A Systematic Review of the Distribution and Prevalence of Viruses Detected in the *Peromyscus maniculatus* Species Complex (Rodentia: Cricetidae)

Authors:

Ally Finkbeiner

Nicholas School of the Environment

Duke University

ORCID: 0009-0000-2189-2544

Ahmad Khatib

School of Life Sciences

Arizona State University

ORCID: 0009-0009-4661-9961

Nathan Upham

School of Life Sciences

Arizona State University

ORCID: 0000-0001-5412-9342

*Beckett Sterner

School of Life Sciences

Arizona State University

ORCID: 0000-0001-5219-7616

*Corresponding author: bsterne1@asu.edu

Abstract: The North American Deermouse, *Peromyscus maniculatus*, is one of the most widespread and abundant mammals on the continent. It is of public health interest as a known host of several viruses that are transmissible to humans and can cause illness, including the acute respiratory disease Hantavirus pulmonary syndrome (HPS). However, recent taxonomic studies indicate that *P. maniculatus* is a complex of multiple species, raising questions about how to identify and interpret three decades of hantavirus monitoring data. We conducted a systematic review investigating the prevalence and spatial distribution of viral taxa detected in wild populations allocated to *P. maniculatus*. From the 49 relevant studies published from 2000 to 2022, we extracted and analyzed spatial occurrence data to calculate weighted populational prevalences for hantaviruses. We found that detection efforts have been concentrated in the Western United States and Mexico with a focus on the spread of Sin Nombre virus (*Orthohantavirus sinnombreense*), the primary causative agent of HPS. There are significant gaps in the existing literature both geographically and in regard to the types of viruses being sampled. These results are significantly impacted by a recent taxonomic split of *P. maniculatus* into four species, and we were able to update 94% of hantavirus observations to reflect this change. Investigating the uncertain, and likely multiple, phylogenetic histories of these viral hosts should be a key emphasis of future modeling efforts.

Author Summary:

Understanding the interactions of viruses with mammal hosts is critical for monitoring disease spread and identifying species or geographic areas at high risk for future zoonotic disease outbreaks. However, much of the data scientists have collected about viruses in mammals remains fragmented in published literature and is not available in an up-to-date, standardized format online. We conducted a systematic review of virus observations in one of the most abundant, widespread group of rodent species in North America, the *Peromyscus maniculatus* (North American Deermouse) species complex. We extracted and analyzed the spatial distribution of 62,421 observations reporting the results of tests for viruses in the Hantaviridae, Arenaviridae, and Flaviviridae families. We find that conclusions about reservoir host status may be impacted by recent taxonomic proposals that split *P. maniculatus* into four distinct species, and we show how the taxonomic identifications of 94% of the reported observations can be updated using expert-generated maps of the ranges for the proposed new species. We also highlight the uneven geographic sampling of viruses in the rodent populations and major gaps that remain in our knowledge.

1. Introduction

The North American Deermouse, *Peromyscus maniculatus* (J. A. Wagner, 1845) is one of the most abundant and widespread mammals native to the continent (1). The genus *Peromyscus* sec. Mammal Diversity Database v1.12.1 (2–4) includes 82 recognized living species (3) and has been the subject of extensive research in ecology, development, genetics, evolution, and epidemiology (1,5–10). *P. maniculatus* is of particular interest to the public health sector and has been traditionally well-studied because it is known to be a source of several diseases that are communicable to humans (9,11). Research has been particularly focused on the North American Deermouse's role as the primary host of Sin Nombre virus (SNV), an RNA virus in the genus *Orthohantavirus* which was first identified in 1993 and was recently renamed *Orthohantavirus sinnombreense* by the International Committee on the Taxonomy of Viruses (12–15). SNV and other New World hantaviruses are the causative agents of Hantavirus pulmonary syndrome (HPS), an acute respiratory disease with a case fatality rate of 60% at the time of the first outbreaks (8). From 1993 to 2021, a total of 850 cases of HPS have been reported in the United States (16), with most cases occurring in the Southwest (Figure 1). Despite the relative rarity of HPS cases compared to other hantavirus illnesses, the sudden emergence of the disease and potential for mutation has spurred intensive efforts to characterize the causative viral transmission pathways (17–20). Modeling the risks of hantavirus disease and novel zoonoses remains an urgent research objective that depends on comprehensive, up-to-date information about pathogen prevalence and the ecological factors influencing disease spread (21,22).

[FIGURE 1 GOES HERE]

This paper presents a systematic review of the existing literature on virus detection studies of wild *P. maniculatus* in order to summarize findings from the last two decades and identify potential areas of further research. We focus on a spatial analysis of sampling effort and prevalence in studies that tested for hantaviruses and arenaviruses. Our results thereby add to the evidence base for ongoing research into the co-evolution of North American rodents with viral pathogens.

In addition, we discuss how study results are relevant for the reservoir host status of *P. maniculatus* considering recently proposed taxonomic revisions. Multiple genetic studies indicate *Peromyscus maniculatus* is a species complex (23). It was recently proposed to split *P. maniculatus* into between four and six species-level lineages (2,24,25), though this remains controversial (9). According to the Mammal Diversity Database v1.12.1 (2–4), the *P. maniculatus* species complex is composed of four species that are distributed as follows: (i) throughout Northern and Central Mexico, referred to *P. labecula* D. G. Elliot, 1903; (ii) in southern California and Baja California, referred to *P. gambelii* (S. F. Baird, 1858); (iii) across the continental United States west of the Mississippi River and into northern Canada, referred to *P. sonoriensis* (Le Conte, 1853); and (iv) east of the Mississippi River until the Atlantic Ocean and north until the Hudson Bay, referred to *P. maniculatus* sensu stricto (J. A. Wagner, 1845) (Figure 2). The genetic distinctiveness of *P. gambelii* and *P. sonoriensis* from each other and from *P. keeni* in the Pacific Northwest (9) supports the existence of multiple species within what has been typically referred to as "*P. maniculatus*" in ecological and biomedical studies. The latter identifications are now referrable to the continentally distributed *P. maniculatus* sensu lato, and thus require re-interpretation relative to these alternative taxonomic hypotheses.

[FIGURE 2 GOES HERE]

While hantavirus prevalence is the primary focus of biomedical research on wild-caught *P. maniculatus*, some evidence suggests that rodents in this group may potentially be susceptible to infection by other zoonotic viruses connected with human fatalities, including SARS-CoV-2, flaviviruses associated with tick-borne encephalitis, and arenaviruses (26–28). Members of this deermouse species complex can be found in every terrestrial ecosystem across the North American continent, including peridomestic areas resulting in frequent contact with humans. Numerous field studies have been conducted on North American Deermice to determine virus prevalence and distribution, especially in regions where outbreaks have previously occurred. These field studies have detected novel viruses and genetic sequencing has allowed researchers to build up-to-date virus and host phylogenies, but, as we show here, the approach to zoonotic virus monitoring in deermice has not been consistent or standardized.

2. Results

Studies have sampled populations in the *P. maniculatus* species complex across 11 different U.S. states, six states in Mexico, and one Canadian province since 2000. Considering 62,663 unique observations across 49 studies, we are able to assign geographic locations to 62,421 (99.6%) at the county (U.S.) or province level (Canada and Mexico). This includes observations that we split evenly across counties when studies only reported aggregate results for two or more counties. While we focus only on observations related to the *P. maniculatus* complex, we note that the full extracted dataset in Table S3 includes observations from over 100 rodent species and that about a third of these had positive test results for an arenavirus, flavivirus

or hantavirus. This highlights the taxonomic scope of rodent-virus “dark data” published in scientific articles that is not readily available for open science modeling and research (29).

Scopus returns the largest number of potential papers compared to other databases, regardless of the type of search query used in the literature review (Table S1). PubMed Central produces more potential papers than PubMed when the search queries are broad, though the latter returns nearly 5 times as many papers when the search query includes virus-related terms as well as host species names. We find that for this particular species, inclusion of common names and junior synonyms in search queries has only marginal effects (<1%) on the total number of citations returned across all three databases. However, this result excludes some additional references that are returned by searching for the common names “deer mice” and “deer mouse,” which are not recognized by MDD and are sometimes also used to refer to the whole genus *Peromyscus*. A supplementary search with “deer mice” and “deer mouse” identified three additional papers meeting our inclusion criteria. The small effect of including synonyms on search results may be due to the relatively recent application of alternative species names to populations in the *P. maniculatus* species complex (24,25). Studies of the influence of including synonyms on search results for other taxa have found larger effects (30,31).

