Multiple spillovers and onward transmission of SARS-Cov-2 in free-living and captive White-tailed deer (*Odocoileus virginianus*)

Authors: Suresh V. Kuchipudi,[1#] Meera Surendran-Nair[1], Rachel M. Ruden[2,3], Michelle Yon[4], Ruth H. Nissly[1], Rahul K. Nelli[3], Lingling Li[4], Bhushan M. Jayarao[4], Kurt J. Vandegrift[5], Costas D. Maranas[6], Nicole Levine[7], Katriina Willgert[8], Andrew J. K. Conlan[8], Randall J. Olsen[9,10], James J. Davis[11], James M. Musser[9,10], Peter J. Hudson[5], and Vivek Kapur[7#]

[1]*Animal Diagnostic Laboratory and Huck Institutes of Life Sciences, Penn State, University Park, PA 16802, USA*
[2]*Wildlife Bureau, Iowa Department of Natural Resources, Des Moines, Iowa, USA,*
[3]*Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, Iowa, USA,*
[4]*Animal Diagnostic Laboratory, Department of Veterinary and Biomedical Sciences, The Pennsylvania State University, PA,16802, USA*
[5]*The Center for Infectious Disease Dynamics, Department of Biology and Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, 16802, USA*
[6]*Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, 16802, USA*
[7]*Department of Animal Science and Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, 16802, USA*
[8]*Disease Dynamics Unit (DDU), Department of Veterinary Medicine, University of Cambridge, UK,*
[9]*Laboratory of Molecular and Translational Human Infectious Disease Research, Center for Infectious Diseases, Department of Pathology and Genomic Medicine, Houston Methodist Research Institute and Houston Methodist Hospital, Houston, TX 77030, USA*
[10]*Departments of Pathology and Laboratory Medicine and Microbiology and Immunology, Weill Cornell Medical College, NY 10021, USA*
[11]*University of Chicago Consortium for Advanced Science and Engineering, University of Chicago and Division of Data Science and Learning, Argonne National Laboratory, Argonne, Illinois, USA,*

Address correspondence to: Suresh V. Kuchipudi, E-mail: skuchipudi@psu.edu and Vivek Kapur, E-mail: vkapur@psu.edu

Abstract

Many animal species are susceptible to SARS-CoV-2 and could potentially act as reservoirs, yet transmission in non-human free-living animals has not been documented. White-tailed deer (*Odocoileus virginianus)*, the predominant cervid in North America, are susceptible to SARS-CoV-2 infection, and experimentally infected fawns transmit the virus to other captive deer. To test the hypothesis that SARS-CoV-2 may be circulating in deer, we evaluated 283 retropharyngeal lymph node (RPLN) samples collected from 151 free-living and 132 captive deer in Iowa from April 2020 through December of 2020 for the presence of SARS-CoV-2 RNA. Ninety-four of the 283 deer (33.2%; 95% CI: 28, 38.9) samples were positive for SARS-CoV-2 RNA as assessed by RT-PCR. Notably, between Nov 23, 2020 and January 10, 2021, 80 of 97 (82.5%; 95% CI 73.7, 88.8) RPLN samples had detectable SARS-CoV-2 RNA by RT-PCR. Whole genome sequencing of the 94 positive RPLN samples identified 12 SARS-CoV-2 lineages, with B.1.2 (n = 51; 54.5%), and B.1.311 (*n* = 19; 20%) accounting for ~75% of all samples. The geographic distribution and nesting of clusters of deer and human lineages strongly suggest multiple zooanthroponotic spillover events and deer-to-deer transmission. The discovery of sylvatic and enzootic SARS-CoV-2 transmission in deer has important implications for the ecology and long-term persistence, as well as the potential for spillover to other animals and spillback into humans. These findings highlight an urgent need for a robust and proactive "One Health" approach to obtaining a better understanding of the ecology and evolution of SARS-CoV-2.

One-Sentence Summary: SARS-CoV-2 was detected in one-third of sampled White-tailed deer in Iowa between September 2020 and January of 2021 that likely resulted from multiple human-to-deer spillover events and deer-to-deer transmission.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the cause of coronavirus disease 2019 (COVID-19) in humans, is a novel coronavirus in the genus *Betacoronavirus* (subgenus *Sarbecovirus*) (1). SARS-CoV-2 was first identified in Wuhan, China, toward the end of 2019 (2), and has caused a pandemic with 246 million COVID-19 cases and 4.9 million deaths globally as of October 23, 2021 (3). The virus continues to evolve with growing concern for the emergence of new variants. SARS-CoV-2 uses the host angiotensin-converting enzyme 2 (ACE-2) receptor to enter cells (4). ACE-2 receptors are well conserved in multiple vertebrate species including humans (5). Unsurprisingly, computational analyses predict high binding affinities of SARS-CoV-2 to ACE-2 in multiple animal species indicating that they could be susceptible to infection (5). Amongst these, three species of cervids, the Père David's deer (*Elaphurus davidianus*), reindeer (*Rangifer tarandus*), and white-tailed deer (*Odocoileus virginianus*) (6).

The widespread and global dissemination of SARS-CoV-2 among humans provides opportunities for spillovers into non-human hosts (7). Indeed, SARS□CoV□2 infections have been documented in dogs, cats, zoo animals (i.e. tigers and lions) and farmed mink (8, 9). In principle, SARS-CoV-2 infection of an animal host could result in it becoming a reservoir that drives the emergence of new variants with risk of spillback to humans. This type of transmission cycle has been described to occur among workers on mink farms (10). However, widespread SARS-CoV-2 transmission in a free-living animal species has not yet been documented.

Our study was prompted by a recent report that 40% of free-living white-tailed deer in the USA had antibodies against SARS-CoV-2 (11). In addition, there is limited evidence of SARS-CoV-2 transmission among experimentally infected deer in controlled captive conditions (6). To test the hypothesis that infection and subsequent transmission of SARS-CoV-2 of deer occurs in nature, we assayed 283 retropharyngeal lymph node (RPLN) samples collected from free-living and captive deer in Iowa from April 2020 through January 2021. We discovered that one-third of the WTD sampled had SARS-CoV-2 nucleic acid in their RPLN. We then sequenced the SARS-CoV-2 genomes present in all positive samples and found that the genomes represented multiple lineages that corresponded to viral genotypes circulating contemporaneously in humans. In the aggregate the results are consistent with a model of multiple independent human-to-deer

transmission events and deer-to-deer transmission. Our findings raise the possibility of reverse zoonoses, especially in exurban areas with high deer density. The study also highlights the potential risks and knowledge gaps associated with the continued evolution of SARS-CoV-2 in other animal hosts and call for the implementation of enhanced surveillance programs to identify potential reservoir species at the animal-human interface.

