ViBioM 2024

International Virus Bioinformatics Meeting

28. – 30. May 2024 | Leuven, Belgium

European VIRUS BIOINFORMATICS Center

KU LEUVEN

ULB UNIVERSITÉ LIBRE DE BRUXELLES

Abstract Book

https://evbc.uni-jena.de/events/vibiom2024/

ViBioM 2024

International Virus Bioinformatics Meeting

28. – 30. May 2024 | Leuven, Belgium

European VIRUS BIOINFORMATICS Center

KU LEUVEN

ULB UNIVERSITÉ LIBRE DE BRUXELLES

Contents

Programme Schedule 4

Abstracts 8

T01 1984 . . . 8
T02 Disentangling the evolutionary history of varidnaviruses . . . 8
T03 Phylogenetic Evidence for Viral Eukaryogenesis . . . 8
T04 Prior considerations for tests of temporal signal in ancient pathogen genomics . . . 8
T05 Exploration of the Updated Diversity of Helicase Superfamilies 1 and 2 Reveals New Families and Presence of Multiple Transfer Events Between Viruses and Cells . . . 9
T06 Anellovirus in humans, do we live happily ever after? . . . 10
T07 The prevalence of back-and-forth substitutions in the evolutionary landscape of human viruses . . . 10
T08 Characterization of HIV-1 Genome in A Long-Term Experimental Evolution Study . . . 10
T09 Unleashing molecular lego inspired by bacteriophage evolution . . . 11
T10 Phold: Using Protein Structural Homology & Protein Language Models to Enhance Phage Genome Annotation . . . 11
T11 Mapping glycoprotein structure reveals defining events in the evolution of the Flaviviridae . . . 12
T12 Characterisation and monitoring of SARS-CoV-2 and future emergent viruses using protein language models 12
T13 Adapting broad protein language models to viruses . . . 13
T14 Explaining Influenza A Neuraminidase Classification Results from a Biological Perspective using Interpretable Machine Learning Models . . . 13
T15 Improved prediction of protein-protein interactions using a next sentence cross-encoder . . . 14
T16 A system view on phage-host interactions using transcriptomics and acetyl-proteomics . . . 14
T17 Decoding the blueprint of filovirus entry through large-scale binding assays and machine learning . . . 14
T18 The Intriguing World of Unknowns: exploration of Polycipiviridae genomes, Phylogenetics and Host Specificity . . . 15
T19 Towards a global virus genomic surveillance system . . . 15
T20 Shotgun metagenomics on air: Longitudinal surveillance of viruses in a daycare center . . . 15
T21 Molecular characteristics of a localized influenza A(H1N1)pdm09 outbreak in South Africa, 2020 . . . 16
T22 Representative full-genome alignments of viral clades - a semi-automated approach incorporating amino acids, nucleotides & RNA secondary structures . . . 17
T23 detectEVE: Illuminating the hidden viral footprint in genomic data . . . 17
T24 varVAMP: automated pan-specific primer design for tiled full genome sequencing of highly diverse viral pathogens . . . 18
T25 CLIMADE: Pandemic Preparedness for Arboviruses in a time of Climate Change . . . 18
T26 Hendra virus in Australian bats show a diffuse spatio-temporal structure . . . 18
T27 Sequencing error inspires novel phylogenetic approaches to sequence data . . . 19

Poster Sessions 20

Organising Committee

- **Simon Dellicour**, Université Libre de Bruxelles, Belgium
- **Franziska Hufsky**, Friedrich Schiller University Jena, Germany
- **Liana Kafetzopoulou**, KU Leuven, Belgium
- **Philippe Lemey**, KU Leuven, Belgium
- **Manja Marz**, Friedrich Schiller University Jena, Germany
- **Jelle Matthijnssens**, KU Leuven, Belgium

Scientific Program Committee

- **Arif Nur Muhammad Ansori**, Universitas Airlangga, Indonesia
- **Katy Brown**, University of Cambridge, UK
- **Anamarija Butkovic**, Institut Pasteur, France
- **Daniel Depledge**, Hannover Medical School, Germany
- **Evangelia-Georgia Kostaki**, National and Kapodistrian University of Athens, Greece
- **Philippe Le Mercier**, Swiss Institute of Bioinformatics, Switzerland
- **Cédric Lood**, University of Oxford, UK
- **Alejandro Matía**, Stanford University, US
- **Ingrida Olendraite**, University of Cambridge, UK
- **Arli Aditya Parikesit**, Indonesia International Institute for Life Sciences, Indonesia
- **Rhys Parry**, The University of Queensland, Australia
- **Petri Susi**, University of Turku, Finland

Monday, May 27

19:00		**Welcome Get Together** MALZ, Brusselsestraat 51, 3000 Leuven, Belgium

Tuesday, May 28

09:00 – 09:30		**Welcome and opening remarks** Manja Marz (EVBC Director) Jelle Matthijnssens, Philippe Lemey, Simon Dellicour (Local Organizers)
09:30 – 10:00	**T01**	*Opening talk invited by VIROINF* **1984** Alexander Gorbalenya | Leiden University, Netherlands
10:00 – 10:30		Coffee break
Session 1: Viral Evolutionary History		
10:30 – 11:00	**T02**	*Keynote talk* **Disentangling the evolutionary history of varidnaviruses** Mart Krupovic | Institut Pasteur, France
11:00 – 11:20	**T03**	**Phylogenetic Evidence for Viral Eukaryogenesis** Sangita Karki | Virginia Tech, Blacksburg, United States
11:20 – 11:40	**T04**	**Prior considerations for tests of temporal signal in ancient pathogen genomics** Sebastian Duchene | Institut Pasteur, Paris, France
11:40 – 12:00	**T05**	**Exploration of the Updated Diversity of Helicase Superfamilies 1 and 2 Reveals New Families and Presence of Multiple Transfer Events Between Viruses and Cells** Anamarija Butkovic | Institut Pasteur, Paris, France
12:00 – 13:30		Lunch break
13:30 – 14:00		**Poster Pitches A**
14:00 – 15:30		**Poster Session A**
15:30 – 16:00		Coffee break

Session 2: Evolution of Human Viruses		
16:00 – 16:30	**T06**	*Keynote talk* **Anellovirus in humans, do we live happily ever after?** Lia van der Hoek | University of Amsterdam, Netherlands
16:30 – 16:50	**T07**	**The prevalence of back-and-forth substitutions in the evolutionary landscape of human viruses** Florian Poulain | University of Liege, Liège, Belgium
16:50 – 17:10	**T08**	**Characterization of HIV-1 Genome in A Long-Term Experimental Evolution Study** Ali Movasati | University of Zurich / University Hospital Zurich, Zurich, Switzerland

Wednesday, May 29

09:30 – 10:00	**T09**	*Keynote talk* **Unleashing molecular lego inspired by bacteriophage evolution** Yves Briers | Ghent University, Belgium

