8-20-2021

Metagenomic characterization of swine slurry in a North American swine farm operation

A. Ramesh

E. S. Bailey

V. Ahyong

C. Langelier

M. Phelps

See next page for additional authors

Authors

A. Ramesh, E. S. Bailey, V. Ahyong, C. Langelier, M. Phelps, N. Neff, R. Sit, C. Tato, J. L. DeRisi, A. E. Greer, and G. C. Gray

scientific reports

Check for updates

OPEN

Metagenomic characterization of swine slurry in a North American swine farm operation

Akshaya Ramesh[1,2,4,12], Emily S. Bailey[3,4,12]✉, Vida Ahyong[5], Charles Langelier[5,6], Maira Phelps[5], Norma Neff[5], Rene Sit[5], Cristina Tato[5], Joseph L. DeRisi[5,7], Annette G. Greer[8] & Gregory C. Gray[3,9,10,11]

Modern day large-scale, high-density farming environments are inherently susceptible to viral outbreaks, inadvertently creating conditions that favor increased pathogen transmission and potential zoonotic spread. Metagenomic sequencing has proven to be a useful tool for characterizing the microbial burden in both people, livestock, and environmental samples. International efforts have been successful at characterizing pathogens in commercial farming environments, especially swine farms, however it is unclear whether the full extent of microbial agents have been adequately captured or is representative of farms elsewhere. To augment international efforts we performed metagenomic next-generation sequencing on nine swine slurry and three environmental samples from a United States of America (U.S.A.) farm operation, characterized the microbial composition of slurry, and identified novel viruses. We assembled a remarkable total of 1792 viral genomes, of which 554 were novel/divergent. We assembled 1637 *Picobirnavirus* genome segments, of which 538 are novel. In addition, we discovered 10 new viruses belonging to a novel taxon: porcine *Statoviruses*; which have only been previously reported in human, macaques, mouse, and cows. We assembled 3 divergent *Posaviruses* and 3 swine *Picornaviruses*. In addition to viruses described, we found other eukaryotic genera such as *Entamoeba* and *Blastocystis*, and bacterial genera such as *Listeria*, *Treponema*, *Peptoclostridium* and *Bordetella* in the slurry. Of these, two species *Entamoeba histolytica* and *Listeria monocytogenes* known to cause human disease were detected. Further, antimicrobial resistance genes such as tetracycline and MLS (macrolide, lincosamide, streptogramin) were also identified. Metagenomic surveillance in swine fecal slurry has great potential for novel and antimicrobial resistant pathogen detection.

In modern large-scale farming environments, humans and dynamic populations of livestock are often in frequent and prolonged close contact. Swine are recognized to harbor multiple pathogens which may spillover to humans[1,2]. In such farms, viruses are readily transmitted within the farm and across species. The transmission occurs through direct contact and environmental pathways such as via aerosol, feces, and water[3–7]. Globalization of the swine industry has contributed to the emergence and global spread of pathogens of swine: porcine epidemic diarrhea virus spread from China to the United States of America (U.S.A.) in 2013, affecting 50% of the U.S.A. breeding herds; and African swine fever that emerged in 2007 and re-introduced in China in 2018 which resulted in the culling of over 300 million pigs and direct economic losses of $141 billion U.S.A. dollars[8,9]. Of particular concern is the movement of zoonotic pathogens between livestock and farm workers and the threat of epidemic pathogen transmission to neighboring communities, exemplified by the H1N1 "swine flu" pandemic in 2009 that

[1]Weill Institute for Neurosciences, University of California, San Francisco, CA 94158, USA. [2]Department of Neurology, University of California, San Francisco, CA 94158, USA. [3]Division of Infectious Diseases, Duke University School of Medicine, Durham, NC, USA. [4]Julia Jones Matthews Department of Public Health, Texas Tech University Health Sciences Center, Abilene, TX, USA. [5]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA. [6]Division of Infectious Diseases, University of California San Francisco, San Francisco, CA 94158, USA. [7]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158, USA. [8]Department of Bioethics and Interdisciplinary Studies, Brody School of Medicine, North Carolina Agromedicine Institute, East Carolina University, Greenville, NC, USA. [9]Duke Global Health Institute, Duke University, Durham, NC, USA. [10]Emerging Infectious Disease Program, Duke-NUS Medical School, Singapore, Singapore. [11]Global Health Center, Duke Kunshan University, Kunshan, China. [12]These authors contributed equally: Akshaya Ramesh and Emily S. Bailey. ✉email: emily.bailey@ttuhsc.edu

originated from influenza A viruses circulating in pig populations[1,2,10–13]. These examples highlight the need to build a complete unbiased global picture of swine pathogens to understand patterns of emergence and spread[13]. Unbiased metagenomic sequencing is an important and proven tool to preemptively investigate the potential disease-causing microbial burden of livestock environments[14–16].

A 2018 review article, analyzing over 57,000 publications over the last 50 years, has identified the top 40 priority pathogens (including zoonotic) for swine diseases, with viruses (40%) and bacteria (37.5%) accounting for a majority of the pathogens[9]. Swine farms have been the subject of prior metagenomic sequencing efforts, mostly outside North America[17–21]. These investigations have yielded a plethora of rich information, especially with respect to possible viral pathogens, and suggest that additional exploratory screening should take place[17–21].

While obvious, engaging agribusinesses in emerging or reemerging pathogens surveillance is challenging[22]. Major industry objections that must be overcome include the biosecurity risks of permitting researchers to enter farms, the harm that specimen collection may cause the animals, and the often unspoken concerns that surveillance may reveal occupational hazards that could damage revenue[23]. Some of these objections can be met using noninvasive metagenomic surveillance techniques[24,25].

In this study, we build upon previous efforts using metagenomic sequencing methods to examine swine fecal slurry samples from a U.S.A. farm for molecular evidence of pathogens. Our primary goal was to determine if indirect and noninvasive swine slurry sampling could yield robust microbial detections in order to support pathogen surveillance at the human-animal interface in North America. We focus our results on the viral families identified, given that they account for the highest number of pathogens in the priority swine (and zoonotic) pathogen list, and global economic impact[8,13,26]. Here, we report that even limited sampling can reveal a rich, highly informative pathogen landscape, replete with both known and novel viral entities.

