Date: December 2024

Final report

From negative to positive - New diagnostic to determine the aetiology of unknown viral infections in horses

Project number: H-20-47-555

Project period: 2021-01-01 – 2024-12-31

Main applicant:
Anne-Lie Blomström, Swedish University of Agricultural Sciences, anne-lie.blomstrom@slu.se

Co-applicant(s):
Miia Riihimäki, Swedish University of Agricultural Sciences
Siv Hanchen-Olsen, Norwegian University of Life Sciences
Anna Myykänen, University of Helsinki
Sanni Hansen, University of Copenhagen

Part 1: Detailed summary

Bakgrund och syfte
Virusinfektioner hos hästar har betydande inverkan på djurvälfärd och kan leda till stora ekonomiska förluster. Ett av de första tecknen på en virusinfektion är utvecklingen av feber och att i detta tidiga stadium kunna identifiera orsaken bakom infektionen skulle leda till snabb och korrekt behandling samt möjliggöra implementering av lämpliga biosäkerhetsåtgärder för att stoppa ytterligare spridning. Men feber och andra kliniska symptom på viral infektion, t.ex. neurologiska störningar, förblir ofta odiagnostiserade.

I detta projekt har vi därför använt så kallad virusmetagenomik för att undersöka den virala orsaken bakom feberfall (endast feber och feber kombinerat med diarré) samt bakom neurologiska störningar. Virusmetagenomik har använts eftersom denna metod gör det möjligt för oss att simultant identifiera alla virus som finns i ett kliniskt prov utan att i förväg specificera vilket virus som vi söker efter. För att även kunna bidra till snabbare korrekt diagnostik vid exempelvis akuta virusutbrott av okänd orsak samt inte vara beroende av stora och dyra sekvenseringsmaskiner har vi även titta på användningen av Nanopore (MinION) sekvensering. Detta är en liten portabel sekvensator som kan kopplas in i en dator via USB. Därmed hoppas vi att kunskapen från detta projekt kan bidra till att förbättra hästhälsan och kontrollera spridningen av dessa virusinfektioner.

Metoder
Totalt har 275 häst provtagits i denna studie, alla hästar faller i en av 4 kategorier: (1) feber, (2) feber och diarré, (3) neurologiska symptom samt (4) hästar som inte uppvisar tecken på infektion (kontrollhästar). Från alla dess hästar har blodprover och nässvabbar tagits, från kategori 2 togs även fekalprover och från en av hästarna i kategori 4 mottogs hjärnvävnad. Den virala metagenomisk analysen består kortfattat (se del 2 av rapporten för detaljerad

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

beskrivning) av följande steg: provtagning, eventuell homogenisering av prov, filtrering för att bli av med bakterier, extraktion av DNA och RNA, ospecifik amplifiering av arvsmassan, storskaligsekvensering och slutligen bioinformatiska analyser för att identifiera virus. Efter virusmetagenomiken har, i vissa fall, ytterligare analyser körts för att karakterisera hela/stora delar av den virala arvsmassan. Även kvantitativ realtids-PCRer (qPCR) har utvecklats för vissa av virusen, dessa qPCRer har sedan används för att screena de individuella proverna.

Resultat och slutsatser
Genom dessa studier har vi påvisat förekomsten av ett antal virus inte bara för första gången i Skandinavien utan även i Europa. I hästar med feber identifierades i nässvabbarna samt i serum ett antal såväl DNA som RNA virus. Förekomsten av vissa av dessa virus såsom olika ekvina herpesvirus (EHV-2, -5 och -4) är sedan tidigare kända virala agens i såväl Skandinavien som i hela världen. Av dessa var EHV-2 och EHV-5 de mest förekommande virusen i nässvabbarna, detta gällde oavsett hälsostatus (feber- vs kontrollhästar). Mera intressant var att vi identifierade virusmedlemmar ur familjerna *Flaviviridae* (ekvint pegivirus), *Anelloviridae (*ekvint torque teno virus 1 och 2), *Parvoviridae* (ekvint copivirus och ekvint parvovirus CSF) och *Circoviridae* (ekvint circovirus 1) av vilka merparten av dessa virus endast har identifierats tidigare i USA och i ett fall även i Kina. Eftersom dessa inte beskrivits här tidigare så har vi genetiska karakteriserat deras arvsmassa och jämför med virus från US och Kina. Dessa studier har visat på en hög sekvenslikhet särskilt för de virus med DNA arvsmassa. Genom qPCR analyser har vi sett att 1.6% - 9.4% av hästarna som analyserats (feber- och kontrollhästar) är positive för dessa olika virus. För ekvint circovirus och ekvint torque teno virus sågs en högre positivitet hos hästar med feber.

I hästar med feber och diarré så var det primära virusfyndet i avföringsproverna ekvint coronavirus. Detta är en viral patogen som även sedan tidigare är känd för att orsaka feber och diarré hos hästar runt om i världen. Förutom coronavirus identifierades även virursekvenser som uppvisade låg sekvenslikhet (<30% på proteinnivå) till olika picornavirus. Tyvärr var dessa sekvenser få och vi har ännu inte lyckats karakterisera genomet av detta/dessa väldigt divergenta virus.

