# Chapter 1: What does pathology have to do with disease ecology? Linking pathogenesis to viral spillover

## Introduction

Interspecies pathogen transmission is a multifactorial process with both extrinsic drivers, and intrinsic pathogen and host factors playing major roles in the process (Plowright et al. 2017, Becker et al. 2019, Ellwanger and Chies 2021). Viral traits, such as having an envelope (Valero-Rello and Sanjuán 2022) and being able to replicate in the host cytoplasm (Pulliam and Dushoff 2009) likely play important roles in cross species transmission.. Recent work has looked at the genotypes of viruses as a predictor of zoonotic potential(Mollentze et al. 2021). Host factors often studied have been phylogenetic distance between hosts (Guth et al. 2019)and phylogenetic aggregation(Park 2019). Phylogeny is likely to capture a broad range of host biological traits that facilitate cross species transmission. Both host and viral factors have been included in predictive models of viral spillover (Olival et al. 2017). Although these host and viral factors are important spillover drivers, they must be viewed through the lens of ecological viability of a spillover event happening and these extrinsic drivers are just as important (Engering et al. 2013, Borremans et al. 2019). While there has been great attention paid to investigating specific pathogen and host traits involved in cross species transmission, there has been little research on the interplay between host and pathogen in disease spillover models (Figure 1). Recent advances in metagenomics has allowed for the discovery of large numbers of viruses in animals (Cui et al. 2023). Whilst these are often viewed as potential risks for cross species transmission. Understanding how pathogenesis relates to the viability of these viruses becoming a threat is an important piece of information that can help further direct research into these newly discovered pathogens. We believe there is a knowledge gap when it comes to the role of disease pathogenesis in cross species transmission between wildlife, domestic animals, and humans.

The potential role of infectious disease pathogenesis, defined as the manner in which a pathogen infects, replicates within a host, and is transmitted between hosts in predictive models of cross-species disease transmission was raised by Pulliam et al, but it has proven to be a difficult question to answer(Pulliam and Dushoff 2009). For instance, comparative analysis of viral-host receptor distribution may provide insights to the probability of inter-species viral transmission. The distribution of the ACE2 receptor in humans is postulated to lead to the multisystem pathology of SARS-CoV-2, but there is a lack of detailed studies in other species (Ruiz-Aravena et al. 2022). The outcome of infections can be drastically different depending on the anatomy of receptor distribution, for example, the ACE2 upper respiratory distribution in humans compared to DPP4 lower respiratory may partly explain differing transmissibility between SARS and MERS in humans (Widagdo 2016, Hou et al. 2020). In addition to distribution of receptors, virulence mismatches in hosts can shape the outcomes of cross-species transmission, with increased virulence potentially acting as a limiting factor preventing onward transmission(Mollentze et al. 2020). Pathogenesis therefore may be playing a role in the outcome of cross species transmission events. Major factors influencing pathogenesis include the cellular receptor used by the virus for *attachment* to host cells and the distribution of those receptors within different tissues and organs(Norkin 2010). The cellular-level mechanisms of pathogenesis are known in well-studied viral systems such as influenza (Bender and Small Jr 1992, Zambon 1999, 2001, Fukuyama and Kawaoka 2011). However, much of that detailed knowledge is likely to be absent in many cases, therefore cell tropism and organ systems affected can serve as a proxy to study how infectious disease pathogenesis is related to cross species transmission. Additionally, the method of within-host spread of the pathogen may play an important role in determining whether disease spread can occur in the ‘recipient’ host species.

Ecological Drivers

Pathogenesis

Viral

Traits

Host

Traits

Figure 1: representation of the factors involved in cross species disease transmission.

The important steps in a viral lifecycle within a host are :

1. Primary transmission
2. **Entry and local replication**
3. **Dissemination within host**
4. **Secondary replication**
5. Shedding/secondary transmission

All these steps can shape pathogenesis in some way. However, the major stages that play a role in pathogenesis are the within-host stages of **entry and local replication, dissemination within the host, and secondary replication**. Pathogenesis therefore plays a key role both in the link from primary transmission to infection and to that of secondary transmission. This is also the case when it comes to cross species transmission or spillover of viral pathogens. Knowledge of viral-host interactions can infer if transmission to a new host will be productive (Escudero-Pérez et al. 2023). Similarly, dissemination within a host and organ tropism can predict shedding and likely secondary transmission (Louten 2016). Clearly the pathogenesis of a viral infection could play a critical role in the disease ecology and transmission dynamics of a pathogen and determine cross species transmission potential.

### Viral Entry and local replication

The initial “decision” for a virus upon encountering a new potential host is that of receptor selection and attachment. Contacting a compatible receptor allows the virus to attach to and gain entry to the host cell and consequently access the cellular machinery it requires for its replication cycle. Viruses that can infect multiple hosts tend to use evolutionarily conserved receptors(Woolhouse et al. 2012). A good example of this would be rabies virus, which uses the nACh-receptor which is very conserved across all mammals, and rabies has been shown to possess the ability to infect all eutherian mammals(Marston et al. 2018). Another option is to have multiple potential receptors that can be used for viral attachment to a host cell. Foot-and-mouth disease virus has at least three potential integrin receptors along with Fc receptors it can use, giving it the ability to infect most known cloven-hoofed mammals (Duque and Baxt 2003). This ability to attach to multiple receptors increases the likelihood of encountering a suitable receptor when the virus contacts a novel host.

### Dissemination within host

With regards to viral dissemination within a host and the potential role of this in cross species transmission to other hosts, canine distemper can serve as an illustrative example of the importance of within host dissemination and cross species transmission. Canine distemper virus is considered a multi-cell and multi-host pathogen that can infect three different types of host cells; epithelial, lymphoid, and neurological cells (Rendon-Marin et al. 2019). During the first stages of infection within the host, resident dendritic cells and alveolar macrophages in the respiratory tract are infected along with other cells which express CD150 in the alveolae. Infected cells carry the virus to the draining lymph node where resident activated T-cells and B-cells are infected through the CD150 receptor, resulting in virus amplification and the initiation of primary viremia (Rendon-Marin et al. 2019). The virus then disseminates to secondary lymphoid organs and is spread through the entire immune system, and finally disseminates to brain, liver, skin, gastrointestinal tract, genitals, and respiratory mucosal surfaces. This systemic spread, particularly to the respiratory mucosa, results in the potential for rapid transmission dynamics through close contact populations. This strategy of using the lymphatic system to spread systemically is almost certainly involved in this virus’ ability to infect a wide range of species (Beineke et al. 2015).

### Secondary Replication

In addition to receptor selection, a large part of cellular pathogenesis and within-host infection dynamics are controlled by the distribution of the receptor used by the virus within or across organ systems that allow secondary replication at distant sites. The most studied system of receptor distribution across different host species is that of influenza virus. Avian and human influenza strains preferentially use differently terminated sialic acid receptors with SAα2,3Gal (avian receptor) and SAα2,6Gal (mammalian receptor) terminated saccharides distributed in the upper respiratory tract of birds and humans, respectively(Kumlin et al. 2008). However, both types of receptor are also present in the upper respiratory tract of pigs giving rise to the pigs as a “mixing vessel” theory of flu recombination and evolution (Ma et al. 2008). The pig upper respiratory tract expresses both of these receptors, allowing for coinfection with an avian and human strain and recombination to produce a highly pathogenic strain that is more transmissible in people than an avian strain, such as was the case in the 2009 swine flu epidemic (Vijaykrishna et al. 2010).Tissue tropism of avian Influenza has been shown to influence spillover from wild birds to pigs (Zhang et al. 2020). Additionally, distribution of receptors in human spillover hosts is important to cross-species influenza transmission Humans possess SAα2,3Gal receptors, but only in their lower respiratory tract (de Graaf and Fouchier 2014). So, while people do occasionally become infected with avian influenza, this subtle difference in receptor distribution plays a huge role, both in disease pathogenesis and in transmissibility of infection. The presence of the virus and subsequent replication in the lower respiratory tract results in a much more severe infection with higher morbidity and mortality than a typical human strain of flu. Additionally, the fact that the virus cannot replicate in the upper respiratory tract makes it difficult for the virus to be transmitted via respiratory aerosol.

Waterfowl are considered as the primary wild reservoir of avian influenza strains and their role in this partly comes down to receptor distribution. The SAα2,3Gal receptors used by the influenza virus are present in large amounts in the intestinal tract of many species of migratory waterfowl (Costa et al. 2012). Waterfowl belonging to the *Anatidae* family (ducks, geese, and swans) are the primary reservoir of all 16 hemagglutinin and 9 neuraminidase subtypes of avian influenza viruses (Hansbro et al. 2010). Migration by these birds results in the inoculation of waterways with live influenza virus which is relatively stable in the water (Blagodatski et al. 2021). Additionally, there is also the potential for free-ranging domestic fowl to be exposed to faeces containing avian influenza from these birds. This illustrates how pathological features of infection and ecology can interact to directly influence viral transmission dynamics.

These are some well-studied examples of where infection pathogenesis plays a significant role in the ability to infect a novel host. The question arises if similar examples exist in the broader literature of pathological traits influencing a virus’s ability to infect novel species and whether general patterns exist that can at least partially explain the role that host-pathogen interaction or ‘pathogenesis’ traits may play in cross species transmission, this is a seldom studied area of cross species transmission and there are likely significant gaps in the available literature.