Session 3: Viral Protein Structure		
10:00 – 10:20	**T10**	**Phold: Using Protein Structural Homology & Protein Language Models to Enhance Phage Genome Annotation** George Bouras | The University of Adelaide, Adelaide, Australia
10:20 – 10:40	**T11**	**Mapping glycoprotein structure reveals defining events in the evolution of the Flaviviridae** Jonathon Mifsud | University of Sydney, Sydney, Australa
10:40 – 11:00	**T12**	**Characterisation and monitoring of SARS-CoV-2 and future emergent viruses using protein language models** Kieran Lamb | MRC-University of Glasgow Centre for Virus Research, Glasgow, Scotland

11:00 – 11:30		Coffee break

Session 4: Machine Learning for Viral Proteomics		
11:30 – 11:50	**T13**	**Adapting broad protein language models to viruses** Spyros Lytras | University of Tokyo, Tokyo, Japan
11:50 – 12:10	**T14**	**Explaining Influenza A Neuraminidase Classification Results from a Biological Perspective using Interpretable Machine Learning Models** Lynn V. Reuss | University of Applied Sciences Mittweida, Mittweida, Germany

12:10 – 12:30	**T15**	**Improved prediction of protein-protein interactions using a next sentence cross-encoder** Dan Liu | University of Glasgow, Glasgow, United Kingdom
12:30 – 14:00		Lunch break
14:00 – 14:30		**Poster Pitches B**
14:30 – 16:00		**Poster Session B**
16:00 – 16:30		Coffee break
Session 5: Virus-host-interaction		
16:30 – 17:00	**T16**	*Keynote talk* **A system view on phage-host interactions using transcriptomics and acetyl-proteomics** Vera Van Noort | KU Leuven, Netherlands
17:00 – 17:20	**T17**	**Decoding the blueprint of filovirus entry through large-scale binding assays and machine learning** Gorka Lasso | Albert Einstein College of Medicine, New York, United States
17:20 – 17:40	**T18**	**The Intriguing World of Unknowns: exploration of Polycipiviridae genomes, Phylogenetics and Host Specificity** Ingrida Olendraite | University of Cambridge, Cambridge, United Kingdom
19:00		**Conference Dinner** Restaurant Mykene, Muntstraat 44, 3000 Leuven, Belgium

Thursday, May 30

Session 6: Surveillance / Epidemiology I		
09:30 – 10:00	**T19**	*Keynote talk* **Towards a global virus genomic surveillance system** Nathan Grubaugh | Yale School of Public Health, United States
10:00 – 10:20	**T20**	**Shotgun metagenomics on air: Longitudinal surveillance of viruses in a daycare center** Mustafa Karatas | KU Leuven, Leuven, Belgium

Time	ID	Talk
10:20 – 10:40	**T21**	**Molecular characteristics of a localized influenza A(H1N1)pdm09 outbreak in South Africa, 2020** Dikeledi Kekana | National Institute for Communicable Diseases, Johannesburg, South Africa
10:40 – 11:10		Coffee break
Session 7: Viral sequence analysis		
11:10 – 11:30	**T22**	**Representative full-genome alignments of viral clades - a semi-automated approach incorporating amino acids, nucleotides & RNA secondary structures** Sandra Triebel | Friedrich Schiller University Jena, Jena, Germany
11:30 – 11:50	**T23**	**detectEVE: Illuminating the hidden viral footprint in genomic data** Nadja Brait | University of Groningen, Groningen, Netherlands
11:50 – 12:10	**T24**	**varVAMP: automated pan-specific primer design for tiled full genome sequencing of highly diverse viral pathogens** Jonas Fuchs | Institut für Virologie, Universtitätsklinikum Freiburg, Freiburg, Deutschland
12:10 – 13:50		Lunch break
Session 8: Surveillance / Epidemiology II		
13:50 – 14:20	**T25**	*Keynote talk* **CLIMADE: Pandemic Preparedness for Arboviruses in a time of Climate Change** Houriiyah Tegally | Stellenbosch University, South Africa
14:20 – 14:40	**T26**	**Hendra virus in Australian bats show a diffuse spatio-temporal structure** Claude Yinda Kwe | NIAID/NIH, Hamilton, USA
14:40 – 15:00	**T27**	**Sequencing error inspires novel phylogenetic approaches to sequence data** Chris Illingworth | MRC-University of Glasgow Centre for Virus Research, Glasgow, UK
15:00 – 15:45		**Annual Meeting of the EVBC**
15:45 – 16:15		**Closing ceremony / awards**

T01: 1984

Alexander Gorbalenya | Leiden University, Netherlands

No abstract available.

Session 1: Viral Evolutionary History

T02: Disentangling the evolutionary history of varidnaviruses

Mart Krupovic | Institut Pasteur, France

No abstract available.

T03: Phylogenetic Evidence for Viral Eukaryogenesis

Sangita Karki | Virginia Tech, Blacksburg, United States

The details surrounding the evolution of complex cells is one of the most enduring mysteries in biology. Recent evidence has suggested that Asgard Archaea are the closest cellular relatives of eukaryotes, but the evolution of cellular complexity in early eukaryotes remains unclear. In particular, phylogenies of eukaryotic DNA polymerases and RNA polymerases often support a 3-domain topology that is not consistent with an origin from Asgard Archaea. Here we present exhaustive phylogenetic analysis of eukaryotic DNA and RNA polymerases and show that several key components of these enzymes are derived from the ancestors of modern giant viruses. Specifically, we show that the eukaryotic delta polymerase (EukPolDelta), one of the key processive polymerases required for genome replication in all eukaryotes, clusters within an ancient viral clade, strongly supporting a viral origin. We also show that that EukPolEpsilon, the other processive polymerase required for genome replication, is derived from Asgard Archaea, providing a direct link between early eukaryotes, Asgard archaea, and giant viruses. Lastly, we provide an exhaustive phylogenetic analysis of eukaryotic RNA polymerases and show that RNA polymerase II, which is responsible for mRNA transcription in eukaryotes, is also derived from the ancestors of modern giant viruses, consistent with previous research (Guglielmini et al, PNAS, 2019). In total, our results are consistent with a model of viral eukaryogenesis put forward by Philip Bell (Front. In Microbiol., 2022) in which eukaryotes emerged from an Asgard Archaea that was infected with an ancient giant virus that encoded a processive polymerase Delta and a multi-subunit RNA polymerase.