I en av nässvabbarna från hästarna med neurologiska symptom identifierades ett nytt ekvint papillomavirus. Även om vi inte nödvändigtvis anser att detta virus är kopplat till sjukdomsbilden så valde vi att genetiskt karakterisera det då detta papillomavirus inte tidigare beskrivits. Dessa analyser visar att dess virusarvsmassa skiljer så pass mycket genetiskt från kända virus att det troligen utgör ett nytt genus. Ett fynd till från de neurologiska hästarna av intresse är att vi i den undersökta hjärnvävnaden från en häst identifierade ekvint torque teno virus 1. Detta virus har aldrig identifierats i hjärnvävnad från häst tidigare. Det human motsvariga torque teno viruset har dock en gång identifierats i hjärnvävnad hos en individ (ej neurologiska symptom) och detta viruset har även identifierats i human CSF prover från encefalitfall.

Slutligen har vi visat att Nanopore sekvensering (MinION) är ett bra alternativ till Illuminas högkapacitetssekvensering och har en högkänslighet med möjlighet att identifiera även virus som det endast finns få partiklar av. Med denna metod kan man förutom att detektera även genetiskt karakteriser de identifierande virusen. Detta möjliggör användning av denna teknik vid exempelvis akuta virusutbrott.

Part 2: Main report (max. 10 pages)

Introduction

Viral infection of horses pose a major threat to the equine welfare and cause great economical losses for the equine industry worldwide. Infections of different viral agents can lead to a wide spectrum of clinical symptoms and different outcomes ranging from sub-clinical infections to life-threating conditions are observed. One of the first warning signs of a viral infection is the development of high fever [1-3]. If we could already at this stage identify the causative agent, it would enable us to more rapidly set in correct treatment and give a more accurate prognosis as well as decrease the viral spread to other horses. Unfortunately, most often the causative agent of fever cannot be identified with the common diagnostic measurements and thus remains unknown.

Neurological disorders in horses are often life-threating conditions and accurate diagnostic tests are necessary for correct treatment and prognosis [3]. Even if these disorders are not as common as the clinical presentation of fever alone, the outcome is often dramatic, as was seen during the EHV-1 outbreak in 2018 in Sweden. These horses at first often display signs of fever but then also develop mild to severe clinical neurological signs. Although we know of several viruses causing encephalitis in horses worldwide [3], such as herpesvirus, different flaviviruses, orthobornavirus etc, for the majority of the cases in Sweden and globally the virus causing these diseases remains unknown.

Therefore, in this project we aimed to identify viruses that cause fever and/or neurological signs but that are not recognized with common diagnostic tests. To do this we used an approach called viral metagenomics that allows us to simultaneously identify all viruses in a particular sample without targeting specific viruses as well as identify previously unknown viruses. This approach combines high-throughput sequencing (HTS) with advanced data analysis (bioinformatics) [4]. In addition, we evaluated if this approach is something that could be used in a clinical setting to allow for a rapid response during, for example, an infectious disease outbreak of unknown cause.

Material and methods

Sampling

Horses (no. 275) included in this study have been divided in to the following categories: fever (>38.5°C); fever (>38.5°C) with diarrhoea; neurological signs and controls. The latter are horses not displaying any signs of a viral infections. All the horses have been sampled no later than 48 hours after arrival/or after fever peak.

From all the horses, blood samples were taken in order to obtain serum and plasma. Nasal swabs were also taken from all horses. One of the neurological horses included were euthanized and was autopsied in collaboration with the Swedish Veterinary Agency and the pathology section at SLU and from this horse also brain tissue was obtained. All samples have been stored in -80°C to maintain the integrity of the virus particles and its nucleic acid.

Sample preparation metagenomics

All the samples were processed through different pre-preparations steps prior to nucleic acid extractions.

For the fecal and tissue samples homogenization was performed. The fecal samples were placed in ZR BashingBead tube containing DNA/RNA Shield and homogenized at +4°C

using a Vortex Genie 2. The tubes were then centrifuged and the supernatant collected. The tissue samples were placed in Precellys CK14 tubes containing sterile PBS and homogenized at +4°C using a Cryolys Evolution. The tubes were then centrifuged and the supernatant collected.

In order to remove bacteria, all the samples (serum, nasal swab buffer, as well as the supernatant from the fecal samples and from the tissues) were filtrated through a 0.45µm filter. The filtrate from each sample was divided in to two aliquots in order to allow a specific RNA and DNA virus pathway, respectively. To the RNA aliquots, TRIzol was added and RNA extracted as described below. The DNA aliquots were subjected to RNAse and DNase treatment with the aim to reduce the host (horse) nucleic acid so that the overall proportion of viral nucleic acid would increase.

Metagenomic nucleic acid (RNA and DNA) extraction
DNA was extracted from the nuclease-treated DNA aliquots using Genejet DNA purification kit.