Results and Discussion

SARS-CoV-2 RNA is found in a substantial fraction free-living and captive WTD across Iowa with strong evidence of temporal clustering

A total of 283 RPLN samples recovered from wild and captive WTD in Iowa between April 8, 2020 and January 6, 2021 were analyzed by RT-PCR for the presence of SARS-CoV-2 RNA signatures using the OPTI Medical SARS-CoV-2 RT-PCR assay targeting the N1 and N2 genes in a CLIA-approved laboratory ((12); Supplementary Table 1). The sampling period closely follows the trajectory of the pandemic in Iowa (with the first human case reported on March 8, 2020) through the peak during the second week of November 2020 ((3); Figure 1).

Sources and characteristics associated with the 283 RPLN samples are summarized in Table 1. Overall, 94 of the 283 (33.2%; 95% CI: 28, 38.9) RPLN samples were found to be positive for SARS-CoV-2 RNA. Since a majority (n = 261, 91.9%) of the 283 RPLN samples in our study were harvested from deer from September through December of 2020, a period that coincided with the regular deer hunting season in Iowa that started on September 19, 2020 and ended January 10, 2021, we explored temporal trends in recovery of SARS-CoV-2 RNA from deer in the sample set. The results show that all 17 RPLNs from deer collected during April through August 2020 were negative for the presence of SARS-CoV-2 RNA (Table 1). The first positive identification of SARS-CoV-2 in deer was on September 28, 2020 in an animal harvested on a game preserve in South Eastern Iowa (Figure 2; Supplementary Table 1). This was closely followed by a second positive sample identified in a free-living deer killed in a road accident on September 30, 2020, approximately 300 miles away in Woodbury County on the Western border of the State (Supplementary Table 1). In total, two the 39 samples collected in September were positive for SARS-CoV-2 RNA (5.1%: 95% CI: 1.4, 16.9). Similarly, in October 2020, four of 70 (5.7%; 95% CI 2.2, 13.8) RPLNs recovered from WTD, one each from Des Moines and Pottawattamie counties, and two from Woodbury, were found to be positive for SARS-CoV-2

RNA. Coinciding with the major peak of infection in humans in Iowa, the positivity rate in deer rapidly increased with 22 of 77 RPLN samples (27.8%; 95% CI 19.2, 38.6) harboring SARS-CoV-2 RNA in November 2020, and 61 of 75 samples (81.3%; 95% CI: 71.1, 88.5) positive in December of 2020 (Figure 1). During the second week of January 2021 at the end of the regular hunting season, all five deer RPLN samples were positive for SARS-CoV-2 RNA (83.3%; 95% CI 43.6, 97). Notably, in the seven-week period starting Nov 23, 2020 through the end of hunting season on January 10, 2021, 80 of 97 (82.5%; 95% CI 73.7, 88.8) deer RPLN samples from across the State were positive for SARS-CoV-2 RNA by RT-PCR. Importantly, the results show a high viral copy number in deer RPLN samples (Median 106,000 viral copies per ml; ranging from 268 to 5.4 x 10^8 copies per ml) (Supplementary Figure 1), suggesting that many of the animals likely had a very high viral load.

Ecological associations and risk factors of presence SARS-CoV-2 RNA recovered from free-living and captive deer in Iowa

An exploratory analysis of potential risk factors for the identification of SARS-CoV-2 RNA in deer was carried out. No statistically significant differences (95% level) were found in the proportion of positive tests by age or sex within this sample (Table 1, Supplementary Figure 2). However, this study was not designed or powered to probe this question with rigorously so that further investigations may be warranted.

The deer included in the study were either free-living on public lands or in peri-urban environments (n = 151) or were resident in nature or hunting preserves in captive settings (n = 132). Notably, the results suggest that the proportion of free-living deer positive for SARS-CoV-2 (44.4%; 95% CI: 36.4, 52.3) was significantly higher ($z = 4.3$; $p < 0.0001$) than in the deer living within preserves / captive settings (20.4%; 95% CI 13.7, 27.4). However, these results may be confounded in that four times as many RPLN were harvested from free-living deer (n = 52) as were from captive deer (n = 13) during December 2020 when the virus positivity rate was at its peak in Iowa deer herd. Hence, further studies are needed to assess the true significance or reasons for the observed differences in prevalence between free-living deer and those in managed environments. For instance, to better assess the risk of spillover and transmission, it would be important to understand whether deer in managed settings may be less stressed, have

different nutritional status, or otherwise exhibit behaviors that influence opportunities for spillover or within-herd transmission of SARS-CoV-2 compared with those that are free-living.

Finally, we explored regional differences in observed SARS-CoV-2 positivity among deer at the county level across the State. Figure 2A shows the widespread regional distribution of positive samples recovered from deer in Iowa and illustrates the temporal growth in SARS-CoV-2 positivity as the year progressed. The study identified 10 counties with at least one positive sample (Table 1). The largest number of RPLN samples represented in the collection were from a single game preserve (Preserve 2; Fig 2B) in Southeastern Iowa. Overall, 23 of the 112 deer RPLN samples from this preserve were found to be positive for SARS-CoV-2 RNA, with the first positive in September and the second in October 2020, and 11 of 38 deer sampled in November and all 10 deer sampled in December 2020, suggesting a rapidly increasing herd-level prevalence.

Seven counties had at least 10 samples collected, with all 11 specimens from Allamakee county being found to be SARS-CoV-2 positive, as were 21 of the 28 samples collected from Appanoose county (Table 1; Supplementary Table 1). In contrast, none of the 9 samples collected from Black Hawk county were positive, and nor were the 6 RPLN samples from Henry county. While the exact reasons for this heterogeneity in PCR positive response rates are unknown, the timing of collection in relation to the SARS-CoV-2 spread in deer may play a role. For instance, the samples from Henry county were collected during April and May of 2020 during the early of the pandemic, and well before the first positive sample was identified in deer. Similarly, all 9 RPLN samples tested from Black Hawk country were collected prior to mid-November peak of reported SARS-CoV-2 cases in humans in Iowa. Together, these results suggest widespread presence of SARS-CoV-2 RNA in deer across the State of Iowa, with strong evidence of temporal clustering.

