. Transmission Routes	21
2.3. Metagenomics	23
2.3.1. Viral Metagenomics.....	24
2.3.1.1. Shotgun Sequencing	25
3. MATERIALS AND METHODS.....	26
3.1. Type of the Study	26
3.2. Time and Place of the Study.....	26
3.3. Materials of the Study.....	27

3.4.	Variables of the Study	27
3.5.	Tools for Data Collection	28
3.5.1.	Materials used in the research.....	28
3.5.2.	Experimental Methods.....	33
3.5.2.1.	Sample Collection.....	33
3.5.2.2.	Sample Preparation.....	35
3.5.2.3.	Nucleic Acid Isolation and Concentration.....	36
3.5.2.4.	cDNA Synthesis.....	37
3.5.2.4.1.	<i>First-Strand cDNA Synthesis</i>	37
3.5.2.4.2.	<i>Second-Strand cDNA Synthesis</i>	38
3.5.2.5.	cDNA purification.....	38
3.5.2.6.	PCR Enrichment.....	39
3.5.2.7.	Agarose Gel Electrophoresis.....	40
3.5.2.8.	PCR Purification.....	41
3.5.2.9.	Quality Control.....	41
3.5.2.10.	Library Preparation.....	42
3.5.2.11.	Next Generation Sequencing.....	44
3.5.2.12.	Sequence Analysis and Assembly.....	44
3.5.2.13.	Taxonomic Assignment of Sequence Reads via BLAST.....	44
3.5.2.14.	Taxonomic Assignment of Sequence Reads via Kraken 2.....	45
3.5.2.15.	Taxonomic Assignment of the Contigs that were Obtained by de novo Assembly.....	45
3.6.	Study Plan Calendar	46
3.7.	Data Evaluation	47
3.8.	Limitations of the Study	47
3.9.	Ethics Committee Approval	47
3.10.	Legal Permits	48
4.	RESULTS	48
4.1.	Sample Collection	48
4.2.	Molecular Detection of Viral cDNAs in Bat Fecal Samples	50
4.3.	Sequence Analysis	51
4.3.1.	Taxonomic Assignment of Sequence Reads via BLAST.....	52
4.3.2.	Taxonomic Assignment of Sequence Reads via Kraken 2.....	56
4.3.3.	Taxonomical assignment of the Contigs Obtained by de novo Based Assembly.....	59
4.3.4.	Comparison of Three Approaches that were Used for Taxonomic Assignment.....	63
5.	DISCUSSION	69
6.	CONCLUSION AND FUTURE DIRECTIONS	77
7.	REFERENCES	78
8.	APPENDIX 1: CV	98

LIST OF TABLES

Table 1. List of plastic and glass consumables	28
Table 2. List of kits, reagents, solutions, chemicals, and antibiotics	30
Table 3. List of equipment	32
Table 4. List of primers used for amplification	33
Table 5. Components of the viral transport medium.	34
Table 6. Content of the reaction mixture	39
Table 7. PCR conditions	40
Table 8. Bat species recorded in sampling locations	49
Table 9. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Cilingoz Cave at Family Level via BLAST, NCBI Taxonomy ad ICTV databases.	53
Table 10. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Sefer Yitigi Cave at Family Level via BLAST, NCBI Taxonomy ad ICTV databases.	55
Table 11. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Sarihidir Tunnel at Family Level via BLAST, NCBI Taxonomy ad ICTV databases.	55
Table 12. Taxonomic Assignment of Viral Sequences in Bat Feces from Cilingoz Cave at Family Level via Kraken 2, NCBI Taxonomy ad ICTV databases.....	57
Table 13. Taxonomic Assignment of Viral Sequences in Bat Feces from Sefer Yitigi Cave at Family Level via Kraken 2, NCBI Taxonomy ad ICTV databases.....	57
Table 14. Taxonomic Assignment of Viral Sequences in Bat Feces from Sarihidir Tunnel at Family Level via Kraken 2, NCBI Taxonomy ad ICTV databases.....	58
Table 15. Taxonomic Assignment of the Contigs Obtained from Viral Sequences in Bat Feces from Cilingoz Cave by <i>de novo</i> Based Assembly at Family Level via BLAST, NCBI Taxonomy ad ICTV databases.	60
Table 16. Taxonomic Assignment of the Contigs Obtained from Viral Sequences in Bat Feces from Sefer Yitigi Cave by <i>de novo</i> Based Assembly at Family Level via BLAST, NCBI Taxonomy ad ICTV databases.	61
Table 17. Taxonomic Assignment of the Contigs Obtained from Viral Sequences in Bat Feces from Sarihidir Tunnel by <i>de novo</i> Based Assembly at Family Level via BLAST, NCBI Taxonomy ad ICTV databases.	62

LIST OF FIGURES

Figure 1. Bat species in Türkiye.	7
Figure 2. Transmission routes of bat-borne viruses.	22
Figure 3. Sampling locations of the bat fecal samples.....	34
Figure 4. dscDNAs after PCR enrichment.....	51
Figure 5. Comparison of Three Approaches that were Used for Taxonomic Assignment of Viral Sequences Obtained from Bat Feces Collected from Cilingoz Cave in Tekirdag in Northwest Anatolia.	64
Figure 6. Comparison of Three Approaches that were Used for Taxonomic Assignment of Viral Sequences Obtained from Bat Feces Collected from Sefer Yitigi in Burdur in Central Anatolia.....	66
Figure 7. Comparison of Three Approaches that were Used for Taxonomic Assignment of Viral Sequences Obtained from Bat Feces Collected from Sarihidir Tunnel in Nevsehir in Central Anatolia.	68

LIST OF ABBREVIATIONS

°C: Celsius degree

µg: microgram

µl: microliter

µM: micromolar

bp: Base pair

U: unit

rpm: Revolutions per minute

(-) ssRNA: negative-sense single-stranded RNA

BLAST: Basic Local Alignment Search Tool

BLT: Bead-Linked Transposomes

BSL2: Biosafety Level 2

cDNA: Complementary DNA

DBatVir: the Database of Bat-associated Viruses

DNA: Deoxyribonucleic acid

dNTP: Deoxynucleoside triphosphate

dscDNA: Double-stranded complementary DNA

dsDNA: Double-stranded DNA

dsRNA: Double-stranded RNA

EPM: Enhanced PCR Mix

EtOH: Ethanol

EUROBATS: Agreement on the Conservation of Populations of European Bats

ICTV: International Committee on Taxonomy of Viruses

MEGA X: Molecular Evolutionary Genetics Analysis Version X

MERS-CoV: Middle East Respiratory Syndrome Coronaviruses

NCBI: National Center for Biotechnology Information

NGS: Next Generation Sequencing

NIH: National Institutes of Health

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

RdRp: RNA dependent RNA polymerase

RNA: Ribonucleic acid

ROS: Reactive Oxygen Species

RSB: Resuspension Buffer

SARS- CoV-2: Severe Acute Respiratory Syndrome Coronaviruses 2

SARS-CoV: Severe Acute Respiratory Syndrome Coronaviruses

SPB: Sample Purification Beads

TAE: Tris-Acetic acid-EDTA

TB1: Tagmentation Buffer 1

TLR: Toll like receptor

TSB: Tagment Stop Buffer

TWB: Tagment Wash Buffer

WHO: World Health Organization

ACKNOWLEDGEMENT

I would like to sincerely thank to my advisor Asst. Prof. Zeynep Ahsen Koer for her support and patience starting from my summer internship to graduate studies. I am thankful for her guidance and great assistance in my thesis which could not have been achieved without her support. Her guidance has helped me to become a better scientist; therefore, I will be grateful to her for the rest of my life. I feel very lucky to have an opportunity to work with a stunning PI ever. I would also want to thank to all Koer lab members, Feray Yılmaz (best researcher ever), Muhammet Ekin Azbazdar (the most stabile one), Mert Dikmenoğulları (cell culture monster), Emre Mert Asar (the most dimer one) and Yavuz Mercan (the most former member) for their support, help and friendship.

I would like to present my gratitude to Prof. Dr. Raşit Bilgin and members of the field work for collecting the bat fecal samples which were used in this thesis. I would like to thank to Dokuz Eylul University (DEU) for their financial support for this project which could not be conducted without their support. I also want to thank DEU BAP staff, especially Elvan Özyılmaz for her help, patience and support during purchasing process.

At last, I would like to thank my father, Kamil Durak, my mother, Gülay Durak, my brother Eren Durak, my sister-in-law Ece Durak, my little brother Tutku, my lovely nephew Sirius and my grandfather, Zeynel Durak for being with me every time that I need them and their infinite love. I would like to dedicate this thesis to my mother and father, I could not be here without their support. I will always be grateful to my dear friends, Enes Furkan Aydın and Sezin Eyupustaoglu for their wise advises, endless energies, and being with me whenever I need them, even we have been living at different time zones. I also would like to thank to my oldest friends Bengisu Kantekin and Şilan Mete for their love and being with me through thick and thin.

ABSTRACT

Identification of Virome Content in Fecal Samples of Bats in Northwest and Central Anatolia

Öykü Durak, Dokuz Eylul University, Izmir International Biomedicine and Genome Institute, Department of Genome Sciences and Molecular Biotechnology, Molecular Biology and Genetics Master Program, 15 July Health and Art Campus, Balçova, 35340, Izmir/Türkiye

Bats carry various virus families that infect vertebrates, invertebrates, plant, fungi, or bacteria. Viruses may be shed via several routes and transmitted to humans. In current literature, there is insufficient information about bat-related viruses in Türkiye. In this study, feces of bats were collected from three different locations in the Northwest and Central Anatolia of Türkiye and investigated via metagenomics approach to profile viruses shed in feces. Then, virome content of bats were identified using three different bioinformatics approaches in the analysis of sequencing data. According to our results, different virus families could be detected by different bioinformatics approach and the virome content of bat feces was affected by bat species in terms of dietary choice, microbiome, and interaction with other animals. In Cilingoz Cave, virus families that infect bacteria, vertebrates, invertebrates, plant, and algae were detected. In Sefer Yitigi Cave, virus families that infect, bacteria, invertebrates, plant, and fungi were detected. In Sarihidir Cave, virus families that infect bacteria, vertebrates, invertebrates, archaea, and fungi were detected. While the highest viral diversity was detected in Sefer Yitigi Cave by BLAST before and after *de novo* assembly, the highest viral diversity was detected in Sarihidir Tunnel by Kraken 2. Overall, bacteriophages were the most dominant virus families by each bioinformatics approach for all regions. The sequence reads that hit virus families infecting vertebrates were investigated further. It was found that none of them carry risk for zoonosis. Such studies are crucial for monitoring emerging viruses.

Keywords: Bat virome, metagenomics, *Myotis myotis*, *Myotis capaccinii*, *Miniopterus schreibersii*, *Miniopterus pallidus*, *Rhinolophus euryale*, *Rhinolophus ferrumequinum*

ÖZET

Kuzeybatı ve Orta Anadolu'daki Yarasaların Dışkı Örneklerinde Virom İçeriğinin Belirlenmesi

Öykü Durak, Dokuz Eylül Üniversitesi, İzmir Uluslararası Biyotıp ve Genom Enstitüsü, Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Moleküler Biyoloji ve Genetik Yüksek Lisans Programı, 15 Temmuz Sağlık ve Sanat Kampüsü, Balçova, 35340, İzmir/TÜRKİYE

Yarasalar bakterileri, bitkileri, mantarları, omurgalı ve omurgasız hayvanları enfekte edebilen çeşitli virüs ailelerini bünyelerinde barındırırlar. Bu virüsler çeşitli yollarla etrafa saçılabilir ve insanlara bulaşabilirler. Güncel verilerde, Türkiye'de yaşayan yarasalarla ilişkili virüs aileleri hakkında yeterli bilgi bulunmamaktadır. Bu çalışmada, yarasa dışkısında bulunan virome içeriğinin belirlenebilmesi için Türkiye'nin Kuzeybatı ve Orta Anadolu bölgesinde yer alan üç farklı konumdan yarasa dışkı örnekleri toplanmış ve metagenomik yaklaşımla incelenmiştir. Son olarak, elde edilen diziler üç farklı biyoinformatik yöntemle analiz edilerek yarasa dışkılarında bulunan virom içeriği belirlenmiştir. Sonuç olarak, farklı biyoinformatik yöntemlerle farklı virüs aileleri tespit edilebildiği gibi, yarasa dışkı viromunun yarasa türlerine ve bu türlerin beslenme seçimlerine, mikrobiyomlarına ve diğer hayvanlarla iletişimlerine göre değişebildiği gözlemlenmiştir. Çilingoz Mağarası'nda, bakterileri, bitkileri, algleri, omurgalı ve omurgasız hayvanları enfekte edebilen virüs aileleri tespit edilmiştir. Sefer Yitiği Mağarası'nda ise bakterileri, bitkileri, mantarları, ve omurgasız hayvanları enfekte edebilen virüs aileleri tespit edilmiştir. Son olarak, Sarıhıdır Tüneli'nde bakterileri, arkeaları, mantarları, omurgalı ve omurgasız hayvanları enfekte edebilen virüs aileleri tespit edilmiştir. En yüksek viral çeşitlilik, dizilerin *de novo* yaklaşımla birleştirilmesinden önce ve sonra BLAST yöntemi ile Sefer Yitiği Mağarası'nda tespit edilirken, Kraken 2 yöntemi ile Sarıhıdır Tüneli'nde tespit edilmiştir. Kullanılan üç farklı biyoinformatik yöntemle de bakteriyofajlar her bölgede en baskın virüs ailesi olarak tespit edilmiştir. Omurgalıları enfekte edebilen virüs ailelerine isabet eden dizi okumaları detaylı olarak incelenmiştir. Bu dizi okumalarının eşleştiği virüslerin hiçbirisi zoonoz riski taşımamaktadır. Yeni ortaya çıkabilecek virüslerin takibi için bu tür çalışmalar oldukça önemli bir yere sahiptir.

Anahtar Sözcükler: Bat viromu, metagenomiks, *Myotis myotis*, *Myotis capaccinii*, *Miniopterus schreibersii*, *Miniopterus pallidus*, *Rhinolophus euryale*, *Rhinolophus ferrumequinum*

1. INTRODUCTION

1.1. Statement and Importance of the Problem

Bats are the second largest mammalian group classified under the Chiroptera order. Their roost selection, flying capacity, and feeding behaviors may differ in different bat species. They are observed in every continent except Antarctica. Among nesting locations, Türkiye is one of the hotspot locations for several bat species. According to the studies carried out in Türkiye, 39 different bat species that belong to the *Pteropodidae*, *Emballonuridae*, *Rhinolophidae*, *Vespertilionidae*, and *Molossidae* families have been recorded in Türkiye (Benda and Horacek, 1998; Çoraman *et al.*, 2013).

Recent studies showed that bats are the natural reservoir for many virus families such as *Coronaviridae*, *Rhabdoviridae*, *Orthomyxoviridae*, *Filoviridae*, and *Paramyxoviridae*. They can carry more than one virus family at the same time with few or no symptoms. Even though bats share similar immune systems with mammals, the pathogenicity of viruses on bats is different from other animals. Different evolutionary adaptations in their immune system might be resulted from their high body temperature, high metabolic rate, and hibernation behavior. In addition, bats live approximately 25-35 years which is long life span for small mammals. Having such a long life might contribute to persistent viral infection and new adaptation in immune system for viral clearance.

Bats have critical roles in ecosystem. They maintain the ecological balance by contributing to the prey-predator cycle, pollination, seed dispersal, and soil fertility via high mineral source in bat guano. The balance between bats and nature might be disturbed by humans due to agricultural needs, hunting, search for treasure, tourism, urbanization, habitat fragmentation, and extreme usage of pesticides. Disturbance of this balance might cause transmission of viruses from bats to humans, which might cause outbreaks. The transmission could be either via direct contact with body fluids of bats (saliva, feces, blood), vector borne, intermediate hosts or indirect contact with contaminated environmental sources. Viral metagenomics studies have been significantly contributing to the identification and detection of emerging/re-emerging viruses for almost a decade. With recent developments in molecular

biology and omics technologies, the causative agents of outbreaks can be detected and characterized immediately. Unfortunately, the number of studies about bats and bat related viruses in Türkiye are insufficient. Therefore, screening of known viral reservoirs like bats are important to be aware of viruses in bat populations in Türkiye and their potential risk for zoonosis. This study will be the first study investigating the virome content in the fecal samples of bats that inhabit in Northwest and Central Anatolian regions of Türkiye using metagenomics approach.

1.2. Aim of the Study

The aim of this study is to identify the virome profile in bat feces that are collected from three different sampling locations in the Northwest and Central Anatolia region of Türkiye via metagenomics approach.

For this purpose, it is aimed to:

- a. Concentrate and inactivate viral particles found in bat fecal samples and eliminate free nucleic acids from the samples
- b. Obtain high quality viral nucleic acids from each sample
- c. Synthesize and enrich cDNA libraries with random primers
- d. Prepare genomic libraries from synthesized cDNA libraries
- e. Sequence cDNA libraries via next-generation sequencing
- f. Identify the virome profile in bat feces using bioinformatic tools.

The following steps were followed to achieve these goals:

- a. Viral particles in fecal samples were concentrated via ultracentrifugation and treated with DNase and RNase cocktail to remove the free nucleic acids in samples. Then, viral particles were inactivated using a specific reagent for safety purposes.
- b. Total viral nucleic acid isolation was performed to obtain high-quality nucleic acids.
- c. The first and second strands of cDNA were synthesized consecutively and were amplified with primers covering a part of RNA-dependent RNA-polymerase (RdRp).

- d. Genomic libraries were prepared for each location separately and sequenced by Illumina NextSeq500 platform.
- e. Paired-end sequence reads were investigated using three different approaches: 1) taxonomic identification of sequence reads using Basic Local Alignment Search Tool (BLAST) on CLC Genomics Workbench 22.0.1. (QIAGEN) 2) taxonomic identification of sequence reads via Kraken 2 (Wood *et al.*, 2019) 3) taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on BLAST at National Center for Biotechnology Information (NCBI).
- f. Paired-end sequence reads before or after assembly were investigated in databases and classified according to their family rank.

1.3. Hypothesis of the study

It is known that bats carry various number of virus families, some of which are shed via feces. Metagenomics studies are mainly descriptive studies with no hypothesis. Thus, in this non-hypothesis driven study, we aimed to identify the virus families in the feces of bat species inhabiting in Northwest (*Myotis capaccinii*, *Miniopterus schreibersii*, *Rhinolophus euryale*) and Central Anatolia (*Myotis capaccinii*, *Miniopterus schreibersii*, *Rhinolophus ferrumequinum*, *Miniopterus pallidus*, *Myotis myotis*) via three different bioinformatics approaches. This will be the first study of its own kind that has been conducted in Türkiye.

2. REVIEW OF THE LITERATURE

2.1. Bats

Bats are the second largest mammalian group classified under the order of Chiroptera that includes 1400 species representing 20% of the mammalian species (Li *et al.*, 2010). The Chiroptera order is divided into suborders Megachiroptera and Microchiroptera (Wu *et al.*, 2016). Individuals who belong to Megachiroptera are distinguished by having a large body size

and predominantly feed on fruits. On the other hand, individuals belonging to Microchiroptera have an echolocating ability and predominantly feed on insects (Sun *et al.*, 2020).

Bats are capable of flying long distances which results in wide range of distribution and diversity in the ecosystem. They are detected on all continents except Antarctica. Therefore, their impacts on the biodiversity and the ecosystem from rainforests to deserts are indispensable piece of nature (Jones *et al.*, 2009). Türkiye is one of the locations providing roost areas that protect bats from harsh environment conditions and predator attacks during mating, resting, pregnancy, and nursing juvenile bats. Bats may select abandoned mines, rock crevices, ruins, tunnels, tree hollows, and abandoned buildings as roosts (Yük *et al.*, 2019). Their reproductive strategies, feeding, and roost choice differ in different bat species (Jones *et al.*, 2009).

2.1.1. Bat Species in Türkiye

Türkiye is rich in biodiversity and is located on a wide range of landscape from the Blacksea Region to the Mediterranean Region and from the Eastern Anatolian Region to the Aegean Region. Under the order Chiroptera, 39 different bat species belonging to the *Pteropodidae*, *Emballonuridae*, *Rhinolophidae*, *Vespertilionidae*, and *Molossidae* families can be found in different ecogeographic regions of Türkiye (Benda and Horacek, 1998; Çoraman *et al.*, 2013; Yorulmaz *et al.*, 2020). Among these bats, one of the species feeds on fruits, and the other bat species feed on insects (Benda and Horacek, 1998; Çoraman *et al.*, 2013). According to the literature and field studies that have been conducted in Türkiye, 1902 individuals were recorded until 2020 (Yorulmaz *et al.*, 2020).

In Türkiye, bats are usually spotted in caves, ruins, man-made structures, tunnels, bridges, forested areas, and rocky locations (Albayrak, 2012). Overall, approximately 40.000 caves have been found in Türkiye (Albayrak, 2012). Some of the caves in Burdur, Antalya, and Tokat are visited by tourists which affects the dynamics of the bat populations (Albayrak, 2012).



Figure 1. Bat species in Türkiye. A) *Miniopiterus schreibersii* B) *Miniopiterus pallidus* C) *Myotis capacinii* D) *Myotis myotis* E) *Rhinolophus euryale* F) *Rhinolophus ferrumequinum* (Adapted from: UNEP/EUROBATS, 2022a).

Some of the bat species recorded in Türkiye are listed as *Miniopiterus schreibersii*, *Miniopiterus pallidus*, *Myotis capacinii*, *Myotis myotis*, *Rhinolophus euryale* and *Rhinolophus ferrumequinum* (Figure 1).

2.1.1.1. *Miniopiterus schreibersii*

The common name for *Miniopiterus schreibersii* is ‘Schreibers’ Bent-winged Bat’. They prefer wood-rich areas, mines, karst caves, and cellars as a roost (UNEP/EUROBATS, 2022a). Schreibers’ bent-winged bats are the cave-dwelling species that generally consume lepidoptera larvae, moths, spiders, and beetles (UNEP/EUROBATS, 2022a). They migrate more than one time in a year and can fly long distances. The average distance that they can fly is between 40-100 km. According to the IUCN Red List, their condition is near threatened (UNEP/EUROBATS, 2022a). Up to now, they have been recorded in south-western Europe, the Balkans, and the Caucasus (UNEP/EUROBATS, 2022a).

2.1.1.2. *Miniopterus pallidus*

The common name for *Miniopterus pallidus* is ‘Pale Bent-winged Bat’ (UNEP/EUROBATS, 2022b). They generally prefer to live in tunnels, caves, and mines sharing the same roost with the species that belong to *Rhinolophus* and *Myotis* (GBIF, 2022). For a long time, *M. pallidus* was thought to be a subspecies of *M. schreibersii* that shows differences due to geographical distribution (UNEP/EUROBATS, 2022b; GBIF, 2022). Later, *M. pallidus* was ranked as a species based on the comparison of nuclear DNA and mitochondrial DNA. On the IUCN Red List, their condition is not evaluated (UNEP/EUROBATS, 2022b). Up to now, they have been recorded in the Europe, Northern Asia, and Southern Asia (ITIS, 2022).

2.1.1.3. *Myotis capaccinii*

The common name for *Myotis capaccinii* is ‘Long-fingered Bat’. They prefer caves in limestone areas and anywhere in the underground site as a roost (UNEP/EUROBATS, 2022c). Long-fingered bats are cave-dwelling species and can spend the whole year in caves or mines feeding on flies, fish, water boatmen, and moths (UNEP/EUROBATS, 2022c). They are able to fly approximately 40- to- 50 km (UNEP/EUROBATS, 2022c). According to the IUCN Red List, their condition is vulnerable (UNEP/EUROBATS, 2022c). Up to present, they have been observed on the Mediterranean coasts and in eastern Türkiye (UNEP/EUROBATS, 2022c).

2.1.1.4. *Myotis myotis*

The common name for *Myotis myotis* is ‘Greater Mouse-eared Bat’ that generally prefers the woodlands as roosting site, although they might prefer caves, mines, and cellars in the winter season (UNEP/EUROBATS, 2022d). They are one of the cave-dwelling species (UNEP/EUROBATS, 2022d). Their echolocation, high advanced hearing, and smelling abilities provide advantages during hunting for ground beetles, beetle larvae, spiders, and grasshoppers (UNEP/EUROBATS, 2022d). Greater mouse-eared bats are regional migrants and have different roosts in winter and summer. Generally, the distance between their winter

and summer roosts are between 50-100 km (UNEP/EUROBATS, 2022d). According to the IUCN Red List, their condition is stable. Up to now, they have been recorded in western, central, and southern Europe, as well as in Anatolia (UNEP/EUROBATS, 2022d).

