UNIVERSITÁ DEGLI STUDI DI VERONA

DEPARTMENT OF BIOTECHNOLOGY GRADUATE

SCHOOL OF NATURAL AND ENGINEERING

SCIENCES

DOCTORAL PROGRAM IN BIOTECHNOLOGY

CYCLE XXXVI

TITLE OF THE DOCTORAL THESIS

Development and application of innovative strategies for identification and characterization of viruses

Coordinator:	Professor Matteo Ballottari
Tutor:	Professor Nicola Vitulo
Co-tutor:	Doctor Isabella Monne
Co-tutor:	Doctor Alice Fusaro

Doctoral Student: Ambra Pastori

Development and application of innovative strategies for identification and characterization of viruses

Ambra Pastori

PhD Thesis

Verona, 03/2024

Well, I must endure the presence of a few caterpillars if I wish to become acquainted with the butterflies.

-Antoine de Saint-Exupéry

Table of contents

Sommario

Come testimoniato dalla pandemia COVID-19, la capacità di identificare rapidamente e caratterizzare geneticamente i virus emergenti, con un focus su quelli con impatto zoonotico, è sempre più importante per la salute pubblica. I virus hanno tassi di mutazione elevati e possono essere coinvolti in frequenti eventi di ricombinazione e riassortimento. Queste proprietà evolutive portano ad un'alta diversità genetica e all'emergenza di nuove varianti che possono avere un impatto zoonotico, virulenza e diffusione più elevati. In uno scenario come questo, disporre di protocolli di laboratorio e di pipeline bioinformatiche capaci di identificare e caratterizzare tempestivamente varianti virali emergenti è cruciale in piani di early warning volti a prevenire o controllare tempestivamente nuove minacce infettive.

Questa tesi si concentra sull'identificazione e caratterizzazione dei dati di sequenziamento genetico, sviluppando un flusso di lavoro completo per l'analisi dei dati NGS, partendo dai dati grezzi generati con differenti approcci di sequenziamento, fino alla creazione della sequenza consenso e all'applicazione di metodi filogenetici e filodinamici per studiare le dinamiche evolutive. Per la caratterizzazione di SARS-CoV-2, agente causativo della pandemia COVID-19, sono stati sviluppati protocolli ad hoc ed è stata condotta una comparazione tra le tecnologie di sequenziamento di seconda (MiSeq, Illumina) e terza (MinION, Nanopore) generazione. Le tecnologie di sequenziamento di seconda generazione e il loro protocollo sono stati applicati all'analisi genetica di SARS-CoV-2 identificato in un gatto domestico, rivelando un'infezione con la variante B.1.177.

Il miglioramento, l'ottimizzazione e l'applicazione delle analisi bioinformatiche dei dati NGS sono stati realizzati anche per l'identificazione e caratterizzazione del genoma completo di Hantavirus rilevato nei topi.

Inoltre, utilizzando un approccio di sequenziamento diretto, così detto “untargeted”, è stato possibile caratterizzare il genoma di due virus a DNA nei pappagalli in Italia: il Polyomavirus e un possibile nuovo Circovirus.

Sono stati inoltre applicati approcci bioinformatici anche per l’analisi e lo studio del genoma del virus influenzale aviario ad alta patogenicità (sottotipo H5) (HPAI).

In particolare, attraverso la generazione del genoma completo di centinaia di virus H5 HPAI identificati durante le ondate epidemiche italiane del 2020-2023 che hanno colpito volatili domestici e selvatici e la successiva applicazione di analisi filogenetiche e filogeografiche bayesiane, è stato possibile rilevare la co-circolazione di diversi genotipi virali, ricostruire le dinamiche di diffusione del patogeno sul territorio nazionale ed esplorare il ruolo svolto dai diversi ospiti coinvolti nell'infezione.

Ma lo studio del genoma dei virus HPAI non si è limitato alla situazione nazionale e alle specie aviarie. Una grave epidemia causata ancora una volta da virus H5 HPAI ha colpito i gatti in Polonia nell'estate del 2023. Le analisi molecolari ed evolutive applicate al genoma completo dei virus coinvolti nell'infezione hanno svelato la circolazione di un virus con mutazioni capaci di aumentare l'adattamento del virus aviario al mammifero e ha consentito di chiarire che tutti gli animali coinvolti nell'evento epidemico erano stati esposti ad una comune fonte di infezione, verosimilmente di origine alimentare.

I risultati generati nel presente progetto di dottorato sottolineano la strategicità del dato genetico e delle analisi bioinformatiche per comprendere le dinamiche di emergenza, evoluzione e disseminazione dei virus e per fornire informazioni chiave nella valutazione del rischio associato ai patogeni virali.

Abstract

As testified by the COVID-19 pandemic, the ability to rapidly identify and genetically characterize emerging viruses, with focus on ones with zoonotic impact, is increasingly crucial for public health. Viruses have high mutation rates and may be involved in frequent recombination and reassortment events. These evolutionary properties lead to high genetic diversity and the emergence of new variants that may have higher zoonotic impact, virulence, and increased spread. In a scenario like this, having laboratory protocols and bioinformatic pipelines capable of promptly identifying and characterizing emerging viral variants is crucial in early warning plans aimed at preventing or promptly controlling new infectious threats.

This thesis focuses on the identification and characterization of genetic sequencing data, developing a comprehensive workflow for NGS data analysis, starting from raw data generated using different sequencing approaches to the consensus sequence creation and the phylogenetic and phylodynamic methods application to study evolutionary dynamics. For the characterization of SARS-CoV-2, causative agent of the COVID-19 pandemic, specific protocols were developed, and a comparison was conducted between second-generation sequencing technologies (MiSeq, Illumina) and third-generation sequencing technologies (MinION, Nanopore). Second-generation sequencing technologies and their protocol were applied to the genetic analysis of SARS-CoV-2 identified in a domestic cat, revealing an infection with the B.1.177 variant.

Improvements, optimizations, and the application of bioinformatic analyses of NGS data were also accomplished for the identification and characterization of the complete genome of Hantavirus detected in mice. Additionally, using an "untargeted" direct sequencing approach, the genome of two DNA viruses in parrots in Italy, namely Polyomavirus and a possible new Circovirus, could be characterized.

Bioinformatic approaches were also applied to the analysis and study of the genome of highly pathogenic avian influenza virus (subtype H5) (HPAI). Specifically, by generating the complete genome of hundreds of H5 HPAI viruses identified during

the Italian epidemic waves of 2020-2023 that affected domestic and wild birds and subsequently applying Bayesian phylogenetic and phylogeographic analyses, it was possible to reveal the co-circulation of different viral genotypes, reconstruct the pathogen's spread dynamics across the national territory, and explore the role played by different hosts in the infection.

However, the study of the genome of HPAI viruses was not limited to the national situation and avian species. A severe epidemic, caused once again by H5 HPAI viruses, affected cats in Poland in the summer of 2023. Molecular and evolutionary analyses applied to the complete genome of the viruses involved in the infection revealed the circulation of a virus with mutations capable of increasing the adaptation of the avian virus to mammals and clarified that all animals involved in the epidemic event had been exposed to a common source of infection, presumably of food origin.

The results generated in this PhD project emphasize the strategic importance of genetic data and bioinformatic analyses in understanding the dynamics of emergence, evolution, and dissemination of viruses and providing key information in the assessment of the risk associated with viral pathogens.

1 Background and aim of the research project

DNA and RNA virus populations are characterized by high mutation rates and many of them engage in frequent recombination and reassortment events, leading to the creation of novel genotypes. These evolutionary properties result in a high genetic diversity, which has a detrimental impact on the ability to identify and characterize the new and constantly emerging RNA viruses. Historically, molecular tools as polymerase chain reaction (PCR) or conventional non-molecular techniques, such as microscopy, enzyme-linked immunosorbent assay (ELISA) and virus isolation were used for virus identification, but they have several limitations [1]. In particular, isolation in cell culture is unsuitable for viruses for which no permissive cell line is known and molecular identification of viruses is mostly based on nucleic acids amplification protocols which are in need of genetic information for each tested pathogen and can generally be conducted only for a selected number of viruses due to laboratory practical difficulties. Furthermore, novel viruses are not identified with these techniques. The recent advances in metagenomic processes for virus discovery in clinical samples have opened new opportunities to understand the aetiology of unexplained illnesses and to increase the possibility of identifying emerging or re-emerging viruses. The combination of the genetic information obtained from the unbiased sequencing of new viruses with the development of targeted protocols to characterize their whole genome offers unprecedented opportunities for the improvement of the laboratory diagnostic flowchart. The unique properties of each sequencing technology commercially available affect the way the data are assembled as well as the completeness and accuracy of the resulting genomes. Specifically, Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe), where I am employed as bioinformatician at Viral Genomics and Transcriptomics Laboratory, is equipped with both second (Illumina MiSeq and NextSeq 550) and third (MinION sequencer from Oxford Nanopore) generation platforms. Illumina MiSeq and NextSeq 550 produce very high quality, high accuracy, but short, up to 300bp and 150bp, respectively, reads with high costs. It takes days to obtain the results and the sample preparation is complex [2]. The MinION sequencer from Oxford Nanopore is an evolving technology that produces long-read sequencing data with low equipment cost. It belongs to the third

generation platforms category, it has relatively simple sample preparation, flexible run times and portability, even if it has lower accuracy than second generation platforms.

The major aim of this research project is to develop a diagnostic flowchart for rapid identification and characterization of RNA viruses detected in animal species with a focus on zoonotic viruses using both platforms for routine analysis and research. Furthermore, the integration between the results generated by the two platforms can give more robust results and more informative data. The relatively higher error rates of MinION can be used in conjunction with lower error rate (2 x 300bps) Illumina data that are too short to cover, for example, multiple informative variants in samples composed of two or more genomes [3]. As testified by the recent epidemics and pandemic, the ability to rapidly identify and genetically characterize emerging viruses, with focus on the ones with zoonotic impact, is increasingly important for the public health. The unique properties of viruses result in a high genetic diversity and emergence of new variants that can have higher zoonotic impact, virulence and spread. This leads to the development of a*d hoc* improved laboratory protocols and bioinformatics pipelines for the genetic identification and characterization of viruses. Specifically, this project focused on improvement, optimization and application of the bioinformatics analyses of NGS data for viral characterization obtained from untargeted approaches in MiSeq platform, which allowed the identification not only of RNA, but also of DNA viruses such as Polyomavirus and a possible novel Circovirus in parrots in Italy. Furthermore, the complete genome of a Hantavirus from mice was characterized. In the last part of my PhD project, the pipelines developed for consensus sequences generation were used and downstream phylogenetic and network analyses were performed, with identification of genotypes and mutations, as well as Bayesian evolutionary analysis, using BEAST cross-platform program.

In 1996, high pathogenic avian influenza (HPAI) A(H5) virus of the A/goose/Guangdong/1/1996 (Gs/Gd) lineage was identified for the first time in poultry in China. In the spring of 2005, this strain spread reached for the first time Europe and Africa during the wild bird autumn migration [4]. Since then, the HPAI A(H5) virus of the Gs/Gd lineage has widely spread across Asia, Europe, Africa

and North America, becoming a serious global threat for the poultry industry, endangered wild bird species and representing a zoonotic risk for human health. The focus was on the analysis on the HPAI A(H5) viruses collected during the 2020-2023 Italian epidemic waves, and HPAI A(H5N1) viruses responsible of outbreaks of HPAI AIV H5N1 in cats in Poland (June-July 2023).

References

1. Grubaugh, N. D., Ladner, J. T., Lemey, P., Pybus, O. G., Rambaut, A., Holmes, E. C., & Andersen, K. G. (2019). Tracking virus outbreaks in the twenty-first century. Nature Microbiology, 4(1), 10–19. https://doi.org/10.1038/S41564-018-0296-2

2. Schadt, E. E., Turner, S., & Kasarskis, A. (2010). A window into third-generation sequencing. https://doi.org/10.1093/hmg/ddq416

3. Lindberg, M. R., Schmedes, S. E., Curtis Hewitt, F., Haas, J. L., Ternus, K. L., Kadavy, D. R., & Budowle, B. (2016). A Comparison and Integration of MiSeq and MinION Platforms for Sequencing Single Source and Mixed Mitochondrial Genomes. https://doi.org/10.1371/journal.pone.0167600

4. Olsen, B.; Munster, V.J.; Wallensten, A.; Waldenström, J.; Osterhaus, A.D.M.E.; Fouchier, R.A.M. Global patterns of influenza a virus in wild birds. Science 2006, 312, 384–8, doi:10.1126/science.1122438.

2 Introduction

A virus is a microscopic organism and it consists of proteins and nucleic acid, DNA or RNA as genetic material. Viruses do not contain ribosomes, mitochondria or other cell-like organelles, and since they cannot replicate without the metabolic processes of the host cell, they are genetic parasites that need to infect hosts as humans, plant, animals, bacteria and fungi to reproduce. Their mode of replication makes viruses unique: burst of thousands of virus particles from a single virus over a short time. Before the invention of the electron microscope (1931), they were impossible to visualize. The first viruses to be visualized were bacteriophage and they appeared to have a head and tail-like structure, through it the nucleic acid entered the bacterial cell. Animal viruses were described as spherical or rod-shaped; they were bound to receptors and were taken up by the cell. Studies of viral replication indicate that most viruses self-assemble as a result of interactions between the viral proteins to form a viral capsid that interacts with the nucleic acid to form the whole [1].

2.1 RNA Virus

RNA viruses have RNA as genetic material, single-stranded RNA or double stranded RNA. Viruses may exploit the presence of RNA-dependent RNA polymerases (RdRp) for replication of their genomes or, in retroviruses, with two copies of single strand RNA genomes, reverse transcriptase produces viral DNA which can be integrated into the host DNA under its integrase function [2]. The RNA genome is used as template for synthesis of additional RNA strands. The replicase complex usually consists of the RdRp and RNA-helicases, to unwind base-paired regions of the RNA genome, and NTPases, to supply energy for the polymerization process. The genomes of RNA viruses have one or more open reading frames that encode the viral proteins and also regions of RNA that do not code for protein and they are conserved (non-coding regions, NCRs, or untranslated regions, UTRs). NCRs are present at 5’ and 3’ ends of all RNA genomes and they direct genome replication. An important feature of RNA viruses is that many exist in nature as quasispecies, a group of closely related, but nonidentical genomes. The RdRps of RNA viruses control RNA synthesis error rates. RdRp mutations can have

impact on fidelity and this measurably affects virus fitness [3]. RNA viruses can be positive or negative-strand. The genomes of positive-strand RNA viruses are functional mRNAs and they are translated shortly after penetration into the host cell. They produce the RdRp and other proteins necessary for synthesis of additional viral RNAs. Positive-strand RNA viruses, for genome replication, use large complexes of cellular membranes and they actively modify host cell membranes. On the other hand, negative-strand RNA viruses, the ambisense RNA viruses, and double-stranded RNA viruses need to trascript their genome by viral proteins, RdRp included [3].

2.1.1 Virus Recombination

The phenomenon of Influenza A recombination plays a pivotal role in shaping viral diversity and it contributes to the evolution of viruses. There are diverse mechanisms through which viruses undergo recombination, such as template switching during replication and reassortment in segmented viruses that can be detected and studied through computational algorithms and experimental approaches. The ability of influenza A viruses to undergo genetic reassortment introduces significant variability in their genetic makeup, leading to the emergence of novel strains with unique combinations of genetic material. This process of recombination occurs when different influenza A viruses infect the same host cell, enabling the exchange of genetic segments. The resulting progeny virus carries a mosaic of genetic material derived from the parental strains, contributing to the vast diversity observed in influenza A viruses. Recombination is a driving force behind the constant emergence of new influenza A virus variants, allowing the virus to evade host immune responses and potentially gain new biological properties, such as altered transmissibility or virulence. Understanding influenza A virus recombination is crucial for anticipating and responding to changes in the virus's genetic makeup. It poses challenges for vaccine development, as the circulating strains may differ from those included in current vaccines. Additionally, monitoring recombination events is essential for public health surveillance, enabling timely responses to emerging strains that may pose increased risks of transmission or severity. The evolutionary consequences of the dynamic process of influenza A

virus recombination are shaping viral populations, driving genetic diversity, and potentially influencing viral adaptation and pathogenicity [4].

2.1.2 Influenza A Virus

Influenza virus is a negative stranded RNA and it belongs to the Orthomyxoviridae family, which is divided into five genera: Influenzavirus A, Influenzavirus B, Influenzavirus C, Isavirus and Thogotovirus. Viruses of the Influenzavirus A genus are the only known to infect birds, that are a natural reservoir of influenza A viruses. Influenza A viruses are classified into subtypes based on the antigenic relationships of the surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA), in any combination. To date, 17 HA subtypes (H1– H17) and 10 NA subtypes (N1–N10) have been recognized [5].

2.1.2.1 Biology of Influenza A Virus

Influenza A virus genome is segmented and this characteristic facilitates the virus reassortment and the high rate of mutation, that makes the virus rapidly adaptive, since it uses its own error-prone RNA polymerase during replication [6]. The RNA genome of the influenza A virus is short (~13.5 Kb), single stranded and "negative sense": the viral RNA segments serve as templates for mRNAs transcription [7]. Influenza A virus possesses a genome formed by eight segments, which are essential for efficient virus replication and they encode for 12 viral proteins (Figure 1). Segment 1 is the largest and it is the basic polymerase protein 2 (PB2) that encodes for a cap-binding protein which is essential for viral mRNA synthesis [8]. Segment 2, the basic polymerase protein 1 (PB1) encodes for the main polymerase protein that is involved in chain elongation. It also encodes for the PB1-F2 protein, in alternate reading frame, that plays a role in cell apoptosis [9]. Segment 3 encodes for the acid polymerase protein (PA) that is required for viral transcription, replication and transport to the nucleus but its specific function remains partially unclear [10]. As for segment 2, segment 3 contains a second open reading frame (X-ORF), via ribosomal frameshifting. Its product, PA-X, functions to repress cellular gene expression. The kinetics of the global host response, which notably includes increases in inflammatory, apoptotic and T lymphocyte-signaling

pathways, changes with PA-X loss of expression [11]. Segment 4, HA gene, encodes for hemagglutinin (HA), the main surface glycoprotein, which binds to host sialic acid receptors. It allows the entrance of the virus into the cell and fuses viral and endosomal membranes to release viral nucleocapsids into the cell cytoplasm [12]. The haemagglutinin undergoes post-translational cleavage by a protease dividing the protein into the HA1 and the HA2. It consists of two distinct domains: a globular head formed by the HA1 that contains a conserved receptor binding site and five antigenic regions A-E and a fibrous stem formed by both HA1 and HA2 [13; 14]. Segment 5, the nucleoprotein, provides the core of the ribonucleoprotein (RNP) complex binding the vRNA with the polymerase complex. NP controls the switching of the RNA polymerase from transcription to replication [15]. Segment 6, NA gene, encodes for neuraminidase (NA) glycoprotein, which is involved in cleaving the sialic acid receptor to facilitate viral release from infected cells [16]. Segment 7, the MP protein, encodes for two proteins: matrix protein 1 (M1) and matrix protein 2 (M2), due to two overlapping reading frames, that are translated from spliced mRNAs [17]. M1 protein is involved in functions related to the assembly and disassembly of viral particles, the transport of the RNPs to the nucleus, the nuclear export of viral proteins and viral morphology [18], while M2 protein is essential for viral infection of host cells [19]. Segment 8 is the smallest viral segment and it contains two overlapping reading frames. NS1 and NS2 are derived from alternatively spliced RNAs that are transcribed from segment 8. The NS1 is the main non-structural protein and interacts with host cell processes to promote viral protein synthesis. NS1 also sequesters dsRNA to prevent induction of the host interferon response [20]. NS2 is a structural component of the viral particle and it is frequently called the “nuclear export protein” (NEP) for its role in transporting newly synthesized RNP from the nucleus to the cytoplasm [21]. In addition, NS2 interacts with nucleoporins and was suggested to serve as an adaptor between vRNPs and the nuclear pore complex [21]. A role of NS2 in the regulation of influenza virus transcription and replication has also been proposed [22]. The influenza A virus particle is 80-120 nm in diameter and it is formed by a lipid envelope derived from the host's cell membrane during viral budding [23]. M2 ion channel protein and the HA and NA surface glycoproteins are embedded in the

envelope. The viral matrix protein (M1) is under the lipid bilayer. Inside the virion, the viral nucleoprotein (NP) in association with a polymerase complex consisting of PB2, PB1 and PA forming ribonucleoprotein (RNP) particles encloses each genome segment. The globular head of the viral HA binds the sialic acid (N-acetyl-neuraminic acid) receptors found on host cells to initiate the infection of host cells, followed by viral entry by endocytosis. In mammals, influenza viruses typically infect cells binding to a sialic acid linked to galactose by a α2,6-linkage, while avian influenza viruses typically infect the avian species preferentially binding to a α2,3-linked sialic acid [24; 25]. The acid environment induces a pH-dependent conformational change in the HA that causes the viral and endosomal membranes to fuse and release viral RNPs into the cytoplasm [26; 27]. The HA0 precursor molecule must be cleaved into HA1 and HA2 subunits by a host cell protease in order to the fusion to occur. Several amino acids at the cleavage site determine the cleavability of HA0. Viral infection happens to host cells where proteases that recognize the cleavage site are found. After the virus entries the cell, H ions flood through the M2 ion channel, causing the separation of the M1 from the RNPs in order to allow the transport of the RNPs from the cytoplasm into the nucleus. In the nucleus, the virus uses host cell machinery during its transcription and replication.

The viral NP and the RNA polymerase complex (PB1, PB2 and PA) are required for both transcription and replication [28] which happen in the nucleus of infected host cells. Influenza A virus uses its own error-prone RNA polymerase, for this reason frequent mutations occur. Viral replication produces a cRNA, which is a full length copy of the vRNA, and serves as a template for the sythesis of additional vRNAs. After transcription, the mRNAs are transported to the cytoplasm and there they are translated by the host ribosome. The viral proteins NP, PB2, PB1 and PA just synthesized are transported back to the nucleous where they bind to vRNA and form RNPs for new viral particles. The HA, NA and M2 are transported via the Golgi apparatus to the cell membrane and there they aggregate with the RNPs and M1 proteins to form new viral particles that will be released from the cell upon NA

cleavage of host sialic acid receptors [29].

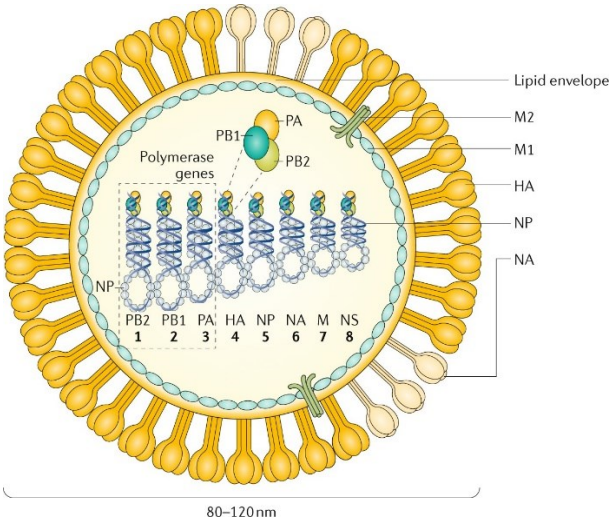


Figure 1. Structure of Influenza A Virus (Source: Krammer, F., Smith, G.J.D., Fouchier, R.A.M. et al. Influenza. Nat Rev Dis Primers 4, 3 (2018). https://doi.org/10.1038/s41572-018-0002-y).

2.1.3 SARS-CoV-2

Coronavirus Study Group (CSG) of the International Committee on Virus Taxonomy, based on phylogenetic studies of related coronaviruses, renamed the virus, initially named novel coronavirus (2019-CoV) by World Health Organization (WHO), as Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This virus is the responsible of Coronavirus Disease 2019 (COVID-19) (Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020), the disease started in December 2019 in Wuhan, Hubei Province, China and it is characterize by dry cough, high body temperature, breathe shortness, and pneumonia [30]. Coronaviridae family is divided into four genera: alfa, beta, gamma, and delta [31]. SARS-CoV-2 belongs to beta coronavirus and COVID-19 is the third zoonotic outbreak of beta-CoVs [32].

2.1.3.1 Biology of SARS-CoV-2

SARS-CoV-2 has round or elliptic and often pleomorphic form, with a diameter of approximately 60–140 nm. It consists of four components: spike (S), nucleocapsid (N), membrane (M), and envelop (E) (Figure 2) [33]. The reason for the global

spread and severity of this disease is the nature prone to mutation of this novel virus [34; 35]. The RNA virus is enveloped, positive-sense, single-stranded with 3′ and 5′ cap structure. The size of its genome is approximately 30 kb in length [36]. The genome sequence of SARS-CoV-2 consists of approximately 29,903 nucleotides, linear ssRNA with 14 open reading frames (Orfs). The proteins encoded by the Spike, Envelop, Membrane, and Nucleocapsid gene are the structural proteins [37; 38], and the proteins encoded by the Orf3a, Orf3b, Orf6, Orf7a, Orf7b, Orf8, Orf9b, Orf9c, and Orf10 in SARS-CoV-2 represent the accessory proteins, which provide a collective advantage in infection and pathogenesis [37; 39; 40]. Spike protein mediates the entry of SARS-CoV-2 into the host cell and this protein is composed of three domains, the N-terminal domain with unit S1 and S2, the cytoplasmic C-terminal domain, and a transmembrane domain [40; 41].

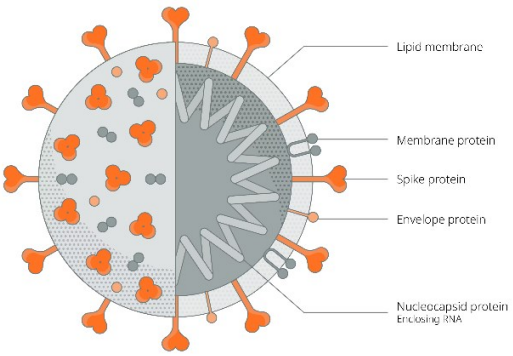


Figure 2. Structure of SARS-CoV-2 (Source: https://www.abcam.com/).

2.1.4 Dobrava-Belgrade virus

Hantaviruses is a group of zoonotic viruses transmitted from animal reservoirs directly to humans, with no involvement of arthropod vectors, with mortality rates ranging from 12 to 40% [42]. These viruses are often found in rodents, bats, and insectivores. In Europe, species like the yellow-necked mouse, black-striped field mouse, and bank vole host specific strains of hantaviruses. Humans get infected through the inhalation of aerosols or dust particles contaminated with virus-

containing rodent excreta. The Dobrava-Belgrade virus (DOBV) was identified in the yellow-necked mouse (*Apodemus flavicollis*) in Italy, specifically within the mountainous region of the Udine province, adjacent to Austria and Slovenia, during an event of enhanced mortality in wild mice and voles. This finding occurred during a period of heightened mortality among wild mice and voles. Hantaviruses are circulating in Italy in rodents and humans since 2000, but this is the first virological confirmation of the infection.

2.1.4.1 Biology of Dobrova-Belgrade virus

Hantaviruses are enveloped viruses within the Hantaviridae family, characterized by segmented, negative-sense, single-stranded RNA. The Orthohantavirus genus, as classified by the International Commission on the Taxonomy of Viruses (ICTV), includes 23 species, with approximately 30 other viruses for which a complete characterization is not available yet (ICTV Hantaviridae). Each species of Hantavirus seems to be specifically associated with its own reservoir host. In Europe, the main reservoir species are the wild yellow-necked mouse (*Apodemus flavicollis*), the black-striped field mouse (*Apodemus agrarius*) and the bank vole (*Myodes glareolus*), maintaining two different strains of the Dobrava-Belgrade virus (DOBV) and Puumala virus (PUUV), respectively. In addition, the grey rat (*Rattus norvegicus*) potentially maintains Seoul virus (SEOV) globally, sporadically reported outside Asia, including a single identification in France [43]. The DOBV has three-segmented genome consisting of the S (small), M (medium) and L (large) segments which encode the nucleocapsid protein (N), a precursor protein giving rise to the two envelope glycoproteins Gn and Gc, and the RNA-dependent RNA polymerase (RdRp), respectively [44].

2.2 DNA Virus

DNA viruses are a diverse group of viruses that contain double-stranded DNA (dsDNA) as their genetic material. They infect a wide range of organisms, including humans, animals, plants, and bacteria. These viruses have evolved various strategies to enter host cells, replicate their genetic material, produce new viral components, and exit the host cell to infect other cells. DNA viruses carry their genetic

information in the form of double-stranded DNA and this DNA can vary in size, structure, and complexity based on the virus type [45]. They have a protective protein coat called a capsid that surrounds their genetic material. Some DNA viruses also possess an outer envelope derived from the host cell membrane. They are classified into different families based on their genomic structure, mode of replication, and other characteristics. DNA viruses attach to specific receptors on the surface of host cells and enter through different mechanisms, such as membrane fusion or endocytosis. Once inside the host cell, the virus releases its genetic material into the cell's nucleus (for viruses that replicate in the nucleus) or cytoplasm (for viruses that replicate there). The viral DNA uses the host cell's machinery to replicate and transcribe new viral components that are synthesized and assembled into complete viruses within the host cell. After assembly, newly formed viruses exit the host cell either by causing cell lysis (destroying the cell) or by budding through the cell membrane, acquiring an envelope in the process. DNA viruses cause a wide spectrum of diseases in humans and other organisms. These diseases can range from mild, self-limiting conditions to severe illnesses. Some DNA viruses have the ability to establish latent infections, where the viral genome remains in the host cell without actively replicating. Periodically, the virus may reactivate, leading to recurrent infections.

2.2.1 Aves Polyomavirus

Polyomaviruses are DNA viruses found in mammals and birds. While in mammals they often do not cause severe illness, in birds they can lead to serious conditions. One specific polyomavirus, APyV, formerly known as budgerigar fledgling disease virus (BFDV), was first identified in budgerigars and caused disease and high mortality in nestlings and young Psittaciformes [46]. The symptoms of APyV infection include feather abnormalities, abdominal distention, and even sudden death [47] and other clinical signs, such as subcutaneous hemorrhage, weakness, anorexia, and apathy, are also found in birds with this disease [48]. Different species of this virus have been found in various bird species, like geese, canaries, and finches [49]. Microscopy for inclusion bodies or electron microscopy for icosahedral virus particles are the methods to identify this virus or probes from APyV (Aves polyomavirus) [50]. When the family Polyomaviridae was created it

included only a handful of polyomavirus species, whose members had all been discovered by the early 1980s, but by September 2015, more than 1200 complete polyomavirus genome sequences representing roughly 100 genetically and biologically distinct polyomaviruses had been deposited in public databases. In 2016 the Polyomaviridae Study Group of the International Committee on Taxonomy of Viruses et al divided Polyomaviridae in 4 genera, to date there are 6 genera (Alpha-, Beta, Delta-, Epsilon, Gamma- and Zetapolyomavirus) that together include 117 species.

2.2.1.1 Biology of Aves Polyomavirus

The virions of Polyomavirus are non-enveloped double-stranded DNA and are approximately 40–45 nm in diameter. Their icosahedral capsid consists of 72 capsomers arranged in a skewed (T = 7d) lattice configuration, 50 nm in diameter (Figure 3). Each virion carries a single molecule of circular dsDNA, with the genomic size of 4981-5421 bp. The coding capacity of Gammapolyomavirus genomes is of 9 proteins, with predicted sizes ranging from 7 to 88 kDa based on nucleic acid sequences (Figure 4). Upon transcription from one side of the viral origin of DNA replication (ORI), mRNA synthesis encodes the early proteins (Figure 4). These nonstructural proteins, referred to as tumor (T) antigens, disrupt cell cycle regulation and, in certain instances, prompt cellular transformation or tumor development. Virions binding to cellular receptors undergo cellular uptake and transportation to the nucleus, where genome transcription and replication occur. During a productive infection, transcription of the viral genome is divided into an early (replication) and a late (virion assembly/exit) stage. Transcription of the early and late coding regions is controlled by separate promoters through the binding of specific transcription factors and cis-acting elements [51]. Proteins are expressed from the two pre-mRNA by alternative splicing. All genes are transcribed by host RNA pol II. To date, nine avian polyomavirus species have been classified into the Gammapolyomavirus genus of the Polyomaviridae family (Table 1). The natural host of Gammapolyomavirues are birds. The characteristic of this genus is that members infect only birds. Some cause severe illness and even death, but

oncogenicity has not been observed. Pyrrhula Pyrrhula Polyomavirus and Cracticus Polyomavirus belong to the Gammapolyomavirus genus (Table 1).

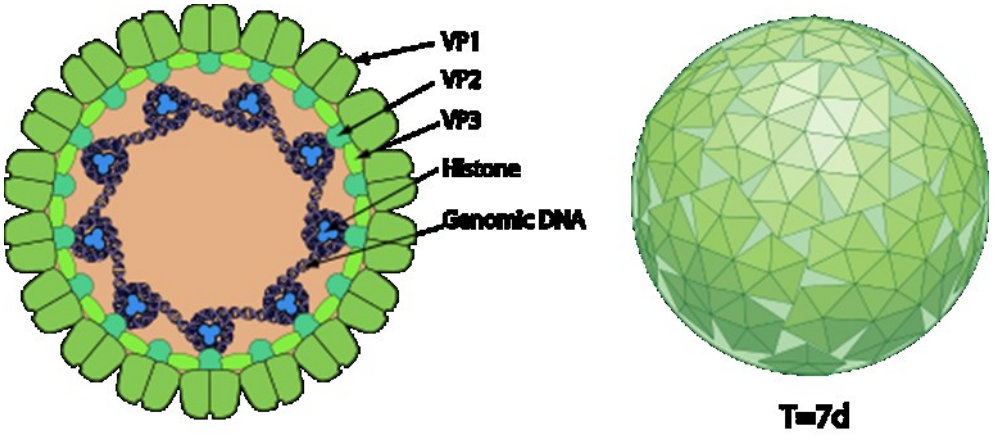


Figure 3. The virion has a non-enveloped capsid with a T=7d icosahedral symmetry, about 50 nm in diameter (Source: https://viralzone.expasy.org/7838).

Table 1. Species in genus Gammapolyomavirus (https://ictv.global/report/chapter/polyomaviridae/polyomaviridae/gammapolyomavirus).

Genus	Species	Virus Name	Isolate	Accesion	Available sequence	Abbrev.
Gammapolyomaviru	*Gammapolyomavirus anseris*	goose hemorrhagic polyomavirus	Germany 2001	AY140894	Complete genome	GHPV
Gammapolyomaviru	*Gammapolyomavirus avis*	budgerigar fledgling disease virus		AF241168	Complete genome	BFDV
Gammapolyomaviru	*Gammapolyomavirus corvi*	crow polyomavirus		DQ192570	Complete genome	CpyV
Gammapolyomaviru	*Gammapolyomavirus cratorquatus*	butcherbird polyomavirus	AWH19840	KF360862	Complete genome	Butcherbird
Gammapolyomaviru	*Gammapolyomavirus egouldiae*	Erythrura gouldiae polyomavirus 1	1209	KT302407	Complete genome	EgouPyV1
Gammapolyomaviru	*Gammapolyomavirus lonmaja*	Hungarian finch polyomavirus	14534/2011	KX756154	Complete genome	HunFPyV
Gammapolyomaviru	*Gammapolyomavirus padeliae*	Adélie penguin polyomavirus	Crozier_2012	KP033140	Complete genome	ADPyV
Gammapolyomaviru	*Gammapolyomavirus pypyrrhula*	finch polyomavirus		DQ192571	Complete genome	FpyV
Gammapolyomaviru	*Gammapolyomavirus secanaria*	canary polyomavirus	Ha09	GU345044	Complete genome	CaPyV

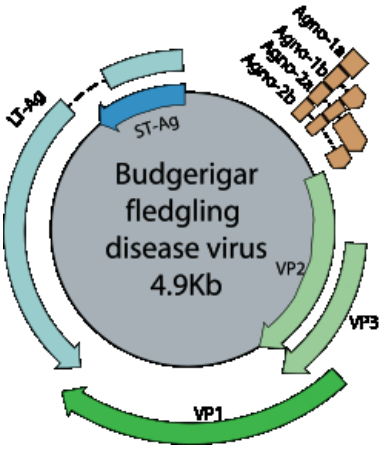


Figure 4. The genome of APyV (Source: https://viralzone.expasy.org/7838).

2.2.2 Circovirus

The Circovirus genus, belonging to the Circoviridae family, finds natural hosts in birds like pigeons, ducks, and pigs. This virus is categorized as a single-stranded DNA virus (ssDNA) and encompasses 49 species, among which there are the Psittacine beak and feather disease virus (BFDV) and canary Circovirus [51; 52]. The primary culprit behind a significant dermatological ailment in various Psittaciformes is the Beak and Feather Disease Virus (BFDV), posing a global threat [53]. Its widespread distribution results from international trade, legal and illegal, involving exotic Psittaciformes [54; 55]. BFDV demonstrates a remarkable ability to adapt among closely related hosts, facilitating its rapid dissemination worldwide [56; 57]. This disease manifests as symmetric feather dystrophy and immunosuppression, affecting numerous species of psittacine and even distant avian orders. Its transmission occurs through droppings, feather dander, close contact, aerosols, and fomites, rendering all Psittaciformes susceptible to infection [48]. Circoviruses have also been associated with a condition called "black spot" in neonatal canaries in Europe [58]. Subsequently, a newly identified circovirus, termed canary circovirus (CaCV), was detected in adult canaries. These birds exhibited a brief illness marked by symptoms such as dullness, anorexia, and feather abnormalities, ultimately resulting in their demise [59; 60; 61].

2.2.2.1 Biology of Circovirus

Circoviruses exhibit small virions, typically measuring around 20–25 nm in diameter. These virions possess a non-enveloped structure, appearing spherical and showcasing T=1 icosahedral symmetry. Comprising 60 capsid subunits, they encapsulate the viral circular single-stranded DNA. Within infected cells in diagnostic samples, mature virions are found either as independent particles or arranged in a linear "strings of pearls" configuration in cell-free specimens. The genome of circoviruses consists of a singular molecule of circular, single-stranded ambisense DNA, spanning approximately 1.7–2.3 kb in size. These viruses stand as the smallest known DNA viruses that can infect mammals. Their compact genome is notably reduced to the absolute essentials, serving only two fundamental functions: the replication and packaging of the viral genetic material [62].

2.2.2.2 Psittacine Circovirus

The sequence of the Beak and Feather Disease Virus (BFDV) comprises seven open reading frames (ORFs), consisting of three within the viral strand and four within two complementary strands, situated within the repetitive sequence [63; 64; 60]. However, among these ORFs, only two, namely ORF C1 and ORF V1, undergo significant structural and genetic alterations [57]. ORF V1, positioned on the viral strand, is thought to encode the replication-associated protein (Rep), crucial for the process of rolling circle propagation (RCR). Conversely, ORF C1, found in the complementary strand, encodes an essential structural element of the viral capsid known as the cap gene. Typically, the functional domain of the rep gene remains relatively unchanged, initiating RCR to generate the single-stranded DNA sequence, thereby facilitating cloning sequences by binding to the host cell. Within the sequence's intergenic region, a theoretical stem-loop structure carrying a conserved nanonucleotide sequence motif (TAGTATTAC) has been pinpointed. This motif represents the binding site for circoviruses to engage with their host DNA polymerase, a crucial interaction facilitating viral replication [57].

2.2.2.3 Canary circovirus

In 2001, Phenix et al. conducted the cloning and sequencing of the circular, single-stranded DNA genome of a newly discovered circovirus in canaries, named Canary Circovirus (CaCV). The analysis of the sequence revealed a genome size of 1952 nucleotides with the potential to encode three viral proteins, including the hypothetical capsid and replication-associated (Rep) proteins. Upon examination, the CaCV genome displayed its closest sequence resemblance (58.3% nucleotide identity) to another recently identified circovirus, the Columbid Circovirus (CoCV). Conversely, it exhibited more distant relationships with two strains of porcine circovirus (PCV1 and PCV2), Beak and Feather Disease Virus (BFDV), and a newly isolated Goose Circovirus (GCV), sharing approximately 46.8–50.9% nucleotide identity with these viruses. Similar to other members within the Circovirus genus, various nucleotide structures and amino acid motifs, believed to play roles in virus replication, were identified on the putative viral strand.

Furthermore, phylogenetic analysis conducted on both the capsid and Rep protein-coding regions provided additional support for the closer relationship between CaCV, CoCV, and BFDV, while showcasing a more distant relation to GCV, PCV1, and PCV2 [59].

