ViBioM

International Virus Bioinformatics Meeting

Vilnius, Lithuania | 18-20 May 2026

ABSTRACT BOOK

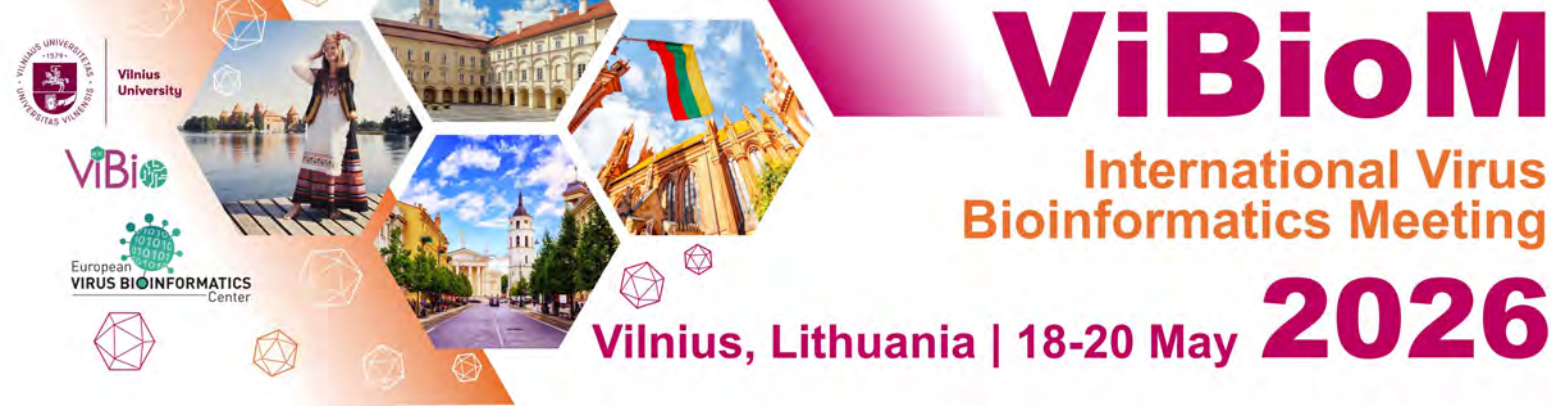
https://evbc.uni-jena.de/events/vibiom2026/

ViBioM

International Virus Bioinformatics Meeting

Vilnius, Lithuania | 18-20 May 2026

Contents



ViBioM

International Virus Bioinformatics Meeting

Vilnius, Lithuania | 18-20 May 2026

Organizers

European Virus Bioinformatics Center (EVBC), Jena, Germany
University Vilnius, Vilnius, Lithuania
Non-profic association ViBio e. V., Jena, Germany

European VIRUS BIOINFORMATICS Center

VILNIAUS UNIVERSITETAS 1579 UNIVERSITAS VILNENSIS

Vilnius University

e.V. ViBio

Organizing Committee

- **Gytis Dudas**, Vilnius University, Vilnius, Lithuania
- **Maria Fabisch**, European Virus Bioinformatics Center, FSU Jena, Germany
- **Manja Marz**, European Virus Bioinformatics Center, FSU Jena, Germany
- **Ingrida Olendraitė**, RdRp Summit, European Virus Bioinformatics Center

Scientific Committee

- **Joana Abrantes**, University of Porto, Portugal
- **Arif Nur Muhammad Ansori**, Universitas Airlangga, Surabaya, Indonesia
- **Evangelia Georgia Kostaki**, National and Kapodistrian University of Athens, Greece
- **Alejandro Matia**, Stanford University, Stanford, CA, USA
- **Maximilian K. Nocke**, Ruhr University Bochum, Germany
- **Sofia Paraskevopoulou**, Robert Koch Institute, Berlin, Germany
- **Sunit K. Singh**, Institute of Medical Sciences, Banaras Hindu University, Varanasi, India