The primary serological method used across all studies is ELISA, though RT-PCR and IFAT, among other tools, are also performed (Table S2). SNV is the primary target of 40 studies. One study reports results for Monongahela virus, which is now part of SNV according to recent taxonomic updates by the ICTV (32). Two studies report on the arenavirus species Whitewater Arroyo mammarenavirus (WWAV) and Amapari virus (AMAV), and one study reports on the flavivirus Powassan virus (POWV). Additionally, five studies explicitly sought to detect *Hantaviridae* or *Arenaviridae* antibodies at the family level and do not report results at the virus

species level. The seroprevalence ranges greatly across studies from 0% to 100% of individuals sampled in each subpopulation. Sampling effort varies as well, with the smallest study reporting 5 unique observations and the largest reporting 11,391.

We find that interpreting the specificity of reported SNV antibody detection results involves making implicit background assumptions that are not typically stated in the metadata associated with existing species interaction databases. Of the 40 studies reporting antibody results for SNV, 38 use detection methods targeting the nucleocapsid protein on the S segment as the antigen. The remaining two antibody studies (33,34) used a test based on antigen sourced from the lysate of Vero E6 cells infected by Prospect Hill Virus strain PHV1 (35). Since the nucleocapsid protein is widely conserved among hantaviruses, antibody tests based on this protein in one hantavirus species will also be broadly cross-reactive with other hantaviruses. However, if SNV is the only hantavirus species actually present in *P. maniculatus*, then the antibody tests can still be interpreted as detecting SNV.

To investigate the conclusion that SNV is the only hantavirus species detected in the *P. maniculatus* species complex to date, we analyzed the phylogenetic history of all available hantavirus nucleotide sequences identified with *P. maniculatus* as the host (Figures S1–S3). For each segment (S, M, and L), all hantavirus sequences clustered most closely with the SNV reference sequence compared to other hantavirus reference sequences. For the M and S segments, bootstrap support of 98 and 86, respectively, supported the clustering of *P. maniculatus*-derived sequences with the SNV reference. For the L segment, a bootstrap value of 100 supports the exclusion of 25 non-SNV reference sequences from the SNV reference and the *P. maniculatus*-derived sequences. The specificity of reported antibody tests is also supported by evidence from studies using PCR tests, which target both the S and M segments and should

therefore be sufficient to distinguish SNV from other sequences. In particular, the interpretation of antibody tests as detecting SNV is confirmed by 8 studies that used PCR tests to confirm positive hantavirus antibody results (Supplementary Table 2).

With these results in mind, we find that studies reporting SNV results span the longitudinal range from California to Indiana (Figure 3). One or more studies about SNV-*P. maniculatus* interactions has been published in almost every year since 2000. Other viruses are sampled less frequently and at a lower volume, with a maximum number of observations of 353 for both WWAV and AMAV collected in a single study (Table S2). POWV and arenavirus are investigated in only a single study each. In New Mexico, one study reported finding POWV in *P. maniculatus*, but precise details on location are unavailable (36). In Texas, multiple studies focused on detecting arenaviruses in *P. maniculatus*, including WWAV and AMAV.

Since 2000, sampling has been concentrated on the western half of the U.S. despite the fact that rodents in the *P. maniculatus* species complex can be found across the entirety of North America (Figures 2 and 3). Counties in Montana, Utah, and Colorado had the highest average sampling effort, with fifteen different counties hosting more than 1,500 unique observations each. This increased level of observations is partially due to multiple studies being carried out in the same county, as some locations are sites of more than one study while others are only sampled once within the last 23 years (Table S2). In studies that took place over multiple counties, all collection sites are represented. One study is published from Indiana and another from West Virginia, but the rest of the U.S. Midwest and East Coast are not sampled for *P. maniculatus* during the focal interval of 2000 to 2022.

[FIGURE 3 GOES HERE]

Sampling effort is more evenly distributed across geographic locations than across virus types. In terms of the number of studies conducted in each state, Montana has twelve while Nevada, Arizona, West Virginia, Oregon, and Indiana each boast one (Table S2). However, the number of studies alone does not accurately represent the total sampling effort in a geographic area; both West Virginia and Oregon only host one study, but the former has a total of 15 observations while the latter has 3,175. Montana has the highest number of observations at 29,857, while Colorado and Utah have the second and third highest counts with 11,794 and 7,889, respectively.

As the number of reported observations increases between 2000 and 2022, so do the number of locations sampled. However, the total number of localities sampled across this time period is only 22 within 11 U.S. states as well as Mexico and Canada, which does not accurately cover the entire range of rodents within the *P. maniculatus* species complex.

The dominance of studies aimed at detecting SNV appears to correlate with historical data on the risk of HPS (Figure 4). There is a moderate positive relationship between the number of reported human HPS cases in a location and the sampling effort in that location (coefficient=0.06, standard error=0.008, no intercept; R^2=0.52; p-value < 0.0001). This trend, however, is driven strongly by a handful of data points. Notably, Montana and Utah are overrepresented in terms of sampling effort compared to the number of historic HPS cases in those locations (studentized residuals of 4.7 and 3.7, respectively). On the other hand, Alberta, Arizona, and Washington are undersampled relative to HPS cases (studentized residuals of -2.2, -2.1, -1.8, respectively).

[FIGURE 4 GOES HERE]

The highest average seroprevalences are found in New Mexico and California, with Rio Arriba, Cibola, and Inyo counties all reporting greater than 40% of individuals as positive (Table S2). The majority of counties hovers between seropositivity levels of 10% and 30% of sampled rodents (Figure 5). Several counties in southern California, which fall within the range of *P. gambelii*, find no rodents with virus antibodies among those sampled (Table S2, Figure 5).

The total number of unique virus species confirmed across all studies is four (Table S2). The relationship between the number of observations and the number of virus species identified in the literature is not linear. This is because the overwhelming majority of studies are focused on SNV antibodies and thus add to the volume of observations without introducing a novel virus.

Most of the gathered studies are conducted within the range of the proposed species *P. sonoriensis*, which covers most of the western U.S. including the aforementioned states of Montana, Utah, and Colorado (Figure 6). Sampling is also conducted within the ranges of three other putative species within the *P. maniculatus* species complex: *P. gambelii, P. labecula,* and *P. maniculatus* (sensu stricto). The arenavirus studies correspond with the range of *P. labecula*, while the SNV and hantavirus studies are concentrated within the range of *P. sonoriensis*.

Based on the non-overlapping expert species range maps (i.e., reciprocally allopatric), we are able to assign 94% (58,845 out of 62,404) of all observations unambiguously to one of the four species (Table S4). The other 6% are ambiguous with respect to two or more species. We assign an observation unambiguously to a species if the boundaries of the corresponding county or province fall entirely within the range of that species. In some cases, the county or province boundaries fall partly or wholly outside the range of all four species, which we assign

unambiguously to a species only if the county or province boundaries overlapped with a single species. Of the 94% of observation that we could assign, *P. sonoriensis* received by far the most assignments (90%) compared to 6% for *P. gambelii*, 2% for *P. maniculatus* sensu stricto, and <1% for *P. labecula*.

[FIGURES 5 AND 6 GO HERE]

3. Discussion

Zoonotic viruses in the species complex *P. maniculatus* have been non-systematically monitored over the last 20+ years. In scientific literature published since 2000, we found reports of 62,663 unique observations of wild *P. maniculatus* (sensu lato) interacting with four different virus species from three viral families. This body of research covers 11 unique U.S. states as well as regions in Mexico and Canada, but it does not reflect the entirety of the range of the *P. maniculatus* species complex, indicating that there are significant gaps in the literature within the 2000 – 2022 period.

Sampling effort varied significantly by both geographical location and virus genotype, with Sin Nombre virus (SNV) overrepresented in the literature relative to other virus families, especially in the states of Montana and Colorado. With 82% of the studies reporting on SNV, the SNV prevalence in *P. maniculatus* over the last 22 years is likely the best representation of hantaviral sharing dynamics in this group of rodents. Results on Powassan virus, Whitewater Arroyo Virus, and Amapari virus are informative in that they indicate the presence of other virus families within the *P. maniculatus* complex. However, further conclusions cannot be drawn about the distribution of these virus families without additional targeted sampling.

This systematic review suggests that, over time, even as more populations referrable to the *P. maniculatus* species complex are being sampled for zoonotic viruses, new viruses are not being discovered. While at first glance this may suggest that virus sampling for *P. maniculatus* is adequate, it more likely indicates that sampling efforts have not been sufficiently dedicated to discovering other known or novel viruses (7). Instead, researchers have been targeting Sin Nombre virus using antibody and PCR tests given their focus on understanding the public health risks posed by the distribution and prevalence of certain known viruses. Hantaviruses in general were sampled far more than arenaviruses or flaviviruses, which reflects public health concerns relative to HPS cases. However, both arenaviruses and flaviviruses are known to be capable of spreading to humans (37,38), suggesting that increased sampling effort dedicated to these viral families will provide valuable public health information.