Here, we synthesize available literature on cross-species transmission events along with viral pathogenesis data to determine how aspects of viral pathogenesis affect cross-species transmission in RNA viruses.

The objectives of this study are two-fold; firstly, we investigate the availability and quality of existing virus pathogenesis data as it currently stands and how it is suited to answering questions related to pathogenesis and cross species transmission and where gaps in this knowledge may lie. Secondly we aimed to analyze the available host-virus pathogenesis data and identify pathological traits associated with different types of cross species transmission events, in particular whether these events involve humans as a host or spillover species.

## Methods

### Data acquisition and overview

A search of the PubMed database was performed as described below. These search results were then screened using the *metagear* package in R according to the PRISMA guidelines fig 3.**.** Following this a secondary search was conducted to provide known pathological data for each virus.

The resulting data were imported into R Studio (version 2022.12.0+353) (Posit team 2022).. A detailed description of data analysis is contained in the scripts within the project repository (https://github.com/JJWilson1991/Pathogenesis\_project). All analyses were conducted in the R programming environment (version 4.2.0.)(R Core Team 2022). References to packages in this methods section indicate specific packages used within the R environment to perform analyses.

### Literature search

A literature search was performed of the PubMed data base according to Prisma guidelines(Page et al. 2021). The search terms described in figure 2 were used with the results being screened for eligibility using the Metagear package in R(Lajeunesse 2016). Eligibility criteria described in table 1.

**Figure 2: List of search terms used in PubMed search for gathering initial list of literature.**

(((RNA virus) AND (((((Spillover) OR (cross-species transmission)) OR (host-switching)) OR (interspecies transmission)) OR (zoonosis))) NOT (Covid)) NOT (Review[Publication Type])

Following inclusion in the study, data relating to the cross-species transmission event were recorded from each article, listed in table 2. The pathological traits recorded from the secondary search are also listed in table 2.events were defined as zoonotic, anthroponotic or not zoonotic. “Not zoonotic” being defined as a virus which is not considered to infect humans. An “anthroponotic” event is defined as a virus which has been transmitted from a human reservoir to another species. Spillover is defined as  the transmission of a parasite from one host species to another, regardless of whether onwards transmission in the recipient host is successful(Borremans et al. 2019). Reservoir host is defined as a species responsible for shedding the parasite and causing a spillover exposure event, either by shedding the parasite into the environment or through direct contact with the recipient host(Borremans et al. 2019). Spillover host is defined as a species that is infected by a parasite originating from a different host species (Borremans et al. 2019).

Table 2: List of data recorded from each qualifying piece of literature on a cross species transmission event and the corresponding pathological data recorded from reference material.

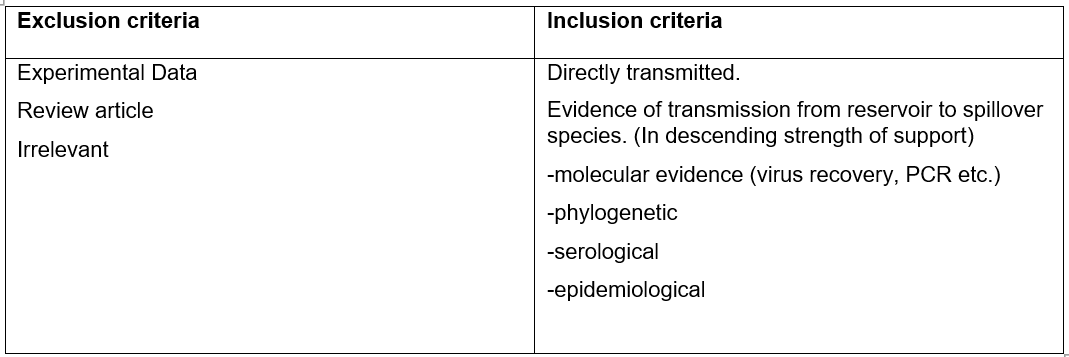
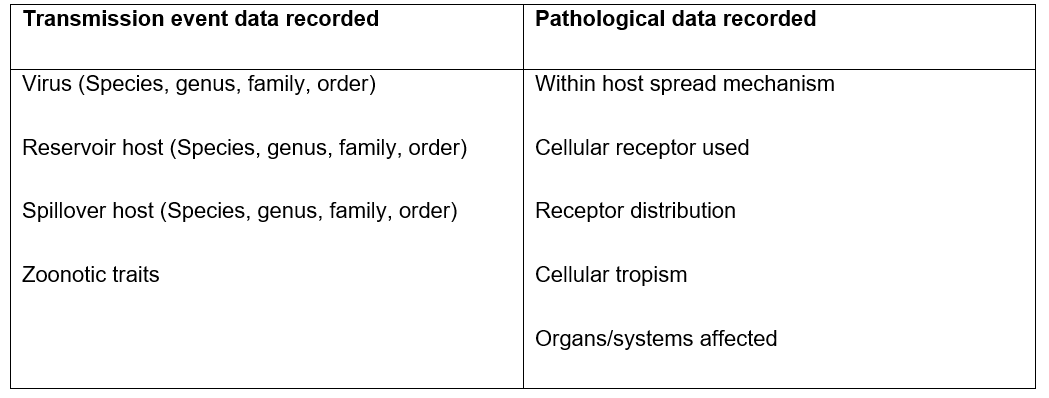


Table 1: list of Exclusion and inclusion criteria for literature search screening

### Multiple correspondence Analysis

Records identified from\*:

Databases (n = 1) (NCBI)

Records removed *before screening*:

Duplicate records removed (n =0)

Records removed for other reasons (n =0)

Records screened.

(n =7258)

Records excluded\*\*

(n = 6578)

Reports sought for retrieval.

(n =680)

Reports not retrieved.

(n =0)

Reports assessed for eligibility.

(n =680)

Reports excluded:

N=491

Studies included in review.

(n = 189)

**Identification of studies via databases and registers**

**Identification**

**Screening**

**Included**

Figure 3: Flow diagram for screening and reviewing search results for this study based on PRISMA guidelines.

To analyze the relationship between pathological parameters, multiple correspondence analysis (MCA) was conducted. MCA is a descriptive technique designed to measure correspondence between the rows and columns in tables of data. The objective of MCA is to visualize the relationship of categorical variables. Correspondence analysis is used to explore the relationship between variables by comparison with distance in multiple dimension space. The first two dimensions can usually explain most of the variation seen in the data. Following data processing a total of n=213 rows along with 31 columns were uses for MCA. The MCA was also repeated with just the unique viruses (n=52), to remove bias in the correspondence caused by viruses which occur often in the literature in spillover events between multiple different species e.g., influenza. Multiple correspondence analysis was conducted on the data using the package *FactoMineR* (Le 2008)*.*

### Hierarchical cluster analysis

We used hierarchical cluster analysis to group viruses into pathologically similar clusters. Specifically, we constructed a dissimilarity matrix using the Gower distance (Gower 1971) with the `daisy` function from the `cluster` package (Maechler 2022). This matrix was then used with the ‘hclust’ function in the ‘stats’ package(R Core Team 2022) to perform agglomerative clustering on this distance matrix. We next visualized the results of our hierarchical cluster analysis as a dendrogram with the ‘dendextend’ and `factoextra` packages(Galili 2015, Kassambara and Mundt 2017).

## Results

|  |  |
| --- | --- |
| **Pathology variable** | **Unknown data proportion** |
| Cellular Receptor | 14/52 (27%) |
| Within-host spread mechanism | 14/52 (27%) |
| Cell tropism | 5/52 (10%) |
| Organ systems affected | 3/52 (6%) |

A total of 189 articles were entered into the database and following processing, this resulted in n=213 unique entries of virus-reservoir host-spillover host interactions. This was comprised of n=52 unique viruses. The unknown data for pathology-related variables for each unique virus are summarized in table 3.

Table 3: following searches of databases and reference material described in methods, the proportions of viruses with missing data for their cellular receptors, spread mechanisms, cellular tropism and organ system affected are described.

The most frequent reservoir host was humans (34/213) followed by pigs and dogs (supplementary). Spillover hosts was dominated by humans with 103/213 recorded. The nature of interactions between reservoir hosts and spillover hosts is represented in figure 3. From the 52 viruses recorded there were 33 different cellular receptors plus a further 14 unknown receptors.

