

**Article Summary Line:** The causative agent of a mass mortality event in passerines remains inconclusive after metagenomic high-throughput sequencing with results prompting further investigation into non-pathogenic causes.

**Running Title:** Metagenomic insights into a wildlife mortality event

**Keywords:** Metagenomics, Wildlife, Mortality Event, Avian, Rapid response

**Title:** Metagenomics for pathogen detection during a wildlife mortality event in songbirds

**Authors:** Lusajo Mwakibete, Sabrina S. Greening, Katrina Kalantar, Vida Ahyong, Eman Anis, Erica A. Miller, David B. Needle, Michael Oglesbee, W. Kelley Thomas, Joseph L. Sevigny, Lawrence M. Gordon, Nicole M. Nemeth, C. Brandon Ogbunugafor, Andrea J. Ayala, Seth A. Faith, Norma Neff, Angela M. Detweiler, Tessa Baillargeon, Stacy Tanguay, Stephen D. Simpson, Lisa A. Murphy, Julie C. Ellis, Cristina M. Tato, Roderick B. Gagne

**Author Affiliations:**

Chan Zuckerberg Biohub, San Francisco, CA, USA (L. Mwakibete, V. Ahyong, N. Neff, A. M. Detweiler, C.M. Tato)

Chan Zuckerberg Initiative, Redwood City, CA, USA (K. Kalantar)

Department of Pathobiology, Wildlife Futures Program, University of Pennsylvania School of Veterinary Medicine, New Bolton Center, Kennett Square, PA, USA (S. S. Greening, E. Anis, E. A. Miller, L. A. Murphy, J. C. Ellis, R. B. Gagne)

Department of Pathobiology, PADLS New Bolton Center, University of Pennsylvania School of Veterinary Medicine, New Bolton Center, Kennett Square, PA USA (E. Anis, L. A. Murphy)

Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA (C. B. Ogbunugafor, A. J. Ayala)

University of New Hampshire, New Hampshire Veterinary Diagnostic Lab, Durham, NH, USA

(D. B. Needle, T. Baillargeon, S. Tanguay)

Hubbard Center for Genome Studies, University of New Hampshire, Durham, NH, USA (J. L.

Sevigny, L. M. Gordon, S. D. Simpson, W. K. Thomas)

Infectious Diseases Institute, The Ohio State University, Columbus, OH, USA (M. Oglesbee, S.

A. Faith)

Southeastern Cooperative Wildlife Disease Study and Department of Pathology, College of

Veterinary Medicine, University of Georgia, Athens, GA, USA (N. M. Nemeth)

Department of Pathology, College of Veterinary Medicine, University of Georgia, Athens, GA,

USA (N. M. Nemeth)

## ABSTRACT

Mass mortality events in wildlife can be indications of an emerging infectious disease. During the spring and summer of 2021, hundreds of dead passernines were reported across the eastern US. Birds exhibited a range of clinical signs including swollen conjunctiva, ocular discharge, ataxia, and nystagmus. As part of the diagnostic investigation, high-throughput metagenomic next-generation sequencing was performed across three molecular laboratories on samples from affected birds. Many potentially pathogenic microbes were detected, with bacteria comprising the largest proportion; however, no singular agent was consistently identified, with many of the detected microbes also found in unaffected (control) birds, and thus considered to be subclinical infections. Congruent results across laboratories have helped drive further investigation into alternative causes including environmental contaminants and nutritional deficiencies. This work highlights the utility of metagenomic approaches in investigations of emerging diseases and provides a framework for future wildlife mortality events.

## INTRODUCTION

Wildlife health and diversity are under increasing threats from a multitude of sources, with disease emergence in wildlife having the potential to affect the health of humans and domesticated species (1–3). The development of successful mitigation strategies for emerging infectious diseases in wildlife is often limited by the ability to identify the etiologic agent (4). For example, in May 2015, a mass mortality event was observed in central Kazakhstan in which over half of all saiga antelopes (*Saiga tatarica*) were lost prior to the identification of the etiologic agent and before any mitigation measures could be implemented (5).

Rapid advances in high-throughput sequencing technologies have seen a rise in the number of genomic approaches being applied in disease investigations alongside more traditional techniques, such as histopathology, bacterial culture, virus isolation, and PCR tests (6,7). One such approach is metagenomic next-generation sequencing (mNGS); a culture-independent untargeted technique that can be used to analyze all nucleic acids (i.e., DNA or RNA) within a biological sample. Untargeted approaches, such as mNGS, are unbiased when it comes to capturing all the microbes within a clinical sample, as the majority of microbes can be identified in the absence of a priori assumption (8). This ability is an advantage particularly when the etiologic agent is unknown, and untargeted approaches are increasingly being used to identify pathogenic agents in disease outbreaks affecting humans and livestock (9–11), although it remains relatively uncommon in wildlife (12–15).

Here we highlight the recent use of mNGS to investigate a wildlife mortality event that began in late May 2021 when reports of sick and dead birds were received across the eastern USA. The majority of reports involved nestling and juvenile passerine species including the Common Grackle (*Quiscalus quiscula*), Blue Jay (*Cyanocitta cristata*), European Starling

(*Sturnus vulgaris*), American Robin (*Turdus migratorius*), and Northern Cardinal (*Cardinalis cardinalis*), as well as limited reports of non-passeriform avian species that presented with similar clinical signs (i.e., swollen conjunctiva, crusty ocular discharge, head tilt, ataxia, hind limb paresis, and nystagmus). Several diagnostic laboratories launched investigations focused on identifying an etiologic agent using common diagnostic techniques from across multiple disciplines including pathology, virology, microbiology, parasitology, and toxicology (16). Findings from these investigations failed to identify a causative agent but were able to rule out common pathogens and toxicants previously associated with mass avian mortality, including *Salmonella* spp., *Chlamydia* spp., avian influenza viruses, West Nile virus, herpesvirus, *Trichomonas* spp., coccidiosis, and numerous pesticides (17). Several *Mycoplasma* spp. were detected in diseased conjunctiva of some affected birds (unpub. data), but detections were inconsistent, and these bacterial species are commonly detected in non-diseased birds (18).

To further investigate this event, three diagnostic laboratories, namely University of New Hampshire (UNH) Veterinary Diagnostic Laboratory in collaboration with Hubbard Center for Genome Studies, University of Pennsylvania's Wildlife Futures Program (WFP) in collaboration with Chan Zuckerberg Biohub San Francisco (CZ Biohub SF), and the Infectious Disease Institute at Ohio State University (IDI) undertook mNGS approaches to assist in the detection of a causative agent. Similar mNGS approaches have been previously used in response to mortality events in wild and captive avian species (19–21). Here, we describe the mNGS approaches undertaken by each lab and demonstrate how concurrent approaches can be helpful when investigating the primary cause of a mass mortality event in wildlife.