2.3 Whole-genome sequencing

Whole-genome sequencing (WGS) is a comprehensive technique used to determine the complete genetic DNA sequence of an organism. Unlike targeted sequencing methods that focus on specific regions, WGS covers the entire genome, providing a holistic view of the genetic blueprint of an organism. It involves the sequencing of both coding (exons) and non-coding (introns, intergenic regions) portions of the genome, enabling a deeper understanding of genetic variations, structural alterations, and functional elements. WGS has been applied in different fields and it has revolutionized clinical diagnostics, enabling the identification of disease-causing mutations, understanding inherited disorders, and guiding personalized treatment strategies. Studying genomes across species aids in understanding evolutionary relationships, speciation events, and genetic adaptations. Over the years, WGS technologies have evolved rapidly, driven by innovations in sequencing platforms, data analysis pipelines, and cost reduction. From first-generation Sanger sequencing to next-generation sequencing (NGS) and the emerging third-generation sequencing technologies, each iteration has increased sequencing throughput, reduced turnaround time, and lowered per-base costs, democratizing access to genomic information.

2.3.1 Sanger technology

Sanger sequencing, developed by Frederick Sanger in the late 1970s, revolutionized DNA sequencing. This method is also called “first generation sequencing” and it is based on the synthesis of a complementary DNA template in the presence of natural 2’-deoxynucleotides (dNTPs) and of modified nucleotides (2’,3’-dideoxynucleotides, ddNTPs) that terminate chain elongation by DNA polymerases [65]. Sanger sequencing involves four separate reactions, each with a small amount of ddNTPs labeled with a fluorescent marker [66]. These reactions create a set of DNA fragments of varying lengths, each ending with a specific ddNTP. The

fragments are separated by size using high-resolution gel electrophoresis, allowing the determination of the DNA sequence based on the order in which the terminated fragments appear, in fact shorter DNA strands migrate through the gel more quickly and exit the capillary first. Longer strands take longer to exit. A laser excites the fluorescent labels, and a detector reads the emission of light, generating a chromatogram that represents the DNA sequence [67]. This technology allowed scientists to sequence DNA with high accuracy and it is highly reproducible and accessible, paving the way for various applications in genomics, molecular biology, and medical diagnostics. However, Sanger sequencing is limited in throughput and cost-effectiveness compared to newer, high-throughput sequencing methods like next-generation sequencing (NGS). NGS techniques have largely replaced Sanger sequencing for large-scale genomic projects due to their ability to generate vast amounts of sequence data rapidly and at a lower cost. Nonetheless, Sanger sequencing remains valuable for targeted sequencing, validation of NGS results, and small-scale applications requiring high accuracy.

2.3.2 Next-Generation Sequencing

Next-Geneation Sequencing (NGS, high-throughput sequencing) refers to a set of advanced technologies that revolutionized DNA and RNA sequencing methodologies and it is known as the “second generation sequencing”. Millions of DNA or RNA fragments are sequenced rapidly and in parallel, allowing researchers to decipher genetic information with unprecedented speed and depth compared to traditional Sanger sequencing. They generate vast amounts of sequencing data in a single run facilitating comprehensive analysis of genomes, transcriptomes, and epigenomes. A key quality of the NGS technologies is versatility, they accommodate various applications, including whole-genome sequencing, exome sequencing, RNA sequencing (RNA-Seq), ChIP sequencing (ChIP-Seq), metagenomics, and more. The sequencing costs per base have been reduced, making large-scale genomic studies more accessible and affordable. NGS enables the detection of rare genetic variants, structural variations, single nucleotide polymorphisms (SNPs), and gene expression profiles with high sensitivity and resolution. NGS data necessitates sophisticated bioinformatics tools and pipelines

for processing, analysis, and interpretation due to the massive volume and complexity of sequencing data generated.

2.3.2.1 Illumina technology

Illumina is a leader in NGS technologies, offering a range of sequencing platforms known for their high-throughput and accuracy. Illumina's hallmark technology, Sequencing by Synthesis (SBS), uses reversible dye terminators to sequentially identify nucleotides as they are incorporated into DNA strands, enabling massively parallel sequencing of millions of DNA fragments on a single flow cell. Illumina platforms, such as the NovaSeq, HiSeq, and NextSeq series, generate vast amounts of sequencing data in a single run, facilitating comprehensive genomic analysis, large-scale projects, and population-scale studies. Illumina systems provide high-quality sequencing data with low error rates, allowing for accurate detection of genetic variations, structural changes, single nucleotide polymorphisms (SNPs), and complex genomic features. The platforms also offer flexibility in read lengths, supporting various applications. Over time, Illumina has reduced sequencing costs per base, making NGS more accessible for researchers, clinicians, and institutions, contributing to the scalability and affordability of genomic studies. Illumina's sequencing technology allows unbiased sequencing without prior knowledge of the complete DNA content in a sample, but at the same time has the flexibility to allow for targeted sequencing. NGS uses DNA polymerase to integrate fluorescently labeled building blocks into a DNA template, identifying nucleotides via fluorophore excitation. However, NGS extends this process across millions of fragments simultaneously, a critical difference from sequencing a single DNA fragment. The workflow involves four main steps:

1. Library Preparation: DNA or cDNA samples are fragmented and ligated with adapters (Figure 5A) for sequencing library preparation. Alternatively, "tagmentation" that is the combination of fragmentation and ligation reaction in a single step, enhances efficiency of the library preparation process (www.illumina.com/documents/products/datasheets/datasheet_nextera_dn

a_sample_prep.pdf). Adapter-ligated fragments are then amplified via PCR and purified.

2. Cluster Generation: The library is loaded into a flow cell where fragments are captured using complementary oligos, generating distinct clusters through amplification, and preparing templates for sequencing (Figure 5B).
3. Sequencing: Illumina SBS technology uses reversible terminators to detect single bases as they integrate into DNA strands (Figure 5C). This minimizes bias and errors [68; 69], providing highly accurate sequencing across repetitive regions and homopolymers.
4. Data Analysis: Sequence reads are aligned to a reference genome (Figure 5D), allowing various analyses.

Paired-end (PE) sequencing marks a significant advancement in NGS technology, involving the sequencing of both ends of DNA fragments within a library and aligning these forward and reverse reads as read pairs. This method offers twice the number of reads compared to single-end sequencing for the same library preparation effort and time. Aligning sequences as read pairs enhances read alignment accuracy and enables the detection of indels, which is not feasible with single-read data [70]. PE sequencing also allows for the analysis of differential read-pair spacing, facilitating the removal of PCR duplicates, a common artifact arising during library preparation. Moreover, it leads to a higher count of SNV (single nucleotide variant) calls post read-pair alignment [70] (www.illumina.com/documents/products/datasheets/datasheet_nextera_dna_sample_prep.pdf). Library preparation for Illumina NGS has evolved significantly. Current protocols, like Nextera® XTDNA Library Preparation, have slashed library prep time to under 90 minutes (www.illumina.com/documents/products/datasheets/datasheet_nextera_xt_dna_sample_prep.pdf). PCR-free and gel-free kits are available, offering superior coverage in traditionally challenging regions like AT/GC-rich segments, promoters, and homopolymeric regions (www.illumina.com/documents/products/datasheets/datasheet_truseq_dna_pcr_free_sample_prep.pdf). Multiplexing has remarkably increased both data output per sequencing run and sample throughput in NGS over time (Figure 6). This method

involves pooling and sequencing numerous libraries simultaneously in a single sequencing run. Each DNA fragment within multiplexed libraries is tagged with unique index sequences during library preparation, enabling the identification and sorting of each read before final data analysis. However, sequencing reads from pooled libraries must undergo computational identification and sorting, a process known as demultiplexing, before final data analysis. Index misassignment, a known issue in NGS technologies since the inception of sample multiplexing, poses a challenge. The misassignment of the index assigned to the libraries can lead to misalignment and inaccuracies in sequencing results [71]. Addressing these challenges is crucial for ensuring the accuracy and reliability of multiplexed NGS data.

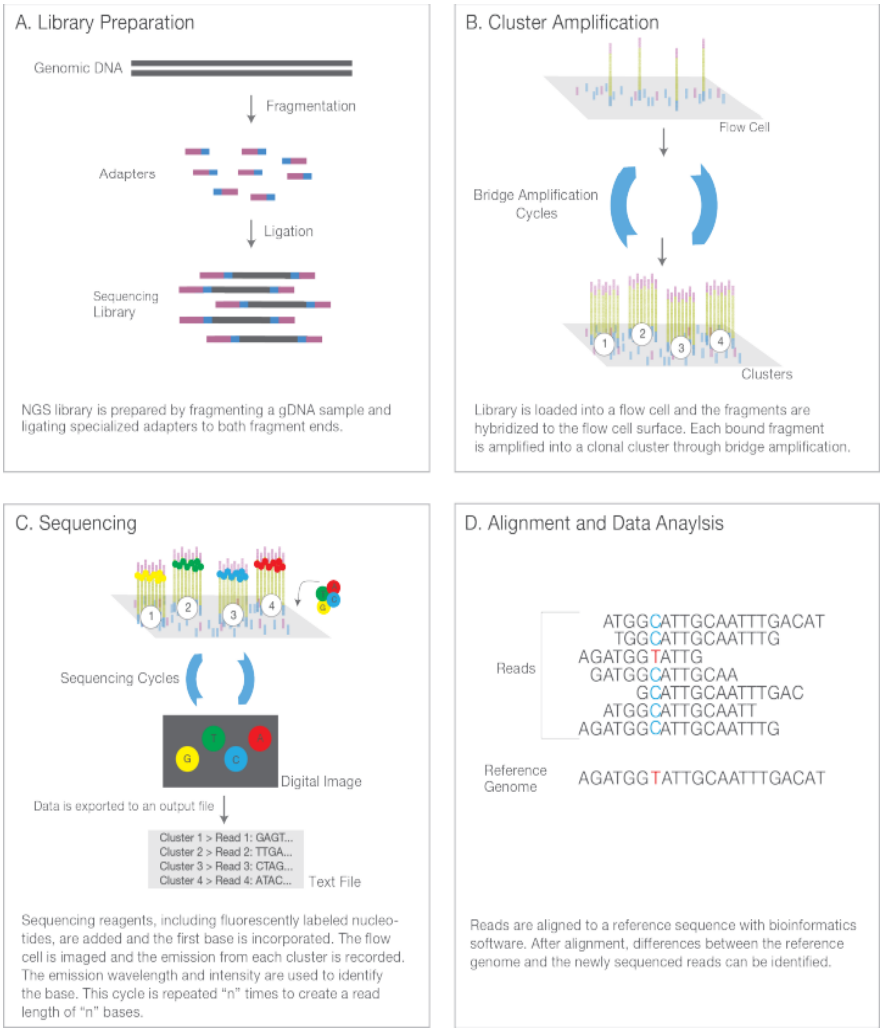


Figure 5. Next-Generation Sequencing Chemistry Overview—Illumina NGS includes four steps: **(A)** library preparation, **(B)** cluster generation, **(C)** sequencing, **(D)** alignment and data analysis. (Source: https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf).

NGS technology not only generates immense data output but also boasts remarkable flexibility and scalability. Sequencing systems cater to various study scales, from small labs to large genome centers. Illumina's NGS sequencers are designed with adaptable run configurations. This flexibility empowers researchers to customize runs according to their specific study requirements, utilizing the instrument best suited for their needs (https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf).

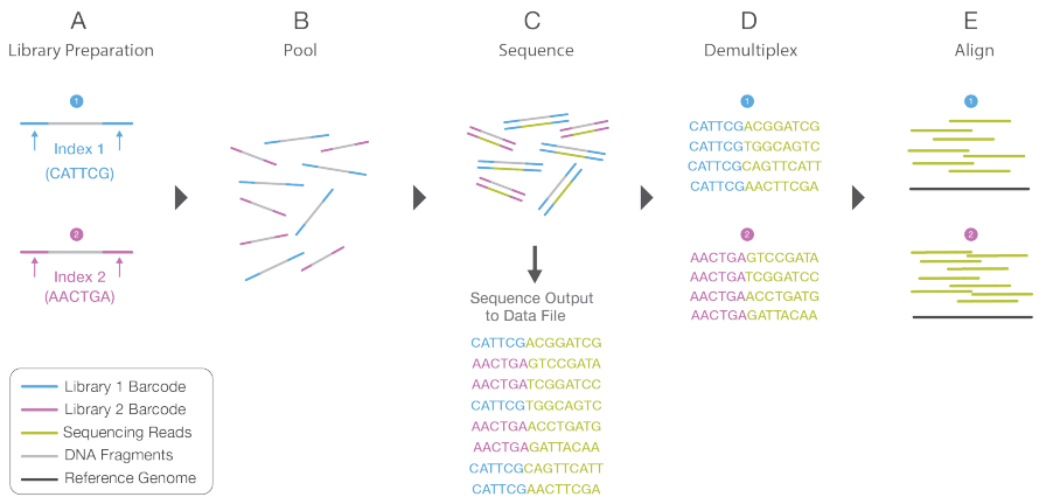


Figure 6. Library Multiplexing Overview. **(A)** Unique index sequences are added to two different libraries during library preparation. **(B)** Libraries are pooled together and loaded into the same flow cell lane. **(C)** Libraries are sequenced together during a single instrument run. All sequences are exported to a single output file. **(D)** A demultiplexing algorithm sorts the reads into different files according to their indexes. **(E)** Each set of reads is aligned to the appropriate reference sequence. (Source: https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf)

2.3.2.1.1 lllumina MiSeq and NextSeq

Illumina MiSeq and NextSeq 550 are next-generation sequencing platforms developed by Illumina and present at Istituto Zooprofilattico Sperimentale delle Venezie in the Viral Genomics and Transcriptomics Laboratory. They have been designed to offer flexibility, scalability, versatility, efficiency and high-throughput sequencing capabilities suitable for various applications in genomics and molecular biology research and to deliver accurate and reliable sequencing data. They are accurate and sensitive, in fact, they provide high-quality sequencing data with low error rates, allowing researchers to sequence multiple samples concurrently and obtain extensive genomic information in a single run. The MiSeq platform generates up to 15-25 million reads per run while NextSeq 550 is a higher throughput platform capable of generating an output from 130 million to 1 billion reads per run. The read length of the two platforms is different: MiSeq offers read length up to 2x300 base pairs for paired-end sequencing while NextSeq 550 supports up to 2x150bp or 2x300bp paired-end reads. MiSeq technology is often used for targeted amplicon sequencing, targeted resequencing, metagenomics and small transcriptomics studies, while NextSeq 550 is suited for a wider range of applications including whole-genome sequencing, RNA sequencing, exome sequencing, ChIP-Seq and larger-scale projects due to its higher throughput. MiSeq is a smaller and more compact benchtop sequencer, occupying less space and requiring less infrastructure, while NextSeq 550 might require more space and infrastructure to accommodate its size and output. The MiSeq is suitable for smaller-scale studies and targeted applications, while the NextSeq 550 is more appropriate for larger-scale and diverse sequencing projects requiring higher throughput.

2.3.2.2 Nanopore technology

Nanopore technology, “third generation sequencing”, is an innovative approach used in DNA and RNA sequencing, offering unique advantages in the field of genomics. It operates by passing DNA or RNA strands through nanoscale pores, measuring changes in electrical current as individual nucleotides traverse these pores. Nanopores are small, nanometer-sized apertures embedded in a membrane.

When a DNA or RNA molecule passes through a nanopore, it causes characteristic disruptions in the ionic current, which are measured and interpreted to determine the DNA or RNA sequence. Nanopore sequencing allows for real-time detection and analysis of the nucleotide sequence. This feature enables rapid data generation and immediate insights into the sequenced molecule. Nanopore sequencing platforms, such as those offered by Oxford Nanopore Technologies, are known for their capability to generate long read lengths, facilitating the study of complex genomic regions, repetitive sequences, and structural variations. Some nanopore sequencing devices, like the MinION from Oxford Nanopore, are portable and pocket-sized, allowing sequencing experiments to be conducted in various settings, including fieldwork or resource-limited environments. Nanopore technology is very versatile and it supports multiple sequencing applications, including whole-genome sequencing, targeted sequencing, metagenomics, RNA sequencing (RNA-Seq), and more. Continuous advancements in nanopore sequencing technology aim to improve accuracy, throughput, and versatility, making it a promising tool for various research applications, clinical diagnostics, and personalized medicine. Nanopore technology continues to evolve and has contributed significantly to genomics research by providing accessible, real-time, and portable sequencing solutions with potential applications across diverse scientific fields.

2.3.2.2.1 MinION

The MinION is a portable, pocket-sized, and USB-powered nanopore sequencing device developed by Oxford Nanopore Technologies. It is part of nanopore sequencing platforms and offers several unique features: it uses nanopore technology and it has compact size and portability, allowing researchers to conduct sequencing experiments in various environments, including remote locations or field settings. It provides rapid sequencing and real-time data analysis, enabling immediate insights into the sequenced DNA or RNA and it generates long reads, aiding in resolving complex genomic regions, detecting structural variations, and facilitating de novo assembly of genomes. The MinION is utilized across various research domains, including genomics, metagenomics, transcriptomics, and pathogen detection due to its versatility and portability. It has contributed to

democratizing sequencing technologies, making them more accessible to smaller laboratories, field researchers, and educational institutions due to its relatively lower cost compared to larger sequencing platforms. The MinION has been instrumental in advancing research by providing portable and real-time sequencing capabilities, allowing for applications in diverse scientific disciplines and enabling researchers to explore biological samples in new and innovative ways.

2.4 Bioinformatics

Bioinformatics plays a pivotal role in sequencing data analysis, providing the computational framework essential for interpreting the vast volumes of genetic information generated by high-throughput sequencing technologies. The diversity and complexity of viral genomes as well as the evolving nature of viral populations pose challenges in sequencing viruses. Viruses exhibit a wide range of genome sizes, structures, and replication strategies that makes it challenging to develop universal sequencing methods that can efficiently capture all viral genomes. They can also evolve rapidly due to high mutation rates and recombination events so the viral populations are constantly changing. Isolating viral genetic material from the complex biological samples containing host DNA and RNA, without contamination, can be significantly difficult. In many infections, viral genetic material may be present at low concentrations, making the detection and sequencing of the viral genome accurately very challenging. Analyzing viral sequence data requires specialized bioinformatics tools and pipelines that need to be developed and maintained to handle the diversity of viral genomes. Additionally, interpreting the massive amounts of data generated from high-throughput sequencing technologies requires sophisticated algorithms and computational resources. In this context, the development of bioinformatics pipelines handles various stages of analysis, from initial quality control and pre-processing to alignment, variant calling, and downstream interpretation. These tools facilitate the identification of genetic variations, such as single nucleotide polymorphisms (SNPs), insertions, deletions, and structural variations, enabling researchers to unravel the genetic basis of diseases, explore evolutionary relationships, and understand complex biological processes. Algorithms like BLAST (Basic Local Alignment Search Tool) are

employed for sequence similarity searches, while tools like Smith-Waterman perform local sequence alignment to identify regions of similarity between DNA, RNA, or protein sequences. De novo assembly algorithms reconstruct genomic sequences from short reads obtained through sequencing, enabling the reconstruction of entire genomes. Most of the platforms have integrated data analysis tools and many bioinformatics softwares have been developed and accessible online. The available tools for data analysis usually have user-friendly interface, making them accessible to users with limited programming skills or bioinformatics background. They are typically hosted on servers so users do not need to install any software or deal with dependencies and they are quick and easy to use since they are designed to be straightforward and require minimal setup time but they present limited flexibility, scalability, data size limitations and privacy and security concerns. The development of in-house pipeline, frequently in Python, R or command-line tools provide greater flexibility to design custom analysis pipeline. The increasing volume of biological data with the advent of high-throughput technologies poses challenges in analysing big data efficiently. The development of in-house pipeline used on a scalable and responsive computational infrastructure and resources as high-performance computing clusters can overcome this limitation and can offer efficient data handling and processing capabilities, allowing large data sets analysis and complex manipulations and computations of data. For data security, sensitive and proprietary data cannot be uploaded to online web tools for analysis. The development of in-house pipeline allows the overcome of this limitation and additionally they can be documented, version-controlled and shared enabling reproducibility and reusability of analysis.

2.5 Metagenomics

Metagenomics is a powerful approach in the field of virology that allows the study of viral communities within complex environmental or clinical samples. Analyzing viruses through metagenomics involves the direct sequencing of genetic material from a sample, without the need for isolation or cultivation of individual viral particles. The samples can be collected from the environment or from clinical sources. The nucleic acids, DNA and/or RNA, are then extracted from the collected

samples. The preparation of metagenomic libraries is performed by randomly fragmenting the extracted nucleic acids and attaching sequencing adapters. This step enables the subsequent high-throughput sequencing of the mixed genetic material using NGS technologies. Specific bioinformatics tools and pipelines are used to analyse the output starting from quality control to taxonomic classification until functional annotation and assembly. This can provide a more complete view of individual viral genomes present in the sample. Metagenomic analyses enables the discovery of novel viruses, understanding viral diversity, their interactions with host organisms, and their potential roles in different environments.

2.6 Molecular epidemiology

2.6.1 Phylogenetic and Evolutionary analysis

The increasing availability of viral genome sequences, the improvement in sequencing methods and the statistical methodologies used to analyse the data have helped to explore the field of evolutionary biology of viral genomes. Phylogenetic analyses involve studying the evolutionary relationships among different species, genes, or populations by examining their shared ancestry and divergence over time. These analyses are represented using diagrams like phylogenetic trees or networks, illustrating the evolutionary connections and patterns of ancestry among the sequences being investigated. Phylogenetic analyses rely on information extracted from genetic material such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA) or protein sequences [72] and they encompass various methods used to study the evolutionary relationships. The computation of the matrix of genetic distances between pairs of aligned nucleotide sequences provides a measure of the similarity between two sequences. The p-distance is the proportion of different sites between two aligned sequences, but this underestimates the d-distance (true genetic distance). There are different models of substitution, that describe changes over evolutionary time, as the simplest one parameter Jukes-Cantor model which assumes that all nucleotides occur in equal proportions and the probabilities of each nucleotide substitutions are equal [73]. A more complex model, GTR model (general time reversible), is used to calculate the likelihood of phylogenetic trees and it is a statistically better fit to the patterns of sequence evolution in this thesis

[74]. It is the most general neutral, independent, finite-sites, time-reversible model possible. The frequency of the four nucleotide bases and the rate of substitution between nucleotide pairs can vary in this model, as the third codon positions in protein coding sequences mutate faster than the first two positions. The Γ-distribution accommodates for varying degree of rate heterogeneity. The various methods used to study the evolutionary relationships are of different common types. Maximum Likelihood (ML) and Bayesian Inference are statistical methods used to estimate the most likely tree topology given the observed data and model assumptions. ML methods outperform other methods in term of accuracy, according to simulation studies [75; 76] but the computational intensity is high with a large number of sequences. Bayesian inference of phylogeny combines the information in the prior and in the data likelihood to create a posterior probability of trees, which is the probability that the tree is correct given the data, the prior and the likelihood model. However, while ML seeks the tree that maximizes the probability of observing the data given that tree, Bayesian analysis seeks the tree that maximizes the probability of the tree given the data and the model of evolution. However, it is usually not possible to calculate the posterior probabilities of all the trees analytically. The Markov chain Monte Carlo (MCMC) method, as implemented by the MrBayes program can be used for this purpose [77]. Distance-Based Methods calculate the evolutionary distance between each pair of operational taxonomic units (OUT), using a model of nucleotide substitution to produce a pairwise distance matrix and then infer the phylogenetic relationship from this matrix. The Neighbour-Joining distance method is the most commonly used. Parsimony Methods, as Maximum Parsimony, identify the tree that requires the fewest evolutionary changes to explain the observed differences among sequences. Parsimony operates by selecting the trees that minimise the number of evolutionary steps, by an algorithm, required to explain the data. Molecular Clock Methods estimate divergence times based on the assumption of a constant or variable rate of evolution across lineages. Phylogeographic Analyses focus on studying the geographical distribution and migration patterns of organisms or pathogens through time. Coalescent Methods are used to model the stochastic process of genetic lineages coalescing backward in time to a common ancestor. Phylogenetic

Networks represent reticulate evolution for scenarios where tree-like structures are insufficient. Phylogenomic Analyses utilize genome-scale data to infer evolutionary relationships, often providing higher resolution compared to single-gene analyses. The preferred method for constructing trees, especially when evolutionary rates vary among lineages, are Maximum Likelihood and Bayesian methods, both of which are employed in the studies within this thesis. Confidence levels for each node in the phylogeny are statistically evaluated using the bootstrapping method [78]. This method assesses the reliability of a phylogenetic structure by determining the proportion of "pseudoreplicate" datasets supporting each node. These datasets are created by randomly sampling the original character matrix (alignment columns) to generate new matrices of the same size as the original. These proportions, presented as percentages, serve as a measure for the reliability of individual branches in the optimal tree, with percentages greater than 70 considered statistically significant. The foundation of phylogenetic analysis relies on identifying discrete clades of genetically related viral isolates that group together on a tree and share a common ancestor. Clades, also known as genetic clusters, are separated topologically by long branches and are considered statistically robust when bootstrap values exceed 70 [78]. Bayesian statistical methods for phylogeography is the approach used to study the geographical spread and evolution of pathogens and other organisms and it has revolutionized the analysis of spatial and temporal aspects of evolutionary processes. It is based on integrating molecular data, evolutionary models, and geographical information to reconstruct the migration patterns and historical dispersion of species [79].

2.7 Istituto Zooprofilattico Sperimentale delle Venezie

The Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe) is a health institution dedicated to public health objectives, conducting activities related to control and surveillance, scientific research, and specialized services in the fields of animal health, food safety, and zoonotic diseases. Established by national and regional laws as a technical-scientific tool within the National Health Service, it serves the Ministry of Health, Regions and Autonomous Provinces, Local Health Authorities, independent veterinary professionals, livestock sector operators, food

companies, and citizens. The IZSVe operates within the Veneto Region, the Autonomous Region of Friuli Venezia Giulia, and the Autonomous Provinces of Trento and Bolzano. It is part of a network of 10 Zooprophylactic Institutes with similar functions and various areas of expertise distributed throughout Italy. Situated in Legnaro (Padua), IZSVe also encompasses 10 territorial laboratories spread across different provinces within its jurisdiction and it collaborates with major national and international organizations focused on animal health and food safety. It serves as a national and international reference center on specific topics for the Ministry of Health, the World Organisation for Animal Health (OIE), and the Food and Agriculture Organization of the United Nations (FAO). The mission of IZSVe is centered on improving public health and contributing to the socio-cultural and economic advancement of the community through control and research activities in food safety and animal welfare. The primary objectives of IZSVe's institutional mandate include providing specialized diagnostic services for animals, conducting analytical assessments of food products, performing official analyses on livestock and food products, implementing epidemiological surveillance plans, undertaking scientific research in veterinary medicine and food safety, organizing training activities for professionals in the veterinary and food sectors. It engages in scientific communication focused on health risks related to animal contact and food consumption, and it promotes for animal welfare in the productive world, biomedical scientific research, and society in general. The complex structure SCS5-Research and Innovation of the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe) deals with virology in the field of animal health, aiming to create and enhance knowledge and expertise in this area through scientific research, technological innovation, and cooperation at both national and international levels. Specifically, applied and experimental research is used on the etiology, pathogenesis, and prophylaxis of infectious and contagious diseases in animals, as well as diseases at the human-animal interface. Basic and targeted research is applied to advance knowledge in veterinary hygiene and health, following designated programs and also through collaborations with Italian and foreign universities and research institutions. These efforts are undertaken upon request from the State, Regions, Autonomous Provinces, and public or private entities. The

activities carried out by National and International Reference Centers affiliated with DSBIO are particularly in the fields of Avian Influenza/Newcastle Disease (AI/ND), Rabies and Lyssaviruses. This includes international cooperation and training. Technical and scientific collaboration are made with domestic and foreign veterinary institutions. The development and implementation of molecular methods, bioinformatics, and next-generation genomic sequencing within virology focuses on host-pathogen-environment interactions. The Viral Genomics and Transcriptomics Laboratory is engaged in the design and execution of research projects applied to the field of infectious diseases in animals and diseases at the human-animal interface, focusing particularly on investigating the factors influencing the virulence and transmissibility of pathogenic microorganisms. It also envolved in the development of applied genetics research lines and projects in the field of species and population traceability. It provides sequencing and bioinformatics service to various laboratories within IZSVe and external national and international users. Finally, it is the leader laboratory in the development and implementation of methodologies for viral genomic characterization and transcriptomic analysis.

References

1. Taylor MW. What Is a Virus? Viruses and Man: A History of Interactions. 2014 Jul 22:23–40. doi: 10.1007/978-3-319-07758-1_2. PMCID: PMC7122971.

2. Poltronieri P, Sun B, Mallardo M. RNA Viruses: RNA Roles in Pathogenesis, Coreplication and Viral Load. Curr Genomics. 2015 Oct;16(5):327-35. doi: 10.2174/1389202916666150707160613. PMID: 27047253; PMCID: PMC4763971.

3. Payne S. Introduction to RNA Viruses. Viruses. 2017:97–105. doi: 10.1016/B978-0-12-803109-4.00010-6. Epub 2017 Sep 1. PMCID: PMC7173417.

4. Pérez-Losada M, Arenas M, Galán JC, Palero F, González-Candelas F. Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. Infect Genet Evol. 2015 Mar;30:296-307. doi: 10.1016/j.meegid.2014.12.022. Epub 2014 Dec 23. PMID: 25541518; PMCID: PMC7106159.

5. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol. 2012

Nov 27;12:181. doi: 10.1186/1471-2288-12-181. PMID: 23185978; PMCID: PMC3552766.

6. Forrest HL, Webster RG. Perspectives on influenza evolution and the role of research. Anim Health Res Rev. 2010 Jun;11(1):3-18. doi: 10.1017/S1466252310000071. PMID: 20591210.

7. Lamb R A, Krug R A. Orthomyxoviridae: the viruses and their replication. In: Fields B N, Knipe D M, Howley P M, editors. Fields virology. third ed. Philadelphia, Pa: Lippincott-Raven Publishers; 1996. pp. 1353–1395.

8. Plotch S.J., Bouloy M. , R.M. Krug. Transfer of 5′ terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitroProc. Nat. Acad. Sci. USA, 76 (1979), pp. 1618-1622

9. Zamarin D., Ortigoza M.B., Palese P. Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice J. Virol., 80 (2006), pp. 7976-7983

10. Cheung TK, Poon LL. Biology of influenza a virus. Ann N Y Acad Sci. 2007 Apr;1102:1-25. doi: 10.1196/annals.1408.001. PMID: 17470908.

11. Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, Xiao YL, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL, Dalton RM, Lo A, Efstathiou S, Atkins JF, Firth AE, Taubenberger JK, Digard P. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. Science. 2012 Jul 13;337(6091):199-204. doi: 10.1126/science.1222213. Epub 2012 Jun 28. PMID: 22745253; PMCID: PMC3552242.

12. Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem. 2000;69:531-69. doi: 10.1146/annurev.biochem.69.1.531. PMID: 10966468.

13. Wiley DC, Skehel JJ. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. Annu Rev Biochem. 1987;56:365-94. doi: 10.1146/annurev.bi.56.070187.002053. PMID: 3304138.

14. Weis W, Brown JH, Cusack S, Paulson JC, Skehel JJ, Wiley DC. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature. 1988 Jun 2;333(6172):426-31. doi: 10.1038/333426a0. PMID: 3374584.

15. Shapiro GI, Krug RM. Influenza virus RNA replication in vitro: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. J Virol. 1988

Jul;62(7):2285-90. doi: 10.1128/JVI.62.7.2285-2290.1988. PMID: 2453679; PMCID: PMC253375.

16. Palese P, Schulman JL, Bodo G, Meindl P. Inhibition of influenza and parainfluenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA). Virology. 1974 Jun;59(2):490-8. doi: 10.1016/0042-6822(74)90458-9. PMID: 4364826.

17. Lamb RA, Lai CJ, Choppin PW. Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proc Natl Acad Sci U S A. 1981 Jul;78(7):4170-4. doi: 10.1073/pnas.78.7.4170. PMID: 6945577; PMCID: PMC319750.

18. Martin K, Helenius A. Transport of incoming influenza virus nucleocapsids into the nucleus. J Virol. 1991 Jan;65(1):232-44. doi: 10.1128/JVI.65.1.232-244.1991. PMID: 1985199; PMCID: PMC240510.

19. Pinto LH, Holsinger LJ, Lamb RA. Influenza virus M2 protein has ion channel activity. Cell. 1992 May 1;69(3):517-28. doi: 10.1016/0092-8674(92)90452-i. PMID: 1374685.

20. Kochs G, Koerner I, Thiel L, Kothlow S, Kaspers B, Ruggli N, Summerfield A, Pavlovic J, Stech J, Staeheli P. Properties of H7N7 influenza A virus strain SC35M lacking interferon antagonist NS1 in mice and chickens. J Gen Virol. 2007 May;88(Pt 5):1403-1409. doi: 10.1099/vir.0.82764-0. PMID: 17412966.

21. O'Neill RE, Talon J, Palese P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. EMBO J. 1998 Jan 2;17(1):288-96. doi: 10.1093/emboj/17.1.288. PMID: 9427762; PMCID: PMC1170379.

22. Robb NC, Smith M, Vreede FT, Fodor E (2009) NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. J Gen Virol 90: 1398–14

23. Kates M, Allison AC, Tyrell DA, James AT. Origin of lipids in influenza virus. Cold Spring Harb Symp Quant Biol. 1962;27:293-301. doi: 10.1101/sqb.1962.027.001.027. PMID: 13958354.

24. Carroll SM, Higa HH, Paulson JC. Different cell-surface receptor determinants of antigenically similar influenza virus hemagglutinins. J Biol Chem. 1981 Aug 25;256(16):8357-63. PMID: 6167577.

25. Tumpey TM, Maines TR, Van Hoeven N, Glaser L, Solórzano A, Pappas C, Cox NJ, Swayne DE, Palese P, Katz JM, García-Sastre A. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. Science. 2007 Feb 2;315(5812):655-9. doi: 10.1126/science.1136212. PMID: 17272724.

26. Skehel JJ, Bayley PM, Brown EB, Martin SR, Waterfield MD, White JM, Wilson IA, Wiley DC. Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc Natl Acad Sci U S A. 1982 Feb;79(4):968-72. doi: 10.1073/pnas.79.4.968. PMID: 6951181; PMCID: PMC345880.

27. Bullough PA, Hughson FM, Skehel JJ, Wiley DC. Structure of influenza haemagglutinin at the pH of membrane fusion. Nature. 1994 Sep 1;371(6492):37-43. doi: 10.1038/371037a0. PMID: 8072525.

28. Herz C, Stavnezer E, Krug R, Gurney T., Jr Influenza virus, an RNA virus, synthesizes its messenger RNA in the nucleus of infected cells. Cell. 1981;26:391–400

29. Baigent SJ, McCauley JW. Influenza type A in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. Bioessays. 2003 Jul;25(7):657-71. doi: 10.1002/bies.10303. PMID: 12815721.

30. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020 Feb 15;395(10223):507-513. doi: 10.1016/S0140-6736(20)30211-7. Epub 2020 Jan 30. PMID: 32007143; PMCID: PMC7135076.

31. Woo PC, Lau SK, Lam CS, Lau CC, Tsang AK, Lau JH, Bai R, Teng JL, Tsang CC, Wang M, Zheng BJ, Chan KH, Yuen KY. Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J Virol. 2012 Apr;86(7):3995-4008. doi: 10.1128/JVI.06540-11. Epub 2012 Jan 25. PMID: 22278237; PMCID: PMC3302495.

32. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, Lu G, Qiao C, Hu Y, Yuen KY, Wang Q, Zhou H, Yan J, Qi J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell. 2020 May 14;181(4):894-904.e9. doi: 10.1016/j.cell.2020.03.045. Epub 2020 Apr 9. PMID: 32275855; PMCID: PMC7144619.

33. Cascella M., Rajnik M., Cuomo A., Dulebohn S.C., Di Napoli R. StatPearls StatPearls Publishing; 2020. Features, evaluation and treatment coronavirus (COVID-19)

34. Khailany RA, Safdar M, Ozaslan M. Genomic characterization of a novel SARS-CoV-2. Gene Rep. 2020;19:100682. doi:10.1016/j.genrep.2020.100682

35. Ray D, Salvatore M, Bhattacharyya R, Wang L, Du J, Mohammed S, Purkayastha S, Halder A, Rix A, Barker D, Kleinsasser M, Zhou Y, Bose D, Song P, Banerjee M, Baladandayuthapani V, Ghosh P, Mukherjee B. Predictions, role of interventions and effects of a historic national lockdown in India's response to the COVID-19 pandemic: data science call to arms. Harv Data Sci Rev. 2020;2020(Suppl 1):10.1162/99608f92.60e08ed5. doi: 10.1162/99608f92.60e08ed5. Epub 2020 Jun 9. PMID: 32607504; PMCID: PMC7326342.

36. Kumar M., Kuroda K., Patel A.K., Patel N., Bhattacharya P., Joshi M., Joshi C.G. Decay of SARS-CoV-2 RNA along the wastewater treatment outfitted with Upflow Anaerobic Sludge Blanket (UASB) system evaluated through two sample concentration techniques. Sci. Total Environ. 2020;754:142329. doi: 10.1016/j.scitotenv.2020.142329.

37. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, O'Meara MJ, Rezelj VV, Guo JZ, Swaney DL, Tummino TA, Hüttenhain R, Kaake RM, Richards AL, Tutuncuoglu B, Foussard H, Batra J, Haas K, Modak M, Kim M, Haas P, Polacco BJ, Braberg H, Fabius JM, Eckhardt M, Soucheray M, Bennett MJ, Cakir M, McGregor MJ, Li Q, Meyer B, Roesch F, Vallet T, Mac Kain A, Miorin L, Moreno E, Naing ZZC, Zhou Y, Peng S, Shi Y, Zhang Z, Shen W, Kirby IT, Melnyk JE, Chorba JS, Lou K, Dai SA, Barrio-Hernandez I, Memon D, Hernandez-Armenta C, Lyu J, Mathy CJP, Perica T, Pilla KB, Ganesan SJ, Saltzberg DJ, Rakesh R, Liu X, Rosenthal SB, Calviello L, Venkataramanan S, Liboy-Lugo J, Lin Y, Huang XP, Liu Y, Wankowicz SA, Bohn M, Safari M, Ugur FS, Koh C, Savar NS, Tran QD, Shengjuler D, Fletcher SJ, O'Neal MC, Cai Y, Chang JCJ, Broadhurst DJ, Klippsten S, Sharp PP, Wenzell NA, Kuzuoglu-Ozturk D, Wang HY, Trenker R, Young JM, Cavero DA, Hiatt J, Roth TL, Rathore U, Subramanian A, Noack J, Hubert M, Stroud RM, Frankel AD, Rosenberg OS, Verba KA, Agard DA, Ott M, Emerman M, Jura N, von Zastrow M, Verdin E, Ashworth A, Schwartz O, d'Enfert C, Mukherjee S, Jacobson M, Malik HS, Fujimori DG, Ideker T, Craik CS, Floor SN, Fraser JS, Gross JD, Sali A, Roth BL, Ruggero D, Taunton J, Kortemme T, Beltrao P, Vignuzzi M, García-Sastre A, Shokat KM, Shoichet BK, Krogan NJ. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020

Jul;583(7816):459-468. doi: 10.1038/s41586-020-2286-9. Epub 2020 Apr 30. PMID: 32353859; PMCID: PMC7431030.

38. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. Methods Mol Biol. 2015;1282:1-23. doi: 10.1007/978-1-4939-2438-7_1. PMID: 25720466; PMCID: PMC4369385.

39. Li F, Li W, Farzan M, Harrison SC. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science. 2005 Sep 16;309(5742):1864-8. doi: 10.1126/science.1116480. PMID: 16166518.

40. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020 May;581(7807):215-220. doi: 10.1038/s41586-020-2180-5. Epub 2020 Mar 30. PMID: 32225176.

41. Ziebuhr J. Molecular biology of severe acute respiratory syndrome coronavirus. Curr Opin Microbiol. 2004 Aug;7(4):412-9. doi: 10.1016/j.mib.2004.06.007. PMID: 15358261; PMCID: PMC7108451.

42. Leopardi S, Drzewnioková P, Baggieri M, Marchi A, Bucci P, Bregoli M, De Benedictis P, Gobbo F, Bellinati L, Citterio C, Monne I, Pastori A, Zamperin G, Palumbo E, Festa F, Castellan M, Zorzan M, D'Ugo E, Zucca P, Terregino C, Magurano F. Identification of Dobrava-Belgrade Virus in Apodemus flavicollis from North-Eastern Italy during Enhanced Mortality. Viruses. 2022 Jun 7;14(6):1241. doi: 10.3390/v14061241. PMID: 35746712; PMCID: PMC9229784.

43. Heyman P., Plyusnina A., Berny P., Cochez C., Artois M., Zizi M., Pirnay J.P., Plyusnin A. Seoul hantavirus in Europe: First demonstration of the virus genome in wild Rattus norvegicus captures in France. Eur. J. Clin. Microbiol. Infect. Dis. 2004;23:711–717. doi: 10.1007/s10096-004-1196-3.

44. Tsergouli K, Papadopoulou E, Tsioka K, Papa A. Molecular epidemiology of Dobrava-Belgrade virus in Greece. Infect Genet Evol. 2018 Oct;64:9-12. doi: 10.1016/j.meegid.2018.06.007. Epub 2018 Jun 7. PMID: 29885476.