The RNA was extracted from RNA aliquots using a combination of the Trizol and Genejet RNA purification kit protocol. In order to remove potential DNA contaminants, the samples were treated with DNase and at the same time concentrated using the RNeasy MinElute Cleanup Kit. As a large proportion of the RNA is host (and potential bacterial) ribosomal RNA (rRNA) the RNA was subjected to rRNA depletion using rRNA depletion Ribo-Zero Plus rRNA Depletion Kit.

Nucleic acid labelling and random PCR (rPCR)
Due to the different steps aimed to remove host and bacterial nucleic acid the concentration of the extracted nucleic acid was too low to sequence directly and therefore needed to be amplified prior to high-throughput sequencing.

For the RNA, two approaches were used, Ribo-SPIA and random PCR (rPCR). Initially the plan was to use Ribo-SPIA, which is an isothermal strand-displacement amplification process that uses a DNA/RNA chimeric random SPIA primer, DNA polymerase and RNase H. However, after some failed attempts to get this approach to work for serum (most likely due to low RNA concentrations) we switched to using rPCR instead. In the rPCR approach, all the nucleic acid is during the cDNA synthesis labelled with a tag-sequence on both ends and then a PCR is run using a primer complementary to the tag-sequence. This approach allows all nucleic acid in the samples, regardless of its genetic background, to be amplified simultaneously.

rPCR was also used to amplify the extracted DNA. In this case the DNA was labelled with the previous mentioned tag-sequence during a klenow fragment reaction. The labelled DNA was then amplified through rPCR a described above.

From this step, the remining steps were identical for the RNA and DNA samples.

All the rPCR products were purified using Genejet PCR purification kit.

High-throughput sequencing
The metagenomic samples in this study have been sequenced through one of the following high-throughput sequencing approaches: Illumina or Nanopore (MinION) sequencing.

For the Illumina sequencing, sequencing libraries was constructed using Nextera XT DNA Library Preparation Kit and barcoded with Nextera XT Index Kit. The libraries was sequenced using the NextSeq500.

For the Nanopore sequencing, sequencing libraries were constructed using either the Native Barcoding Kit 24 V14 or Ligation Sequencing Kit V14. The libraries were sequenced on flow cells (R10.4.1) for approximately 18 hours using live basecalling.

Bioinformatic analysis
Oxford Nanopore Technologies provides an open-source EPI2ME platform containing different analysis workflows that can be run either real-time i.e. as the sequences are being produced or after the sequencing is completed. Hence, the initial analysis of all the Nanopore sequences was processed through this platform using the "What is in my pot" WIMP workflow and this provides an annotation of each read.

In addition, all the sequence reads (both Illumina and Nanopore) were processed though our own standard metagenomic sequence analysis workflow as described below.

CLC genomic workbench was used to perform quality check of all raw reads, to trim and remove bad quality sequence data as well as very short reads, to perform *de novo* assembly in order to create longer contigs of overlapping sequences as well as make sequence lists containing the created contigs as well as the remaining unassembled sequences (singletons).

The sequence list (contigs/singletons) were uploaded to Uppmax and Diamond, a rapid sequence aligner for protein and translated DNA searches, analysis was performed in order to annotate the sequences. The Diamond results were visualized using Megan, a taxonomic binning tool for second-generation metagenomic sequencing projects.

The Diamond/Megan results/output was further analysed/confirmed using blastn and blastx searches of viral hits of interest.

For virus sequences showing >70% sequence similarity to viral genomes available in GenBank a reference genome was downloaded to the CLC genomic workbench and a reference assembly was performed in order to investigate the genome coverage, identity and if possibly obtain the complete viral genome present in our sample/s.

Genetic characterization
In some cases, the CLC genomic reference assembly was used to design primers to cover the gaps of the genome. The primers were used to, by PCR, amplify the missing regions. The PCR products were purified and sent for Sanger sequencing (Macrogen Europe), in some cases the amplified PCR products were also sequenced by Nanopore technology. The sequenced reads where then further processed using CLC genomic workbench.

qPCR screening
To screen the individual samples for selected viruses detected through the high-throughput sequencing and allow an relative comparison of viral amount between the different samples, quantitative real-time PCR systems (qPCR) were set up. In short, primers were designed based on the HTS reads. To account for possible sequencing errors blast alignments were investigated to find appropriate primer regions.

Nucleic acid from all the samples used in the qPCR screen was extracted using GeneJET Viral DNA and RNA Purification Kit. This allowed simultaneous extraction of both RNA and DNA as well as a comparison of relative viral amounts between the different samples.

The qPCR was run using iTaq universal SYBR® Green one-step kit (RNA viruses) and iTaq universal SYBR® Green kit (DNA viruses).