T04: Prior considerations for tests of temporal signal in ancient pathogen genomics

Sebastian Duchene | Institut Pasteur, Paris, France

Molecular clock methods are routinely used to infer evolutionary rates and timescales. However, these methods always require some external source of information known as a calibration, which can come in the form of genome sampling

times. Importantly, the window of time of the sequence sampling times should be sufficiently wide as to capture a measurable amount of molecular evolutionary change. For example, rapidly evolving viruses tend to require shorter windows of time than those that are slowly evolving. There exist a range of statistical tests to determine whether the sequence sampling times are sufficiently informative for reliable molecular clock calibration. The premise of such approaches, known as tests of temporal signal, is that the sequence sampling times improve the fit of the phylogenetic model. In the absence of temporal signal, the molecular clock calibration may be misleading or uninformative, and thereby may produce spurious estimates of evolutionary rates and times. We conducted analyses of a range of data sets of ancient viral outbreaks to investigate the performance of tests of temporal signal under a Bayesian framework. We also conducted extensive simulation experiments to assess the impact of the tree prior in the sensitivity and specificity of these tests. Our results demonstrate that the tree prior can have a substantial impact, not just whether temporal signal is detected, but on the precision and accuracy of the resulting estimates. This situation stems from the fact that there are interactions between parameters in the full Bayesian phylogenetic model that are poorly understood, which we investigate here using prior predictive simulations. Finally, we provide new guidelines and methods to improve the reliability of tests of temporal signal and for choosing parameter priors.

T05: Exploration of the Updated Diversity of Helicase Superfamilies 1 and 2 Reveals New Families and Presence of Multiple Transfer Events Between Viruses and Cells

Anamarija Butkovic | Institut Pasteur, Paris, France

Superfamily 1 and 2 helicases (S1H and S2H) are prevalent in all life domains and their viruses. These helicases are crucial for DNA replication, repair, and gene expression. S1H and S2H are evolutionarily related but they show unique conserved motifs, indicating specialized functions. In this work, we explored the global diversity of S1H and S2H across all biomes. Detailed comparison of protein sequences and structures allowed us to define six new helicase families, of which 5 consisted of numerous viral sequences, with unique arrangements of conserved motif and domain organizations, suggestive of distinct functions. Phylogenetic analysis revealed frequent transfers between cells, large DNA viruses, and phages within both superfamilies, highlighting a complex evolution driven by host-virus interactions. Bacterial defense helicases, not involved in restriction-modification, were found in both S1H and S2H superfamilies and show a close relationship to virus-encoded helicases, suggesting transfer events between them. While helicases from eukaryote infecting viruses, such as herpesviruses, originated from different sources. Unlike DNA virus helicases, RNA virus S1H and S2H appeared monophyletic. Structural comparisons revealed the conservation of the ancestral RecA-like ATPase core and high variability of terminal domains which likely determine the specific functions of different helicases. RNA virus helicases are smaller than DNA virus helicases likely due to different evolutionary pressures, as RNA viruses replicate rapidly in the cytoplasm using virus-encoded polymerases, while DNA viruses use host machinery in the nucleus, possibly needing more complex helicases. The beta loop in some RNA helicases and the large alpha loop in a newly identified helicase family related to restriction enzymes are thought to assist in nucleic acid separation, specific to RNA helicases and restriction enzymes, respectively.

Session 2: Evolution of Human Viruses

T06: Anellovirus in humans, do we live happily ever after?

Lia van der Hoek | University of Amsterdam, Netherlands

No abstract available.

T07: The prevalence of back-and-forth substitutions in the evolutionary landscape of human viruses

Florian Poulain | University of Liege, Liège, Belgium

It has been largely reported that viruses are known entities with the fastest evolution. However, despite a very high mutation rate, various observations report that viral sequences evolve little over time when compared over several hundred years. To explain this paradox, a model named the "Prisoner of War" has recently been proposed. It establishes that viral evolution is constrained by the host environment, and viral evolution rates tend to converge towards that of the host over long time scales. To explore the dynamics of viral mutation aggregation, we inferred 692 phylogenetic trees and extracted for each their evolutionary landscape. This landscape corresponds to the set of mutations of a tree distributed according to the type and context of the mutation. The study of these evolutionary landscapes has allowed us to observe a mirror effect which corresponds to a symmetry between a type of mutation and its reversion in the same context. Among the 295 viral species tested, approximately 90% exhibit such a mirror effect. It is interesting to note that the species with the most significant lack of mirror effect are the emerging viruses such as SARS-CoV-2 and MPXV, indicating that such an effect is only found in viral genomes fully adapted to their host. Among the over 3 million mutations identified, more than 20% are involved in back-and-forth substitution. This proportion, however, seems largely underestimated, and the prevalence of these reversions seems to explain the presence of the mirror effect. Based on these observations, we propose a global dynamic of back-and-forth substitution reversion, which potentially contributes to the "Prisoner of War" effect and further enhances our understanding of viral evolution. Viruses seem to adapt very quickly but evolve slowly due to constant reversion of substitutions.

T08: Characterization of HIV-1 Genome in A Long-Term Experimental Evolution Study

Ali Movasati | University of Zurich / University Hospital Zurich, Zurich, Switzerland

One of the popular laboratory methods to mimic evolution is long-term experimental evolution (LTEE). In our ongoing LTEE study, we have been passaging HIV-1 NL4_3 in two human T-cell leukemia cell lines, namely MT-2 and MT-4, for over 1100 generations. At certain intervals, the viral genome is extracted and sequenced using the Illumina technology. So far we have observed stark differences in the evolutionary dynamics between viral populations passaged on MT-2 and MT-4 lines. For instance, the rate of mutation supply is higher for MT-4 viruses compared to MT-2 (0.32 minority mutation per generation vs 0.19); in turn leading to higher within-population diversity (as evident by mean Shannon entropy of 0.44 compared to 0.28 at transfer 570). On the contrary, the fixation rate is higher for MT-2 compared to MT-4 propagating virus (0.12 vs 0.06 fixated mutation per generation). These differences are most likely stemming from distinct environmental pressures that the virus is exposed to. In addition, evolution in all four viral lines appears more repeatable than expected

(P-value $< 10^{-16}$). We observe a total of 738 unique majority mutations across the four evolution lines throughout the experiment; out of which 571 are only seen in one line (0.77%). Of note, the evolution parallelism is more pronounced among the MT-4 than MT-2 viral lines (18% VS 11.9% shared majority mutations between replicates). Lastly, in all viral lines and across the viral genome, dN/dS values consistently decrease over time, pointing to selection playing a more important role in driving evolution in the initial phase of the experiment. This genotypic observation is corroborated by preliminary fitness data acquired from dual competition assays between adapted viral populations and the ancestral virus, where fitness gains are observed only in the first 100 passages. In summary, using the genomic sequencing data of our LTEE we are able to identify interesting dynamics in HIV-1 genomic evolution.

T09: Unleashing molecular lego inspired by bacteriophage evolution

Yves Briers | Ghent University, Belgium

No abstract available.