45. Kazlauskas D, Varsani A, Koonin EV, Krupovic M. Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. Nat Commun. 2019 Jul 31;10(1):3425. doi: 10.1038/s41467-019-11433-0. PMID: 31366885; PMCID: PMC6668415.

46. Johne R, Müller H. Avian polymavirus in wild birds: genome analysis of isolates from Falconiformes and Psittaciformes. Arch Virol. 1998;143(8):1501-12. doi: 10.1007/s007050050393. PMID: 9739329.

47. Fehér E, Kaszab E, Bali K, Hoitsy M, Sós E, Bányai K. A novel gammapolyomavirus in a great cormorant (Phalacrocorax carbo). Arch Virol. 2022 Aug;167(8):1721-1724. doi: 10.1007/s00705-022-05478-8. Epub 2022 May 28. PMID: 35633392; PMCID: PMC9234025.

48. Philadelpho NA, Chacón RD, Diaz Forero AJ, Guimarães MB, Astolfi-Ferreira CS, Piantino Ferreira AJ. Detection of aves polyomavirus 1 (APyV) and beak and feather disease virus (BFDV) in exotic and native Brazilian Psittaciformes. Braz J Microbiol. 2022 Sep;53(3):1665-1673. doi: 10.1007/s42770-022-00785-3. Epub 2022 Jun 29. PMID: 35767215; PMCID: PMC9433492.

49. Circella E, Caroli A, Marino M, Legretto M, Pugliese N, Bozzo G, Cocciolo G, Dibari D, Camarda A. Polyomavirus Infection in Gouldian Finches (Erythrura gouldiae) and Other Pet Birds of the Family Estrildidae. J Comp Pathol. 2017 May;156(4):436-439. doi: 10.1016/j.jcpa.2017.01.006. Epub 2017 Feb 21. PMID: 28238305.

50. Rinder, M., Schmitz, A., Peschel, A. et al. Identification and genetic characterization of polyomaviruses in estrildid and fringillid finches. Arch Virol 163, 895–909 (2018). https://doi.org/10.1007/s00705-017-3688-3

51. King A.M.Q. Family – Circoviridae. In: Fauquet C.M., Mayo M.A., Maniloff J., Desselberger U., Ball L.A., editors. Virus Taxonomy. Elsevier; San Diego: 2012. pp. 343–349

52. Chaves Hernández AJ. Poultry and Avian Diseases. Encyclopedia of Agriculture and Food Systems. 2014:504–20. doi: 10.1016/B978-0-444-52512-3.00183-2. Epub 2014 Aug 21. PMCID: PMC7152037.

53. Todd D. Circoviruses: immunosuppressive threats to avian species: a review. Avian Pathol. (2000) 29:373–94. doi: 10.1080/030794500750047126

54. Fogell DJ, Martin RO, Groombridge JJ. Beak and feather disease virus in wild and captive parrots: an analysis of geographic and taxonomic distribution and methodological trends. Arch Virol. (2016) 161:2059–74. doi: 10.1007/s00705-016-2871-2

55. Fogell DJ, Martin RO, Bunbury N, Lawson B, Sells J, McKeand AM, et al. Trade and conservation implications of new beak and feather disease virus detection in native and introduced parrots. Conserv Biol. (2018) 32:1325–35. doi: 10.1111/cobi.13214

56. Raidal SR, Peters A. Psittacine beak and feather disease: ecology and implications for conservation. Emu-Austral Ornithol. (2018) 118:80–93. doi: 10.1080/01584197.2017.1387029

57. Dolatyabi S, Peighambari SM, Razmyar J. Molecular detection and analysis of beak and feather disease viruses in Iran. Front Vet Sci. 2022 Dec 1;9:1053886. doi: 10.3389/fvets.2022.1053886. PMID: 36532332; PMCID: PMC9751380.

58. Goldsmith, T. L. (1995). Documentation of passerine circoviral infection. In Proceedings of the annual conference of the Association of Avian Veterinarians (pp. 349-350) Philadelphia PA, USA.

59. Phenix, K.V., Weston, J.H., Ypelaar, I., Lavazza, A., Todd, D., Wilcox, G.E. & Raidal, S.R. (2001) Nucleotide sequence analysis of a novel circovirus of canaries and its relationship to other members of the genus Circovirus of the family Circoviridae. Journal General Virology, 82, 2805-2809.

60. Todd D, Scott ANJ, Fringuelli E, Shivraprasad HL, Gavier-Widen D, Smyth JA. Molecular characterization of novel circoviruses from finch and gull. Avian Pathol. (2007) 36:75–81. doi: 10.1080/03079450601113654

61. Todd Daniel. Molecular characterisation of novel circoviruses from finch and gull. Avian Pathology, 2007, 36 (01), pp.75-81. ff10.1080/03079450601113654ff. ffhal-00540066f

62. Fenner's Veterinary Virology; Fifth Edition • 2016 DOI https://doi.org/10.1016/C2013-0-06921-6

63. Trinkaus K, Wenisch S, Leiser R, Gravendyck M, Kaleta EF. Psittacine beak and feather disease infected cells show a pattern of apoptosis in psittacine skin. Avian Pathol. (1998) 27:555–61. doi: 10.1080/03079459808419383

64. Crowther RA, Berriman JA, Curran WL, Allan GM, Todd D. Comparison of the structures of three circoviruses: chicken anemia virus, porcine circovirus type 2, and beak and feather disease virus. J Virol. (2003) 77:13036–41. doi: 10.1128/JVI.77.24.13036-13041.2003

65. 111 Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, et al. Nucleotide sequence of bacteriophage phi X174 DNA. Nature. 1977;265: 687–695. doi:10.1038/265687a0

66. 112 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977;74: 5463–5467. Available: https://www.ncbi.nlm.nih.gov/pubmed/271968

67. Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. third Edition, Vol. 1, Cold Spring Harbor Laboratory Press, New York.

68. Ross MG, Russ C, Costello M, et al. Characterizing and measuring bias in sequence data. Genome Biol. 2013;14(5):R51.

69. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nat. 2008;456(7218):53–59.

70. Nakazato T, Ohta T, Bono H. Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive. PLoS One. 2013;8(10):e77910.

71. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic AcidsRes. 2012:2513–2524.

72. Moret, B.M.E., Warnow, T.: Reconstructing optimal phylogenetic trees : a challenge in experimental algorithmics. In: Experimental Algorithmics, LNCS, pp. 163–180 (2002)

73. Jukes, T.H. and Cantor, C.R. (1969) Evolution of Protein Molecules. In: Munro, H.N., Ed., Mammalian Protein Metabolism, Academic Press, New York, 21-132.http://dx.doi.org/10.1016/B978-1-4832-3211-9.50009-7

74. Yang Z. Estimating the pattern of nucleotide substitution. J Mol Evol. 1994 Jul;39(1):105-11. doi: 10.1007/BF00178256. PMID: 8064867.

75. Kuhner MK, Felsenstein J. 1995. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol Biol Evol May;12(3):525.

76. Huelsenbeck JP 1995. Performance of phylogenetic methods in simulation. Syst Biol 44: 17-48.

77. Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

78. Felsenstein J 1985. Confidence limits on phylogenies: an approach using bootstrap. Evolution 39: 783-791.

79. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS Comput Biol. 2009 Sep;5(9):e1000520. doi: 10.1371/journal.pcbi.1000520. Epub 2009 Sep 25. PMID: 19779555; PMCID: PMC2740835.

3 Results

Due to the multiple emergencies that the Viral Genomics and Transcriptomics Laboratory has had to face in recent years, we have developed protocols for analyzing various viruses using different sequencing methodologies. Starting from a new protocol for the analysis of SARS-CoV-2, including a comparison of the two sequencing platforms, Illumina and MinION, to the publication of the case of SARS-CoV-2 infection in a cat in the North of Italy. The development of various pipelines also extends to DNA viruses, with the identification of a possible new genotype of Circovirus using a metagenomic approach. Research also encompasses the identification of Dobrava-Belgrade orthohantavirus. As the National Reference Centre for Avian Influenza (AI), numerous research projects are conducted in our laboratory in this area. These projects range from the phylogenetic analysis of the HPAI A(H5) epidemic waves from 2020 to 2023 in Italy to the transmission of avian influenza A(H5N1) to mammals, specifically to cats, in Poland during summer 2023. In this section all the studies and the pubblications out of these studies are reported.

3.1 SARS-CoV-2

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is a novel betacoronavirus identified as the causative agent of the COVID-19 pandemic. Its genome, a single-stranded positive-sense RNA of approximately 30 kilobases in length, encodes several structural proteins, non-structural proteins, and accessory proteins. Among these, the spike (S) protein, responsible for receptor binding and viral entry into host cells, particularly through interaction with the angiotensin-converting enzyme 2 (ACE2) receptor, plays a pivotal role in infectivity [1]. The genetic composition of SARS-CoV-2 exhibits inherent variability, leading to the emergence of different viral lineages characterized by specific mutations across its genome. These genetic variations have implications for viral transmissibility, disease severity, and the efficacy of diagnostics, therapeutics, and vaccines. Continuous genomic surveillance has been crucial in tracking the evolution and

spread of different SARS-CoV-2 variants, aiding in understanding their impact on disease dynamics and public health interventions [2].

3.1.1 Long amplicon-based protocol for SARS-CoV-2: Illumina VS ONT

Abstract

In response to the urgent need for rapid identification and characterization of SARS-CoV-2 during the COVID-19 pandemic, innovative, cost-effective, and high-throughput sequencing techniques were developed and applied. The different sequencing methods available are unable to accurately identify large indels, large structural variants or to determine haplotypes. Here we developed techniques that encompassed both targeted and untargeted sequencing approaches utilizing second (Illumina) and third (Oxford Nanopore Technologies) generation sequencing platforms for virus characterization starting from different diagnostic matrices (swab and saliva) also in samples with low viral load (up to 30ct*), allowing comprehensive identification and characterization of SARS-CoV-2 in a one-health framework. Overall, this study highlights the efficacy and robustness of the developed targeted sequencing protocols for SARS-CoV-2 characterization, offering insights into viral population dynamics and evolution in clinical settings.

*ct = viral load obtained by Real-time RT-PCR for E gene using oligonucleotides developed by Corman et al. [11].

Introduction

In January 2020 the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was identified as the causative agent of COVID-19 (Coronavirus Disease 2019). SARS-CoV-2 is an enveloped, single stranded positive-sense RNA virus of the family Coronaviridae, genus Betacoronavirus [3]. Its poly-adenylated genome of about 30-kb consists of two open reading frames (ORFs) for viral nonstructural proteins (Nsps) and nine small ORFs that encode structural proteins, including surface (S), envelope (E), membrane (M), nucleocapsid (N) proteins, and accessory genes [4]. Timely whole genome sequencing of SARS-CoV-2 has been proven as a powerful tool for early detection of introductions of known variants and for

characterization of emerging variants, as demonstrated by the United Kingdom sequencing programme (https://www.gov.uk/government/publications/SARS-cov-2-variant-of-concern-diagnostic-assurance), which enables the identification of the first Variant of Concern (VOC). Moreover, genomes analyses can support contact tracing and can be used to identify mutations which can i) affect the performance of diagnostic assays, ii) increase virus virulence or transmissibility, or iii) change the virus antigenic properties. With the beginning of the vaccination campaign in several countries around the world, early detection of possible vaccine escape mutants has assumed a high priority. World Health Organization designed specific criteria for SARS-CoV-2 sample selection for which sequencing should be prioritized, among which: i) samples from re-infected individuals or from infected individuals vaccinated for SARS-CoV-2; ii) samples from immunocompromised patients with prolonged shedding, especially when receiving antibody therapy against SARS-CoV-2; iii) samples collected from clusters of infection in order to investigate transmission events or to evaluate the efficacy of infection control procedures [5]. Given the relatively low evolutionary rate of SARS-CoV-2, which is currently estimated to be 1×10^{-3} nucleotide substitutions per site per year [6], in all these scenario characterization of minority variants and/or of quasispecie haplotypes can help reconstructing transmission dynamics of highly related viruses and promptly identifying potential viral escape mutants. At the time of the research, several methods were available for sequencing of SARS-CoV-2 from clinical samples, but the COVID-19 global pandemic has requested the scientific community for the rapid virus identification and early characterization. The development and application of rapid, low costs and high throughput innovative targeted and untargeted sequencing techniques for Illumina and ONT platforms were necessary for the analysis in a one-health prospective. Amplicon-based methods is a cost-effective way to achieve complete genome characterization. Among them, the ARTIC protocol (https://artic.network/ncov-2019) and commercial kits based on short amplicons are the most commonly used (ECDC, 2021); however, these methods are not adequate for accurate detection of large indels, large structural variants or to determine haplotypes. Moreover, coverage may be not homogeneous as a consequence of differences in primer efficiency or

presence of variants in the primer annealing regions, which can affect an accurate analysis of minority variants (ECDC, 2021) [7]. Here we describe a long amplicon-based protocol for SARS-CoV-2 sequencing on second (Illumina) and third (Oxford Nanopore) generation sequencing platforms and evaluate its accuracy and analytical validity on dilutions of cultured isolates and clinical samples. We showed that this protocol can be applied to detect intra-individual single nucleotide variations. Illumina technology produces very high quality, high accuracy, but short (up to 300bp and 150bp, respectively) reads with high costs. It takes days to obtain the results and the sample preparation is complex [8]. The MinION sequencer from Oxford Nanopore is an evolving technology that produces long-read sequencing data with low equipment cost. It belongs to the third generation platforms category; it has relatively simple sample preparation, flexible run times and portability, even if it has lower accuracy than second generation platforms [9].

Methodology

222 SARS-CoV-2 genomes available in GISAID Epicov repository on 8th March 2020 were downloaded and aligned for primer design. Primal scheme software (https://primalscheme.com/) was used to design a set of 22 primers (Table 1) to cover the complete coding regions (29260 nucleotides). This set was built with the aim of obtaining 11 amplicons (the first 10 of about 3 kb and last one of about 1.6 kb) with overlapping edges of about 200 bp (Figure 1). SuperScript™ IV VILO™ Master Mix with ezDNase™ Enzyme (ThermoFisher) was used to remove contaminating genomic DNA (ezDNAse activity) and to perform retrotranscription (SSIV VILO enzyme activity) following the manufacturer instructions. Target amplification was obtained using Platinum™ SuperFi™ Green PCR Master Mix and the specific primers pairs used in singleplex. In detail, the optimized PCR protocol was set as follows: 5 µM primers, 10 µl 2X Plat SuperFi Green Mastermix, 1.5 µl of cDNA in a final volume of 20 µl; thermal profile was: 98°C for 2 min followed by 40 cycles of 98°C for 15 s, 65°C for 30 s, and 72°C for 2 min, with a final extension at 72°C for 5 min. 1 % Agarose gel electrophoresis (Figure 3) was set up to confirm the presence of all the 11 PCR products. In addition to a targeted approach, an untargeted specific method SISPA (Sequence-independent, single

primer amplification) has been tested, that contemplate the use of tailed random primer and tailed oligodT [10]. Two distinct sequencing platforms have been evaluated, namely the Illumina MiSeq platform and the Nanopore sequencing device MinION (ONT). Bioinformatics analysis has been performed for raw data generated by both the platforms, using pipelines developed in house for the project that provide reads filtering and trimming, reads alignment, consensus sequence generation and analysis of variants.

Table 1. Primer sets designed for full genome amplification of SARS-CoV-2 genome. Genome position refers to Wuhan-Hu-1 (NC_045512.2).

Amplicon	Primer Name	Sequence	Genome position
1	Covid19_1F	ACCAACCAACTTTCGATCTCTTGT	31 - 54
	Covid19_2R	AATCAATGCCCAGTGGTGTAAGT	2932 - 2954
2	Covid19_3F	GACACTGTGATAGAAGTGCAAGGT	2747 - 2770
	Covid19_4R	AGTTCATACTGAGCAGGTGGTG	5691 - 5712
3	Covid19_5F	AAGAGTCTTGAACGTGGTGTGT	5500 - 5521
	Covid19_6R	TGACTTTTTGCTACCTGCGCAT	8364 - 8385
4	Covid19_7F	TCGGCAAGGGTTTGTTGATTCA	8176 - 8197
	Covid19_8R	AGCTACAGTGGCAAGAGAAGGT	11188 - 11209
5	Covid19_9F	ACACACCACTGGTTGTTACTCAC	11000 - 11022
	Covid19_10R	AATACCAGCATTTCGCATGGCA	14022 - 14043
6	Covid19_11F	GGCATTTTGATGAAGGTAATTGTGACA	13835 - 13861
	Covid19_12R	GCTCTTGTGGCACTAGTGTAGG	16948 - 16969
7	Covid19_13F	ACCGAAATTATGTCTTTACTGGTTATCGT	16766 - 16794
	Covid19_14R	CAGTGAGTGGTGCACAAATCGT	19960 - 19981
8	Covid19_15F	AATGTAGCATTTGAGCTTTGGGC	19774 - 19796
	Covid19_16R	CCCTGGAGCGATTTGTCTGACT	22780 - 22801
9	Covid19_17F	TAACGCCACCAGATTTGCATCT	22588 - 22609
	Covid19_18R	TGCAGTAGCGCGAACAAAATCT	25470 - 25491
10	Covid19_19F	TGCTGTAGTTGTCTCAAGGGCT	25280 - 25301
	Covid19_20R	ACTGCCAGTTGAATCTGAGGGT	28330 - 28351
11	Covid19_21F	ACCTTTTACAATTAATTGCCAGGAACC	28145 - 28171
	Covid19_21R	TAGGCAGCTCTCCCTAGCATTG	29769 - 29790

SARS-CoV-2 clinical samples

Five clinical samples (2 nasal swabs and 3 saliva) which tested positive by Real-time PCR targeting the E gene [11] were selected for testing the sequencing protocol using both Illumina and Oxford Nanopore technologies. Specifically, the

two nasal swabs have ct of 25.88 and 28.84, while the three saliva have ct of 25.03, 27.11 and 29.54.

Viral clones

Two viral clones of SARS-CoV-2, (labeled as pS and pL), isolated through limiting dilutions and then amplified, have been quantified and genetically characterized. For these two clones the whole genome consensus sequences have been generated by targeted amplification protocol and sequenced on an Illumina MiSeq; SNPs and INDELs identified have been confirmed by Sanger sequencing. The two pS and pL clones differ to each other for 6 consensus nucleotide variants (2 in ORF1ab, 2 in Spike and 2 in ORF3a) and for a deletion of 4 nucleotides in ORF7a of sample pL. Based on their genome, pL and pS clones have been mixed in proportion 25:75 (mix-pS/pL) (Table 2). Furthermore, pS clone and/or mix-pS/pL have been properly diluted in matrix in order to obtain biological samples with medium (~ 26ct) and low (~29ct) viral load (Figure 2). pS clone has been spiked in nasal swab in VTM medium, while the mix of two viral clones were spiked in nasal swab in VTM medium and in saliva.

Illumina and Oxford Nanopore sequencing

For each sample 5 µl of the ten 3 kb amplicons and 2.5 µl of the smallest one (1.6 kb) were combined in a sterile tube and the entire volume of amplicons pool was purified using Agencourt AMPure XP beads (Beckman Coulter™) (beads to sample ratio of 1.8x). Each purified sample amplicons pool was quantified using the Qubit 1X dsDNA HS (High-Sensitivity) Assay Kit and diluted to the concentration required by the sequencing kit used. Illumina sequencing libraries were prepared using Nextera XT DNA Library Preparation kit following the manufacturer's instructions and sequencing reactions was performed using platform Miseq and 250PE mode. Nanopore sequencing was made using Ligation Sequencing Kit and Native Barcoding following the manufacturer's instructions. Sequencing was performed on the MinION device.

Bioinformatic analyses

For MiSeq sequencing data the read quality was assessed by using FastQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Illumina Nextera XT adaptors sequences (Illumina, San Diego, CA, USA) were clipped from reads with scythe v0.991 (https://github.com/vsbuffalo/scythe) and the adaptors were trimmed with cutadapt v2.10 [12] that was used also to filter the raw data by removing reads with Q score below 20 and length below 80 nucleotides. Complete genomes were generated through a reference-based approach using BWA v0.7.12 (https://github.com/lh3/bwa) [13] and the alignments were processed with Picard-tools v2.1.0 (http://picard.sourceforge.net) and GATK v3.5 (https://github.com/moka-guys/gatk_v3.5) [13;14;15]. LoFreq v2.1.2 (https://github.com/CSB5/lofreq) [16] was used to call Single Nucleotide Polymorphisms (SNPs). The consensus sequences were obtained by in-house script with a minimum coverage of 10X and 50% of allele frequency as threshold for the base calling. Pangolin COVID-19 Lineage Assigner application (https://pangolin.cog-uk.io/) [17] has been used to assign virus lineage. In order to detect viral subpopulations, minority variants analysis was conducted. The BAM alignment files were parsed using the DiversiTools program (http://josephhughes.github.io/btctools/) to determine the average base-calling error probability and to identify the frequency of polymorphisms at each site relative to the reference used for the alignment. Only polymorphisms with a frequency above 2% identified in positions with a minimum coverage of 500 were taken in consideration. For Oxford Nanopore sequencing data, the read quality and reads filtering were assessed by FastQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and by NanoFilt [18]. Complete genome sequences were generated through a reference-based approach using Minimap2 [19;20]. For the variant calling step, Medaka tool from Oxford Nanopore Technologies (https://github.com/nanoporetech/medaka) was used. The consensus sequences were obtained using BCFtools consensus [21]. All the samples analyzed by targeted amplification and by SISPA have been sequenced on Illumina platform. Sequencing by MinION device has been performed only on samples amplified by targeted RT-PCR.

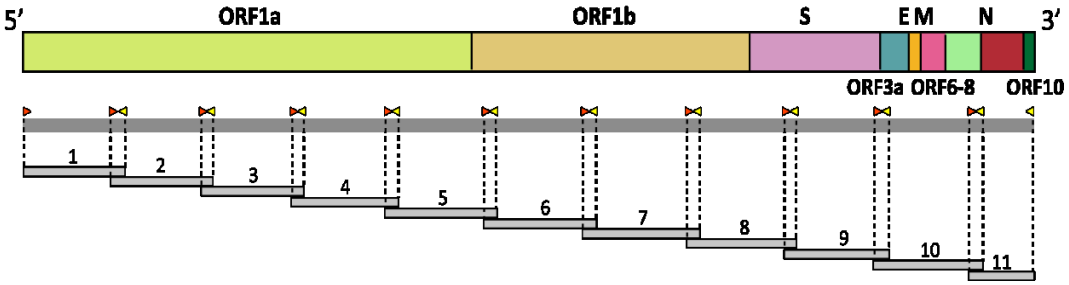


Figure 1. Schematic representation of 11 specific amplicons of around 200 bp in SARS-CoV-2 genome.

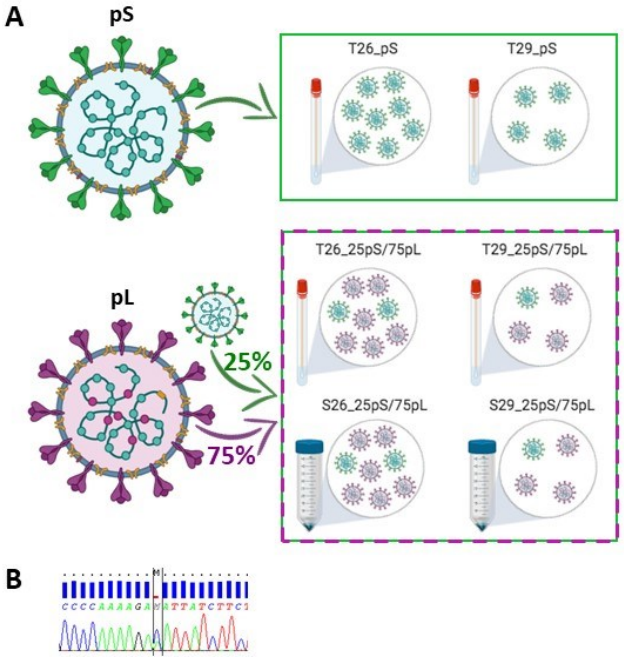


Figure 2. Panel A. pL and pS clones (large and small, respectively) and the mixed 25:75 (mix-pS/pL). pS clone and/or mix-pS/pL have been properly diluted in matrix in order to obtain biological samples with medium (~ 26ct) and low (~29ct) viral load. pS clone has been spiked in nasal swab in VTM medium, while the mix of two viral clones were spiked in nasal swab in VTM medium and in saliva.

Panel B. Sanger sequencing verifies in the mix-pL/pS samples the presence of a double peak in the electropherograms at the positions differing between the two isolates. Schematic representation of 11 specific amplicons of around 200 bp in SARS-CoV-2 genome.

Results

Description of samples and genome amplification

Two purified SARS-CoV-2 isolates by limiting dilutions, namely pS and pL, were used to create the reference samples. For this purpose, the two isolates were

quantified and genetically characterized. The genome of the two strains differs for six nucleotides and a four nucleotides deletion (Table 2).

Table 2. Nucleotide differences between SARS-CoV-2 reference clones pS e pL.
* numbering based on the reference genome Wuhan-Hu-1 (NC_045512.2)
** 4 nucleotides deletion compared to the reference genome Wuhan-Hu-1 (NC_045512.2)

	Position*	pL	pS
ORF1ab	2479	A	C
	11095	A	C
Spike	22205	G	C
	23615	A	C
ORF3a	25470	G	A
	25959	T	C
ORF7a	27694-27697	-**	TTTC

In order to assess the performance of the NGS assay on different type of specimens (nasal swabs and saliva) with different viral loads and to evaluate its accuracy in detecting multiple viral subpopulations, we generated six reference samples (Figure 2A). Specifically, pS was diluted in negative nasal swabs (NS) to obtain biological samples with ct of ~26 (NS26_pS) and ~29 (NS29_pS). In addition, a mix 75:25 of the two isolates pL and pS (mix-pL/pS) were diluted in negative nasal swabs (NS) and saliva (Sa) to obtain four samples with ct of ~26 (NS26_ mix-pL/pS and Sa26_ mix-pL/pS) and ~29 (NS29_ mix-pL/pS and Sa29_ mix-pL/pS). The viral load was confirmed for each sample by quantitative Real-time RT-PCR. Sanger sequencing was used to verify in the mix-pL/pS samples the presence of a double peak in the electropherograms at the positions differing between the two isolates (Figure 2B). Besides the reference samples, a total of 5 clinical samples (2 nasal swabs and 3 saliva) with ct ranging from 25,03 to 29,54 were selected for testing the sequencing protocol. The genome was efficiently amplified for both the reference and clinical samples. Moreover, 459 SARS-CoV-2 clinical samples belonging to different PANGO lineages including all the variants-of-concern (VOCs) B.1.1.7, P.1, P.1.1 and B.617.1 and B.617.2 [22] have been efficiently amplified using this protocol during our genome surveillance activities.

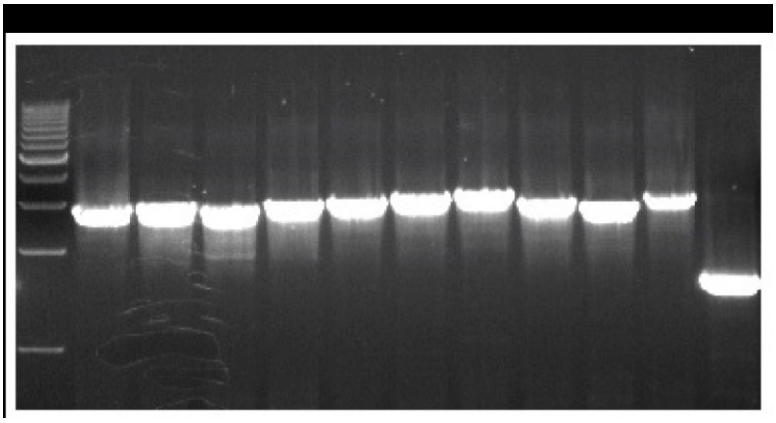


Figure 3. Agarose gel electrophoresis (1%) of the eleven RT-PCRs products for SARS-CoV-2 whole genome amplification. M: 1 Kb Plus DNA Ladder; lanes 1-11: PCR products.

Analysis of reference samples

To assess the accuracy of high throughput sequencing assay we first sequenced reference samples on Illumina MiSeq and ONT MinION.

Illumina sequencing

Following filtering, a total number of paired end reads ranging from 607010 to 856320 for each sample were generated, with more than 99% of the reads aligning to SARS-CoV-2 genome and an average coverage depth ranging from 4581x to 6626x (Figure 4, Table 4). We obtained complete genomes (99.46% of the length of the reference genome Wuhan-Hu-1 and 100% of the cds) with a minimum 10x depth threshold for all the analysed samples.

Table 3. Genome number of reads, sequence length, average coverage for reference samples obtained by Illumina and ONT sequencing.

	Illumina						ONT					
Sample	T26_pS	T29_pS	T26_25pS/75pL	S26_25pS/75pL	T29_25pS/75pL	S29_25pS/75pL	T26_pS	T29_pS	T26_25pS/75pL	S26_25pS/75pL	T29_25pS/75pL	S29_25pS/75pL
# TOT raw reads	857468	796424	729360	698350	608080	827174	268000	272000	268000	280000	228000	280000
# of passing filter reads	856320	795020	727718	697416	607010	825856	3501	3049	2875	8558	2982	4421
Passing filter reads aligned to SARS-CoV-2	852650	790326	722360	694250	603206	821228	2547	2249	2274	2700	2480	2675
% SARS-CoV-2 reads	99,57	99,41	99,26	99,54	99,37	99,44	72,75	73,76	79,09	31,55	83,16	60,51
Sequence length (10x)	29743	29744	29743	29743	29743	29743	29752	29752	29752	29752	29752	29752
% sequenced genome (10x)	99,46	99,47	99,46	99,46	99,46	99,46	99,49	99,49	99,49	99,49	99,49	99,49
Average coverage depth	6626	6021	5561	5345	4581	6294	232	214	201	227	201	212,65
Minimum coverage depth (cds region)	947	469	1522	1773	806	2064	81	96	59	72	59	94
Maximum coverage depth (cds region)	12233	12194	8794	8820	7889	10219	612	563	493	499	493	598

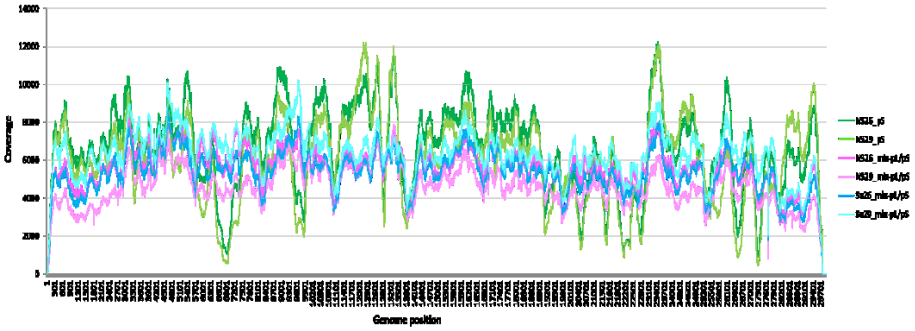


Figure 4. Genome coverage profiles for reference samples obtained by Illumina sequencing.

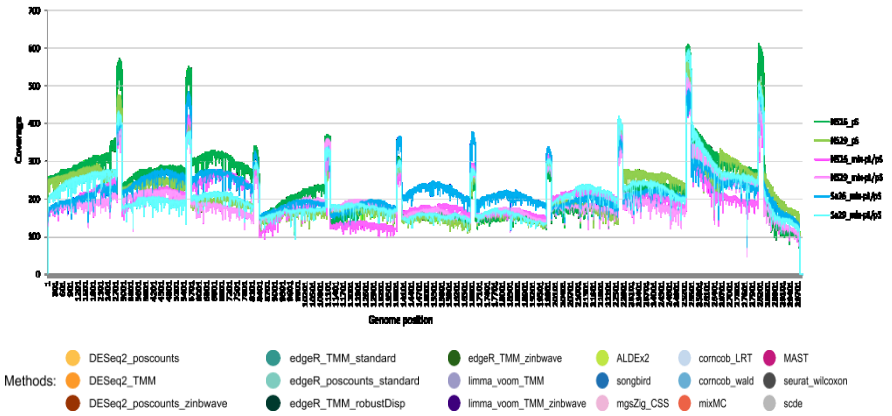


Figure 5. Genome coverage profiles for reference samples obtained by Oxford Nanopore sequencing.

ONT sequencing

The number of filtered reads ranged from 2249 to 2700, with 31.55%-83.17% of the reads aligning to SARS-CoV-2 genome for each sample. We obtained complete genome sequences from all reference samples (99.50% of complete genome length and 100% of cds covered with at least 10x coverage), with an average coverage depth ranging from 201x to 232x (Figure 5, Table 3). The sequences generated by the two sequencing methodologies differ, in the cds, in 1 to 2 positions (9812 and/or 11075) where the sequences generated by ONT sequencing present the deletion of T nucleotide (Table 4). These differences are in homopolymeric regions.

Analysis of clinical samples

Five clinical samples with a different viral load were selected for testing the sequencing protocol on Illumina MiSeq and ONT MinION.

Illumina sequencing

For each sample, a total number of filtered reads ranging from 375214 to 553398 was obtained, of which at least 90% aligned with SARS-CoV-2 genome. The average coverage depth ranged from 2991x to 4338x (Figure 6) and did not correlate with viral load. We obtained complete genome sequences for all samples, with at least 99,46% of complete genome length and 100% of cds sequenced.

Table 4. Nucleotide differences between consensus sequences of SARS-CoV-2 reference and clinical samples generated by Illumina MiSeq and ONT MinION. The colours indicate the nucleotides where the consensus sequences generated by MiSeq and MinION technologies differ (orange for the reference samples where we know that MiSeq is correct and green for clinical samples). The presence of double population and the percentage of presence of the most present nucleotide is represented in grey. In light blue are indicated the positions where isolate pS differs from isolate pL.

Sample	203	2479	3874	7309	9812	10252	11075	11095	13458	22205	23615	25273	25470	25791	25959	27694	27695	27696	27697	28625
hCoV-19/Italy/PD20VIR19351-P4-S/2020	**C**	**C**	**C**	**A**	**T**	**T**	**T**	**C**	**C**	**C**	**C**	**G**	**A**	**C**	**C**	**T**	**T**	**T**	**C**	**G**
hCoV-19/Italy/PD 20VIR19351-P4-L/2020	C	A	C	A	T	T	T	C	C	G	A	G	G	C	T	-	-	-	-	G
T26 pS MiSeq	C	C	C	A	T	T	T	C	C	C	C	G	A	C	C	T	T	T	C	G
T26 pS MinION	C	C	C	A	T	T	T	C	C	C	C	G	A	C	C	T	T	T	C	G
T29 pS MiSeq	C	C	C	A	T*	T	T*	C	C	C	C	G	A	C	C	T	T	T	C	G
T29 pS MinION	C	C	C	A	-*	T	-*	C	C	C	C	G	A	C	C	T	T	T	C	G
T26 25pS/75pL MiSeq	C	A	C	A	T	T	T*	A	C	G(70.4%)	A(75.6%)	G	G(71.3%)	C	T(71.8%)	-	-	-	-	G
T26 25pS/75pL MinION	C	A	C	A	T	T	-*	A	C	G	A	G	G	C	T	-	-	-	-	G
T29 25pS/75pL MiSeq	C	A	C	A	T*	T	-*	T	C	G	A(66.7%)	G	G	C	T	-	-	-	-	G
T29 25pS/75pL MinION	C	A	C	A	-*	T	T	A	C	G	A	G	G	C	T	-	-	-	-	G
S26 25pS/75pL MiSeq	C	A	C	A	T*	T	T	A(74.2%)	C	G(70.4%)	A(69.8%)	G	G(64.7%)	C	T(66.7%)	-	-	-	-	G
S26 25pS/75pL MinION	C	A	C	A	-*	T	N	A	C	G	A	G	G	C	T	-	-	-	-	G
S29 25pS/75pL MiSeq	C	A	C	A	T*	T	T*	A(72.6%)	C	G(72.9%)	A(60.8%)	G	G(68.6%)	C	T(69.7%)	-	-	-	-	G
S29 25pS/75pL MinION	C	A	C	A	-*	T	-*	A	C	G	A	G	G	C	T	-	-	-	-	G
S hCoV-19/Italy/TV20VIR4894-17/2020	C (53.7%)	A	C (61.9%)	A	T	T*	T	C	T (55.7%)	G	C	G	A	C	T	T	T	C	C	G
S hCoV-19/Italy/TV 20VIR4894-17/2020 MinION	T	A	T	A	T	-*	T	C	T	G	C	G	A	C	T	T	T	C	C	G
S hCoV-19/Italy/TV 20VIR4894-20/2020	C	A	C	A	T	T	T*	C	C	G	C	G	A	C	T	T	T	C	C	G
S hCoV-19/Italy/TV 20VIR4894-20/2020 MinION	C	A	C	A	T	T	-*	C	C	G	C	G	A	C	T	T	T	C	C	G
S hCoV-19/Italy/TV 20VIR4894-74/2020	C	A	C	A	T	T*	T	C	C	G	C	G	A	C	T	T	T	C	C	G
S hCoV-19/Italy/TV 20VIR4894-74/2020 MinION	C	A	C	A	T	-*	T	C	C	G	C	G	A	C	T	T	T	C	C	G
T hCoV-19/Italy/VE20COV908-43/2020	C	A	C	T(51.5%)	T	T	T*	C	C	G	C	G	A	T	T	T	T	C	C	G
T hCoV-19/Italy/VE20COV908-43/2020 MinION	C	A	C	T	T	T	-*	C	C	G	C	G	A	T	T	T	T	C	C	G
T hCoV-19/Italy/VE20CO908-5/2020	C	A	C	A	T	T	T*	C	C	G	C	G	A	T	T	T	T	T	C	G
T hCoV-19/Italy/VE20CO908-5/2020 MinION	C	A	C	A	T	T	-*	C	C	G	C	G	A	T	T	T	T	T	C	G

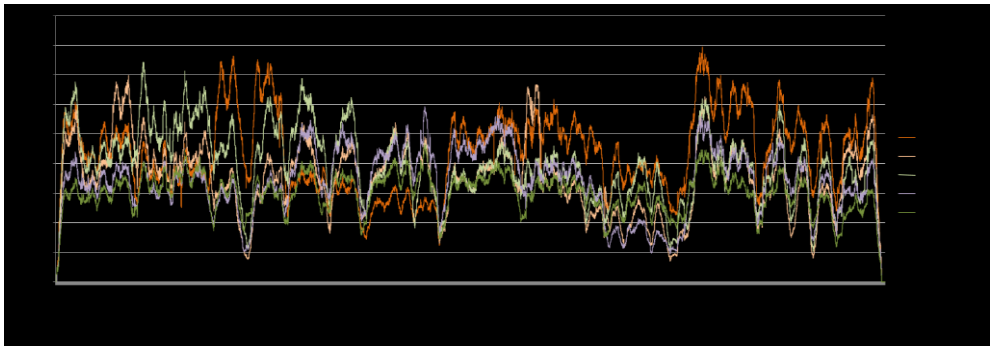


Figure 6. Genome coverage profiles for clinical samples obtained by Illumina sequencing.

ONT sequencing

Following filtering, a total number of reads ranging from 2130 to 3472 was obtained, with 57.46%-84.62% of reads aligning to SARS-CoV-2 genome. The minimum 10x coverage has been reached in at least 99.50% of the complete genome for each clinical sample, with 100% of cds sequenced. The average coverage depth ranged from 152 to 283 (Figure 7).

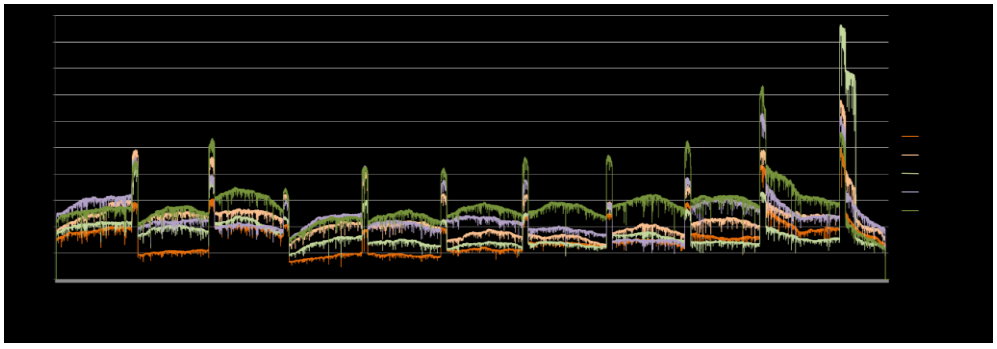


Figure 7. Genome coverage profiles for clinical samples obtained by Oxford Nanopore sequencing.

The sequences generated by the two sequencing methodologies differ, in the cds, in 1 position (10252 or 11075) where the sequences generated by ONT sequencing present the deletion of T nucleotide (Table 4). At consensus level, sample 20VIR4894-17 presents 2 more differences in positions 203 and 3874 where there is a double population (Table 4). These differences are in homopolymeric regions.