2.1.1.5. *Rhinolophus euryale*

The common name for *Rhinolophus euryale* is ‘Mediterranean Horseshoe Bat’ that prefers underground sites, cellars and roofs as roosting sites in summer, and caves in winter (UNEP/EUROBATS, 2022e). They feed on small moths, crane flies, and small beetles (UNEP/EUROBATS, 2022e). According to recent updates, they stay at the same roost in all seasons. In addition, *R. euryale* colonies live in the 50 km range (UNEP/EUROBATS, 2022e). According to the IUCN Red List, their condition is near threatened (UNEP/EUROBATS, 2022d). Up to now, they have been recorded in the Mediterranean covering from Mediterranean coast of Europe, Balkans, Hungary, Slovakia, west coast of Anatolia to northwest Africa. They are also observed in the Caucasus, Iran, and Iraq (UNEP/EUROBATS, 2022e).

2.1.1.6. *Rhinolophus ferrumequinum*

The common name for *Rhinolophus ferrumequinum* is ‘Greater Horseshoe Bat’ (UNEP/EUROBATS, 2022f). Greater horseshoe bats have a very special echolocation ability (UNEP/EUROBATS, 2022f). They feed on beetles, moths, and other insects (UNEP/EUROBATS, 2022f). Up to now, they have been detected in areas covering from North Africa to southern Europe including Anatolia (UNEP/EUROBATS, 2022f). Greater Horseshoe Bats prefer to stay in same location. Colonies of *R. ferrumequinum* are found in 20- to- 50km distance from each other (UNEP/EUROBATS, 2022f). Individuals of this species are affected by the usage of toxic pesticides, habitat loss and disturbance of roosts which cause a decrease in the number of individuals (UNEP/EUROBATS, 2022f). They are under the danger of extinction in Europe and their numbers are decreasing according to the IUCN Red List (UNEP/EUROBATS, 2022f).

2.1.2. Ecological Niche

Bats, especially cave-dwelling species, are accepted as keystone species due to their vital role in the ecosystem with their position in prey-predator cycle, pollination, seed dispersal, and soil fertility.

Bats, as a predator, hunt at night and are nourished by a wide range of food sources: insects, seeds, frogs, nectar, small mammals, fish, fruits, and blood of mammals and birds (Kunz *et al.*, 2011). The feeding behavior of bats varies based on gender, age, time of feeding, species, accessibility to prey, and reproductive status which contribute to the partition of resources due to differences in body size and shape of wings (Fukui *et al.*, 2009). On the other hand, bats are hunted by amphibians, birds like owls and hawks, reptiles like snakes, and mammals like raccoons and opossums (Speakman *et al.*, 1991; Fenton *et al.*, 1995).

In addition to their role in the food chain, bats serve in pollination and seed dispersal. Because bats consume different parts of plants such as nectar and pollen, they can actively take role in seed dispersal and pollination (Muscarella and Fleming, 2007). By this way, they contribute to species diversity and the regeneration of tropical forests. Especially members of *Pteropodidae*, *Phyllostomidae*, and fruit-eating bats have an impact on seed dispersal and pollination of ecological and economically important plants like columnar cacti (Kunz *et al.*, 2011).

Last but not least, bats play an important role in the distribution of nutrients from high fertile habitats such as close proximity of lake or river to poor fertile habitats such as deserts or arid landscapes, especially via their fecal material. Thus, dispersion of bat feces to different habitats is important for the ecosystem. The bat feces consist of seeds, pulp, insect residuals, or arthropod residuals (Ferreira, 2019). In addition, fecal material accumulates on the ground of caves that are used as roosting sites by bats (Ferreira, 2019; Yök *et al.*, 2019). In time, accumulated fecal material becomes harden and the content of it is highly rich in mineral. This accumulated fecal material is known as guano (Ferreira and Martins, 1999). The guano is composed of a high amount of nitrogen, phosphorus, and potassium and a small amount of

calcium, magnesium, chlorine, sodium, and iron (Altıntaş *et al.*, 2005). Therefore, bat guano is a highly demanded material as an agricultural fertilizer. The researchers showed the effect of guano on the economically important plant species. In a study, plants treated with guano grew better compared to the plants treated with chemicals (Sothearen *et al.*, 2014). Due to high demand on guano as a fertilizer, guano mining has become popular in many countries, which makes caves a target for guano miners. Guano mining is actively pursued in Türkiye since 2006 (Altıntaş *et al.*, 2005; Yök *et al.*, 2019). The price for bat guano may change according to the composition of ingredients. For instance, more than 332 tons of guano are collected from the Magarawa region in South Africa every year. The estimated income from this bat guano trade is almost \$28000 per year (Harouna *et al.*, 2021). Thus, the guano trading is one of the main income sources for the local economy in Magarawa region. On the other hand, improper guano mining activities in caves disrupts the cave fauna and increases the risk of pathogen transmission from guano to humans.

Guano has a significant influence on the cave ecosystem. The content of guano is dynamic and affected by the dietary habits of bat species, thus the composition of guano varies in different caves. The bat guano provides a microhabitat for fungi, bacteria, arthropods, and lichens. In addition, some of the invertebrates and cave-dwelling salamander groups are dependent on the accumulated guano on the cave floor, which maintains the balance in cave fauna and flora (mosses, vascular plants, etc.) (Monro *et al.*, 2018; Ferreira, 2019). It also reflects the gut microbiome (including virome) of the bats residing in the caves (Ferreira, 2019).

2.1.3. The Mechanism Against Pathogens

Bats are unique organisms hosting multiple virus families at the same time with no sign of disease or infection. The distinctive immune response of bats against viral infections are evolved under the influence of co-existence with viruses (Banerjee *et al.*, 2020). Recent transcriptomic and genomic studies show similarities between innate immune systems of bats and humans. For instance, Toll like receptor 3 (TLR3) is involved in the detection of viral RNA in humans as well as in *Eptesicus fuscus* (Big brown bat) (Kawai *et al.*, 2006; Banerjee *et al.*, 2017). Sequence analyses indicate homology in the immune related genes between bats and humans. Transcriptomic study using blood and tissue samples (spleen, thymus lymph node, and bone marrow) from *Pteropus alecto* (Australian black flying fox) was conducted to understand the immune response (Papenfuss *et al.*, 2012). They found that 3.5% of the genes were immune associated in bats, while 7% of the genes were immune associated in humans (Kelley *et al.*, 2005; Papenfuss *et al.*, 2012). Further studies might be necessary to identify the immune associated genes in bats or in a different perspective, all of the immune-related genes have already been discovered. On the other hand, these genes might be already discovered and the lower number of genes than humans might be the indication of the novelty. While bats may control the pro-inflammatory response level via specific mechanisms, other mammals may show excessive immune response against viral infections, which might increase the severity of the disease (Banerjee *et al.*, 2020). Bats have gone through different evolutionary steps compared to other mammals. For instance, in one study the researchers showed that PYHIN (PYRIN and HIN domain) gene family does not exist in *Pteropus alecto* (Black flying fox) and *Myotis davidii* (David's myotis) (Ahn *et al.*, 2016). This gene family is known to have a role in immune response mechanisms against exogenous or own DNA to trigger the inflammatory pathways in mammals (Ahn *et al.*, 2016). Bats can increase their metabolic capacity thirty-four times compared to their resting stage. The production of reactive oxygen species (ROS) increases along with metabolic rate (Thomas and Suthers, 1972). These metabolic products give damage to the structure of DNA by adding extra double bonds to bases and removing hydrogen atom from thymine base at the methyl group (Cooke *et al.*, 2003; Ahn *et al.*, 2016). Mutations in DNA structure can be repaired by DNA repair mechanisms. However, excessive amount of ROS might alter the gene expression in these pathways (Cooke *et al.*, 2003; Ahn *et al.*, 2016).

Increased body temperature induces the immune system, activates mobilization of immune cells, and eliminates viral and bacterial infections in vertebrates and invertebrates. Body temperature of bats is high as a consequence of flying (Ahn *et al.*, 2016). The viral infections trigger the DNA damage *in vitro* and *in vivo* (Li *et al.*, 2015). Persistent infection and flying long distances might drive the adaptation in immune system of bats (Ahn *et al.*, 2016). High body temperature, metabolic rate and long-life span of bats might contribute to the adaptations in the immune system and viral clearance (Calisher *et al.*, 2006).

2.2. Bat-Borne Viruses

Aside from their vital role in the ecosystem and their unique immune system, bats stand out with their hospitality to variety number of virus families at the same time showing mild or no symptoms. Molecular techniques proved that bats are the host for a significant number of bacterial and viral pathogens (Li *et al.*, 2010; Zheng *et al.*, 2018; Wang *et al.*, 2019; Li *et al.*, 2021). However, the virome structure is dynamic between individuals and affected by many factors such as ecological habitats, dietary habits, and social behavior (Li *et al.*, 2010).

2.2.1. Viruses in Bat Blood

Viral nucleic acids detection and virus isolation methods are established from blood samples of bats. The infectious filovirus, Lloviu virus, was isolated from the blood sample of *M. schreibersii* in Hungary (Kemenesi *et al.*, 2022). In one study, the circulation of Lloviu virus were shown in the colonies of *M. schreibersii*. They concluded that *M. schreibersii* is host for Lloviu virus (Kemenesi *et al.*, 2022). To the best of our knowledge, there is no bat origin virus that is spread via contact with bat blood. However, one study showed that the fruit bats can carry arthropod-related viruses in their blood (Bennet *et al.*, 2019). They concluded that this might be a result of bitten by arthropods or feed on fruits that are infested by arthropods. They also speculate that some of the virus families can spread from the gastrointestinal system to the circular system (Bennet *et al.*, 2019).

2.2.2. Viruses in Bat Saliva

Bats can shed viruses via their saliva and transmitted through open wounds or by biting (CDC, 2022a). In one study, 138 saliva specimens were taken from *Desmodus rotundus* (common vampire bat) and *Molossus molossus* (velvety free-tailed bat) species in two seasons (Salmier *et al.*, 2017). Mostly phages that belong to *Siphoviridae* and *Podoviridae* were detected in saliva samples. In addition, viruses that belong to *Baculoviridae*, *Iridoviridae*, *Potyviridae*, *Virgaviridae*, *Herpesviridae*, *Anelloviridae*, *Circoviridae*, *Parvoviridae* and *Polyomaviridae* were identified. They showed similarities between sequences obtained from fecal and saliva samples suggesting evidence for oral-fecal transmission (Salmier *et al.*, 2017). Apart from these virus families, rabies, the most common virus from the *Rhabdoviridae* family in bat populations, is one of the dangerous viruses that can be transmitted to humans via direct contact with the nervous tissue and/or saliva from infected animals (Streicker *et al.*, 2012; CDC, 2022b). The lethality caused by rabies is varied based on several parameters: animal species, host immune status, a variant of the virus, exposure severity, and concentration of inoculum (Streicker *et al.*, 2012). Rabies infection in humans affects the nervous system, changes the characteristic behavior, and mostly results in death. However, some of the infected animals might show mild symptoms (Streicker *et al.*, 2012; CDC, 2022c). *Desmodus rotundus* (common vampire bat) is known as the main reservoir of rabies. These bats feed on mammalian blood. Therefore, they spread the virus via biting animals and humans (Streicker *et al.*, 2012). In South America, the number of people bitten by vampire bats is higher than the number of people bitten by other animals. In Latin America, the population of vampire bats has been increasing along with the number of livestock in the agricultural area, as these bats generally feed on the blood of livestock. This might increase the possibility of disease transmission from bats to livestock, and even humans (Streicker *et al.*, 2012).

2.2.3. Viruses in Bat Urine

Bats can shed viruses via urination and transmit the viruses to humans (CDC, 2022d). In one study, 100 urine samples were taken from *Eonycteris spelaea* (cave nectar bat) (Mendenhall *et al.*, 2019). After metagenomics analysis, they detected viral genome sequences

that belong to *Adenoviridae*, *Astroviridae*, *Betaflexiviridae*, *Chrysoviridae*, *Coronaviridae*, *Dicistroviridae*, *Flaviviridae*, *Herpesviridae*, *Microviridae*, *Myoviridae*, *Paramyxoviridae*, *Partitiviridae*, *Podoviridae*, *Polymaviridae*, *Poxviridae*, *Reoviridae*, *Retroviridae*, *Siphoviridae* and *Totiviridae* families in urine samples (Mendenhall *et al.*, 2019). In their results, majority of the sequences were from *Retroviridae* (25.38%), *Dicistroviridae* (20.59%) and *Coronaviridae* (15.22%) families (Mendenhall *et al.*, 2019).

Another virus that can transmit through urine is Hendra virus from *Paramyxoviridae* family that have caused infections in humans. Cases showed that the Hendra virus can infect humans and other terrestrial mammals causing severe effects on the lung and central nervous system (Breed, 2008). Hendra viruses can be transmitted to humans through body fluids of intermediate hosts which are horses. Horses may be infected by the viruses in the urine or saliva of flying foxes (CDC, 2022d). Hendra viruses were first isolated from the lungs of sick horses, therefore, for a long time, horses were accepted as the natural reservoirs of Hendra viruses (Breed, 2008). Then, with the developments in sequencing, flying foxes belonging to the *Pteropus* genus of the *Pteropodidae* family were identified as the natural reservoir of Hendra viruses (Breed, 2008). In an epidemiologic study conducted in Australia during the Hendra virus outbreak in 2008, eleven cases of virus transmission from bats to horses were detected. Moreover, transmission from horses to human were also observed (Field *et al.*, 2010).

2.2.4. Viruses in Bat Feces

Bats harbor many DNA and RNA viruses that can infect plants, insects, bacteria, and vertebrates, which comprise the large part of the gut virome (Li *et al.*, 2010; Yinda *et al.*, 2018; Zheng *et al.*, 2018; Wang *et al.*, 2019; Li *et al.*, 2021). The dietary choices, habitats, and environmental stress might affect the species diversity, species distribution and frequency in microbial communities in gut (Edenborough *et al.*, 2020). Microbiomes play a significant role in viral infection in mammals influencing host immune response. In mice, the infectivity and pathogenicity of the poliovirus is increased with the existence of gut microbiome (Kuss *et al.*, 2011; Edenborough *et al.*, 2020). Therefore, the virome might be varied in bat species.

According to the research that were conducted in multiple countries at different time intervals, viruses that belong to *Adenoviridae*, *Astroviridae*, *Caliciviridae*, *Coronaviridae*, *Flaviviridae*, *Papillomaviridae*, *Paramyxoviridae*, *Parvoviridae*, *Picornaviridae*, *Polyomaviridae*, and *Reoviridae* families were detected in fecal samples of different bat species (Yinda *et al.*, 2018; MendeHall *et al.*, 2019; Bolatti *et al.*, 2020; Hardmeier *et al.*, 2021). In one study, the researchers investigated 4000 rectal and pharyngeal swabs belonging to horseshoe bats (*Rhinolophus*), mouse-eared bats (*Myotis*), and bent winged bats (*Miniopterus*) in China (Wu *et al.*, 2016). They detected many viruses from *Adenoviridae*, *Herpesviridae*, *Papilomaviridae*, *Polyomaviridae*, *Reoviridae*, *Hepadnaviridae*, *Retroviridae*, *Circoviridae*, *Parvoviridae*, *Astroviridae*, *Caliciviridae*, and *Flaviviridae* families as well as several novel viruses from the *Coronaviridae*, and *Paramyxoviridae* families (Wu *et al.*, 2016).

2.2.4.1. Plant Viruses

Plant viruses are divided into 73 genera and 49 families. *Bromoviridae*, *Closteroviridae*, *Luteoviridae*, and *Potyviridae* cover the high proportion of plant virus families and possess positive-sense RNA genome (Gergerich *et al.*, 2006). Small number of plant viruses in *Bunyaviridae* and *Rhabdoviridae* families carry negative-sense RNA as a genetic material (Gergerich *et al.*, 2006). *Caulimoviridae* (pararetrovirus) family possess dsDNA genomes. *Reoviridae* with dsRNA and *Geminiviridae* with ssDNA are the largest families in their genome classification (Gergerich *et al.*, 2006).

Viral sequences belonging to plant viruses were detected in bat fecal samples in many studies. In one study, plant viruses belonging to *Luteoviridae*, *Secoviridae*, *Tymoviridae*, and *Partitiviridae* families were detected in the guano samples of insectivorous bats in California and Texas (Li *et al.*, 2010). They speculated that the high number of sequences from plant viruses are associated with the dietary choice of ingested insects which consumes plants (Li *et al.*, 2010). In another study, 86 fecal samples that belong to *Cynopterus sphinx*, *Miniopterus schreibersii*, *Rousettus leschenaultii*, *Hipposideros larvatus*, *Rhinolophus blythi* and *Scotophilus kuhlii* in four different provinces in China were analyzed and sequenced (Zheng *et*

al., 2018). According to sequence analysis, plant viruses were classified under the *Luteoviridae*, *Bromoviridae*, *Bunyaviridae*, and *Potyviridae* (Zheng *et al.*, 2018).

2.2.4.2. Insect Viruses

Just like other organisms, there are many virus families that can infect insects (Prasad, 2016). The dietary habits reflect the viral content in fecal samples. Therefore, viruses that infect insects have been detected in the fecal samples of insectivorous bats. These viruses possess either DNA or RNA genome. For instance, *Cypovirus*, *Dicistrovirus*, *Nodavirus*, and *Tetravirus* have RNA genome; while *Densovirus*, *Entomopoxvirus*, *Ascovirus*, *Iridovirus*, *Nudivirus*, *Hytrosavirus*, *Iflaviruses*, and *Baculovirus* contain DNA genome (Prasad, 2016).

In rectal and pharyngeal swab samples of *Myotis ricketti*, *Rhinolophus ferrumequinum*, *Myotis myotis*, *Rhinolophus sinicus*, *Hipposideros armiger*, *Hipposideros pomona*, *Hipposideros cineraceus*, *Tylonycteris robustula*, *Rhinolophus affinis* and *Miniopterus schreibersii*, viruses belonging to *Baculoviridae*, *Iflaviridae*, *Dicistroviridae*, *Tetraviridae*, and *Densovirinae* families were detected (Wu *et al.*, 2016). In California and Texas in USA, guano samples were collected from *Tadarida brasiliensis*, *Myotis velifer*, *Nycticeus humeralis*, and *Perimyotis subflavus* (Li *et al.*, 2010). In that study, viruses from *Dicistroviridae*, *Alphanodavirus*, *Iflaviridae*, *Tetraviridae*, and *Densovirinae* families were detected (Li *et al.*, 2010). In another study conducted in Northern California in USA, viruses from *Alphatetraviridae*, *Ascoviridae*, *Baculoviridae*, *Bidnaviridae*, *Iflaviridae*, *Genomoviridae*, *Hytrosaviridae*, *Iridoviridae*, *Metaviridae*, *Nodaviridae*, *Nudiviridae*, *Peribunyaviridae*, *Permutotetraviridae*, *Polucipiviridae* and *Polydnaviridae* families were detected in the guano samples of *Myotis californicus* and *Myotis yumanensis* (Li *et al.*, 2021). In China, viruses belonging to *Parvoviridae*, *Potyviridae*, *Baculoviridae*, and *Polydnaviridae* families were detected in the fecal samples of *Cynopterus sphinx*, *Miniopterus schreibersii*, *Rousettus leschenaultii*, *Hipposideros larvatus*, *Rhinolophus blythi* and *Scotophilus kuhlii* (Zheng *et al.*, 2018).

2.2.4.3. Fungal Viruses

Recently, bats have suffered from one of the fungus-related diseases called white-nose syndrome. This fungus spreads on the bat skin and cause starvation, dehydration, and even death, while bats are hibernating (USGS, 2022). Fungal viruses, in other words mycoviruses, are widely common in fungal species (Ghabrial and Suzuki, 2009). Some of the mycovirus families such as *Narnaviridae*, *Chrysoviridae*, and *Hypoviridae* can only infect fungal host (Ghabrial and Suzuki, 2009). However, members of *Metaviridae*, *Pseudoviridae*, *Reoviridae*, *Totiviridae*, and *Partitiviridae* families are not restricted in only one host and can infect fungi, plants, animals, or protozoa (Ghabrial and Suzuki, 2009). Therefore, fungal viruses can be detected in bat fecal samples. For instance, viruses from *Chrysoviridae*, *Hypoviridae*, *Partitiviridae*, and *Totiviridae* families were detected in rectal and pharyngeal swab samples of *Myotis ricketti*, *Rhinolophus ferrumequinum*, *Myotis myotis*, *Rhinolophus sinicus*, *Hipposideros armiger*, *Hipposideros pomona*, *Hipposideros cineraceus*, *Tylonycteris robustula*, *Rhinolophus affinis* and *Miniopterus schreibersii* (Wu *et al.*, 2016). In another study, viruses that belong to *Chrysoviridae* and *Narnaviridae* families were detected in the fecal samples of *Cynopterus sphinx*, *Miniopterus schreibersii*, *Rousettus leschenaultii*, *Hipposideros larvatus*, *Rhinolophus blythi* and *Scotophilus kuhlii* (Zheng *et al.*, 2018).

Viruses in *Alphaflexiviridae*, *Endornaviridae*, *Gammaflexiviridae*, *Narnaviridae*, *Barnaviridae*, *Hypoviridae*, and *Metaviridae* families possess positive-sense RNA genome (Ghabrial and Suzuki, 2009; Ghabrial *et al.*, 2015). On the other hand, viruses belonging to *Reoviridae*, *Totiviridae*, *Partitiviridae*, *Megabirnaviridae*, and *Chrysoviridae* families carry dsRNA genome (Ghabrial and Suzuki, 2009; Ghabrial *et al.*, 2015). Rhizidiovirus is classified under dsDNA viruses (Ghabrial and Suzuki, 2009). *Mycomononegaviridae* family own negative-sense RNA genome (Ghabrial *et al.*, 2015).

2.2.4.4. Vertebrate Viruses

Approximately 75-80% of the infectious diseases are originated from animals. Bats are known as the reservoirs for many viruses, such as Coronaviruses, Marburg virus, and Ebola viruses, that infect vertebrates (Wu *et al.*, 2016).

Nipah viruses are one of the vertebrate viruses that originated from bats. Nipah viruses have non-segmented negative-stranded RNA genome. They are enveloped and classified under paramyxoviruses (Banerjee *et al.*, 2019). Nipah viruses were first detected during an outbreak in Malaysia in 1999. Back then, pigs were thought to be the reason for human infections. Thus, more than 1 million pigs were killed to prevent the spread of viruses. Further studies showed that Nipah viruses are found in Malayan flying fox (*Pteropus vampyrus*) and island flying fox (*Pteropus hypomelanus*) (Johara *et al.*, 2001). Nipah virus-associated outbreaks occurred five times in Bangladesh during 2001-2005. Contrary to the outbreak in Malaysia, human to human transmission was observed in Bangladesh (ICDDR *et al.*, 2004). The sequence data indicated that virus transmission occurred via an intermediate host in Malaysia. However, sequences obtained from the viruses that caused outbreaks in Bangladesh and Malaysia show 0.8% difference (Harcourt *et al.*, 2005). Therefore, in Bangladesh, horizontal transmission in humans is thought to be the reason for the spread of Nipah virus (Breed, 2008).

Another vertebrate virus originated from bats is Marburg viruses from the *Filoviridae* family. They are suggested to originate from Egyptian rousette bat (*Rousettus aegyptiacus*) (Towner *et al.*, 2009). They carry their genomic information with negative sense ssRNA genome. These viruses can infect primates and mostly result in death. Marburg virus was first discovered during an outbreak in Europe in 1967. The source of this outbreak was the African green monkeys which were transported from Uganda to Europe. The transmission route of Marburg virus is unknown. Nevertheless, according to the cases in Uganda, direct contact with the bat fecal material might be one of the transmission routes (CDC, 2021).

Another virus from the *Filoviridae* family is Ebola viruses. These viruses were first detected in 1976 (Koch *et al.*, 2020). The natural host of this virus is thought to be bats from

Pteropodidae family (WHO, 2022). This virus has caused several outbreaks in many countries in Africa. Infected animals can spread Ebola virus via their body fluids (feces, saliva, urine, amniotic fluid, and semen) (CDC, 2022).

Coronaviruses are positive sense ssRNA viruses classified under the *Coronaviridae* family. Some members of coronaviruses circulate in humans. The natural host of human coronaviruses (NL63 and 229E) were thought to be African trident bat (*Triaenops afer*) and *Hipposideros* bats (Tao *et al.*, 2017). On the other side, some coronaviruses are the causative agents of outbreaks. In 2002, Severe Acute Respiratory Syndrome Coronaviruses (SARS-CoV) were identified for the first time in human in southern China (Drosten *et al.*, 2003). The most likely origin of this virus was Horseshoe bats (*Rhinolophus*) which is natural reservoir for SARS-related CoV (Li *et al.*, 2005; Hu *et al.*, 2017). SARS-CoV transmission to human occurs via an intermediate host, masked palm civet (*Paguma larvata*) (Graham *et al.*, 2013). Another outbreak was caused by Middle East Respiratory Syndrome Coronaviruses (MERS-CoV) which was isolated in 2012 in Saudi Arabia (Zaki *et al.*, 2012). MERS-CoV was spread into many countries. The transmission to human happens via an intermediate host which is dromedary camel (Milne-Price *et al.*, 2014). MERS-CoV might also be originated from bats. In one study, they inoculated Jamaican fruit bats (*Artibeus jamaicensis*) with MERS-CoV. They showed that MERS-CoV can infect Jamaican fruit bats which supports the hypothesis that the original source of MERS-CoV could be the Jamaican fruit bats (Munster *et al.*, 2016). Severe Acute Respiratory Syndrome Coronaviruses 2 (SARS-CoV-2) that was reported in China 2019 has been the causative agent of recent COVID-19 pandemic worldwide (Li *et al.*, 2020). At the beginning of the outbreak, the causative agent could not be identified. Further investigation led to seafood market and animals (WHO, 2021). Up to now, the closest coronavirus sequence (RaTG13) was obtained from the Horseshoe bats (*Rhinolophus*) in China. The sequence homology between the genomes of SARS-CoV-2 and RaTG13 was 96% (Zhou *et al.*, 2020). This information might enlighten the bat origin of SARS-CoV-2. Other studies also suggested that pangolins might be the possible intermediate host from bat to human (Lam *et al.*, 2020; Xiao *et al.*, 2020). However, the amino acid similarities between the receptor binding domains of spike proteins were found to be 89.2% for SARS-CoV-2 and RaTG13 (Liu *et al.*, 2021). The original source of SARS-CoV-2 is still unknown.

