



**37. OTDK Biológia Szekció
Országos Tudományos Diákköri Konferencia**



**Molecular evolution of filoviruses:
Insights from new sequence data of the Lloviu virus**

Láng Levente

Biotechnology BSc II. year

Supervisors:

Dr. Gábor Kemenesi

Associate professor

Dr. Zsófia Lanszki

Assistant professor

PTE, Szentágothai Research center, National Laboratory of Virology



UNIVERSITY OF PÉCS
NATIONAL LABORATORY OF
VIROLOGY

Pécs, 2025

Table of content

1. INTRODUCTION	4
2. AIMS OF THE STUDY	8
3. MATERIALS AND METHODS.....	9
3.1 SAMPLE COLLECTION	9
3.2. NUCLEIC ACID EXTRACTION	10
3.3 QRT-PCR DIAGNOSTIC	11
3.4 VIRAL GENOME SEQUENCING.....	11
3.5 BIOINFORMATIC PROCESSING OF THE SEQUENCE DATA	12
3.6 PHYLOGENETIC TREE AND SIMILARITY PLOT ANALYSIS.....	13
6. RESULTS AND DISCUSSION	15
7. CONCLUSIONS	20
8. ACKNOWLEDGEMENT	20
9. REFERENCES	21

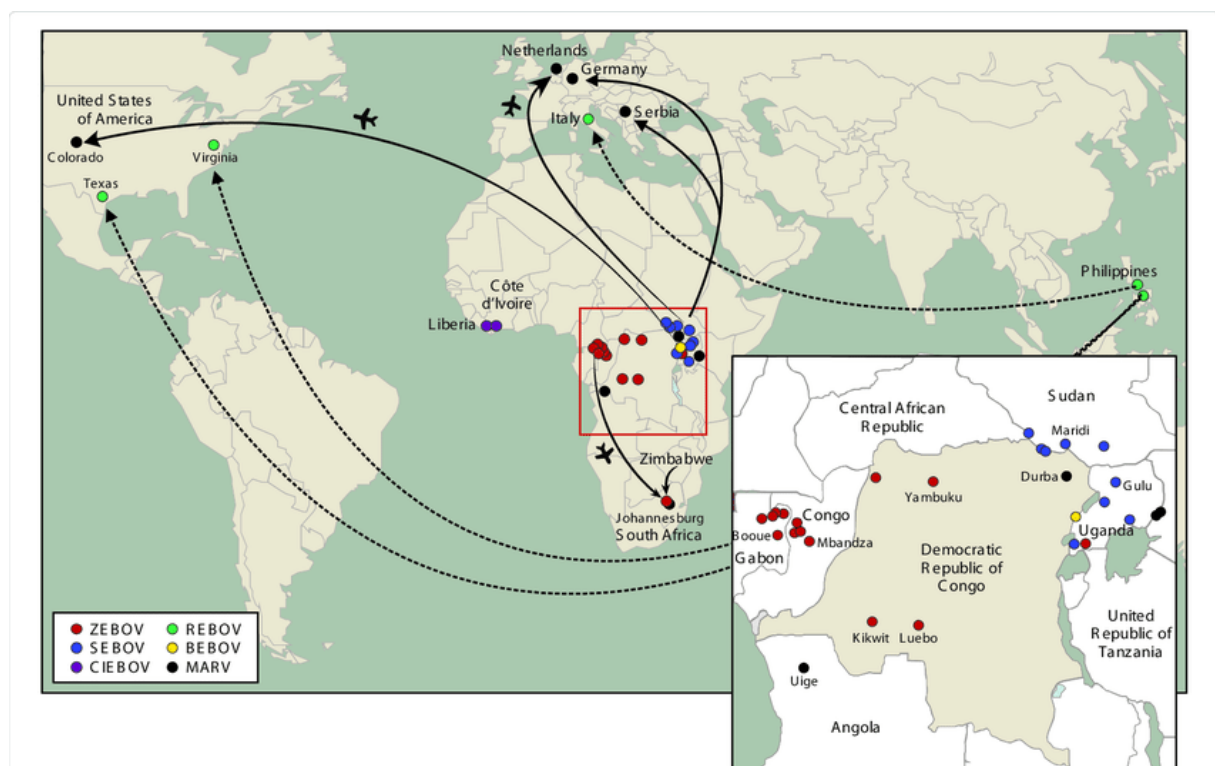
Abstract

Cuevavirus lloviuense (LLOV) is a member of the *Filoviridae* family, which was discovered in 2002 in Spain during a bat die-off event. Since the first detection the presence of the virus had been confirmed in many other European countries including Hungary in 2016 at a similar bat die-off event. Its host animal is the Schreiber's long-fingered bat present with a wide distribution range over Central and Southern Europe. The virus might have spillover potential based on its ability to infect human cell lines; however, its pathogenicity spectrum is unknown yet. Since it has a close functional and genetic relation to well known filoviruses (Marburg – and ebolaviruses in particular), the work described in the thesis is focused on the molecular evolutionary analysis of the virus. Phylogenetic analyses and different mutation patterns were examined to obtain a better understanding of the virus's evolution. Phylogenetic trees established during the work strengthens the hypothesis of the virus's enzootic property. Moreover, by studying the genetic structure of the virus carried by the bats also gives us information about its host mammals, their movement in the area, being the only viable reason for the sequences of the viral genomes to be so similar even though collected from different sampling locations. LLOV's evolutionary relationship with other filoviruses was examined also in details. Comparison between the well-known filoviruses like Marburg and Ebola's and LLOV's evolutionary pattern was found to be slightly different. The reason behind this is the smaller sampling interval of LLOV compared to Marburg and Ebola which had been sampled and studied for several decades more than LLOV. A similarity plot analysis was carried out which compares the same nucleotide positions in the compared genomes and visualizes which region of the virus's genome accumulates the most mutations. For these examinations to be carried out we had to prepare the samples which at the end by means of the consensus sequence is not only useful for our work on evolutionary studies but also serves as great data for other areas of virology such like synthetic virology.