Only 22% of U.S. states were represented in the literature even though the *P. maniculatus* species complex is widespread across the continental U.S. except for the Southeast, where other *Peromyscus* species are found (e.g., *P. polionotus*, *P. leucopus*). While Indiana and West Virginia each received only one study, the rest of the U.S. Midwest and East Coast has remained unsampled for *P. maniculatus*-to-hantavirus interactions despite regional virus detection studies being conducted in other rodent species over the same time period (39–41). HPS disease cases have been reported only sparingly in states east of the Mississippi since 1993 (Figure 1), so it is unsurprising that recent sampling efforts have not been dedicated to these regions where the risk of disease transmission to humans is sufficiently low (Figure 3). Interestingly, a handful of states where a significant number of hantavirus disease cases have historically been reported—Washington, Wyoming, North Dakota, and South Dakota—were not sampled for *P. maniculatus*

occurrences in any of the studies, even though new hantavirus disease cases have occurred in each of these places within the last decade (16).

Among the geographic locations that were sampled (Figure 3), most populations fall largely within the subdivided range of *P. sonoriensis*, and the dominant virus over this range was SNV (Figure 6). Surprisingly, only two sampling locations are now associated with *P. maniculatus* sensu stricto: Marion County, Indiana, and Randolph County, West Virginia. The latter was found positive for Monongahela virus, which is now grouped taxonomically with SNV, and the former was a positive for a generic hantavirus ELISA test.

Seroprevalence varied greatly across geographic regions, and the associated variation in sampling effort makes it difficult to draw conclusions about regional trends of virus prevalence. Several counties in Montana, Utah, and Colorado reported seropositivity of greater than 30%, while numerous studies conducted in counties in California, New Mexico, and Texas found zero positive individuals. Since sampling effort was highly inconsistent across all locations, the seropositivity results should be considered as reflections of not only the prevalence of hantavirus in the given region but also of the number of rodents sampled and the time period over which sampling took place. Additionally, many of the studies we included investigated demographic and ecological factors that may influence Sin Nombre prevalence, including rodent age and sex, biotic interactions with other rodent species, and abiotic effects from temperature and rainfall (e.g., 42–44,34). More consistent and even sampling efforts across the wider range of *P. maniculatus* would be necessary to more conclusively investigate the geographic trends in virus prevalence.

We also note that virus taxonomy is changing in parallel with the proposed revisions to *P. maniculatus* (45). The recent taxonomic union of Sin Nombre virus and New York virus, for

example, would implicate both *P. maniculatus* and *P. leucopus* as reservoir hosts for different strains of a single virus species whose range spans from the Pacific to Atlantic coasts of North America. This viral taxonomic change is based on hierarchical genetic clustering, in part due to the high similarity of nucleoprotein and glycoprotein amino acid sequences of Sin Nombre and New York viruses. However, a human HPS patient infected with the New York variant showed no serologic reactivity to the Sin Nombre glycoprotein, potentially indicating different seroneutralization responses in humans (46).

Further topics for research include a deeper investigation into the evidence for particular *Peromyscus* species as reservoir hosts for hantavirus and other species. Scientific definitions of reservoir host status vary significantly and prioritize different types of biological relationships and evidence, ranging from simple detection in a host to persistent pathogen maintenance with or without serious symptoms or a history of co-evolution (47–53). Historically, hantaviruses were thought to closely co-evolve with single host species responsible for indefinitely maintaining the pathogen in the environment. Recent analyses have added nuance to this picture, showing that a number of hantavirus species infect multiple rodent hosts. More generally, biologists increasingly define reservoir hosts as composed of meta-populations or ecological assemblages of multiple species. Our results provide evidence that *P. maniculatus* is not the reservoir host for Sin Nombre virus in the narrow sense of being the sole biological species responsible for the pathogen's maintenance. Instead, evidence suggests that Sin Nombre virus has multiple reservoir host species, especially in light of the taxonomic union of Sin Nombre and New York viruses. More broadly, the *Peromyscus* genus is likely not monophyletic (54–56), indicating the need for an expanded survey of what is known about viral pathogens in other North American rodent species. In this respect, we lack an up-to-date, comprehensive analysis of the evidence for rodent

reservoirs of hantaviruses that is consistent with leading frameworks for assessing future zoonotic disease risk (57).

Hantavirus pulmonary syndrome and other diseases spread via zoonotic viruses pose an ongoing risk to human health in the United States and surrounding areas. Rodents within the *P. maniculatus* species complex have been identified as hosts for the viruses that cause HPS and other human diseases, but the recent scientific literature on virus prevalence in this group of species has many gaps, both in terms of the types of viruses being sampled and the geographic regions in which sampling is occurring. Viral sampling has been uneven relative to the known human HPS cases. Taxonomic changes in the *P. maniculatus* species complex have large (and quantifiable) impacts on our knowledge of especially SNV prevalence and thus also HPS risk. Therefore, future work is needed to establish a systematic framework for sampling wild rodent hosts of hantaviruses, both to optimize the allocation of resources and identify regions and populations that pose disproportionate threats to human health.

4. Materials and Methods

We conducted a systematic review of existing literature reporting observations of zoonotic viruses in the *Peromyscus maniculatus* species complex. The focal taxon was *P. maniculatus* sensu lato, and the outcomes were positive or negative test results using a range of detection methods. Examples of this include antibody tests, Polymerase Chain Reaction (PCR) tests, and genetic sequencing. Three databases, Scopus, PubMed, and PubMed Central, were used to find existing scientific literature on this topic. Search queries were conducted in August 2022 with the common and scientific names of *P. maniculatus* as well as nomenclatural synonyms and the keywords “virus”, “viral”, and “viruses” (Table 1). Through these parameters, 448 papers were identified as potential candidates for the systematic review.

We also added 6 ad hoc papers from other sources. Three of these (58–60) were found when we ran a supplementary search in all three databases at the reviewers suggestion to identify any papers that used the common name “deer mice,” “deermice,” “deer mouse,” or “deermouse.” To avoid duplication of effort, we constructed these searches to return only papers that used either common name but did not mention “Peromyscus maniculatus.” For example, the PubMed search query was: “(virus OR viral OR viruses) AND (“Deermouse” OR “Deermice” OR “Deer mouse” OR “Deer Mice”) NOT (“Peromyscus maniculatus”).”

Database	**Search Query Used**	**Search Settings Used**
PubMed	(virus OR viral OR viruses) AND ("Peromyscus maniculatus" OR "Peromyscus gracilis" OR "Peromyscus bairdi" OR "Peromyscus abietorium" OR "Peromyscus nubiterrae" OR "Peromyscus argentatus" OR "Peromyscus eremus" OR "Peromyscus anticostiensis" OR "Peromyscus plumbeus" OR "Peromyscus gambelii" OR "Peromyscus labecula" OR "Peromyscus arcticus" OR "Peromyscus sonoriensis" OR "Eastern Deermouse" OR "North American Deermouse")	Title/Abstract
PubMed Central		Abstract OR Body-All Words OR Title
Scopus	"Peromyscus maniculatus" OR "Peromyscus gracilis" OR "Peromyscus bairdi" OR "Peromyscus abietorium" OR "Peromyscus nubiterrae" OR "Peromyscus argentatus" OR	TITLE-ABSTRACT-KEY

	"Peromyscus eremus" OR "Peromyscus anticostiensis" OR "Peromyscus plumbeus" OR "Peromyscus gambelii" OR "Peromyscus labecula" OR "Peromyscus arcticus" OR "Peromyscus sonoriensis" OR "Eastern Deermouse" OR "North American Deermouse" AND virus OR viral OR viruses	

Table 1. Search terms used in conducting the systematic review of existing literature across three separate databases. Common names as well as scientific names for species within the *P. maniculatus* species complex were included.

We used the CADIMA web tool to review articles for inclusion (61). We uploaded lists of identified papers from each source to CADIMA, after which we removed duplicates, resulting in 253 remaining papers. Each paper was manually reviewed by AF for inclusion, and a second person (BS) was consulted on unclear cases. Inclusion criteria were that the article must report new primary data about the results of testing wild *P. maniculatus* for virus occurrence or infection. For reasons of scope, we excluded lab studies based on artificial infection experiments as well as papers published prior to 2000. We excluded 192 papers for one or more of the following reasons: not having a full text available in English through interlibrary loan or subscription services at the author's home institution, not meeting the inclusion criteria, or being published before the cutoff period. This left 49 papers for data extraction from the systematic review process (Figure 7).