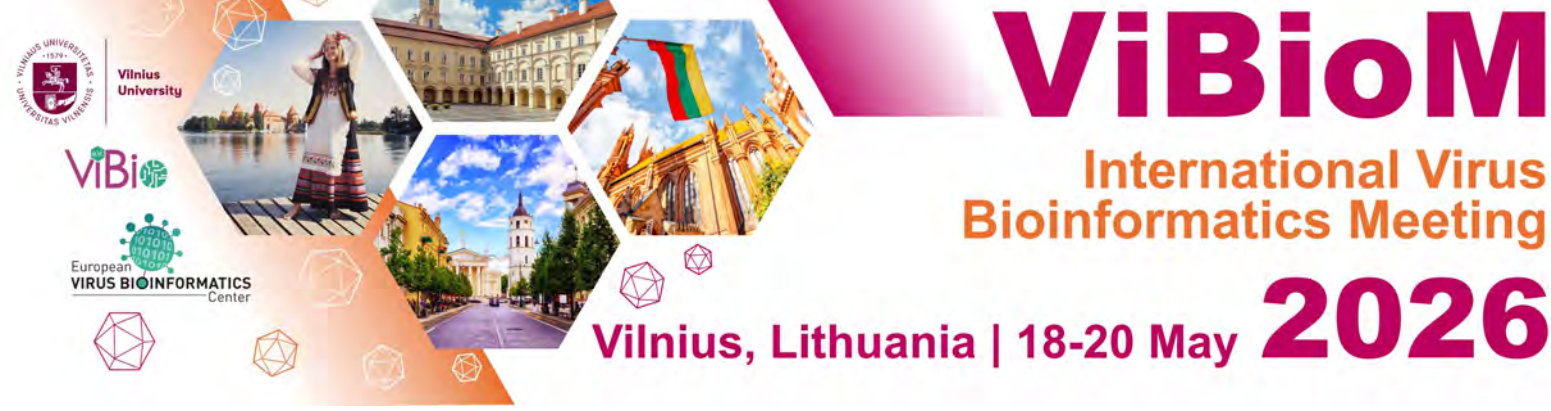
ViBioM

International Virus Bioinformatics Meeting

Vilnius, Lithuania | 18-20 May 2026

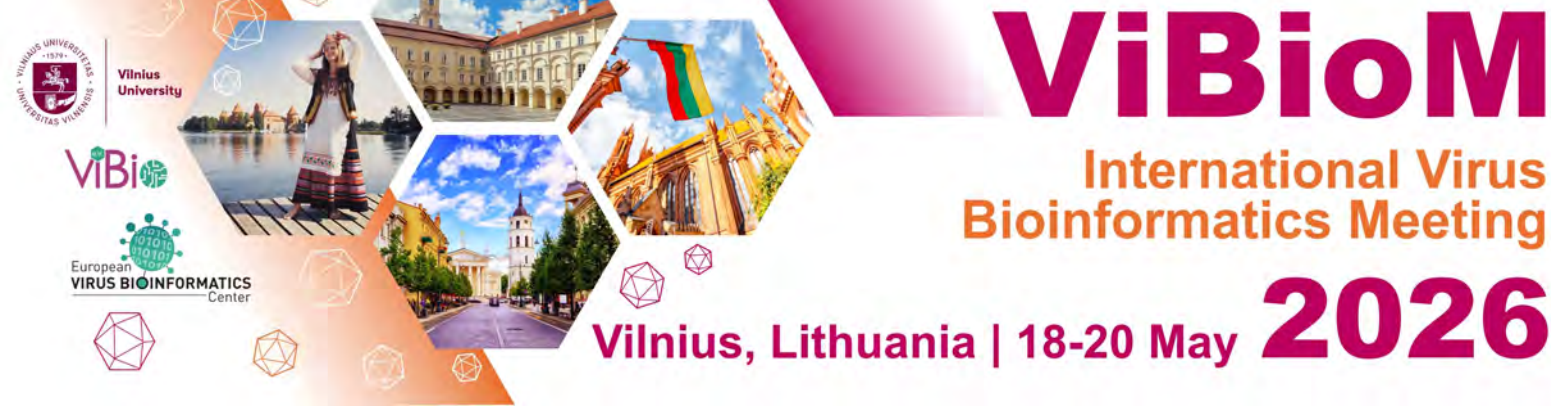
Programme Monday – May 18

08:30	Registration Desk Opens
09:00–09:30	**Welcome and Opening Remarks** Manja Marz, Gytis Dudas, Ingrida Olendraitė
Session 1	**Machine Learning Methods**
09:30–09:50	**T01 – VTT-Net: Learning Viral Tissue Tropism Using Graph Neural Networks** Haley Stone | University of Glasgow, UK
09:50–10:10	**T02 – Phylogenetically structured machine learning allows for interpretable predictions of paramyxovirus hosts** James Herzig | University of Glasgow, UK
10:10–10:30	**T03 – Family-level classification of viral contigs using deep learning** Emma Soufir | Université de Montpellier, France
10:30–11:00	Coffee Break
Session 2	**Metagenomics**
11:00–11:30 *Keynote Talk*	**T04 – Logan: Planetary-scale assembly of DNA/RNA sequencing data and its applications in virology** Rayan Chikhi | Institut Pasteur, Paris, France
11:30–12:00 *Keynote Talk*	**T05 – Illuminating Viral Dark Matter** Ingrida Olendraitė | EVBC, RdRp Summit
12:00–13:30	Lunch Break
13:30–14:00	**Poster Pitches/Snapshots A**
14:00–15:30	**Poster Session A**
15:30–16:00	Coffee Break
Session 3	**Mutational Effects**
16:00–16:20	**T06 – Using in-silico DMS to identify conservation and change in viral proteins** Robert Strange | University of Glasgow, UK
16:20–16:40	**T07 – V-gTK and V-gDB: A Modular Framework for Interpreting Mutation in Viral Genome** Joseph Hughes | University of Glasgow, UK



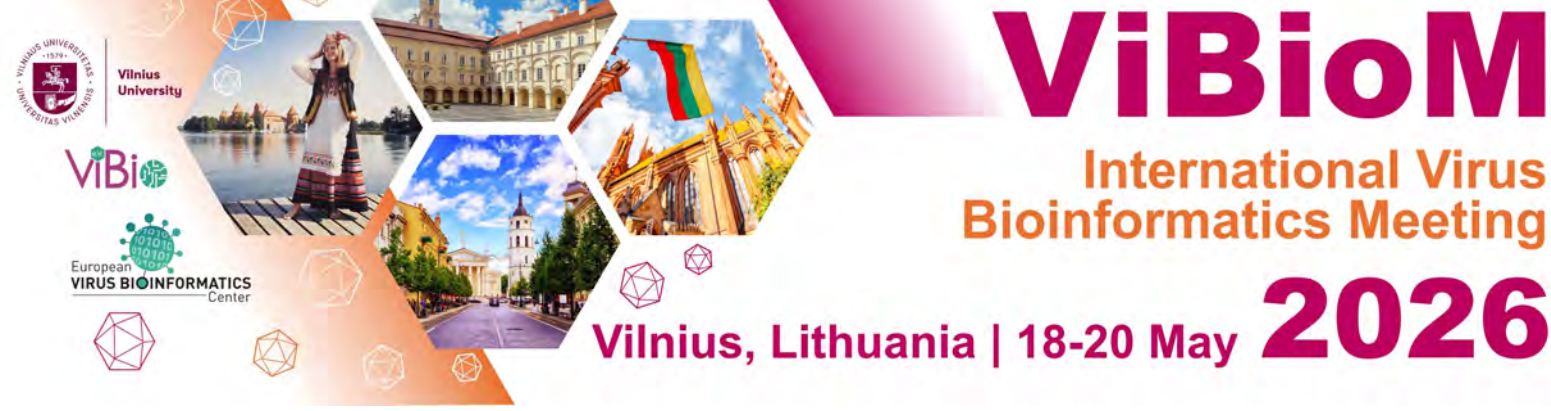
Programme Tuesday – May 19

08:30–09:00	**Announcements & Group Photo**
Session 4	**Virus Proteomics**
09:00–09:30 *Keynote Talk*	**T08 – Comparative Proteomic and Structural Analysis of Newly Discovered (Pro)phages within the Phylum Cyanobacteriota** Darius Kazlauskas | Vilnius University, Lithuania
09:30–09:50	**T09 – Application of AlphaFold in virology: from predicting protein structures to modelling protein-protein interactions** Ulad Litvin | University of Glasgow, UK
09:50–10:10	**T10 – Variable evolutionary routes for the gain and loss of the upstream open reading frame in enteroviruses** Katy Brown | University of Cambridge, UK
10:10–10:30	**T11 – Benchmarking sequence, embedding and structure-based methods for eukaryotic virus protein annotation** Lander De Coninck | KU Leuven, Belgium
10:30–11:00	Coffee Break
Session 5	**Ancient Virus Research**
11:00–11:20	**T12 – Exploring historical retroviral transmission patterns in vertebrates using endogenous retroviruses** Emma Harding | University of Oxford, UK
11:20–11:40	**T13 – A network-based approach to characterizing Endogenous Viral Elements in vertebrate genomes** Laura Muñoz-Baena | University of Oxford, UK
11:40–12:00	**Virus Bioinformatics + nf-core Hybrid Collaborative Hackathon: Outcomes** Shahram Saghaei | Friedrich Schiller University Jena, Germany
12:00–13:30	Lunch Break
13:30–13:45	**Poster Pitches/Snapshots B**
13:45–15:30	**Poster Session B**
15:30–16:00	Coffee Break
16:00–16:45 *Keynote Talk*	**T14 – VirusREvolution: Decoding tools for virus research** Manja Marz | Friedrich Schiller University Jena, Germany
19:00	**Conference Dinner**



Programme Wednesday – May 20

09:15–09:30	**Announcements**
Session 6	**Phylogenomics**
09:30–10:00 *Keynote Talk*	**T15 – Thought over throughput: a public data-fueled deep dive into orthomyxovirus biology** Gytis Dudas | Vilnius University, Lithuania
10:00–10:20	**T16 – Using phylogenetic and phylogeographic approaches to explore the role of reassortment in the diversification of host specificity and associated spread of panzootic HSNI influenza** Will Harvey | Roslin Institute, Edinburgh, UK
10:20–10:50	Coffee Break
Session 7	**RNA Biology**
10:50–11:20 *Keynote Talk*	**T17 – An RNA-centric perspective on human pathogenic RNA viruses** Mathias Munschauer | Medical Faculty Heidelberg University, Germany
11:20–11:40	**T18 – Integrating RNA structure and protein interactions to uncover the mechanisms of viral and cellular ires function** Riccardo Delli Ponti | Italian Institute of Technology, Genoa, Italy
11:40–12:00	**T19 – RdRpCATCH: A unified resource for RNA virus discovery using viral RNAdependent RNA polymerase profile Hidden Markov models** Dimitris Karapliafis | Wageningen University & Research, The Netherlands
12:00–13:30	Lunch Break
13:30–14:30	**Annual Meeting of the EVBC** *open to all interested ViBioM participants*
14:30–15:00	**Closing Ceremony / Awards**



Session 1: Machine Learning Methods

T01 – VTT-Net: Learning Viral Tissue Tropism Using Graph Neural Networks

Haley Stone | University of Glasgow, Glasgow, United Kingdom

Viral tissue tropism is a defining determinant of disease manifestation, transmission routes, and host range, and plays a central role in viral evolution and cross-species emergence. While large host-virus datasets are increasingly available, modelling virus-tissue interactions remains difficult due to difficulty in extrapolating from in-vitro experiments to in-vivo, and limited sampling coverage. We present VTT-Net (Viral Tissue Tropism Network), a graph neural network for modelling virus-tissue interactions and predicting tissue-level tropism. In this study, VTT-Net is applied to human tissues, representing tissue states and viral identities within a unified interaction framework. The network learns tissue-specific tropism patterns directly from observed virus-tissue associations, capturing relationships across tissues, encoded with broad-sample gene expression data, under sparse and uneven supervision. VTT-Net leverages structured relationships between tissues and viruses to integrate interaction information across tissues, enabling tissue-level tropism inference even when direct evidence is sparse. The framework is agnostic to viral taxonomy and tissue system, allowing tropism patterns to be learned across heterogeneous biological contexts. VTT-Net produces tissue-level permissivity scores for virus-tissue pairs and supports downstream analysis stratified by tissue and viral taxonomy. Although demonstrated here in the context of human tissue tropism, the framework is designed to be extensible to additional host species as appropriate data become available. VTT-Net provides a quantitative framework for analysing viral tissue specificity, supporting comparative studies of tropism evolution, host adaptation, and cross-system infection patterns.

