University of Nottingham
UK | CHINA | MALAYSIA

Assessing the risk of wild rodents as a potential source of emerging virus infections

Joseph Graham Chappell MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2019

i. Abstract

Rodent-bourne viral zoonoses are a clinically significant group of viruses, however, their prevalence in the United Kingdom is not well-understood. Of particular importance are the *Hantaviridae* and the *Picornaviridae*.

This thesis describes the screening of rodent tissue collected from four sites in the United Kingdom, as well as sites in Poland and Egypt, for Orthohantaviruses, picornaviruses and orthobornaviruses. Two strains of Tatenale orthohantavirus were detected in field voles (*M agrestis*) captured at two sites in the United Kingdom. One of these viruses represents a novel strain and was detected in a location with no previous detection of Orthohantaviruses. Additionally, four species of picornaviruses were detected in *Mus musculus, Rattus norvegicus* and *Myodes glareolus*. Additional high-throughput sequencing recovered the first complete coding genome of Tatenale virus, which allowed characterisation of the genome and confirmed its status as a novel species. The serological screening of captive non-human primates for evidence of orthohantavirus infection showed evidence of both acute and past infections, which may have implications for the healthcare of these animals, and the conservation efforts of the species. Similarly, there was serological evidence of acute infection in vulnerable human cohorts, which has implications for the healthcare of these individuals and the local population. Finally, high-throughput sequencing of samples from humans with neurological disease of unknown aetiology, recovered evidence of human pegivirus infection

in patients with neurological disease, a controversial and emerging topic in clinical virology currently.

ii. Publications

Tsoleridis T, **Chappell JG**, Onianwa O, Marston DA, Fooks AR, Monchatre-Leroy E, Umhang G, Müller MA, Drexler JF, Drosten C, Tarlinton RE, McClure CP, Holmes EC And Ball JK. 2019. Shared Common Ancestry of Rodent Alphacoronaviruses Sampled Globally. Viruses. 11(2).

Chappell JG, Byaruhanga T, Tsoleridis T, Ball JK, McClure CP. 2019. Identification of Infectious Agents in High-Throughput Sequencing Datasets is Easily Achievable Using Free, Cloud-Based Bioinformatics Platforms. Journal of Clinical Microbiology. (In Press).

Chappell JG, Tsoleridis T, Onianwa O, Drake G, Ashpole, I, Dobbs, P, Edema W, Kumi-Ansah F. Bennett M, Tarlinton RE, Ball JK and McClure CP. 2020. Retrieval of the Complete Coding Sequence of the UK-Endemic Tatenale Orthohantavirus Reveals Extensive Strain Variation and Supports Its Classification as a Novel Species. Viruses, 12, 454.

iii. Oral Presentations

Investigation of hantavirus prevalence in wild rodent populations in the United Kingdom. 6th European Congress of Virology, Hamburg. 2016.

High-Throughput Sequencing of Patients with Symptoms of Unknown Etiology. Microbiology Society of Annual Conference, Belfast. 2019.

Investigation and Genome Characterisation of Tatenale Orthohantavirus. Microbiology Society of Annual Conference, Belfast. 2019.

iv. Poster Presentations

Investigation of hantavirus prevalence in British rodent populations. Microbiology Society of Annual Conference, Edinburgh. 2017.

Investigation of zoonotic picornavirus prevalence in Polish vole populations. Microbiology Society of Annual Conference, Birmingham. 2018

v. Acknowledgements

Firstly, I would like to thank my supervisors, Professor Jonathan Ball, Dr Rachael Tarlinton and Dr Patrick McClure for their exceptional support and guidance throughout my PhD.

Many thanks to Dr Theocharis Tsoleridis, Dr Richard Urbanowicz, Dr Barnabas King and Dr Alex Tarr for providing technical assistance and guidance.

I would also like to thank Professor Jerzy Behnke, Professor Malcolm Bennett, Phillipa Dobbs, Gabby Drake and Ian Ashpole for providing rodent samples, and Professor Will Irving and Dr Gemma Clark for providing clinical human samples. I am also very grateful to Professor Albert Rizvanov for providing the Orthohantavirus ELISA kits.

I'd like to thank my friends and family for their continued support throughout my entire academic career, for which I'm incredibly grateful. I am also incredibly grateful to my dad and late mum for their support and guidance over the years, giving me the confidence to pursue a career in science.

Finally, I would like to thank my partner Hollie, whose unwavering support has kept me going over the course of this PhD.

vi. Abbreviations

AA	Amino Acid
ADLV	Adler virus
ANDV	Andes virus
BoDV-1	Borna disease virus 1
CaV	Cardiovirus
CDS	Coding sequence
CSF	Cerebrospinal fluid
DOBV	Dobrava-Belgrade Virus
EBL	Endogenous bornavirus-like elements
EDTA	Ethylenediaminetetraacetic-acid preserved blood
EID	Emerging infectious diseases
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HCPS	Hantavirus cardiopulmonary syndrome
HF	Hantavirus Fever
HFRS	Haemorrhagic fever with renal syndrome
HOKV	Hokkaido virus
HPS	Hantavirus pulmonary syndrome
HTNV	Hantaan virus
HTS	High-throughput sequencing
HuV	Hunnivirus
ICTV	International Committee on Taxonomy of Viruses
IVDA	Intra-venous drug abuser
KHAV	Khabarovsk Virus
L	'Large' segment of orthohantaviruses
LV	Ljungan virus
M	'Medium' segment of orthohantaviruses

NCR	Noncoding region
NE	Nephropathia epidemica
Nr	Non-redundant protein
Nt	Nucleotide
ORF	Open reading frame
PCR	Polymerase chain reaction
PUUV	Puumala virus
qPCR	Quantitative polymerase chain reaction
RV	Rosavirus
RT-PCR	Reverse-transcription polymerase chain reaction
S	'Small' segment of orthohantaviruses
SEOV	Seoul virus
SNV	Sin Nombre virus
TATV	Tatenale virus
TULV	Tula virus
UK	United Kingdom
VSBV-1	Variegated squirrel bornavirus 1
YUJV	Yuanjiang virus

vii. Table of Contents

viii. List of Figures

ix. List of Tables

1. Introduction

1.1. Zoonoses

1.1.1. Terminology

Zoonotic infections are defined as 'any disease or infection that is naturally transmissible from vertebrate animals to humans [1]. They are transmissible either through direct (bites, handling infected animals, etc.) or indirect (inhalation of infected fluids, contaminated water, etc.) exposure. Whilst much of the current concern focuses on animal to human transmission, human-animal transmission (zooanthroponosis, reverse-zoonoses) and animal-animal transmission also pose a significant threat to many protected animal species [2]. It has been estimated that of the known human pathogens, around 58% are zoonotic, including viral, bacterial, fungal, protozoal and helminth species [3].

Emerging infectious diseases (EID) are those which have been recently introduced into human or animal populations, whilst re-emerging infections are those which have previously been described in humans or animals but are rapidly increasing in incidence or host range [4]. Around 73% of emerging and re-emerging pathogens in humans are considered to be zoonotic in origin [5]; the primary host species for these pathogens encompass both domesticated (cattle, pigs, poultry etc.) as well as wild animals (rodents, bats, primates etc.). It is the latter that represents the most significant (and increasing) risk in the context of EID events [6]; with the majority of 'spillover events' - the point in which a pathogens crosses into a novel host -,

occurring during the interaction between human and animals. Increasing contact between humans and wildlife, through habitat encroachment and the increasing exploitation of wildlife as exotic pets, for example, is allowing for transmission events to occur with greater regularity [7]

1.1.2. Reservoirs of Zoonoses

A reservoir host is defined as an organism in which a pathogen can be permanently maintained, from which the pathogen is transmitted into a defined population; humans in the case of a zoonotic pathogen [8].

Zoonotic pathogens can be found in a broad range of animal species, including birds [9], amphibians [10] and reptiles [11], the majority, however, are found in mammalian species. The class *Mammalia* comprises 84% of known reservoir species [12], of these, rodents (Order: Rodentia) are amongst the largest source of potential zoonoses, second only to bats (Order *Chiroptera*). At least 217 species of rodents have been identified as a reservoir for 66 zoonotic pathogenic species. Of these rodents, 79 were considered 'hyper-reservoirs', hosting between 2 and 11 pathogens [13]. Despite numerous studies, it is still unclear why rodents harbour a disproportionately high number of infectious agents. As rodents are commonly used model organisms, sampling bias could likely a factor. Another possibility is the sheer abundance of rodents compared to other mammalian species, with rodents accounting for around 40% of all extant mammalian species [14], especially considering the global distribution of some species, such as the brown rat. There is a strong

positive correlation between reservoir status and the total number of zoonoses, with species that have shorter gestational periods and more frequent litters [13]; reproductive traits which many species of rodents possess. It has also been suggested that since rodents are more closely related to humans than reservoir species of other taxa (bats, for example), that features such cell receptors between the two should be more conserved, assuming the evolutionary patterns of these receptors follow the patterns of the whole genome evolution; meaning that any infectious agents would be more able to cross the species barrier [15].

1.1.3. Viral Zoonoses

Viral pathogens account for approximately 44% of novel emerging pathogens [16]. These infections can lead to sporadic, isolated infections in humans as a result of multiple spillover events, with a high mortality rate but limited transmission; Hendra virus [17] and Nipah virus [18], for example. Other viral zoonoses, however, have led to widespread epidemics and pandemics, leading to substantial deaths, social disruption and extensive economic burden; Human immunodeficiency virus (HIV-1), originating in chimpanzees, currently infects 36.9 million people globally and has led to approximately 34 million deaths [19]. The recent West-African Ebola virus outbreak is another example of the effects that viral zoonoses can have, with around 28000 infected individuals, 11310 of whom died [20], costing approximately $6 billion [21].

Predicting which viral species are likeliest to cross into human populations is difficult and unreliable [22]. Despite the general unpredictability of zoonotic capability, there are certain traits which are essential in assessing this capability; certain virus taxa may be predisposed to zoonotic transmission, the host range of the virus may influence potential transmission events, the transmission route of a virus may influence the likelihood of spillover, the evolvability of a virus will influence the likelihood of host adaptation [23]. A three-stage model of zoonotic virus emergence has been proposed by Morse *et al* [24]. The first stage is pre-emergence, where the dynamics of transmission change (due to ecological or socio-economic factors) and result in virus expansion within its reservoir host and/or increase in host range. The second stage is a localised emergence of the virus into a human population with little to no human-human transmission. The third and final stage is full pandemic emergence, where the virus has sustained human-human transmission and has spread internationally. Wolfe *et al* [25] created a more comprehensive five-stage model of the stages of zoonotic microbe emergence. His proposed first stage is the microbe in its natural reservoir with no human infection; Stage two occurs with transmission from the animal reservoir into humans but without any human-human transmission. Stage three occurs when there is a limited human-human transmission, which results in small, localised outbreaks of human infection. Stage four happens when infection in humans can be sustained for long periods without any interaction with the animal reservoir host; and stage five follows when humans become the

primary host of the virus, with no animal intermediary. Most species of viruses remain in the first stage of these models and very rarely progress to stage three or further.

Most viruses that do progress this far are those with RNA genomes. The increased cross-species transmissibility and subsequently increased representation of RNA viruses in zoonotic spillover events are thought to be largely a result of the rapid evolutionary potential of RNA viruses. The increased nucleotide substitution rate [26], lack of proof-reading enzymes and rapid replication cycles [27] means that mutations can accumulate quickly within the genome. This results in a greater genomic instability than DNA based viruses [28], and a significantly increased ability to adapt to novel hosts. Examples of zoonotic viral outbreaks between 2003 and 2013 are shown in Figure 1. There are several prominent examples of RNA virus families which contain zoonotic viruses. The *Orthomyxoviridae*, for example, contains the species *Influenza virus A,* which is transmitted from birds and several mammals and is responsible for causing respiratory disease. The *Rhabdoviridae* is another notable family; the often-fatal neurological disease 'Rabies' is caused by the species *Rabies lyssavirus* which can be transmitted from bats and dogs. Three families in particular should be considered for further investigation; *Hantaviridae, Bornaviridae* and *Picornaviridae*. *Hantaviridae* contains several known zoonoses that can cause viral haemorrhagic fever; though they are generally understudied in the United Kingdom. Viruses in the *Bornaviridae* family have long been suspected to be zoonotic, though it

has only been in recent years that this theory has been proven correct. Finally, most species of the family *Picornaviridae*, whilst clinically relevant, are not considered to be zoonotic in nature. The genus *parechovirus* may be an exception to this, with evidence suggesting some species may be a rodent-bourne zoonoses, warranting further investigation into the genus.

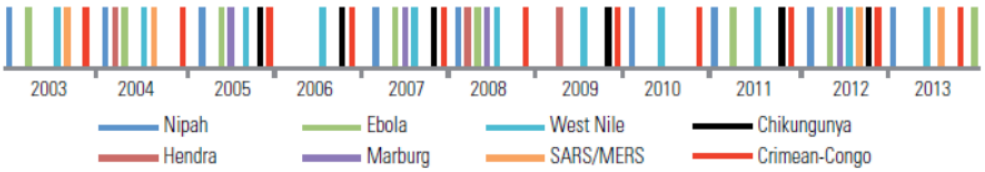

Figure 1 Summary of viral outbreaks with a zoonotic origin. Adapted from Wang & Crameri (2014) [346]

1.2. *Hantaviridae*

1.2.1. Taxonomy

The *Bunyavirales* are a particularly large viral order, consisting of 12 families. The family *Hantaviridae* contains four subfamilies based on host class; *Acantavirinae* and *Agantavirinae* from fish, *Repantavirinae* from reptiles, and *Mammantavirinae* from mammals. Except for *Mammantavirinae*, these sub-families contain only a single genus and species in the case of *Agantavirinae* and *Reptavirinae*, and a single genus with three species in *Acantavirinae*; though this limited taxonomic diversity is likely a result of limited sampling of potential fish and reptile reservoir species. *Mammantavirinae* contains four genera, *Loanvirus, Mobatvirus, Thottimvirus* and *Orthohantavirus*. As with the sub-families, these genera contain viruses associated with different host orders; *Loanvirus* in *Chiroptera* (Bats), *Mobatvirus* in *Chiroptera* and

Soricomorpha (Moles), *Thottimvirus* in *Soricomorpha* (Shrews) and *Orthohantavirus* in *Rodentia*. *Orthohantavirus* contains 36 recognized species, of which Hantaan virus (HTNV), is considered to be the type species by the International Committee on Taxonomy of Viruses (ICTV) (Figure 2).

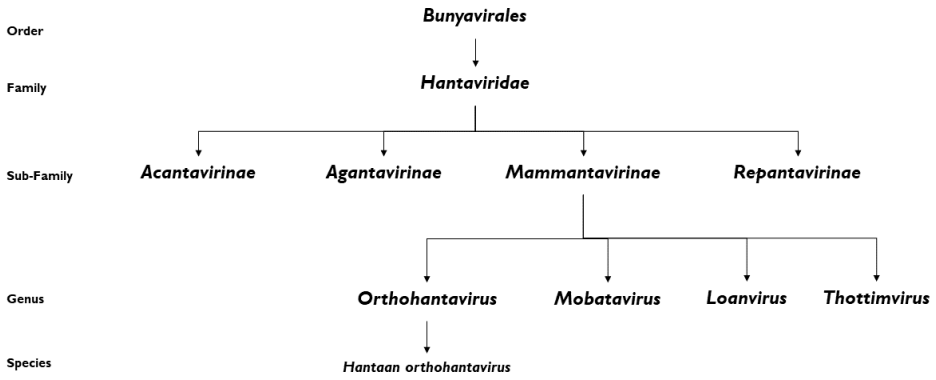

Figure 2 The Taxonomy of the Orthohantavirus genus, Hantaan orthohantavirus is included as the type species.

1.2.2. Genome

The *Hantaviridae* consist of a negative sensed single-stranded RNA genome with a length of approximately 12.5 kilobases, depending on species. These genomes consist of 3 segments, with each of these containing an open reading frame (ORF) flanked by non-coding regions (NCR) at both the 3' and 5' ends. These segments are the S (Small), encoding a nucleocapsid protein (N), M (Medium), encoding

two glycoproteins (Gn & Gc) and L, (Long) segment encoding an RNA-dependent RNA polymerase [29] (Figure 3).

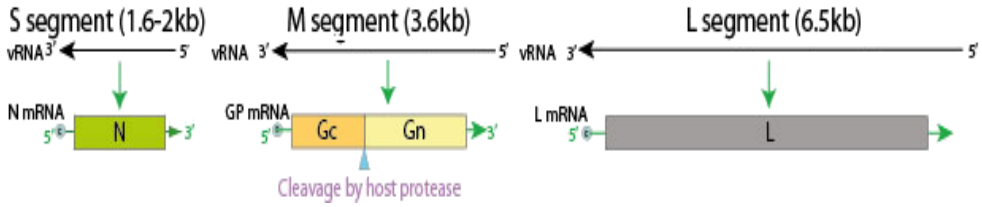

Figure 3 Structure of the Hantavirus genome. The S segment encodes a nucleocapsid (N), the M segment encodes two glycoproteins (Gc and Gn) and the L segment encodes an RNA polymerase (L) Adapted from [347]

1.2.3. Virions

Like other *Bunyavirales*, orthohantavirus virions are spherical particles, approximately 100nm in diameter and enveloped in a 5nm bilayered membrane (Figure 4). They can be distinguished from other Bunyaviridae under an electron microscope by their distinctive grid-like structure within the interior of the particle [30].

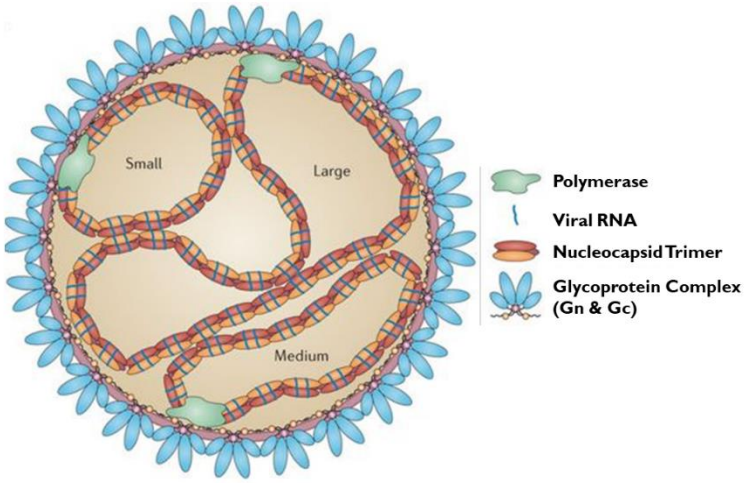

Figure 4 A Schematic of a typical orthohantavirus virion, showing each of the four encoded proteins. Adapted from Vaheri *et al* (2013).

1.2.4. Virus Entry and Replication

Integrins are thought to be the primary cellular receptors for infection in humans, Integrin β_3, in particular, is associated with pathogenic orthohantavirus species such as Seoul orthohantavirus (SEOV), Puumala orthohantavirus (PUUV), HTNV [31] and Sin Nombre virus (SNV) [32]. Integrin β_2 [33] and protocadherin-1 (PCDH1) [34] have also shown to be likely receptors for entry. The receptors involved in cellular entry in reservoir species is poorly understood, however. *In vitro* evidence has shown that infectivity in cells derived from reservoir species is maintained in the absence of Integrin β_3, indicating that alternative mechanisms of entry exist in rodents [35].

Following attachment to the receptor, the virions of most species enter the cell through clathrin-coated vesicles. Once in the cell, the virion is uncoated from the clathrin and fuses into an endosome, mediated by the PH-induced conformational changes. Viral ribonucleoproteins are released from the endosome and transcribed, they are then transported to the Golgi, where replication occurs before budding out of the Golgi and being transported to the plasma membrane for release [36] (Figure 5).

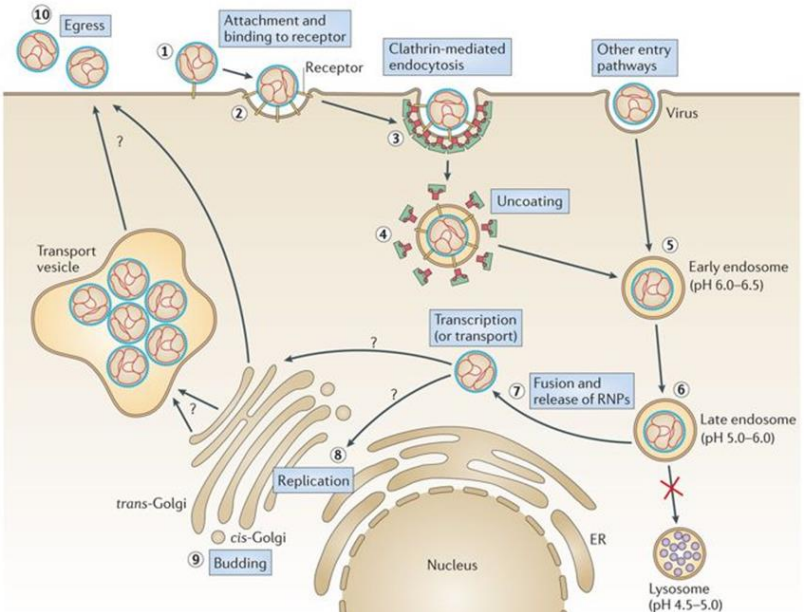

Figure 5 Diagram of the replication pathway of orthohantavirus. RNP, Ribonucleoprotein. Adapted from Vaheri et al (2013).

1.2.5. Reservoirs

Hantaviridae host species are primarily rodent species (*Rodentia*), and until recently, these were considered to be the only reservoirs of these viruses. However, hantavirus reservoir species have been identified in other taxons including bats (Order *Chiroptera*) [37][38], shrews [39][40] and moles [41] (Order *Eulipotyphyla*). Interestingly, these newly identified reservoirs belong to the same taxonomic superorder, *Laurasiatheria* (which doesn't include rodents), which has been identified as the likeliest reservoirs of the ancestral hantavirus [42]. As of the latest ICTV recognised taxonomy release (2014), there are 24 species of hantaviruses [43], though it is believed that the number of species likely exceeds 50. Within the majority of reservoir species, hantaviruses infection results in achronic but asymptomatic infection, though it may result in a minor drop in reproductive fitness [44]. It is thought that this pattern of infection is due to the immune response

of these reservoir species, in particular the regulatory T-cell response. Experimental inactivation of regulatory T-cells in SEOV infected laboratory rats reduced the amount of detectable viral RNA, suggesting that regulatory T-cells can mediate the persistence of SEOV [45]. Similarly, the expression of tumour necrosis factor-alpha (TNF-α) has been shown to influence Puumala virus persistence [46].

One of the core dogmas in the study of hantaviruses is that each hantavirus species is associated with a single reservoir species. This is supported by phylogenetic evidence that hantaviruses cluster together based on their reservoir species (Figure 6) with viruses diverging similarly to the reservoir, displaying a pattern of co-speciation between each virus and their reservoir. One theory suggests that an ancestral hantavirus-infected a common ancestor of the various reservoir species and virus and reservoir species co-evolved. [47]. Additionally, experimental evidence shows that when a hantavirus reservoir species is infected with an atypical virus, the host can mount an effective immune response, entirely clearing the infection [48], meaning they are unable to persist in species other than the reservoir.

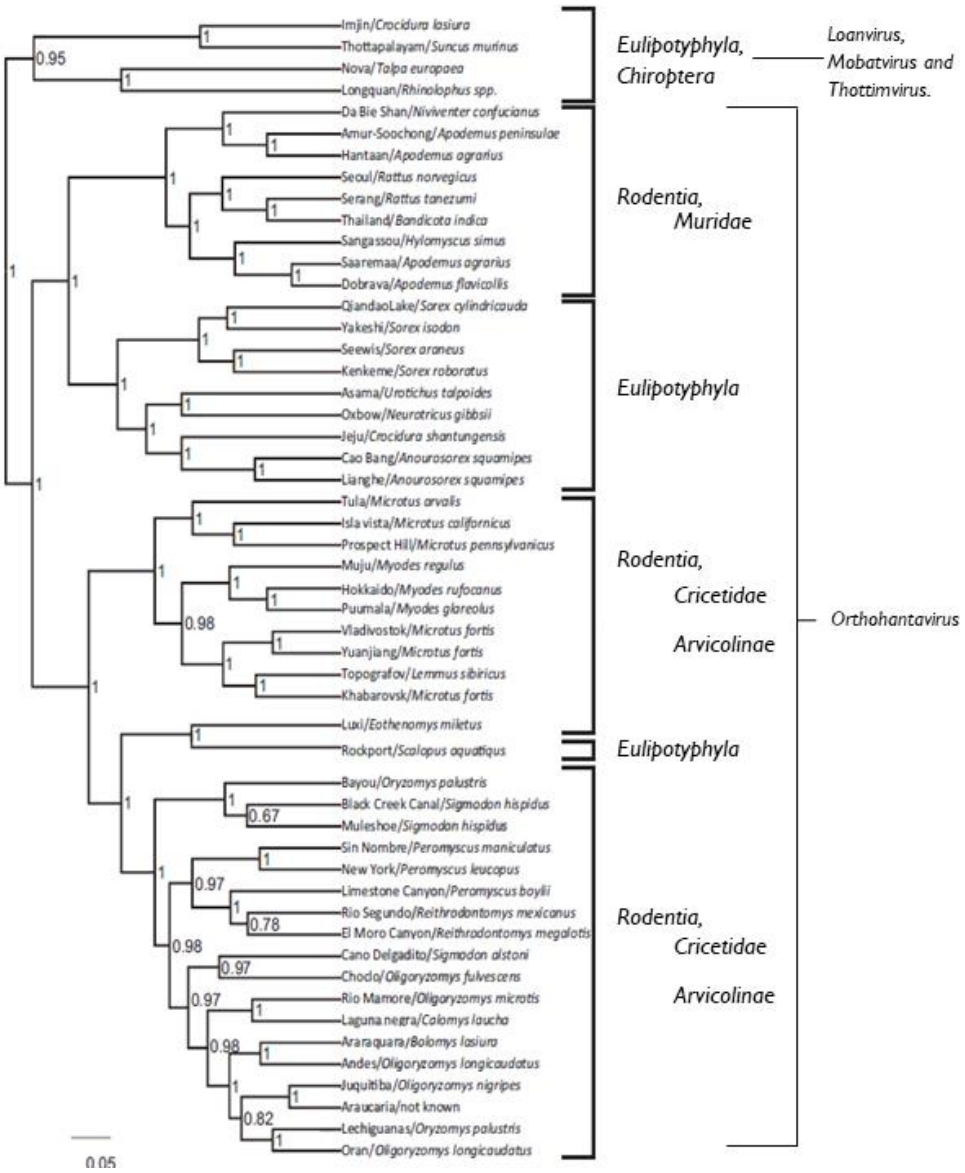

Figure 6 Phylogenetic tree of the S-segments of representative Mammantatavirinae species, with their primary reservoir species. Adapted from [46]

Studies into this apparent co-speciation have brought the validity of this theory into question; Ramsden *et al.* [49] considered that the divergence timescales of host species and virus species were far too dissimilar and that the apparent similarity in phylogenies as a result of recent preferential host-switching. Cross-species transmission has been demonstrated for several insectivorous [50] and rodent [51]

associated Orthohantaviruses into species other than the reservoir. Which, with evidence that Puumala (PUUV) virus lineages and their bank vole (*Myodes glareolus*) hosts [52] have not co-diverged, raises the possibility that hantavirus transmission between sympatrically occurring, related hosts might be more common than previously thought, and overlooked until recently [53].

As these viruses are mostly limited to a single reservoir species, they are effectively limited to the geographic range of each host, which combined with the lack of human to human transmission (except for Andes virus, ANDV [54]), means that human hantavirus infections are also limited to the range of the typical reservoir host locations.

1.2.6. Transmission

Hantaviridae is unique amongst the *Bunyavirales* in that they are not transmitted through an intermediary arthropod vector. They are instead transmitted through contact with bodily fluids; commonly the faeces, urine and saliva, of an infected host species [55].

Transmission from a reservoir species into humans is thought to occur indirectly, through the inhalation of dried and aerosolised faeces and urine. In favourable conditions, such as a cool, stable temperature, shed Hantaviruses can survive outside of a host for a considerable amount of time. Wild-type Puumala virus, for example, has been shown to maintain infectivity for 12-15 days at room temperature, whilst cultured Puumala and Tula viruses can remain infectious for 18 days at 4°C[56]. This ability for environmental

persistence is thought to be critical for viral transmission dynamics, where many of the reservoir species undergo strong density fluctuations. During periods of low population densities, the frequency of direct contact between individual animals is lower, reducing the opportunities for transmission, which in turn can lead to localised extinction of the virus [56]

This environmental persistence results in infectious virions contaminating areas and objects that are frequently interacted with by humans. For example, contaminated dust disturbed through sweeping or simply walking [57], can cause the virions to become aerosolised and inhaled. Alternatively, handling contaminated firewood, animal bedding or food can also disturb virions, leading to infection. Subsequently, people working in professions with much more frequent contact with rodent-infested areas are more likely to become infected; profession such as cleaning, construction [58] farming or forestry [59] have much higher exposure and incidences of disease.

In addition to indirect exposure to the excreted virus, infection through contact with saliva - for example through rodent bites, grooming and sharing food - is also a known route for intraspecies transmission, [60]. It is well understood that certain new-world hantavirus species, such as Sin Nombre and Andes, oral transmission through saliva is thought to be the primary route of intraspecies transmission. Viral RNA has been detected in saliva and lung, but not in urine of the rodent hosts [61], which suggests that viral shedding via urine is unlikely and therefore unlikely to be a primary route of

transmission between rodents. Oral transmission has also been implicated in transmission into humans [62], though this route appears to be rare.

1.2.7. Human Disease

Infection in humans can lead to one of two illnesses depending on the virus; infection with Eurasian species can result in hantavirus haemorrhagic fever with renal syndrome (HFRS), whilst those in the Americas cause hantavirus pulmonary syndrome (HPS), also referred to as hantavirus cardiopulmonary syndrome (HCPS) [63]

HFRS is the more prevalent and milder of the two forms of illnesses. Following infection, an incubation period of 10 days to 6 weeks is followed by a febrile phase with nonspecific symptoms, including myalgia, headache, abdominal pain, inflammation and rashes. Following this, patients can enter a hypotensive phase, presenting with vascular leakage, associated with thrombocytopenia and shock. The following oliguric phase is characterised by renal failure, proteinuria and occasionally hypertension and pulmonary oedema. Whilst uncommon, HFRS can lead to death, particularly in elderly patients or those with pre-existing conditions. The severity of the illness can also vary depending on the particular infecting species. Dobrava-Belgrade virus (DOBV) and Hantaan virus (HTNV) have mortality rates of 12% [64] 5-15% [65] respectively, whilst Puumala virus (PUUV) is associated with a mild form of the disease, known as nephropathia epidemica (NE), and is only fatal in 0.4% of cases [66]. Most reported HFRS cases are in China; between 1950 and 2007

there were a total of 1,557,622 diagnoses and 46,427 deaths (3%), peaking in 1982 with 115,804 cases and gradually decreasing to 37,814 in 2000 and 11,248 in 2007. The causative agents of these infections are likely due to Seoul virus (SEOV) and HTNV [67]. In Europe, Russia accounts for most HFRS diagnoses, almost all of which are the milder NE illness, averaging 7,476 cases annually. NE is also prevalent in other northern European countries, 2,981 diagnosed annually in Scandinavia, 526 in Germany, 270 in Belgium and 110 in France [68]. Other HFRS causing hantaviruses circulating within Europe include DOBV, which is found almost exclusively in the Balkan peninsula, central Europe and the Black Sea regions. Whilst it is much more pathogenic than the other hantaviruses, cases are much more sporadic [69].

HPS is a much more severe form of Hantavirus infection, though less prevalent than HFRS. It was first identified following a succession of hospitalisations in the southwestern United States in 1993 and linked to Sin Nombre virus (SNV). From its initial detection, there have been 659 reported cases with a mortality rate of 36% in clinically diagnosed cases[70], caused primarily by SNV. Like HFRS, HPS initially presents with non-specific 'flu-like' symptoms lasting for around five days; this leads into a severe phase with the onset of a rapidly developing diffuse, non-cardiogenic pulmonary oedema, hypotension and shock [71]. Cardiogenic shock also often develops in fatal cases. In South and Central America, HPS is more prevalent than in North America. By 2013, approximately 1600 cases had been

diagnosed in Brazil, 1300 in Argentina [72] and a smaller number of cases reported in many other countries. Andes virus (ANDV) is one of the more prevalent South American hantavirus species, circulating within the *Oligoryzomys* genus of rodents (Pygmy Rice Rats) and is the only known hantavirus to be transmitted between humans [54].

There have been hypotheses that HFRS and HPS are in fact only variations of the same disease. Three Swedish cases of severe PUUV infection were diagnosed, with all three presenting with pulmonary, rather than renal disease and satisfying the criteria for HPS [73]. Many other reported diagnoses of HFRS in Europe have had pulmonary involvement [74–76], though this had often been dismissed as a complication of infection rather than a characteristic symptom. Clement *et al* [77] have suggested redefining HFRS and HPS (and HCPS) into a single all-encompassing disease, Hantavirus Disease.

1.3. *Bornaviridae*

1.3.1. Taxonomy

The Bornaviridae is a small viral family, consisting of three genera, *Carbovirus, Cultervirus* and *orthobornavirus*. *Carbovirus* is associated with snakes, whilst *Cultervirus* is associated with fish. *orthobornavirus* contains 8 ICTV-recognised species and is associated with a range of host species. Morphologically, they form enveloped, spherical virions between 40 to 190 nm in diameter. They have a negative-sense, single-stranded RNA genome, around 8.9kb in length (the smallest amongst the *Mononegalovirales* order) that encodes 6 proteins; a nucleoprotein (P40, N), a phosphoprotein (p24, P), protein p10 (X), a

matrix protein (p16, M), a glycoprotein (p57, G) and an RNA polymerase (L) (Figure **7**)

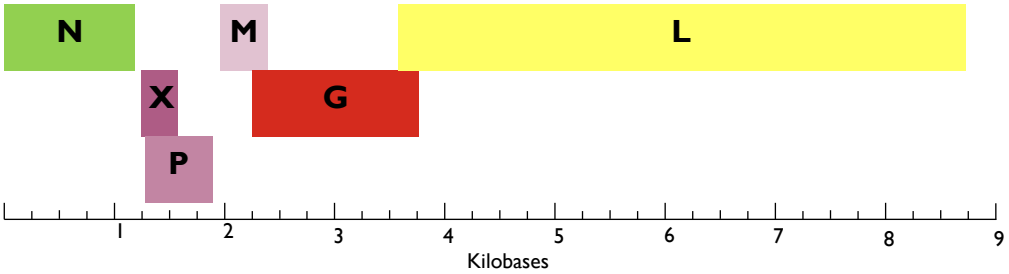

Figure 7 Genome Structure of Mammalian 2 bornavirus (VSBV), measured in kilobases. N, Nucleoprotein; X, P10; P, Phosphoprotein; M, Matrix; G, Glycoprotein; L, RNA polymerase.

The six major open reading frames are expressed from three transcription units. The first unit is monocistronic, encoding the N protein. The second transcription unit is bicistronic and encodes both the P and X proteins. The final unit is tricistronic and encodes the G, M and L proteins [78].

The *Bornaviridae* genome is highly conserved, with a similarity often greater than 95% between isolates. It is thought that this is a result of a strict selection pressure due to a complex chain of transmission events that includes multiple host species, this would result in multiple hosts exerting constraints, resulting in an increased likelihood that mutations will result in a loss of fitness [79].

1.3.2. Reservoir Species

Reservoir species associated with orthobornaviruses include birds (Passeriform 1 orthobornavirus, Passeriform 2 orthobornavirus, Psittaciform 1 orthobornavirus, Psittaciform 2 orthobornavirus & Waterbird 1 orthobornavirus), snakes (Elapid 1 orthobornavirus) and

mammals (Mammalian 1 orthobornavirus & Mammalian 2 orthobornavirus). These viruses do not always infect their natural hosts asymptomatically. In some species of birds, for example, Psittaciform 1 orthobornavirus is linked to proventricular dilation disease (PDD), which causes an often fatal nonsuppurative ganglioneuritis of the gastrointestinal tract [80][81].

1.3.2.1. Mammalian 1 orthobornavirus

Mammalian 1 orthobornavirus (also referred to as Borna disease virus, BoDV-1) has been found to infect a wide range of mammalian species, including rodents [82], shrews [83], horses, dogs, pigs, sheep, goats, cows, rabbits [84] and cats [85]. Classical disease (Borna disease) has been described primarily in sheep and horses in southern Germany, typically manifesting as a neurological disorder, with non-purulent meningoencephalitis accompanied by neurobehavioral alterations and paralysis. The mortality rate of the disease is estimated at between 60-80% within 5 weeks of becoming symptomatic. Spontaneous recovery has been observed and is associated with permanently altered behaviour and occasional recurrence of severe encephalitis [79].

The definitive reservoir species for BoDV-1 has yet to be identified, though it is widely believed to be a small mammal. The reservoir is the bicoloured shrew (*Crocidura leucodon*) [80,81][86][87]; both BoDV-1-specific antibodies and nucleic acid having being detected in screened animals. However, it is unknown if this represents the sole reservoir species for BoDV-1. Whilst the bicoloured shrew is

primarily distributed across Eastern, Southern and Central Europe and Western Asia, evidence of BoDV-1 infection has been found in regions where this shrew is not present, for example, Iceland [88], Japan [89] and Scandinavia [90]. Though other potential reservoirs are unknown, there is a significant volume of evidence that points to a small mammalian host of BoDV-1, for example, rural and free-roaming cats and those that are known to hunt rodents are at a sevenfold increased risk of BoDV-1 infection [91] which would suggest that they are likely contracting BODV-1 from their prey. Additionally, focal epidemics of Classical Borna disease in Southern Germany in livestock occur independently of both location and the affected species. These occur at intervals of between two and five years, which is suspected to be correlated to the fluctuation in populations of rodent species [92]. Immunocompetent rats can be experimentally infected intranasally with BoDV-1 [93]. Experimental infection typically results in neurological disease and had a fatality rate of 90% one week after onset of symptoms

BoDV-1 is also capable of horizontal transmission in brown rats (Via urine) [94] and vertical transmission in house mice [95]. In addition, offspring of wild-caught, Finnish bank voles (*M glareolus*) have also been shown to be permissive for BoDV-1 and are capable of shedding the virus in faeces and urine [96].

An alternative explanation for the discrepancy between the distribution of seropositive animals and the suspected shrew reservoir is several serologically cross-reactive viruses. Classical Borna disease is

geographically restricted, whilst other asymptomatic, seropositive animals are not, supporting the hypothesis of multiple viruses, with these hypothetical species resulting in a different or no clinical outcome.

1.3.2.2. Mammalian 2 orthobornavirus

Mammalian 2 orthobornavirus, also referred to as Variegated squirrel Bornavirus (VSBV-1), was identified in 2015 following a series of lethal meningoencephalitis cases in squirrel breeders. Screening of the squirrel breeding colonies detected VSBV-1 in variegated squirrels (*Sciurus variegatoides*) [97]. Following the initial discovery, screening of captive squirrels detected VSBV-1 in three further species of squirrels, Prevost's squirrels (*Callosciurus prevostii*), Finlayson's squirrels (*C. finlaysonii*) and Asiatic striped squirrels (*Tamiops swinhoei*) located in Germany, Croatia [98] and the Netherlands [99].

Importantly, it remains to be seen whether these squirrel species are the definitive VSBV reservoir. Investigation of VSBV-1 positive variegated and Prevost's squirrels shows a similar pattern of virus distribution in tissue as with BoDV-1 in bicoloured white-toothed shrews [100], which with the lack of any observable disease would suggest that the squirrels may be the true reservoir for VSBV-1. Variegated squirrels are native to Central America, whilst Prevost's squirrel, Finlayson's squirrel and the Asiatic striped squirrel are native to South-East Asia; so it is unlikely that VSBV-1 had been independently imported with multiple squirrel species. A more likely explanation is that VSBV-1 was imported in a single species and

transmitted into others during captivity. Alternatively, VSBV-1 may have been transmitted into captive squirrel populations from another source, such as a rodent or shrew. Given that the only other known mammalian orthobornavirus, BoDV-1, is endemic in shrews within Southern Germany, and most VSBV-1 detections have also been in Germany, it is possible that VSBV-1 also has a similar, restricted geographic origin.

1.3.3. Human orthobornavirus Infection

1.3.3.1. Mammalian 1 orthobornavirus

Until recently, human infection with BODV-1 was a highly controversial topic. Many studies have reported increased BODV-1 seroprevalence (Between 0 and 48% [101]) in patients with psychiatric disorders, such as bipolar disorder [102], major depressive disorder [103], autism [104], schizophrenia, multiple sclerosis and Parkinson's disease [105]. However, the validity of these studies has been questioned due to the reliability of certain assays, specifically ELISA-based sandwich assays [106], in which the potential presence of cross-reactive antibodies may cause false-positive results [107]. Studies that claimed to have detected viral RNA have also come under scrutiny, as the highly conserved genome means that it is difficult to differentiate contamination with positive controls and a genuine sequence [108]. A US-based case-control study, using standardised methods, found no association between psychiatric illness with BODV-1 antibodies or nucleic acids [109], further refuting the claimed causal relationship

between psychiatric disorders and BoDV-1. Endogenous bornavirus-like elements (EBL) are DNA sequences found in many vertebrate genomes, including humans, which are derived from the mRNA of ancient bornaviruses, likely as a result of their replication within the nucleus of infected cells [110]. This ability to integrate within human genomes also raises the possibility of novel methods of pathogenicity, such as mutagenicity [111]. The presence of Long Interspersed Element 1 (L1) retrotransposons in the brain [112], which are likely responsible for the integration of EBLs in the human genome [113], exposes a potential route for the alleged causative relationship between BoDV-1 and psychiatric conditions.

The first definitive association of BoDV-1 and human disease was not until 2018. A 25-year-old man in southern Germany was hospitalised with severe fever and headache, which later developed into encephalitis. No infectious agents were detected in the normal panel of tests; post-mortem HTS sequencing of the brain tissue recovered complete BoDV-1 genome. Subsequent PCR, immunohistochemistry and indirect immunofluorescence results supported the HTS findings [114]. A similar report published in the same issue described the transplantation of kidneys and liver from a brain-dead patient (from southern Germany) to three patients. The donor recipients subsequently developed neurological syndromes; two of which developed cerebral atrophy (and encephalitis in one) and died 6-7 months post-transplantation. The third developed leukoencephalitis and slowly recovered. High-throughput sequencing

(HTS) of one of the recipients detected BODV-1, an identical sequence was then detected in the second recipient and serological evidence of BODV-1 infection was detected in the third. [115]. Retrospective screening of frozen brain biopsies from cases of neurological disease of suspected infectious origin revealed a BODV-1 positive sample from a woman (also from southern Germany) diagnosed with Guillain–Barré-syndrome (GBS)/acute polyradiculitis in 1996 [116], which shows that spillover of BODV-1 into humans is not a recent event.

Both incidences of human BODV-1 infection occurred in southern Germany, within the typical geographic range for BODV-1 infections for animals.

1.3.3.2. Mammalian 2 orthobornavirus

Between 2011 and 2013, three breeders of variegated squirrels developed progressive encephalitis or meningoencephalitis reported in Saxony, Germany, ultimately resulting in the death of all three patients [97]. At the time of death, no infectious agent could be detected using microscopy, serology, culturing or molecular methods. HTS of the patients and squirrels revealed VSBV-1.

Following this initial association of VSBV-1 and neurological disease, there have been several studies retroactively screening patients who diagnosed with a neurological disease with an unknown etiology and a history of contact with exotic squirrel species. Attempts were made to assess the risk of VSBV-1 in those with frequent contact with exotic

squirrels, in particular those involved in trading these animals. Two breeders who had died in 2005 and 2006 of meningoencephalitis had the same clinical characteristics as the three cases identified in 2015, and so it is possible that these two men were infected with VSBV-1, though no samples remained to confirm this. Serological screening identified a seropositive breeder, who had experienced severe fever and a year-long period of headaches, lethargy and myoclonus, but subsequently recovered. Contact tracing of these breeders showed that there was a trade of animals between the breeder who died in 2005 and two of the breeders who died in 2013 [117]

One such case was a zoo caretaker who developed fatal limbic encephalitis two years prior to the discovery of VSBV-1, in Northern Germany. Viral sequence recovered from the patient was most similar to a Prevost's squirrel associated VSBV-1, a species which was present at the zoo [118].

It is unclear what the exact route of transmission was, though high RNA load from the oropharyngeal swab and reports of the patients being bitten by their squirrels would suggest that infection was transmitted through the squirrel's bite.

1.4. *Picornaviridae*

1.4.1. Taxonomy

The Picornaviridae are a diverse family of non-enveloped, positive-sense RNA viruses, composed of 35 genera and 80 species. These viruses are known to cause a significant number of infections in

humans and animals. picornaviruses typically infect the gastrointestinal tract causing gastrointestinal illness, occasionally spreading into other organs where they can cause a range of diseases [119], including respiratory, hepatic and neurological disease. The tropism of these viruses, the clinical presentations and outcomes of infection vary significantly between genera, species and even serotype. The *Enterovirus* genus in humans, for example, can cause mild and sub-clinical respiratory disease in the case of *Rhinovirus A* [120], paralysis with *Enterovirus C* (Poliovirus) and encephalitis with *Enterovirus A* [121].

They are also a significant cause of disease in both wild and domestic animals [122]. *Foot-and-mouth disease virus* (*Apthovirus genus*) for example causes serious morbidity in domesticated cattle and sheep. An outbreak in the United Kingdom in 2001 led to the culling of 6.5 million animals [123] and cost the British economy approximately $8 billion [124].

1.4.2. Parechovirus

Similar to other picornaviruses, parechoviruses are small non-enveloped viruses, approximately 30nm in diameter. They have a monopartite genome 7.3kb in length, encoding one polyprotein from a single open reading frame, flanked by untranslated regions (UTR) at both the 5' and 3' end. This polyprotein is subsequently cleaved into three structural proteins (VP0, VP3 & VP1) and seven non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C & 3D) [125] (Figure 8). Recombination of the genome has also been widely reported, with the break-point often occurring at either the 5'UTR-VP0 or VP1-2A junctions [126].

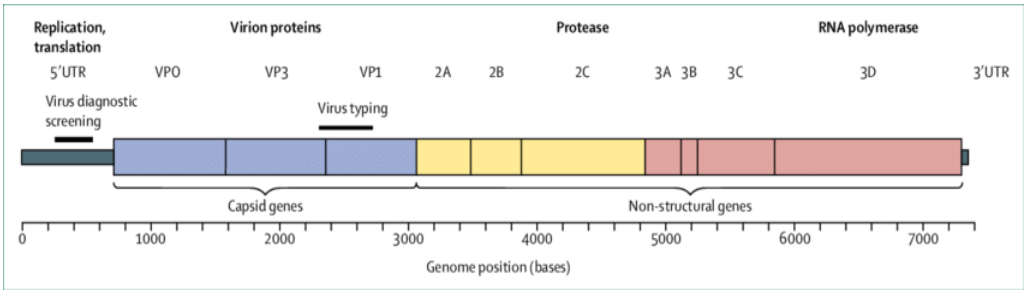

Figure 8 Schematic of the Parechovirus genome, showing the length of the typical virus, and the sizes and locations of each of the genes Adapted from [348]

The genus contains four recognised species, Parechovirus A, B, C and D. Parechovirus A, or Human Parechovirus (HPeV), is a routinely diagnosed, clinically important species, which can be further divided into 19 genotypes. Infection is typically reported in infants in the first year of life, with HPeV-1 the most prevalent genotype, causing mild gastroenteritis or respiratory illness [127]. HPeV-3 is a more clinically

important strain associated with fever and meningitis in infants under 8-weeks old [128]. Infections are globally distributed, and outbreaks of different genotypes are often seasonal [129]. Genotypes 1 and 6 of HPeV have also been detected in monkeys with diarrhoea [130], though a definitive causal link between infection and disease was not established. There is no evidence that HPeV has an origin in monkey species. A more likely scenario is that infection in monkeys is a result of zooanthroponosis, where monkeys have become infected through contact with humans. It does, however, raise the possibility that infected monkeys could transmit HPeV back into humans.

Parechovirus C (Sebokele virus, SEBV) was first detected in 1972 in African wood mice (*Hylomyscus sp.*), from Central Africa [131]. It was then fully sequenced using HTS in 2013 [132] and was found to be novel enough to create a new species. Phylogenetic analyses showed that SEBV was most closely related to the Parechovirus B species, Ljungan virus (LV).

Metagenomic analyses of ferrets (*Mustela putorius furo*) in 2013 recovered the genome of a novel Parechovirus, this virus was named ferret parechovirus (FPeV). The polyprotein of this virus was found to be less than 43% identical at the amino acid level to either HPeV or LV, and so was divergent enough to constitute a novel species, *Parechovirus D* [133]. A parechovirus recently detected in common pipistrelle bats (*Pipistrellus pipistrellus)* from China was identified as a distant second member of Parechovirus D, this virus was 72.4%

identical at the amino acid level to the FPeV, across the entire polyprotein [134].

1.4.3. Parechovirus B, Ljungan Virus

The fourth species of Parechovirus, Parechovirus B or Ljungan virus (LV) was originally identified in Swedish bank voles (*M. glareolus*) in 1999 [135] and has since been detected in predominantly rodent species, including yellow-necked mice (*Apodemus flavicollis*) [136], wood mice (*Apodemus sylvaticus*), house mice (*Mus musculus*), field voles (*Microtus agrestis)* [137] and Eurasian red squirrels (*Sciurus vulgaris*) [138]. A novel strain of LV has also been reported in gulls, though as samples were faeces collected from the ground, a rodent origin through dietary or environmental contamination cannot be ruled out [139]. LV has mostly been reported in bank voles from Scandinavia and Finland, though it is globally distributed, with strains detected in continental Europe [138], United Kingdom [137], USA [140] and Japan [139].

Whilst there is no formal genotyping criteria for LV, the enterovirus genotyping criteria of >75% nucleotide or >88% amino acid similarity in the VP1 gene has been used by many studies [139]. There are five genotypes of LV, which have strong geographic associations; genotypes 1 and 2 are associated with the European rodent strains and genotypes 3 and 4 with American rodent strains, whilst the fifth strain was detected in bird faeces in Japan. However, there is a paucity in the sequence data available for LV, with only ten complete genomes available.

Ljungan infection has been linked to several diseases in rodents, for example, type-1 diabetes in both captive [141] and wild bank voles [142] . In *in vivo* studies, LV infection and induced stress has been associated with increased intrauterine foetal death (IUFD) and foetal malformations [143] in laboratory mice, both the experimentally infected mouse and their offspring, suggesting the effects of LV infection on reproductive capability are persistent [144].