[FIGURE 7 GOES HERE]

Data extraction entailed reading each scientific paper in full and extracting relevant information on sampling time, sampling location, number of hosts, various host identifiers, detection method used, material sampled, and the results of observations into a spreadsheet (Table S3). Information on all species captured and sampled was recorded for future use, though this study focuses solely on detection presented in *P. maniculatus* (Table S2). Summary results for *P. maniculatus* are available in Table S4.

To visually represent the patterns discovered in the extracted data, we created a series of maps using the Tableau software program (62). Spatially coded results were taken from the data extraction spreadsheet and translated into the maps on either the county, province, or state level, depending on whether the sample was taken in the United States, Canada, or Mexico, respectively.

Sampling effort was represented by the number of observations, which are defined as the number of times a unique host was tested for a unique virus using a specific method. Hence, if an individual rodent was tested by ELISA for both Whitewater Arroyo virus (WWAV) and Amapari virus (AMAV), this was recorded as two separate observations. Similarly, there would be multiple observations reported if an individual rodent was tested for the same virus multiple times in a recapture study. This way of individuating observations therefore provides a more fine-grained basis for collecting and analyzing test results. In the case of counties that were the site of multiple studies, the observations across each study were summed to create a single value which was then reflected in the map.

In the rare instance where one county was sampled for multiple virus species, the higher-order classification of virus species was used to create the maps. For example, if one study

reported sampling for SNV and a separate study conducted in the same county tested for hantavirus antibodies but not for a specific virus species within that family, the county was represented as a hantavirus county on the map. However, all results for each type of virus tested are available in Table S2.

Virus prevalence was determined as the percentage of antibody-positive *P. maniculatus* reported by each study. For sampling locales with multiple studies and thus multiple prevalence results to consider, the reported prevalence results were taken and weighted by sampling effort in order to avoid over-representing results from studies that tested a smaller number of rodents. The county estimates were calculated according to the following equation:

$$P_c = \Sigma w_s \times p_s$$

The weights w_s are the number of rodents tested for that pathogen in each study divided by the sum of all rodents tested. Thus, the weights relative to a county sum to 1. The weight for each study is multiplied by the seroprevalence proportion p_s from each study, calculated as the number of positive tests divided by the total (63). Weighting is important to account for the relative precision of different studies in estimating the mean prevalence of the pathogen in the host population.

To assess the richness and distribution of the viruses detected across the selected literature, we processed the data into cumulative measures of sampling effort by virus type and geographic location. Geographic regions were analyzed at the state level for the United States and the country level for any other sampling locales. To track how sampling effort has been distributed over time, the relationship between the number of observations reported and (a) the number of viruses identified and (b) the number of unique locations sampled were also examined.

To assess the feasibility of accurately distinguishing SNV from other hantaviruses using DNA sequence data, we downloaded 79 sequences of hantavirus with *Peromyscus maniculatus* as the host from the ZOVER database on 4 Nov 2024. We then aligned those empirical sequences to verified NCBI RefSeq sequences from 28 hantavirus species (downloaded 8 Nov 2024), separately partitioning the M, S, and L segments for alignment. Alignments were made using MAFFT v7.490 software (64) and maximum likelihood trees were constructed using RAxML v8.2.12 (65), specifying the GAMMA+CAT model performing 100 bootstrap replicates.

Acknowledgements

Our thanks to DeeAnn Reeder, Jorrit Poelen, and Donat Agosti for sharing their insights and suggestions on data extraction and synthesis methods. We'd also like to recognize several other Arizona State University undergraduates who have assisted this project along the way, including Linnea Donovan, Gwen Calaro, and Tanishq Jain. The project stemmed from Arizona State University President's Special Initiative Funds to BS and has been supported by National Institutes of Health grants 5R01GM152813-02 and 1R21AI164268-01/02 to BS and NU.

References

1. Bedford NL, Hoekstra HE. Peromyscus mice as a model for studying natural variation. eLife. 2015 Jun 17;4:e06813.

2. Mammal Diversity Database. Peromyscus maniculatus (ASM Mammal Diversity Database #1002307). 2024. Available from: https://www.mammaldiversity.org/taxon/1002307

3. Mammal Diversity Database. Mammal Diversity Database. Zenodo; 2024. Available from: https://zenodo.org/records/10595931

4. Burgin CJ, Colella JP, Kahn PL, Upham NS. How many species of mammals are there? J Mammal. 2018 Feb 1;99(1):1–14.

5. Shorter KR, Crossland JP, Webb D, Szalai G, Felder MR, Vrana PB. *Peromyscus* as a Mammalian Epigenetic Model. Genet Res Int. 2012 Mar 7;2012:e179159.

6. Vrana PB, Shorter KR, Szalai G, Felder MR, Crossland JP, Veres M, et al. Peromyscus (deer mice) as developmental models. WIREs Dev Biol. 2014;3(3):211–30.

7. Milholland MT, Castro-Arellano I, Suzán G, Garcia-Peña GE, Lee TE, Rohde RE, et al. Global Diversity and Distribution of Hantaviruses and Their Hosts. EcoHealth. 2018 Mar 1;15(1):163–208.

8. Guralnick R, Hantak MM, Li D, McLean BS. Body size trends in response to climate and urbanization in the widespread North American deer mouse, Peromyscus maniculatus. Sci Rep. 2020 Jun 1;10(1):8882.

9. Boria RA, Blois JL. Phylogeography within the Peromyscus maniculatus species group: Understanding past distribution of genetic diversity and areas of refugia in western North America. Mol Phylogenet Evol. 2023 Mar 1;180:107701.

10. Chapman OS, McLean BS. Seasonal and sex-specific changes in the gastrointestinal tracts of Peromyscus maniculatus. J Mammal. 2023 Dec 1;104(6):1364–76.

11. Chen RX, Gong HY, Wang X, Sun MH, Ji YF, Tan SM, et al. Zoonotic Hantaviridae with Global Public Health Significance. Viruses. 2023 Aug;15(8):1705.

12. Nichol ST, Spiropoulou CF, Morzunov S, Rollin PE, Ksiazek TG, Feldmann H, et al. Genetic Identification of a Hantavirus Associated with an Outbreak of Acute Respiratory Illness. Science. 1993 Nov 5;262(5135):914–7.

13. Childs JE, Colby L, Krebs JW, Strine T, Feller M, Noah D, et al. Surveillance and Spatiotemporal Associations of Rabies in Rodents and Lagomorphs in the United States, 1985-1994. J Wildl Dis. 1997 Jan;33(1):20–7.

14. Ubelaker J, Rácz G. A Checklist of Parasites of Peromyscus maniculatus in North America. MANTER J Parasite Biodivers [Internet]. 2022 Aug 2 [cited 2024 May 23]; Available from: https://digitalcommons.unl.edu/manter/20/

15. Postler T, Bradfute S, Calisher C, Klingström J, Laenen L, Maes P, et al. Rename all species in the family to comply with the ICTV-mandated binomial format (Bunyavirales: Hantaviridae). International Committee on the Taxonomy of Viruses; 2021. Report No.: 2021.013M. Available from: https://ictv.global/ictv/proposals/2021.013M.Hantaviridae_sprename.zip

16. Centers For Disease Control. Reported Cases of Hantavirus Disease. 2023. Available from: https://web.archive.org/web/20240307003345/https://www.cdc.gov/hantavirus/surveillance/index.html

17. MacNeil A, Ksiazek TG, Rollin PE. Hantavirus Pulmonary Syndrome, United States, 1993–2009. Emerg Infect Dis. 2011 Jul;17(7):1195–201.

18. Vigueras-Galván AL, López-Pérez AM, García-Peña GE, Rico-Chávez O, Sarmiento-Silva RE, Suzán G. Current Situation and Perspectives on Hantaviruses in Mexico. Viruses. 2019 Jul;11(7):642.

19. Avšič-Županc T, Saksida A, Korva M. Hantavirus infections. Clin Microbiol Infect. 2019 Apr 1;21:e6–16.

20. Kuhn JH, Schmaljohn CS. A Brief History of Bunyaviral Family Hantaviridae. Diseases. 2023 Mar;11(1):38.

21. Mills JN, Amman BR, Glass GE. Ecology of Hantaviruses and Their Hosts in North America. Vector-Borne Zoonotic Dis. 2010 Aug;10(6):563–74.

22. Mull N, Carlson CJ, Forbes KM, Becker DJ. Virus isolation data improve host predictions for New World rodent orthohantaviruses. J Anim Ecol. 2022;91(6):1290–302.

23. Dragoo JW, Lackey JA, Moore KE, Lessa EP, Cook JA, Yates TL. Phylogeography of the deer mouse (Peromyscus maniculatus) provides a predictive framework for research on hantaviruses. J Gen Virol. 2006;87(7):1997–2003.