Viruses belonging to *Adenoviridae*, *Nodaviridae*, *Polyomaviridae*, *Rotaviridae*, *Poxviridae*, *Bunyaviridae*, *Herpesviridae*, *Papillomaviridae*, *Retroviridae*, *Circoviridae*, *Rhabdoviridae*, *Astroviridae*, *Flaviridae*, *Coronaviridae*, *Picornaviridae*, and *Parvoviridae* families that can infect vertebrates were also detected in rectal swab, pharyngeal swab, and fecal samples of bats (Li *et al.*, 2010; Wu *et al.*, 2016; Li *et al.*, 2021).

2.2.4.5. Bacteriophages

Bacteriophages are the viruses that use bacteria as hosts. Bacteriophages are classified under the *Caudovirales* order. Members of *Siphoviridae*, *Podoviridae*, *Myoviridae* families carry double-stranded DNA genome. *Inoviridae* and *Microviridae* families have single-stranded DNA genome. In bats, eukaryotic and prokaryotic viruses comprise the large part of the gut virome (Yinda *et al.*, 2018). Bacteriophages were associated with the microbiome of the bats (Zheng *et al.*, 2018). Therefore, phages in fecal material reflect the microbial community in the gut (Zheng *et al.*, 2018).

Bacteriophages in the *Siphoviridae*, *Podoviridae*, *Myoviridae*, and *Microviridae* families were detected in the guano samples of *Tadarida brasiliensis*, *Myotis velifer*, *Nycticeus humeralis*, *Perimyotis subflavus*, *Myotis californicus* and *Myotis yumanensis* (Li *et al.*, 2010; Li *et al.*, 2021). Members that belong to *Siphoviridae* and *Myoviridae* families were found in the fecal samples of *Cynopterus sphinx*, *Miniopterus schreibersii*, *Rousettus leschenaultii*, *Hipposideros larvatus*, *Rhinolophus blythi* and *Scotophilus kuhlii* (Zheng *et al.*, 2018). In another study, *Inoviridae* and *Microviridae* families were detected in rectal and pharyngeal swab samples of bats (Wu *et al.*, 2016).

2.2.5. Transmission Routes

Increase in human population results in the invasion of the zones that belonged to wild animals. Pathogen transmission occurs through two different ways: direct and indirect contact. Pathogens can infect and transmit to other animals and human populations via direct contact

with animals or body fluids (saliva, urine, feces and mucous) (CDC, 2022). In some cultures, humans share the same habitat with wild animal and interfere with wildlife by hunting or cutting trees (Parashar *et al.*, 2000). On the other hand, interaction with contaminated materials, intermediate host, or vectors are the disease transmission routes via indirect contact. For instance, consuming food and water that could be contaminated with infectious agents is one of the transmission routes via indirect contact. One of the six people in America consumes food which is contaminated by fecal material of infected animals (CDC, 2022). The other indirect contact route is an intermediate host (Figure 2). For instance, Nipah virus was detected in humans who worked at or lived in pig farms in Malaysia during 1998 to 1999. The phylogenetic analysis confirmed the homology of the viral sequences that were obtained from pigs and humans (Parashar *et al.*, 2000). The last indirect contact route is vector-borne transmission. Insects are extraordinary carriers for infectious agents. The vectors like flea, ticks, and mosquitoes can spread disease agents from one individual to another without getting infected (CDC, 2022). According to the World Health Organization (WHO), approximately 1 million human cases with fatality had been recorded until 2022 (WHO, 2022).

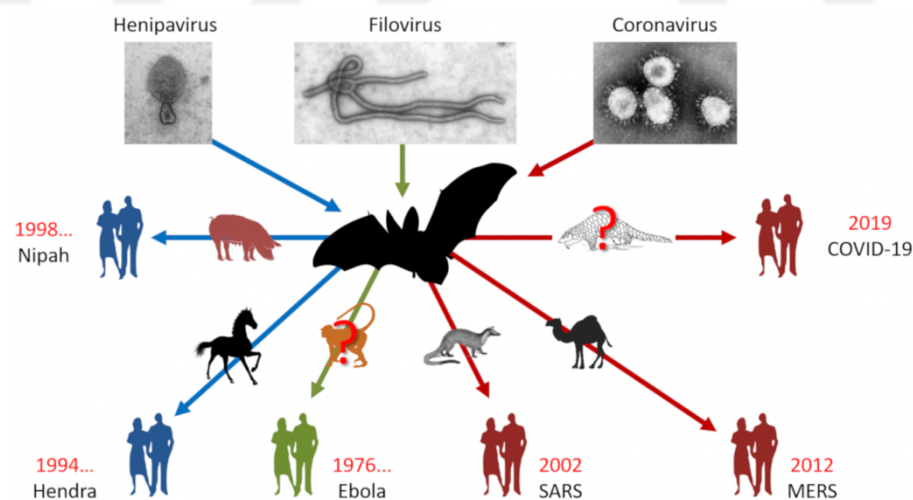


Figure 2.Transmission routes of bat-borne viruses. Bat-borne viruses have been transmitted via intermediate species (Adapted from: Gèrald, 2022).

In conclusion, disease transmission might occur through several routes such as contaminated surfaces, water or food, as well as body fluids and vectors (CDC, 2022). When

the pathogens overcome species barrier, contagious disease agents might be transmitted from human to human via several routes like air droplets or blood. Once the pathogens are established in the population, they can easily spread from country to country due to international travels, tourism, and global trade. Human populations have invaded many wildlife habitats, including the territories of bats. These mostly result from agricultural needs, touristic purposes, hunting, and rumors about treasures in caves. The bat roosts are vulnerable to disturbance by tourism, urbanization, and habitat fragmentation. Therefore, regular surveillance studies are crucial for the detection of novel pathogens that circulate in wildlife populations. Metagenomics approach provide a valuable resource for such surveillance studies.

2.3. Metagenomics

Until late 1970s, microbiological studies were limited to morphological identification and culturing. In 1977, Carl Woese presented the idea of using 16S ribosomal RNA (16S rRNA) molecule to investigate the evolutionary relationship between organisms because 16S rRNA changes slowly in time (Woese and Fox, 1977). Using 16S rRNA as a genetic marker along with the development of Sanger sequencing opened a new area in microbiology field to discover new metabolic products and genes (Sanger *et al.*, 1977; Woese and Fox, 1977). Since then, the developments in sequencing technologies have contributed to our understanding in genetics and its direct or indirect effect on the phenotype of organisms.

Maxam Gilbert and Sanger sequencing are the first-generation sequencing methods that were discovered in the same year. Nowadays, Maxam Gilbert sequencing is not preferred because of the toxic and radioactive agents used during process (Kchouk *et al.*, 2017). On the other hand, Sanger sequencing is a cost-effective and fast method. However, sequencing the targeted pieces in genomes along with poor read quality especially at the beginning and at the end of sequences are the main limitations of Sanger sequencing. The second-generation technologies are based on the synthesis or ligation. One of these is pyrosequencing that is based on pyrophosphate detection during synthesis. 454 sequencing platform, also uses pyrosequencing approach, is the first revolutionized technology in the genomics area

(Margulies *et al.*, 2005). In terms of yield and read length, 454 Sequencing Technology and Ion Torrent are similar technologies. Apart from other second-generation sequencing technologies, Illumina platforms (MiSeq, NextSeq etc.) have been used for several purposes. For instance, Large Whole Genome Sequencing, Small Whole Genome Sequencing, Targeted Gene Sequencing, 16S Metagenomic Sequencing, and Metagenomic Profiling (shotgun metagenomics, metatranscriptomics) are some of the applications that can be achieved by Illumina platforms. The third-generation sequencing technologies (MinION and Nanopore) provide rapid results with longer reads compared to the second-generation technologies. However, the high error rate is problem that should need to be overcome.

The sequencing technologies facilitate the identification of microorganisms. For a long time, microorganisms could only be characterized by culturing methods. With the high throughput sequencing, omics studies have been developed. Metagenomics provides a way to identify microbial community in a variety of samples, such as water, cheese, yogurt, human skin, fecal samples, wastewater, or soil, without using culture methods (NIH, 2022).

2.3.1. Viral Metagenomics

Bacteria and viruses can be pathogenic to human beings and these infections can spread from human to human that may cause an epidemic or a pandemic (Mohsin *et al.*, 2021). Even though the immune system, antivirals, and vaccines contribute to the fight against microorganisms, microbes can evolve fast (Mohsin *et al.*, 2021). Almost 75% of the emerging infectious diseases results from viruses that are originated from animals (Wang *et al.*, 2020).

Emergence of new variants, spillover from hosts, and close interactions between animals and humans increase the possibility of outbreaks (Mohsin *et al.*, 2021). The fact that pathogen transmission from animals to humans is inevitable due to global climate change and urbanization indicates the importance of surveillance studies in wildlife to detect novel disease agents. For many years, cell culturing, serological tests, immunological tests, electron microscopy, and polymerase chain reaction (PCR) methods were the main detection tools in

the field of virology (Mohsin *et al.*, 2021). Such methodologies are used for known viruses which limits novel virus detection (Mohsin *et al.*, 2021). Metagenomics introduces an alternative way to study the known and unknown microorganisms from environmental samples without culturing or serological assays using high throughput DNA sequencing technologies (Mohsin *et al.*, 2021). There have been many metagenomics studies in the literature that were conducted by researchers in different countries to identify the composition of the microbial community in human organs, sewage, ocean, soil, and fecal samples.

2.3.1.1. Shotgun Sequencing

Shotgun sequencing is a next generation sequencing (NGS) method that have been used to detect the entire genetic material of organisms in metagenomics studies without targeting specific region on the genome. This method consists of three steps: 1) sample collection and processing 2) sequencing of libraries 3) analysis of sequence reads (Quince *et al.*, 2017). Whole genome can be fragmented by various ways (acoustic shearing, nebulization, enzymatic shearing, etc.) based on the sequencing technology and fragmented reads can be identified in different databases after sequencing (Solonenko and Sullivan, 2013). Short sequence reads are taxonomically classified using bioinformatics tools. Therefore, this method provides an opportunity for the fast and accurate detection of novel genes and organisms, which also contributes to the detection of emerging viruses (Mohsin *et al.*, 2021). In one study using metagenomics approach, 54 virus families including dsDNA viruses, ssDNA viruses, dsRNA viruses, and ssRNA viruses were detected in bat fecal samples that were collected in China (Zheng *et al.*, 2018). In a study conducted with French insectivorous bats, viruses from *Retroviridae*, *Herpesviridae*, *Bunyaviridae*, *Poxviridae*, *Flaviviridae*, *Reoviridae*, *Bornaviridae* and *Picobirnaviridae* families along with novel viruses were detected in bat specimens (liver, brain, and lungs) (Dacheux *et al.*, 2014). In another study, virus families (*Parvoviridae*, *Picornaviridae*, *Totiviridae*, *Mononegavirus*, *Reoviridae*, *Nairoviridae*, *Phenuiviridae*, *Tymovirus*, and *Retroviridae*) were identified in different organs (brain, kidney, heart, salivary gland, intestine, liver, spleen, and lung) of German bats via metagenomics approach (Kohl *et al.*, 2021). Novel viruses, such as *Alphapapillomavirus*, *Betaretrovirus*,

Alpharetrovirus, *Varicellovirus*, *Cyprinivirus*, *Chlorovirus*, and *Cucumovirus* were discovered via shotgun metagenomics (Zheng *et al.*, 2018).

Bats are unique organisms hosting multiple virus families at the same time with no sign of disease or infection. Shotgun metagenomics is one of the methods for the detection of all viral organisms in samples without targeting specific region. However, taxonomic assignment is one of the challenges of viral metagenomics. Even sequencing technologies have had a remarkable improvement, insufficient viral sequences in databases restrict the classification of viruses. Therefore, in this thesis, bat fecal virome was investigated at virus family level via three different bioinformatics approach in samples collected from Cilingoz Cave in Tekirdag in Northwest of Anatolia, Sefer Yitigi Cave in Burdur in Central Anatolia, and Sarihidir Tunnel in Central Anatolia via shotgun metagenomics approach.

3. MATERIALS AND METHODS

3.1. Type of the Study

This thesis study is a viral metagenomics study from bat feces. In this thesis, fecal samples were collected from Cilingoz Cave (Tekirdag in Northwest Anatolia), Sefer Yitigi Cave (Burdur in Central Anatolia) and Sarihidir Tunnel (Nevsehir in Central Anatolia) in Türkiye to identify the virome profile in bat fecal samples in each location.

3.2. Time and Place of the Study

Collection of bat fecal samples was conducted from Cilingoz Cave in Tekirdag in 2018, from Sefer Yitigi Cave in Burdur and Sarihidir Tunnel in Nevsehir in 2020. After sample collection, all experimental steps were performed at the Emerging Viral Diseases Laboratory at Izmir Biomedicine and Genome Center, 15 July Health and Art Campus of Dokuz Eylul

University. All experimental steps up to inactivation were conducted in a Class II A2 Biosafety Cabinet at a Biosafety Level 2 (BSL2) laboratory between April 2021 and March 2022. Bioinformatic analyses were performed in June-August 2022.

3.3. Materials of the Study

In this thesis, bat fecal samples collected from Cilingoz Cave in Tekirdag in 2018, from Sefer Yitigi Cave in Burdur and Sarihidir Tunnel in Nevsehir in 2020 were used for analysis.

3.4. Variables of the Study

Independent variables of this study are the sampling locations, amount of sample collected, bioinformatics approach used for analysis.

The dependent variables of this study are the number of taxonomically assigned sequence reads, bat species at sampling locations, and virus families identified in bat species at each location and by each bioinformatics approach.

3.5. Tools for Data Collection

3.5.1. Materials used in the research

Table 1. List of plastic and glass consumables

Material	Brand	Catalog Number
PES Filter 0.45µm	Sigma-Aldrich	I.094.05.008
PES Filter 0.22µm	Sigma-Aldrich	I.094.07.008
50ml Centrifuge Tubes	Greiner	227261
Ultracentrifuge Tubes	Beckman Coulter	326814
Titanium Tube Cap Assembly	Beckman Coulter	341968
1.5 ml Tubes	Thermo Fisher Scientific	MBP-3448
1.5 ml DNA Lo-binding Tubes	Eppendorf	0030108051
0.2 ml Tubes	Thermo Fisher Scientific	AB0620
2 ml Nunc Cryotubes (external thread)	Greiner	126278
Tube Stand	Isolab	079.11.016
10 µl Micropipette Tips	Neptune	BIO-2040
Filtered Micropipette Tips (10 µl/20 µl/200 µl/1000 µl)	Neptune	BIO-BT10/20/200/1000
Serological Pipets (1 ml/2 ml/5 ml/10 ml/25 ml/50 ml)	SPL	I.083.03.001/002/005/010/025/050
1 L/500 ml/250 ml/100 ml Glass Bottles	Isolab	I.061.01.90/1/500/250/100

Graduated Cylinder 1 L/500 ml/250 ml/100 ml/ 50ml	Isolab	S.015.01.90/1/500/250/100/ 50
Gloves	BRP	S218099709NCZB



Table 2. List of kits, reagents, solutions, chemicals, and antibiotics

Material	Brand	Catalog Number
QIAamp MinElute Virus Spin Kit (QIAamp MinElute Columns, Collection Tubes (2 ml), Buffer AL, Buffer AW1 (concentrate), Buffer AW2 (concentrate), Buffer AVE (tubes with purple caps), Protease Resuspension Buffer, Carrier RNA (tubes with red caps), QIAGEN® Protease)	QIAGEN	57704
High Sensitivity DNA Kit (DNA Chips, Electrode Cleaner, Syringe Kit, Syringe, Spin Filters, High Sensitivity DNA Ladder, High Sensitivity DNA Markers 35/10380 bp, High Sensitivity DNA Dye Concentrate, High Sensitivity DNA Gel Matrix)	Agilent	AGT-5067-4626
Wizard® SV Gel and PCR Clean-Up System (Wizard SV Minicolumns, Collection Tubes, Membrane Wash Solution, Membrane Binding Solution, Nuclease-Free Water)	Promega	A9281
The SuperScript™ III First Strand cDNA Synthesis Kit (Oligo(dT) ₂₀ (50 µM), Random Hexamers (50 ng/µl), 10X RT Buffer, 0.1 M DTT, 25 mM Magnesium Chloride, 10 mM dNTP Mix, SuperScript™ III RT (200 U/µl), RNaseOUT (40 U/µl), <i>E. coli</i> RNase H (2 U/µl), DEPC-Treated Water, Total HeLa RNA (10 ng/µl), Sense Control Primer (10 µM), Antisense Control Primer (10 µM))	Invitrogen™	18080051
Invitrogen™ Second Strand cDNA Synthesis Kit (Second Strand Enzyme Mix, 5X Second Strand Reaction Mix, 0.5 M EDTA, pH 8.0, RNase I, 10 U/µL nuclease-free Water)	Invitrogen™	A48570

DNA/RNA Shield™	Zymo Research	R1100-50
RNase A	Macherey-Nagel™	740505.50
RNase-Free DNase I	New England Biolabs	M0303S
Q5® High-Fidelity 2X Master Mix (Q5 High-Fidelity 2X Master Mix, 10 µM Forward Primer, 10 µM Reverse Primer, Nuclease-Free Water)	New England Biolabs	M0492S
Invitrogen™ UltraPure™ Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v)	Thermo Fisher Scientific	15593031
Ammonium Acetate Solution	Sigma-Aldrich	A2706-100ML
PBS Tablets	Sigma-Aldrich	P4417
TAE Buffer 50X	BioShop	TAE222.1
Agarose	BioMax	HS-8000
Gel Loading Dye 6X	QIAGEN	239901
Safe View™	Abm	G108
100bp DNA Ladder	New England Biolabs	N3231S
1 kb DNA Ladder	New England Biolabs	N0468S
Ethanol (96%)	Alkomed	-
Micro-Chem Plus	NCL®	0255-29
Glycerol	Sigma-Aldrich	G2025
Penicillin G	Sigma-Aldrich	P3032-10MU
Streptomycin	Sigma-Aldrich	S9137
Gentamycin Sulfate	Sigma-Aldrich	G1914
Nystatin	Sigma-Aldrich	N6261-500KU
Polymyxin B	Sigma-Aldrich	P4932-1MU

Table 3. List of equipment

Equipment	Brand	Model
Ultracentrifuge	Beckman Coulter	L-100XP
-86°C freezer	Eppendorf	HEF U410
-20°C freezer	Bosch	GSN51AW30
+4°C fridge	Bosch	KSV36AI31
Thermal Cycler	Thermo Scientific TM	SimpliAmp
Gel Electrophoresis System	Thermo Scientific TM	Owl TM EasyCast TM B1A Mini Gel Electrophoresis Systems
Gel Electrophoresis System Power Supply	Bio-Rad	PowerPac TM
Gel Imaging Device	Bio-Rad	Molecular Imager Gel Doc XR System (170-8170)
Microwave	BEKO	MWB 200 EXN
Microcentrifuge	Thermo Scientific TM	MicroCL 17R
Fume Hood	Kötterman	2-401
Biological Safety Cabinets	Thermo Scientific TM	TS Safe 2020 Class IIA2 BSC
Waterbath	Nuve	NB 9
Magnetic Stirrer	Thermo Scientific TM	Cimarec SP88857105
Vortex Mixer	Thermo Scientific TM	LP Vortex Mixer
Balance	Sartorius	Entris 822-1S
Mini Centrifuge	Thermo Scientific TM	MySpin 6 Sprout-Mini Centrifuge
Micropipette Set (2µl-20µl; 20µl-200µl; 100-1000µl)	GILSON®	NEO Starter Pipet Kit
Micropipette Set (2µl-20µl; 20µl-200µl; 100-1000µl)	Eppendorf	Research® Plus
Micropipette (0.5µl-10µl)	Eppendorf	Research® Plus

Table 4. List of primers used for amplification

Primers	Sequence (5' to 3')	Melting Temperature (T _m)
K8N	GACCATCTAGCGACCTCCAC	55
Random hexamer	NNNNNN	-

3.5.2. Experimental Methods

3.5.2.1. Sample Collection

Bat fecal samples were collected from Cilingoz Cave (Tekirdag) in 2018, Sefer Yitigi Cave (Burdur) in 2020, and Sarihidir Tunnel (Nevsehir) in 2020 (Figure 3) with the permissions of the General Directorate of Nature Conservation and National Parks of the Republic of Türkiye, Ministry of Agriculture and Forestry (permission numbers: 72784983-488.04-178077 for 2018 and 693887891 for 2020). Approximately 25-30 pieces of bat feces were collected either from the harp trap or from the cave floor and placed into one tube (per location) containing 15 ml viral transport medium (Table 5). Then, the samples were transferred to the Emerging Viral Diseases Lab at Izmir Biomedicine and Genome Center by cold chain. Upon arrival, the samples were immediately transferred to a -86°C freezer and stored until further steps.



Figure 3. Sampling locations of the bat fecal samples. Sampling locations are shown in yellow circles. The first location is Cilingoz Cave in Tekirdag where fecal samples of *Myotis capacinii*, *Miniopterus schreibersii* and *Rhinolophus euryale* were collected. The second location is Sarihidir Tunnel in Nevşehir where fecal samples of *Miniopterus pallidus*, *Myotis myotis* and *Myotis capacinii* were collected. The third location is Sefer Yitigi Cave in Burdur where fecal samples of *Miniopterus schreibersii*, *Myotis myotis*, *Myotis capacinii* and *Rhinolophus ferrumequinum* were collected.

Table 5. Components of the viral transport medium.

Material	Amount (in 1L)
Glycerol (Autoclaved)	500ml
PBS (Autoclaved)	Up to 1L
Penicillin	1x 10 ⁶ U/L
Streptomycin	200 mg/L
Gentamicin Sulfate	150 mg/L
Polymyxin B	5000 U/L
Nystatin	25000 U/L

3.5.2.2. Sample Preparation

Samples were named after the location where they were collected (samples name: Cilingoz, Sarihidir, and Sefer Yitigi). As an internal control, 15 ml sterile Phosphate Buffered Saline (PBS) in a tube was placed in Class II A2 biosafety cabinets without cap for 1 hour to detect any potential lab-associated contamination, if any. After 1 hour incubation, PBS was taken and stored in a -20 °C freezer until next step. Then, the bat fecal samples from three different regions which were collected in tubes containing viral transport medium were placed in a water bath at 37°C for a quick thaw.

Samples were then transferred into a Class II A2 biosafety cabinet before their caps were opened. All samples were vortexed vigorously until they were homogenized. Then, the samples were centrifuged three times at 4.400 rpm for 10 minutes, 15 minutes, and 5 minutes, respectively. After each centrifugation, the supernatant was taken and transferred into a new 50 ml centrifuge tubes for the next centrifugation.

Supernatants were passed through 0.45 µm and 0.22 µm PES filters to remove the eukaryotic cells and bacteria found in the supernatant. However, the final volume of three samples varied at the end of centrifugation and filtration steps due to viscosity of each sample. This difference might result from the composition of the fecal samples in each location. For instance, the fecal sample collected from Cilingoz Cave was more viscous and the debris could not be precipitated well enough with centrifugation. This resulted in the clogging of the filtrates, therefore loss of sample volume in varying degrees in each sample during sample preparation step. The final volume of the samples from Sarihidir, Sefer Yitigi and Cilingoz was 10.5 ml, 7.5 and 2.5 ml, respectively. The final volume of each sample was adjusted to 10.5 ml by adding sterile PBS.

Filtrates for each sample were then carefully transferred into separate ultracentrifuge tubes with titanium caps to avoid the possibility of spillover during centrifugation. The samples were centrifuged at 184.000xg for 3 hours at 4°C. The pellets were resuspended in 169 µl

nuclease-free water to complete the final reaction volume up to 200 µl. Resuspended samples were simultaneously treated with 2 µl DNase I in 20 µl DNase I buffer and 9 µl RNase A at 37°C for 1h for enzymatic activity and then at 75°C for 10 minutes for enzyme inactivation. After incubation, 25 µl protease was added to each sample and incubated at room temperature for 30 minutes to inhibit the nuclease activities. Finally, 600 µl “DNA/RNA Shield™” was added to inactivate the viruses and to preserve the viral nucleic acids released during inactivation process. All procedures in sample preparation step were performed using aseptic techniques at a BSL2 laboratory which is used only by the members of the Emerging Viral Diseases Laboratory. All material was discarded after they were decontaminated in 5% Micro-Chem Plus solution.