Characterization of minority variants

To assess whether this method was able to detect viral subpopulations, we tested it on 75:25 mix-pL/pS reference samples and we looked for the presence of minority variants at genomic positions where the two viral isolates differed. This analysis was performed from Illumina data only, as MinION has a much higher error rate [23]. For mix samples with ct of ~26, minority variants were correctly identified at all expected positions, with allele frequency ranging from 0,23 to 0,30 for the mix diluted in negative nasal swab, and from 0,21 to 0,35 for the one diluted in saliva (Figure 8). By looking at minority variants throughout the genome, we confirmed the accuracy of this protocol for their detection, as no variants were found with frequency higher than 0,06. For mix samples with ct of ~ 29, expected minority variants were identified with a frequency ranging from 0,19 to 0,33 and from 0,2 to 0,39 for nasal swabs and saliva samples, respectively.

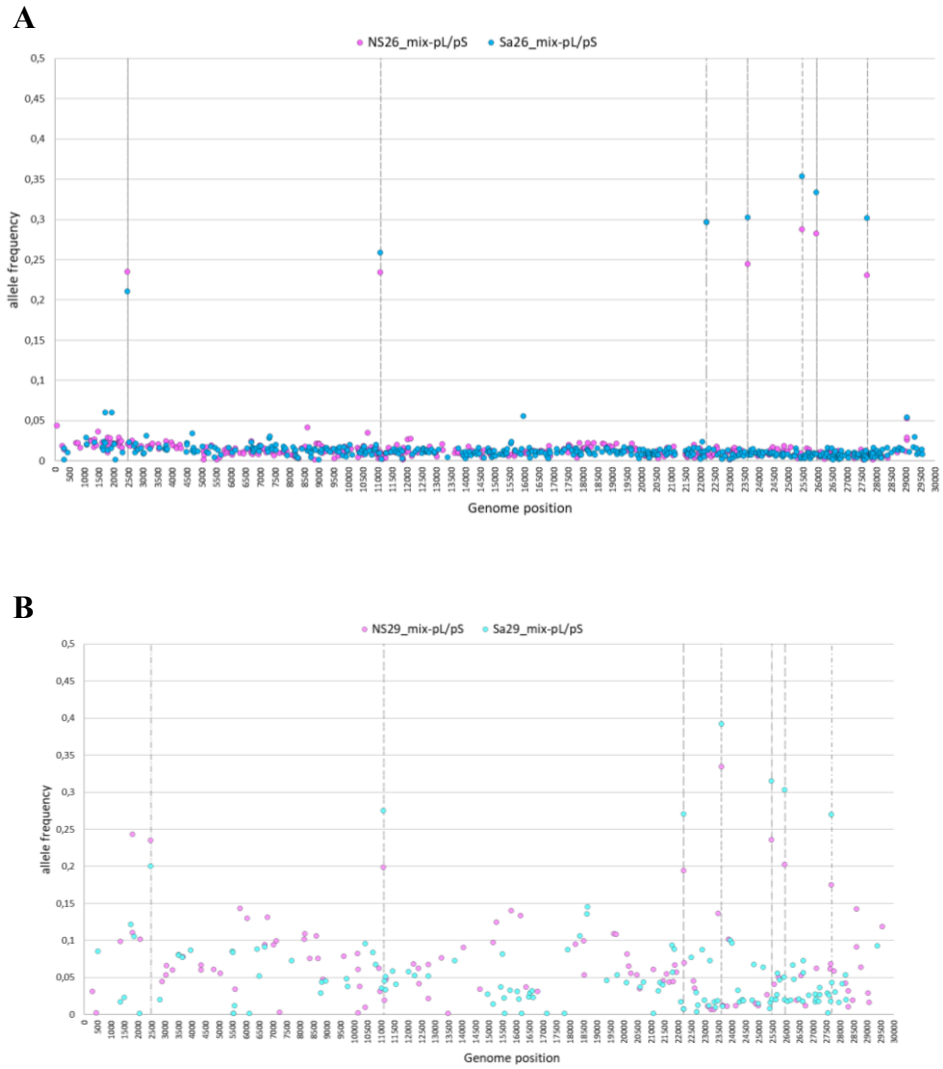


Figure 8. Minority variant allele frequencies of mix samples with ct of ~26 (A) and ~29 (B) from Illumina data. The dashed lines indicate the genomic positions where the two viral isolates used for the mix differ, and thus where the variant frequency is expected around 0,25.

However, other variants were observed throughout the genome, with a frequency of up to 0,24 (NS) and 0,15 (Sa), indicating that identification of minority variants with this protocol in samples with low viral load should be confirmed using multiple replicates. Overall, the results of the validation approach applied on reference materials show higher performance of targeted approach in comparison to SISPA approach: in the first case the percentage of viral reads mapped to the viral genome is higher than 99% while of the reads generated by SISPA approach less than 1% is SARS-Cov-2. For all the samples analyzed by targeted amplification, the protocol

has good performance since whole genome average depth of coverage is high and uniform, from different matrices (swab and saliva) and with samples with a viral load from 18ct to 30ct. This allowed to study the intra-host variability of the viral population and to detect a double population in all the positions where the two clones differ. Differently, with SISPA approach we obtained a depth of coverage of the genome (minimum coverage for each position is 10X) of 67.52%-98.25% for reference samples and of 33.52%-99.09% for clinical samples. For reference samples, with known consensus sequence, the targeted protocol allowed the generation of the correct consensus sequence for all the samples. Differently, 1 to 6 nucleotide differences for each sample, with respect to the expected sequences, have been identified when using the SISPA protocol. Regarding the nanopore sequencing, only the samples generated by targeted amplification have been tested on MinION platform, since the results obtained using the SISPA protocol were not completely satisfying. Specifically, 6 reference samples and 5 clinical samples, previously sequenced by Illumina platform (gold standard), have been sequenced also by third generation platform MinION. As expected, the mapping read length is from 94 to 16884. In order to analyze the data sequenced by MinION device, we adapted a bioinformatics pipeline for viral nanopore sequencing data (https://github.com/artic-network/fieldbioinformatics). This workflow starts with raw data in FASTQ format and includes read filtering, primer trimming, amplicon coverage normalization, mapping with minimap2 [24], variant calling with medaka (https://github.com/nanoporetech/medaka) and consensus building with bcftools (https://samtools.github.io/bcftools/bcftools.html). For every reference and clinical sample, we obtain 100% of consensus sequence with a coverage higher than 100X for the majority of the samples processed, with an average coverage for sample between 201-232 for the reference samples (Figure 5) and 152-283 for clinical samples (Figure 7). Compared to the consensus sequence generated by Illumina sequencing, the sequences generated by MinION platform are correct in 99.9%-100% of the genome positions. The differences (Table 4) are in homopolymeric regions (T nucleotide repeated 5, 6 and 8 times), where MinION technology sometimes identified a nucleotide less. In summary, the protocols developed for the targeted sequencing of SARS-CoV2 allowed the virus characterization starting

from different diagnostic matrices (swab and saliva) tested on both Illumina MiSeq and MinION platform and with samples with variable viral load (18ct- 30ct).

Discussion

In order to test whether the protocol described in this work could offer greater coverage uniformity, we calculated the Inter-Quartile Range (IQR) for four reference samples (NS26_pS, NS29_pS, NS26_mix-pL/pS and NS29_mix-pL/pS) processed with target-based protocol or ARTIC protocol on Illumina MiSeq platform. The IQR is used to measure coverage variability across the genome, meaning that a high IQR indicates low coverage uniformity across the genome, and vice versa (https://emea.illumina.com/). Three out of four samples tested (NS29_pS, NS26_mix-pL/pS and NS29_mix-pL/pS) showed greater uniformity of coverage when sequenced with the target-based protocol compared to sequencing with the ARTIC protocol, confirming that sequencing based on short amplicons often results in high coverage variability. This analysis allowed to confirm pros and cons of MinION technology application. The easy acquisition of the device, the portability and the small size promote the use of this technology in multiple situations. Another crucial benefit of this approach is the rapidity since it allows to obtain results in short time, although the library preparation is not so easy neither quick. As expected, according to the technology properties, comparing the consensus sequences obtained by Illumina sequencing to the ones generated by MinION platform the differences are in homopolymeric regions. The strategic MinION protocol allows to process few samples, as it requires assigning a lower number of reads per sample to obtain the complete genome with good coverage, thanks to its homogeneity. The skills/expertise acquired on these sequencing approaches lead to the application of the in-house target sequencing approach to clinical samples.

References

1. Zhou P, Yang XL, Wang XG, et al. 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.

2. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020 Mar;579(7798):265-269. doi: 10.1038/s41586-020-2008-3.

3. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020 Apr;5(4):536-544. doi: 10.1038/s41564-020-0695-z. Epub 2020 Mar 2. PMID: 32123347; PMCID: PMC7095448.

4. Khailany RA, Safdar M, Ozaslan M. Genomic characterization of a novel SARS-CoV-2. Gene Rep. 2020;19:100682. doi:10.1016/j.genrep.2020.100682

5. WHO. SARS-CoV-2 genomic sequencing for public health goals: Interim guidance, 8 January 2021. Geneva: WHO; 2021. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-genomic_sequencing-2021.1

6. Candido DS, Claro IM, de Jesus JG, Souza WM, Moreira FRR, Dellicour S, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. Science. 2020;369:1255–60. doi: 10.1126/science.abd2161.

7. Xiao, M., Liu, X., Ji, J. et al. Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. Genome Med 12, 57 (2020). https://doi.org/10.1186/s13073-020-00751-4

8. Schadt, E. E., Turner, S., & Kasarskis, A. (2010). A window into third-generation sequencing. https://doi.org/10.1093/hmg/ddq416

9. Jain, M., Olsen, H.E., Paten, B. et al. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 17, 239 (2016). https://doi.org/10.1186/s13059-016-1103-0

10. Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., Afonso, C., Zhang, X., Anderson, N. G., Ghedin, E., & Spiro, D. J. (2008). Viral genome sequencing by random priming methods. https://doi.org/10.1186/1471-2164-9-5

11. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, Mulders DG, Haagmans BL, van der Veer B, van den Brink S, Wijsman L, Goderski G, Romette JL, Ellis J, Zambon M, Peiris M, Goossens H, Reusken C, Koopmans MP, Drosten C. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 2020 Jan;25(3):2000045. doi: 10.2807/1560-7917.ES.2020.25.3.2000045. Erratum in: Euro Surveill. 2020 Apr;25(14): Erratum in:

Euro Surveill. 2020 Jul;25(30): Erratum in: Euro Surveill. 2021 Feb;26(5): PMID: 31992387; PMCID: PMC6988269.

12. MARTIN, Marcel. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, [S.l.], v. 17, n. 1, p. pp. 10-12, may 2011. ISSN 2226-6089. Available at: <https://journal.embnet.org/index.php/embnetjournal/article/view/200>. Date accessed: 02 jan. 2024. doi:https://doi.org/10.14806/ej.17.1.200.

13. Van Der Auwera, G.; Carneiro, M.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinforma. 2013, 43, 10.1–10.33.

14. Depristo, M.A.; Banks, E.; Poplin, R.E.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; Del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 2011, 43, 491–498.

15. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010, 20, 1297–1303.

16. Wilm, A.; Aw, P.P.K.; Bertrand, D.; Yeo, G.H.T.; Ong, S.H.; Wong, C.H.; Khor, C.C.; Petric, R.; Hibberd, M.L.; Nagarajan, N. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 2012, 40, 11189–11201.

17. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol. 2020 Nov;5(11):1403-1407. doi: 10.1038/s41564-020-0770-5. Epub 2020 Jul 15. PMID: 32669681; PMCID: PMC7610519.

18. De Coster W, D'Hert S, Schultz D, Cruts M and Van BroeckhovenC. NanoPack: visualizing and processing long-read sequencing dataBioinformatics, Volume 34, Issue 15, August 2018, Pages 2666–2669, https://doi.org/10.1093/bioinformatics/bty149

19. Heng Li, Minimap2: pairwise alignment for nucleotide sequences, Bioinformatics, Volume 34, Issue 18, September 2018, Pages 3094–3100, https://doi.org/10.1093/bioinformatics/bty191

20. Heng Li, New strategies to improve minimap2 alignment accuracy, Bioinformatics, Volume 37, Issue 23, December 2021, Pages 4572–4574, https://doi.org/10.1093/bioinformatics/btab705

21. Danecek P, Bonfield JK, et al. Twelve years of SAMtools and BCFtools. Gigascience (2021) 10(2):giab008

22. WHO, Weekly epidemiological update on COVID-19 - 20 April 2021 (https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---20-april-2021)

23. Laver T, Harrison J, O'Neill PA, et al. Assessing the performance of the Oxford Nanopore Technologies MinION. Biomol Detect Quantif. 2015;3:1-8. doi:10.1016/j.bdq.2015.02.001

24. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018 Sep 15;34(18):3094-3100. doi: 10.1093/bioinformatics/bty191. PMID: 29750242; PMCID: PMC6137996.Lai, C.C.; Shih, T.P.; Ko, W.C.; Tang, H.J.; Hsueh, P.R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. Int. J. Antimicrob. Agents 2020, 55, 105924.

3.1.2 SARS-Cov-2 Natural Infection in a Symptomatic Cat: Diagnostic, Clinical and Medical Management in a One Health Vision

The work described in this chapter is taken from: Natale A, Mazzotta E, Mason N, Ceglie L, Mion M, Stefani A, Fincato A, Bonfante F, Bortolami A, Monne I, Bellinati L, Guadagno C, Quaranta E, Pastori A, Terregino C. SARS-Cov-2 Natural Infection in a Symptomatic Cat: Diagnostic, Clinical and Medical Management in a One Health Vision. Animals (Basel). 2021 Jun 1;11(6):1640. doi: 10.3390/ani11061640. PMID: 34205893; PMCID: PMC8227534.

Abstract

Despite the reported increase in SARS-CoV-2-infected pets, the description of the clinical features from natural infection and the medical follow up in symptomatic pets is still not sufficiently documented. This study reports the case of an indoor cat that displayed respiratory signs and a gastrointestinal syndrome, following the COVID-19 diagnosis of his owners. Thoracic radiographies were suggestive of bronchial pneumonia, while blood tests were indicative of a mild inflammatory process. Nasal and oropharyngeal swabs tested positive through RT-qPCR assays targeting SARS-CoV-2 genes 14 days after his owners tested positive for the virus. Nasal swabs persisted to be RT-qPCR positive after 31 days. Serology confirmed the presence of antibodies through ELISA, electrochemiluminescence analysis and plaque reduction neutralization test, recording a high antibody titre after 31 days. The cat improved after medical treatment and clinically recovered. This study suggests that exposure to SARS-CoV-2 could lead to a natural infection with bronchial pneumonia in cats along with a possible prolonged persistence of SARS-CoV-2 RNA in the upper airways, albeit at a low level. The cat developed neutralizing antibodies, reaching a high titre after 31 days. Further descriptions of SARS-CoV-2 naturally infected pets, their medical management and diagnostic findings would be useful to enhance knowledge about COVID-19 in susceptible animals.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first reported in Wuhan, Hubei Province, China, in December 2019, and was confirmed to have caused Coronavirus Disease (COVID-19) [1,2]. The current SARS-CoV-2 pandemic possibly originated from an animal reservoir, most likely from bats and/or pangolins [3–8]. Since its first identification, it has been demonstrated that SARS-CoV-2 can naturally and experimentally infect several animal species, including companion animals such as cats and dogs [3,9–13]. Human to animal transmission has been reported in domestic, peri-domestic, wild and zoo animals [3,14–23]. As a matter of fact, the association between humans and animals, including companion animals, livestock and wildlife species, raises concerns about the potential risk of SARS-CoV-2 transmission from COVID-19 human patients to animals ("reverse zoonosis"), and about the potential role that infected animals could play in perpetuating the spread of the disease [17,24–26]. A case of COVID-19 human-to-animal and subsequent animal-to-human transmission has been described in Danish mink workers, although further investigations are needed to define this circumstance [22,27,28]. Evidence of exposure to SARS-CoV-2 in cats and dogs from SARS-CoV-2-infected people have been reported [4,16,29–31]. In pets, clinical findings ranged from asymptomatic to mild respiratory or gastrointestinal symptoms [32]. It has been described that cats naturally or experimentally infected with SARS-CoV-2 are able to transmit the virus to other cats within two days after the contact, and that the shedding of the virus most likely occurs through the respiratory and gastrointestinal tract [17,33–35]. It could be supposed that the virus localization in the respiratory tract may vary during the clinical phase of the disease or may depend on the clinical form, the age and the presence of concomitant systemic conditions [17,35]. Experimentally, the replication of SARS-CoV-2 in the nose and throat and a consequent inflammation pathology deeper in the lower respiratory tract (massive lesions in the nasal and tracheal mucosa epithelia and lungs) was reported in young cats [9]. Differently, a recent study has reported that in sub-adult experimentally infected cats the epithelial cells of the trachea and bronchi seemed non-permissive to SARS-CoV-2 replication, even if the SARS-CoV-2 RNA detection with RT-qPCR throughout the respiratory tract tested

positive, particularly in the upper airways [17]. Furthermore, a recent study has described that SARS-CoV-2 effectively replicated in the upper respiratory tract in cats, and that the virus had cleared from the lungs within 6 days post-infection, even when asymptomatic. Moreover, histopathologic examination demonstrated chronic lung sequelae in cats even a month after viral clearance (histiocytic bronchiolitis with occlusive plugs, peribronchiolar fibrosis and thickening of alveolar septa). In addition, it revealed that after initial infection with SARS-CoV-2, cats were protected from reinfection, with no virus replication in the respiratory organs and no additional lung damage [36]. Recently, Hamer et al. delivered an epidemiological assessment of natural SARS-CoV-2 infections, including virus isolation, among serially tested cats and dogs in households with confirmed human COVID-19 cases in Texas (USA) [31], investigating the presence of SARS-CoV-2 through molecular and serological analyses. No particular clinical symptoms were detected in the dogs and cats enrolled in the study. In serological screenings, the prevalence of anti-SARS-CoV-2 antibodies in cats from Germany, Italy, Croatia, France, and China ranged from 0.69% to 23.5% [16,37–41]. Moreover, in a recent study from Texas, reporting the investigations on 17 cats from COVID-19-affected households [31], it emerged that only 41.2% of the tested animals presented neutralizing antibodies. In most cases, virus-neutralizing antibodies were reported and viral genome sequencing did not reveal any nucleotides coding for the spike protein following human-to-animal transmission [32,40,42–46]. The cat population enrolled for experimental studies is normally represented by cats without underlying health conditions, differently from the pet cat population presented to veterinarians. Human patients with underlying clinical conditions, or immunocompromised humans, were shown to have a higher risk of developing severe clinical disease when infected with SARS-CoV-2 [47], and a previous report from Spain in felines suspected the contribution of comorbidities to the clinical outcome in a cat that was found to be SARS-CoV-2 RT-qPCR positive using nasal swabs while suffering from severe respiratory distress and thrombocytopenia. After the cat was euthanized and a necropsy conducted, it was diagnosed with feline hypertrophic cardiomyopathy, severe pulmonary oedema and thrombosis [48]. Moreover, a case of a cat with symptomatic SARS-CoV-2 infection while suffering

from intestinal B-cell lymphoma was reported in Northern Italy [49]. Although most of the experimentally or naturally SARS-CoV-2-infected cats were reported as being asymptomatic or mildly symptomatic, in this study, we investigated the presumptive SARS-CoV-2 infection in a cat with a mild-to-severe respiratory syndrome and gastrointestinal signs infected by COVID-19-positive owners. Despite the increase in cases reporting about SARS-CoV-2-infected pets, the description of the clinical features after natural infection and medical follow up in symptomatic cases is still not well documented [50]. Providing information such as the clinical presentation, medical management and diagnostic findings would be useful to enhance knowledge about COVID-19 disease in susceptible animals.

Materials and Methods

Sampling

According to our estimation of the "day zero", the owners of a 10-year-old European shorthair neutered male cat, 5.8 kg body weight, were both diagnosed positive and symptomatic for SARS-CoV-2 by the public health service. The cat was the only pet in the household and lived exclusively indoors. He was regularly vaccinated against Feline Calicivirus (FPV), Feline Herpes Virus Type 1 (FHV-1) and Feline Panleucopenia Virus (FPV); on Day 14, he was referred to a veterinary clinic to receive medical consultation. No previous history of respiratory or gastrointestinal illness was reported. The owners revealed that the cat had started to show apathy, anorexia, cough, respiratory distress and vomiting for 7 days before veterinary examination (Day 7). His clinical condition was getting worse. The cat was not taking any medication at the time of the visit, if we exclude the monthly administration of fipronil/(S)methoprene/eprinomectina/praziquantel (Broadline® spot on for cats, Boehringer Ingelheim Vetmedica GmbH, Ingelheim/Rhein Germany) for the prevention of ecto–endo parasite infestation. The cat was subjected to clinical investigation on Days 14 and 31. The veterinarians wore specific personal protection equipment (facial mask, gloves, face shield and gowns) to visit the cat, and followed any other measure recommended by the international guidelines to prevent the risk of infection and spread of COVID-19 disease [51–54]. Two saliva samples were self-collected by the two veterinarians on Day 31, for

SARS-CoV-2 RT-qPCR analysis. A serum sample was also checked for SARS-CoV-2 serological evaluation. The local public health service submitted the owners' nasal swab samples collected on Day 7 to our laboratories to compare the viral RNA genomes.

Time of the Study

Day 0: the SARS-CoV-2 first diagnosis of the owners; first molecular test by public health laboratory services.

Day 7: onset of clinical symptoms in the cat; owners' second molecular test by public health laboratory services.

Day 14: first clinic, radiographic (thoracic radiographies) and laboratory investigations; first samples setting: serum, blood in K3-EDTA, nasal (N), oropharyngeal (OP) and rectal (R) swabs.

Day 31: second clinic, radiographic (thoracic radiographies) and laboratory investigations; second samples setting: serum, blood in K3-EDTA, nasal (N), oropharyngeal (OP) and rectal (R) swabs. Saliva and serum samples collected from the two veterinarians

Molecular Investigation

Nucleic Acid Extraction and Qualitative Real-Time Rt-Pcr Analyses Swabs collected from the cat on Days 14 and 31 were screened by molecular protocols for the presence of viral pathogens: N and OP swabs were tested for FHV-1, FCV and SARS-CoV-2, while R swabs were analysed for FPV [55–58], Feline Coronavirus (FCoV) (VetMAX™ FIP Dual IPC Kit, Laboratoire Service International, Lissieu, France, Applied Biosystems by Thermo Fisher Scientific, Waltham, MA, USA) and SARS-CoV-2. Nucleic acids from the N, OP and R swabs were extracted on the KingFisherTM Flex Purification System (Thermo Fisher Scientific) using the MagMAXTM Pathogen RNA/DNA kit (Applied Biosystems, by Thermo Fisher Scientific), according to the low-cell-content sample suggested by the manufacturer's instructions. RNAs were subjected to the SARS-CoV-2 RT-qPCR

protocol [59], targeting fragments of the E, N and RdRP genes on a CFX 96 Deep Well Real time PCR system (Bio-Rad Laboratories, Inc., Singapore).

A universal heterologous control RNA, referred to as 'Intype IC-RNA' (Indical Bioscience GmbH, Leipzig, Germany), was added to each sample in the extraction step with a ratio of 1:10 of the total elution volume and amplified by using the primers and probes as per Hoffman et al. [60], in order to check the efficiency of the RNA extraction and validate each negative result. Negative and positive controls were included in each run. The Ct values equal or superior to 40.0 were considered negative. Results were generated with Bio-Rad CFX Maestro 1.1 software (Bio-Rad Laboratories). Molecular tests on the saliva samples of the two veterinarians were performed using the same method of the RNA extraction and RT-qPCR protocol [59]. A RT-qPCR ssRNA was performed on an EDTA blood sample to exclude an ongoing Feline Leukemia Virus (FeLV) infection [55–58].

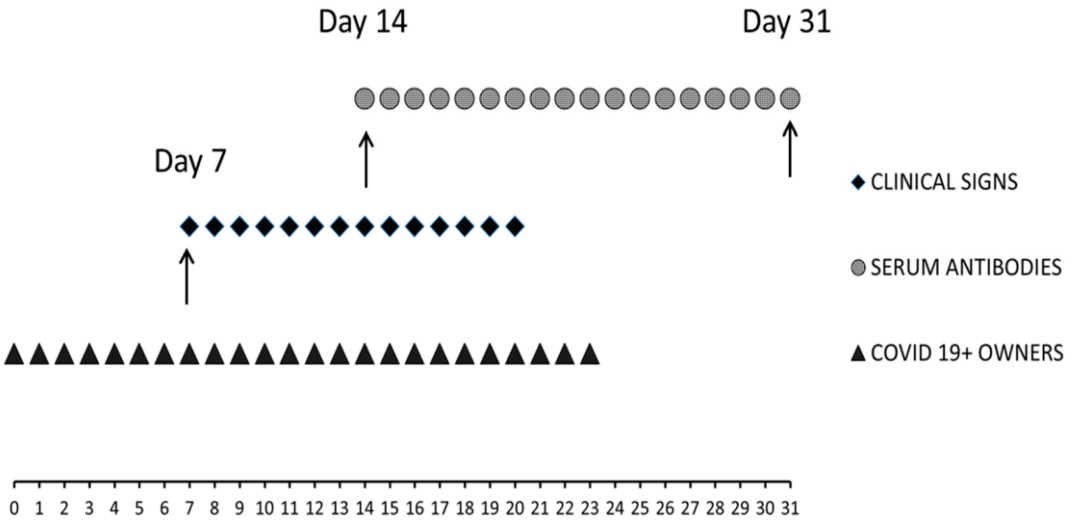


Figure 1. Schematic representation of the clinical case events. The cat started presenting respiratory and gastrointestinal signs on Day 7. On Day 14 occurred the first veterinary examination (sample setting and thoracic radiographies) and the beginning of medical treatment administration. On Day 21, the cat no longer presented clinical symptoms. On Day 31, the second veterinary consult (sample setting and chest radiographies) was carried out. Specific serum antibodies were detected starting from Day 14 (ELISA, ECLIA and PRNT).

The cat's specific serological response against SARS-CoV-2 was investigated at Days 14 and 31 (see Figure 1) by means of two ELISA commercial kits, an electrochemiluminescence immunoassay (ECLIA) and plaque reduction neutralization test (PRNT). Feline Immunodeficiency Virus (FIV) was ruled out performing an immunochromatographic commercial KIT (SNAP FIV/FeLV Combo Test, IDEXX Europe, Hoofddorp, The Netherlands) on the cat's serum. The serological response against SARS-CoV-2 of the two veterinarians following the clinical case was investigated only on Day 31, through an ELISA commercial kit (ID.vet Innovative diagnostics, Grabels, France), ECLIA and PRNT.

Electrochemiluminescence Immunoassay (ECLIA)

For the in vitro quantitative determination of antibodies (including IgG) to the SARS-CoV-2 spike (S) protein, the receptor-binding domain (RBD) in the serum of the Elecsys® Anti-SARS-CoV-2 S immunoassay (Roche Diagnostics International AG, Rotkreuz, Switzerland) on a Cobas e601 analyser was used (Roche Diagnostics International AG, Rotkreuz, Switzerland). The assay is a one-step double antigen sandwich assay. A result of 0.8 U/mL has to be considered as reactive. Analytical performances of the method were evaluated according to the CLSI EP15-A3 protocol [61]. The test has been developed for human testing, but the double-antigen method is species-independent.

Elisa

ID Screen®—SARS-CoV-2 Double Antigen, Grabels, France ID.vet Innovative diagnostics (ELISA KIT 1) detects antibodies against the nucleocapsid (N) protein. Following the manufacturer's instructions, we evaluated the ratio between the optical density (OD) of the sample (S) and OD of the positive control (P), as the SP% value.

The sample is considered negative with SP% 50%, positive with SP% 60% and doubtful when between them. The test is validated for multi-species use, as the double-antigen method is species-independent. The ERADIKITTM COVID-19 multi-species and total Ig, IN3 Diagnostic kit, Torino Italia (ELISA KIT 2), is an indirect ELISA for total IgG anti SARS-CoV-2. Following the manufacturer's

instructions, the sample is considered negative with an SP% < 20% and positive with an SP% ≥ 20%. The test is validated for multi-species use.

Plaque Reduction Neutralization Test (Prnt)

PRNT assays were performed in a Biosafety Level 3 laboratory using a SARS-CoV-2 isolate, as previously described [62]. In brief, serum samples were heat-inactivated (56°C for 30 min) and 2-fold diluted in Dulbecco modified Eagle medium (DMEM). Serum dilutions were mixed with an equal volume (1:1) of a virus solution containing approximately 25 focus-forming units (FFUs) of SARS-CoV-2 and incubated for 1 h at 37°C in a 5% CO2 incubator. Fifty microliters of the virus–serum mixtures were added to the confluent monolayers of Vero E6 cells, in 96-wells plates and incubated for 1 h at 37°C, in a 5% CO2 incubator to allow for the infection of the cells. A total of 100 µL of an overlay solution made of minimum essential medium (MEM) with 2% foetal bovine serum (FBS, Sigma, Saint Louis, MO, USA), penicillin (100 U/mL, Sigma, Saint Louis, MO, US), streptomycin (100 U/mL, Sigma, Saint Louis, MO, US) and 0.8% carboxy methyl cellulose (CMC, Sigma, Saint Louis, MO, USA) were then added to each well after inoculum removal. After 26 h of incubation, the overlay was removed, and the cells were fixed with a 4% paraformaldehyde (PFA) solution. Visualization of plaques was obtained with an immunocytochemical staining method using an anti-dsRNA mouse monoclonal antibody (J2, 1:10,000; Scicons, Sziràk, Hungary) for 1 h, followed by 1 h incubation with peroxidase-labelled goat anti-mouse antibodies (1:1000; Jackson ImmunoResearch Inc., West Grove, PA, USA) and a 7 min incubation with the True Blue (KPL, Gaithersburg, MD, USA) peroxidase substrate. FFUs were counted after acquisition of pictures on a flatbed scanner. The neutralization titre was defined as the reciprocal of the highest dilution resulting in a reduction of the control plaque count >50% (PRNT50).

Sequencing Analysis

Complete genome sequencing was performed on the RNAs extracted from the cat's OP swab sampled on Day 14 and from the nasal swabs of the two owners sampled on Day 7, using an Illumina MiSeq platform (Illumina, San Diego, CA, USA) and

an in-house protocol for target amplification. After trimming and filtering for quality, reads were aligned against the reference genome (GenBank: NC_045512.2) using BWA-mem [63,64]. The sequence was deposited in GISAID under accession number EPI_ISL_962892. The virus lineage was assigned according to the PANGOLIN application, (https://pangolin.cog-uk.io/, Rambaut et al., 2020) (accessed on 23 December 2020) [65,66].

Clinical Procedures

After physical examination, the cat was sedated with medetomidine (Domitor®, ZOETIS Italia S.r.l., Rome, Italy) and propofol (PropoVet®, ZOETIS Italia S.r.l., Rome, Italy) via intramuscular and intravenous injections, respectively, to carry out thoracic radiographies (right lateral and dorso-ventral radiographic projections) and to collect blood (K3EDTA and serum), OP, N and R swabs. The same procedures were carried out on Day 31. All the procedures were performed for diagnostic purposes only with the owners' informed consent.

Haematology and Biochemistry

The first sample of the cat's whole blood in EDTA (Day 14) was analysed by IDEXX VetConnect® PLUS Laboratories at the veterinary clinic; the second one (day 31) was processed at by with the Sysmex XN1000-V analyser (Sysmex Europe GmbH, Norderstedt, Germany). Biochemistry was performed on the serum on Days 14 and 31 (Table 2) through the Cobas c501 clinical chemistry analyser with a related kit (Roche Diagnostics GmbH, Mannheim, Germany).

Results

The clinical exam performed on Day 14 revealed a normal body temperature (38.5°C), no alteration of the oral or conjunctival mucosae membranes, mild retromandibular lymph nodes enlargement, mild dehydration status, normal abdomen examination and no cardio-circulatory alterations. The respiratory tract evaluation revealed a positive tracheal cough and an increase in pulmonary respiratory effort with moderate bronchial and pulmonary sounds. Haematology showed a mild decrease in red blood cells (RBC) and increased reticulocytes. A

mild decrease in platelet count (PLT), aggregation and large platelets were detected at the blood smear examination. Serum biochemical parameters showed a mild increase in serum calcium (Ca) (2.90 mmol/L (2.26–2.73 mmol/L)) with normal total protein and albumin values. Mild hyperglycaemia (Glu) (13.9 mmml/L (3.16–8.88 mmol/L)) was observed and a mild decrease in alkaline phosphatase (ALP) (<5 U/L (6–46 U/L)) and of cholinesterase (1245 U/L (1749–2905 U/L)) were reported. Serum protein electrophoresis showed a mild increase in beta 2 protein fraction (6.1 mmol/L (3–4.7 mmol/L)). An increase in haptoglobin (99 mg/dL (18–74 mg/dL)) was described; conversely, serum amyloid A (SAA) was within the laboratory reference ranges (<0.5 µg/mL (0–9 µg/mL)) (Tables 1–3). The OP, N and R swabs analysed through real-time RT-qPCR assay targeting the SARS-CoV-2 RNA nucleoprotein and envelope protein genes tested positive for the OP and N swabs and negative for the R swab (Table 4). The tests for other possible viral pathogens, such as FCV (N and OP swabs), FHV-1 (N and OP swabs), FCoV and FPV (R swabs), were negative, as well as for the FIV and FeLV assays [55–58]. The cat's serum sample tested negative using ELISA KIT 1 and positive using ELISA KIT 2. The ECLIA showed a positive value of 47.20 U/mL and the PRNT revealed a positive result with a titre of 1:5120 (Table 5). The thoracic radiographies revealed a mild-to-severe bronchial pattern and a diffuse interstitial lung pattern (Figure 2). The clinical diagnosis was bronchial pneumonia potentially due to the SARS-CoV-2 viral infection. Because of the complex respiratory clinical condition, the cat received amoxicillin (Vetrimoxin® Paste, Ceva Salute Animale, Agrate Brianza, Italy) at 10 mg/kg orally twice a day to prevent possible pulmonary bacterial complications and prednisone (Prednicortone®, Dechra, Torino, Italy) 1 mg/Kg orally once a day for 10 days. At the end of the 10 day period, the dose of prednisone was progressively decreased until it was totally suspended on Day 15 of the medical treatment [67]. The clinical status of the cat started to improve on the third day of treatment, and a week later the cat no longer presented any respiratory or gastrointestinal symptoms. Fifteen days after the first examination, the veterinarian evaluated the clinical status of the cat. At this time, both owners tested negative for SARS-CoV-2 by the public health services laboratories. The results of the second set of samples' blood tests (Day 31), thoracic radiographies and N, PO

and R swabs are described in Figure 1. At this time, the owners reported the cat showed no clinical signs. The physical examination was unremarkable. The thorax radiographies showed a normal lung pattern and very low bronchial aspect. The haematology reported a low increase in PLT; aggregation and large platelets were detected at the blood smear examination under optic microscopy. The biochemical analysis reported mild hyperglycaemia (Glu) (11.7 mmml/L (3.16–8.88 mmol/L)) and a low increase in alanine aminotransferase (ALT) (77 U/L (19–71 U/L)) was reported. Serum protein electrophoresis showed a mild increase in the alpha 2 protein fraction (13.4 mmol/L [5.6–10.6 mmol/L]) (Tables 1–3). The RT-qPCR assay performed on the OP, N and R swabs gave a negative result for the OP and R swabs, while the N swab turned out to be slightly positive. The serology for SARS-CoV-2 tested positive by both ELISA kits. Furthermore, a relevant increase in production of SARS-CoV-2 IgG antibodies was reported through the ECLIA assay (1598 U/mL). The PRNT confirmed that the cat had developed neutralizing antibodies against SARS-CoV-2 (1:2560) (Tables 4 and 5). The cat had completely recovered after 15 days (Day 31) and did not show any residual respiratory or gastrointestinal signs. The veterinarians tested negative for SARS-CoV-2. A complete genome sequencing was performed on the RNA extracted from the cat's OP. The virus was assigned to the lineage B.1.177 [65,66], a common lineage in humans in Italy (sequence data available in GISAID as of 14 January 2021). Furthermore, none of the mutations that have occurred to date in the SARS-CoV-2 spike following human-to-animal transmission has been identified. Only the partial genome was obtained from the owners' samples, most probably as a consequence of the low viral titre at the time of the sampling. However, the data generated allowed to cover 93% of the genome, which was sufficient to confirm the clustering of the sequence within the same lineage.

Discussion

Concomitantly with the outbreaks of COVID-19 disease, a relevant number of SARS-CoV-2 naturally infected cats have been reported by the Office International des Epizooties (OIE) in different geographic areas (United States, Latin America, Spain, Switzerland, the Netherlands, Germany, France and China) [68–70]. Several

studies have recently reported natural human-to-pet SARS-CoV-2 transmission in close contact conditions (COVID-19-positive households). The reported clinical features were classified as asymptomatic or mildly symptomatic (lethargy, sneezing); thus, the prevalence of SARS-CoV-2 infection in cats, especially for the asymptomatic cases, may be underestimated [31]. The present study describes a natural human-to-cat SARS-CoV-2 transmission. The cat showed respiratory and gastrointestinal syndromes, even if we may not exclude a mild–low concurrent bacterial infection. This may justify the response to the medical treatment, but it could also represent the natural course of the COVID-19 viral infection. Thoracic radiographic alterations suggestive of bronchial pneumonia were observed. Such abnormalities likely indicate a bronchial inflammatory process in association with pulmonary inflammatory infiltrate or fibrotic tissue. Recently, possible prolonged and persistent pulmonary sequelae in SARS-CoV-2 infected cats have been reported [36]. Even if the pathogenic events and consequences of SARS-CoV-2 in cats have yet to be comprehensively recorded, it has been observed that cats seem to possess a SARS-CoV-2 pulmonary receptor binding model similar to the one in humans. This would explain the susceptibility of felines to SARS-CoV-2 and may describe both the development of the respiratory syndrome and the chest radiographic abnormalities [71,72]. The mild regenerative normocytic normochromic anaemia (mild decrease in RBC) and increased reticulocytes, reported on Day 14, may be consistent with an ongoing inflammatory process [73]. The mild hyperglycaemia detected, similar in both the Day 14 and Day 31 samples, was probably due to the administration of the sedative drugs and to the stressful overall situation [74,75]. The serum electrophoresis increase in the beta 2 protein fraction was probably connected with the clinical inflammatory status and the immunologic response, such as antibodies production, particularly the immunoglobulin complement fraction [76]. The mild decrease in ALP and in cholinesterase was likely due to an impaired liver function consequent to the gastrointestinal symptoms, such as vomit and loss of appetite [77,78]. The increase in haptoglobin reflected an ongoing inflammatory process [76,79], whereas the Feline SAA was possibly within the laboratory reference ranges considering that Feline SAA concentrations increase early during inflammation, usually in

concomitance with other clinical signs (e.g., fever) or increase in the haematological parameter, such as leukocytosis [80]. At the time of the second sampling set (Day 31), a mild increase in PLT value was found, which may be related both to the inflammatory process and, more likely, to the PLT activation and aggregation during the blood sampling (pseudothrombocytopenia) [81,82]. A low increase in ALT and in the serum protein electrophoresis alpha 2 protein fraction would be related both to the medical treatment and the systemic inflammation [77,83]. The cat showed clinical signs 7 days after the owners had been confirmed positive for COVID-19 and were symptomatic themselves (Day 7). In this study, the cat showed a clinical recovery and developed neutralizing antibodies from Day 14, reaching a high antibody titre after 31 days. Such a titre may be considered protective for a reinfection, as previously reported [16,32,35,36,39,41,84]. The RT-qPCR assay performed on OP, N and R swabs resulted positive for the N and OP swabs on Day 14, negative for the OP swab and weak positive for the N swab on Day 31, respectively. The R swabs both tested negative, suggesting a rapid clearance of the virus from the intestinal tract. In experimental studies, cats stop shedding the virus within 10 days [3,9,35], while more recent studies reported that the SARS-CoV-2 experimentally infected cats may have a prolonged period of oral and nasal viral shedding not accompanied by clinical signs, and are capable of direct contact transmission to other cats [35,42]. In agreement with a recent paper [31], we report a prolonged persistence of SARS-CoV-2 infection (N swabs positive in RT-qPCR on Day 31), although the presence of the viral agent is low and presumably not sufficient to infect other susceptible subjects. Although we reported a very low positivity in the last RT-qPCR assay (N swab), it was not possible to perform further diagnostic procedures as the owners denied their consent. As only limited information is available so far on the potential viral shedding routes, it would be beneficial to investigate for SARS-CoV-2 in those cats referred to veterinary clinics for respiratory or gastrointestinal symptoms developed after or concomitantly to their COVID-19-positive owners. Furthermore, SARS-CoV-2 diagnostics in cats living in particular situations and environments, such as dense housing conditions, close contact with elderly people, cat rescue or breeding centres, would be crucial. The recently emerged variants (B.1.1.7. and B.1.351) may have a fitness advantage

associated with mutations in the spike protein, which are suspected to lead to an increase in human-to-human transmissibility and more effective replication [65,85]. Possible changes in the susceptibility of animals in the context of these new variants should be evaluated [86,87]. Fortunately, with reference to this specific case, SARS-CoV-2 genomic mutations which may possibly be involved in the animal-to-human transmission have not been reported so far, although the diagnostic and clinical surveillance, as well teaching how to implement preventions measures, are issues of utmost importance.