Results and discussion

Sequencing output – general overview

For each of the sequenced sample/pool of samples approximately between 1-5 million sequence reads were generated and processed through the previously described metagenomic pathway. The bioinformatic analysis showed that despite the pre-processing steps included, with the aim to remove host and bacterial nucleic acid, a large proportion of the annotated reads were of eukaryotic (host) and bacterial origin. This was not unexpected as the same pattern is shown in other metagenomic studies and merely highlights the need for high-throughput sequencing in this type of samples. The percentage of viral reads in the different samples varied from <1% to up to towards 30% of the annotated reads. In addition, a large number of reads remained unclassified, this is also often seen in different viral metagenomic studies and could account for e.g., bacterial and viral reads to divergent to be classified through the approached used as well as to non-coding host nucleic acid.

In this report we will focus on the specific viral findings in the different disease categories (fever; fever and diarrhea and; neurological signs). More specifically, although phages (viruses infecting bacteria) and so-called giant viruses were found among the identified viruses, we will focus only on viruses infecting eukaryotes as these are most relevant for disease.

Metagenomics – Horses with fever

For the horses with fever, metagenomic analysis was performed on both nasal swabs and serum. Nasal swabs were used as respiratory viruses often cause fever and as indirect viral transmission through aerosol and droplet transmission is a common transmission route of viruses between different horses. In addition, the use of serum allows the identification of potential viremic viruses.

Comparing the nasal swabs to the serum samples the proportion of viral reads were in most cases higher in the nasal swab pools compared to the serum sample pools. In the serum samples, most often the annotated reads only consisted of <1% viral reads. Although virus overlap was seen between these sample types it was a clear difference in the different viromes observed depending on sample type.

Despite the sometimes-low proportions of viral reads, a number of different RNA and DNA viruses were identified in the serum and nasal swabs.

Virus families in the nasal swabs: *Orthoherpesviridae, Anelloviridae, Genomoviridae, Papillomaviridae, Parvoviridae, Retroviridae, Riboviria and Picornavirales.*

Virus families in the serum samples: *Anelloviridae, Genomoviridae, Parvoviridae, Flaviviridae and Retroviridae.*

A selection of these viruses is described below.

Orthoherpesviridae
In all the nasal swabs, equine herpesviruses (EHVs) were the most abundant viruses and the identified EHVs classified in two genera *Percavirus* and *Varicellovirus*. The EHVs of most veterinary significance, EHV-1 and EHV-4, classifies within genus *Varicellovirus*. EHV-1 is one of the most important pathogens of horses causing primarily respiratory disease but can also cause abortions and neurological disease [5]. EHV-4 cause acute respiratory disease in horses [6]. In genus *Percavirus*, one find EHV-2 and EHV-5 and these two are widespread viruses in the equine population and this common presence makes it difficult to associate them to disease. In our samples, EHV-2 and EHV-5 were the most abundant herpesviruses identified through the viral metagenomic analysis and out of these two EHV-2 reads were the most abundant. Equine herpesviruses are dsDNA viruses with a large genome of almost 200 000 bp and although no complete genome was obtained a large portion of the genome was covered.

Although, we have not gone through all complete EHV-2 proteins many such e.g. the envelope glycoprotein N, virion protein G52 and deoxyuridine triphosphatase showed a high (up to 100%) amino acid (a.a) identity to the closest publicly available EHV-2. EHV-4 was also identified, but in fewer samples and with less genome sequencing depths. EHV-1 reads were only identified in one sequencing pool. The latter sequencing pool consisted of samples from Denmark and from here only a few number of EHV-1 reads were detected.

Anelloviridae
In both nasal swabs and serum samples, reads classifying to family *Anelloviridae* and genus *Mutorquevirus* was identified in the samples sequenced from horses with fever. In the Swedish samples all of the anellovirus reads showed closest similarity to *Torque Teno equus virus 1* (eqTTV1) and classified as species Mutorquevirus equid1. No equine TTVs have previously been identified in Europe. In fact, it has previously only been described in the US. The complete genome was obtained and it showed a high nucleotide sequence similarity (98.6%) to the complete genome eqTTV1from a horse in the US [7]. Interestingly, anellovirus reads were also identified in serum samples from horses with fever in Denmark but in these another species namely *Torque teno equus virus 2* was identified. Also this species has only been identified in the US before [8]. Unfortunately, only few reads were obtained just covering a smaller portion of the genome, although to a high similarity (>98%). Hence, very few studies have been done on equine TTVs and this is the first detection of the virus outside of the US. The high sequence similarity between the Scandinavian eqTTVs and those from the US show a high genetic conservation both geographically and over time. At present no data is available on a potential pathogenicity of TTVs and thus renders further investigations.

Circoviridae
In one of the serum pools from the fever horses, equine circovirus 1 (eqCV1) was detected. This virus has, as eqTTV, only been detected in the US where it was detected on a horse with hepatitis and fever [9]. The genome from the circovirus detected in Sweden showed high similarity (98.56% nt identity and 100% aa identity) to that described from the US. As for eqTTV, no knowledge about the potential impact of eqCV1 on equine welfare is known due to the limited studies on this virus. Circoviruses have many been studied in pigs as porcine circovirus type 2 (PCV-2) is recognized as a significant pathogen connected to different porcine circovirus disease complexes [10]. Although PCV-2 is considered the main pathogen

in these diseases, one has also investigated the role of co-infection with porcine TTV something that is interesting also here as co-infection of eqTTV1 and eqCV1 was observed.