Evidence has also indicated that LV may be a zoonosis and play a role in the development of several diseases in humans. Ljungan virus was initially discovered after a series of healthy Swedish hikers developed fatal myocarditis [145]. These deaths, along with rates of Guillain-Barre syndrome and type-1 diabetes, matched fluctuations in vole populations, with peak incidence following peak population density [146]. Vole populations were screened for potential zoonoses and LV was detected. Blood from the deceased hikers was serologically screened for LV-antibodies, four of the five screened were positive [135]. LV has since been associated with several more conditions including diabetes, foetal malformation and sudden infant death syndrome (SIDS) [147]. Studies into foetal disorders have detected LV in 50% [148] to 100% [149] of IUFD cases, 90% of hydrocephaly cases and 55% of anencephaly cases [150] tested via PCR, whilst no LV was detected in a cohort of healthy pregnancies. The association of foetal syndromes and LV is controversial, however, due to small cohorts used and the lack of confirmation for the specific diagnoses in some patients [151].

Seroprevalence of LV in humans is as high as 38% in endemic regions such as Finland [152]. Subsequent studies have supported these prevalence rates, as well as showing that most infections occur during childhood, and seroprevalence was greater in highly-urbanised areas, which may indicate a possible role of transmission between humans [153]. This seroprevalence in both the general population and cohorts of patients with disease indicates that LV likely has zoonotic capability.

1.5. Viral Zoonoses in the United Kingdom

1.5.1. Orthohantavirus

Orthohantaviruses are suspected to have been a cause of disease in the UK as far back as 1485, with a series of epidemics known as the 'English Sweating Sickness' [154]. There are currently four known hantavirus reservoir species found in the UK; the brown rat (*R. norvegicus*), bank vole (*M. glareolus*) yellow-necked mouse (*A. flavicollis*) and field vole (*M. agrestis*), which could potentially harbour Seoul virus, Puumala virus, Dobrava-Belgrade virus and Tatenale virus, respectively, though Puumala and Dobrava virus has yet to be reported in the UK. The reservoir for Tula virus, the European common vole (*M. arvalis*) is also present in the UK, however, its geographic distribution is restricted to the Orkney Islands to the north of Scotland. Sero-epidemiological studies of the human UK population showed seropositivity rates of around 2% to Seoul virus in Northern Ireland [155]. Seroprevalence was increased in populations with frequent contact with rodents; Yorkshire farmworkers had

seropositivity of 7.6% [156], whilst pet rat owners and breeders had seropositivity of 32.9% [157]. Several acute hantavirus infections have been identified across the UK, in Glasgow (1988) [158], Somerset (1991) [159], Sheffield (1993) [160] and Nottingham (1997) [161], though there was no isolation of the causative virus or detection of viral RNA. In December 2011, a 59-year-old farm resident from the Humberside region of northern England presented with fever, chills and a cough, and was subsequently confirmed to have a hantavirus infection. Following trapping and screening of rodents on his farm, a novel strain of Seoul virus was recovered from two rats; this strain was named 'Humber', after the region in which it was recovered; this was the first Seoul virus to be identified in the UK [162]. In November 2012, another suspected HFRS patient was identified in northern Wales, unlike the previous patient, there was no reported contact with wild rodents, he did, however, own pet rats [163]. In addition to the hospitalised patient, epidemiological questioning of the original rat breeder in Cherwell, Oxfordshire, revealed he that may have had an undiagnosed HFRS in the previous year (later confirmed serologically). Genomic analysis of the Seoul virus isolated from the euthanized rats, a novel strain named 'Cherwell', showed a high degree of homology with the Humber strain; for the S segment was 97.29% similar at the nucleotide level and 99.77% at the amino acid level, the M segment was 96.47% and 97.97% similar, whilst the L segment was 96.19% and 99.12% similar. The Cherwell strain was isolated again in another HFRS patient in Gloucester, who also kept pet rats, the isolated virus was found to have a 100% homology to the

original strain. SEOV has also been detected in brown rats captured at several pig farms across northern England, all but one of these viruses were highly similar to the Humber strain, previously detected in northern England [164].

Whilst Seoul virus is the only species associated with confirmed cases of HFRS in the UK, it is possible that other hantavirus species may also responsible for the disease, with Puumala virus as a potential candidate. The host species for PUUV, the bank vole, is ubiquitous across northern and western Europe, including the UK (except for Ireland) and whilst PUUV is responsible for many infections in regions like Scandinavia [165], Finland [166] and Russia [167], it has yet to be reported in the UK.

Incidence of nephropathia epidemica (NE) in northern Europe is strongly linked to the complex population dynamics of the bank vole reservoir of PUUV [168]. These dynamics are strongly influenced by predator-prey relationships in Scandinavia [169], whereas in Western Europe these dynamics are typically associated with masting [170]. Masting is an event in which trees (typically beech and oak) simultaneously produce an over-abundance of seeds and nuts, this glut of food results in an explosion in the population numbers of the bank voles [171], this results in an increased PUUV prevalence in the voles and therefore increased opportunity for spillover into humans [172].

Beech woodlands in the UK are much smaller and more fragmented than those located in continental Europe [173], though masting

events do still occur and are as predictable as those in continental Europe. It is hypothesised that fragmentation of these woodlands results in 'islands' of isolated and discontinuous vole populations; causing reduced contact between vole populations from different woodlands and prevents the critical host density necessary for the virus to maintain itself in voles [174]. As viral spillover is associated with the vole populations reaching a critical threshold, the inability to reach this threshold would significantly reduce spillover opportunities. Alternatively, the often sub-clinical nature of NE infection [175] means that it is likely that many cases remain undiagnosed or misdiagnosed.

Misdiagnosis of hantavirus infection is thought to be relatively common, particularly in western Europe and the UK where, until recently, the infection was thought to be uncommon, and therefore often not considered by clinicians. Leptospirosis, a bacterial infection transmitted through rodent excreta, has several overlapping clinical manifestations with HFRS, such as high fever, acute renal failure and thrombocytopenia; this has led to a suspicion that many leptospirosis diagnoses are likely to be a hantavirus infection or a co-infection of the two. This was confirmed in the Netherlands [176] and Belgium [177], the latter of which serologically confirmed twice the number of HFRS cases than leptospirosis in hospitalised patients sharing these overlapping symptoms for both diseases.

A recently discovered, putative novel species of *Orthohantavirus*, Tatenale virus (TATV) has been identified in the UK. A single PCR

positive field vole (*M. agrestis*) was detected in 2013 in north-west England [178], a further 8 isolates were detected in 2017 in field voles trapped in northern England; giving a 16.7% prevalence in this study [179]. Currently, only a small fragment of the L and S segments have been successfully amplified from these strains, however, the complete S and M segments are required for new species to be recognised by the ICTV. Both phylogenies of the L: and S segments show that TATV clusters with other vole-borne Orthohantaviruses. Blood taken from the positive vole showed reactivity to a PUUV antigen in an indirect fluorescent antibody test, which indicates that TATV appears to be serologically cross-reactive with PUUV [178]. This would suggest that serological evidence of human PUUV infection in the United Kingdom [156] may be a result of TATV infection.

1.5.2. Orthobornavirus

Evidence of orthobornavirus infection in the UK remains sparse, and there has yet to be any molecular detection of BODV or VSBV in humans or rodents. BODV-1, however, has been detected in UK cats with 'staggering disease' a neurological condition characterised by a non-suppurative meningoencephalitis, similar to Borna disease in horses [180]. A seroprevalence study found between 2.3% and 3.1% of those who worked or lived on farms had antibodies to BODV-1, though there was no apparent correlation with the development of neurological disease [181]. Red squirrels from the UK have been screened for the presence of VSBV-1, though all were negative [99]. However, VSBV-1 has not been detected in red squirrels, even those

that were screened in regions where classical Borna Disease is endemic, which combined with the low samples sizes of UK squirrels tested (N=20) is not sufficient evidence to rule out the presence of VSBV in the UK, though it remains unlikely

The presence of BODV-1 antibodies and RNA in these populations would suggest that it is likely that BODV-1 and/or another yet unidentified bornavirus is currently circulating within the UK.

1.5.3. Parechovirus B

Human Parechovirus infection is widely reported in the UK [182]. However, there has been little research into the prevalence of LV in the UK, with only a single published study. The study sampled rodents taken from Kielder forest in northern England, and detected an overall prevalence of 24.4%, across all species. LV was detected in four species, house mice (26.1%), wood mice (19.7%), field voles (26.5%) and bank voles (27%), using a PCR assay targeting the 5'UTR [137].

1.5.4. Rodent Reservoirs

Rodents are ubiquitous throughout the United Kingdom, several species of which are known reservoirs of viruses with zoonotic potential. Many others have yet to be associated with any zoonotic viruses, or any virus at all, which is likely to be the result of a paucity in screening studies.

Brown rats were first introduced into the UK in the 18th century, having originated in China and Mongolia [183] and have since spread

throughout the entirety of the UK. Whilst they can be found in rural areas, they are considered to be a commensal animal and thrive in areas with significant human activity [184]. In the early 20th century, the population of brown rats was estimated to be as high as 40 million, however by 2018 the population estimate had significantly dropped to approximately 7 million [185], likely due to the increase in usage of pesticides. This large population as well as its propensity towards urban areas results in increased contact with humans, which in turn increases the likelihood of transmitting zoonotic viruses such as Seoul orthohantavirus and Hepatitis E virus [186], making the brown rat an important target for screening for viral zoonoses.

Several species of mice are present throughout the UK, the most common of which belong to either the *Mus* or *Apodemus* genus. House mice (*M. musculus*) and wood mice (*A. sylvaticus*) are the most prevalent species. Much like the brown rat, house mice are considered commensal 'pest' species and populations are much higher in urban areas. Population estimates from 2018 put the British population at approximately 5.2 million [185], with the population steadily decreasing since the 1970's. Wood mice are significantly more abundant in the UK, with an estimated population of 39.6 million. They are more abundant in woodland habitats [187], however, they are found in urban areas aswell, where they are considered a pest species. The yellow-necked mouse (*Apodemus flavicollis*) is a close relative of the wood mouse, they are, however, much less common in the UK. They share a similar preference for woodland habitats,

though their host range is restricted largely to southern England and central and eastern Wales [188]. Population sizes in the UK are similarly restricted in comparison to wood mice, with an estimated size of approximately 1.5 million. These species of mice are associated with known viral zoonoses globally, for example, house mice [189] and wood mice [190] are reservoirs of lymphocytic choriomeningitis virus (LCMV), a zoonotic virus which can cause severe viral haemorrhagic fever. The yellow-necked mouse is also the known reservoir of a viral zoonoses, Dobrava-Belgrade orthohantavirus.

There are four species of voles present in the UK; Field vole (*Microtus agrestis*), bank vole (*Myodes glareolus*), water vole (*Arvicola amphibious*) and a subspecies of common vole, the Orkney vole (*Microtus arvalis orcadensis*). Bank and field voles are amongst the most abundant mammal species present in the UK; the field vole population has been estimated at 59.9 million, whilst bank voles have an estimated population of 27.4 million. Water and Orkney voles are considerably rarer, with a population size of approximately 132 thousand and 1 million, respectively [185]. Except for the Orkney vole, these species are distributed throughout mainland UK; the range of the Orkney vole is limited to only the Orkney Islands, of the northern coast of Scotland and the Channel Island of Guernsey. Preferred habitat varies slightly between vole's species; water voles primarily inhabit the banks of various waterways [191], bank voles populations are denser in deciduous woodland, but also extend into heavily vegetated grassland habitats [192]. Field and Orkney voles, however, predominate in

rough grassland and anthropogenic habitats such as meadows and field-margins [193]. Both field and bank voles are becoming more common in urban areas, particularly in gardens, likely as a result of increasing human encroachment on the vole's natural habitat. These species of voles have been associated with several zoonotic viruses, though not necessarily in the UK; for example Cowpox virus is found in field voles [194], Puumala orthohantavirus in bank voles [195] and Tula orthohantavirus in common and water voles [196]

1.6. Molecular Detection of Viral Zoonoses

1.6.1. Polymerase Chain Reaction and Sanger Sequencing

Perhaps the most commonly used technique for detection of infectious agents is Polymerase chain reaction (PCR) [197]. There have been many variations of this technique, designed to achieve different objectives, such as quantitative PCR (qPCR) and digital PCR (dPCR) [198] to quantify any amplified product, or reverse-transcription PCR (RT-PCR) to amplify a target from an RNA sample.

Sanger sequencing is a widely used sequencing technology, which measures the selective incorporation of fluorescently or radioactively labelled chain-terminating ddNTPs into a PCR amplified template sequence.

These techniques when used for diagnostics can be cheap, simple to run and have reproducible results. However, the oligonucleotide primers used in these assays must be designed to target a single or a

conserved motif in several sequences meaning only a limited range of sequence diversity can be detected. This can be mitigated by multiplexing primers in a single reaction, which by adding different primers will broaden the range of targets at the expense of potentially reducing sensitivity. Alternatively, primers can be designed to target several aligned sequences by adding degenerate nucleotides to cover any mismatched bases in the consensus [199].

A significant drawback to using these techniques for the discovery of novel viruses, or viruses presenting in a novel way (such as a recent spillover of a known virus into a novel host), is that the targeted approach biases them towards only looking for viruses that would be expected to be found in each particular sample [200]. For example, researchers may not look for a virus associated with respiratory disease in a CSF sample, or a virus associated with rodent reservoirs in bat samples.

1.6.2. High-Throughput Sequencing

High-throughput sequencing (HTS), or 'next-generation sequencing' (NGS), are terms that encompass several modern sequencing technologies and platforms that can sequence massive numbers of reads in a single sequencing run. Early forms of these technologies were first used for virus discovery in 2002, to detect marine viruses from seawater [201], and in human samples in 2003 [202]. Initially, the costs of these sequencing platforms were prohibitively expensive for widespread screening purposes, but have become significantly

cheaper in recent years, resulting in a massive increase in the number of studies utilizing them for virus discovery.

Short-read sequencing platforms such as the Illumina MiSeq and HiSeq are amongst the most commonly used for microbe detection. Sequencing of samples using Illumina technology involves preparing a sequencing library. DNA or cDNA is randomly fragmented into homogenous fragments, the size of which can vary depending on the specific sequencing machine or needs of the user. Adaptors are then ligated onto the 5' and 3' ends of the fragments and the samples amplified using indexing oligonucleotide primers, which uniquely tags each sample and allows multiplexing of several samples into a single library. This library is loaded onto a flow cell containing surface-bound oligonucleotides complementary to the ligated adaptors. The fragments bind to the oligonucleotides and are clonally amplified using bridge amplification to produce a cluster of approximately 1000 copies, with each cluster resulting in a single read. The fragments are sequenced using 'Sequencing-by-synthesis', where polymerases are used to add a fluorescently labelled nucleotide complementary to the first base in the fragment, the flow-cell is then read, and the process repeats until the fragment has been completely sequenced. Paired-end sequencing involves sequencing both the forward and reverse directions of the fragment and aligning them to create a read-pair; this allows for more accurate sequencing results. Once sequenced, the data can be demultiplexed by sorting the reads by the sequence of the indexing primers [203]. Depending on the purpose of the sequencing,

reads can be mapped onto a reference genome, which is useful when looking for a specific virus for genotyping or similar studies. Alternatively, the reads can be *de novo* assembled into contigs without any reference genome, which is useful for detection of unknown or novel viruses.

Whilst Sanger sequencing and Illumina, representing the first and second generation of sequencing technologies respectively [204] have been used to great effect for virus discovery and analysis; the 'long-read' technologies of the third generation of sequencing have become increasingly frequent [205]. One of the most common platforms in this generation is the Oxford Nanopore MinION. Sequencing libraries are prepared in a similar process to Illumina sequencing; sample DNA is sheared and end-repaired to create blunt ends to the DNA, an A is added to the 3' end and adapters ligated to both ends of the DNA. Samples are then passed through a nanopore via a motor protein. A voltage is applied to the nanopore containing membrane, creating a measurable ion current flow. Each nucleotide causes a characteristic change in the magnitude and pattern of current which allows for base-calling [206]. This allows for ultra-long reads to be sequenced, with single reads of 10,589 base pairs and higher being recorded [207], which would not be possible on platforms such as Illumina. Whilst the accuracy [208] and depth of the reads are not as great as those from Illumina, the MinIon has a significant advantage in that is portable and can sequence in real-time [209], allowing it to be easily and rapidly deployed in field laboratories [210]. This has

been successfully used for diagnosis and surveillance of samples during virus outbreaks such as Zika [211] and Ebola [212,213].

1.7. Aims of the study

The initial aim of this project was to assess the role of British rodents as reservoirs of viral zoonoses and viruses with zoonotic potential, with a focus on viruses in the *Hantaviridae, Picornaviridae* and *Bornaviridae* families. This would be achieved through the collection of rodent tissue and screening using RT-PCR with degenerate, pan-species primers. These methods would also be extended to Polish and Egyptian cohorts or rodents, to assess the prevalence of similar viruses. Secondary to the primary aim was the recovery of the complete genome and characterisation of any viruses detected using high-throughput sequencing.

A secondary aim of this project was to screen several cohorts of undiagnosed, human clinical samples to investigate the possibility that viruses with a zoonotic origin may be present in these samples. These samples will be screened using high-throughput sequencing methods and any viruses discovered will be further characterised.

Finally, the seroprevalence of orthohantavirus infection in human and non-human primate populations in the East Midlands region of England will be assessed using ELISA serological assays. This will give further insight into the levels of exposure to orthohantaviruses in this region.

2. Materials and Methods

2.1. Screening of Rodent Tissue for Viral Zoonoses Using Degenerate Primer RT-PCR

2.1.1. Collection Sites and Samples

2.1.1.1. United Kingdom

Rodents were captured at four sites in the United Kingdom. The first two sites were Twycross zoo and Sutton Bonnington of the University of Nottingham campus, both are in Leicestershire, in the East Midlands region of England. The third site was Chester Zoo, in Cheshire, North-Western England. The fourth site was the island of Jersey, in the English Channel. These locations are shown in Figure 9.

Samples from Twycross zoo and Chester zoo were captured as part of routine pest-control measures between 2013 - 2014, and 2013 – 2016 respectively. Those from Sutton-Bonnington were collected from pest-control measures or recovered from feline-predation between 2014. Samples from Jersey had been hit by vehicles or found injured and taken to a veterinarian who subsequently euthanised them. They were collected in 2008 and 2011

Necropsies were performed on samples from Sutton Bonnington and Twycross zoo; heart, brain, liver, kidney, gut, lung and spleen were collected and stored in RNALater (Sigma Aldrich, St. Louis, MO, USA) at -70°c. Samples from Chester zoo and Jersey had all been necropsied offsite. Rodents from Chester zoo had lung, liver, kidney,

gut and spleen collected, and were stored at -70°c. Rodents from Jersey had Spleen collected and were stored without preservative at -70°c.

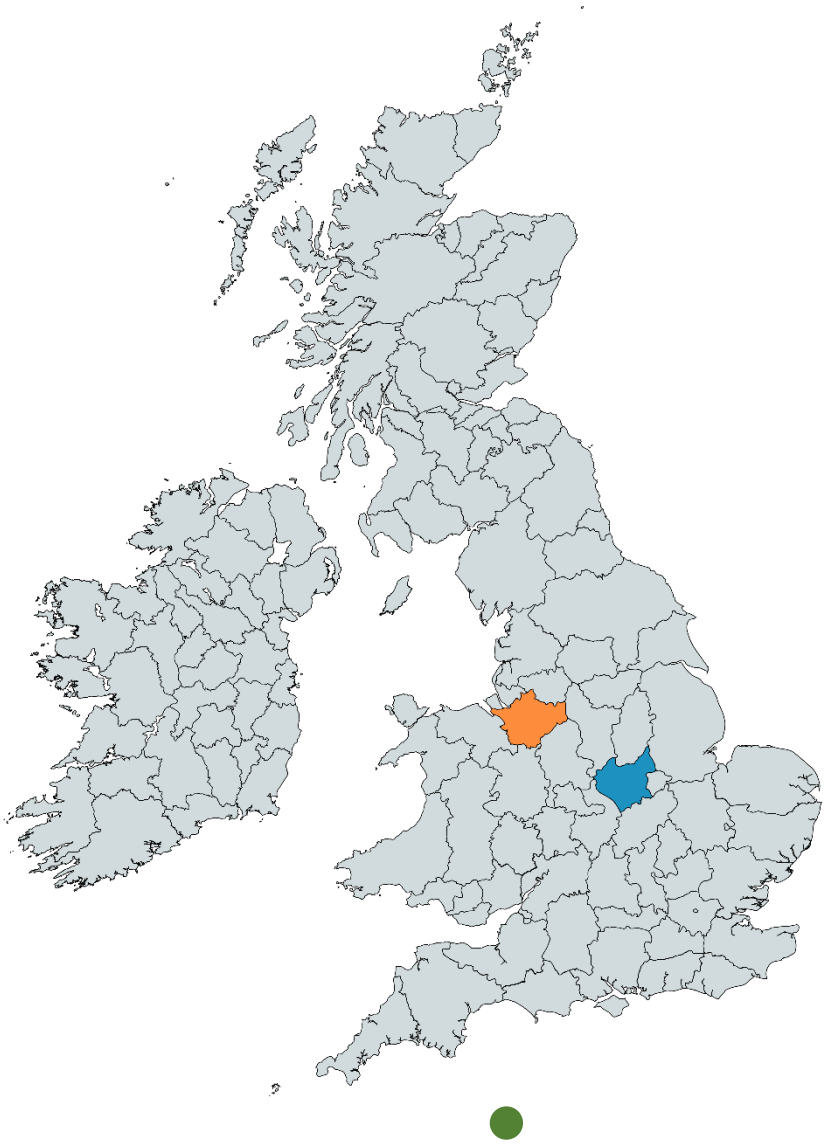

Figure 9 A Map of the United Kingdom, showing the locations in which rodents were sampled. Leicestershire is highlighted in blue, Cheshire is highlighted in orange and Jersey is represented by a green circle.

2.1.1.2. Poland

A cohort of Bank voles (*Myodes glareolus*) was captured at three adjacent sites in the Mazury lakes district in the Warmian-Masurian Voivodeship, North-Eastern Poland (Figure 10).

Animals were live-captured, euthanised and processed as part of previous studies [214]. Collection of samples was approved by the Ethics Commission for Experiments on Animals of the M. Nencki Institute of Experimental Biology of the Polish Academy of Sciences (Warszawa, Poland). The samples used in this study were collected in August and September 2010.

Only the livers of these animals were collected, which were preserved in ethanol and stored at -70°c.

Figure 10 A Map of Poland, the Warmian-Masurian Voivodeship is highlighted in blue, the Mazury lakes district is marked in orange.

2.1.1.3. Egypt

A cohort of Cairo spiny mice (*Acomys cahirinus)* was collected from sites near St Catherine, in the South Sinai Governorate, Egypt (Figure 11). Animals were live-captured and euthanised for a previous study [215]. Liver samples were collected and stored at-70°c.

Figure 11 A Map of Egypt, with the South Sinai Governorate highlighted in blue, and Saint Catherine highlighted with an orange circle.

2.1.1.4. Number of Rodent Samples Collected

The total number of rodents collected at each site, with the organs available for each species are shown in Table 1. Also shown is the percentage of organ samples that were GapDH positive.

Location	Species	Animal count	Organ	Number	GAPDH+ (%)
Twycross Zoo	Brown Rat	71	Kidney	71	100
			Lung	67	100
	House Mouse	208	Kidney	205	91.52
			Lung	208	92.86
	Field Vole	2	Kidney	2	100
			Lung	2	100
Chester Zoo	Brown Rat	8	Lung	8	100
			Liver	2	100
	Field Vole	109	Lung	104	100
			Liver	83	100
	Wood Mouse	81	Lung	78	100
			Liver	39	100
	Bank Vole	3	Lung	3	100
			Liver	1	100
Jersey	Red Squirrel	22	Spleen	22	100
Sutton Bonnington	Field Vole	11	Liver	11	100
Poland	Bank Vole	293	Liver	293	100
Egypt	Spiny Mouse	150	Gut	150	100

Table 1 Table of the number of rodents sampled at each site, and the organs that were collected for each species.

2.1.2. Processing of Rodent Tissue

2.1.2.1. RNA Extraction

RNA was extracted from 2 mm^3 sections of tissue using the GenElute Mammalian Total RNA Miniprep Kit (Sigma Aldrich), and eluted in 50µl of elution buffer, following the manufacturer's instructions.

RNA extractions were stored at -70°c.

RNA was quantified using 2µl of extracted RNA on either a Nanodrop 2000 (Twycross zoo, Sutton Bonnington, Jersey, Poland, Egypt rodent cohorts) or Nanodrop One spectrophotometer (Thermofisher Scientific) (Chester zoo cohort).

2.1.2.2. cDNA Synthesis

First-strand cDNA was synthesised from total RNA using a RevertAid RT Reverse Transcription Kit (Thermo Scientific). Reactions were carried out in a 20µl mixture containing 0.5µl random hexamers (0.4 µg/µl), 4.5 µl 5x Reaction Buffer, 1µl RevertAid Reverse Transcriptase, 2µl of 10 mM dNTPs and 10µl RNA. As maximum total RNA input was 5µg (500ng/µl), samples exceeding the threshold were diluted with DEPC-treated H2O to a final concentration of 500ng/µl.

Reactions were incubated at 25°C for 10 minutes, followed by 42°C for 60 minutes and a final deactivation period of 10 minutes at 70°C. cDNA was stored at -25°C

2.1.3. Positive Controls

2.1.3.1. Orthohantavirus

Initial positive controls used for screening were plasmid vectors encoding the L segment (Polymerase) of Seoul virus, was provided by the late Professor Richard M Elliot (MRC-University of Glasgow, Centre for Virus Research). These plasmids were generated for use during a previous project, prior to the onset of this project.

After the first recovery of Tatenale virus in a field vole from Twycross zoo, the cDNA from this animal was used as a positive control for further screening.

2.1.3.2. orthobornavirus

Initially, the positive control for the bornavirus assay used a plasmid containing a synthesised Variegated Squirrel Bornavirus-1 (VSBV-1) glycoprotein. Borna Disease Virus-1 and VSBV-1 RNA were provided by Dr Bernd Hoffman (Friedrich-Loeffler-Institut, Germany), which was used for subsequent positive controls.

2.1.3.3. Parechovirus

The Parechovirus assays used human parechovirus (HPeV) positive cDNA extracted from clinical samples, obtained from surplus diagnostic samples provided by the NHS. These samples were completely anonymised, and ethical approval had been obtained for extended viral diagnostics to be carried out

2.1.4. Polymerase Chain Reaction (PCR)

2.1.4.1. PCR Reaction Components

PCR reactions were kept consistent across all PCR assays. Each reaction contained 1.25 µl Mg2+ free 10x PCR buffer, 0.06 µl of HotStarTaq DNA Polymerase (Qiagen), 0.5 µl of 2.5 mM dNTP (Sigma Aldrich), 0.5 µl of forward and reverse primer (10 Pmol/µl), 9.19 µl of water was then added for a total volume of 12µl. 0.5µl of cDNA was then added to each reaction.

2.1.4.2. GapDH cDNA Validation

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was targeted to validate the quality of generated cDNA. Previously designed primers targeting the conserved regions of the mammalian GAPDH gene were used [216];

GAPDH-F 5'-CCATCTTCCAGGAGCGAGA-3'

GAPDH-R 5'-GCCTGCTTCACCACCTTCT-3'.

The Cycling conditions for this assay included an initial heat activation of 95°C for 15 minutes, followed by 55 cycles of 94°C for 20 seconds, 55°C for 20 seconds and 72°C for 30 seconds, and a final extension of 72°C for ten minutes. These primers produced a 571bp amplicon.

2.1.4.3. Cytochrome B Species Identification

Published primers targeting the mitochondrial Cytochrome B gene were used to genetically speciate rodent samples. Primer sequences were

CytoB-F 5'-TGAGGBGCYACAGTWATYACAAAC-3'

CytoB-R 5'-CGYAGGATDGCRTATGCRAATA-3'

The cycling conditions were 95°C for 15 minutes, followed by 40 cycles of 95°C for 20 seconds, 51°C for 20 seconds and 72°C for 2 minutes, with an extension of 72°C for 10 minutes.

2.1.4.4. Orthohantavirus Screening Assays

Pan-Orthohantavirus primers were taken from Klempa et al [217];

HANL F: 5'-ATGTAYGTBAGTGCWGATGC-3'

HANL-R: 5'-AACCADTCWGTYCCRTCATC-3'.

Cycling conditions were modified from the article; final conditions were 94°C for 15 minutes, followed by 55 cycles of 94°C (30 seconds), 53°C (50 seconds) and 72°C (30 seconds), finished with 72°C (10 minutes). These primers produce an amplicon of 452bp.

The second pair of pan-Orthohantavirus primers, Han-Semi, was designed in-house by Dr Theocharis Tsoleridis. The sequences of these primers were;

HanSemi-F 5'-GAATATATATCNTAYGGDGGDGA-3'

HanSemi-R 5'-CTGGTGACCAYTTNGTNGCAT-3'.

These primers targeted the L segment and produced an amplicon of 178bp. Cycling conditions were 94°C for 15 minutes, followed by 55 cycles of 94°C, 51°C and 72°C for 20 seconds each, with a final extension of 72°C for 10 minutes.

2.1.4.5. Orthobornavirus Screening Assays

Three pairs of primers were created, each targeting the same region of the glycoprotein, but varying in levels of introduced degeneracy. D-Bor was fully degenerate to all available complete orthobornavirus genomes available. SD-Bor was semi-degenerate, designed to target

only mammal-associated viruses, allowing reduced degeneracy. Finally, VSBV-1 was specific to the reference sequence for *Mammalian 2 bornavirus* (Accession number LN713681).

DBor primers were;

DBorF 5'- GAYGCNTGGGANGAYTGYGARAT-3'

DBorR 5'- CCNCCNADCCANGCNGCCCA-3'.

SDBor primers were;

SDBorF 5'- GATGCATGGGAAGAYTGYGARAT-3'

SDBorR 5'- CCTCCAADCCANGCNGCCCA-3'.

VSBV-1 primers were;

VSBV-1F 5'-GATGCATGGGAAGATTGTGAGAT-3'

VSBV-1R 5'- CCTCCAATCCATGCTGCCCA-3'.

All primers shared cycling conditions, 94°C for 15 minutes, followed by 55 cycles of 94°C (15 seconds), 53°C (30 seconds) and 72°C (30 seconds), finished with 72°C (10 minutes). Amplicon sizes were 632bp.

2.1.4.6. Parechovirus Screening Assay

A Pan-Parechovirus primer pair was adapted from Nix *et al* [218], the primers AN345 and AN344 were renamed qPAR-F and qPAR-R, respectively. They target a conserved region of 5' untranslated region,

producing a 194bp amplicon from human parechoviruses and 204bp from ljungan viruses. Sequences of these primers were;

qPAR-F 5'-GTAACASWWGCCTCTGGGSCCAAAAG-3'

qPAR-R 5'-GGCCCCWGRTCAGATCCAYAGT-3'

The cycling conditions for these primers were 95°C for 15 minutes, followed by 55 cycles of 95°C for 20 seconds, 65°C for 30 seconds and 72°C for 20 seconds, with a final extension of 72°C for 10 minutes.

2.1.4.7. Gradient PCR

Gradient PCRs were used to optimise the PCR assays. This method involved using eight identical 12.5 µl reactions placed across the heating block of a thermocycler capable of generating a gradient of annealing temperatures, such that each reaction had a hotter annealing stage than the last.

Cycle conditions of primers were identical to those used in the original assays, except for the annealing temperature. A range of annealing temperatures was used, depending on the primer. Temperatures gradually increased with each reaction.

2.1.5. Gel Electrophoresis

PCR products were visualised using gel electrophoresis. 2% TAE agarose gels with 5% ethidium bromide. 12.5µl of the PCR reaction was mixed with 2.5µl of loading dye, 6µl of this mix was loaded onto

the gel, and run for 36 minutes at 90 volts. Completed gels were visualised using UV light.

2.1.6. Sequencing and Sequence Analysis

'Positive' PCR samples (Those with appropriately sized amplicons) were diluted to 1ng/µl per 100bp and sent with 3.2 pmol/µl of the appropriate forward primer to SourceBioscience (Nottingham, UK), where they were Sanger sequenced.

Chromatograms of the sequencing were received and analysed in FinchTV (Geospiza Inc). Using BLAST (Basic Local Alignment Search Tool), samples were compared against sequence data in GenBank using the BLASTn function for nucleotide comparison and BLASTx function for translated nucleotide comparison).

2.1.7. Analysis of Primers

2.1.7.1. Orthohantavirus Primer

All complete rodent-bourne Orthohantavirus L segments were downloaded from the NCBI GenBank database, using the 'Virus Pathogen Resource' (ViPR) [219] (Appendix 1). Sequences were opened in MEGA7 and were aligned based on codons, using the MUSCLE function.

Han-L and Han-Semi primers were then mapped onto the aligned sequences using the 'Map Primers' function of Geneious Prime.

2.1.7.2. Orthobornavirus Primers

Complete orthobornavirus glycoprotein sequences were downloaded from the NCBI GenBank database (Appendix 2) and aligned based on codons, using the MUSCLE function of MEGA7.

2.1.8. Statistical Analyses of Picornavirus Infected Bank Voles

SPSS (Version 23) was used to analyse the cohort of Polish bank vole livers. A hierarchical log-linear analysis model, incorporating age (adolescent, adolescent-mature or mature), sex (male or female) and capture location (Talty, Pilchy or Uriwalt) as variables were used. This analytical approach used stepwise regression, specifically backwards elimination, in which interactions between each variable were examined and the most statistically insignificant interaction is removed. Multiple steps were run, until only statistically significant interactions remain. A chi-squared test was included in the model to further ensure no relationships were incorrectly rejected.

2.1.9. Phylogenetic Analyses

2.1.9.1. Orthohantavirus HAN-L Amplicon

All complete L segment coding sequences of vole associated Orthohantaviruses, along with the partial L sequences from Tatenale virus strain B41 and Kielder were downloaded using VIPR.

These sequences along with the Norton-Juxta and Upton-Heath HAN-L amplicons were aligned by codon using the MUSCLE function of MEGA7. A best-fit substitution model was used on the alignment to determine the most appropriate substitution model.

A maximum-likelihood tree was generated using the T92+G+I substitution model, from the aligned L sequences. One thousand replicates were used to determine bootstrap support.

2.2. Screening of Rodent Tissue for Viral Zoonoses Using Degenerate Primer RT-PCR

2.2.1. Evaluation of Enrichment Methods

An enrichment protocol was adapted from Daly *et al* [220]. A 2mm^3 section of tissue was suspended in 250µl of iced PBS. The tissue was then homogenized using a Ribolyser and snap-frozen on dry-ice for 2 minutes and thawed on ice. A total of three freeze-thaw cycles were used on each sample. The resulting homogenate was spun down at 600g for ten minutes, at 4°C. The homogenate was then removed from the pelleted cellular debris, and 30 units of DNase (Promega) added, the homogenate was then incubated at 37°C for 30 minutes. RNA was then extracted using the GenElute Mammalian Total RNA Miniprep Kit, and cDNA synthesised using the protocol described in 2.1.2.2

The PCR Reactions and cycling parameters used are described in 2.1.4.

2.2.2. Library Preparation

2.2.2.1. RNA Quality Control

Total RNA was extracted from animal tissue using the GenElute Mammalian Total RNA Miniprep Kit.

Samples were run on an Agilent 4200 Tapestation (Agilent), using an RNA ScreenTape to quantify the yield and assess the quality of the extracted RNA. Quality of the RNA was expressed as an RNA Integrity Number (RIN), with 1 being the lowest quality (and most degraded), and 10 the highest (Least degraded).

2.2.2.2. Library Preparation

Libraries were prepared using NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (New England Biolabs, Ipswich, Massachusetts, USA) and NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs), using the unmodified protocol.

A total of 1μg of RNA for each sample was used. RNA was host ribosomal RNA (rRNA) depleted, and then sequencing libraries created following the unmodified protocol provided. The fragmentation stage of protocol varied, depending on the RIN of the RNA input. RINS of each sample, the fragmentation time and their unique index number are shown in (**Table 2**). Reactions were optimised for insert sizes of approximately 200bp. Seven amplification cycles were used for all samples.

Sample	RIN	Fragmentation Time (Minutes)	Index
KE4D	3.2	7	1
KE16D	2.4	7	2
KF5D	2.6	7	3
Swm19/20	4.7	7	4
E120	8.3	15	6
E293	8.1	15	7
KE4F	1.5	0	8

Table 2 Samples submitted for HTS library preparation. The RIN required fragmentation time and their unique index number are included.

2.2.2.3. Assessing Library Quality

Completed libraries were analysed using a high-sensitivity DNA ScreenTape on an Agilent 4200 Tapestation.

2.2.2.4. High-Throughput Sequencing

Completed libraries were sent to Source BioScience (Nottingham), where they were analysed for quality control purposes, and sequenced on an Illumina HiSeq. Paired reads of 2 × 150 bp were produced.

2.2.3. Mapping Reads Directly to a Reference Sequence

2.2.3.1. Mapping of Tatenale Orthohantavirus Reads

The reference sequences for Puumala virus were downloaded from GenBank. The accession numbers of these sequences were NC_005225 (L), NC_005223 (M) and NC_005224 (S).

The reads from the KE4D and KE4F indexes were mapped to these sequences using Geneious Prime. The Geneious mapper tool was used, with medium-low sensitivity. Iterative fine-tuning was used, with five iterations of mapping reads onto the consensus of the previous iteration.

2.2.3.2. Mapping of Picornavirus Reads

E120 was mapped to the reference sequence for rosavirus-2 (NC_024070.1). E293 reads were mapped to NC_003976.2, the reference genome for Ljungan virus. KE16D reads were mapped to the genome of Boone cardiovirus (JQ864343.1). KF5D reads were mapped to the reference genome for Norway rat hunnivirus (NC_025675).

The Geneious mapper tool, in Geneious Prime, was used to map reads to the respective reference genomes. Five iterations were used to increase sensitivity.

2.2.4. Processing of HTS Datasets

2.2.4.1. Merging Paired Reads and Filtering Low-Quality Reads

Geneious Prime software was used to process the HTS datasets. The Paired reads were linked using the 'Set-Paired Reads' function, with the forward/reverse (inward pointing) parameters used. Paired reads were then merged using the BBMerge plugin with default settings.

Low-quality reads were trimmed using the BBDuk plugin, which removed adaptors, low-quality sections of reads (minimum quality of

6) and low complexity regions. Reads under 40bp, post-trimming were discarded.

2.2.4.2. Removal of Host Sequences

The Reference genome assemblies of each of the sampled animals were downloaded from GenBank. The GenBank genome assembly accession numbers used for each sample are shown in **Table** . All Samples used the correct genome for each animal species, except for the swan sample. The genome assembly used to filter host sequence for swans was the genome assembly of mallard (*Anas platyrhynchos*), as no assembly was available for swan.

Sample	Species	Reference Genome
E120	Bank Vole	GCA_001305785.1
E293	Bank Vole	GCA_001305785.1
KE4D	Field Vole	GCA_001305995.1
KE4F	Field Vole	GCA_001305995.1
KE16D	Brown Rat	GCA_000001895.4
KF5D	House Mouse	GCA_000001635.8
Swm19/20	Swan	GCA_003850225.1

Table 3 The GenBank accession numbers of the host genome assemblies used to filter out host sequence from each dataset.

The quality trimmed reads were mapped to the host genome assembly, using the Geneious mapper with medium-low sensitivity and five iterations. Reads mapping to the host reference were discarded.

2.2.4.3. *De novo* Assembly

Quality and host filtered reads were *de novo* assembled into contigs. The Geneious assembler was used, with medium-low sensitivity. The specific setting was the default for the sensitivity used, this included a maximum of 10% gaps per read and a maximum gap size of 2, a word length of 18, the index word length was 13, words that were repeated more than 100 times were ignored. The maximum permissible mismatches per read were 20%, and the maximum ambiguity was 4 per read.

2.2.4.4. Reference Database

Custom BLAST databases were created for each virus taxonomic order (Table 4). All complete viruses within each order as of March 2018 were downloaded from GenBank. Viruses from the *Retroviridae* family were not included. Databases for *Tymovirales*, *Caudovirales* and *Ligamenvirales* were not created as viruses in these orders are not associated with vertebrates.

An additional database of the RNA viruses polymerase proteins was obtained from a colleague, containing 2741 sequences.

Viral Order	Number of Sequences
Bunyavirales	10,650
Herpesvirales	13,334
Mononegavirales	28,628
Nidovirales	31,866
Picornavirales	7,791
Non-Ordered Viruses	671,634

Table 4 Custom BLAST databases created from all complete sequence within a viral order, and those that had not been assigned a taxonomic order. The number of sequences within each database are also shown

2.2.5. Recovery of Upton-Heath TATV

2.2.5.1. Primer Design

The Norton-Juxta strain of Tatenale virus was used as a reference for designing a series of overlapping primers to amplify the coding sequence of Upton-Heath.

The primers for the S Segment are shown in Table 5, the CDS of M in Table 6 and the CDS of L in Table 7.

Primer	Sequence (5'-3')	Product Size (bp)
TATV S 1 F	TAGTAGTAGACTCCTTGAAGAGC	705
TATV S 705 R	ACCCATGACAGGACTTACAA	
TATV S 586 F	CCTACTGCACAATCAACGAT	731
TATV S 1316 R	ATCTCCTTCACCTTCTGATCA	
TATV S 1168 F	ATGGGAATACAACTCGACCA	666
TATV S 1833 R	TAGTAGTATGCTCCTTGAAAAGC	

Table 5 List of primers used to amplify the S segment the Upton-Heath strain of Tatenale virus.

Primer	Sequence (5'-3')	Product Size (bp)
TATV M 1F	ATGGAACAAAGAAGTACAGTTTGTCT	800
TATV M 800 R	TACAGTGGTTCAGAGCTCCC	
TATV M 715 F	ACACTTGTCCAGAAAACAGA	803
TATV M 1517 R	AGCCAACCAAAACAAAATGT	
TATV M 1366 F	ACACTGGTTATAGGGCAATG	825
TATV M 2190 R	CCCTAAATGTTGAATTTCTGCA	
TATV M 2089 F	TACAGACGTGAACTCCAAAA	790
TATV M 2878 R	GGTCCTTCAATGAACTGTCT	
TATV M 2748 F	TGGTAATACAGTCTCTGGGT	603
TATV M 3351 R	CACGACCATCCARTTHCCA	699
TATV M 3447 R	CTATGGCTTGTGCTCTTTATAAT	

Table 6 List of primers used to amplify the CDS of the M segment the Upton-Heath strain of Tatenale virus.

Primer	Sequence (5'-3')	Product Size (bp)
TATV L 37 F	ATGGAGAAATACAGAGAGATTCA	777
TATV L 813 R	AAAAGCAAAATCATGGTCCC	
TATV L 678 F	ATTTCATGTCACAGGTCCAA	783
TATV L 1460 R	AGATGCCATGCTGTAGTTTT	
TATV L 1339 F	AGTCAATCCATAAGTCACCC	799
TATV L 2137 R	ATACTCCACTTGCTCCAATT	
TATV L 1844 F	CAGTGTTTGCCTTCCATTTT	824
TATV L 2667 R	AGCCTTTTGATGTTTTGTCTG	
TATV L 2499 F	TTTCAGCCAGACACGAAATA	808
TATV L 3306 R	GAAACAGTCTAACTCAGGGA	
TATV L 3178 F	GTCTCAGCAATGATTAAAGGC	848
TATV L 4025 R	AACTCTCCTAATCTGTCATGC	
TATV L 3828 F	TTTAGGTGGTGATGGGTCTA	850
TATV L 4677 R	ACCTAGTTCCCTGTAGACTC	
TATV L 4537 F	AACCCAATAACTGCTGTCAT	840
TATV L 5376 R	CCTCATACCACACCTCAAAA	
TATV L 5203 F	AAGCTATACGAAGGGGATCT	720
TATV L 5922 R	GCTTGACATTGAACTGACTG	
TATV L 5735 F	TCATGAAAGGGTTGCCTG	767
TATV L 6501 R	TTAATAAAAGGAAGAGGCTGAATC	

Table 7 List of primers used to amplify the CDS of the L segment the Upton-Heath strain of Tatenale virus.

2.2.5.2. PCR Conditions

The RNA extraction and cDNA synthesis protocols used were identical to 2.1.2.1 and 2.1.2.2, respectively. PCR reactions consisted of the same reagents and concentrations as those described in 2.1.4.1.

Unless otherwise stated, the cycling conditions for these primers was an initial heat activation of 95°C for 15 minutes, followed by 55 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute, and a final extension of 72°C for ten minutes.

PCR amplicons were run on a 2% agarose gel and sent to Source Bioscience for Sanger sequencing.

2.2.5.3. Analysis of Complete TATV

The coding sequences of the Upton-Heath and Norton-Juxta strains of TATV were codon aligned, along with other complete vole-associated Orthohantaviruses, using the MUSCLE function of MEGA7.

Maximum-likelihood phylogenetic trees were then produced, using the most appropriate substitution model, as determined by a best-fit substitution model.

2.2.6. BLAST Searching Datasets Against Virus Databases

Each of the trimmed and filtered datasets was BLASTN searched against each of the nucleotide databases of virus sequences, and BLASTx searched against the Nr protein database of RNA virus polymerases.

A Maximum E-Value of 10^{-5} was applied, to reduce the number of false-positives. A word-size of six was used.

2.3. Investigating Seroprevalence of Orthohantavirus In Humans and Captive Non-Human Primates

2.3.1. Sample Collection

2.3.1.1. Human Cohort

Human sera were selected from a pre-existing cohort of Hepatitis C virus (HCV) positive patients. The chosen samples had been taken at several healthcare centres in Nottingham that provide needle-exchange services and support for homeless and intravenous drug-using populations. 181 samples were selected.

2.3.1.2. Non-Human Primate Cohort

Non-human primate (NHP) sera were obtained from a previously established biobank of animal samples at Twycross Zoo, Leicestershire. These samples were primarily taken when animals were sedated during routine medical treatment. Ethical approval for investigating orthohantavirus prevalence was approved by Twycross zoo. The NHP samples used are shown in Table 8. A total of 56 samples were obtained.

Species		Samples
Chimpanzee	*Pan troglodytes*	24
Bonobo	*Pan paniscus*	8
Gorilla	*Gorilla gorilla*	3
Orangutan	*Pongo pygmaeus*	2
Allen's swamp monkey	*Allenopithecus nigroviridis*	1
Crowned guenon	*Cercopithecus pogonias*	1
De Brazza monkey	*Cercopithecus neglectus*	5
Diana monkey	*Cercopithecus diana*	4
Lesser spot nosed guenon	*Cercopithecus petaurista*	2
L'hoest monkey	*Allochrocebus lhoesti*	2
Lowe's Geunon	*Cercopithecus lowei*	2
Roloway monkey	*Cercopithecus roloway*	2

Table 8 Number of samples of each species used for serological screening of orthohantavirus.

2.3.2. IgM and IgG ELISA

VectoHanta-IgM and VectoHanta-IgG ELISA kits (Vector-Best, Novosibirsk, Russia) were used, following the provided protocol. These assays use Hantaan virus, Seoul virus, Puumala virus and Dobrava virus antigens.

A sample was considered positive if the optical density was greater than, or equal to the average optical density of the negative controls plus 0.2.

Samples were run in duplicate, and any positive results were repeated.

2.3.3. PCR Screening of IgM Positive Sera

The Han-Semi and Han-L PCR assay described in 2.1.4.4 were used to screen the IgM positive sera.

2.4. High-Throughput Sequencing of Undiagnosed Clinical Samples

2.4.1. Sample Collection

Surplus nucleic acid extractions, derived from samples that had been taken for diagnostic purposes and no longer required, were collected. These extractions were collected between 2015 and 2018, as part of ongoing collaborative research projects between Nottingham University Hospitals NHS Trust (NUH). All identifying information associated with these samples was removed prior to collection, and so were completely anonymised. The use of these samples for extended molecular diagnosis of infectious agents had been ethically approved prior to the start of this project.

The chosen samples were all PCR-negative using routine PCR-based viral diagnostics. NUH screening panels used an AusDiagnostic 'Viral CSF' assay (Ausdiagnostics, Chesham, United Kingdom) to screen CSF samples, this kit targets Human herpesvirus 1/2/3/4/5 and Enterovirus. The viral targets of the assays used to screen respiratory and blood samples were unknown.

No information was available regarding whether samples from each patient had also been referred for bacterial, fungal or parasitic diagnostics, and so infection with non-viral infectious agents was unknown.

Cerebrospinal fluid (CSF), EDTA-treated blood and respiratory samples were collected; respiratory samples included bronchoalveolar lavages, nasopharyngeal aspirates, endotracheal tube secretions and throat swabs.

2.4.2. Pooling of Samples

2.4.2.1. CSF Samples

CSF samples were pooled into groups of ten samples. Aliquots of twenty of these pools were then further combined into a larger, final pool. Two final pools, each containing two-hundred samples was created.

Due to the low nucleic acid concentration of the CSF samples, each of the pools of ten was consecutively loaded onto the column-based RNeasy MinElute Cleanup kit (Qiagen). The concentrated, membrane-bound RNA was eluted into the final pools, so as to exceed the minimum input of the library preparation protocol.

The samples were pooled in this way to facilitate the PCR screening of any potential positive results. The pools allowed fewer PCR reactions to be performed to determine which sample was positive, and thus conserving nucleic acid.

2.4.2.2. EDTA Blood Samples

Samples were grouped into pools of ten. Aliquots of Eight of these pools were then pooled again, into a final pool of 80 samples.

2.4.2.3. Respiratory Samples

Samples were grouped into pools of ten, and aliquots of ten of these pools were combined again into a larger pool of one hundred. One pool contained a total of one hundred samples was created, and a second pool contained 70 samples.

2.4.3. Tapestation Analyses of RNA

The final pooled samples were analysed on an Agilent TapeStation, using an RNA ScreenTape.

2.4.4. Library Preparation

Libraries were prepared using NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (New England Biolabs, Ipswich, Massachusetts, USA) and NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs), using the unmodified protocol. The total amount of RNA inputted for each sample, the RIN, the fragmentation times and the unique indexing primer used are shown in Table 9.

Sample	RNA Input (ng/µl)	RIN	Fragmentation Time (Minutes)	Index
CSF-1	136.8	4.5	8	1
CSF-2	25.2	4.5	8	2
EDTA-1	1000	4.5	8	3
Resp-1	444	4.5	8	4
Resp-2	1000	5.5	8	5

Table 9 Samples submitted for HTS library preparation. The RNA input, RIN, required fragmentation time and their unique index number are included.

2.4.5. Processing the HTS Datasets

2.4.5.1. Merging Paired Reads and Filtering Low-Quality Reads

The methods used to merge paired reads and filter low-quality reads are described in section 2.2.4.1.

2.4.5.2. Removal of Host Reads

The GRCh38 human reference genome (GCA_000001405.15) and transcripts were downloaded from GenBank.

The representative genome for *Bos Taurus* (GCA_002263795.2) was downloaded to filter the zebu reads from Resp-2.

The methods used to remove host genome sequences from the datasets are described in section 2.2.4.2

2.4.5.3. *De novo* Assembly

De novo assembly of filtered reads was done following the methods described in 2.2.4.3

2.4.6. BLAST Searching of HTS Datasets Against Virus Databases

The methods used are described in 2.2.6.