## METHODS

### *Sample collection and processing*

During the 2021 mass mortality event in passerines, three labs independently collected samples for mNGS (Figure 1). To summarize, WFP in collaboration with CZ Biohub SF collected whole eye (including conjunctiva) and brain samples from 94 birds including 86 suspected cases and 8 controls in addition to lung, cloacal bursa, and heart blood from 28 birds (20 cases and 8 controls). All suspect cases were selected based on the presence of swollen conjunctiva, eye lesions, and/or crusty ocular discharge. For the WFP samples, suspect cases were all fledglings belonging to one of five species (American Robin, Blue Jay, Common Grackle, Northern Cardinal, and Northern Mockingbird (*Mimus polyglottos*)). Control birds of the same species were sourced from rehabilitation centers in the months following the mortality event (from September to November 2021) where they presented in good nutritional condition with no clinical signs of illness and had died or were euthanized due to acute traumatic injuries (i.e., vehicle collisions or window strikes).

In comparison, researchers at UNH collected conjunctiva and ear canal tissue from 103 birds (all suspected cases). These birds were a mixture of adults and fledglings submitted by a number of collaborators including Yale University for which 14 species were identified (American Robin, Blue Jay, Common Grackle, Northern Cardinal, Eastern Phoebe (*Sayornis phoebe*), Rose-breasted Grosbeak (*Pheucticus ludovicianus*), Rusty Blackbird (*Euphagus carolinus*), European Starling, House Finch (*Haemorhous mexicanus*), Tufted Titmouse (*Baeolophus bicolor*), Cooper's Hawk (*Accipiter cooperii*), Eastern Screech Owl (*Megascops asio*), Sharp-shinned Hawk (*Accipiter striatus*), and Mourning Dove (*Zenaida macroura*)) in addition to birds that were not identified to species level but were known to belong to one of the

following broader taxonomic groups: finch (*Fringillidae*), pigeon/dove (*Columbidae*), thrush (*Turdidae*), sparrow (*Passeridae*), crow (*Corvidae*), or blackbird (*Icteridae*). The majority of these birds were found dead with only a small number having been euthanized at wildlife rehabilitation facilities. Birds outside the passeriform order were treated as suspect cases using the same criteria as suspect cases within the passeriform order (i.e., swollen conjunctiva, eye lesions, and/or crusty ocular discharge).

Lastly, IDI collected brain tissue from 4 birds (all suspected cases). These birds were a mixture of adults and fledglings belonging to one of four species (American Robin, Blue Jay, House Sparrow, and Mourning Dove). All birds were found alive with common clinical signs (swollen conjunctiva, ocular exudate, crusty eyes, and ataxia), and either died during transport or were euthanized at wildlife rehabilitation facilities. Additional details regarding sample collection and processing by each lab are provided in the Appendix.

### ***Metagenomic next-generation sequencing***

At each diagnostic laboratory, different metagenomic approaches were used for extraction, library preparation, and sequencing as summarized in Table 1. Further details are also provided in the Appendix. Some of the major differences between the approaches included the use of different sample types and the extraction of RNA (by WFP/CZ Biohub SF), DNA (by UNH/Yale), or both (by IDI). Following sequencing, the mNGS bioinformatic analysis across the laboratories remained the same with each utilizing the CZ ID metagenomic pipeline - an open-source sequencing analysis platform for identifying microbial sequences within a metagenomic dataset (<http://czid.org>, v6.8). The pipeline removes the avian and human host using STAR (22) and Bowtie2 (23), trims adapters using Trimmomatic (24), filters low-quality reads using PriceSeq (25), filters low-complexity sequences using LZW and identifies duplicate

reads using czid-dedup (<https://github.com/chanzuckerberg/czid-dedup>). The remaining reads are queried on the NCBI nucleotide (NT) and non-redundant protein (NR) databases utilizing GSNAP-L (26) and RAPSearch2 (27), respectively, to determine the microbes (28).

To account for background contamination, 24 water controls (used when sequencing the WFP/CZ Biohub SF samples) were selected on CZ ID to create a mass-normalized background model for processing their respective samples. Significant microbial hits were called from the normalized unique reads per million (rPM) that mapped to specific species and genera which passed the following threshold filters: z-score  $\geq 1$  (to denote significant presence in the sample compared to water background), NT rPM  $\geq 10$  (i.e.,  $\geq 10$  nucleotide reads per million mapping to specific taxa), NR rPM  $\geq 5$  (i.e.,  $\geq 5$  protein reads per million mapping to specific taxa) and average base pair nucleotide alignment  $\geq 50$  base pairs (i.e.,  $\geq 50$  average nucleotide reads alignment mapping to specific taxa). To further increase the validity of the microbial hits, select samples were run on the CZ ID Consensus Genome pipeline (v3.4.7), to assess the genomic coverage and ensure the number of reads was adequate to obtain consensus genomes. An example of this analysis looking at the validity of West Nile virus detected in a single bird is provided in the Appendix.

### ***Microbial composition analysis***

To investigate the microbial composition, samples were first filtered to remove background taxa (i.e., those present in water controls) by eliminating taxa with a z-score  $\leq 1$ . The proportion of microbes belonging to the taxonomic categories: archaea, bacteria, eukaryotes, or viruses, were reported per respective sample in addition to the two most commonly detected taxa per sample. Further analyses were conducted across the WFP/CZ Biohub SF samples to evaluate differences between the cases and controls. For these analyses, sample reports containing

taxonomic relative abundance data for all samples were downloaded from CZ ID and imported into R statistical software (*v*4.2.1; (29)). To investigate the difference in the abundance (NT rPM) of microbes in the eye and brain between the cases and the control group, a Wilcoxon rank-sum test was performed. The significantly differentially abundant microbe genera with *p*-values <0.01 were reported between the groups. Alpha (Simpsons) and beta (Bray Curtis) diversity measures were also calculated using the R package vegan (*v*2.5; (30)) to further investigate the microbial diversity in the eye and brain samples both within and between the case and control groups. The statistical significance in the alpha and beta diversity metrics were evaluated using a Mann-Whitney U test and Permutational Multivariate Analysis of Variance (PERMANOVA) analysis, respectively.

#### ***Data availability***

The SRA files of non-host reads for the WFP/CZ Biohub SF and UNH/Yale samples have been deposited with links to BioProject accession numbers PRJNA909835 and PRJNA961153, respectively, in the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>).