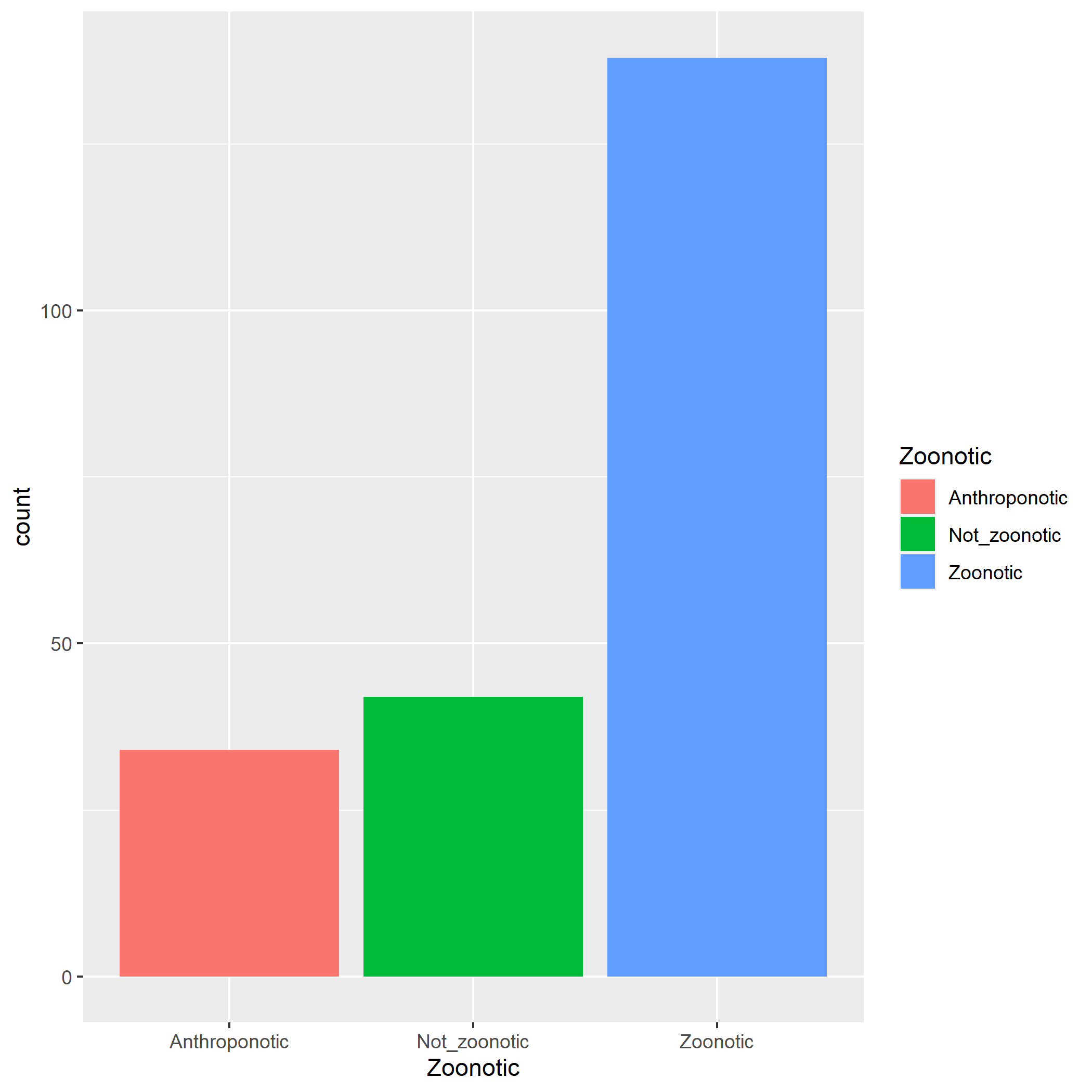


Fig 4: Proportion of each individual cross species transmission event that was recorded as being anthroponotic, zoonotic or neither.

### Multiple correspondence analysis

The first two dimensions of the MCA explain 35% of the variance in the data. The major variables which contribute to the first two dimensions are shown in supplementary figure 1. The MCAfor individual host interactions (Figure 5) clusters epithelial spread and respiratory pathology with anthroponosis, suggesting an association between these variables. There also appears to be an association between lymphocytic tropism and not being zoonotic. 95% confidence ellipses for “anthroponotic” cluster separately from “not zoonotic”, suggesting pathological features differ (Figure 5).

The MCA was repeated with the unique viruses in the dataset (Figure 5). Unique viruses being a viral species which has been involved in a spillover event, so that there is one entry per virus. Whereas in the original search and analysis many viruses had multiple entries as they were involved in multiple spillover events involving different species. In this case gastrointestinal pathology is associated with “not zoonotic” in addition to lymphocyte tropism. The “zoonotic” ellipses for the unique virus MCA have the “zoonotic” ellipse as a smaller ellipse within a larger ellipse for “not zoonotic” (Figure 5). This suggests that the zoonotic interactions have a narrower range of pathologies compared to the not zoonotic.

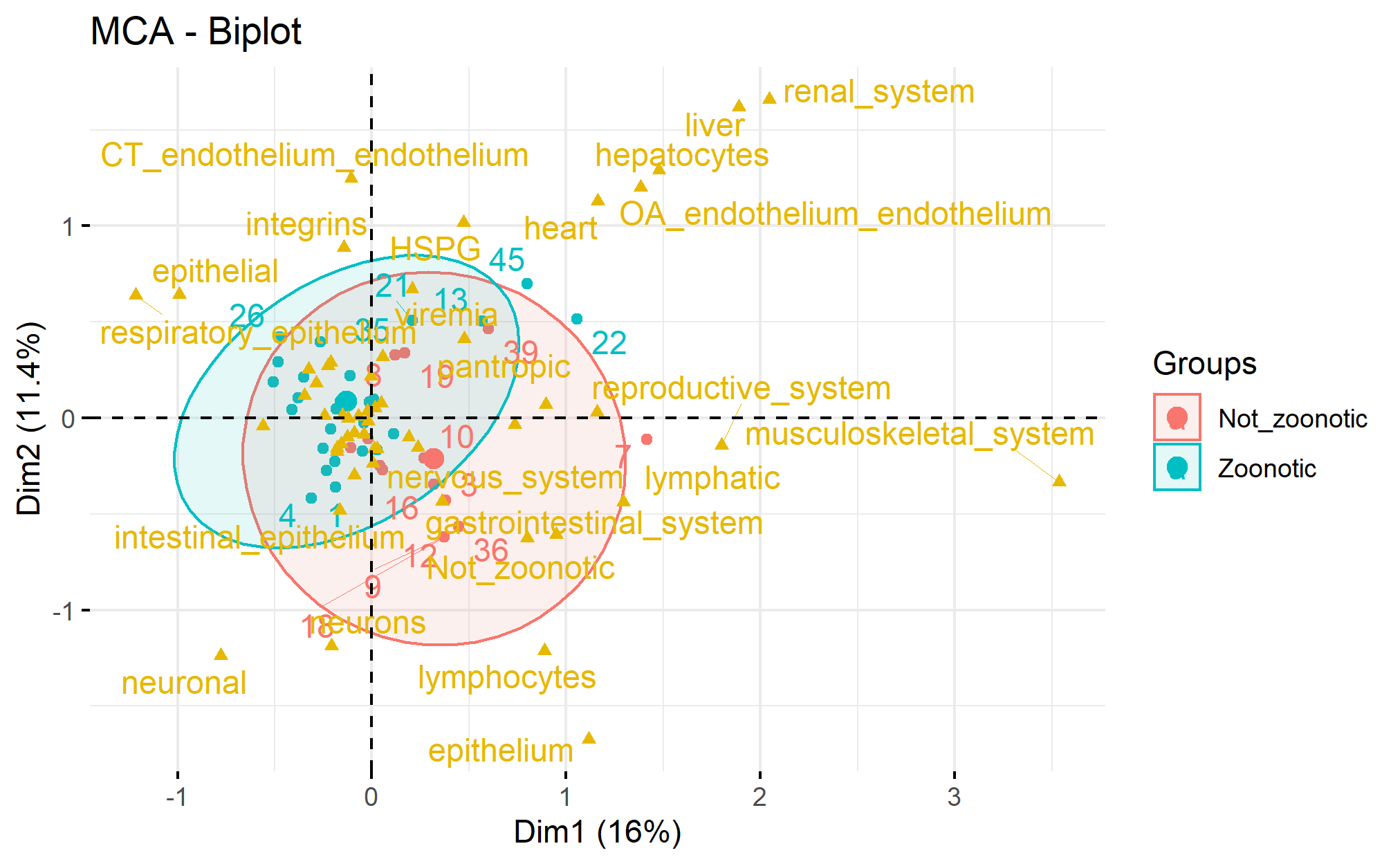
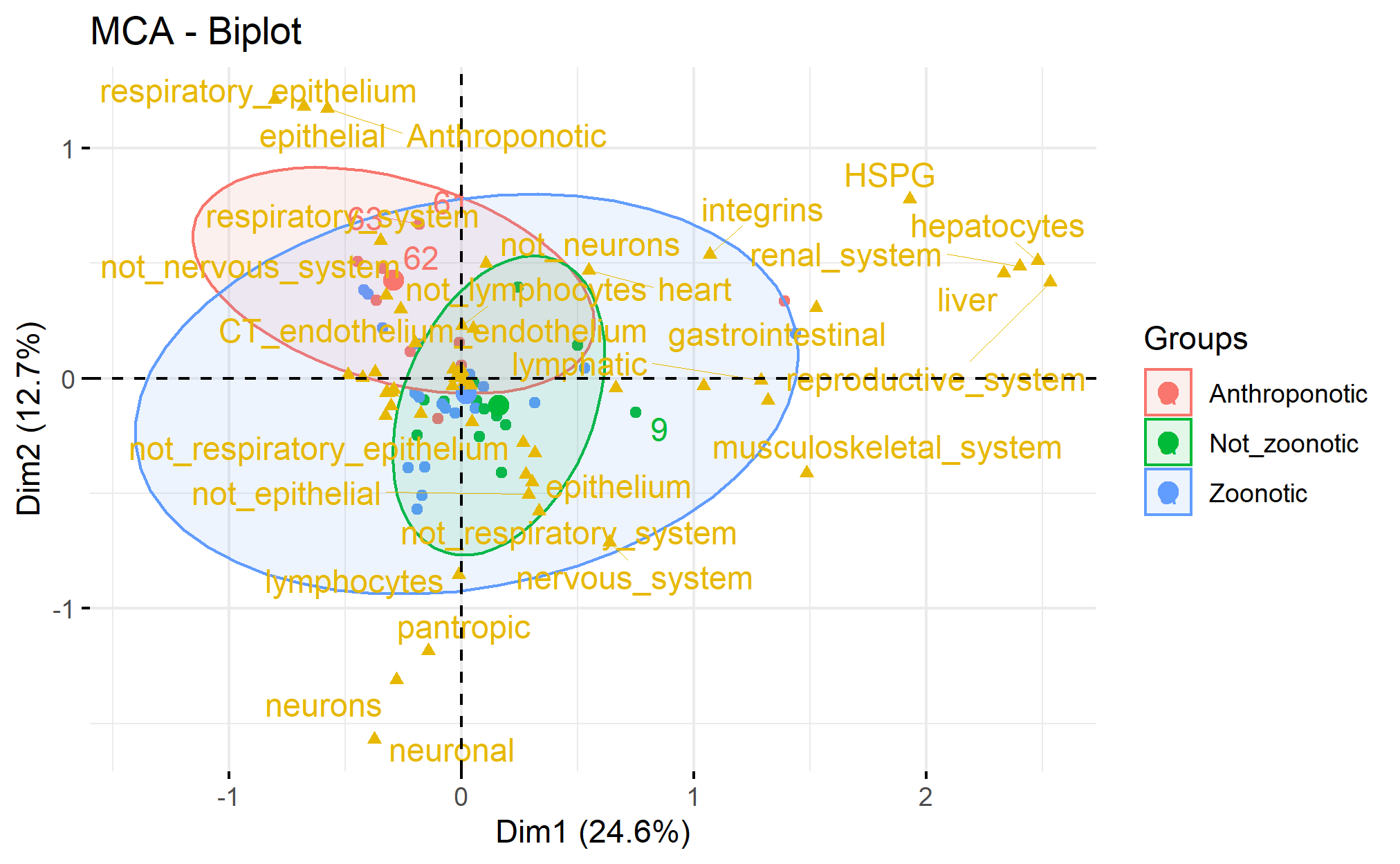