24. Bradley R, Lindsey L. Mitochondrial DNA sequence data indicate evidence for multiple species within Peromyscus maniculatus. Spec Publ Mus Tex Tech Univ. 2019;

25. Greenbaum IF, Honeycutt RL, Chirhart SE. Taxonomy and Phylogenetics of the Peromyscus maniculatus Species Group. Spec Publ Mus Tex TECH Univ. 2019;

26. Griffin, Chan, M., Tailor, N., Mendoza, E. J., Leung, A., Warner, B. M., et al. SARS-CoV-2 infection and transmission in the North American deer mouse. Nature Communications. 2021;12(1):3612.

27. Ebel GD. Update on Powassan Virus: Emergence of a North American Tick-Borne Flavivirus. Annu Rev Entomol. 2010 Jan 7;55(Volume 55, 2010):95–110.

28. Reignier T, Oldenburg J, Flanagan ML, Hamilton GA, Martin VK, Cannon PM. Receptor use by the Whitewater Arroyo virus glycoprotein. Virology. 2008 Feb 20;371(2):439–46.

29. Upham NS, Poelen JH, Paul D, Groom QJ, Simmons NB, Vanhove MPM, et al. Liberating host–virus knowledge from biological dark data. Lancet Planet Health. 2021 Oct 1;5(10):e746–50.

30. Guala GF. The Importance of Species Name Synonyms in Literature Searches. PLOS ONE. 2016 Sep 14;11(9):e0162648.

31. Remsen D. The use and limits of scientific names in biological informatics. ZooKeys. 2016 Jan 7;550:207–23.

32. Kuhn J, Bradfute S, Calisher C, Klempa B, Klingström J, Laenen L, et al. Reevaluate and reorganize family Hantaviridae (order Bunyavirales). International Committee on the Taxonomy of Viruses; 2023. Report No.: 2023.035M.

33. Biggs JR, Bennett KD, Torrez-Martinez N, Hjelle BL. Sin Nombre Virus Antibody Prevalence in Rodents of North-Central New Mexico. Southwest Nat. 2000;45(1):61–6.

34. Danforth ME, Messenger S, Buttke D, Weinburke M, Carroll G, Hacker G, et al. Long-Term Rodent Surveillance after Outbreak of Hantavirus Infection, Yosemite National Park, California, USA, 2012. Emerg Infect Dis. 2020 Mar;26(3):560–7.

35. Childs JE, Ksiazek TG, Spiropoulou CF, Krebs JW, Morzunov S, Maupin GO, et al. Serologic and Genetic Identification of Peromyscus maniculatus as the Primary Rodent Reservoir for a New Hantavirus in the Southwestern United States. J Infect Dis. 1994;169(6):1271–80.

36. Deardorff ER, Nofchissey RA, Cook JA, Hope AG, Tsvetkova A, Talbot SL, et al. Powassan Virus in Mammals, Alaska and New Mexico, USA, and Russia, 2004–2007. Emerg Infect Dis. 2013 Dec;19(12):2012–6.

37. Charrel RN, de Lamballerie X. Zoonotic aspects of arenavirus infections. Vet Microbiol. 2010 Jan 27;140(3):213–20.

38. Pierson TC, Diamond MS. The continued threat of emerging flaviviruses. Nat Microbiol. 2020 Jun;5(6):796–812.

39. Lyubsky S, Gavrilovskaya I, Luft B, Mackow E. Histopathology of Peromyscus leucopus naturally infected with pathogenic NY-1 hantaviruses: pathologic markers of HPS viral infection in mice. Lab Investig J Tech Methods Pathol. 1996 Mar;74(3):627–33.

40. Luong LT, Vigliotti BA, Campbell S, Comer JA, Mills JN, Hudson PJ. Dynamics of Hantavirus Infection in *Peromyscus leucopus* of Central Pennsylvania. Vector-Borne Zoonotic Dis. 2011 Nov;11(11):1459–64.

41. Mlera L, Meade-White K, Saturday G, Scott D, Bloom ME. Modeling Powassan virus infection in Peromyscus leucopus, a natural host. Ebel GD, editor. PLoS Negl Trop Dis. 2017 Jan 31;11(1):e0005346.

42. Calisher CH, Wagoner KD, Amman BR, Root JJ, Douglass RJ, Kuenzi AJ, et al. Demographic Factors Associated with Prevalence of Antibody to Sin Nombre Virus in Deer Mice in the Western United States. J Wildl Dis. 2007 Jan 1;43(1):1–11.

43. Previtali MA, Lehmer EM, Pearce-Duvet JMC, Jones JD, Clay CA, Wood BA, et al. Roles of human disturbance, precipitation, and a pathogen on the survival and reproductive probabilities of deer mice. Ecology. 2010 Feb;91(2):582–92.

44. Carver S, Kuenzi A, Bagamian KH, Mills JN, Rollin PE, Zanto SN, et al. A temporal dilution effect: hantavirus infection in deer mice and the intermittent presence of voles in Montana. Oecologia. 2011 Jul;166(3):713–21.

45. Laenen L, Vergote V, Calisher CH, Klempa B, Klingström J, Kuhn JH, et al. Hantaviridae: Current Classification and Future Perspectives. Viruses. 2019 Sep;11(9):788.

46. Hjelle B, Lee SW, Song W, Torrez-Martinez N, Song JW, Yanagihara R, et al. Molecular linkage of hantavirus pulmonary syndrome to the white-footed mouse, Peromyscus leucopus: genetic characterization of the M genome of New York virus. J Virol. 1995 Dec;69(12):8137–41.

47. Ashford RW. What it takes to be a reservoir host. Belg J Zool. 1997;127(Supplement 1):85–90.

48. Haydon DT. Identifying Reservoirs of Infection: A Conceptual and Practical Challenge. Emerg Infect Dis. 2002 Dec;8(12):1468–73.

49. Go YY, Balasuriya UBR, Lee C kyo. Zoonotic encephalitides caused by arboviruses: transmission and epidemiology of alphaviruses and flaviviruses. Clin Exp Vaccine Res. 2014 Jan;3(1):58–77.

50. Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. Proc Natl Acad Sci. 2015 Jun 2;112(22):7039–44.

51. Kuno G, Mackenzie JS, Junglen S, Hubálek Z, Plyusnin A, Gubler DJ. Vertebrate Reservoirs of Arboviruses: Myth, Synonym of Amplifier, or Reality? Viruses. 2017 Jul;9(7):185.

52. Forbes KM, Sironen T, Plyusnin A. Hantavirus maintenance and transmission in reservoir host populations. Curr Opin Virol. 2018 Feb 1;28:1–6.

53. Salkeld D, Hopkins S, Hayman D. Reservoir hosts. In: Salkeld D, Hopkins S, Hayman D, editors. Emerging Zoonotic and Wildlife Pathogens: Disease Ecology, Epidemiology, and Conservation. Oxford University Press; 2023. Available from: https://doi.org/10.1093/oso/9780198825920.003.0007

54. Platt RN II, Amman BR, Keith MS, Thompson CW, Bradley RD. What Is Peromyscus? Evidence from nuclear and mitochondrial DNA sequences suggests the need for a new classification. J Mammal. 2015 Aug 3;96(4):708–19.

55. Steppan SJ, Schenk JJ. Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. PLOS ONE. 2017 Aug 16;12(8):e0183070.

56. Sullivan KAM, Platt RN, Bradley RD, Ray DA. Whole mitochondrial genomes provide increased resolution and indicate paraphyly in deer mice. BMC Zool. 2017 Aug 18;2(1):11.

57. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, et al. Pathways to zoonotic spillover. Nat Rev Microbiol. 2017 Aug;15(8):502–10.

58. Douglass RJ, Kuenzi AJ, Williams CY, Douglass SJ, Mills JN. Removing Deer Mice from Buildings and the Risk for Human Exposure to Sin Nombre Virus. Emerg Infect Dis. 2003 Mar;9(3):390–2.

59. Torres-Pérez F, Wilson L, Collinge SK, Harmon H, Ray C, Medina RA, et al. Sin Nombre Virus Infection in Field Workers, Colorado, USA. Emerg Infect Dis. 2010 Feb;16(2):308–10.

60. Núñez JJ, Fritz CL, Knust B, Buttke D, Enge B, Novak MG, et al. Hantavirus Infections among Overnight Visitors to Yosemite National Park, California, USA, 2012. Emerg Infect Dis. 2014 Mar;20(3):386–93.

61. Kohl C, McIntosh EJ, Unger S, Haddaway NR, Kecke S, Schiemann J, et al. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. Environ Evid. 2018 Feb 1;7(1):8.