3.5.2.3. Nucleic Acid Isolation and Concentration

Viral nucleic acids were isolated from each sample (PBS- internal control, Ciligoz, Sarihidir, and Sefer Yitigi) using “QIAamp MinElute Virus Spin Kit” (QIAGEN) using the manufacturer’s instructions with minor modifications. First, 225 µl of each sample was transferred into a tube containing 200 µl AVL buffer (including 28 µg/ml carrier RNA) to disrupt the viral envelope and reveal the viral nucleic acids. Tubes were vortexed for 15 seconds for mechanical disruption and incubated at room temperature for 15 minutes for efficient lysis. After incubation, tubes were briefly centrifuged at 8000 rpm for the removal of drops on the lid. Then, 250 µl of 96% ethanol (EtOH) was added into each tube and homogenized by a brief vortex. The solution was incubated at room temperature for 15 seconds. Then, the solution of each sample was loaded into MinElute columns and incubated for 1 min to increase the nucleic acid adsorption. Tubes were centrifuged at 8000 rpm for 1 minute and flowthroughs were discarded. 500 µl Wash I buffer was added to each column and centrifuged at 8000 rpm for 1 minute. Collection tubes were removed and replaced with new ones. Then, 500 µl Wash II buffer was added to each column and centrifuged at 8000 rpm for 1 minute. Collection tubes were removed and replaced with new ones. Finally, 500 µl of 96% EtOH was loaded into each column and centrifuged at 8000 rpm for 1 minute. Flowthroughs were discarded. To remove

the EtOH remains, tubes were centrifuged at 13300 rpm for 3 minutes. Next, they were incubated at room temperature for 1 minute and spun for 1 minute at 13300 rpm. Finally, 150 µl of elution buffer was added to each column to elute. Nucleic acids were collected in sterile tubes by centrifugation at 8000 rpm for 1 minute.

Next, to concentrate the amount of genetic materials, nucleic acids were precipitated using ethanol precipitation method. First, 150 µl of 96% EtOH was mixed with 15 µl sodium acetate followed by the addition of 375 µl of 96% EtOH. Next, the solution was placed in a -86°C freezer overnight. After overnight incubation, the solution was centrifuged at 17000xg for 30 minutes at +4°C. Then, the supernatant was removed and suspended with 200 µl of 80% EtOH. Tubes were centrifuged at 10000xg for 10 minutes at +4°C. Supernatants were removed, and tubes were incubated for 15 minutes at room temperature to evaporate residual EtOH. Finally, nucleic acids were resuspended in 25 µl nuclease-free water.

3.5.2.4. cDNA Synthesis

3.5.2.4.1. First-Strand cDNA Synthesis

First strand cDNA synthesis was completed using “SuperScript™ III First-Strand Synthesis System” (Invitrogen™) according to the manufacturer’s manual with minor modifications. First, 8 µl nucleic acid from each sample was incubated at 95°C for 3 minutes, then at 4 °C for 2 minutes to denature the double-stranded RNA or DNA. Then, 1 µl random hexamer (50 ng/µl) and 1 µl 10 mM dNTP were added onto the denatured nucleic acids. The samples were incubated at 65°C for 5 min to anchor the random hexamers to the template. Next, the samples were incubated at 4°C for 2 minutes to cool down. The reaction mixture was prepared by adding 2 µl 10X RT Buffer, 4 µl 25 mM MgCl₂, 2 µl 0.1 M DTT, 1 µl RNaseOUT (40 U/µl), and 1 µl SuperScript™ III RT (200 U/µl) to the template anchored with random hexamers. Then the reactions were incubated at 25°C for 10 minutes for annealing, 50°C for 50 minutes for replication and 85°C for 5 minutes to end the reaction. After incubation steps, the tubes containing reaction mixtures were placed on ice for 2 minutes. Finally, 1 µl RNase H was

added into each reaction and incubated at 37°C for 20 minutes followed by at 4°C for 2 minutes to remove the RNA strand on DNA-RNA complex.

3.5.2.4.2. Second-Strand cDNA Synthesis

Second strand cDNA synthesis was performed immediately after the first strand cDNA synthesis using “Invitrogen™ Second Strand cDNA Synthesis Kit” (Invitrogen™) according to the manufacturer’s manual. Entire reaction from the first strand cDNA synthesis of each sample was used as a template. 54 µl nuclease-free water, 20 µl 5X Second Strand Reaction Mix and 5 µl Second Strand Enzyme Mix was mixed and transferred directly into tube containing the first strand cDNA. Then, the reaction mixture for the second strand cDNA synthesis was incubated at 16 °C for 1 hour.

3.5.2.5. cDNA purification

Double-stranded cDNAs obtained from each sample were purified to remove the remnants of reactions. Therefore, the samples were mixed with UltraPure™ Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) (Invitrogen™) in 1:1(v/v) ratio and vortexed for 20 seconds under the fume hood. Then, the samples were centrifuged at 16000xg for 5 minutes to separate the nucleic acids from organic compounds. After centrifugation, 75 µl was taken from the aqueous phase without touching the layer and transferred into a new 1.5 ml tube. Then, to remove the phenol:chloroform:isoamyl alcohol from the samples and concentrate the nucleic acids simultaneously, each cDNA was mixed with ammonium acetate solution and 100% EtOH at 1:0.5 (v/v) and 1:2.5 (v/v) ratios, respectively. Then the mixture was incubated at -20°C for overnight. After overnight incubation, the samples were centrifuged at 16000xg at 4°C for 30 minutes. The supernatant was carefully removed without touching the pellet. Then, the pellet was resuspended with 150 µl of 70% EtOH by multiple pipet movements. Sample was centrifuged at 4°C for 2 minutes at 16000xg and supernatant was removed. The tubes were left open at an upside-down position at room temperature for 15

minutes to get rid of the remaining EtOH. Finally, the samples were resuspended in 20 μ l nuclease-free water.

3.5.2.6. PCR Enrichment

To enrich the amount of nucleic acids, double-stranded cDNAs were amplified via PCR using K8N primers targeting 20 nucleotide long region of RdRp gene (Wu *et al.*, 2012; Wu *et al.*, 2016; Zheng *et al.*, 2018). K8N primer can hybridize with both DNA strand at the proper temperature. Therefore, only one primer was used for PCR enrichment. To obtain 50 μ l reaction volume, a master mix were prepared by adding 14 μ l nuclease-free water, 25 μ l Q5 High-Fidelity 2X Master Mix (NEB) and 1 μ l K8N primer as shown in Table 6. Then, the samples were amplified on the thermal cycler according to the conditions in Table 7. Briefly, the reaction conditions were as follows: one cycle at 98 °C for 30 seconds for initial denaturation, 35 cycles of denaturation at 98 °C for 10 seconds, annealing at 55 °C for 30 seconds and elongation at 72 °C for 2 minutes, followed by one cycle of final extension at 72 °C for 2 minutes.

Table 6. Content of the reaction mixture

Reagents	50 μ l reaction	Final Concentration
Nuclease-free water	Up to reaction volume	-
Q5 High-Fidelity 2X Master Mix	25 μ l	1X
K8N primer	1 μ l	0.5 μ M
Template	10 μ l	<1,000 ng

Table 7. PCR conditions

Step	Temperature	Time	Cycle number
Initial denaturation	98°C	30 seconds	1
Denaturation	98°C	10 seconds	35
Annealing	55°C	30 seconds	
Extension	72°C	2 minutes	
Final extension	72°C	10 minutes	1

3.5.2.7. Agarose Gel Electrophoresis

Following PCR enrichment, the amplicons were visualized using agarose gel electrophoresis. 100 ml 1X TAE buffer was measured using graduated cylinder and transferred into a glass bottle. Then, 2 gr agarose was weighed on the balance and added into TAE buffer to obtain a 2% agarose gel. After dissolving agarose in a microwave, 5 µl Safe View was added and gently mixed. Then, the agarose gel was solidified in a casting tray with a comb. The comb was gently removed from the solidified gel, which then was transferred into a gel tank which was filled with 1X TAE buffer.

0.5 µl of 6X loading dye per PCR product was placed on a piece of parafilm. Then, 2.5 µl PCR product was mixed with loading dye by pipetting and loaded into a well on the gel. Samples were run on the agarose gel at 95 V for 45 minutes. The amplicons were visualized using the Gel Imaging Device (Bio-Rad).

3.5.2.8. PCR Purification

To remove the PCR leftovers, the samples were purified using “Wizard SV Gel and PCR Clean-Up System” (Promega) according to the manufacturer’s protocols. First, the membrane binding solution was added to 37.5 µl amplicon in 1:1 ratio (v/v) and gently mixed by tapping. Then, the mixture was transferred into a column and incubated for 3 minutes at room temperature. The column was centrifuged at 16000xg for 1 minute. Then, 700 µl membrane wash solution was added and centrifuged at 16000xg for 1 minute. Next, 500 µl membrane wash solution was added and centrifuged at 16000xg for 5 minutes. The flowthrough was discarded after each centrifugation step. The column was centrifuged at 16000xg for 1 minute to remove the residual EtOH. Then, the column was placed into a new 1.5 ml tube. Finally, 50 µl nuclease-free water was added on the column and incubated for 1 minute. Then, purified nucleic acids were collected by centrifugation at 16000xg for 1 minute. After purification step, the samples were immediately sent to a company for sequencing.

3.5.2.9. Quality Control

The quantity and purity values were detected via Qubit (Thermo Fisher, ABD) by the sequencing company. In this process, fluorescent dyes were attached to desired product and detected by the Qubit fluorometers. First, 190 µl working solution was added into 1.5 ml tubes for two Qubit standards and completed up to 200 µl volume adding 10 µl standards. For each sample, 198 µl working solution and 2 µl sample were added into 1.5 ml tubes. All samples were mixed by vortex and incubated at room temperature for 2 minutes. Then, quantity and purity were measured via Qubit. Samples above the quantity and purity limit required for library preparation were used.

3.5.2.10. Library Preparation

Samples with enough quantity and purity were tagged, amplified, cleaned up and pooled for the construction of libraries using “Illumina DNA Prep kit” (Illumina) according to the manufacturer’s instructions. This step was performed by the sequencing company responsible for the next generation sequencing.

Bead-Linked Transposomes (BLT) and Tagmentation Buffer 1 (TB1) were used for tagging the DNA with adapter sequences. Samples were loaded into each well of 96-well PCR plate. Meanwhile, BLT was vortexed for 10 seconds for resuspension. Then, 11 µl BLT and 11 µl TB1 were mixed and vortexed for the preparation of tagmentation master mix. Master mix was divided into wells that included equal amount of samples (20 µl per reaction) and mixed with several micropipette movements. Plate was sealed and incubated at 55°C for 15 minutes on a thermal cycler.

Tagment Stop Buffer (TSB) and Tagment Wash Buffer (TWB) were used to wash the adapter-tagged DNA. For that, 10 µl TSB was added into the reaction and pipetted several times for resuspension of the beads in the reaction. Then, the plate was sealed and incubated at 37 °C for 15 minutes on a thermal cycler. After incubation, the plate was transferred on the magnetic stand and waited until reaction solution appeared clear approximately 3 minutes. Then, the supernatant was removed. The plate was transferred from the magnetic stand to the bench. 100 µl TWB was added to the beads attached to DNA and pipetted until resuspension of beads. The plate was transferred on the magnetic stand and waited until reaction solution appeared clear approximately 3 minutes again. Then, the plate was removed from the magnetic stand and 100 µl TWB was added to the beads attached to DNA and pipetted until resuspension of beads. This step was performed twice.

Enhanced PCR Mix (EPM) and index adapters (index 1- i7 and index 2 – i5) were used for the amplification of tagmented DNA. 22 µl EPM and 22 µl nuclease-free water was mixed (PCR master mix) via vortex and centrifuged at 280xg for 10 seconds. The supernatant was

removed from the wells on the plate from the tagmentation step on the magnetic stand. Then, the plate was transferred onto the bench and 40 µl PCR master mix was added into each well immediately. After that, the plate was centrifuged at 280xg for 3 seconds. Then, index adapters were added for each sample and mixed at 1600 rpm for 1 minute. The plate was centrifuged at 280xg for 30 seconds and amplified via PCR. Briefly, the reaction conditions were as follows: one cycle at 68 °C for 3 minutes, one cycle of initial denaturation at 98 °C for 3 minutes, 5 cycles of denaturation at 98 °C for 3 minutes, annealing at 62 °C for 30 seconds and elongation at 68 °C for 2 minutes, followed by one cycle of final extension at 68 °C for 1 minute.

Sample Purification Beads (SPB) and Resuspension Buffer (RSB) were used for cleaning up the libraries. After PCR amplification, the plate was centrifuged at 280xg for 1 minute for the collection of the samples at the bottom. Plate was transferred on the magnetic stand for approximately 5 minutes and waited until reaction solution appeared clear. 45 µl supernatant was taken from the plate and transferred into a new plate. Then, 40 µl nuclease-free water and 45 µl SPB which was vortexed and inverted several times were added respectively into each well that included supernatant. The plate was incubated at room temperature for 5 minutes and transferred on the magnetic stand until obtaining clear reaction. Then, 15 µl SPB was added to each well. All the samples were collected from the wells and transferred into wells on a new plate. The plate was incubated at room temperature for 5 minutes. At the end of incubation, the plate was transferred on the magnetic stand and waited until the liquid appears clear. The supernatant was discarded, and the wells were washed twice with 80% EtOH. To remove leftover EtOH, the plate was incubated on the magnetic stand for 5 minutes. Then, the plate was transferred on the bench and 32 µl RSB was added to each well and mixed with several pipetting. The plate was incubated at room temperature for 2 minutes and kept on the magnetic stand for 2 minutes. 30 µl sample was collected from each well and the quantity was checked via Qubit before sequencing process.

3.5.2.11. *Next Generation Sequencing*

Libraries were first denatured, then diluted before sequencing. Prepared libraries were sequenced on the Illumina NextSeq™ 500 platform by the sequencing company.

3.5.2.12. *Sequence Analysis and Assembly*

150 bp-long, paired-end sequence reads were received from the sequencing company as raw data (in FastQ format). The low-quality sequence reads, and adapters were removed using Trimmomatic v.0.32 (Bolger *et al.*, 2014). All the sequences obtained from three locations were submitted to NCBI GenBank with a BioProject ID: PRJNA875251. The BioSample accession numbers for the Cilingoz Cave, Sefer Yitigi Cave and Sarihidir Tunnel were SAMN30609335, SAMN30609336 and SAMN30609337, respectively. Then, sequence reads were investigated using three different approaches: 1) taxonomic identification of sequence reads using BLAST on CLC Genomics Workbench 22.0.1. 2) taxonomic identification of sequence reads via Kraken 2 (Wood *et al.*, 2019), 3) taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on NCBI BLAST.

3.5.2.13. *Taxonomic Assignment of Sequence Reads via BLAST*

11.623.082 paired-end sequence reads (4.078.204 reads for Cilingoz Cave, 3.863.402 reads for Sefer Yitigi Cave, 3.681.836 reads for Sarihidir Tunnel) were directly searched in BLAST database by CLC Genomics Workbench 22.0.1. with default parameters (Max target sequences: 100; Expect threshold: 0.05; Word size: 11; Match/Mismatch Scores: 2,-3; Gap costs: Existence:5 Extension:2). Then, BLAST hit table was generated that included the first BLAST hit for each sequence, E-value, alignment score and the best-aligned organism. Then, the best-aligned organism was searched in the NCBI Taxonomy to find information about the taxonomical classification of viruses. Assigned organisms were also searched in the databases

of the International Committee on Taxonomy of Viruses (ICTV) to obtain information about the host and nucleic acid type.

3.5.2.14. *Taxonomic Assignment of Sequence Reads via Kraken 2*

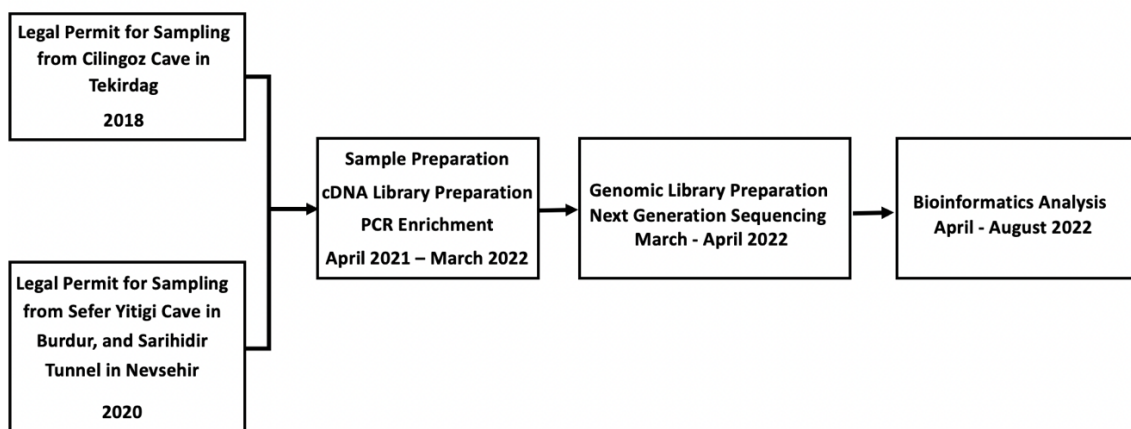
11.623.082 paired-end sequence reads (4.078.204 reads for Cilingoz Cave, 3.863.402 reads for Sefer Yitigi Cave, 3.681.836 reads for Sarihidir Tunnel) were taxonomically labelled via Kraken 2 (Wood *et al.*, 2019) program. This program investigates the k-mers in the sequences. Then, assigned k-mers are used to determine the lowest common ancestor for each sequence. A k-mer value of 31 (31 nucleotide-long sequences) was used in taxonomic assignment of the sequences in this study. After taxonomic assignment, the data table was generated. This table included the family, genus, and species information of assigned viral sequences. In this thesis virus families were investigated. Because of the short length of sequences, species level assignment might not be precise as family level assignment. Assigned organisms were also searched in the databases of the NCBI Taxonomy for family assignment and the ICTV to find information on the host and nucleic acid type.

3.5.2.15. *Taxonomic Assignment of the Contigs that were Obtained by de novo Assembly*

11.623.082 paired-end sequence reads (4.078.204 reads for Cilingoz Cave, 3.863.402 reads for Sefer Yitigi Cave, 3.681.836 reads for Sarihidir Tunnel) were assembled by *de novo* approach to construct contigs using the Microbial Genomics Module of CLC Genomics Workbench 22.0.1. Each contig (838 contigs for Cilingoz Cave, 777 contigs for Sefer Yitigi Cave and 1760 contigs for Sarihidir Tunnel) were screened in selected databases in BLAST with default parameters (Max target sequences: 100; Expect threshold: 0.05; Word size: 11; Match/Mismatch Scores: 2,-3; Gap costs: Existence:5 Extension:2). The databases were selected according to their relatedness with the study material to increase the hit numbers for assigned sequences and to avoid mismatches from bacterial or eukaryotic organisms. Thus, metagenome-related and virus-related databases were mainly selected. These specific databases were as follows: ‘viral metagenome (taxid:1070528), viruses (taxid:10239), RNA viruses and

viroids (taxid:2559587), RNA viruses (taxid:2559587), negative-strand RNA viruses (taxid:2497569), single-stranded DNA viruses (taxid:2731342), negative-sense genome single-stranded RNA viruses (taxid:11157), unclassified sequences (taxid:12908), unclassified Siphoviridae (taxid:196894), unclassified (taxid:12908), unclassified DNA viruses (taxid:2204151), unclassified RNA viruses ShiM-2016 (taxid:1922348), dsRNA viruses (taxid:35325), unclassified dsRNA viruses (taxid:39780), dsRNA virus environmental sample (taxid:1075826), metagenomes (taxid:408169), viral metagenome (taxid:1070528), ecological metagenomes (taxid:410657), geographical metagenomes (taxid:410657), marine metagenome (taxid:408172), fecal metagenome (taxid:1861841), soil metagenome (taxid:410658), plant metagenome (taxid:1297885), insects (taxid:6960)'. After obtaining BLAST results, a data table was constructed that included the first BLAST hit for each sequence, E-value, alignment score and the best aligned organism. Then, each assigned organism was searched in the NCBI Taxonomy to obtain information about the viral family. Assigned organisms were also searched in the databases of the ICTV to find information on the host and nucleic acid type.

3.6. Study Plan Calendar



3.7. Data Evaluation

Sequences were evaluated separately for each region. The sequences were searched in databases; and the relationships between these viruses and the observed bat species in the region were investigated.

3.8. Limitations of the Study

The main limitation of this thesis is insufficient viral sequences in the publicly available databases. Therefore, some of the sequence reads could not be taxonomically labeled even at a family level. Another limitation is that virome analysis could not be achieved individually for each bat or each species in the populations. In addition, samples were collected in different years (2018 and 2020) and from locations in long distances. Thus, these factors limit the comparison of viruses in each sampling locations in terms of beta and alpha diversity. Another limitation is high number of bacteriophages in feces due to high bacterial content in the gut microbiome. This abundancy might result in bias towards DNA viruses. To eliminate this, cDNAs were enriched by a PCR with a primer targeting RdRp gene of the RNA viruses. However, the low viral load in the samples was another limitation as the initial nucleic acid concentration in each sample was 0.27 ng/μl, 0.18 ng/μl, 0.15 ng/μl for Cilingoz Cave, Sefer Yitigi Cave and Sarihidir Tunnel, respectively. This might be due to the amount or condition of the fecal samples as well as loss of some viruses during centrifugation and/or filtration for the removal of debris and larger organisms.

3.9. Ethics Committee Approval

This thesis study was approved by the Non-interventional Research Ethics Committee of İzmir Biomedicine and Genome Center (IBG-GOEK) (No: 2020-031).

3.10. Legal Permits

Bat fecal samples were collected from Cilingoz Cave in Tekirdag in 2018, Sefer Yitigi Cave in Burdur, and Sarihidir Tunnel in Nevsehir in 2020 with the scientific research permits (No: 72784983-488.04-178077 for 2018 and No: 693887891 for 2020) granted by the General Directorate of Nature Conservation and National Parks, Republic of Türkiye Ministry of Agriculture and Forestry. The analysis of the collected samples was carried out at the Emerging Viral Diseases Laboratory at Izmir Biomedicine and Genome Center with the scientific research permit (No: 75362993-972) granted from the General Directorate of Food and Control, Republic of Türkiye Ministry of Agriculture and Forestry on 04.11.2019.

4. RESULTS

4.1. Sample Collection

Bat fecal samples were collected as explained in Section 3.5.2.1. The samples used in this study resembles the bat populations in Northwest (Cilingoz Cave, Tekirdag) and Central Anatolia region (Sefer Yitigi Cave, Burdur and Sarihidir Tunnel, Nevsehir) in 2018 and 2020, respectively. The species in bat colonies varied in different locations. Totally, six different bat species were observed in all locations during sample collection. The information on the observed species per location were given in Table 8. *Myotis capaccinii*, *Miniopterus schreibersii* and *Rhinolophus euryale* were observed in Cilingoz Cave. *Miniopterus schreibersii*, *Myotis myotis*, *Myotis capaccinii* and *Rhinolophus ferrumequinum* were observed in Sefer Yitigi Cave. *Miniopterus pallidus*, *Myotis myotis* and *Myotis capaccinii* were observed in Sarihidir Tunnel. *Myotis capaccinii* was recorded in three locations. *Myotis myotis* was recorded in Sefer Yitigi Cave and Sarihidir Tunnel.

Table 8. Bat species recorded in sampling locations

Location	Family	Species	Common name	Diet
Cilingoz Cave, Tekirdag (2018)	<i>Vespertilionidae</i>	<i>Myotis capacinii</i>	Long-fingered bat	Insectivorous
		<i>Miniopterus schreibersii</i>	Common bent-winged bat	Insectivorous
	<i>Rhinolophidae</i>	<i>Rhinolophus euryale</i>	Mediterranean horseshoe bat	Insectivorous
Sefer Yitigi Cave, Burdur (2020)	<i>Vespertilionidae</i>	<i>Miniopterus schreibersii</i>	Common bent-winged bat	Insectivorous
		<i>Myotis myotis</i>	Greater mouse-eared bat	Insectivorous
		<i>Myotis capacinii</i>	Long-fingered bat	Insectivorous
	<i>Rhinolophidae</i>	<i>Rhinolophus ferrumequinum</i>	The greater horseshoe bat	Insectivorous
Sarihidir Tunnel, Nevsehir (2020)	<i>Vespertilionidae</i>	<i>Miniopterus pallidus</i>	Pale bent-winged bat	Insectivorous
		<i>Myotis myotis</i>	Greater mouse-eared bat	Insectivorous
		<i>Myotis capacinii</i>	Long-fingered bat	Insectivorous

4.2. Molecular Detection of Viral cDNAs in Bat Fecal Samples

Before isolating the viral nucleic acids, samples were exposed to several steps of centrifugation and filtration to eliminate bacteria, eukaryotic organisms, and/or environmental residues from the samples. Then, viral particles in the samples were concentrated via ultracentrifugation followed by nuclease treatment to remove free nucleic acids in the environment. Viral particles were inactivated and preserved in a solution to protect the total nucleic acids in each sample. To prevent the denaturation of total nucleic acids in each sample, especially RNA, first and second strands of cDNA were synthesized and purified immediately after nucleic acid isolation.