Session 3: Viral Protein Structure

T10: Phold: Using Protein Structural Homology & Protein Language Models to Enhance Phage Genome Annotation

George Bouras | The University of Adelaide, Adelaide, Australia

It has long been known that a protein's structure is more conserved and more related to its function than its amino acid sequence. As most phage proteins do not have a known function, using structural information promises to super-charge our understanding of phage genomes and phage protein function.
Until recently it was cost- and time-prohibitive to predict and compare protein structures. However, recent advances of in silico structure prediction following Alphafold2 and Colabfold have allowed protein structures to be predicted directly from sequences, albeit at significant computational cost. The advent of the Foldseek algorithm has also allowed for rapid and large-scale protein comparisons. Further, the development of protein language models like ESM-2 and ProstT5 has led to more computationally lightweight and faster inference of structural information directly from sequences, compared to Alphafold2.
Here we present Phold, which combines the protein language model ProstT5 with Foldseek and a database of Colabfold-generated proteins to annotate phages with far more sensitivity than existing approaches. Using a database of nearly a million phage protein structures, Phold significantly increases the number of phage proteins with a functional annotation compared to our existing phage genome annotation tool Pharokka. Phold yields particular improvement on phages that are less similar to those that have previously been characterised.
Phold is an open-source software program installable on MacOS and Linux and takes only a few minutes to annotate a single phage on a laptop. If a GPU is available, Phold scales seamlessly to metagenomic-scale datasets.

T11: Mapping glycoprotein structure reveals defining events in the evolution of the Flaviviridae

Jonathon Mifsud | University of Sydney, Sydney, Australia

Viral glycoproteins drive membrane fusion in enveloped viruses and determine host range, tissue tropism and pathogenesis. Despite their importance, there is a fragmentary understanding of glycoproteins within the Flaviviridae; for many species the glycoproteins have not yet been identified, for others, such as the hepaciviruses, the molecular mechanisms of membrane fusion remain uncharacterised. Here, we combine comprehensive phylogenetic analyses with systematic protein structure prediction to survey glycoproteins across the entire Flaviviridae. We discover class-II fusion systems, homologous to the orthoflavivirus E glycoprotein, in most species, including highly-divergent jingmenviruses and large genome flaviviruses. However, the E1E2 glycoproteins of the hepaci-, pegi- and pestiviruses are structurally distinct, may represent a novel class of fusion mechanism, and are strictly correlated with infection of vertebrate hosts. By mapping glycoprotein distribution onto the underlying phylogeny we reveal a complex history of evolutionary events that have shaped the diverse virology and ecology of the Flaviviridae.

T12: Characterisation and monitoring of SARS-CoV-2 and future emergent viruses using protein language models

Kieran Lamb | MRC-University of Glasgow Centre for Virus Research, Glasgow, Scotland

The SARS-CoV-2 pandemic is a primary example of how the spillover of a novel human virus can rapidly progress from a localised outbreak to a global pandemic. In the early stages of an outbreak, data is scarce and what's available can be exceedingly valuable. Experimental data is time consuming to produce and is often not available until much later stages of these fast developing situations. Any information on the propensity for a new virus to evolve is important but this usually requires sequence alignment data. Protein language models (PLMs) like ESM-2 use millions of protein sequences to learn the properties of amino acid sequences. ESM-2 can now fold protein sequences into their 3D structure, with no alignment or other information about the sequence necessary. As such, PLMs clearly contain information about protein structure and evolutionary constraint.
Here, we describe how ESM-2 can be used to represent meaningful information about a novel virus from sequences at various stages of an outbreak. We show using the SARS-CoV-2 pandemic how these models could have been applied; both in the early stages before experimental observations and in later stages for monitoring and horizon scanning. Using PLM enabled in-silico deep mutational scanning (DMS) we demonstrate that PLMs describe fundamental properties of viral proteins. We show that PLMs can differentiate between phenotypically different viral protein sequences and even demonstrate their ability to identify mutational epistasis. These model outputs can supplement the available information, and we make a case for their future application and use in outbreaks or pandemics to come.

Session 4: Machine Learning for Viral Proteomics

T13: Adapting broad protein language models to viruses

Spyros Lytras | University of Tokyo, Tokyo, Japan

The field of artificial intelligence has recently experienced a boom in popularity and commercial applications, particularly in the form of large language models (LLMs). LLMs are state-of-the-art machine learning algorithms trained on a large text corpus (set of words in the context of sentences) that aim to identify fundamental properties of grammar and meaning. Similarly, LLMs can be trained in any other corpus of contextualised characters, such as amino acids in proteins. Evolutionary Scale Modeling (ESM) is a popular example of such a model, trained on over 60 million protein sequences covering all known biological diversity. Given any protein sequence, this model can infer how likely it is for each amino acid to be present at each protein site. We questioned whether we could focus ESM's attention on specific protein groups, while retaining the model's broader understanding of protein biology. To test this, we fine-tuned ESM-2 with a masked-learning task using a comprehensive set of more than 15,000 unique influenza A hemagglutinin (IAV HA) protein sequences. We find that, in our IAV HA-specific model, the distribution of inferred probabilities for how likely each amino acid change is in a given HA sequence (summarised as the probabilities' Shannon entropy for each site) correlates well with the per-site entropy in the sequence alignment of all HA proteins. This is not the case when using the original ESM-2 model. Hence, our model can predict 'mutable' sites in any IAV HA protein without the need for a sequence alignment, and provide probabilities for which amino acid change is more likely to occur in each site. This approach can be implemented for any virus protein and paves the way for efficient, alignment-free variant forecasting of human viral pathogens.

T14: Explaining Influenza A Neuraminidase Classification Results from a Biological Perspective using Interpretable Machine Learning Models

Lynn V. Reuss | University of Applied Sciences Mittweida, Mittweida, Germany

The surface protein neuraminidase (NA) of the Influenza A Virus is responsible, among other things, for the release of virions from the cell and is thus of interest in pharmacological research. Depending on the protein structure and genetic relationship, the NA protein can be distinguished into nine different NA subtypes which in turn belong to one of two NA groups. In this study, we analyzed the amino acid sequences of the NA head domain by means of an interpretable machine learning (ML) classification method known as Generalized Matrix Learning Vector Quantization (GMLVQ). The aim of this work was to achieve the classification of NA by subtype and by group based on the amino acid sequences. For ML purposes, we generated numerical data from the biological sequences using descriptive features. Using this procedure, NA groups and subtypes can be classified with a high accuracy, for example with an accuracy for the binary classification by NA group of 99.73%, the data seems linear separable. These findings could facilitate the identification of group affiliation of unknown NA, but more importantly interpretable ML models give an insight into the most relevant amino acids to differentiate between protein sequences, which characteristics can then be analyzed in more detail, thus turning one's attention to the biological explanation. In particular, the GMLVQ classification results are accomplished by design through the impact of the features on the class separation, which finally enabled us to estimate the importance of the amino acids. This is the main advantage of the GMLVQ-classifier compared to other state-of-the-art bioinformatic

machine learning methods. In the presentation, we will explain the methods and results in detail and demonstrate how interpretable ML models can advance biological knowledge/research.

T15: Improved prediction of protein-protein interactions using a next sentence cross-encoder

Dan Liu | University of Glasgow, Glasgow, United Kingdom

No abstract available.

Session 5: Virus-host-interaction

T16: A system view on phage-host interactions using transcriptomics and acetyl-proteomics

Vera Van Noort | KU Leuven, Netherlands

No abstract available.