T02 – Phylogenetically structured machine learning allows for interpretable predictions of paramyxovirus hosts

James Herzig | University of Glasgow, Glasgow, United Kingdom

Identifying reservoir hosts is a notorious challenge that impedes surveillance and control measures for zoonotic virus prevention. Numerous approaches have been taken to host prediction, including trait-based models, network inference and models utilising genomic features. The latter leverages the fact that viral genomes adapt to their host, with adaptation likely occurring through mimicry of host genomes and adaptations to evade host-specific innate immune responses. However, a consistent problem when building supervised learning models for host prediction arises from phylogenetic structure. Both host preference and genomic features are highly correlated with viral ancestry; as a result, models utilising genomic features learn to reconstruct phylogeny. Feature importance therefore typically measures which features optimally reconstruct viral phylogeny, rather than true and generalisable signals of host adaptation. In order to resolve this issue, we have developed a host prediction model incorporating explicit representation of phylogeny by utilising graph neural networks (GNNs). GNNs are an implementation of neural network architecture that makes use of pairwise message passing between nodes in a graph to make predictions informed by heterogeneously structured data. We cluster viral genomes to generate a representative set covering 189 viral species in the family Paramyxoviridae. We train GNNs using a graph representation of the phylogenetic relations between these sequences. We demonstrate that genomic features do provide predictive power additional to that inferable solely from graph structure. Thorough feature ablation and tests with alternate

graph structures reveal that simple compositional features are predictive and that sparse representations of phylogeny result in optimal predictions. Our modelling approach results in improved host prediction accuracy while allowing for identification of genome features that optimally represent host adaptation.

T03 – Family-level classification of viral contigs using deep learning

Emma Soufir | Université de Montpellier, Montpellier, France

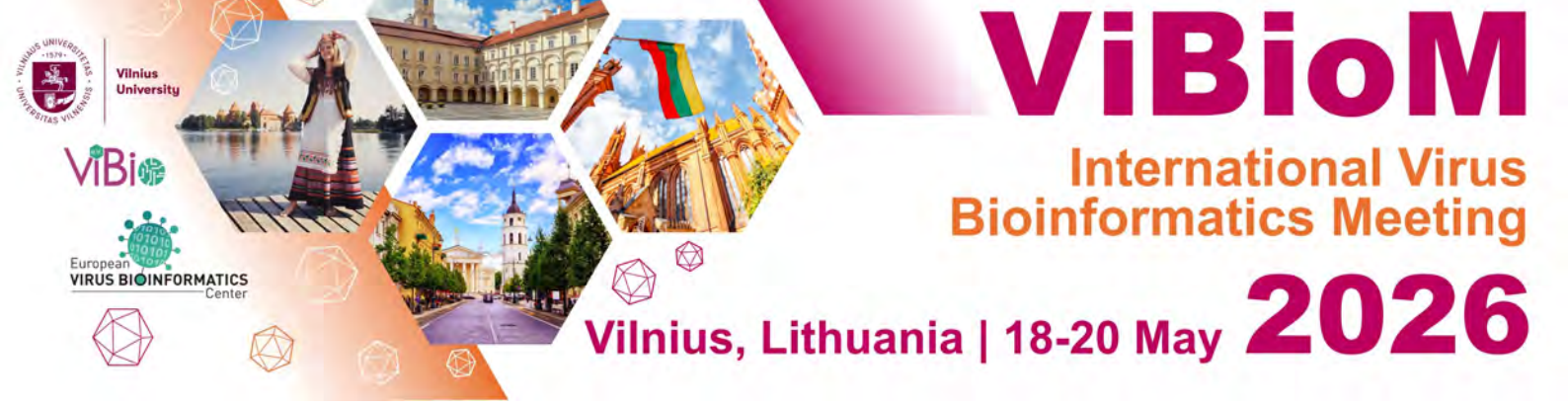
Metagenomics applied to viruses has greatly improved the ability to study virus diversity. However, this approach has also come with its own challenges. Among others, sequence assemblies often result in fragmented genomes which complicate taxonomic identification using traditional methods. Indeed, most taxonomic classification methods rely heavily on sequence similarities against reference databases or marker gene detection, and therefore depend on the completeness and quality of these databases. The limited coverage of viral diversity in these databases often leads to inaccurate or unresolved annotations. At the same time, recent advances in deep learning, particularly the use of large language models with genomic sequences, offer promising solutions for sequence-based classification tasks. Here, we present a method for family-level classification of viral contigs using a fine-tuned DNABERT-2 model. For training, we generated a curated viral sequence database. We used the ICTV MSL to define the taxonomy and focused specifically on eukaryotic infecting viruses. Additional sequences were also retrieved to expand the dataset. Genomes were cut into 1 000 b fragments to mimic contig-like sequences. To better reflect real scenarios, we investigated several training strategies designed to simulate the presence of previously unseen species, thereby evaluating the model's ability to generalize beyond known taxa. We also explored different training optimizations, like data augmentation, to enhance predictive performance and robustness. We provide a benchmark comparing multiple methods, like Kraken, evaluated on simulated and real metagenomic datasets. Our results highlight the potential of transformer-based models for accurate viral taxonomic classification in metagenomic studies. We also analyze the errors made by the model to better understand its limitations and strengths in comparison with traditional methods.

Session 2: Metagenomics

T04 – Logan: Planetary-scale assembly of DNA/RNA sequencing data and its applications in virology (*Keynote Talk*)

Rayan Chikhi | Institut Pasteur, Paris, France

Petabytes of valuable DNA sequencing data reside in public repositories, roughly doubling in size every three years. They contain a wealth of genetic information about viruses, bacteria, animals, humans. We have developed two bioinformatics cloud infrastructures, Serratus and Logan, to perform petabase-scale sequence analysis. In particular, such an infrastructure is useful for applications in virology. With Serratus, we analyzed 10 petabases of public RNA-seq samples and discovered 10x more RNA viral species than previously known (Edgar et al, Nature, 2022). In Logan, we are making Earth's sequencing data more accessible by reducing its size by 100x without significant loss of information. In this talk, I will give an overview of Logan and recent progress on improving its accessibility, and present some of the work we have done on mining RNA viruses and satellites.



T05 – Illuminating Viral Dark Matter (*Keynote Talk*)