62. Tableau Version 23.2.9. 2023. Available from: https://www.tableau.com/

63. Bergeri I, Whelan MG, Ware H, Subissi L, Nardone A, Lewis HC, et al. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: A systematic review and meta-analysis of standardized population-based studies. PLOS Med. 2022 Nov 10;19(11):e1004107.

64. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013 Apr 1;30(4):772–80.

65. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014 May 1;30(9):1312–3.

66. Warner BM, Dowhanik S, Audet J, Grolla A, Dick D, Strong JE, et al. Hantavirus Cardiopulmonary Syndrome in Canada. Emerg Infect Dis. 2020 Dec;26(12):3020–4.

Figures

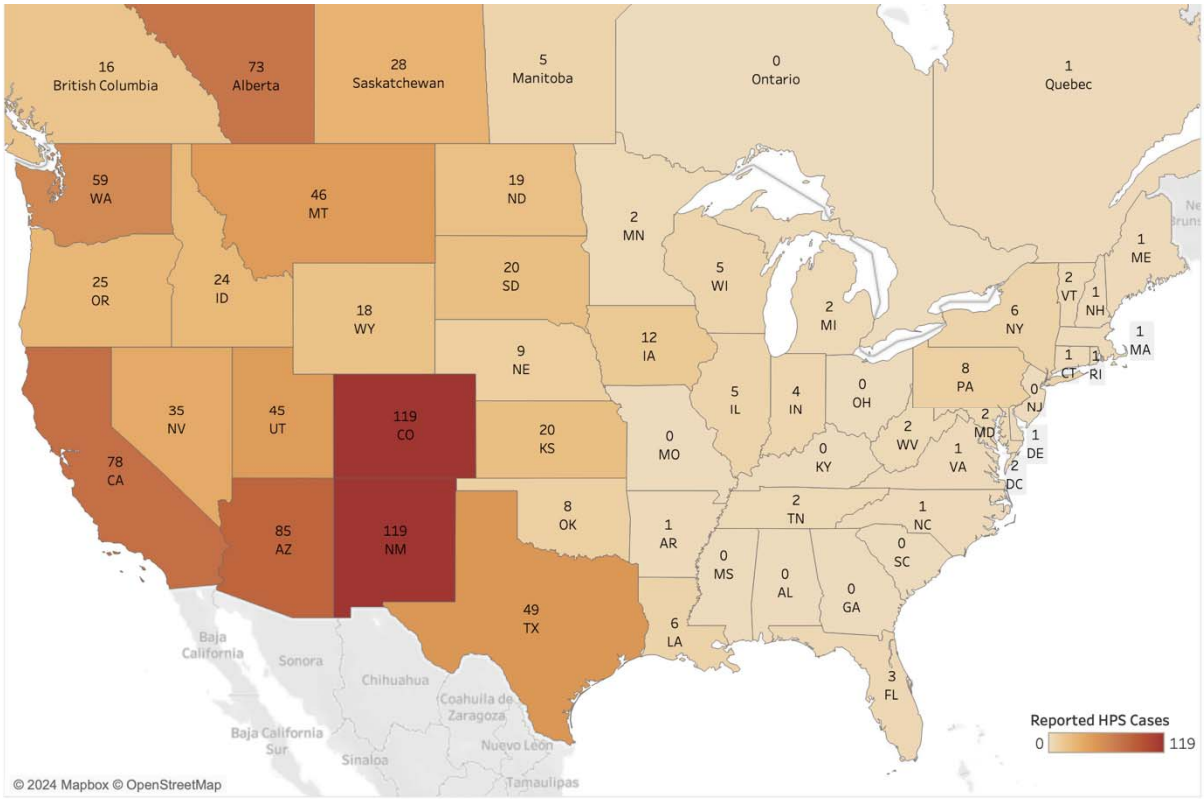

Figure 1. Map of hantavirus pulmonary syndrome (HPS; also called hantavirus cardiopulmonary syndrome disease) human cases reported from 1993 to 2021 in the U.S. and to 2020 in Canada (16,66).

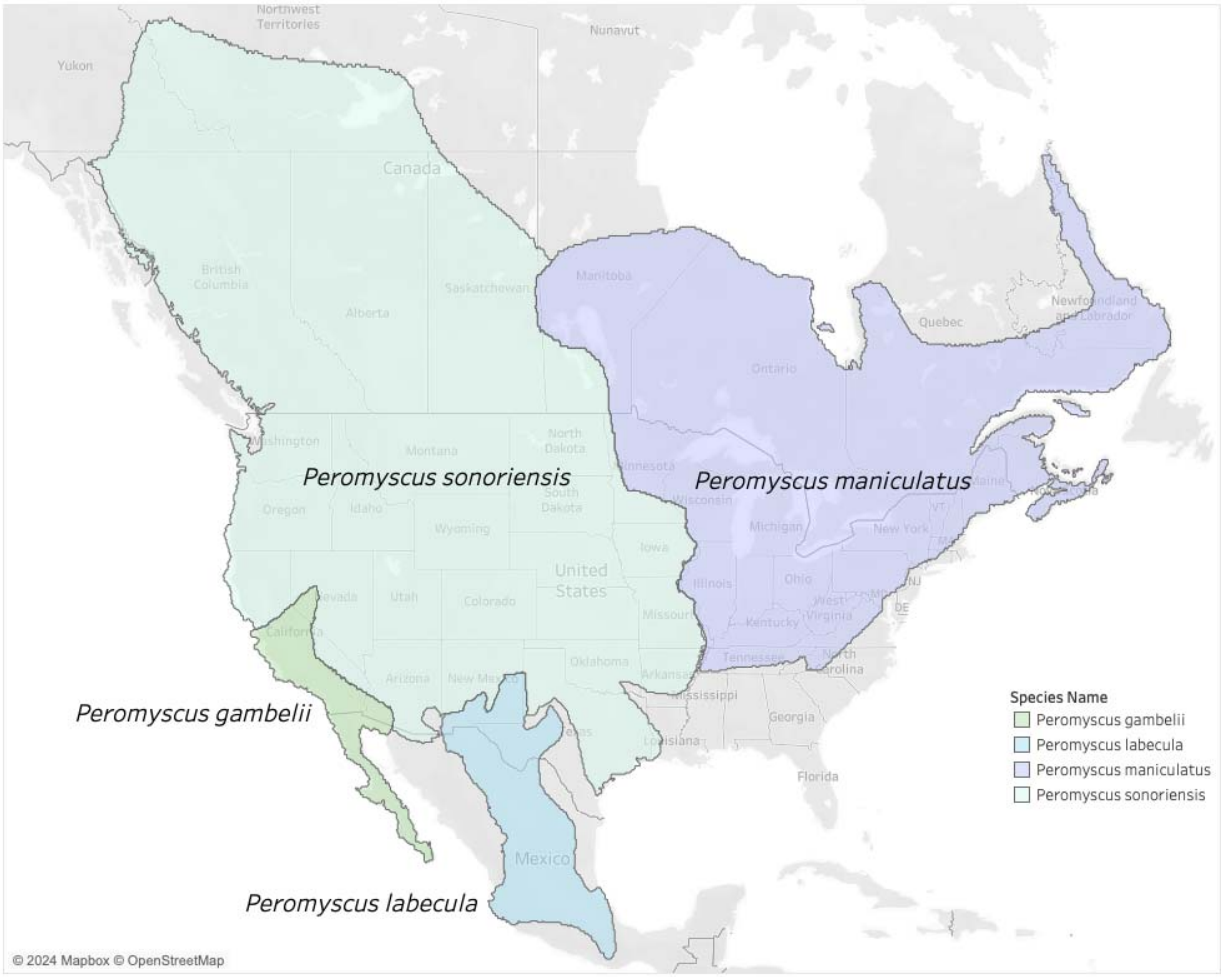

Figure 2. Proposed taxonomic split of the *Peromyscus maniculatus* species complex (24,25) into four new species as recognized by the Mammal Diversity Database (3). Note that the name *P. maniculatus* sensu stricto applies east of the Mississippi River in the U.S. and *P. sonoriensis* applies to most of the Southwest where the initial outbreak of HPS was first detected.