Results

Overview of mNGS results. Nine fecal slurry samples were sequenced at an average depth of 71,585,514 reads (Interquartile Range (IQR): 53,008,460–94,664,576), with an average of 22,402,963 non-host reads (IQR: 16,368,802–26,021,652). Three environmental samples were sequenced at an average depth of 36,654,247 reads (IQR: 8,048,069–53,255,240), with an average of 12,818,139 non-host reads (IQR: 1,986,958–18,978,324). The water control was sequenced at a depth of 920,074 reads.

A total of 270 microbial genera meeting our criteria were found in the slurry samples. Figure 1A shows a breakdown of the microbial composition of the pig slurry. Bacteria were the most prevalent kingdom (54.81%), followed by Eukaryota (38.52%), Viruses (4.04%) and Archaea (2.59%). Figure 1B represents the top 25 most abundant genera within each kingdom that were identified. The most prevalent bacteria genera were *Oscilibacter* and *Treponema*. Among viruses, *Posavirus*, *Picobirnavirus* and *Mamastrovirus* were the most prevalent genera. As mentioned in the introduction, we have focused our results mainly on the viruses identified in the swine slurries, and offer a brief description of the microbial compositions of other kingdoms. Supplemental Table 1 has detailed breakdown of all the microbial genera found across all nine slurry and 3 farm environment (environment and aerosol) samples.

Bacteria and antimicrobial resistance genes. The most prevalent bacterial genera present across all samples (nine slurry + three environment) are: *Veillonella* (predominantly *V. parvula*), *Roseburia* (predominantly *R. hominis* and *R. intestinalis*), *Phascolarctobacterium* (predominantly *P. faecium*), *Hungateiclostridium* (predominantly *H. clariflavum* and *H. thermocellum*), *Flavonifractor* (predominantly *F. plautii*), *Coprococcus* (predominantly *C. catus*), *Barnesiella* (predominantly *B. viscericola*), *Listeria* (predominantly *L. monocytogenes*), *Chlamydia* (predominantly *C. suis*), *Mycoplamsa* (predominantly *M. fermentans* and *M. capricolum*), and *Treponema* (predominantly *T. brennaborense*) (Fig. 1B, Supplemental Table 1). In addition to these, other notable bacterial genera identified included *Peptoclostridium* (predominantly *P. acidaminophilum*) and *Bordetella* (predominantly *B. bronchialis*) found in 75% and 50% of the samples, respectively.

In addition to bacteria, we detected a diversity of antimicrobial resistance genes.

The most prevalent class of antibiotic resistant genes detected in the nine slurry (and 3 environmental) samples were tetracycline and MLS (macrolide, lincosamide, streptogramin) both of which were detected in 100% of the samples using a 10% allele coverage level (Table 1).

Eukaryotes. Eukaryotic genera found in these swine slurry samples (nine slurry + three environment) included *Trichomonas* (predominantly *T. vaginalis*) (25%), *Entamoeba* (predominantly *E. histolytica)* (41.66%), and *Blastocystis* (predominantly *B. hominis)* (75%) (Supplemental Table 1). Additionally, plant genera such as *Lupinus*, *Capsicum*, *Zea, Saccharomyces* and *Arachis* were also present in over 50% of the samples.

Viruses. *Picobirnavirus. Picobirnavirus* was the most prevalent genus in the slurry samples (Fig. 1B). A total of 638 RNA-dependent RNA polymerase (RdRP) segments and 1033 capsid/ORF segments greater than 1 KB were assembled across all nine slurry samples; no viral genomes could be assembled from the farm environment samples (viral reads to *Picobirnavirus* were identified in two of three environment samples). Phylogenetic analysis of all complete RdRP segments in this study (354/638) and all complete *Picobornavirus* genomes from NCBI indicates that this genus was highly diverse, and belong to genogroups I and II. The porcine picobirnaviruses were not limited by host range (Supplemental Fig. 1). BLASTN and BLASTX percent identities of all RdRP segments had closest GenBank relatives with percent identities between 70.96 and 100% (mean: 85.93%, 50 segments had no BLASTN hits) and 48.1–99.56% (mean: 84.07%, 3 segments had no BLASTX hits), respectively. Of these, 31/683 RdRP genomes were divergent, and less than 75% identical (lower identity between BLASTN and BLASTX) to their closest sequence relative on GenBank. Capsid/ORF gene segments, had BLASTN and