Ingrida Olendraitė | EVBC, RdRp Summit

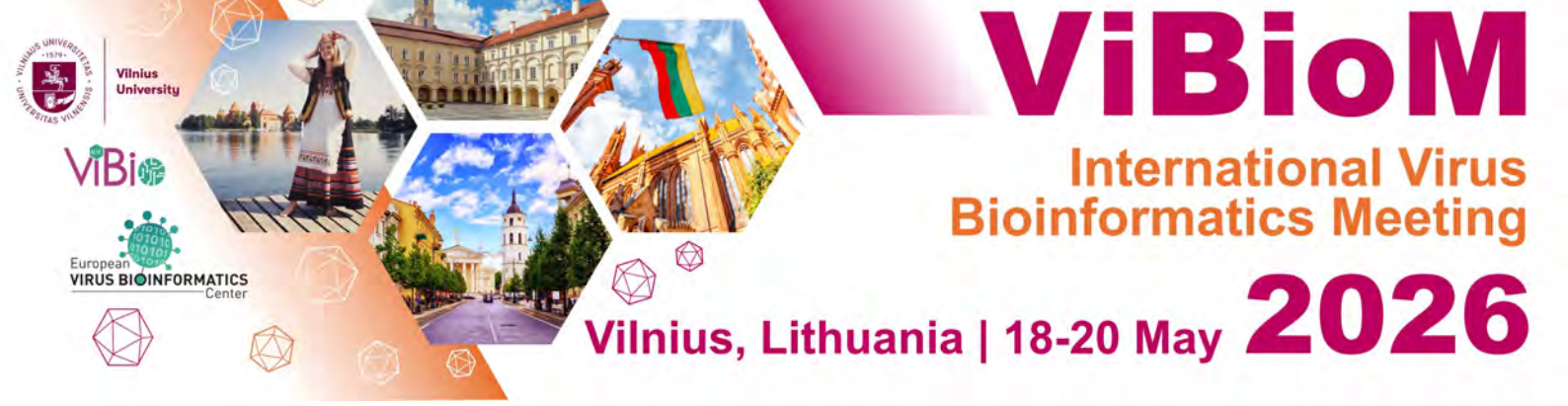
Over the past decade, large-scale sequencing and mining of public datasets have transformed RNA virus discovery, revealing vast numbers of candidate viruses through identification of RNA-dependent RNA polymerase (RdRp) sequences. However, finding more RdRp sequences is only the beginning. Much less attention has been given to turning these large-scale findings into high-quality assemblies and genomes, reliable annotations, robust alignments and evolutionary interpretation, realistic virus-host associations, and computational results ready for experimental validation. In this talk, I will illustrate my approach to deeper data mining with an example of a novel mononegavirales-like virus in which we identified unusual splicing in the RdRp core region. By detailed interrogation of the contig, we showed that this represented genuine splicing-dependent RdRp expression. In addition, while expanding the number of sequences in the Polycipiviridae family by over 100-fold overall, this increase did not illuminate the unknown biology, but instead uncovered substantial hidden diversity. These examples show that we are nowhere near saturation of virus discovery and, at the same time, we are reaching a turning point in our research community, with a need to pause and investigate rather than focus primarily on detection. Finally, I will highlight the “RdRp Summit” efforts to tackle the field issues discussed above. The initiative has evolved into a strong collaborative community for RNA virus discovery researchers, where people share practical advice, discuss tools and datasets, work on collaborative tools and joint publications.

Session 3: Mutational Effects

T06 – Using in-silico DMS to identify conservation and change in viral proteins

Robert Strange | University of Glasgow, Glasgow, United Kingdom

Predicting mutational impact for viruses can inform phenotypic and drug-resistance studies, vaccine design and variant characterisation. In-silico deep mutational scanning (DMS) offers an inexpensive alternative to experimental DMS for assessing functional consequence. However, in previous work we’ve shown a lack of correlation between in-silico and experimental DMS variant effect predictions. To investigate further we compare in-silico and experimental methods for predicting Variant of Concern (VOC) mutations across the SARS-CoV-2 pandemic. We demonstrate, that in-silico and experimental DMS show comparable performance, varying based upon assay and model. While the performance of methods is similar, they predict an alternative set of mutations across the sequence. The discordant predictions are non-random and reflective of how different methods infer mutational effect. Alignment based and protein language model in-silico methods act as a ‘evolutionary assay’, assessing where in the sequence change is readily accommodated. Experimental DMS captures context specific mechanisms but may misjudge mutations in regions not assessed by the chosen assay. By combining the experimental and in-silico results, 98% of VOC mutations were able to be identified, whereas only 58% were identified within an in-silico method and 47% in a single experimental assay. Our findings confirm that in-silico methods can be used effectively to predict the emergence of mutations within VOCs. Importantly in-silico methods find a comparable but alternate set of VOC mutations to experimental methods, suggesting the in-silico DMS should not be directly compared to experimental results. Combining both experimental and in-silico results significantly improves VOC mutation prediction.



T07 – V-gTK and V-gDB: A Modular Framework for Interpreting Mutation in Viral Genome

Joseph Hughes | University of Glasgow, Glasgow, United Kingdom

The rapid expansion of viral genome sequencing has created new opportunities to investigate viral evolution, host adaptation, and antiviral resistance. However, deriving meaningful biological insight requires computational frameworks that account for viral genome architecture and evolutionary relationships in the context of mutations. Platforms such as GLUE and Nextstrain have demonstrated the value of integrated systems for viral sequence analysis, particularly in genomic epidemiology and outbreak surveillance. Here we present the Virus Genome Toolkit (V-gTK) and Virus Genome Database (V-gDB), a modular and reproducible ecosystem for managing and interpreting viral genome data with a focus on biologically meaningful mutations. V-gTK provides tools for end-to-end management of viral genomes, including data retrieval, metadata validation, sequence filtering, alignment, and phylogenetic analysis. The framework emphasizes reproducibility through standardized input formats, version control, and transparent data provenance, while enabling collaborative curation by virologists and bioinformaticians. V-gDB complements this infrastructure by supporting the creation of searchable databases linking genome sequences with curated mutation data, reference phylogenies, and contextual metadata. The framework is implemented in Python and integrates established bioinformatics tools including BLAST, MAFFT, NextAlign, IQ-TREE, and EPA-NG. Metadata validation incorporates controlled vocabularies such as United NationsM49 country codes and the National Center for Biotechnology Information Taxonomy database to ensure consistent and high-quality datasets. We demonstrate the versatility of the system using two examples: HCV and Rabies virus. These case studies illustrate how V-gTK and V-gDB enable the construction of durable, mutation-aware viral genome resources that support deeper interpretation of viral genomic variation.

Session 4: Virus Proteomics

T08 – Comparative Proteomic and Structural Analysis of Newly Discovered (Pro)phages within the Phylum Cyanobacteriota (*Keynote Talk*)

Darius Kazlauskas | Vilnius University, Vilnius, Lithuania

The phylum Cyanobacteriota, comprising the classes Cyanobacteriia, Vampirovibrionia, and Sericytochromatia, represents a critical lineage of bacterial diversity. While some lytic cyanophages infecting Cyanobacteriia are well-characterized, our knowledge of (pro)phages in the non-photosynthetic Vampirovibrionia and Sericytochromatia remains remarkably sparse. To address this gap, we conducted a comprehensive bioinformatic survey across more than 2,500 genome assemblies from GTDB and NCBI, integrated with viral sequences from Prophage-DB, MTVGD, and ICTV. Using five prophage identification tools we identified approximately 23,000 viral contigs, from which 5,000 sequences were selected for in-depth analysis after CheckV quality filtering. Proteome-based clustering alongside characterized viruses revealed 70 distinct clusters (≥5 members). Notably, 11 of the 14 largest clusters contained no previously characterized members, highlighting a vast reservoir of “dark matter” within the Cyanobacteriota virome. These clusters exhibited strong host-taxonomic specificity, suggesting localized viral spread. To explore the evolutionary history of these viruses, we performed phylogenetic analyses of highly conserved proteins, including terminases, capsids, and integrases, supplemented by AlphaFold3 structural modeling. Our results demonstrate that the genomic compositions of lysogenic and lytic phages typically do not overlap. Furthermore, we observed

a distinct divergence in auxiliary metabolic genes (AMGs); unlike lytic T4/T7-like phages, Cyanobacteriota prophages rarely carry photosynthetic genes, reflecting their specialized ecological niches. This study provides a new framework for understanding the diversity, proteomic composition, and evolution of viruses infecting the Cyanobacteriota phylum.