Whole genome based phylogeographic and phylogenomic analyses provide evidence of multiple likely reverse zoonotic spillover events of SARS-CoV-2 to deer and deer-to-deer transmission

To begin to understand the genomic diversity of SARS-CoV-2 associated with free-living and captive deer, we characterized the complete SARS-CoV-2 genomes from all 94 deer RPLN positive for the presence of viral RNA. A high-level of sequencing coverage was obtained, and

Pangolin version 3.1.11 (*https://github.com/cov-lineages/pangoLEARN*, last accessed October 27, 2021) was used to identify SARS-CoV-2 lineages using a previously described analysis pipeline (13-15). The automated vSNP pipeline (https://github.com/USDA-VS/vSNP) was next used to identify SNPs and construct phylogenetic trees in the context of 84 additional publicly available animal origin SARS-CoV-2 isolates as well as from 372 SARS-CoV-2 isolates identified from humans in Iowa during this same period (Supplementary Table 2). All newly sequenced SARS-CoV-2 consensus genomes from deer RPLN are deposited in GISAID and raw reads submitted to NCBI's Short Read Archive (BioProject Number PRJNA776532).

The analysis identified a total of 12 SARS-CoV-2 lineages amongst the 94 samples from deer, with two lineages, B.1.2 (n = 51; 54.5%), and B.1.311 (*n* = 19; 20%) accounting for ~ 75% of all samples. Together with the next two most abundant lineages, B.1 (*n* = 7) and B.1.234 (*n* = 6), these four lineages were geographically widely distributed and represented ~ 88% of SARS-CoV-2 circulating within deer in Iowa (Table 2). While the number of SARS-CoV-2 sequences from Iowa is not very high (only 372 sequences are available) during this period in publicly available data, it is noteworthy that the B.1.2 was also the most abundant (~ 43.5%) SARS-CoV-2 lineage circulating in humans in Iowa. In contrast, the B.1.311 lineage representing about one fifth of the isolates from deer was relatively poorly represented (~1.6%) amongst the publicly available SARS-CoV-2 from humans in Iowa (Supplementary Table 2). Conversely, the second most abundant SARS-CoV-2 lineage in humans, B.1.565, accounting for ~ 8.1% of available sequences, was not identified amongst the sampled deer. However, given the lack of representativeness of the sampling from both humans and deer, we urge caution in overinterpreting these differences in the apparent prevalence of SARS-CoV-2 lineages between deer and sympatric human hosts.

The temporal and geographic patterns of clustering of SARS-CoV-2 lineages, together with phylogenetic analyses, provide strong evidence of multiple likely zooanthroponotic spillovers from humans to deer. The results provide evidence of near simultaneous recovery of multiple SARS-CoV-2 lineages within temporally and geographically restricted deer herds at various locations throughout the State. For instance, the results show that a vast majority (22 of 23) of samples recovered from deer in a game preserve in Southeastern Iowa was represented by the B.1.2 lineage. The single outlier isolate was represented by lineage B.1.311 and was recovered

from a deer at this preserve on November 24, 2020, in synchrony with 12 additional specimens harboring the B.1.2 lineage. Similarly, in the Yellow River State Forest area in Allamakee county, during the 5 days between December 5 and December 9, all 11 hunter-killed deer were positive for the presence of SARS-CoV-2 RNA in their RPLNs. Nine of these samples were represented by the B.1.2 lineage. The two outliers, represented by B.1.311 and B.1.459 lineages were recovered from another hunter-killed deer on the same date, December 8, 2020, as was another animal harboring the B.1.2 lineage. All these deer were hunted within a few miles of each other. A third example of the near synchronous recovery of genetically distinctive SARS-CoV-2 lineages from infected deer is from the Volga River State Recreation Area in Fayette county where 4 lineages were recovered from the RPLN of hunter-killed deer within a two-mile radius in the three days spanning December 7 through 9, 2020. A final example is from the Lake Rathbun area of Appanoose county, where on December 5, 2020, of 10 positive RPLNs harvested from hunter-killed deer, there were 5 distinctive lineages represented within a 5 mile radius – lineage B.1.311 was recovered from 6 deer RPLNs, and one each of lineages B.1.362, B.1.240, B.1.400, and B.1. Together, these findings strongly suggest multiple point sources of spillover of distinct SARS-CoV-2 lineages to captive and free-living deer during this time in Iowa.

Recent evidence suggests that experimentally infected deer readily transmit the virus to other susceptible WTD between 3-5 days post infection, and the virus can be recovered from the palatine tonsils and RPLNs of infected animals for up to 21 days post-exposure (6). However, evidence of deer-to-deer transmission of the virus in free-living WTD has not yet been documented. To explore evidence for potential sylvatic transmission in free-living WTD, we applied a molecular epidemiologic approach to explore the temporal patterns of recovery of SARS-CoV-2 lineages from free-living WTD to identify potential evidence of deer-to-deer transmission. One example is evident from the Lake Rathbun area of Appanoose county, where lineage B.1.311 was predominant amongst deer representing 14 of 21 positive samples. The first RPLN sample harboring a B.1.311 lineage was recovered on December 5, 2020 from a hunter-killed WTD from this area, this was followed by additional recoveries of B.1.311 on Dec 8 and then again on January 2, 2021 and January 9, 2021 – more than a month apart – and with high viral loads in the lymph nodes suggestive of active infection. Together with the observation that the B.1.311 lineage was less frequently reported (based on available sequences) from humans in

Iowa as well as in WTD in other IA counties, the results are suggestive of the continued circulation of this lineage amongst WTD in free-living settings. However, it is essential to note that in the absence of a comprehensive longitudinal study of circulating lineages, it is not possible to formally exclude the possibility that B.1.311 was circulating within humans or other hosts in this area, and the deer were repeatedly exposed to the same point source(s) of infection.

