



## ORIGINAL ARTICLE

# Genetic characterization of *Streptococcus equi* subspecies *zooepidemicus* associated with high swine mortality in the United States

Xuhua Chen<sup>1</sup> | Nubia Resende-De-Macedo<sup>1</sup> | Panchan Sitthicharoenchai<sup>1</sup>  |  
Orhan Sahin<sup>1</sup> | Eric Burrough<sup>1</sup> | Maria Clavijo<sup>1</sup> | Rachel Derscheid<sup>1</sup> | Kent Schwartz<sup>1</sup> |  
Kristina Lantz<sup>2</sup> | Suelee Robbe-Austerman<sup>2</sup> | Rodger Main<sup>1</sup> | Ganwu Li<sup>1</sup> 

<sup>1</sup>Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, IA, USA

<sup>2</sup>National Veterinary Services Laboratories, Ames, Iowa

## Correspondence

Ganwu Li, Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, IA 50011, USA.  
Email: liganwu@iastate.edu

## Funding information

Swine Health Information Center, Grant/Award Number: SHIC#19-236

## Abstract

High mortality events due to *Streptococcus equi* subspecies *zooepidemicus* (*Streptococcus zooepidemicus*) in swine have not previously been reported in the United States. In September and October 2019, outbreaks with swine mortality up to 50% due to *S. zooepidemicus* septicaemia were reported in Ohio and Tennessee. Genomic epidemiological analysis revealed that the eight outbreak isolates were clustered together with ATCC 35246, a Chinese strain caused outbreaks with high mortality, also closely related to three isolates from human cases from Virginia, but significantly different from an outbreak-unrelated swine isolate from Arizona and most isolates from other animal species. Comparative genomic analysis on two outbreak isolates and another outbreak-unrelated isolate identified several genomic islands and virulence genes specifically in the outbreak isolates only, which are likely associated with the high mortality observed in the swine population. These findings have implications for understanding, tracking and possibly preventing diseases caused by *S. zooepidemicus* in swine.

## KEYWORDS

bacterial pathogens, emerging diseases, veterinary epidemiology

## 1 | INTRODUCTION

*Streptococcus equi* subspecies *zooepidemicus* (*Streptococcus zooepidemicus*), a beta-haemolytic and Gram-positive bacterium, is most frequently isolated as an opportunistic pathogen of horses in the upper respiratory and lower genital tracts. It can also cause infections in a wide range of other animal species, including cats, ruminants, pigs, monkeys, dogs and guinea pigs (Blum et al., 2010; Byun, Yoon, Woo, Jung, & Joo, 2009; Fan, Wang, Tang, & Lu, 2008; Gruszynski et al., 2015; Sharp, Prince, & Gibbens, 1995;

Soedarmanto, Pasaribu, Wibawan, & Lammler, 1996). *S. zooepidemicus* is zoonotic, with reported transmission from horses, dogs and guinea pigs to humans (Pelkonen et al., 2013) leading to either severe invasive diseases (bacteraemia, septic arthritis, pneumonia and meningitis) or benign diseases such as pharyngitis, with potential to trigger acute post-streptococcal glomerulonephritis (APSGN) (Torres et al., 2018). The human patients usually acquire the bacteria through direct contact with infected animals or consumption of contaminated animal products such as milk or cheese.

Xuhua Chen and Nubia Resende-De-Macedo are contributed to this work equally.

The pathogenesis of *S. zooepidemicus* is not fully understood; however, virulence factors and the host immune status appear to have roles in the development of disease. The M protein, a member of M/M-like protein family, is a surface-associated protein and a classical virulence factor in group A *Streptococcus* (GAS; e.g. *S. pyogenes*) (Metzgar & Zampolli, 2011). The M-like protein SzP of *S. zooepidemicus* was reported to contribute to the virulence in animal model studies (Kelly et al., 2006; Ma et al., 2012; Walker & Timoney, 1998), and a second M-like protein SzM has been shown to bind fibrinogen, activate plasminogen, inhibit phagocytosis and serve as a protective antigen for vaccination (Velineni & Timoney, 2013). Very recently, BifA, a Fic domain-containing protein, was shown to disrupt the blood–brain barrier integrity by activating moesin in endothelial cells (Ma et al., 2019). In addition, capsular and other surface polysaccharides are important streptococcal virulence factors, and several superantigen genes (*seeM*, *szeL* and *szeM*) (Alber, El-Sayed, Estoepongastie, Lammler, & Zschock, 2005; Paillot, Darby, et al., 2010; Paillot, Robinson, et al., 2010) and several genes associated with anti-phagocytic capability in *S. zooepidemicus* may be involved in the pathogenicity (Xu et al., 2019).