T09 – Application of AlphaFold in virology: from predicting protein structures to modelling protein-protein interactions

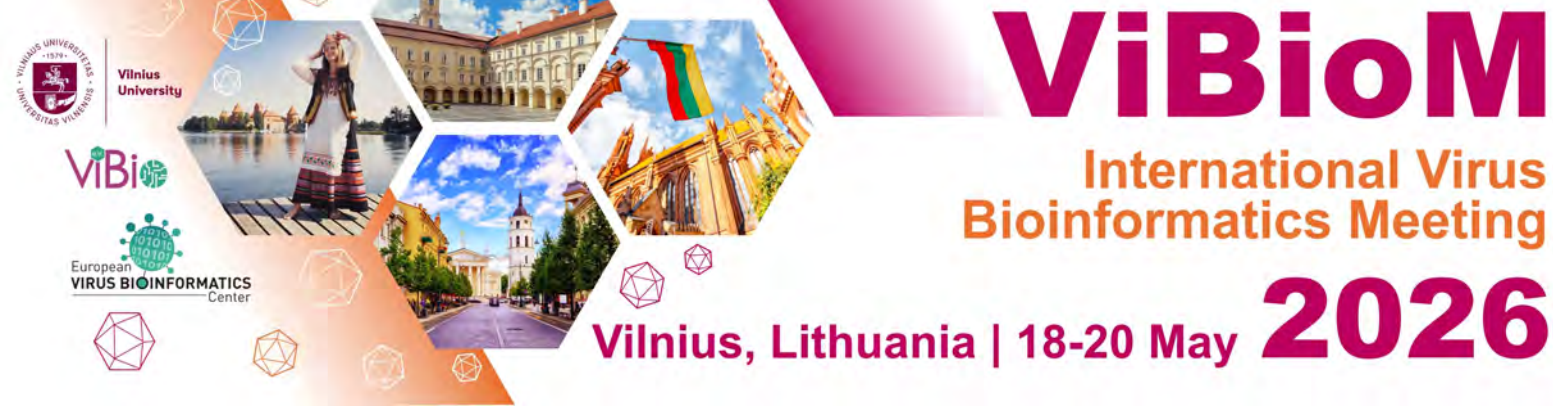
Ulad Litvin | University of Glasgow, Glasgow, United Kingdom

Viruses have been engaged in a continuous evolutionary arms race with their hosts for billions of years. This persistent selective pressure, combined with high mutation rates, drives rapid viral sequence divergence, often obscuring phylogenetic relationships and limiting the effectiveness of sequence-based bioinformatic approaches. In contrast, protein tertiary structures, because of their central role in function, tend to evolve more slowly and are therefore often more informative for identifying distant homologues and investigating deep evolutionary relationships. Although experimental structure determination for all known viral proteins is impractical due to technical and resource constraints, structure prediction tools such as AlphaFold provide a powerful alternative. Here, we present Viro3D (available at viro3d.cvr.gla.ac.uk and alphafold.ebi.ac.uk), a structural database containing ColabFold and ESMFold predictions for 85,000 proteins from 4,400 human and animal viruses. This resource can be used to expand functional annotation and to explore deep evolutionary relationships across viral families. In addition, using the advanced modelling capabilities of AlphaFold3, we modelled thousands of putative protein-protein interactions (PPIs) between 30 human antiviral factors, including members of the OAS, IFITM, and TRIM protein families, and proteins from 10 medically important viruses, such as HIV, Chikungunya virus, and Zika virus. This analysis highlights both the strengths and the current limitations of AlphaFold3 for large-scale PPI modelling. Overall, protein structure prediction tools such as AlphaFold demonstrate substantial value for virology research, enabling large-scale modelling of viral proteomes and systematic exploration of virus-host protein interactions.

T10 – Variable evolutionary routes for the gain and loss of the upstream open reading frame in enteroviruses

Katy Brown | University of Cambridge, Cambridge, United Kingdom

Enteroviruses are highly infectious and environmentally stable human and animal pathogens represented by both enteric and respiratory species. While many infections are mild or asymptomatic, some serotypes can cause severe and even fatal diseases with symptoms ranging from fever, hand foot and mouth disease, myocarditis, viral meningitis, encephalitis, acute hemorrhagic conjunctivitis and acute flaccid paralysis. Many enteroviruses utilise two open reading frames (ORFs) to translate their proteins, the upstream protein (UP) and the main polyprotein. We have analysed the relationship between the presence or absence of the upstream ORF (uORF) with host, environmental and virus-specific factors. By performing a detailed comparative genomic and phylogenetic analysis of available enterovirus sequences and associated data, we elucidate the presence and nature of selection within the uORF region, conservation and predicted properties of UP, as well as associations with infection route, host, tropism and disease complications.



T11 – Benchmarking sequence, embedding and structure-based methods for eukaryotic virus protein annotation

Lander De Coninck | KU Leuven, Leuven, Belgium

The advent of protein language models (pLMs) and structural prediction promises new avenues for protein annotation through structure homology. Earlier work (eg. Phold) has shown protein strucural homology annotation works well for prokaryotic viruses. To leverage these new techniques for protein annotation of eukaryotic viruses, we benchmarked four recent embedding- and structure-based methods in comparison with the gold standard of sequence-based annotation. We selected 10% of the entries in ICTV's VMR v40.2 that did not have bacteria or archaea as designated host (n=1,054), and subsequently retrieved all associated proteins from NCBI (n=11,360). Boltz-2 structure prediction of these proteins produced generally high-quality models (median pLDDT 71.3). Using the structures and protein sequences of the entries in the Big Fantastic Virus Database (BFVD), we queried the VMR proteins with Blastp, Diamond, and MMseqs2 (sequence), TEA and ProstT5-3Di (embedding) and Foldseek and Reseek (structure). Sequence homology-based methods showed a high overlap in hits, whereas the other methods had low intra- and inter-method agreement in hits between themselves and the sequence-based methods. Additionally, all methods improved the number of informative categories (ie. not "hypothetical protein") for the VMR proteins, with embedding and structure methods producing the largest increases. However, when evaluated against proteins with informative GenBank annotations as ground truth (n=2,938), embedding- and structure-based methods agreed less often with the ground truth than sequence-based methods (67% and 76% vs 83%, respectively). Overall, sequence homology remains the most accurate approach for protein annotation of eukaryotic viruses at present.

Session 5: Ancient Virus Research

T12 – Exploring historical retroviral transmission patterns in vertebrates using endogenous retroviruses

Emma Harding | University of Oxford, Oxford, United Kingdom

Retroviruses have co-existed with vertebrates for hundreds of millions of years, having untold effects on their population ecology and evolutionary trajectories. Today, retroviruses are known for their zoonotic risk and oncogenic properties, posing threats as viruses of increasing importance to global health. Endogenous retroviruses (ERVs), formed through integration of the retroviral DNA into the host genome, allow us to study retroviral history through millions of years, uncovering patterns of host and virus competition, co-evolution and historical transmission events. We mined 4,155 vertebrate genomes for ERVs using the BLAST-based Nextflow tool HI-FEVER and classified each ERV using custom Hidden Markov Models. Proviruses were reconstructed from their constituent genes, and their age of integration was estimated through prediction of their LTR divergence with LTRHarvest. Using these methods we identified over 14 million ERVs and over 60,000 proviruses throughout Vertebrata. To focus on relatively recent events, we clustered proviruses with >97% LTR identity with MMSeqs2 to collapse them into viral lineages, based on whole genomes, polymerases and envelopes. Through a combination of phylogenetic and network analysis we explore this data to uncover the diversity of recently circulating retroviruses, historical host jumps, predict host reservoir species and explore viral characteristics that differentiate generalist lineages from specialists. We

identify an abundance of recent retroviruses with a broad host range in artiodactyls, primates, passeriformes and testudines, suggesting these animals may harbour yet-undiscovered circulating retroviruses with low transmission barriers. Interestingly, rodents or bats are not as represented in recent transmission data as may be expected. Finally, we correlate ERV numbers, diversity and transmissions with host life history traits to uncover factors that influence the proliferation and movement of retroviruses over time.