My own contributions to the research was helping in sample collection, processing of 376 samples, including the nucleic acid extraction and qRT-PCR reaction. I also took part in the library preparation and NGS sequencing process and the subsequent bioinformatic processing of the raw data. This was then followed by the different phylogenetic and mutation pattern analysis, which I also took part in.

1. Introduction

Bats (Chiroptera) are the second largest group of mammals and the only one that are capable of true flight. Their „superpower” is, that they are extremely resistant to pathogens, with that being exceptionally good vectors for several diseases. Main reason for their extraordinary immune system is their high metabolic rate caused by flying (Brook et al. 2015). Most abundant pathogens carried by bats are viruses (Chen et al. 2014) some being able to cause deadly diseases in humans such as SARS-like coronaviruses (Lin-Fa et al. 2006) or Marburgvirus (Swanepoel et al. 2007). The latter one with several other viruses for example Ebolavirus Zairevirus and Lloviu virus compose the family: *Filoviridae*. Zoonotic filoviruses has been causing outbreaks in Africa since the end of the 1960's. Largest outbreak was the 2014 West-African Ebola outbreak where the *Orthoebolavirus zairense* (EBOV) caused the death of at least 11.000 people setting up a mortality rate of around 40% (Kaner et al. 2016). Currently an effective FDA approved vaccine only exists against Zaire ebolavirus, which is one of the 5 known human pathogenic filovirus species. Research on filoviruses thus, is very important including ecological, evolutionary studies and drug development, so future outbreaks would be more preventable or controllable.



1. Figure: Filovirus outbreaks and cases worldwide (Marzi et al. 2011)

According to 2025 data of the ICTV (International committee on Taxonomy of Viruses) the *Filoviridae* family contains 9 genera: *Oblavirus*, *Striavirus*, *Thamnovirus*, *Cuevavirus*, *Dianlovirus*, *Orthoebolavirus*, *Orthomarburgvirus*, *Tapjosvirus* and *Loebvirus*. Five species are known to be human pathogenic, these are: *Orthomarburgvirus marburgense* (MARV), *Orthoebolavirus sudanense* (SUDV), *Orthoebolavirus zairense* (EBOV), *Orthoebolavirus bundibogyoense* (BDBV), *Orthoebolavirus taiense* (TAFV), and there are viruses which are able to infect mammals including humans, but not cause a disease, like the *Orthoebolavirus restonense* (RESTV) (Pappalardo et al. 2017). Filoviruses got their name from their physical appearance, looking like small rods, filaments (Booth et al. 2015), and have a nonsegmented negative-sense single stranded RNA genome, that codes for 7 monocistronic mRNAs that will translate into 6 structural and one polymerase enzyme protein (Mühlberger 2007). Due to their relatively short and simple genome it is very easy to conduct research with. Next generation sequencing methods (NGS) makes it possible to determine the virus's complete genome for further molecular virus-host interaction research or evolutionary research, but gaining genetic data can help us find potential drug targets too.

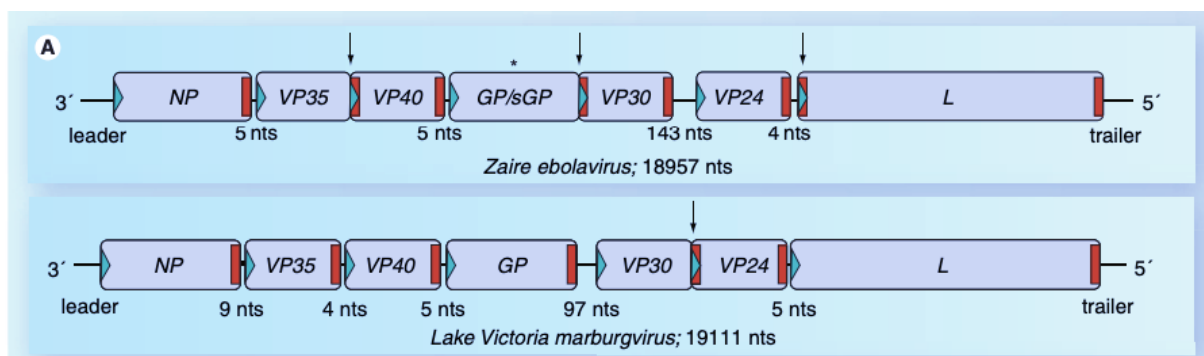


Figure 2. Schematic diagram of Ebolavirus (EBOV) and Marburgvirus (MARV) genomes. Mühlberger et al. 2007)

Cuevavirus lloviuense (LLOV)

The *Cuevavirus lloviuense* is one of the filovirus species, in fact the only present in European bat colonies. First detection happened in 2003 in a Spanish cave (*Cueva del Lloviu*) when researchers were examining a significant bat die-off event (Negredo et al. 2011). Since then, the virus presence had been confirmed in several other European countries (Toth et al. 2023) (Goletic et al. 2023) and the host animal is known, to be Schreiber's long-fingered bat (*Miniopterus schreibersii*). Isolation of the virus happened in 2020 (Kemenesi et al. 2022) and it was very important for future infection experiments, until that researchers had to work with

a recombinant virus (where this kind of gain of function experiments are legal) made with sequences from Ebola virus to complement the end of the LLOV genome due to the very high similarity. Lloviu is able to infect African green monkey kidney (VERO E6) and human cell lines (human neuroblastoma: SH-SY5Y, human hepatocyte carcinoma: Hep G2, human lung adenocarcinoma: HCC-78, human colon caecinoma: HCT 116) (Kemenesi et al. 2022) raising the spillover potential of the microbe. The most important aspect, the pathogenicity is unknown, however results from both *in vitro* and *in vivo* infection experiments suggests we do not have to worry (Hume et al. 2022, Fletcher et al. 2023,2024). The genome of LLOV is almost entirely known, there are only small gaps in the intergenic regions. Our research team is out every season collecting samples from bats and testing for the virus. We have a handful of sequence data now from a variety of countries, mainly from East and Central-Europe.

Molecular evolutionary studies of filoviruses

Phylogenetics is a fast-developing area of molecular evolutionary research. Information gained from phylogeny can be helpful in determining the pattern of spreading of a virus either within population or geographically, approximating the time of origin of a new virus strain or in predicting and preventing pandemics (Lam et al. 2010).