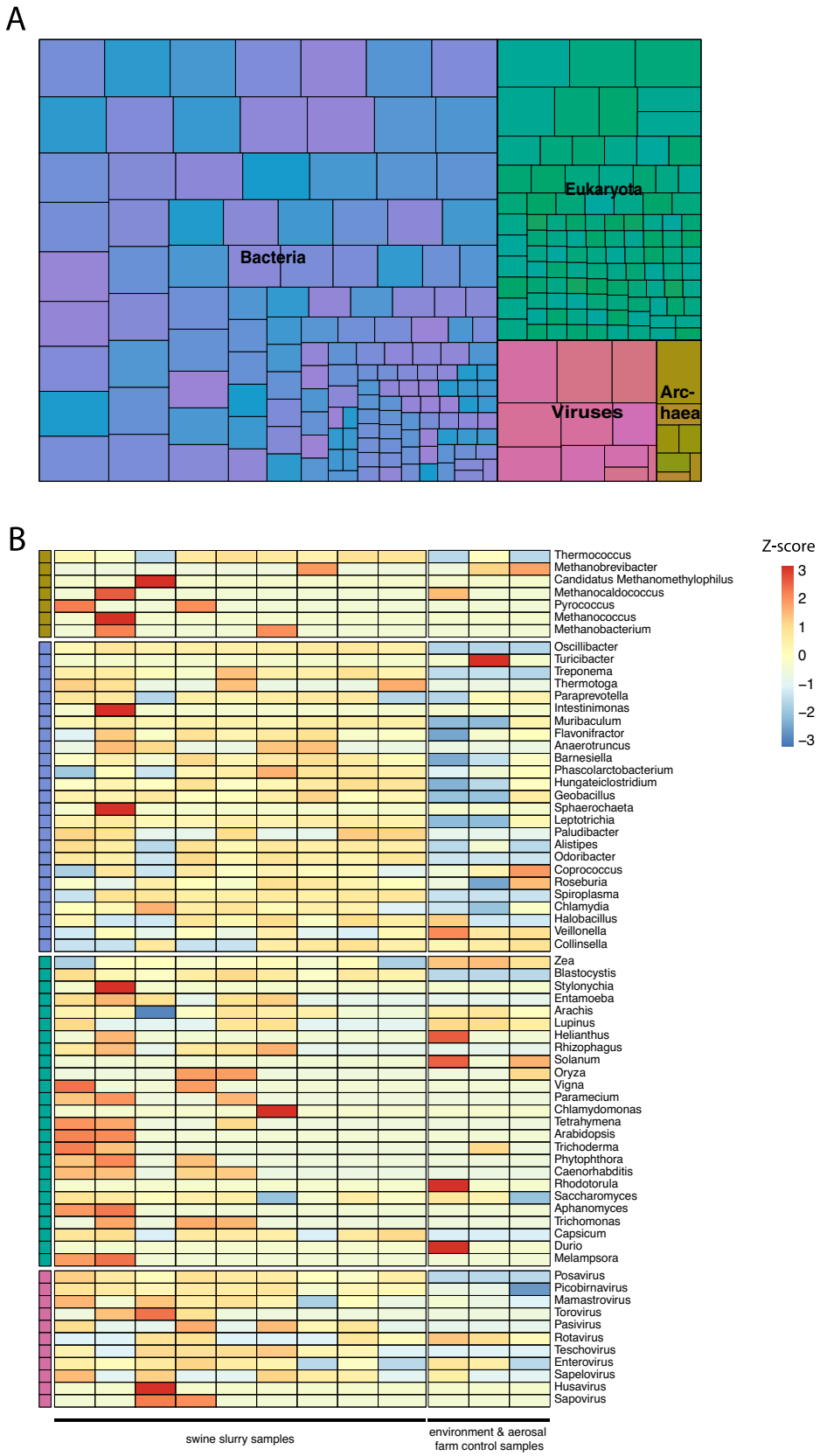

Figure 1. Microbial composition of swine slurry: (**A**) Treemap of microbial kingdoms (genus count) identified across all the nine swine slurry samples. (**B**) Top 25 most abundant genera (combined NT + NR reads per million) within each kingdom identified in the slurry samples.

AMR gene	# Gene Alleles detected in samples (% prevalence in samples)	Mean rpM
Tet	67 (100%)	80
MLS	66 (100%)	82
AGly	29 (92%)	79
Bla	17 (100%)	78
Phe	3 (25%)	88
Nim	2 (17%)	102
Sul	1 (8%)	66

Table 1. Antimicrobial resistance genes detected in 9 slurry samples and 3 environmental samples. *AMR* antimicrobial resistance, *rpM* reads per million mapped reads, *tet* tetracycline, *MLS* macrolide, lincosamide, streptogramin, *AGly* aminoglycoside, *Bla* beta-lactamas, *Phe* phenicol, *Nim* nitromidazole, *Sul* sulfa.

BLASTX percent identities of 63.4–100% (mean: 76.4%, 17 segments had no BLASTN hits) and 19.26–93.1% (mean: 42.2%, 5 segments had no BLASTX hits), respectively. 507/1033 capsid/ORF segments were divergent.

Posavirus. Followed by *Picobirnavirus*, *Posavirus* was most prevalent in the slurry samples (Fig. 1B, Supplemental Table 1). A total of 22 genomes, grouped into 10 clusters (Fig. 2A), were assembled across the nine samples. No reads to *Posavirus* were detected across farm environment samples. Among these, 17 genomes had > 85% identities (BLASTN, BLASTX) compared to existing *Posavirus* genomes on GenBank: *Posavirus* 1 (5/17), *Posavirus* strain 8805 (5/17), *Posavirus* strain 11,038 (4/17), *Posavirus* 3 (1/17), *Posavirus* 4 (1/17) and *Posavirus* strain 10,835 (1/17). Of the remaining five genomes, 2/22 were most closely related to *Posavirus* 2 (2/22) with BLASTN and BLASTX of ~ 74% and ~ 78% respectively. The other 3/22 genomes, were highly divergent, with no BLASTN hits, and closest BLASTX hits were ~ 50% [mean nucleotide % identity: 50.2% (range: 50.15–50.39%); mean amino acid % identity: 50.2% (range: 39.89–39.96% 40.07%)] to their closest relative on GenBank (KT833072.1). These 3 divergent genomes had pairwise identities of 49.46, 57.99 and 55.05% similar to each other, and a mean of 26.57% (IQR: 23.07–30.58%) to the other 19 *Posavirus* genomes at the nucleotide level.

Porcine astrovirus. A total of 36 genomes, of Porcine astrovirus types 1, 2, 4 and 5 were assembled across the nine samples. While reads to porcine astrovirus were identified in all three farm environment samples, no genomes could be assembled. BLASTN and BLASTX of these viral genomes to the NCBI server revealed that the percent identities of their respective closest genomes to GenBank had a mean of 90.4% and 90.04%, with a range between 88.56–92.49% and 63–96.9%, respectively.