2.4.7. Mapping of Human Pegivirus Reads to a Reference Sequence

The reference genome for Human Pegivirus-1/GB Virus C (NC_001710) was downloaded from GenBank. Reads from each of the five datasets were mapped directly to this reference sequence, using the Geneious mapper tool in Geneious Prime. A medium-low

sensitivity and Iterative fine-tuning were used, with five iterations of mapping reads onto the consensus of the previous iteration.

2.4.8. PCR Screening of Samples for Human Pegivirus-1

2.4.8.1. Primers, PCR Reactions and Cycling Conditions

The reagents and concentrations used are described in 2.1.4.1.

The following primers were used to amplify a 291bp fragment of the NS5B gene;

GBV-C 8435 F: 5'- CCTGGGGAAATACTATGCCT-3'

GBV-C 8726 R: 5'-TGATGCATGATATGAGGGCT-3'

Cycling conditions were 94°C for 15 minutes, followed by 55 cycles of 94°C (30 seconds), 55°C (30 seconds) and 72°C (30 seconds), finished with 72°C (10 minutes).

The PCR products were visualised on an agarose gel, and sent for Sanger sequencing, using the methods described in 2.1.5 and 2.1.6.

2.4.8.2. Samples

Each of the five final pools was screened, to confirm the detection of HPgV-1. The constituent sub-pools of the positive pools were then screened. The individual samples from each of the positive sub-pools were subsequently screened to identify which samples were HPgV-1 positive.

3. Screening of Rodent Tissue for Viral Zoonoses Using Degenerate Primer RT-PCR

3.1. Introduction

Zoonoses are pathogens that are transmitted from an animal reservoir or intermediary, and into humans where they may result in disease. Viral zoonoses are of particular importance, as they are responsible for several severe pandemics and outbreaks, including the recent West African Ebola virus outbreak, the emergence of HIV from primates and numerous avian and swine influenza epidemics. Not all zoonotic transmission events result in severe outbreaks; the majority often result in a mild, sub-clinical illness which can be misdiagnosed as a more common virus or go undiagnosed completely. Rodents are one of the largest reservoirs of potentially zoonotic viruses, due in part to their population sizes, geographic ubiquity and proximity to human populations. Therefore, screening rodent populations for viral families with known zoonotic species will further the known distribution of these viruses.

There have been numerous techniques developed to detect the presence of viruses in samples. Molecular detection techniques such as Polymerase-Chain Reaction (PCR) are the most widely used method of virus detection currently. Its popularity is in part due to the low per-sample cost, simplicity of use and its adaptability to

almost all types of pathogens. Typically, PCR requires a known sequence to which the primers are designed to target, this can be adapted to improve detection using degenerate primers. When many divergent target sequences are used, there is unlikely to be a region with identical nucleotide sequence; by introducing degeneracies into the primer sequences, multiple sequences can be targeted with a single primer.

This chapter describes the processing and screening of rodent tissue from the United Kingdom, Poland and Egypt. Two-step reverse transcription PCR, with degenerate, pan-species primers were used to screen for known viral zoonoses, including orthohantavirus, orthobornavirus and picornavirus. This work resulted in the discovery of multiple strains of Tatenale orthohantavirus and four picornavirus species from four species of rodents.

3.2. Results

3.2.1. Collection and processing of rodent and small mammal tissue.

Cohorts of wild rodents were trapped and collected at four sites in the United Kingdom, and one site each in Poland and Egypt. Tissue samples were taken from 958 animals, with multiple organs sampled where available. Rodent species included brown rat (Rattus norvegicus), field vole (Microtus agrestis), bank vole (Myodes glareolus), wood mouse (Apodemus sylvaticus), house mouse (Mus musculus), red squirrel (Sciurus vulgaris) and Cairo Spiny Mouse (Acomys cahirinus).

RNA was extracted from tissue collected from the rodents, and cDNA synthesis using random hexamers. Spectrophotometry was used to assess the yield and quality of the extracted RNA. A PCR assay targeting the Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a housekeeping gene, was used to validate the quality of the extracted RNA and synthesised cDNA. GAPDH was successfully amplified in almost all samples, indicating that the cDNA tested was usable for viral PCR screening.

3.2.2. Analysis of the Polish Bank Vole Cohort

Extensive information was provided for the Polish bank vole cohort, allowing analysis into any associations between virus infection and host data.

Voles were categorised into 3 age brackets; juvenile, adolescent and mature. Across all capture sites, there were 110 juvenile (37.7%), 44 adolescent (15.1%) and 138 mature (47.3%) voles. The proportion of ages varied by site; juvenile voles were the majority in Urwitalt and Pilchy, whilst mature voles predominated in Talty (Figure 12). The number of voles by sex was almost identical overall, with a slight bias towards males (52.3%). Urwitalt and Talty had a larger proportion of males than Pilchy, where females made up the majority (Figure 13).

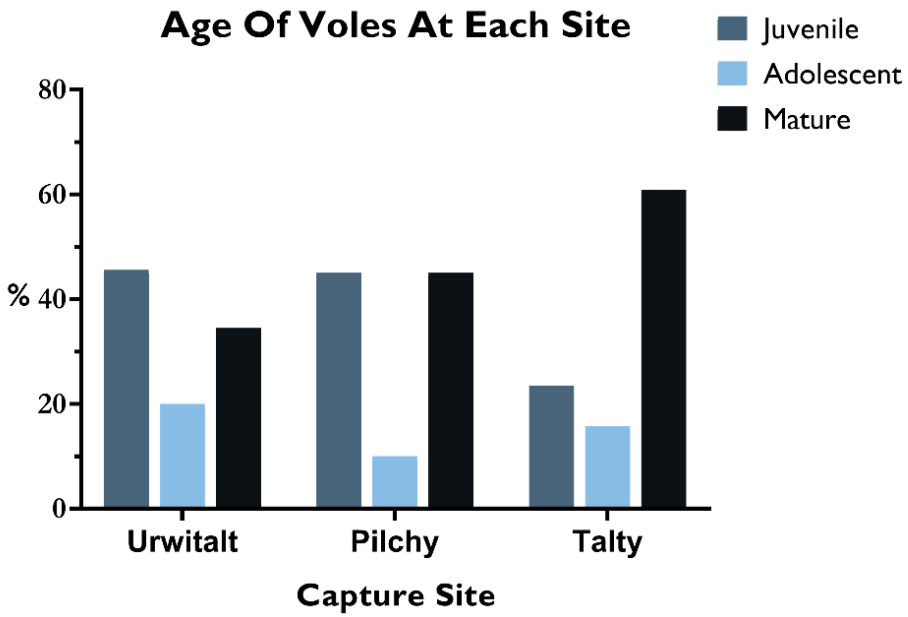

Figure 12 A Summary of the ages of bank voles that were captured at each of the three sites in Poland. The ages of the voles are categorised into three groups. Bars show the number of voles of each age category as a percentage of the total number of voles at each site.

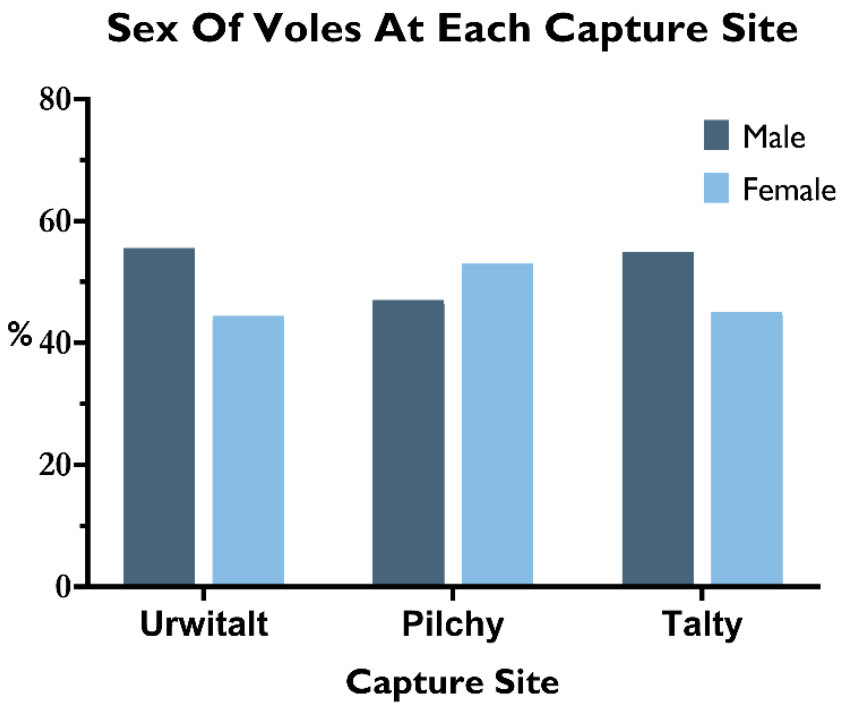

Figure 13 A summary of the sex of bank voles captured at each of the three sites in Poland. Bars show the number of voles of each sex as a percentage of the total number of voles at each site.

3.2.3. Detection of an orthohantavirus in field voles

3.2.3.1. Analysis of pan-orthohantavirus degenerate primer sets

Two pairs of degenerate primers were sourced from previous studies aiming to detect orthohantaviruses. These were a published primer pair, HAN-L, and a pair, HAN-Semi, that were designed in-house by a previous PhD student, Theocharis Tsoleridis. Both assays targeted the RNA Polymerase gene, on the L segment.

The HAN-L Primer pair were mapped onto the consensus sequence of an alignment of orthohantaviruses (Figure 14). The forward primer has degenerate nucleotides at position 6, 9 and 15; Position 6 and 15 degeneracies match the degeneracies on the consensus sequence, whilst position 9 requires an 'N' rather than a 'B' for full coverage. Additional degeneracies occur at positions 3, 12, 14, 15, 16, 18; however, the Sequence Logos show that these positions are a result of a variance in a low number, often single, of sequences. The degenerate nucleotides in positions 6, 9 and 12 of the reverse primer require an N for full coverage. Like the forward primer, there are degeneracies not covered by the primer which are due to minor variants. The sequences responsible for the extra degeneracy in the consensus sequence were all reported after this primer-pair were created, so at the time of design, they would have matched 100% to all known sequences.

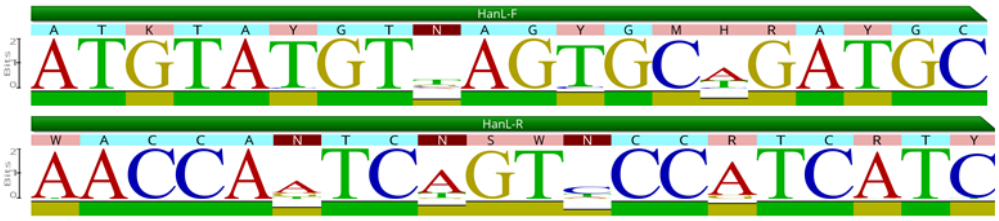

Figure 14 Sequence Logo graph of the HAN-L primer set mapped onto an alignment of orthohantavirus sequences

The Han-Semi primer set were mapped to the same alignment consensus (Figure 15). The degeneracies at positions 12, 15 and 18 in the forward primer match completely those in the consensus, whilst the 'D' at position 21 requires changing to 'N' to match all variants.

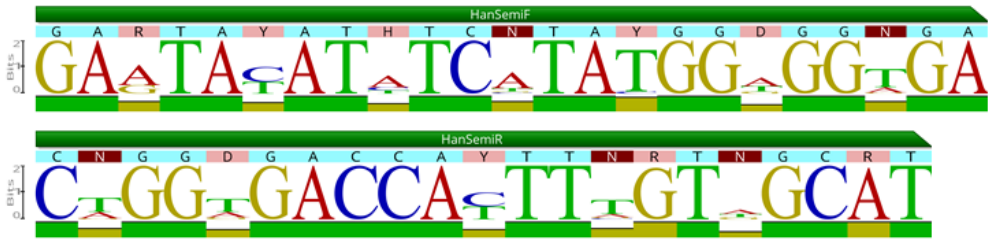

Figure 15 Sequence Logo graph of the HAN-Semi primer set mapped onto an alignment of orthohantavirus sequences

3.2.3.2. Optimisation and validation of orthohantavirus assay

Positive controls for the Pan-orthohantavirus PCR screening assay were originally sourced from a previous screening project within the research group. This control was a plasmid containing the complete coding sequence of the L segment of a Seoul Virus isolate.

The initial cycling parameters for the HAN-L primer PCR were an initial denaturation of 94°C for 2 minutes, followed by 55 cycles of 94°C for 15 seconds, a variable temprature annealing of 50 seconds and an extension of 68°C for 30 seconds, finished with a final extension of 68°C for 5 minutes [157]. Cycling optimisation was performed using a step-wise annealing temperature gradient, ranging from 49.7°C to 68.2°C, with a plasmid with a copy number of 10^4 was used as a positive. Bands were observed until 64°C, with the strongest bands at 54°C.

A second PCR to determine the sensitivity of the assay was performed. Samples containing plasmid copy numbers of 10^0, 10^1, 10^2, 10^3, 10^4, 10^5, 10^6 and 10^8. The assay produced detectable bands for input template copy number as low as 10^2. To simulate a positive sample from an infected animal, bank vole liver RNA that had been screened and subsequently orthohantavirus-negative, was spiked with SEOV 10^1, an annealing temperature of 60°C was used, to reduce potential mispriming. Weak bands were observed in 2 of the 3 spiked samples. Taking into account the lower concentration of positive control material, a second annealing temperature gradient was performed on a SEOV 10^1 plasmid, annealing temperatures of between 49.1°C and 60.9°C were used. Bands were observed at all temperatures apart from 60.9°C, which showed that annealing temperatures of 60°C and higher were unsuitable for samples with a low concentration of target material. Annealing temperatures were reduced to 54°C and the spiked-sample PCR assay repeated; strong bands were observed in all spiked samples.

The HAN-Semi assay was sourced after the screening of animal tissue had successfully detected Orthohantavirus, and so all optimisation was performed on positive-tissue derived RNA, rather than plasmids. Initial screening-detected orthohantavirus without the need for an annealing temperature gradient.

Screening serially-diluted orthohantavirus positive kidney and lung tissue with the optimised cycling parameters showed that HAN-L could only detect viral RNA at undiluted concentrations in the

kidney, and a dilution of 10^{-1} in the lungs. In contrast, HAN-Semi could detect RNA at dilutions of 10^{-2} and 10^{-3} in the kidney and lungs; indicating improved sensitivity

3.2.3.3. Detection of orthohantavirus in field vole tissue

Initial screening of the lungs and kidneys of the Twycross cohort used the HAN-L primer assay. The mouse and rat samples were all negative for the presence of orthohantavirus viral RNA. Both the lung and kidney tissue of the field vole were positive for orthohantavirus, tissue was rescreened using HAN-Semi to confirm the result. sequencing and BLAST analysis showed that the sample was closely related to Tatenale orthohantavirus.

For subsequent screening assays, we used the Han-Semi assay, due to its improved sensitivity compared to the HAN-L assay. Screening of the Chester zoo cohort revealed that all wood mice, brown rats and bank voles were orthohantavirus negative. However, 12 Field voles (10%) were positive for orthohantavirus; these results were confirmed with the larger HAN-L assay. Sequencing of the larger HAN-L amplicon revealed a Tatenale orthohantavirus, of a different strain to that isolated from Twycross.

Screening of the red squirrels from Jersey, Cairo spiny mice from Egypt and the field voles from Sutton Bonnington did not detect the presence of orthohantavirus

The bank voles from the Polish cohort were not screened, as they had been previously screened two separate occasions, as part of a previous MSc. projects and were all found to be orthohantavirus negative.

3.2.3.4. Analysis of the HAN-L Tatenale orthohantavirus Amplicon

The 452bp HAN-L assay derived sequences recovered from the Chester zoo field voles were highly conserved, with 5 of the sequences being completely identical; the most divergent two viruses were 97.69% identical at the nucleotide level. The Twycross zoo strain was genetically distinct from the Chester zoo strains, with a nucleotide similarity of 86.3-89.3%. When compared to previous strains, B41 and Kielder-1/-2, the Chester strains were 94.9-96.9% identical to B1 and 84-86.6% identical to both Kielder strains. The Twycross strain shared a similar similarity to B41 as it did with the Chester strains, 87.3% and an 86.5-86.8% similarity to both Kielder strains.

To reflect the geographic origin of these viruses, the Chester zoo strains were named 'Upton-Heath', whilst the Twycross zoo strain was named 'Norton-Juxta'.

The Norton-Juxta and an Upton-Heath HAN-L amplicon were aligned with all other complete vole associated orthohantavirus L sequences. The 3 published HAN-L derived Tatenale virus sequences were also added to this alignment. A Maximum-Likelihood phylogenetic tree was generated from this alignment. Norton-Juxta and Upton-Heath both clustered with the previous Tatenale strains,

Upton-Heath with B41 and Norton-Juxta forming a separate clade equidistant from B41 and Kielder (Figure 16). The Tatenale viruses form a discrete clade; the closest related orthohantaviruses are Tula virus and Prospect Hill virus.

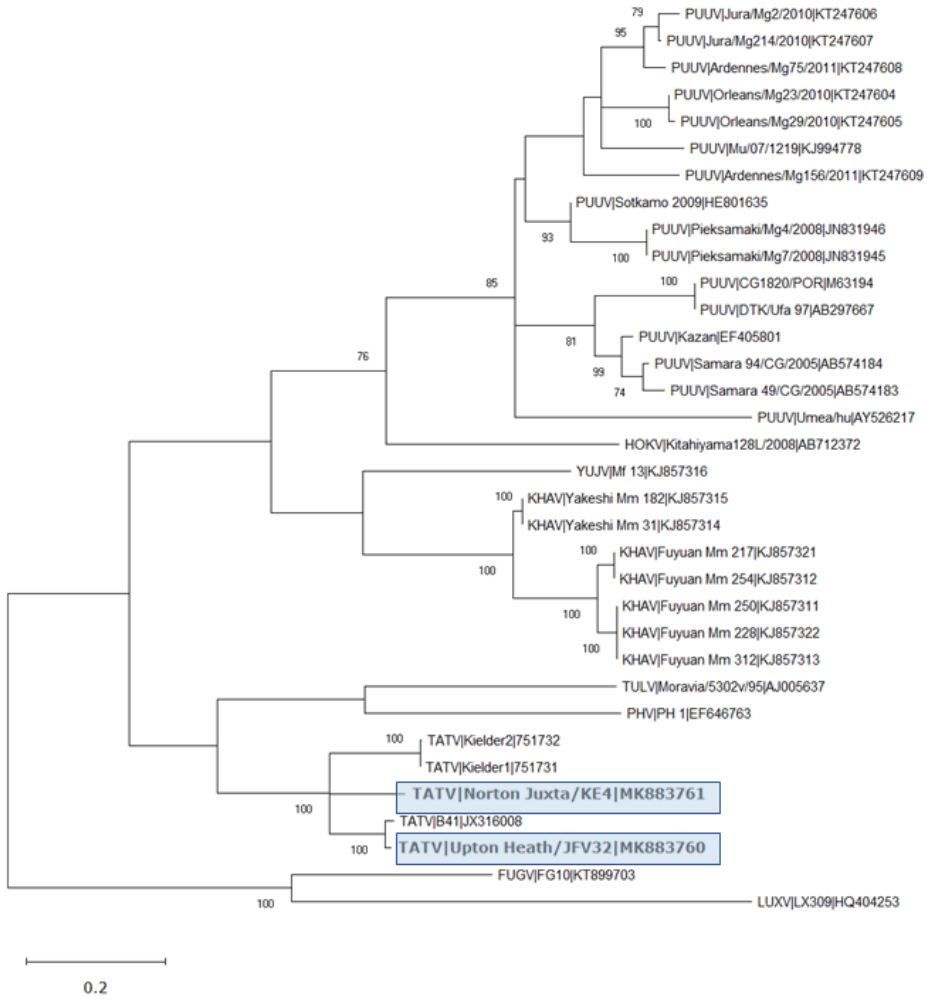

Figure 16 Phylogenetic relationship of Tatenale virus with previous Tatenale isolates and other vole-associated orthohantavirus species. Representative partial sequences were obtained for the L segments. Maximum Likelihood trees were created using the T92+G+I model, branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 100 replicates. Tatenale virus strains are highlighted in boldface and a blue box. Sequences are shown with species name, strain name and the GenBank accession number. PUUV, Puumala virus; HOKV, Hokkaido virus; FUSV, Fusong virus; YUJV, Yuanjiang virus; KHAV, Khabarovsk virus; TATV, Tatenale virus; PHV, Prospect Hill virus; ILV, Isla Vista virus; TULV, Tula virus; ADLV, Adler virus; LUXV, Luxi virus; FUGV, Fugong virus.

3.2.4. Screening of Rodents for Bornavirus

3.2.4.1. Design and selection of a Bornavirus assay

Three pairs of PCR primers of varying degeneracy were designed to amplify a conserved region of the glycoprotein gene of mammalian orthobornaviruses. Both primers were designed on the same positions of the gene, with one primer pair (D-Bor) being completely degenerate to match all sequences, the second (SD-Bor) using fewer degenerate bases at the more conserved positions and the third (VSBV-1) was specific to Mammalian 2 orthobunyavirus (Variegated squirrel bornavirus)

At the time of design (2015), there were 15 complete Bornavirus genomes available on GenBank; 13 Mammalian 1 orthobornavirus and 2 mammalian 2 orthobornavirus sequences.

The D-Bor primers were completely degenerate at all variable positions, with a total of 6 degeneracies in the forward primer, and 5 in the reverse. The SD-Bor primers contained fewer degeneracies, 3 in both primers. The VSBV-1 pair contained no degeneracies.

3.2.4.2. Optimisation and Validation of Pan-orthobornavirus primers.

A plasmid encoding a Mammalian 2 orthobornavirus glycoprotein gene was produced and used as a positive control for the orthobornavirus screening assay.

To establish the sensitivity of the primers, serial dilutions of the glycoprotein-containing plasmid were created. Neat plasmid and

dilutions of 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, and 10^{-10} were screened; D-Bor and SD-Bor were able to amplify the target down to a dilution of 10^{-6}, whilst the VSBV-1 primers could amplify dilutions of 10^{-8}.

To determine an optimal annealing temperature for the D-Bor, SD-Bor and VSBV-1 primers, an annealing temperature gradient PCR was performed. Annealing temperatures used were 55.2°C, 56.6°C, 58.5°C, 60.8°C, 63.5°C, 65.8°C, 67.5°C and 68.8°C, and a plasmid diluted 10^{-4} was used as a positive. The D-Bor primers were able to amplify the target at annealing temperatures up to 60.8°C, whilst the SD-Bor and VSBV-1 primers amplified a product at all annealing temperatures.

Three Rat kidney cDNA samples were spiked with the 10^{-4}, 10^{-6} and 10^{-8} dilutions of the glycoprotein plasmid, and screened with each of the three orthobornavirus primers. The D-Bor primers were only able to amplify one of the three spiked samples at an annealing temperature of 55°C. The SD-Bor primers could consistently amplify the spiked cDNA in each dilution of the three samples, at annealing temperatures of 55°C, 60°C and 63°C. At 65°C annealing, however, it could only amplify the 10^{-4} and 10^{-6} dilutions. The VSBV-1 primers were able to amplify each of the spiked samples with all three dilutions at 63°C annealing (The only annealing temperature tested).

3.2.4.3. Screening rodent tissue

Lung and kidney extracts from the rodents captured at Twycross Zoo were firstly screened using both the D-Bor and SD-Bor primers. All

samples screened were PCR-negative for orthobunyavirus. The liver samples from the Chester Zoo, Poland and Egypt cohorts, and the spleen samples from the Jersey cohort were screened with only the SD-Bor primer assay. All samples tested negative for orthobunyavirus.

3.2.5. Screening of Rodents for Parechovirus

3.2.5.1. Pan-Parechovirus assay

A Pan-Parechovirus assay was sourced from Nix et al [218], which targets the 5' untranslated region of all known (at the time of publication) parechoviruses.

3.2.5.2. Optimisation and validation of the pan-picornavirus assay

Human Parechovirus (HPeV) samples were sourced from a repository of surplus clinical diagnostic nucleic acids and used as positive controls for the Parechovirus screening assay.

An Annealing temperature gradient was performed on an HPeV positive clinical sample, to determine the optimal annealing temperature. Annealing temperatures of 54.2°C to 66.9°C were used; the assay successfully amplified the target at all annealing temperatures.

3.2.5.3. Detection of multiple picornaviruses in Rodent tissue

Screening of the Twycross cohort yielded PCR-products corresponding to the expected size of parechovirus amplicon in the

kidneys of 5 brown rats and 6 house mice. However, sequencing showed that these were PCR products generated from other *Picornaviridae* genera. The rat samples were a different *genus* of *Picornaviridae*, cardiovirus. These amplicons were most closely related to the Boone cardiovirus strain (JQ864242.1), with a similarity of 96%. The Mouse samples contained a second *picornaviridae* genus, hunnivirus. These were moderately divergent from their closest related sequences; with 76% similarity to strain 05VZ-75-RAT099 (KT944214).

In the Polish cohort of bank voles, 97 samples were PCR positive; sequencing of the amplicons showed that the samples were a mix of Ljungan virus (Parechovirus-B) and rosavirus, a further *Picornaviridiae genus*. To differentiate the rosavirus and Ljungan virus positives, a rosavirus specific assay was used to rescreen all positive samples. In total there were 80 Ljungan virus-positive voles (27.2%) and 17 rosavirus positives (6.14%). All Parechovirus and rosavirus negative samples were GAPDH positive.

No further parechovirus or similar picornaviruses were detected with this assay in any of the other rodent cohorts.

3.2.5.4. Analysis of Parechovirus

The overall prevalence of Ljungan Virus (LV) was 27.2%. Analysis of the relationship between LV infection and host attributes showed several statistically significant correlations. Prevalence of LV in voles from Tałty (18.3%) was significantly lower (P=0.005) than those from

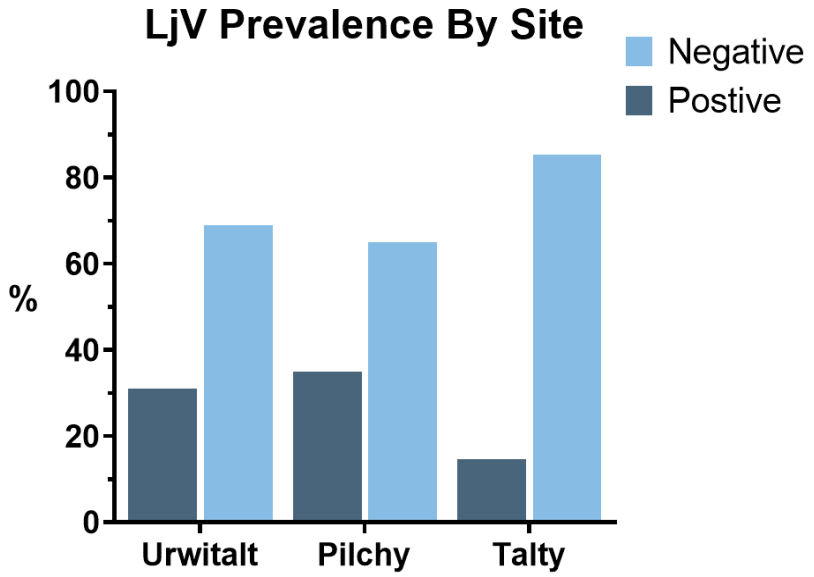

Figure 17 Summary of the PCR prevalence of Ljungan virus in bank voles captured at each of three sites in Poland.

the Pilchy (42.7%) and Uriwitalt (39%) sites (**Figure 17**). Prevalence of LV was also significantly associated with the age of voles. Juvenile and adolescent voles across all sites were 34.5% and 31.8% LV positive, respectively. Mature voles, however, were only 18.1% LV positive, a statistically significant difference (P=0.004) (**Figure 18**).

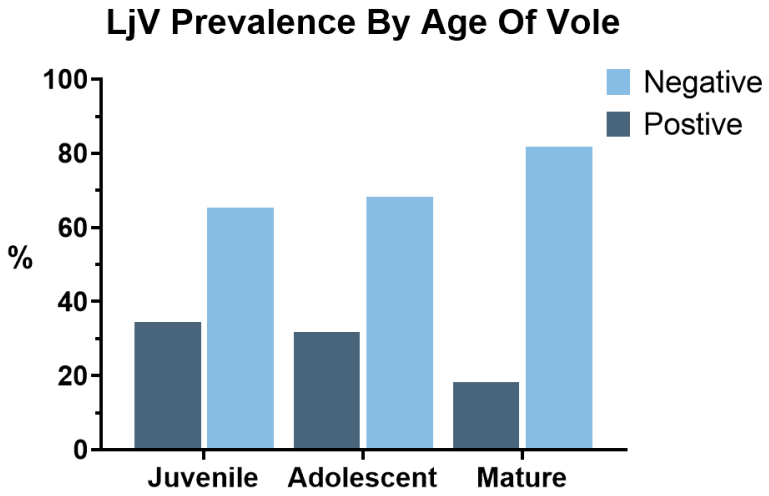

Figure 18 The PCR prevalence of Ljungan virus in bank voles of different age groups.

A Phylogenetic tree created with several qPAR amplicons shows that they cluster into three clades, which suggests that they may belong to different genotypes (**Figure 19**). PMgE289 and PMgE293 were both captured at the Talty site, whilst PMgE2, PMgE13 and PMgE31 were captured at Urwitałt. The Talty sequences cluster into one clade, whilst the Urwitałt sequences cluster into two other clades, which would suggest that these genotypes are more prevalent in different geographic locations.

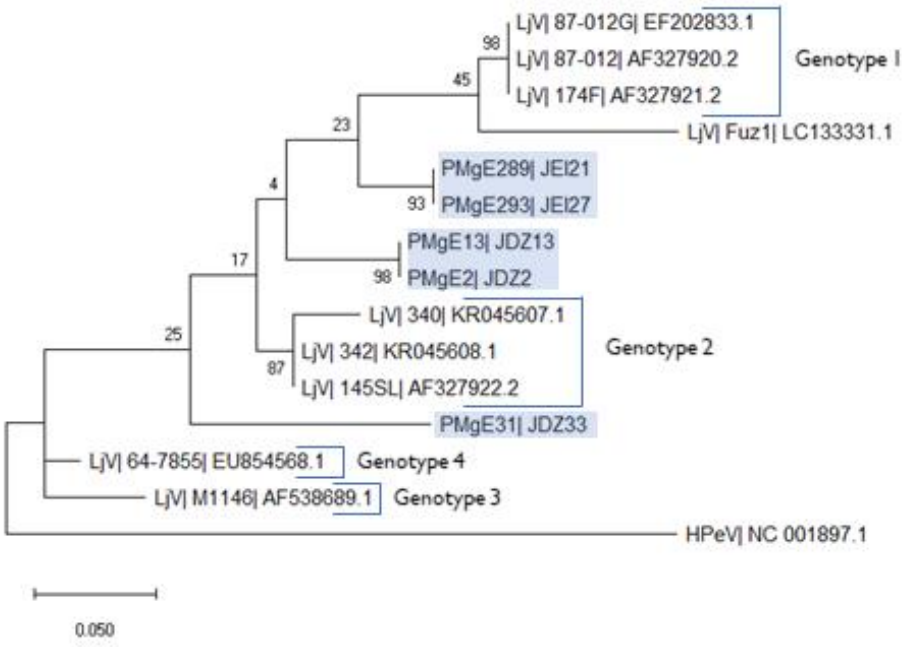

Figure 19 Phylogenetic tree of the qPAR LV amplicons with previous LV isolates. Maximum Likelihood trees were created using the K2+G model, branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 100 replicates. LV sequences from this study are highlighted with a blue box. Sequences are shown with species name, strain name and the GenBank accession number. HPeV, Human parechovirus.

Due to the short sequence length and high homology between the sequences, the generated phylogenetic tree has low bootstrapping probability scores, and cannot be used to draw any definite conclusions regarding the phylogeny of these viruses.

Primers were then designed to retrieve further sequence of the Ljungan Virus (LV), particularly coding sequence. Initial primers were designed to target a 142bp section of VP0 and a 537bp section of the VP1 gene, for both human parechovirus and Ljungan virus. Optimisation using annealing temperature gradients was performed on these primers, but they failed to amplify the target sequence. LV specific versions of these primers were then designed to improve the specificity, but again they failed to amplify the target. A published group of primers were taken from Nix *et al* [221] and had the annealing temperatures optimised; these again failed to amplify any target. Several other primers were designed targeting different genes (Table 2), all of which failed to amplify any further sequence.

3.2.5.5. Analysis of rosavirus

A Total of 17 (6.14%) bank voles were positive for Rosa virus. Urwitalt had the most infected voles (n=8), followed by Pilchy (N=6) and Talty (N=3), however, this difference in virus prevalence was not statistically significant (**Figure 20**).

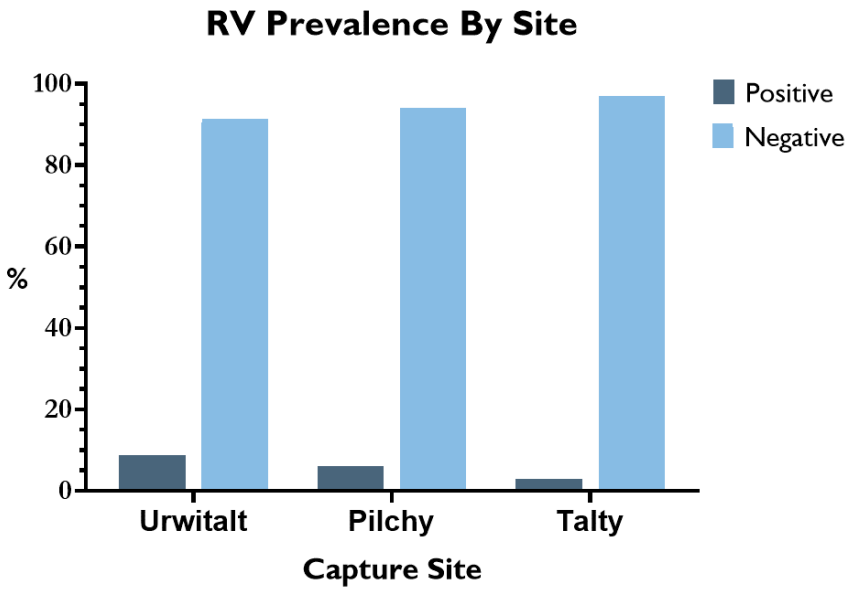

Figure 20 Prevalence of Rosavirus in bank voles captured at each of the three sites in Poland

rosavirus prevalence was similar between voles in each age bracket; the lowest prevalence was in mature voles (5%), with prevalence increasing to 6.3% in juvenile voles and 6.9% in adolescents (**Figure 21**).

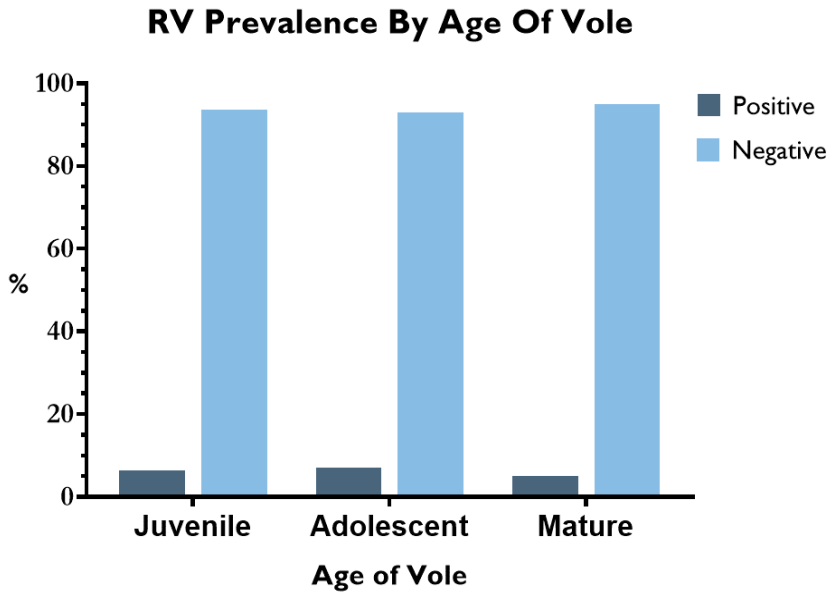

Figure 21 Prevalence of Rosavirus in bank voles of different ages.

Phylogenetic analyses of the 330bp amplicon of rosavirus 3DPol shows that the two sequenced samples, 'PMgE31' and 'PMgE2' both cluster with the rosavirus A species, specifically rosavirus A1 (The type species) (**Figure 22**).

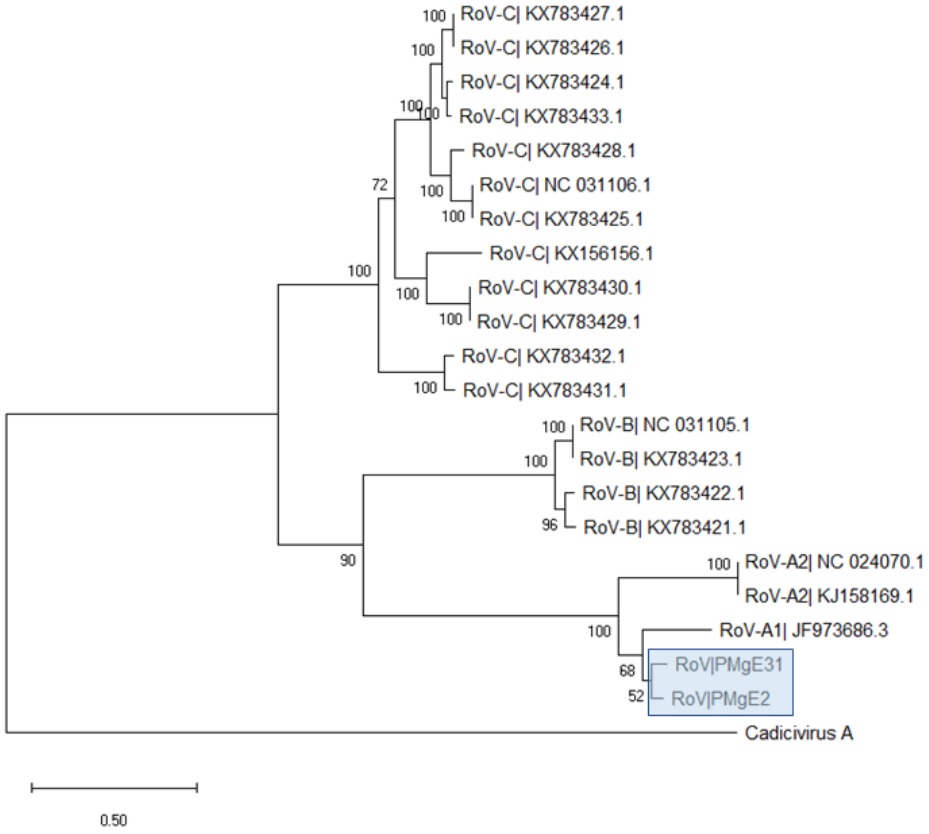

Figure 22 Phylogenetic tree of the RoV. Maximum Likelihood trees were created using the K2+G model, branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 100 replicates. RoV sequences from this study are highlighted with a blue box. Sequences are shown with species name and the GenBank accession number.

3.2.5.6. Analysis of cardiovirus

Five rat kidneys were positive for cardiovirus (7%). Two of the five were sequenced, the closest related species was Boone cardiovirus (JQ864242.1), with an identity of 96.6%. Phylogenetic analysis of the small fragment of the sequence also supports this, with the two sequences clustering with Boone cardiovirus in the cardiovirus C clade (**Figure 23**)

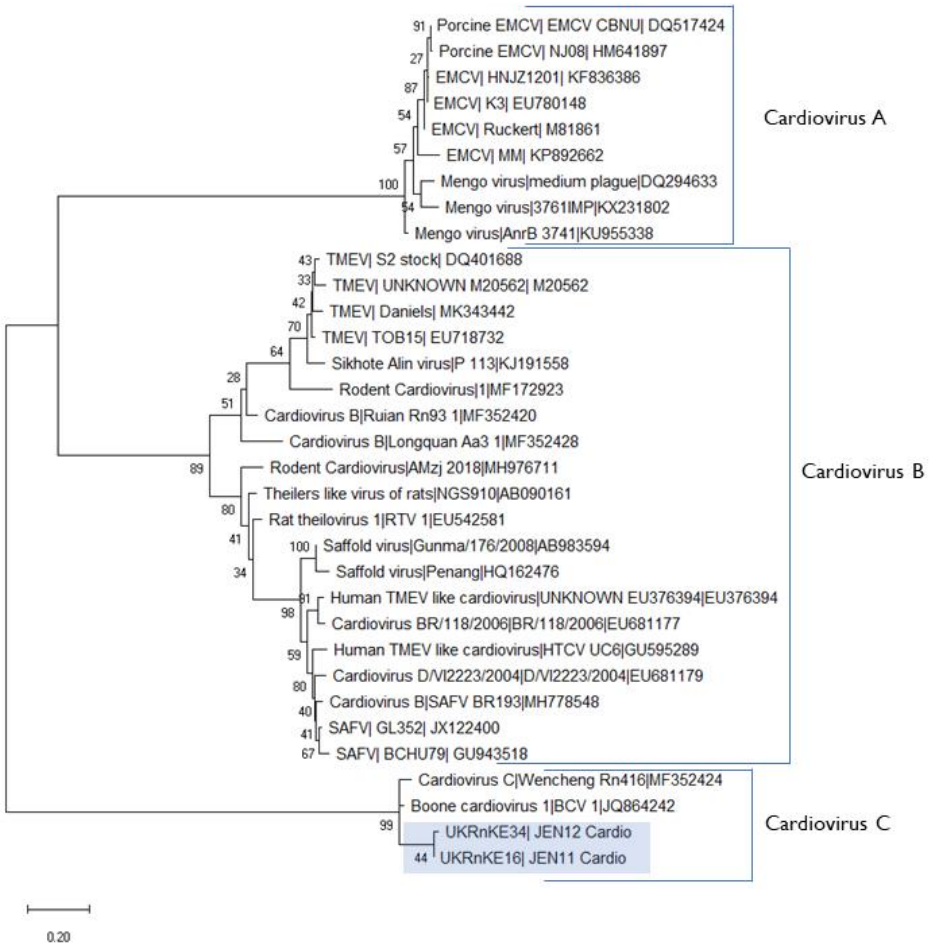

Figure 23 Phylogenetic tree of qPAR CaV amplicons, and published sequences. Maximum Likelihood trees were created using the K2+G model, branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 100 replicates. CaV sequences from this study are highlighted with a blue box. Sequences are shown with species name, strain name and the GenBank accession number.

3.2.5.7. Analysis of the hunnivirus

Six of the house mice were positive for HuV. A phylogenetic tree of the qPAR amplicon shows that the mouse-derived HuV broadly clusters with other rodent-derived hunniviruses, though it forms a novel clade within this group.

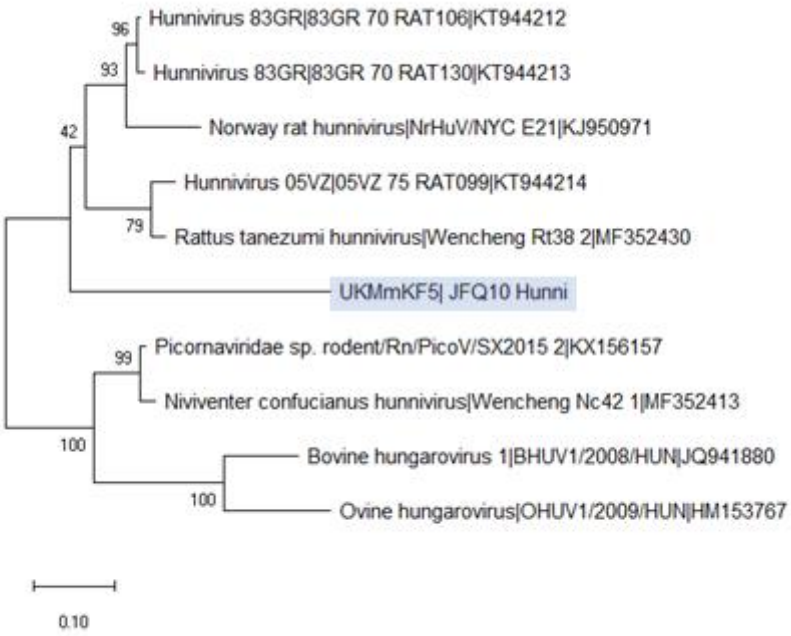

Figure 24 Phylogenetic tree of qPAR amplified HuV sequences, and published HuV sequences. Maximum Likelihood trees were created using the K2+G model, branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 100 replicates. HuV sequences from this study are highlighted with a blue box. Sequences are shown with species name, strain name and the GenBank accession number.

3.3. Discussion

3.3.1. Hantavirus

3.3.1.1. Primers

The Primers that were initially chosen, HAN-L were designed by Klempa et al [217] in 2006 and were subsequently adopted in numerous studies of orthohantavirus prevalence in rodents [222], shrews [223], bats [224], moles [225] and humans [226]. Since then, the diversity of orthohantaviruses has increased with the discovery of novel species, and it becoming evident that these primers may not detect all hantavirus strains.

HAN-L primers have been used to detect Seoul virus in British rats [164], this particular strain would likely be highly similar to any circulating within the Rats screened in this study and so would also likely be detected if present. To date, no orthohantavirus has been detected in house mice, field mice, spiny mice or red squirrels; as such it is impossible to say whether any orthohantaviruses would be detected by the primers used in this study.

3.3.1.2. Detection of TATV

An orthohantavirus closely related to Tatenale orthohantavirus was detected in a single field vole from the Twycross zoo cohort and field voles from the Chester zoo cohort. Tatenale orthohantavirus has been detected twice, in Chester in 2013 [178] and the Scottish-English border region in 2017 [179].

Prevalence of Tatenale virus in the voles from the Chester cohort was 10%; greater than the previously reported prevalence of B41 in field voles that had been captured at the village of Tattenhall, 8.4 miles away. Given the degree of similarity of the viruses from Chester and Tattenhall, this variance in prevalence is unlikely to be a result of a genetic factor such as increased transmissibility. The Field voles captured at Chester zoo were predominately collected between October 2013 and November 2014 (with a small number caught in late 2016 and 2015), whilst those screened in Pounder *et al* were captured between 2009 and 2011. Puumala orthohantavirus, a closely related species has a well-studied relationship between the population dynamics of its reservoir, the Bank Vole (*Myodes glareolus*), and Puumala virus infection. Bank Voles undergo a resource-dependent temporal population density cycle, during which Puumala virus prevalence fluctuates and is at its highest during years of increasing or peak vole density [227]. Field vole populations in Northern England exhibit a similar multi-annual cyclic pattern, these fluctuations are irregular with cycles of 3 or 4 years being described [228]. These cyclic population dynamic have also been reported in Scotland and Southern England [229]; both exhibit a regular 3-4 year cycle. Tatenale virus likely exhibits a similar pattern of prevalence following the increasing and decreasing densities of the Field vole population, which could be an explanation for the variation in the prevalence between the populations screened here and previously. Thomason *et al* [179] reported a much higher prevalence of 16.7% in the Kielder Forest populations. These voles were captured in 2015, as were the

Chester Zoo cohort, this would mean that there is unlikely to be any significant differences between the population densities as described above. The Kielder strains of Tatenale virus are genetically distinct from the Upton-Heath (and B41) strains, which could result in this difference in prevalence; for example, Kielder could have greater potential for transmissibility. The geography of the sample sites also differs; Chester Zoo is situated amongst agricultural and urban land, whereas the Kielder sites consisted of fragmented grassy clearings surrounded by heavily forested areas. Therefore, access to food, levels of predation and breeding site availability will differ between the two sites, affecting population densities and therefore change the probability of transmission.

The previous studies both used the same degenerate pan-hantavirus primer set used in this study (HAN-L), this 452bp fragment of the polymerase, and a small fragment of the S-segment were the only sequences that were published and as such means that there is limited genetic information regarding this virus. The second degenerate primer set, HAN-Semi, was targeted at a region upstream of the HAN-L target. There is no published sequence available for this region of B41 or Kielder, which prevents any comparison of these regions of the polymerase.

The Upton-Heath strain, detected in voles from the Chester zoo cohort, were highly similar to the B41 strain, also detected in field voles from Chester. Given the close geographic proximity and highly similar sequences, it is likely that both viruses are of the same strain.

The Norton-Juxta strain is divergent from both the B41/Upton-Heath and the Kielder strains, indicating that is a third, novel strain of Tatenale orthohantavirus.

Full-length sequence from multiple segments would allow a more accurate phylogeny to be analysed. Accepted species demarcation of orthohantavirus requires full-length coding sequence of both the S and M segments; a 7% or greater amino acid difference in both the S and M segments [230], or stricter criteria of greater than 10% AA identity in the S and greater than 12% in the M segment [231]. As PCR assays failed to recover any sequence data from the S or M segments, Tatenale virus remains unable to be accurately speciated

This is only the third reported discovery of Tatenale virus, and so it remains unclear whether the field vole represents the definitive reservoir of this virus. The accepted dogma of hantaviruses being associated with a single host reservoir would suggest that the field vole is indeed the primary reservoir. However, the closely related Tula virus has been known to infect multiple Microtus hosts, which would suggest that Tatenale may also have a capability for host promiscuity. Though as the geographic range of field voles do not overlap with the only other UK *Microtus* species (European common vole, *Microtus arvalis*) this is unlikely.

As field voles are distributed throughout mainland Great Britain and can be found in both rural and urban environments, there will likely be many opportunities for transmission of a hantavirus between the

reservoirs and human, if Tatenale virus is capable of transmission into humans.

3.3.2. Bornavirus

3.3.2.1. No Detection of orthobornavirus

There was no detectable presence of orthobornavirus RNA in any of the rodent tissue samples screened during the study. Given the scarcity in the reports of Bornavirus infection in the UK and the comparatively limited sample size used it is perhaps unsurprising that no orthobornavirus has been detected so far. Additionally, naturally occurring Bornavirus infection has yet to be reported in any of the species screened. Successful experimental infection of rats with BoDV-1 has been reported on several occasions, which demonstrates that they are susceptible to orthobornavirus infection and may act as a potential if improbable reservoir. Infection studies have shown that levels of detectable orthobornavirus RNA in experimentally infected immunocompromised rodents are relatively high, particularly in the blood. However, these levels were found to significantly drop in animals with functioning immune systems, which would suggest that the presumably healthy rodents used in this study would have reduced viremia. Experimental infection of mice with BoDV-1 have reported the complete absence of detectable BoDV-1 in blood, which again suggests little to no viremia in immune-competent rodents. As these viruses are primarily neurotrophic, brain tissue or cerebrospinal fluid would be the most suitable sample to screen. Brain tissue was only available for a small minority of the rodents in the Twycross Zoo

cohort and none of the other cohorts. Due to the Twycross Zoo samples being snap-trapped for pest-control, the samples had potentially been dead for several hours or even days before being collected by staff, and therefore the brain tissue had substantially degraded; RNA extraction was attempted on four samples, but as predicted, the tissue was poor quality and resulted in a negligible yield and poor quality of the RNA extracted.