T13 – A network-based approach to characterizing Endogenous Viral Elements in vertebrate genomes

Laura Muñoz-Baena | University of Oxford, Oxford, United Kingdom

Endogenous viral elements (EVEs) are viral sequences integrated into host germlines and inherited as genomic fossils of past infections. They provide evidence of long-term virus–host interactions, yet large-scale comparative frameworks to interpret how EVEs are shared across hosts remain limited. We used HI-FEVER to identify non-retroviral EVEs across nearly 4,000 vertebrate genomes, reconstructing 55,640 viral-like proteins and retaining more than 10,000 high-confidence loci. We focused on the most abundant viral families and clustered reconstructed proteins at 50–90% identity using MMseqs2. We built host networks linking species that shared EVEs within clusters and quantified connections across taxonomic levels. To distinguish shared ancestry from host switching, we inferred orthology using flanking regions and synteny. Most cross-order edges in Bornaviridae and Circoviridae corresponded to orthologous insertions inherited before host diversification. In contrast, many Parvoviridae edges (e.g., between primates and rodents, or bats and artiodactyls) lacked shared genomic context and did not mirror host phylogeny, consistent with independent germline insertions following historical host jumps. Together, these results reveal distinct evolutionary regimes among viral families. Filoviridae and Chuviridae form tightly clustered networks largely restricted to single host orders, whereas Parvoviridae and Bornaviridae show broader cross-class distributions, with Parvoviridae displaying the strongest signal of host switching. Circoviridae connectivity is largely explained by deep ancestral integrations. These patterns demonstrate that EVE-derived host networks capture family-specific modes of persistence and host exchange across vertebrate evolution.

Tuesday Afternoon Talk

T14 – VirusREvolution: Decoding tools for virus research (*Keynote Talk*)

Manja Marz | Friedrich Schiller University Jena, Jena, Germany

The continued emergence and re-emergence of viruses cause global outbreaks and exposes fundamental gaps in our understanding of virus genomes, structures, evolution, and host interactions. Their rapid adaptation and diversity challenge existing analytical frameworks. Current bioinformatic, phylogenetic, and imaging approaches remain limited and insufficiently integrate genomic, structural, and functional data. This highlights the need for advanced tools to characterise pathogenic potential, tissue tropism, and replication strategies in real time. The VirusREvolution project addresses these challenges by developing integrated bioinformatic and photonic tools specifically tailored for viruses. By combining computational, optical, and experimental methods, the consortium aims to facilitate the systematic, high-resolution analysis of virus sequences, structures, interactions, and infection dynamics. Research is organised into three synergistic areas: (A) sequence- and regulation-focused bioinformatics, (B) structure-, function-, and interaction-based

computational tools, and (C) photonic technologies for analysing virus morphology, entry, and signatures, as well as host responses, at high spatial and temporal resolution. Tool development is guided by application to model systems, including SARS-CoV-2 and bacteriophage N4, and subsequently extended across diverse RNA and DNA viruses. By integrating computational, experimental, and optical approaches, VirusREvolution establishes a framework for linking iterative tool development and application. This accelerates hypothesis-driven research into virus identification, classification, and function, while generating scalable, transferable technologies for the broader community. The long-term vision is a comprehensive platform that connects genomic, structural, and phenotypic data, enabling rapid, data-driven responses to emerging virus threats and fundamentally transforming how viruses are studied across disciplines. The presentation will outline key objectives and projects within this framework, highlighting avenues for scientific exchange and collaboration.

Session 6: Phylogenomics

T15 – Thought over throughput: a public data-fueled deep dive into orthomyxovirus biology (*Keynote Talk*)

Gytis Dudas | Vilnius University, Vilnius, Lithuania

Orthomyxoviruses are a moderately sized group of RNA viruses with segmented genomes that includes a number of known vertebrate pathogens. Despite ever higher throughput and cheaper sequencing these days, orthomyxoviruses remain poorly understood owing to frequent inability to identify their segments as such amongst metatranscriptomic “dark matter”, and are virtually uncharacterised outside of influenza and infectious salmon anaemia viruses. Over the last few years our lab has used a combination of computational, experimental and molecular biology approaches to fix this situation. We made use of advances in protein structure prediction combined with sensitive phylogenetic methods to establish the relationships between known orthomyxoviruses to explore the evolution of their surface protein use. A series of ancient horizontal gene transfers for a class III viral membrane fusion protein called gp64 can be inferred, including their directionality involving orthomyxoviruses and unrelated virus groups. Additionally, we identified and recovered a number of orthomyxovirus genome segments from public sequence datasets that would otherwise constitute metatranscriptomic “dark matter”. We propose that most orthomyxoviruses found in arthropods have genomes comprised of eight segments with hints that tick hosts select for reduced genomes, and identify many blindspots where orthomyxovirus genome composition cannot be determined at this time. Finally, we’ve been developing Wuhan mosquito virus 6 (WuMV-6) – a common and globally distributed orthomyxovirus of Culex mosquitoes – into an experimental and evolutionary model system. Our experiments on WuMV-6, its population dynamics, and evolution so far present at least a few exciting paradoxes whose resolution will tell us much about how wild arthropod RNA viruses “experience” the world.

T16 – Using phylogenetic and phylogeographic approaches to explore the role of reassortment in the diversification of host specificity and associated spread of panzootic HSNI influenza

Will Harvey | Roslin Institute, Edinburgh, United Kingdom

Since 2021, subclade 2.3.4.4b A(H5N1) high pathogenicity avian influenza (HPAI) viruses have undergone changes in ecology and epidemiology, causing a panzootic of unprecedented scale in birds with transmission in mammals raising concerns about zoonotic potential. Influenza viruses readily exchange gene segments via reassortment, a mechanism that can drive dramatic phenotypic change. We use phylogenetic approaches to explore apparent shifts in seasonality and host range finding clear differences between lineages generated by reassortment. Focusing on Europe, we show reassortment has influenced evolutionary dynamics producing high fitness genotypes with enhanced transmission and contrast successful generalist genotypes with a specialist (EA-2022-BB) adapted to birds of the order Charadriiformes. We then use phylogeographical analyses to explore how this diversification of host specificity impacts geographical spread finding reassortant-specific drivers, thus the genetic diversification of the viral population has implications for spatiotemporal risk mapping. We discuss these results in the context of complementary laboratory data and consider the opportunities for further divergence of viral traits presented by a diverse viral population with endemicity in different host populations.

Session 7: RNA Biology

T17 – An RNA-centric perspective on human pathogenic RNA viruses (*Keynote Talk*)

Mathias Munschauer | Medical Faculty Heidelberg University, Heidelberg, Germany

RNA viruses pose a significant burden to human health. Identifying host cell factors that bind and regulate viral RNA during infection is crucial for understanding how viruses hijack host cells, subvert host processes, and evade innate immune defense mechanisms. We recently introduced a cutting-edge suite of RNA interactomics techniques to resolve interactions between distinct viral RNA species and the host cell proteome at unprecedented resolution. Here, we describe the data-driven discovery of several previously unknown host-dependency mechanisms that uncover unique viral strategies to produce and utilize RNA in an infected cell. Our approach charts a path toward a systemscale characterization of RNA-centric regulatory mechanisms utilized by virus or host and identifies pathways for therapeutic exploitation.

T18 – Integrating RNA structure and protein interactions to uncover the mechanisms of viral and cellular ires function