## **RESULTS**

### ***Investigating the microbial composition***

No single pathogenic microbe was identified across all the cases, with the most commonly detected microbes varying across diagnostic/research laboratories (Table 2). The species-level distribution consisted mainly of bacterial microbes in both cases and controls with the post-filtering species-level distribution ranging from 43.76% to 95.67% bacterial (mean = 73.08%) across all the samples (Appendix Figures 1,2,3). In addition to the most commonly

detected microbes, the presence of other microbes known to be pathogenic to avian species also varied across laboratories. For instance, across the WFP/CZ Biohub SF eye and brain samples, *Avibacterium* spp. (including *A. paragallinarum*, *A. endocarditidis*, *A. volantium*, and *A. avium*) and *Mycoplasma* spp. (including *M. gallisepticum*, *M. pneumoniae*, and *M. mycoides*) were detected in a large proportion of the cases (72.1% and 57.0% respectively) and controls (25.0% and 12.5% respectively), while both *Avibacterium* spp. and *Mycoplasma* spp. were not detected in any of the IDI cases and only in a small proportion from UNH/Yale (21.4% and 6.8% respectively).

*Plasmodium* spp. were also detected in 20.4% of all samples collected (9.7% cases and 65.6% controls) by WFP/CZ Biohub SF but only in a limited number of samples across the other laboratories (Table 3). In comparison, other microbes were detected in the UNH/Yale and IDI samples that were not found in the WFP/CZ Biohub SF; for instance, canarypox virus was detected in 7.62% and 22.2% of the samples, respectively. Furthermore, a comparison of the microbes detected in DNA or RNA libraries at IDI revealed that *Burkholderia cenocepacia*, *Bifidobacterium* spp., and *Prevotella melaninogenica* were only detected in DNA libraries, whilst *Cutibacterium acnes* and *E. coli* were detected in both DNA and RNA libraries.

Differences in the microbial taxa between case and control samples of the eye and brain were also detected in the WFP/CZ Biohub SF samples. Specifically, the microbial genera of *Mycoplasma* spp., *Campylobacter* spp., and *Avibacterium* spp. were detected at significantly higher levels ( $p$ -values  $<0.01$ ) in cases (Figure 2, Table 4). In the diversity analyses, we observed marginally higher, though insignificant differences in Simpson's alpha diversity ( $p$ -value = 0.09) species richness in cases versus controls. Meanwhile, the Bray-Curtis beta diversity was

significantly different between cases and controls ( $p$ -values  $<0.001$ ), suggesting distinct microbial profiles across the two groups.

## DISCUSSION

Here, we describe the metagenomic approaches used to investigate the presence of potential etiologic agent(s) responsible for a mass mortality event in passerines in the eastern USA. After concurrent investigations across multiple diagnostic labs using an array of both targeted (17) and untargeted approaches, no singular pathogenic microbe was identified that would account for the observed morbidity and mortality in these birds, and to date, the results remain inconclusive.

Given the rapid onset and short time period of the event, it is likely that if a pathogen was the primary driver, it would have been detected in a larger percentage of samples. Variations in detection rates due to factors such as the disease stage at which the cases presented, and the tissues collected for sampling are unlikely sufficient to explain the lack of detection of a consistent pathogen across samples. Due to the unpredictable nature of wildlife health events and the reliance on opportunistic sampling, particularly early in the event, some of these factors are easier to control than others. For example, in multi-species events, such as the one reported here, some species may be overrepresented if they thrive in urbanized areas close to humans who can observe the event. The lack of observations from remote regions often results in geographical biases in the samples (31). Temporal biases are also likely, with samples coming from multiple sources and consisting of both morbid animals and animals that were dead for some time, often rendering them suboptimal for diagnostic purposes and raising concerns when comparing samples. Other factors are more easily controlled, such as the tissues selected for sampling, how

those tissues are processed, and what diagnostic tests are performed. In this study, whole carcasses and/or tissues of birds suspected to be involved in the outbreak were evaluated at numerous veterinary and wildlife diagnostic laboratories with the utilization of cross-disciplinary diagnostic techniques including pathology, virology, microbiology, parasitology, and toxicology (16). The tissues collected and protocols used by each group varied according to a multitude of factors including resource limitations, time constraints, and funding. This highlights a need for a minimum set of standards to help guide wildlife investigations and ensure some level of consistency across different working groups. A unified framework would also help facilitate collaboration in large multi-state events.

Employing the use of unbiased mNGS for this type of investigation has previously revealed many pathogenic agents including novel pathogens (32–34). Metagenomic findings in this study identified several bacterial pathogens significantly more in cases compared to controls; however, these were deemed unlikely drivers of the mortality event. Foremost, none appeared in a large percentage of samples across groups. In addition, characteristics of the pathology of the specific bacteria identified were inconsistent across birds and with the observed clinical signs. Nevertheless, given the uniqueness of the event, care must be taken in the interpretation of the bacterial pathogens detected. For instance, *Mycoplasma* spp. were found at significantly higher levels in eye and brain samples in cases versus controls. This result could explain the conjunctivitis reported on examination (35); however, *Mycoplasma* spp. are also a common commensal bacterium in many avian species (18). Thus, it is possible the detection of this microbe was due to opportunistic and/or subclinical infections rather than being a primary causative agent - particularly given it was still found in a large percentage of control birds. In

addition, *Mycoplasma* spp. were not consistently found across laboratories. Together, this suggests *Mycoplasma* spp. was not primarily responsible for the mortality event.

*Avibacterium* spp. were also detected at significantly higher levels in the cases than the controls; however, the clinical signs characteristic of *Avibacterium* spp. infections are more consistent with respiratory disease (i.e., mouth breathing, swollen sinuses, and nasal discharge (36,37), which were not observed. However, these clinical signs are typically described for chickens (*Gallus gallus domesticus*), with infections in wild avian species rarely reported and not well described; therefore, we cannot be certain that wild avian species would experience the same clinical signs. For example, a recent study reported severe periocular swelling, periocular skin crusting, fibrinous sinusitis, and conjunctivitis in wild turkeys infected with a novel clade of *Avibacterium* (38). The presence of multiple pathogenic agents also makes it difficult to determine the contribution of different microbes. For instance, 45.3% of cases for which both eye and brain were assessed were coinfecte with *Mycoplasma* and *Avibacterium* spp. This coinfection has also not been well described and thus, the resulting clinical signs are unknown; however, without more control cases or experimental infection trials, it is difficult to draw further conclusions. Yet, despite the limited number of controls, the beta diversity analysis did reveal a significant difference in the microbial compositions between the clinical cases and controls, suggesting that the former contained a different microbial profile which could have contributed to the morbidity and mortality.

Although no causative agent for the mortality event was identified, findings highlight the potential of using non-targeted approaches, such as mNGS, to help describe the microbial community circulating in wild populations including both pathogenic and non-pathogenic microbes (39,40). For example, the detection of *Avibacterium paragallinarum* in this study is

important to document as it is known primarily as a respiratory disease affecting chickens and is rarely described to cause disease in wild avian species (41). Generating baseline data and understanding how it changes over time further enables downstream comparisons to be made between diseased and healthy individuals and supports diagnostic responses to future mortality events. For example, the finding of *Mycoplasma* spp. in the control birds highlights the importance of using controls to help elucidate potential causes of disease in wildlife (42,43). Further, with the added health challenge posed by continual changes in the environments in which these birds live (e.g., landscape, climate, accumulation of potentially toxic substances), baseline microbes could be used as an indicator of the changing health status of an animal population (44).