Parvorviridae
Virus in the family *Parvoviridae*, genus *Copiparvovirus* were identified in all Swedish serum pools as well as in one nasal swab pool. It was also identified in serum samples from Denmark. The parvovirus (equine copivirus (eqCopV)) identified in Swedish horses was for the first time identified in a horse with respiratory signs in the US and has before this study only been detected in the US [11, 12] and China [13]. The complete coding part of the genome was obtained showing 92-94% nucleotide sequence similarity to the previous mentioned viruses. In the samples from Denmark another parvovirus, namely equine parvovirus CSF (EqPV-CSF), was also identified. This virus was identified for the first time in the US from a horse displaying neurological signs [11]. However, few viral reads were obtained and the complete genome could not be obtained. Also these viruses are understudied and looking at parvoviruses in other species we find both known pathogens as well as those to who no disease can be linked.

Flaviviridae
The metagenomic analysis identified a flavivirus in the genus *Pegivirus* in all the serum pools from the horses with fever in Sweden but not in the pools from the healthy pools neither in any of the samples from the other countries. The complete coding genome was characterised (10 035 nt), for the identified equine copivirus (eqCopV) and it coded for a 3 344 amino acid long polyprotein. The identified eqCopV showed on nucleotide and protein level a 90-91% and 98-99% identity, respectively, to viruses from the US [14, 15] and China [16]. In the samples from Denmark, the complete genome of an equine Hepacivirus (eqHV) was obtained from the serum samples. This genome showed most similarity (94% nt identity) to an equine isolate from Europe (MH028004.). EqHVs are viruses that have the potential to cause chronic hepatitis and liver failure in horses.

qPCR screening
The screening of individual horses revealed a low positivity rate for most of these viruses (Table 1). The exception was for EHV2 and 5 that displayed an overall total high positivity rate for both viruses of 66% and 56%, respectively. For the horses with fever, 73% and 65% of the screened horses were positive compared to 56% and 44% of the control horses. For the other viruses (EqTTV, EqPV, eqCopV, EqCV1), investigated with qPCR screening, an overall positivity rate was between 1.6% - 9.4%. In general, a high ct-value (>30) was observed for the viruses detected in the serum indicating low levels of viruses in the samples. The exception was the eqTTV1 positive samples that had an average ct-value of 24. In the nasal swabs, the eqTTV1 ct-values varied between 24 – 37. In regards to EHV-2 and EHV-5, the ct-values varied between 18-39 and 19-39, respectively, indicating a great variation of virus levels in the different samples. In all of the cases, apart from two serum sample, only one virus was identified per sample. In the serum samples with a co-infection both eqTTV1 and EqCV1 as well as EqPV and eqCopV, respectively was identified. In the nasal swab a co-infection between EHV2 and EHV-5 was commonly observed.

Table 1. Positivity rate (%) of qPCR screened individual horses

	Serum			**Nasal swabs**		
	Overall	*Fever*	*Control*	*Overall*	*Fever*	*Control*
EHV2	n/a	n/a	n/a	65.6	73.0	55.6
EHV5	n/a	n/a	n/a	56.3	64.9	44.4

EqTTV1	6.3	8.1	3.7	7.8	13.5	0
EqPV	4.7	2.7	7.4	n/a	n/a	n/a
eqCopV	9.4	5.4	14.8	n/a	n/a	n/a
EqCV1	1.6	2.7	0	n/a	n/a	n/a

Metagenomics – Horses with fever and diarrhea

For these horses, viral metagenomics were performed on fecal material and the high-throughput sequencing yielded around 4 million sequencing reads per pool. In the diarrheic horses approximately 2.5% of the classified reads were viral reads while 78% classified as archea and 18% as bacteria. The corresponding values for the control were <1% (viruses), 70% (archea) and 28% (bacteria). Majority of the identified viruses were phages but a few eukaryotic viruses were identified as described below.

Coronaviridae

The main eukaryotic virus finding was coronavirus and these viruses were identified in both horses with diarrhea as well as in the control horses. A low number of reads from the samples mapped to different regions of a reference equine coronavirus genome, however, the reads were very few and no complete genome was obtained Blast analysis of one of the created contigs showed a 90% nucleotide similarity to the closest match (an equine coronavirus from a horse in the US). Equine coronavirus is a well-known pathogen that can cause both fever and diarrhea [17]. It is globally widespread and known to be present in Europe including Scandinavia.

Other viruses

In addition to coronavirus sequences, a few sequences/contigs matched with low amino acid similarity (around 30%) to the RNA-dependent RNA polymerase of viruses within the picornavirales. Unfortunately, the reads were too few for us to be able to genetically characterise these novel virus(es) and hence these requires additional attention.