Fig 6: multiple correspondence analysis biplot for the n=52 unique RNA viruses from the literature search. The yellow points represent the variables. 95% confidence ellipses for cross species transmission events classified as zoonotic, or not zoonotic have been drawn over the individual observations in the plot.

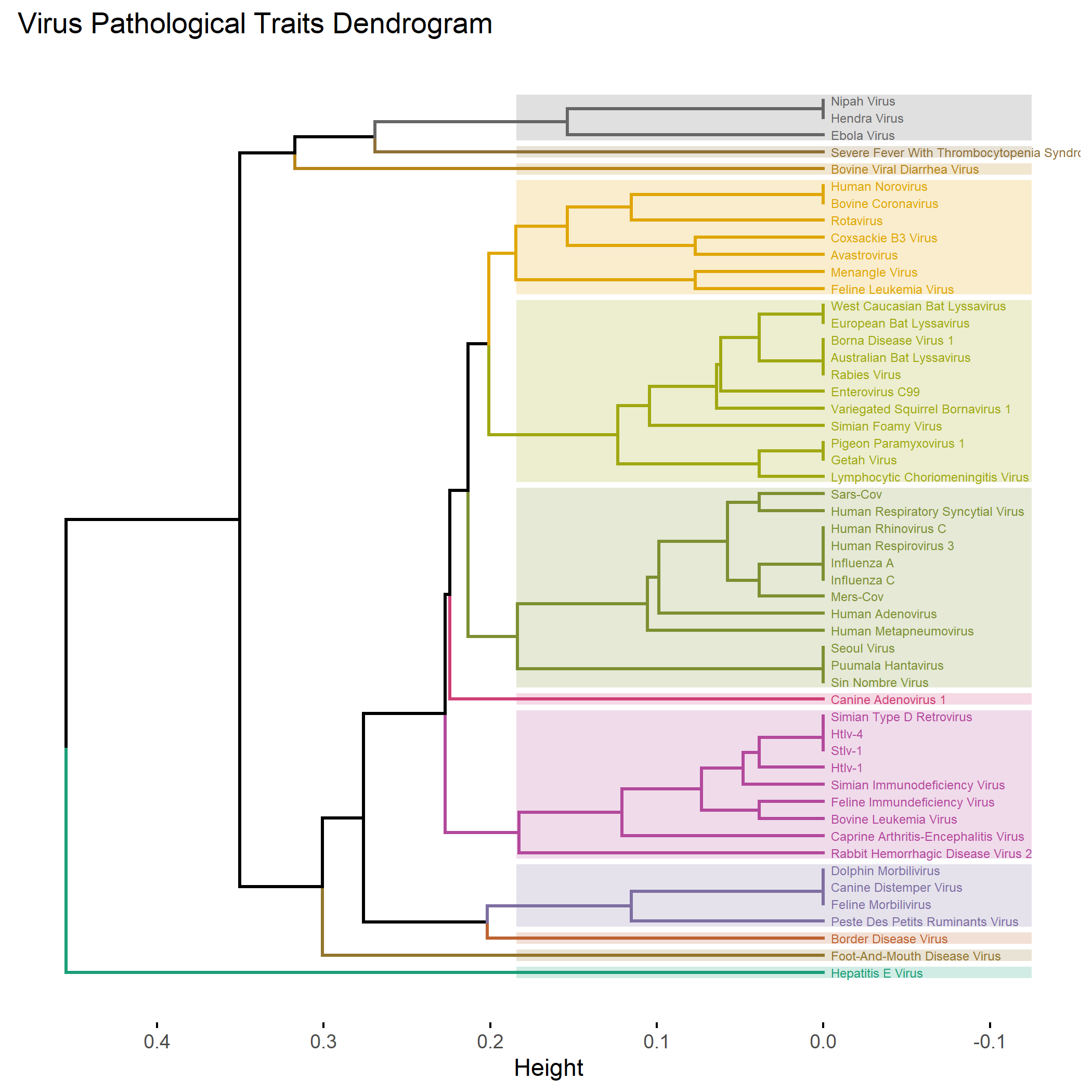
Fig 5: multiple correspondence analysis biplot for the n=213x individual observations of cross species transmission of RNA viruses from the literature search. The yellow points represent the variables. 95% confidence ellipses for cross species transmission events classified as zoonotic, anthroponotic or not zoonotic have been drawn over the individual observations in the plot.



### Hierarchical cluster analysis

When hierarchical cluster analysis is performed on the second data set of 52 viruses and their typical pathogenesis they tend to cluster into distinct groups based on pathology. Six main clusters were formed, with five of these based on pathology and morbilliviruses forming their own cluster. There were also an additional six “orphan viruses” that did not cluster with any others based on pathology (Figure 7).

.



Neurological

Gastrointestinal

Respiratory

Endothelial

Immunological

Fig 7: Dendrogram generated from 52 unique viruses and their pathological data.. To generate the dendrogram, each pathway was assigned codes reflecting its content; these data were then used to generate a distance matrix. We performed a hierarchical cluster analysis on the matrix to group similar pathological data. The dendrogram was cut to divide the tree into 12 clusters, each of which is represented by a different color. The six major clusters are named based on the predominant pathology in each (in addition to the morbilivirus cluster)

Morbilliviruses

## Discussion

### Overview

The current study investigated the nature of available pathogenesis data as it relates to cases of cross species RNA virus transmission events. The study identified some major gaps in the scientific literature particular at a molecular level when it comes to understanding the cellular level pathogenesis of cross species transmission events. Secondly, we analyzed the results of a literature search for cross-species transmission events in RNA viruses using multiple correspondence analysis and identified trends in the pathological traits associated with these infections, particularly as they related to spillover events involving humans. Thirdly, hierarchical cluster analysis identified the existence of “pathological groups” of viruses which cluster based on their pathogenesis traits which tend to have other similar traits, especially hosts involved and transmission mechanisms.

### Data availability

With respect to the availability of pathogenesis data there was a trend between the microscopic level of the data and level of unavailability, with microscopic traits like cellular receptor being less available than relatively more coarse scaled data such as organ systems affected (Fig X). The percentage of unknowns for pathological data increased the finer scale (cell or molecular level) the level of the pathological variable, with cellular receptors being the most unknown. Whilst one of the aims was to study cellular receptor trends from the 52 viruses recorded there were 33 different cellular receptors for virus infections known plus a further 14 unknown receptors. While many of the cellular receptors for viral entry in the host are unknown, equally many of the recorded receptors are putative or just one piece of the cellular attachment and entry puzzle making this difficult to analyze. Some basic types of receptors associated with viral-host interaction did occur regularly for integrins and HSPGS, suggesting that these may be a conserved group of receptors for a number of viruses, and are known to be widely spread receptors in hosts. Recent work has also begun to look at the potential role of sialic acid receptors as a conserved receptor involved in zoonotic infections (Kuchipudi et al. 2021). As this may be a promising area of future investigation, the major barrier highlighted by this study is the absence of available data at the *cellular receptor level*. Additionally, there may be even less understanding of cellular receptors in viruses involved in novel spillover events, compared to other diseases which have been endemic in humans or certain domestic species for a longer period of time, as they have been less studied.