Though mNGS is a powerful tool for detecting potential microbial pathogens, variability in experimental protocols, such as the sampling procedures, processing, and data analyses, can lead to artifacts that may result in incorrect conclusions and false negative detections (45,46). Each of the three laboratories in this study used a different experimental design, which enabled us to compare results and make recommendations on the best practices for conducting an mNGS investigation in response to a wildlife mass mortality event. Here we suggest a framework (Figure 3) to assist in addressing some of the limitations of this study for future explorations.

Firstly, developing a strong case definition is pivotal to help guide sample selection. In the early stages of the investigation, a case definition may be primarily based on the range of clinical signs present in the affected population; however, as more cases are identified it is important to revisit the case definition and include information regarding both clinical, laboratory, and pathologic characteristics as well as information on the affected individuals (e.g., species, age, etc.), and any geographical or temporal characteristics. A strong case definition also

helps in the selection of high-quality controls from the same population as the cases to provide background microbiota data from healthy specimens when performing downstream comparative analyses. Controls should only be selected if their cause of death is known and unrelated (e.g., they show no clinical and/or pathological signs of disease and/or are from a different area/population). Water controls during sample extraction should also be implemented to remove background contamination during the library preparation, as well as mock samples that may be used to spot contamination that may occur during sample collection and processing.

During sample collection, utilizing a nucleic acid stabilizer is crucial to assisting with retaining the total nucleic acid integrity and inactivating any infectious agents. We suggest starting with RNA libraries for the initial pass to reveal all actively replicating microbes within the samples that may be contributing to an active infection, while also allowing the capture of RNA viruses. However, if the RNA libraries yield no results, DNA libraries can be performed. For microbe detection in mNGS data, several steps can be taken to help increase the confidence in microbe detection further downstream in the analysis, including the implementation of threshold filters to help validate and increase confidence in the presence of detected pathogens, incorporating controls to compare microbial profiles and abundances between the case and control groups, and utilizing metadata to aid in the logical assessment of potentially pathogenic microbes that have been detected. It is also important to consider that any reference-based assembly could miss novel pathogens currently unavailable in the reference database, and further interrogation of the reads may be required (47,48). Our suggested framework for mNGS approaches (Figure 3) helps ensure that steps have been taken to minimize artifacts in the data. Here, parallel analyses across three diagnostic laboratories revealed no single pathogen associated with the 2021 mass mortality event in passerines. Findings suggest that the underlying

mortality is not due to pathogenic microorganisms and have guided the investigation to refocus time and resources on other potential factors, such as dietary deficiencies, to explain the mortality event.

## ACKNOWLEDGMENTS

We are grateful to all wildlife rehabilitators, biologists, veterinarians and veterinary staff, volunteers, members of the public, and responding state wildlife agencies who assisted with sample collection. We also want to extend our thanks to the CZ Biohub SF Genomics team for the sequencing support, the Wildlife Futures team for sampling assistance, and the faculty and staff at the Pennsylvania Animal Diagnostic Laboratory System-New Bolton Center who contributed to the sample collection, diagnostics, and shipping. S.S.G. was supported by the Robert J. Kleberg, Jr. and Helen C. Kleberg Foundation, while A. J. A was supported by an NSF Postdoctoral Fellowship in Biology (Award #2010904). UNH sequencing was funded through the USDA NIFA NH00694 awarded to W. K. Thomas.

## ADDRESS FOR CORRESPONDENCE

**Roderick Gagne**, Department of Pathobiology, Wildlife Futures Program, University of Pennsylvania School of Veterinary Medicine, New Bolton Center, 382 W. Street Road, Kennett Square, PA 19348, USA. [rgagne@vet.upenn.edu](mailto:rgagne@vet.upenn.edu)

## REFERENCES

1. Smith KF, Acevedo-Whitehouse K, Pedersen AB. The role of infectious diseases in biological conservation. *Anim Conserv*. 2009 Jan;12(1):1–12.

2. Heard MJ, Smith KF, Ripp K, Berger M, Chen J, Dittmeier J, et al. Increased threat of disease as species move towards extinction. *Conserv Biol.* 2013 Dec;27(6):1378–88.
3. Cunningham AA, Daszak P, Wood JLN. One Health, emerging infectious diseases and wildlife: two decades of progress? *Philos Trans R Soc Lond B Biol Sci.* 2017 Jul;372(1725):e20160167.
4. Stallknecht, DE. Impediments to wildlife disease surveillance, research, and diagnostics. In: Childs JE, Mackenzie JS, Richt JA (eds) *Wildlife and Emerging Zoonotic Diseases: The Biology, Circumstances and Consequences of Cross-Species Transmission.* Current Topics in Microbiology and Immunology; Springer, Berlin, Germany; 2007 vol 315.
5. Kock RA, Orynbayev M, Robinson S, Zuther S, Singh NJ, Beauvais W, et al. Saigas on the brink: Multidisciplinary analysis of the factors influencing mass mortality events. *Sci Adv.* 2018 Jan;4(1):eaao2314.
6. Blanchong JA, Robinson SJ, Samuel MD, Foster JT. Application of genetics and genomics to wildlife epidemiology. *J Wildl Managem.* 2016;80(4):593–608.
7. Lipkin WI. The changing face of pathogen discovery and surveillance. *Nat Rev Microbiol.* 2013;11(2):133–41.
8. Simner PJ, Miller S, Carroll KC. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin Infect Dis.* 2018 Feb;66(5):778–88.
9. Bohl JA, Lay S, Chea S, Ahyong V, Parker DM, Gallagher S, et al. Discovering disease-causing pathogens in resource-scarce Southeast Asia using a global metagenomic pathogen monitoring system. *Proc Natl Acad Sci USA.* 2022 Mar;119(11):e2115285119.
10. Greninger AL, Waghmare A, Adler A, Qin X, Crowley JL, Englund JA, et al. Rule-Out

outbreak: 24-hour metagenomic next-generation sequencing for characterizing respiratory virus source for infection prevention. *J Pediatr Infect Dis Soc*. 2017 Jun;6(2):168–72.