In a separate study, VSBV-1 was recovered from multiple organs of asymptomatic variegated squirrel; RT-qPCR showed a greater viral load in the kidneys and a similar viral load in the lungs than the brains of these squirrels. This would suggest that kidneys and lungs are suitable organs to screen for infection with VSBV-1 or related orthobornaviruses and that had any of the rodents been infected, it would have been detected using the RT-PCR screening in this study.

This is perhaps a defining distinction in the tropism of shrew-associated orthobornaviruses like BoDV-1 and rodent-associated orthobornavirus like VSBV-1. It is likely that if any orthobornaviruses were present in the rodents screened, that they would share the same tissue tropism as VSBV-1 and that the lung and kidney tissues were indeed suitable targets.

Given the limitations in the sample types and sizes used in this study, particularly in the British rodent cohorts, it is premature to rule out the possibility of orthobornaviruses circulating in the United Kingdom. Further studies would benefit from using live-caught and euthanised rodents to limit the degradation of tissue samples.

Expanding from degenerate PCR screening, seroprevalence studies would be incredibly useful to investigate past infections with orthobornavirus.

3.3.3. Picornavirus

3.3.3.1. Selection of Screening Primers

Four genera of *Picornaviridae* were detected in the rodents screened in this study; *Parechovirus, rosavirus, cardiovirus* and *hunnivirus*. The primers selected from Nix *et al* [218] were designed to amplify only *Parechovirus A* and *Parechovirus B*, however. Nix *et al* had screened multiple picornaviruses to determine the specificity of the primers, these included 2 cardiovirus species (encephalomyocarditis virus and Theiler's murine encephalomyelitis virus) and 4 enteroviruses (coxsackievirus A16, echovirus 30, coxsackievirus A24 and enterovirus 68); it is unclear why the primers had failed to amplify cardio viruses in the previous study but had amplified them here.

The primers were designed to target a section of the 5' untranslated region (UTR) which encodes an internal ribosomal entry site (IRES), which are highly conserved. picornaviruses have one of five classes of IRES [232]; *Parechovirus* [233]*, rosavirus* [234]*, cardiovirus* [235] and *hunnivirus* [236] all have type II IRES sites. The similarity between each species IRES would explain why these primers were amplifying non-parechovirus picornaviruses.

3.3.3.2. Detection of Ljungan Virus

This is the first detection of Ljungan virus in Poland, and Eastern Europe in general; expanding it's known range in Europe from Fennoscandia, Denmark [237] and Southern Europe [138]. The three sampling sites used are geographically isolated from one another, Talty and Uriwalt are 6 linear kilometres from one another, whilst Uriwalt and Pilchy are 21 kilometres apart, each of these sites is also separated by lakes and motorways. As direct interactions between voles from each of these sites is likely constrained, but LV prevalence remains consistent between at least two of the sites, it would suggest that LV is well established in the region and likely to be prevalent in other vole populations. The Talty site had a significantly higher proportion of samples from mature voles; mature voles have a significantly lower prevalence of LV than juvenile specimens, this would likely explain the lower prevalence of LV infection at the Talty site.

Little is known of interactions between LV and wild bank vole reservoirs. The lower prevalence of LV infection in mature voles would suggest that infection may not be persistent, with the voles clearing infection before reaching adulthood. This is similar to human parechoviruses, in which most infections occur before the age of 2 and are subsequently cleared [238]. Fevola et al [239] recently reported similar findings, lower LV prevalence in mature voles, whilst studying multiple populations of voles in Finland. They also found a

significant difference in LV prevalence between males and females, with males having a higher LV prevalence.

The sequenced isolates cluster into 3 genotypes, 2 of which appear to be novel. As these sequences are only a small fragment from a single region of the genome, the accuracy of the phylogenetic tree cannot be determined. Recombination events have been described in both HPeV [126] and LV [240], this further emphasises the need for sequencing of other regions, as different parts of the genome may have different evolutionary origins.

3.3.3.3. Detection of rosavirus in Polish Bank Voles

RV has been detected in 6.14% of screened bank voles. The RV genus is primarily associated with rodent reservoir hosts, including mice [234] and multiple species of rats [241] The RV detected in this study appears to be closely related to rosavirus A species; the closest related strain is M-7, which was isolated from the faeces of *Peromyscus crinitus* in 2010. All RV strains detected so far have neem detected in the USA [242], Asia [241], Africa [243] and Hungary [244]. These findings represent the first detection of a rosavirus in Poland and the first detection in a species of vole. Given that RV was detected in multiple samples, it is likely that bank voles are a true reservoir and the positive voles were not spillover infections from another source.

M-7 strain of RV-A has yet to be associated with any disease in host species, and there is no evidence to suggest that the RV detected in this study had any pathogenic effect on the bank vole hosts. A related

strain, rosavirus A2, has been detected in the faeces of children with diarrhoea [243], though a causative link between virus and disease wasn't established. These viruses may be rodent-borne like other rosavirus', and the discovery in humans represent a spillover event; the low prevalence of RV-A2 in the human cohorts (0.5%) would suggest that humans are not a primary reservoir for these viruses.

There has been no research into the population dynamics of natural RV infection in rodents. There was no significant difference between the prevalence of RV in different ages of voles, as there was in LV infection. This may be explained by persistent, chronic infection in the voles, though more research would be required to validate this hypothesis.

3.3.3.4. Detection of cardiovirus in British Brown Rats

The *cardiovirus* genus contains three species, with a number of these having a rodent reservoir, such as Theiler's Murine Encephalomyelitis Virus (TMEV) in mice [245] and Boone cardiovirus in rats [242]. These viruses are known to be pathogenic in their host species, causing neurological disease [246]. Importantly, several viruses closely related to the rodent-borne viruses are thought to be pathogenic in humans, for example, Saffold virus [247] is closely related to TMEV [248], and has been detected in children with respiratory [249] and gastrointestinal disease [250]. cardioviruses have also been implicated in cases of meningitis and sudden infant death syndrome [251].

The cardioviruses detected in this study were most closely related to cardiovirus C, Boone cardiovirus (BcV). Boone cardiovirus was first isolated in 2012 [252] and again in 2014 [242] in brown rats. It is unknown whether BcV is capable of zoonotic transmission into humans or even if it is pathogenic within its rat hosts. This is the first isolation of a cardiovirus in the UK and the first isolation of Boone cardiovirus outside of the USA.

3.3.3.5. Detection of hunnivirus in British House Mice

hunnivirus, formerly Hungarovirus, has been detected in cows, sheep [253], rats [242] and cats [254]. They were originally detected in Hungary, and later in Northern Ireland, USA, China, Turkey and Vietnam. The sequences detected in this study represent the first detection of a HuV in a house mouse and the first detection of HuV in England. This expands the known geographic distribution of these viruses and further widens the known range of hosts. It is unknown whether house mice constitute a true reservoir for HuV or are merely a dead-end host; further screening and detection of HuV in *M. musculus* samples and screening known hosts from the same locations, without detection of HuV from the same locations would strengthen the evidence of M. musculus as a primary reservoir.

To date, there is no evidence associating hunnivirus with any pathogenicity within their host reservoirs. Lu et al [254] detected a novel hunnivirus in single diarrheic cat and suggested a causal link between the two, however, as the virus was detected in faeces it is

possible that this was not a true infection, but rather a dietary contaminant from predation on an infected rodent. Additionally, there has been no evidence suggesting any capability of zoonotic transmission into humans.

The 5'UTR of HuV is highly similar to that of Parechoviruses, both in the nucleotide sequence and the secondary folding structure [236], the most closely related IRES to that of HuV is HPeV-3. A close evolutionary history between the two *genera* has been suggested, and that modular recombination of the UTR may have occurred in ancestral viruses. This relationship helps to further explain why the supposedly Parechovirus-specific primers amplified HuV targets.

A complete phylogenetic analysis of the detected HuV wasn't possible, due to the small fragment of available sequence data, and the complete lack of any CDS. Speciation of picornaviruses requires >33% amino acid identity in P1 and >36% amino acid identity in the non-structural proteins 2C+3 CD [255].

3.4. Conclusions

In this chapter, several species of rodents were screened for zoonotic viruses. The detection of Tatenale virus in Twycross and Chester Zoos represent only the third detection of these viruses. This has extended the known geographic range of TATV from Northern and North-Western England, strengthening the evidence that TATV is endemic throughout England. The recovery of a novel TATV strain

'Norton-Juxta', increases the known diversity of the species and is further evidence of its endemicity.

The detection of several picornavirus genera, with the Parechovirus assay, demonstrates and increased usefulness for virus-discovery, though diminishes its suitability as a clinical diagnostic assay. The known geographic range of LV and RV has been extended into Eastern Europe, whilst CaV is now known to be present in the UK.

There remains no molecular evidence of orthobornavirus in the United Kingdom.

4. High-Throughput Sequencing of Virus-Positive Rodents and Recovery of Complete Tatenale Orthohantavirus Coding Sequence

4.1. Introduction

The Previous chapter details the RT-PCR screening of rodents from multiple locations, for zoonotic viruses. Two strains of Tatenale orthohantavirus were recovered from field voles, and four picornaviruses from brown rats, bank voles and house mice. Complete genomes of these viruses were not able to be recovered with degenerate primer-based PCR, and so were sequenced using unbiased high-throughput sequencing (uHTS).

uHTS, or Next-Generation Sequencing (NGS), is an increasingly popular method for both virus discovery and recovery of the viral genome. The costs of sequencing with this technology have decreased significantly [256,257], which has led to an increase in the frequency of its use.

4.2. Results

4.2.1. Sample Selection

Nine samples of animal tissue were chosen for Illumina HiSeq uHTS. These samples were two bank vole livers positive for Ljungan and rosavirus, A house mouse kidney positive for hunnivirus, a brown rat positive for cardiovirus, A lung and kidney from a Tatenale orthohantavirus positive field vole. Two Swan livers with no known viral infections were also included (Table 10). Additionally, A rabbit gut fluid sample and a Field vole gut sample were added to the samples in this HTS run but were analysed as part of a separate project.

Sample	Species	Organ	Known Virus Positive	Library
KE4	Field Vole	Kidney	Tatenale orthohantavirus	1
		Lung	Tatenale orthohantavirus	8
KE16	Brown Rat	Kidney	Cardiovirus	2
KF5D	House Mouse	Kidney	Hunnivirus	3
E120	Bank Vole	Liver	Rosavirus, Coronavirus	6
E293	Bank Vole	Liver	Ljungan Virus	7
Swm19	Swan	Liver	None	4
Swm20				

Table 10 Samples that were selected for high-throughput sequencing. The species, tissue type, the virus (if known) and the index number assigned to each sample are shown.

4.2.2. Evaluation of Enrichment Methods

To determine whether enrichment methods would improve the recovery of viral RNA, samples were subjected to a published enrichment protocol, with and without the addition of RNase, and compared to unenriched samples.

To compare the amount of host genome between enriched and unenriched samples, a GAPDH PCR was run on dilutions from both enriched and unenriched TATV-positive field vole kidney. GAPDH was detectable at a dilution of 10^{-2} in the unenriched samples and 10^{-1} in the enriched sample, indicating that GAPDH had been removed from the enriched RNA. The same samples were screened for orthohantavirus using the HAN-Semi primers; Orthohantavirus was detectable at dilutions of 10^{-1} in enriched samples, whilst only undiluted in unenriched samples. These results show that the enrichment process can successfully reduce the amount of host genomic material, concentrating any viral nucleic acid in the sample; this effect only improved the signal by one log, however.

The picornavirus positive samples, E120, E293, KE16D and KF5D, were enriched using the same method. Whilst GAPDH was detectable in each of these samples, picornavirus was not consistently detected in the enriched samples

Preliminary library preparation of KE4F and KE4D without enrichment showed that orthohantavirus was detectable at increased dilutions, 10^{-3} in KE4D and 10^{-4} in KE4F.

It was decided that the ribosomal RNA depletion that is implemented in the library preparation process was sufficient to enrich viral RNA in the sample and that further libraries would not be enriched.

4.2.3. Tapestation Analysis of RNA and Libraries

The RNA was analysed before creating the HTS libraries, as RNA concentration and RIN (RNA Integrity Number) influences the protocol. The libraries were then reanalysed before sequencing, to confirm sufficient yield cDNA concentration and whether correct fragment sizes had been created (Table 11)

	Pre-Library Creation		Post-Library Creation	
Library	**RNA Concentration** (ng/µl)	**RIN**	**cDNA Concentration** (pg/µl)	**Peak Fragment size (bp)**
KE4D	282	3.2	35.1	285
KE4F	138	1.5	2410	280
KE16D	735	2.4	80.4	258
KF5D	432	2.6	24.3	266
E120	468	8.3	35.1	285
E293	750	8.1	124	284
Swm19	772	4.7	196	251

Table 11 Results of Tapestation analysis of the pooled RNA, and the cDNA sequencing libraries following rRNA depletion. RIN is the RNA Integrity Number, and indicates the quality of RNA

4.2.4. Recovery and Analysis of Complete TATV

4.2.4.1. HTS of Norton-Juxta TATV

The kidney and lung of the TATV positive Field vole captured at Twycross zoo were subjected to uHTS. A total of 62,191,960 reads were sequenced from the lung tissue library; after pair-merging and trimming, 27,279,217 reads remained. These reads were then mapped to reference genomes; 3,786,262 reads were left in the final processed dataset. The processed reads were mapped directly onto the reference sequences for S, M and L segments of Puumala orthohantavirus (PUUV).

A total of 30,229 reads, representing 0.79% of filtered reads mapped onto the reference L sequence over five iterations. There was 100% coverage of the 5' untranslated region (UTR) and coding sequence (CDS), and partial 3' UTR. The Recovered CDS of TATV was 6465 nucleotides in length (2155 amino acids).

Mapping the reads to the M segment reference retrieved 32,806 (0.87%) reads over five iterations. Reads mapped to 100% of the reference sequence, recovering complete 5' UTR, CDS and partial 3'UTR. The CDS of M was 3447 nucleotides (1149AA) and the 5'UTR was 40 nucleotides

31,671 reads were mapped to the S segment reference, approximately 0.84% of the total reads. These mapped across 100% of the reference, recovering complete CDS and both the 5' and 3' UTR's. The CDS of the recovered S was 1302 nucleotides (434 AA), and the 5' UTR and 3' UTR being 42 and 489 nucleotides, respectively.

4.2.4.2. Recovery of Upton-Heath TATV

HTS libraries were prepared before the discovery of TATV in the Chester zoo cohort of field voles, and so did not include an Upton-Heath TATV positive sample. To recover the genome of this strain, the PCR primer walking technique was used, in which a series of overlapping primers were designed using the Norton-Juxta strain as a reference. The amplicons from these PCR reactions were then Sanger sequenced and mapped to the Norton-Juxta reference to assemble the full CDS.

The Complete L CDS was recovered using 15 pairs of primers (Figure 25) and complete S CDS with 3 primer pairs (Figure 27). Five pairs of primers were initially designed for the M segment (Figure 26), however, only partial CDS was recovered, with 90bp missing from the 3' end of the CDS.

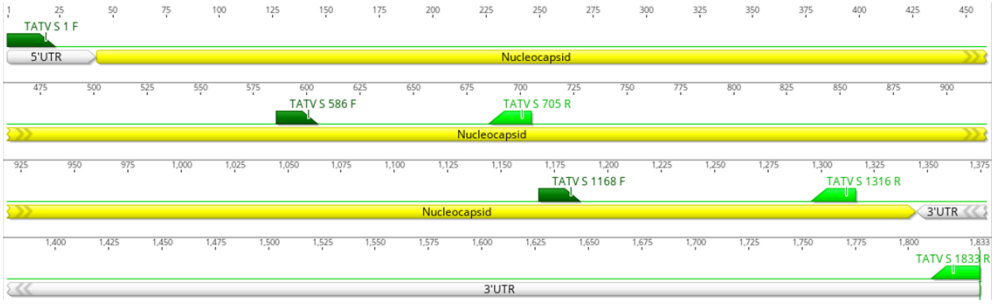

Figure 27 Primers designed to recover Upton-Heath TATV S CDS and UTR, mapped onto the CDS of Norton-Juxta S segment.

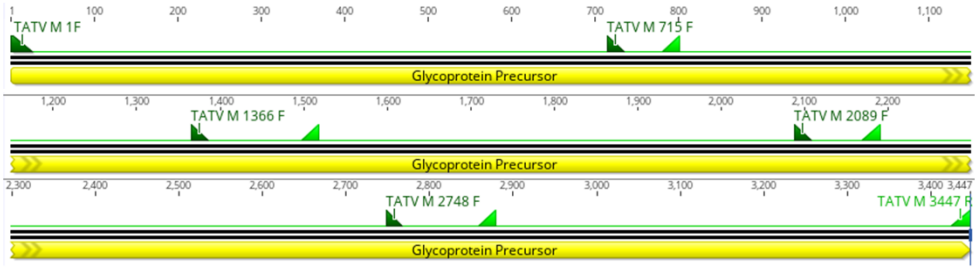

Figure 26 Primers desinged to recover Upton-Heath TATV M CDS, mapped onto the CDS of Norton-Juxta M segment.

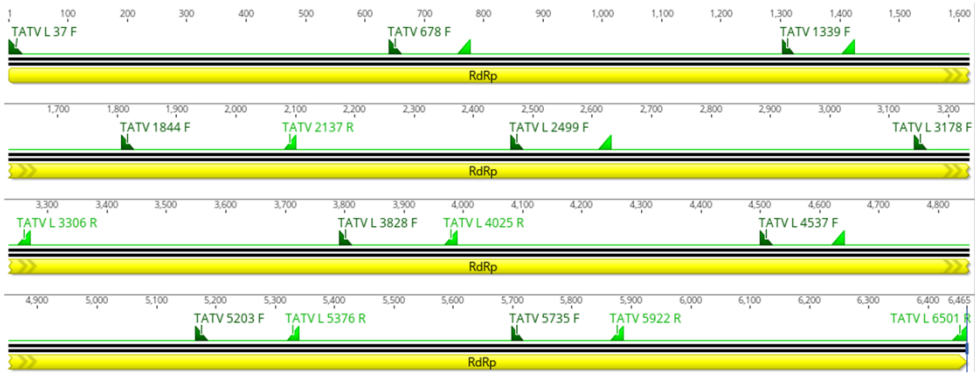

Figure 25 Primers designed to recover Upton-Heath TATV L CDS, mapped onto the CDS of Norton-Juxta L segment.

4.2.4.3. Analysis of Complete TATV CDS

The CDS of the L segments of both strains showed a similarity of 90.6% across the 6465 nucleotides and 99% across the 2154 amino acids. Phylogenetic analysis of the TATV L CDS and other *Arvicolinae*-associated orthohantaviruses show that TATV clusters with Puumala, Khabarovsk and Yuanjiang orthohantaviruses, but forms a distinct, novel clade from them (Figure 28). This phylogeny is supported by strong bootstrapping values. The closest related orthohantavirus was Khabarovsk virus, with both strains having a similarity of 78% at the nucleotide level, and 90% at the amino acid level. A comparable degree of similarity was also seen to the other *Arvicolinae*-associated species; 77.7% (90.4% aa) to Yuanjiang virus, 77.9/77.6% nt (88.1/87.5% aa) to Puumala virus and 76.9% nt (86.5% aa) to Tula virus.

Whilst a small section of the M CDS of Upton-Heath wasn't recovered, the similarity between the two strains is unlikely to differ significantly than that of complete CDS. At the nucleotide level, the two strains were 91.3% similar; at the amino acid level, the similarity was 99.1%. The phylogenetic tree shows a similar pattern of clustering as the L CDS, with TATV again clustering with Puumala virus, Yuanjiang virus and Khabarovsk virus. By contrast, the M CDS forms a clade with Khabarovsk, Yuanjiang and Fusong viruses, with strong bootstrapping values. Khabarovsk is the closest related orthohantavirus, with a nucleotide similarity of 76.4% (Norton-Juxta) and 77.1% (Upton-Heath), and amino acid similarities of 87.5% and

87.8%. Similarities to Yuanjiang virus at nucleotide level were 75.3/75.7% (86.5% aa), Puumala was 74.8/75.5% at nucleotide (84.7/84.6% aa) and 74.7/75% (*/83.5% aa) to Tula virus.

Similarities between the CDS of the S from both strains is 94.2% at the nucleotide level, and 99.1% at amino acids level. Phylogenetic analysis of the CDS showed an almost identical topology as the phylogeny of M. TATV clusters with the Khabarovsk virus, Yuanjiang virus and Fusong virus clade; forming a novel lineage. Both the Norton-Juxta and Upton-Heath strains were most similar to Khabarovsk 79.2% nt (89.4% aa). Similarities to other orthohantavirus species were 78.9% at nucleotide (88.2%) to Yuanjiang virus, 78.7/78.9% nucleotide similarity (88% aa) to Fusong virus, 77.9/78.4% (87.8% aa) to Puumala virus and 74/73% (81.8/80.4% aa) to Tula virus.

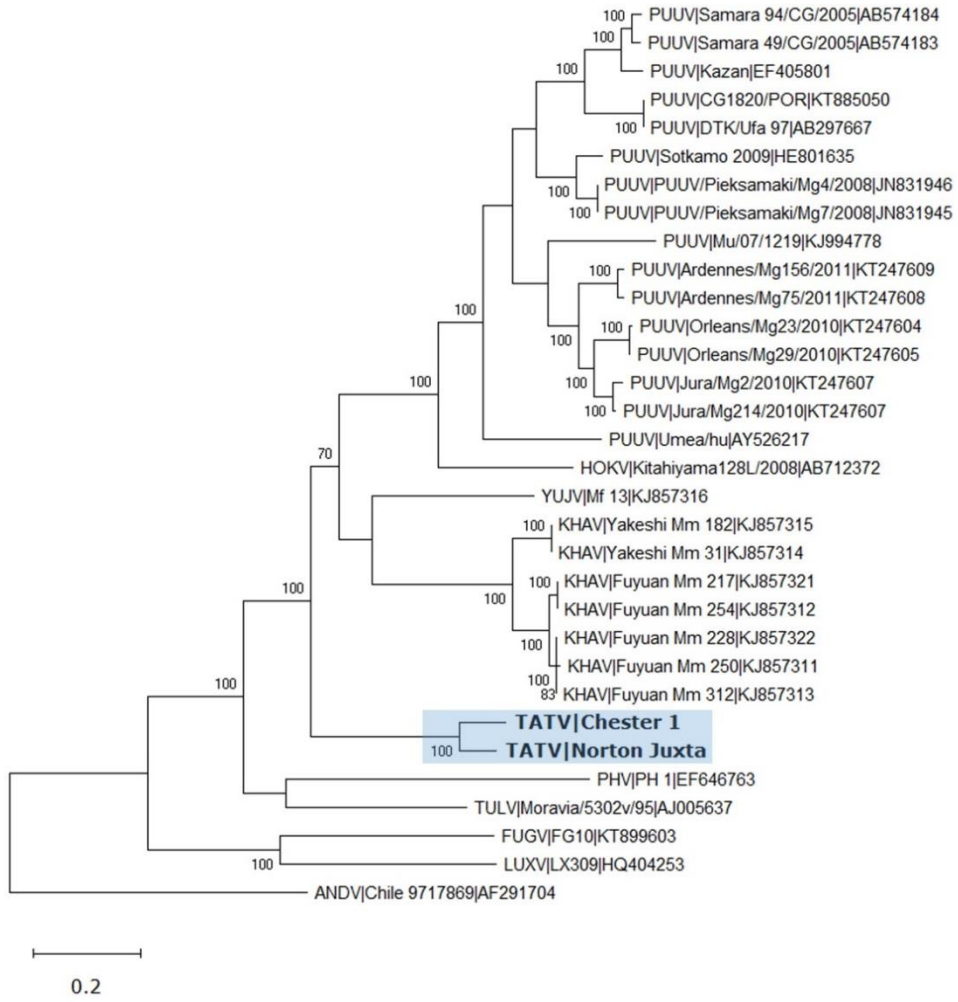

Figure 28 Phylogenetic relationship of the L segment of Tatenale virus with other vole-associated orthohantavirus species. Representative complete coding sequences were used. Maximum Likelihood trees were created with a GTR+G+I model [349]. Branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 1000 replicates. Tatenale virus strains are highlighted in boldface and a blue box. Sequences are shown with the species name, strain name and the GenBank accession number. PUUV, Puumala virus; HOKV, Hokkaido virus; FUSV, Fusong virus; YUJV, Yuanjiang virus; KHAV, Khabarovsk virus; TOPV, Topografov virus; TATV, Tatenale virus; PHV, Prospect Hill virus; ILV, Isla Vista virus; TULV, Tula virus; ADLV, Adler virus; LUXV, Luxi virus; FUGV, Fugong virus; ANDV, Andes virus

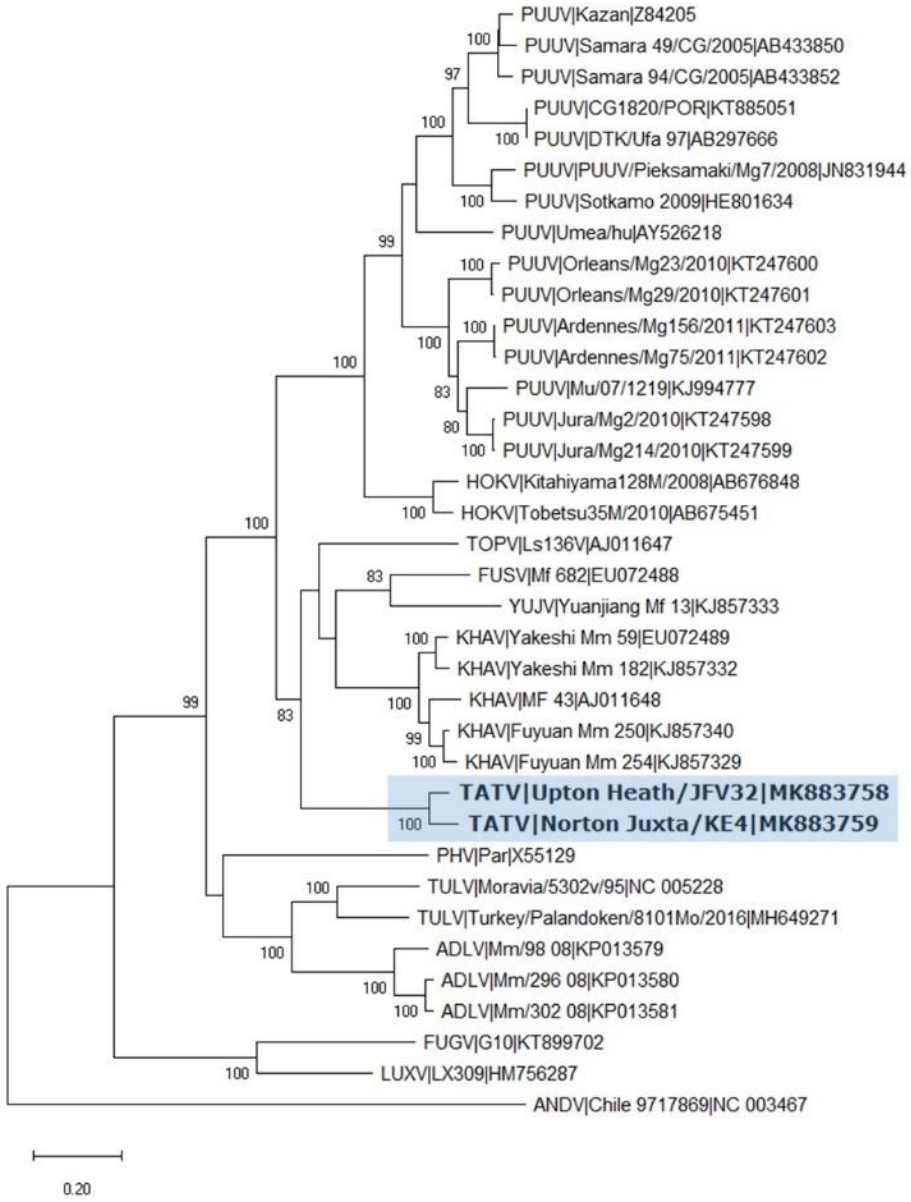

Figure 29 Phylogenetic relationship of the M segment of Tatenale virus with other vole-associated orthohantavirus species. Representative complete coding sequences were used. Maximum Likelihood trees were created with a GTR+G+I model. Branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 1000 replicates. Tatenale virus strains are highlighted in boldface and a blue box. Sequences are shown with the species name, strain name and the GenBank accession number. PUUV, Puumala virus; HOKV, Hokkaido virus; FUSV, Fusong virus; YUJV, Yuanjiang virus; KHAV, Khabarovsk virus; TOPV, Topografov virus; TATV, Tatenale virus; PHV, Prospect Hill virus; ILV, Isla Vista virus; TULV, Tula virus; ADLV, Adler virus; LUXV, Luxi virus; FUGV, Fugong virus; ANDV, Andes virus

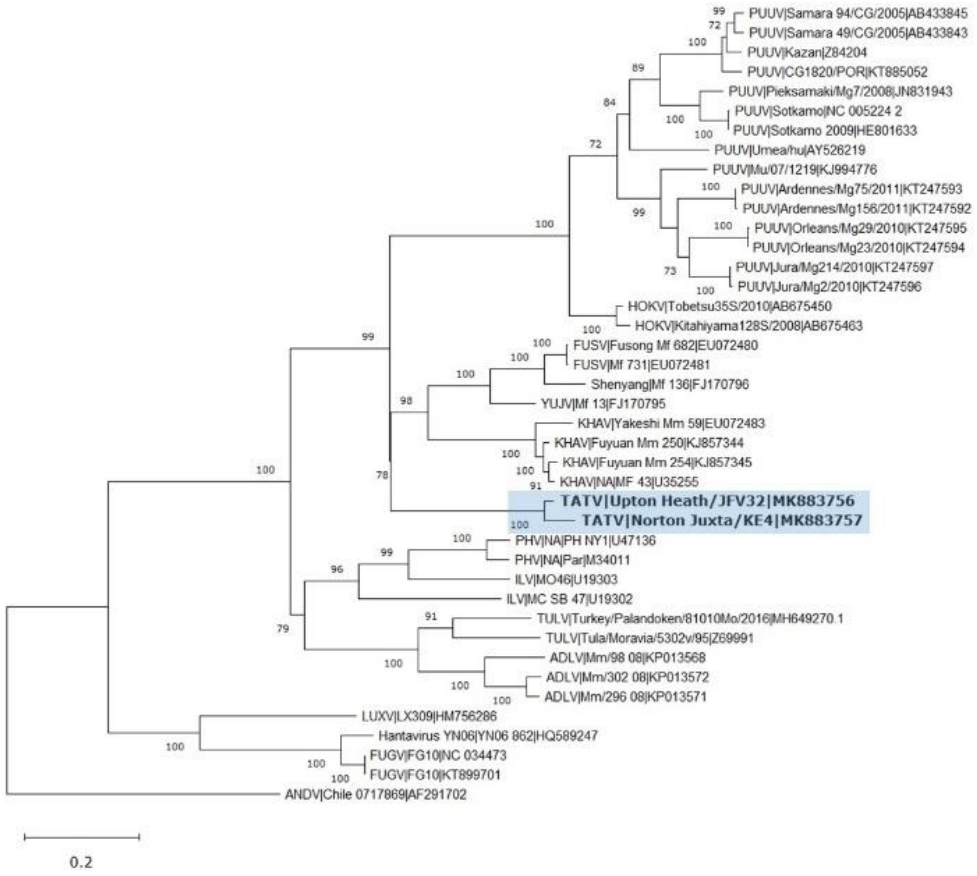

Figure 30 Phylogenetic relationship of the S segment of Tatenale virus with other vole-associated orthohantavirus species. Representative complete coding sequences were used. Maximum Likelihood trees were created with a GTR+G+I model. Branch lengths are drawn to a scale of nucleotide substitutions per site. Numbers above individual branches show bootstrap support after 1000 replicates. Tatenale virus strains are highlighted in boldface and a blue box. Sequences are shown with the species name, strain name and the GenBank accession number. PUUV, Puumala virus; HOKV, Hokkaido virus; FUSV, Fusong virus; YUJV, Yuanjiang virus; KHAV, Khabarovsk virus; TOPV, Topografov virus; TATV, Tatenale virus; PHV, Prospect Hill virus; ILV, Isla Vista virus; TULV, Tula virus; ADLV, Adler virus; LUXV, Luxi virus; FUGV, Fugong virus; ANDV, Andes virus

4.2.5. Retrieving picornavirus Genome from Positive Rodent Samples

4.2.5.1. KE16D, Cardiovirus Positive Rat

The Rat kidney library had 52,224,266 reads. Pairing, trimming merging the reads resulted in 22,189,279 reads. Mapping these reads to the *R. norvegicus* genome and transcriptome removed 21,545647 reads, leaving 597,406 non-host reads (2.7%).

These reads were mapped to the complete CDS sequence for Boone cardiovirus (JQ864343.1) (**Figure 31**); A single set of paired-reads were mapped to the genome. Both reads were 150 bp in length and overlapped one another to make a 189 bp contig. This contig was 2.2% of the total length of the reference sequence, corresponding to positions 5895-6084, the contig mapped with 94.7% pairwise identity to the reference. The contig was then BLASTN searched against the NCBI nt confirming that Boone cardiovirus was the closest match, with Boone cardiovirus 2 92.59% identical and cardiovirus C 92.06% identical.

Figure 31 KE16D Reads mapped to the Boone Cardiovirus Genome

4.2.5.2. E293, Ljungan Positive Bank vole

A total of 61,790,654 reads were obtained from the E293 library; pairing, merging and trimming of the sequencing resulted in 27,627,191 reads. 26,380,091 reads were mapped to a reference genome for bank voles, leaving 1,246,442 unmapped reads (4.5%).

These non-host mapping reads were mapped directly to the reference genome for LV (NC_003976.2) to retrieve any Ljungan virus reads from the dataset. However, no such reads were detected in the data.

4.2.5.3. E120, Rosavirus Positive Bank Vole

There were 65,928,072 paired reads sequenced from the E120 library. The reads were merged, into 30,934,726 reads and then quality trimmed which resulted in a final read number of 30,842,294. These were mapped to NC_024070.1, rosavirus 2, no rosavirus hits were retrieved from this dataset, however.

4.2.5.4. KF5D, Hunnivirus Positive House Mouse

There were 66,682,076 initial reads; merging and trimming these tread left 29,976,320 reads. These reads were mapped to the reference genome and transcriptome for *M. musculus* and the mitochondrial reference sequences. This removed 27,071,650 reads, leaving 2,904,670 non-host reads (9.69%).

Mapping to hunnivirus reference sequences retrieved no further HuV sequences.

4.2.6. Virus Discovery Using HTS

4.2.6.1. SWN19/20, Swan Liver

Sequencing of the swan liver resulted in 56,823,438 paired reads, these were trimmed and merged into 23,892,051 reads. Host mapping removed 20,836,235 reads, leaving 3,055,816 reads in the final dataset (12.7%).

These reads were *De novo* assembled into 408,681 contigs, which were then BLASTn searched against custom nucleotide databases containing all complete viral genomes for each virus order. They were also BLASTx searched against a database containing the amino acid sequences of RNA virus polymerases.

No viral sequences were detected in these samples in either the BLASTn or BLASTx results.

4.2.6.2. KE4, Kidney and Lung

The host-filtered reads were *de novo* assembled, 1,506,754 of the reads were assembled into 101,595 contigs.

BLASTn and BLASTx against virus databases and RNA-Virus polymerase database returned only orthohantavirus sequences.

4.2.6.3. Rat, Mouse and Bank Vole samples

The processed and trimmed sequences from the bank vole liver samples E120 and E293, the Rat kidney KE16D and the mouse kidney KF5D were each individually *de novo* assembled.

KE16 produced 145,488 contigs, KF5D produced 729,201 contigs, E120 produced 338,915 contigs and E293 produced 244,107 contigs.

BLASTn and BLASTx searches failed to detect any additional legitimate viral reads detected in any of these samples.

4.3. Discussion

4.3.1. Library Preparation

This was the first analyses of these samples derived RNA, aside from nanodrop spectrophotometry; the cost and time and time required for Agilent Tapestation analysis prohibited its use for all RNA extracted for RT-PCR screening in the previous chapter. This analysis highlights how RNA quality derived from different tissues can vary significantly. The field vole sample KE4 had RNA extracted from both its lungs and kidneys, these tissues would have had an almost identical time from the death of the animal to processing and storage at -80°c and a similar number of freeze-thaw cycles. The kidney sample, KE4D, had a RIN and RNA yield almost double that of the lung sample, KE4F. Lung tissue is much more exposed to the environment than kidney and would, therefore, degrade at an accelerated rate. Furthermore, lung tissue can contain a significant amount of fibrous tissue, such as collagen, depending on which section of the lung is sampled; fibrous tissue can significantly reduce the amount of RNA extracted from samples.

The quality of some libraries was poor, and ideally would have not have been selected for sequencing, and instead have the sample index re-prepared. However, given the expense of the rRNA depletion kits and the library preparation kits, it was decided to proceed with the samples.

4.3.2. Recovery of Full-Length TATV CDS

The complete CDS of all 3 segments of Norton-Juxta and almost complete Upton-Heath CDS that were sequenced in this study is the first reported recovery of full coding sequences for Tatenale orthohantavirus. This is also the first sequence of the M segment to date. Only two small fragments of the L and the S segments had been successfully sequenced, from which preliminary phylogenetic analysis has been obtained, both in the previous chapter and other publications [178,179]. The phylogenetic trees derived from the ~450bp fragment of the L segment show that TATV is closest related to Tula virus, and forms a novel clade from Puumala/Khabarovsk virus and Tula/Prospect Hill virus. The Phylogeny of the complete L CDS differs from the previous phylogenies, with TATV being much more closely related to Khabarovsk virus than Tula virus, which is consistent across all three complete CDS.

In addition to establishing an accurate phylogeny, recovery of full CDS establishes TATV as a novel species of orthohantavirus, which was suspected but was unable to be verified with the limited amount of sequence data available. Species demarcation of viruses is often contentious, with many papers reporting novels species, without providing sufficient sequence data to support these claims. The International Committee on Taxonomy of Viruses' (ICTV) [258] latest taxonomic proposal for the *Hantavirus* genus removed the species taxonomic status of 7 claimed species, whilst conferring species status to 24 viruses [259]. These changes are a result of

tightening of the demarcation criteria; an amino acid difference of greater than 7% in the complete S and M segments had been suggested in a previous ICTV release [230], whilst a more recent proposal requires stricter criteria of >10% amino acid difference in the S segment and >12% in the M [231]. The Norton-Juxta strain of TATV satisfies both sets of criteria completely, whilst the S and partial M segments of the Upton-Heath strain are sufficiently divergent but do not meet the required criteria due to the incompleteness of M.

Whether TATV can infect humans, and its potential pathogenicity remains unknown. Recovery of the complete CDS, particularly the M segment, will allow in vitro studies into determining the potential of TATV as a human pathogen. Pseudotyping, for example, is a technique in which the glycoprotein of a virus is inserted into a lentiviral backbone with a reporter gene (a pseudotyped virus) and transfected into cell cultures; A technique which has been used previously with hantaviruses [260]. This experiment would examine the virus's ability to enter human cells; if it is unable to infect human cells *in vitro*, then its zoonotic capability would be unlikely. Whilst Pseudotyping would answer the question of whether TATV can infect human cells *in vitro*, it would not provide a definitive answer to TATV's pathogenicity. Serological experiments using human samples to assay past exposure to TATV would be a logical step towards determining infectivity. However, the serological cross-reactivity between certain species of orthohantaviruses that are closely related

to TATV [261]; may throw doubt onto any results of these experiments; TATV-specific antibodies must also be shown to be unreactive to other orthohantavirus to conclusively prove TATV-infection in seropositive samples. Ultimately, the only conclusive way to prove the infectivity of TATV in humans would be to screen human samples and detect viral RNA. If the positive samples were associated with clinical symptoms, then TATV could also be shown to have a pathogenic effect on humans. A Recent example of the evidence required to establish the pathogenicity of an orthohantavirus n example of this would be Tula virus, which was disputedly suspected to be zoonotic [262] until the virus was finally recovered from a patient hospitalised with Hantavirus Disease [263].

4.3.3. High-Throughput Sequencing of Picornavirus Positive Rodent Samples

4.3.3.1. Fragment of a CaV recovered.

The sequencing of the CaV-positive rat only recovered a single further fragment of the genome. This fragment is most identical to the same published sequence as the qPAR amplicon, which supports the previously generated phylogeny. This also confirms that the sequenced virus is indeed a cardiovirus and not a different picornavirus species with a highly similar IRES sequence.

4.3.3.2. No Recovery of LV or RV from Bank Vole Kidney, or HuV from House Mouse Kidney

High-Throughput Sequencing of KF5, E120 and E293 failed to retrieve any viral genome of their associated picornavirus; HuV, RV

and LV, respectively. The precise reason the genomes were not sequenced cannot be completely established; RNA quality, library preparation, sequencing and data analysis are all potential steps which may have negatively affected the recovery of the virus genome.

It is unlikely that poor RNA quality was a factor in the failure to detect rosavirus, as the RIN of this sample was 8.3, the highest of the samples in this HTS run. Similarly, the Ljungan Virus positive sample (E293) had a RIN of 8.1, the second-highest of the samples. KF5D, the hunnivirus positive, had a much lower RIN of 2.6, which may have influenced the results. The RIN measured in the samples is predominantly an indicator of the quality of host RNA present in the sample. However, if a sample has been processed and stored well enough to preserve the host RNA, then viral RNA would presumably have been reasonably well-preserved.

Choice of the tissue sample may also have affected virus recovery. There is surprisingly little information available regarding the organ tropism of LV, RV and HuV; Many of the isolates are recovered from blood samples, or only a single organ, which makes the comparison of viral loads in organs difficult. In the case of E120 and E293, however, only a single organ was available for screening, and so choosing the best organ was a moot point.

A potential reason for failure to recover any LV sequence from the sample could be that the viral load was low enough that viral genome was not present in sufficient quantity in the RNA used to prepare the sequencing index to be detected.

Multiplexing of several indexes into a single sequencing library will reduce the total number of reads for each library, assuming each index is loaded in a uniform concentration. If fewer indexes had been pooled into the library, then perhaps the additional reads available to these picornavirus libraries would have recovered the genome, or fragments thereof.

4.3.4. Virus Discovery with HTS

4.3.4.1. Investigation of Swan Liver

The swan tissue was included in this sequencing as an opportunistic sample, tissue was collected from two swans which had died in the Clumber Park, a country park located in Nottinghamshire. An ongoing die-off of swans in the park and the recent outbreak of avian influenza A H5NE throughout Europe, which had been detected in swans [264], had led to an initial concern of an outbreak in these birds. Influenza screening by the Animal and Plant Health Agency (APHA) were negative, however.

No legitimate viral sequences were detected in these samples using the in-house analytic pipeline.

4.3.4.2. Rodent Samples

A number of additional viruses were detected in KE4D and E120, as well as two other rodent samples included in the sequencing run, the results and analyses of which were part of a separate project [265]. These included picornaviruses, Paramyxoviruses and Rotaviruses.

No additional viruses were detected in KE4F, KE16D, KF5D or E293. It is not surprising that these samples were negative for any further viruses, as the sequencing of only a few animals means that the chances of finding a positive it reduced. Many studies using HTS for virus discovery sample a much greater number of animals, whether through pooling multiple samples into a single index [266] or by increasing the number of indexes [267]. However, as the main purpose of this HTS run, with the exception of the swan samples, was to recover the complete genomes of the orthohantavirus and picornaviruses identified in chapter one. The sample number was therefore limited, to increase the number of reads for each of the samples.

4.3.4.3. Bioinformatics Pipeline

The choice of bioinformatic pipeline strongly affects the outcome of the analysis, leading to false negatives or even false positives [268]. The choice of bioinformatic pipeline can be influenced by several factors, including hardware constraints, familiarity with software and the sequencing technology used. The pipeline used in this chapter was run on a Windows operating system (OS) based computer with only a moderate amount of computing power available, narrowing down the usable pipelines. Most published bioinformatic pipelines require a Unix-based OS, such as macOS or Linux, whilst there is a dearth of available software for Windows, meaning that the Geneious platform used was one of the only accessible software's available. Whilst limited, the Geneious platform has successfully been used for virus

discovery in previous studies [269] and was able to recover almost-complete genome of the Tatenale virus, which shows that it is sufficient for recovering viral sequences from HTS datasets.

Conclusions

This chapter describes the use of HTS to recover the complete genome of Tatenale orthohantavirus. This recovery of the complete coding sequence for Norton-Juxta strain and the almost-complete coding sequence of Upton-Heath strain represents the first complete coding sequence for TATV and the first sequencing of the M segment. This has allowed the speciation of TATV, using the ICTV approved species criteria. This will allow further characterisation of TATV.

Whilst small fragments of a cardiovirus were detected, the further sequence wasn't recovered for LV, RV or HuV, precluding any further work into characterising these viruses.

5. Investigating Seroprevalence of Orthohantaviruses In Humans and Captive Non-Human Primates

5.1. Introduction

Molecular evidence of orthohantavirus has been reported in British rodents on multiple occasions [164], however only serological evidence of infection in humans have been reported [156,162]. Infection with Eurasian orthohantaviruses such as Puumala virus (PUUV) and Seoul virus (SEOV) often causes mild, subclinical infection, which the patient may not seek medical attention, precluding any diagnosis. Furthermore, unlike the persistent infection in rodents, viremia in humans is often short-lived [270], meaning that viral RNA is only detectable for a short-period following onset of symptoms, and using a PCR diagnostic assay alone may potentially lead to false-negatives. Serological screening of suspected orthohantavirus infections, alongside PCR, is often used [271]. IgM response to infection begins with the onset of symptoms and is used to diagnose acute infection, before gradually decreasing over the course of weeks [63]. IgG begins shortly after the acute stage of infection, and persists long-term, allowing for retrospective diagnosis of previous infection [272].

A Human cohort consisting of hepatitis c positive individuals with a 'chaotic' lifestyle was chosen; chaotic lifestyle was defined as IV drug users with an increased likelihood of homelessness or rough-living. Those with a chaotic lifestyle have greater exposure to rodents, and

therefore are more likely to become infected with an orthohantavirus. A cohort of non-human primates (NHP), from a zoo where Tatenale orthohantavirus has been detected, were also selected, to determine whether the circulating Tatenale virus (or a second, undetected species) are causing infection.

5.2. Results

5.2.1. Orthohantavirus Seroprevalence in Humans

5.2.1.1. IgM Seroprevalence

Samples from the hepatitis C positive, chaotic lifestyle cohort had an IgM seroprevalence of 14.4%. Healthy blood donors, however, had an IgM prevalence of 4.8% (Figure 32). Due to the small population size for the blood donor cohort (N=21), compared to the population size of the chaotic lifestyle cohort (N=181), no statistical significance could be inferred.

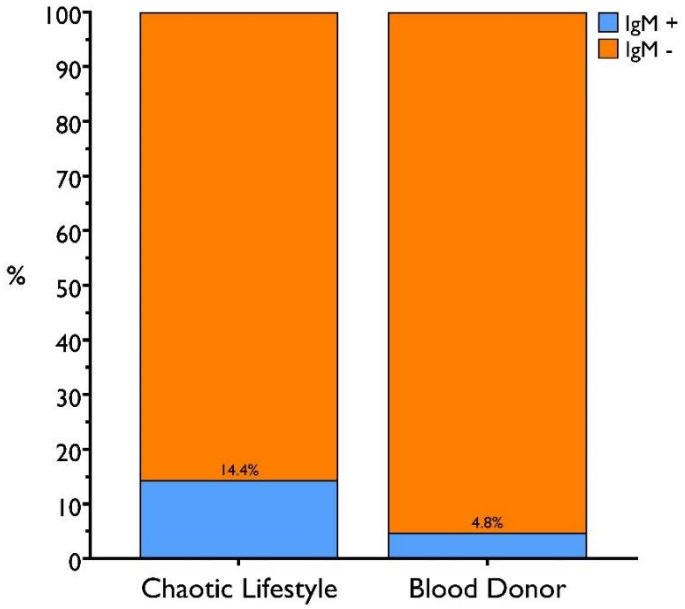

Figure 32 Seroprevalence of orthohantavirus IgM in healthy blood donors, and individuals with a 'Chaotic' lifestyle (Hepatitis C positive, IV drug users).

IgM seroprevalence in the chaotic lifestyle cohort was similar between samples from 2014, 2015 and 2017, though was much lower in 2016 (Figure 33). Blood donors were all sampled during the same period. There was no significant difference (P= 0.6171) in the IgM prevalence between male and female samples from the Chaotic lifestyle (Figure 34)

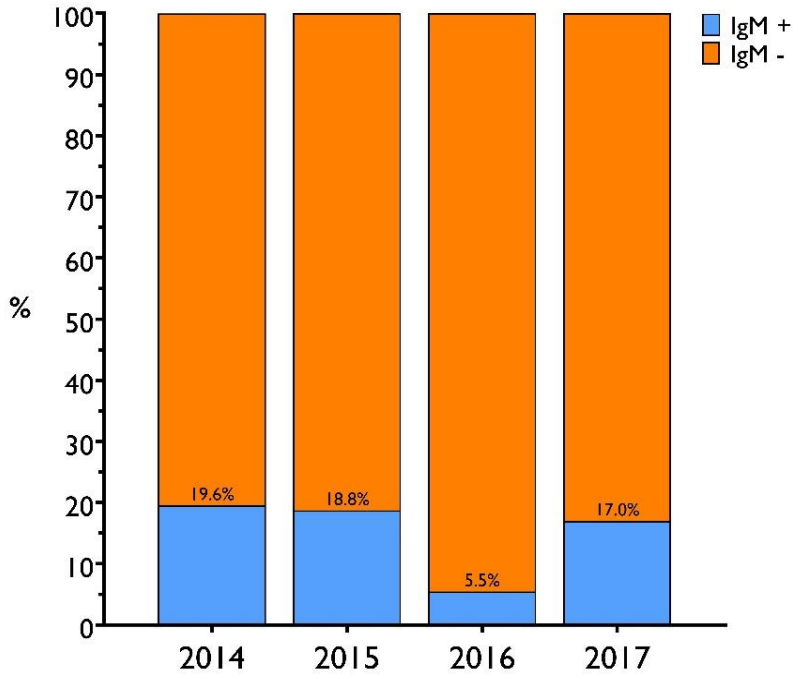

Figure 33 Seroprevalence of orthohantavirus IgM in individuals with a 'Chaotic' lifestyle (Hepatitis C positive, IV drug users), by the sample year.