Sapelovirus, Enterovirus G, Teschovirus A, Swine picornavirus and Pasivirus A. Forty-nine genomes across four different genera in the Family *Picornaviridae* were identified in the nine samples. Specifically, eight genomes of Enterovirus G representing five different enterovirus types were present. Interestingly, only three of eight genomes also had a papain-like cysteine protease sequence (PL-CP) in the 2C–3A junction region. Thirteen genomes of *Teschovirus* A were identified across all nine samples, representing eight different types. Four *Sapelovirus* A genomes, all consisting of a single serotype were assembled in four different samples. Four genomes of swine picornavirus (% identities between genomes: 58.54–99.9% (median: 58.79%)), were assembled across three different samples. Three of the four swine *picornavirus* genomes, were less than 75% identical at the NT level, [VP1 (mean: 61.74, range: 57.38–70.32%), 2C: (mean: 68.23, range: 67.93–68.85%), 3C: (mean: 60.92, range: 58.63–65.33%), 3D: (mean: 67.38, range: 67.14–67.86%)] compared to their closest swine picornavirus genomes in GenBank (LC113907), classifying them as a novel species. Of these three novel species, 2 were almost identical (pairwise identity 99.27%) to each other. Finally, 19 *Pasivirus* A genomes across five samples were identified, with % identities between genomes 70.55–99.6% (median: 75.45%). No divergent viruses were detected in the other viral families: *Sapelovirus*, *Enterovirus* G, *Teschovirus* A, and *Pasivirus* A. While no genomes were assembled from the farm environmental samples, reads to these genus of viruses were identified in the control environment samples as well (see Supplemental Table 1 for further breakdown).

Statovirus. Ten porcine *Statoviruses*, a novel taxon described previously only in human, macaques, mouse, and cows, was identified across five slurry samples (Fig. 2B). Of these, 5/10 were > 99% similar to each other at the NT level and 2/10 were 90% similar to each other at the NT level, with the other 3/10 genomes sharing between 75 and 85% percent identity.

Other viral genera. Additionally, other viral genera of interest that were identified in the pig slurry samples included *Sapovirus, Torovirus, Husavirus, Rotavirus*. Of these, there was sufficient sequencing coverage to assemble one *Sapovirus* and three *Torovirus* genomes successfully.

Finally, a large number of reads (63,043.25–158,443.0213 NT rPM) were labelled as "uncategorized", (i.e. taxa with neither family nor genus classification were found in all samples), including in the water control (23,680.98 rPM). Due to our exclusion criteria, microbial genera identified in the water control were not reported in the samples. It is likely that further analyses of these "uncategorized" reads, will be revealing, and identify further previously unidentified and uncharacterized microbes.

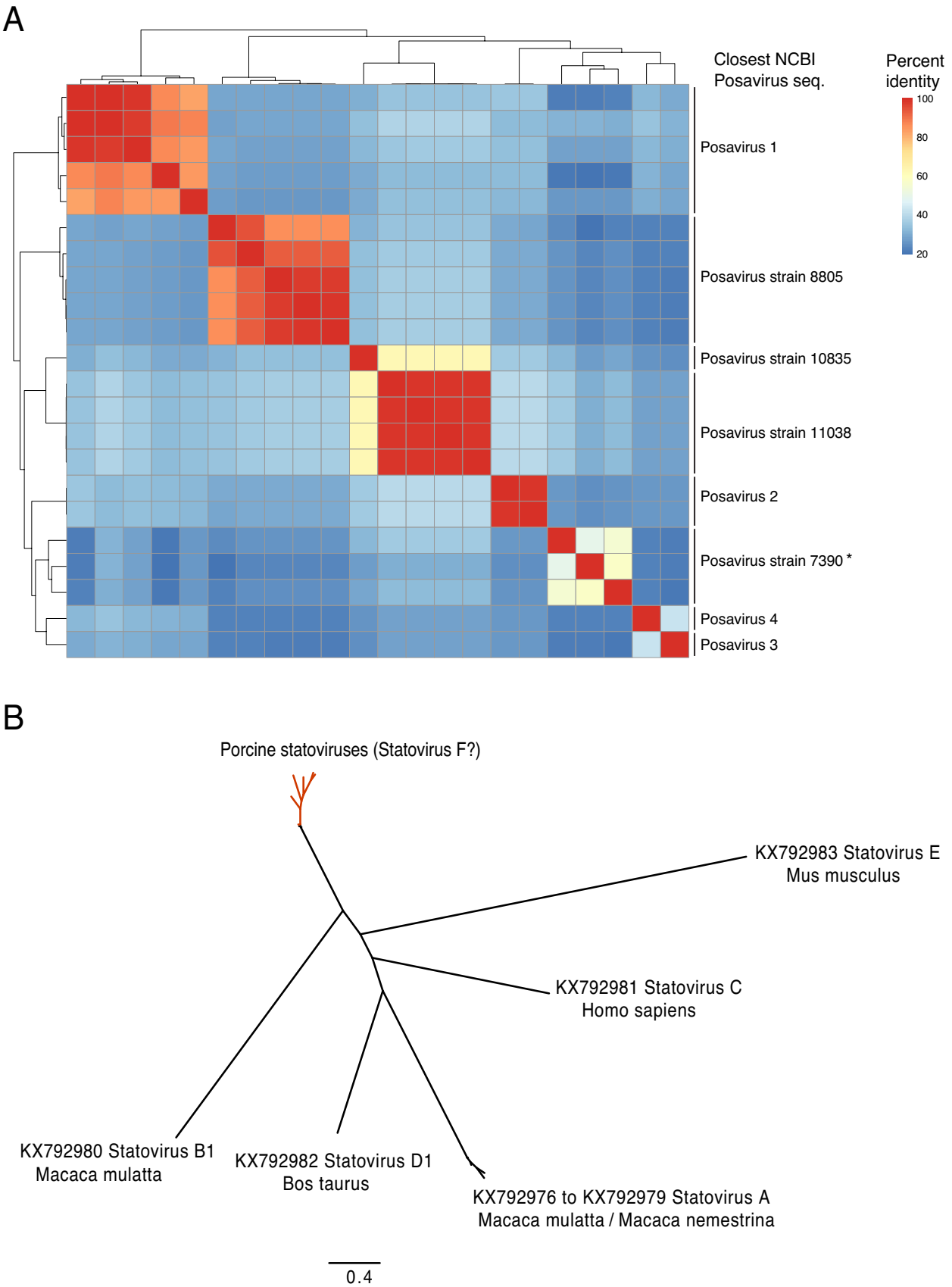

Figure 2. (**A**) Similarity matrix of *Posavirus* assembled in this study. (**B**) Maximum Likelihood phylogenetic tree of a novel taxon of porcine *Statovirus* genomes assembled in this study.