T17: Decoding the blueprint of filovirus entry through large-scale binding assays and machine learning

Gorka Lasso | Albert Einstein College of Medicine, New York, United States

Filoviruses within the genera Ebolavirus and Marburgvirus cause sporadic but increasingly frequent outbreaks of fatal hemorrhagic fever in humans in Africa. Multiple pieces of evidence suggest that bats are important hosts. However, for many filoviruses, the specific animals that act as hosts remain unknown, limiting our efforts to forecast regions or interfaces where spillover is most likely. The cellular Niemann Pick C1 protein (NPC1) mediates the interaction with the viral glycoprotein (GP), acting as an essential entry receptor for all filoviruses, and amino acid variation at the virus-receptor interface is proposed to play a major role in dictating viral susceptibility and species-specific tropism. We reasoned that GP-NPC1 binding affinities for different filoviruses and bat species can inform high-resolution models predicting potential filovirus hosts.
We have performed experimental binding studies across seven filovirus GPs and 82 NPC1 orthologs. Our results revealed that GPs and NPC1s show differential binding propensities that do not correlate with amino acid sequence identity or phylogeny. We integrated these binding studies with machine learning (ML) and amino acid properties to deepen our understanding of the genetic determinants of viral susceptibility and develop a model that predicts binding affinity. Our model predicts binding affinity with an R2 of 0.69-0.75 and reveals amino acid changes in GP that likely play a pivotal role by affecting NPC1 binding non-specifically. Moreover, our model captures general physicochemical principles that dictate GP-NPC1 binding, which prompted us to predict GP-NPC1 binding affinities across filoviruses and other bats in endemic and non-endemic areas. This work represents the largest effort to date to characterize filovirus GP binding to its cellular receptor, and it provides a computational tool that can inform on susceptible animals and the zoonotic potential of known and novel filoviruses.

T18: The Intriguing World of Unknowns: exploration of Polycipiviridae genomes, Phylogenetics and Host Specificity

Ingrida Olendraite | University of Cambridge, Cambridge, United Kingdom

In 2007, a novel polycistronic picorna-like virus (SINV-2) was found in the invasive fire ant. Ten years later we proposed a new family, Polycipiviridae, comprising at the time 15 diverse viruses among three genera infecting various arthropods.The unique polycistronic RNA genome organisation harbours five main ORFs (capsid proteins in separate ORFs and one long polyprotein). Recently we have expanded the family to over 400 unique sequences and proposed 12 additional genera with even variations of genome organisation. However, the comparative genomic analysis faced various challenges. Firstly, we were able to identify and fix multiple quality issues in publicly available nucleotide sequences, including likely assembly problems. We also analysed the phylogenetic tree robustness within the family by comparing different strategies and algorithms. We investigated the inconsistencies and challenges of low support of between-genera topologies when proposing new monophyletic genera. In addition, we investigated the issue of novel polyprotein cleavage site identification. We used AlphaFold2 to predict separate polyprotein domains and this aided the prediction of novel putative cleavage sites, which differ in sequence motif for different genera. Finally, making use of the greatly expanded set of polycipivirus sequences, we observed a striking variety of sources and putative hosts. However we show that the great taxonomic variety represents not an impressive range of true hosts but rather misannotation of datasets. With more information available we propose a more precise host taxonomic group than previously thought: Neoptera (infraclass of Insects).

Session 6: Surveillance / Epidemiology I

T19: Towards a global virus genomic surveillance system

Nathan Grubaugh | Yale School of Public Health, United States

No abstract available.

T20: Shotgun metagenomics on air: Longitudinal surveillance of viruses in a daycare center

Mustafa Karatas | KU Leuven, Leuven, Belgium

Surveillance of the environment can provide insights into pathogenic and environment-related viruses. In this study, we aimed to identify a spectrum of viruses and provide genomic insights into their presence in indoor air samples, potentially identifying a novel method for environmental surveillance.
From January 2022 to December 2022, we collected 41 air samples using an AerosolSense active air sampler (Thermo Fisher Scientific) for two hours, during operational hours of the daycare center. VLPs were extracted and random amplified using The Novel Enrichment Technique of Viromes (NetoVIR) protocol. High-throughput sequencing was performed on an Illumina NovaSeq 6000 system. We identified viruses using EsViritu, a tool for mapping reads against reference genomes of human and animal viruses as well as using de-novo assembly pipeline ViPER, BLASTn and DIAMOND.

Alongside the expected human respiratory viruses (e.g. rhinoviruses), our metagenomic analysis revealed the presence of a wide spectrum of viruses, including plant, fungi, and insect-related viruses, namely viruses from families of Alphaflexiviridae, Dicistroviridae, Totiviridae. However, human polyomaviruses were consistently detected in all but seven samples, with a notable prevalence of WU Polyomavirus during the summer and Human polyomavirus 10 (HPyV10) throughout the rest of the year. Human and animal enteric viruses such as Rotavirus A, F, G, Sapovirus, Human Mastadenovirus, and astroviruses were identified in 15 out of 41 samples. The retrieved genomic information allowed the species and subspecies-level identification of several viruses, including rhinoviruses.
Beyond the expected detection of respiratory viruses, shotgun metagenomics on air samples has shown a potential for the surveillance of respiratory, skin-related and enteric viruses as well as the identification of plant, fungal and animal-related viruses.

T21: Molecular characteristics of a localized influenza A(H1N1)pdm09 outbreak in South Africa, 2020

Dikeledi Kekana | National Institute for Communicable Diseases, Johannesburg, South Africa

Haemagglutinin(HA) and neuraminidase(NA) are surface glycoproteins that mutate frequently through antigenic drift. A cluster of A(H1N1)pdm09 infections from a localised outbreak were detected in the Western Cape from week 2 to 15 of 2020, outside of the normal influenza season. Nonpharmaceutical interventions (NPIs) for COVID-19 implemented from week 15, led to low influenza detection for 2020, with only one case from Gauteng in June and a resurgence later in 2021. We sequenced influenza viruses from the outbreak to compare to viruses that circulated in previous season and following the easing of NPIs. Influenza PCR-positive specimens were collected through syndromic respiratory illness surveillance in South Africa. Specimens with Ct <30 were sequenced using Illumina. H1N1 sequences from 2020 (n=70 of 135 cases) were compared to sequences from 2018-2023(n=281). Clades and mutations were determined using NextClade while N-glycosylation sites(NLGs) were predicted using NetNGlyc1.0. Bayesian analyses was achieved using BEAST. 2020 viruses clustered within the 5a.1(57%) and 5a.2(43%) clades. Seven substitutions were introduced within the HA receptor binding site: K130N, N156K(antigenic site Sa), L161I(Sa) and V250A in viruses belonging to clade 5a.2, D187A(Sb), Q189E(Sb), K209M(Sa) in viruses belonging to clade 5a.1 which were also detected in viruses circulating in 2021(5a.1) and 2022(5a.2a.1). Reversion of several conserved mutations was observed in the NA gene from 2020 onwards. The substitutions rate for the HA gene was lowest during 2020 compared to 2018-2022, characteristic of a localised outbreak. Unlike the HA gene, high variability of NLGs was observed in the NA gene in the 2020 cluster. Influenza viruses from the 2020 localised outbreak had unique mutational and glycosylation profiles compared to viruses circulating pre-2020 and continued to circulate when influenza activity returned in 2021 and 2022. Importation might explain genetic variability in 2020 virus.