Finally, to better visualize phylogenetic relationships amongst circulating SARS-CoV-2 originating in free-living and captive deer, we generated an SNP-based maximum likelihood tree including available human and animal lineage isolates (Fig. 3). As evident from the branching patterns of the phylogram, the results highlight the presence of multiple independent but closely related SARS-CoV-2 lineages circulating amongst deer in Iowa, as well as provide strong evidence for transmission within deer as many of the genomes from individual deer shared complete genomic identity (no SNPs) or differed by between 1 and 5 SNPs. The results also highlight several branches with shared human and deer origin SARS-CoV-2 isolates circulating in IA, that are related to but distinct from the isolates previously identified outbreaks from animals such as farmed mink or otters or other domesticated animal species. Hence, taken together, the results of these analyses provide strong evidence of multiple spillover events and the circulation of these strains within free-living and captive deer.

Broader implications of our findings for the ecology of SARS-CoV-2

Most viruses causing disease in humans originated in animals and many are capable of transmitting among multiple species (16, 17). The ability to infect a range of host species is a risk factor for disease emergence (18, 19). Despite this knowledge, reservoir host(s) are rarely identified and studied. Indeed, the reservoir(s) of SARS-1, SARS-CoV-2 and MERS-CoV are still not known. There have been numerous cases of isolated human-to-animal transmission of SARS-CoV-2 involving companion, farmed, and zoo animals since the COVID-19 pandemic began (8, 9, 20, 21). Our study is the first to provide evidence of widespread dissemination of SARS-CoV-2 into any free-living species, in this case white tailed deer. Our results suggest that deer have the potential to emerge as a major reservoir host for SARS-CoV-2, a finding that has important implications for the virus genomic diversity and future trajectory of the pandemic.

A key question is, what are be the consequences of the deer acting as a reservoir of SARS-CoV-2? When pathogens infect a single host species, the population dynamics are intrinsically unstable, and an outbreak spreads rapidly through a population and then fades out as hosts either develop immunity or die from the infection. The outbreak's speed depends on the basic reproductive number (R_0) and the generation time of the infection, but this changes when a pathogen is a generalist and infects multiple host species. In this instance, the dynamics are dominated by what occurs within the reservoir host, defined as the host which can maintain the infection and from which infection is transmitted to other hosts (22).

A reservoir host can facilitate viral mutation and increased pathogenicity to the original host. Animal reservoirs are also rife with a plethora of other CoVs with which SARS-CoV-2 will be afforded the opportunity to recombine, opening potential avenues for the acquisition of genes that might serve to increase virulence, transmissibility, pathogenicity, and immune evasion (23). Evidence of some of these exist in the Denmark mink spillover event where the Y435F substitution (which conferred increased affinity of the spike protein to human ACE2) evolved after human to mink transmission (24-26). Animal reservoirs can provide a refuge outside of a largely immune/vaccinated human population and thus represents a looming threat of re-emergence into humans (23). In addition, human viruses under immune pressure may diverge, whereas those circulating in wild reservoirs may not and reintroduction of these viruses into a future human could lead to outbreaks in young human humans who were never exposed to the ancestral pandemic strains (23). This is the most likely scenario observed in the 2009 A-H1N1 (swine flu) pandemic where the virus was related both to the pandemic 1918 strain and to strains circulating in the early 20th century (23, 27).

Predicting how the utilization of a new host species by a virus can affect virulence in the primary host is not simple. In theory, in a single-host system pathogens evolve to an optimum determined by a trade off with transmission, but this becomes more complicated in multiple-host systems (28). With the infection apparently spreading so quickly through the deer population as seen in our study, this could potentially result in fade out with insufficient susceptible deer recruits to sustain the infection within the deer population alone. Alternatively, with large annual birth cohorts or invasion into areas in which deer have not previously been infected, the virus may continue to spread among susceptible deer or circulate with the deer population. In general, while

the dynamics in these multi-host systems can be complex, they often result in more stable dynamics with multiple reservoir hosts, and the pathogens that utilize many hosts can be at a selective advantage since they are not lost during periods soon after the fade out.

Finally, the discovery of sylvatic and enzootic transmission in a substantial fraction of free-living deer has important implications for the natural ecology and long-term persistence of the SARS-CoV-2 pandemic, including through spillover to other free-living or captive animals and potential for spillback to human hosts, and highlights an urgent need for expanded active surveillance of potential wildlife reservoirs. The White-tailed deer is the most abundant wild cervid species in the United States with an estimated 25 million individuals. Deer hunting is the most popular type of hunting in the United States and contributed over $20 billion to the US GDP and supported more than 300,000 jobs in 2016. Given the social and economic importance of deer to the US economy, even while experimental evidence suggest that SARS-CoV-2 infected deer remain largely asymptomatic, the clinical outcomes and health implications of SARS-CoV-2 infection in free-living deer are unknown and warrant further investigation.

Study limitations

The study has several limitations: The RPLN samples tested were from only one State in the USA, and the sampling was not uniform throughout the State of Iowa. In addition, the deer samples tested in this study were from 2020, representing the early part of the pandemic, a time before the global dissemination of the highly successful Alpha and Delta variants. Hence, testing of more recent deer samples and robust longitudinal sampling will provide critically needed information on the role of deer as a SARS-CoV-2 reservoir or variant generator. While we lack incontrovertible evidence for spill back (reverse zoonosis) from deer to humans or other free-living animals, our findings warrant a heightened awareness of the potential risks associated with human contact with free-living or captive deer. Further, our results highlight an urgent need for the continued surveillance of SARS-CoV-2 in deer and other susceptible wild and peridomestic mammals.

Concluding comments

To better predict the emergence of the next pandemic and control infectious diseases with pandemic and panzootic potential, understanding the human–animal molecular and ecological

interface and their relevance to infection transmission and disease are essential (9). Hence, we call for an urgent need to implement a more proactive and robust "One Health" approach to better understand the ecology and evolution of SARS-CoV-2.

Materials and Methods

Samples:

The Iowa Department of Natural Resources (DNR) routinely collects medial retropharyngeal lymph nodes (RPLNs) from white-tailed deer across Iowa for its statewide Chronic Wasting Disease (CWD) surveillance program. Tissue samples were collected by trained field staff. Paired RPLNs were then removed and placed into separate Whirl-Paks with corresponding sample identification numbers and frozen at -20°F in a standard chest or standing freezer. A total of 283 RPLN samples collected between April 2020 to January 2021 were studied (Supplementary Table 1). An additional 60 RPLN archived samples from the 2019 deer hunting season were included as process negative control samples (Supplemental Table 2).