Discussion

Modern high-density swine farming practices facilitate novel pathogen emergence and horizontal transmission, and increase the risk of zoonotic spillover to other species, including humans[27]. This is due to multiple factors including crowded livestock conditions, the frequent introduction of immunologically naïve young animals, the frequent movement of animals[28], and environmental transmission pathways. As most high-density animal operations process animal waste in large lagoons, these pools pose an additional transmission risk to neighboring farm and communities through ground water seepage and large rainwater runoffs. GPS mapping of animal production locations in North Carolina can demonstrate that the proximity of animal facilities can pose public health risks of viral spread from one farm to another through simple measures such as feed delivery[29–31]. Communication between farmers regarding biosecurity and emerging viruses is essential to the public health of animals and humans working in production environments as well as surrounding communities.

Our pilot metagenomic study of swine slurry captures the virome, both known and novel, as well as microbes across the tree of life, revealing tremendous diversity and complexity. Given that viruses account for the largest number of swine pathogens, and their global economic impact, we have focused the results of our study on this kingdom[8,13]. In this study alone, we report 554 novel viruses of unknown pathogenicity and zoonotic potential. Further, *Entamoeba histolytica* and *Listeria monocytogenes*, known human pathogens of concern were identified. We also demonstrated how fecal slurry can be used to examine antimicrobial resistance genes in bacteria prevalent in swine.

In this study, *Picobirnaviruses* represented a substantial fraction (30%, 1671 genomes) of the viral sequences. *Picobirnaviruses* are non-enveloped double-stranded RNA viruses, typically consisting of two segments and are often found in the fecal matter of a wide range of species, sometimes associated with diarrhea[32]. A large number of diverse *Picorbirnaviruses* (n = 1236) have also been previously reported in camel feces[33]. This study adds 1671 genomes, to the existing 2540 genomes of *Picobirnaviruses* currently publicly available in NCBI, greatly increasing our sequence knowledge of this group of viruses within the Durnavirales order. The true host of *Picobirnaviruses*, whether it be eukaryotic (including fungal), or prokaryotic, remains unresolved[34]. Whether *Picobirnaviruses* are causes, or indirect markers of enteric disease also remains controversial.

Posaviruses (Porcine stool-associated RNA viruses) are highly diverse members of *Picornavirales* and widely observed and first discovered in fecal samples from pigs[35–37]. In addition to previously known *Posavirus* types, we identified three highly divergent *Posaviruses*. Further, we report the first porcine *Statovirus* (Stool Associated Tombus-like virus); to-date these viruses have only been reported in the gastrointestinal tract of four other mammals: humans, macaques, cows and mice[38]. As is the case with *Picobirnaviruses,* the true host and pathogenicity of these viruses are unclear.

Unlike the above mentioned viral genera, there is evidence of porcine host infection by Astroviruses, another positive-sense single stranded non-enveloped virus that has been studied extensively[35,39,40]. Aside from enteric and respiratory infection, porcine astrovirus type 3 in particular has been associated with encephalomyelitis in pigs[41]. However in this study, we only detected porcine astrovirus types 1, 2, 4 and 5.

In addition to viral families discussed above, other microbial species that are known to be pathogenic to humans were detected. *E. histolytica*, a human pathogen associated with intestinal and extraintestinal infections, responsible for 5 million infections annually, was identified in 41.66% of the samples[42]. Further, *Listeria monocytogenes*, a serious foodborne illness with a high mortality and hospitalization rates was identified in all slurry samples studied[43]. We also detected antimicrobial resistance genes in these samples. The genes identified have been implicated in both human disease and food safety[27,44]; however due to a lack of a understanding of how genomic detection of these genes compares to the gold-standard, the significance of these findings are unclear[45].

Our exploratory pilot study has important limitations. First, only a small number (nine) of slurry samples were collected from two farms. Second, these samples were collected over a limited time period of 6 months; larger and wider sampling is likely to reveal temporal dynamics of the microbial species. Finally, we have focused our analyses and discussion only on viral families and further analyses of the other microbial families (including "uncategorized" reads) is likely to yield a more complex and diverse microbial portrait. Nonetheless, from two moderately-sized farms with an average of 3800 head of swine, we identified a rich and diverse microbial landscape, including 554 novel viruses using our small sample set. More extensive and longitudinal studies of fecal slurry from a large US farm (20,000 head) or a megafarm in China (> 84,000 head)[46] would likely yield even greater diversity. Just as studies of human sewage have assisted public health officials in understanding polio virus[47] and SARS-CoV-2 transmission[48], our results suggest that periodic assessments of swine farm fecal slurry might be an effective noninvasive approach to novel pathogen surveillance at industrialized swine farms. Such surveillance should be attractive to the swine industry as an early warning method for swine pathogen incursions[49] and additionally, assist public health officials in assessing possible swine zoonoses public health threats[50].

Materials and methods

Site enrollment. During the fall of 2018, A North Carolina, USA, swine farm with two geographic locations was identified to participate in this preliminary study. Each barn had 12 pens with a center hallway and a fully slatted concrete floor over a deep pit to hold feces, urine, and waste water.at Location #1 and slightly larger at location #2. This farm held on average 3800 head of swine during a year. Pits were emptied up to three times per year and recharged with recycled water. Farm personnel collected up to two (6 oz. syringes) slurry samples per week from the pit of each production location.

This study was granted exemption from review status by the Institutional Animal Care and Use Committee at Duke University on the grounds that the research did not include direct sample collection from animals.

Sample collection and processing. Slurry is defined as the feces and urine from pigs and the waste water used to remove the urine and feces from the pig pens[51]. Nine slurry samples from two swine barns containing finishing pigs were collected from approximately 5–10 cm below the surface of pits. Three environmental samples (two aerosol and one surface swab) were also collected and were frozen at – 20 °C until shipped to our laboratory (a maximum of 24 h). Aerosol sample and environmental swabs were collected as previously described[52]. Frozen samples and completed surveys were transported overnight to the Duke One Health Research Laboratory. Dates and pre-assigned sample numbers were used for sample tracking.