Session 7: Viral sequence analysis

T22: Representative full-genome alignments of viral clades - a semi-automated approach incorporating amino acids, nucleotides & RNA secondary structures

Sandra Triebel | Friedrich Schiller University Jena, Jena, Germany

Multiple sequence alignments (MSAs) of biological sequences provide insights into evolutionary relationships by showing the conservation of primary sequences and secondary structures. Computing an optimal MSA for an input set greater than three sequences is an NP-hard problem. Heuristics such as the progressive alignment approach can reduce the CPU time to polynomial time. For viruses thousands of genomes are available and it remains unclear which are representative for a specific species. Aligning complete viral genomes poses further challenges due to genetic diversity by error-prone replicases and length variability by deletion of non-essential genes or insertion of host sequences.
We developed ViralClust to select useful sets of representative viral genomes using k-mer distributions, followed by dimension reduction with PCA and clustering with HDBSCAN. One highlight of our approach is the identification of conserved regions ("anchors") on amino acid level to reduce the runtime of the full Sankoff algorithm from O(n[illegible]) to short sequences computable in $O(n^4)$. We aligned subsequences between anchors on nucleotide level, to include information about RNA secondary structures.
Exemplarily for pestiviruses, we selected 59 of 1,685 available complete genomes. Our tool anchorna identified 75 anchors, of which 16 were sufficient to divide the ∼12,000 nt long genome into subsequences, which were then aligned with LocARNA. Finally, the alignments were concatenated and served as input for a final RNA secondary structure prediction. Throughout the process, manual curation in collaboration with virologists ensures a high-quality alignment.
Our workflow generated a full genome alignment of pestiviruses, confirming the presence of known RNA secondary structures and revealing previously undocumented ones. Our goal is to further automate the process to efficiently align viral clades such as influenza A viruses, filoviruses, phleboviruses, and coronaviruses at whole-genome level.

T23: detectEVE: Illuminating the hidden viral footprint in genomic data

Nadja Brait | University of Groningen, Groningen, Netherlands

Endogenous viral elements (EVEs) are remnants of viral genetic material integrated into the host genome. Retroviruses contribute to the majority of EVEs due to their genomic integration during their life cycle, however the latter can also arise from non-retroviral RNA or DNA viruses, then collectively known as non-retroviral (nr)EVEs. However, detecting nrEVEs poses challenges due to their sequence and genomic structural diversity, unlike retroviral elements that often have flanking LTRs facilitating their identification. This contributes to the absence of specific tools designed for nrEVEs detection.
Here, we introduce detectEVE, a user-friendly and efficient tool designed for the accurate identification of nrEVEs within genomic assemblies. detectEVE deviates from other nrEVE detection pipelines, which usually classify sequences in a more rigid manner as either virus associated or not. Instead, we implemented a scaling system assigning confidence scores to hits in protein sequence similarity searches, based on factors such as bit score distributions and search hints associated with viral/EVE, non-viral, false-positive, and retroviral characteristics, allowing for higher sensitivity and specificity. Deployed as a conda package and combined with Snakemake's workflow automation, we ensure easy installation, an isolated working environment, and resource-efficient parallel execution, allowing for rapid large-scale analyses.

Our tool can help to fill current gaps in both host-associated fields and virus-related studies. This includes (i) enhancing genome annotations with metadata for EVE loci, (ii) conducting large-scale paleo-virological studies to explore deep viral evolutionary histories, and (iii) aiding in the identification of actively expressed EVEs in transcriptomic data, reducing the risk of misinterpretations between exogenous viruses and EVEs.

T24: varVAMP: automated pan-specific primer design for tiled full genome sequencing of highly diverse viral pathogens

Jonas Fuchs | Institut für Virologie, Universtitätsklinikum Freiburg, Freiburg, Deutschland

Since the SARS-CoV-2 pandemic, full genome sequencing of viruses from patient specimens has significantly increased and greatly impacted viral epidemiology, diagnostics and basic research. Particular tiled amplicon sequencing in which the viral genome is first amplified in overlapping fragments via PCR enables time- and cost-efficient surveillance of viral pathogens. However, designing primers for viral pathogens with high genomic variability is challenging. Therefore, we developed a novel bioinformatic command-line tool, varVAMP, which allows pan-specific tiled primer design based on multiple sequence alignments.
In contrast to other primer design tools that also rely on primer3, varVAMP integrates degenerate nucleotides into primer sequences and uses a graph-based machine learning to achieve pan-specificity and full genome coverage. Based on currently available genomic sequences, we used varVAMP to design pan-specific tiled primer schemes for SARS-CoV-2, human hepatitis E virus, rat hepatitis E virus, hepatitis A virus, Borna-disease-Virus, and poliovirus and performed Illumina sequencing. For all tested virus isolates and patient samples, we were able to recover near-to-complete genomic sequences using cost-effective wet-lab protocols and computational pipelines already established for SARS-CoV-2.
Here, we present the first primer design suite tailored for viral pan-specific amplicon design. Provided a careful prior data selection, varVAMP greatly simplifies manual and often laborious pan-specific primer design, yielding primers that do not need extensive wet-lab evaluation. The here designed and evaluated primer schemes are a proof-of-principle for the bioinformatic method but could be already used as a cost-effective alternative to non-targeted sequencing approaches.

Session 8: Surveillance / Epidemiology II

T25: CLIMADE: Pandemic Preparedness for Arboviruses in a time of Climate Change

Houriiyah Tegally | Stellenbosch University, South Africa

No abstract available.

T26: Hendra virus in Australian bats show a diffuse spatio-temporal structure

Claude Yinda Kwe | NIAID/NIH, Hamilton, USA

The Henipavirus genus includes viruses that are both zoonotic and highly pathogenic, including the Hendra virus (HeV) and Nipah virus (NiV). These viruses can cause respiratory distress and fatal encephalitis, with case fatality rates ranging

from 40-100%. Despite the availability of an effective vaccine since 2012, equine spillover of HeV has continued since its discovery in 1994. Spillover of HeV occurs through close association between the natural bat reservoir and horses, and then onward transmission to humans in close contact with the horse. To evaluate the temporal and spatial risks of HeV spillover in eastern subtropical Australia, we sampled five bat roosts monthly for approximately three years and screened 9853 samples for Hendra virus. We then sequenced full genomes of positive samples with higher viral loads using next-generation sequencing techniques. Our data showed that viral RNA could be detected year-round, but high viral loads predominantly occurred in winter (June-September). We recovered forty-eight bat HeV genomes (currently only 15 full genomes are available in GenBank). We classified these, and nine additional horse HeV genomes from other past outbreaks, into four lineages with additional cryptic lineages. We also found that some bat strains are closely related to those causing fatal outbreaks in humans and horses, while others constitute novel lineages. Moreover, we showed that individual roosts have multiple circulating lineages, but with each genetic lineage covering a much larger spatial footprint, probably reminiscent of bat mobility patterns. Our results suggest that bats, which fly long distances and often aggregate in mixed-species roosts, may maintain a more diverse population of HeV variants than was previously known. This project was funded by the Intramural research Program of NIAID, NIH.