An additional area of interest in cross species transmission is that of codon usage bias in viruses (CUB). As CUB varies per virus and the host species they are adapted to this may provide important information regarding abilities to infect new species (Bahir et al. 2009). However the relative lack of studies conducted in this area emphasize that cellular and molecular factors are often under investigated in the context of pathogenesis sand cross species transmission.

### Pathological trends

In the hierarchical cluster analysis, whilst there was some clustering based on viral phylogeny, this was not strict, with some phylogenetically related viruses not clustering together, but instead clustering based on the pathology induced. With the HCA it is interesting to note that there are still several “orphan viruses” that do not cluster well, highlighting part of the challenge here in that the interactions between virus and host that produces pathology is complex and can produce some unique outcomes. Could the clustering of viruses in groups be used to predict whether certain viruses in the group will be of risk to hosts which are infected by other viruses in the group? For example, the respiratory group all involve humans as reservoir or spillover hosts. Could the neurological group, which contains several viruses known to infect humans, in fact all possess this ability based on their shared pathological traits?

Those viruses with primary neuronal tropism have been known to have a broad host range (Brunker 2018). The neurological cluster of the dendrogram featured more variety in terms of hosts and viruses than other clusters. The receptor for rabies virus and possibly other lyssa viruses is the nicotinic acetylcholine receptor which is widely conserved between mammalian species.(Le Novere and Changeux 1995, Gotti and Clementi 2004) This feature of the pathogenesis of this virus helps to explain its broad host range and ease of transmission into novel species.

With the exception of influenza, respiratory tropic viruses tended to be transmitted between more closely related species. High barriers to infection due to mucosal immunity may require more conserved receptors i.e., between more related species for successful cross species transmission(Sato and Kiyono 2012).

Despite endotheliotropic viruses being associated with cross species transmission events causing severe illness, such as hemorrhagic fever viruses (LeDuc 1989), there were not many records with endothelial pathology and there was not a strong correspondence with any other factors. This may be due to the difficulty in accessing this receptor due to its location, making it a more specialized virus trait.

Lymphoid tropism was associated with viruses categorized as ”not zoonotic”. This greater host range may be related to co-opting the host immune system to bypass some of these interspecific barriers or perhaps there may be conserved immunological virus receptors between species.

There are some outlying variables in the MCA e.g., muscle, caused by individual recordings, which are likely not important to overall trends in pathogenesis but highlight the importance of referring to the actual data when interpreting MCA.

### Human related features

64% of articles involved humans as either reservoir or spillover host highlighting the inherit human bias in cross species transmission research. Humans are in close proximity to spillover events involving them as reservoir or spillover hosts and can report symptoms etc. They also have a vested interest in research involving these diseases that affect them directly. This is reflected in the four most common viruses in the dataset; influenza A, rabies, hepatitis E and simian foamy virus, which are all known to infect humans. In addition to cases involving humans themselves, 66% of records involved what would be considered a domestic species (either livestock or pet) as reservoir or spillover host, with only 8% of records involving neither humans nor domestic species. This demonstrates the inherent bias that exists in this type of data with either the proximity to these species resulting in reporting biases or the interest in surveillance of diseases in these species.

In addition to domesticated species, whereby contact is likely an important factor, the other major group of species with which humans exchanged viruses are primates. This is likely due to lower barriers to cross species transmission in this case due to phylogenetic relatedness (Olival et al. 2017, Guth et al. 2019)

The association between anthroponosis, those cases of virus transmission from human to other species and respiratory pathology may have to do with human behaviors. Respiratory pathology is generally caused by viruses with a respiratory droplet mode of transmission(Wang et al. 2021). Other methods of transmission are less likely from humans due to human behaviors related to hygiene reducing the risk of faeco-oral transmitted pathogens and those transmitted by direct contact(Penakalapati et al. 2017). This was reflected in the MCA where zoonosis and anthroponosis formed a subset within a larger more variable group containing cross species events not involving humans. This may also reflect a narrower range of pathologies. The occurrence of the “anthroponotic” ellipsis as a smaller ellipsis within the “zoonotic” ellipsis in the MCA for the individual host MCA and the “zoonotic” ellipsis as a smaller ellipsis within a larger “not zoonotic” in the unique virus MCA seems to reflect this (Figures 5 and 6). The importance of respiratory pathology in humans was shown again in the HCA with the respiratory cluster including only viruses with humans as spillover or reservoir hosts. Much of the driving force behind this trend is the repeated high profile anthroponotic spillover of human respiratory viruses into great apes (Muehlenbein 2016, Devaux et al. 2019). The strength of this association suggests that whilst hygiene measures are in place to prevent spillover of other pathogens that require closer contact it appears ecotourism and other activities that encroach on wild primates are still getting close enough to transmit respiratory pathogens,