The gut microbiome consists of various bacteriophage families which are generally DNA viruses. Thus, to avoid bias towards DNA phages and DNA viruses in fecal samples, cDNAs were also enriched with K8N primer to increase the possibility of detection of RNA viruses. K8N primer, which targets a region of RdRp, have been used in metagenomics studies (Wu *et al.*, 2012; Wu *et al.*, 2016; Zheng *et al.*, 2018). The final amplicon yield before library preparation for sequencing was visualized on a 2% agarose gel (Figure 4).

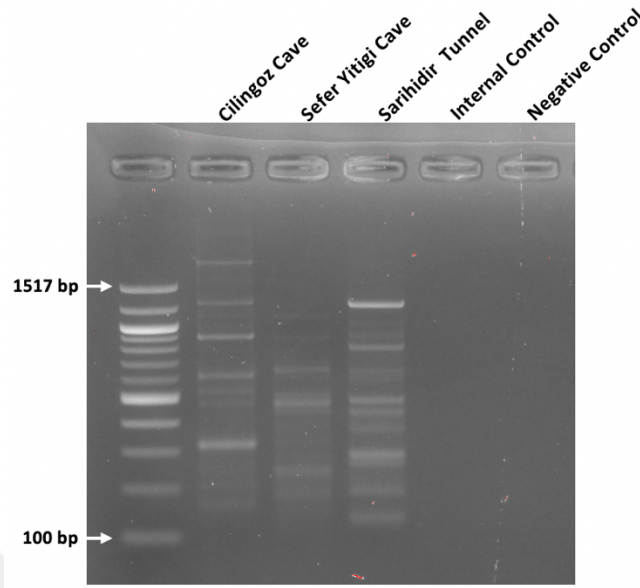


Figure 4. dscDNAs after PCR enrichment. First and second strand synthesis of cDNA was followed by a PCR enrichment using K8N primer targeting RdRp to enrich RNA viruses, as well. Purified dscDNAs were visualized on a 2% agarose gel that was run at 90 volt for 45 minutes. Samples were Cilingoz Cave, Sefer Yitigi Cave, and Sarihidir Tunnel from left to right. Sterile PBS that was kept open in the biosafety cabinet was used as an internal control for any laboratory-associated contamination and sterile PBS was used as a negative control. 100 bp DNA ladder was used as a reference.

4.3. Sequence Analysis

Genomic libraries were prepared from cDNAs that were enriched with K8N primer for each location except the internal and the negative control due to no amplicon (Figure 4). The concentrations of enriched cDNAs before genomic library preparation were 48.8 ng/ μ l, 36.5 ng/ μ l, 62.7 ng/ μ l for Cilingoz, Sefer Yitigi and Sarihidir Tunnel, respectively. Then, the genomic libraries were sequenced on the Illumina NextSeq500 platform. After removal of adapter and low-quality sequences, a total of 11.623.082 paired-end sequence reads were generated (4.078.204 reads for Cilingoz, 3.863.402 reads for Sefer Yitigi, 3.681.836 reads for Sarihidir Tunnel). Paired-end sequence reads were investigated using three different

approaches: 1) taxonomic identification of sequence reads using BLAST on CLC Genomics Workbench 22.0.1. 2) taxonomic identification of sequence reads via Kraken 2 (Wood *et al.*, 2019), 3) taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on NCBI, BLAST.

4.3.1. Taxonomic Assignment of Sequence Reads via BLAST

The paired sequence reads were directly searched in the BLAST database by CLC Genomics Workbench 22.0.1. with the default parameters (Max target sequences: 100; Expect threshold: 0.05; Word size: 11; Match/Mismatch Scores: 2,-3; Gap costs: Existence:5 Extension:2). According to the BLAST results, a data table was generated that included the first BLAST hit for each sequence, E-value, alignment score and the best aligned organism. Then, each assigned organism was searched in the NCBI Taxonomy to obtain information about the viral family. Assigned organisms were also searched in the databases of the ICTV to find information on the host and nucleic acid type.

4,078,204 paired-end sequence reads were recorded for the bat fecal samples collected from Cilingoz Cave. After BLAST search, 4420 hits (0.11% of total reads) were obtained. Among those hits, 96.83% could be assigned taxonomically, while no information was obtained for 3.17% of the hits. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 9. Overall, viral sequences were grouped in 8 different virus families as follows: *Autographiviridae*, *Myoviridae*, *Siphoviridae*, *Podoviridae*, *Polydnenviriformidae*, *Baculoviridae*, *Iridoviridae* and *Parvoviridae*. The viral families could not be assigned for 11.7% of the hits using NCBI Taxonomy and ICTV, therefore, they were classified as ‘Unclassified Viruses’.

The majority of the viral sequences from Cilingoz Cave were detected as bacteriophage families (*Autographiviridae*, *Siphoviridae*, *Myoviridae* and *Podoviridae*) along with unclassified groups (Uncultured Caudovirales Phage, Unclassified Caudovirales, Unclassified Bacterial Viruses). Based on the sequences with taxonomical identification, *Morganella phage*

in the *Autographiviridae* family that infects *Morganella sp.* was the most common virus in bat fecal samples in Cilingoz Cave. Aside from bacteriophages, invertebrate viruses were the second most abundant group (Table 9). Several sequence reads that were similar to *Turdus hortulorum* Iteradensovirus from *Parvoviridae* was the only virus that can infect vertebrates among all taxonomically labelled sequences. *Parvoviridae* was the only virus family that can infect both invertebrates and vertebrates that was observed in Cilingoz Cave. All viral sequences were found to belong dsDNA viruses, except for *Parvoviridae* which are ssDNA viruses (Table 9).

Table 9. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Cilingoz Cave at Family Level via BLAST, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Cilingoz Cave	<i>M. schreibersii</i> <i>M. capacinii</i> <i>R. euryale</i>	<i>Autographiviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Podoviridae</i>	dsDNA	Bacteria
		Uncultured Caudovirales Phage	dsDNA	Bacteria
		<i>Polydnaviriformidae</i>	dsDNA	Invertebrates
		<i>Baculoviridae</i>	dsDNA	Invertebrates
		<i>Iridoviridae</i>	dsDNA	Invertebrates
		<i>Parvoviridae</i>	ssDNA	Vertebrate-Invertebrates
		Unclassified caudovirales	dsDNA	Bacteria
		Unclassified bacterial viruses	dsDNA	Bacteria
		Unclassified viruses	*	*

3.863.402 paired-end sequence reads were recorded for the bat fecal samples collected from Sefer Yitigi Cave. After BLAST search, 1500 hits (0.04% of total reads) were obtained. Among those hits, 14.93% could be assigned taxonomically, while no information was obtained for 85.07% of the hits. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 10. Overall, viral sequences were grouped in 10 different virus families as follows: *Siphoviridae*, *Herelleviridae*, *Myoviridae*, *Podoviridae*, *Inoviridae*, *Ackermannviridae*, *Phycodnaviridae*, *Polymycoviridae*, *Ascoviridae*, *Baculoviridae*. The viral families could not be assigned for 6.7% of the hits using NCBI Taxonomy and ICTV, therefore, they were classified as ‘Unclassified Viruses’.

The majority of the viral sequences from Sefer Yitigi Cave were detected as bacteriophage families (*Siphoviridae*, *Herelleviridae*, *Myoviridae*, *Podoviridae*, *Inoviridae* and *Ackermannviridae*) along with unclassified groups (Unclassified Bacterial Viruses and Uncultured Caudovirales Phage). Aside from bacteriophage families, two virus families (*Ascoviridae* and *Baculoviridae*), which infect invertebrates, were recorded. *Phycodnaviridae* and *Polymycoviridae* families were also detected that infect algae and fungi, respectively. Majority of the viral sequences were found to belong dsDNA viruses, however, there was also virus families that belong to ssDNA and dsRNA viruses (Table 10).

Table 10. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Sefer Yitigi Cave at Family Level via BLAST, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Types	Host Range
Sefer Yitigi Cave	<i>M. schreibersii</i> <i>M. myotis</i> <i>M. capaccinii</i> <i>R. ferrumequinum</i>	<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Herelleviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Unclassified Bacterial Viruses</i>	dsDNA	Bacteria
		<i>Uncultured Caudovirales Phage</i>	dsDNA	Bacteria
		<i>Podoviridae</i>	dsDNA	Bacteria
		<i>Inoviridae</i>	(+)ssDNA	Bacteria
		<i>Ackermannviridae</i>	dsDNA	Bacteria
		<i>Phycodnaviridae</i>	dsDNA	Algae
		<i>Polymycoviridae</i>	dsRNA	Fungi
		<i>Ascoviridae</i>	dsDNA	Invertebrates
		<i>Baculoviridae</i>	dsDNA	Invertebrates
		<i>Unclassified viruses</i>	*	*
		<i>unclassified dsDNA viruses</i>	dsDNA	*

3.681.836 paired-end sequence reads were recorded for the bat fecal samples collected from Sarihidir Tunnel. After BLAST search, 760 hits (0.02% of total reads) were obtained. Among those hits, 76.22% could be assigned taxonomically, while no information was obtained for 23.73% of the hits. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 11. Overall, viral sequences obtained from Sarihidir Tunnel were grouped in 3 different virus families as follows: *Siphoviridae*, *Myoviridae* and *Zierdtviridae*. All of the viral sequences from Sarihidir Tunnel were detected as bacteriophage families (*Siphoviridae*, *Myoviridae* and *Zierdtviridae*). All of the viral sequences were found to belong to dsDNA viruses (Table 11).

Table 11. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Sarihidir Tunnel at Family Level via BLAST, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Sarihidir Tunnel	<i>M. myotis</i>	<i>Siphoviridae</i>	dsDNA	Bacteria
	<i>M. capaccinii</i>	<i>Myoviridae</i>	dsDNA	Bacteria
	<i>M. pallidus</i>	<i>Zierdtviridae</i>	dsDNA	Bacteria

4.3.2. Taxonomic Assignment of Sequence Reads via Kraken 2

Taxonomic assignment of sequence reads was performed using Kraken 2 (Wood *et al.*, 2019) program with default parameters for each sample collected from Cilingoz Cave, Sefer Yitigi Cave, and Sarihidir Tunnel. Then, each assigned organism was searched in the NCBI Taxonomy to obtain information about the viral family. Assigned organisms were also searched in the databases of the ICTV to find information on the host and nucleic acid type.

4,078,204 paired-end sequence reads were recorded for the bat fecal samples collected from Cilingoz Cave. After taxonomic assignment via Kraken 2 (Wood *et al.*, 2019), 23% could be assigned taxonomically, while no information was obtained for 67% of the sequence reads. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 12. Overall, viral sequences obtained from Cilingoz Cave were grouped in 7 different virus families as follows: *Podoviridae*, *Siphoviridae*, *Myoviridae*, *Microviridae*, *Iflaviridae*, *Dicistroviridae*, and *Parvoviridae*.

The majority of the viral sequences were assigned to bacteriophage families (*Podoviridae*, *Siphoviridae*, *Myoviridae* and *Microviridae*). Among the sequences, *Morganella phage* in the *Podoviridae* family that infects *Morganella sp* was the most common virus in bat fecal samples in Cilingoz Cave. Aside from bacteriophage families, invertebrate viruses (*Iflaviridae* and *Dicistroviridae*) were the second most abundant group. According to ICTV database, *Parvoviridae* family was the only one that may infect invertebrates and vertebrates. Sequences that were similar to viruses under *Parvoviridae* were also investigated separately and found similar to the viral sequences that infect invertebrates. Thus, vertebrate viruses were not identified in the sample collected from Cilingoz Cave. Majority of the viral sequences were found to belong dsDNA viruses, however, there was also virus families that belong to ssDNA and ssRNA viruses (Table 12).

Table 12. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Cilingoz Cave at Family Level via Kraken 2, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Cilingoz Cave	<i>M. capaccinii</i> <i>M. schreibersii</i> <i>R. euryale</i>	<i>Podoviridae</i>	dsDNA	Bacteria
		<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Microviridae</i>	ssDNA	Bacteria
		<i>Iflaviridae</i>	ssRNA	Invertebrates
		<i>Dicistroviridae</i>	ssRNA	Invertebrates
		<i>Parvoviridae</i>	ssDNA	Invertebrates-Vertebrates

3.863.402 paired-end sequence reads were recorded for the bat fecal samples collected from Sefer Yitigi Cave. After taxonomic assignment via Kraken 2 (Wood *et al.*, 2019), 14% could be assigned taxonomically, while no information was obtained for 86% of the sequence reads. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 13. Overall, viral sequences obtained from Sefer Yitigi Cave were grouped in 7 different virus families as follows: *Myoviridae*, *Siphoviridae*, *Herelleviridae*, *Baculoviridae*, *Malacoherpesviridae*, *Partitiviridae* and *Tombusviridae*.

The majority of the viral sequences were detected as bacteriophage families (*Myoviridae*, *Siphoviridae* and *Herelleviridae*). The second most abundant virus families were invertebrate viruses (*Baculoviridae* and *Malacoherpesviridae*). Only one virus family (*Tombusviridae*) that infects plants was recorded. In addition, *Partitiviridae* family which can infect fungi or plants was detected. Majority of the viral sequences were found to belong dsDNA viruses, however, there were also virus families that belong to dsRNA and ssRNA viruses (Table 13).

Table 13. Taxonomic Assignment of Viral Sequences in Bat Feces Collected from Sefer Yitigi Cave at Family Level via Kraken 2, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Sefer Yitigi Cave	<i>M. schreibersii</i> <i>M. myotis</i> <i>M. capaccinii</i> <i>R. ferrumequinum</i>	<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Herelleviridae</i>	dsDNA	Bacteria
		<i>Baculoviridae</i>	dsDNA	Invertebrates
		<i>Malacoherpesviridae</i>	dsDNA	Invertebrates
		<i>Partitiviridae</i>	dsRNA	Plants-Fungi
		<i>Tombusviridae</i>	(+)ssRNA	Plants

3.681.836 paired-end sequence reads were recorded for the bat fecal samples collected from Sarihidir Tunnel. After taxonomic assignment via Kraken 2 (Wood *et al.*, 2019), 8% could be assigned taxonomically, while no information was obtained for 92% of the sequence reads. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 14. Overall, viral sequences obtained from Sarihidir Tunnel were grouped in 8 different virus families as follows: *Podoviridae*, *Myoviridae*, *Siphoviridae*, *Herelleviridae*, *Lipothrixviridae*, *Baculoviridae*, *Parvoviridae* and *Adenoviridae*.

The majority of the viral sequences were detected as bacteriophage families (*Podoviridae*, *Myoviridae*, *Siphoviridae* and *Herelleviridae*). *Lipothrixviridae* was the only virus family in the sample collected from Sarihidir Tunnel, which infects archaea. Only one virus family, *Baculoviridae*, that infects invertebrates was detected. *Parvoviridae* family that can infect invertebrates and vertebrates was also detected. Sequence reads that were similar to the viruses under *Parvoviridae* were found to be similar to the viruses that infect invertebrates. Bat adenovirus from *Adenoviridae* was the only vertebrate virus that infects *Pipistrellus pipistrellus* (common pipistrelle bat) and *Pipistrellus nathusii* (Nathusius's pipistrelle bat). Majority of the viral sequences were found to belong dsDNA viruses, however, there was also one virus family that belong to ssDNA viruses (Table 14).

Table 14. Taxonomic Asssignment of Viral Sequences in Bat Feces Collected from Sarihidir Tunnel at Family Level via Kraken 2, NCBI Taxonomy ad ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Sarihidir Tunnel	<i>M. myotis</i> <i>M.capaccinii</i> <i>M.pallidus</i>	<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Podoviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Herelleviridae</i>	dsDNA	Bacteria
		<i>Lipothrixviridae</i>	dsDNA	Archaea
		<i>Baculoviridae</i>	dsDNA	Invertebrates
		<i>Parvoviridae</i>	ssDNA	Vertebrates-Invertebrates
		<i>Adenoviridae</i>	dsDNA	Vertebrates

4.3.3. Taxonomical assignment of the Contigs Obtained by de novo Based Assembly

Paired-end sequences were assembled by *de novo* approach to form contigs using the Microbial Genomics Module of CLC Genomics Workbench 22.0.1. Each contig was screened in selected databases on virus sequences in BLAST as explained in Section 3.5.2.15. According to the BLAST results, a data table was generated that included the first BLAST hit for each sequence, E-value, alignment score and the best aligned organism. Then, each assigned organism was searched in the NCBI Taxonomy to obtain information about the viral family. Assigned organisms were also searched in the databases of the ICTV to find information on the host and nucleic acid type.

838 contigs were constructed from the 4,078,204 paired-end sequence reads obtained from the sample collected from Cilingoz Cave. After BLAST search, 69% could be assigned taxonomically, while no information was obtained for 31% of the hits. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 15. Overall, viral sequences were grouped in 10 different virus families as follows: *Autographiviridae*, *Siphoviridae*, *Podoviridae*, *Myoviridae*, *Parvoviridae*, *Caliciviridae*, *Herpesviridae*, *Iflaviridae*, *Tymoviridae* and *Phycodnaviridae*. Unassigned sequences using the NCBI Taxonomy and ICTV were classified under the ‘Unclassified Viruses’.

The majority of the viral sequences were detected as bacteriophage families (*Autographiviridae*, *Siphoviridae*, *Podoviridae* and *Myoviridae*) along with unclassified groups (*Uncultured Caudovirales Phage* and *Unclassified Bacterial Viruses*). Among the bacteriophage sequences, *Morganella phage* in the *Autographiviridae* family that infects *Morganella sp* was the most common virus in bat fecal samples in Cilingoz Cave. *Parvoviridae* was the only virus family that can infect both invertebrates and vertebrates in the sample collected from Cilingoz Cave. Sequence reads were found to be similar to *Turdus hortulorum* Iiteradensovirus from *Parvoviridae*. Aside from these, virus families that infect vertebrates (*Caliciviridae* and *Herpesviridae*) were detected, but with low reliability based on high e-values. Sequences for *Stealth virus* from *Herpesviridae* and bat calicivirus from *Caliciviridae* were detected in Cilingoz Cave. In addition, virus families that infect invertebrates (*Iflaviridae*

family), plants (*Tymoviridae* family) and algae (*Phycodnaviridae* family) were recorded in the sample from Cilingoz Cave. Majority of the viral sequences were found to belong dsDNA viruses, however, there was also virus families that belong to ssDNA and ssRNA viruses (Table 15).

Table 15. Taxonomic Assignment of the Contigs Obtained from Viral Sequences in Bat Feces Collected from Cilingoz Cave by *de novo* Based Assembly at Family Level via BLAST, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Cilingoz Cave	<i>M. schreibersii</i> <i>M. capacinii</i> <i>R. euryale</i>	<i>Autographiviridae</i>	dsDNA	Bacteria
		<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Podoviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Parvoviridae</i>	ssDNA	Vertebrate-Invertebrate
		<i>Caliciviridae</i>	(+)ssRNA	Vertebrate
		<i>Herpesviridae</i>	dsDNA	Vertebrate
		<i>Iflaviridae</i>	(+)ssRNA	Invertebrate
		<i>Tymoviridae</i>	dsDNA	Plants
		<i>Phycodnaviridae</i>	dsDNA	Algae
		Uncultured Caudovirales Phage	dsDNA	Bacteria
		Unclassified Bacterial Viruses	dsDNA	Bacteria
		Unclassified Sequence	*	*

777 contigs were constructed from 3.863.402 paired-end sequence reads obtained from the sample collected from Sefer Yitigi Cave. After BLAST search, 33% could be assigned taxonomically, while no information was obtained for 67% of the hits. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 16. Overall, viral sequences were grouped in 13 different virus families as follows: *Siphoviridae*, *Podoviridae*, *Myoviridae*, *Autographiviridae*, *Inoviridae*, *Poxviridae*, *Iridoviridae*, *Parvoviridae*, *Chrysoviridae*, *Partitiviridae*, *Polymycoviridae*, *Hypoviridae* and *Totiviridae*. Unassigned sequences using NCBI Taxonomy and ICTV were classified under the ‘Unclassified Viruses’.

The majority of the viral sequences were detected as bacteriophage families (*Siphoviridae*, *Podoviridae*, *Myoviridae*, *Autographiviridae* and *Inoviridae*) along with

Uncultured Caudovirales Phage. Aside from bacteriophage families, viral families that may infect vertebrates and invertebrates (*Poxviridae*, *Iridoviridae* and *Parvoviridae*) were detected. Among those, sequence reads that hit BeAn 58058 virus from *Poxviridae* was the only one that infects vertebrates in Sefer Yitigi Cave. In addition to these, virus families that infect fungi (*Polymycoviridae*, *Hypoviridae* and *Totiviridae*) were detected. *Chrysoviridae* and *Partitiviridae* families that can infect plant and fungi were also recorded. Majority of the viral sequences were found to belong dsDNA viruses, however, there was also virus families that belong to ssDNA, dsRNA and ssRNA viruses (Table 16).

Table 16. Taxonomic Assignment of the Contigs Obtained from Viral Sequences in Bat Feces Collected from Sefer Yitigi Cave by *de novo* Based Assembly at Family Level via BLAST, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Sefer Yitigi Cave	<i>M. schreibersii</i> <i>M. myotis</i> <i>M. capaccinii</i> <i>R. ferrumequinum</i>	<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Podoviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Autographiviridae</i>	dsDNA	Bacteria
		<i>Inoviridae</i>	ssDNA	Bacteria
		<i>Poxviridae</i>	dsDNA	Vertebrates-Invertebrates
		<i>Iridoviridae</i>	dsDNA	Vertebrates-Invertebrates
		<i>Parvoviridae</i>	ssDNA	Vertebrates-Invertebrates
		<i>Chrysoviridae</i>	dsRNA	Fungi-Plant
		<i>Partitiviridae</i>	dsRNA	Fungi-Plant
		<i>Polymycoviridae</i>	dsRNA	Fungi
		<i>Hypoviridae</i>	(+)ssRNA	Fungi
		<i>Totiviridae</i>	dsRNA	Fungi
		Uncultured Caudovirales phage	dsDNA	Bacteria
		Unclassified RNA Viruses	RNA	Invertebrate
		Unclassified dsDNA Viruses	*	*
		Unclassified Sequences	*	*

1760 contigs were constructed from 3.681.836 paired-end sequence reads obtained from the sample collected from Sarihidir Tunnel. After BLAST search, 40% could be assigned taxonomically, while no information was obtained for 60% of the hits. The viral sequences were summarized based on their family, nucleic acid type, and host range as shown in Table 17. Overall, viral sequences were grouped in 8 different virus families as follows: *Siphoviridae*, *Myoviridae*, *Podoviridae*, *Autographiviridae*, *Chaseviridae*, *Inoviridae*, *Parvoviridae*, and

Totiviridae. Unassigned sequences using the NCBI Taxonomy and ICTV were classified under the ‘Unclassified Viruses’.

The majority of the viral sequences were detected as bacteriophage families (*Siphoviridae*, *Myoviridae*, *Podoviridae*, *Autographiviridae*, *Chaseviridae* and *Inoviridae*) along with the unclassified groups (unclassified Caudovirales, uncultured Caudovirales phage) in bat fecal samples in Sarihidir Tunnel. Aside from bacteriophage families, *Parvoviridae* that infects vertebrates and invertebrates were detected. Sequence reads that hit *Parvoviridae* family were found to be similar to *Grus japonensis ambidensovirus* from *Parvoviridae* that infects *Grus japonensis* (bird) were detected in Sarihidir Tunnel. In addition to these viral families, *Totiviridae* that infects fungi was detected. Majority of the viral sequences were found to belong dsDNA viruses, however, there was also virus families that belong to ssDNA and dsRNA viruses (Table 17).

Table 17. Taxonomic Assignment of the Contigs Obtained from Viral Sequences in Bat Feces Collected from Sarihidir Tunnel by *de novo* Based Assembly at Family Level via BLAST, NCBI Taxonomy and ICTV databases.