RNA extraction

RPLN tissues were processed by adding 3ml UTM (Copan) to a whirl-pak bag containing the tissue and placing the bag in the stomacher on a high setting for 120 seconds. Liquid volume was recovered and centrifuged at 3,000 rpm for 5 minutes to pellet cellular debris. 400 µL of the RPLN tissue homogenate supernatant was used for viral RNA extraction with a KingFisher Flex machine (ThermoFisher Scientific) with the MagMAX Viral/Pathogen extraction kit (ThermoFisher Scientific) following the manufacturer's instructions.

Detection of SARS-CoV-2 viral RNA by RT-PCR

The presence of SARS-CoV-2 nucleic acid was assessed by a real-time reverse transcription-polymerase chain reaction (RT-PCR) assay using the OPTI Medical SARS-CoV-2 RT-PCR kit following the manufacturer's instructions on an ABI 7500 Fast instrument (ThermoFisher Scientific). The OPTI Medical SARS-CoV-2 RT-PCR assay detects two different targets in the gene encoding viral nucleocapsid (N) protein coding region. The assay is highly sensitive with a limit of detection of 0.36 copies/µl. The internal control RNase P (RP) was used to rule out

human contamination. We generated a standard curve using SARS-CoV-2 RNA with a known copy number. Using the standard curve, viral RNA copies per milliliter of tissue homogenate were calculated.

To ensure assay specificity, a subset of 25 positive and 25 negative samples were additionally tested with the ThermoFisher TaqPath kit (ThermoFisher Scientific) targeting the SARS-CoV-2 ORF1ab, N gene, and S gene (Ref). The results were concordant with both assays. Further, to ensure samples were not inadvertently contaminated with human origin tissue or fluids during harvesting or processing, all samples were tested and found negative for the presence of human RNaseP. As a final check of assay specificity, none of the 60 RPLN samples collected in 2019 prior to the first reported case in humans in the United States were found positive for the presence of SARS-CoV-2 RNA.

SARS-CoV-2 Genome Sequencing

Total RNAs extracted from RPLN samples was used for sequencing the whole genomes of SARS-CoV-2 as previously described (13-15). Briefly, libraries were prepared according to version 4 of the ARTIC nCoV-2019 sequencing protocol (https://artic.network/ncov-2019, last accessed October 27, 2021). We used a semi-automated workflow that employed BioMek i7 liquid handling workstations (Beckman Coulter Life Sciences, Indianapolis, IN) and MANTIS automated liquid handlers (FORMULATRIX, Bedford, MA). Short sequence reads were generated with a NovaSeq 6000 instrument (Illumina, San Diego, CA). To ensure a very high depth of coverage, the RPLN sequencing libraries were prepared in duplicate and sequenced with a SP 300 cycle reagent kit.

SARS-CoV-2 Genome Sequence Analysis and Identification of Variants

Viral genomes were assembled with the BV-BRC SARS-Cov2 assembly service (The Bacterial and Viral Bioinformatics Resource Center, *https://www.bv-brc.org/app/comprehensivesars2analysis*, last accessed October 27, 2021, requires registration)(29). The One Codex SARS-CoV-2 variant calling and consensus assembly pipeline was used to assemble all sequences (GitHub, *https://github.com/onecodex/sars-cov-2.git*, last accessed October 27, 2021) using default parameters and a minimum read depth of three. Briefly, the pipeline uses seqtk version 1.3-r116 for sequence trimming

(GitHub, *https://github.com/lh3/seqtk.git*, last accessed October 27, 2021); minimap version 2.1 (*https://github.com/lh3/minimap2*, last accessed October 27, 2021) for aligning reads against reference genome Wuhan-Hu-1 (*https://www.ncbi.nlm.nih.gov/nuccore/1798174254*; NC_045512.2)(30); samtools version 1.11 (*http://www.htslib.org*, last accessed October 27, 2021) for sequence and file manipulation(31); and iVar version 1.2.2 (*https://github.com/andersen-lab/ivar/releases*, last accessed October 29, 2021) for primer trimming and variant calling(32). Genetic lineages, VOCs, and VOIs were identified on the basis of genome sequence data and designated by Pangolin version 3.1.11 (*https://github.com/cov-lineages/pangoLEARN*, last accessed October 27, 2021) with pangoLEARN module 2021-08-024 (SARS-CoV-2 lineages, *https://cov-lineages.org/pangolin.html*, last accessed October 27, 2021). SNPs identified using vSNP (https://github.com/USDA-VS/vSNP,) SNP analysis program.

QGIS mapping software version 3.16.10 was used to portray/visualize the geographic distribution of the RPLN samples.

Funding

The study is funded by the Huck institutes of the life sciences (SVK and VK), US Department of Agriculture National Institute of Food and Agriculture (NIFA) Award 2020-67015-32175, USFWS Wildlife and Sport Fish Restoration Program grant awards: FY20 F19AF00434 (19/20 samples), FY21 F20AF00309 (20/21 samples), and FY22 F21AF01914 (21/22 samples) and The Department of Natural Resources Fish and Game Protection Fund. This project was also partly supported by the Houston Methodist Academic Institute Infectious Diseases Fund; and supported in whole or in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. 75N93019C00076 (J.J.D.)

Author contributions

S.V.K conceptualized the study; S.V.K, V.K and R.M.D designed the project; M.S.N, M.Y, R.H.N, L.L, B.M.J, K.J.V, C.D.M, N.L, R.K.N, K.W, A.J.K.C, R.J.O, J.J.D, J.M.M and P.J.H performed research and or analysis. All authors contributed to writing the manuscript.

Competing interests

Authors declare that they have no competing interests.