Conclusions

The diagnosis of SARS-CoV-2 infection in pets such as cats would be extremely important (i) to provide appropriate veterinary care for the infected animals; (ii) to guarantee adequate protection of veterinary staff and pet owners; and (iii) to apply quarantine measures to prevent transmission between pets, people and potentially susceptible animals. Even though the viral shedding from pets does not appear sufficient to infect other family members or other animals, the usual precautionary measures should urgently be considered as part of the global control efforts and One Health approach. There is currently no evidence that cats play a significant role in human infection and in the spread of the virus to humans. The recently emerged variants (B.1.1.7. and B.1.351) may raise concerns about the possible involvement of susceptible species in new mutations [70] and about the chance of severe clinical signs in animals [87]. Thus, reverse zoonosis is possible if infected owners expose their pets to the virus, particularly during the acute phase of the infection. It is important that pet owners are educated to adopt the precautionary measures to avoid human-to-cat SARS-CoV-2 transmission [88]. Preventing interspecies transfer of an emergent pathogen is essential to decrease the risk of emerging mutations that could affect the transmissibility or effectiveness of the countermeasures, and is also needed to safeguard pet welfare and discourage animal abandonment.

Table 1. Haematology performed from K3EDTA blood samples on Days 14 and 31, respectively. The first sample was analysed by IDEXX VetConnect® PLUS Laboratories, and the second (Day 31) was performed at the SCS3 Laboratory Medicine of the IZSVE with the Sysmex XN1000-V analyser (Sysmex Europe GmbH, Norderstedt, Germany).

Haematology	IDEXX VetConnect® PLUS	Laboratory Medicine (IZSVE)
	Day 14	Day 31
RBC (M/µL)	6.9 (7.1–11.5)	7.69 (5.1–10)
Hgb (g/dL)	9.4 (10.3–16.2)	11.1 (8–15)
Hct (%)	32.6 (28.2–52.7)	31.4 (30.0–45.0)
MCV (fL)	47.2 (39–56)	40.8 (39–55)
MCH (pg)	13.6 (12.6–16.5 pg)	14.4 (13.0–17.0)
MCHC (g/dL)	28.8 (28.5–37.8)	35.4 (30.0–36.0)
Reticulocytes (K/µL)	8.3 *	-
RDW (%)	-	17.3 (16.4–21.7)
PLT (K/µL)	133(155–641)	430 (100–400)
MPV (fL)	-	13.3 (8.1–15.4)
WBC (K/µL)	7.1 (3.9–19)	7.24 K/µL (5–19)
NEUT (K/µL)	4.52 (2.62–15.17)	3.79 (1.80–14.80)
LYMPH (K/µL)	1.98 (0.85–5.85)	3.29 (1.10–8.60)
MONO (K/µL)	0.28 (0.04–0.53)	0.09 (0.05–0.80)
EO (K/µL)	0.28 (0.09–2.18)	0.06 (0.05–2.30)
BASO (K/µL)	0 (0.09–2.18)	0.01 (0.00–0.80)

* Reticulocytes/uL < 50,000 is normal in case of a non-anaemic patient; <50,000 is inadequate in case of an anaemic patient; 50,000–75,000 is mild regeneration; 75,000–175,000 is moderate regeneration; and >175,000 is marked regeneration.

Table 2. Biochemistry performed on Days 14 and 31, respectively. Analyses were performed on serum through the Cobas c501 clinical chemistry analyser with a related kit (Roche Diagnostics International AG, Rotkreuz, Switzerland) at the SCS3 Laboratory Medicine of the IZSVE.

BIOCHEMISTRY	Day 14	Day 31	Reference Values
Haptoglobin [1]	99	18	18–74 mg/dL
Serum Amyloid A [2]	<5.0	<5.0	0–9 µg/mL
Total Proteins	68 g/L	72 g/L	62–80 g/L
Albumin	38 g/L	40 g/L	30–47 g/L
Globuline	30 g/L	32 g/L	22–47 g/L
Ratio A/G	1.28	1.40	1.07–1.87
Urea Nitrogen	10.0	7.3	4.8–12.6 mmol/L
Creatinine	150	122	66–178 µmol/L
Glucose	13.9	11.7	3.2–8.9 mmol/L
Cholesterol	4.37	4.79	1.35–6.09 mmol/L
Triglycerides	0.58	4.21	0–2.48 mmol/L
Total Bilirubine	<2.5	<2.5	0–8.55 µmol/L
Direct Birubine	<1.5	<1.5	0–2.56 µmol/L
Unconj Bilirubine	0	0	0–6.5 µmol/L
AST	15	29	0–61 U/L
ALT	27	77	19–71 U/L
ALP	<5	17	6–46 U/L
GGT	<3	<3	1–5 U/L
Cholinesterase	1245	1781	1749–2905 U/L
CK	11	141	0–305 U/L
Calcium	2.90	2.38	2.26–2.73 mmol/L
Phosphorus	1.43	1.03	0.94–1.98 mmol/L
Magnesium	0.92	0.88	0.79–1.07 mmol/L
Sodium	152	151	141–168 mmol/L
Potassium	4.37	4.29	3.55–5.15 mmol/L
Chlorine	114	112	103–126 mmol/L
Iron	74 µg/dL	82	68–215 µg/dL
Uibc	168 µg/dL	*	105–205 µg/dL
Tibc	242 µg/dL	*	222–423 µg/dL
Saturated Transferrin	30.6%	*	20–56%

* Sample insufficient. Analysis not performed. AST (serum glutamic oxaloacetic transaminase), ALT (serum glutamic pyruvic transaminase), ALP (alkaline phosphatase), ALT (gamma glutamil transferase), CK (creatin kinase). [1] Haptoglobin was measured with a Tridelta PHASE Haptoglobin Assay (Tridelta Development Limited, Maynooth, County Kildare, Ireland) on a Cobas c501 analyser. [2] Serum amyloid A was measured with a multispecies VET-SAA kit (Eiken Chemical Co Ltd., Tokyo, Japan) on a Cobas c501 analyser.

Table 3. Serum protein electrophoresis (Minicap, Sebia Italia S.r.l., Firenze, Italy) on Days 14 and 31, respectively. A mild increase in alpha 2 globulin and beta 1 globulin on Day 14 was reported, consistent with the inflammatory response and the antibodies production.

Serum Protein Electrophoresis	Day 14	Day 31	Reference Values
Albumin (%)	56.2	58.4	52.4–66.2
Alpha 1 (%)	1.3	1.9	0.8–1.9
Alpha 2 (%)	15.6	18.6	7.4–15.4
Beta 1 (%)	9.0	5.7	4.5–6.2
Beta 2 (%)	5.5	5.8	4.5–8
Gamma (%)	12.4	9.6	8.5–24.2
Albumin (g/L)	38.2	42.0	35.7–48.7
Alpha 1 (g/L)	0.9	1.4	0.6–1.3
Alpha 2 (g/L)	10.6	13.4	5.6–10.6
Beta 1 (g/L)	6.1	4.1	3–4.7
Beta 2 (g/L)	3.7	4.2	3.2–5.8
Gamma (g/L)	8.4	6.9	5.1–18.3
A/G Ratio	1.28	1.40	1.07–1.87

Table 4. Results of the real-time RT-qPCR performed on Days 14 and 31, using oropharyngeal, nasal and rectal swabs collected at Days 14 and 31, targeting the SARS-CoV-2 RNA.

	Day 14				Day 31			
-	Ct Values			Conclusive Laboratory Diagnosis	Ct Values			Conclusive Laboratory Diagnosis
Swab	E Gene	N Gene	RdRp Gene	-	E Gene	N Gene	RdRp Gene	-
OP	30.14	36.38	39.60	Positive	n.d.	n.d,	n.d,	Negative
N	27.83	34.47	36.00	Positive	36.00	n.d.	n.d.	Positive
R	n.d.	n.d.	n.d.	Negative	n.d.	n.d.	n.d.	Negative

OP—oropharyngeal swab; N—nasal swab; R—rectal swab; n.d.—not detected; N gene—nucleocapsid protein; E gene—envelope protein gene; RdRp gene—RNA-dependent RNA polymerase gene.

Table 5. Serological assays performed at different times: first cat's blood sample collected 7 days after the beginning of the respiratory and gastrointestinal symptoms on Days 14 and 31.

SEROLOGY SARS-CoV-2	Day 14	Day 31	Reference Ranges
ELISA KIT 1	NEGATIVE	POSITIVE (68%)	Cut-off $\geq$ 60%
ELISA KIT 2	POSITIVE (33.6%)	POSITIVE (20.8%)	Cut-off $\geq$ 20%
ECLIA	47.20 U/mL	1598 U/mL	POSITIVE $\geq$ 0.8 U/mL
PNRT	1:5120	1:2560	<1:10

ELISA KIT 1, ELISA KIT 2, SARS-CoV-2 ECLIA (electrochemiluminescence: Elecsys anti SARS-CoV-2 S double antigen assay for the detection of IgG antibodies against coronavirus RBD spike protein, Roche Diagnostics), and PRNT (plaque reduction neutralization test).

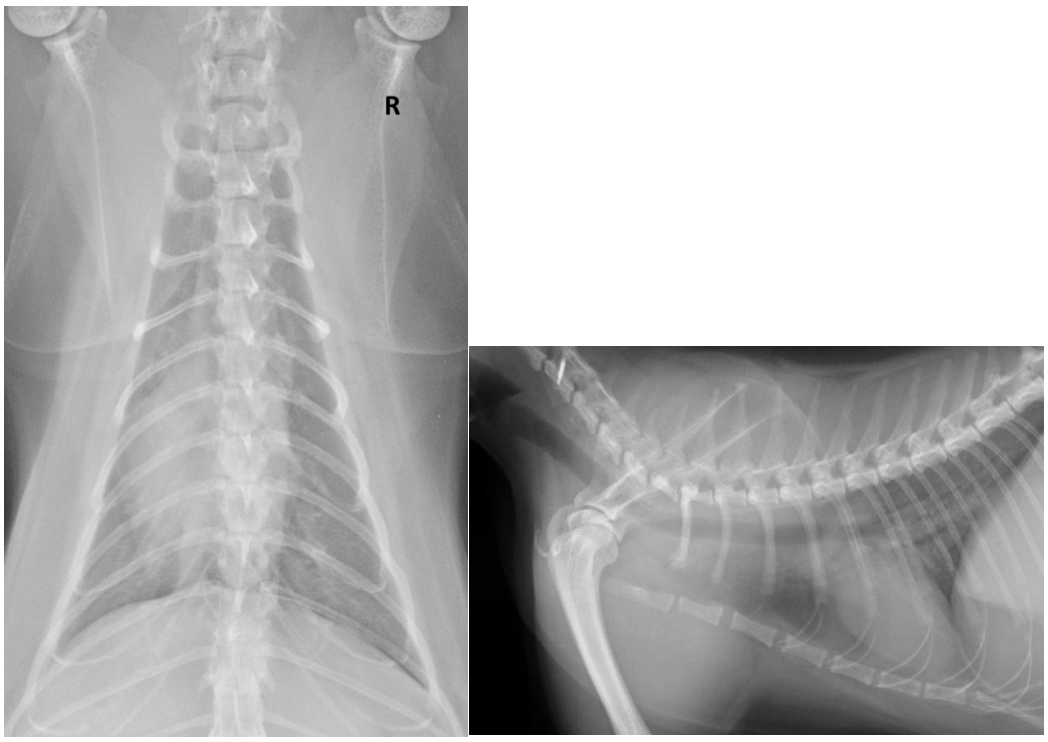


Figure 2. Thoracic radiographic study on Day 14: dorsoventral (DV) and right lateral radiographies (RL). The radiologic findings show a mild-to-severe diffuse bronchial pattern and a mild diffuse interstitial lung pattern. R = right side.

References

1. Lai, C.C.; Shih, T.P.; Ko, W.C.; Tang, H.J.; Hsueh, P.R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. Int. J. Antimicrob. Agents 2020, 55, 105924.

2. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020, 395, 497–506.

3. Abdel-Moneim, A.S.; Abdelwhab, E.M. Evidence for SARS-CoV-2 Infection of Animal Hosts. Pathogens 2020, 9, 529.

4. Decaro, N.; Lorusso, A. Novel human coronavirus (SARS-CoV-2): A lesson from animal coronaviruses. Vet. Microbiol. 2020, 244, 108693.

5. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.; et al. A new coronavirus associated with human respiratory disease in China. Nature 2020, 579, 265–269.

6. Ge, X.-Y.; Li, J.-L.; Yang, X.-L.; Chmura, A.A.; Zhu, G.; Epstein, J.H.; Mazet, J.K.; Hu, B.; Zhang, W.; Peng, C.; et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 2013, 503, 535–538.

7. Xiao, K.; Zhai, J.; Feng, Y.; Zhou, N.; Zhang, X.; Zou, J.-J.; Li, N.; Guo, Y.; Li, X.; Shen, X.; et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. Nat. Cell Biol. 2020, 583, 286–289.

8. Vale, B.D.; Lopes, A.P.; Fontes, M.D.C.; Silvestre, M.; Cardoso, L.; Coelho, A.C. Bats, pangolins, minks and other animals— villains or victims of SARS-CoV-2? Veter. Res. Commun. 2021, 45, 1–19.

9. Shi, J.; Wen, Z.; Zhong, G.; Yang, H.; Wang, C.; Huang, B.; Liu, R.; He, X.; Shuai, L.; Sun, Z.; et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2. Science 2020, 368, 1016–1020.

10. Muñoz-Fontela, C.; Dowling, W.E.; Funnell, S.G.P.; Gsell, P.-S.; Riveros-Balta, A.X.; Albrecht, R.A.; Andersen, H.; Baric, R.S.; Carroll, M.W.; Cavaleri, M.; et al. Animal models for COVID-19. Nature 2020, 586, 509–515.

11. Hobbs, E.C.; Reid, T.J. Animals and SARS-CoV-2: Species susceptibility and viral transmission in experimental and natural conditions, and the potential implications for community transmission. Transbound. Emerg. Dis. 2020, 13885.

12. Kumar, R.; Harilal, S.; Al-Sehemi, A.G.; Pannipara, M.; Behl, T.; Mathew, G.E.; Mathew, B. COVID-19 and Domestic Animals: Exploring the Species Barrier Crossing, Zoonotic and Reverse Zoonotic Transmission of SARS-CoV-2. Curr. Pharm. Des. 2021, 27, 1194–1201.

13. Kim, Y.-I.; Kim, S.-G.; Kim, S.-M.; Kim, E.-H.; Park, S.-J.; Yu, K.-M.; Chang, J.-H.; Lee, S.; Casel, M.A.B.; Um, J.; et al. Infection and Rapid Transmission of SARS-CoV-2 in Ferrets. Cell Host Microbe 2020, 27, 704–709.e2.

14. Hossain, G.; Javed, A.; Akter, S.; Saha, S. SARS-CoV-2 host diversity: An update of natural infections and experimental evidence. J. Microbiol. Immunol. Infect. 2021, 54, 175–181.

15. Gibbons, A. Captive gorillas test positive for coronavirus. Science 2021.

16. Patterson, E.I.; Elia, G.; Grassi, A.; Giordano, A.; Desario, C.; Medardo, M.; Smith, S.L.; Anderson, E.R.; Prince, T.; Patterson, G.T.; et al. Evidence of exposure to SARS-CoV-2 in cats and dogs from households in Italy. Nat. Commun. 2020, 11, 1–5.

17. Gaudreault, N.N.; Trujillo, J.D.; Carossino, M.; Meekins, D.A.; Madden, D.W.; Indran, S.V.; Bold, D.; Balaraman, V.; Kwon, T.; Libanori Artiaga, B.; et al. SARS-CoV-2 Infection, Disease and Transmission in Domestic Cats Running Title: SARS-CoV-2 in Domestic Cats. bioRxiv 2020. Available online: https://www.biorxiv.org/content/10.1101/2020.08.04.235002v1 (accessed on 6 May 2021).

18. Bosco-Lauth, A.M.; Jeffrey Root, J.; Porter, S.M.; Walker, A.E.; Guilbert, L.; Hawvermale, D.; Pepper, A.; Maison, R.M.; Hartwig, A.E.; Bielefeldt-Ohmann, H.; et al. Survey of Peridomestic Mammal Susceptibility to SARS-CoV-2 Infection. BioRxiv 2021. Available online: https://www.biorxiv.org/content/10.1101/2021.01.21.427629v1 (accessed on 6 May 2021).

19. Palmer, M.V.; Martins, M.; Falkenberg, S.; Buckley, A.; Caserta, L.C.; Mitchell, P.K.; Cassmann, E.D.; Rollins, A.; Zylich, N.C.; Renshaw, R.W.; et al. Susceptibility of White-Tailed Deer (Odocoileus Virginianus) to SARS-CoV-2. BioRxiv 2021. Available online: https://www.biorxiv.org/content/10.1101/2021.01.13.426628v1 (accessed on 6 May 2021).

20. McAloose, D.; Laverack, M.; Wang, L.; Killian, M.L.; Caserta, L.C.; Yuan, F.; Mitchell, P.K.; Queen, K.; Mauldin, M.R.; Cronk, B.D.; et al. From People to Panthera: Natural SARS-CoV-2 Infection in Tigers and Lions at the Bronx Zoo. mBio 2020, 11.

21. Gryseels, S.; De Bruyn, L.; Gyselings, R.; Calvignac-Spencer, S.; Leendertz, F.H.; Leirs, H. Risk of human-to-wildlife transmission of SARS-CoV-2. Mammal Rev. 2021, 51, 272–292.

22. Munnink, B.B.O.; Sikkema, R.S.; Nieuwenhuijse, D.F.; Molenaar, R.J.; Munger, E.; Molenkamp, R.; van der Spek, A.; Tolsma, P.; Rietveld, A.; Brouwer, M.; et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. Science 2021, 371, 172–177.

23. Mykytyn, A.Z.; Lamers, M.M.; Okba, N.M.A.; Breugem, T.I.; Schipper, D.; Doel, P.B.V.D.; van Run, P.; van Amerongen, G.; de Waal, L.; Koopmans, M.P.G.; et al. Susceptibility of rabbits to SARS-CoV-2. Emerg. Microbes Infect. 2021, 10, 1–7.

24. Leroy, E.M.; Gouilh, M.A.; Brugère-Picoux, J. The risk of SARS-CoV-2 transmission to pets and other wild and domestic animals strongly mandates a one-health strategy to control the COVID-19 pandemic. One Health 2020, 10, 100133.

25. Newman, A.; Smith, D.; Ghai, R.R.; Wallace, R.M.; Torchetti, M.K.; LoIacono, C.; Murrell, L.S.; Carpenter, A.; Moroff, S.; Rooney, J.A.; et al. First Reported Cases of SARS-CoV-2 Infection in Companion Animals—New York, March–April 2020. MMWR. Morb. Mortal. Wkly. Rep. 2020, 69, 710–713.

26. Hernández, M.; Abad, D.; Eiros, J.M.; Rodríguez-Lázaro, D. Are Animals a Neglected Transmission Route of SARS-CoV-2? Pathogens 2020, 9, 480.

27. Jo, W.K.; De Oliveira-Filho, E.F.; Rasche, A.; Greenwood, A.D.; Osterrieder, K.; Drexler, J.F. Potential zoonotic sources of SARS-CoV-2 infections. Transbound. Emerg. Dis. 2020, 13872.

28. Bonilla-Aldana, D.K.; Cardona-Trujillo, M.C.; García-Barco, A.; Holguin-Rivera, Y.; Cortes-Bonilla, I.; Bedoya-Arias, H.A.; Patiño-Cadavid, L.J.; Tamayo-Orozco, J.D.; Paniz-Mondolfi, A.; Zambrano, L.I.; et al. MERS-CoV and SARS-CoV Infections in Animals: A Systematic Review and Meta-Analysis of Prevalence Studies. Preprints 2020, 28 (Suppl. 1), 71–83. Available online: https://www.preprints.org/manuscript/202003.0103/v2 (accessed on 6 May 2021).

29. Stout, A.E.; André, N.M.; Jaimes, J.A.; Millet, J.K.; Whittaker, G.R. Coronaviruses in cats and other companion animals: Where does SARS-CoV-2/COVID-19 fit? Vet. Microbiol. 2020, 247, 108777.

30. Costagliola, A.; Liguori, G.; D'Angelo, D.; Costa, C.; Ciani, F.; Giordano, A. Do Animals Play a Role in the Transmission of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)? A Commentary. Animals 2020, 11, 16.

31. Hamer, S.A.; Pauvolid-Corrêa, A.; Zecca, I.B.; Davila, E.; Auckland, L.D.; Roundy, C.M.; Tang, W.; Torchetti, M.; Killian, M.L.; Jenkins-Moore, M.; et al. Natural SARS-CoV-2 Infections, Including Virus Isolation, among Serially Tested Cats and Dogs in Households with Con-firmed Human COVID-19 Cases in Texas, USA. bioRxiv 2020. Available online: https://www.biorxiv.org/ content/10.1101/2020.12.08.416339v1 (accessed on 6 May 2021).

32. Sailleau, C.; Dumarest, M.; Vanhomwegen, J.; Delaplace, M.; Caro, V.; Kwasiborski, A.; Hourdel, V.; Chevaillier, P.; Barbarino, A.; Comtet, L.; et al. First

detection and genome sequencing of SARS-CoV-2 in an infected cat in France. Transbound. Emerg. Dis. 2020, 67, 2324–2328.

33. Halfmann, P.J.; Hatta, M.; Chiba, S.; Maemura, T.; Fan, S.; Takeda, M.; Kinoshita, N.; Hattori, S.-I.; Sakai-Tagawa, Y.; Iwatsuki- Horimoto, K.; et al. Transmission of SARS-CoV-2 in Domestic Cats. N. Engl. J. Med. 2020, 383, 592–594.

34. Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020, 579, 270–273.

35. Bosco-Lauth, A.M.; Hartwig, A.E.; Porter, S.M.; Gordy, P.W.; Nehring, M.; Byas, A.D.; VandeWoude, S.; Ragan, I.K.; Maison, R.M.; Bowen, R.A. Experimental infection of domestic dogs and cats with SARS-CoV-2: Pathogenesis, transmission, and response to reexposure in cats. Proc. Natl. Acad. Sci. USA 2020, 117, 26382–26388.

36. Chiba, S.; Halfmann, P.J.; Hatta, M.; Maemura, T.; Fan, S.; Armbrust, T.; Swartley, O.M.; Crawford, L.K.; Kawaoka, Y. Protective Immunity and Persistent Lung Sequelae in Domestic Cats after SARS-CoV-2 Infection. Emerg. Infect. Dis. 2021, 27, 660–663.

37. Zhang, Q.; Zhang, H.; Gao, J.; Huang, K.; Yang, Y.; Hui, X.; He, X.; Li, C.; Gong, W.; Zhang, Y.; et al. A serological survey of SARS-CoV-2 in cat in Wuhan. Emerg. Microbes Infect. 2020, 9, 2013–2019.

38. Deng, J.; Liu, Y.; Sun, C.; Bai, J.; Sun, J.; Hao, L.; Li, X.; Tian, K. SARS-CoV-2 Serological Survey of Cats in China before and after the Pandemic. Virol. Sin. 2020, 35, 846–848.

39. Stevanovic, V.; Vilibic-Cavlek, T.; Tabain, I.; Benvin, I.; Kovac, S.; Hruskar, Z.; Mauric, M.; Milasincic, L.; Antolasic, L.; Skrinjaric, A.; et al. Seroprevalence of SARS-CoV-2 infection among pet animals in Croatia and potential public health impact. Transbound. Emerg. Dis. 2020, 13924.

40. Fritz, M.; Rosolen, B.; Krafft, E.; Becquart, P.; Elguero, E.; Vratskikh, O.; Denolly, S.; Boson, B.; Vanhomwegen, J.; Gouilh, M.A.; et al. High prevalence of SARS-CoV-2 antibodies in pets from COVID-19+ households. One Health 2020, 11, 100192.

41. Michelitsch, A.; Hoffmann, D.; Wernike, K.; Beer, M. Occurrence of Antibodies against SARS-CoV-2 in the Domestic Cat Population of Germany. Vaccines 2020, 8, 772.

42. Gaudreault, N.N.; Trujillo, J.D.; Carossino, M.; Meekins, D.A.; Morozov, I.; Madden, D.W.; Indran, S.V.; Bold, D.; Balaraman, V.; Kwon, T.; et al. SARS-CoV-2 infection, disease and transmission in domestic cats. Emerg. Microbes Infect. 2020, 9, 2322–2332.

43. Garigliany, M.; Van Laere, A.-S.; Clercx, C.; Giet, D.; Escriou, N.; Huon, C.; Van Der Werf, S.; Eloit, M.; Desmecht, D. SARS-CoV-2 Natural Transmission from Human to Cat, Belgium, March 2020. Emerg. Infect. Dis. 2020, 26, 3069–3071.

44. Barrs, V.R.; Peiris, M.; Tam, K.W.; Law, P.Y.; Brackman, C.J.; To, E.M.; Yu, V.Y.; Chu, D.K.; Perera, R.A.; Sit, T.H. SARS-CoV-2 in Quarantined Domestic Cats from COVID-19 Households or Close Contacts, Hong Kong, China. Emerg. Infect. Dis. 2020, 26, 3071–3074.

45. Hosie, M.J.; Hofmann-Lehmann, R.; Hartmann, K.; Egberink, H.; Truyen, U.; Addie, D.D.; Belák, S.; Boucraut-Baralon, C.; Frymus, T.; Lloret, A.; et al. Anthropogenic Infection of Cats during the 2020 COVID-19 Pandemic. Viruses 2021, 13, 185.

46. Klaus, J.; Meli, M.; Willi, B.; Nadeau, S.; Beisel, C.; Stadler, T.; Egberink, H.; Zhao, S.; Lutz, H.; Riond, B.; et al. Detection and Genome Sequencing of SARS-CoV-2 in a Domestic Cat with Respiratory Signs in Switzerland. Viruses 2021, 13, 496.

47. Yang, J.; Zheng, Y.; Gou, X.; Pu, K.; Chen, Z.; Guo, Q.; Ji, R.; Wang, H.; Wang, Y.; Zhou, Y. Prevalence of Comorbidities and Its Effects in Patients Infected with SARS-CoV-2: A Systematic Review and Meta-Analysis. Int. J. Infect. Dis. 2020, 94, 91–95.

48. Segalés, J.; Puig, M.; Rodon, J.; Avila-Nieto, C.; Carrillo, J.; Cantero, G.; Terrón, M.T.; Cruz, S.; Parera, M.; Noguera-Julián, M.; et al. Detection of SARS-CoV-2 in a Cat Owned by a COVID-19 affected Patient in Spain. Proc. Natl. Acad. Sci. USA 2020, 117, 24790–24793.

49. Klaus, J.; Palizzotto, C.; Zini, E.; Meli, M.L.; Leo, C.; Egberink, H.; Zhao, S.; Hofmann-Lehmann, R. SARS-CoV-2 Infection and Antibody Response in a Symptomatic Cat from Italy with Intestinal B-Cell Lymphoma. Viruses 2021, 13, 527.

50. De Morais, H.A.; dos Santos, A.P.; do Nascimento, N.C.; Kmetiuk, L.B.; Barbosa, D.S.; Brandão, P.E.; Guimarães, A.M.S.; Pettan-Brewer, C.; Biondo, A.W. Natural Infection by SARS-CoV-2 in Companion Animals: A Review of Case Reports and Current Evidence of Their Role in the Epidemiology of COVID-19. Front. Vet. Sci. 2020, 7, 591216.

51. World Health Organization. Rational Use of Personal Protective Equipment for Coronavirus Disease 2019 (COVID-19) and Considerations during Severe Shortages; WHO: Geneva, Switzerland, 2020.

52. World Health Organization. Home Care for Patients with Suspected Novel Coronavirus (NCoV) Infection Presenting with Mild Symptoms and Management of Contacts; WHO: Geneva, Switzerland, 2020.

53. European Centre for Disease Prevention and control. Cloth Masks and Mask Sterilisation as Options in Case of Shortage of Surgical Masks and Respirators Use of Cloth Face Masks for Protection against COVID-19 in Clinical Settings. 2020. Available online: https://www.ecdc.europa.eu/sites/default/files/documents/Cloth-face-masks-in-case-shortage-surgical-masks-respirators2020-03-26.pdf (accessed on 6 May 2021).

54. European Centre for Disease Prevention and control. Infection Prevention and Control for COVID-19 in Healthcare Settings. 2021. Available online: https://www.ecdc.europa.eu/en/publications-data/infection-prevention-and-control-and-preparedness-covid-19-healthcare-settings (accessed on 6 May 2021).

55. Hofmann-Lehmann, R.; Huder, J.B.; Gruber, S.; Boretti, F.; Sigrist, B.; Lutz, H. Feline Leukaemia Provirus Load during the Course of Experimental Infection and in Naturally Infected Cats. J. Gen. Virol. 2001, 82.

56. Meli, M.L.; Berger, A.; Willi, B.; Spiri, A.M.; Riond, B.; Hofmann-Lehmann, R. Molecular Detection of Feline Calicivirus in Clinical Samples: A Study Comparing Its Detection by RT-QPCR Directly from Swabs and after Virus Isolation. J. Virol. Methods 2018, 251.

57. Hussein, I.T.M.; Field, H.J. Development of a Quantitative Real-Time TaqMan PCR Assay for Testing the Susceptibility of Feline Herpesvirus-1 to Antiviral Compounds. J. Virol. Methods 2008, 152.

58. Decaro, N.; Elia, G.; Martella, V.; Desario, C.; Campolo, M.; Di Trani, L.; Tarsitano, E.; Tempesta, M.; Buonavoglia, C. A Real-Time PCR Assay for Rapid Detection and Quantitation of Canine Parvovirus Type 2 in the Feces of Dogs. Vet. Microbiol. 2005, 105.

59. Corman, V.M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D.K.W.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M.L.; et al. Detection of 2019 Novel Coronavirus (2019-NCoV) by Real-Time RT-PCR. Eurosurveillance 2020, 25.

60. Hoffmann, B.; Depner, K.; Schirrmeier, H.; Beer, M. A Universal Heterologous Internal Control System for Duplex Real-Time RT-PCR Assays Used in a Detection System for Pestiviruses. J. Virol. Methods 2006, 136.

61. Neil Carey, R.; Paul Durham, A.; Hauck, W.W.; Kallner, A.; Kondratovich, M.V.; Guy Middle, J.; Pierson-Perry, J.F.; Smith, M.B.; Srinivasan, A. EP15-A3 User Verification of Precision and Estimation of Bias. In Approved Guideline, third ed.; Clinical and Laboratory Standards Institute: Pittsburg, PA, USA, 2014; Volume 34.

62. Padoan, A.; Bonfante, F.; Pagliari, M.; Bortolami, A.; Negrini, D.; Zuin, S.; Bozzato, D.; Cosma, C.; Sciacovelli, L.; Plebani, M. Analytical and Clinical Performances of Five Immunoassays for the Detection of SARS-CoV-2 Antibodies in Comparison with Neutralization Activity. EBioMedicine 2020, 62, 103101.

63. Houtgast, E.J.; Sima, V.M.; Bertels, K.; Al-Ars, Z. Hardware Acceleration of BWA-MEM Genomic Short Read Mapping for Longer Read Lengths. Comput. Biol. Chem. 2018, 75, 54–64.

64. Houtgast, E.J.; Sima, V.M.; Bertels, K.; Al-Ars, Z. Comparative Analysis of System-Level Acceleration Techniques in Bioinformatics: A Case Study of Accelerating the Smith-Waterman Algorithm for BWA-MEM. In Proceedings of the 2018 IEEE 18th International Conference on Bioinformatics and Bioengineering, BIBE 2018, Taichung, Taiwan, 29–31 October 2018.

65. Volz, E.; Mishra, S.; Chand, M.; Barrett, J.C.; Johnson, R.; Hopkins, S.; Gandy, A.; Rambaut, A.; Ferguson, N.M. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from Linking Epidemiological and Genetic Data. medRxiv 2021.

66. O'Toole, Á.; Hill, V.; McCrone, J.T.; Scher, E.; Rambaut, A. Pangolin COVID-19 Lineage Assigner Phylogenetic Assignment of Named Global Outbreak LINeages. Available online: https://pangolin.cog-uk.io/ (accessed on 23 December 2020).

67. Trzil, J.E. Feline Asthma. Vet. Clin. N. Am. Small Anim. Pract. 2020, 50, 375–391.

68. World Health Organization. Coronavirus Disease (COVID-19) Pandemic. Available online: https://www.who.int/emergencies/ diseases/novel-coronavirus-2019 (accessed on 6 May 2021).

69. OIE. World Organization for Animal Health. Available online: https://www.oie.int/en/scientific-expertise/specific-information-and-recommendations/questions-and-answers-on-2019novel-coronavirus/ (accessed on 6 May 2021).

70. Promed International Society for Infectious Disease Homepage. Available online: https://promedmail.org/ (accessed on 6 May 2021).

71. Wu, L.; Chen, Q.; Liu, K.; Wang, J.; Han, P.; Zhang, Y.; Hu, Y.; Meng, Y.; Pan, X.; Qiao, C.; et al. Broad Host Range of SARS-CoV-2 and the Molecular Basis for SARS-CoV-2 Binding to Cat ACE2. Cell Discov. 2020, 6, 68.

72. Zhao, X.; Chen, D.; Szabla, R.; Zheng, M.; Li, G.; Du, P.; Zheng, S.; Li, X.; Song, C.; Li, R.; et al. Broad and Differential Animal Angiotensin-Converting Enzyme 2 Receptor Usage by SARS-CoV-2. J. Virol. 2020, 94.

73. Von Roedern, M.; Buriko, Y.; Prittie, J.; Lamb, K. Investigation of Iron Status and Markers of Inflammation in Anaemic and Non-Anaemic Hospitalised Cats. J. Small Anim. Pract. 2017, 58.

74. Rand, J.S.; Kinnaird, E.; Baglioni, A.; Blackshaw, J.; Priest, J. Acute Stress Hyperglycemia in Cats Is Associated with Struggling and Increased Concentrations of Lactate and Norepinephrine. J. Vet. Intern. Med. 2002, 16, 123.

75. Bouillon, J.; Duke, T.; Focken, A.P.; Snead, E.C.; Cosford, K.L. Effects of Dexmedetomidine on Glucose Homeostasis in Healthy Cats. J. Feline Med. Surg. 2020, 22, 344–349.

76. Paltrinieri, S. The Feline Acute Phase Reaction. Vet. J. 2008, 177, 26–35.

77. Willard, M.D.; Tvedten, H. Small Animal Clinical Diagnosis by Laboratory Methods, 5th ed.; Saunders: St. Louis, MO, USA, 2011.

78. Fernandez, N.J.; Kidney, B.A. Alkaline Phosphatase: Beyond the Liver. Vet. Clin. Pathol. 2007, 36, 223–233.

79. Cerón, J.J.; Eckersall, P.D.; Martínez-Subiela, S. Acute Phase Proteins in Dogs and Cats: Current Knowledge and Future Perspectives. Vet. Clin. Pathol. 2005, 34, 85–99.

80. Kajikawa, T.; Furuta, A.; Onishi, T.; Tajima, T.; Sugii, S. Changes in Concentrations of Serum Amyloid A Protein, A1-Acid Glycoprotein, Haptoglobin, and C-

Reactive Protein in Feline Sera Due to Induced Inflammation and Surgery. Vet. Immunol. Immunopathol. 1999, 68, 91–98.

81. Marly-Voquer, C.; Riond, B.; Jud Schefer, R.; Kutter, A.P.N. Reference Values for Rotational Thromboelastometry (ROTEM) in Clinically Healthy Cats. J. Vet. Emerg. Crit. Care 2017, 27, 185–192.

82. Solbak, S.; Epstein, S.E.; Hopper, K. Influence of Needle Gauge Used for Venipuncture on Measures of Hemostasis in Cats. J. Feline Med. Surg. 2019, 21, 143–147.

83. Khelik, I.A.; Berger, D.J.; Mochel, J.P.; Seo, Y.-J.; Palerme, J.-S.; Ware, W.A.; Ward, J.L. Clinicopathologic, Hemodynamic, and Echocardiographic Effects of Short-Term Oral Administration of Anti-Inflammatory Doses of Prednisolone to Systemically Normal Cats. Am. J. Vet. Res. 2019, 80, 743–755.

84. Zhang, Q.; Zhang, H.; Huang, K.; Yang, Y.; Hui, X.; Gao, J.; He, X.; Li, C.; Gong, W.; Zhang, Y.; et al. SARS-CoV.2 Neutralizing Serum Antibodies in Cats: A Serological Investigation. bioRxiv Prepr. Serv. Biol. 2020. Available online: https://www.biorxiv.org/content/10.1101/2020.04.01.021196v1 (accessed on 6 May 2021).

85. Volz, E.; Hill, V.; McCrone, J.T.; Price, A.; Jorgensen, D.; O'Toole, Á.; Southgate, J.; Johnson, R.; Jackson, B.; Nascimento, F.F.; et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. Cell 2021, 184, 64–75.e11.

86. Davis, M.F.; Innes, G.K. The Cat's in the Bag: Despite Limited Cat-to-Cat SARS-CoV-2 Transmission, One Health Surveillance Efforts Are Needed. J. Infect. Dis. 2021.

87. Ferasin, L.; Fritz, M.; Ferasin, H.; Legros, V.; Leroy, E.M. Myocarditis in Naturally Infected Pets with the British Variant of COVID-19. bioRxiv 2021, 2021.03.18.435945. Available online: https://www.biorxiv.org/content/10.1101/2021.03.18.435945v1 (accessed on 6 May 2021).

88. Centers for Disease Control and Prevention. Interim Guidance for Public Health Professionals Managing People with COVID-19 in Home Care and Isolation Who Have Pets or Other Animals. Available online: https://www.cdc.gov/coronavirus/2019-ncov/animals/interimguidance-%0Amanaging-people-in-home-care-and-isolation-who-have-pets.html (accessed on 6 May 2021).

3.2 DNA Viruses

Known viruses can be identified by a wide range of techniques: traditional, molecular and NGS methods [1]. NGS, using a metagenomics approach, are the only able to detect novel viruses. The rapid diagnosis of novel viruses is of extreme importance for the selection of prevention and treatment strategies. The DNA viruses have mutation rates between 10^{-6} to 10^{-8} mutations per base per generation, while the average mutation rates for RNA viruses fall in the range of 10^{-3} to 10^{-6} mutations incorporated per nucleotide copied [2]. The companion bird deaths in 2021 in the North Central Italy gave the opportunity to investigate different DNA species present in bird species and to solve discrepancies in identification of viruses using different diagnostic techniques.

3.2.1 Polyomavirus and Novel Circovirus: Untargeted approach using Miseq technology

Abstract

A wave of fatal incidents among companion birds struck several regions in North Central Italy between February and June 2021, affecting various bird species. Despite standard diagnostic efforts, including necropsies and molecular tests, the causative agent remained unidentified, leading to an in-depth investigation utilizing next-generation sequencing (NGS) technology. Analysis of organ homogenates through electron microscopy initially suggested Circoviridae-like particles, though specific PCR tests failed to confirm these findings. Employing a metagenomic approach revealed the presence of Pyrrhula Pyrrhula Polyomavirus and Corvus/Cracticus Polyomavirus in select samples. Notably, discrepancies in avian Polyomavirus detection were attributed to primer mismatches, highlighting the need for precise primer design. Further RNA analysis exposed genetic variations in Beak and Feather Disease Virus (Circovirus), emphasizing the dynamic nature of viral evolution. Phylogenetic analysis led to the understanding of the different strain and the development of a specific protocol for rapid and accurate detection of the novel Circovirus strain. This investigation unveiled the limitations of conventional diagnostics and underscored the importance of complementary methodologies in identifying elusive pathogens. The study emphasizes the necessity for refined

molecular tools to adapt to evolving viral genomes, highlighting the dynamic nature of viral evolution and its impact on diagnostic accuracy. Moreover, the discovery of novel viral sequences highlighted the significance of comprehensive metagenomic approaches in uncovering previously unknown viral species or strains. The study's implications extend beyond diagnostics, emphasizing the need for adaptive strategies to monitor and manage viral outbreaks in companion bird populations. In conclusion, the integration of electron microscopy, metagenomics, and refined primer design proved pivotal in delineating the diverse viral landscape in these bird populations. This study advocates for the integration of multiple diagnostic modalities and continuous refinement of molecular tools for accurate and comprehensive viral identification in avian veterinary pathology.