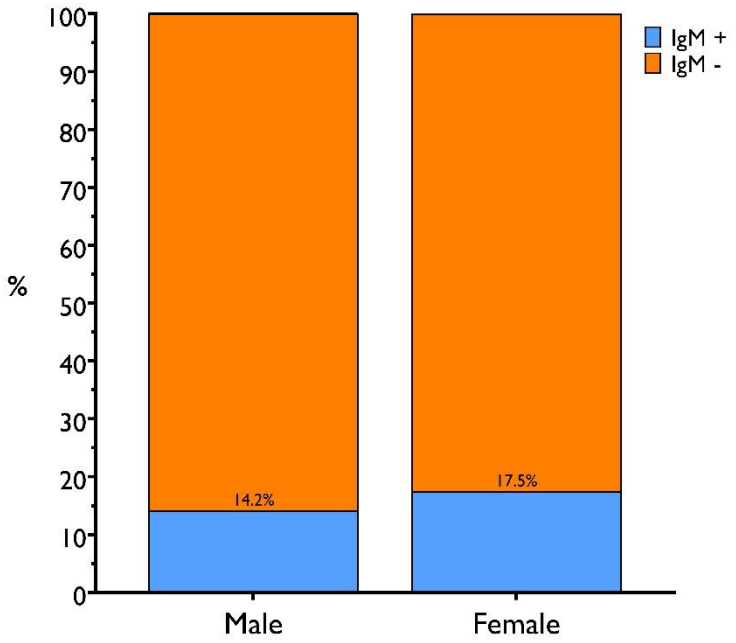

Figure 34 Seroprevalence of orthohantavirus IgM in individuals with a 'Chaotic' lifestyle (Hepatitis C positive, IV drug users), by sex.

5.2.1.2. IgG Seroprevalence

IgG seroprevalence was similar between the Chaotic lifestyle and healthy blood donor cohorts, 3.9% and 4.8% respectively (Figure 35). IgG prevalence in the samples taken during different years (Figure 38) There was no IgG detected in chaotic lifestyle cohort was not significantly different (P=0.513969) between any of the female samples from the chaotic lifestyle cohort (Figure 38).

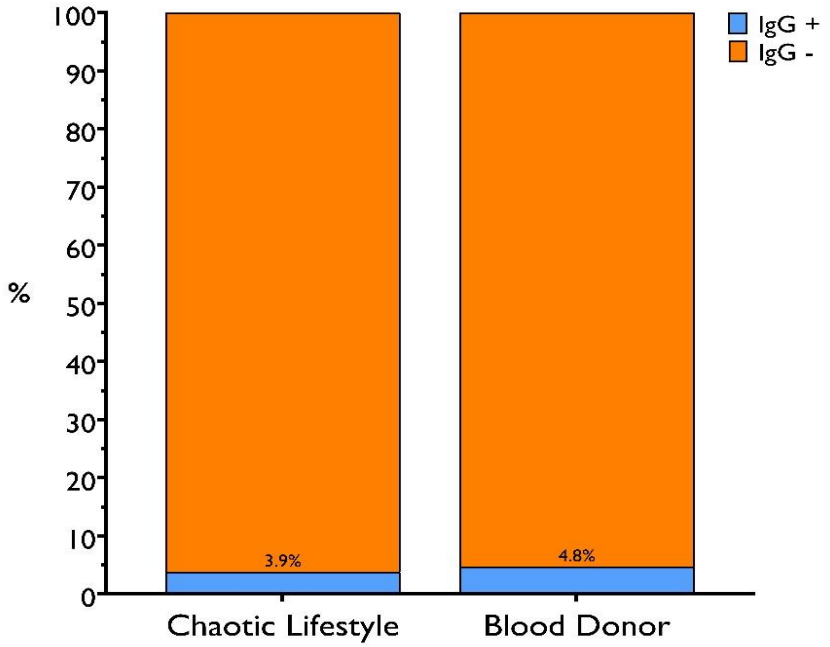

Figure 35 Seroprevalence of Orthohantavirus IgG in healthy blood donors, and individuals with a 'Chaotic' lifestyle (Hepatitis C positive, IV drug users). The seroprevalence is shown as a percentage of the total number of samples used.

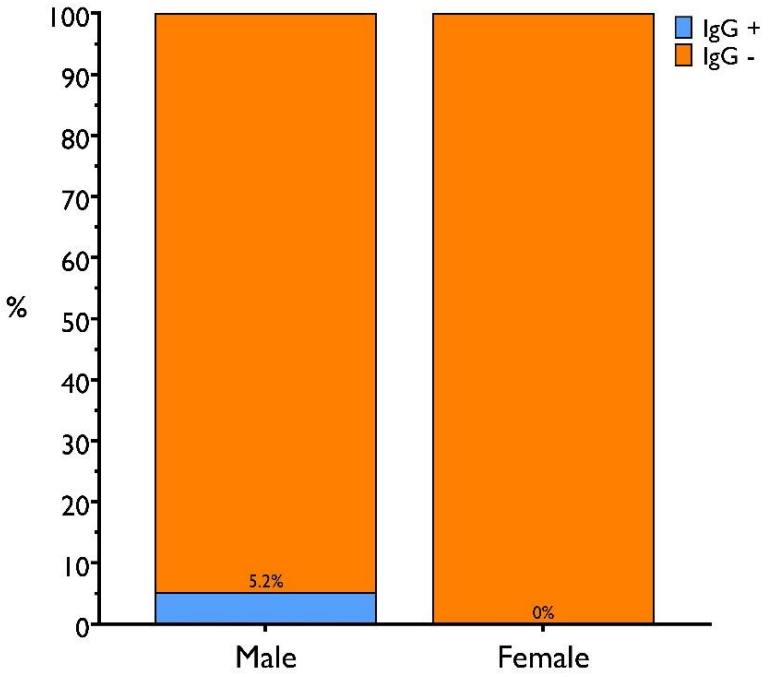

Figure 36 Seroprevalence of Orthohantavirus IgG in individuals with a 'Chaotic' lifestyle (Hepatitis C positive, IV drug users), by sex. The seroprevalence is shown as a percentage of the total number of samples used.

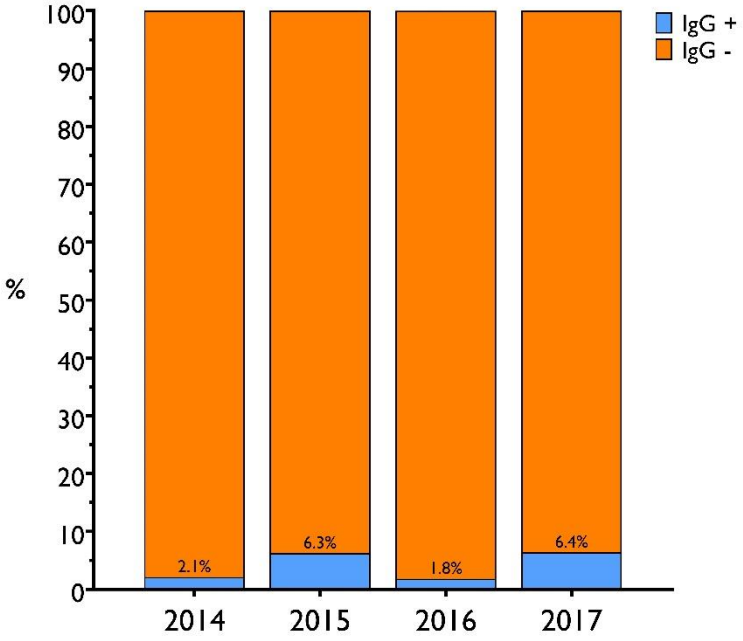

Figure 37 Seroprevalence of Orthohantavirus IgG in individuals with a 'Chaotic' lifestyle (Hepatitis C positive, IV drug users), by the sample year. The seroprevalence is shown as a percentage of the total number of samples used.

5.2.2. Orthohantavirus Seroprevalence in Non – Human Primates

5.2.2.1. IgM Seroprevalence

Several species of primates were screened for IgM prevalence.

Orthohantavirus IgM was only detected in Chimpanzee's and

Bonobo's, with each species having a 12.5% prevalence (Figure 40)

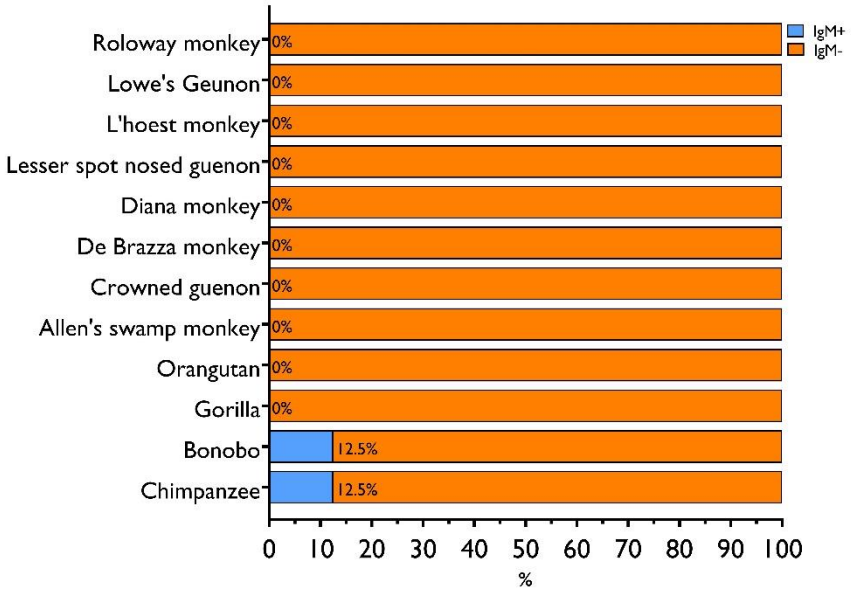

Figure 40 The Seroprevalence of Orthohantavirus IgM in several species of captive primates. The seroprevalence is shown as a percentage of the total number of animals used.

5.2.2.2. IgG Seroprevalence

Orthohantavirus IgG was detected only in Chimpanzee's and

Bonobo's, the prevalence was 58.3% and 75% respectively (**Figure 41**).

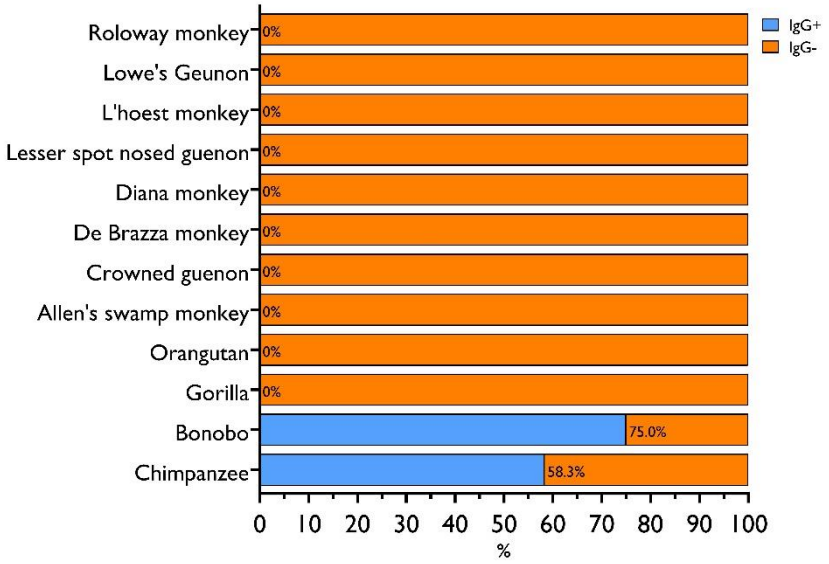

Figure 41 Seroprevalence of Orthohantavirus IgG in several species of captive primates. The seroprevalence is shown as a percentage of the total number of animals used.

5.2.3. PCR Screening of IgM positive Sera

Human and NHP sera that were IgM positive were screened for orthohantavirus RNA using RT-PCR, however, none were detected in any of the samples.

5.3. Discussion

5.3.1. Human Serology

5.3.1.1. Selection of Samples

Samples from the 'Chaotic lifestyle' cohort had an IgM prevalence almost three times higher than that of the healthy blood donor control, though due to the small cohort size of blood donors, no statistical significance was inferred. As the IgM response is short-lived [63], screening for them is generally used to diagnose acute infection, and not for assessing prevalence. However, patients with suspected haemorrhagic fever with renal syndrome in Northern Ireland were screened for IgM and detected a prevalence of only 2.1% [155]. This prevalence was much lower than detected in this study, however, both studies were twenty years apart, and in completely different regions, which may explain this disparity in prevalence.

Prevalence of orthohantavirus-specific IgG in the 'Chaotic lifestyle' cohort was 3.9%. Previous serological studies have reported various IgG prevalence's, which appear to reflect the relative risk of exposure to rodents. For example, owners of pet rats have an IgG seroprevalence of 34.1% [273] and farmers a prevalence of 7.6% [156], whilst seroprevalence in random blood donors was 3.3%.

The human samples for the serology screening were taken from an existing cohort of hepatitis C positive patients. These patients had all received treatment from one of several centres that provides needle-exchange, homeless outreach support and support for drug-users. There was no data available that would indicate whether each patient

was homeless or not, however intravenous drug use (IVDU) is prevalent amongst homeless populations [274], and would suggest that many of the sampled individuals are also homeless. Homelessness is associated with poor health status, sanitation and increased time spent outdoors [275], factors which increase exposure to zoonoses, particularly those associated with rodents. Seroprevalence of orthohantaviruses within populations with 'chaotic' lifestyles, i.e. the homeless and IV drug users, are not significantly different than the baseline prevalence in the United States [276,277]. This suggests that a chaotic lifestyle may not be a significant factor in the increased prevalence of orthohantaviruses in the 'chaotic lifestyle' cohort. As all the samples in this cohort were chronically infected with Hepatitis C, susceptibility to infection from orthohantavirus may be increased.

IgM prevalence is significantly higher than the prevalence of IgG. Given the short-lived IgM response, an IgM prevalence lower than IgG was predicted. A possible explanation for the large IgM prevalence could be that samples were taken shortly after the instigation of a chaotic-lifestyle, and therefore had only recently become infected. Alternatively, a recent increase in transmission events, possibly as a result of an increase in orthohantavirus positive rodents in the region, would cause an increased IgM seroprevalence, whilst IgG would remain consistent. IgM positive samples had no matched samples from later time points, and therefore the greater

IgM prevalence would not be reflected in a greater IgG prevalence in samples from later years

5.3.2. Non-Human Serology

Chimpanzees and Bonobos were both found to be IgG and IgM positive. Research into orthohantavirus infection in non-human primates is sparse. Experimental infection of primates has demonstrated susceptibility to orthohantavirus infection in macaques [278,279] and chimpanzees [280]. Serological evidence of orthohantavirus was detected in captive macaques in Germany [281].

This is the first detection of natural orthohantavirus infection in both chimpanzees and bonobos. The ELISA used in this study detects several species of orthohantavirus but does not differentiate between them, meaning that the causative species is unknown. Currently, only two species have been detected in the UK, Seoul virus and Tatenale virus. Rodents from Twycross zoo were screened for orthohantavirus using an RT-PCR assay, all rats (the reservoir species for Seoul virus) were negative, whilst Tatenale virus was detected in field voles. This would suggest that Tatenale virus, and not Seoul virus, is infecting these primates.

5.3.3. Limitations of the Seroprevalence Study

5.3.3.1. Human Cohorts

Samples from the 'Chaotic lifestyle' cohort were all chronically infected with hepatitis C, the healthy blood donor cohort was not,

however. There is a possibility that the increased seroprevalence of IgM is a result of false-positives, due to interference from the hepatitis C infection. A more appropriate comparative cohort would be hepatitis C positive individuals with no history of a chaotic lifestyle, which would clarify whether the hepatitis C or chaotic lifestyle is the factor increasing orthohantavirus seroprevalence. Identifying a suitable cohort is problematic, however, as IV drug use is the single greatest risk factor for hepatitis C in England, with 90% of all laboratory-confirmed cases reporting IV drug use [282]. This severely limits the pool of candidates available for screening.

Access to matched samples for positive samples was also limited, preventing the retrospective screening of IgG positive samples to determine when infection occurred and the screening of samples from later time-points of IgM positive individuals.

5.3.3.2. Technical Limitations of the ELISA

The ELISAs performed in this chapter were intended as a preliminary study into the seroprevalence of orthohantaviruses. The ELISA assays used can detect multiple orthohantavirus species, though it is not able to differentiate between these species. Commercial indirect immunofluorescence assays can differentiate between these species [283] and are typically used in conjunction with ELISA [157]; however, these were unavailable for used during this project.

This ELISA kit is designed for clinical diagnosis in humans and has not been validated in non-human primates. Therefore, the possibility

remains that the reactivity in the NHP cohort may be false positives. Positive results from rescreening these samples with other serological assays would support the validity of these ELISA results.

5.4. Conclusions

In this chapter, an orthohantavirus ELISA assay was used to screen human and non-human primate cohorts. The results of these serological assays show that there is increased exposure of orthohantaviruses in at-risk human populations and endangered captive non-human primates. Whilst the limitations in obtaining appropriate control cohorts have precluded any definitive conclusions from these results, they have highlighted a potential ……

6. High-Throughput Sequencing of Undiagnosed Clinical Samples

6.1. Introduction

A significant number of clinical samples from patients with a suspected viral illness screened using routine diagnostics, fail to detect any viruses. In particular, patients with neurological disease such as meningitis or encephalitis often remain undiagnosed [284,285]; 80% of neurological samples referred for viral screening in Nottingham University Hospitals (NUH) are negative.

Diagnostic kits are often PCR-based, and either viral species-specific primers are used, or a multiplexed to detect wider range viruses, that tends to result in similar clinical presentation, i.e. neurological, respiratory or gastrointestinal disease. Whilst these assays are highly sensitive to their targets, their specificity is often narrow, which requires clinicians to select assays based what they believe is the likeliest virus, considering factors such as clinical presentation of the patient, geographic area and previous clinical experience. This targeted approach is therefore heavily biased towards viruses that are considered to be common and which present clinically in a typical manner. Novel viruses or unlikely viruses, such as a common pathogen presenting in an atypical manner or a zoonotic virus, are unlikely to be detected using these methods.

High-Throughput sequencing (HTS) significantly reduces the effects of targeting bias. These technologies do not require the virus-specific primers that are necessary for PCR-based diagnostics, they are instead capable of sequencing any genomic material in the sample. The expense, turn-around time for sequencing and complexity associated with the preparation and data analysis of these platforms have largely precluded routine use in diagnostic laboratories. However, HTS has become increasingly popular for the retrospective screening of undiagnosed clinical samples [286,287] and for patients with severe, undiagnosed disease [288]. Several cloud-based analytic platforms have recently been developed, to offset the requirement for access to powerful, local computers. These have successfully been used to detect previously undiagnosed viruses in clinical samples [287,289,290]

This Chapter describes the collection, and subsequent HTS of cohorts of undiagnosed cerebrospinal fluid, respiratory and blood samples. It also compares the efficacy of cloud-based analytic platforms with offline analytic pipelines, in processing an artificial HTS dataset.

6.2. Results

6.2.1. Selection of Samples

6.2.1.1. Cerebrospinal Fluid Samples

Four hundred CSF samples were collected and added in equal volumes into pools of ten samples. Twenty of these ten-sample pools were then equally combined into a larger pool, containing 200 samples in total; this resulted in two final pools, CSF-1 and CSF-2. These CSF samples were all screened using the NUH standard neurological virus panel, they were collected between August 2016 and January 2018. The patient ages of these samples are shown in Figure 42, samples from patients under the age of 10 comprised around half of all samples in both pools, whereas patients over 80 were the smallest grouping in both pools.

Samples were chosen primarily from patients with neurological conditions, the most common of the neurological clinical features was meningitis, followed by encephalitis, seizure, headaches, meningoencephalitis and epilepsy (Figure 43).

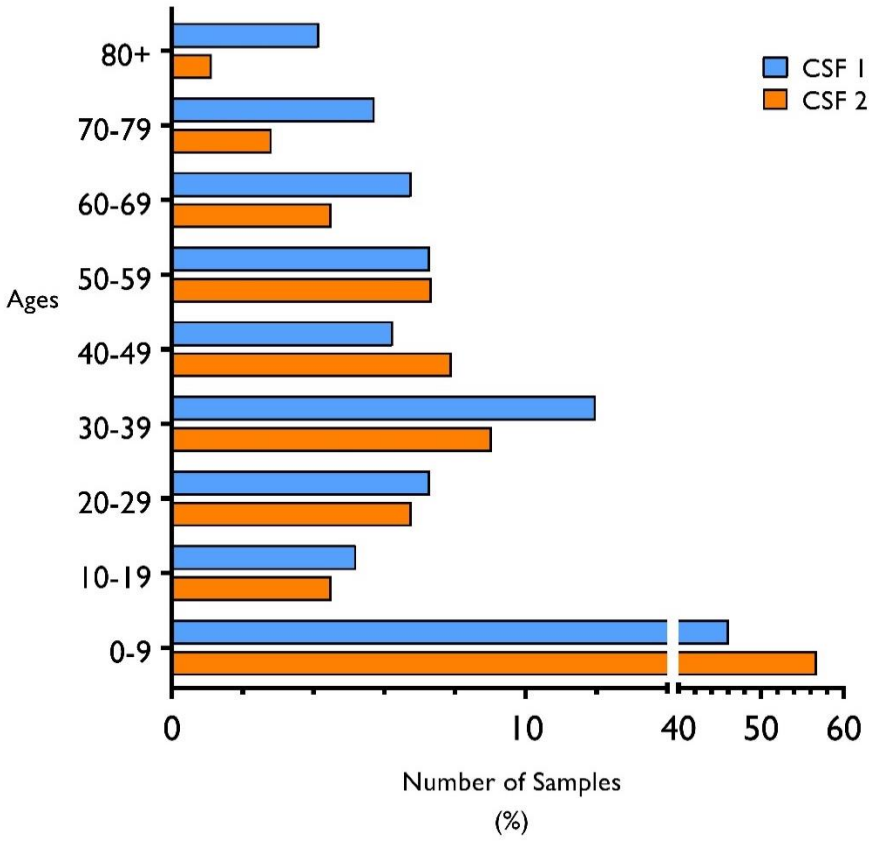

Figure 42 The ages of the patients at the time of sampling, whose samples were used to create CSF pools 1 and 2, given as a percentage of the total number of patients.

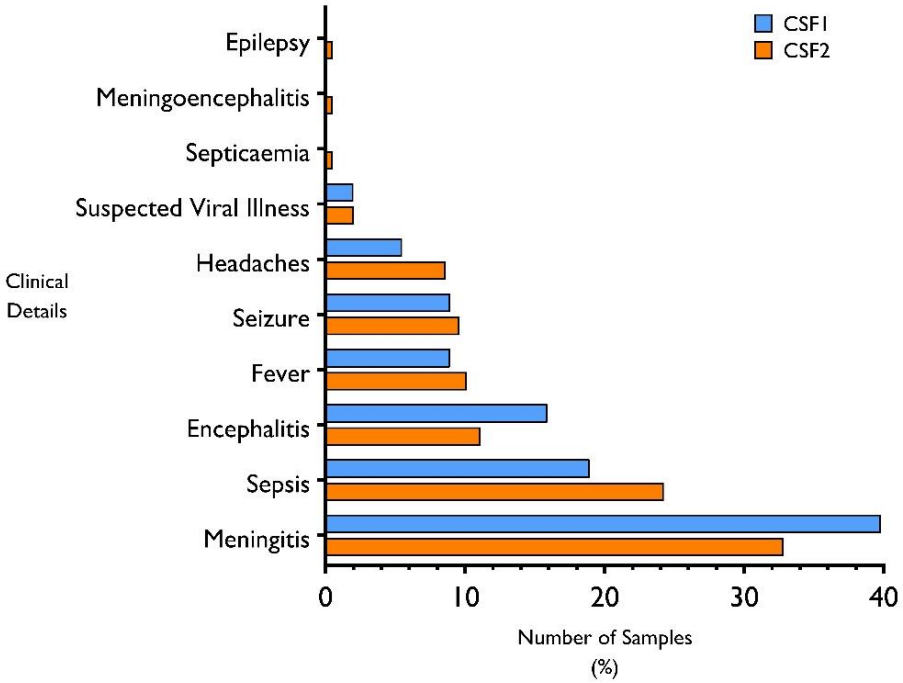

Figure 43 The Clinical details provided for each of the samples used to create CSF pools 1 and 2. The number of each clinical detail is given as a percentage of total number of samples.

6.2.1.2. **Blood Samples**

Eighty EDTA-preserved blood samples collected between January 2016 and December 2017 were pooled into eight groups of ten samples each, these pools were then combined into a single large pool, EDTA-1. Samples taken from patients between 0-9, 10-19 and 40-49 made up the majority of the EDTA-1 pool, with around 15% of the pool each. The 20-29 and 30-39 ages were the smallest age brackets, comprising only 5% each (Figure 44).

Samples were selected from patients with a broad range of clinical details, though the majority of the samples were taken from patients with a fever (Figure 45).

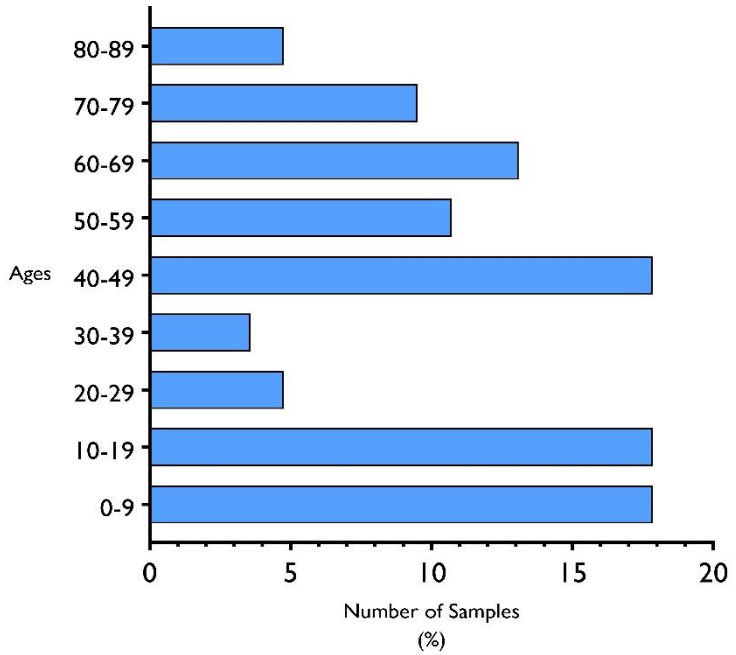

Figure 44 The ages of the patients at the time of sampling, whose samples were used to create the EDTA1 pool, given as a percentage of the total number of patients.

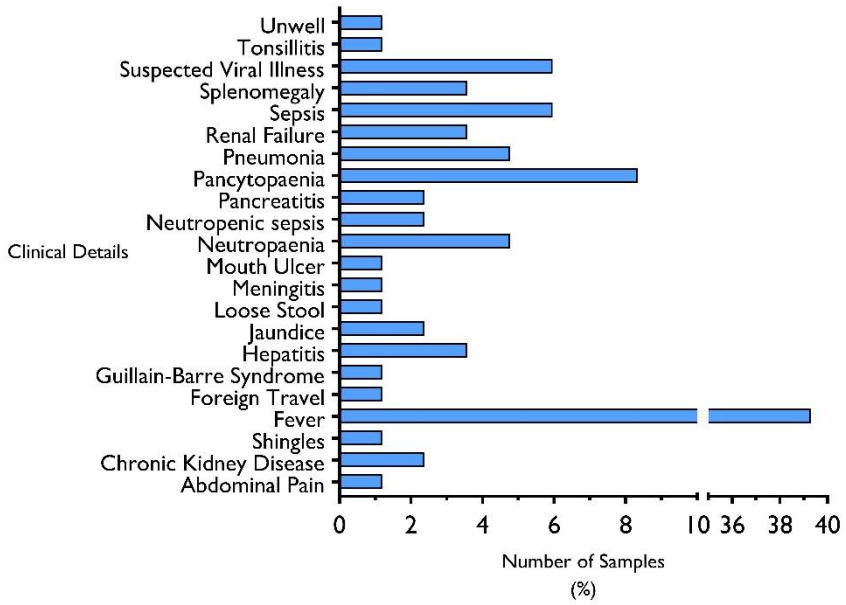

Figure 45 The Clinical details provided for each of the samples used to create the EDTA-1 pool. The prevalence of each clinical detail is given as a percentage of total number of samples.

6.2.1.3. Respiratory Samples

Respiratory samples from 170 patients collected between June 2014 and December 2017 were pooled into seventeen groups of ten. These groups were then combined to created two larger pools, Resp-1 containing ten pools and Resp-2 containing seven pools. Most samples were throat swabs, 93% of Resp-1 and 87% of Resp-2, whilst the remainder of Resp-1 included nasopharyngeal swabs (6%) and endotracheal tube secretions (1%), whilst Resp-2 was nasopharyngeal swabs (13%). Meningitis was the most common clinical feature for samples in both pools, followed by sepsis and febrile illness; the remainder of the samples encompassed a range of clinical details in smaller numbers (Figure 46). Samples from patients under the age of ten were the most common in both pools; the distribution of other age groups differed slightly between each pool (Figure 47).

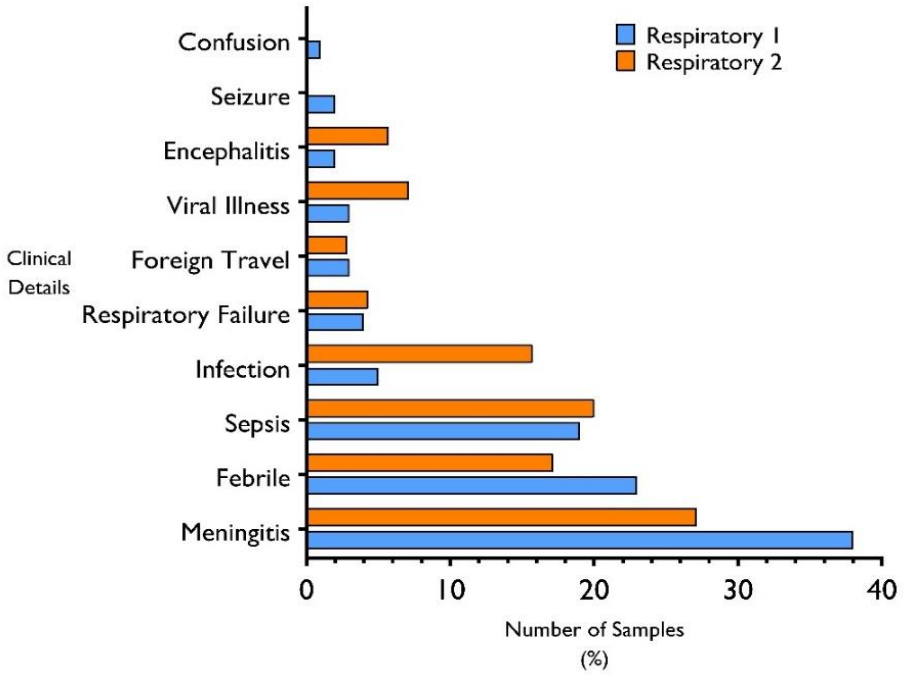

Figure 46 The Clinical details provided for each of the samples used to create the Resp-1 and Resp-2 pools. The prevalence of each clinical detail is given as a percentage of total number of samples.

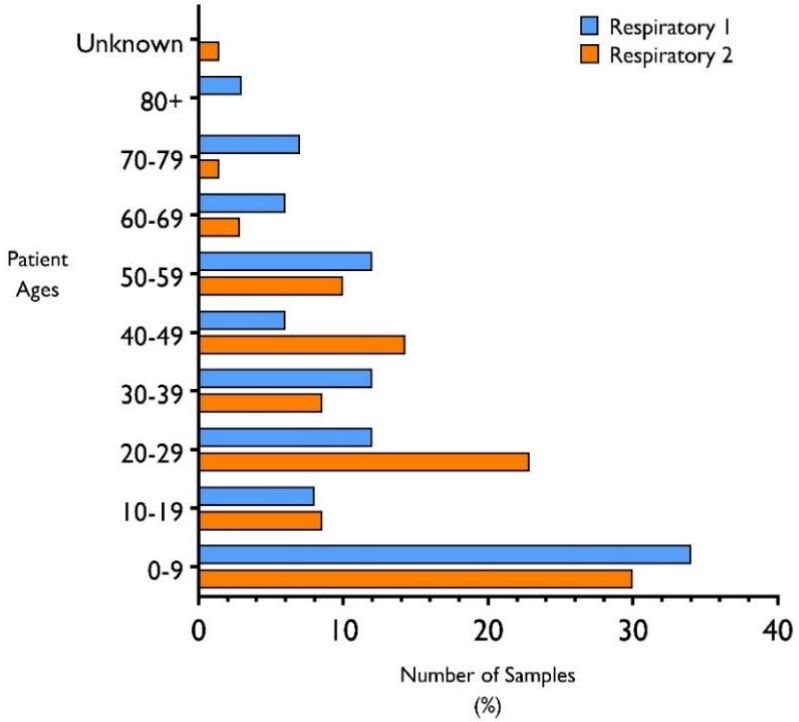

Figure 47 The ages of the patients at the time of sampling, whose samples were used to create the Resp-1 and Resp-2 pools, given as a percentage of the total number of patients.

6.2.2. Tapestation Analysis of Pooled RNA and Libraries

The Pooled RNA was analysed before creating the HTS libraries, as RNA concentration and RIN (RNA Integrity Number) influences the protocol. The libraries were then reanalysed before sequencing, to confirm sufficient yield cDNA concentration and whether correct fragment sizes had been created (Table 12)

	Pre-Library Creation		Post-Library Creation	
Library	**RNA Concentration** (ng/µl)	**RIN**	**cDNA Concentration** (ng/µl)	**Peak Fragment size (bp)**
CSF-1	11.4	4.5	0.47	274
CSF-2	2.1	4.5	0.33	256
EDTA-1	104	4.5	0.24	281
Resp-1	37	4.5	0.94	279
Resp-2	302	5.5	0.4	336

Table 12 Results of Tapestation analysis of the pooled RNA, and the cDNA sequencing libraries following rRNA depletion. RIN is the RNA Integrity Number, and indicates the quality of RNA

6.2.3. *In silico* Processing of Sequenced Libraries

The sequenced libraries were initially processed to remove low-quality and host genomic sequences, the results of this processing are shown in Table 13.

Library	Raw reads	Merged Paired-Reads	Post QC and Host Filtration (% of Raw Reads)	Number of Contigs Created
CSF-1	45,361,108	22,680,554	4,527,639 (9.98%)	136,497
CSF-2	55,584,272	21,082,479	7,784,986 (14%)	281,070
EDTA-1	35,761,102	15,388,741	11,442,721 (31.99%)	
Resp-1	60,692,430	27,592,279	3,819,653 (6.29%)	173,553
Resp-2	244,808,674	111,751,983	11,201,496 (4.97%)	279,331

Table 13 Number of reads sequenced from each library, and the number of low-quality and human reads removed during processing. Also included is the number of contigs generated from the filtered reads.

6.2.4. Virus discovery

6.2.4.1. Human Pegivirus

Reads matching Human Pegivirus (HPgV) were detected in the CSF-1, CSF-2 and EDTA pools.

BLASTx searching of the EDTA-1 contigs against the RNA polymerase AA database detected nine HPgV positive contigs. Mapping of unassembled reads against the reference genome for HPgV-1 detected 29 positive reads. These reads mapped to multiple regions, covering 38.4% of the reference genome; the overall pairwise identity of these reads to the reference genome was 91.6% (Figure 48).

A single read from the CSF-1 pool mapped to the HPgV reference genome. This read was 85bp in length, corresponding to positions 8599-8684, a region encoding the NS5B protein, the read was 91.8% identical to the reference sequence. BLASTn searching showed that

its closest related strain was a genotype-2 isolate of HPgV, with an identity of 95.5%. (Figure 49).

Two reads from the CSF-2 pool were mapped to the reference coding sequence for HPgV. The first read was an 85bp sequence corresponding to positions 6513-6598, in the NS5A gene; the read was 91.8% identical to the reference sequence. The read was subsequently BLASTn searched against the GenBank nucleotide collection, where the most identical (97.6%) to three genotype 2 HPgV isolates. The second mapped read was 213bp in length and mapped to positions 7912-8053, part of the NS5B gene. This read was 92% identical to the reference genome, and when BLASTn searched, was 95.8% identical to two genotype-2 isolates, different from those matching read 1 (Figure 50).

Figure 48 Coverage of EDTA-1 reads mapping to the HPgV reference genome (NC_001710.1)

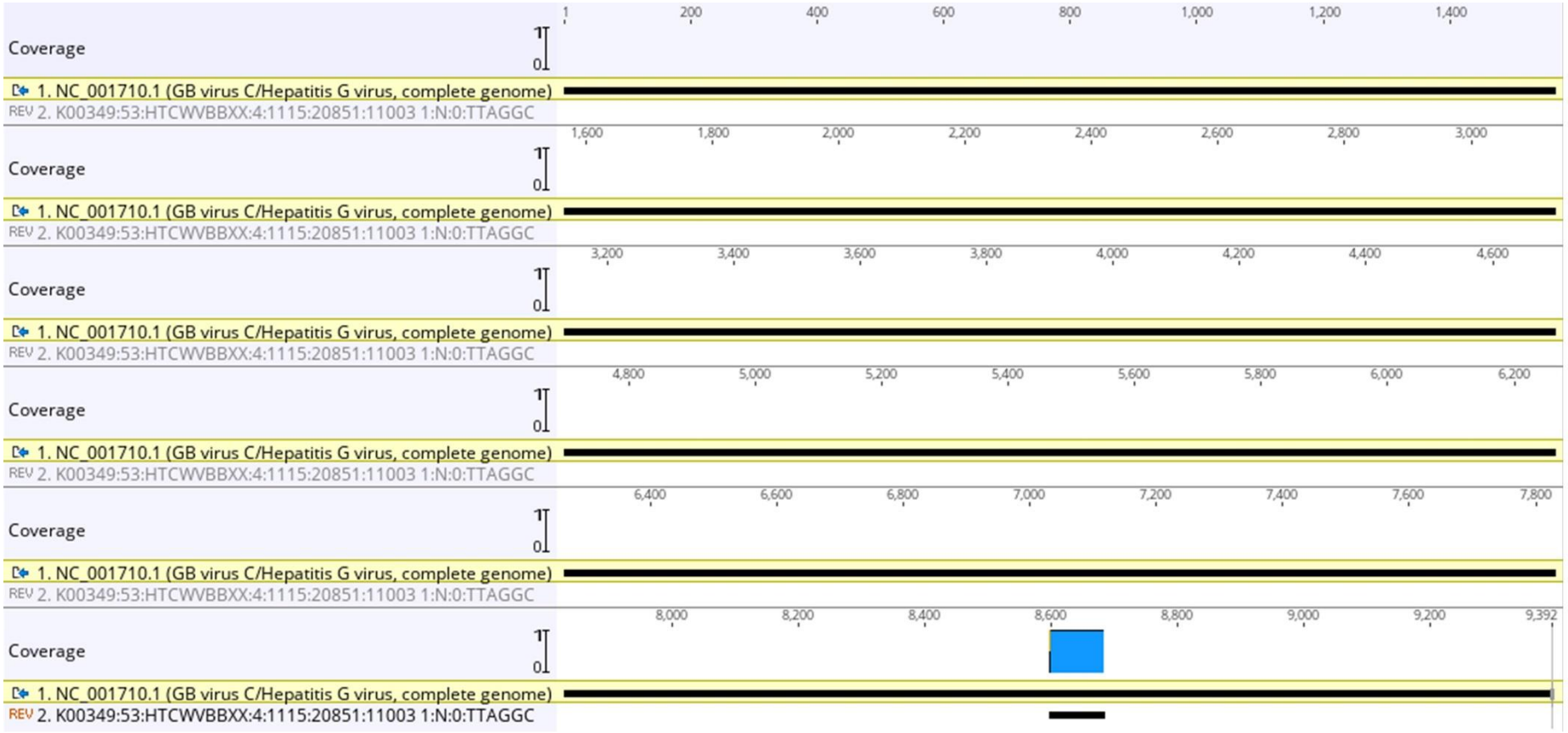

Figure 49 Coverage of CSF-1 reads mapping to the HPgV reference genome (NC_001710.1)

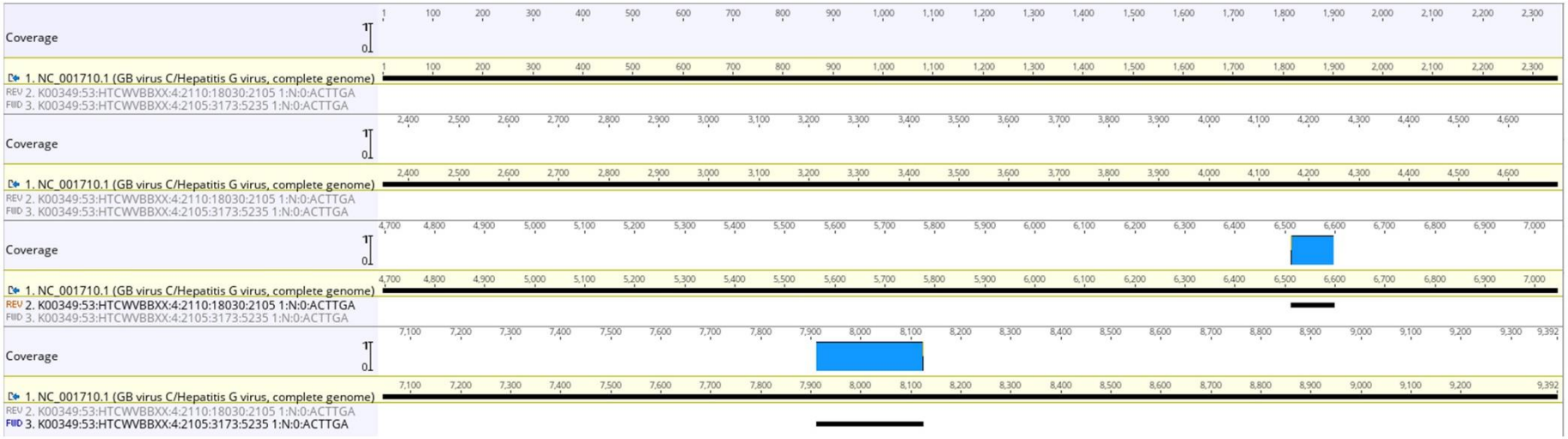

Figure 50 Coverage of CSF-2 reads mapping to the HPgV reference genome (NC_001710.1)

6.2.5. Screening Samples for Pegivirus

6.2.5.1. CSF-1

Constituent samples from CSF-1 were screened with HPgV-1 specific primers.

A sample from sub-pool 5 was positive. This patient was a 35-year-old female, with 'Headaches' as the clinical detail associated with the sample (#17X617313). Further clinical information associated with the sample revealed that the patient has complained of severe headaches for nine days, vomiting, neck stiffness and blurring of vision. Analysis of the CSF showed a red-blood-cell (RBC) count of 16, a Peripheral blood mononuclear cell (PMNC) count of 10, a lymphocyte count of 112 and a protein level of 476 ng/μl. These results led to the diagnosis of viral meningitis, the patient subsequently made a full recovery and was discharged.

A second HPgV positive patient, from sub-pool 14 was a 52-year-old female, and like the first patient, had 'Headaches' (#17X612694). The CSF had an RBC count of 2, a lymphocyte count of 1, a protein level of 323 ng/μl and a glucose level of 3.8 mmol/l., no information for the PMNC levels were available. This patient was also diagnosed with suspected viral meningitis.

6.2.5.2. CSF-2

The sub-pools used to create CSF-2 were also screened and HPgV was detected in a single sub-pool, pool 21. However, several of the individual samples that made up this sub-pool were used completely

in the creation of the sequencing library, and no HPgV was detected in any of the remaining samples.

6.2.5.3. EDTA-1

A total of five samples from the EDTA-1 sub-pools were positive.

The HPgV-1 positive patients had all presented with different clinical disorders. The positive samples from sub-pool 4 was a 17-year-old female presenting with pancytopenia following a liver transplant 2.5 years before the sample collection (#16X617094), and a 35-year-old female who had been admitted with unexplained splenomegaly which had subsequently resolved (#17X602891). Two samples from sub-pool 6 were positive, a 35-year-old female who had received a liver transplant

1.5 years before sampling, with no further available clinical available (#17X614574), and an 83-year-old male who was immunosuppressed due to peripheral T-cell lymphoma, presenting with biliary sepsis, abnormal liver function tests and a six-week history of fevers (#17X615088). The fifth HPgV-1 positive sample was in sub-pool 7, from a 76-year-old female who was receiving dialysis treatment following a failed renal transplant, which was received 3 years before sampling (#17X616426).

6.2.5.4. Phylogenetic Analysis

The PCR products from the PCR screening were sequenced, and a 225bp fragment of the NS5B RNA polymerase gene of HPgV. All HPgV sequences from the samples were divergent from one another, with the most similar strains being 90.1% identical to one another (Table 14).

		CSF		EDTA				
		17X 617313	17X 612694	16X 617094	17X 602891	17X 614574	17X 615088	17X 616426
CSF	17X617313	100	-	-	-	-	-	-
	17X612694	88.5	100	-	-	-	-	-
EDTA	16X617094	83.9	85.3	100	-	-	-	-
	17X602891	87.9	88.9	87.7	100	-	-	-
	17X614574	87.7	92.6	85.3	87.7	100	-	-
	17X615088	86.5	83.9	85.3	86.1	84.5	100	-
	17X616426	90.1	90	83.5	86.9	90.1	85.1	100

Table 14 A Distance matrix comparing the nucleotide similarity of the 225bp HPgV-1 NS5B amplicons of samples from the CSF-1, CSF-2 pools and EDTA-1 pools.

6.2.6. Evaluation of Cloud-based Bioinformatics Platforms.

6.2.6.1. IDseq

The analysis took 92 minutes from initiating the sample upload, to the returning of a report of the results of the mapped reads. Of the 6,339,908 reads in the dataset, 1,362,725 of these reads passed host filtering, 21.5% of the total dataset. Of these non-host reads, 996,855 reads (73.2%) were mapped to sequences from the bacterial

nucleotide (nt) databases. All four of the viruses added to the dataset were detected by IDseq, in either the nucleotide or non-redundant protein databases. The sensitivity of the IDseq analysis was lower than the median of the published protocols, though the heavily divergent avian bornavirus was detected unlike in the majority of other protocols (Table 15)

6.2.6.2. Genome Detective

Genome-Detective identified and removed 6,290,069 reads as non-viral hits, 99% of the dataset. The analysis was completed in 16

		Sensitivity (%)				Time (Hours)
		Torque teno virus	*Human herpesvirus 1*	*Measles virus*	*Avian bornavirus*	
Participant 1-13 Median (Range)		100 (0-102)	99 (10-400)	100 (0-140)	0 (0-100)	15.5 (3-216)
IDseq	**Nt**	59	56	69	0	1.5
	Nr	59	55	70	53	
Genome Detective		100	84	82	41	0.25

Table 15 A Comparison of the viral reads identified by IDseq and Genome-Detective, and the COMPARE virus proficiency tests that were presented in Brinkmann *et al.*

minutes. Genome Detective is designed for only analysis of viruses, and so did not filter bacterial hits separately to human sequences. All four viruses were detected in the dataset; detection of Torque teno virus was as sensitive as the median of the tested protocols, whilst Human herpesvirus 1 and Measles virus had a comparatively lower sensitivity (Table 15).

6.2.7. Reanalysing Clinical HTS Datasets Using IDseq

6.2.7.1. Quality Control and Host Filtering

Libraries were submitted to IDseq for reanalysis. The QC and host trimming results are shown in (**Table 16**)

Library	Number of Reads	Passed Quality Control	Passed Host Filtration (% of Raw Reads)	Total Runtime
CSF-1	45,361,108	81.65%	13,100,430 (28.88%)	4 Hours 8 Minutes
CSF-2	55,584,272	43.84%	554,920 (0.98%)	2 Hours 40 Minutes
EDTA-1	35,761,102	70.15%	107,340 (0.3%)	2 Hours 11 Minutes
Resp-1	60,692,430	86.74%	24,946,471 (41.1%)	4 Hours 22 Minutes

Table 16 Number of reads filtered by quality and host genome of the libraries submitted to IDseq for reanalysis

6.2.7.2. Detected Viruses

IDseq detected several virus species across each of the pools (**Figure 51**).

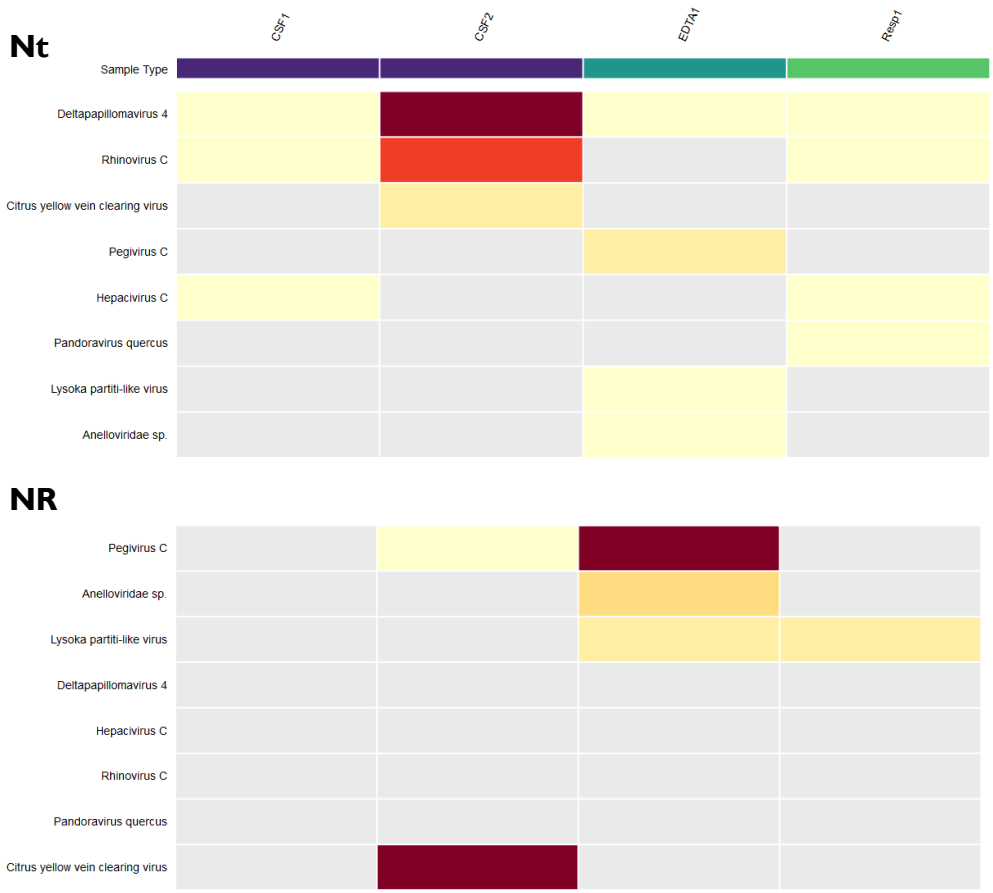

Figure 51 IDseq generated heatmap of reads mapping to virus species at nucleotide (Nt) or non-redundant proteins (NR) levels. Darker colours indicate more reads.

Deltapapillomavirus 4 was detected in all four of the pools, CSF-2 had 320 reads, the most in any of the pool; CSF-1 had 10 reads, EDTA-1 had 4 and Resp-1 had 2.