## Conclusion

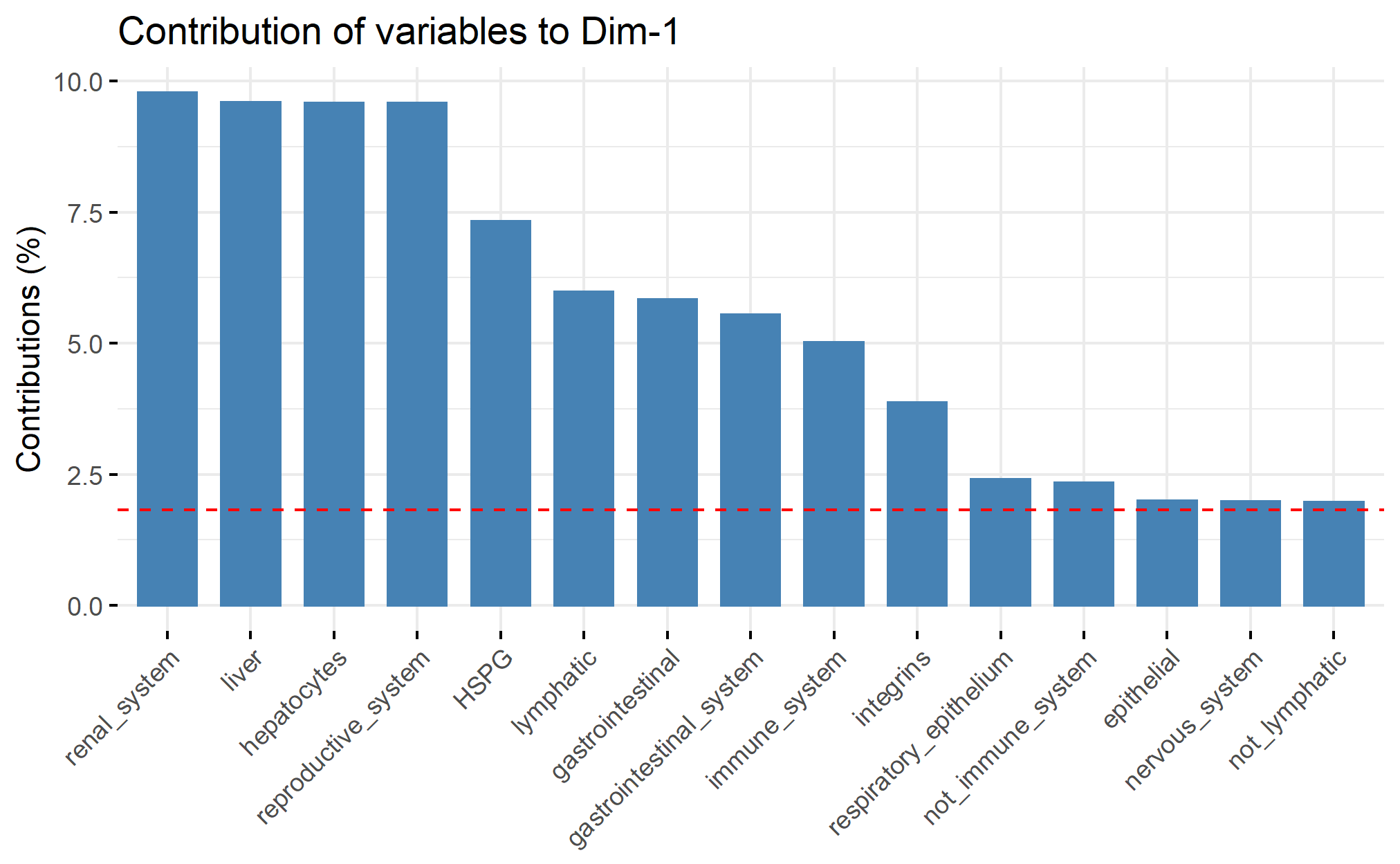
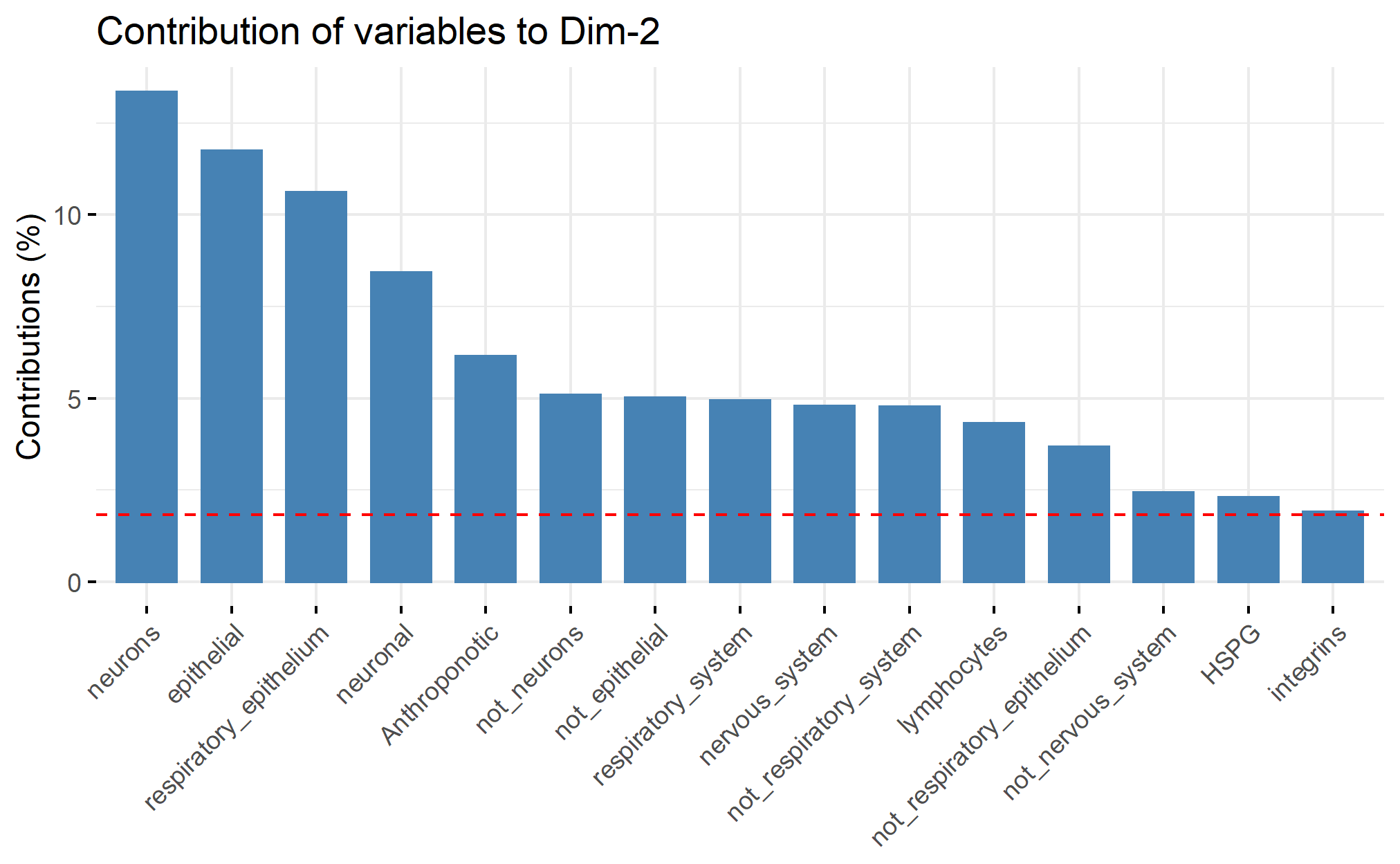
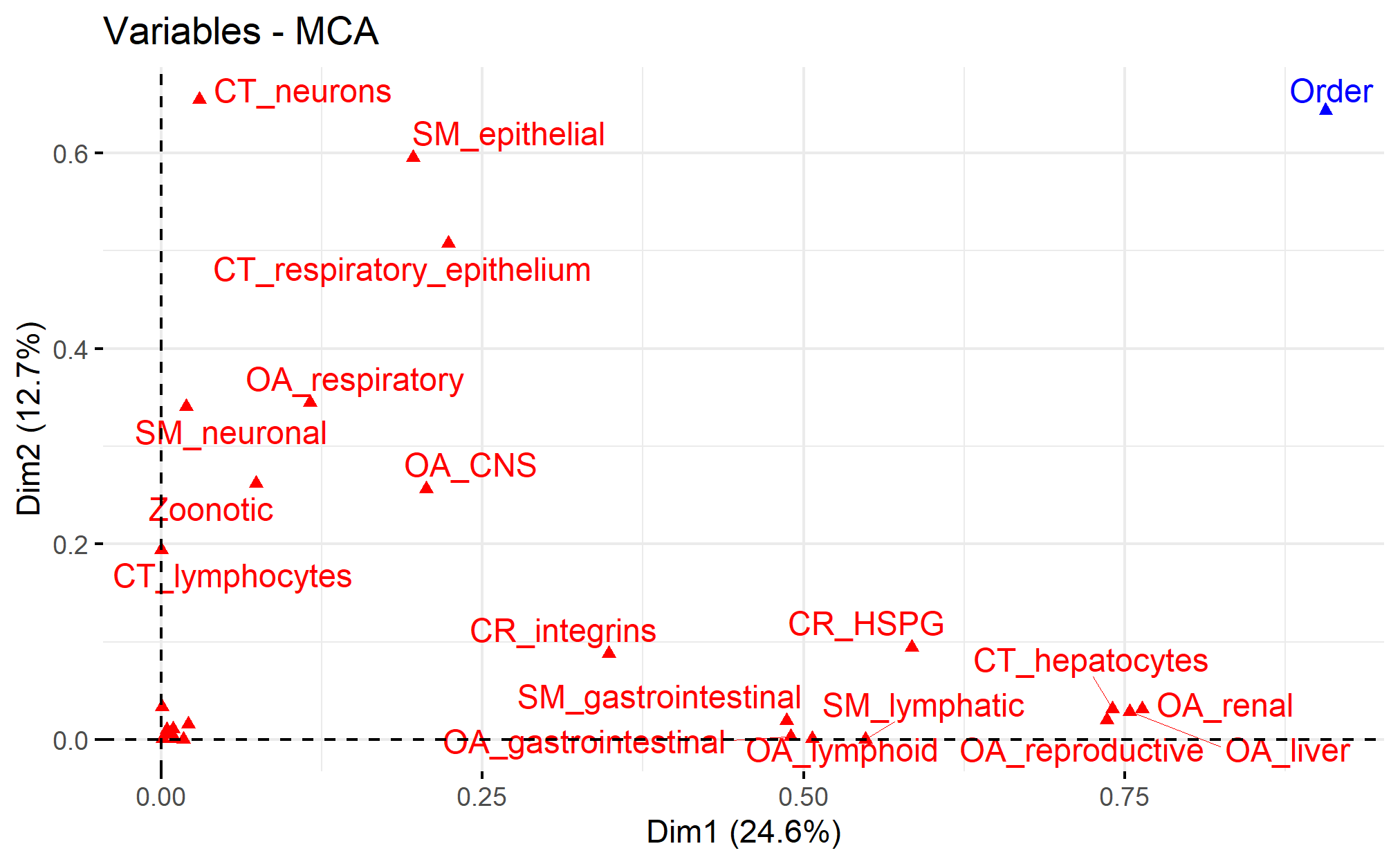
There are two main points to take from this study. Firstly, there appear to be trends in pathogenesis data at the virus-host interface which correspond to certain types of spillover event, particularly those events involving humans. Specifically, viruses involved in cross species transmission events tend to group together based on certain pathological/pathogenesis traits which often corresponded with types of hosts involved. The trends that have emerged from this study demonstrate that further work on cellular and molecular pathology of cross species transmission events is merited and should be better integrated with disease ecology and epidemiological approaches.

Secondly, this study highlights the current absence of high-quality and fine-scale virus-host pathogen interaction data in some areas, particularly at the cellular/subcellular level. This calls for more collaboration and focused efforts among the scientific community to identify receptors which may be indicative of spillover risk. Elucidating the potential role of pathogenesis in cross species transmission can raise the possibility of trying to include this data into predictive frameworks of cross species transmission risk.