Location	Bat Species	Virus Families	Nucleic Acid Type	Host Range
Sarihidir Tunnel	<i>M. myotis</i> <i>M. capaccinii</i> <i>M. pallidus</i>	<i>Siphoviridae</i>	dsDNA	Bacteria
		<i>Myoviridae</i>	dsDNA	Bacteria
		<i>Podoviridae</i>	dsDNA	Bacteria
		<i>Autographiviridae</i>	dsDNA	Bacteria
		<i>Chaseviridae</i>	dsDNA	Bacteria
		<i>Inoviridae</i>	ssDNA	Bacteria
		<i>Parvoviridae</i>	ssDNA	Vertebrate-Invertebrate
		<i>Totiviridae</i>	dsRNA	Fungi
		Uncultured Caudovirales Phage	dsDNA	Bacteria
		Unclassified Caudovirales	dsDNA	Bacteria
		Unclassified Bacterial Viruses	dsDNA	Bacteria
		Unclassified DNA Viruses	DNA	*
		Unclassified Sequences	*	*

4.3.4. Comparison of Three Approaches that were Used for Taxonomic Assignment

To investigate the consistency between the results obtained via different bioinformatics analyses, paired-end sequence reads were taxonomically labelled using three different approaches: 1) taxonomic identification of sequence reads using BLAST on CLC Genomics Workbench 22.0.1. 2) taxonomic identification of sequence reads via Kraken 2 (Wood *et al.*, 2019) 3) taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on NCBI, BLAST. The detailed comparison of the results using these three approaches on the viral sequences obtained from the bat feces collected from Cilingoz Cave, Sefer Yitigi Cave, and Sarihidir Tunnel was given below.

In the bat feces from Cilingoz Cave, four virus families (*Myoviridae*, *Siphoviridae*, *Podoviridae* and *Parvoviridae*) could be identified by all three approaches (Figure 5). Three of those families were bacteriophages (*Myoviridae*, *Siphoviridae* and *Podoviridae*), while *Parvoviridae* was the virus family that infects both invertebrates and vertebrates. *Autographviridae* was the only bacteriophage family that was detected before and after *de novo* assembly in BLAST, but not via Kraken 2. *Iflaviridae* that infects invertebrates were detected by Kraken 2 approach and BLAST after *de novo* assembly. On the other hand, *Microviridae* (the bacteriophage family), and *Dicistroviridae* (invertebrate virus) were detected only by Kraken 2 approach. Two virus families (*Caliciviridae* and *Herpesviridae*) which infect vertebrates, one virus family (*Tymoviridae*) which infects plants and one virus family (*Phycodnaviridae*) which infects algae were only detected after *de novo* assembly of the contigs from sequence reads. Three virus families (*Polydnaviriformidae*, *Baculoviridae* and *Iridoviridae*) which infect invertebrates were detected by BLAST search of sequence reads (1st approach).

Interestingly, *Morganella phage* was the most common virus that was detected via three approaches. However, *Morganella phage* was assigned to different families in BLAST and Kraken 2 database. While *Morganella phage* was classified under the *Autographviridae* family in BLAST database, same phage was classified under the *Podoviridae* in Kraken 2 database.

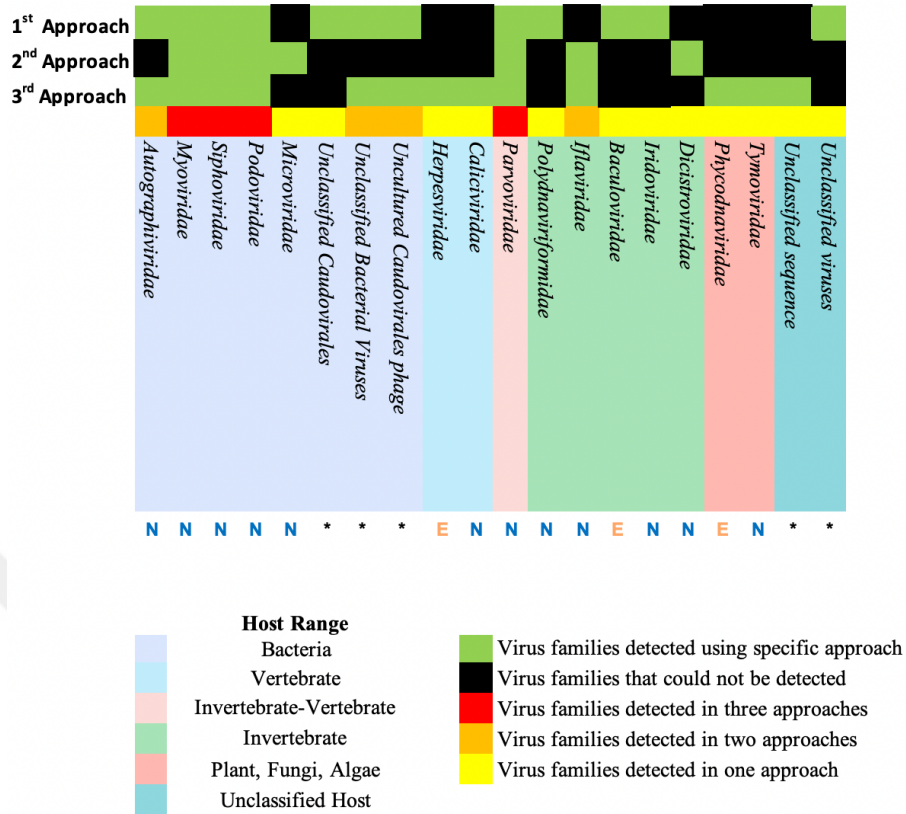


Figure 5. Comparison of Three Approaches that were Used for Taxonomic Assignment of Viral Sequences Obtained from Bat Feces Collected from Cilingoz Cave in Tekirdag in Northwest Anatolia. The viral sequences were analyzed using three different approaches. 1st approach: taxonomic identification of paired-end sequence reads using BLAST on CLC Genomics Workbench 22.0.1. 2nd approach: taxonomic identification of sequence reads via Kraken 2. 3rd approach: taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on NCBI, BLAST. The color codes in the heatmap are as follows: Green: virus families that were detected using a specific approach; Black: virus families that could not be detected using the same approach; Red: virus families that were detected via all three approaches; Orange: virus families that were detected via two approaches; Yellow: virus families that were detected via only one approach. N indicates the non-enveloped virus families and E indicates the enveloped virus families. Viral families were also color coded based on their host range as shown in the figure. * denotes the virus families with no envelope information.

In the bat feces from Sefer Yitigi Cave, two bacteriophage families (*Myoviridae* and *Siphoviridae*) could be identified by all three approaches (Figure 6). Two virus families (*Podoviridae* and *Inoviridae*) that infect bacteria and one virus family (*Polymycoviridae*) that infects fungi were detected by BLAST before and after *de novo* assembly. *Herelleviridae* that infects bacteria and *Baculoviridae* that infects invertebrates were detected by BLAST and Kraken 2 approaches. *Partitiviridae* that infects fungi and plants were identified by Kraken 2 approach and BLAST after *de novo* assembly.

One virus family (*Ascoviridae*) that infects invertebrates and one virus family (*Phycodnaviridae*) that infects algae were detected only by BLAST approach. *Malacoherpesviridae* that infects invertebrates and *Tombusviridae* that infects plants were only detected by Kraken 2 approach. Two virus families (*Ackermannviridae* and *Autographviridae*) that infect bacteria, one virus family that infects vertebrates (*Poxviridae*), one virus family (*Iridoviridae*) that infects invertebrates, one virus family (*Parvoviridae*) that infects invertebrates and vertebrates, two virus families (*Hypoviridae* and *Totiviridae*) that infect fungi and one virus family (*Chrysoviridae*) that infects fungi and plant were only detected by BLAST after *de novo* assembly.

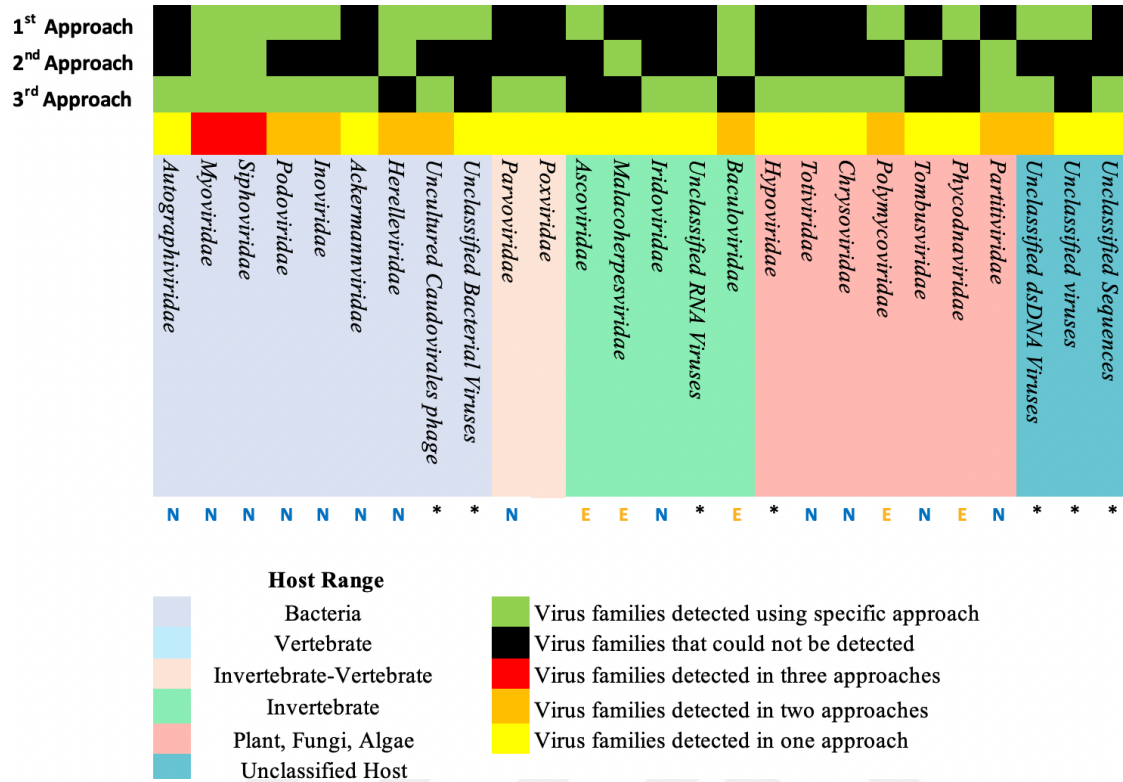


Figure 6. Comparison of Three Approaches that were Used for Taxonomic Assignment of Viral Sequences Obtained from Bat Feces Collected from Sefer Yitigi Cave in Burdur in Central Anatolia. The viral sequences were analyzed using three different approaches. 1st approach: taxonomic identification of paired-end sequence reads using BLAST on CLC Genomics Workbench 22.0.1. 2nd approach: taxonomic identification of sequence reads via Kraken 2. 3rd approach: taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on NCBI, BLAST. The color codes in the heatmap are as follows: Green: virus families that were detected using a specific approach; Black: virus families that could not be detected using the same approach; Red: virus families that were detected via all three approaches; Orange: virus families that were detected via two approaches; Yellow: virus families that were detected via only one approach. N indicates the non-enveloped virus families and E indicates the enveloped virus families. Viral families were also color coded based on their host range as shown in the figure. * denotes the virus families with no envelope information.

In the bat feces from Sarihidir Tunnel, two bacteriophage families (*Myoviridae* and *Siphoviridae*) could be identified by all three approaches (Figure 7). One virus family (*Podoviridae*) that infects bacteria and one virus family (*Parvoviridae*) that infect invertebrates and vertebrates were detected by Kraken 2 approach and BLAST after *de novo* assembly.

Three virus families (*Autographiviridae*, *Chaseviridae* and *Inoviridae*) that infect bacteria and one virus family (*Totiviridae*) that infects fungi were only detected by BLAST after *de novo* assembly. One virus family (*Herelleviridae*) that infects bacteria, one virus family (*Adenoviridae*) that infect vertebrates, one virus family (*Baculoviridae*) that infects invertebrates and one virus family (*Lipothrixviridae*) that infects archaea were detected only by Kraken 2 approach. *Zierdtviridae* that infects bacteria was detected by BLAST before contig assembly.

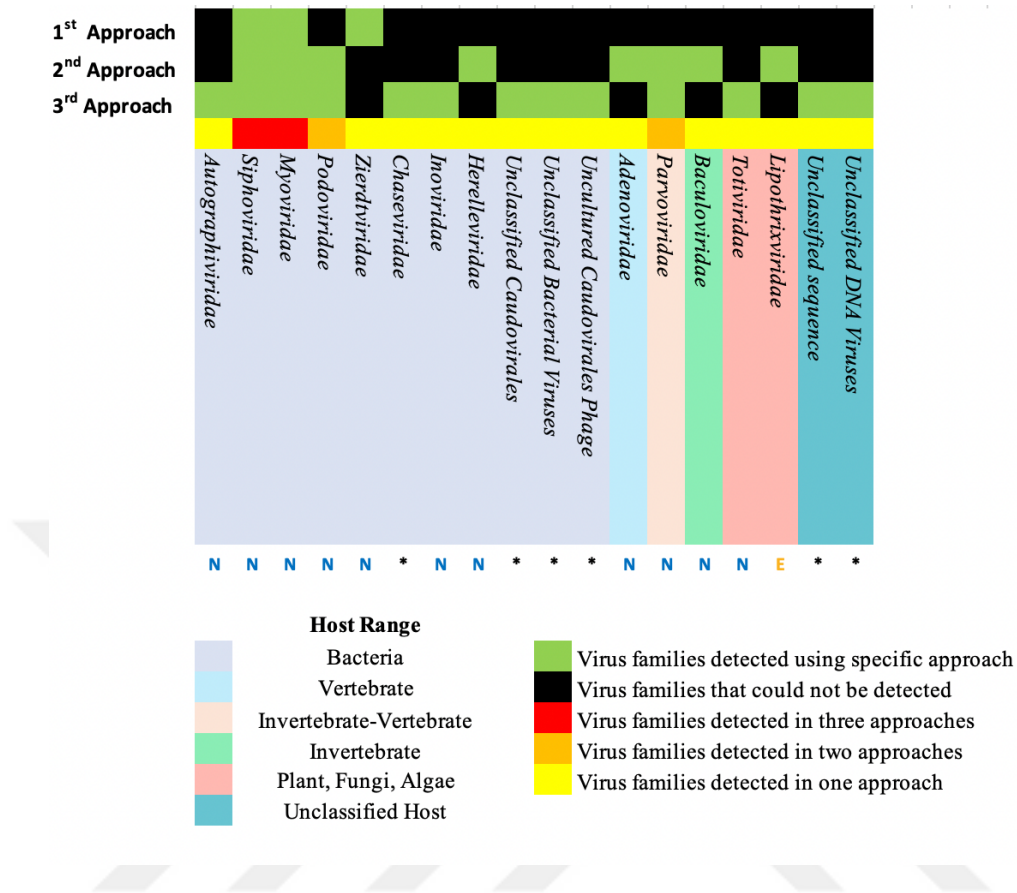


Figure 7. Comparison of Three Approaches that were Used for Taxonomic Assignment of Viral Sequences Obtained from Bat Feces Collected from Sarihidir Tunnel in Nevsehir in Central Anatolia. The viral sequences were analyzed using three different approaches. 1st approach: taxonomic identification of paired-end sequence reads using BLAST on CLC Genomics Workbench 22.0.1. 2nd approach: taxonomic identification of sequence reads via Kraken 2. 3rd approach: taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for virus sequences on NCBI, BLAST. The color codes in the heatmap are as follows: Green: virus families that were detected using a specific approach; Black: virus families that could not be detected using the same approach; Red: virus families that were detected via all three approaches; Orange: virus families that were detected via two approaches; Yellow: virus families that were detected via only one approach. N indicates the non-enveloped virus families and E indicates the enveloped virus families. Viral families were also color coded based on their host range as shown in the figure. * denotes the virus families with no envelope information.

5. DISCUSSION

This thesis study was performed for the investigation of virome profile in bat fecal samples in Northwest and Central Anatolia via metagenomics approach. To investigate this, the fecal samples of bats were collected from three different locations resembling Northwest Anatolia (Cilingoz Cave) and Central Anatolia (Sefer Yitigi Cave and Sarihidir Tunnel). To identify the viruses that could be found in bat feces, nucleic acids were converted to dsDNA followed by PCR enrichment with a primer targeting a region of RdRp of RNA viruses to avoid bias towards DNA viruses due to high bacterial content in gut microbiome. Then, genomic libraries were prepared and sequenced on the Illumina NextSeq500 platform. Finally, the paired-end sequence reads were classified using three different approaches to investigate the consistency between the results obtained via different bioinformatics analyses: 1) taxonomic identification of sequence reads using BLAST 2) taxonomic identification of sequence reads via Kraken 2 (Wood *et al.*, 2019) 3) taxonomic identification of the contigs that were obtained by *de novo* assembly using selected databases for viruses.

In literature there are many studies using different bioinformatics approaches in metagenomics analysis. Kraken 2 approach is most widely used for the taxonomic identification of short-length sequences compared to BLAST (Sabatier *et al.*, 2020; Zhang *et al.*, 2021; Vitney *et al.*, 2021; Mona *et al.*, 2021; Crook *et al.*, 2021). The accuracy of BLAST and Kraken 2 is similar (Wood and Lu, 2020), although the algorithms used by each method differ. The BLAST program is an algorithm that finds the best alignment for unknown sequences for the taxonomic assignment (Altschul *et al.*, 1990). On the other hand, Kraken 2 program uses k-mers to assign taxonomic labels without alignment (Wood and Lu, 2020). However, Kraken 2 provides some advantages over BLAST. For instance, use of k-mers helps to identify the lowest common ancestor in Kraken 2 and the database can be customized based on the data to be analyzed. However, BLAST tool on CLC Genomics Workbench 22.0.1 allows aligning the sequences against all available sequences in databases. Although selection of certain databases is an option on BLAST on NCBI, this option is not provided by CLC Genomics Workbench. On the other hand, an input file with millions of sequence reads is hard to handle on BLAST on NCBI.

Therefore, Kraken 2 is faster than BLAST which is an advantage for the analysis of a metagenomic data (Wood and Lu, 2020).

In this study, the paired-end sequence reads that were obtained from the bat fecal samples from three different locations were first analyzed via BLAST tool on CLC Genomics Workbench 22.0.1 or Kraken 2. Although the algorithms of each method are different, same virus families could be detected by each method in each sample. For instance, four virus families (*Myoviridae*, *Siphoviridae*, *Podoviridae* and *Parvoviridae*) in Cilingoz Cave, two virus families (*Myoviridae* and *Siphoviridae*) in Sefer Yitigi Cave, and two virus families (*Myoviridae* and *Siphoviridae*) in Sarihidir Tunnel were detected in common by these two approaches. On the other hand, there were some virus families in each sample that were detected either by BLAST or Kraken 2. In Cilingoz Cave, four virus families (*Autographiviridae*, *Baculoviridae*, *Polydnviriformidae* and *Iridoviridae*) could only be detected by BLAST, while three virus families (*Microviridae*, *Iflaviridae* and *Dicistroviridae*) could only be detected by Kraken 2 (Figure 5). In Sefer Yitigi Cave, five virus families (*Podoviridae*, *Inoviridae*, *Ascoviridae*, *Polymycoviridae* and *Phycodnaviridae*) could be detected by BLAST, while four virus families (*Poxviridae*, *Malacoherpesviridae*, *Partitiviridae* and *Tombusviridae*) could be detected by Kraken 2 (Figure 6). In Sarihidir Tunnel, one virus family (*Zierdtviridae*) could be detected by BLAST, while six virus families (*Podoviridae*, *Herelleviridae*, *Adenoviridae*, *Parvoviridae*, *Baculoviridae* and *Lipothrixviridae*) could be detected by Kraken 2 (Figure 7).

Constructing contigs from the paired-end sequence reads by *de novo* approach and searching in selected viral databases significantly increased the efficiency of taxonomic labelling in terms of the number of the sequences used and the number of virus families detected. After contig assembly, ten, thirteen and eight virus families were identified in Cilingoz Cave, Sefer Yitigi Cave and Sarihidir Tunnel, respectively (Figure 5-7). The highest difference in the number of virus families were observed in Sarihidir Tunnel after contig assembly as compared to the taxonomic identification of short-length sequences via BLAST. This could be due to low percentage of short length sequences in Sarihidir Tunnel that could be taxonomically assigned (0.02%) compared to that in other locations (0.11% for Cilingoz Cave and 0.04% for Sefer Yitigi Cave). For instance, bat calicivirus were detected only by BLAST

after *de novo* which might indicate that the longer sequences increase the percentage of identification. Overall, four, two and two virus families were common by three approaches in Cilingoz Cave, Sefer Yitigi Cave and Sarihidir Tunnel, respectively (Figure 5-7). For instance, *Morganella phage* was detected as the most abundant virus in bat fecal samples collected from Cilingoz Cave by three approaches. However, due to the differences between databases used by each program, *Morganella phage* was assigned to the *Autographviridae* family by BLAST, while it was assigned to the *Podoviridae* family by Kraken 2.

Apart from the virus families that were detected in common by three approaches, different virus families were identified depending on different bioinformatics approach used. Overall, in Cilingoz Cave, eight virus families were identified by BLAST, seven virus families were identified by Kraken 2, ten virus families were identified by BLAST after *de novo* assembly. In Sefer Yitigi Cave, ten virus families were identified by BLAST, seven virus families were identified by Kraken 2, thirteen virus families were identified by BLAST after *de novo* assembly. In Sarihidir Tunnel, three virus families were identified by BLAST, eight virus families were identified by Kraken 2, eight virus families were identified by BLAST after *de novo* assembly.

The widest host range was detected in Sefer Yitigi Cave by three approaches. The host range in Cilingoz Cave was restricted in bacteria, vertebrates and invertebrates by BLAST and Kraken 2. In addition to these hosts, plant and algae viruses were detected after *de novo* assembly. The smallest host range (only bacteria) was detected in Sarihidir Tunnel by BLAST. The host range expanded to bacteria, vertebrates, archaea and invertebrates by Kraken 2 and bacteria, vertebrates, fungi, and invertebrates by *de novo* assembly.

The differences in viral diversity in each location could also be resulted from several reasons. First, bat species inhabiting at each location could affect the content of the virome in feces. *M. capacinii*, *M. schreibersii* and *R. euryale*, which were observed during sample collection in Cilingoz Cave, are known to carry several virus families that infect vertebrates. In previous studies, two virus families (*Coronaviridae* and *Paramyxoviridae*) were detected in feces of *M. capacinii* (Drexler *et al.*, 2012; Ar *et al.*, 2018), while four virus families

(*Coronaviridae*, *Adenoviridae*, *Astroviridae* and *Parvoviridae*) were detected in feces of *M. schreibersii* (Chu *et al.*, 2008a; Li *et al.*, 2010; Ar *et al.*, 2018), and one virus family (*Coronaviridae*) was detected in feces of *R. euryale* (Drexler *et al.*, 2010). Interestingly, *Parvoviridae* was the only virus family that can infect vertebrates in Cilingoz Cave which might be shed through feces of *M. schreibersii* (Li *et al.*, 2010). In Sefer Yitigi Cave, *M. myotis*, *M. capacinii*, *M. schreibersii*, and *R. ferrumequinum* were recorded. In previous studies, virus families that infect vertebrates were identified in these bat species, as well. For instance, one virus family (*Coronaviridae*) were detected in feces of *R. ferrumequinum* (Ar *et al.*, 2018), four virus families (*Coronaviridae*, *Paramyxoviridae*, *Adenoviridae* and *Astroviridae*) were detected in feces of *M. myotis* (Drexler *et al.*, 2011; Ar *et al.*, 2018), while two families (*Coronaviridae* and *Paramyxoviridae*) were detected in fecal sample of *M. capacinii* (Drexler *et al.*, 2012; Ar *et al.*, 2018), and four virus families (*Coronaviridae*, *Adenoviridae*, *Astroviridae* and *Parvoviridae*) were detected in *M. schreibersii* (Chu *et al.*, 2008a; Drexler *et al.*, 2011; Ar *et al.*, 2018). None of those viral families were detected in Sefer Yitigi Cave. In Sarihidir Tunnel, *M. myotis*, *M. capacinii*, and *M. pallidus* were recorded. In previous studies, four virus families (*Coronaviridae*, *Paramyxoviridae*, *Adenoviridae* and *Astroviridae*) that infect vertebrates were detected in feces of *M. myotis* (Drexler *et al.*, 2011; Ar *et al.*, 2018), two virus families (*Coronaviridae* and *Paramyxoviridae*) were detected in feces of *M. capacinii* (Drexler *et al.*, 2012; Ar *et al.*, 2018) and, two virus families (*Coronaviridae* and *Astroviridae*) were detected in fecal and rectal swab samples of *M. pallidus* in China (Chu *et al.*, 2008a; Chu *et al.*, 2008b). Among all sequences, bat adenovirus from *Adenoviridae* family was the only one being detected in Sarihidir Tunnel which might be shed through feces of *M. myotis* (Drexler *et al.*, 2011).