Data and materials availability

All SARS-CoV-2 consensus genomes are deposited in GISAID and raw reads submitted to NCBI's Short Read Archive (BioProject Number: PRJNA776532)

Acknowledgments

We thank Abhinay Gontu, Shubhada Chothe, Padmaja Jakka and Abirami Ravichnadran from Kuchipudi lab, Penn State for their help with tissue homogenization. Vincent Nelson, Lindsey LaBella and Corey Price from the Penn State CLIA lab for help with sample processing. We greatly acknowledge Matthew Ojeda Saavedra, Sindy Pena, Kristina Reppond, Madison N. Shyer, Jessica Cambric, Ryan Gadd, Rashi M. Thakur, Akanksha Batajoo, and Regan Mangham for their technical assistance in the genome sequencing. We are grateful to Drs Suelee Robbe-Austerman, Tod Stuber and Kristina Lantz, USDA NVSL for their invaluable assistance and wise counsel throughout this investigation.

References Cited

1. V. Coronaviridae Study Group of the International Committee on Taxonomy of, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544 (2020).
2. V. J. Munster, M. Koopmans, N. van Doremalen, D. van Riel, E. de Wit, A Novel Coronavirus Emerging in China - Key Questions for Impact Assessment. *N Engl J Med* **382**, 692-694 (2020).
3. J. H. U. Medicine (2020) Coronavirus Resource Center.
4. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J Virol* **94** (2020).
5. J. Damas *et al.*, Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc Natl Acad Sci U S A* **117**, 22311-22322 (2020).
6. M. V. Palmer *et al.*, Susceptibility of White-Tailed Deer (Odocoileus virginianus) to SARS-CoV-2. *Journal of Virology* **95**, e00083-00021 (2021).
7. S. R. Weiss, S. Navas-Martin, Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol. Mol. Biol. Rev.* **69**, 635-664 (2005).
8. N. Oreshkova *et al.*, SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Euro Surveill* **25** (2020).

9. E. C. Hobbs, T. J. Reid, Animals and SARS-CoV-2: Species susceptibility and viral transmission in experimental and natural conditions, and the potential implications for community transmission. *Transbound Emerg Dis* **68**, 1850-1867 (2021).
10. B. B. Oude Munnink *et al.*, Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **371**, 172-177 (2021).
11. J. C. Chandler *et al.*, SARS-CoV-2 exposure in wild white-tailed deer (Odocoileus virginianus). *bioRxiv* 10.1101/2021.07.29.454326, 2021.2007.2029.454326 (2021).
12. M. C. Smithgall, M. Dowlatshahi, S. L. Spitalnik, E. A. Hod, A. J. Rai, Types of Assays for SARS-CoV-2 Testing: A Review. *Lab Med* **51**, e59-e65 (2020).
13. R. J. Olsen *et al.*, Trajectory of Growth of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Variants in Houston, Texas, January through May 2021, Based on 12,476 Genome Sequences. *Am J Pathol* **191**, 1754-1773 (2021).
14. S. W. Long *et al.*, Sequence Analysis of 20,453 Severe Acute Respiratory Syndrome Coronavirus 2 Genomes from the Houston Metropolitan Area Identifies the Emergence and Widespread Distribution of Multiple Isolates of All Major Variants of Concern. *Am J Pathol* **191**, 983-992 (2021).
15. S. W. Long *et al.*, Molecular Architecture of Early Dissemination and Massive Second Wave of the SARS-CoV-2 Virus in a Major Metropolitan Area. *mBio* **11** (2020).
16. M. E. J. Woolhouse, S. Gowtage-Sequeria, Host range and emerging and reemerging pathogens. *Emerg. Infect. Dis* **11**, 1842-1847 (2005).
17. L. H. Taylor, S. M. Latham, M. E. J. Woolhouse, Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B-Biological Sciences* **356**, 983-989 (2001).
18. S. Cleaveland, M. K. Laurenson, L. H. Taylor, Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **356**, 991-999 (2001).
19. P. Daszak, A. A. Cunningham, A. D. Hyatt, Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science* **287**, 443-449 (2000).
20. E. I. Patterson *et al.*, Evidence of exposure to SARS-CoV-2 in cats and dogs from households in Italy. *Nat Commun* **11**, 6231 (2020).
21. G. A. Calvet *et al.*, Investigation of SARS-CoV-2 infection in dogs and cats of humans diagnosed with COVID-19 in Rio de Janeiro, Brazil. *PLoS One* **16**, e0250853 (2021).
22. D. T. Haydon, S. Cleaveland, L. H. Taylor, M. K. Laurenson, Identifying reservoirs of infection: a conceptual and practical challenge. *Emerg Infect Dis* **8**, 1468-1473 (2002).
23. A. Telenti *et al.*, After the pandemic: perspectives on the future trajectory of COVID-19. *Nature* **596**, 495-504 (2021).
24. B. B. O. Munnink *et al.*, Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **371**, 172-177 (2021).
25. N. Oreshkova *et al.*, SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Eurosurveillance* **25**, 2-8 (2020).
26. R. Bayarri-Olmos *et al.*, The SARS-CoV-2 Y453F mink variant displays a pronounced increase in ACE-2 affinity but does not challenge antibody neutralization. *The Journal of biological chemistry* **296**, 100536 (2021).

27. B. Manicassamy *et al.*, Protection of Mice against Lethal Challenge with 2009 H1N1 Influenza A Virus by 1918-Like and Classical Swine H1N1 Based Vaccines. *Plos Pathogens* **6** (2010).
28. M. E. Woolhouse, L. H. Taylor, D. T. Haydon, Population biology of multihost pathogens. *Science* **292**, 1109-1112 (2001).
29. J. J. Davis *et al.*, The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res* **48**, D606-D612 (2020).
30. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
31. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
32. N. D. Grubaugh *et al.*, An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol* **20**, 8 (2019).