11. Miller RR, Montoya V, Gardy JL, Patrick DM, Tang P. Metagenomics for pathogen detection in public health. *Genome Med*. 2013 Sep;5(9):81-95.
12. Fitak RR, Antonides JD, Baitchman EJ, Bonaccorso E, Braun J, Kubiski S, et al. The expectations and challenges of wildlife disease research in the era of genomics: forecasting with a horizon scan-like exercise. *J Hered*. 2019 May;110(3):261–74.
13. Zylberberg M, Van Hemert C, Dumbacher JP, Handel CM, Tihan T, DeRisi JL. Novel picornavirus associated with avian keratin disorder in Alaskan birds. *mBio*. 2016 Jul;7(4):e00874-16.
14. Retallack H, Okihiro MS, Britton E, Sommeran SV, DeRisi JL. Metagenomic next-generation sequencing reveals *Miamiensis avidus* (*ciliophora: scuticociliatida*) in the 2017 epizootic of leopard sharks (*triakis semifasciata*) in San Francisco Bay, California, USA. *J Wildl Dis*. 2019 Apr;55(2):375–86.
15. Ko KKK, Chng KR, Nagarajan N. Metagenomics-enabled microbial surveillance. *Nat Microbiol*. 2022;7(4):486–96.
16. Richards BJ, Nemeth NM, Lewis NJ, Needle DB, Sevigny JL, Oglesbee M, et al. The 2021 songbird mortality event – A summary of findings and discussion of ongoing investigations [Internet]. Conference Proceedings presented at: 70th Annual International Wildlife Disease Association Conference; 2022 Jul; Madison, Wisconsin, US. Available from: <https://www.wildlifedisease.org/PersonifyEbusiness/Events/Past-International-Conferences>

17. USGS National Wildlife Health Center. Update for the Association of Fish and Wildlife Agencies from the USGS National Wildlife Health Center [Internet]. USGS National Wildlife Health Center; 2021 Sep. Available from: <https://www.usgs.gov/media/files/update-association-fish-and-wildlife-agencies-sept-2021>
18. Sawicka A, Durkalec M, Tomczyk G, Kursa O. Occurrence of *Mycoplasma gallisepticum* in wild birds: A systematic review and meta-analysis. PLoS One. 2020 Apr;15(4):e0231545.
19. Pankovics P, Boros Á, Phan TG, Delwart E, Reuter G. A novel passerivirus (family *Picornaviridae*) in an outbreak of enteritis with high mortality in estrildid finches (*Uraeginthus* sp.). Arch Virol. 2018 Apr;163(4):1063–1071.
20. Papineau A. Targeted enrichment and viral metagenomics in the detection of livestock and wildlife viruses [Master of Science]: University of Manitoba, Winnipeg, Canada; 2019.
21. Chang WS. Metagenomic Applications in Virus Discovery, Ecology, and the Surveillance of Australian Wildlife [Doctor of Philosophy]. University of Sydney, Sydney, Australia; 2021.
22. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan;29(1):15–21.
23. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012 Apr;9(4):357–9.
24. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug;30(15):2114–2120.

25. Ruby JG, Bellare P, Derisi JL. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. *G3 (Bethesda)*. 2013 May;3(5):865–880.
26. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*. 2010 Apr;26(7):873–881.
27. Ye Y, Choi JH, Tang H. RAPSearch: a fast protein similarity search tool for short reads. *BMC Bioinformatics*. 2011 May;12(1):159.
28. Kalantar KL, Carvalho T, de Bourcy CFA, Dimitrov B, Dingle G, Egger R, et al. IDseq- An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience*. 2020 Oct;9(10):giaa111.
29. R Core Team. R: A language and environment for statistical computing [Internet]. R Foundation for Statistical Computing, Vienna, Austria; 2018. Available from: <https://www.R-project.org/>
30. Oksanen J, Blanchet GB, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community ecology package [Internet]. R package version 2.5-7; 2020. Available from: <https://CRAN.R-project.org/package=vegan>
31. Wobeser GA. Essentials of Disease in Wild Animals. 1st ed. Ames, Iowa, USA: Blackwell Publishing; 2006.
32. Michel AL, Van Heerden H, Crossley BM, Al Dahouk S, Prasse D, Rutten V. Pathogen detection and disease diagnosis in wildlife: challenges and opportunities. *Rev Sci Tech*. 2021 Jun;40(1):105–18.
33. Fagre AC, Cohen LE, Eskew EA, Farrell M, Glennon E, Joseph MB, et al. Assessing the risk of human-to-wildlife pathogen transmission for conservation and public health. *Ecol Lett*. 2022 Mar;25:1534–1549.

34. Daszak P, Cunningham AA, Hyatt AD. Emerging infectious diseases of wildlife-- threats to biodiversity and human health. *Science*. 2000 Jan;287(5452):443–449.
35. Fischer JR, Stallknecht DE, Luttrell P, Dhondt AA, Converse KA. Mycoplasmal conjunctivitis in wild songbirds: the spread of a new contagious disease in a mobile host population. *Emerg Infect Dis*. 1997 Mar;3(1):69–72.
36. Blackall PJ. Infectious Coryza: Overview of the disease and new diagnostic options. *Clin Microbiol Rev*. 1999 Oct;12(4):627–32.
37. Paudel S, Ruhnau D, Wernsdorf P, Liebhart D, Hess M, Hess C. Presence of *Avibacterium paragallinarum* and histopathologic lesions corresponds with clinical signs in a co-infection model with *Gallibacterium anatis*. *Avian Dis*. 2017 Jul;61(3):335–40.
38. Ellis JS, MacGlover CA, Sondgeroth KS, Brown D, Daniels JB, Fox KA. Novel avibacterium species associated with sinusitis and conjunctivitis in a merriam's wild turkey (*meleagris gallopavo merriami*) flock in Colorado, USA. *J Wildl Dis*. 2022 Oct;58(4):725-734.
39. Jurburg SD, Buscot F, Chatzinotas A, Chaudhari NM, Clark AT, Garbowski M, et al. The community ecology perspective of omics data. *Microbiome*. 2022 Dec;10(1):225-238.
40. Martin TC, Visconti A, Spector TD, Falchi M. Conducting metagenomic studies in microbiology and clinical research. *Appl Microbiol Biotechnol*. 2018 Oct;102(20):8629–8646.
41. Nsengimana O, Habarugira G, Ojok L, Ruhagazi D, Kayitare A, Shyaka A. Infectious coryza in a grey crowned crane (*Balearica regulorum*) recovered from captivity. *Vet Med Sci*. 2022;8(2):822–6.
42. Van Hemert C, Pearce JM, Handel CM. Wildlife health in a rapidly changing North:

focus on avian disease. *Front Ecol Environ.* 2014 Dec;12(10):548–56.

43. Ryser-Degiorgis MP. Wildlife health investigations: needs, challenges and recommendations. *BMC Vet Res.* 2013 Nov;9(1):223-236.

44. Sun F, Chen J, Liu K, Tang M, Yang Y. The avian gut microbiota: Diversity, influencing factors, and future directions. *Front Microbiol.* 2022 Aug;13:e934272.

45. Wooley JC, Ye Y. Metagenomics: facts and artifacts, and computational challenges. *J Comput Sci Technol.* 2009 Jan;25(1):71–81.

46. Teeling H, Glöckner FO. Current opportunities and challenges in microbial metagenome analysis- A bioinformatic perspective. *Brief Bioinform.* 2012 Nov;13(6):728–42.

47. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat Biotechnol.* 2014 Aug;32(8):822–828.