Riccardo Delli Ponti | Italian Institute of Technology, Genoa, Italy

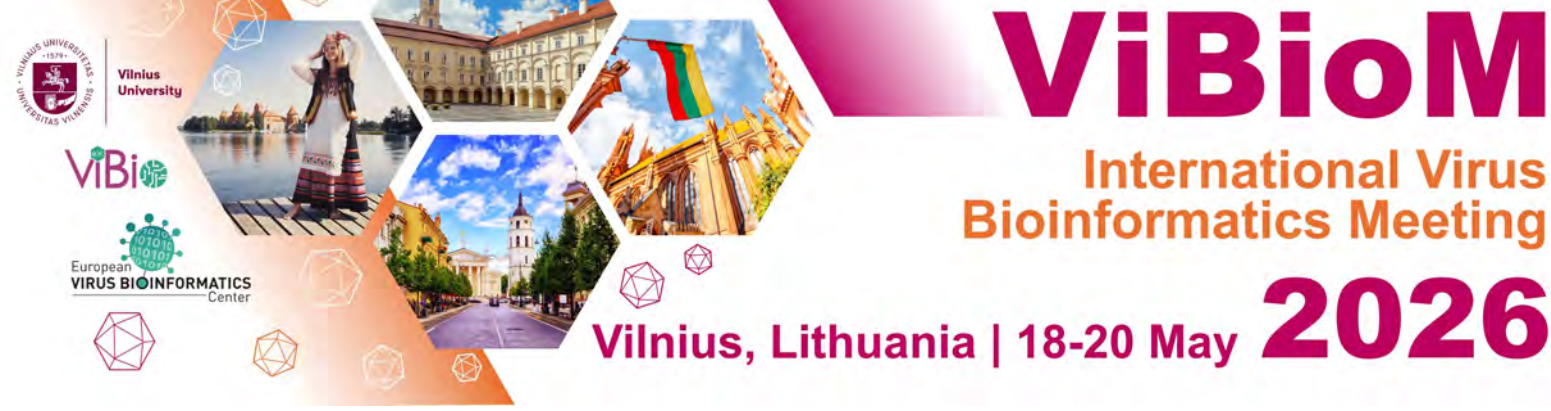
RNAs fold into complex structures that critically influence gene expression. A prominent class of regulatory elements resides in the 5′ untranslated region (5′ UTR), where internal ribosome entry sites (IRESs) promote cap-independent translation by directly engaging the ribosome. First discovered in viral genomes, IRESs have been classified into four types according to their structural compactness and factor requirements. While viral IRESs are well studied, cellular IRESs remain poorly understood: they display limited sequence conservation, reduced structural compactness, and variable dependence on auxiliary RNA-binding proteins

known as IRES trans-acting factors (ITAFs). Whether their activity relies mainly on RNA structure or protein assistance remains unresolved. Here, we present a computational framework that combines in silico mutagenesis and RNA–protein interaction profiling to investigate IRES mechanisms and guide the design of synthetic elements.

T19 – RdRpCATCH: A unified resource for RNA virus discovery using viral RNAdependent RNA polymerase profile Hidden Markov models

Dimitris Karapliafis | Wageningen University & Research, Wageningen, The Netherlands

Recent advances in large-scale sequence mining expanded our knowledge of RNA virus diversity. Most genome mining approaches for detecting RNA viruses that encode RNA-dependent RNA polymerase (RdRp) rely on identifying this conserved protein, by employing profile Hidden Markov Models (pHMMs) to scan sequencing datasets. Recently, several new pHMM databases for RdRp detection have been released, each following distinct design principles. However, their relative performance is unclear and their accessibility to users without specialized computational expertise is limited. Here we introduce the RdRp Collaborative Analysis Tool with Collections of pHMMs (RdRpCATCH: https://github.com/dimitris-karapliafis/RdRpCATCH), developed to consolidate publicly available RdRp pHMM resources into a single, accessible platform. RdRpCATCH enables the scanning of (meta)transcriptomic assemblies to discover RNA viruses and provides subsequent taxonomic annotation of detected contigs. A comparative analysis of RdRp pHMM databases reveals that most are highly effective at detecting known diversity of RNA viruses while minimizing false positives, supporting their joint use within RdRpCATCH. RdRpCATCH is distributed as both a conda package and a web server application (https://rdrpcatch.bioinformatics.nl), facilitating access for researchers with diverse expertise. By integrating multiple pHMM resources, this unified framework addresses fragmentation in the field and reduces technical barriers to enable comprehensive viral discovery.



Poster Sessions

Highlighted posters will be pitched in a snapshot presentation immediately before the poster session.

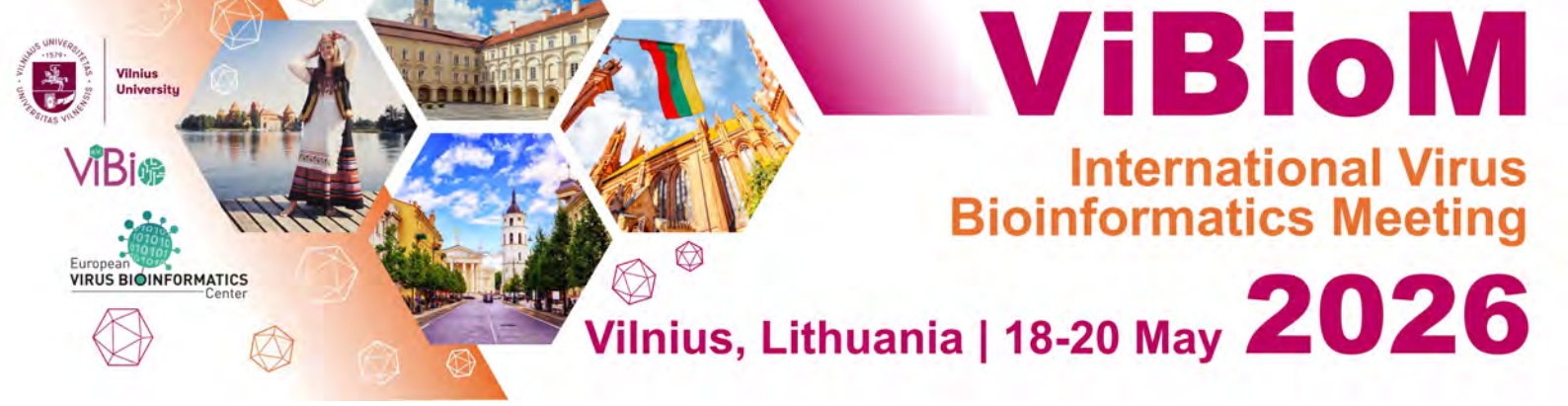
Poster Session A

A01	**Fishing metagenomic dark matter to recover missing viral segment, case study with quaranjaviruses** Emre Mert Asar | Vilnius University, Vilnius, Lithuania
A02	**Predicting bat-borne virus spillover in livestock from their genome sequences** Bailey Atkinson | University of Glasgow, Glasgow, United Kingdom
A03	**Virus identification using HTS in pea and sugar beet: A basis for future resistance breeding studies** Karima Ben Mansour | Swedish University of Life Sciences, Uppsala, Sweden
A04	**First detection of fig cryptic virus in Croatia revealed by nanopore sequencing of *Ficus carica*** Nina Buljević | University of Zagreb Faculty of Agriculture, Zagreb, Croatia
A05	**Early Intra-Host Evolution of Hepatitis E Virus in Asymptomatic Blood Donors: Foundations for a Clinical Course Prediction Tool** Nadi Dixit | Ruhr University Bochum, Bochum, Germany
A07	**Pregnancy and Early-Life Gut Virome in the Lifelines NEXT cohort: Origin, Persistence, Influencing Factors and Health Implications** Asier Fernández Pato | University Medical Center Groningen, Groningen, The Netherlands
A08	**FluMut: open-source tool for mutation surveillance in Avian Influenza Viruses** Edoardo Giussani | Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro, Italy
A09	**A Machine Learning approach for Alphainfluenzavirus influenzae (H3N2) host prediction based on virus-host genetic features** João Pedro Guimarães | University of Campinas – UNICAMP, Campinas, SP, Brazil
A10	**Protocol Bias Shapes Viral Detection and Assembly in Public Bat Viromes** Dominika Kadlečková | Charles University, Prague, Czechia
A11	**Context-Dependent Mutation Rates and Fitness Landscapes in RNA Viruses** Alexander Kuznetsov | University of Basel, Biozentrum, Basel, Switzerland