Table 1: Population Demographics of White-tailed deer population in IA screened for SARSCoV-2 surveillance

Category		Total samples (n)	SARSCoV2 RT-PCR positive	95% CI (Two-tailed Wilson Method)	p value	PCR positives Month				
						SEP	OCT	NOV	DEC	JAN
Gender	Male (M)	179	59	26.1 - 39.9	0.86 [M, F]	2	1	17	39	
	Female (F)	103	35	24.8- 43.1			3	5	22	5
	Unknown	1								
	Total (N)	283	94			2	4	22	61	5
Age	Adult (A)	250	82	26.9 - 38.6	0.59 [A, Y]	2	4	22	50	4
	Yearling (Y)	32	12	20.7 - 54.3					11	1
	Fawn	1								
	Total (N)	283	94			2	4	22	61	5
Location type	Captive*	131	27	13.7 - 27.5	<0.0001	1	1	14	11	
	Free living*	152	67	36.1 - 51.8		1	3	8	50	5
	Total (N)	283	94			2	4	22	61	5
Category		**Total samples (n)**	**SARSCoV2 RT-PCR positive**	**Proportion/ Prevalence (%)**	**95% CI** (Two-tailed Wilson Method)	**PCR positives Month**				
						SEP	OCT	NOV	DEC	JAN
County	Allamakee	11	11	100	74.1 - 100				11	
	Jasper	4	4	100	51.0 - 100				4	
	Polk	7	6	85.7	48.7 - 97.4				6	
	Appanoose	28	21	75	56.6 - 87.3				16	5
	Fayette	10	6	60	31.3 -83.2				6	
	Jefferson	8	4	50	21.5 - 78.5			3	1	
	Woodbury	46	13	28.3	17.3 - 42.6	1	2	7	3	
	Pottawattamie	22	5	22.8	10.1 - 43.4		1	1	3	
	Des Moines	112	23	20.5	1.5 - 28.9	1	1	11	10	
	Dubuque	12	1	8.3	1.5 - 35.4				1	
	Black Hawk	9		0	0 - 29.9					
	Dickinson	1		0	0 - 79.4					
	Henry	6		0	0 - 39.0					
	Jackson	3		0	0 - 56.2					
	Keokuk	1		0	0 - 79.4					
	Van Buren	1		0	0 - 79.4					
	Washington	1		0	0 - 79.4					
	Webster	1		0	0 - 79.4					
	Total (N)	283	94			2	4	22	61	5

*Captive: White-tailed deer samples obtained from preserve areas. Free-living: Either free-living on public lands or in peri-urban environments

Table 2: Temporal and spatial distribution of SARSCoV-2 lineages detected in White-tailed deer

Lineage		IA Counties										Month (Sep 2020 -Jan 2021)				
	Total samples (n)	Allamakee	Appanoose	Des Moines	Dubuque	Fayette	Jasper	Jefferson	Polk	Pottawattamie	Woodbury	SEP	OCT	NOV	DEC	JAN
B.1	7		1			1	4				1			1	6	
B.1.1	1							1						1		
B.1.119	2		1						1						2	
B.1.2	51	9		22	1	2		3		5	9	2	4	18	27	
B.1.234	6		1			2			3						5	1
B.1.240	1		1												1	
B.1.264	1								1						1	
B.1.311	19	1	14	1		1					2			2	14	
B.1.362	2		2												1	
B.1.400	2		1						1						2	
B.1.459	1	1													1	
B.1.596	1										1				1	
Total (N)	94	11	21	23	1	6	4	4	6	5	13	2	4	22	61	

Figure Legends

Figure 1: Epidemic curve showing SARS-CoV-2 weekly cases (per 100,000) in humans and the monthly change in SARS-CoV-2 positivity in White-tailed deer in Iowa. The histogram represents the progression in weekly reported cases (per 100,000 individuals; left axis). The percentage (and 95%) CI of White-tailed deer found positive for the presence of SARS-CoV-2 RNA (red line; right axis) appear to closely follows the trajectory of the human pandemic. The first identified positive sample on September 28, 2020 is marked, as is the start of the regular hunting season on September 19, 2020 and end on January 10, 2021.

Figure 2: Temporal and spatial distribution of SARSCoV-2 positive samples from white-tailed deer in IA. The 94 SARS-CoV-2 positive RPLN were from geographically dispersed sites and showed strong temporal clustering in frequency of detection of positive samples in deer. **A.** Monthly snapshots showing number and location of SARS-CoV-2 positive cases identified in deer in Iowa. **B.** Close-up view of increase in number of positive cases in deer from September through December 2020 within three exemplar regions with different sampling intensities and size of area sampled. Each filled black circle represents a negative test result, and each filled red circle represents a positive test result for presence of SARS-CoV-2 RNA in deer RPLN.

Figure 3: Whole genome SNP-based phylogenies of 94 SARS-CoV-2 genomes recovered from free-living and captive White-tailed deer in Iowa. Whole genome sequences from the 94 newly sequenced SARS-COV-2 genomes from deer RPLN were analyzed in the context of 84 additional publicly available animal origin SARS-CoV-2 isolates and 372 human SARS-CoV-2 isolates circulating in Iowa during this same period. The sequences were screened for quality thresholds, and SNP positions called against the SARS-CoV-2 reference were determined together with SNP alignments that were used to assemble a maximum-likelihood phylogenetic trees using RAxML. The results show several genetically distinct clusters of animal and human SARS-CoV-2 lineages circulating within the IA deer herd, suggesting multiple likely spillover events from humans to deer. Several branches with shared human and deer origin SARS-CoV-2 isolates circulating in IA were observed whereas found genetically distinct from the isolates from previous outbreaks in farmed mink and otters, but close clustering among deer isolates and SASRS-CoV-2 genomes recovered humans in Iowa.

Supplementary Figure 1: Temporal changes in distribution of SARS-CoV-2 viral genome copy numbers in White-tailed deer RPLNs. As the positivity proportion among the collected samples increased over the months of collection depicted on X axis, the viral copy numbers (y-axis) increased in a range of 268 to 5.4 x 10^8 copies/ml with a median of 106,000 viral copies/ml.

Supplementary Figure 2: Estimation plots of gender and age associated effects on the distribution of Ct values from RPLN SARS-CoV-2 positive samples. The results show no significant difference between the sexes (**A**) or age (**B**) in SARS-CoV-2 proportion of positive samples.