Although *S. zooepidemicus* was reported to cause epizootic outbreaks in swine resulting in significant economic losses in China and Indonesia (Ma et al., 2011; Soedarmanto et al., 1996), previous isolation of *S. zooepidemicus* from clinically ill pigs has been rather limited in the United States in past decades. The first high mortality event from *S. zooepidemicus* in North America was reported very recently in Canada in March 2019 (<https://www.biorxiv.org/content/10.1101/812636v2>). From late September to early October of 2019, three cases of cull sows and feeder pigs from Ohio and Tennessee were submitted to the Veterinary Diagnostic Laboratory at Iowa State University (ISU-VDL). High mortality ranging from 10% to 50% in groups of pigs was reported over the period of 8–10 days at the buying station in Ohio and similar high mortality (922 out of 2,222 sows in lairage) from an abattoir in Tennessee. The clinical signs included sudden death, weakness, lethargy and high fever. Splenomegaly and haemorrhagic lymph nodes were the most consistent macroscopic findings. Microscopic lesions were consistent with acute bacterial septicaemia. A laboratory diagnosis of *S. zooepidemicus* septicaemia was given, which was corroborated by histopathology, PCR and bacterial culture. To genetically characterize *S. zooepidemicus* strains associated with high mortality and gain insights into the epidemiology of these highly unusual and unexpected outbreaks, we performed whole-genome sequencing on eight isolates from the Ohio and Tennessee outbreaks, another outbreak-unrelated swine isolate from Arizona, and 15 *S. zooepidemicus* isolates from other animal species. Three full-length complete genome sequences were further assembled, and genomic epidemiological and comparative genomic analyses were conducted.

## 2 | MATERIALS AND METHODS

### 2.1 | *Streptococcus zooepidemicus* isolates

In total, twenty-four *S. zooepidemicus* isolates were whole-genome-sequenced and included in this bacterial genomic

epidemiological study. Among them, eight isolates (OH-71905, TN-74097, NVSLTN-LUNG1, NVSLTN-LUNG2, NVSLTN-LUNG3, NVSLTN-LIVER4, NVSLTN-TB1 and NVSLTN-TC1) were from Ohio and Tennessee outbreaks, another swine isolate (AZ-45470) from a case not related to the outbreaks and 15 isolates from different animal species (6 from equine, 3 from feline, 3 from guinea pig, 1 from canine, 1 from caprine and 1 from chinchilla). Another 24 strains from different countries and years with complete or draft genomes that were publicly available from GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) were included in the study. The detailed information of all 48 *S. zooepidemicus* strains is summarized in Tables S1 and S2.

### 2.2 | DNA extraction and library preparation

A single colony of each *S. zooepidemicus* strain was inoculated into Tryptic soy broth (TSB) with 10% bovine serum and incubated at 37°C with overnight shaking. ChargeSwitch gDNA Mini Bacteria Kit (Life Technologies) was used to extract genomic DNA from *S. zooepidemicus* cells following the manufacturer's guidelines. DNA quality was determined by NanoDrop (Thermo Fisher Scientific), and accurate concentration was measured by Qubit fluorometer double-stranded DNA high-sensitivity (dsDNA HS) kit (Life Technologies). At the National Veterinary Service Laboratories (NVSL), a loop of *S. zooepidemicus* from an overnight culture grown on blood agar was suspended in 400 µL of TE buffer containing 10 mM of lysozyme and incubated for 2 hr at 37°C. The entire volume was extracted using a Promega Maxwell RSC Whole Blood DNA Kit on the Promega Maxwell RSC 48 (Promega). Accurate DNA concentration was measured by Qubit fluorometer double-stranded DNA high-sensitivity (dsDNA HS) kit (Life Technologies). Indexed genomic libraries were prepared at both laboratories by using Nextera XT DNA Library Prep Kit (Life Technologies) for subsequent sequencing.