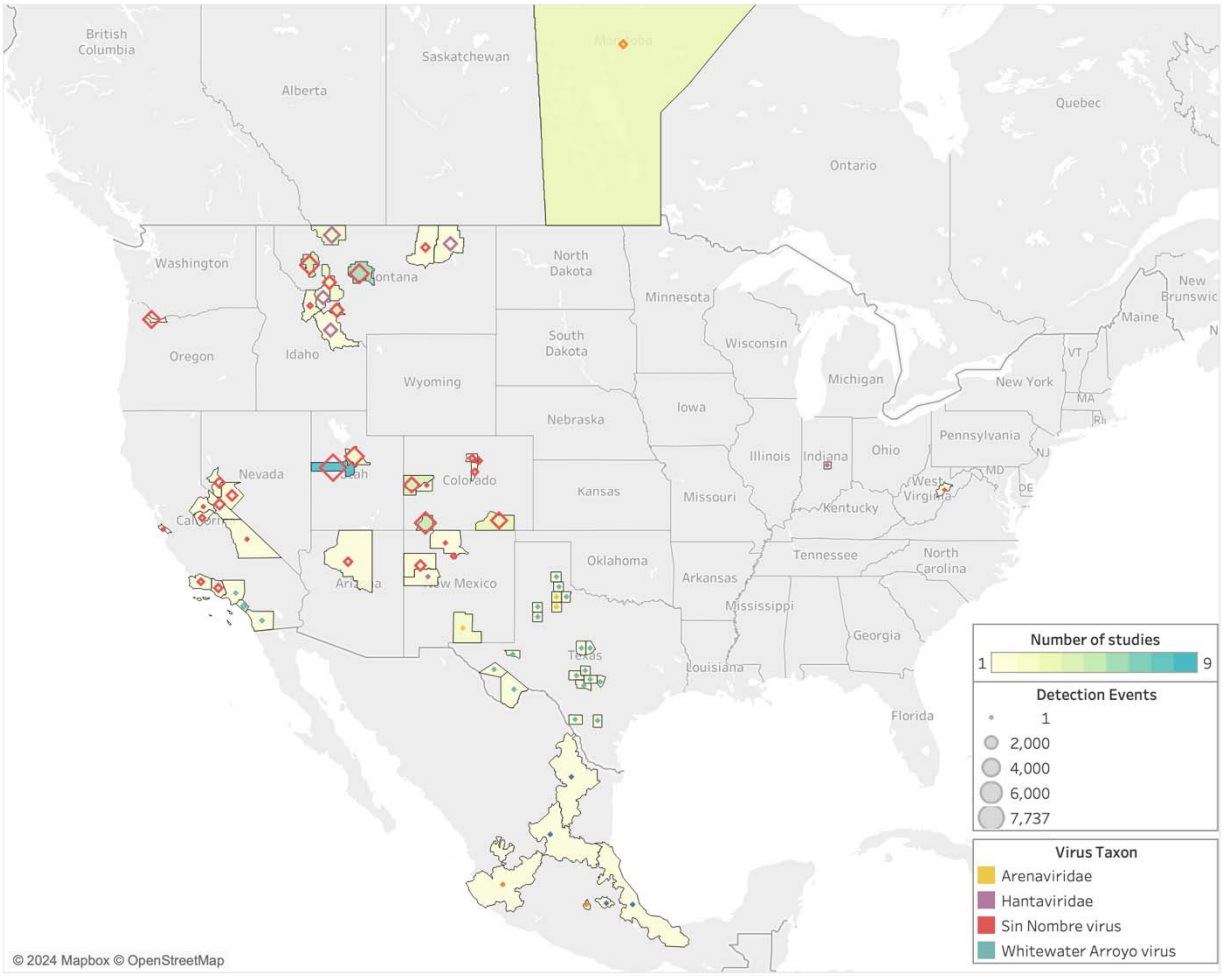

Figure 3. Geographic distribution of sampling effort and zoonotic virus type tested for in *P. maniculatus* field studies. Legend refers to the number of observations in each county, province, or state, where observations are defined as the number of unique hosts tested for a unique virus. Hantavirus and arenavirus refer to studies in which a specific virus species was not identified.

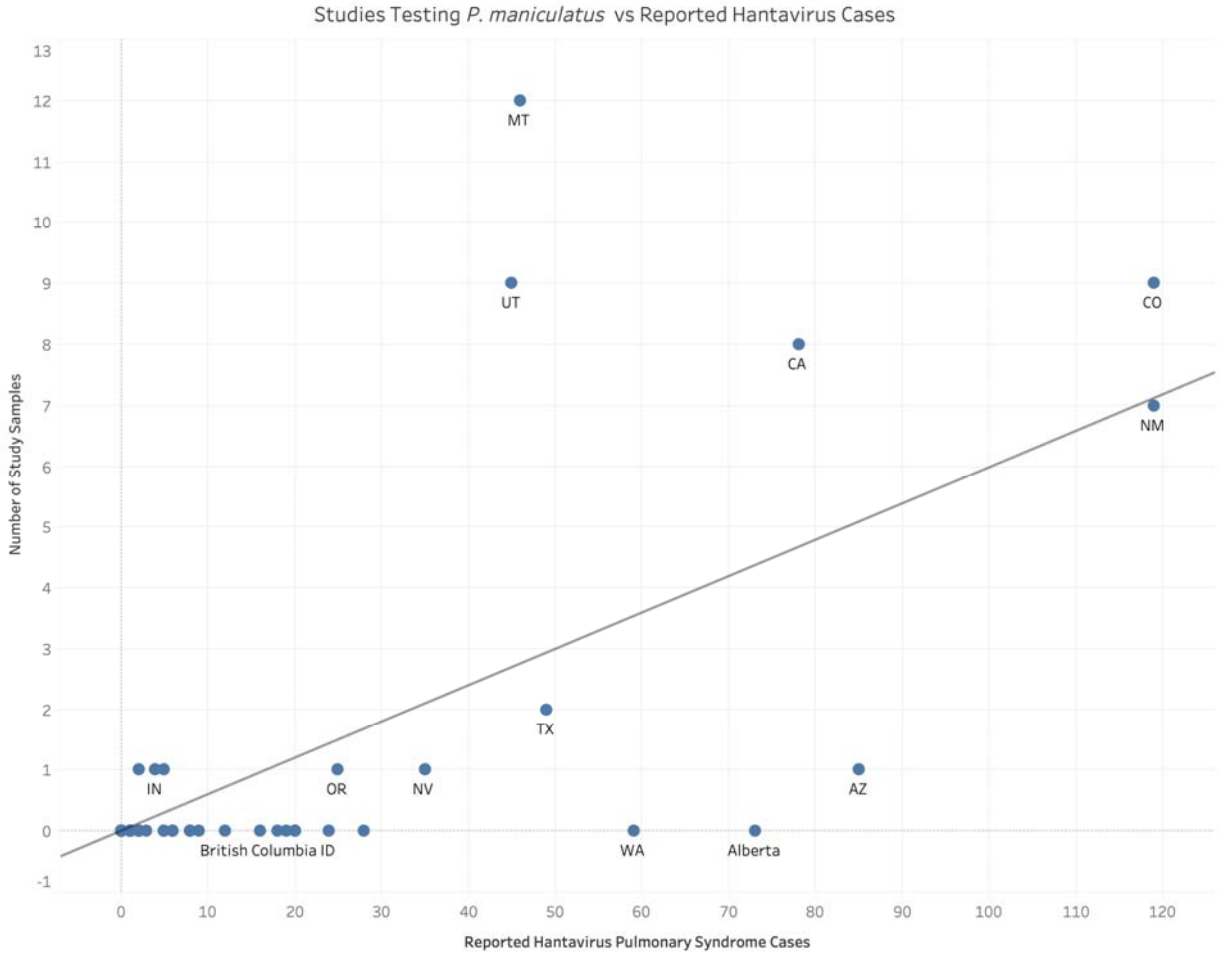

Figure 4: Comparison of field studies published since 2000 in which testing for viruses in *P. maniculatus* was conducted against the total number of reported human hantavirus cases across states and Canadian provinces.