Viruses retain significant adaptivity impelled by mutation rates and interactions with their hosts. We can study mutational rate by analysing the product of it, the genetic variations of a given organism. Sometimes different genetic variants behave differently or built up differently by structural means. The host switching capability of the virus, including transferring from an animal host to human is also known as zoonosis is poorly understood, however the evolvability of the virus, including high mutability might be connected to it. This property of pathogens can be thoroughly screened by evolutionary studies, helping researchers to understand and maybe prevent outbreaks in the future (Morse et al. 2012) .

With advanced techniques of phylogeny, researchers can determine the approximate time of a virus strain's origin. This type of analysis is called time calibrated phylogenetic analysis, and it is based on extremely complex mathematical principles. Previous time estimations had already been carried out for species of the *Filoviridae* family (Suzuki and Gojobori 1997, Carrol et al. 2013), however most of them did not include the more recently discovered and comparatively neglected genera's species. In the last few years discovery of several novel filovirus species was made (Hierweger et al. 2017, Chen et al. 2019,), but since then, not much information about them were published. By involving them in similar analyses

like we do on Marburg, Ebola and Lloviu we could gain important information about them, for example their origin, zoonotic potential.

2. Aims of the study

The primary aim of our research is to identify new LLOV positive samples year by year within a framework of a surveillance program. With the deployment of a mobile laboratory and the help of different molecular biological techniques, we are not only able to carry out the real-time detection of the virus, but the measurement of the prevalence and distribution of the pathogen in various bat populations. Enough data is crucial before conducting further analyses to achieve more precise and reliable results.

Data derived from the positive samples are used in different types of evolutionary analyses, to fill in the gaps of our knowledge of LLOV, addressing critical questions like the origin of the virus and its relationship with other filoviruses.

Another key goal of our research is to establish a complete and accurate genome of the Lloviu virus, because currently the known sequence is still gappy at some regions. By achieving this goal, we help the work of the field of synthetic virology and the development of new diagnostic tools, vaccines or antivirals.

3. Materials and methods

Altogether I processed (nucleic acid extraction, PCR) 376 samples in the year of 2024 since I joined the research group; including 122 blood samples and 254 other samples, mainly bat flies. We are in possession of all necessary permits and ethical declaration concerning the live sampling of bats.

3.1 Sample collection

Hungarian samples were collected from 2016 February until 2024 October every month, excluding breeding and rearing season as well as hibernation period, from Northern Mountain range (Zemplén mountains) and the Transdanubian Hills (Villány Mountains). Moreover, in September of 2022 from Bulgaria, Serbia and Romania (Transylvania), while in October only from Slovakia and in 2023 from Bulgaria, Serbia and Bosnia-Herzegovina additional samples were collected. Primarily individuals of Schreibers's long-finger bats were sampled, however in some cases other bat species, habitats of the same cave as the Schreiber's bats were sampled as well. Bats were captured by mist net or harp trap during their emergence at dawn and by hand-nets during the day. Also, Schreibers's bat carcasses were collected from several locations during annual colony checks. Identification of the species were carried out following morphological identification keys⁺. Samples taken were the following: blood, guano, urine and ectoparasites (bat flies and ticks), if present.

Blood was collected by stabbing a capillary in the tail membrane of the animal with a 27-gauge needle and a Minivette POCT 100 µl (Sarstedt) capillary. The approximately 50-100 µl blood was then transferred into a 0,5 ml eppendorf tube. For the collection of ectoparasites we used sterile (this happened by immersing them in 96% ethanol and burning it off) tweezers for each individual and put them in a 2 ml tube one by one. Urine was collected via a Minivette POCT 100 µl (Sarstedt) capillary and transferred into a 0,5 ml eppendorf tube. Guano samples were either transferred into a 1,5 ml eppendorf tube filled with PBS or into a 2 ml tube. Separation of blood happened by a centrifuge on the site of sampling. To the sediment we added 200 µl DMEM (Dulbecco's Modified Eagle medium) or EMEM (Eagles Minimal Essential Medium) and after thorough suspension were divided (100-100 µl) into two 0,5 ml eppendorf tube.

All samples were stored in liquid nitrogen until further experiments. In some cases we could set up a „field laboratory” and from one set of the previously mentioned blood samples we extracted nucleic acid and ran a qRT-PCR reaction to test for viral presence on site.

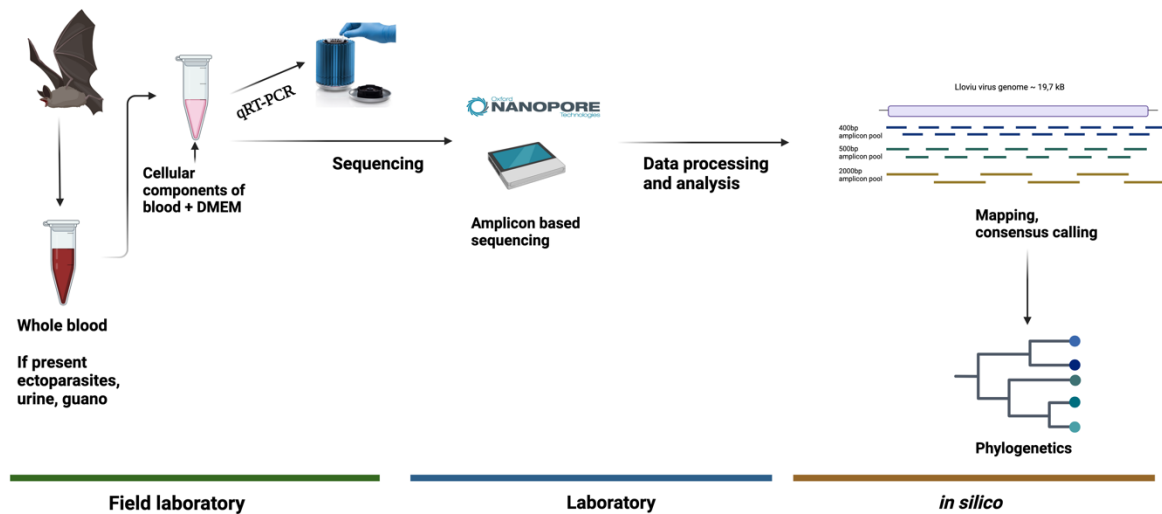


Figure 3. workflow from field to the laboratory (made with BioRender)