Slurry samples were diluted by methods previously described[25] and genomic DNA was extracted using the Zymo Research *Quick*-DNA Miniprep Kit (Cat. No. D3024). Extracted samples were shipped to the Chan-Zuckerberg Biohub (San Francisco, California) and stored at – 80 °C until processed by molecular methods.

Library preparation and metagenomic next-generation sequencing (mNGS). RNA was extracted and Libraries were prepared using NEBNext Ultra II RNA Library Prep Kit for Illumina (Cat. No. E7770) using Mayday et al. (2019)[53].

External RNA Controls 103 Consortium (ERCC) [ThermoFisher, catalog no 4456740] collection spike-in controls were used in all samples. Libraries were pooled and sequenced on a NextSeq to generate 150 base pair, paired-end sequences.

Microbe identification and bioinformatic analysis. Microbial pathogens were identified from raw sequencing reads using IDseqV3.9, a cloud-based, open-source bioinformatics platform recently described in Ramesh et al. (2019), Saha et al. (2019), and Kalantar et al. (2020)[54–56]. Only microbial genera present at over 10 rpM, both at the nucleotide (NT) and protein (NR) levels were reported. Additionally, to control for any potential background contaminants, all microbial genera identified in the water control were excluded.

To more comprehensively characterize the genomes of viruses in the metagenomic dataset, genomes of identified microbes were assembled de novo using St. Petersburg genome assembler (SPAdes)[57] and annotated using Geneious v10.3.2. Contigs assembled that were 50% or greater than the Reference genome were included in the downstream analysis. For the *Picobirnavirus* genus, contigs greater than 1 KB were included in the analysis. For a given viral species, the assembled genomes were then aligned using the default settings in MUSCLEv3.8.1551. ModelTest-NGv0.1.5 was used to identify the best-fitting evolutionary model for each viral species. We reconstructed a maximum-likelihood phylogeny using RAxML-ngv0.6.0 using default settings (bootstrap = 200).

Criteria for identification of novel or divergent virus. Picobirnavirus virus was categorized as novel/divergent if they were less than 75% identical to their closest relative on GenBank (BLASTN and BLASTX)[58]. *Posavirus* were classified as divergent if their BLASTN and BLASTX identities were below 50%[35]. Novel species in swine picornavirus was identified based on ICTV guidelines[59].

Antimicrobial resistance (AMR) assessment. To identify antimicrobial resistance genes present in the metagenomic dataset, SRST2 was used[60]. AMR genes with at least 10% allele coverage were considered positive.

Ethics approval. This study was granted exemption from review status by the Institutional Animal Care and Use Committee at Duke University on the grounds that the research did not include direct sample collection from animals.

Data availability

All of the raw data generated for this study is available PRJNA683083. Assembled genomes can be found at MW977024—MW977661, MW977662—MW978694, and MW504477-MW504597.

Received: 17 June 2021; Accepted: 30 July 2021
Published online: 20 August 2021

References

1. Uddin Khan, S., Atanasova, K. R., Krueger, W. S., Ramirez, A. & Gray, G. C. Epidemiology, geographical distribution, and economic consequences of swine zoonoses: A narrative review. *Emerg. Microb. Infect.* **2**(12), e92. https://doi.org/10.1038/emi.2013.87 (2013).
2. Smith, T. C. *et al.* Emerging swine zoonoses. *Vector Borne Zoonotic Dis.* **11**(9), 1225–1234. https://doi.org/10.1089/vbz.2010.0182 (2011).
3. Ma, M. J. *et al.* Evidence for cross-species influenza a virus transmission within swine farms, China: A one health, prospective cohort study. *Clin. Infect. Dis.* **66**(4), 533–540. https://doi.org/10.1093/cid/cix823 (2018).
4. Anderson, B. D. *et al.* Prospective surveillance for influenza A virus in Chinese swine farms. *Emerg. Microb. Infect.* **7**(1), 87. https://doi.org/10.1038/s41426-018-0086-1 (2018).
5. Yugo, D. M. & Meng, X. J. Hepatitis E virus: Foodborne, waterborne and zoonotic transmission. *Int. J. Environ. Res. Public Health* **10**(10), 4507–4533. https://doi.org/10.3390/ijerph10104507 (2013).
6. Vonesch, N. *et al.* Emerging zoonotic viral infections of occupational health importance. *Pathog. Dis.* https://doi.org/10.1093/femspd/ftz018 (2019).
7. Mattison, K. *et al.* Human noroviruses in swine and cattle. *Emerg. Infect. Dis.* **13**(8), 1184–1188. https://doi.org/10.3201/eid1308.070005 (2007).
8. EFSA Panel on Animal Health and Welfare *et al.* Scientific opinion on the assessment of the control measures of the category A diseases of Animal Health Law: African swine fever. *EFSA J.* **19**(1), e06402. https://doi.org/10.2903/j.efsa.2021.6402 (2021).
9. VanderWaal, K. & Deen, J. Global trends in infectious diseases of swine. *Proc. Natl. Acad. Sci. U. S. A.* **115**(45), 11495–11500. https://doi.org/10.1073/pnas.1806068115 (2018).