Secondly, social habits of bat species inhabiting at each location could affect the content of the virome in feces. Bats can interact and share caves with other animals which might be potential for the interspecies transmission of viruses (Thomas *et al.*, 2012). In this study, virus families which can infect and be carried by other animals were also detected. For example, *Turdus hortulorum* Iteradensovirus, which was previously identified in the fecal sample of grey-backed thrush, was detected in Cilingoz Cave (NCBI, 2022). According to Trakus, grey-backed thrush have not seen in Türkiye before. This might indicate that *Turdus hortulorum*

Iteradensovirus may be carried by other bird species in Türkiye or might be introduced to birds from bats. Interestingly, bat calicivirus from *Caliciviridae*, which was previously detected in the fecal sample of *Myotis daubentonii*, was detected in Cilingoz Cave (Hardmeier *et al.*, 2021). In Sarihidir Tunnel, bat adenovirus that infects *Pipistrellus pipistrellus* and *Pipistrellus nathusiix* was detected. *Pipistrellus pipistrellus* and *Pipistrellus nathusiix* were frugivorous bat species observed in Türkiye (EUROBATS, 2022g; EUROBATS, 2022h). These two viruses show that bats in Cilingoz Cave and Sarihidir Tunnel might interact with the other bat species which might result in the introduction of new viral families in the population, especially by *M. schreibersii* (flying range:40-100 km) or *M. capacinii*, as they are seasonal migrants (EUROBATS, 2022a; EUROBATS, 2022c). On the other hand, *Myotis daubentonii*, *Pipistrellus pipistrellus* and/or *Pipistrellus nathusiix* might live in Cilingoz Cave or Sarihidir Tunnel and leave or migrate before sample collection date.

Another factor that could affect the content of the virome in feces is dietary habits of bat species inhabiting at each location. All bats observed in the sampling locations are insectivorous. Therefore, invertebrate virus families, can be detected in fecal samples of bats (Li *et al.*, 2010; Bolatti *et al.*, 2020). As expected, this group of viruses were detected as the second abundant group in this study. For instance, *Dicistroviridae* was detected in the fecal samples collected from Cilingoz Cave. This virus family was also detected in another study analyzing the feces of *R. ferrumequinum* and *M. schreibersii* (Wu *et al.*, 2016). In addition, *Baculoviridae* that infects invertebrates was detected in Sefer Yitigi Cave and Sarihidir Tunnel. Invertebrates are known to be a dietary choice of *R. ferrumequinum* and *M. myotis* (Andreas *et al.*, 2012). In this study, aside from invertebrate viruses, plant viruses were also detected in Sefer Yitigi Cave and Sarihidir Tunnel. These virus families might reflect the dietary choice of insects that are consumed by bats (Bolatti *et al.*, 2020).

Moreover, abiotic features of roosts could affect the content of the virome in feces. Temperature, humidity, and soil structure are the abiotic factors that can vary from cave to cave (Oliveria *et al.*, 2017). These abiotic factors can affect the stability of the microorganisms in the feces. For instance, rabies is sensitive to ultra-violet light which lead to inactivation (CDC,

2022h). In addition, light sensitive microorganisms were detected in the cave microhabitat (Oliveria *et al.*, 2017). The effect of abiotic factors was not assessed in this study.

Finally, microbiome of bat species inhabiting at each location could affect the content of the virome in feces. Because bat fecal material resembles the gut microbiome, which is mostly dominated by bacteria, high number of bacteriophage families were expected in each sample (Li *et al.*, 2010; Šimić *et al.*, 2020). However, RNA viruses that could be shed through bat feces are quite important, as 89% of human infections with RNA viruses were transmitted from animals to humans (Woolhouse *et al.*, 2013). Thus, to avoid a bias towards DNA viruses in the samples, K8N primer targeting a region of RdRp, which has been used in metagenomics studies, was used to enrich the cDNAs in terms of RNA viruses, as well (Wu *et al.*, 2012; Wu *et al.*, 2016; Zheng *et al.*, 2018). Using this enrichment method, RNA virus families such as *Polymycoviridae*, *Iflaviridae*, *Dicistroviridae*, *Partitiviridae*, *Caliciviridae*, *Chrysoviridae*, *Hypoviridae*, *Totiviridae* and *Tombusviridae* could be detected in the samples. However, bacteriophage families were abundant in all three samples (Figure 5, Figure 6, and Figure 7). This might be resulted from several reasons. First reason of this abundance might be related with attachment of K8N primer on bacteriophage (*Siphoviridae*, *Myoviridae* and *Podoviridae*) genome, especially on the genes that mostly express hypothetical proteins. This might increase the copy number of genome in enrichment step. However, the abundancy of bacteriophages was reported in other studies, as well (Li *et al.*, 2010; Šimić *et al.*, 2020). Therefore, this might not be the only reason. Another and main reason of the bacteriophage abundance in feces is that bacteriophages are representation of the microbiome of bat species. For instance, *Morganella phage* infecting *Morganella morganii*, which was detected as the most abundant virus in Cilingoz Cave, was previously detected in fecal samples of *M. capacinii* (Bella *et al.*, 2003). Other reason might be the stability of bacteriophages which mostly do not have envelope (non-enveloped). Their stabilization might be higher than the other viruses. Because enveloped viruses are more sensitive to environmental conditions than non-enveloped viruses (Firquet *et al.*, 2015). As our data showed that the number of virus families without envelope is higher than the virus families with envelope (12 non-enveloped and 3 enveloped virus families in Cilingoz Cave, 13 non-enveloped and 6 enveloped virus families in Sefer Yitigi Cave, 11 non-enveloped and 1 enveloped virus families in Sarihidir Tunnel) (Figure 5).

Enveloped viruses might be disrupted during homogenization and nuclease treatment steps which are suggested steps for the removal of free nucleic acids in samples in metagenomics. Such disruptions might result in release of viral genome which probably eliminates the genomic material of enveloped viruses during nuclease treatment. Therefore, this might be the reason for the low number of virus families with envelope. Another reason might be adsorption of viruses on fecal material. After homogenization, samples from each location were centrifuged and passed through filters to remove the debris and bacterial or eukaryotic organisms. On the other hand, this step might cause elimination of not only the giant viruses but also other viruses. Because viruses can be adsorbed by solid materials based on their isoelectric point and size (Dowd *et al.*, 1997; Davis *et al.*, 2006). It was shown that elemental composition of sediments such as Aluminum (Al), Calcium (Ca), Copper (Cu), Iron (Fe), Potassium (K), Molybdenum (Mo), Oxygen (O), Silicon (Si), Titanium (Ti) and Zinc (Zn) might increase the adsorption capacity of viruses in environmental sources (Azbazdar *et al.*, 2022). It was also shown that feces of *Megaderma spasma* (lesser false vampire bat) and *Taphozous melanopogon* (insectivorous, black-bearded tomb bat) contain similar elements (Johnson *et al.*, 2020). Therefore, viruses might be adsorbed on bat feces and eliminated during centrifuge and filtration steps.

More than 60% of the paired-end sequence reads could not be identified by three of the bioinformatic approaches in this study. This situation was reported in many studies in literature (Leeuwen *et al.*, 2010; Dacheux *et al.*, 2014; Wu *et al.*, 2016; Šimić *et al.*, 2020). For instance, in a study conducted in Croatia, only 2.52% (39,527 viral sequence /1,565,543 total sequence) of the sequences were identified as viral sequences, 14.2% of which was recorded as unclassified (Šimić *et al.*, 2020). The problem in the classification of viruses is mostly resulted from lack of conserved region in viral genomes. Therefore, surveillance and metagenomics studies are significant tools to discover novel microorganisms and identify unclassified organisms. The number of viral sequences in databases will increase with more surveillance and metagenomics studies. In this study, there are many sequences that could not be identified as viral sequences or assigned taxonomically. Some of these sequences might belong to novel viruses. Improvement in viral databases will help to reanalyze these unassigned sequences in the future.

Screening and identification of viruses in their natural reservoirs are crucial for the detection of the viruses with zoonotic potential. In current literature there is not enough information about bats and bat-related viruses in Türkiye. Therefore, bat feces were screened via metagenomics analysis to determine the viruses shed in the feces of bat populations residing in Türkiye. This thesis is the first study that identifies the virome profile in the fecal samples of bats that inhabit in Northwest and Central Anatolian regions of Türkiye by metagenomics approach.



6. CONCLUSION AND FUTURE DIRECTIONS

Even though bats harbor various virus families that infect vertebrates, invertebrates, plant, fungi, and bacteria at the same time, they show no or few symptoms. Viruses can be transmitted from bats to humans either via direct contact with body fluids of bats (saliva, feces, blood), vector borne, intermediate hosts or indirect contact with contaminated environmental sources. There is no sufficient information about bat-related viruses in Türkiye. For that reason, virome content in feces of bats inhabiting in Northwest and Central Anatolian regions of Türkiye were investigated. In Cilingoz Cave, virus families that infect bacteria, vertebrates, invertebrates, plant, and algae were detected. In Sefer Yitigi Cave, virus families that infect, bacteria, vertebrates, invertebrates, plant, and fungi were detected. In Sarihidir Tunnel, virus families that infect bacteria, vertebrates, invertebrates, archaea, and fungi, were detected. Overall, bacteriophages were the most abundant and invertebrate viruses were the second most abundant virus families that were identified in the feces of bats in all three locations by each bioinformatics approach used. None of the vertebrate viruses identified in this study carry zoonotic risk.

The number of detected virus families in bat feces varied based on the bioinformatics approaches. While the highest viral diversity was detected in Sefer Yitigi Cave by BLAST (1st and 3rd) approach, the highest viral diversity was detected in Sarihidir Tunnel by Kraken 2 (2nd approach). Among all approaches for all regions, *de novo* assembly approach stands out with the advantage in viral diversity and host range. On the other hand, more than 60% of the sequence reads could not be identified by any of the bioinformatic approaches. This is mainly resulted from lack of virus sequences in databases. Therefore, omics studies such as metagenomics should be encouraged as they may significantly contribute to the discovery of novel microorganisms without culturing methods.

7. REFERENCES

Ahn M, Cui J, Irving A. Unique Loss of the PYHIN Gene Family in Bats Amongst Mammals: Implications for Inflammasome Sensing. *Sci Rep*. 2016; 6, 2172 <https://doi.org/10.1038/srep21722>

Albayrak I. Ecosystem of Cave. *Biyoloji Bilimleri Araştırma Dergisi*, 2012;5: 61-64.

Altıntaş A, Konaş T, Yıldız G, Erkal N. Yarasa dışkısı (bat guano) mineral düzeyleri, Ankara Üniversitesi Veterinerlik Fakültesi Dergisi, 2005; 52, 1-5.

Andreas M, Reiter A, Benda P. Prey Selection and Seasonal Diet Changes in the Western Barbastelle Bat (*Barbastella barbastellus*). *Acta Chiropterologica*. 2012;14. 81-92. 10.3161/150811012X654295.

APS (Advancing the Science of Plants Pathology) Introduction to Plant Viruses, the Invisible Foe. *The Plant Health Instructor*, 2006; DOI: 10.1094/PHI-I-2006-0414-01

Ar Gouilh M, Puechmaille SJ, Diancourt L, Vandebogaert M, Serra-Cobo J, Lopez Roig M, Brown P, Moutou F, Caro V, Vabret A, Manuguerra JC; EPICOREM consortium. SARS-CoV related Betacoronavirus and diverse Alphacoronavirus members found in western old-world. *Virology*. 2018 Apr;517:88-97. doi: 10.1016/j.virol.2018.01.014. Epub 2018 Feb 23. PMID: 29482919; PMCID: PMC7112086.

Altschul S, Gish W, Miller W, Myers E, Lipman D: Basic local alignment search tool. *J Mol Biol*. 1990, 215: 403-410.

Azbazdar ME, Akıncı G, Güler E, Koçer ZA. The effect of sediment composition and polyethylene glycol precipitation on the detection limit of H6N2 influenza virus in sediment samples. *Marine and Life Sciences*. 2022;4 (1), 53-62. DOI: 10.51756/marlife.1091169

- Baker ML, Schountz T, Wang LF. Antiviral immune responses of bats: a review. *Zoonoses Public Health*. 2013; 60:104–16. <http://dx.doi.org/10.1111/j.1863-2378.2012.01528.x>
- Banerjee A, Baker ML, Kulcsar K, Misra V, Plowright R, Mossman K. Novel Insights Into Immune Systems of Bats. *Front. Immunol*. 2020;11:26. doi: 10.3389/fimmu.2020.00026
- Banerjee A, Rapin N, Bollinger T, Misra V. Lack of inflammatory gene expression in bats: a unique role for a transcription repressor. *Sci Rep*. 2017;7:2232. doi: 10.1038/s41598-017-01513-w
- Banerjee S, Gupta N, Kodan P, Mittal A, Ray Y, Nischal N, Soneja M, Biswas A, Wig N. Nipah virus disease: A rare and intractable disease. *Intractable Rare Dis Res*. 2019
- Bennett AJ, Bushmaker T, Cameron K, Ondzie A, Niama FR, Parra HJ, Mombouli JV, Olson SH, Munster VJ, Goldberg TL. Diverse RNA viruses of arthropod origin in the blood of fruit bats suggest a link between bat and arthropod viromes. *Virology*. 2019
- Bella C, Piraino C, Caracappa S, Fornasari L, Violani C, Zava B. Enteric microflora in Italian Chiroptera. *J. Mt. Ecol*. 2003; 7. 221-224.
- Biek R, Walsh PD, Leroy EM, Real LA. Recent common ancestry of Ebola Zaire virus found in a bat reservoir. *PLoS Pathog*. 2006;2:e90.
- Bolatti EM, Zorec TM, Montani ME, Hosnjak L, Chouhy D, Viarengo G, Casal PE, Barquez RM, Poljak M, Giri AA: A preliminary study of the virome of the South American free- tailed bats (*Tadarida brasiliensis*) and identification of two novel mammalian viruses. *Viruses*.2020; 12:422.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*,2014;btu170.
- Breed AC. Paramyxoviruses in Bats. *Zoo and Wild Animal Medicine*. 2008:225–cp3. doi: 10.1016/B978-141604047-7.50031-2. Epub 2009 May 19. PMID: PMC7152176.

Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall, AM, Mead D, Azam F, Rohwer F. Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences of the United States of America*, 2002; 99(22)

Buchler ER. Food transit time in *Myotis lucifugus* (Chi- Chiroptera: Vespertilionidae). *Journal of Mammalogy*, 1975; vol. 54, pp. 985–990, 1975.

Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. Bats: important reservoir hosts of emerging viruses. *Clin Microbiol Rev.* 2006 ;Jul;19(3):531-45. doi: 10.1128/CMR.00017-06. PMID: 16847084; PMCID: PMC1539106.

CDC (Centers for Disease Control and Prevention) How is rabies transmitted? <https://www.cdc.gov/rabies/transmission/index.html>. Accessed on :22.07.2022

CDC (Centers for Disease Control and Prevention) What are the signs and symptoms of rabies?<https://www.cdc.gov/rabies/symptoms/index.html>. Accessed on :22.07.2022

CDC (Centers for Disease Control and Prevention). Ebola (Ebola Virus Disease). <https://www.cdc.gov/vhf/ebola/about.html>. Accessed on: 03.08.2022

CDC (Centers for Disease Control and Prevention). Marburg(Marburg Virus Disease).<https://www.cdc.gov/vhf/marburg/index.html>. Accessed on:03.08.2022

CDC (Centers for Disease Control and Prevention). Transmission. <https://www.cdc.gov/vhf/hendra/transmission/index.html>. Accessed on:22.07.2022

CDC (Centers for Disease Control and Prevention) Zoonotic Diseases.[https://www.cdc.gov/onehealth/basics/zoonotic-diseases.html#:~:text=Zoonotic%20diseases%20\(also%20known%20as,spread%20between%20animals%20and%20people.&text=Animals%20provide%20many%20benefits%20to,home%20and%20away%20from%20home](https://www.cdc.gov/onehealth/basics/zoonotic-diseases.html#:~:text=Zoonotic%20diseases%20(also%20known%20as,spread%20between%20animals%20and%20people.&text=Animals%20provide%20many%20benefits%20to,home%20and%20away%20from%20home). Accessed on: 22.07.2022

CDC (Centers for Disease Control and Prevention) Rabies Exposure in Healthcare Settings. https://www.cdc.gov/rabies/specific_groups/hcp/exposure.html. Accessed on: 16.08.2022

CLC Genomics Workbench 22.0.1, CLC Microbial Genomics Module 20.1 (<https://digitalinsights.qiagen.com/>)

Cooke MS, Evans MD, Dizdaroglu M, Lunec J. Oxidative DNA damage: mechanisms, mutation, and disease. *The FASEB Journal*. 2003; 17: 1195-1214. <https://doi.org/10.1096/fj.02-0752rev>

Dacheux L, Cervantes-Gonzalez M, Guigon G, Thiberge JM, Vandebogaert M, Maufrais C, Caro V, Bourhy H. A preliminary study of viral metagenomics of French bat species in contact with humans: identification of new mammalian viruses. *PLoS One*. 2014 Jan 29;9(1):e87194. doi: 10.1371/journal.pone.0087194. PMID: 24489870; PMCID: PMC3906132.

Davis JA, Farrah SR, Wilkie AC. Adsorption of viruses to soil: impact of anaerobic treatment. *Water Sci Technol*. 2006;54(3):161-7. doi: 10.2166/wst.2006.464. PMID: 17037148.

DBatVir(the Database of Bat-associated Viruses).<http://www.mgc.ac.cn/DBatVir/> Accessed on: 17.08.2022

De Oliveira C, Gunderman L, Coles C, Lochmann J, Parks M, Ballard E, Glazko G, Rahmatallah Y, Tackett A, Thomas, D. 16S rRNA Gene-Based Metagenomic Analysis of Ozark Cave Bacteria. *Diversity*. 2017; 9. 31. 10.3390/d9030031.

Dowd SE, Pillai SD, Wang S, Corapcioglu MY. Delineating the specific influence of virus isoelectric point and size on virus adsorption and transport through sandy soils. *Appl Environ Microbiol*. 1998 ;Feb;64(2):405-10. doi: 10.1128/AEM.64.2.405-410.1998. PMID: 9464373; PMCID: PMC106058.

Drexler JF, Corman VM, Wegner T, Tateno AF, Zerbinati RM, Gloza-Rausch F, Seebens A, Müller MA, Drosten C. Amplification of emerging viruses in a bat colony. *Emerg Infect Dis*.

2011 Mar;17(3):449-56. doi: 10.3201/eid1703.100526. PMID: 21392436; PMCID: PMC3165994.

Drexler JF, Corman VM, Müller MA, Maganga GD, Vallo P, Binger T, Gloza-Rausch F, Cottontail VM, Rasche A, Yordanov S, Seebens A, Knörnschild M, Oppong S, Adu Sarkodie Y, Pongombo C, Lukashev AN, Schmidt-Chanasit J, Stöcker A, Carneiro AJ, Erbar S, Maisner A, Fronhoffs F, Buettner R, Kalko EK, Kruppa T, Franke CR, Kallies R, Yandoko ER, Herrler G, Reusken C, Hassanin A, Krüger DH, Matthee S, Ulrich RG, Leroy EM, Drosten C. Bats host major mammalian paramyxoviruses. *Nat Commun.* 2012; Apr 24;3:796. doi: 10.1038/ncomms1796. Erratum in: *Nat Commun.* 2014;5:3032. Cottontail, Veronika M [added]; Knörnschild, Mirjam [added]. PMID: 22531181; PMCID: PMC3343228.

Drexler JF, Gloza-Rausch F, Glende J, Corman VM, Muth D, Goettsche M, Seebens A, Niedrig M, Pfefferle S, Yordanov S, Zhelyazkov L, Hermanns U, Vallo P, Lukashev A, Müller MA, Deng H, Herrler G, Drosten C. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *J Virol.* 2010 Nov;84(21):11336-49. doi: 10.1128/JVI.00650-10. Epub 2010 Aug 4. PMID: 20686038; PMCID: PMC2953168.

Drosten C, Günther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA, Berger A, Burguière AM, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra JC, Müller S, Rickerts V, Stürmer M, Vieth S, Klenk HD, Osterhaus AD, Schmitz H, Doerr HW. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med.* 2003 May 15;348(20):1967-76. doi: 10.1056/NEJMoa030747. Epub 2003 Apr 10. PMID: 12690091.

Edenborough KM, Mu A, Mühldorfer K, Lechner J, Lander A, Bokelmann M, Couacy-Hymann E, Radonic A, Kurth A. Microbiomes in the insectivorous bat species *Mops condylurus* rapidly converge in captivity. *PLoS One.* 2020 Mar 20;15(3):e0223629. doi: 10.1371/journal.pone.0223629. PMID: 32196505; PMCID: PMC7083271.

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Miniopterus pallidus*. https://www.eurobats.org/about_eurobats/protected_bat_species/miniopterus_pallidus. Accessed on: 12.06.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Miniopterus schreibersii*. https://www.eurobats.org/about_eurobats/protected_bat_species/miniopterus_schreibersii. Accessed on: 12.06.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Myotis capaccinii*. https://www.eurobats.org/about_eurobats/protected_bat_species/myotis_capaccinii. Accessed: 12.06.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Myotis myotis*. https://www.eurobats.org/about_eurobats/protected_bat_species/myotis_myotis. Accessed: 12.06.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Rhinolophus euryale*. https://www.eurobats.org/about_eurobats/protected_bat_species/rhinolophus%20euryale. Accessed on: 12.06.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Rhinolophus ferrumequinum*. https://www.eurobats.org/about_eurobats/protected_bat_species/rhinolophus_ferrumequinum. Accessed on: 12.06.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Pipistrellus nathusii*. https://www.eurobats.org/about_eurobats/protected_bat_species/pipistrellus_nathusii. Accessed on: 16.08.2022

EUROBATS (Agreement on the Conservation of Populations of European Bats). *Pipistrellus pipistrellus*. https://www.eurobats.org/about_eurobats/protected_bat_species/pipistrellus_pipistrellus. Accessed on: 16.08.2022

Fenton MB. Constraint and flexibility—bats as predators, bats as prey. Symposia of the Zoological Society of London, 1995; vol. 67, pp. 277–289

Ferreira RL, Martins RP. Trophic structure and natural history of bat guano invertebrate communities, with special reference to Brazilian caves, 1999; Trop. Zool. 12 (2), 231–252. <https://doi.org/10.1080/03946975.1999.10539391>.

Ferreira RL. Guano communities. 2019; Encyclopedia of Caves

Field H, Schaaf K, Kung N, Simon C, Waltisbuhl D, Hobert H, Moore F, Middleton D, Crook A, Smith G, Daniels P, Glanville R, Lovell D. Hendra virus outbreak with novel clinical features, Australia. Emerging infectious diseases, 2010; 16(2), 338–340. <https://doi.org/10.3201/eid1602.090780>

Firquet S, Beaujard S, Lobert PE, Sané F, Caloone D, Izard D, Hober D. Survival of Enveloped and Non-Enveloped Viruses on Inanimate Surfaces. Microbes Environ. 2015;30(2):140-4. doi: 10.1264/jsme2.ME14145. Epub 2015 Apr 3. PMID: 25843687; PMCID: PMC4462923.

Fukui D, Okazaki K, Maeda K. Diet of three sympatric insectivorous bat species on Ishigaki Island, Japan. Endangered Species Research, 2009; vol. 8, no. 1-2, pp. 117–128

GBIF (Global Biodiversity Information Facility). *Miniopterus pallidus* Thomas, 1907. <https://www.gbif.org/species/7500714>. Accessed on:22.07.2022

Geda MK, Balakrishnan, M. Ecological and Economic Importance of Bats. ISRN Biodiversity. 2013. 1-9. 10.1155/2013/187415.

Gérald L. Bats and viruses or how to live together in harmony. <https://www.encyclopedie-environnement.org/en/life/bats-viruses-how-live-together-harmony/>. Accessed on:01.08.2022

Gergerich RC, Dolja VV. Introduction to Plant Viruses, the Invisible Foe. The Plant Health Instructor. 2006; DOI: 10.1094/PHI-I-2006-0414-01

Ghabrial SA, Castón JR, Jiang D, Nibert ML, Suzuki N. 50-plus years of fungal viruses. *Virology*. 2015 May;479-480:356-68. doi: 10.1016/j.virol.2015.02.034. Epub 2015 Mar 13. PMID: 25771805.

Ghabrial SA, Suzuki N. Fungal viruses, In B. W. J. Mahy and M. H. V. Van, 2009

Graham RL, Donaldson EF, Baric RS. A decade after SARS: strategies for controlling emerging coronaviruses. *Nat Rev Microbiol* 11, 836–848, 10.1038/nrmicro3143 (2013).

Hardmeier I, Aeberhard N, Qi W, Schoenbaechler K, Kraettli H, Hatt JM, Fraefel C, Kubacki J: Metagenomic analysis of fecal and tissue samples from 18 endemic bat species in Switzerland revealed a diverse virus composition including potentially zoonotic viruses. *PLoS One* 2021, 16:e0252534.

Harcourt BH, Lowe L, Tamin A, Liu X, Bankamp B, Bowden N, Rollin PE, Comer JA, Ksiazek TG, Hossain MJ, Gurley ES, Breiman RF, Bellini WJ, Rota PA. Genetic characterization of Nipah virus, Bangladesh, 2004. *Emerg Infect Dis*. 2005 Oct;11(10):1594-7. doi: 10.3201/eid1110.050513. PMID: 16318702; PMCID: PMC3366751.