## Supplementary Materials

|  |  |  |  |
| --- | --- | --- | --- |
| **Virus\_Name** | **Order** | **Family** | **Genus** |
| Australian Bat Lyssavirus | Mononegavirales | Rhabdoviridae | Lyssavirus |
| Avastrovirus | Stellavirales | Astroviridae | Avastrovirus |
| Border Disease Virus | Amarillovirales | Flaviviridae | Pestivirus |
| Borna Disease Virus 1 | Mononegavirales | Bornaviridae | Orthobornavirus |
| Bovine Coronavirus | Nidovirales | Coronaviridae | Betacoronavirus |
| Bovine Leukemia Virus | Ortervirales | Retroviridae | Deltaretrovirus |
| Bovine Viral Diarrhea Virus | Amarillovirales | Flaviviridae | Pestivirus |
| Canine Adenovirus 1 | Rowavirales | Adenoviridae | Mastadenovirus |
| Canine Distemper Virus | Mononegavirales | Paramyxoviridae | Morbillivirus |
| Caprine Arthritis-Encephalitis Virus | Ortervirales | Retroviridae | Lentivirus |
| Coxsackie B3 Virus | Picornavirales | Picornaviridae | Enterovirus |
| Dolphin Morbilivirus | Mononegavirales | Paramyxoviridae | Morbillivirus |
| Ebola Virus | Mononegavirales | Filoviridae | Ebolavirus |
| Enterovirus C99 | Picornavirales | Picornaviridae | Enterovirus |
| European Bat Lyssavirus | Mononegavirales | Rhabdoviridae | Lyssavirus |
| Feline Immundeficiency Virus | Ortervirales | Retroviridae | Lentivirus |
| Feline Leukemia Virus | Ortervirales | Retroviridae | Gammaretrovirus |
| Feline Morbilivirus | Mononegavirales | Paramyxoviridae | Morbillivirus |
| Foot-And-Mouth Disease Virus | Picornavirales | Picornaviridae | Aphthovirus |
| Getah Virus | Martellivirales | Togaviridae | Alphavirus |
| Hendra Virus | Mononegavirales | Paramyxoviridae | Henipavirus |
| Hepatitis E Virus | Hepelivirales | Hepeviridae | Hepevirus |
| HTLV-1 | Ortervirales | Retroviridae | Deltaretrovirus |
| HTLV-4 | Ortervirales | Retroviridae | Deltaretrovirus |
| Human Adenovirus | Rowavirales | Adenoviridae | Mastadenovirus |
| Human Metapneumovirus | Mononegavirales | Pneumoviridae | Metapneumovirus |
| Human Norovirus | Picornavirales | Caliciviridae | Norovirus |
| Human Respirovirus 3 | Mononegavirales | Paramyxoviridae | Respirovirus |
| Human Rhinovirus C | Picornavirales | Picornaviridae | Enterovirus |
| Influenza A | Articulavirales | Orthomyxoviridae | Alphainfluenzavirus |
| Influenza C | Articulavirales | Orthomyxoviridae | Gammainfluenzavirus |
| Lymphocytic Choriomeningitis Virus | Bunyavirales | Arenaviridae | Mammarenavirus |
| Menangle Virus | Mononegavirales | Paramyxoviridae | Pararubulavirus |
| MERS-CoV | Nidovirales | Coronaviridae | Betacoronavirus |
| Nipah Virus | Mononegavirales | Paramyxoviridae | Henipavirus |
| Peste Des Petits Ruminants Virus | Mononegavirales | Paramyxoviridae | Morbillivirus |
| Pigeon Paramyxovirus 1 | Mononegavirales | Paramyxoviridae | Avulavirus |
| Puumala Hantavirus | Bunyavirales | Hantaviridae | Orthohantavirus |
| Rabbit Hemorrhagic Disease Virus 2 | Picornavirales | Caliciviridae | Lagovirus |
| Rabies Virus | Mononegavirales | Rhabdoviridae | Lyssavirus |
| Human Respiratory Syncytial Virus | Mononegavirales | Pneumoviridae | Orthopneumovirus |
| Rotavirus | Reoviridae | Reoviridae | Rotavirus |
| SARS-CoV | Nidovirales | Coronaviridae | Betacoronavirus |
| Seoul Virus | Bunyavirales | Hantaviridae | Orthohantavirus |
| Severe Fever With Thrombocytopenia Syndrome Virus | Bunyavirales | Phenuiviridae | Bandavirus |
| Simian Foamy Virus | Ortervirales | Retroviridae | Spumavirus |
| Simian Immunodeficiency Virus | Ortervirales | Retroviridae | Lentivirus |
| Simian Type D Retrovirus | Ortervirales | Retroviridae | Betaretrovirus |
| Sin Nombre Virus | Bunyavirales | Hantaviridae | Orthohantavirus |
| STLV-1 | Ortervirales | Retroviridae | Deltaretrovirus |
| Variegated Squirrel Bornavirus 1 | Mononegavirales | Bornaviridae | Orthobornavirus |
| West Caucasian Bat Lyssavirus | Mononegavirales | Rhabdoviridae | Lyssavirus |

Supplementary table 1: list of virus name, order, family, and genus for each unique virus included in the database of spillover events.



Supplementary Figure 1: contributions of each variable to the variance shown in dimensions 1 and 2 of the multiple correspondence analysis are shown in figure A , with greater distance from the origin demonstrating a greater contribution to the variance. The scree plots B and C, show individual contribution of variables to each dimension.

## References

BAHIR, I., M. FROMER, Y. PRAT, ANDM. LINIAL. 2009. Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. Mol Syst Biol 5: 311.

BECKER, D. J., A. D. WASHBURNE, C. L. FAUST, J. R. C. PULLIAM, E. A. MORDECAI, J. O. LLOYD-SMITH, ANDR. K. PLOWRIGHT. 2019. Dynamic and integrative approaches to understanding pathogen spillover. Philos Trans R Soc Lond B Biol Sci 374: 20190014.

BEINEKE, A., W. BAUMGÄRTNER, ANDP. WOHLSEIN. 2015. Cross-species transmission of canine distemper virus—an update. One Health 1: 49-59.

BENDER, B., ANDP. SMALL JR. 1992. Influenza: pathogenesis and host defense.*In* Proceedings: Seminars in respiratory infections. pp. 38-45.

BLAGODATSKI, A., K. TRUTNEVA, O. GLAZOVA, O. MITYAEVA, L. SHEVKOVA, E. KEGELES, N. ONYANOV, K. FEDE, A. MAZNINA, E. KHAVINA, S. J. YEO, H. PARK, ANDP. VOLCHKOV. 2021. Avian Influenza in Wild Birds and Poultry: Dissemination Pathways, Monitoring Methods, and Virus Ecology. Pathogens 10.

BORREMANS, B., C. FAUST, K. R. MANLOVE, S. H. SOKOLOW, ANDJ. O. LLOYD-SMITH. 2019. Cross-species pathogen spillover across ecosystem boundaries: mechanisms and theory. Philos Trans R Soc Lond B Biol Sci 374: 20180344.

BRUNKER, K. M., N. 2018. Rabies Virus. Trends Microbiol 26: 886-887.

COSTA, T., A. J. CHAVES, R. VALLE, A. DARJI, D. VAN RIEL, T. KUIKEN, N. MAJO, ANDA. RAMIS. 2012. Distribution patterns of influenza virus receptors and viral attachment patterns in the respiratory and intestinal tracts of seven avian species. Vet Res 43: 28.

CUI, X., K. FAN, X. LIANG, W. GONG, W. CHEN, B. HE, X. CHEN, H. WANG, X. WANG, P. ZHANG, X. LU, R. CHEN, K. LIN, J. LIU, J. ZHAI, D. X. LIU, F. SHAN, Y. LI, R. A. CHEN, H. MENG, X. LI, S. MI, J. JIANG, N. ZHOU, Z. CHEN, J.-J. ZOU, D. GE, Q. YANG, K. HE, T. CHEN, Y.-J. WU, H. LU, D. M. IRWIN, X. SHEN, Y. HU, X. LU, C. DING, Y. GUAN, C. TU, ANDY. SHEN. 2023. Virus diversity, wildlife-domestic animal circulation and potential zoonotic viruses of small mammals, pangolins and zoo animals. Nature Communications 14: 2488.

DE GRAAF, M., ANDR. A. FOUCHIER. 2014. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. EMBO J 33: 823-841.

DEVAUX, C. A., O. MEDIANNIKOV, H. MEDKOUR, ANDD. RAOULT. 2019. Infectious disease risk across the growing human-non human primate interface: A review of the evidence. Frontiers in Public Health 7: 305.

DUQUE, H., ANDB. BAXT. 2003. Foot-and-mouth disease virus receptors: comparison of bovine alpha(V) integrin utilization by type A and O viruses. J Virol 77: 2500-2511.

ELLWANGER, J. H., ANDJ. A. B. CHIES. 2021. Zoonotic spillover: Understanding basic aspects for better prevention. Genet Mol Biol 44: e20200355.

ENGERING, A., L. HOGERWERF, ANDJ. SLINGENBERGH. 2013. Pathogen-host-environment interplay and disease emergence. Emerg Microbes Infect 2: e5.

ESCUDERO-PÉREZ, B., A. LALANDE, C. MATHIEU, ANDP. LAWRENCE. 2023. Host&ndash;Pathogen Interactions Influencing Zoonotic Spillover Potential and Transmission in Humans. Viruses 15: 599.

FUKUYAMA, S., ANDY. KAWAOKA. 2011. The pathogenesis of influenza virus infections: the contributions of virus and host factors. Current opinion in immunology 23: 481-486.

GALILI, T. 2015. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics 31: 3718-3720.

GOTTI, C., ANDF. CLEMENTI. 2004. Neuronal nicotinic receptors: from structure to pathology. Prog Neurobiol 74: 363-396.

GOWER, J. C. 1971. A General Coefficient of Similarity and Some of Its Properties. Biometrics 27: 857-871.