Metagenomics – Horses with neurological signs

For the horse with neurological signs, the metagenomic analysis was done on nasal swabs and serum samples with the exception of one horse that was euthanized and from this sample brain tissue material was used.

Eutanized horse - Brain tissue

In the euthanized horse with neurological signs an autopsy was performed. No pathological typical signs of viral infection was observed in the investigated brain tissue, but despite this material was collected to confirm this finding through viral metagenomics. As expected, a low percentage (<1%) of the classified reads were of viral origin. All of the classified viral reads were eqTTV1, i.e. the virus previously discussed under the “Fever horse section”. The near-complete genome was obtained and this displayed high sequence identity to the other available eqTTV1s. This is the first identification of eqTTV in the brain tissue of a horse. In humans, TTV have been detected in the CSF of encephalitis cases [18-20] as well as been detected in the brain tissue of a deceased person with no abnormal brain pathology [21]. Hence, the potential role of TTV in encephalitis remains unclear and require further investigations.

Serum and nasal swabs from neurological horses

In the serum and nasal swabs from the horses displaying neurological signs very few eukrayotic viruses were identified and none that is believed to be the cause of the signs. Apart

from phages, only EHV-2 and papillomavirus sequences were identified. Although we do not believe that the papillomavirus is involved in the neurological clinical picture, as some of the identified viral reads in one of the horses displayed a low amino acid similarity to known equine papilloma viruses (eqPapV) we decided to attempt to genetically characterise the complete genome.

Novel equine papillomavirus
Using a combination of PCR, sanger sequencing and Nanopore sequencing the complete genome of a novel eqPapV was characterised. The genome is 7 763 nucleotides long and the expected ORFs could be identified. The most conserved part of the genome was that covering the E1 showing up to 68% nucleotide and 60% amino acid similarity to other eqPapPV sequences in GenBank. Other genes such as e.g., E6 are to divergent to render any similarity hits when investigating the nucleotide sequence with blastn and looking at the protein in level the identity to the closest eqPapV is just 42%. Based on these analysis we propose the this virus should probably be placed in its own genus.

Use of Minion for metagenomics
Nanopore sequencing has the advantage that it does not require any large expensive sequencing machine but the sequencer is small and portable, is connected to a computer through USB and, if access to internet is available, it can provide real-time sequence data. In order to investigate its suitability for viral metagenomics in clinics and less well-equipped labs we first used it on “easy” samples. Through these initial investigations we used viral isolates obtained through cell culture and showed that we could recover all expected viruses including their complete genomes at a high sequence depth. However, these types of samples have a low background compared to what is seen in clinical samples therefor, the next step was to use relevant material and, in this case, we used nasal swabs.

These nasal swabs were processed through our viral metagenomic pathway and sequenced using both Illumina (the HTS platform most often used in viral metagenomics) and Nanopore sequencing using MinION. In summary, all of the viruses identified through the Illumina sequencing was also identified through the Nanopore sequencing. EHV-2 was identified in both of the samples, while EHV-5 was only identified through the Nanopore sequencing. In addition, papillomavirus sequences were identified in both datasets.

Using EPI2ME real-time annotation of the produced Nanopore sequences provide a real-advantage over Illumina sequencing if considering a clinical setting where a rapid result is needed. The first herpesvirus sequences could be identified in the first hours of sequencing and requires no advanced bioinformatic skills from the personal. However, we do believe that in-house bioinformatics should also be used in parallel as this allows the set-up of workflows with less background (false positive viral hits) as well as the identification of more divergent viruses.

We could observe that the Nanopore sequencing depth (around 4 million reads/run) and the metagenomic workflow used in our studies had a high-sensitivity and could identify low level viruses in the samples (ct-values down to 30). In addition, complete viral genomes could be obtained.

Conclusions
Through these studies, we have identified and genetically characterized several DNA and RNA viruses infecting horses, both those exhibiting clinical signs (fever, fever and diarrhoea,

neurological symptoms) and, in some cases, apparently healthy individuals. Notably, many of these viruses have been detected and characterized for the first time in Scandinavia and, in most cases, Europe. Among our findings, a novel papillomavirus was identified and characterized in a horse with neurological signs, and eqTTV1 was observed for the first time in brain tissue of horse. While several viruses were detected in both diseased and control horses, a trend was observed indicating higher levels or a greater presence of some viruses in horses with fever. However, further studies are needed to establish their potential role in disease.

Certain viruses, such as eqTTVs, warrant further investigation, particularly regarding their association with neurological signs and other disease cases suspected to involve viral infections. Additionally, circoviruses, given their known significance in other species such as pigs, require focused studies to explore their potential impact on horses potentially initially focusing on circovirus infections in young horses.

Our research also demonstrated that Nanopore sequencing is a viable alternative to Illumina high-throughput sequencing. This technology enables rapid detection and annotation of equine viral metagenomes without the need for large, costly sequencing equipment.