Hu B, Zeng LP, Yang XL, Ge XY, Zhang W, Li B, Xie JZ, Shen XR, Zhang YZ, Wang N, Luo DS, Zheng XS, Wang MN, Daszak P, Wang LF, Cui J, Shi ZL. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog*. 2017 Nov 30;13(11):e1006698. doi: 10.1371/journal.ppat.1006698. PMID: 29190287; PMCID: PMC5708621.

ICDDR. Person-to-person transmission of Nipah virus during outbreak in Faridpur District. *Health Science Bulletin*. 2004; 2:5–9.

ITIS (Integrated Taxonomic Information System). *Miniopterus pallidus*. https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=947260#null) Accessed on:22.07.2022

Johara M, Field H, Rashdi A, Morrissy C, Heide B, Rota P. Serological evidence of infection with Nipah virus in bats (order Chiroptera) in Peninsular Malaysia. *Emerg Infect Dis.* 2001; 7. 439-441.

Johnson J, Vincent M. Tracing heavy metals in urban ecosystems through the study of bat guano - a preliminary study from Kerala, India. *Journal of Threatened Taxa.* 2020;12(10): 16377–16379. <https://doi.org/10.11609/jot.6225.12.10.16377-16379>

Jones G, Jacobs D, Kunz TH, Wilig MR, Racey PA. *Carpe Noctem: the importance of bats as bioindicators.* *Endangered Species Research*, 2009; vol. 8, pp. 3–115,

KA Harouna, AD Tidjani, AG Fanna, G Yadji. Bats Guano from Magarawa in Niger: Exploitation, Marketing and Peasant Perception. *International Journal of Innovation and Applied Studies.* 2021

Kawai T, Akira S. Innate immune recognition of viral infection. *Nat Immunol.* 2006;7:131–7. doi: 10.1038/ni1303

Kchouk M, Gibrat JF, Elloumi M. Generations of Sequencing Technologies: From First to Next Generation. *Biology and Medicine.* 2017; 09. 10.4172/0974-8369.1000395.

Kelley J, de Bono B, Trowsdale J. IRIS: a database surveying known human immune system genes. *Genomics.* 2005; 85:503–11. doi: 10.1016/j.ygeno.2005.01.009

Kemenesi G, Tóth GE, Mayora-Neto M, Scott S, Temperton N, Wright E, Mühlberger E, Hume AJ, Suder EL, Zana B, Boldogh SA, Görföl T, Estók P, Lanszki Z, Somogyi BA, Nagy Á, Pereszlényi CI, Dudás G, Földes F, Kurucz K, Madai M, Zeghib S, Maes P, Vanmechelen B, Jakab F. Isolation of infectious Lloviu virus from Schreiber's bats in Hungary. *Nat Commun.* 2022;Mar 31;13(1):1706. doi: 10.1038/s41467-022-29298-1. PMID: 35361761; PMCID: PMC8971391.

Koch LK, Cunze S, Kochmann J, Klimpel S. Bats as putative Zaire ebolavirus reservoir hosts and their habitat suitability in Africa. *Sci Rep.* 2020 Aug 31;10(1):14268. doi: 10.1038/s41598-020-71226-0. PMID: 32868789; PMCID: PMC7459104.

Kohl C, Brinkmann A, Radonić A, Dabrowski PW, Mühldorfer K, Nitsche A, Wibbelt G, Kurth A. The virome of German bats: comparing virus discovery approaches. *Sci Rep.* 2021; Apr 1;11(1):7430. doi: 10.1038/s41598-021-86435-4. PMID: 33795699; PMCID: PMC8016945.

Kunz TH, de Torrez EB, Bauer D, Lobova T, Fleming TH. Ecosystem services provided by bats. *Annals of the New York Academy of Sciences*, 2011; vol. 1223, no. 1, pp. 1–38

Kuss SK, Best GT, Etheredge CA, Pruijssers AJ, Frierson JM, Hooper LV, Dermody TS, Pfeiffer JK. Intestinal microbiota promote enteric virus replication and systemic pathogenesis. *Science.* 2011 Oct 14;334(6053):249-52. doi: 10.1126/science.1211057. PMID: 21998395; PMCID: PMC3222156.

Lam TT, Jia N, Zhang YW, Shum MH, Jiang JF, Zhu HC, Tong YG, Shi YX, Ni XB, Liao YS, Li WJ, Jiang BG, Wei W, Yuan TT, Zheng K, Cui XM, Li J, Pei GQ, Qiang X, Cheung WY, Li LF, Sun FF, Qin S, Huang JC, Leung GM, Holmes EC, Hu YL, Guan Y, Cao WC. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature.* 2020 Jul;583(7815):282-285. doi: 10.1038/s41586-020-2169-0.

Leeuwen M, William M, Koraka P, Simon J, Smits S, Osterhaus A. Human Picobirnaviruses Identified by Molecular Screening of Diarrhea Samples. *Journal of clinical microbiology.* 2010; 48. 1787-94. 10.1128/JCM.02452-09.

Li L, Victoria JG, Wang C, Jones M, Fellers GM, Kunz TH, Delwart, E. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *Journal of virology*, 2010; 84(14), 6955–6965. <https://doi.org/10.1128/JVI.00501-10>

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med*. 2020 Mar 26;382(13):1199-1207. doi: 10.1056/NEJMoa2001316. Epub 2020 Jan 29. PMID: 31995857; PMCID: PMC7121484.

Li Y, Altan E, Reyes G, Halstead B, Deng X, Delwart E. Virome of Bat Guano from Nine Northern California Roosts. *J Virol*. 2021; Jan 13;95(3):e01713-20. doi: 10.1128/JVI.01713-20. PMID: 33115864; PMCID: PMC7925108.

Li Y, Ge X, Hon CC, Zhang H, Zhou P, Zhang Y, Wu Y, Wang LF, Shi Z. Prevalence and genetic diversity of adeno-associated viruses in bats from China. *J Gen Virol*. 2010 Oct;91(Pt 10):2601-9. doi: 10.1099/vir.0.020032-0. Epub 2010 Jun 23. PMID: 20573859.

Li N, Parrish M, Chan TK, Yin L, Rai P, Yoshiyuki Y, Abolhassani N, Tan KB, Kiraly O, Chow VT, Engelward BP. Influenza infection induces host DNA damage and dynamic DNA damage responses during tissue regeneration. *Cell Mol Life Sci*. 2015 Aug;72(15):2973-88. doi: 10.1007/s00018-015-1879-1. Epub 2015 Mar 26. PMID: 25809161; PMCID: PMC4802977.

Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S, Wang LF. Bats are natural reservoirs of SARS-like coronaviruses. *Science*. 2005 Oct 28;310(5748):676-9. doi: 10.1126/science.1118391. Epub 2005 Sep 29. PMID: 16195424.

Liu K, Pan X, Li L, Yu F, Zheng A, Du P, Han P, Meng Y, Zhang Y, Wu L, Chen Q, Song C, Jia Y, Niu S, Lu D, Qiao C, Chen Z, Ma D, Ma X, Tan S, Zhao X, Qi J, Gao GF, Wang Q. Binding and molecular basis of the bat coronavirus RaTG13 virus to ACE2 in humans and other

species. *Cell*. 2021 Jun 24;184(13):3438-3451.e10. doi: 10.1016/j.cell.2021.05.031. Epub 2021 May 24. PMID: 34139177; PMCID: PMC8142884.

Liu H, Zhu J, Hu Q, Rao X. *Morganella morganii*, a non-negligent opportunistic pathogen. *Int J Infect Dis*. 2016; Sep;50:10-7. doi: 10.1016/j.ijid.2016.07.006.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bembien LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005 Sep 15;437(7057):376-80. doi: 10.1038/nature03959. Epub 2005 Jul 31. Erratum in: *Nature*. 2006 May 4;441(7089):120. Ho, Chun He [corrected to Ho, Chun Heen]. PMID: 16056220; PMCID: PMC1464427.

Mbehang Nguema PP, Onanga R, Ndong Atome GR, Obague Mbeang JC, Mabika Mabika A, Yaro M, Lounnas M, Dumont Y, Zohra ZF, Godreuil S, Bretagnolle F. Characterization of ESBL-Producing Enterobacteria from Fruit Bats in an Unprotected Area of Makokou, Gabon. *Microorganisms*. 2020; Jan 19;8(1):138. doi: 10.3390/microorganisms8010138.

Mendenhall IH, Wen DLH, Jayakumar J, Gunalan V, Wang L, Mauer-Stroh S, Su YCF, Smith GJD. Diversity and Evolution of Viral Pathogen Community in Cave Nectar Bats (*Eonycteris spelaea*). *Viruses*. 2019; 11(3):250. <https://doi.org/10.3390/v11030250>

Milne-Price S, Miazgowicz KL, Munster VJ. The emergence of the Middle East Respiratory Syndrome coronavirus. *Pathog Dis*. 2014;71, 119–134, 10.1111/2049-632X.12166

Mohsin H, Asif A, Fatima M, Rehman Y. Potential role of viral metagenomics as a surveillance tool for the early detection of emerging novel pathogens. *Archives of microbiology*, 2021; 203(3), 865–872. <https://doi.org/10.1007/s00203-020-02105-5>

Monchatre-Leroy E, Boué F, Boucher JM, Renault C, Moutou F, Ar Gouilh M, Umhang G. Identification of Alpha and Beta Coronavirus in Wildlife Species in France: Bats, Rodents, Rabbits, and Hedgehogs. *Viruses*. 2017 Nov 29;9(12):364. doi: 10.3390/v9120364. PMID: 29186061; PMCID: PMC5744139.

Monro AK, Bystriakova N, Fu L, Wen F, Wei Y. Discovery of a diverse cave flora in China. *PLoS ONE*. 2018; 13(2): e0190801. <https://doi.org/10.1371/journal.pone.0190801>

Mora M, Wicaksono WA, Egamberdieva D, Krause R, Martinez JL, Cernava T, Berg G. Explorative assessment of coronavirus-like short sequences from host-associated and environmental metagenomes. *Sci Total Environ*. 2021; Nov 1;793:148494. doi: 10.1016/j.scitotenv.2021.148494. Epub 2021 Jun 24. PMID: 34328954; PMCID: PMC8222970.

Munster VJ, Adney DR, van Doremalen N, Brown VR, Miazgowiec KL, Milne-Price S, Bushmaker T, Rosenke R, Scott D, Hawkinson A, de Wit E, Schountz T, Bowen RA. Replication and shedding of MERS-CoV in Jamaican fruit bats (*Artibeus jamaicensis*). *Sci Rep*. 2016; 22;6:21878. doi: 10.1038/srep21878. PMID: 26899616; PMCID: PMC4761889.

Munster VJ, Adney DR, van Doremalen N, Brown VR, Miazgowiec KL, Milne-Price S, Bushmaker T, Rosenke R, Scott D, Hawkinson A, de Wit E, Schountz T, Bowen RA. Replication and shedding of MERS-CoV in Jamaican fruit bats (*Artibeus jamaicensis*). *Sci Rep*. 2016 Feb 22;6:21878. doi: 10.1038/srep21878. PMID: 26899616; PMCID: PMC4761889.

Muscarella R, Fleming TH. The role of frugivorous bats in tropical forest succession. *Biological Reviews*. 2007; vol. 82, no. 4, pp. 573–590,

NIH(National Human Genome Research Institute). Metagenomics. <https://www.genome.gov/genetics-glossary/Metagenomics#:~:text=Metagenomics%20is%20the%20study%20of,or%20in%20a%20water%20sample>. Accessed: 13.06.2022

NCBI (National Center for Biotechnology Information). MAG: *Turdus hortulorum* Iteradensovirus isolate gbt104par01 genomic sequence, sequence. <https://www.ncbi.nlm.nih.gov/nuccore/2015373096>. Accessed on: 17.08.2022

O'Shea TJ, Cryan PM, Cunningham AA, Fooks AR, Hayman DTS, Luis AD, Peel AJ, Plowright RK, Wood JLN. Bat flight and zoonotic viruses. *Emerg Infect Dis*, 2014;20: 741-745.

Prasad V. Ecofriendly Pest Management for Food Security. *Insect Viruses*. 2016; 411–442. doi:10.1016/b978-0-12-803265-7.00013-0

Papenfuss AT, Baker ML, Feng ZP, Tachedjian M, Crameri G, Cowled C, Ng J, Janardhana V, Field HE, Wang LF. The immune gene repertoire of an important viral reservoir, the Australian black flying fox. *BMC Genomics*. 2012 Jun 20;13:261. doi: 10.1186/1471-2164-13-261. PMID: 22716473; PMCID: PMC3436859.

Parashar UD, Sunn LM, Ong F, Mounts AW, Arif MT, Ksiazek TG, Kamaluddin MA, Mustafa AN, Kaur H, Ding LM. Case-control study of risk factors for human infection with a new zoonotic paramyxovirus, Nipah virus, during a 1998-1999 outbreak of severe encephalitis in Malaysia. *J. Infect. Dis*. 2000;181:1755–1759.

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol*. 2017 Sep 12;35(9):833-844. doi: 10.1038/nbt.3935. Erratum in: *Nat Biotechnol*. 2017 Dec 8;35(12):1211. PMID: 28898207.

Rumberia C, Ommeh S, Obanda V, Agwanda B, Gakuya F, Ge X, Yang X, Wu LJ, Zohaib A, Hu B, Shi ZL. Molecular detection of viruses in Kenyan bats and discovery of novel astroviruses, caliciviruses and rotaviruses. *Virologica Sinica*. 2017; 32. 10.1007/s12250-016-3930-2.

Pinus M, Müller HE. Enterobakterien bei Fledertieren (Chiroptera) [Enterobacteria of bats (Chiroptera) (author's transl)]. *Zentralbl Bakteriologie A*. 1980 Aug;247(3):315-22. German. PMID: 7424248.

Regenmortel (ed.), *Encyclopedia of virology*, 2008; 3rd ed., vol. 2. Elsevier, Oxford, United Kingdom

Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011 Jul 20;475(7356):348-52. doi: 10.1038/nature10242. PMID: 21776081.

Sabatier M, Bal A, Destras G, Regue H, Quéromès G, Cheynet V, Lina B, Bardel C, Brengel-Pesce K, Navratil V. Comparison of nucleic acid extraction methods for a viral metagenomics analysis of respiratory viruses *Microorganisms*. 2020; 8 p. 1539

Salmier A, Tirera S, de Thoisy B, Franc A, Darcissac E, Donato D, Bouchier C, Lacoste V, Lavergne A. Virome analysis of two sympatric bat species (*Desmodus rotundus* and *Molossus molossus*) in French Guiana. *PLoS One*. 2017 Nov 8;12(11):e0186943. doi: 10.1371/journal.pone.0186943. PMID: 29117243; PMCID: PMC5695591.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 1977; 74, 5463–5467. doi: 10.1073/pnas.74.12.5463

Šimić I, Zorec TM, Lojkić I, Krešić N, Poljak M, Cliquet F, Picard-Meyer E, Wasniewski M, Zrnčić V, Čukušić A, Bedeković T. Viral Metagenomic Profiling of Croatian Bat Population Reveals Sample and Habitat Dependent Diversity. *Viruses*. 2020 Aug 14;12(8):891. doi: 10.3390/v12080891. PMID: 32824037; PMCID: PMC7472731.

Speakman JR. The impact of predation by birds on bat populations in the British Isles. *Mammal Review*, 1991; vol. 21, no. 3, pp. 123–142,

Streicker DG, Recuenco S, Valderrama W, Gomez Benavides J, Vargas I, Pacheco V, Condori RE, Montgomery J, Rupprecht CE, Rohani P, Altizer S. Ecological and anthropogenic drivers of rabies exposure in vampire bats: implications for transmission and control. *Proc Biol Sci*. 2012; Sep 7;279(1742):3384-92. doi: 10.1098/rspb.2012.0538.

Solonenko SA, Sullivan MB. Preparation of metagenomic libraries from naturally occurring marine viruses. *Methods Enzymol*. 2013; 531 pp. 143-165

Sun DL, Gao YZ, Ge XY, Shi ZL, Zhou NY. Special Features of Bat Microbiota Differ from Those of Terrestrial Mammals. *Frontiers in microbiology*, 2020; 11: 1040. <https://doi.org/10.3389/fmicb.2020.01040>

Sothearen T, Furey NM, Jurgens JA. Effect of bat guano on the growth of five economically important plant species. *Journal of Tropical Agriculture*. 2014; vol. 52, no. 2, pp. 169–173,

Tang XC, Zhang JX, Zhang SY, Wang P, Fan XH, Li LF, Li G, Dong BQ, Liu W, Cheung CL, Xu KM, Song WJ, Vijaykrishna D, Poon LL, Peiris JS, Smith GJ, Chen H, Guan Y. Prevalence and genetic diversity of coronaviruses in bats from China. *J Virol*. 2006 Aug;80(15):7481-90. doi: 10.1128/JVI.00697-06. PMID: 16840328; PMCID: PMC1563713.

Tao Y, Shi M, Chommanard C, Queen K, Zhang J, Markotter W, Kuzmin IV, Holmes EC, Tong S. Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History. *J Virol.* 2017 Feb 14;91(5):e01953-16. doi: 10.1128/JVI.01953-16. PMID: 28077633; PMCID: PMC5309958.

Tao Y, Shi M, Chommanard C, Queen K, Zhang J, Markotter W, Kuzmin IV, Holmes EC, Tong S. Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History. *J Virol.* 2017 Feb 14;91(5):e01953-16. doi: 10.1128/JVI.01953-16. PMID: 28077633; PMCID: PMC5309958.

Thomas SP, Suthers RA. The physiology and energetics of bat flight. *J. Exp. Biol.* 1972; 57, 317–335.1972.

Thomas DJ, Boyd M, Crowell KM, Curtwright AE, Foll MN, Kuehl MM, McQueen CM, Middaugh R, Moore VM, Moreno M. A biological inventory of Meacham Cave (independencecounty, Arkansas). *J. Ark. Acad. Sci.* 2012,66, 126–13

USGS(United States Geological Survey)New Study Unveils the Impacts of the Disease on Native Bat Populations.<https://www.usgs.gov/news/national-news-release/white-nose-syndrome-killed-over-90-three-north-american-bat-species#:~:text=White%2Dnose%20syndrome%20is%20a,dehydration%2C%20starvation%20and%20often%20death>. Accessed on: 15.06.2022

Wang LF, Anderson DE, Mackenzie JS, Merson, MH. From Hendra to Wuhan: what has been learned in responding to emerging zoonotic viruses. *Lancet* (London, England), 2020; 395(10224), e33–e34. [https://doi.org/10.1016/S0140-6736\(20\)30350-0](https://doi.org/10.1016/S0140-6736(20)30350-0)

WHO (World Health Organization). Genuine intersectoral collaboration is needed to achieve better progress in vector control. <https://www.who.int/news/item/11-04-2022-genuine-intersectoral-collaboration-is-needed-to-achieve-better-progress-in-vector-control>. Accessed: 03.08.2022

WHO (World Health Organization). Zoonoses. 2020; <https://www.who.int/news-room/fact-sheets/detail/zoonoses#:~:text=A%20zoonosis%20is%20an%20infectious,food%2C%20water%20or%20the%20environment>. Accessed: 12.06.2022

WHO. 2021. *WHO-convened global study of origins of SARS-CoV-2*. WHO. Available from: <https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part>.

Witney AA, Aller S, Strang BL. Metagenomic profiling of placental tissue suggests DNA virus infection of the placenta is rare. *J Gen Virol*. 2021 Nov;102(11):001677. doi: 10.1099/jgv.0.001677. PMID: 34723784; PMCID: PMC8742990.

Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A*, 1977; 74, 5088–5090. doi: 10.1073/pnas.74.11.5088

Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Bio*. 2019; 20, 257 <https://doi.org/10.1186/s13059-019-1891-0>

Wood DE, J Lu. About Kraken. The Center for Computational Biology at Johns Hopkins University.2020

Woolhouse ME, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. *Emerg Infect Dis*, 2005; 11,12:1842.

Wu Z, Ren X, Yang L, Hu Y, Yang J, He G, Zhang J, Dong J, Sun L, Du J, Liu L, Xue Y, Wang J, Yang F, Zhang S, Jin Q. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. *J Virol*. 2012 Oct;86(20):10999-1012. doi: 10.1128/JVI.01394-12. Epub 2012 Aug 1. PMID: 22855479; PMCID: PMC3457178.

Wu Z, Yang L, Ren X, He G, Zhang J, Yang J, Qian Z, Dong J, Sun L, Zhu Y, Du J, Yang F, Zhang S, Jin Q. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *The ISME journal*. 2016;10(3), 609–620. <https://doi.org/10.1038/ismej.2015.138>

Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, Li N, Guo Y, Li X, Shen X, Zhang Z, Shu F, Huang W, Li Y, Zhang Z, Chen RA, Wu YJ, Peng SM, Huang M, Xie WJ, Cai QH, Hou FH, Chen W, Xiao L, Shen Y. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature*. 2020 Jul;583(7815):286-289. doi: 10.1038/s41586-020-2313-x. Epub 2020 May 7. Erratum in: *Nature*. 2021 Dec;600(7887):E8-E10. PMID: 32380510.

Yinda CK, Ghogomu SM, Conceição-Neto N, Beller L, Deboutte W, Vanhulle E, Maes P, Ranst MV, Matthijnssens J. Cameroonian fruit bats harbor divergent viruses, including rotavirus H, bastroviruses, and picobirnaviruses using an alternative genetic code. *Virus Evolution*, 2018; 4(1)

Yorulmaz T, Arslan N. Current Status of The Bats in Türkiye with Their Ecogeographic Distributions and Recommendations for National Conservation Status (Mammalia: Chiroptera). *Fresenius Environmental Bulletin*, 2020; 29:6691-6706.

Yuk Y, Albayrak I, Ergene A. A contribution to the microbiological characterization of bat guano. *Munis Entomology Zoology*, 2019;14 (1): 188-191

Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*. 2012 Nov 8;367(19):1814-20. doi: 10.1056/NEJMoa1211721. Epub 2012 Oct 17. Erratum in: *N Engl J Med*. 2013 Jul 25;369(4):394. PMID: 23075143.

Zhang F, Zuo T, Yeoh YK, Cheng FWT, Liu Q, Tang W, Cheung KCY, Yang K, Cheung CP, Mo CC, Hui M, Chan FKL, Li CK, Chan PKS, Ng SC. Longitudinal dynamics of gut bacteriome, mycobiome and virome after fecal microbiota transplantation in graft-versus-host

disease. *Nat Commun.* 2021 Jan 4;12(1):65. doi: 10.1038/s41467-020-20240-x. PMID: 33397897; PMCID: PMC7782528.

Zheng XY, Qiu M, Guan WJ, Li JM, Chen SW, Cheng MJ, Huo ST, Chen Z, Wu Y, Jiang LN, Chen Q. Viral metagenomics of six bat species in close contact with humans in southern China. *Arch Virol*, 2018 Jan;163(1):73-88. doi: 10.1007/s00705-017-3570-3.

Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020 Mar;579(7798):270-273. doi: 10.1038/s41586-020-2012-7.

8. APPENDIX 1: CV

ÖYKÜ DURAK

Narlıdere/İZMİR

oykudurakk@gmail.com

EDUCATION

Dokuz Eylül University

2019-2022

Izmir International Biomedicine and Genome Institute,
Department of Genome Sciences and Molecular Biotechnology
Molecular Biology and Genetics M.Sc. Program (English) - Izmir, TURKEY
GPA: 3.94/4.00

Thesis Project: Identification of Virome Content in Fecal Samples of Bats in Northwest and Central Anatolia

Istanbul Technical University

2014-2019

Faculty of Science and Letters
Department of Molecular Biology and Genetics B.Sc. Program (English) – Istanbul, TURKEY
GPA: 3.20/4.00,
Honor List 2015-2016
Fall-Honor List 2018-2019
Thesis Project: Analysis of New Variants on Autoinflammatory Genes in *MEFV* Negative FMF Patients

RELEVANT EXPERIENCES

Master Thesis Project

2019-2022

Emerging Viral Diseases Lab, Izmir Biomedicine and Genome Center, IBG
• Identification of Virome Content in Fecal Samples of Bats in Northwest and Central Anatolia

TUBITAK (The Scientific and Technological Research Council of Turkey) 1001 Project – Bursar

2021-2022

TUBITAK (The Scientific and Technological Research Council of Turkey) STAR - Bursar

2020-2021

Master Intern Researcher Scholarship Program

Internship

07.2018-08.2018

Emerging Viral Diseases Lab, Izmir Biomedicine and Genome Center, IBG

SKILLS AND TECHNIQUES

Laboratory: Viral Nucleic Acid Extraction, cDNA Synthesis, Cell Culture Methods (HUH7, MDCK), Transfection & Transduction (Lentivirus), Virus Propagation in Embryonic Chicken Eggs, Hemagglutination Assay, Gene cloning, EID50, RT-PCR, PCR, Fecal sample collection in Aquatic fieldwork, Experiences in BSL2+ and BSL2

Bioinformatics: R studio (beginner)

Language: English (B1-B2), Turkish (Native), German(A1)