48. Yang C, Chowdhury D, Zhang Z, Cheung WK, Lu A, Bian Z, et al. A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. *Comput Struct Biotechnol J.* 2021 Jan;19:6301–6314.

## TABLES

**Table 1.** Metagenomic next-generation sequencing (mNGS) approaches taken across three diagnostic laboratories in response to the 2021 mass mortality event in passerines

Institution	Samples (no. states & species)	Tissues	Extraction kit	Library prep	Sequencing platform	Mean pair end read per sample
NovaSeq 6000						
at paired						
WFP* & CZB SF†	86 cases¶ (2 states, 5 species)	Eye/ brain (n = 94)	Quick RNA	NEBNext Ultra	150bp.	65.7 million
		Bursa (n = 24)	Pathogen	II Library Prep	(Batch 1)	(batch 1)
		Lung (n = 24)	MagBead kit	Kit	+	+
IDI-OSU‡	8 controls¶ (3 states, 4 species)	Heart blood (n= 24)	(Zymo Research)	(New England Biolabs)	NextSeq 2000	5.5 million
					at paired	(batch 2)
					150bp.	
(Batch 2)						
UNH§ & Yale	4 cases (1 state, 4 species)		Maxwell RSC	DNA prep kit	NextSeq2000	
			Tissue DNA kit	(Illumina)	at paired-end	58.3 million
			(Promega)	+	100 bp (DNA)	(RNA)
		Brain (n = 8) #	+	RNA	+	+
			Maxwell RSC	Enrichment	NextSeq 2000	16.3 million
			Simpy RNA kit	prep kit	at paired	(DNA)
	103 cases (10 states, 18 species)		(Promega)	(Illumina)	150bp. (RNA)	
			MagMax DNA	HyperPrep kit		
		Conjunctiva + ear	Multi-Sample	(Kapa Biosystem	NovaSeq 6000	
		(n = 103)	Ultra 2.0 kit	Inc.)	at paired 250	7.9 million
			(Applied Biosystems)		bp	

\*WFP: Wildlife Futures Program, University of Pennsylvania. †CZB SF: Chan Zuckerberg Biohub San Francisco. ‡IDI OSU: Infectious Disease Institute at Ohio State University. §UNH: University of New Hampshire. ¶ Numbers representative for eye/ brain samples only. Lung, bursa, and heart blood were sampled from 16 cases and 8 control birds. #Brain samples were taken from the right and left hemispheres in four birds.

**Table 2.** Most frequently detected species taxa per respective sample type and status (i.e., case vs control) following metagenomic next-generation sequencing (mNGS) bioinformatic analysis using the CZ ID metagenomic pipeline (v6.8)

Institution	Sample Type	Group	Most Detected Taxa (n, % per respective group)
WFP* & CZB SFT	Eye + Brain	Control	<i>Delftia acidovorans</i> (n = 3, 37.5%)
		(n = 8)	<i>Escherichia coli</i> (n = 3, 37.5%)
			<i>Pasteurella multocida</i> (n = 3, 37.5%)
		Case	<i>Escherichia coli</i> (n = 67, 77.9%)
		(n=86)	<i>Avibacterium paragallinarum</i> (n = 53, 61.6%)
			<i>Pasteurella multocida</i> (n = 51, 59.3%)
	Bursa	Control	<i>Besnoitia besnoiti</i> (n = 2, 25.0%)
		(n=8)	<i>Clostridium perfringens</i> (n = 2, 25.0%)
			<i>Neospora caninum</i> (n = 2, 25.0%)
			<i>Rosa chinensis</i> (n = 2, 25.0%)
			<i>Toxoplasma gondii</i> (n = 2, 25.0%)
		Case	<i>Enterococcus faecalis</i> (n = 14, 87.5%)
		(n=16)	<i>Escherichia coli</i> (n = 13, 81.3%)
	Heart Blood	Control	<i>Cyclospora cayetanensis</i> (n = 2, 25.0%)
		(n=8)	<i>Plasmodium spp.</i> (n = 2, 25.0%)
			<i>Delftia acidovorans</i> (n = 2, 25.0%)
		Case	<i>Clostridium perfringens</i> (n = 7, 43.8%)
		(n=16)	<i>Escherichia spp.</i> (n = 7, 43.8%)

**Table 2 continued**

Institution	Sample Type	Group	Most Detected Taxa (n, % per respective group)
	Lung	Control (n=8)	<i>Cyclospora cayetanensis</i> (n = 3, 37.5%) <i>Pasteurella multocida</i> (n = 2, 25.0)
		Case (n=16)	<i>Escherichia coli</i> (n = 13, 81.3%) <i>Enterococcus faecalis</i> (n = 8, 50.0%) <i>Clostridium perfringens</i> (n = 8, 50.0%)
UNH§ & Yale	Conjunctiva + ear	Case (n=103)	<i>Escherichia coli</i> (n = 49, 47.6%) <i>Enterococcus faecalis</i> (n = 38, 35.0%)
IDI - OSU‡	Eye + Brain	Case (n=8)	<i>Cutibacterium acnes</i> (n = 6, 75.0%) <i>Bifidobacterium breve</i> (n = 4, 50.0%) <i>Bifidobacterium longum</i> (n = 4, 50.0%) <i>Burkholderia cenocepacia</i> (n = 4, 50.0%) <i>Escherichia coli</i> (n = 4, 50.0%) <i>Prevotella melaninogenica</i> (n = 4, 50.0%)

\*WFP: Wildlife Futures Program, University of Pennsylvania. †CZB SF: Chan Zuckerberg Biohub San Francisco. ‡IDI OSU: Infectious Disease Institute at Ohio State University. §UNH: University of New Hampshire.

**Table 3.** *Plasmodium spp.* and *P. relictum* detection across birds following metagenomic next-generation sequencing (mNGS) bioinformatic analysis using the CZ ID metagenomic pipeline (v6.8)

Institution	Sample Type	Group	Microbe	
			<i>Plasmodium spp.</i>	<i>P. relictum</i>
WFP* & CZB SF†	Eye + Brain	Control	4	0
		Case	7	1
	Bursa	Control	2	0
		Case	4	1
	Heart Blood	Control	6	2
		Case	2	2
	Lung	Control	7	1
		Case	2	1
	Total Samples		34	8
	Total Birds		15	5
UNH§ & Yale	Conjunctiva + ear	Case	1	0
		Total Samples	1	0
	Total Birds		1	0
IDI - OSU‡	Eye + Brain	Case	0	0
		Total Samples	0	0
	Total Birds		0	0