T27: Sequencing error inspires novel phylogenetic approaches to sequence data

Chris Illingworth | MRC-University of Glasgow Centre for Virus Research, Glasgow, UK

Genome sequence data holds exceptional value for understanding virus evolution. Phylogenetic models are used on a regular basis to understand patterns of transmission and viral epidemics. However, phylogenetic models usually regard sequences as being correct, neglecting the role of errors in the process of generating a consensus sequence. Using SARS-CoV-2 sequence data we here show that, where the expected amount of evolution between samples is low, and the rate of sequencing error is non-negligible, accounting for error is potentially important. We describe methods for reconstructing networks of viral transmission, and for understanding the within-host evolution of viruses, which incorporate sequencing error into models involving phylogenetic reconstruction. Our results provide insight into the spread of SARS-CoV-2 infection on hospital wards, and the complex nature of within-host SARS-CoV-2 evolution.

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

Poster Session A

A01	**Assessing distinctiveness profiling for predicting SARS-CoV-2 variant success** A S M Rubayet Ul Alam | University of Glasgow, Glasgow, United Kingdom
A04	**A novel Illumina iSeq100-based next-generation sequencing (mNGS) pipeline for virus discovery, genomics and outbreak prevention/management** Daniel Cadar | Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany
A05	**GUT VIROME IS ASSOCIATED WITH SUBSEQUENT CELIAC DISEASE IN TWO EUROPEAN PROSPECTIVE BIRTH COHORTS** Katerina Chuda | Charles University, Second Faculty of Medicine, Prague, Czechia
A06	**Genome Detective dengue subtyping tool** Sara Cleemput | Emweb, Herent, Belgium
A07	**Cameroonian onchocerciasis vectors (Diptera: Simuliidae) harbor a plethora of novel (RNA) viruses** Lander De Coninck | KU Leuven, Leuven, Belgium
A08	**Virus-host interplay in polar sea ice** Tatiana Demina | Helsinki University, Helsinki, Finland
A09	**Evolution of Crassvirales in the context of their hosts** Mikhail Fofanov | Friedrich Schiller University Jena / Utrecht University, Jena, Germany
A10	**Phylogeographic history of SARS-CoV-2 variants in Spain** Pilar Gallego-García | Universidade de Vigo, Vigo, Spain
A11	**Navigating sampling bias in discrete phylogeographic analysis: assessing the performance of an adjusted Bayes factor** Fabiana Gambaro | ULB, Brussels, Belgium
A12	**Extrahepatic replication and genomics signatures of the hepatitis E virus in the kidney** André Gömer | Ruhr University Bochum, Bochum, Germany
A13	**Epidemiology and global spread of emerging tick-borne Alongshan virus** Saskia Janshoff | Ruhr University Bochum, Bochum, Germany
A14	**Surveillance of dengue and other arboviruses in mosquitoes from Dar es Salaam, Tanzania** Silvan Hälg | Swiss TPH, Allschwil, Switzerland
A15	**Unravelling the Virome from Different Anatomical Sites in Psoriasis Patients** Dominika Kadlečková | Charles University, Faculty of Science, Prague, Czechia
A16	Kijin Kim | Simon Fraser University, Richmond, Canada

A17	**Viralgenie, a pipeline for viral metagenome sequencing data: applications in Lassa virus research** Joon Klaps | KULeuven, Leuven, Belgium
A18	**Detecting recombination events in closely related viruses from sequence data** Sarah Krautwurst | Friedrich Schiller University Jena, Jena, Germany
A19	**ViralZone enzymes: improved overview of the viral reaction landscape.** Philippe Le Mercier | Swiss Institute of Bioinformatics, Geneva, Switzerland
A20	**Data and Resources of the International Committee on Taxonomy of Viruses: ICTV** Elliot Lefkowitz | University of Alabama at Birmingham, Birmingham, Alabama, USA
A21	**Data and resources of the Bacterial and Viral Bioinformatics Resource Center (BV-BRC)** Elliot Lefkowitz | University of Alabama at Birmingham, Birmingham, Alabama, USA
A22	**Unveiling Ribes virome: A Comprehensive Exploration Study** Rebeka Ludviga | Latvian Biomedical Research and Study centre, Riga, Latvia
A23	**Sex and viruses: specific virome structure between males and females depending on mosquito species** Côme MOREL | CIRAD / UMR ASTRE, Montferrier-sur-Lez, France
A24	**Diversity and evolution of the deltavirus-like agents in termites** Jose Gabriel Nino Barreat | University of Oxford, Oxford, United Kingdom
A25	**Revealing dynamic genomic rearrangements in hepatitis E virus infection using Hyper-Eins** Maximilian Klaus Nocke | Ruhr University Bochum, Bochum, Germany
A26	**Insights into the gut microbiota development in healthy Bangladeshi infants: exploring the gut virome during early life.** Maria Ioanna Papadaki | KU Leuven, Leuven, Belgium
A27	**Phage hunting at the air-water boundary of the Central Arctic** Janina Rahlff | Linnaeus University, Kalmar, Sweden
A28	**Virome analyses of metatranscriptome data using NeoRdRp** Shoichi Sakaguchi | Osaka Medical and Pharmaceutical University, Takatsuki, JAPAN
A29	**Dissecting the gene expression dynamics during early HCMV infection** Lygeri Sakellaridi | University of Regensburg, Würzburg, Germany
A30	**Determinants of Species-Specific Hepatitis E Virus Pathogenicity** Leyla Sirkinti | Ruhr University/ Medical Virology, Bochum, Germany
A31	**Paint4IRAS. Development of biocidal paints against viruses and bacteria in healthcare settings.** David Talavera Cortés | Universitat de València, Valencia, Spain
A32	**Exploration the virome of European mosquitoes** Gabor Toth | Bernhard-Nocht Institute for Tropical Diseases, Hamburg, Germany
A33	**Phosphoproteomic analysis of Aedes aegypti Aag-2 cells infected with the arboviruses MAYV and CHIKV** Anna Fernanda Vasconcellos | University of Brasilia, Brasília, Brazil