ViBioM International Virus Bioinformatics Meeting 2026

Vilnius, Lithuania | 18-20 May

A12	**Metagenomic De Novo Assembly of Viral Nanopore Field Samples** Mia Le | Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany
A13	**ViMOP: A field-ready pipeline for viral metagenomic nanopore sequencing** Mia Le | Outbreak Preparedness and Response, Bernhard-Nocht-Institute for Tropical Medicine, Hamburg, Germany
A14	**Current and Future AI-based Resources Provided by the BV-BRC and ICTV** Elliot Lefkowitz | University of Alabama at Birmingham, Birmingham, United States
A15	**Analyzing feature importance in machine learning models uncovers genotype discriminators for hepatitis E virus** Maximilian Nocke | Department of Translational and Computational Infection Research, Ruhr University Bochum, Bochum, Germany
A16	**Exploring the whitefly microbiome: insights into the composition of a complex ecological community** Laura Patiño Medina | Wageningen University & Research, Wageningen, The Netherlands
A17	**Spider webs as natural samplers for metagenomic surveillance of viral diversity** Anja Pecman | National Institute of Biology, Ljubljana, Slovenia
A18	**Systematic characterization of major and minor variants in Enterovirus D68 reveals extensive low-frequency diversity in respiratory samples (2023–2025)** Greta Romano | Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
A19	**Genin2: an open-source ML-based A(H5) influenza genotyping tool** Alessandro Sartori | Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro (PD), Italy
A20	**Genomic characterization of a genetically engineered HCMV strain for vaccine production** Hanno Schmidt | University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany
A21	**Design and Assessment of PANDEVIR capture panel for viral characterization in environmental samples** Maria Tarradas-Alemany | Universitat de Barcelona, Barcelona, Spain
A22	**Inferring antigenic escape mutations in GII.4 noroviruses** A S M Rubayet Ul Alam | MRC-University of Glasgow Centre for Virus Research (CVR), Glasgow, United Kingdom
A23	**Genomic Mutations and DNA Methylation Define Divergent Phage Resistance Strategies in *P. aeruginosa*** Laurynas Vaitkus | Vilnius University, Vilnius, Lithuania
A24	**Mining public metagenomic databases for the presence of AAV2** Lukas Visockas | University of West London, London, United Kingdom



Poster Session B

B01 **Sub-Consensus Variation in Influenza A across UK Charadriiformes (2023–2026)**
Ben Clifton | Animal Plant and Health Agency, Woking, United Kingdom

B02 **suvtk: making viral genome submission FAIRly easy**
Lander De Coninck | KU Leuven, Leuven, Belgium

B03 **New thermophilic bacteriophages and their proteins**
Monika Dębińska | University of Gdansk, Gdansk, Poland

B04 **Back to the past: Evolutionary insights into historical human astrovirus**
Génesis Dela | National Laboratory of Virology, Szentagothai Research Centre, University of Pécs, Pécs, Hungary

B05 **A Single-Cell Transcriptomics Approach to Compare Coronavirus Infection in Human and Camelid Primary Airway Epithelial Cells**
Vera Flück | University of Bern, Bern, Switzerland

B06 **Parallel evolution of comparable large complexity in two sister RNA virus orders**
Alexander Gorbalenya | Leiden University Medical Center, Leiden, The Netherlands

B07 **Structural analysis of host insertions in the polyproline region of the hepatitis E virus pORF1 polyprotein**
Nicolas Jeanne | CHU Purpan Virology Laboratory, Toulouse, France

B08 **Multi-omics analysis of long-term coevolution of *Streptococcus thermophilus* and its lytic phage**
Yana Karnitskaya | Vilnius University, Vilnius, Lithuania

B09 **Finetuning protein language models with viral proteins**
Kieran Lamb | MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom

B10 **ViralZone: Linking genomes, enzymes, structure and function**
Philippe Le Mercier | Swiss Institute of Bioinformatics (SIB), Geneva, Switzerland

B11 **First Report of Parahenipavirus in UK Shrews (*Sorex araneus*) Highlights Accessory X Locus Plasticity**
Dan Maskell | Animal and Plant Health Agency (UK), Addlestone, United Kingdom

B12 **Whole genome phylogenetics reveals dominance of Eurasian Avian-like and reassortment with pandemic H1N1 in Swiss Pig herds**
Mike Mwanga | Universtity of Bern, Bern, Switzerland

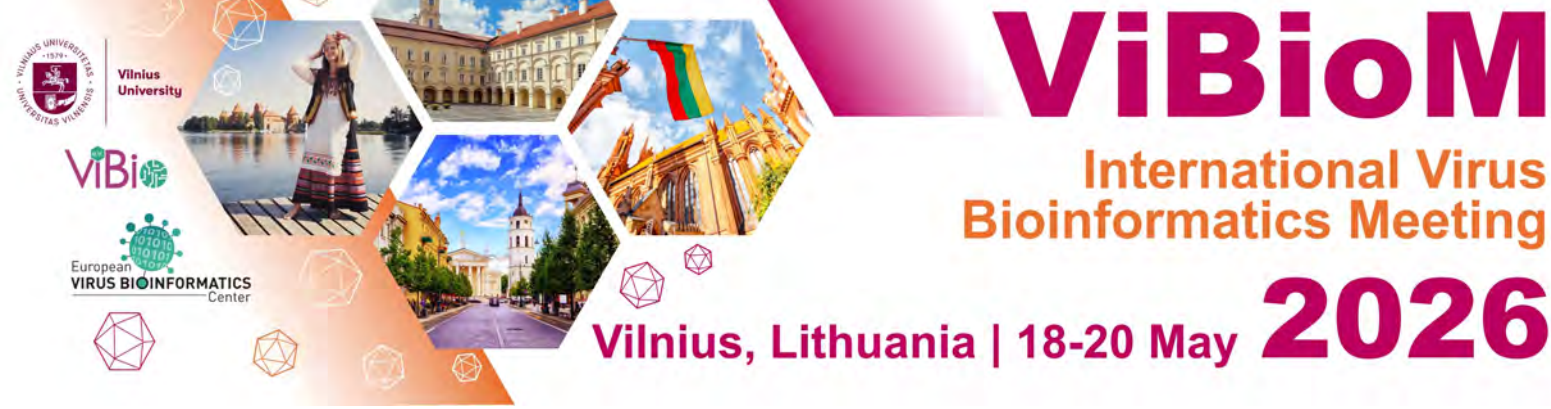
B13 **Pathoplexus: Building an equitable pathogen database**
Anna Parker | Swiss TPH; SIB; Pathoplexus; Basel, Switzerland

ViBioM

International Virus Bioinformatics Meeting

Vilnius, Lithuania | 18-20 May 2026

B14	**Poxvirus Gene Repertoire Evolution Driven by de novo Gene Birth, Gene Duplication, and Horizontal Gene Transfer** Arnon Plianchaisuk | The Institute of Medical Science, The University of Tokyo, Tokyo, Japan
B15	**A NeoRdRp-based Three-Step Pipeline for High-Precision RdRp Detection** Shoichi Sakaguchi | Osaka Medical and Pharmaceutical University, Takatsuki, Japan
B16	**Which viruses betray mosquito movements?** Martynas Smicius | Vilnius University, Vilnius, Lithuania
B17	**Production of Nanobody to Canine Parvovirus VP2 Protein** Hasan Emre Tali | Istanbul University-Cerrahpasa, Istanbul, Turkey
B18	**Preliminary Studies on the Detection and Molecular Characterisation of Parainfluenza Virus-5 in Dogs in Istanbul** Hasan Emre Tali | Istanbul University-Cerrahpasa, Istanbul, Turkey
B19	**What is facilitating a global virus pandemic in mosquitoes?** Aistė Židonytė | Vilnius University, Vilnius, Lithuania
B20	**The consequences of extinct viruses returning to human populations** Andrius Zimnickas | Vilnius University, Vilnius, Lithuania
B21	**Resolving prophage diversity and functional cargo in plant-associated *Erwinia billingiae* using long-read sequencing** Nikita Zrelovs | Latvian Biomedical Research and Study Centre, Riga, Latvia
B22	**Evolution and Diversity in Papillomaviruses Found in Wild Felids** Ayla Zustra | Arizona State University, Tempe, United States
B23	**Vertebrate infecting viruses found in deceased bobcat (*Lynx rufus*)** Ayla Zustra | Arizona State University, Tempe, United States
B24	**Exploring the evolution of Influenza-like viruses: Sequence and structure-based analysis of PB1 and glycoproteins** Ausrine Zvirblyte | Vilnius University, Vilnius, Lithuania
B25	**SpaXSeq & TIGO: GUI Tools for Viral Spatial Transcriptomics and scRNA-Seq Trajectory Inference** Quan Gu | MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom



Sponsors

GO VILNIUS

viruses

FEBS openbio

evbc@uni-jena.de

https://evbc.uni-jena.de/