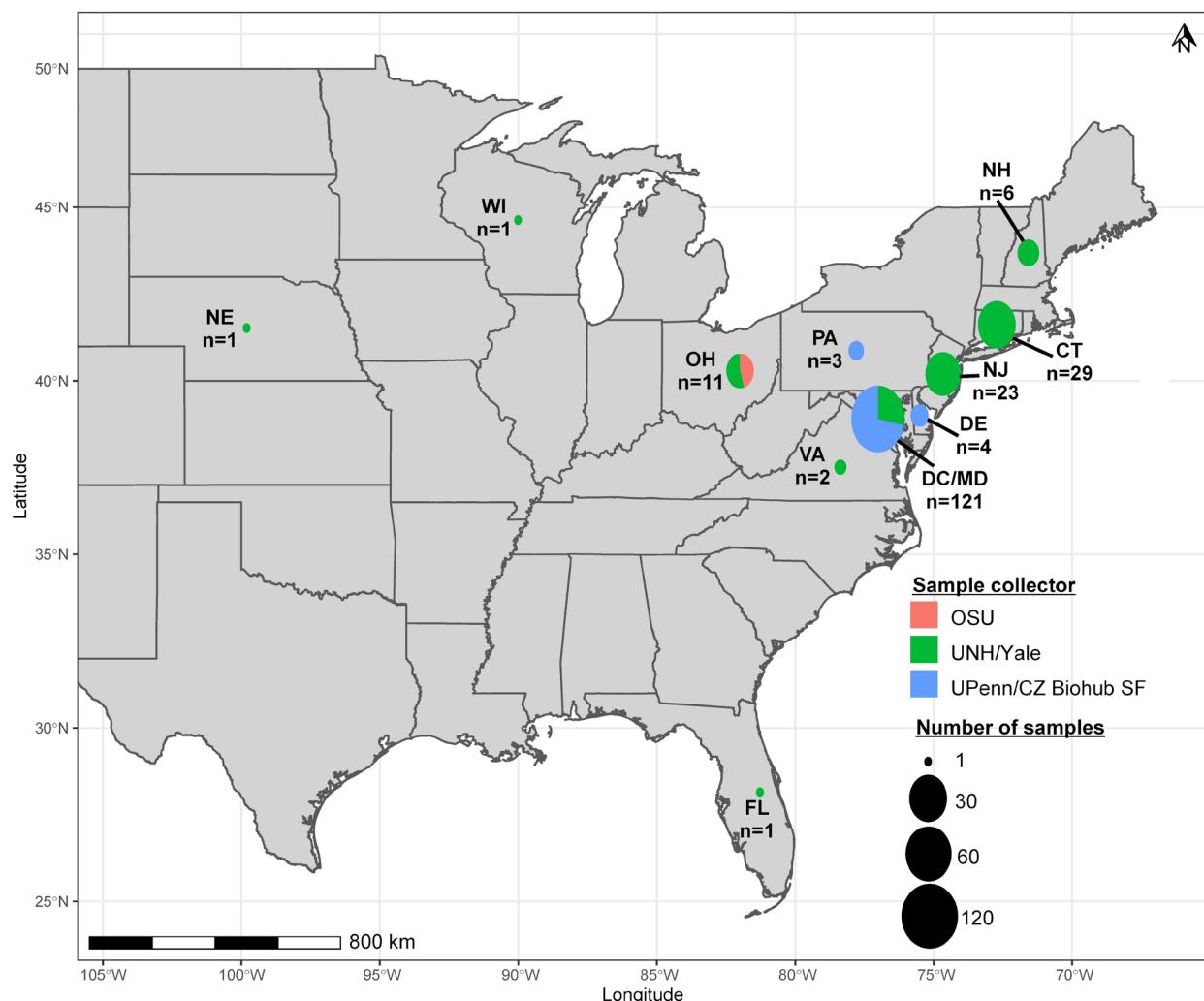
\*WFP: Wildlife Futures Program, University of Pennsylvania. †CZB SF: Chan Zuckerberg Biohub San Francisco. ‡IDI OSU: Infectious Disease Institute at Ohio State University. §UNH: University of New Hampshire.

**Table 4.** Wilcoxon rank test values between cases and controls utilizing NT rPM values from eye/brain samples collected by the Wildlife Futures Program, University of Pennsylvania, and Chan Zuckerberg Biohub

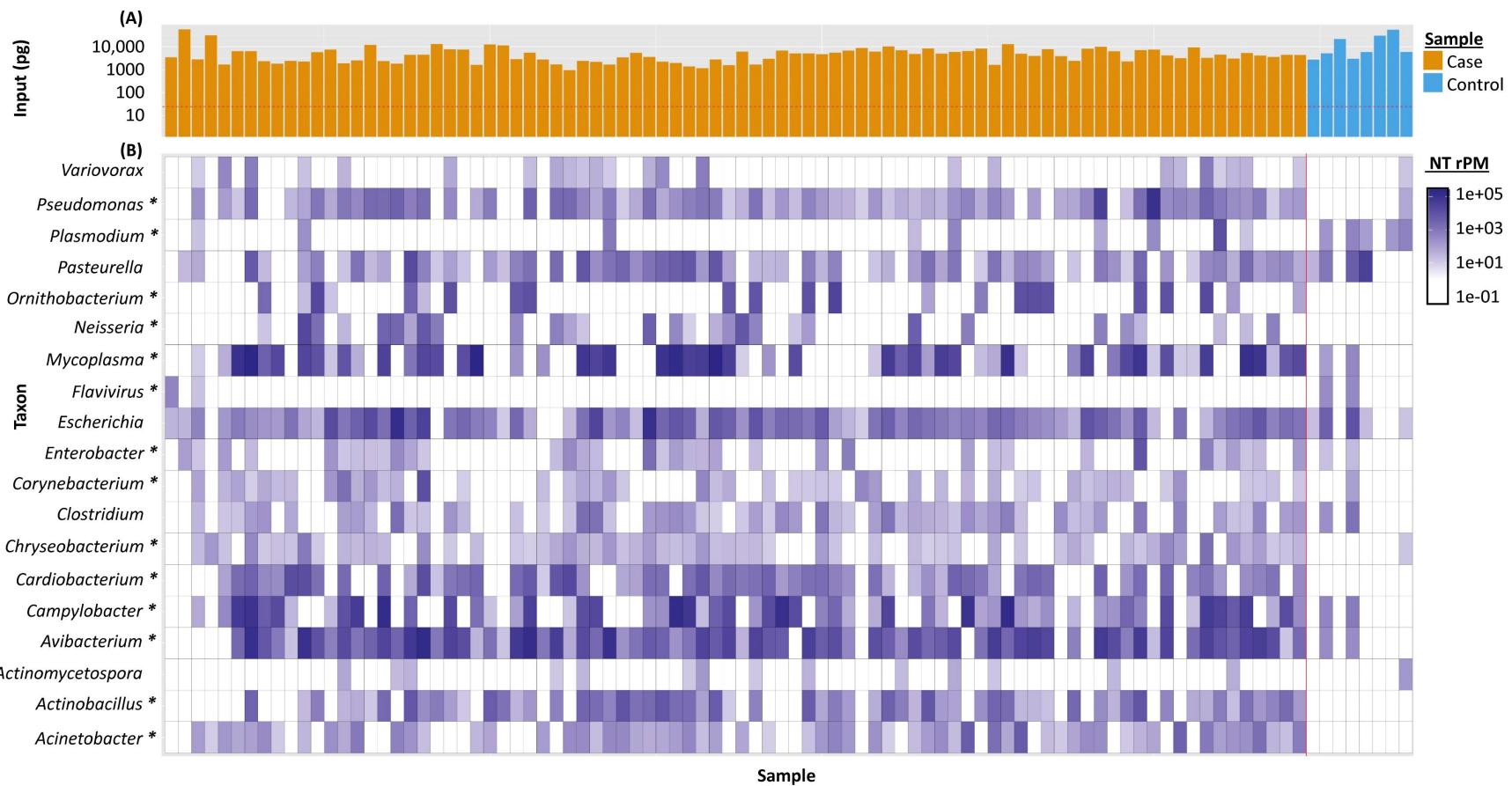
Microbe Genera	p-value
<i>Acinetobacter</i> *	9.38E-06
<i>Actinobacillus</i> *	5.78E-05
<i>Actinomycetospora</i>	1.14E-05
<i>Avibacterium</i> *	8.04E-05
<i>Campylobacter</i> *	4.15E-03
<i>Cardiobacterium</i> *	6.51E-05
<i>Chryseobacterium</i> *	9.78E-05
<i>Clostridium</i>	3.88E-02
<i>Corynebacterium</i>	1.55E-03
<i>Enterobacter</i> *	9.09E-03
<i>Escherichia</i>	2.16E-02
<i>Flavivirus</i> *	2.75E-06
<i>Mycoplasma</i> *	1.14E-03
<i>Neisseria</i> *	8.32E-04
<i>Ornithobacterium</i> *	1.09E-04
<i>Pasteurella</i>	5.83E-01
<i>Plasmodium</i> *	4.16E-04
<i>Pseudomonas</i> *	8.28E-05
<i>Variovorax</i>	1.06E-02