Introduction

A series of fatal episodes affecting companion birds in various regions of Italy, specifically in the North Central regions (Province of Bolzano, Torino, Treviso, Vicenza, Verona and Brescia) happened from February to June 2021. These outbreaks affected different species of birds within the families of Psittacidae, Estrildidae and Fringillidae (i.e Melopsittacus undulatus, Psittacus erithacus, Erythrura gouldiae, Spinus cucullatus and Serinus canaria). Despite standard necropsy protocols and infectious disease diagnosis, the cause of these fatalities remained elusive upon electron microscopy, prompting a deeper investigation using NGS technology. To identify potential viral agents, the organ homogenates underwent rigorous analysis, including virological molecular test and conventional PCR, classical virological technique as electron microscopy and, in the end, metagenomic approaches and new protocols were developed. Interestingly, electron microscopy revealed the presence of virus-like particles resembling Circoviridae, while specific PCR tests failed to detect these viruses. To resolve the discrepancies, a metagenomic approach was developed, leading to the discovery of Pyrrhula Pyrrhula Polyomavirus and Corvus/Cracticus Polyomavirus in certain samples. Further examination of the avian Polyomavirus primers highlighted the presence of mismatches for the detection of the viruses, explaining the failed detections. Subsequent RNA analysis of Psittacines unveiled the presence of Beak and Feather

Disease Virus (Circovirus), displaying slight genetic variations among samples from different locations. This variation was elucidated through phylogenetic analysis, prompting the design of a specific protocol for rapid detection of the novel Circovirus in subsequent screenings. This comprehensive investigation not only shed light on the diversity of viruses present but also highlighted the importance of different technology, specifically with an untargeted approach for the discovery of novel species. The findings provide crucial insights into the complexity of viral infections in companion birds, potentially aiding in better diagnostic protocols and mitigation strategies.

Methodologies

Sample analysis

Conventional PCR

Livers and/or intestine were submitted for virological molecular tests and according to the bird species, specific conventional PCR for Canary circovirus (CaCV) was applied in fringillidae birds and Psittacine Circovirus (PBFD) and avian Polyomavirus (APV) in Psittacines. The DNA was extracted from the livers and/or intestine of these samples using High Pure PCR Template Preparation Kit (Roche Diagnostics). AmpliTaq Gold with GeneAmp Twelve Paq 10X PCR Buffer II (Applied Biosystems) with GeneAmp dNTP Mix with dTT (Applied Biosystems) was used for the amplification of 4 µl of DNA. The primers used for the detection of Canary circovirus (CaCV) are: 5'-TATACGGCAGAAGAAGAAGC-3' (Operon/MWG) 10µM and 5'-TTGTAGAACTTCGATCCTTCC-3' (Operon/MWG) 10µM, (1,25 µl each for reaction). The primers used for the detection of Avian Polyomavirus (APV) are: POLY-A 5'-CTTATGTGGGAGGCTGCAGTGTT-3' (Operon/MWG) 100µM and POLY-B 5'-TACTGAAATAGCGTGGTAGGCCTC-3' (Operon/MWG) 100µM, (0,5 µl each for reaction). For the detection of Psittacine Beak and Feather Disease (PBFD) the primers are: PBFD-2F 10µM 5'-AACCCTACAGACGGCGAG-3' (Operon/MWG) 10µM and Primer PBFD-4R 10µM 5'-

GTCACAGTCCTCCTTGTACC-3' (Operon/MWG), 10µM (1,25 µl each for reaction).

Electron Microscopy

An organ homogenate was obtained by grinding 1 g of tissue sample in 5 mL of sterile distilled water (1:5 w/v) using a small mortar and pestle with the aid of fine sterile quartz. The homogenate was first frozen at -20°C, thawed in a water bath at 37°C for 15', and then centrifuged at 3,500 rpm or 2,205× g for 15' and the supernatant at 15,000 rpm or 24,133× g for 15' for clarification. Once 90 µL of the second supernatant were vialed, a formvar/carbon supported copper grid (Electron Microscopy Sciences Formvar/Carbon Copper Grid 200 Mesh) was placed flat on the bottom of the vial. This preparation was spun at high-speed (28−30 psi at 95,000 rpm or 100,000× g) for 15' (Beckman Air-Driven Ultracentrifuge Airfuge). The grid was stained with 10 µL of 2% phosphotungstic acid (PTA) (pH 7); PTA was left on the grid for a few seconds (8−10 sec). The grid was then examined under an EM 208S transmission electron microscope (Philips) and virus particles observed with Megaview III and measured using iTEM software (Olympus SIS).

NGS

The five samples underwent whole-genome sequencing. The DNA and RNA were extracted from liver and intestine of five carcasses using QIAsymphony DSP Virus/Pathogen Kit for the DNA extraction and the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany) for the RNA extraction. Due to the need to identify all the viruses present in these samples, sequencing was carried out using a metagenomic approach. Starting from purified RNA a double-stranded cDNA was produced using the Maxima H Minus Double-Stranded cDNA Synthesis kit (Thermo Scientific™), in parallel 1 ng of genomic DNA was used to specifically analyze the DNA viral component. In both cases sequencing libraries were generated using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA, USA) and processed on the Illumina Miseq platform with the MiSeq reagent kit V2 (2x250 bp PE) (Illumina, San Diego, CA, USA).

Bioinformatic analysis

The DNA was extracted from liver and intestine of the carcasses and analyzed in MiSeq Illumina. In order to analyze these data, an in-house pipeline was developed. From the raw data, in FASTQ format, generated by MiSeq Illumina platform we used Genome Detective online tool (https://www.genomedetective.com/) for a first investigation of the raw data. Blastn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed and the consensus sequence generated by Genome Detective and the best blast hits have been used as reference sequence for a reference-based approach. Cutadapt (Martin, 2011; https://cutadapt.readthedocs.io/en/stable/) has been used to trim and filter the reads for quality. FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) has been used for quality analysis of raw and filtered FASTQ files. In order to map the reads against the reference genome, we use the software package BWA (http://bio-bwa.sourceforge.net/) using the algorithm BWA-MEM, that is fast and accurate. An in-house script based on LoFreq (https://csb5.github.io/lofreq/) has been developed for variant calling and another script for consensus sequence generation.

Results

Livers and/or intestine were submitted for virological molecular test and the specific conventional PCR for Canary circovirus (CaCV) applied in fringillidae birds and Psittacine Circovirus (PBFD) could not reveal the presence of these viruses, as the avian Polyomavirus (APV) in Psittacines. These samples were also submitted to classical virological technique by means of negative staining electron microscopy. Negative staining electron microscopy of organs homogenates revealed the presence of numerous virus particles, single or in small groups, morphologically resembling to Circoviridae (Figure 1).

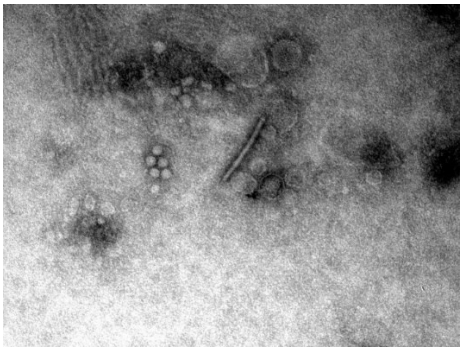


Figure 1. Electron microscopy of organs homogenates.

In order to elucidate the cause of these discrepancies in laboratory results of electron microscopy and those obtained from the available conventional PCR for Circoviridae, a metagenomic approach was applied to these biological specimens.

The investigation started with the analysis of DNA from the two samples from Bolzano (Table 1).

Table 1. Samples analyzed with untargeted approach in MiSeq platform.

sample	host	location	source	method	Virus
21VIR2609-7	Erythura gouldiae	Bolzano	liver	DNA + RNA sequencing	Polyomavirus
21VIR2281-9	Erythura gouldiae	Bolzano	intestine	DNA + RNA sequencing	Polyomavirus
21VIR1149-3	Psittacus erithacus	Torino	liver	RNA sequencing	Circovirus
21VIR2659-12	Psittacine	Macerata	liver	RNA sequencing	Circovirus
21VIR2198-7	Melopsittacus undulatus	Vicenza	liver	RNA sequencing	Circovirus

This bioinformatics analysis using the untargeted approach (metagenomics approach) revealed the presence of Pyrrhula Pyrrhula Polyomavirus and, interestingly, one of the two samples contains also Corvus/Cracticus Polyomavirus (21VIR2281-9). The analysis shows a coinfection of two different species of Polyomavirus in this sample. To prove the presence of these viruses we sequenced also the RNA from the same samples that confirmed these results. We could reconstruct the complete consensus sequence of Pyrrhula Pyrrhula Polyomavirus from both analyses. We examined the primers utilized in the diagnostic in use at the laboratory to detect and amplify the genome of avian Polyomavirus (APV) and a high number of mismatch (3 mismatch in Poly-A primer and 7 mismatch in Poly-B_rc) were identified, that justifies the unsuccessful detection of these viruses. Then

we analyzed RNA from 3 Psittacines (Table 1) and we could obtain 60-93% of the whole genome of Beak and Feather Disease Virus (Circovirus) for all of them. Two of them (21VIR2659-12 and 21VIR2198-7) are identical, where present, and identical to the reference sequence MK387075.1. Interestingly, the consensus sequence of the sample from Turin (21VIR1149-3) has around 70 nucleotide differences from the others (coverage 93%) revealing the presence of a possible novel Circovirus. The phylogenetic analysis performed clearly shows these differences (Figure 3).

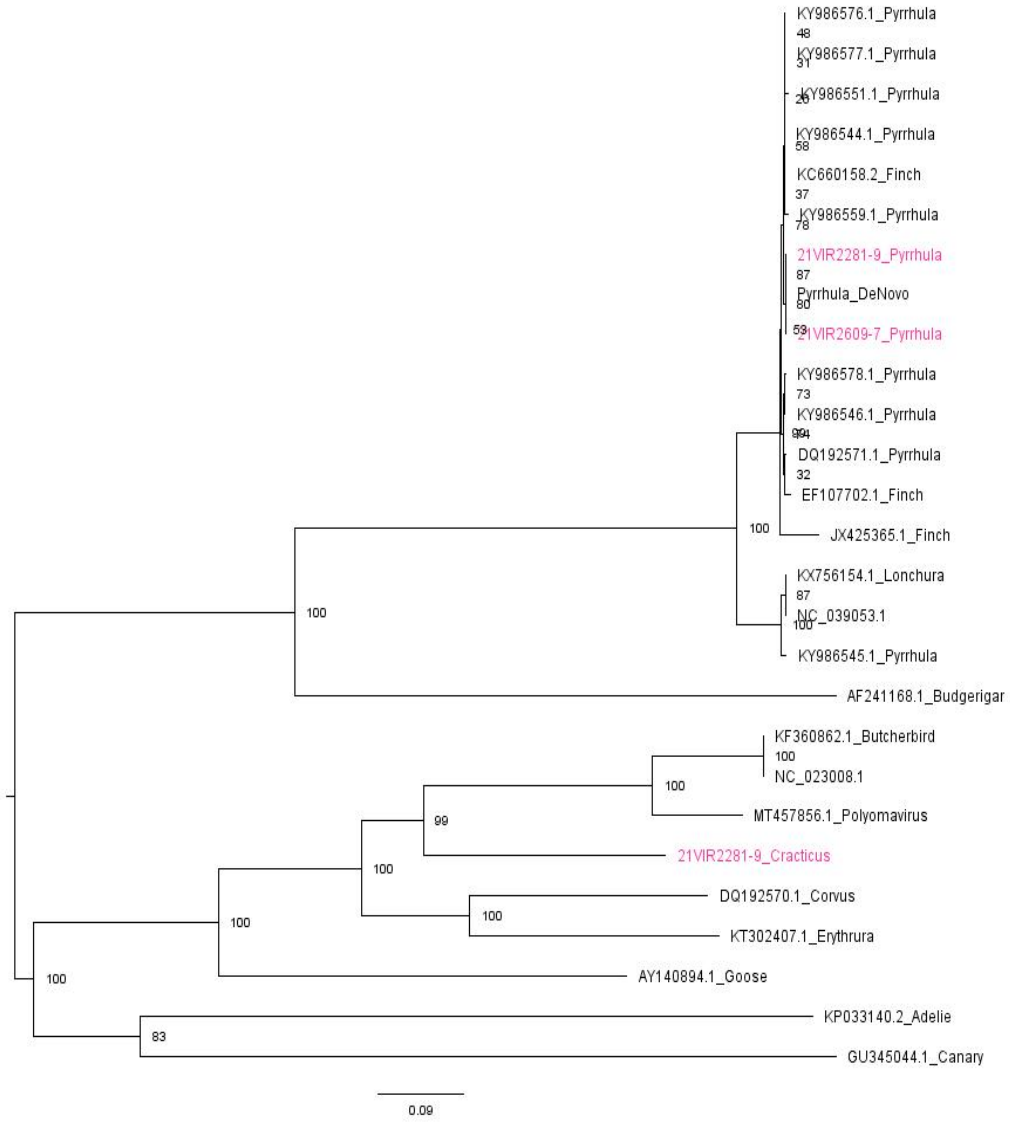


Figure 2. Phylogenetic analysis of Polyomavirus. In pink the Pyrrhula Pyrrhula Polyomavirus that cluster together with the other Pyrrhula Pyrrhula Polyomavirus present in literature and the Corvus/Cracticus Polyomavirus identified in 21VIR2281-9 sample.

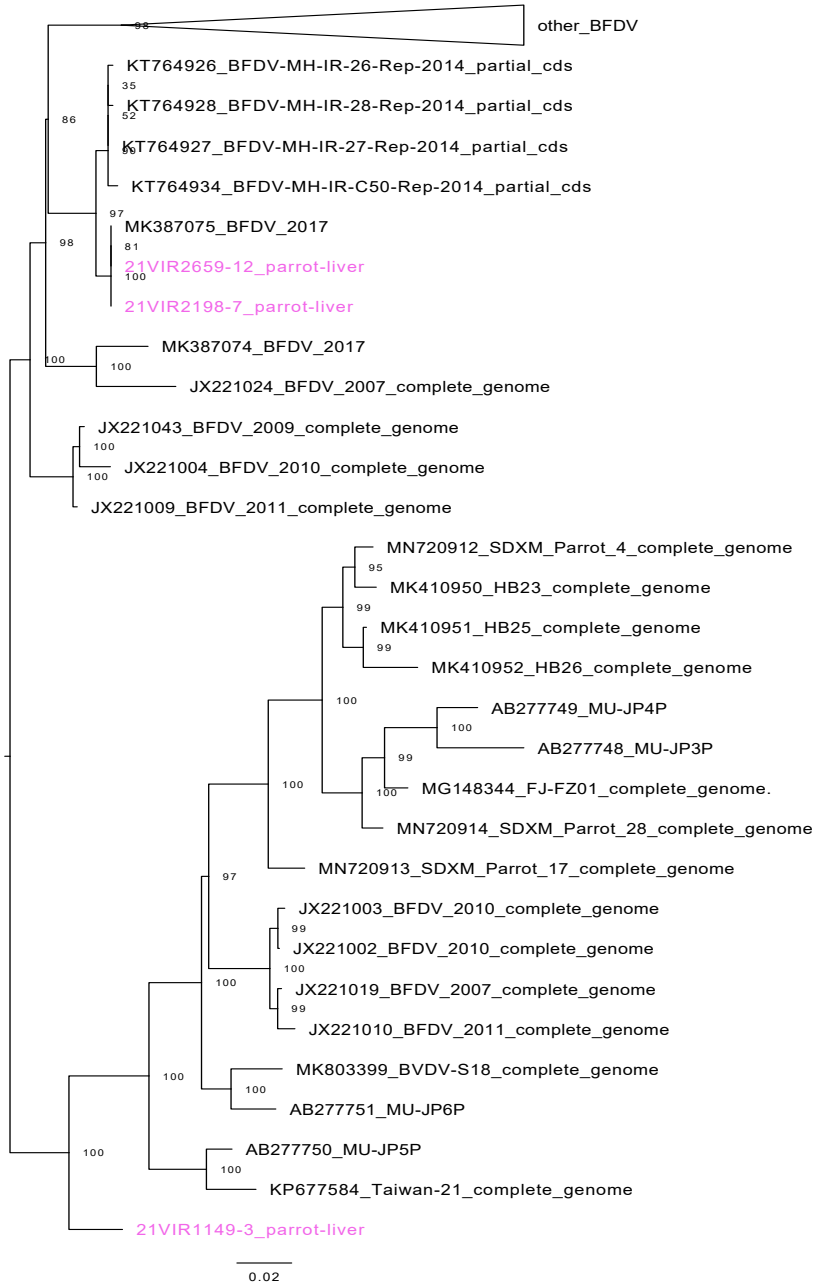


Figure 3. Phylogenetic analysis of Circovirus. The tree shows that sample 21VIR1149-3 (in pink) clusters far from the two other samples (in pink) and from the best 100 sequences of its blast.

In order to perform a rapid detection of this possible novel Circovirus, two couples of specific primers were degnided (Table 2) for conventional PCR using the Primal scheme software (https://primalscheme.com/). These primers are based on the consensus sequence of the possible novel Circovirus sample to perform PCR and rapidly screen all the other samples.

Table 2. Primer designed for novel Circovirus

Primer Direction	Primer Sequence	% GC	Tm	Length	start	stop	Amplicon length
forward	AGCTAAGGGCAGTGATGCG	57.9	58.8	19	368	386	469
reverse	TAAACTCCACGAAGGCGCCT	55	59.3	20	818	837	
forward	AAACCCTTGCTAAGATACCAT	38.1	54	21	1497	1517	302
reverse	GAAGCAAACTACCGCTGCT	52.6	56.7	19	1781	1799	

New protocol for PCR

A new protocol was developed for Master PCR, nested PCR and One Step PCR for novel Circovirus detection at IZSVe.

Preparation method of Master Mix: CircoNew1

While the enzymes are kept refrigerated, the reagents are thowed and mixed, by centrifuging them. The reaction mix is prepared following the scheme outlined below (Table 3).

Table 3: Reagents and primers concentration for PCR

Reagent	**Final concentration**	**µl x 1 REACTION**
Sterile ultrapure water	/	13.75
10X PCR Buffer (Platinum Taq DNA Polymerase Kit Invitrogen cod. 10966-034)	1X	2.5
$MgCl_2$ 50 mM (Platinum Taq DNA Polymerase Kit Invitrogen cod. 10966-034)	1.5 mM	0.75
dNTPs mix 10 mM (Applied Biosystem)	0.5 mM	1.25
Primers: CircoNew1 FOR 10 µM5′ AGTAAGGGCAGTGATGCG 3′	300 nM	0.75
Primers: CircoNew1 REV 10 µM 5′ TAAACTCCACGAAGGCGCCT 3′	300 nM	0.75
Platinum Taq 5 U/µl (Platinum Taq DNA Polymerase Kit Invitrogen cod. 10966-034)	1.25 U	0.25
Total VOLUME		20
DNA sample VOLUME		5
FINAL VOLUME of reaction		25

The prepared mix is vortexed and centrifuged, then the mix is distributed in 20 µl aliquots into 0.2 ml PCR tubes and the DNA is added into their respective wells (Table 4).

Table 4: Temperature and time of PCR steps

Activation of Taq polymerase	Denaturation	Annealing	Extension	Final elongation	Cooling
95°C	94°C	**56°C**	72°C	72°C	4°C
5 min	30 sec	**45 sec**	45sec	5 min	∞
	45 cycles				

Discussion

The investigation into the fatal episodes affecting companion birds in the north central regions of Italy presented a complex scenario where initial diagnostic methods failed to pinpoint the causative agents. The identification of Pyrrhula Pyrrhula Polyomavirus and Corvus/Cracticus Polyomavirus (Figure 2) through metagenomic approaches revealed the virus presence that eluded standard diagnostic techniques. This underscores the limitations of conventional PCR testing, emphasizing the necessity of complementary methodologies in unveiling elusive pathogens within avian populations. The discrepancy between electron microscopy findings of Circoviridae-like particles and the failure of specific PCR tests to detect Canary circovirus (CaCV), Psittacine Circovirus (PBFD), and avian Polyomavirus (APV) warrants attention. The elucidation of mismatches in the avian Polyomavirus primer sequences offers a crucial insight, indicating the need for meticulous primer design and constant adaptation to evolving viral genomes. This finding underscores the importance of periodic reassessment and refinement of molecular diagnostic tools to ensure their efficacy in viral identification. The genetic variability observed in Beak and Feather Disease Virus (Circovirus) among samples from different regions highlights the dynamic nature of viral evolution and its implications for diagnostic accuracy. The phylogenetic analysis delineating genetic differences among these strains emphasizes the necessity of precise

molecular tools tailored to detect such variations. The development of a specific protocol that uses primers based on the identified consensus sequences presents a promising approach for enhancing the rapidity, sensitivity and accuracy of PCR-based diagnostics in identifying the novel strain of Circovirus. Furthermore, the discovery of novel viral sequences highlight the importance of comprehensive metagenomic approaches in uncovering previously unknown viral species or strains. The application of these advanced techniques not only broadens our understanding of viral diversity but also highlights the potential existence of novel viruses. The implications of these findings extend beyond diagnostic methodologies. They highlight the necessity of continuous vigilance, adaptive diagnostic strategies, and a holistic approach to comprehend the intricate dynamics of viral infections in companion bird populations. This study offers valuable insights into the challenges associated with viral diagnostics and emphasizes the imperative need for multifaceted approaches to effectively monitor and manage viral outbreaks. In conclusion, the combination of electron microscopy, metagenomic analyses, and refined primer design proved instrumental in unraveling the diverse viral landscape in the investigated bird populations. This study serves as a testament to the necessity of integrating multiple diagnostic modalities and refining molecular tools for accurate and comprehensive viral identification in avian veterinary pathology.

References

1. Shifman O, Cohen-Gihon I, Beth-Din A, Zvi A, Laskar O, Paran N, Epstein E, Stein D, Dorozko M, Wolf D, Yitzhaki S, Shapira SC, Melamed S, Israeli O. Identification and genetic characterization of a novel Orthobunyavirus species by a straightforward high-throughput sequencing-based approach. Sci Rep. 2019 Mar 4;9(1):3398. doi: 10.1038/s41598-019-40036-4. PMID: 30833612; PMCID: PMC6399452.

2. Domingo E, García-Crespo C, Lobo-Vega R, Perales C. Mutation Rates, Mutation Frequencies, and Proofreading-Repair Activities in RNA Virus Genetics. Viruses. 2021 Sep 21;13(9):1882. doi: 10.3390/v13091882. PMID: 34578463; PMCID: PMC8473064.

3.3 Dobrova-Belgrade Virus

DOBV is a RNA virus of the genus Orthohantavirus, a group of pathogens maintained in nature by different species of rodents, but also by small insectivorous mammals and bats. Hantaviruses are important because some of them are zoonotic, meaning that they can transmit between animals and people, causing from mild to serious illnesses. Human cases are reported in different parts of the world and especially in Asia, while they are less frequent in Europe. To date, only a few cases have been reported in Italy, all contracted in neighbouring Slovenia and associated almost exclusively with two viral species: Puumala and Dobrava-Belgrade Orthohantavirus. As both these viruses circulate widely in wild rodents in Croatia and Slovenia, the search for Hantavirus was essential in the samples of the two respective reservoir species, the bank vole (*Myodes glareoulus*) and the yellow-necked mouse (*Apodemus flavicollis*), collected in northeastern Italy. In spring 2021, a large number of individuals belonging to these species were submitted to the IZSVe laboratories where a large spectrum screening was carried out. The identification of Dobrava-Belgrade was conducted in collaboration with the Italian National Health Institute (Istituto Superiore di Sanità).

3.3.1 Identification of Dobrava-Belgrade Virus in *Apodemus flavicollis* from North-Eastern Italy during Enhanced Mortality

The work described in this chapter is taken from: Leopardi S, Drzewnioková P, Baggieri M, Marchi A, Bucci P, Bregoli M, De Benedictis P, Gobbo F, Bellinati L, Citterio C, Monne I, Pastori A, Zamperin G, Palumbo E, Festa F, Castellan M, Zorzan M, D'Ugo E, Zucca P, Terregino C, Magurano F. Identification of Dobrava-Belgrade Virus in Apodemus flavicollis from North-Eastern Italy during Enhanced Mortality. Viruses. 2022 Jun 7;14(6):1241. doi: 10.3390/v14061241. PMID: 35746712; PMCID: PMC9229784.

Abstract

Hantaviruses include several zoonotic pathogens that cause different syndromes in humans, with mortality rates ranging from 12 to 40%. Most commonly, humans get infected through the inhalation of aerosols or dust particles contaminated with

virus-containing rodent excreta. Hantaviruses are specifically associated with the host species, and human cases depend on the presence and the dynamics of reservoir hosts. In this letter, we report the identification of Dobrava-Belgrade virus (DOBV) in the yellow-necked mouse (*Apodemus flavicollis*) from Italy. The virus was detected in the mountainous area of the province of Udine, bordering Austria and Slovenia, during an event of enhanced mortality in wild mice and voles. Despite serological evidence in rodents and humans that suggested the circulation of hantaviruses in Italy since 2000, this is the first virological confirmation of the infection. Phylogenetic analyses across the whole genome of the two detected viruses confirmed the host-specificity of DOBV sub-species and showed the highest identity with viruses identified in Slovenia and Croatia from both A. flavicollis and humans, with no signs of reassortment. These findings highlight the need for ecologists, veterinarians and medical doctors to come together in a coordinated approach in full compliance with the One Health concept.

Introduction

Hantaviruses are enveloped viruses belonging to the family Hantaviridae with segmented, negative sense, single-stranded RNA. The genus Orthohantavirus includes 23 species classified by the International Commission on the Taxonomy of Viruses (ICTV), plus about 30 other viruses for which a complete characterization is not available yet [1]. In contrast to other viruses of this family, hantaviruses are transmitted directly from animal reservoirs with no arthropod vectors involved. Most hantaviruses have been described in rodents (order Rodentia, families Cricetidae, Muridae, Nesomyidae and Sciuridae) and new species are increasingly found in bats (order Chiroptera) and insectivores (order Soricomorpha) [2]. Each species of Hantavirus seems to be specifically associated with its own reservoir host. In Europe, the main reservoir species are the wild yellow-necked mouse (*Apodemus flavicolli*s), the black-striped field mouse (*Apodemus agrarius*) and the bank vole (Myodes glareolus), maintaining two different strains of the Dobrava-Belgrade virus (DOBV) and Puumala virus (PUUV), respectively. In addition, the grey rat (Rattus norvegicus) potentially maintains Seoul virus (SEOV) globally, sporadically reported outside Asia,

including a single identification in France [3]. Hantaviruses, particularly the several species associated with rodents, are known as zoonotic pathogens that cause acute febrile diseases in humans with a mortality rate ranging from 12% for haemorrhagic fever with renal syndrome (HFRS) to 40% hantavirus cardiopulmonary syndrome (HCPS). Although most cases are reported from China, Europe notifies more than 9000 severe cases each year, with several others likely remaining undetected [4,5]. These cases are mainly related to Puumala and Dobrava-Belgrade viruses, with mortality usually lower than 1% and ranging from 5 to 15%, respectively [3,4,6]. As humans are exposed mostly by inhaling viral particles aerosolized from the saliva, urine and/or faeces of rodents, occupations and activities such as working in forestry, farming and road maintenance strongly increase the likelihood of viral infection [5]. In addition, the incidence of human cases seems to be highly correlated with the dynamics of the populations of the reservoir species [6]. In particular, the massive increase in population density caused by an abundance of food during the mast years of beech (Fagus sylvatica) and oak (Quercus robur) trees can enhance the circulation of hantaviruses in the wild during the following year and increase the risk of human exposure [6–9]. The fact that mast-seeding events are linked to environmental factors, such as higher temperatures and seasonal changes in rainfall, suggests that the burden of hantavirus disease might increase as a consequence of climate change [10]. In this context, there is a major need for an integrated approach between ecologists, veterinarians and medical doctors, in full compliance with the One Health concept [11,12]. In Italy, only eight cases of Hantavirus disease have been described, mostly from tourists or transboundary workers, suggesting they contracted the infection abroad. Among these, most had a clinical presentation of HFRS, but three cases related to infection with the new world hantavirus Sin Nombre virus, showed typical signs of HCPS [5]. To date, no native cases of hantavirus disease have been described in Italy, despite the proximity to endemic countries and the presence of both the wild yellow-necked mouse, the black-striped field mouse and the bank vole [5]. However, antibodies neutralizing either PUVV or DOBV have been historically detected in both humans and rodents, particularly in the mountainous areas bordering Austria and Slovenia, suggesting that these pathogens are actively circulating at least in some parts of the

country [5,12,13]. The present communication reports the first molecular finding of Dobrava-Belgrade orthohantavirus in *Apodemus flavicollis* from north-eastern Italy in Summer 2021.

Materials and Methods

Between June and July 2021, the diagnostic laboratories at the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe) received 35 rodent carcasses from the mountainous area bordering Austria and Slovenia, in the province of Udine, that were necropsied. Twenty-one samples consisting of pools of liver and lungs from one to five individuals were sent to the Department of Infectious Disease (IDD) at the Italian National Institute of Health (ISS) for molecular detection. Further analysis was conducted at the IZSVe on individual kidney samples to determine the percentage of positive samples using a nested RT-PCR [14] and to confirm the host species through sequencing of the cytochrome oxidase I, as described elsewhere [15]. All samples positive for the detection of Orthohantavirus spp. were initially characterized through Sanger sequencing before attempting whole genome sequencing of viruses from individual samples. No further analyses were performed on the pooled samples, because they included more than one positive individual that could have challenged correct reconstruction of the sequences. We used a target approach amplifying the three segments using four pairs (one for S, one for M and two for L) of DOBV specific primers modified from Taylor et al., [16] (Table 1). The reaction was performed in a two-step RT-PCR reaction using SuperScript™ IV Reverse Transcriptase (Invitrogen, Waltham, MA, USA) and Platinum™ SuperFi II Green PCR Master Mix (Invitrogen). Positive samples were sequenced using next generation sequencing implemented in Illumina MiSeq with Reagent Nano Kit v2 (500 cycles).

Table 1. Primers used for targetted amplification of the complete genome of DOBV.

Segment	Primer Orientation	Sequence (5′→3′)	bp
S	Sense Antisense	TAGTAGTAKRCTCCCTAAARAGCACTAYAC TAGTAGTAGRCTCCCTAAAAAGACATTCAGGAAGC	1673
M	Sense Antisense	TAGTAGTAGRCTCCGCAAGAAAYAG TAGTAGTAKGCTCCGCARGATATAG	3664
L1	Sense Antisense	TAGTAGTAGACTCCGGAAGAGACARAYTAC CATYCCKACACCRAAAAGAGATGAAC	3253
L2	Sense Antisense	GATAACTCAGCTAARTTYAGAAGRTTCAC TAGTAGTATGCTCCGGAAAATGAAAATRAAT	3351

To obtain complete genomes for the viruses, we first assessed the quality of raw reads using FastQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, accessed on 26 January 2022) and removed the reads related to the Illumina Nextera XT adaptors (Illumina, San Diego, CA, USA) using scythe v0.991 (https://github.com/vsbuffalo/scythe, accessed on 26 January 2022). We then used cutadapt v2.10 to trim the adaptors, designed specifically for Hantavirus, and to filter the raw reads with a Q score below 30 and a length below 80 nucleotides [17]. Complete genomes were generated through a reference-based approach using BWA v0.7.12 (https://github.com/lh3/bwa, accessed on 26 January 2022) [18]. We then used loFreq v2.1.2 (https://github.com/CSB5/lofreq, accessed on 26 January 2022) [19] to call single nucleotide polymorphisms (SNPs) and applied an in-house script to obtain the consensus sequences, setting 10 as the minimum coverage and 50% of allele frequency as the threshold for base calling. Viral sequences obtained in the study were aligned with reference strains using G-INS-1 and default parameters implemented in Mafft [20]. Maximum likelihood (ML) nucleotide phylogenetic trees were inferred using PhyML (version 3.0) implemented in Seaview employing the GTR+Γ4 substitution model, a heuristic SPR branch-swapping algorithm and 100 bootstrap reiterations [21]; the obtained trees were edited online for graphical display using iTOL (version 6.0) [22].

Results

Between June and July 2021, the diagnostic laboratories at the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe) received 35 rodent carcasses collected from the mountainous area bordering Austria and Slovenia, in the province of Udine. Necropsies showed no significant macroscopic lesions. Genetic analyses classified the hosts as *Apodemus flavicollis* (n = 22) or Myodes glareolus (n = 8), whereas five analyses provided no interpretable sequences, likely due to the advanced decay stage of the carcasses. Among the twenty-one liver-lung pools, two tested positive for Orthohantavirus spp. Further analysis conducted on individual kidney samples confirmed the positivity of four individuals of A. flavicollis, with a percentage of positivity of 11.7% in the yellow-necked mice. All voles tested negative. Sanger sequencing of the PCR product attributed the infection to Dobrava-Belgrade virus. Target next generation sequencing performed from individual samples provided the complete genomes of two viruses, with an average coverage ranging respectively between 2967 and 3152. All consensus sequences were submitted to GenBank under accessions n. OM677634-OM677639, raw data were submitted to SRA with accession n. PRJNA807277, and BioSamples SAMN25965836 and SAMN25965837. Maximum likelihood phylogenetic reconstructions using whole genome sequences demonstrated that the Italian DOBV clustered within variants found in A. flavicollis and humans from Croatia and Slovenia in all three segments (Figure 1A–C) and showed no signs of genetic reassortment. Highest identity was found with strains from A. flavicollis in Slovenia (SLO/Af BER, with an identity of 97.5% in L and 97.8% in M) and Croatia (Croatia_Gerovo/Af968/2008, with an identity of 98.7% in S and no sequences available for the remaining segments).

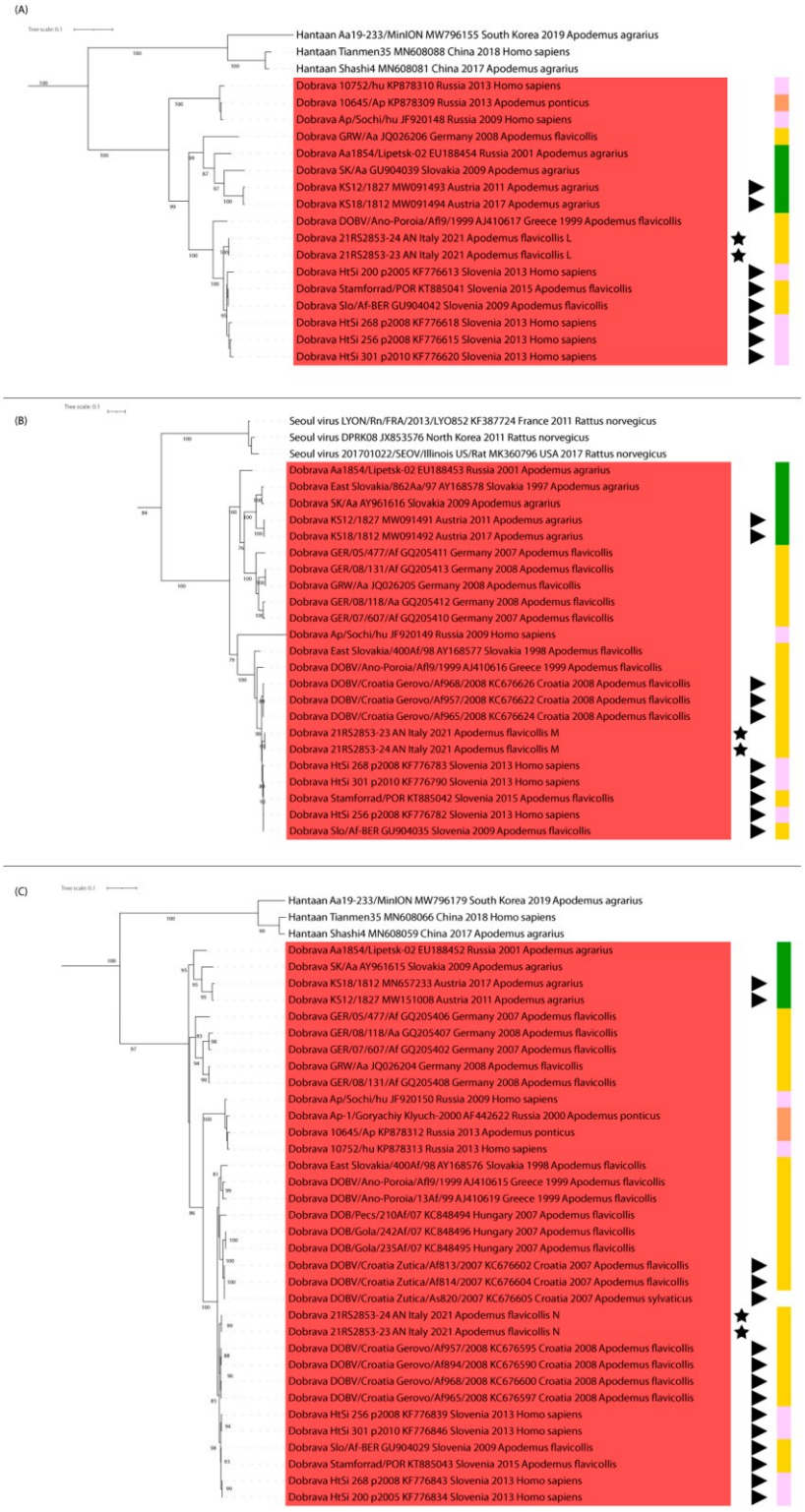


Figure 1. ML phylogenetic trees of the whole sequences for segments L (A), M (B) and S (C). The red box shows the viruses belonging to the specie Dobrava-Belgrade orthohantavirus. Original sequences are marked with a star, and sequences from countries neighbouring the study area are indicated with arrows. Coloured boxes represent the host in which the strains were detected, namely A. flavicollis (yellow), A. ponticus (orange), A. agrarius (green) and human infections (pink).

Discussion and Conclusions

The fact that hantaviruses are specifically associated with a certain host species makes the presence of these reservoirs the most relevant risk factor for human infection. Other than the grey rat, all rodents known as hantavirus reservoirs in Europe are present in north-eastern Italy, including both *A. flavicollis* and *A. agrarius*, that have been found to maintain different sub-species of Dobrava-Belgrade orthohantavirus [23]. In this short report, we describe the identification of DOBV in the yellow-necked mouse from the Italian mountainous area bordering Austria and Slovenia. Phylogenetic analyses confirmed that our sequences grouped only with variants from the yellow-necked mouse, confirming the host-specificity of DOBV sub-species [23]. Despite serological evidence in rodents and humans the suggested the circulation of hantaviruses in Italy since 2000, this is the first virological confirmation of the infection [5,13,24]. Previous studies questioned why hantaviruses should be restricted to a limited portion of the distribution of their hosts [25]. Indeed, our findings suggest that negative findings in certain areas could be due to a lack of or ineffective surveillance rather than the true absence of the pathogen. Indeed, the number of rodents received from north-eastern Italy in 2021 is the highest reported from the laboratory over the past 5 years, with average submissions between 2017 and 2020 being about 15 individuals with 1 single sample analysed from the province of Udine before 2021. Such an increase in rodent diagnostic activity was associated with enhanced mortality reported from north-eastern Italy. In Italy, previous efforts in active surveillance were pivotal in supporting the circulation of both DOBV and PUVV using serology but failed to identify the virus for over 20 years. This is not new in the study of pathogens in their wild reservoir, where they often circulate at low prevalence, making serology the most effective tool for monitoring the infection in the absence of viral detection [26]. Periodic epidemic waves have been largely described for both PUVV and DOBV in the case of relevant variations in population dynamics [6–8,27]. In our case, positive findings could have been favoured by increased prevalence within the reservoir due to increased population density, likely associated with the high production of seed in 2020, reported as a mast year across Europe [28]. This hypothesis is consistent with the enhanced mortality of rodents experienced in the

area that was likely associated with exceeding the carrying capacity for the species rather than to the viral infection. However, it is crucial to notice that all the samples collected in 2012 after a similar event tested negative, which further highlights the challenges related to the study of pathogens in wildlife and might suggest changes in the ecology of hantaviruses in Italy that can be only unravelled through longitudinal active and passive surveillance. In addition, genetic data from our investigation show that the increased mortality was not peculiar to mice but also involved voles, despite the absence of virus detected in M. glareoulus, the reservoir species for PUVV. Since serological evidence suggests that PUVV might circulate in the country as well [13,24], our negative result could indicate that the sample size of bank voles was too low to detect the virus, or it could be related to differences in the dynamics of the two viral species in the area. Numerous studies have revealed that outbreaks of hantavirus disease in humans are often associated with a high prevalence of the pathogen within the reservoir [4,11,12]. In this context, Slovenia reported several cases of PUVV over spring-summer in 2021 and compatible syndromes were also recorded in the provinces of Trieste and Udine, with a single case confirmed by molecular approaches that had contracted the infection in neighbouring Slovenia. Our study, together with previous serological evidence [5,12,13], confirms the circulation of DOBV in the wild yellow-necked mice in north-eastern Italy. Data presented herein encourage both active and passive surveillance plans in wild mice, taking into account population dynamics to better understand hantavirus ecology in thec areas bordering their putative distribution, with the final aim of mitigating the risk of infection in humans. Indeed, this first identification in the reservoir emphasizes the risk of occurrence in humans of mild and severe forms of hantavirus disease in Italy, HFRS and HCPS included. It also reveals the need to conduct further research on hantaviruses in Italy, and to increase the awareness of physicians and local populations regarding the transmission pathways of the viruses and the available diagnostic methods. Equally important is the need to implement and integrate effective surveillance systems.