Rhinovirus C was detected in both CSF pools and Resp-1. There were 210 reads in CSF-2, 10 in CSF-1 and 2 in Resp-1. Pegivirus C (HpGV) was detected in only the EDTA-pool when reads were compared to the nucleotide database, with 29 matched reads. HPgV

was detected in CSF-2 when the *non-redundant* protein database was used, with 2 reads mapping. Pegivirus A was detected in the CSF-2 and EDTA-1 pools, with 33 Nt and 16 NR matching reads in the EDTA-1 pool and 4 Nt reads in CSF-2. Other virus hits were Citrus Yellow Vein Clearing Virus in CSF-2 (44 Nt hits and 42 NR hits), Pandoravirus quercus in Resp-1 (6 Nt hits), Hepacivirus C in CSF-1 (2 Nt hits) and Resp-1 (9 Nt hits), lysoka partiti-like virus in EDTA-1 (2 Nt hits, 3 NR hits) and Resp-1 (6 NR hits), and an unclassified Anneloviridae in EDTA-1 (10 hits to both Nt and NR.

6.3. Discussion

6.3.1. Samples Selection

This study was designed to make use of the surplus samples that were no longer required by the NHS diagnostic laboratory. The samples used were surplus nucleic acids extracts that had been sent for virus PCR screening but were negative. Ethical approval for these samples had been granted previously, for extended diagnosis

The use of surplus samples allowed for access to a massive cohort of samples that would have otherwise been disposed of. However, as all samples were extracted nucleic acids, there was no opportunity to choose extraction protocols suited to the requirements of HTS preparation. Additionally, several samples had only minute amounts of nucleic acid remaining, which meant that these samples were used completely in the creating the sequencing library, with no opportunity to re-extract and could not be used for follow-up screening.

These samples had only a limited amount of clinical details available initially, this included age, sex, sample type, sampling date, results of virus screening and a very brief clinical detail summary (often only a single word). Further details for samples had to be requested from a clinician, which was excessively time-intensive to use on the entire database of candidate samples and was reserved for virus-positive samples.

As these samples had all been submitted for virus screening, the assumption had been made that each of the patients had symptoms

characteristic of a viral infection, and thus were likeliest to be virus-positive. However, whilst information was available for the results of viral PCR assays, no information was initially available for any other diagnostic assays that had been used on these samples, or samples taken from the same patients. Which means that patients may have been positive for bacterial, fungal or parasitic pathogens, and were the etiological agent of the patient's disease.

6.3.2. Detection of Human Pegivirus

6.3.2.1. Background of Human Pegivirus

Human Pegivirus (HPgV) was detected in the CSF-1 and EDTA pools.

HPgV is a member of the *Pegivirus C* species and was previously classified as Hepatitis G or GBV-C [291]. They are single-stranded, positive sensed RNA viruses, in the *Flaviviridae* family. Their genome is typical of a *Flaviviridae*, a single open reading frame (ORF) encoding two structural proteins E1 and E2, and six non-structural proteins NS2, NS3, NS4A/B and NS5A/B, flanked by a long 5' UTR and 3' UTR. The complete genome is approximately 9.3kb in length.

HPgV was discovered in non-human primates who had developed hepatitis following injection of sera from a surgeon who had hepatitis of unknown aetiology [292]. This led to the hypothesis that HPgV was a causative agent of hepatitis, like the closely related Hepatitis C virus (HCV). Later studies, however, failed to establish a causal relationship between hepatitis and HPgV [293].

6.3.2.2. HPgV positive EDTA Blood Samples

Five patients of the EDTA-1 pool were found to be positive for HPgV-1 infection. Prevalence of HPgV has been shown to be remarkably high in healthy cohorts of blood donors, between 1-5% in cohorts from developed countries and 20% in developing countries [294]. Prevalence in the blood-donors in the UK has been reported at 1% in Scotland [295], and 2.25% in Nottingham [296]. The

prevalence in the cohort of samples used in this study was 6.25%, much higher than previous reports. Those with exposure to contaminated blood, such as injecting drug users, have a much higher incidence of HPgV-1 [297]. Whilst no information was available regarding any high-risk behaviours such as drug injection was available for these samples, most had a relevant clinical history. Three of the positive samples had all recently received solid-organ transplants, two received livers and one a kidney (which was rejected). Both kidney and liver transplantations, as well as heart and bone-marrow transplants, have been associated with increased HPgV prevalence in recipients [298–300], though there was no correlation between infection and different outcomes in any of the recipients. Human Pegivirus is considered a blood-borne virus and transmission occur primarily through contact with contaminated blood, though sexual [301] and mother-child [302,303] transmission has been reported. Contaminated transfusion blood is a known route of transmission of HPgV, and has been associated with infection of a liver recipient during transplantation [304]. It is impossible to determine how and when the three organ recipients became infected with HPgV, but given the prevalence of the virus in blood-donors and the lack of screening [305], transmission could have occurred as a result of transfusing contaminated blood during transplantation.

One of the patients was diagnosed with peripheral T-cell lymphoma. Human Pegivirus viremia has been associated with an increased incidence of lymphoma, an association which persists when infection

risk-factors are controlled for [306,307]. In the UK, 10% of patients attending a lymphoma clinic had HPgV positive sera [296].

It is unclear by what mechanisms HPgV infection contributes to the development of lymphoma. Increased levels of cytokines are associated with lymphoma development and have been shown to be higher in HPgV-positive lymphoma patients [308]. Furthermore, HPgV-1 is a lymphotropic virus [309], causing disruption of lymphocytes and negatively influences immune-surveillance, another known risk factor for the development of lymphomas [310]. Whether HPgV infection influenced the development of lymphoma in the patient is unknown, as is whether infection preceded lymphoma at all. No further samples were available for this patient from earlier dates, which may have determined whether infection preceded the diagnosis of cancer. It should be emphasised that only correlation can be shown with this, single positive sample, and not causation.

6.3.2.3. HPgV in the CSF

Human Pegivirus was detected in two samples in the second CSF pool. Both patients were adult women, who had initially presented with headaches, and were subsequently diagnosed with meningitis. Whilst HPgV is typically associated with blood-borne infection, several reports have detected HPgV is patients with neurological disease. Kriesel *et al* detected HPgV in the brain of a woman who had died of primary-progressive multiple sclerosis [311]. Liu *et al* later detected HPgV in the CSF of a patient with HIV-1, cerebral toxoplasmosis and fungal encephalitis. This patient, however, had a

severely compromised blood-brain barrier, which may have allowed the transfer of the virus from the blood into the CSF, and therefore may not represent a true neurological infection [312]. Human Pegivirus was detected in a patient with severe encephalitis using a metagenomic pan-microbial array, which failed to detect any other known pathogens in the sample [313]. Retrospective screening of patients with encephalitis of unknown origin detected HPgV in three patients. Further analysis of matched serum samples from these three patients showed that the sequences from viruses in the CSF and serum differed slightly, which indicates that viruses were replicating independently from one another and CSF infection is unlikely to be a result of a compromised blood-brain barrier [314].

This is the first association of HPgV and meningitis. Importantly, this association cannot be determined to be causal, given the small sample number. Both CSF samples contained a small number of red blood cells, the low count is unlikely to be the result of a traumatic lumbar puncture but may be a result of a compromised blood-brain barrier. Further samples from either of the patients were unavailable, which precluded any further analysis into whether HPgV was detectable in blood, and whether viruses from the two samples were genetically distinct from one another, as was reported in Bukowska-Ośko *et al* [314].

6.3.4. Evaluation of Cloud-based Bioinformatics Platforms.

The results of the comparison between cloud-based platforms and more traditional analytic pipelines show that online platforms can accurately identify viral genomes in HTS datasets. The proficiency testing by Brinkmann *et al* [268] explored several virus discovery pipelines, which encompassed numerous software and ranges of computational power. Whilst these cloud-based platforms were not as sensitive as some of the pipelines using high-specification computational power, the processing times were significantly faster than any of the traditional pipelines.

Despite the loss in sensitivity, all viruses were nevertheless detected, including the divergent avian bornavirus, which shows that these platforms are suitable for virus-discovery. Whilst they may not be of immediate use or interest to a trained bioinformatician with access to sufficient computational processing power, the simplicity of these platforms will allow researchers with little or no background into bioinformatics or funds to acquire sufficiently powerful computers to fully exploit the potential of HTS for virus discovery.

Several studies have begun to use these platforms, IDseq in particular, for virus discovery in retrospective clinical samples [287,290].

6.3.5. Analysing HTS Datasets with IDseq

6.3.5.1. Quality Control and Host Filtration

IDseq QC was significantly faster than the in-house developed geneious-based pipeline. The analysis of reads took a matter of hours

to complete, rather than the days that were required for the geneious pipeline. The substantial increase in computational power used by IDseq was the most significant factor in this reduction of runtime; where the geneious pipeline only used a single PC with moderate specifications, IDseq, however, employs a series of cloud-based software packages, scalable to the size of the dataset, and allow for the simultaneous analysis of different portions of the dataset [287].

IDseq was able to filter more host and low-quality reads from CSF-2 and EDTA-1 than geneious but was filtered less from CSF-1 and Resp-1. Different software packages are used by IDseq than geneious, which would account for the variation of filtered reads.

6.3.5.2. Pegivirus C

Pegivirus C, or HPgV, was detected in CSF-2 and EDTA-1, but not in CSF-1 where it had previously been detected with the geneious pipeline. However, only a single HPgV read had been detected in CSF-1, and so may have been missed due to chance. The number of reads mapping to the HPgV Nt in CSF-2 and EDTA-1 were the same as those mapping using geneious, which would suggest that there is no difference in the sensitivity of the two pipelines.

6.3.5.3. Human Rhinovirus C

Rhinovirus C was detected in both CSF pools and Resp-1. Human Rhinovirus C (HRV-C) is a member of the *enterovirus* genus in the *Picornaviridae* family and was initially detected in 2007 [315]. There is a global distribution of these viruses, with prevalence ranging from 1.4

to 30.9%, though the prevalence of >5% is typical [316]. They have been detected primarily in respiratory samples, where they have been associated with respiratory diseases such as acute lower respiratory tract disease [317] and bronchiolitis [318]. Detection of HRV-C in the Resp-1 pool would likely represent a typical respiratory infection in one, or several, of the patients. Detection in CSF has been reported in patients without neurological symptoms [319], as well as in patients with meningitis, though a causal link between the two was not established [320]. Detection in both CSF pools is therefore not unexpected, though whether the HRV-C was the etiological agent of the disease in the positive patients is unclear.

6.3.5.4. Hepacivirus C

Hepatitis C (HCV) is the sole virus in the *Hepacivirus C* genus [321], it is a well-known and characterised human pathogen. Globally, approximately 142 million people are infected, and infection can lead to several liver diseases such as cirrhosis and liver cancer [322]. HCV was detected in the CSF-1 pool; detection of HCV in the CSF has been reported [323], though is uncommon in acute infections. Neurological disorders have been reported in 50% of chronic HCV cases [324], and HCV infection is associated with severe neurological diseases such as encephalitis [325] and encephalomyelitis [326]. Whether the HCV-positive patient (or patients) were diagnosed with neurological disease is unknown, though as many of the samples pooled into CSF-1 had neurological symptoms, it is possible that they were. As initial clinical information for samples was limited, it is

possible that known HCV-positive samples were included in the sequencing pools. Hepatitis C is not included in the standard neurological viral diagnostic assay, and so HCV positive patients would still appear as a virus-negative sample.

HCV was also detected in a respiratory pool. HCV is typically not associated with the respiratory system, so whether this represents a respiratory infection is unclear. It is likelier that contamination of the sample has occurred, for example, if the patient had a cut in their mouth, then blood may have contaminated the throat swab or lavage.

6.3.5.5. Lysoka partiti-like virus

Detected in both the EDTA and Resp-1 pools, Lysoka partiti-like virus is a picobirnavirus (PBV) that was detected in a Straw-coloured fruit bat (*Eidolon helvum)* from Cameroon [327]. Picobirnaviruses are double-stranded RNA viruses and have been detected in humans. Human infection is typically associated with gastrointestinal disease [328], however, it has also been detected in the respiratory tract of patients with unexplained respiratory disease [329]. Whether the picobirnaviruses from the respiratory samples are responsible for disease in the patients from Resp-1 is unknown; no evidence of a causal association between picobirnavirus and respiratory disease has been proven, and so it is likelier that the viruses from Resp-1 are commensal.

Detection of PBV in human blood has not previously been reported. Bovine picobirnavirus has been detected in the plasma of cattle,

though contamination from faeces on the skin was suggested [330]; faecal contamination of the skin site used to draw blood should be unlikely, assuming correct protocols were followed. Cross-contamination between sequencing libraries is another possibility.

Despite the reads mapping to a bat-associated genotype of PBV, it is unlikely that this a bat zoonoses. PBVs show significant sequence divergence, compared to other genera with similar genomes, and that phylogenies of gene fragments are not necessarily representative of the phylogeny of the virus as a whole [331]. The further genome of the PBV in these samples needs to be sequenced to accurately determine the strain circulating within these samples.

6.3.5.6. *Anelloviridae*

Anelloviruses are small, circular DNA viruses [332]. An Anellovirus was detected in the EDTA-1 pool, the reads mapped to unclassified anellovirus species. Anelloviruses are highly prevalent in humans [333], and whilst they are found in healthy populations[334], several species have been associated with disease in humans, such as fever [335] and hepatitis [336]. Given the prevalence of Anelloviruses in healthy humans and the lack of any further information regarding which of the samples are positive, this strain of anellovirus cannot be associated with a particular disease.

.

6.3.5.7. Potential Contaminants

Citrus yellow vein clearing virus CYVCV was detected in CSF2. CYVCV is an RNA virus, responsible for disease in citrus trees [337]. There is currently no published evidence of CYVCV infection anything other than plants, and no evidence of any plant virus causing disease in humans [338]. It is highly likely that CYVCV is a contaminant, though whether contamination occurred during the collection of CSF, processing of the nucleic acid extractions or creation of the sequencing library is unknown.

Pandoravirus quercus is a 'Giant' DNA virus, which typically infects amoeba [339]. Similar amoeba-associated giant viruses have been detected in humans and associated with respiratory disease [340]. It is possible that the Pandoravirus quercus detected here represents a true infection, though detection of this virus indicates the likely presence of its typical host *Acanthamoeba*, which has previously been detected in patients with atypical pneumonia [341]. Presence of pandoravirus in the samples could simply be a result of environmental contamination, and the virus unrelated to the disease in the patient.

Deltapapillomavirus 4, or Bos Taurus papillomavirus (BPV), was detected in all four pools, with the majority in CSF-2. BPV typically infects ruminants, causing cutaneous papillomas and bladder and gastrointestinal carcinomas [342]. As papillomaviruses are typically associated with a dermatological disease, they are rarely detected in CSF, which combined with the lack published reports of transmission of BPV into humans, suggests that BPV detection is a result of

contamination. Zebu (*Bos indicus*) blood pools were included in the Resp-2 pool, and cross-contamination between pools may have occurred during the creation of libraries.

6.4. Conclusions

In this chapter, several cohorts of undiagnosed clinical samples were high throughput sequenced with an Illumina HiSeq, with the aim of recovering novel, misdiagnosed or zoonotic viruses. Human Pegivirus was detected in the CSF of several patients who had been diagnosed with viral meningitis. Whilst a causal relationship cannot be proven to between HPgV-1 and meningitis at this time, this initial association between the two may lead to further investigations, where a causal relationship could be established. Several other HPgV-1 strains were detected in the blood of immunocompromised patients, including one with lymphoma.

Several other viruses were detected in these samples, though due to time constraints, theses were unable to be followed up for further investigation.

The effectiveness of cloud-based bioinformatics platforms was compared with several custom pipelines, using an *in-silico* dataset. This demonstrated that these online platforms are an effective tool for researchers with limited access to funds for high-specification local computing power, or experience in bioinformatics.

7. Final Conclusions

The primary aim of this project was to assess the prevalence of viruses with zoonotic potential in cohorts of British, Polish and Egyptian rodents, and to characterise any detected viruses. An orthohantavirus, Tatenale virus, was detected in field voles captured at Chester zoo, several miles from the site where Tatenale virus was initially detected in 2013. A single field vole from Twycross zoo was also positive for Tatenale virus, this extended the known range of the virus into the East Midlands region of England. There were no orthohantaviruses detected in any other species of rodents captured at these sites, or the sites in Poland or Egypt. High-throughput sequencing recovered the complete coding sequence of the Twycross zoo strain of Tatenale virus (Norton-Juxta), whilst almost complete coding sequence of the Chester zoo strain (Upton-Heath) was recovered with primer-walking PCR. Comparison of these strains showed both strains were 90.6%, 94.1% and 91.3% similar at nucleotide level across the L, S and M segments respectively. The Upton-Heath strain was highly similar to the first recovered strain of Tatenale virus (B41), with a similarity of 98.7% and 96.9% across fragments of S and L; given the similarities in the genome and geographic origin, they are likely circulating within the same populations of vole. The Norton-Juxta strain, however, is dissimilar to all previously detected strains, 87.5% and 93.9% similarity across the S and L, to the closest relative. Furthermore, this is the first recovery of the complete coding sequence of Tatenale virus, which has confirmed that Tatenale virus is a novel species of

orthohantavirus. Previous sequence data of Tatenale virus was insufficient to meet the ICTV criteria for assigning species status, which required complete S and M segments. In addition to the detection of Tatenale virus, viruses from four genera of picornaviruses were detected; Ljungan virus and rosavirus in Polish bank voles, cardiovirus in British brown rats and hunnivirus in British house mice. High-throughput sequencing failed to recover complete genomes, precluding any attempts into fully characterising these viruses.

A secondary aim was to investigate the seroprevalence of orthohantaviruses in non-human primates and humans in the East Midlands. There was evidence of previous orthohantavirus infection in chimpanzee's and bonobos from Twycross Zoo, but not Gorillas, Orangutans or several species of Guenon. Twycross zoo is the site in which the Norton-Juxta strain of TATV was recovered, suggesting that TATV may have infected these primates, though the lack of specificity in the ELISA assay precludes confirmation of a causative species. Additionally, orthohantavirus specific IgM antibodies were detected in a cohort of intravenous drug abusers with an unstable lifestyle in Nottingham, demonstrating that orthohantavirus infection is prevalent in at-risk populations.

An additional secondary aim of the project was to investigate undiagnosed clinical CSF, blood and respiratory samples for the presence of novel or atypically presenting viral pathogens. Samples of the same type were pooled, and high throughput sequenced. Multiple

pathogens were detected in these samples, the most significant of these was human Pegivirus. Pegivirus was detected in several patients, most notably in the CSF of patients who had been diagnosed with viral meningitis. However, further work would be required to demonstrate a causative relationship between Pegivirus and meningitis. Furthermore, the reads from these samples, as well as an artificially simulated clinical HTS library, were reanalysed with an online analytic pipeline, IDSeq. This pipeline was able to analyse the reads to a comparable degree of accuracy as the in-house pipeline and was able to do so significantly quicker, demonstrating that it would be a useful tool for researchers with limited funding and bioinformatics expertise.

8. Future Work

There are several sections of this project that with further funding and time could be greatly developed upon. The pathogenicity of TATV remains unknown, as the previous paucity of sequence data has precluded any in-silico investigation. With the recovery of the sequence of both glycoproteins, it is now possible to explore cellular entry and tropism of the viruses. The Pseudotyping process, in which glycoproteins of interest are expressed in non-replicative recombinant retrovirus backbones, allows the investigation of viral entry without the associated risk of infection and therefore does not require the increased biosafety requirements of live-virus work. Pseudotyping work in orthohantaviruses has been well described [260,343] which shows that this technique is a feasible route for exploring the possibility of Tatenale infectivity in human cells. Further serological screening of orthohantaviruses is another future direction for this work. The orthohantavirus nucleocapsid is often used as a target for ELISA's [344,345], so the now complete nucleocapsid sequence of Tatenale orthohantavirus can be used for serological studies to determine whether TATV had infected the seropositive non-human primates or humans.

Repeating the high-throughput sequencing of the picornavirus samples, to retrieve full sequence, would allow further characterisation of these viruses. It is unclear why the initial high-throughput sequencing failed; re-extracting samples and sequencing the samples on an Illumina HiSeq again may recover further sequence

data, alternatively sequencing the samples on a different platform, such as a MinIon, may recover further sequence data.

9. References

1. WHO | Zoonoses. WHO. World Health Organization; 2017; Available: https://www.who.int/topics/zoonoses/en/

2. Messenger AM, Barnes AN, Gray GC. Reverse zoonotic disease transmission (zooanthroponosis): a systematic review of seldom-documented human biological threats to animals. PloS one. Public Library of Science; 2014;9: e89055. doi:10.1371/journal.pone.0089055

3. Bueno-MarÃ- R, Almeida APG, Navarro JC. Editorial: Emerging Zoonoses: Eco-Epidemiology, Involved Mechanisms, and Public Health Implications. Frontiers in Public Health. Frontiers; 2015;3: 157. doi:10.3389/fpubh.2015.00157

4. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. Nature. Nature Publishing Group; 2004;430: 242–9. doi:10.1038/nature02759

5. Woolhouse MEJ, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. Emerging infectious diseases. 2005;11: 1842–7. doi:10.3201/eid1112.050997

6. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. Nature. Nature Publishing Group; 2008;451: 990–3. doi:10.1038/nature06536

7. Chomel BB, Belotto A, Meslin F-X. Wildlife, Exotic Pets, and Emerging Zoonoses. Emerging Infectious Diseases. Centers for Disease Control and Prevention; 2007;13: 6. doi:10.3201/EID1301.060480

8. Haydon DT, Cleaveland S, Taylor LH, Laurenson MK. Identifying reservoirs of infection: A conceptual and practical challenge. Emerging Infectious Diseases. Centers for Disease Control and Prevention (CDC); 2002. pp. 1468–1473. doi:10.3201/eid0812.010317

9. Reed KD, Meece JK, Henkel JS, Shukla SK. Birds, Migration and Emerging Zoonoses: West Nile Virus, Lyme Disease, Influenza A and Enteropathogens. Clinical Medicine & Research. 2003;1: 5–12. doi:10.3121/cmr.1.1.5

10. Briones V, Téllez S, Goyache J, Ballesteros C, del Pilar Lanzarot M, Domínguez L, et al. Salmonella diversity associated with wild reptiles and amphibians in Spain. Environmental microbiology. 2004;6: 868–71. doi:10.1111/j.1462-2920.2004.00631.x

11. Burnham BR, Atchley DH, DeFusco RP, Ferris KE, Zicarelli JC,

Lee JH, et al. Prevalence of fecal shedding of Salmonella organisms among captive green iguanas and potential public health implications. Journal of the American Veterinary Medical Association. 1998;213: 48–50. Available: http://www.ncbi.nlm.nih.gov/pubmed/9656023

12. Plourde BT, Burgess TL, Eskew EA, Roth TM, Stephenson N, Foley JE. Are disease reservoirs special? Taxonomic and life history characteristics. Baker ML, editor. PLOS ONE. Public Library of Science; 2017;12: e0180716. doi:10.1371/journal.pone.0180716

13. Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. Proceedings of the National Academy of Sciences. 2015;112: 201501598. doi:10.1073/pnas.1501598112

14. Wilson DE, Reeder D. Mammal Species of the World: A Taxonomic and Geographic Reference [Internet]. Baltimore: JHU Press; 2005. Available: https://books.google.com/books?id=JgAMbNSt8ikC&pgis=1

15. Luis AD, Hayman DTS, O'Shea TJ, Cryan PM, Gilbert AT, Pulliam JRC, et al. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proceedings Biological sciences / The Royal Society. 2013;280: 20122753. doi:10.1098/rspb.2012.2753

16. Taylor LH, Latham SM, Woolhouse ME. Risk factors for human disease emergence. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2001;356: 983–9. doi:10.1098/rstb.2001.0888

17. Field H, de Jong C, Melville D, Smith C, Smith I, Broos A, et al. Hendra virus infection dynamics in Australian fruit bats. PloS one. Public Library of Science; 2011;6: e28678. doi:10.1371/journal.pone.0028678

18. Chua KB, Goh KJ, Wong KT, Kamarulzaman A, Tan PS, Ksiazek TG, et al. Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia. Lancet. 1999;354: 1257–9. doi:10.1016/S0140-6736(99)04299-3

19. World Health Organisation. HIV/AIDS [Internet]. World Health Organization; 2015 [cited 1 May 2016]. Available: http://www.who.int/mediacentre/factsheets/fs360/en/

20. World Health Organisation. Ebola data and statistics [Internet]. World Health Organization; 2016 [cited 1 May 2016]. Available: http://apps.who.int/gho/data/view.ebola-sitrep.ebola-summary-20160427?lang=en

21. Smith G, Cruz M, Himelein K, Over M, Ferreira FHG, Evans D, et al. The economic impact of Ebola on sub-Saharan Africa : updated estimates for 2015. 2015; 1–18. Available: http://documents.worldbank.org/curated/en/2015/01/23831803/economic-impact-ebola-sub-saharan-africa-updated-estimates-2015

22. Woolhouse MEJ, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. Trends in Ecology & Evolution. 2005;20: 238–244. doi:10.1016/j.tree.2005.02.009

23. Woolhouse MEJ, Brierley L, McCaffery C, Lycett S. Assessing the Epidemic Potential of RNA and DNA Viruses. Emerging infectious diseases. Centers for Disease Control and Prevention; 2016;22: 2037–2044. doi:10.3201/eid2212.160123

24. Morse SS, Mazet JA, Woolhouse M, Parrish CR, Carroll D, Karesh WB, et al. Prediction and prevention of the next pandemic zoonosis. The Lancet. Elsevier; 2012;380: 1956–1965. doi:10.1016/S0140-6736(12)61684-5

25. Wolfe ND, Dunavan CP, Diamond J. Origins of major human infectious diseases. Nature. Nature Publishing Group; 2007;447: 279–83. doi:10.1038/nature05775

26. Drake JW, Holland JJ. Mutation rates among RNA viruses. Proceedings of the National Academy of Sciences. 1999;96: 13910–13913. doi:10.1073/pnas.96.24.13910

27. Moya A, Holmes EC, González-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. Nature Reviews Microbiology. Nature Publishing Group; 2004;2: 279–288. doi:10.1038/nrmicro863

28. Shackelton LA, Holmes EC. The evolution of large DNA viruses: combining genomic information of viruses and their hosts. Trends in Microbiology. Elsevier Current Trends; 2004;12: 458–465. doi:10.1016/J.TIM.2004.08.005

29. Guu TSY, Zheng W, Tao YJ. Bunyavirus: structure and replication. Advances in experimental medicine and biology. 2012;726: 245–66. doi:10.1007/978-1-4614-0980-9_11

30. McCaughey C, Hart CA. Hantaviruses. Journal of medical microbiology. Microbiology Society; 2000;49: 587–99. doi:10.1099/0022-1317-49-7-587

31. Gavrilovskaya IN, Brown EJ, Ginsberg MH, Mackow ER. Cellular entry of hantaviruses which cause hemorrhagic fever with renal syndrome is mediated by beta3 integrins. Journal of virology. 1999;73: 3951–9. Available: http://www.ncbi.nlm.nih.gov/pubmed/10196290

32. Gavrilovskaya IN, Shepley M, Shaw R, Ginsberg MH, Mackow ER. beta3 Integrins mediate the cellular entry of hantaviruses that cause respiratory failure. Proceedings of the National Academy of Sciences of the United States of America. National Academy of Sciences; 1998;95: 7074–9. doi:10.1073/pnas.95.12.7074

33. Raftery MJ, Lalwani P, Krautkrämer E, Peters T, Scharffetter-Kochanek K, Krüger R, et al. β2 integrin mediates hantavirus-induced release of neutrophil extracellular traps. Journal of Experimental Medicine. Rockefeller University Press; 2014;211: 1485–1497. doi:10.1084/JEM.20131092

34. Jangra RK, Herbert AS, Li R, Jae LT, Kleinfelter LM, Slough MM, et al. Protocadherin-1 is essential for cell entry by New World hantaviruses. Nature. Nature Publishing Group; 2018;563: 559–563. doi:10.1038/s41586-018-0702-1

35. Müller A, Baumann A, Essbauer S, Radosa L, Krüger DH, Witkowski PT, et al. Analysis of the integrin β3 receptor for pathogenic orthohantaviruses in rodent host species. Virus Research. Elsevier; 2019;267: 36–40. doi:10.1016/J.VIRUSRES.2019.04.009

36. Vaheri A, Strandin T, Hepojoki J, Sironen T, Henttonen H, Mäkelä S, et al. Uncovering the mysteries of hantavirus infections. Nature Reviews Microbiology. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2013;11: 539–550. doi:10.1038/nrmicro3066

37. Arai S, Nguyen ST, Boldgiv B, Fukui D, Araki K, Dang CN, et al. Novel bat-borne hantavirus, Vietnam. Emerging Infectious Diseases. 2013;19: 1159–1161. doi:10.3201/eid1907.121549

38. Weiss S, Witkowski PT, Auste B, Nowak K, Weber N, Fahr J. Hantavirus in Bat, Sierra Leone. Emerging infectious diseases. 2012;18: 159–161. doi:10.4161/pri.1.1.3828

39. Klempa B, Fichet-Calvet E, Lecompte E, Auste B, Aniskin V, Meisel H, et al. Novel hantavirus sequences in Shrew, Guinea. Emerging Infectious Diseases. 2007;13: 520–522. doi:10.3201/eid1303.061198

40. Arai S, Song JW, Sumibcay L, Bennett SN, Nerurkar VR, Parmenter C, et al. Hantavirus in northern short-tailed shrew, United States. Emerging Infectious Diseases. 2007;13: 1420–1423. doi:10.3201/eid1309.070484

41. Kang HJ, Bennett SN, Sumibcay L, Arai S, Hope AG, Mocz G, et al. Evolutionary insights from a genetically divergent hantavirus harbored by the European common mole (Talpa europaea). PloS

one. Public Library of Science; 2009;4: e6149. doi:10.1371/journal.pone.0006149

42. Witkowski PT, Drexler JF, Kallies R, Ličková M, Bokorová S, Mananga GD, et al. Phylogenetic analysis of a newfound bat-borne hantavirus supports a laurasiatherian host association for ancestral mammalian hantaviruses. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2016;41: 113–119. doi:10.1016/j.meegid.2016.03.036

43. International Committee on Taxonomy of Viruses. ICTV Virus Taxonomy [Internet]. 2014. Available: http://www.ictvonline.org/virusTaxonomy.asp

44. Kallio ER, Voutilainen L, Vapalahti O, Vaheri A, Henttonen H, Koskela E, et al. Endemic hantavirus infection impairs the winter survival of its rodent host. Ecology. 2007;88: 1911–1916. doi:10.1890/06-1620.1

45. Easterbrook JD, Zink MC, Klein SL. Regulatory T cells enhance persistence of the zoonotic pathogen Seoul virus in its reservoir host. Proceedings of the National Academy of Sciences of the United States of America. 2007;104: 15502–7. doi:10.1073/pnas.0707453104

46. Guivier E, Galan M, Salvador AR, Xuéreb A, Chaval Y, Olsson GE, et al. Tnf-α expression and promoter sequences reflect the balance of tolerance/resistance to Puumala hantavirus infection in European bank vole populations. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2010;10: 1208–17. doi:10.1016/j.meegid.2010.07.022

47. Plyusnin A, Sironen T. Evolution of hantaviruses: Co-speciation with reservoir hosts for more than 100MYR. Virus Research. Elsevier B.V.; 2014;187: 22–26. doi:10.1016/j.virusres.2014.01.008

48. Spengler JR, Haddock E, Gardner D, Hjelle B, Feldmann H, Prescott J. Experimental Andes virus infection in deer mice: characteristics of infection and clearance in a heterologous rodent host. PloS one. Public Library of Science; 2013;8: e55310. doi:10.1371/journal.pone.0055310

49. Ramsden C, Holmes EC, Charleston MA. Hantavirus evolution in relation to its rodent and insectivore hosts: No evidence for codivergence. Molecular Biology and Evolution. 2009;26: 143–153. doi:10.1093/molbev/msn234

50. Kang HJ, Bennett SN, Dizney L, Sumibcay L, Arai S, Ruedas LA, et al. Host switch during evolution of a genetically distinct hantavirus

in the American shrew mole (Neurotrichus gibbsii). Virology. 2009;388: 8–14. doi:10.1016/j.virol.2009.03.019

51. Schlegel M, Klempa B, Auste B, Bemmann M, Schmidt-Chanasit J, Büchner T, et al. Dobrava-Belgrade virus spillover infections, Germany. Emerging Infectious Diseases. 2009;15: 2017–2020. doi:10.3201/eid1512.090923

52. Nemirov K, Leirs H, Lundkvist Å, Olsson GE. Puumala hantavirus and Myodes glareolus in northern Europe: No evidence of co-divergence between genetic lineages of virus and host. Journal of General Virology. 2010;91: 1262–1274. doi:10.1099/vir.0.016618-0

53. Reusken C, Heyman P. Factors driving hantavirus emergence in Europe. Current Opinion in Virology. Elsevier B.V.; 2013;3: 92–99. doi:10.1016/j.coviro.2013.01.002

54. Martinez VP, Bellomo C, San Juan J, Pinna D, Forlenza R, Elder M, et al. Person-to-person transmission of Andes virus. Emerging infectious diseases. 2005;11: 1848–53. doi:10.3201/eid1112.050501

55. Plyusnin A, Morzunov SP. Virus Evolution and Genetic Diversity of Hantaviruses and Their Rodent Hosts. Springer, Berlin, Heidelberg; 2001. pp. 47–75. doi:10.1007/978-3-642-56753-7_4

56. Kallio ER, Klingström J, Gustafsson E, Manni T, Vaheri A, Henttonen H, et al. Prolonged survival of Puumala hantavirus outside the host: evidence for indirect transmission via the environment. Journal of General Virology. Microbiology Society; 2006;87: 2127–2134. doi:10.1099/vir.0.81643-0

57. Richardson KS, Kuenzi A, Douglass RJ, Hart J, Carver S. Human exposure to particulate matter potentially contaminated with sin nombre virus. EcoHealth. NIH Public Access; 2013;10: 159–65. doi:10.1007/s10393-013-0830-x

58. de St Maurice A, Ervin E, Schumacher M, Yaglom H, VinHatton E, Melman S, et al. Exposure Characteristics of Hantavirus Pulmonary Syndrome Patients, United States, 1993-2015. Emerging infectious diseases. Centers for Disease Control and Prevention; 2017;23: 733–739. doi:10.3201/eid2305.161770

59. Groen J, Gerding MN, Jordans JGM, Clement JP, Nieuwenhuijs JHM, Osterhaus ADME. Hantavirus infections in The Netherlands: epidemiology and disease. Epidemiology and Infection. Cambridge University Press; 1995;114: 373–383. doi:10.1017/S0950268800058003

60. Glass GE, Childs JE, Korch GW, LeDuc JW. Association of intraspecific wounding with hantaviral infection in wild rats (

Rattus norvegicus). Epidemiology and Infection. Cambridge University Press; 1988;101: 459–472. doi:10.1017/S0950268800054418

61. Padula P, Figueroa R, Navarrete M, Pizarro E, Cadiz R, Bellomo C, et al. Transmission Study of Andes Hantavirus Infection in Wild Sigmodontine Rodents. Journal of Virology. American Society for Microbiology Journals; 2004;78: 11972–11979. doi:10.1128/JVI.78.21.11972-11979.2004

62. Schultze D, Lundkvist Å, Blauenstein U, Heyman P. Tula Virus Infection Associated with Fever and Exanthema After a Wild Rodent Bite. European Journal of Clinical Microbiology and Infectious Diseases. 2002;21: 304–306. doi:10.1007/s10096-002-0705-5

63. Jonsson CB, Figueiredo LTM, Vapalahti O. A global perspective on hantavirus ecology, epidemiology, and disease. Clinical microbiology reviews. 2010;23: 412–41. doi:10.1128/CMR.00062-09

64. Avsic-Zupanc T, Petrovec M, Furlan P, Kaps R, Elgh F, Lundkvist A. Hemorrhagic fever with renal syndrome in the Dolenjska region of Slovenia--a 10-year survey. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 1999;28: 860–5. doi:10.1086/515185

65. Wichmann D, Grone H-J, Frese M, Pavlovic J, Anheier B, Haller O, et al. Hantaan Virus Infection Causes an Acute Neurological Disease That Is Fatal in Adult Laboratory Mice. Journal of Virology. 2002;76: 8890–8899. doi:10.1128/JVI.76.17.8890-8899.2002

66. Hjertqvist M, Klein SL, Ahlm C, Klingstrom J. Mortality rate patterns for hemorrhagic fever with renal syndrome caused by Puumala virus. Emerging infectious diseases. 2010;16: 1584–6. doi:10.3201/eid1610.100242

67. Zhang Y-Z, Zou Y, Fu ZF, Plyusnin A. Hantavirus infections in humans and animals, China. Emerging infectious diseases. 2010;16: 1195–203. doi:10.3201/eid1608.090470

68. Watson DC, Sargianou M, Papa A, Chra P, Starakis I, Panos G. Epidemiology of Hantavirus infections in humans: A comprehensive, global overview. Critical Reviews in Microbiology. Taylor & Francis; 2013; Available: http://www.tandfonline.com/doi/abs/10.3109/1040841X.2013.783555

69. Papa A, Bojovic B, Antoniadis A. Hantaviruses in Serbia and

Montenegro. Emerging infectious diseases. 2006;12: 1015–8. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3373030&tool=pmcentrez&rendertype=abstract

70. CDC. Annual U.S. HPS Cases and Case-fatality - Hantavirus [Internet]. 2016 [cited 4 May 2016]. Available: http://www.cdc.gov/hantavirus/surveillance/annual-cases.html

71. Duchin JS, Koster FT, Peters CJ, Simpson GL, Tempest B, Zaki SR, et al. Hantavirus pulmonary syndrome: a clinical description of 17 patients with a newly recognized disease. The Hantavirus Study Group. The New England journal of medicine. 1994;330: 949–55. doi:10.1056/NEJM199404073301401

72. Figueiredo LTM, Souza WM de, Ferrés M, Enria DA. Hantaviruses and cardiopulmonary syndrome in South America. Virus research. 2014;187: 43–54. doi:10.1016/j.virusres.2014.01.015

73. Rasmuson J, Andersson C, Norrman E, Haney M, Evander M, Ahlm C. Time to revise the paradigm of hantavirus syndromes? Hantavirus pulmonary syndrome caused by European hantavirus. European Journal of Clinical Microbiology & Infectious Diseases. 2011;30: 685–690. doi:10.1007/s10096-010-1141-6

74. Antoniades A, Grekas D, Rossi CA, LeDuc JW. Isolation of a Hantavirus from a Severely Ill Patient with Hemorrhagic Fever with Renal Syndrome in Greece. Journal of Infectious Diseases. 1987;156: 1010–1013. doi:10.1093/infdis/156.6.1010

75. Clement J, Colson P, McKenna P. Hantavirus pulmonary syndrome in New England and Europe. The New England journal of medicine. 1994;331: 545–6; author reply 547-8. Available: http://www.ncbi.nlm.nih.gov/pubmed/8041425

76. Vollmar P, Lubnow M, Simon M, Müller T, Bergler T, Alois P, et al. Hantavirus cardiopulmonary syndrome due to Puumala virus in Germany. Journal of Clinical Virology. 2016;84: 42–47. doi:10.1016/j.jcv.2016.10.004

77. Clement J, Maes P, Lagrou K, Van Ranst M, Lameire N. A unifying hypothesis and a single name for a complex globally emerging infection: Hantavirus disease. In: European Journal of Clinical Microbiology and Infectious Diseases [Internet]. 2012 [cited 23 Mar 2016] pp. 1–5. doi:10.1007/s10096-011-1456-y

78. Lipkin WI, Briese T, Hornig M. Borna disease virus – Fact and fantasy. Virus Research. Elsevier B.V.; 2011;162: 162–172. doi:10.1016/j.virusres.2011.09.036

79. Kinnunen PM, Palva A, Vaheri A, Vapalahti O. Epidemiology and

host spectrum of Borna disease virus infections. The Journal of general virology. Microbiology Society; 2013;94: 247–62. doi:10.1099/vir.0.046961-0

80. Sassa Y, Horie M, Fujino K, Nishiura N, Okazaki S, Furuya T, et al. Molecular epidemiology of avian bornavirus from pet birds in Japan. Virus genes. 2013;47: 173–7. doi:10.1007/s11262-013-0913-3

81. Weissenböck H, Bakonyi T, Sekulin K, Ehrensperger F, Doneley RJT, Dürrwald R, et al. Avian bornaviruses in psittacine birds from Europe and Australia with proventricular dilatation disease. Emerging Infectious Diseases. 2009;15: 1453–1459. doi:10.3201/eid1509.090353

82. Carbone KM, Duchala CS, Griffin JW, Kincaid AL, Narayan O. Pathogenesis of Borna disease in rats: evidence that intra-axonal spread is the major route for virus dissemination and the determinant for disease incubation. J Virol. 1987;61: 3431–3440. Available: http://jvi.asm.org/content/61/11/3431.short

83. Nobach D, Bourg M, Herzog S, Lange-Herbst H, Encarnação JA, Eickmann M, et al. Shedding of Infectious Borna Disease Virus-1 in Living Bicolored White-Toothed Shrews. PloS one. Public Library of Science; 2015;10: e0137018. doi:10.1371/journal.pone.0137018

84. Zhang L, Wang X, Zhan Q, Wang Z, Xu M, Zhu D, et al. Evidence for natural Borna disease virus infection in healthy domestic animals in three areas of western China. Archives of virology. 2014;159: 1941–9. doi:10.1007/s00705-013-1971-5

85. Wensman JJ, Jäderlund KH, Holst BS, Berg M. Borna disease virus infection in cats. Veterinary journal (London, England : 1997). 2014;201: 142–9. doi:10.1016/j.tvjl.2013.12.012

86. Hilbe M, Herrsche R, Kolodziejek J, Nowotny N, Zlinszky K, Ehrensperger F. Shrews as reservoir hosts of borna disease virus. Emerging infectious diseases. 2006;12: 675–7. doi:10.3201/eid1204.051418

87. Bourg M, Herzog S, Encarnação JA, Nobach D, Lange-Herbst H, Eickmann M, et al. Bicolored whitetoothed shrews as reservoir for Borna disease virus, Bavaria, Germany. Emerging Infectious Diseases. 2013;19: 2064–2066. doi:10.3201/eid1912.131076

88. Björnsdóttir S, Agustsdóttir E, Blomström A-L, Oström I-LÖ, Berndtsson LT, Svansson V, et al. Serological markers of Bornavirus infection found in horses in Iceland. Acta veterinaria Scandinavica. BioMed Central; 2013;55: 77. doi:10.1186/1751-0147-55-77

89. Someya A, Fukushima R, Yoshida M, Tanahashi Y, Prapeuk T, Iizuka R, et al. A study on Borna disease virus infection in domestic cats in Japan. The Journal of veterinary medical science / the Japanese Society of Veterinary Science. 2014;76: 1157–60. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4155199&tool=pmcentrez&rendertype=abstract

90. Wensman JJ, Jäderlund KH, Gustavsson MH, Hansson-Hamlin H, Karlstam E, Lilliehöök I, et al. Markers of Borna disease virus infection in cats with staggering disease. Journal of feline medicine and surgery. 2012;14: 573–82. doi:10.1177/1098612X12446638

91. Berg A-L, Reid-Smith R, Larsson M, Bonnett B. Case control study of feline Borna disease in Sweden. Veterinary Record. BMJ Publishing Group Limited; 1998;142: 715–717. doi:10.1136/vr.142.26.715

92. Dürrwald R, Kolodziejek J, Muluneh A, Herzog S, Nowotny N. Epidemiological pattern of classical Borna disease and regional genetic clustering of Borna disease viruses point towards the existence of to-date unknown endemic reservoir host populations. Microbes and infection / Institut Pasteur. 2006;8: 917–29. doi:10.1016/j.micinf.2005.08.013

93. Kupke A, Becker S, Wewetzer K, Ahlemeyer B, Eickmann M, Herden C. Intranasal Borna Disease Virus (BoDV-1) Infection: Insights into Initial Steps and Potential Contagiosity. International journal of molecular sciences. Multidisciplinary Digital Publishing Institute (MDPI); 2019;20. doi:10.3390/ijms20061318

94. Sauder C, Staeheli P. Rat Model of Borna Disease Virus Transmission: Epidemiological Implications. Journal of Virology. 2003;77: 12886–12890. doi:10.1128/JVI.77.23.12886-12890.2003

95. Okamoto M, Hagiwara K, Kamitani W, Sako T, Hirayama K, Kirisawa R, et al. Experimental vertical transmission of Borna disease virus in the mouse. Archives of virology. 2003;148: 1557–68. doi:10.1007/s00705-003-0120-y

96. Kinnunen PM, Inkeroinen H, Ilander M, Kallio ER, Heikkilä HP, Koskela E, et al. Intracerebral Borna disease virus infection of bank voles leading to peripheral spread and reverse transcription of viral RNA. PloS one. Public Library of Science; 2011;6: e23622. doi:10.1371/journal.pone.0023622

97. Hoffmann B, Tappe D, Höper D, Herden C, Boldt A, Mawrin C, et al. A Variegated Squirrel Bornavirus Associated with Fatal Human Encephalitis. New England Journal of Medicine. Massachusetts Medical Society; 2015;373: 154–162.

doi:10.1056/NEJMoa1415627

98. Schlottau K, Hoffmann B, Homeier-Bachmann T, Fast C, Ulrich RG, Beer M, et al. Multiple detection of zoonotic variegated squirrel bornavirus 1 RNA in different squirrel species suggests a possible unknown origin for the virus. Archives of Virology. Springer Vienna; 2017;162: 2747–2754. doi:10.1007/s00705-017-3432-z

99. Schlottau K, Jenckel M, Van den Brand J, Fast C, Herden C, Höper D, et al. Variegated squirrel bornavirus 1 in squirrels, Germany and the Netherlands. Emerging Infectious Diseases. Centers for Disease Control and Prevention; 2017;23: 477–481. doi:10.3201/eid2303.161061

100. Petzold J, van den Brand JMA, Nobach D, Hoffmann B, Hoffmann D, Fast C, et al. Distribution of zoonotic variegated squirrel bornavirus 1 in naturally infected variegated and Prevost's squirrels. Scientific reports. Nature Publishing Group; 2019;9: 11402. doi:10.1038/s41598-019-47767-4

101. Chalmers RM, Thomas DR, Salmon RL. Borna disease virus and the evidence for human pathogenicity: a systematic review. QJM : monthly journal of the Association of Physicians. 2005;98: 255–74. doi:10.1093/qjmed/hci039

102. Barrantes-Rodríguez X, Silva S, Macaya G, Bonilla JA. Detection of Borna Virus Disease by real-time RT-PCR in Costa Rican equines and humans. Archives of Clinical Microbiology. iMedPub; 2012;3. Available: http://www.acmicrob.com/abstract/detection-of-borna-virus-disease-by-realtime-rtpcr-in-costa-rican-equines-and-humans-0.html

103. Mazaheri-Tehrani E, Maghsoudi N, Shams J, Soori H, Atashi H, Motamedi F, et al. Borna disease virus (BDV) infection in psychiatric patients and healthy controls in Iran. Virology journal. BioMed Central; 2014;11: 161. doi:10.1186/1743-422X-11-161

104. Honda T, Sofuku K, Matsunaga H, Tachibana M, Mohri I, Taniike M, et al. Prevalence of antibodies against Borna disease virus proteins in Japanese children with autism spectrum disorder. Microbiology and Immunology. John Wiley & Sons, Ltd (10.1111); 2018;62: 473–476. doi:10.1111/1348-0421.12603

105. Zhang L, Xu M-MM-M, Zeng L, Liu S, Liu X, Wang X, et al. Evidence for Borna disease virus infection in neuropsychiatric patients in three western China provinces. In: European Journal of Clinical Microbiology & Infectious Diseases [Internet]. Apr 2013 [cited 21 Oct 2015] pp. 621–627. doi:10.1007/s10096-013-1996-4

106. Wolff T, Heins G, Pauli G, Burger R, Kurth R. Failure to detect Borna disease virus antigen and RNA in human blood. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology. Elsevier; 2006;36: 309–11. doi:10.1016/j.jcv.2006.05.005

107. Wolff T, Burger R, Kurth R. Re: Absence of Borna virus in human blood. Journal of Clinical Virology. Elsevier; 2006;36: 314. doi:10.1016/j.jcv.2006.06.003

108. Dürrwald R, Kolodziejek J, Nowotny N. Borna disease virus (BDV) sequences derived from plasma samples of Australian cats contain multiple sequencing errors and are otherwise almost identical to strain V, a commonly used BDV laboratory strain. Microbes and Infection. 2006;8: 1421–1422. doi:10.1016/j.micinf.2006.02.005

109. Hornig M, Briese T, Licinio J, Khabbaz RF, Altshuler LL, Potkin SG, et al. Absence of evidence for bornavirus infection in schizophrenia, bipolar disorder and major depressive disorder. Molecular psychiatry. Macmillan Publishers Limited; 2012;17: 486–93. doi:10.1038/mp.2011.179

110. Horie M, Kobayashi Y, Suzuki Y, Tomonaga K. Comprehensive analysis of endogenous bornavirus-like elements in eukaryote genomes. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2013;368: 20120499. doi:10.1098/rstb.2012.0499

111. Feschotte C. Virology: Bornavirus enters the genome. Nature. 2010;463: 39–40. doi:10.1038/463039a

112. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, et al. L1 retrotransposition in human neural progenitor cells. Nature. 2009;460: 1127–31. doi:10.1038/nature08248

113. Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, Oshida T, et al. Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature. Macmillan Publishers Limited. All rights reserved; 2010;463: 84–7. doi:10.1038/nature08695

114. Korn K, Coras R, Bobinger T, Herzog SM, Lücking H, Stöhr R, et al. Fatal Encephalitis Associated with Borna Disease Virus 1. New England Journal of Medicine. Massachusetts Medical Society; 2018;379: 1375–1377. doi:10.1056/NEJMc1800724

115. Schlottau K, Forth L, Angstwurm K, Höper D, Zecher D, Liesche F, et al. Fatal Encephalitic Borna Disease Virus 1 in Solid-Organ Transplant Recipients. New England Journal of Medicine. 2018;379: 1377–1379. doi:10.1056/NEJMc1803115

116. Coras R, Korn K, Kuerten S, Huttner HB, Ensser A. Severe bornavirus-encephalitis presenting as Guillain–Barré-syndrome. Acta Neuropathologica. Springer Berlin Heidelberg; 2019;137: 1017–1019. doi:10.1007/s00401-019-02005-z

117. Tappe D, Frank C, Homeier-Bachmann T, Wilking H, Allendorf V, Schlottau K, et al. Analysis of exotic squirrel trade and detection of human infections with variegated squirrel bornavirus 1, Germany, 2005 to 2018. Eurosurveillance. European Centre for Disease Prevention and Control; 2019;24: 1800483. doi:10.2807/1560-7917.ES.2019.24.8.1800483

118. Tappe D, Schlottau K, Cadar D, Hoffmann B, Balke L, Bewig B, et al. Occupation-Associated Fatal Limbic Encephalitis Caused by Variegated Squirrel Bornavirus 1, Germany, 2013. Emerging Infectious Diseases. 2018;24: 978–987. doi:10.3201/eid2406.172027

119. Tapparel C, Siegrist F, Petty TJ, Kaiser L. picornavirus and enterovirus diversity with associated human diseases. Infection, Genetics and Evolution. Elsevier; 2013;14: 282–293. doi:10.1016/J.MEEGID.2012.10.016

120. Landry ML, Lu X. Rhinoviruses. Manual of Clinical Microbiology, 11th Edition. American Society of Microbiology; 2015. pp. 1551–1564. doi:10.1128/9781555817381.ch89

121. Lum LCS, Wong KT, Lam SK, Chua KB, Goh AYT, Lim WL, et al. Fatal enterovirus 71 encephalomyelitis. The Journal of Pediatrics. Mosby; 1998;133: 795–798. doi:10.1016/S0022-3476(98)70155-6

122. Arzt J, Juleff N, Zhang Z, Rodriguez LL. The Pathogenesis of Foot-and-Mouth Disease I: Viral Pathways in Cattle. Transboundary and Emerging Diseases. Blackwell Publishing Ltd; 2011;58: 291–304. doi:10.1111/j.1865-1682.2011.01204.x

123. Haydon DT, Kao RR, Kitching RP. The UK foot-and-mouth disease outbreak — the aftermath. Nature Reviews Microbiology. Nature Publishing Group; 2004;2: 675–681. doi:10.1038/nrmicro960

124. Knight-Jones TJD, Rushton J. The economic impacts of foot and mouth disease - what are they, how big are they and where do they occur? Preventive veterinary medicine. Elsevier; 2013;112: 161–73. doi:10.1016/j.prevetmed.2013.07.013

125. Harvala H, Simmonds P. Human parechoviruses: Biology, epidemiology and clinical significance. Journal of Clinical Virology. 2009;45: 1–9. doi:10.1016/j.jcv.2009.03.009

126. Chieochansin T, Vichiwattana P, Korkong S, Theamboonlers A,

Poovorawan Y. Molecular epidemiology, genome characterization, and recombination event of human parechovirus. Virology. 2011;421: 159–166. doi:10.1016/j.virol.2011.09.021

127. Olijve L, Jennings L, Walls T. Human Parechovirus : an Increasingly Recognized Cause of. Clinical Microbiology Reviews. 2018; 1–17. Available: https://doi.org/10.1128/CMR.00047-17.