3.2. Nucleic acid extraction

The extraction of LLOV RNA was performed by following the manufacturer's instructions using Direct-Zol RNA MiniPrep kit (Zymo) and QIAamp Viral RNA kit (Quiagen). For the Direct-Zol extraction the 100 µl samples were transferred into 300 µl TRI reagent for inactivation, then we added 400 µl ethanol (96 V/V%), after we measured this 800 µl solution onto the columns. In the next step we washed the samples twice with 400 µl RNA Pre-wash solution. Each washing step was followed by a centrifugation at 14,000g. With 700 µl RNA wash buffer we washed the samples again. The nucleic acid samples were eluted in 40 µl of nuclease-free water instead of the recommended 50 µl to maintain a higher concentration.

In case of QIAamp extraction 140 µl sample was added to 560 µl lysis buffer (AVL) to which after a 10-minute incubation we added 560 µl ethanol (96 V/V%) to fully inactivate the samples. The washing steps were performed in a vacuum pump system. We measured the solutions onto the columns and under a constant vacuum we added 750 µl of the first wash buffer (AW1) then the same amount of the second wash buffer (AW2). For elution we used 50 µl of the kit's elution buffer (AVE) and added 1,5 µl RNase-out to prevent the degradation of the extracted fragile RNA. The RNA samples are stored at 4°C until the PCR reaction.

3.3 qRT-PCR diagnostic

The qRT-PCR was optimized for 20 µl reaction, 15 µl master mix + 5 µl RNA sample. For this we used Brilliant III ultra-fast PCR kit (Agilent technologies), 10 µl 2x buffer, 1 µl RT/RNase block, 0,5 µl (500nM) forward and reverse primer, 0,25 probe (250nM), 2,55 µl nuclease-free water, and 0,2 µl DTT per reaction. The reaction consists of a reverse transcription step, initial denaturation and a two-step amplification: 50°C 10 minutes, 95°C 3 minutes, (95°C 5 seconds, 60°C 30 seconds) x50 cycle. The detection of the fluorescence of the probe happened at the 60°C step. The primers and the probes were designed to bind the 110 base long conserved region of the RdRp (L) gene. The reaction was performed on MyGo Pro, MyGo Mini and MIC qPCR machines (Toth et al. 2023).

Sequence of the primers used for amplification:

- LLOV-Fw-scr: 5'-AAGCATTTCCGAGTAATATGATGGTTG-3'
- LLOV-Rev-scr1: 5'-TACATGGTCTCCTAGATTGCCCTG-3'

Sequence of the TaqMan probe used for detection:

- LLOV-Prob-scr1: 5'-**FAM**-CCTGATGAAGGAGAGTTTCTTTCTG-**ZEN**-3'

3.4 Viral genome sequencing

From the extracted nucleic acid, the Oxford Nanopore Technologies MinION sequencer system was used for genome determination. The technique we used is an „ARTIC-like” amplicon based sequencing method which ensures greater coverage and accuracy according to previously used protocols (protocols.io 2025).

From the isolated RNA first, we synthesized cDNA with the help of SuperScript IV enzyme (Invitrogen, USA) and RandomHexamer mix (Invitrogen, USA). The cDNA then got amplified by 3 sets of primers (LLOV_400bp pool, LLOV_500bp pool, LLOV_2000bp pool; primers were: 400bp, 500 bp and 2000bp long), the enzyme used was the Q5 Hot Start HF Polymerase. The PCR protocol was the following: 98°C 30 sec, 24-40 cycles 98°C 15 sec, 65°C 5 minutes. Cycle number determination was based on the previous qRT-PCR reactions CT number, by adding 3 to it.

The amplicon PCR product was cleaned with 0,6x AMPure beads (AMPure XP beads, Beckman Coulter, USA). The washing steps were done using 70% ethanol, then the samples

were eluted in nuclease-free water. The concentration of the nucleic acid in our samples were measured by using Qubit dsDNA HS Assay Kit (Invitrogen, USA) and the Qubit 4 fluorometer (ThermoFisher). This step was followed by the end preparation of the amplicon sequences by NEBNext Ultra II End Repair/dA-Tailing Module (New England Biolabs, USA). The next step was the barcoding of the amplicon sequences (SQK-NBD114.96, ONT) this was ligated by NEBNext® Ultra™ II Ligation Module (New England Biolabs, USA). Because of the sequences having their unique barcode sequences we can measure the content of all tubes into one for further reactions. Another cleaning with AMPure beads happened, however the elution this time was the sequencing kits elution buffer (EB, SQK-NBD114.96, ONT). The last step was ligating the sequencer adapter NA (Native Adapter) with the help of Quick T4 ligase (New England Biolabs, USA). The prepared library again went under an AMPure beads cleaning, but this time for the washing steps we used the sequencing kits short-fragment buffer. The sequencing was run on R10.4.1 flowcells for 24-72 hours. Minimal quality score was set to Q7 for the fastbasecalling during the run.

3.5 Bioinformatic processing of the sequence data

Every bioinformatic process ran under Ubuntu 22.04 operating system. The first step of the processing of the raw data was basecalling. This was done by using Guppy basecaller (v6.5.7) software with a configuration file: dna_r10.4.1_e8.2_400bps_sup.cfg specific to the flowcell and sequencing rate. For the superaccurate basecalling a minimum quality score of Q10 was set. Next step was demultiplexing which happened with the use of Guppy barcoder (v6.5.7). In this step we sorted the sequences into folders of different barcodes. The trimming of the barcode was done by this software as well.