A34	**A deep learning pipeline for Bacteriophage detection** Rajitha Yasas Wijesekara | University Medicine Greifswald, Greifswald, Germany
A35	**VAZyMolO-2 - How did you annotate your new viral genome without it ?** Vincent WILDE | CNRS-UMR725 (AFMB), Marseille, France
A36	**Using protein language models to investigate the evolutionary processes leading to virus attenuation** Francesca Young | MRC University of Glasgow Centre for Virus Research, Glasgow, UK
A37	**Mriyaviricetes: a new class of viruses related to Nucleocytoviricota** Natalya Yutin | NCBI, Bethesda, USA
A38	**Are C-terminal anchor (CTA) endolysins a thing? Identification of possible CTA endolysins across sequences of uncultivated bacteriophages** Nikita Zrelovs | Latvian Biomedical Research and Study Centre; University of Latvia, Riga, Latvia
A39	**Genomic monitoring to unravel the emergence and maintenance of arboviruses in the Netherlands** Emmanuelle Münger | Erasmus MC, Rotterdam, Netherlands

Poster Session B

B01	**The Virome of the Western Honey Bee in Europe** Nikolas Basler | KU Leuven, Leuven, Belgium
B02	**Identification of tandem repeats of an endogenous RdRp-like element in multiple species of butterfly and moth.** Katy Brown | University of Cambridge, Cambridge, UK
B03	**On the importance of assessing topological convergence in Bayesian phylogenetic inference** Marius Brusselmans | KU Leuven, Leuven, Belgium
B04	**Virus metagenomics reveals high diversity of novel newlaviruses in Canadian seals** Marta Canuti | University of Copenhagen, Copenhagen, Denmark
B05	**Phage genome architecture and GC content: Structural genes and where to find them** Ritam Das | Aero-Aquatic Virus Research Group, Friedrich Schiller University Jena, Jena, Germany
B06	**BLMPred: predicting linear B-cell epitopes using pre-trained protein language models and machine learning** Barnali Das | Technical University of Munich, Munich, Germany
B07	**Disentangling viral strains using long reads** Roland Faure | Université Libre de Bruxelles, Brussels, Belgium
B08	**Evaluation of Different Extraction Techniques of Skin Samples for Virome Analysis** Lisa Faye | KU LEUVEN, Leuven, Belgium

B09	**VILOCA: Local haplotype reconstruction and mutation calling for short- and long-read viral sequencing data** Lara Fuhrmann | ETH Zurich, Basel, Switzerland
B10	**Exploring the human virome in chronic liver disease with integrated viral and bacterial metagenomics data** Emilio Rafael Garcia Rios | TUM, Freising, Germany
B11	**Assessing the genetic stability of phage therapeutic products through deep sequencing** Nathalie Goeders | Sciensano, Brussels, Belgium
B12	**Coupling homology search and network analysis for OTU clustering in eukaryotic viromes** Serafin Gutierrez | CIRAD, Montpellier, France
B13	**Bioinformatic analysis of clinical samples from an HEV-infected pregnant woman using metagenomics** Marta Ibañez-Lligoña | Vall d'Hebron Institute of Research (VHIR), Barcelona, Spain
B14	**Viral metagenomic workflow for identification of viruses from bats** Camille Johnston | Statens Serum Institut, Copenhagen, Denmark
B15	**Detection of RNA viruses in Dutch natural plant ecosystems** Dimitris Karapliafis | Wageningen University & Research, Wageningen, Netherlands
B16	**Phage-Antibiotic Synergy: definition, evaluation & prediction** Eliška Kučerová | Masaryk University, Brno, Czech Republic
B17	**Catching them all: Assessing the diversity and evolution of Iflaviridae genomes found in publicly available Lepidoptera sequencing data** Anne Kupczok | Wageningen University, Wageningen, Netherlands
B18	**RNAswarm: Differential Analysis of RNA-RNA Interactions in Influenza A Virus** Gabriel Lencioni Lovate | Friedrich Schiller University Jena, Jena, Germany
B19	**Spread.gl: Visualising Pathogen Dispersal in a High-performance Browser Application** Yimin Li | Rega Institute, Leuven, Belgium
B20	**Improved prediction of protein-protein interactions using a next sentence cross-encoder** Dan Liu | University of Glasgow, Glasgow, United Kingdom
B22	**Ancient viral discoveries through ancient metagenomic data** Luca Nishimura | The University of Tokyo, Minato-ku, Japan
B23	**The Intriguing World of Unknowns: exploration of Polycipiviridae genomes, Phylogenetics and Host Specificity** Ingrida Olendraite | University of Cambridge, Cambridge, United Kingdom
B24	**Impact of genetic reassortment on the structure of RNA genomic segments of Influenza A viruses** Rithu Paul Stansilaus | University of Strasbourg, Strasbourg, France

B25	**Transmission-mediated adaptation of virulence in Deformed Wing Virus** Harshit Kumar Prajapati | Freie Universität Berlin, Berlin, Germany
B26	**Genomic Echoes: Exploring Challenges and Opportunities in Identifying Endogenous Viral Elements** Muriel Ritsch | Friedrich Schiller University Jena, Jena, Germany, Jena, Germany
B27	**Exploring the in-silico molecular docking method to study the interactions of novel entry inhibitors against multiple HIV subtypes.** ARADHANA SINGH | SOUTH ASIAN UNIVERSITY, NEW DELHI, INDIA
B29	**Assessment of COVID-19 contact tracing network accuracy via phylogenetic analysis of community-level SARS-CoV-2 genomic data** Jonathan Thibaut | Catholic University of Leuven (KU Leuven), Leuven, Belgium
B30	**Control of human anelloviruses by cytosine to uracil genome editing** Anne Timmerman | Amsterdam UMC, Amserdam, Netherlands
B31	**Dissecting the unknown specificity of cell wall-binding domains of phage endolysins** Roberto Vázquez | Ghent University, Ghent, Belgium
B32	**From water to wild plants: exploring plant virus diversity of various sample types through high-throughput sequencing data analysis.** Lana Vogrinec | National Institute of Biology, Ljubljana, Slovenia
B33	**Viral community diversity decreases from the bulk soil to the rhizosphere** Lingyi Wu | Utrecht University, Utrecht, Netherlands
B34	**Dissect antiviral immunity in pluripotent stem cells via single-cell transcriptomics** Qing Yang | Fred Hutchinson Cancer Center, Seattle, United States
B35	**Identification of a tomato brown rugose fruit virus mutant isolate overcoming virus-specific resistance in new resistant tomato cultivar** Zafeiro Zisi | Laboratory of Viral Metagenomics, KU Leuven & Scientia Terrae, Leuven, Belgium
B36	**Comparative Evaluation of Bioinformatic Pipelines for Full-Length Viral Genome Assembly** Levente Zsichla | Eötvös Loránd University, Budapest, Hungary
B37	**Parallel evolution of Drosophila C virus independent of host RNAi status** Oscar Morales Lezcano | Radboudumc, Nijmegen, Netherlands
B38	**Fragmentation of AlphaFold structures for improved interaction prediction** Stefaan Verwimp | KU Leuven, Leuven, Belgium
B39	**Comparison of R9 and novel R10 Nanopore flow cells for mixed multiplexed viral amplicon sequencing** David Nieuwenhuijse | Erasmus MC, Rotterdam, Netherlands