10. Forgie, S. E. *et al.* Swine outbreak of pandemic influenza A virus on a Canadian research farm supports human-to-swine transmission. *Clin. Infect. Dis.* **52**(1), 10–18. https://doi.org/10.1093/cid/ciq030 (2011).
11. Poljak, Z. *et al.* Prevalence of and risk factors for influenza in southern Ontario swine herds in 2001 and 2003. *Can. J. Vet. Res.* **72**(1), 7–17 (2008).
12. Borkenhagen, L. K. *et al.* High risk of influenza virus infection among swine workers: Examining a dynamic cohort in China. *Clin. Infect. Dis.* **71**(3), 622–629 (2020).
13. Smith, G. J. *et al.* Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**(7250), 1122–1125. https://doi.org/10.1038/nature08182 (2009).
14. Kubacki, J., Fraefel, C. & Bachofen, C. Implementation of next-generation sequencing for virus identification in veterinary diagnostic laboratories. *J. Vet. Diagn. Invest.* **33**(2), 235–247. https://doi.org/10.1177/1040638720982630 (2021).
15. Belak, S., Karlsson, O. E., Blomstrom, A. L., Berg, M. & Granberg, F. New viruses in veterinary medicine, detected by metagenomic approaches. *Vet. Microbiol.* **165**(1–2), 95–101. https://doi.org/10.1016/j.vetmic.2013.01.022 (2013).
16. Rubiola, S., Chiesa, F., Dalmasso, A., Di Ciccio, P. & Civera, T. Detection of antimicrobial resistance genes in the milk production environment: Impact of host DNA and sequencing depth. *Front. Microbiol.* **11**, 1983. https://doi.org/10.3389/fmicb.2020.01983 (2020).
17. Jiang, H., Fang, S., Yang, H. & Chen, C. Identification of the relationship between the gut microbiome and feed efficiency in a commercial pig cohort. *J. Anim. Sci.* https://doi.org/10.1093/jas/skab045 (2021).
18. Li, Y., Zhu, Y., Wei, H., Chen, Y. & Shang, H. Study on the diversity and function of gut microbiota in pigs following long-term antibiotic and antibiotic-free breeding. *Curr. Microbiol.* **77**(12), 4114–4128. https://doi.org/10.1007/s00284-020-02240-8 (2020).
19. Wang, C. *et al.* Characterization of the pig gut microbiome and antibiotic resistome in industrialized feedlots in China. *mSystems* https://doi.org/10.1128/mSystems.00206-19 (2019).
20. Wang, C. *et al.* Characteristics of gut microbiota in pigs with different breeds, growth periods and genders. *Microb. Biotechnol.* https://doi.org/10.1111/1751-7915.13755 (2021).
21. Zhang, R. M. *et al.* Distribution patterns of antibiotic resistance genes and their bacterial hosts in pig farm wastewater treatment systems and soil fertilized with pig manure. *Sci. Total Environ.* **758**, 143654. https://doi.org/10.1016/j.scitotenv.2020.143654 (2021).
22. Gray, G. C. M. J. To succeed, one health must win animal agriculture's stronger collaboration. *Clin. Infect. Dis.* https://doi.org/10.1093/cid/ciz729 (2019).
23. Gray, G. C. & Merchant, J. A. Pigs, pathogens, and public health. *Lancet Infect. Dis.* **18**(4), 372–373. https://doi.org/10.1016/S1473-3099(18)30158-0 (2018).
24. Anderson, B. D., Lednicky, J. A., Torremorell, M. & Gray, G. C. The use of bioaerosol sampling for airborne virus surveillance in swine production facilities: A mini review. *Front. Vet. Sci.* **4**, 121. https://doi.org/10.3389/fvets.2017.00121 (2017).
25. Bailey, E. S. *et al.* A feasibility study of conducting surveillance for swine pathogens in slurry from North Carolina swine farms. *Sci. Rep.* **10**(1), 10059. https://doi.org/10.1038/s41598-020-67313-x (2020).
26. Mason-D'Croz, D. *et al.* Modelling the global economic consequences of a major African swine fever outbreak in China. *Nat. Food* **1**(4), 221–228. https://doi.org/10.1038/s43016-020-0057-2 (2020).
27. Smith, T. C. *et al.* Geographic distribution of livestock-associated *Staphylococcus aureus* in the United States. *Microb. Infect.* **20**(6), 323–327. https://doi.org/10.1016/j.micinf.2018.05.004 (2018).
28. Nelson, M. I. *et al.* The emergence and evolution of influenza A (H1alpha) viruses in swine in Canada and the United States. *J. Gen. Virol.* **98**(11), 2663–2675. https://doi.org/10.1099/jgv.0.000924 (2017).
29. Stoddard, E. A. & Hovorka, A. Animals, vulnerability and global environmental change: The case of farmed pigs in concentrated animal feeding operations in North Carolina. *Geoforum* **100**, 153–165 (2019).
30. Handan-Nader, C. & Ho, D. E. Deep learning to map concentrated animal feeding operations. *Nat. Sustain.* **2**(4), 298–306 (2019).
31. Bucini, G. *et al.* Risk attitudes affect livestock biosecurity decisions with ramifications for disease control in a simulated production system. *Front. Vet. Sci.* **6**, 196. https://doi.org/10.3389/fvets.2019.00196 (2019).
32. Kashnikov, A. Y., Epifanova, N. V. & Novikova, N. A. Picobirnaviruses: Prevalence, genetic diversity, detection methods. *Vavilovskii Zhurnal Genet Selektsii* **24**(6), 661–672. https://doi.org/10.18699/VJ20.660 (2020).
33. Woo, P. C. *et al.* Metagenomic analysis of viromes of dromedary camel fecal samples reveals large number and high diversity of circoviruses and picobirnaviruses. *Virology* **471–473**, 117–125. https://doi.org/10.1016/j.virol.2014.09.020 (2014).
34. Ghosh, S. & Malik, Y. S. The true host/s of picobirnaviruses. *Front. Vet. Sci.* **7**, 615293. https://doi.org/10.3389/fvets.2020.615293 (2020).
35. Shan, T. *et al.* The fecal virome of pigs on a high-density farm. *J. Virol.* **85**(22), 11697–11708. https://doi.org/10.1128/JVI.05217-11 (2011).
36. Munnink, B. B. O. *et al.* Characterization of Posa and Posa-like virus genomes in fecal samples from humans, pigs, rats, and bats collected from a single location in Vietnam. *Virus Evol.* **3**(2), vex022. https://doi.org/10.1093/ve/vex022 (2017).
37. Aoki, H. *et al.* Phylogenetic analysis of novel posaviruses detected in feces of Japanese pigs with posaviruses and posa-like viruses of vertebrates and invertebrates. *Arch. Virol.* **164**(8), 2147–2151. https://doi.org/10.1007/s00705-019-04289-8 (2019).
38. Janowski, A. B. *et al.* Statoviruses, A novel taxon of RNA viruses present in the gastrointestinal tracts of diverse mammals. *Virology* **504**, 36–44. https://doi.org/10.1016/j.virol.2017.01.010 (2017).
39. Luo, Z. *et al.* Multiple novel and prevalent astroviruses in pigs. *Vet. Microbiol.* **149**(3–4), 316–323. https://doi.org/10.1016/j.vetmic.2010.11.026 (2011).
40. Qin, Y. *et al.* Molecular epidemiology and viremia of porcine astrovirus in pigs from Guangxi province of China. *BMC Vet. Res.* **15**(1), 471. https://doi.org/10.1186/s12917-019-2217-x (2019).
41. Arruda, B. *et al.* Porcine astrovirus type 3 in central nervous system of swine with polioencephalomyelitis. *Emerg. Infect. Dis.* **23**(12), 2097–2100. https://doi.org/10.3201/eid2312.170703 (2017).
42. Project GWP. Entamoeba histolytica (2017). https://www.waterpathogens.org/book/entamoeba-histolytica.
43. Aministration UFaD. Get the facts about Listeria (2020). https://www.fda.gov/animal-veterinary/animal-health-literacy/get-facts-about-listeria#statistics.
44. Jia, S. *et al.* Fate of antibiotic resistance genes and their associations with bacterial community in livestock breeding wastewater and its receiving river water. *Water Res.* **124**, 259–268. https://doi.org/10.1016/j.watres.2017.07.061 (2017).
45. Su, M., Satola, S. W. & Read, T. D. Genome-based prediction of bacterial antibiotic resistance. *J. Clin. Microbiol.* https://doi.org/10.1128/JCM.01405-18 (2019).
46. World's largest pig farm in China will 'produce' 2 million pigs per year 2020 (cited 2021 March 18); https://www.theanimalreader.com/2020/12/11/worlds-largest-pig-farm-in-china-will-produce-2-million-pigs-per-year/.
47. Akhtar, R. *et al.* Genetic epidemiology reveals 3 chronic reservoir areas with recurrent population mobility challenging poliovirus eradication in Pakistan. *Clin. Infect. Dis.* **71**(7), e58–e67. https://doi.org/10.1093/cid/ciz1037 (2020).
48. Martin, J. *et al.* Tracking SARS-CoV-2 in sewage: Evidence of changes in virus variant predominance during COVID-19 pandemic. *Viruses* https://doi.org/10.3390/v12101144 (2020).
49. Wang, Q., Vlasova, A. N., Kenney, S. P. & Saif, L. J. Emerging and re-emerging coronaviruses in pigs. *Curr. Opin. Virol.* **34**, 39–49. https://doi.org/10.1016/j.coviro.2018.12.001 (2019).
50. Centers for Disease Control and Prevention (CDC). Swine-origin influenza A (H3N2) virus infection in two children–Indiana and Pennsylvania, July–August 2011. *MMWR Morb. Mortal. Wkly. Rep.* **60**(35), 1213–1215 (2011).