Relevance for the practical horse sector incl. recommendations

Viral infections in horses can significantly impact their welfare, reduce performance, and result in economic losses for horse owners and the broader equine industry. To address these challenges, it is essential to identify the viruses circulating within the horse population, including both those causing apparent disease and those responsible for subclinical infections.

This project has significantly expanded our understanding of equine viruses circulating in Europe, identifying and characterizing several viruses for the first time in the region. Furthermore, we have developed multiple qPCR systems, which can be employed to screen large numbers of horses for these viruses. These tools enable researchers and veterinarians to validate the role of these specific viruses in various disease complexes and assess their impact on horse performance. Where relevant, these viruses can also be incorporated into routine diagnostic protocols, improving disease detection and management.

While viral metagenomics is increasingly applied in human disease outbreak investigations, its use in veterinary medicine, particularly for horses, remains limited. Nanopore sequencing, with its portability and real-time virus detection capabilities, represents a valuable innovation for the horse sector. Our findings demonstrate that viral metagenomics and Nanopore sequencing can simultaneously detect multiple viruses from the same sample with high sensitivity and would this be a valuable tool to include in emerging disease outbreaks.

References

1. Chapman GE, Baylis M, Archer D, Daly JM: **The challenges posed by equine arboviruses**. *Equine Vet J* 2018, **50**(4):436-445.
2. Gilkerson JR, Bailey KE, Diaz-Mendez A, Hartley CA: **Update on viral diseases of the equine respiratory tract**. *Vet Clin North Am Equine Pract* 2015, **31**(1):91-104.
3. Lecollinet S, Pronost S, Coulpier M, Beck C, Gonzalez G, Leblond A, Tritz P: **Viral Equine Encephalitis, a Growing Threat to the Horse Population in Europe?** *Viruses* 2019, **12**(1).
4. Blomstrom AL: **Viral metagenomics as an emerging and powerful tool in veterinary medicine**. *Vet Q* 2011, **31**(3):107-114.

5. Oladunni FS, Horohov DW, Chambers TM: **EHV-1: A Constant Threat to the Horse Industry**. *Front Microbiol* 2019, **10**:2668.
6. Patel JR, Heldens J: **Equine herpesviruses 1 (EHV-1) and 4 (EHV-4)--epidemiology, disease and immunoprophylaxis: a brief review**. *Vet J* 2005, **170**(1):14-23.
7. Li L, Giannitti F, Low J, Keyes C, Ullmann LS, Deng X, Aleman M, Pesavento PA, Pusterla N, Delwart E: **Exploring the virome of diseased horses**. *J Gen Virol* 2015, **96**(9):2721-2733.
8. Fisher M, Nebroski M, Davies J, Janzen E, Sullivan D, Lung O: **Discovery and comparative genomic analysis of a novel equine anellovirus, representing the first complete Mutorquevirus genome**. *Sci Rep* 2023, **13**(1):3703.
9. Hui A, Altan E, Slovis N, Fletcher C, Deng X, Delwart E: **Circovirus in Blood of a Febrile Horse with Hepatitis**. *Viruses* 2021, **13**(5).
10. Opriessnig T, Karuppannan AK, Castro A, Xiao CT: **Porcine circoviruses: current status, knowledge gaps and challenges**. *Virus Res* 2020, **286**:198044.
11. Altan E, Li Y, Sabino-Santos G, Jr., Sawaswong V, Barnum S, Pusterla N, Deng X, Delwart E: **Viruses in Horses with Neurologic and Respiratory Diseases**. *Viruses* 2019, **11**(10).
12. Pusterla N, James K, Barnum S, Delwart E: **Investigation of Three Newly Identified Equine Parvoviruses in Blood and Nasal Fluid Samples of Clinically Healthy Horses and Horses with Acute Onset of Respiratory Disease**. *Animals (Basel)* 2021, **11**(10).
13. Xu Y, Zhang X, Peng P, Liu Y, Yu M, Xie L: **First detection and genomic characterization of a new parvovirus in equines in China**. *Arch Virol* 2022, **167**(7):1607-1609.
14. Kapoor A, Simmonds P, Cullen JM, Scheel TK, Medina JL, Giannitti F, Nishiuchi E, Brock KV, Burbelo PD, Rice CM *et al*: **Identification of a pegivirus (GB virus-like virus) that infects horses**. *J Virol* 2013, **87**(12):7185-7190.
15. Tomlinson JE, Wolfisberg R, Fahnoe U, Sharma H, Renshaw RW, Nielsen L, Nishiuchi E, Holm C, Dubovi E, Rosenberg BR *et al*: **Equine pegiviruses cause persistent infection of bone marrow and are not associated with hepatitis**. *PLoS Pathog* 2020, **16**(7):e1008677.
16. Lu G, Fu C, Huang J, Xu H, Wu P, Ping X, Li S: **Molecular characterization of a genetically divergent equine pegivirus strain identified in China**. *Arch Virol* 2018, **163**(1):249-252.
17. Pusterla N, Vin R, Leutenegger C, Mittel LD, Divers TJ: **Equine coronavirus: An emerging enteric virus of adult horses**. *Equine Vet Educ* 2016, **28**(4):216-223.
18. Jurasz H, Bukowska-Osko I, Rydzanicz M, Popiel M, Dzieciatkowski T, Bakula-Grzadka K, Paciorek M, Makowiecki M, Horban A, Laskus T *et al*: **Torque teno virus (TTV) Infection in Patients with Encephalitis**. *Int J Mol Sci* 2024, **25**(20).
19. Kang YJ, Zhou MF, Huang W, Deng C, Yan G, Lu ZH: **Identification of a novel torque teno mini virus in cerebrospinal fluid from a child with encephalitis**. *Virol Sin* 2017, **32**(6):541-544.
20. Perlejewski K, Pawelczyk A, Bukowska-Osko I, Rydzanicz M, Dzieciatkowski T, Paciorek M, Makowiecki M, Caraballo Cortes K, Grochowska M, Radkowski M *et al*: **Search for Viral Infections in Cerebrospinal Fluid From Patients With Autoimmune Encephalitis**. *Open Forum Infect Dis* 2020, **7**(11):ofaa468.
21. Kraberger S, Mastroeni D, Delvaux E, Varsani A: **Genome Sequences of Novel Torque Teno Viruses Identified in Human Brain Tissue**. *Microbiol Resour Announc* 2020, **9**(37).