With the help of NanoPlot (v1.41) we can do a quality check, although we set the minimum quality score to Q10 at this step by seeing that most of the reads exceeded Q12 we decided to go on from this point with only those reads. This was followed by filtering, done by NanoFilt (v.2.0.8). We filtered by quality score (Q12) and lengths of the reads. In case of 400bp reads minimum length was set to 300bps, for 500 bps long reads minimum was 400bps and for 2000 bps long reads minimum was set to 1000 bps. An additional 25-25 bps were trimmed off of the ends getting rid of overlapping primer regions. Mapping happened with minimap2 (v2.24) with the reference being the Spanish 2003 sample (NC_016144).

An additional polishing step was applied with the use of medaka (v2.0.1) software.

This software is responsible for creating the consensus sequence from the mapped sequence, while getting rid of all systemic errors possible made by the sequencer by taking into account the flowcell type. The model we used was: r1041_e82_400bps_sup_g615.

The last step is manual correction. Here, we go through the consensus sequence manually locating any errors and removing parts where the coverage is not minimum 20 reads.

3.6 Phylogenetic tree and similarity plot analysis

For the establishment of the large all-filovirus containing phylogenetic tree (figure no 4.) we assembled a database containing the GenBank identification number, genus, species, sampling location, sampling date, host and the fasta file of the sequences with the help of NCBI virus's large database. Mostly complete genomes were collected (except from species having only 1-2 sequence datas on NCBI) to contribute to more accurate results of analyses. We also made sure to use sequences with known sampling location and date to be able to carry out analyses like geo-ecological studies or a time calibrated phylogeny in the future. The software we used for the study was RAxML-NG (v1.2.2). This software is a phylogenetic tree inference tool which uses maximum-likelihood method. It offers great speed, flexibility and a user-friendly interface. For the construction of the tree firstly after uploading the files we ran a test to see what model the software recommends. We set the bootstrap number to 1000 and started the test run. The model selected was the general time reversible with additional gamma parameter set to the default 4 categories (GTR+G4) model. It took several hours to complete the analysis due to the large dataset, but eventually we were left with plentiful of files from which the .xml file was used further. For visualization we used iTOL (Interactive Tree Of Life), an online tool for displaying phylogenetic trees.

For the phylogenetic tree containing the LLOV genomes we were consistent with the softwares and after the test run the same model was the recommended one (GTR+G4). Only difference was the bootstrap number for what we used 10.000. This change however was not significant, with the softwares live tracking of the process we saw the same outcome from the ~500th bootstrap tree, but because of the much smaller dataset still notably less resources and time was necessary. Similarly, iTOL was used for visualization. All the fasta files of the sequences are part of the research group's own database (except the reference sequence), only the reference sequence (Negredo et al. 2011) and the sequence of the Italian sample (Toth et al. 2023) is available on NCBI.

The similarity plot analysis was made by using SimPlot++ software. The sequences for the analysis were selected so not all sequences are from one pandemic but also not too far in time in case of EBOV, while LLOV sequences were randomly selected. We selected a simple but useful model, the Kimura 2-parameters, which not only detects if the given nucleotide is the same or not but the type of mutation (transition or transversion). Other parameters were set manually, the sliding window to 200bps and the steps to 20bps.

6. Results and discussion

Altogether I processed (nucleic acid extraction, PCR) $n=376$ samples that were collected in the year of 2024. Based on the type of samples $n=122$ was blood sample and $n=254$ other kind of sample, mainly batflies, out of blood samples $n=4/376$ were positive for LLOV. Rest of the sequences used for the analyses were already established in the research groups own database or a central database of NCBI. The sequencing of these samples were carried out and the consensus sequence was included in the following analyses.

As figure no 4. demonstrates LLOV genetically falls between Marburg and Ebolavirus. It's genome contains parts originated from Marburg virus, but also have similarities to Ebola virus. It is also rather visible, that LLOV exhibits relatively small amount of evolution unlike Marburg and Ebolaviruses. This is partly due to the lack of data, mostly because of the immensely narrow interval of sample collection. Note that the very first sequence data is originated from the first sample found in 2003; since then only 22 years passed and not in all we were able to find a positive sample. An additional similarly simple reason is the property of certain species of Marburg and Ebola genus to be able to infect humans and thus cause epidemics. While LLOV is steadily circulating in European bat colonies, EBOV and MARV has onset several outbreaks in which the viruses were exposed to much higher chances of mutation with each cross infection between people.