51. Ziemer, C. J. *et al.* Fate and transport of zoonotic, bacterial, viral, and parasitic pathogens during swine manure treatment, storage, and land application. *J. Anim. Sci.* **88**(13 Suppl), E84–E94. https://doi.org/10.2527/jas.2009-2331 (2010).
52. Anderson, B. D. *et al.* Environmental bioaerosol surveillance as an early warning system for pathogen detection in North Carolina swine farms: A pilot study. *Transbound. Emerg. Dis.* **68**(2), 361–367. https://doi.org/10.1111/tbed.13683 (2021).
53. Mayday, M. Y., Khan, L. M., Chow, E. D., Zinter, M. S. & DeRisi, J. L. Miniaturization and optimization of 384-well compatible RNA sequencing library preparation. *PLoS ONE* **14**(1), e0206194. https://doi.org/10.1371/journal.pone.0206194 (2019).
54. Ramesh, A. *et al.* Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLoS ONE* **14**(6), e0218318. https://doi.org/10.1371/journal.pone.0218318 (2019).
55. Saha, S. *et al.* Unbiased metagenomic sequencing for pediatric meningitis in Bangladesh reveals neuroinvasive chikungunya virus outbreak and other unrealized pathogens. *MBio* https://doi.org/10.1128/mBio.02877-19 (2019).
56. Kalantar, K. L. *et al.* IDseq-An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience* https://doi.org/10.1093/gigascience/giaa111 (2020).
57. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**(5), 455–477. https://doi.org/10.1089/cmb.2012.0021 (2012).
58. Krishnamurthy, S. R. & Wang, D. Extensive conservation of prokaryotic ribosomal binding sites in known and novel picobirnaviruses. *Virology* **516**, 108–114. https://doi.org/10.1016/j.virol.2018.01.006 (2018).
59. Gorbalenya, A. E., Knowles, N. J., Lindberg, M., Oberste, S., Palmenberg, A. C., Reuter, G., Simmonds, P., Skern, T., Tapparel, C., Wolthers, K., Woo, P. & Zell, R. Picornaviridae (2017). https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/picornaviridae.
60. Inouye, M. *et al.* SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genom. Med.* **6**(11), 90. https://doi.org/10.1186/s13073-014-0090-6 (2014).

Acknowledgements

This study was supported by pilot funds from the North Carolina Agromedicine Institute, an interinstitutional collaboration of East Carolina University, North Carolina State University, and North Carolina A&T State University and by Duke University discretionary funding (Gray PI). We thank Sarah Paust, Calvin Wang, Christine Wang, Julie Zemke, and Anfal Abdelgadir for assisting with sample extraction.

Author contributions

E.S.B. collected samples and conducted initial professing. E.S.B. and V.A. extracted samples and prepared viral libraries. V.A., C.L., M.P., N.N., R.S., C.T., and J.L.D. supported metagenomic study design and N.N., R.S. performed sequencing. A.R. conducted metagenomic analyses. E.S.B. and A.R. wrote the manuscript and A.G.G., J.L.D., and G.C.G. edited the final version. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-95804-y.

Correspondence and requests for materials should be addressed to E.S.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.