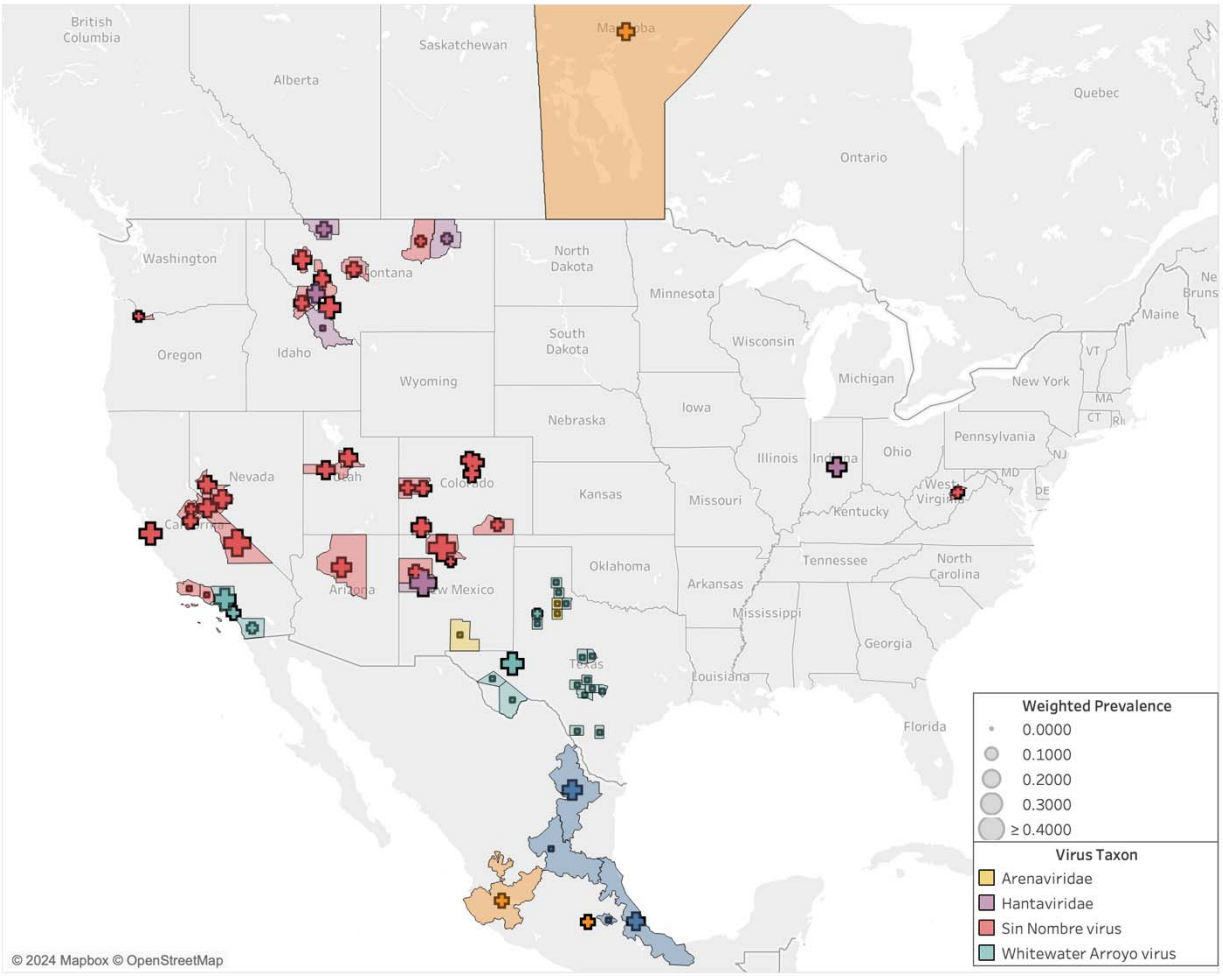

Figure 5. Seroprevalence for viruses studied in *P. maniculatus* by United States county, Canadian province, and Mexican state. In areas in which there was more than one detection study conducted, the average seroprevalence across studies was weighted by sampling effort.

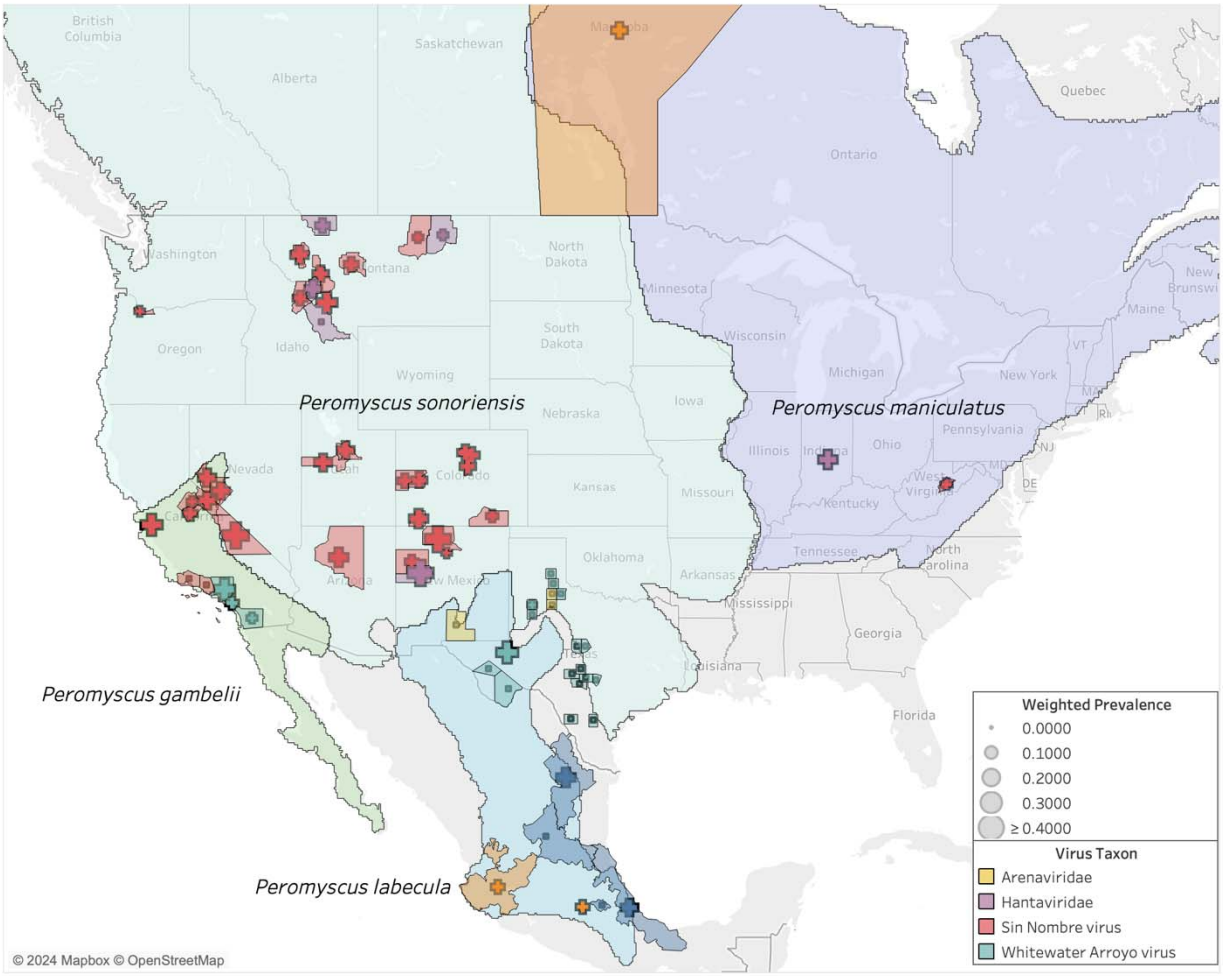

Figure 6. Seroprevalence compared to host ranges in the *P. maniculatus* species complex.

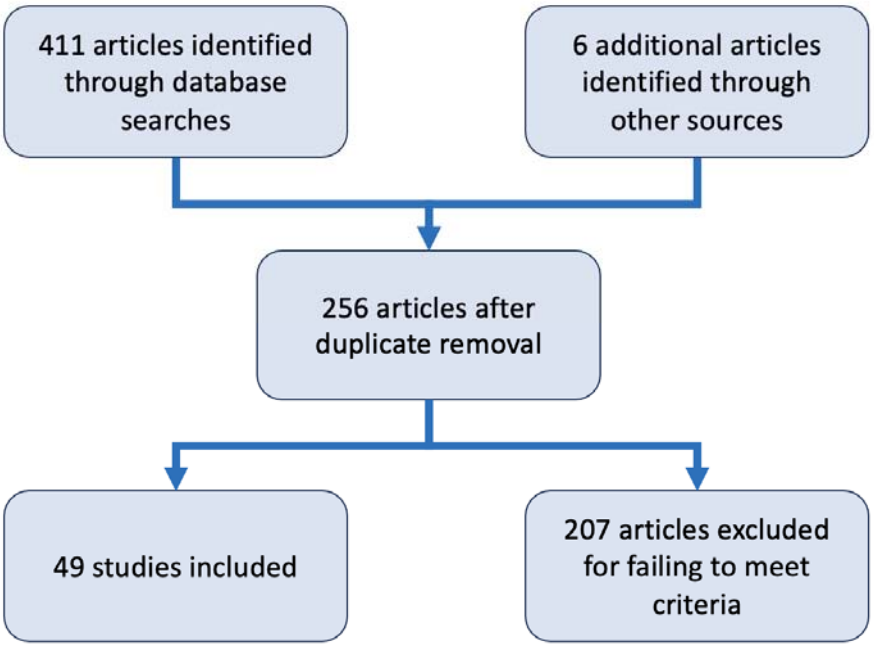

Figure 7. Diagram summarizing the steps of the systematic review done in CADIMA that led to the final set of 49 that were included.