Part 3: Result dissemination

State all result dissemination from the financed project into the appropriate section, including information as indicated in each section. Additional rows can be added to the table.

Scientific publications, *published*	Anne-Lie Blomström, Annika Källse and Miia Riihimäki; Identification of viruses in horses with fever in Sweden; BMC Veterinary Research; https://bmcvetres.biomedcentral.com/articles/10.1186/s12917-025-04613-2
Scientific publications, *submitted*	Anne-Lie Blomström, Sanni Hansen, and Miia Riihimäki; Identification and whole genome characterization of a novel equine papillomavirus (submitted to Virus Genes March 2025)
Scientific publications, *manuscript*	Anne-Lie Blomström, Sanni Hansen, Marlene Pinto, Harindranath Choletti and Miia Riihimäki; Viral metagenomic analysis of horses with fever in Denmark
	Anne-Lie Blomström, Sanni Hansen, Marlene Pinto, Harindrantath Choletti and Miia Riihimäki;Viruses in horse with diarrhea
Conference publications/ presentations	
	Anne-Lie Blomström, Harindranath Choletti, Siv Hanche-Olsen, Ingunn Risnes Hellings, Sanni Hansen and Miia Riihimäki; 2024; Using viral metagenomics to identify viruses associated with fever in horses in Scandinavia; The EPIZONE 16th Annual Meeting; Uppsala, Sweden; 25-27 September; https://www.epizone-eu.net/en/home/am-2024.htm
	Anne-Lie Blomström, Harindranath Choletti, Siv Hanche-Olsen, Ingunn Risnes Hellings, Sanni Hansen and Miia Riihimäki; 2024; Using viral metagenomics to identify viruses associated with fever in horses in Scandinavia; The twelfth International Equine Infectious Disease Conference (IEIDCXII); Deauville, France; 30th September – 4th October; https://ieidc.org/
Other publications, *media etc.*	
	Project presentation in Swedish Veterinary Journal , June 2021, vol 73, pages 6-9
	Virus upptäckta för första gången hos svenska hästar – VeterinärMagazinet: https://www.veterinarmagazinet.se/2025/04/virus-upptackta-for-forsta-gangen-hos-svenska-hastar/

	Ny upptäckt – flera virus för första gången hos svenska hästar – Ridsport: https://www.tidningenridsport.se/ny-upptackt-flera-virus-for-forsta-gangen-hos-svenska-hastar/
	Nya virus upptäckta hos svenska hästar – kan ge bättre smittskydd: https://www.hippson.se/forskning/nya-virus-upptackta-hos-svenska-hastar-kan-ge-battre-smittskydd
Oral communication, *to horse sector, students etc.*	From negative to positive-a new diagnostic to determine the aetiology of unknown viral infections; 9th June 2021; University Animal hospital, UDS, Equine clinic
	Using viral metagenomics to identify viruses associated with fever in horses in Scandinavia; 27th November 2024; joint meeting for researchers and employees at SLU and the Swedish Veterinary Agency (SVA) working with different aspects of horse welfare.
	Viral infectious disease of horses, yearly (2022-2024) within the course Allmänn sjukdomslära, veterinary students year 2
	Using viral metagenomics to identify; 25th October 2024; The section of serology and virology at the Swedish Veterinary Agency (SVA); From negative to positive - New diagnostic to determine the aetiology of unknown viral infections in horses
	Presentation for Equine Science students at SLU (Wången, Flyinge, Strömsholm), yearly (2022-2024)
	Project presentation Hippocampus 24/9, 2024 digital lecture, HästSverige
	Provect presentation, Swedish Shetland pony society: SLU Uppsala November 9. 2024
Student theses	
Other	