\* p-values <0.01

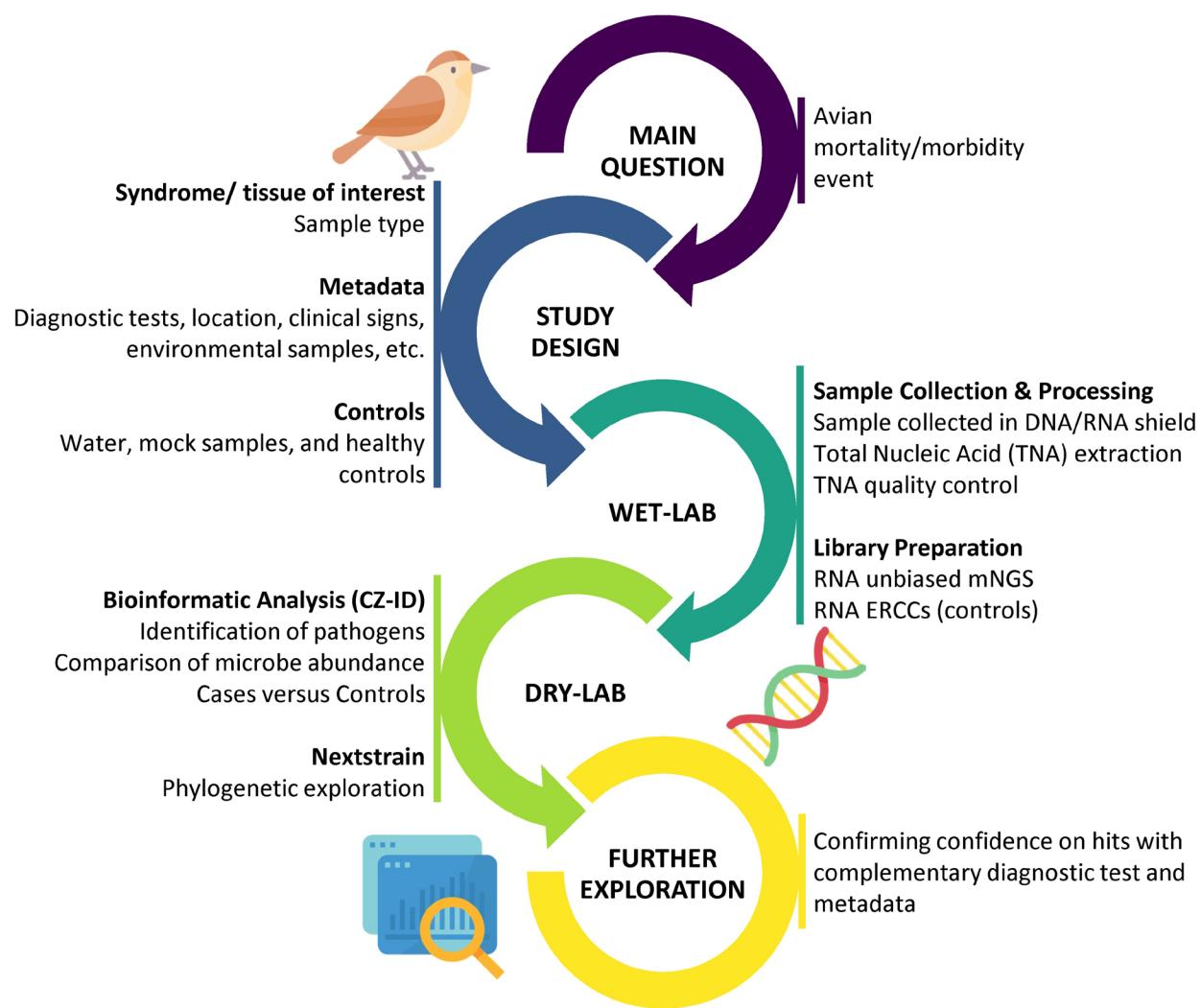
## FIGURES



**Figure 1.** Map showing the distribution of sampled birds (both cases and controls) that were included in the mNGS analyses by state (n=number of birds). The size of the pie chart is proportional to the number of samples while the color of the pie chart indicates the laboratory to which the samples were sent. UPenn: University of Pennsylvania, CZ Biohub SF: Chan Zuckerberg Biohub San Francisco, OSU: Ohio State University, UNH: University of New Hampshire, Yale: Yale University.



**Figure 2.** Microbes detected by mNGS in eye and brain samples sorted by genus. Panel (A) denotes sample input (pg) in each sample, whilst the color denotes if the sample was derived from a suspected case (orange) or control (blue). Panel (B) denotes each microbe genus. The asterisk (\*) denotes significantly different ( $p < 0.01$ ) microbe presence between the cases (low + high) and controls detected in the sample post-filtering.



**Figure 3.** Potential framework for metagenomic next-generation sequencing (mNGS) in wild avian species.