### 2.3 | Genome sequencing and assembly

Bacterial genomes were sequenced using the Illumina MiSeq platform (Illumina) with 250x2 read length in the NGS Unit of the ISU-VDL or at the NVSL. Low-quality raw reads and adapters were filtered and trimmed by Seqtk and Trimmomatic-0.36 (Bolger, Lohse, & Usadel, 2014). The filtered reads were tested for quality by FastQC and were assembled utilizing SPAdes 3.13.1-Darwin (Bankevich et al., 2012). Assembly quality was evaluated using QUAST (Mikheenko, Pribelski, Saveliev, Antipov, & Gurevich, 2018) to determine N50, longest contigs, total length of contigs, GC content and other parameters as appropriate. Contigs with low average depth ( $\leq 2$ ) or small coverage ( $\leq 500$ ) were removed from further genomic analysis. The genome sequencing and assembly data were deposited at NCBI under BioProject Accession Numbers PRJNA588803 (VDL) and PRJNA591128 (NVSL).

## 2.4 | Phylogenetic analysis

Single nucleotide polymorphisms (SNPs) of 48 *S. zooepidemicus* isolates were identified by running kSNP3 with standard mode. The optimal k-mers size was calculated by Kchooser program, and the whole-genome phylogeny was analysed based on identified core genome SNPs (Gardner, Slezak, & Hall, 2015). Sequence type (ST) based on 7 highly conserved housekeeping genes (*arc*, *nrdE*, *proS*, *spi*, *tdk*, *tpi* and *yqil*) (Webb et al., 2008) was assigned for each *S. zooepidemicus* genome according to the PubMLST *S. zooepidemicus* database (<http://pubmlst.org/szooepidemicus>) (Jolley, Chan, & Maiden, 2004). Tree Of Life (iTOL, <https://itol.embl.de/>) (Letunic & Bork, 2019) was used for display, manipulation and annotation on the base of core SNPs tree.

## 2.5 | Genome gap closure and comparative genomic analysis

Three genomes of *S. zooepidemicus* including two outbreak strains of OH-71905 from Ohio and TN-74097 from Tennessee, and another outbreak-unrelated swine isolate AZ-45470 from Arizona were further sequenced using Oxford Nanopore Technologies (Oxford, United Kingdom) GridIONx5 in the DNA facility at ISU to generate longer reads for genome gap closure. The full-length circular genome sequences were obtained by hybrid assembly combining both Illumina short reads and Nanopore long reads using Unicycler v0.4.8 (Wick, Judd, Gorrie, & Holt, 2017). Three complete genome sequences are available at NCBI under BioProject Accession Number PRJNA588803. Comparative genomic studies were performed with BLAST Ring Image Generator (BRIG) (Alikhan, Petty, Ben Zakour, & Beatson, 2011) to generate a circular genomic map using three closed genome sequences in addition to two complete sequences of control strains ATCC 35246 and CY (Nanjing) (GenBank Accession Number CP002904.1 and CP006770.1). The circular graphical map was plotted to feature GC skew, GC content and predicted genomic islands (GIs) along with genome comparisons.

## 2.6 | Virulence gene and genomic island identification

Putative virulence genes were retrieved from genome sequences according to previous publications (Alber et al., 2005; Bergmann et al., 2019; Kittang et al., 2017; Ma et al., 2019; Rash et al., 2014). The prediction of genomic island (GI) was based on IslandPath-DIMOB, SIGI-HMM, IslandPick and Islander using IslandViewer 4 (Bertelli et al., 2017). Coding sequences (CDS) in every GI were annotated by NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016). The distribution of putative virulence genes and proportion of CDS of GIs in all of the *S. zooepidemicus* isolates were determined in R using pheatmap package. All gene sequences including putative virulence genes and CDSs

in GIs were identified by local Blast+ (2.9.0 version) choosing BLASTn option.