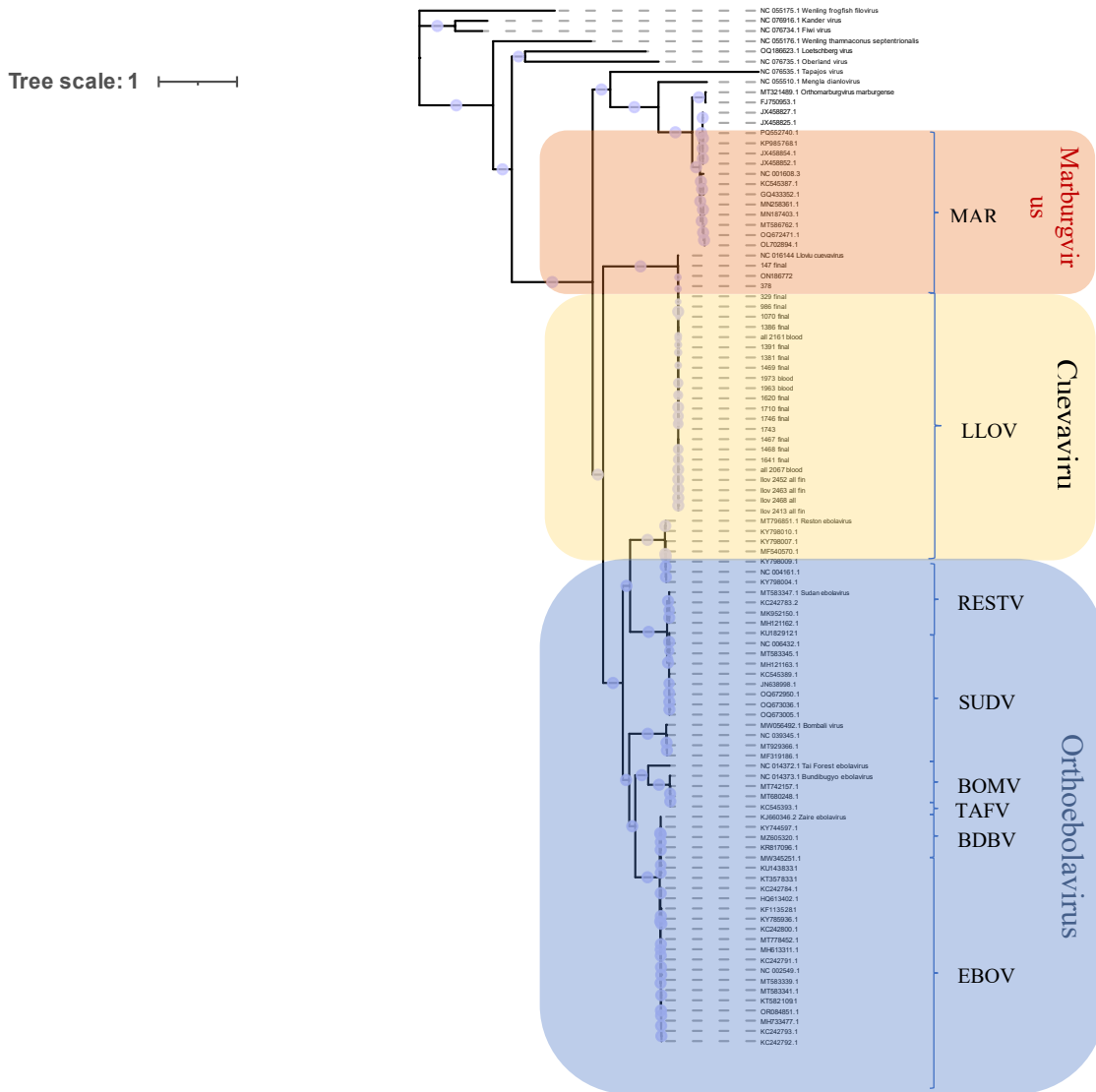


Figure 4. Phylogenetic tree including every filovirus species, for an overall visualisation of their evolution. The tree is a maximum likelihood tree, using GTR+G4 model with 1000 bootstraps. For outgroup we used the Wenling frogfish filovirus.

Figure 5. contains every sequence from all positive samples for LLOV. By comparing the branch lengths to the tree scale, the small amount of evolution is visible on this figure as well. However we see the sequences derived from samples collected from different countries mix with each other. This similarity in the genomes is caused by the mixing of the bat colonies. In such small area, as Central-Europe the relatively small distances between countries is not

causing any problem for the bats to fly through. It is principally characteristic for the mating season. This pattern proves us the virus's enzootic behaviour, meaning the virus had been present in Europe for very long time, stably circulating in the bat populations living here. Future plan of our research team is to acquire samples from locations much further from Central-Europe, for example Spain where the first detection of the virus happened. Involving those samples in phylogenetic studies we expect a similar pattern of evolution, creating a much more separated clade similar to RAVN strain in case of Marburg virus (Peterson and Holder 2012).

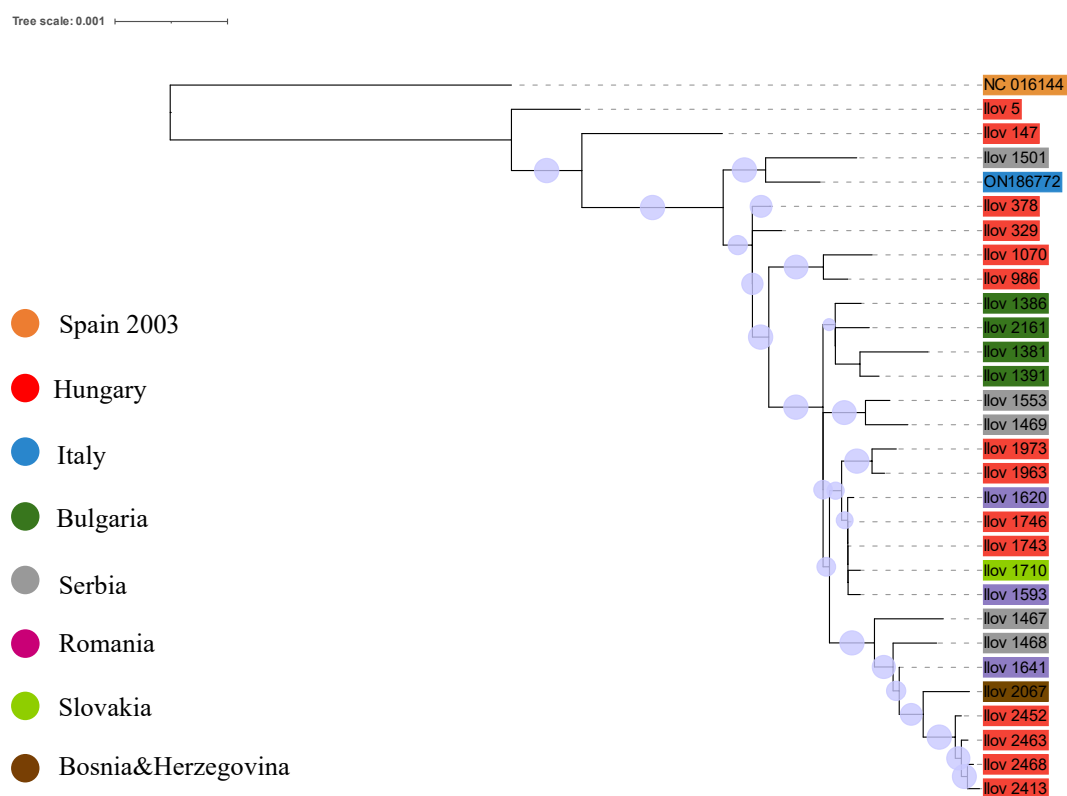


Figure 5. Phylogenetic tree of LLOV from own data. Samples were collected: 2016-2024. Maximum likelihood tree, we used GTR+G4 model, 10.000 bootstraps. Only the bootstrap values over 0,7 is visualized. Outgroup is the 2003 Spanish sample.