Figure 1

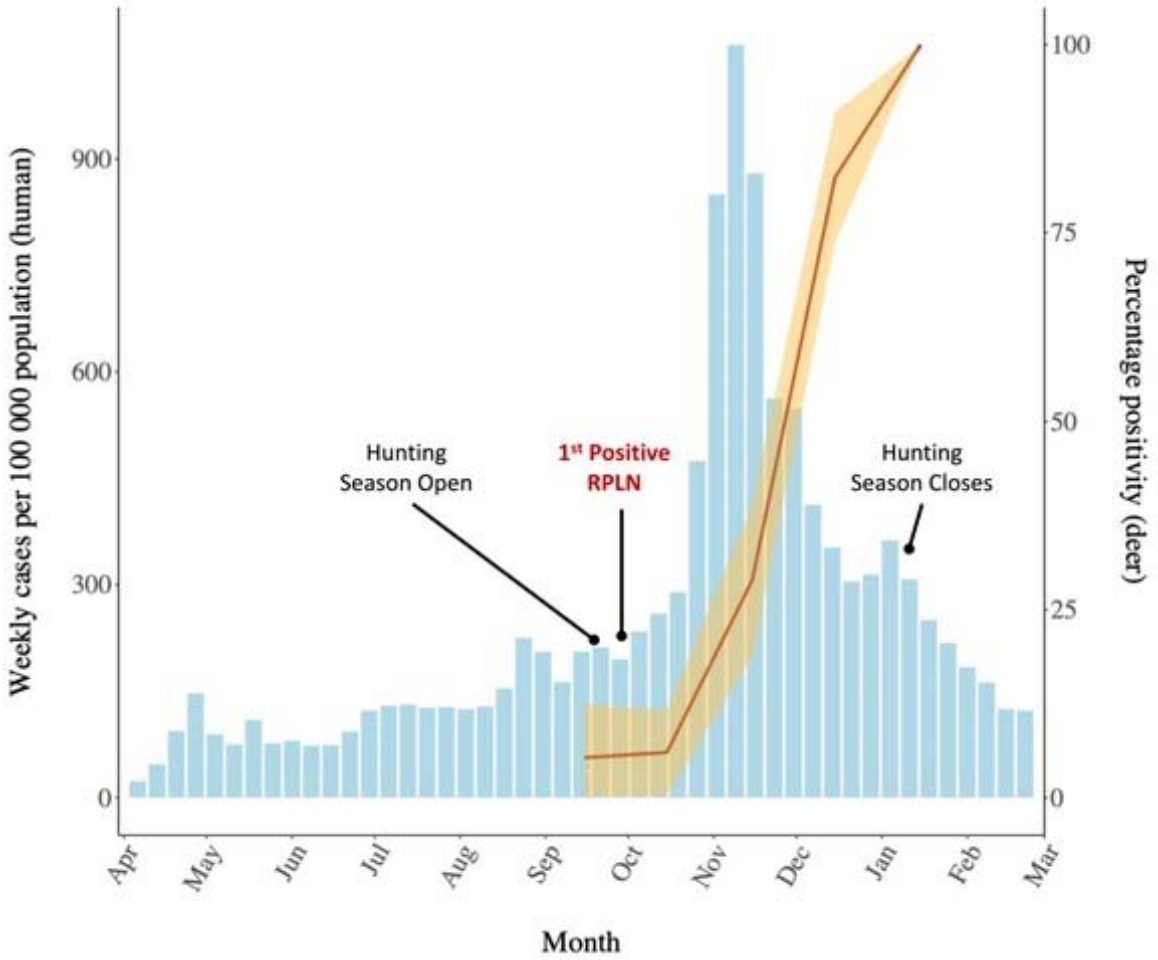

Figure 2

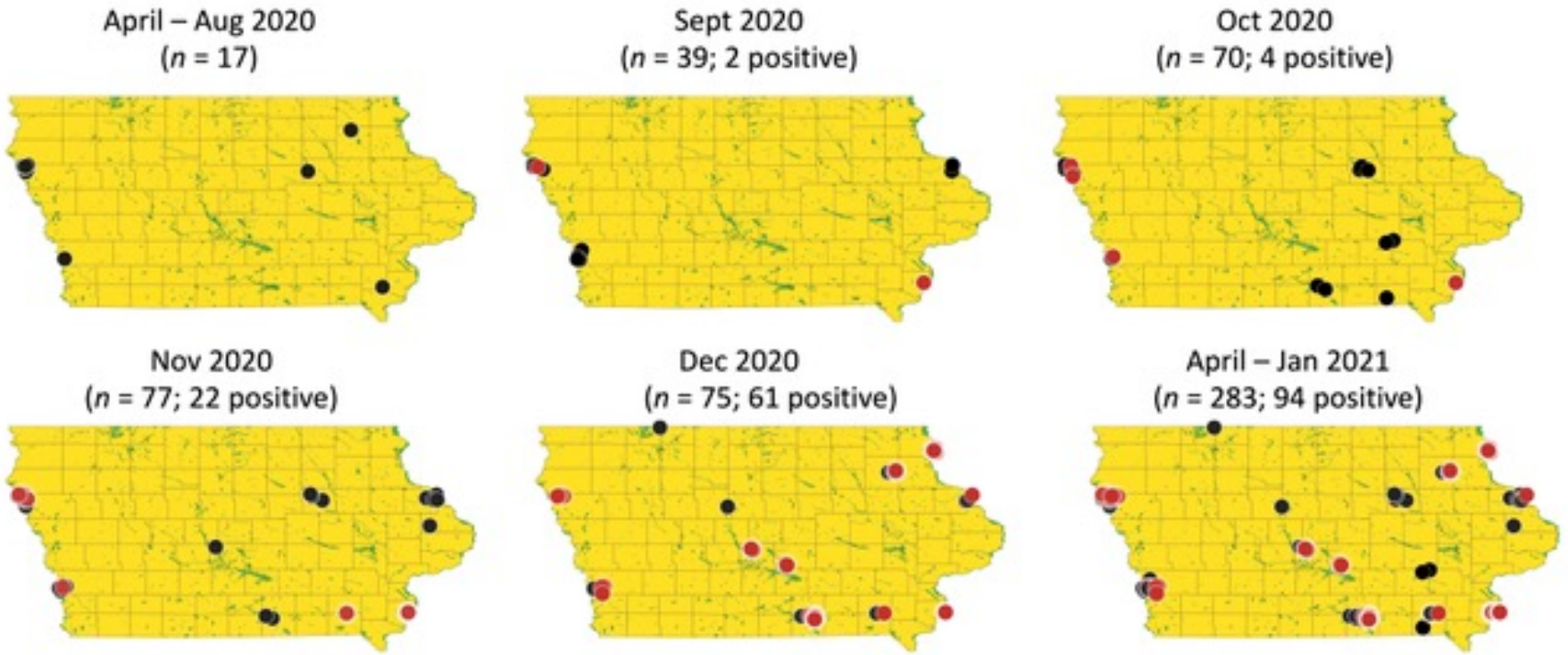

Figure 3