GUTH, S., E. VISHER, M. BOOTS, ANDC. E. BROOK. 2019. Host phylogenetic distance drives trends in virus virulence and transmissibility across the animal-human interface. Philos Trans R Soc Lond B Biol Sci 374: 20190296.

HANSBRO, P. M., S. WARNER, J. P. TRACEY, K. E. ARZEY, P. SELLECK, K. O'RILEY, E. L. BECKETT, C. BUNN, P. D. KIRKLAND, D. VIJAYKRISHNA, B. OLSEN, ANDA. C. HURT. 2010. Surveillance and analysis of avian influenza viruses, Australia. Emerg Infect Dis 16: 1896-1904.

HOU, Y. J., K. OKUDA, C. E. EDWARDS, D. R. MARTINEZ, T. ASAKURA, K. H. DINNON, T. KATO, R. E. LEE, B. L. YOUNT, ANDT. M. MASCENIK. 2020. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. Cell 182: 429-446. e414.

KASSAMBARA, A., ANDF. MUNDT. 2017. Package ‘factoextra’. Extract and visualize the results of multivariate data analyses 76.

KUCHIPUDI, S. V., R. K. NELLI, A. GONTU, R. SATYAKUMAR, M. SURENDRAN NAIR, ANDM. SUBBIAH. 2021. Sialic Acid Receptors: The Key to Solving the Enigma of Zoonotic Virus Spillover. Viruses 13.

KUMLIN, U., S. OLOFSSON, K. DIMOCK, ANDN. ARNBERG. 2008. Sialic acid tissue distribution and influenza virus tropism. Influenza and other respiratory viruses 2: 147-154.

LAJEUNESSE, M. J. 2016. Facilitating systematic reviews, data extraction, and meta-analysis with the metagear package for R. . Methods in Ecology and Evolution 7: 323-330.

LE NOVERE, N., ANDJ. P. CHANGEUX. 1995. Molecular evolution of the nicotinic acetylcholine receptor: an example of multigene family in excitable cells. J Mol Evol 40: 155-172.

LE, S., JOSSE, J., HUSSON, F. 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25: 1-18.

LEDUC, J. W. 1989. Epidemiology of Hemorrhagic Fever Viruses. Reviews of Infectious Diseases 11: S730-S735.

LOUTEN, J. 2016. Virus Transmission and Epidemiology. Essential Human Virology: 71-92.

MA, W., R. E. KAHN, ANDJ. A. RICHT. 2008. The pig as a mixing vessel for influenza viruses: Human and veterinary implications. J Mol Genet Med 3: 158-166.

MAECHLER, M., ROUSSEEUW, P., STRUYF, A., HUBERT, M., HORNIK, K. 2022. cluster: Cluster Analysis Basics and Extensions. . pp.

MARSTON, D. A., A. C. BANYARD, L. M. MCELHINNEY, C. M. FREULING, S. FINKE, X. DE LAMBALLERIE, T. MULLER, ANDA. R. FOOKS. 2018. The lyssavirus host-specificity conundrum-rabies virus-the exception not the rule. Curr Opin Virol 28: 68-73.

MOLLENTZE, N., S. A. BABAYAN, ANDD. G. STREICKER. 2021. Identifying and prioritizing potential human-infecting viruses from their genome sequences. PLOS Biology 19: e3001390.

MOLLENTZE, N., D. G. STREICKER, P. R. MURCIA, K. HAMPSON, ANDR. BIEK. 2020. Virulence mismatches in index hosts shape the outcomes of cross-species transmission. Proc Natl Acad Sci U S A 117: 28859-28866.

MUEHLENBEIN, M. P. 2016. Disease and human/animal interactions. Annual Review of Anthropology 45: 395-416.

NORKIN, L. C. 2010. Virology: molecular biology and pathogenesis. ASM press

OLIVAL, K. J., P. R. HOSSEINI, C. ZAMBRANA-TORRELIO, N. ROSS, T. L. BOGICH, ANDP. DASZAK. 2017. Host and viral traits predict zoonotic spillover from mammals. Nature 546: 646-650.

PAGE, M. J., J. E. MCKENZIE, P. M. BOSSUYT, I. BOUTRON, T. C. HOFFMANN, C. D. MULROW, L. SHAMSEER, J. M. TETZLAFF, E. A. AKL, S. E. BRENNAN, R. CHOU, J. GLANVILLE, J. M. GRIMSHAW, A. HRÓBJARTSSON, M. M. LALU, T. LI, E. W. LODER, E. MAYO-WILSON, S. MCDONALD, L. A. MCGUINNESS, L. A. STEWART, J. THOMAS, A. C. TRICCO, V. A. WELCH, P. WHITING, ANDD. MOHER. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj 372: n71.

PARK, A. W. 2019. Phylogenetic aggregation increases zoonotic potential of mammalian viruses. Biol Lett 15: 20190668.

PENAKALAPATI, G., J. SWARTHOUT, M. J. DELAHOY, L. MCALILEY, B. WODNIK, K. LEVY, ANDM. C. FREEMAN. 2017. Exposure to Animal Feces and Human Health: A Systematic Review and Proposed Research Priorities. Environ Sci Technol 51: 11537-11552.

PLOWRIGHT, R. K., C. R. PARRISH, H. MCCALLUM, P. J. HUDSON, A. I. KO, A. L. GRAHAM, ANDJ. O. LLOYD-SMITH. 2017. Pathways to zoonotic spillover. Nat Rev Microbiol 15: 502-510.

POSIT TEAM. 2022. RStudio: Integrated Development Environment for R. Posit Software, Boston, MA. pp.

PULLIAM, J. R., ANDJ. DUSHOFF. 2009. Ability to replicate in the cytoplasm predicts zoonotic transmission of livestock viruses. J Infect Dis 199: 565-568.

R CORE TEAM. 2022. R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. pp.

RENDON-MARIN, S., R. DA FONTOURA BUDASZEWSKI, C. W. CANAL, ANDJ. RUIZ-SAENZ. 2019. Tropism and molecular pathogenesis of canine distemper virus. Virol J 16: 30.

RUIZ-ARAVENA, M., C. MCKEE, A. GAMBLE, T. LUNN, A. MORRIS, C. E. SNEDDEN, C. K. YINDA, J. R. PORT, D. W. BUCHHOLZ, Y. Y. YEO, C. FAUST, E. JAX, L. DEE, D. N. JONES, M. K. KESSLER, C. FALVO, D. CROWLEY, N. BHARTI, C. E. BROOK, H. C. AGUILAR, A. J. PEEL, O. RESTIF, T. SCHOUNTZ, C. R. PARRISH, E. S. GURLEY, J. O. LLOYD-SMITH, P. J. HUDSON, V. J. MUNSTER, ANDR. K. PLOWRIGHT. 2022. Ecology, evolution and spillover of coronaviruses from bats. Nat Rev Microbiol 20: 299-314.

SATO, S., ANDH. KIYONO. 2012. The mucosal immune system of the respiratory tract. Curr Opin Virol 2: 225-232.

VALERO-RELLO, A., ANDR. SANJUÁN. 2022. Enveloped viruses show increased propensity to cross-species transmission and zoonosis. bioRxiv: 2022.2007.2029.501861.

VIJAYKRISHNA, D., L. POON, H. ZHU, S. MA, O. LI, C. CHEUNG, G. SMITH, J. PEIRIS, ANDY. GUAN. 2010. Reassortment of pandemic H1N1/2009 influenza A virus in swine. Science 328: 1529-1529.

WANG, C. C., K. A. PRATHER, J. SZNITMAN, J. L. JIMENEZ, S. S. LAKDAWALA, Z. TUFEKCI, ANDL. C. MARR. 2021. Airborne transmission of respiratory viruses. Science 373: eabd9149.

WIDAGDO, W. R., V. S.;SCHIPPER, D.;KOLIJN, K.;VAN LEENDERS, GJLH;BOSCH, B. J.;BENSAID, A.;SEGALÉS, J.;BAUMGÄRTNER, W.;OSTERHAUS, ADME;KOOPMANS, M. P.;VAN DEN BRAND, J. M. A.;HAAGMANS, B. L. 2016. Differential Expression of the Middle East Respiratory Syndrome Coronavirus Receptor in the Upper Respiratory Tracts of Humans and Dromedary Camels. J Virol 90: 4838-4842.

WOOLHOUSE, M., F. SCOTT, Z. HUDSON, R. HOWEY, ANDM. CHASE-TOPPING. 2012. Human viruses: discovery and emergence. Philos Trans R Soc Lond B Biol Sci 367: 2864-2871.

ZAMBON, M. C. 1999. Epidemiology and pathogenesis of influenza. Journal of Antimicrobial Chemotherapy 44: 3-9.

ZAMBON, M. C. 2001. The pathogenesis of influenza in humans. Reviews in medical virology 11: 227-241.

ZHANG, X., F. L. CUNNINGHAM, L. LI, K. HANSON-DORR, L. LIU, K. WATERS, M. GUAN, A. K. OLIVIER, B. S. SCHMIT, J. M. NOLTING, A. S. BOWMAN, M. K. TORCHETTI, T. J. DELIBERTO, ANDX. F. WAN. 2020. Tissue Tropisms of Avian Influenza A Viruses Affect Their Spillovers from Wild Birds to Pigs. J Virol 94.