Similarity plot is an analysis that compares the same nucleotide position in the examined genomes. Kimura 2-parameter model. By using the simplified figure of the genomes of the viruses, like seen on top of the two diagrams we can easily visualize which protein's coding

region is undergone the most mutation. We obtained the expected results, in case of both viruses we see that the glycoprotein coding region accumulates the most mutations. The glycoprotein has the most significant role in receptor binding, thus in the first step of entering a cell, making it require the largest amount of adaptation which happens via mutations. Not only the rate of mutation is high in this region, but the non-synonym/synonym ratio exceeds 1 exclusively in the glycoprotein coding region (Toth et al. 2023), meaning it is under positive selection. Mutations, if causing amino acid change (non-synonym) might result in change of the structure of the glycoprotein, thus changing its ability to bind to receptors. Another reason making the analysis of the GP coding region very important is that, the neutralization serological test is GP based. To maintain accuracy and sensitivity of this reaction we need to keep the sequence of the gene under surveillance.

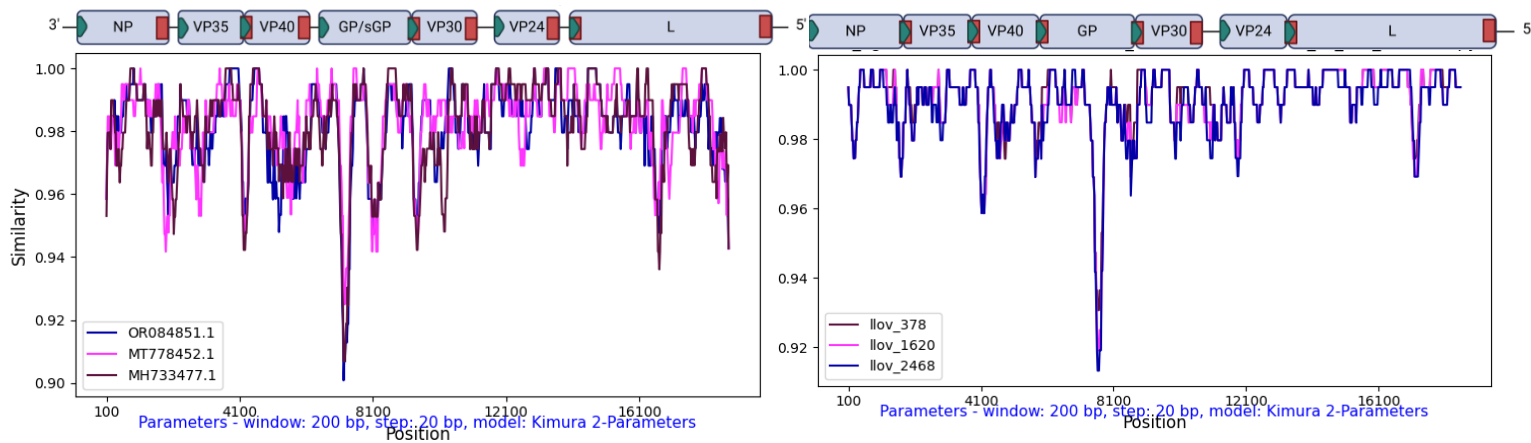


Figure 6. Similarity plot of EBOV and LLOV. Kimura 2-parameter model was used.

4 new genomes were established from samples collected from a different sampling location in Hungary. Although the samples exhibited high ct number (35, 36, 33, 38) shown in the PCR reaction meaning low copy number of the virus we were still able to sequence almost the complete genome, the coding region is completely gapless. This confirms the proper functioning of our sequencing system, enabling in the future to achieve our goal of establishing the complete genome without any gaps.

In recent studies for phylogenetic analysis of filoviruses very often only the reference sequence, the 2003 Spanish genome is used for representing the *Cuevavirus* genus (Makenov et al. 2023, He et al. 2023) Not less than 22 years passed since, while the genome of the virus had undergone several mutations that could have altered the virus's structure or certain behaviors. It is important thus to keep track of the evolution of the microbe on a larger scale, involving the newly obtained sequences in the phylogenetic study of all filoviruses.

7. Conclusions

With our research we strengthened the suggestion of LLOV possessing enzootic property. Based on our results we can say the virus is stably circulating in European populations, however still holding zoonotic potential. We also successfully expanded the knowledge on geographic distribution of the virus, by involving sequences obtained from samples from more and more countries in the phylogenetic analyses. We are able to sequence almost the complete genome of a sample even containing very low copy number of the virus. This fact exhibits how developed our protocol of sequencing is, representing that everything is given for the further improvement and maybe even completing the genome of the virus.

The analysis needed for better understanding of the origin of the virus is still in progress due to its complexity and time requirements. In our future research we would like to expand sampling locations and acquire samples from further away from Central-Europe, and our intentions are to start testing different bat species, these data could serve as a great basis for more complex, comparative analyses.

8. Acknowledgement

I would like to thank my two supervisors Dr. Gábor Kemenesi and Dr. Zsófia Lanszki and Ágota Ábrahám for their tiring and devoted help they provided. Also thanks to all the colleagues of the National Laboratory of Virology who helped at any point of my work. I also owe a thank you to the chiropterologists and nature conservation specialists for their work in sample collection.

9. References

Brook, C. E., & Dobson, A. P. (2015). Bats as ‘special’reservoirs for emerging zoonotic pathogens. *Trends in microbiology*, 23, 172-180.

doi: [10.1016/j.tim.2014.12.004](https://doi.org/10.1016/j.tim.2014.12.004)

Chen, L., Liu, B., Yang, J., & Jin, Q. (2014). DBatVir: the database of bat-associated viruses. *Database*, 2014, bau021.

<https://doi.org/10.1093/database/bau021>

Wang, L. F., Shi, Z., Zhang, S., Field, H., Daszak, P., & Eaton, B. T. (2006). Review of bats and SARS. *Emerging infectious diseases*, 12, 1834.