128. Van Der Sanden S, De Bruin E, Vennema H, Swanink C, Koopmans M, Van Der Avoort H. Prevalence of human parechovirus in The Netherlands in 2000 to 2007. Journal of Clinical Microbiology. American Society for Microbiology Journals; 2008;46: 2884–2889. doi:10.1128/JCM.00168-08

129. Shee A, Weber H. Human Parechovirus Infection in Neonates and Children: An Overview. Journal of Pediatric Infectious Diseases. 2017;12: 099–103. doi:10.1055/s-0037-1602384

130. Shan TL, Wang CM, Cui L, Delwart E, Yuan CL, Zhao W, et al. Human parechovirus infections in monkeys with diarrhea, China. Emerging infectious diseases. Centers for Disease Control and Prevention; 2010;16: 1168–9. doi:10.3201/eid1607.091103

131. Digoutte JP. Sebokele (SEB). The American Journal of Tropical Medicine and Hygiene. The American Society of Tropical Medicine and Hygiene; 1978;27: 424–425. doi:10.4269/ajtmh.1978.27.424

132. Joffret ML, Bouchier C, Grandadam M, Zeller H, Maufrais C, Bourhy H, et al. Genomic characterization of Sebokele virus 1 (SEBV1) reveals a new candidate species among the genus Parechovirus. Journal of General Virology. Microbiology Society; 2013;94: 1547–1553. doi:10.1099/vir.0.053157-0

133. Smits SL, Raj VS, Oduber MD, Schapendonk CME, Bodewes R, Provacia L, et al. Metagenomic Analysis of the Ferret Fecal Viral Flora. Kapoor A, editor. PLoS ONE. Public Library of Science; 2013;8: e71595. doi:10.1371/journal.pone.0071595

134. Zhang C, Sheng J, Zhao Z, Yan C, Tu C, He B. Genomic Characterization of the First Parechovirus in Bats. Virologica Sinica. Springer Singapore; 2019;34: 471–473. doi:10.1007/s12250-019-00108-1

135. Niklasson B, Kinnunen L, Hörnfeldt B, Hörling J, Benemar C, Hedlund KO, et al. A new picornavirus isolated from bank voles *(Clethrionomys glareolus)*. Virology. 1999;255: 86–93. doi:S0042682298995575 [pii]

136. Hauffe HC, Niklasson B, Olsson T, Bianchi A, Rizzoli A, Klitz W. Ljungan Virus Detected in Bank Voles (Myodes glareolus) and Yellow-Necked Mice (Apodemus flavicollis) from Northern Italy.

Journal of Wildlife Diseases. 2010;46: 262–266. doi:10.7589/0090-3558-46.1.262

137. Salisbury AM, Begon M, Dove W, Niklasson B, Stewart JP. Ljungan virus is endemic in rodents in the UK. Archives of Virology. 2014;159: 547–551. doi:10.1007/s00705-013-1731-6

138. Romeo C, Ferrari N, Rossi C, Everest DJ, Grierson SS, Lanfranchi P, et al. Ljungan Virus and an Adenovirus in Italian Squirrel Populations. Journal of Wildlife Diseases. 2014;50: 409–411. doi:10.7589/2013-10-260

139. Mitake H, Fujii Y, Nagai M, Ito N, Okadera K, Okada K, et al. Isolation of a sp. nov. Ljungan virus from wild birds in Japan. Journal of General Virology. Microbiology Society; 2016;97: 1818–1822. doi:10.1099/jgv.0.000508

140. Johansson ES, Niklasson B, Tesh RB, Shafren DR, Rosa APAT da, Lindberg AM. Molecular characterization of M1146, an American isolate of Ljungan virus (LV) reveals the presence of a new LV genotype. Journal of General Virology. Microbiology Society; 2003;84: 837–844. doi:10.1099/vir.0.18792-0

141. Niklasson B, Heller KE, Schønecker B, Bildsøe M, Daniels T, Hampe CS, et al. Development of type 1 diabetes in wild bank voles associated with islet autoantibodies the novel Ljungan virus. Experimental Diabesity Research. Hindawi Publishing Corporation; 2003;4: 35–44. doi:10.1080/15438600303733

142. Niklasson B, Hörnfeldt B, Nyholm E, Niedrig M, Donoso-Mantke O, Gelderblom HR, et al. Type 1 diabetes in Swedish bank voles (Clethrionomys glareolus): Signs of disease in both colonized and wild cyclic populations at peak density. Annals of the New York Academy of Sciences. John Wiley & Sons, Ltd (10.1111); 2003;1005: 170–175. doi:10.1196/annals.1288.020

143. Samsioe A, Feinstein R, Saade G, Sjöholm Å, Hörnfeldt B, Fundele R, et al. Intrauterine death, fetal malformation, and delayed pregnancy in Ljungan virus-infected mice. Birth Defects Research Part B - Developmental and Reproductive Toxicology. Wiley Subscription Services, Inc., A Wiley Company; 2006;77: 251–256. doi:10.1002/bdrb.20083

144. Samsioe A, Sjöholm A, Niklasson B, Klitz W. Fetal death persists through recurrent pregnancies in mice following Ljungan virus infection. Birth defects research Part B, Developmental and reproductive toxicology. NIH Public Access; 2008;83: 507–10. doi:10.1002/bdrb.20169

145. Wesslén L, Påhlson C, Friman G, Fohlman J, Lindquist O,

Johansson C. Myocarditis caused by Chlamydia pneumoniae (TWAR) and sudden unexpected death in a Swedish elite orienteer [Internet]. The Lancet. 1992. pp. 427–428. doi:10.1016/0140-6736(92)91509-7

146. Niklasson B, Hörnfeldt B, Lundman B. Could myocarditis, insulin-dependent diabetes mellitus, and Guillain-Barré syndrome be caused by one or more infectious agents carried by rodents? Emerging infectious diseases. Centers for Disease Control and Prevention; 1998;4: 187–93. doi:10.3201/eid0402.980206

147. Niklasson B, Almqvist PR, Hörnfeldt B, Klitz W. Sudden infant death syndrome and Ljungan virus. Forensic Science, Medicine, and Pathology. Humana Press Inc; 2009;5: 274–279. doi:10.1007/s12024-009-9086-8

148. Niklasson B, Samsioe A, Papadogiannakis N, Kawecki A, Hörnfeldt B, Saade GR, et al. Association of zoonotic Ljungan virus with intrauterine fetal deaths. Birth Defects Research Part A: Clinical and Molecular Teratology. Wiley Subscription Services, Inc., A Wiley Company; 2007;79: 488–493. doi:10.1002/bdra.20359

149. Samsioe A, Papadogiannakis N, Hultman T, Sjöholm Å, Klitz W, Niklasson B. Ljungan virus present in intrauterine fetal death diagnosed by both immunohistochemistry and PCR. Birth Defects Research Part A: Clinical and Molecular Teratology. Wiley Subscription Services, Inc., A Wiley Company; 2009;85: 227–229. doi:10.1002/bdra.20554

150. Niklasson B, Samsioe A, Papadogiannakis N, Gustafsson S, Klitz W. Zoonotic ljungan virus associated with central nervous system malformations in terminated pregnancy. Birth Defects Research Part A - Clinical and Molecular Teratology. Wiley Subscription Services, Inc., A Wiley Company; 2009;85: 542–545. doi:10.1002/bdra.20568

151. Krous HF, Langlois NE. Ljungan virus: A commentary on its association with fetal and infant morbidity and mortality in animals and humans. Birth Defects Research Part A: Clinical and Molecular Teratology. John Wiley & Sons, Ltd; 2010;88: 947–952. doi:10.1002/bdra.20728

152. Jääskeläinen AJ, Kolehmainen P, Voutilainen L, Hauffe HC, Kallio-Kokko H, Lappalainen M, et al. Evidence of ljungan virus specific antibodies in humans and rodents, Finland. Journal of Medical Virology. 2013;85: 2001–2008. doi:10.1002/jmv.23681

153. Jääskeläinen AJ, Voutilainen L, Lehmusto R, Henttoen H, Lappalainen M, Kallio-Kokko H, et al. Serological survey in the Finnish human population implies human-to-human transmission

of Ljungan virus or antigenically related viruses. Epidemiology and Infection. 2016;144: 1278–1285. doi:10.1017/S0950268815002551

154. Heyman P, Simons L, Cochez C. Were the English sweating sickness and the Picardy sweat caused by hantaviruses? Viruses. 2014;6: 151–71. doi:10.3390/v6010151

155. McKenna P, Clement J, Matthys P, Coyle P V., McCaughey C. Serological evidence of hantavirus disease in Northern Ireland. Journal of Medical Virology. 1994;43: 33–38. doi:10.1002/jmv.1890430107

156. Jameson LJ, Newton A, Coole L, Newman ENC, Carroll MW, Beeching NJ, et al. Prevalence of antibodies against hantaviruses in serum and saliva of adults living or working on farms in Yorkshire, United Kingdom. Viruses. Multidisciplinary Digital Publishing Institute; 2014;6: 524–34. doi:10.3390/v6020524

157. Jameson LJ. Seoul hantavirus as a cause of acute kidney injury in the United Kingdom [Internet]. University of Liverpool. 2015. Available: http://repository.liv.ac.uk/2009807/1/JamesonLis_Apr2015_2009807.pdf

158. Kudesia G, Christie P, Walker E, Pinkerton I, Lloyd G. Dual infection with leptospira and hantavirus. Lancet (London, England). 1988;1: 1397. Available: http://www.ncbi.nlm.nih.gov/pubmed/2898071

159. Pether JVS, Jones N, Lloyd G. Acute hantavirus infection. The Lancet. 1991;338: 1025. doi:10.1016/0140-6736(91)91893-Y

160. Rice P, Kudesia G, Leach M. Acute hantavirus infection. Journal of Infection. Elsevier; 1993;27: 342–344. doi:10.1016/0163-4453(93)92532-2

161. Watson AR, Irving WL, Ansell ID. Playing in a scrapyard and acute renal failure. Lancet. 1997;349: 1446. doi:10.1016/S0140-6736(97)01408-6

162. Jameson LJ, Logue CH, Atkinson B, Baker N, Galbraith SE, Carroll MW, et al. The continued emergence of hantaviruses: isolation of a Seoul virus implicated in human disease, United Kingdom, October 2012. Eurosurveillance. European Centre for Disease Prevention and Control; 2013;18: 20344. doi:10.2807/ese.18.01.20344-en

163. McElhinney LM, Marston DA, Pounder KC, Goharriz H, Wise EL, Verner-Carlsson J, et al. High prevalence of Seoul hantavirus in a breeding colony of pet rats. Epidemiology and Infection. 2017;145: 3115–3124. doi:10.1017/S0950268817001819

164. Murphy EG, Williams NJ, Bennett M, Jennings D, Chantrey J, McElhinney LM. Detection of Seoul virus in wild brown rats (Rattus norvegicus) from pig farms in Northern England. Veterinary Record. British Medical Journal Publishing Group; 2019;184: 525–525. doi:10.1136/VR.105249

165. Pettersson L, Boman J, Juto P, Evander M, Ahlm C. Outbreak of Puumala virus infection, Sweden. Emerging infectious diseases. Centers for Disease Control and Prevention; 2008;14: 808–10. doi:10.3201/eid1405.071124

166. Plyusnina A, Razzauti M, Sironen T, Niemimaa J, Vapalahti O, Vaheri A, et al. Analysis of Complete Puumala Virus Genome, Finland. Emerging Infectious Diseases. 2012;18. doi:10.3201/eid1812.120747

167. Davidyuk Y, Kabwe E, Shakirova V, Martynova E V., Ismagilova R, Khaertynova I, et al. Characterization of the Puumala orthohantavirus strains in the northwestern region of the Republic of Tatarstan in relation to the clinical manifestations in hemorrhagic fever with renal syndrome patients. Frontiers in Pharmacology. Frontiers; 2019;10: 970. doi:10.3389/FPHAR.2019.00970

168. Olsson GE, White N, Ahlm C, Elgh F, Verlemyr A-C, Juto P, et al. Demographic Factors Associated with Hantavirus Infection in Bank Voles (*Clethrionomys glareolus*). Emerging Infectious Diseases. 2002;8: 924–929. doi:10.3201/eid0809.020037

169. Hanski I, Hansson L, Henttonen H. Specialist Predators, Generalist Predators, and the Microtine Rodent Cycle. The Journal of Animal Ecology. British Ecological Society; 1991;60: 353. doi:10.2307/5465

170. Piechotowski I, Brockmann SO, Schwarz C, Winter CH, Ranft U, Pfaff G. Emergence of hantavirus in South Germany: rodents, climate and human infections. Parasitology Research. 2008;103: 131–137. doi:10.1007/s00436-008-1055-8

171. Zwolak R, Bogdziewicz M, Rychlik L. Beech masting modifies the response of rodents to forest management. Forest Ecology and Management. Elsevier; 2016;359: 268–276. doi:10.1016/J.FORECO.2015.10.017

172. Reil D, Imholt C, Eccard JA, Jacob J. Beech Fructification and Bank Vole Population Dynamics - Combined Analyses of Promoters of Human Puumala Virus Infections in Germany. Waldenström J, editor. PLOS ONE. Public Library of Science; 2015;10: e0134124. doi:10.1371/journal.pone.0134124

173. Clement J, Maes P, van Ypersele de Strihou C, van der Groen G, Barrios JM, Verstraeten WW, et al. Beechnuts and outbreaks of nephropathia epidemica (NE): of mast, mice and men. Nephrology Dialysis Transplantation. Narnia; 2010;25: 1740–1746. doi:10.1093/ndt/gfq122

174. Bennett E, Clement J, Sansom P, Hall I, Leach S, Medlock JM. Environmental and ecological potential for enzootic cycles of Puumala hantavirus in Great Britain. Epidemiology and infection. 2010;138: 91–8. doi:10.1017/S095026880999029X

175. Markus Brummer-Korvenkontio, Olli V M, Vapalahti O, Henttonen H, Koskela P, Kuusisto P, Vaheri A. Epidemiological Study of Nephropathia epidemica in Finland 1989-96. Scandinavian Journal of Infectious Diseases. 1999;31: 427–435. doi:10.1080/00365549950163941

176. Goeijenbier M, Hartskeerl RA, Reimerink J, Verner-Carlsson J, Wagenaar JF, Goris MG, et al. The hanta hunting study: underdiagnosis of Puumala hantavirus infections in symptomatic non-travelling leptospirosis-suspected patients in the Netherlands, in 2010 and April to November 2011. Eurosurveillance. 2014;19: 27–36. Available: http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=CitingArticles&qid=1&SID=S1vCOCa27fmqH4Q7wIr&page=1&doc=5

177. Clement J, Van Esbroeck M, Lagrou K, Verschueren J, Sunil-Chandra N, Van Ranst M. Letter to the Editor: Leptospirosis versus hantavirus infections in the Netherlands and in Belgium, 2000 to 2014. Eurosurveillance. 2014;19: 20912. doi:10.2807/1560-7917.ES2014.19.38.20912

178. Pounder KC, Begon M, Sironen T, Henttonen H, Watts PC, Voutilainen L, et al. Novel Hantavirus in field vole, United Kingdom. Emerging Infectious Diseases. 2013;19: 673–675. doi:10.3201/eid1904.121057

179. Thomason AG, Begon M, Bradley JE, Paterson S, Jackson JA. Endemic Hantavirus in Field Voles, Northern England. Emerging infectious diseases. Centers for Disease Control and Prevention; 2017;23: 1033–1035. doi:10.3201/eid2306.161607

180. Reeves NA, Helps CR, Gunn-Moore DA, Blundell C, Finnemore PL, Pearson GR, et al. Natural Borna disease virus infection in cats in the United Kingdom. The Veterinary Record. 1998;143: 523–526. doi:10.1136/vr.143.19.523

181. Thomas DR, Chalmers RM, Crook B, Stagg S, Thomas H V, Lewis G, et al. Borna disease virus and mental health: a cross-sectional

study. QJM : monthly journal of the Association of Physicians. 2005;98: 247–54. doi:10.1093/qjmed/hci040

182. Ferreras Antolín L, Kadambari S, Braccio S, Tang JWT, Xerry J, Allen DJ, et al. Increased detection of human parechovirus infection in infants in England during 2016: Epidemiology and clinical characteristics. Archives of Disease in Childhood. BMJ Publishing Group Ltd; 2018;103: 1061–1066. doi:10.1136/archdischild-2017-314281

183. Suckow M, Weisbroth S, Franklin C. The Laboratory Rat [Internet]. The Laboratory Rat. 2006. doi:10.1016/B978-0-12-074903-4.X5000-6

184. Traweger D, Travnitzky R, Moser C, Walzer C, Bernatzky G. Habitat preferences and distribution of the brown rat (Rattus norvegicus Berk.) in the city of Salzburg (Austria): implications for an urban rat management. Journal of Pest Science. Springer Science and Business Media LLC; 2006;79: 113–125. doi:10.1007/s10340-006-0123-z

185. Mathews F, Kubasiewicz LM, Gurnell J, Harrower CA, McDonald RA, Shore RF. A Review of the Population and Conservation Status of British Mammals: Technical Summary [Internet]. 2018. Available: http://publications.naturalengland.org.uk/publication/5636785878597632?category=47019

186. Murphy EG, Williams NJ, Jennings D, Chantrey J, Verin R, Grierson S, et al. First detection of Hepatitis E virus (Orthohepevirus C) in wild brown rats (Rattus norvegicus) from Great Britain. Zoonoses and Public Health. Wiley-VCH Verlag; 2019;66: 686–694. doi:10.1111/zph.12581

187. Attuquayefio DK, Gorman ML, Wolton RJ. Home range sizes in the Wood mouse Apodemus sylvaticus: habitat, sex and seasonal differences. Journal of Zoology. John Wiley & Sons, Ltd; 1986;210: 45–53. doi:10.1111/j.1469-7998.1986.tb03619.x

188. Marsh ACW, Poulton S, Harris S. The Yellow-necked Mouse Apodemus flavicollis in Britain: status and analysis of factors affecting distribution. Mammal Review. Wiley; 2008;31: 203–227. doi:10.1111/j.1365-2907.2001.00089.x

189. N'Dilimabaka N, Berthet N, Rougeron V, Mangombi JB, Durand P, Maganga GD, et al. Evidence of Lymphocytic Choriomeningitis Virus (LCMV) in Domestic Mice in Gabon: Risk of Emergence of LCMV Encephalitis in Central Africa. Journal of Virology. American Society for Microbiology; 2015;89: 1456–1460. doi:10.1128/jvi.01009-14

190. Ledesma J, Fedele CG, Carro F, Lledó L, Sánchez-Seco MP, Tenorio A, et al. Independent lineage of lymphocytic choriomeningitis virus in wood mice (Apodemus sylvaticus), Spain. Emerging Infectious Diseases. 2009;15: 1677–1680. doi:10.3201/eid1510.090563

191. Lawton JH, Woodroffe GL. Habitat and the Distribution of Water Voles: Why are there Gaps in a Species' Range? The Journal of Animal Ecology. JSTOR; 1991;60: 79. doi:10.2307/5446

192. Hutterer R, Kryštufek B, Yigit N, Mitsain G, Palomo LJ, Henttonen H, et al. Myodes glareolus (errata version published in 2017). In: The IUCN Red List of Threatened Species 2016 [Internet]. 2016 [cited 1 Feb 2020]. Available: https://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T4973A22372716.en

193. Kryštufek B, Vohralík V, Zima J, Zagorodnyuk I. Microtus agrestis (Field Vole). In: The IUCN Red List of Threatened Species 2016 [Internet]. 2016 [cited 1 Feb 2020]. Available: https://www.iucnredlist.org/species/13426/115112050#errata

194. Burthe S, Telfer S, Lambin X, Bennett M, Carslake D, Smith A, et al. Cowpox virus infection in natural field vole Microtus agrestis populations: Delayed density dependence and individual risk. Journal of Animal Ecology. John Wiley & Sons, Ltd; 2006;75: 1416–1425. doi:10.1111/j.1365-2656.2006.01166.x

195. Reil D, Rosenfeld UM, Imholt C, Schmidt S, Ulrich RG, Eccard JA, et al. Puumala hantavirus infections in bank vole populations: Host and virus dynamics in Central Europe. BMC Ecology. BioMed Central Ltd.; 2017;17: 9. doi:10.1186/s12898-017-0118-z

196. Schlegel M, Kindler E, Essbauer SS, Wolf R, Thiel J, Groschup MH, et al. Tula virus infections in the Eurasian water vole in central Europe. Vector-Borne and Zoonotic Diseases. 2012;12: 503–513. doi:10.1089/vbz.2011.0784

197. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor symposia on quantitative biology. 1986;51 Pt 1: 263–73. doi:10.1101/sqb.1986.051.01.032

198. Morley AA. Digital PCR: A brief history. Biomolecular Detection and Quantification. Elsevier; 2014;1: 1–2. doi:10.1016/J.BDQ.2014.06.001

199. Compton T. PCR protocols — A guide to methods and applications. PCR Protocols: A Guide to Methods and Applications. Academic Press; 1990. doi:10.1016/0968-

0004(90)90252-7

200. Graf EH, Simmon KE, Tardif KD, Hymas W, Flygare S, Eilbeck K, et al. Unbiased Detection of Respiratory Viruses by Use of RNA Sequencing-Based Metagenomics: a Systematic Comparison to a Commercial PCR Panel. Journal of clinical microbiology. American Society for Microbiology Journals; 2016;54: 1000–7. doi:10.1128/JCM.03060-15

201. Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, Mead D, et al. Genomic analysis of uncultured marine viral communities. Proceedings of the National Academy of Sciences of the United States of America. National Academy of Sciences; 2002;99: 14250–5. doi:10.1073/pnas.202488399

202. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, et al. Metagenomic analyses of an uncultured viral community from human feces. Journal of bacteriology. American Society for Microbiology Journals; 2003;185: 6220–3. doi:10.1128/jb.185.20.6220-6223.2003

203. Illumina inc. An introduction to Next-Generation Sequencing Technology [Internet]. 2017. pp. 1–8. doi:http://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

204. Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE. Landscape of next-generation sequencing technologies. Analytical chemistry. NIH Public Access; 2011;83: 4327–41. doi:10.1021/ac2010857

205. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The Third Revolution in Sequencing Technology. Trends in genetics : TIG. Elsevier; 2018;34: 666–681. doi:10.1016/j.tig.2018.05.008

206. Lu H, Giordano F, Ning Z. Oxford Nanopore MinION Sequencing and Genome Assembly. Genomics, Proteomics & Bioinformatics. Elsevier; 2016;14: 265–279. doi:10.1016/J.GPB.2016.05.004

207. Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nature Biotechnology. Nature Publishing Group; 2018;36: 338–345. doi:10.1038/nbt.4060

208. Laver T, Harrison J, O'Neill PA, Moore K, Farbos A, Paszkiewicz K, et al. Assessing the performance of the Oxford Nanopore Technologies MinION. Biomolecular Detection and Quantification. Elsevier; 2015;3: 1–8. doi:10.1016/J.BDQ.2015.02.001

209. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biology. BioMed Central; 2016;17: 239. doi:10.1186/s13059-016-1103-0

210. Walter MC, Zwirglmaier K, Vette P, Holowachuk SA, Stoecker K, Genzel GH, et al. MinION as part of a biomedical rapidly deployable laboratory. Journal of Biotechnology. Elsevier; 2017;250: 16–22. doi:10.1016/J.JBIOTEC.2016.12.006

211. Faria NR, Sabino EC, Nunes MRT, Alcantara LCJ, Loman NJ, Pybus OG. Mobile real-time surveillance of Zika virus in Brazil. Genome Medicine. BioMed Central; 2016;8: 97. doi:10.1186/s13073-016-0356-2

212. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. Nature. Nature Publishing Group; 2016;530: 228–232. doi:10.1038/nature16996

213. Hoenen T, Groseth A, Rosenke K, Fischer RJ, Hoenen A, Judson SD, et al. Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool. Emerging infectious diseases. Centers for Disease Control and Prevention; 2016;22: 331–4. doi:10.3201/eid2202.151796

214. Grzybek M, Bajer A, Bednarska M, Al-Sarraf M, Behnke-Borowczyk J, Harris PD, et al. Long-term spatiotemporal stability and dynamic changes in helminth infracommunities of bank voles (Myodes glareolus) in NE Poland. Parasitology. 2015;142: 1722–1743. doi:10.1017/s0031182015001225

215. Behnke JM, Bajer A, Behnke-Borowczyk J, Clisham N, Gilbert F, Glover A, et al. Long-term spatiotemporal stability and dynamic changes in helminth infracommunities of spiny mice (*Acomys dimidiatus*) in St. Katherine's Protectorate, Sinai, Egypt. Parasitology. Cambridge University Press; 2019;146: 50–73. doi:10.1017/S0031182018000987

216. Tsoleridis T. Identification of novel alphacoronaviruses in European wildlife [Internet]. University of Nottingham. 2018. Available: http://eprints.nottingham.ac.uk/50287/

217. Klempa B, Fichet-Calvet E, Lecompte E, Auste B, Aniskin V, Meisel H, et al. Hantavirus in African wood mouse, Guinea. Emerging infectious diseases. 2006;12: 838–40. doi:10.3201/eid1205.051487

218. Nix WA, Maher K, Johansson ES, Niklasson B, Lindberg AM, Pallansch MA, et al. Detection of all known parechoviruses by real-time PCR. Journal of Clinical Microbiology. 2008;46: 2519–

2524. doi:10.1128/JCM.00277-08

219. Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, et al. ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic acids research. Oxford University Press; 2012;40: D593-8. doi:10.1093/nar/gkr859

220. Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, et al. A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing. Aguilar P V., editor. PLoS ONE. 2011;6: e28879. doi:10.1371/journal.pone.0028879

221. Nix WA, Maher K, Pallansch MA, Oberste MS. Parechovirus typing in clinical specimens by nested or semi-nested PCR coupled with sequencing. Journal of Clinical Virology. Elsevier B.V.; 2010;48: 202–207. doi:10.1016/j.jcv.2010.04.007

222. Tkachenko EA, Witkowski PT, Radosa L, Dzagurova TK, Okulova NM, Yunicheva Y V., et al. Adler hantavirus, a new genetic variant of Tula virus identified in Major's pine voles (Microtus majori) sampled in southern European Russia. Infection, Genetics and Evolution. Elsevier; 2015;29: 156–163. doi:10.1016/J.MEEGID.2014.11.018

223. Yanagihara R, Song J-W, Arai S, Cook JA, Parmenter C, Sumibcay L, et al. Phylogenetically Distinct Hantaviruses in the Masked Shrew (Sorex cinereus) and Dusky Shrew (Sorex monticolus) in the United States. The American Journal of Tropical Medicine and Hygiene. The American Society of Tropical Medicine and Hygiene; 2008;78: 348–351. doi:10.4269/ajtmh.2008.78.348

224. Arai S, Taniguchi S, Aoki K, Yoshikawa Y, Kyuwa S, Tanaka-Taya K, et al. Molecular phylogeny of a genetically divergent hantavirus harbored by the Geoffroy's rousette (Rousettus amplexicaudatus), a frugivorous bat species in the Philippines. Infection, Genetics and Evolution. Elsevier; 2016;45: 26–32. doi:10.1016/J.MEEGID.2016.08.008

225. Gu SH, Hejduk J, Markowski J, Kang HJ, Markowski M, Połatyńska M, et al. Co-circulation of soricid- and talpid-borne hantaviruses in Poland. Infection, Genetics and Evolution. Elsevier; 2014;28: 296–303. doi:10.1016/J.MEEGID.2014.10.017

226. Hofmann J, Weiss S, Kuhns M, Zinke A, Heinsberger H, Kruger DH. Importation of Human Seoul Virus Infection to Germany from Indonesia. Emerging infectious diseases. Centers for Disease Control and Prevention; 2018;24: 1099–1102. doi:10.3201/eid2406.172044

227. Voutilainen L, Kallio ER, Niemimaa J, Vapalahti O, Henttonen H. Temporal dynamics of Puumala hantavirus infection in cyclic populations of bank voles. Scientific reports. Nature Publishing Group; 2016;6: 21323. doi:10.1038/srep21323

228. Lambin X, Petty SJ, Mackinnon JL. Cyclic dynamics in field vole populations and generalist predation. Journal of Animal Ecology. John Wiley & Sons, Ltd (10.1111); 2000;69: 106–119. doi:10.1046/j.1365-2656.2000.00380.x

229. Tapper S. The Effect of Fluctuating Vole Numbers (Microtus agrestis) on a Population of Weasels (Mustela nivalis) on Farmland. The Journal of Animal Ecology. 1979;48: 603. doi:10.2307/4182

230. ICTV. ICTV Ninth Report; 2009 Taxonomy Release [Internet]. [cited 7 Mar 2019]. Available: https://talk.ictvonline.org/ictv-reports/ictv_9th_report/negative-sense-rna-viruses-2011/w/negrna_viruses/205/bunyaviridae

231. Maes P, Klempa B, Clement J, Matthijnssens J, Gajdusek DC, Krüger DH, et al. A proposal for new criteria for the classification of hantaviruses, based on S and M segment protein sequences. Infection, Genetics and Evolution. 2009;9: 813–820. doi:10.1016/j.meegid.2009.04.012

232. Asnani M, Kumar P, Hellen CUT. Widespread distribution and structural diversity of Type IV IRESs in members of Picornaviridae. Virology. Academic Press; 2015;478: 61–74. doi:10.1016/J.VIROL.2015.02.016

233. Ghazi F, Stanway G, Hughes PJ, Hyypi√§ T. Molecular analysis of human parechovirus type 2 (formerly echovirus 23). Journal of General Virology. Microbiology Society; 1998;79: 2641–2650. doi:10.1099/0022-1317-79-11-2641

234. Phan TG, Vo NP, Simmonds P, Samayoa E, Naccache S, Chiu CY, et al. rosavirus: the prototype of a proposed new genus of the Picornaviridae family. Virus Genes. Springer US; 2013;47: 556–558. doi:10.1007/s11262-013-0968-1

235. Jang SK. Internal initiation: IRES elements of picornaviruses and hepatitis c virus. Virus Research. Elsevier; 2006;119: 2–15. doi:10.1016/J.VIRUSRES.2005.11.003

236. Reuter G, Pankovics P, Knowles NJ, Boros Á. Two closely related novel picornaviruses in cattle and sheep in Hungary from 2008 to 2009, proposed as members of a new genus in the family Picornaviridae. Journal of virology. American Society for Microbiology; 2012;86: 13295–302. doi:10.1128/JVI.01142-12

237. Niklasson B, Hultman T, Kallies R, Niedrig M, Nilsson R, Berggren P-O, et al. The BioBreeding rat diabetes model is infected with Ljungan virus. Diabetologia. Springer-Verlag; 2007;50: 1559–1560. doi:10.1007/s00125-007-0659-3

238. de Crom SCM, Rossen JWA, van Furth AM, Obihara CC. Enterovirus and parechovirus infection in children: a brief overview. European journal of pediatrics. Springer; 2016;175: 1023–9. doi:10.1007/s00431-016-2725-7

239. Fevola C, Rossi C, Rosà R, Nordström Å, Ecke F, Magnusson M, et al. Distribution and Seasonal Variation of Ljungan Virus in Bank Voles (Myodes glareolus) in Fennoscandia. Journal of Wildlife Diseases. Wildife Disease Association; 2017; 2016-06–145. doi:10.7589/2016-06-145

240. Tolf C, Gullberg M, Johansson ES, Tesh RB, Andersson B, Lindberg AM. Molecular characterization of a novel Ljungan virus (Parechovirus; Picornaviridae) reveals a fourth genotype and indicates ancestral recombination. Journal of General Virology. Microbiology Society; 2009;90: 843–853. doi:10.1099/vir.0.007948-0

241. Lau SKP, Woo PCY, Li KSM, Zhang H-J, Fan RYY, Zhang AJX, et al. Identification of Novel rosavirus Species That Infects Diverse Rodent Species and Causes Multisystemic Dissemination in Mouse Model. Delwart EL, editor. PLOS Pathogens. Public Library of Science; 2016;12: e1005911. doi:10.1371/journal.ppat.1005911

242. Firth C, Bhat M, Firth MA, Williams SH, Frye MJ, Simmonds P, et al. Detection of Zoonotic Pathogens and Characterization of Novel Viruses Carried by Commensal Rattus norvegicus in New York City. mBio. American Society for Microbiology; 2014;5: e01933-14-e01933-14. doi:10.1128/mBio.01933-14

243. Lim ES, Cao S, Holtz LR, Antonio M, Stine OC, Wang D. Discovery of rosavirus 2, a novel variant of a rodent-associated picornavirus, in children from The Gambia. Virology. Academic Press; 2014;454–455: 25–33. doi:10.1016/J.VIROL.2014.01.018

244. Boros Á, Orlovácz K, Pankovics P, Szekeres S, Földvári G, Fahsbender E, et al. Diverse picornaviruses are prevalent among free-living and laboratory rats (Rattus norvegicus) in Hungary and can cause disseminated infections. Infection, Genetics and Evolution. 2019;75: 103988. doi:10.1016/j.meegid.2019.103988

245. Roos RP. Pathogenesis of Theiler's Murine Encephalomyelitis Virus-Induced Disease. Molecular Biology of picornavirus. American Society of Microbiology; 2002. pp. 427–435. doi:10.1128/9781555817916.ch34

246. Brahic M, Bureau J-F, Michiels T. THE GENETICS OF THE PERSISTENT INFECTION AND DEMYELINATING DISEASE CAUSED BY THEILER'S VIRUS. Annual Review of Microbiology. Annual Reviews; 2005;59: 279–298. doi:10.1146/annurev.micro.59.030804.121242

247. Tan SZK, Tan MZY, Prabakaran M. Saffold virus, an emerging human cardiovirus. Reviews in Medical Virology. 2017;27: e1908. doi:10.1002/rmv.1908

248. Himeda T, Ohara Y. Saffold virus, a novel human cardiovirus with unknown pathogenicity. Journal of virology. American Society for Microbiology Journals; 2012;86: 1292–6. doi:10.1128/JVI.06087-11

249. Abed Y, Boivin G. New Saffold cardioviruses in 3 children, Canada. Emerging infectious diseases. Centers for Disease Control and Prevention; 2008;14: 834–6. doi:10.3201/eid1405.071675

250. Chiu CY, Greninger AL, Kanada K, Kwok T, Fischer KF, Runckel C, et al. Identification of cardioviruses related to Theiler's murine encephalomyelitis virus in human infections. Proceedings of the National Academy of Sciences. 2008;105: 14124–14129. doi:10.1073/pnas.0805968105

251. Drexler JF, Baumgarte S, Eschbach-Bludau M, Simon A, Kemen C, Bode U, et al. Human cardioviruses, meningitis, and sudden infant death syndrome in children. Emerging infectious diseases. Centers for Disease Control and Prevention; 2011;17: 2313–5. doi:10.3201/eid1712.111037

252. RILEY, Lela K, GOHNDRONE, Judith D, MYLES, Matthew H. BOONE CARDIOVIRUS [Internet]. PCT/US2013/050245, 2014. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2014014767

253. Reuter G, Pankovics P, Knowles NJ, Boros A. Two Closely Related Novel picornaviruses in Cattle and Sheep in Hungary from 2008 to 2009, Proposed as Members of a New Genus in the Family Picornaviridae. Journal of Virology. 2012;86: 13295–13302. doi:10.1128/JVI.01142-12

254. Lu G, Huang M, Chen X, Sun Y, Huang J, Hu R, et al. Identification and genome characterization of a novel feline picornavirus proposed in the hunnivirus genus. Infection, Genetics and Evolution. 2019;71: 47–50. doi:10.1016/j.meegid.2019.03.011

255. Zell R, Delwart E, Gorbalenya AE, Hovi T, King AMQ, Knowles NJ, et al. ICTV Virus Taxonomy Profile: Picornaviridae. Journal of

General Virology. Microbiology Society; 2017;98: 2421–2422. doi:10.1099/jgv.0.000911

256. Sboner A, Mu X, Greenbaum D, Auerbach RK, Gerstein MB. The real cost of sequencing: higher than you think! Genome Biology. BioMed Central; 2011;12: 125. doi:10.1186/gb-2011-12-8-125

257. Muir P, Li S, Lou S, Wang D, Spakowicz DJ, Salichos L, et al. The real cost of sequencing: scaling computation to keep pace with data generation. Genome Biology. BioMed Central; 2016;17: 53. doi:10.1186/s13059-016-0917-0

258. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic acids research. Oxford University Press; 2018;46: D708–D717. doi:10.1093/nar/gkx932

259. Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, et al. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2017). Archives of Virology. Springer Vienna; 2017;162: 2505–2538. doi:10.1007/s00705-017-3358-5

260. Slough MM, Chandran K, Jangra RK. Two Point Mutations in Old World Hantavirus Glycoproteins Afford the Generation of Highly Infectious Recombinant Vesicular Stomatitis Virus Vectors. mBio. American Society for Microbiology (ASM); 2019;10. doi:10.1128/mBio.02372-18

261. Ling J, Vaheri A, Hepojoki S, Levanov L, Jäälskeläinen A, Henttonen H, et al. Serological survey of Seewis virus antibodies in patients suspected for hantavirus infection in Finland; a cross-reaction between Puumala virus antiserum with Seewis virus N protein? Journal of General Virology. Microbiology Society; 2015;96: 1664–1675. doi:10.1099/vir.0.000127

262. Clement J, Frans J, Van Ranst M. Human Tula virus infection or rat-bite fever? European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2003;22: 332–3; author reply 334-5. doi:10.1007/s10096-003-0921-7

263. Reynes JM, Carli D, Boukezia N, Debruyne M, Herti S. Tula hantavirus infection in a hospitalised patient, France, June 2015. Euro surveillance. 2015;20. doi:10.2807/1560-7917.ES.2015.20.50.30095

264. Adlhoch C, Brown IH, Angelova SG, Bálint Á, Bouwstra R, Buda S,

et al. Highly pathogenic avian influenza A(H5N8) outbreaks: protection and management of exposed people in Europe, 2014/15 and 2016. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. European Centre for Disease Prevention and Control; 2016;21. doi:10.2807/1560-7917.ES.2016.21.49.30419

265. Tsoleridis T, Chappell J, Onianwa O, Monchatre-Leroy E, Umhang G, Shi M, et al. Discovery of novel highly divergent RNA viruses in European rodents and rabbits. Access Microbiology. Microbiology Society; 2019;1: 657. doi:10.1099/acmi.ac2019.po0407

266. Shi M, Neville P, Nicholson J, Eden J-S, Imrie A, Holmes EC. High-Resolution Metatranscriptomics Reveals the Ecological Dynamics of Mosquito-Associated RNA Viruses in Western Australia. Journal of virology. American Society for Microbiology Journals; 2017;91: e00680-17. doi:10.1128/JVI.00680-17

267. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, Li C-X, et al. Redefining the invertebrate RNA virosphere. Nature. Nature Publishing Group; 2016;540: 539–543. doi:10.1038/nature20167

268. Brinkmann A, Andrusch A, Belka A, Wylezich C, Höper D, Pohlmann A, et al. Proficiency testing of virus diagnostics based on bioinformatics analysis of simulated in silico high-throughput sequencing datasets. Tang Y-W, editor. Journal of Clinical Microbiology. 2019;57. doi:10.1128/jcm.00466-19

269. Runckel C, Flenniken ML, Engel JC, Ruby JG, Ganem D, Andino R, et al. Temporal Analysis of the Honey Bee Microbiome Reveals Four Novel Viruses and Seasonal Prevalence of Known Viruses, Nosema, and Crithidia. Moritz RFA, editor. PLoS ONE. Public Library of Science; 2011;6: e20656. doi:10.1371/journal.pone.0020656

270. Evander M, Eriksson I, Pettersson L, Juto P, Ahlm C, Olsson GE, et al. Puumala hantavirus viremia diagnosed by real-time reverse transcriptase PCR using samples from patients with hemorrhagic fever and renal syndrome. Journal of clinical microbiology. American Society for Microbiology Journals; 2007;45: 2491–7. doi:10.1128/JCM.01902-06

271. Vaheri A, Vapalahti O, Plyusnin A. How to diagnose hantavirus infections and detect them in rodents and insectivores. Reviews in Medical Virology. John Wiley & Sons, Ltd; 2008;18: 277–288. doi:10.1002/rmv.581

272. Zheng Y, Wei J, Zhou B-Y, Xu Y, Dong J-H, Guan L-Y, et al. Long-term persistence of anti-hantavirus antibodies in sera of patients undergoing hemorrhagic fever with renal syndrome and subjects

vaccinated against the disease. Infectious Diseases. Taylor & Francis; 2016;48: 262–266. doi:10.3109/23744235.2015.1121289

273. Duggan JM, Close R, McCann L, Wright D, Keys M, McCarthy N, et al. A seroprevalence study to determine the frequency of hantavirus infection in people exposed to wild and pet fancy rats in England. Epidemiology and Infection. 2017;145: 2458–2465. doi:10.1017/S0950268817001480

274. Pleace N, Bretherton J. Crisis Skylight : Final Report of the University of York Evaluation. Crisis; 2017; Available: http://eprints.whiterose.ac.uk/112010/

275. Fazel S, Geddes JR, Kushel M. The health of homeless people in high-income countries: descriptive epidemiology, health consequences, and clinical and policy recommendations. Lancet (London, England). Europe PMC Funders; 2014;384: 1529–40. doi:10.1016/S0140-6736(14)61132-6

276. Glass GE, Watson AJ, LeDuc JW, Kelen GD, Quinn TC, Childs JE. Infection with a Ratborne Hantavirus in US Residents Is Consistently Associated with Hypertensive Renal Disease [Internet]. The Journal of Infectious Diseases. Oxford University Press; pp. 614–620. doi:10.2307/30112626

277. Khabbaz RF, Ksiazek TG, Caiaffa WT, Rollin PE, Taylor E, Vlahov D. Seoul Hantavirus Seropositivity Among Injecting Drug Users In Baltimore. Journal of Infectious Diseases. Narnia; 1994;170: 1636–1637. doi:10.1093/infdis/170.6.1636

278. Groen J, Gerding M, Koeman JP, Roholl PJM, van Amerongen G, Jordans HGM, et al. A Macaque Model for Hantavirus Infection. Journal of Infectious Diseases. 1995;172: 38–44. doi:10.1093/infdis/172.1.38

279. Klingström J, Plyusnin A, Vaheri A, Lundkvist A. Wild-type Puumala hantavirus infection induces cytokines, C-reactive protein, creatinine, and nitric oxide in cynomolgus macaques. Journal of virology. American Society for Microbiology (ASM); 2002;76: 444–9. Available: http://www.ncbi.nlm.nih.gov/pubmed/11739712

280. Yanagihara R, Amyx HL, Lee P-W, Asher DM, Gibbs CJ, Gajdusek DC. Experimental hantavirus infection in nonhuman primates. Archives of Virology. Springer-Verlag; 1988;101: 125–130. doi:10.1007/BF01314657

281. Mertens M, Essbauer SS, Rang A, Schröder J, Splettstoesser WD, Kretzschmar C, et al. Non-human primates in outdoor enclosures: Risk for infection with rodent-borne hantaviruses.

Veterinary Microbiology. 2011;147: 420–425. doi:10.1016/j.vetmic.2010.07.018

282. Harris H, Costella A, Harris R, Mandal S, and other contributors. Hepatitis C in England, 2019 report: Working to eliminate hepatitis C as a major public health threat [Internet]. 2019. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/732469/HCV_IN_THE_UK_2018_UK.pdf

283. Lederer S, Lattwein E, Hanke M, Sonnenberg K, Stoecker W, Lundkvist Å, et al. Indirect Immunofluorescence Assay for the Simultaneous Detection of Antibodies against Clinically Important Old and New World Hantaviruses. Reithinger R, editor. PLoS Neglected Tropical Diseases. Public Library of Science; 2013;7: e2157. doi:10.1371/journal.pntd.0002157

284. Granerod J, Tam CC, Crowcroft NS, Davies NWS, Borchert M, Thomas SL. Challenge of the unknown. A systematic review of acute encephalitis in non-outbreak situations. Neurology. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology; 2010;75: 924–32. doi:10.1212/WNL.0b013e3181f11d65

285. Mailles A, Stahl J. Infectious Encephalitis in France in 2007: A National Prospective Study. Clinical Infectious Diseases. Narnia; 2009;49: 1838–1847. doi:10.1086/648419

286. Li T, Mbala-Kingebeni P, Naccache SN, Thézé J, Bouquet J, Federman S, et al. Metagenomic Next-Generation Sequencing of the 2014 Ebola Virus Disease Outbreak in the Democratic Republic of the Congo. Journal of clinical microbiology. American Society for Microbiology Journals; 2019;57: e00827-19. doi:10.1128/JCM.00827-19

287. Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, et al. Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. Lin B, editor. PLOS ONE. Public Library of Science; 2019;14: e0218318. doi:10.1371/journal.pone.0218318

288. Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. The New England journal of medicine. NIH Public Access; 2014;370: 2408–17. doi:10.1056/NEJMoa1401268

289. Vilsker M, Moosa Y, Nooij S, Fonseca V, Ghysens Y, Dumon K, et al. Genome Detective: an automated system for virus

identification from high-throughput sequencing data. Birol I, editor. Bioinformatics. 2019;35: 871–873. doi:10.1093/bioinformatics/bty695

290. Saha S, Ramesh A, Kalantar K, Malaker R, Hasanuzzaman M, Khan LM, et al. Unbiased metagenomic sequencing for pediatric meningitis in Bangladesh reveals neuroinvasive Chikungunya virus outbreak and other unrealized pathogens. bioRxiv. Cold Spring Harbor Laboratory; 2019; 579532. doi:10.1101/579532

291. Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P. The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. The Journal of general virology. Microbiology Society; 2011;92: 233–46. doi:10.1099/vir.0.027490-0

292. Simons JN, Pilot-Matias TJ, Leary TP, Dawson GJ, Desai SM, Schlauder GG, et al. Identification of two flavivirus-like genomes in the GB hepatitis agent. Proceedings of the National Academy of Sciences of the United States of America. National Academy of Sciences; 1995;92: 3401–5. doi:10.1073/pnas.92.8.3401

293. Alter HJ, Nakatsuji Y, Melpolder J, Wages J, Wesley R, Shih JW-K, et al. The Incidence of Transfusion-Associated Hepatitis G Virus Infection and Its Relation to Liver Disease. New England Journal of Medicine. 1997;336: 747–754. doi:10.1056/NEJM199703133361102

294. Bhattarai N, Stapleton JT. GB virus C: the good boy virus? Trends in Microbiology. 2012;20: 124–130. doi:10.1016/j.tim.2012.01.004

295. Blair CS, Davidson F, Lycett C, McDonald DM, Haydon GH, Yap PL, et al. Prevalence, Incidence, and Clinical Characteristics of Hepatitis G Virus/GB Virus C Infection in Scottish Blood Donors. The Journal of Infectious Diseases. Narnia; 1998;178: 1779–1782. doi:10.1086/314508

296. Minton J, Iqbal A, Eskiturk A, Irving W, Davies J. Hepatitis G virus infection in lymphoma and in blood donors. Journal of Clinical Pathology. BMJ Publishing Group; 1998;51: 676–678. doi:10.1136/jcp.51.9.676

297. Miao Z, Gao L, Song Y, Yang M, Zhang M, Lou J, et al. Prevalence and Clinical Impact of Human Pegivirus-1 Infection in HIV-1-Infected Individuals in Yunnan, China. Viruses. Multidisciplinary Digital Publishing Institute (MDPI); 2017;9. doi:10.3390/v9020028

298. De Filippi F, Lampertico P, Soffredini R, Rumi M., Lunghi G, Aroldi A, et al. High prevalence, low pathogenicity in hepatitis G virus in kidney transplant recipients. Digestive and Liver Disease. W.B.