References

1. ICTV. Hantaviridae ICTV Taxonomy Release 2021; ICTV online 2021. Available online: https://talk.ictvonline.org/taxonomy/ (accessed on 1 June 2022).

2. Milholland, M.T.; Castro-Arellano, I.; Suzán, G.; Garcia-Peña, G.E.; Lee, T.E.; Rohde, R.E.; Alonso Aguirre, A.; Mills, J.N. Global Diversity and Distribution of Hantaviruses and Their Hosts. EcoHealth 2018, 15, 163–208.

3. Heyman, P.; Plyusnina, A.; Berny, P.; Cochez, C.; Artois, M.; Zizi, M.; Pirnay, J.P.; Plyusnin, A. Seoul hantavirus in Europe: First demonstration of the virus genome in wild Rattus norvegicus captures in France. Eur. J. Clin. Microbiol. Infect. Dis. 2004, 23, 711–717.

4. Avšicˇ-Županc, T.; Saksida, A.; Korva, M. Hantavirus infections. Clin. Microbiol. Infect. 2019, 21, e6–e16.

5. Riccò, M.; Peruzzi, S.; Ranzieri, S.; Balzarini, F.; Valente, M.; Marchesi, F.; Bragazzi, N.L. Hantavirus infections in Italy: Not reported doesn't mean inexistent. Acta Biomed. Atenei Parm. 2021, 92, 2021324.

6. Vapalahti, O.; Mustonen, J.; Lundkvist, Å.; Henttonen, H.; Plyusnin, A.; Vaheri, A. Hantavirus infections in Europe. Lancet Infect. Dis. 2003, 3, 653–661.

7. Tersago, K.; Verhagen, R.; Vapalahti, O.; Heyman, P.; Ducoffre, G.; Leirs, H. Hantavirus outbreak in Western Europe: Reservoir host infection dynamics related to human disease patterns. Epidemiol. Infect. 2011, 139, 381–390.

8. Reusken, C.; Heyman, P. Factors driving hantavirus emergence in Europe. Curr. Opin. Virol. 2013, 3, 92–99.

9. Prist, P.R.; Prado, A.; Tambosi, L.R.; Umetsu, F.; de Arruda Bueno, A.; Pardini, R.; Metzger, J.P. Moving to healthier landscapes: Forest restoration decreases the abundance of Hantavirus reservoir rodents in tropical forests. Sci. Total Environ. 2021, 752, 141967.

10. Mittler, E.; Dieterle, M.E.; Kleinfelter, L.M.; Slough, M.M.; Chandran, K.; Jangra, R.K. Hantavirus entry: Perspectives and recent advances. Adv. Virus Res. 2019, 104, 185–224.

11. Yon, L.; Duff, J.P.; Ågren, E.O.; Erdélyi, K.; Ferroglio, E.; Godfroid, J.; Hars, J.; Hestvik, G.; Horton, D.; Kuiken, T.; et al. Recent changes in infectious diseases in european wildlife. J. Wildl. Dis. 2019, 55, 3–43.

12. Rizzoli, A.; Tagliapietra, V.; Rosà, R.; Hauffe, H.C.; Marini, G.; Voutilainen, L.; Sironen, T.; Rossi, C.; Arnoldi, D.; Henttonen, H. Recent increase in prevalence of antibodies to Dobrava-Belgrade virus (DOBV) in yellow-necked mice in northern Italy. Epidemiol. Infect. 2015, 143, 2241–2244.

13. Kallio-Kokko, H.; Laakkonen, J.; Rizzoli, A.; Tagliapietra, V.; Cattadori, I.; Perkins, S.E.; Hudson, P.J.; Cristofolini, A.; Versini, W.; Vapalahti, O.; et al. Hantavirus and arenavirus antibody prevalence in rodents and humans in Trentino, Northern Italy. Epidemiol. Infect. 2006, 134, 830–836.

14. Klempa, B.; Fichet-Calvet, E.; Lecompte, E.; Auste, B.; Aniskin, V.; Meisel, H.; Denys, C.; Koivogui, L.; Ter Meulen, J.; Krüger, D.H. Hantavirus in African wood mouse, Guinea. Emerg. Infect. Dis. 2006, 12, 838–840.

15. De Benedictis, P.; Leopardi, S.; Markotter, W.; Velasco-Villa, A. The importance of accurate host species identification in the framework of rabies surveillance, control and elimination. Viruses 2022, 14, 492.

16. Taylor, M.K.; Williams, E.P.; Wongsurawat, T.; Jenjaroenpun, P.; Nookaew, I.; Jonsson, C.B. Amplicon-Based, Next-Generation Sequencing Approaches to Characterize Single Nucleotide Polymorphisms of Orthohantavirus Species. Front. Cell. Infect. Microbiol. 2020, 10, 565591.

17. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 2011, 17, 10–12.

18. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr. Protoc. Bioinform. 2013, 43, 11.10.1–11.10.33.

19. Wilm, A.; Aw, P.P.K.; Bertrand, D.; Yeo, G.H.T.; Ong, S.H.; Wong, C.H.; Khor, C.C.; Petric, R.; Hibberd, M.L.; Nagarajan, N. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 2012, 40, 11189–11201.

20. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002, 30, 3059–3066.

21. Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.F.; Guindon, S.; Lefort, V.; Lescot, M.; et al. Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008, 36, 465–469.

22. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016, 44, W242–W245.

23. Klempa, B.; Schmidt, H.A.; Ulrich, R.; Kaluz, S.; Labuda, M.; Meisel, H.; Hjelle, B.; Krüger, D.H. Genetic Interaction between Distinct Dobrava Hantavirus Subtypes in Apodemus agrarius and A. flavicollis in Nature. J. Virol. 2003, 77, 804–809.

24. Cosseddu, G.M.; Sozio, G.; Valleriani, F.; Di Gennaro, A.; Pascucci, I.; Gavaudan, S.; Marianneau, P.; Monaco, F. Serological Survey of Hantavirus and Flavivirus among Wild Rodents in Central Italy. Vector-Borne Zoonotic Dis. 2017, 17, 777–779.

25. Olsson, G.E.; Leirs, H.; Henttonen, H. Hantaviruses and their hosts in Europe: Reservoirs here and there, but not everywhere? Vector-Borne Zoonotic Dis. 2010, 10, 549–561.

26. Gilbert, A.T.; Fooks, A.R.; Hayman, D.T.S.; Horton, D.L.; Müller, T.; Plowright, R.; Peel, A.J.; Bowen, R.; Wood, J.L.N.; Mills, J.; et al. Deciphering serology to understand the ecology of infectious diseases in wildlife. EcoHealth 2013, 10, 298–313.

27. Heyman, P.; Thoma, B.R.; Marié, J.-L.; Cochez, C.; Essbauer, S.S. In search for factors that drive hantavirus epidemics. Front. Physiol. 2012, 3, 237.

28. Chianucci, F.; Tattoni, C.; Ferrara, C.; Ciolli, M.; Brogi, R.; Zanni, M.; Apollonio, M.; Cutini, A. Evaluating sampling schemes for quantifying seed production in beech (Fagus sylvatica) forests using ground quadrats. For. Ecol. Manag. 2021, 493, 119294.

3.4 Avian Influenza Virus

The molecular epidemiology of Avian Influenza Viruses (AIVs) delves into the genetic composition and evolutionary patterns of these viruses, elucidating their spread, diversification, and impact on avian populations. By analyzing the genetic sequences of AIVs, researchers trace their transmission pathways, identify emerging strains, and understand the potential for interspecies transmission, including zoonotic threats [1]. This field employs genomic techniques to characterize various AIV subtypes, such as H5 and H7, and their subclades, aiding in surveillance, outbreak prediction, and vaccine development [2]. Understanding the genetic makeup and evolutionary dynamics of AIVs is crucial for devising effective control strategies to mitigate their impact on both poultry and public health.

3.4.1 The three epidemic waves in Italy, 2020-2023 of HPAI A(H5) viruses: phylogenetic networks and Bayesian phylogenetic analysis to identify genetic diversity

Abstract

The study focuses on the molecular epidemiology and spread dynamics of Highly Pathogenic Avian Influenza (HPAI) A(H5) viruses, particularly during the 2020-2023 epidemic waves in Italy. The investigation encompasses 547 complete viral genomes collected from poultry, wild birds, and wild carnivores, sequenced using the Illumina MiSeq platform. Phylogenetic analyses through IQtree v1.6.6 and BEAST v1.10.4 identified thirteen distinct genotypes resulting from reassortment events. Notably, the H5N8, H5N5, and H5N1 subtypes circulated, with the C genotype (H5N1-A/Eurasian Wigeon/Netherlands/1/2020-like) predominant during the second epidemic wave. The third wave saw a prevalence of genotypes AB and BB, notably linked to significant poultry outbreaks and mass mortality in wild birds, including foxes. Phylodynamics and network analyses revealed multiple introductions from wild birds to domestic poultry, with turkeys identified as key contributors to inter-species transmission. Verona and Mantua provinces were highlighted as principal sources of viral spread, emphasizing the importance of geographical patterns. Despite sporadic mutations associated with mammalian

adaptation, the continuous emergence of reassortant viruses poses significant concerns for animal and public health within Italy and Europe. This comprehensive genetic investigation elucidates the varied biological characteristics and behaviors of multiple viral genotypes, emphasizing the ongoing threat and need for vigilant surveillance and control measures.

Introduction

Avian Influenza viruses (AIVs) are part of the Alphainfluenzavirus genus in the Orthomyxoviridae family, characterized by their segmented negative-sense RNA structure. Their impact varies based on pathogenicity, that can be either low (LP) or highly pathogenic (HP), particularly in chickens (Gallus gallus domesticus) and other related species. Typically, highly pathogenic strains arise from low pathogenic ones initially introduced by wild aquatic birds into domestic poultry, historically leading to outbreaks (Swayne and Suarez, 2000). However, the H5 lineage HPAIVs, specifically A/goose/Guangdong/1/96 (Gs/GD), have been a persistent threat to both domestic poultry and wild bird populations since 1996. Initially observed in Asian domestic ducks and geese, these viruses spilled over into migratory wild birds, facilitating their dissemination across various continents and the emergence of multiple variants [3]. The World Health Organization (WHO) has established a unified nomenclature system for the hemagglutinin (HA) gene of this H5 Gs/GD lineage, aiding in tracking its phylogenetic evolution (World Health Organization, 2020). The current H5 2.3.4.4 clade exhibits several defined subclades, showcasing significant reassortment with low pathogenic strains from wild birds across multiple flyways. In October 2020, a distinct H5N1 2.3.4.4b virus 4b of the H5 Goose/Guangdong/96 lineage was first detected in wild birds in the Netherlands [4]. Since that time, several European countries have reported H5N1 detections in wild birds prior and up to December 2021 [5]. This virus has a devastating effect on domestic and wild birds and it has repeatedly emerged in Italy. Since 2020, three consecutive epidemic waves have occurred, with the last ones (2021-2022 and 2022-2023) having a devastating impact on poultry industry in the densely populated areas of northern Italy.

Methodology

To elucidate the diffusion dynamics and the genetic diversity of the HPAI A(H5) viruses in Italy, phylodynamics analyses were performed on the complete genomes of 547 viruses identified in poultry/captive birds, wild birds and wild carnivores. Samples were collected from November 2020 to April 2023 and sequences were generated using an Illumina MiSeq platform. Phylogenetic and evolutionary analyses were performed using IQtree v1.6.6, BEAST v1.10.4, and Network 10. The identification of the viruses genotype was based on the phylogenetic trees topology (IQtree 1.6.6). Monophyletic clusters supported by ultrafast bootstrap values > 90 were identified for all the trees. A reference sequence, the potential progenitor of the European HPAI H5 viruses, was selected within each cluster and associated to an ID number. The genotype was assigned to each virus based on the combination of the eight ID (one for each gene segment), defined from the tree topology. Bayesian discrete analysis (BEAST 1.10.4, strict clock, coalescent constant population size model, MCMC chains run for 50 million iterations) of the second wave (October 2021-March 2022) was performed to obtain the Maximum Clade Credibility Tree and using Location and Species as traits. Genetic network of the complete genome of H5N1 viruses belonging to genotype C and BB and collected during the second and third epidemic wave respectively was generated using the Median Joining method implemented in Network 10 for the eight concatenated gene segments of non-reassortant H5N1 viruses.

Results

Three HPAI A(H5) subtype viruses (H5N8, H5N5, H5N1) and thirteen different genotypes, originating from reassortment events, were identified since November 2020 (Figure 1). Few cases, in wild (mainly waterfowls) and domestic birds, were recorded during the first wave (November 2020-February 2021) with the A genotype (H5N8-A/duck/Chelyabinsk/1207-1/2020-like) as the main circulating one (Figure 2i). A high circulation of the virus in poultry farms characterized the second epidemic wave (October 2021-March 2022) (Figure 2ii). During this wave, the most widespread genotype was the C genotype (H5N1-A/Eurasian Wigeon/Netherlands/1/2020-like) (poultry, N=281; wild birds, N=14), which

started to circulate during the first wave and became predominant in the second wave (Figure 1). The majority of the characterized viruses marking the third epidemic wave (since September 2022) belong to the AB (H5N1-A/duck/Saratov/29-02/2021-like) and BB (H5N1-A/Herring gull/France/22P015977/2022-like) genotypes (Figure 1). Genotype AB was identified in most of the poultry outbreaks recorded in September-December 2022 and in wild anseriformes (Figure 2iii), whereas genotype BB was associated to the mass mortality events reported in black-headed gulls in northern Italy starting from December 2022. The BB genotype, as in other European countries, was identified also in two foxes found sick or dead at the beginning of April 2023 (Figure 2iii).

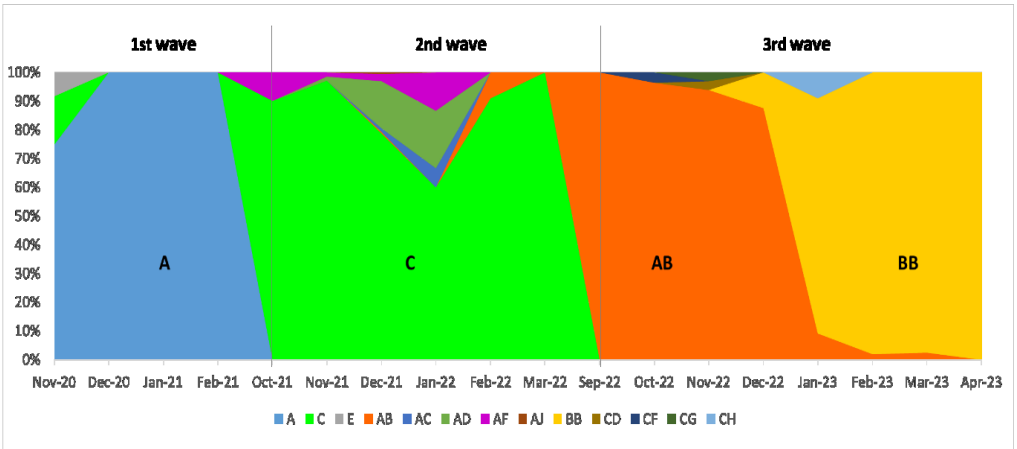


Figure 1. Genetic diversity over time. Genotypes A (H5N8) in the first wave, C (H5N1) in the second wave, AB (H5N1) and BB (H5N1) in the third wave were the most prevalent.

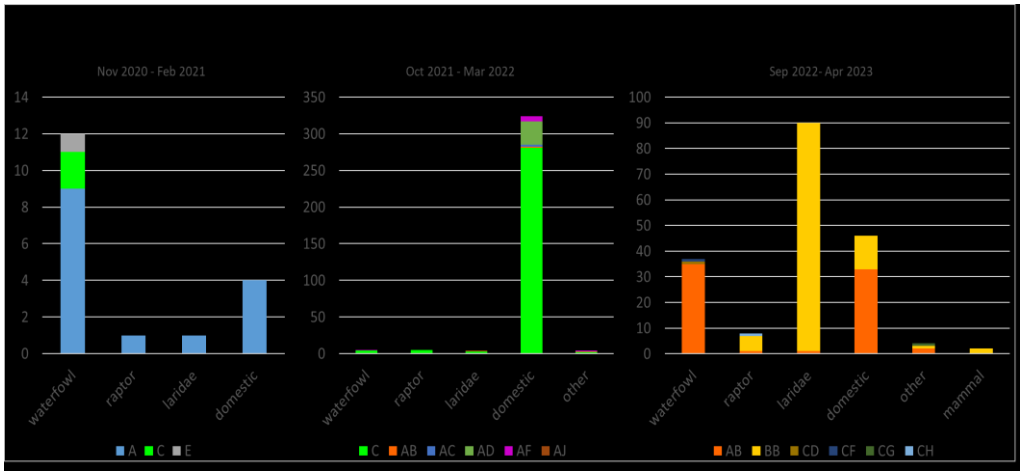


Figure 2. Distribution of genotypes by species for each of the 3 epidemic waves. **i)** first wave**. ii)** second wave**. iii)** third wave.

Since the highest number of cases was recorded during the second epidemic wave (2021-2022), we performed the Phylodynamics analysis to understand the viral spread among the different locations and species involved. From the pylogeographic analyses, many distinct introductions into poultry from wild birds were recorded and also the occurrence of several secondary outbreaks (Figure 3). These results were also confirmed by the network analysis (Figure 5). The province of Verona (VR) and Mantua (MN) were identified as the main source of the virus for the neighboring areas (Figure 3).

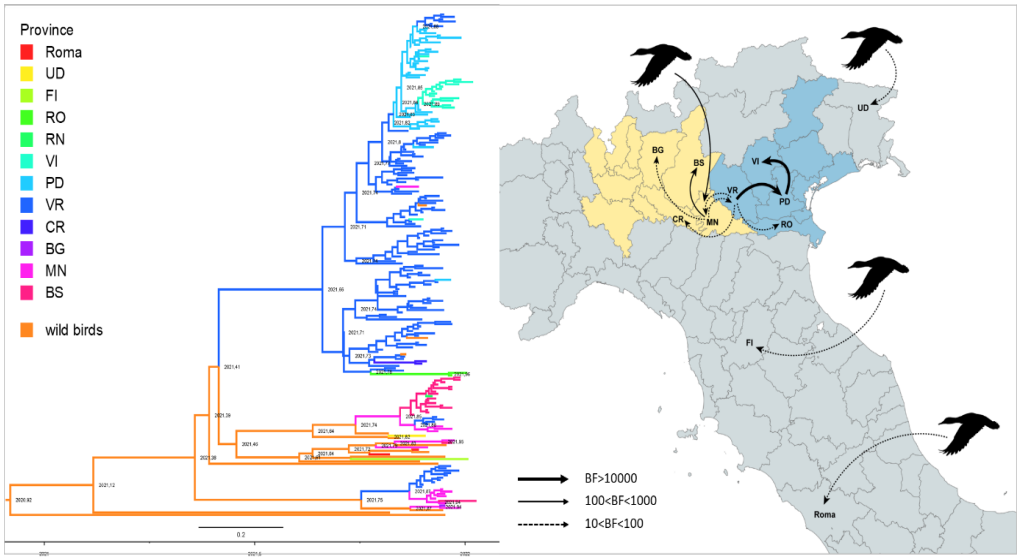


Figure 3. Maximum Clade Credibility Tree obtained from the Bayesian discrete phylogeographic analysis of the second wave (October 2021-March 2022), Each Italian province is associated to a different color. The map shows statistically supported non-zero rates for the HA gene segment of 2.3.4.4B H5 viruses. The thickness of the lines is proportional to the relative strength by which rates are supported. The Lombardy region is in yellow and the Veneto region in blue.

From the phylodynamic analysis, using species as discrete trait, turkeys have been identified as the source that mostly contributed to the viral spread among the different species (Figure 4).

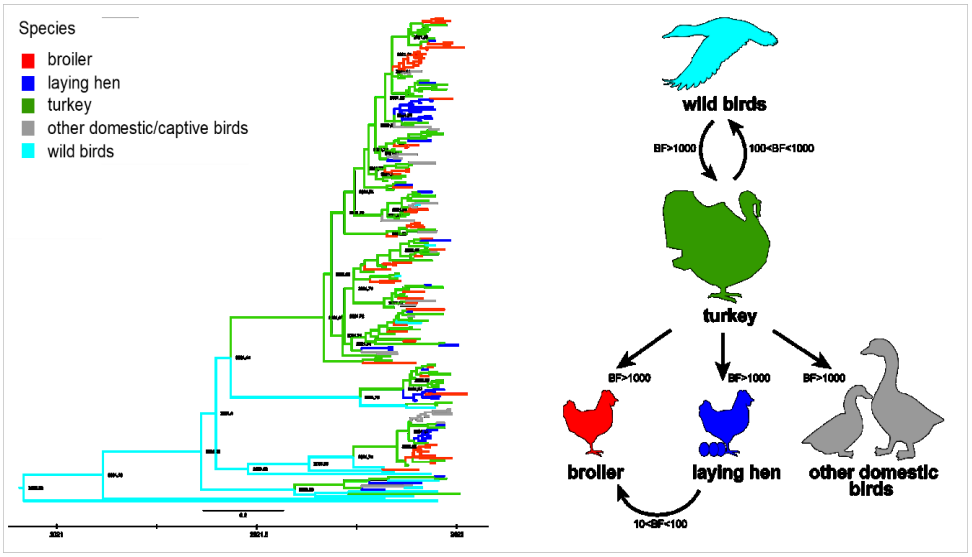


Figure 4. Maximum Clade Credibility Tree obtained from the Bayesian discrete analysis of the second wave (October 2021-March 2022), showing the species contribution to the viral spread.

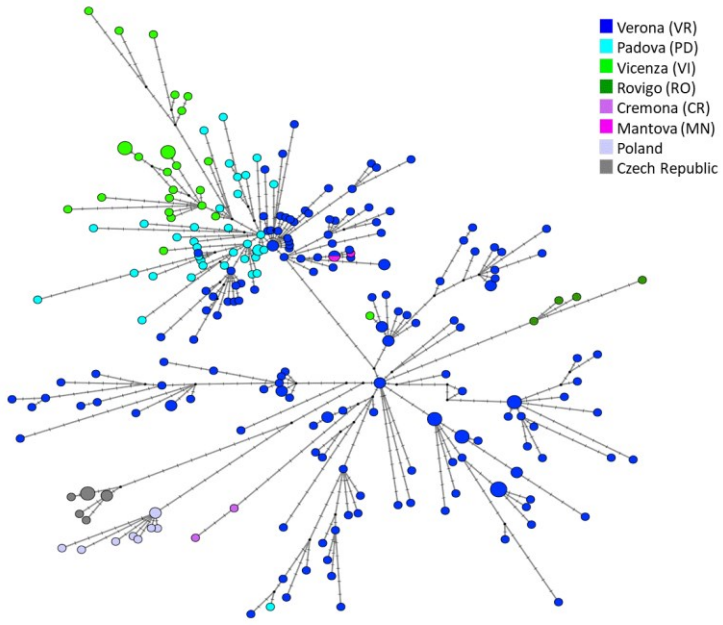


Figure 5. Genetic network of the complete genome of H5N1 viruses belonging to genotype C and collected during the second epidemic wave. Each circle represents one viral variant; the size of the circles is proportional to the number of viruses sharing the same genic composition; the length of the branches is proportional to the number of nucleotide differences between two variants.

During the third epidemic wave, the most prevalent was BB genotype, which caused nine out of the ten poultry outbreaks detected between March and April 2023. From the network analysis performed on the viruses belonging on this genotype (Figure 6), we observed a single introduction from wild birds to domenstic birds.

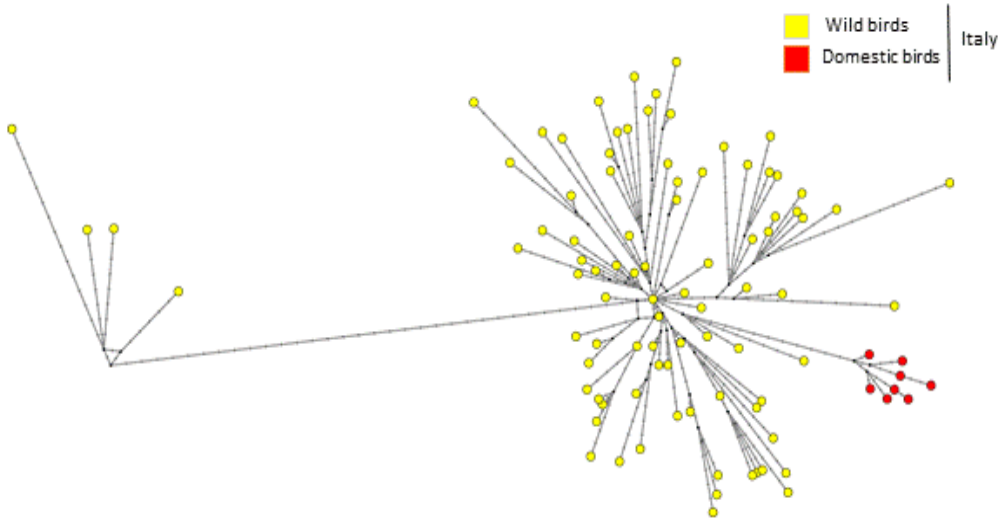


Figure 6. Genetic network of the complete genome of Italian H5N1 viruses belonging to genotype BB and collected during the third epidemic wave. It is generated using the Median Joining method implemented in Network 10 for the eight concatenated gene segments of non-reassortant H5N1 viruses; each circle represents one viral variant; the size of the circles is proportional to the number of viruses sharing the same genic composition; the length of the branches is proportional to the number of nucleotide differences between two variants.

Discussion

In conclusion, the genetic investigation of the HPAI A(H5N1) viruses responsible of the 2020-2023 epidemic in Italy revealed the incursions of multiple viral genotypes with different biological characteristics and behaviors in domestic and wild birds. C genotype accounted for most of the outbreaks in poultry (2021-2022). Mutations associated with mammalian adaptation were only sporadically identified in the analysed viruses from avian hosts. However, the continuous circulation and emergence of new reassortant viruses in Italy and Europe raises concern for animal and public health.

References

1. Kim SM, Kim YI, Pascua PN, Choi YK. Avian Influenza A Viruses: Evolution and Zoonotic Infection. Semin Respir Crit Care Med. 2016 Aug;37(4):501-11. doi: 10.1055/s-0036-1584953. Epub 2016 Aug 3. PMID: 27486732; PMCID: PMC7171714.

2. Charostad J, Rezaei Zadeh Rukerd M, Mahmoudvand S, Bashash D, Hashemi SMA, Nakhaie M, Zandi K. A comprehensive review of highly pathogenic avian influenza (HPAI) H5N1: An imminent threat at doorstep,Travel Medicine and Infectious Disease, Volume 55, 2023,102638, ISSN 1477-8939, https://doi.org/10.1016/j.tmaid.2023.102638. (https://www.sciencedirect.com/science/article/pii/S1477893923000984)

3. Youk S, Torchetti MK, Lantz K, Lenoch JB, Killian ML, Leyson C, Bevins SN, Dilione K, Ip HS, Stallknecht DE, Poulson RL, Suarez DL, Swayne DE, Pantin-Jackwood MJ. H5N1 highly pathogenic avian influenza clade 2.3.4.4b in wild and domestic birds: Introductions into the United States and reassortments, December 2021-April 2022. Virology. 2023 Oct;587:109860. doi: 10.1016/j.virol.2023.109860. Epub 2023 Aug 2. PMID: 37572517.

4. Lewis NS, Banyard AC, Whittard E, Karibayev T, Al Kafagi T, Chvala I, Byrne A, Meruyert Akberovna S, King J, Harder T, Grund C, Essen S, Reid SM, Brouwer A, Zinyakov NG, Tegzhanov A, Irza V, Pohlmann A, Beer M, Fouchier RAM, Akhmetzhan Akievich S, Brown IH. Emergence and spread of novel H5N8, H5N5 and H5N1 clade 2.3.4.4 highly pathogenic avian influenza in 2020. Emerg Microbes Infect. 2021 Dec;10(1):148-151. doi: 10.1080/22221751.2021.1872355. PMID: 33400615; PMCID: PMC7832535.

5. EFSA, Avian influenza overview September – December 2021 https://doi.org/10.2903/j.efsa.2021.7108

3.4.2 Outbreak of highly pathogenic avian influenza A(H5N1) clade 2.3.4.4b virus in cats, Poland, June to July 2023

The work described in this chapter is taken from: Domańska-Blicharz K,Świętoń E, Świątalska A, Monne I, Fusaro A, Tarasiuk K, Wyrostek K, Styś-Fijoł N, Giza A, Pietruk M, Zechchin B, Pastori A, Adaszek Ł,Pomorska-Mól M, Tomczyk G, Terregino C, Winiarczyk S. Outbreak of highly pathogenic avian influenza A(H5N1) clade 2.3.4.4b virus in cats, Poland, June to July 2023. Euro Surveill. 2023 Aug;28(31):2300366. doi: 10.2807/1560-7917.ES.2023.28.31.2300366. PMID: 37535474; PMCID:PMC10401911.

Background: Over a 3-week period in late June/early July 2023, Poland experienced an outbreak caused by highly pathogenic avian influenza (HPAI) A(H5N1) virus in cats. **Aim**: This study aimed to characterise the identified virus and investigate possible sources of infection. **Methods**: We performed next generation sequencing and phylogenetic analysis of detected viruses in cats. **Results**: We sampled 46 cats, and 25 tested positive for avian influenza virus. The identified viruses belong to clade 2.3.4.4b, genotype CH (H5N1 A/ Eurasian wigeon/Netherlands/3/2022-like). In Poland, this genotype was responsible for several poultry outbreaks between December 2022 and January 2023 and has been identified only sporadically since February 2023. Viruses from cats were very similar to each other, indicating one common source of infection. In addition, the most closely related virus was detected in a dead white stork in early June. Influenza A(H5N1) viruses from cats possessed two amino acid substitutions in the PB2 protein (526R and 627K) which are two molecular markers of virus adaptation in mammals. The virus detected in the white stork presented one of those mutations (627K), which suggests that the virus that had spilled over to cats was already partially adapted to mammalian species. **Conclusion**: The scale of HPAI H5N1 virus infection in cats in Poland is worrying. One of the possible sources seems to be poultry meat, but to date no such meat has been identified with certainty. Surveillance should be stepped up on poultry, but also on certain species of farmed mammals kept close to infected poultry farms.

Introduction

Europe, and more recently the Americas, have been experiencing highly pathogenic avian influenza (HPAI) virus infections of an unprecedented scale in wild birds and poultry [1,2]. The most important variations in the course of the 2021/22 epidemic season were the lack of quiescence of the infections during the summer and their continuation into the 2022/23 season. Moreover, the currently circulating Eurasian HPAI A subtype H5N1 virus has been identified in a wide range of wild mammals as foxes, lynxes, skunks, raccoons, bears, otters, polecats, badgers, ferrets, pumas, panthers, oposums, seals, porpoises and sea lions, as well as dolphins. In the vast majority of these cases, mammals, including marine mammals such as seals and porpoises, were carnivorous predators [1]. During the 2022/23 epidemic season in Poland (from 21 September 2022 until 10 July 2023), HPAI H5N1 virus was detected in 93 outbreaks in poultry and 147 outbreaks in wild birds. The first outbreak in Poland occurred on 21 September 2022 in Łódź voivodeship. This single outbreak was followed by a 2.5-month break, as the next HPAI virus infection was detected in early December, followed by 36 further outbreaks in poultry. The infections continued in 2023, with a total of 39 outbreaks in January, 17 in February, then only two outbreaks in March, followed by one outbreak in May and the most recent one in a backyard flock on 1 July. The vast majority of outbreaks were caused by HPAI H5N1 virus CH genotype. The CH (H5N1 A/ EurasianWigeon/Netherlands/3/2022-like) genotype in Poland was identified for the first time in mid-December 2022 and has since then been responsible for 58% of cases in domestic birds and 30% of cases in wild birds (mainly waterfowl) [1]. Between December 2022 and January 2023, it was responsible for several outbreaks in poultry mainly in the Wielkopolskie region, while since February 2023, this genotype has only been identified sporadically in the country (n = 5) when it was replaced with the gull-adapted BB (H5N1-A/gull/ France/22PO15977/2022-like) genotype [3]. The BB genotype was identified in the two most recent poultry outbreaks: in May in a turkey flock and in June in a backyard flock. In wild birds, a total of 147 outbreaks were recorded, 12 in autumn 2022 and 135 in 2023. Outbreaks up to February 2023 were mainly recorded in wild birds of the order Anseriformes (n = 37) such as mute swans, graylag and bean goose, while later

outbreaks included only black-headed gulls and other birds of the Laridae family (n = 96) and were caused by infection with the HPAI H5N1 virus BB genotype. On 4 June 2023, a dead white stork was found, infected with HPAI virus of CH (H5N1_A/Eurasian_Wigeon/ Netherlands/3/2022-like) genotype, and on 20 June 2023 a mute swan where the amount of retrieved virus material was too small to perform whole genome sequencing. In addition to the CH and BB genotypes, the AB genotype (A/duck/Saratov/29–02/2021-like) was also detected in poultry and wild birds between December 2022 and March 2023, but to a lesser extent. Here, we describe an outbreak caused by HPAI H5N1 virus detected in 25 cats from different regions of Poland over a period of 3 weeks between the end of June and the beginning of July 2023.

Methods

Setting

A national programme aimed at detection of HPAI virus infections in poultry and wild birds has been conducted in Poland for 20 years, since 2003. The obligation to conduct these surveys, as well as the detailed manner and mode of control of this disease, derives from the provisions of Commission Delegated Regulation (EU) 2020/689 of 17 December 2019 suplementing Regulation (EU) 2016/429 of the European Parliament and of the Council [4]. It includes both passive surveillance of HPAI virus in poultry and wild birds (testing of dead and sick birds) and active surveillance testing of live poultry from species that do not show typical symptoms following HPAI virus infection, as well as active surveillance of low pathogenic avian influenza virus (serology). It is mandatory that all detected infections are notified, sequencing is performed on representative samples, and the sequences are then forwarded to the European Union Reference Laboratory (EURL) for Avian Influenza in Italy (Istituto Zooprofilattico Sperimentale delle Venezie) [5]. Apart from the aforementioned monitoring and veterinary surveillance in slaughterhouses for the clinical health status of animals, there are no other measures in place to prevent avian influenza entering the food chain. There is no monitoring of avian influenza infections in mammals in Poland.

Epidemiological investigation

Information about a highly fatal disease in cats with respiratory and nervous system signs began to circulate in social media and among cat lovers in mid-June 2023. At the time, all of the information came from the media; there was no complete knowledge, let alone certainty, of what had actually happened. In order to systematise the research and to obtain detailed data on cats, we developed a questionnaire for practitioners at veterinary clinics, in which we asked about various aspects of cat health, behaviour, nutrition, clinical manifestations or gross lesions. Information was also given on which samples to take from dead (whole cats/organ fragments, especially brain and respiratory tissues) or live (throat and rectal swabs) cats. Instructions on how to send samples for testing were also given (refrigerated if anticipated transport was 2–3 days, frozen if longer). A few veterinary clinics sent samples to the laboratory together with questionnaires; such samples were in addition sent directly by cat owners. The information received in the questionnaires as well as interviews with cat owners were thoroughly analysed.

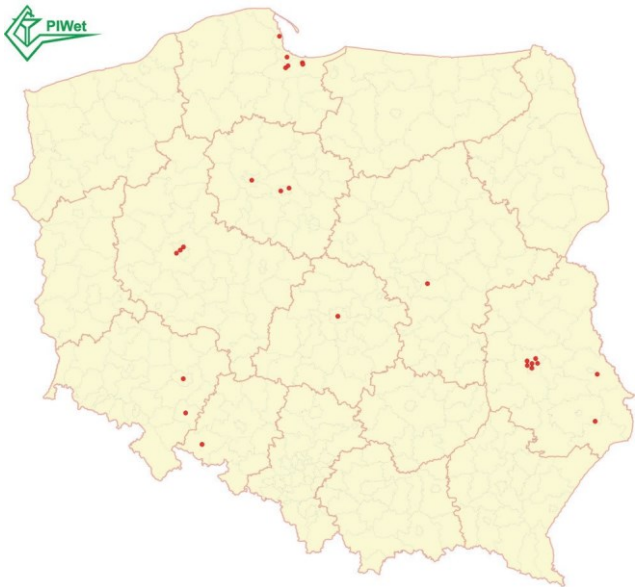


Figure 1. Locations of cats infected with highly pathogenic avian influenza A(H5N1) virus, Poland, June–July 2023 (n = 25) marked as red dots.

Laboratory investigation

Swabs were immersed in phosphate buffered saline (PBS) and organs were homogenised in PBS to 20% weight/volume suspensions. The RNA was extracted using IndiMag Pathogen Kit in IndiMag 48s (Indical Bioscience). Samples were tested with real-time RT-PCR targeting the M gene of influenza A virus [6], and subtyping was performed with primers and probes specific for H5 and N1 genes [7,8]. From each cat, the samples with the highest viral load (n = 23) were subjected to whole genome sequencing as previously described [9]. Libraries were sequenced in NextSeq 550 or iSeq 100 (Illumina). Consensus sequences were generated and compared with previously obtained HPAI H5N1 virus genomes from Poland, as well as sequences from other European countries available in GISAID. Maximum likelihood phylogenetic trees were generated for each genome segment using IQ-TREE [10,11] with 1,000 ultra-fast bootstrap replications and visualised with FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree). The phylogenetic network was generated using the median joining method implemented in NETWORK 10.2.0.0 [12] for the eight concatenated gene segments of all non-reassortant HPAI H5N1 viruses collected in Europe that belonged to the same genotype of the viruses from cats (genotype CH). This allowed us to visualise how the viral genomes are connected on the basis of their genetic similarity.

Results

Epidemiological investigation

The disease was reported from different locations across the vast area of Poland, mainly in big cities (Figure 1). The samples from cats were collected over a time period of a few days, the earliest on 14 June and the latest on 23 June 2023 (Table 1). Analysis of the epidemic curve over time suggests that the peak of infections was around 18–20 June (Figure 2). Of the 46 cats tested during 14–23 June, 25 were infected. The infected cats were of a wide range of ages between 6 weeks and 12 years, different sex and breeds. They were fed a variety of foods, 12 of 25 had fresh, raw poultry meat in their diet, two were fed a biologically appropriate raw food (BARF) diet, and for 11 cats no such data were obtained. Regarding contact with

the environment, six cats were kept only indoors or had limited access to the outside environment, four cats went outside, two were backyard cats and for 13 there was no such information. The questionnaires received showed that for cats fed raw, fresh poultry meat, the first clinical signs (apathy, fever) usually appeared a few days (2–3 days) after meat consumption. The course of the infection ended with the death of the cat (either a natural death or euthanasia). The course of the disease was similar in all cases: loss of appetite, apathy, hypersalivation, fever, dyspnoea (shallow and accelerated breathing), hard and painful abdomen, sometimes incontinence of urine, reddened mucous membranes, trismus, followed by nervous symptoms such as epileptic seizures, increased muscle tension and sometimes stiffness of the limbs. On clinical examination, an exacerbated vesicular murmur and constricted pupils unresponsive to light were noted. Attempts to treat pneumonia with various antibiotics were not successful and the conditions of the animals worsened after 1 or 2 days. In most cases, the animals were euthanised. Post-mortem examination in 11 cats revealed the presence of lesions in every organ, which were congested, sometimes swollen with the presence of bloody fluid.

Laboratory investigation

Due to emerging nervous signs, the cats were initially examined for rabies, and then the samples were screened for the presence of avian influenza virus. Influenza A (H5N1) virus was identified and was present in some organs and tissues with a very high viral load (based on quantification cycle values) (Table 1).

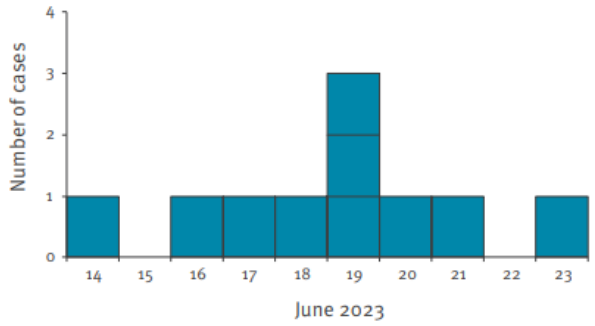


Figure 2. Epidemic curve of cats infected with highly pathogenic avian influenza A(H5N1) virus, Poland, June 2023 (n = 25).