<https://doi.org/10.3201/eid1212.060401>

Swanepoel, R., Smit, S. B., Rollin, P. E., Formenty, P., Leman, P. A., Kemp, A., et al. Paweska, J. T. (2007). Studies of reservoir hosts for Marburg virus. *Emerging infectious diseases*, 13, 1847.

doi: [10.3201/eid1312.071115](https://doi.org/10.3201/eid1312.071115)

Kaner, J., & Schaack, S. (2016). Understanding Ebola: the 2014 epidemic. *Globalization and health*, 12, 1-7.

<https://doi.org/10.1186/s12992-016-0194-4>

Marzi, A., Feldmann, H., Geisbert, T. W., & Falzarano, D. (2011). Vesicular stomatitis virus-based vaccines for prophylaxis and treatment of filovirus infections. *Journal of bioterrorism & biodefense*, .

[10.4172/2157-2526.S1-004](https://doi.org/10.4172/2157-2526.S1-004)

Pappalardo, M., Collu, F., Macpherson, J., Michaelis, M., Fraternali, F., & Wass, M. N. (2017). Investigating Ebola virus pathogenicity using molecular dynamics. *BMC genomics*, 18, 35-45.

<https://doi.org/10.1186/s12864-017-3912-2>

Booth, T. F., Beniac, D. R., Rabb, M. J., & Lamboo, L. L. (2015). Filovirus structure and morphogenesis. In *Biology and Pathogenesis of Rhabdo-and Filoviruses* 427-451.

https://doi.org/10.1142/9789814635349_0017

Mühlberger, E. (2007). Filovirus replication and transcription. *Future virology*, 2, 205-215.

<https://doi.org/10.2217/17460794.2.2.205>

Negredo, A., Palacios, G., Vázquez-Morón, S., González, F., Dopazo, H., Molero, F., et al. (2011). Discovery of an ebolavirus-like filovirus in europe. *PLoS pathogens*, 7, e1002304.

<https://doi.org/10.1371/journal.ppat.1002304>

Tóth, G. E., Hume, A. J., Suder, E. L., Zeghib, S., Ábrahám, Á., Lanszki, Z., et al. Kemenesi, G. (2023). Isolation and genome characterization of Lloviu virus from Italian Schreibers's bats. *Scientific Reports*, 13, 11310.

<https://doi.org/10.1038/s41598-023-38364-7>

Goletic, S., Goletic, T., Omeragic, J., Supic, J., Kapo, N., Nicevic, M., et al. Alic, A. (2023). Metagenomic Sequencing of Lloviu Virus from Dead Schreiber's Bats in Bosnia and Herzegovina. *Microorganisms*, 11, 2892.

<https://doi.org/10.3390/microorganisms11122892>

Hume, A. J., Heiden, B., Olejnik, J., Suder, E. L., Ross, S., Scoon, W. A., et al. Mühlberger, E. (2022). Recombinant Lloviu virus as a tool to study viral replication and host responses. *PLoS pathogens*, 18, e1010268.

<https://doi.org/10.1371/journal.ppat.1010659>

Fletcher, P., Feldmann, F., Takada, A., Crossland, N. A., Hume, A. J., Albariño, C., et al. Marzi, A. (2023). Pathogenicity of Lloviu and Bombali Viruses in Type I Interferon Receptor Knockout Mice. *The Journal of Infectious Diseases*, 228, S548-S553.

<https://doi.org/10.1093/infdis/jiad226>

Fletcher, P., O'Donnell, K. L., Rhoderick, J. F., Henderson, C. W., Okumura, A., Bushmaker, T., et al. Marzi, A. (2024). Lack of Lloviu Virus Disease Development in Ferret Model. *Emerging Infectious Diseases*, 30, 2639.

<https://doi.org/10.3201/eid3012.240818>

Lam, T. T. Y., Hon, C. C., & Tang, J. W. (2010). Use of phylogenetics in the molecular epidemiology and evolutionary studies of viral infections. *Critical reviews in clinical laboratory sciences*, 47, 5-49.

<https://doi.org/10.3109/10408361003633318>

Morse, S. S., Mazet, J. A., Woolhouse, M., Parrish, C. R., Carroll, D., Karesh, W. B., et al. Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *The Lancet*, 380, 1956-1965.

[10.1016/S0140-6736\(12\)61684-5](https://doi.org/10.1016/S0140-6736(12)61684-5)

Suzuki, Y., & Gojobori, T. (1997). The origin and evolution of Ebola and Marburg viruses. *Molecular biology and evolution*, 14, 800-806.

<https://doi.org/10.1093/oxfordjournals.molbev.a025820>

Carroll, S. A., Towner, J. S., Sealy, T. K., McMullan, L. K., Khristova, M. L., Burt, F. J., et al. Nichol, S. T. (2013). Molecular evolution of viruses of the family Filoviridae based on 97 whole-genome sequences. *Journal of virology*, 87, 2608-2616.

<https://doi.org/10.1128/jvi.03118-12>

Peterson, A. T., & Holder, M. T. (2012). Phylogenetic assessment of filoviruses: how many lineages of Marburg virus?. *Ecology and Evolution*, , 1826-1833.

<https://doi.org/10.1002/ece3.297>

Hierweger, M. M., Koch, M. C., Rupp, M., Maes, P., Di Paola, N., Bruggmann, R., et al. Seuberlich, T. (2021). Novel filoviruses, hantavirus, and rhabdovirus in freshwater fish, Switzerland, 2017. *Emerging infectious diseases*, 27, 3082.

[10.3201/eid2712.210491](https://doi.org/10.3201/eid2712.210491)

Makenov, M. T., Le, L. A. T., Stukolova, O. A., Radyuk, E. V., Morozkin, E. S., Bui, N. T., et al. Karan, L. S. (2023). Detection of filoviruses in bats in Vietnam. *Viruses*, 15, 1785. <https://doi.org/10.3390/v15091785>

He, B., Hu, T., Yan, X., Pa, Y., Liu, Y., Liu, Y., .et al. Tu, C. (2024). Isolation, characterization, and circulation sphere of a filovirus in fruit bats. *Proceedings of the National Academy of Sciences*, 121, e2313789121. <https://doi.org/10.1073/pnas.2313789121>

Protocols.io (2025) <https://www.protocols.io/view/lloviu-cuevavirus-sequencing-protocol-q26g78bp3lwz/v1>