Saunders; 2001;33: 477–479. doi:10.1016/S1590-8658(01)80025-6

299. Kallinowski B, Seipp S, Dengler T, Klar E, Theilmann L, Stremmel W. Clinical impact of hepatitis G virus infection in heart and liver transplant recipients. Transplantation proceedings. 2002;34: 2288–91. doi:10.1016/s0041-1345(02)03239-6

300. Corbi C, Traineau R, Esperou H, Ravera N, Portelette E, Benbunan M, et al. Prevalence and clinical features of hepatitis G virus infection in bone marrow allograft recipients. Bone Marrow Transplantation. 1997;20: 965–968. doi:10.1038/sj.bmt.1701005

301. Scallan MF, Clutterbuck D, Jarvis LM, Scott GR, Simmonds P. Sexual transmission of GB virus C/hepatitis G virus. Journal of Medical Virology. John Wiley & Sons, Ltd; 1998;55: 203–208. doi:10.1002/(SICI)1096-9071(199807)55:3<203::AID-JMV4>3.0.CO;2-5

302. Moaven LD, Tennakoon PS, Bowden DS, Locarnini SA. Mother-to-baby transmission of hepatitis G virus. Medical Journal of Australia. John Wiley & Sons, Ltd; 1996;165: 84–85. doi:10.5694/j.1326-5377.1996.tb124854.x

303. Miura T, Takayama S, Taki M, Shimotohno K. Mother-to-child transmission of hepatitis G virus. Pediatrics. American Academy of Pediatrics; 1998;102: 1222–3. doi:10.1542/peds.102.5.1222

304. Izumi T, Sakata K, Okuzaki D, Inokuchi S, Tamura T, Motooka D, et al. Characterization of human pegivirus infection in liver transplantation recipients. Journal of Medical Virology. 2019; jmv.25555. doi:10.1002/jmv.25555

305. NHS Blood and Transplant. Tests we carry out [Internet]. [cited 15 Sep 2019]. Available: https://www.blood.co.uk/the-donation-process/further-information/tests-we-carry-out/

306. Chang CM, Stapleton JT, Klinzman D, McLinden JH, Purdue MP, Katki HA, et al. GBV-C Infection and Risk of NHL among U.S. Adults. Cancer Research. American Association for Cancer Research; 2014;74: 5553–5560. doi:10.1158/0008-5472.CAN-14-0209

307. Krajden M, Yu A, Braybrook H, Lai AS, Mak A, Chow R, et al. GBV-C/hepatitis G virus infection and non-Hodgkin lymphoma: a case control study. International Journal of Cancer. John Wiley & Sons, Ltd; 2010;126: NA-NA. doi:10.1002/ijc.25035

308. Fama A, Xiang J, Link BK, Allmer C, Klinzman D, Feldman AL, et al. Human Pegivirus infection and lymphoma risk and prognosis: a North American study. British Journal of Haematology. John Wiley & Sons, Ltd (10.1111); 2018;182: 644–653.

doi:10.1111/bjh.15416

309. Tucker TJ, Smuts HEM, Eedes C, Knobel GD, Eickhaus P, Robson SC, et al. Evidence that the GBV-C/hepatitis G virus is primarily a lymphotropic virus. Journal of Medical Virology. John Wiley & Sons, Ltd; 2000;61: 52–58. doi:10.1002/(SICI)1096-9071(200005)61:1<52::AID-JMV8>3.0.CO;2-L

310. Pietersma F, Piriou E, van Baarle D. Immune surveillance of EBV-infected B cells and the development of non-Hodgkin lymphomas in immunocompromised patients. Leukemia & Lymphoma. Taylor & Francis; 2008;49: 1028–1041. doi:10.1080/10428190801911662

311. Kriesel JD, Hobbs MR, Jones BB, Milash B, Nagra RM, Fischer KF. Deep Sequencing for the Detection of Virus-Like Sequences in the Brains of Patients with Multiple Sclerosis: Detection of GBV-C in Human Brain. Feng Y, editor. PLoS ONE. Public Library of Science; 2012;7: e31886. doi:10.1371/journal.pone.0031886

312. Liu Z, Zhang Y, Wei F, Xu M, Mou D, Zhang T, et al. Detection of GB virus C genomic sequence in the cerebrospinal fluid of a HIV-infected patient in China: A case report and literature review [Internet]. Epidemiology and Infection. Cambridge University Press; 2016. pp. 106–112. doi:10.1017/S0950268815001326

313. Fridholm H, Østergaard Sørensen L, Rosenstierne MW, Nielsen H, Sellebjerg F, Bengård Andersen Å, et al. Human pegivirus detected in a patient with severe encephalitis using a metagenomic pan-virus array. Journal of Clinical Virology. Elsevier B.V.; 2016;77: 5–8. doi:10.1016/j.jcv.2016.01.013

314. Bukowska-Ośko I, Perlejewski K, Pawełczyk A, Rydzanicz M, Pollak A, Popiel M, et al. Human pegivirus in patients with encephalitis of unclear etiology, Poland. Emerging Infectious Diseases. Centers for Disease Control and Prevention; 2018;24: 1785–1794. doi:10.3201/eid2410.180161

315. Lau SKP, Yip CCY, Tsoi H -w., Lee RA, So L -y., Lau Y -l., et al. Clinical Features and Complete Genome Characterization of a Distinct Human Rhinovirus (HRV) Genetic Cluster, Probably Representing a Previously Undetected HRV Species, HRV-C, Associated with Acute Respiratory Illness in Children. Journal of Clinical Microbiology. 2007;45: 3655–3664. doi:10.1128/JCM.01254-07

316. Lau SKP, Yip CCY, Woo PCY, Yuen K-Y. Human rhinovirus C: a newly discovered human rhinovirus species. Emerging health threats journal. Taylor & Francis; 2010;3: e2. doi:10.3134/ehtj.10.002

317. Linsuwanon P, Payungporn S, Samransamruajkit R, Posuwan N, Makkoch J, Theanboonlers A, et al. High prevalence of human rhinovirus C infection in Thai children with acute lower respiratory tract disease. Journal of Infection. W.B. Saunders; 2009;59: 115–121. doi:10.1016/J.JINF.2009.05.009

318. Jartti T, Lehtinen P, Vuorinen T, Osterback R, van den Hoogen B, Osterhaus ADME, et al. Respiratory picornaviruses and respiratory syncytial virus as causative agents of acute expiratory wheezing in children. Emerging infectious diseases. Centers for Disease Control and Prevention; 2004;10: 1095–101. doi:10.3201/eid1006.030629

319. Lupo J, Schuffenecker I, Morel-Baccard C, Bardet J, Payen V, Kaiser L, et al. Disseminated rhinovirus C8 infection with infectious virus in blood and fatal outcome in a child with repeated episodes of bronchiolitis. Journal of clinical microbiology. American Society for Microbiology Journals; 2015;53: 1775–7. doi:10.1128/JCM.03484-14

320. Pelkonen T, Roine I, Anjos E, Kaijalainen S, Roivainen M, Peltola H, et al. picornaviruses in cerebrospinal fluid of children with meningitis in Luanda, Angola. Journal of Medical Virology. John Wiley & Sons, Ltd; 2012;84: 1080–1083. doi:10.1002/jmv.23304

321. Smith DB, Becher P, Bukh J, Gould EA, Meyers G, Monath T, et al. Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family. Journal of General Virology. 2016;97: 2894–2907. doi:10.1099/jgv.0.000612

322. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators G 2015 D and II and P. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet (London, England). Elsevier; 2016;388: 1545–1602. doi:10.1016/S0140-6736(16)31678-6

323. Laskus T, Radkowski M, Bednarska A, Wilkinson J, Adair D, Nowicki M, et al. Detection and analysis of hepatitis C virus sequences in cerebrospinal fluid. Journal of virology. American Society for Microbiology; 2002;76: 10064–8. doi:10.1128/jvi.76.19.10064-10068.2002

324. Adinolfi LE, Nevola R, Lus G, Restivo L, Guerrera B, Romano C, et al. Chronic hepatitis C virus infection and neurological and psychiatric disorders: an overview. World journal of gastroenterology. Baishideng Publishing Group Inc; 2015;21: 2269–80. doi:10.3748/wjg.v21.i8.2269

325. Seifert F, Struffert T, Hildebrandt M, Blümcke I, Brück W, Staykov D, et al. In vivo detection of hepatitis C virus (HCV) RNA in the brain in a case of encephalitis: evidence for HCV neuroinvasion. European Journal of Neurology. John Wiley & Sons, Ltd (10.1111); 2008;15: 214–218. doi:10.1111/j.1468-1331.2007.02044.x

326. Bolay H, Söylemezoğlu F, Nurlu G, Tuncer S, Vari K. PCR detected hepatitis C virus genome in the brain of a case with progressive encephalomyelitis with rigidity. Clinical Neurology and Neurosurgery. Elsevier; 1996;98: 305–308. doi:10.1016/0303-8467(96)00040-6

327. Yinda CK, Ghogomu SM, Conceição-Neto N, Beller L, Deboutte W, Vanhulle E, et al. Cameroonian fruit bats harbor divergent viruses, including rotavirus H, bastroviruses, and picobirnaviruses using an alternative genetic code. Virus evolution. Oxford University Press; 2018;4: vey008. doi:10.1093/ve/vey008

328. van Leeuwen M, Williams MMW, Koraka P, Simon JH, Smits SL, Osterhaus ADME. Human picobirnaviruses identified by molecular screening of diarrhea samples. Journal of clinical microbiology. American Society for Microbiology Journals; 2010;48: 1787–94. doi:10.1128/JCM.02452-09

329. Smits SL, van Leeuwen M, Schapendonk CME, Schürch AC, Bodewes R, Haagmans BL, et al. Picobirnaviruses in the Human Respiratory Tract. Emerging Infectious Diseases. 2012;18: 1538–1539. doi:10.3201/eid1809.120507

330. Wang H, Li S, Mahmood A, Yang S, Wang X, Shen Q, et al. Plasma virome of cattle from forest region revealed diverse small circular ssDNA viral genomes. Virology Journal. BioMed Central; 2018;15: 11. doi:10.1186/s12985-018-0923-9

331. Knox MA, Gedye KR, Hayman DTS. The Challenges of Analysing Highly Diverse Picobirnavirus Sequence Data. Viruses. Multidisciplinary Digital Publishing Institute (MDPI); 2018;10. doi:10.3390/v10120685

332. Nishizawa T, Okamoto H, Konishi K, Yoshizawa H, Miyakawa Y, Mayumi M. A Novel DNA Virus (TTV) Associated with Elevated Transaminase Levels in Posttransfusion Hepatitis of Unknown Etiology. Biochemical and Biophysical Research Communications. Academic Press; 1997;241: 92–97. doi:10.1006/BBRC.1997.7765

333. De Vlaminck I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, et al. Temporal Response of the Human Virome to Immunosuppression and Antiviral Therapy. Cell. Cell Press; 2013;155: 1178–1187. doi:10.1016/J.CELL.2013.10.034

334. Al-Qahtani AA, Alabsi ES, AbuOdeh R, Thalib L, El Zowalaty ME, Nasrallah GK. Prevalence of anelloviruses (TTV, TTMDV, and TTMV) in healthy blood donors and in patients infected with HBV or HCV in Qatar. Virology Journal. BioMed Central; 2016;13: 208. doi:10.1186/s12985-016-0664-6

335. McElvania TeKippe E, Wylie KM, Deych E, Sodergren E, Weinstock G, Storch GA. Increased Prevalence of Anellovirus in Pediatric Patients with Fever. Jhaveri R, editor. PLoS ONE. Public Library of Science; 2012;7: e50937. doi:10.1371/journal.pone.0050937

336. Gonzales-Gustavson E, Timoneda N, Fernandez-Cassi X, Caballero A, Abril JF, Buti M, et al. Identification of sapovirus GV.2, astrovirus VA3 and novel anelloviruses in serum from patients with acute hepatitis of unknown aetiology. Datta S, editor. PLOS ONE. Public Library of Science; 2017;12: e0185911. doi:10.1371/journal.pone.0185911

337. Loconsole G, Önelge N, Potere O, Giampetruzzi A, Bozan O, Satar S, et al. Identification and Characterization of *Citrus yellow vein clearing virus* , A Putative New Member of the Genus *Mandarivirus*. Phytopathology. The American Phytopathological Society ; 2012;102: 1168–1175. doi:10.1094/PHYTO-06-12-0140-R

338. Balique F, Lecoq H, Raoult D, Colson P. Can Plant Viruses Cross the Kingdom Border and Be Pathogenic to Humans? Viruses. Multidisciplinary Digital Publishing Institute; 2015;7: 2074–2098. doi:10.3390/v7042074

339. Philippe N, Legendre M, Doutre G, Couté Y, Poirot O, Lescot M, et al. Pandoraviruses: Amoeba Viruses with Genomes Up to 2.5 Mb Reaching That of Parasitic Eukaryotes. Science. 2013;341: 281 LP – 286. doi:10.1126/science.1239181

340. Colson P, Aherfi S, La Scola B. The role of giant viruses of amoebas in humans. Current Opinion in Microbiology. Elsevier Current Trends; 2016;31: 199–208. doi:10.1016/J.MIB.2016.04.012

341. Łanocha N, Kosik-Bogacka D, Maciejewska A, Sawczuk M, Wilk A, Kuźna-Grygiel W. The occurrence Acanthamoeba (Free living amoeba) in environmental and respiratory samples in Poland. Acta Protozoologica. 2009;48: 271–279.

342. Daudt C, Da Silva FRC, Lunardi M, Alves CBDT, Weber MN, Cibulski SP, et al. Papillomaviruses in ruminants: An update. Transboundary and Emerging Diseases. John Wiley & Sons, Ltd (10.1111); 2018;65: 1381–1395. doi:10.1111/tbed.12868

343. Ogino M, Ebihara H, Lee B-H, Araki K, Lundkvist A, Kawaoka Y, et al. Use of vesicular stomatitis virus pseudotypes bearing hantaan or seoul virus envelope proteins in a rapid and safe neutralization test. Clinical and diagnostic laboratory immunology. American Society for Microbiology; 2003;10: 154–60. doi:10.1128/cdli.10.1.154-160.2003

344. Ahlm C, Juto P, Stegmayr B, Settergren B, Wadell G, Tärnvik A, et al. Prevalence of Serum Antibodies to Hantaviruses in Northern Sweden as Measured by Recombinant Nucleocapsid Proteins. Scandinavian Journal of Infectious Diseases. Taylor & Francis; 1997;29: 349–354. doi:10.3109/00365549709011829

345. Ulrich R, Meisel H, Schütt M, Schmidt J, Kunz A, Klempa B, et al. [Prevalence of hantavirus infections in Germany]. [Internet]. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz. 2004. doi:10.1007/s00103-004-0858-8

346. Wang L, Crameri G. Emerging zoonotic viral diseases. Rev sci tech Off int Epiz. 2014;33: 569–581. Available: http://www.oie.int/doc/ged/D14089.PDF

347. Hantavirus. In: Viral Zone [Internet]. Available: http://viralzone.expasy.org/all_by_species/213.html

348. Kadambari S, Harvala H, Simmonds P, Pollard AJ, Sadarangani M. Strategies to improve detection and management of human parechovirus infection in young infants. The Lancet Infectious diseases. Elsevier; 2019;19: e51–e58. doi:10.1016/S1473-3099(18)30288-3

349. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Battistuzzi FU, editor. Molecular Biology and Evolution. 2018;35: 1547–1549. doi:10.1093/molbev/msy096

10. Appendices

10.1. Appendix 1

Virus Name	Accession	Virus Name	Accession
Variegated squirrel 1	LN713680.1	Borna disease virus isolate	AF158633.1
Variegated squirrel 1	LN713681.1	Borna disease virus	AY066023.1
Borna disease virus 1 strain	AJ311522.1	Borna disease virus vaccine	DQ680832.1
Borna disease virus (isolate	L76237.1	Aquatic bird bornavirus 1	KF578398.1
Borna disease virus (isolate	L76238.1	Canary bornavirus 3 isolate	KC595273.2
Borna disease virus 1	AB032031.1	Canary bornavirus 1 isolate	KC464471.1
Borna disease virus 1 strain	AB246670.1	Canary bornavirus 2 isolate	NC 027892.1
Borna disease virus 1 strain	AB258389.1	Parrot bornavirus 1 isolate	JX065207.1
Borna disease virus 1 strain	AY114163.1	Parrot bornavirus 1 strain	GU249595.2
Borna disease virus 1 strain	AY114161.1	Parrot bornavirus 2 isolate	FJ620690.1
Borna disease virus 1 strain	AY114162.1	Parrot bornavirus 2 isolate	EU781967.1
Borna disease virus 1 strain	AJ311523.1	Parrot bornavirus 4 isolate	JX065209.1
Borna disease virus 1 strain	AJ311521.1	Parrot bornavirus 4 isolate	JN035148.1
Borna disease virus 2 strain	AJ311524.1	Parrot bornavirus 4 isolate	JN014948.1
Borna disease virus	NC 001607.1	Parrot bornavirus 4 isolate	JN014949.1
Borna disease virus isolate	DQ680833.1	Parrot bornavirus 4 strain	GU249596.2

Table 17 Orthobornavirus glycoprotein sequences used to design pan-species screening primers. Mammalian orthobornavirus sequences are highlighted in blue.

10.2. Appendix 2

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

LETTER TO THE EDITOR

Identification of Infectious Agents in High-Throughput Sequencing Data Sets Is Easily Achievable Using Free, Cloud-Based Bioinformatics Platforms

Joseph G. Chappell,[a] Timothy Byaruhanga,[a,b] Theocharis Tsoleridis,[a] Jonathan K. Ball,[a,c] C. Patrick McClure[a,c]

[a]School of Life Sciences, University of Nottingham, Nottingham, United Kingdom
[b]Uganda Virus Research Institute, Entebbe, Uganda
[c]NIHR Nottingham BRC, Nottingham University Hospitals NHS Trust and the University of Nottingham, Nottingham, United Kingdom

KEYWORDS high-throughput sequencing, bioinformatics, virology

It was with great interest that we read the recent publication by Brinkmann et al. (1) on the comparison of various methodologies for diagnosing viral infections in high-throughput sequencing (HTS) data sets. The authors demonstrated that there is a plethora of workflows and pipelines available to analyze HTS data sets and the choice of technique can lead to different results, even with a uniform proficiency testing data set.

Processing HTS data sets is computationally intensive, may require significant investment, and often necessitates a comprehensive technical background to fully analyze the results. Currently, these requirements can limit the use of HTS, preventing clinicians and researchers with minimal funding or expertise in bioinformatics from exploring and exploiting this powerful technology.

However, several online tools, such as IDseq (2, 3) and Genome Detective (4), have recently been made available for research involving pathogen discovery and identification. The cloud-based nature of these tools removes the requirement for users to have high-specification computers for data processing, and automated identification of microbial sequences reduces the need for any significant background in bioinformatics. HTS data sets, with identifying information removed, are simply uploaded, and annotated sequence matches to potential pathogens are delivered within hours, in a format that can be easily interpreted by those with relevant clinical or academic skills. While IDseq automatically discards any human genomic reads, the submission of data sets containing patient sequences, although anonymized, to third-party platforms necessitates ethical consideration and permission.

We evaluated IDseq and Genome Detective against the simulated *in silico* data set provided by Brinkmann et al. (1). IDseq analysis took 92 min from the initiation of sample uploading to the presentation of the mapped reads, one-half of the time for the fastest participant (participant 1) reported by Brinkmann et al. (1). Of the 6,339,908 reads in the data set, 1,362,725 reads (21.5%) passed host filtering; of those, 996,855 reads (73.2%) mapped to bacterial nucleotide databases (70.3% to nonredundant protein databases). Genome Detective identified and removed 6,290,069 reads (99%) as nonviral hits, completing the analysis in only 16 min. Both platforms detected all four viruses in the data set (Table 1). Detection of Torque teno virus, human herpesvirus 1, and measles virus was not as sensitive as in many of the other participant workflows. However, both IDseq and Genome Detective identified the highly divergent avian orthobornavirus (55% similarity to a reference sequence), whereas 9 of the 13 workflows in the study by Brinkmann et al. (1) did not.

Citation Chappell JG, Byaruhanga T, Tsoleridis T, Ball JK, McClure CP. 2019. Identification of infectious agents in high-throughput sequencing data sets is easily achievable using free, cloud-based bioinformatics platforms. J Clin Microbiol 57:e01386-19. https://doi.org/10.1128/JCM.01386-19.

Editor Alexander J. McAdam, Boston Children's Hospital

Address correspondence to Joseph G. Chappell, joseph.chappell1@nottingham.ac.uk, or C. Patrick McClure, patrick.mcclure@nottingham.ac.uk.

The authors of the published article declined to respond.

Published 22 November 2019

TABLE 1 Comparison of viral reads identified by IDseq and Genome Detective and the COMPARE virus proficiency tests[a]

Method and database	Sensitivity (%)				Time (h)
	Torque teno virus	Human herpesvirus 1	Measles virus	Avian bornavirus	
Proficiency test, participants 1 to 13 (median [range])	100 (0–102)	99 (10–400)	100 (0–140)	0 (0–100)	15.5 (3–216)
IDseq					
Nucleotide	59	56	69	0	1.5
Nonredundant protein	59	55	70	53	
Genome Detective	100	84	82	41	0.25

[a]A sensitivity of >100% indicated false-positive results.

Our results show that both platforms can accurately identify viral genomes in HTS data sets, with little or no prior knowledge of bioinformatic approaches. IDseq has the additional capability to detect bacterial genomes as well as viral genomes. While not as sensitive as some of the other methodologies tested, IDseq and Genome Detective were able to identify all of the infectious agents included in the proficiency data set, in a fraction of the time reported for the other pipelines, and required very little local computational power. IDseq, Genome Detective, and similar free cloud-based online tools will significantly reduce the barrier to entry for exploiting HTS, without the hardware and background required for traditional bioinformatics approaches.

ACKNOWLEDGMENT

This work was supported by a Medical Research Council studentship award (grant 1651320).

REFERENCES

1. Brinkmann A, Andrusch A, Belka A, Wylezich C, Höper D, Pohlmann A, Nordahl Petersen T, Lucas P, Blanchard Y, Papa A, Melidou A, Oude Munnink BB, Matthijnssens J, Deboutte W, Ellis RJ, Hansmann F, Baumgärtner W, van der Vries E, Osterhaus A, Camma C, Mangone I, Lorusso A, Marcacci M, Nunes A, Pinto M, Borges V, Kroneman A, Schmitz D, Corman VM, Drosten C, Jones TC, Hendriksen RS, Aarestrup FM, Koopmans M, Beer M, Nitsche A. 2019. Proficiency testing of virus diagnostics based on bioinformatics analysis of simulated in silico high-throughput sequencing datasets. J Clin Microbiol 57:e00466-19. https://doi.org/10.1128/JCM.00466-19.
2. Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, Egger R, Dimitrov B, Juan Y-F, Sheu J, Wang J, Kalantar K, Langelier C, Ruel T, Mpimbaza A, Wilson MR, Rosenthal PJ, DeRisi JL. 2019. Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. PLoS One 14:e0218318. https://doi.org/10.1371/journal.pone.0218318.
3. Chan Zuckerberg Initiative. 2017. Infectious disease sequencing platform. https://github.com/chanzuckerberg/idseq-web.
4. Vilsker M, Moosa Y, Nooij S, Fonseca V, Ghysens Y, Dumon K, Pauwels R, Alcantara LC, Vanden Eynden E, Vandamme A-M, Deforche K, de Oliveira T. 2019. Genome Detective: an automated system for virus identification from high-throughput sequencing data. Bioinformatics 35:871–873. https://doi.org/10.1093/bioinformatics/bty695.

10.3. Appendix 3

viruses

MDPI

Communication

Retrieval of the Complete Coding Sequence of the UK-Endemic Tatenale Orthohantavirus Reveals Extensive Strain Variation and Supports Its Classification as a Novel Species

Joseph G. Chappell [1], Theocharis Tsoleridis [1], Okechukwu Onianwa [1], Gabby Drake [2], Ian Ashpole [2], Phillipa Dobbs [3], William Edema [1], Frederick Kumi-Ansah [1], Malcolm Bennett [4], Rachael E. Tarlinton [4], Jonathan K. Ball [1,*] and C. Patrick McClure [1]

1 School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK; joseph.chappell1@nottingham.ac.uk (J.G.C.); Patrick.Mcclure@nottingham.ac.uk (C.P.M.)
2 Chester Zoo, Chester, Cheshire CH2 1EU, UK
3 Twycross Zoo, Atherstone, Warwickshire CV9 3PX, UK
4 School of Veterinary Science, University of Nottingham, Sutton Bonnington, Loughborough LE12 5RD, UK
* Correspondence: mrzjb@exmail.nottingham.ac.uk; Tel.: +44-(0)-115-823-0745; Fax: +44-(0)-115-823-0759

Received: 26 February 2020; Accepted: 13 April 2020; Published: 17 April 2020

check for updates

Abstract: Orthohantaviruses are globally distributed viruses, associated with rodents and other small mammals. However, data on the circulation of orthohantaviruses within the UK, particularly the UK-endemic Tatenale virus, is sparse. In this study, 531 animals from five rodent species were collected from two locations in northern and central England and screened using a degenerate, pan-orthohantavirus RT-PCR assay. Tatenale virus was detected in a single field vole (*Microtus agrestis*) from central England and twelve field voles from northern England. Unbiased high-throughput sequencing of the central English strain resulted in the recovery of the complete coding sequence of a novel strain of Tatenale virus, whilst PCR-primer walking of the northern English strain recovered almost complete coding sequence of a previously identified strain. These findings represented the detection of a third lineage of Tatenale virus in the United Kingdom and extended the known geographic distribution of these viruses from northern to central England. Furthermore, the recovery of the complete coding sequence revealed that Tatenale virus was sufficiently related to the recently identified Traemersee virus, to meet the accepted criteria for classification as a single species of orthohantavirus.

Keywords: *Orthohantavirus*; hantavirus; high-throughput sequencing; virus discovery; field vole; United Kingdom

1. Introduction

Orthohantaviruses are a large and diverse genus of viruses, belonging to the *Hantaviridae* family within the order *Bunyavirales*. The genome of orthohantaviruses consists of a linear, negative-sensed and single-stranded RNA, divided into three segments. The large (L) segment encodes a single RNA-dependent RNA polymerase, the medium (M) segment encodes a glycoprotein precursor and the small (S) segment encodes a nucleocapsid protein [1]. Historically, orthohantaviruses have predominantly been associated with rodent reservoir species [2]; however, they have increasingly been detected in other mammalian taxa, such as bats [3], shrews [4] and moles [5]. Each species of orthohantavirus is typically associated with a single reservoir species, where the infection is considered to be persistent and asymptomatic [6].

Several orthohantavirus species are capable of transmission into humans, through the inhalation of aerosolised contaminated excreta [7]. Human infection is thought to result in two forms of the disease, depending on the causative species; old-world species are associated with a primarily renal syndrome known as 'haemorrhagic fever with renal syndrome' (HFRS), whilst new-world species are associated with pulmonary disease, 'hantavirus pulmonary syndrome' (HPS) [8]. However, an overlap of clinical presentations between the two syndromes has led to suggestions that they should be reconsidered as a single clinical syndrome, hantavirus fever (HF) [9]. The severity of HF can vary significantly; Puumala virus (PUUV) infection, for example, causes a mild, often sub-clinical disease [10], whilst new-world species, such as Sin Nombre virus (SNV), has a case fatality rate of 35% [11]. There are four species known to cause HF in Europe; Seoul (SEOV), Dobrava-Belgrade (DOBV), Tula (TULV) and Puumala (PUUV) [12]; the reservoirs associated with these viruses are the brown rat (*Rattus norvegicus*), yellow-necked mouse (*Apodemus flavicollis*)/striped field mouse (*Apodemus agrarius*), common vole (*Microtus arvalis*) and the bank vole (*Myodes glareolus*), respectively. Except for the TULV-associated common vole, which is geographically restricted to the Orkney Islands in Scotland, each of the reservoir species associated with these viruses are present in the United Kingdom (UK). However, of these viruses, only SEOV has been detected in the UK [13].

HF has been reported sporadically throughout the UK, including England [14], Scotland [15] and Northern Ireland [16], though the causative species could not be confirmed due to cross-reactivity of the serological assays used to diagnose the orthohantavirus infections [17]. The first orthohantavirus linked to HF in the UK was in 2011 when a novel strain of SEOV was isolated from wild rats captured on the farm of a patient with suspected HF [18]; SEOV was then detected in pet rats belonging to a patient with serologically confirmed HF in 2013 [19]. Furthermore, a novel vole-associated hantavirus related to TULV and PUUV—Tatenale virus (TATV)—was identified in field voles (*Microtus agrestis*) captured in northwest England in 2013 [20] and again in northern England in 2017 [21]. However, fragments of less than 400 nucleotides were retrieved for two of the three genomic segments, meaning that phylogenetic analysis of this virus was limited. In 2019, an orthohantavirus was detected in German field voles—Traemersee virus (TRAV)—and was suggested to be a strain of Tatenale virus. However, the aforementioned paucity of published TATV sequence data has precluded any accurate comparison between TATV and TRAV [22]. To better understand the prevalence and phylogeny of Tatenale virus, we performed in-depth sampling and analysis of various rodents living in the UK.

2. Materials and Methods

2.1. Samples

Rodents were caught at two semirural sites in the UK: Leicestershire (Site 1, 52.6524° N, 1.5291° W) and Cheshire (Site 2, 53.2273° N, 2.8844° W). Seventy-two rats (*Rattus norvegicus*), 224 mice (*Mus musculus*) and 12 field voles (*Microtus agrestis*) were collected from Site 1 between May 2013 and October 2014. Eight rats, 119 field voles, 93 wood mice (*Apodemus flavicollis*) and 3 bank voles (*Myodes glareolus*) were collected from Site 2 between June 2013 and July 2016.

Rodents were captured as part of routine pest-management at both sites. Ethical approval for collection of rodent tissue had been previously been granted [23] by the University of Nottingham School of Veterinary Science Ethical Panel, reference numbers 1602 151102 and 1786 160518.

2.2. Nucleic Acid Preparation

Sections of lung and kidney tissue, approximately 1 mm^3 were collected, and RNA was extracted using GenElute™ mammalian total RNA miniprep kit (Sigma Aldrich, St Louis, MO, USA), following the provided protocol. RNA was quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). cDNA was synthesised from the RNA using RevertAid reverse transcriptase (ThermoFisher Scientific) following the provided protocol.

2.3. RT-PCR Screening

Two-step RT-PCR was performed on the samples, using a degenerate primer pair (HanSemiF: GAATATATATCNTAYGGDGGDGA and HanSemiR: CTGGTGACCAYTTNGTNGCAT) designed in-house to target a 178 bp region of the L segment of all known hantaviruses. PCR reactions contained 0.5 µL of cDNA, added to 1.25 µL 10× PCR buffer, 0.06 µL of HotStarTaq DNA polymerase (QIAGEN, Hilden, Germany), 0.5 µL 10 mM dNTP's (Sigma Aldrich), 0.5 µL each of forward and reverse primer (10 Pmol/µL), and 9.19 µL of water for a total volume of 12.5 µL. Cycling conditions were 95 °C for 15 min, 55 cycles of 94 °C, 51 °C and 72 °C for 20 s each, followed by 72 °C for 10 min.

A second degenerate pan-hantavirus assay, targeting a different, larger region of the L segment, was used to confirm positive PCR results. Primers were sourced from Klempa et al. [24] (Han-L-F1: ATGTAYGTBAGTGCWGATGC and Han-L-R1: AACCADTCWGTYCCRTCATC); PCR reactions contained the same concentration of reagents and primers; cycling conditions were modified to 95 °C for 15 min, followed by 55 cycles of 94 °C (30 s), 53 °C (50 s) and 72 °C (30 s), finished with a final extension of 72 °C (10 min). This assay produced an amplicon of 452 bp.

2.4. High-Throughput Sequencing

The orthohantavirus positive field vole from Site 1 was selected for high-throughput sequencing (HTS). NEBNext® rRNA depletion kit (Human/Mouse/Rat) (New England Biolabs, Ipswich, MA, USA) with RNA sample purification beads (New England Biolabs, Ipswich, MA, USA) was used to deplete host ribosomal RNA from the sample. Sequencing libraries were then created from the depleted RNA using NEBNext® Ultra™ II directional RNA library prep kit for Illumina (New England Biolabs). Libraries were sent to SourceBioscience (Nottingham, UK) and sequenced with an Illumina HiSeq 4000. Each read length was 2 × 150 bp, and the insert size was 200 bp on average. All generated sequence data were analysed using the Geneious Prime 2019.0.4. Generated reads were mapped to reference orthohantavirus sequences downloaded from GenBank.

2.5. Retrieval of TATV's Complete Coding sequence (CDS) Using PCR Primer-Walking

The complete coding sequences (CDS) of the Tatenale virus retrieved via HTS was used as a reference sequence to design primers to retrieve the CDS of the TATV strain from Site 2, as funding was not available for HTS of both samples.

2.6. Phylogenetic Analysis

Nucleotide sequences for each segment of both TATV strains were aligned with a full-length coding sequence for representative Arvicolinae-associated orthohantaviruses and a non-arvicolinae orthohantavirus (Andes Virus) outgroup, using the MUSCLE function in MEGAX [25]. MEGAX was then used to find the best-fit substitution model for each alignment of sequences, and the model with the lowest Bayesian information criterion scores was considered the most appropriate.

Maximum likelihood trees were created with a GTR+G+I model, using MEGAX software. Robustness was assessed using bootstrap resampling (1000 replicates).

The pairwise evolutionary distance (PED) values of TATV and related orthohantaviruses were calculated using a WAG amino acid substitution model, in the PhyML [26] plugin on Geneious Prime. These calculations were based on a concatenation of the complete coding regions of S and M segments from the same virus.

3. Results

3.1. Detection of Orthohantavirus RNA by RT-PCR

Orthohantavirus RNA was detected in a single field vole from Site 1 (8.3%) and twelve field voles from Site 2 (10%). No orthohantavirus was detected in any house mouse (n = 224), wood mouse (N = 93), brown rat (n = 80) or bank vole (n = 3) samples.

The sequenced HAN-L amplicons from the Site 2 voles were highly conserved; the two most dissimilar were 97.7% identical at the nucleotide level. Comparison of the HAN-L amplicons between the two sites were more divergent, with a nucleotide homology of 86.3% to 89.3% between the single Site 1 virus and Site 2 viruses. We named the strain from Site 1 'Norton-Juxta' and the strain from Site 2 'Upton-Heath', reflecting the geographic origins of the two strains.

BLASTn searches of the 452 bp HAN-L PCR amplicons showed a high level of similarity to Tatenale orthohantaviruses. The Norton-Juxta strain was 87.3% identical to 'Tatenale virus strain B41' and 86.5% identical to 'Kielder hantavirus kld-1' at the nucleotide level, whilst the Upton-Heath Strains were 94.9–96.9% identical to B41 and 84–86.6% identical to kld-1.

3.2. Recovery of Complete TATV CDS

A total of 62,191,960 reads were sequenced from an uHTS library created from the lung tissue of the Norton-Juxta positive field vole. A total of 27,279,217 reads remained following pair-merging and quality processing. Mapping of these reads to reference sequences for each segment resulted in a total of 94,706 reads, representing 2.5% of filtered reads. The complete coding sequence of each segment was recovered. The L segment was 6465 nucleotides in length (Genbank Accession number MK883761), whilst the M segment was 3447 nucleotides (MK883759), and the CDS of S was 1302 nucleotides (MK883757).

Complete CDS of the L (MK883760) and S (MK883756) segments of TATV Upton-Heath was recovered through PCR primer-walking, these sequences were the same length as those for TATV Norton-Juxta. Almost complete CDS of the M segment (MK883758) was recovered, missing 90 nucleotides from the 3' end of the CDS.

3.3. Analysis of Complete TATV CDS

Comparison of the complete L and S segments and almost-complete M segments of the two strains revealed a nucleotide similarity of 90.6%, 94.1%, and 91.3%, respectively. Phylogenetic analysis of the three segments, with complete Arvicolinae-associated orthohantaviruses, showed that both Norton-Juxta and Upton-Heath TATV clustered closely with Traemersee virus, forming a distinct clade, and supported with strong bootstrap values in the L (Figure 1A), M (Figure 1B) and S segments (Figure 1C). Nucleotide and amino acid similarities between both TATV strains and closely related orthohantavirus species are shown in Table 1. Pairwise evolutionary distance (PED) analysis of the concatenated S and M segments of Norton-Juxta and other vole-borne orthohantaviruses showed values of between 0.12 and 0.27. The PED values between Norton-Juxta and TRAV were 0.05.

Comparison with the partial S sequence of TATV-B41 showed a nucleotide similarity of 98.7% with Upton-Heath and 93.9% with Norton-Juxta. Phylogenetic analysis with the partial L and S segments for other TATV strains showed that Norton-Juxta formed a novel lineage in the phylogenies of both segments, whilst Upton-Heath clustered closely with the previous B41 strain (Supplementary Figure S1).

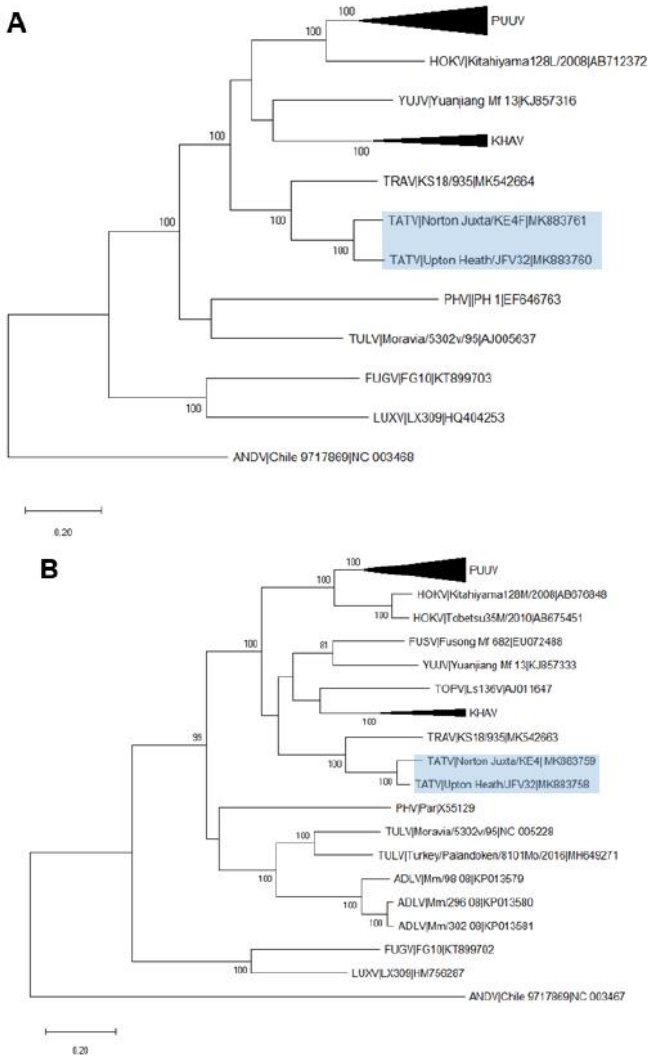

Figure 1. *Cont.*

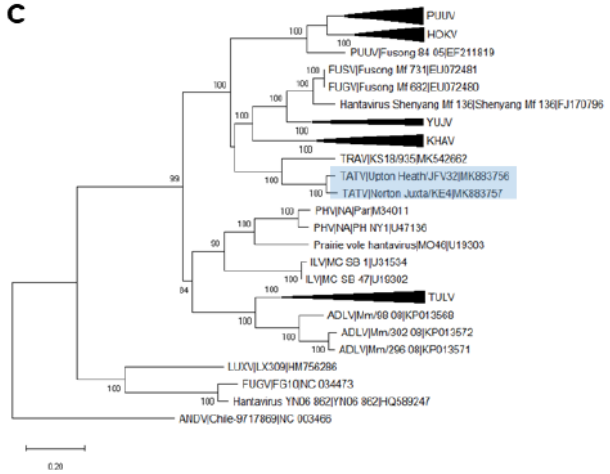

Figure 1. Phylogenetic relationship of Tatenale virus with other vole-associated orthohantavirus species. Representative complete coding sequences were retrieved for each segment; L (**A**), M (**B**) and S (**C**). Maximum likelihood trees were created with a GTR+G+I model, using MEGAX software. Branch lengths were drawn to a scale of nucleotide substitutions per site. L and S trees were based on full-length sequences, whilst the M segment tree was based on the available sequence for the partial Upton-Heath strain. Numbers above individual branches show bootstrap support after 1000 replicates. Tatenale virus strains are highlighted with a blue box. Black triangles represent a compressed species-specific subtree. Sequences are shown with the species name, strain name and the GenBank accession number. PUUV, Puumala virus; HOKV, Hokkaido virus; FUSV, Fusong virus; YUJV, Yuanjiang virus; KHAV, Khabarovsk virus; TOPV, Topografov virus; TATV, Tatenale virus; TRAV, Traemmersee virus; PHV, Prospect Hill virus; ILV, Isla Vista virus; TULV, Tula virus; ADLV, Adler virus; LUXV, Luxi virus; FUGV, Fugong virus; ANDV, Andes virus.

Table 1. The similarity of Norton-Juxta and Upton-Heath strains of Tatenale virus to the closest related strain of the most related species at nucleotide (amino acid) level. Similarities to the M segment of the Upton-Heath strain are based on the available partial sequence. * Indicates no complete sequence data available.

Species (Accession Number)	*S*	*M*	*L*
		Norton-Juxta	
Traemersee	82.7 (96.8)	79.8 (94.2)	81.5 (96.4)
Khabarovsk	79.2 (89.4)	76.4 (87.5)	77.9 (90.9)
Yuanjiang	79.2 (88.5)	75.3 (86.5)	77.7 (90.4)
Fusong	78.7 (88.2)	75.2 (85.9)	-*
Puumala	77.9 (87.8)	74.8 (84.7)	77.9 (88.1)
Hokkaido	78.3 (87.5)	75.5 (84.4)	76.8 (88.5)
		Upton-Heath	
Traemersee	83 (96.5)	80.8 (94.3)	81.5 (96.4)
Khabarovsk	79.9 (88.9)	77.1 (87.8)	78 (90.7)
Yuanjiang	78.9 (88.2)	75.7 (86.5)	77.7 (89.6)
Fusong	78.9 (88)	76 (86.2)	-*
Puumala	78.4 (87.8)	75.5 (84.6)	77.6 (87.5)
Hokkaido	79 (87.8)	75.7 (84.4)	76.7 (87.9)

4. Discussion

This was the first reported recovery of complete coding sequences for TATV in the UK. Based on a genetic divergence, we proposed that this virus represented an additional strain of TATV, tentatively called Norton-Juxta, which extended the known range of TATV from northern to central England. The detection of diverse TATV in field voles, but not other species of rodents sampled from the same sites, strengthened evidence of field voles as the primary reservoir of the virus. The high similarity between the available sequence data for TATV B41 and the corresponding sequence from TATV Upton-Heath, together with the close geographic proximity of the collection sites, indicated that the two viruses might be co-circulating within the same population of field voles.

Orthohantavirus species' demarcation criteria of >7% AA divergence across S and M segments [27], as well as stricter criteria of a PED lower than 0.1 in the concatenated S and M segments [28], have been suggested. As the PED values between the complete Norton-Juxta strain of TATV and TRAV was below the 0.1 speciation threshold, this confirmed that both TATV and TRAV were members of the same viral species, as was hypothesised by Jeske et al. [22]. Though TRAV was the first strain with complete CDS, TATV was detected several years prior and is more established in the literature. Thus, we proposed that the species in which TATV and TRAV belonged to be named Tatenale orthohantavirus.

There is serological evidence of human infection with PUUV- or SNV-like viruses in the UK [29], though there has been no molecular evidence of these viruses in either humans or rodents. A previous study has reported that blood from a vole infected with TATV B41 is cross-reactive with PUUV, which suggests that PUUV/SNV seropositive humans may have been the result of TATV infection [20]. Until now, the paucity of sequence data has precluded significant further investigation of TATV. Recovery of the complete coding sequence for each of the segments, particularly the glycoproteins encoded in the M segment, will allow for in vitro studies to further explore the zoonotic potential of the virus.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1999-4915/12/4/454/s1.

Author Contributions: J.G.C., C.P.M., J.K.B. and R.E.T. conceived and designed experiments. J.G.C., O.O., W.E., T.T. and F.K.-A. performed experiments. G.D., I.A., P.D. and M.B. provided samples. J.G.C. and T.T. analysed HTS data. J.G.C. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Medical Research Council studentship award, grant number 1651320.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Plyusnin, A.; Vapalahti, O.; Vaheri, A. Hantaviruses: Genome structure, expression and evolution. *J. Gen. Virol.* **1996**, *77*, 2677–2687. [CrossRef] [PubMed]
2. Jonsson, C.B.; Figueiredo, L.T.M.; Vapalahti, O. A global perspective on hantavirus ecology, epidemiology, and disease. *Clin. Microbiol. Rev.* **2010**, *23*, 412–441. [CrossRef] [PubMed]
3. Weiss, S.; Witkowski, P.T.; Auste, B.; Nowak, K.; Weber, N.; Fahr, J. Hantavirus in Bat, Sierra Leone. *Emerg. Infect. Dis.* **2012**, *18*, 159–161. [CrossRef] [PubMed]
4. Klempa, B.; Fichet-Calvet, E.; Lecompte, E.; Auste, B.; Aniskin, V.; Meisel, H.; Barrière, P.; Koivogui, L.; ter Meulen, J.; Krüger, D.H. Novel hantavirus sequences in Shrew, Guinea. *Emerg. Infect. Dis.* **2007**, *13*, 520–522. [CrossRef]
5. Laenen, L.; Vergote, V.; Kafetzopoulou, L.E.; Wawina, T.B.; Vassou, D.; Cook, J.A.; Hugot, J.-P.; Deboutte, W.; Kang, H.J.; Witkowski, P.T.; et al. A Novel Hantavirus of the European Mole, Bruges Virus, Is Involved in Frequent Nova Virus Coinfections. *Genome Biol. Evol.* **2018**, *10*, 45–55. [CrossRef]
6. Ermonval, M.; Baychelier, F.; Tordo, N. What Do We Know about How Hantaviruses Interact with Their Different Hosts? *Viruses* **2016**, *8*, 223. [CrossRef]
7. Vaheri, A.; Strandin, T.; Hepojoki, J.; Sironen, T.; Henttonen, H.; Mäkelä, S.; Mustonen, J. Uncovering the mysteries of hantavirus infections. *Nat. Rev. Microbiol.* **2013**, *11*, 539–550. [CrossRef]

8. Macneil, A.; Nichol, S.T.; Spiropoulou, C.F. Hantavirus pulmonary syndrome. *Virus Res.* **2011**, *162*, 138–147. [CrossRef]
9. Clement, J.; Maes, P.; Van Ranst, M. Hemorrhagic Fever with Renal Syndrome in the New, and Hantavirus Pulmonary Syndrome in the old world: Paradi(se)gm lost or regained? *Virus Res.* **2014**, *187*, 55–58. [CrossRef]
10. Hjertqvist, M.; Klein, S.L.; Ahlm, C.; Klingstrom, J. Mortality rate patterns for hemorrhagic fever with renal syndrome caused by Puumala virus. *Emerg. Infect. Dis.* **2010**, *16*, 1584–1586. [CrossRef]
11. MacNeil, A.; Ksiazek, T.G.; Rollin, P.E. Hantavirus pulmonary syndrome, United States, 1993–2009. *Emerg. Infect. Dis.* **2011**, *17*, 1195–1201. [CrossRef] [PubMed]
12. Heyman, P.; Ceianu, C.; Christova, I.; Tordo, N.; Beersma, M.; Joao Alves, M.; Lundkvist, A.; Hukic, M.; Papa, A.; Tenorio, A.; et al. A five-year perspective on the situation of haemorrhagic fever with renal syndrome and status of the hantavirus reservoirs in Europe, 2005–2010. *Eurosurveillance* **2011**, *16*. [CrossRef]
13. Murphy, E.G.; Williams, N.J.; Bennett, M.; Jennings, D.; Chantrey, J.; McElhinney, L.M. Detection of Seoul virus in wild brown rats (Rattus norvegicus) from pig farms in Northern England. *Vet. Rec.* **2019**, *184*, 525. [CrossRef] [PubMed]
14. Watson, A.R.; Irving, W.L.; Ansell, I.D. Playing in a scrapyard and acute renal failure. *Lancet* **1997**, *349*, 1446. [CrossRef]
15. Walker, E.; Boyd, A.J.; Kudesia, G.; Pinkerton, I.W. A Scottish case of nephropathy due to Hantaan virus infection. *J. Infect.* **1985**, *11*, 57–58. [CrossRef]
16. McKenna, P.; Clement, J.; Matthys, P.; Coyle, P.V.; McCaughey, C. Serological evidence of hantavirus disease in Northern Ireland. *J. Med. Virol.* **1994**, *43*, 33–38. [CrossRef] [PubMed]
17. Brus Sjölander, K.; Lundkvist, Å. Dobrava virus infection: Serological diagnosis and cross-reactions to other hantaviruses. *J. Virol. Methods* **1999**, *80*, 137–143. [CrossRef]
18. Jameson, L.J.; Logue, C.H.; Atkinson, B.; Baker, N.; Galbraith, S.E.; Carroll, M.W.; Brooks, T.; Hewson, R. The continued emergence of hantaviruses: Isolation of a Seoul virus implicated in human disease, United Kingdom, October 2012. *Eurosurveillance* **2013**, *18*, 20344.
19. Jameson, L.J.; Taori, S.K.; Atkinson, B.; Levick, P.; Featherstone, C.A.; van der Burgt, G.; McCarthy, N.; Hart, J.; Osborne, J.C.; Walsh, A.L.; et al. Pet rats as a source of hantavirus in England and Wales, 2013. *Eurosurveillance* **2013**, *18*, 20415.
20. Pounder, K.C.; Begon, M.; Sironen, T.; Henttonen, H.; Watts, P.C.; Voutilainen, L.; Vapalahti, O.; Klempa, B.; Fooks, A.R.; McElhinney, L.M. Novel Hantavirus in Field Vole, United Kingdom. Available online: http://wwwnc.cdc.gov/eid/article/19/4/12-1057_article (accessed on 22 January 2016).
21. Thomason, A.G.; Begon, M.; Bradley, J.E.; Paterson, S.; Jackson, J.A. Endemic Hantavirus in Field Voles, Northern England. *Emerg. Infect. Dis.* **2017**, *23*, 1033–1035. [CrossRef]
22. Jeske, K.; Hiltbrunner, M.; Drewes, S.; Ryll, R.; Wenk, M.; Špakova, A.; Petraitytė-Burneikienė, R.; Heckel, G.; Ulrich, R.G. Field vole-associated Traemmersee hantavirus from Germany represents a novel hantavirus species. *Virus Genes* **2019**, *55*, 1–6. [CrossRef] [PubMed]
23. Tsoleridis, T.; Chappell, J.G.; Monchatre-Leroy, E.; Umhang, G.; Shi, M.; Bennett, M.; Tarlinton, R.E.; McClure, C.P.; Holmes, E.C.; Ball, J.K. Discovery and prevalence of divergent RNA viruses in European field voles and rabbits. *Viruses* **2019**, *12*, 47. [CrossRef] [PubMed]
24. Klempa, B.; Fichet-Calvet, E.; Lecompte, E.; Auste, B.; Aniskin, V.; Meisel, H.; Denys, C.; Koivogui, L.; ter Meulen, J.; Krüger, D.H. Hantavirus in African wood mouse, Guinea. *Emerg. Infect. Dis.* **2006**, *12*, 838–840. [CrossRef] [PubMed]
25. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef] [PubMed]
26. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef]
27. ICTV ICTV Ninth Report; Taxonomy Release. 2009. Available online: https://talk.ictvonline.org/ictv-reports/ictv_9th_report/negative-sense-rna-viruses-2011/w/negrna_viruses/205/bunyaviridae (accessed on 7 March 2019).

28. Maes, P.; Klempa, B.; Clement, J.; Matthijnssens, J.; Gajdusek, D.C.; Krüger, D.H.; Van Ranst, M. A proposal for new criteria for the classification of hantaviruses, based on S and M segment protein sequences. *Infect. Genet. Evol.* **2009**, *9*, 813–820. [CrossRef]
29. Jameson, L.J.; Newton, A.; Coole, L.; Newman, E.N.C.; Carroll, M.W.; Beeching, N.J.; Hewson, R.; Christley, R.M. Prevalence of antibodies against hantaviruses in serum and saliva of adults living or working on farms in Yorkshire, United Kingdom. *Viruses* **2014**, *6*, 524–534. [CrossRef]