Since the HPAI H5N1 virus outbreak detected in cats cannot be discussed without outlining the influenza situation in poultry and wild birds in Poland, we also present a map with the locations where these infections occurred in Figure 3, in addition to the data in the Introduction section. The viral sequences obtained from representative bird samples were also submitted to GISAID. At the time of writing this article, we had performed whole genome sequencing on 19 of the HPAI H5N1 virus-positive samples collected from 25 cats (Table 1); sequencing of the remaining positive samples is planned/underway. Their topology indicates that the HPAI H5N1 viruses collected from the cats belonged to the CH (H5N1_A/ Eurasian_Wigeon/Netherlands/3/2022-like) genotype. The sequences of the viruses from cats are highly related to each other and clustered with a virus of the same genotype detected in a white stork in Poland on 4 June 2023 (A/white_stork/Poland/MB244/2023). No clustering by geographical region was observed for the HPAI H5N1 viruses collected from the 25 cats (Figure 4). The HPAI H5N1 viruses collected from these cats differed by 1–12 nucleotides and by 0–8 amino acids (Table 2). Compared with the most closely related viruses collected from birds (A/white_ stork/Poland/MB244/2023), the viruses from cats showed at least four nucleotide and three amino acid mutations (PB2-N82T, PB2-K526R and PB1-P64S).

Molecular markers of virus adaptation in mammals

Amino acid differences identified in the viral proteins of the analysed viruses are detailed in Table 2. In particular, all the viruses possess mutation PB2-E627K, which is an important molecular marker of virus adaptation to mammals [13-23]. The same mutation was present in the H5N1 virus detected in the white stork at the beginning of June (A/white_stork/Poland/MB244/2023). This mutation has rarely been observed in H5N1 viruses collected from birds during the ongoing epidemic wave (0.16% of the viruses from birds) but has frequently been acquired by the virus after transmission to mammals (17% of viruses from mammals) [1]. Moreover, all the viruses from cats possessed mutation PB2-K526R, which is another marker of mammalian adaptation. Of note, the Polish H5N1 viruses from cats gained dual 526R/627K substitutions in the PB2 protein and are the only ones

characterised during the 2.3.4.4b world-wide current epidemic wave showing both mutations.

Discussion

We have recently witnessed changes in the course of the HPAI epidemic, i.e. the spread to the Americas, the persistence of the virus all year round, a greater range of avian hosts, as well as the number of mammals infected. The above facts demonstrate worrying changes in the biological properties of circulating HPAI H5N1 viruses, which confer to the viruses the ability to infect a greater number of wild bird species that are not normally susceptible, expanding the reservoir of infection for poultry [24]. Longer and more efficient seeding of the virus has also been observed in some birds, potentially affecting the virus' greater ability to survive during the summer. Consequently, the amount and pressure of the virus in the environment increases, which in turn increases the risk of its introduction into the poultry population, but also the spread to other animal species – mammals in particular. We describe the detection of HPAI A(H5N1) virus infections in 25 cats during the second half of June in six voivodeships in Poland. Complete genome sequences of 19 HPAI H5N1 virus-positive cats indicate that the viruses belonged to clade 2.3.4.4b, genotype CH (H5N1-A/Eurasian_Wigeon/Netherlands/3/2022-like). The cat viruses were highly related to each other and clustered with a virus of the same genotype detected at the beginning of June in a white stork in Poland. When infections appeared in the cats, it was the BB genotype that was initially suspected as the cause of infection, considering that the EURL had issued a warning stating that the infection in poultry with this genotype could go undetected [1]. In April 2023, it was reported that the infection with the BB genotype could give anomalous disease signs in some poultry species such as turkeys or commercial layers, characterised by low mortality, very low prevalence of infection and almost the total absence of the typical HPAI signs, i.e. egg drop or reduced feed consumption. When there was a definite dominance of this genotype over others detected in wild birds, the veterinary inspectorate and poultry veterinarians in Poland received this information requesting to monitor the health status of poultry flocks very closely. One of the possible sources of infection in cats seems to have

been the consumption of poultry meat. However, from the interviews with cat owners, this is not so clear, as some cats received only specific food and had no contact with the outside environment. On the other hand, we also detected the infection in a stray cat (although it cannot be ruled out that someone had fed it with leftover raw chicken meat). Furthermore, the detection of the CH genotype virus in cats, which had been present in poultry almost 5 months earlier, further obscures the situation. Of course, there are known situations of silent virus infections in commercial poultry, as recently described in broilers in Italy [25], but the BB genotype is now circulating among wild birds and it is this genotype that has been the most suspected in terms of feline infections. However, the occurrence of the CH genotype in poultry would not necessarily need to coincide with its occurrence in cats because freezing of poultry meat is common in the preparation or storage of cat and other pet food. And, as discussed earlier, the surveillance system for influenza infections in poultry may not have detected single outbreaks caused by the CH genotype. It should also be remembered that Poland also imports poultry meat from other countries where the supervision system may be imperfect. Cats found to be infected with the HPAI H5N1 virus suffered from severe outcome of the disease and showed respiratory and nervous symptoms including death, and large amounts of virus genome were detected especially in the brain, but also in the lungs and bronchi. This is similar to recent reports on pathogenicity of A(H5N1) 2.3.4.4b virus lineage in ferrets which targeted the central nervous system causing dramatic neurologic involvement [26]. All the viruses from cats possess two mutations in the PB2 protein, E627K and K526R, which are molecular markers of virus adaptation to mammals [13]. The PB2-E627K mutation has been demonstrated to enhance polymerase activity, virus replication, transmission and, in certain cases, pathogenicity and mortality in mammals [14-23]. The PB2-K526R mutation has in some avian influenza viruses been responsible for human cases (H5N1 and H7N9) and in the majority of the seasonal influenza A(H3N2) viruses [27]. A previous study showed that influenza A(H7N9) viruses possessing both 526R and 627K replicate more efficiently in mammalian (but not avian) cells and in mouse lung tissues, and cause greater mortality in infected mice [28]. Interestingly, the PB2-E627K mutation was present in the influenza A(H5N1) virus

detected in the white stork at the beginning of June. The white stork is a carnivorous bird that feeds on a great variety of food: insects, earthworms, reptiles, amphibians and small mammals. Studies carried out on the population of white storks in Poland have shown that the diet of birds coming to breed was dominated by carp (Cyprinus carpio), European moles (Talpa europaea), voles (Microtus arvalis and agrestis) and earthworms (Lumbricidae), together accounting for 49.9% of the biomass eaten by them. Amphibians accounted for only 4.9% of the biomass they ate [29]. On one hand, it cannot be ruled out that the white stork became infected with the HPAI H5N1 virus containing a mammalian adaptation mutation through the ingestion of an infected mammal. It is known that small mammals such as rodents can serve as mechanical vectors or active shedders of avian influenza viruses [30]. On the other hand, infection of bank voles with the H5 and H7 subtypes of HPAI virus did not cause any symptoms of disease, but reversely resulted in shedding of high amounts of the virus [31,32]. The second mammal-adapted PB2-K526R mutation may have arisen as a result of infection in the cat's body. Nevertheless, it cannot be ruled out that the viruses with both mutations have been silently circulating in the bird population in Poland. To date, the Polish influenza A(H5N1) viruses from cats are the only ones characterised during the current epidemic wave showing both mutations. Additionally, the viruses from the Polish cats have shown 0–12 nucleotide and 0–8 amino acid differences distributed along the entire genome. The number of observed mutations suggests that the cats may have been exposed to multiple sources of infection of highly related viruses. However, we cannot completely exclude that mutations may have been acquired during the intra-host evolution of the virus in each animal.

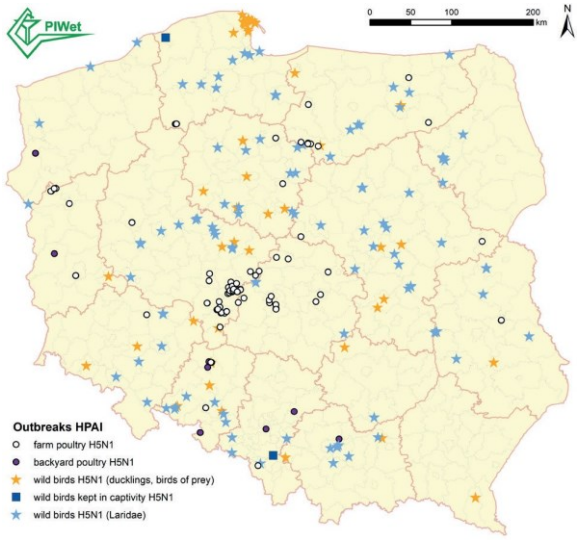
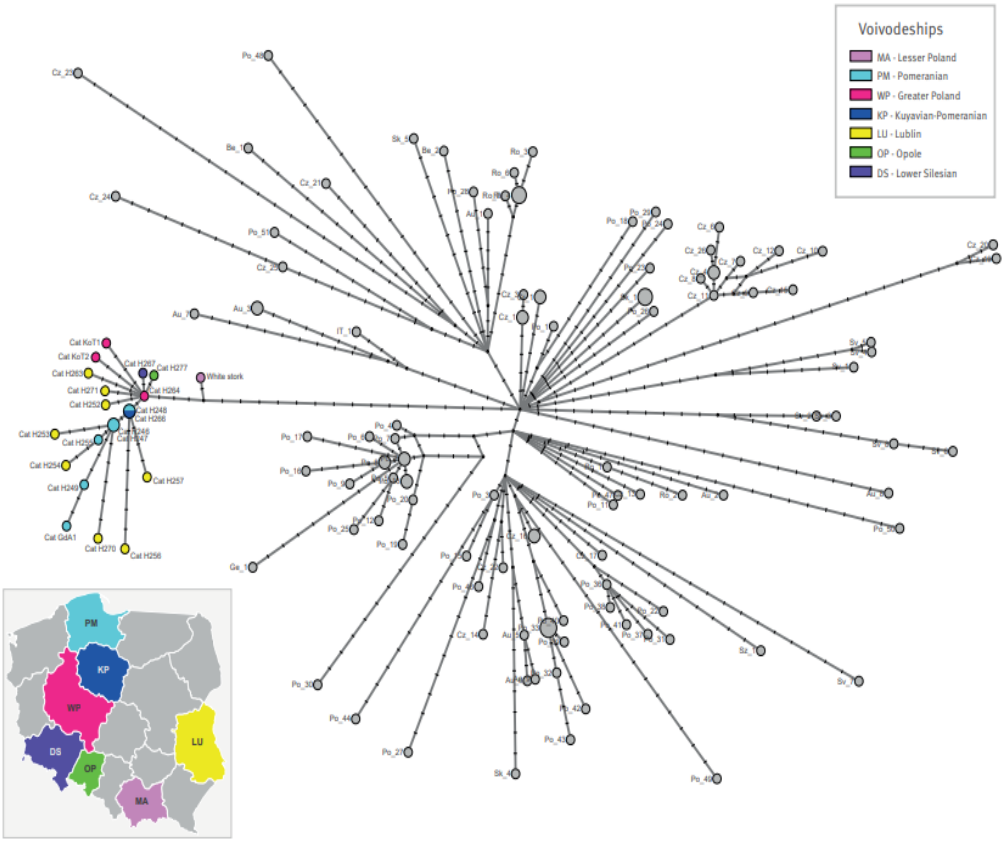


Figure 3. Location of outbreaks of highly pathogenic avian influenza A(H5N1) virus in poultry and wild birds, Poland, during the 2022/23 season.



DS: Lower Silesian; KP: Kuyavian-Pomeranian; LU: Lublin; MA: Lesser Poland; OP: Opole; WP: Greater Poland.

[a] H5N1 A/Eurasian wigeon/Netherlands/3/2022-like genotype from Europe.

Figure 4. Genetic network of complete genome sequences of highly pathogenic avian influenza A(H5N1) virusesa of the CH genotype collected from birds in Europe and from cats (n = 19) in Poland, June–July 2023.

The urgent question at the moment is to identify the direct source of virus infection in cats. The simultaneous detection of highly similar viruses over a vast geographical area undermines the hypothesis of direct transmission from wild birds to cats and points in the direction of an unidentified intermediate food source, e.g. poultry meat contaminated with the virus, that had accidentally entered the cats' food chain. This pathway of transmission requires careful investigation that is currently underway. In view of the high number of infections in cats in Poland and also in view of recent announcements by international institutions (World Health Organization (WHO), World Organisation for Animal Health (WOAH) and Food and Agriculture Organization (FOA)) as well as European authorities (European Centre for Disease Prevention and Control (ECDC), European Food Safety Authority (EFSA) and EURL) that outbreaks of avian influenza in animals pose a threat to humans [33,34], the veterinary inspection in Poland has issued recommendations for cat owners to restrict the outdoor access for animals, stop feeding them with raw poultry meat, and to disinfect surfaces potentially in contact with the bird environment (e.g. shoe soles, terrace surfaces). The occurrence of such outbreaks in cats in other European countries cannot be ruled out either. If the source of infection was meat from a Polish poultry farm, it could potentially be exported to other countries, but there have been no reports on similar events outside of Poland to date.

Conclusions

Recently, there have been a number of worrying changes in the ongoing HPAI outbreak. Another such unusual situation has occurred in Poland – the unprecedented scale of HPAI H5N1 virus infections of cats. Although the most likely source appears to be poultry meat, no such meat has been identified to date. Surveillance of poultry should certainly be enhanced, but also for certain, susceptible species of farmed mammals kept close to infected poultry farms. In addition, it seems reasonable to carry out scientific research into the susceptibility to influenza of other animals, in particular small mammals such as moles or voles. Furthermore, this study highlights the need in Europe to include Mammalia in the group of species posing a considerable risk for the spread of HPAI, in order to

provide health authorities with tools and guidelines for the proper management of such cases.

Table 1 Information on cats positive for highly pathogenic avian influenza A(H5N1) virus, Poland, June–July 2023 (n = 25): https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10401911/table/t1/?report=object only

Table 2 Amino acid differences among the proteins of influenza A(H5N1) viruses collected from cats, Poland, June–July 2023 (n = 19)

Amino acid differences	PB2			PB1					PA				HA		NA		M1		M2	NS1
	443	472	649	47	111	112	211	382	105	277	475	704	11	67	143	450	33	207	55	36
A/domestic cat/Poland/H246-M/2023 H5N1 2023–06–21 Gdynia	K	E	V	H	**I**	**D**	R	N	F	S	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H255-M/2023 H5N1 2023–06–21 Pruszcz	K	E	V	H	**I**	**D**	R	N	F	S	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H256-G/2023 H5N1 2023–06–24 Lublin	K	E	V	H	**I**	E	R	N	F	**S**	A	A	V	**V**	K	**S**	**V**	S	**F**	L
A/domestic cat/Poland/H257-G/2023 H5N1 2023–06–24 Lublin	K	E	V	H	**I**	E	R	N	F	**S**	**T**	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H263-G/2023 H5N1 2023–06–26 Komarow	K	E	V	H	**M**	E	R	N	F	**S**	A	**T**	V	I	K	**G**	A	S	L	L
A/domestic cat/Poland/H264-G/2023 H5N1 2023–06–26 Poznan	K	E	V	H	**M**	E	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H266-W/2023 H5N1 2023–06–19 Bydgoszcz	K	E	V	H	**I**	E	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H267-W/2023 H5N1 Strzelin	K	E	V	H	**M**	E	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H270-W/2023 H5N1 lubelskie	K	E	V	H	**I**	**D**	R	**D**	F	**Y**	A	A	V	I	**R**	S	A	S	L	L
A/domestic cat/Poland/H271-W/2023 H5N1 lubelskie	K	E	V	H	**M**	E	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H277-W1/2023 H5N1 2023–06–26 Namyslow	K	E	V	H	**M**	E	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H249/2023 H5N1 2023–06–22 Gdansk	**R**	**D**	V	H	**I**	**D**	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H248/2023 H5N1 2023–06–15 Pruszcz Gd	K	E	V	H	**M/I**	E	R	N	F	**S**	A	A	V	I	K	S	A	S	L	L

A/domestic cat/Poland/Kot2/2023 H5N1 2023–06–19 Poznan	K	E	**M**	H	M	E	R	N	/	/	/	A	V	I	K	S	A	**N**	L	L
A/domestic cat/Poland/Kot1/2023 H5N1 2023–06–19 Poznan	K	E	V	**Y**	M	E	**K**	N	F	S	A	A	V	I	K	S	A	S	L	**I**
A/domestic cat/Poland/H254/2023 H5N1 2023–06–22 Lublin	K	E	V	H	**I**	**D**	R	N	**L**	S	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H252/2023 H5N1 2023–06–22 Lublin	K	E	V	H	M	E	R	N	F	S	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H247/2023 H5N1 2023–06–20 Gdansk	K	E	V	H	**I**	**D**	R	N	F	S	A	A	V	I	K	S	A	S	L	L
A/domestic cat/Poland/H253/2023 H5N1 2023–06–22 Lublin	K	E	V	H	**I**	**D**	R	N	F	S	A	A	**A**	I	K	S	A	S	L	L

References

1. Adlhoch C, Fusaro A, Gonzales JL, Kuiken T, Mirinaviciute G, Niqueux É, et al. Avian influenza overview March-April 2023. EFSA J. 2023;21(6):e08039. PMID: 37293570

2. Adlhoch C, Baldinelli F, Fusaro A, Terregino C. Avian influenza, a new threat to public health in Europe? Clin Microbiol Infect. 2022;28(2):149-51. https://doi.org/10.1016/j.cmi.2021.11.005 PMID: 34763057

3. Adlhoch C, Fusaro A, Gonzales JL, Kuiken T, Marangon S, Niqueux É, et al. Avian influenza overview June-September 2022. EFSA J. 2022;20(10):e07597. PMID: 36247870

4. European Commission. Commission Delegated Regulation (EU) 2020/689 of 17 December 2019 supplementing Regulation (EU) 2016/429 of the European Parliament and of the Council as regards rules for surveillance, eradication programmes, and disease-free status for certain listed and emerging diseases. Official Journal of the European Union. Luxembourg: Publications Office of the European Union. 3.6.2020:L174/211. Available from: http://data.europa.eu/eli/reg_del/2020/689/oj

5. Government of the Republic of Poland. Rozporządzenie Ministra Rolnictwa i Rozwoju Wsi z dnia 23 marca 2023 r. w sprawie wprowadzenia Krajowego programu mającego na celu wykrycie zakażeń wirusami wywołującymi grypę ptaków (Avian influenza) u drobiu i dzikich ptaków na 2023 r. [Regulation of the Minister of Agriculture

and Rural Development of 23 March 2023 on the introduction of the "National program aimed at detecting infections with viruses causing avian influenza (Avian influenza) in poultry and wild birds for 2023]. Dz.U z dnia 31 marca 2023, poz. 618. 2023. Polish. Available from: https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20230000618/O/D20230618.pdf

6. Spackman E, Senne DA, Myers TJ, Bulaga LL, Garber LP, Perdue ML, et al. Development of a real-time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. J Clin Microbiol. 2002;40(9):3256-60. https://doi.org/10.1128/JCM.40.9.3256-3260.2002 PMID: 12202562

7. Slomka MJ, Pavlidis T, Banks J, Shell W, McNally A, Essen S, et al. Validated H5 Eurasian real-time reverse transcriptase-polymerase chain reaction and its application in H5N1 outbreaks in 2005-2006. Avian Dis. 2007;51(1) Suppl;373-7. https://doi.org/10.1637/7664-060906R1.1 PMID: 17494587

8. Hoffmann B, Hoffmann D, Henritzi D, Beer M, Harder TC. Riems influenza a typing array (RITA): An RT-qPCR-based low density array for subtyping avian and mammalian influenza a viruses. Sci Rep. 2016;6(1):27211. https://doi.org/10.1038/srep27211 PMID: 27256976

9. Śmietanka K, Świętoń E, Kozak E, Wyrostek K, Tarasiuk K, Tomczyk G, et al. Highly pathogenic avian influenza H5N8 in Poland in 2019-2020. J Vet Res (Pulawy). 2020;64(4):469-76. https://doi.org/10.2478/jvetres-2020-0078 PMID: 33367134

10. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268-74. https://doi.org/10.1093/molbev/msu300. PMID: 25371430

11. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol. 2018;35(2):518-22. https://doi.org/10.1093/ molbev/msx281 PMID: 29077904

12. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16(1):37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036 PMID: 10331250

13. Suttie A, Deng Y-M, Greenhill AR, Dussart P, Horwood PF, Karlsson EA. Inventory of molecular markers affecting biological characteristics of avian influenza A

viruses. Virus Genes. 2019;55(6):739-68. https://doi.org/10.1007/s11262-019-01700-z PMID: 31428925

14. Subbarao EK, London W, Murphy BR. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. J Virol. 1993;67(4):1761-4. https://doi.org/10.1128/jvi.67.4.1761-1764.1993 PMID: 8445709

15. Hatta M, Gao P, Halfmann P, Kawaoka Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science. 2001;293(5536):1840-2. https://doi.org/10.1126/ science.1062882 PMID: 11546875

16. Salomon R, Franks J, Govorkova EA, Ilyushina NA, Yen H-L, Hulse-Post DJ, et al. The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04. J Exp Med. 2006;203(3):689-97. https:// doi.org/10.1084/jem.20051938 PMID: 16533883

17. Shinya K, Hamm S, Hatta M, Ito H, Ito T, Kawaoka Y. PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. Virology. 2004;320(2):258-66. https://doi.org/10.1016/j. virol.2003.11.030 PMID: 15016548

18. Steel J, Lowen AC, Mubareka S, Palese P. Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. PLoS Pathog. 2009;5(1):e1000252. https://doi.org/10.1371/journal.ppat.1000252 PMID: 19119420

19. Mehle A, Doudna JA. An inhibitory activity in human cells restricts the function of an avian-like influenza virus polymerase. Cell Host Microbe. 2008;4(2):111-22. https://doi. org/10.1016/j.chom.2008.06.007 PMID: 18692771

20. Hatta M, Hatta Y, Kim JH, Watanabe S, Shinya K, Nguyen T, et al. Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. PLoS Pathog. 2007;3(10):1374-9. https://doi. org/10.1371/journal.ppat.0030133 PMID: 17922570

21. Labadie K, Dos Santos Afonso E, Rameix-Welti MA, van der Werf S, Naffakh N. Host-range determinants on the PB2 protein of influenza A viruses control the interaction between the viral polymerase and nucleoprotein in human cells. Virology. 2007;362(2):271-82. https://doi.org/10.1016/j. virol.2006.12.027 PMID: 17270230

22. Rameix-Welti MA, Tomoiu A, Dos Santos Afonso E, van der Werf S, Naffakh N. Avian Influenza A virus polymerase association with nucleoprotein, but not polymerase

assembly, is impaired in human cells during the course of infection. J Virol. 2009;83(3):1320-31. https://doi.org/10.1128/JVI.00977-08 PMID: 19019950

23. Fornek JL, Gillim-Ross L, Santos C, Carter V, Ward JM, Cheng LI, et al. A single-amino-acid substitution in a polymerase protein of an H5N1 influenza virus is associated with systemic infection and impaired T-cell activation in mice. J Virol. 2009;83(21):11102-15. https://doi.org/10.1128/JVI.00994-09 PMID: 19692471

24. Knief U, Bregnballe T, Alfarwi I, Ballmann M, Brenninkmeijer A, Bzoma S, et al. Highly pathogenic avian influenza causes mass mortality in Sandwich tern (Thalasseus sandvicensis) breeding colonies across northwestern Europe. bioRxiv. 2023:2023.05.12.540367. Preprint.

25. Gobbo F, Zanardello C, Bottinelli M, Budai J, Bruno F, De Nardi R, et al. Silent infection of highly pathogenic avian influenza virus (H5N1) clade 2.3.4.4b in a commercial chicken broiler flock in Italy. Viruses. 2022;14(8):1600. https://doi. org/10.3390/v14081600 PMID: 35893671

26. Kandeil A, Patton C, Jones JC, Jeevan T, Harrington WN, Trifkovic S, et al. Rapid evolution of A(H5N1) influenza viruses after intercontinental spread to North America. Nat Commun. 2023;14(1):3082. https://doi.org/10.1038/s41467-023-38415-7 PMID: 37248261

27. Li J, Ishaq M, Prudence M, Xi X, Hu T, Liu Q, et al. Single mutation at the amino acid position 627 of PB2 that leads to increased virulence of an H5N1 avian influenza virus during adaptation in mice can be compensated by multiple mutations at other sites of PB2. Virus Res. 2009;144(1-2):123-9. https:// doi.org/10.1016/j.virusres.2009.04.008 PMID: 19393699

28. Song W, Wang P, Mok BW, Lau SY, Huang X, Wu WL, et al. The K526R substitution in viral protein PB2 enhances the effects of E627K on influenza virus replication. Nat Commun. 2014;5(1):5509. https://doi.org/10.1038/ncomms6509 PMID: 25409547

29. Kosicki J, Profus P, Dolata P, Tobółka M. Food composition and energy demand of the white stork Ciconia ciconia breeding population. Literature survey and preliminary results from Poland. In: Tryjanowski P, Sparks T, Jerzak L (eds). The white stork in Poland: studies in biology, ecology and conservation. Poznań: Bogucki Wyd. Nauk; 2006. p. 169-83. Available from: https://www.researchgate.net/publication/237325391_Food_

composition_and_energy_demand_of_the_White_Stork_
Ciconia_ciconia_breeding_population_Literature_survey_and_
preliminary_results_from_Poland

30. Velkers FC, Blokhuis SJ, Veldhuis Kroeze EJB, Burt SA. The role of rodents in avian influenza outbreaks in poultry farms: a review. Vet Q. 2017;37(1):182-94. https://doi.org/10.1080/0165 2176.2017.1325537 PMID: 28460593

31. Romero Tejeda A, Aiello R, Salomoni A, Berton V, Vascellari M, Cattoli G. Susceptibility to and transmission of H5N1 and H7N1 highly pathogenic avian influenza viruses in bank voles (Myodes glareolus). Vet Res. 2015;46(1):51. https://doi. org/10.1186/s13567-015-0184-1 PMID: 25963535

32. Shriner SA, VanDalen KK, Mooers NL, Ellis JW, Sullivan HJ, Root JJ, et al. Low-pathogenic avian influenza viruses in wild house mice. PLoS One. 2012;7(6):e39206. https://doi.org/10.1371/ journal.pone.0039206 PMID: 22720076

33. World Health Organization (WHO). Ongoing avian influenza outbreaks in animals pose risk to humans. Situation analysis and advice to countries from FAO, WHO, WOAH. Geneva/Paris/ Rome: WHO/WOAH/FAO; 12 July 2023. Available from: www. who.int/news/item/12-07-2023-ongoing-avian-influenza-outbreaks-in-animals-pose-risk-to-humans

34. European Food Safety Authority; European Centre for Disease Prevention and Control; European Reference Laboratory for Avian Influenza; Adlhoch C, Fusaro A, Gonzales JL, et al. Scientific report: Avian influenza overview April–June 2023. EFSA J. 2023;21(7):8191.

4 Discussions

A series of recent emerging infectious disease outbreaks have underlined the need to better understand which kinds of pathogens are causing disease in animal and human populations. DNA and RNA viruses have enormous diversity [1]. This is the consequence of high rates of evolution [2] due to their mutation rates, and the ability to undergo recombination, resulting in diverse genotypes and the emergence of variants with increased zoonotic impact, virulence, and transmissibility. To follow these outstanding evolutionary properties, the role of rapidly identifying and characterizing emerging viruses, particularly those with zoonotic potential, in safeguarding public health is critical and the development of laboratory protocols and bioinformatics pipelines tailored for the genetic characterization of these viruses is necessary. This thesis revolves around developing comprehensive workflows for next-generation sequencing (NGS) data analysis, ranging from raw data processing to generating consensus sequences and to phylogenetic and phylodynamic methods. These workflows are used to perform the identification and characterization of DNA and RNA viruses for the study of evolutionary dynamics and molecular epidemiology, allowing for a better understanding of infectious diseases, their spread, and pathogen adaptation, with significant impacts on public health, disease control strategies, global health initiatives and advances in scientific understanding. The investigation for SARS-Cov-2 identification and characterization involves the devolopment of NGS methodology based on in-house targeted amplification protocol to sequence SARS-CoV-2 genome from humans using second (Illumina) and third (ONT) generation sequencing technologies. The development of a bioinformatics pipeline for the data analysed using Illumina sequencing platform allowed the correct identification of reads belonging to SARS-Cov-2 using this protocol. The accurate sequencing of SARS-CoV-2 complete genome from clinical samples from different matrices, also with a low viral load and high and uniform depth of coverage, permitted to study the intra-host variability of viral population. A specific pipeline for the data analysed using the MinION technology was developed and the comparison of the two methologies showed that the medium MinION coverage is lower but the differences between the two methodologies were in the homopolymeric regions, that is the weak point of the

portable and rapid MinION instrument. This analysis allowed to understand the role of different sequencing approaches in supporting such studies. The bioinformatics pipeline for the analysis of targeted amplified and sequenced data was applied to study the natural human-to-cat SARS-CoV-2 transmission happened in Italy in 2020 for Identification of Variants and downstream analysis were performed. The diagnosis of SARS-CoV-2 infection in pets, such as cats, is extremely important to provide appropriate veterinary care for the infected animals; to guarantee adequate protection of veterinary staff and pet owners; and to apply quarantine measures to prevent transmission between pets, people and potentially susceptible animals. Even though the viral shedding from pets did not appear sufficient to infect other family members or other animals, the usual precautionary measures should urgently be considered as part of the global control efforts and One Health approach. There were no evidence that cats played a significant role in human infection and in the spread of the virus to humans. Thus, reverse zoonosis is possible if infected owners expose their pets to the virus, particularly during the acute phase of the infection. It is important that pet owners are educated to adopt the precautionary measures to avoid human-to-cat SARS-CoV-2 transmission [3]. Preventing interspecies transfer of an emergent pathogen is essential to decrease the risk of emerging mutations that could affect the transmissibility or effectiveness of the countermeasures, and is also needed to safeguard pet welfare and discourage animal abandonment.

Fatal episodes affecting birds in 2021 in the north central regions of Italy presented the opportunity to test new methodologies in order to discover the causative agents, solving the discrepancies of standard diagnostic methodologies by developing a metagenomic approach. The identification of viruses through metagenomic approaches underscored the limitations of conventional PCR testing, emphasizing the necessity of complementary methodologies in unveiling elusive pathogens within avian populations. The discrepancy between electron microscopy findings of Circoviridae-like particles and the failure of specific PCR tests to detect Canary circovirus (CaCV), Psittacine Circovirus (PBFD), and avian Polyomavirus (APV) warrants attention. The elucidation of mismatches in the avian Polyomavirus primer sequences offered a crucial insight, indicating the need for meticulous primer design and constant adaptation to evolving viral genomes. This finding showed the

importance of periodic reassessment and refinement of molecular diagnostic tools to ensure their efficacy in viral identification. The genetic variability observed in Beak and Feather Disease Virus (Circovirus) highlighted the dynamic nature of viral evolution and its implications for diagnostic accuracy. The phylogenetic analysis delineating genetic differences among these strains emphasized the necessity of precise molecular tools tailored to detect such variations. The development of a specific protocol that uses primers based on the identified consensus sequences presented a promising approach for enhancing the rapidity, sensitivity and accuracy of PCR-based diagnostics in identifying novel strains of Circovirus. Furthermore, the discovery of novel viral sequences underscored the importance of comprehensive metagenomic approaches in uncovering previously unknown viral species or strains. The application of these advanced techniques not only broadened our understanding of viral diversity but also highlighted the potential existence of novel viruses. In conclusion, the amalgamation of electron microscopy, metagenomic analyses, and refined primer design proved instrumental in unraveling the diverse viral landscape in the investigated bird populations. This study served as a testament to the necessity of integrating multiple diagnostic modalities and refining molecular tools for accurate and comprehensive viral identification in avian veterinary pathology.

Hantaviruses, particularly Dobrava-Belgrade orthohantavirus (DOBV), were investigated in rodents in north-eastern Italy. Yellow-necked mice, including *A. flavicollis* and *A. agrarius*, were identified as DOBV reservoirs in the region. In 2021, DOBV presence in yellow-necked mice was confirmed through direct sequencing and phylogenetic analyses. While serological evidence hinted at hantavirus circulation in Italy since 2000, this study marked the first virological confirmation. The increased rodent diagnostic activity in 2021 correlated with elevated mortality in the area. The findings suggested that negative results in certain regions may be due to inadequate surveillance rather than the absence of the pathogen. The study underscored the challenges in detecting pathogens in wildlife and emphasized the need for ongoing active and passive surveillance. Furthermore, the research revealed an upsurge in mortality not only in mice but also in voles, although the reservoir species for Puumala orthohantavirus (PUVV), M. glareolus,

tested negative. The increased mortality might be linked to ecological changes or population density rather than viral infection. The study advocated for continued surveillance to comprehend hantavirus dynamics in the region. Despite reports of PUVV cases in Slovenia in 2021, no evidence was found in Italian bank voles, suggesting the need for further research and enhanced awareness among healthcare professionals and local populations. In conclusion, this study confirmed DOBV circulation in wild mice in north-eastern Italy and underscored the importance of surveillance, research, and awareness to mitigate the risk of hantavirus infection in humans.

Another important focus of this thesis was to genetically investigate the different epidemic waves of HPAI A(H5N1) viruses in Italy happened during 2020-2023 to better undestrand the incursion of multiple viral genotypes with different biological characteristics and behaviors in domestic and wild birds. Bioinformatics workflow and statistics tools were used to study these epidemics waves. During the first wave (November 2020-February 2021) few cases, in wild (mainly waterfowls) and domestic birds, were recorded with the A genotype (H5N8-A/duck/Chelyabinsk/1207-1/2020-like) as the main circulating one. The second epidemic wave (October 2021-March 2022) was characterized by a high circulation of the virus in poultry farms. During this wave, the most widespread genotype was the C genotype (H5N1-A/Eurasian Wigeon/Netherlands/1/2020-like), which started to circulate during the first wave and became predominant in the second wave. The phylodynamics and pylogeographic analyses performed reveal their importance in order to understand the viral spread among the different locations and species involved and many distinct introductions into poultry from wild birds were recorded and also the occurrence of several secondary outbreaks. These results were also confirmed by the network analysis, identifying the province of Verona and Mantua as the main source of the virus for the neighboring areas. These analyses were necessary also to identify the association of genotypes with different hosts as the BB genotypes most present in Laridae during the last epidemic wave (from December 2022 to 2023). However, the continuous circulation and emergence of new reassortant viruses in Italy and Europe raises concern for animal and public health.

Recently, there have been a number of worrying changes in the ongoing HPAI outbreak. These changes in the course of the HPAI epidemics, the different genotypes, the persistence of the virus, the greater range of avian hosts, as well as the number of mammals infected demonstrate worrying changes in the biological properties of circulating HPAI H5N1 viruses, which confer to the viruses the ability to infect a greater number of wild bird species that are not normally susceptible, expanding the reservoir of infection for poultry. Longer and more efficient seeding of the virus has also been observed in some birds, potentially affecting the virus' greater ability to survive during the summer. Consequently, the amount and pressure of the virus in the environment increases, which in turn increases the risk of its introduction into the poultry population, but also the spread to other animal species – mammals in particular. The detection of HPAI A(H5N1) virus infections in 25 cats during the second half of June 2023 in six voivodeships in Poland was performed. The cat viruses were highly related to each other and clustered with a virus of the same genotype detected at the beginning of June in a white stork in Poland. The number of observed mutations suggested that the cats may have been exposed to multiple sources of infection of highly related viruses. However, we could not completely exclude that mutations may had been acquired during the intra-host evolution of the virus in each animal. The urgent question at that moment was to identify the direct source of virus infection in cats. The simultaneous detection of highly similar viruses over a vast geographical area undermined the hypothesis of direct transmission from wild birds to cats and pointed in the direction of an unidentified intermediate food source, e.g. poultry meat contaminated with the virus, that had accidentally entered the cats' food chain. In view of the high number of infections in cats in Poland and also in view of recent announcements by international institutions (World Health Organization (WHO), World Organisation for Animal Health (WOAH) and Food and Agriculture Organization (FOA)) as well as European authorities (European Centre for Disease Prevention and Control (ECDC), European Food Safety Authority (EFSA) and EURL) that outbreaks of avian influenza in animals posed a threat to humans, the veterinary inspection in Poland had issued recommendations for cat owners to restrict the outdoor access for animals, stop feeding them with raw poultry meat, and to disinfect surfaces

potentially in contact with the bird environment. The occurrence of such outbreaks in cats in other European countries could not be ruled out either. If the source of infection was meat from a Polish poultry farm, it could potentially be exported to other countries, but there have been no reports on similar events outside of Poland to date. Surveillance of poultry should certainly be enhanced, but also for certain, susceptible species of farmed mammals kept close to infected poultry farms. In addition, it seems reasonable to carry out scientific research into the susceptibility to influenza of other animals, in particular small mammals such as moles or voles. Furthermore, this study highlights the need in Europe to include Mammalia in the group of species posing a considerable risk for the spread of HPAI, in order to provide health authorities with tools and guidelines for the proper management of such cases.

The results produced in this PhD research underscore the strategic significance of genetic information and bioinformatic examinations for comprehending the dynamics of virus emergence, evolution, and spread. Moreover, they offer crucial insights for evaluating the risks linked with viral pathogens.

References

1. King A.M.Q. Family – Circoviridae. In: Fauquet C.M., Mayo M.A., Maniloff J., Desselberger U., Ball L.A., editors. Virus Taxonomy. Elsevier; San Diego: 2012. pp. 343–349

2. Woolhouse ME, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. Trends Ecol Evol. 2005;20:238–44. 10.1016/j.tree.2005.02.009

3. Centers for Disease Control and Prevention. Interim Guidance for Public Health Professionals Managing People with COVID-19 in Home Care and Isolation Who Have Pets or Other Animals. Available online: https://www.cdc.gov/coronavirus/2019-ncov/animals/interimguidance-%0Amanaging-people-in-home-care-and-isolation-who-have-pets.html (accessed on 6 May 2021)

Ringraziamenti

Portare avanti e completare questo dottorato a distanza di molti anni dalla Laurea Magistrale, lavorando a tempo pieno, è stata un'impresa impegnativa e gratificante che mi ha permesso di superare i miei limiti e di organizzare al meglio le mie energie, offrendomi preziose opportunità di confronto con docenti e colleghi attivi nel settore.

Vorrei esprimere la mia sincera gratitudine al Prof. Nicola Vitulo, il mio supervisore, che con costanza e sostegno è riuscito a guidarmi nonostante i miei molteplici impegni. Anche durante le emergenze causate dalla pandemia e dalle epidemie virali stagionali, le sue parole di incoraggiamento sono state un sostegno fondamentale lungo l'intero percorso. La sua vasta conoscenza e competenza in ambiti anche diversi sono stati un valore inestimabile.

Desidero ringraziare la Dottoressa Isabella Monne, il mio dirigente all'Istituto Zooprofilattico Sperimentale delle Venezie, che cinque anni fa mi ha accolto nel suo gruppo di ricerca e ha sostenuto con entusiasmo questa decisione di intraprendere un percorso formativo così lungo e impegnativo. Un caloroso ringraziamento va anche alla Dottoressa Alice Fusaro, che con dedizione ha coordinato e supervisionato le mie analisi di routine e di ricerca.

La mia gratitudine va a mia madre e mia sorella, il mio tutto, che hanno sopportato il mio nervosismo, le mie insicurezze e le mie paure durante questo percorso.

Un ringraziamento speciale va ad Arya, la piccola palla di pelo nero, per la sua compagnia incondizionata durante le lunghe ore trascorse al computer per completare questa epica opera.

Infine, desidero ringraziare colui che è nel mio cuore e nel mio sangue, che ha visto solo il mio percorso fino alle scuole elementari ma che sarebbe orgoglioso di quanto ho raggiunto.

Acknowledgements

Carrying on and completing this doctorate many years after my Master's degree while working full-time has been a challenging yet rewarding endeavor. It allowed me to surpass my limits, organize my energies efficiently, and provided valuable opportunities for engagement with professors and colleagues active in the field.

I would like to express my sincere gratitude to Professor Nicola Vitulo, my supervisor, who consistently guided and supported me despite my numerous commitments. Even amidst the emergencies caused by the pandemic and seasonal viral outbreaks, his words of encouragement were a fundamental support throughout the entire journey. His extensive knowledge and expertise across various domains have been invaluable.

I wish to thank Dr. Isabella Monne, my director at the Istituto Zooprofilattico Sperimentale delle Venezie, who welcomed me into her research group five years ago and enthusiastically supported my decision to embark on such a lengthy and challenging educational path. A warm thank you also goes to Dr. Alice Fusaro, who diligently coordinated and supervised my routine and research analyses.

My gratitude extends to my mother and sister, my everything, who endured my nervousness, insecurities, and fears throughout this journey.

A special thank goes to Arya, the little black poodle, for her unwavering companionship during the long hours spent at the computer completing this epic endeavor.

Lastly, I want to thank the one who resides in my heart and blood, who witnessed only my journey through elementary school but would be proud of how far I have come.