## 3 | RESULTS

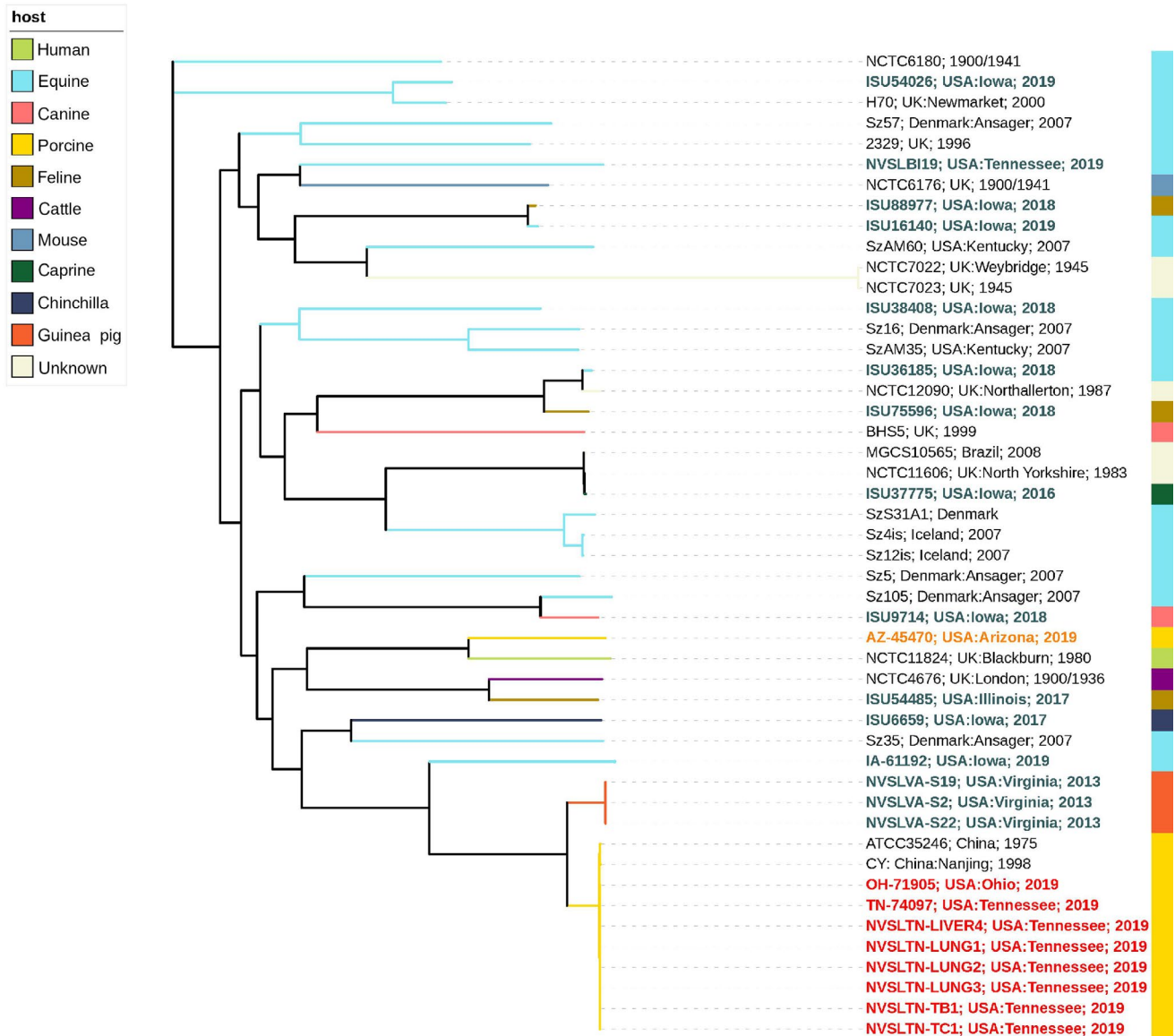
### 3.1 | Phylogenetic characterization of *Streptococcus zooepidemicus*

Whole-genome sequencing was performed on 24 *S. zooepidemicus* isolates. More than 90% of the paired-end reads processed had a Phred score of 35, indicating the high quality of sequencing data. The de novo assembly of raw reads generated contigs with total size ranging from 2.0 Mbp to 2.2 Mbp, and an overall GC content between 40% and 43%, which were similar to those of reference genome ATCC 35246 (2,167,164 bp, 41.65% GC) (Ma et al., 2011, 2013). Thus, the quality of all assembled contigs of 24 isolates was considered sufficient for whole-genome analysis.

Whole-genome phylogenetic analysis based on core genome SNPs was conducted with 48 *S. zooepidemicus* isolates including the 24 isolates sequenced in this study and 24 strains from different countries with publically available genome sequences (Figure 1). A total of 23,659 core genome SNPs were identified from all strains by kSNP3. The phylogenetic tree based on the core genome SNPs indicated a large genetic diversity and the 48 isolates could be clustered into 33 phylogenetic lineages (threshold = 0.01) with only 5 clusters consisting of more than one strain (isolate), while the other 28 clusters contained only one strain (isolate). No obvious geographic or historic distribution differences could be found in our phylogenetic analysis. The eight *S. zooepidemicus* isolates from Ohio and Tennessee outbreaks were indistinguishable based on whole-genome phylogeny, suggesting that they shared a common source. Surprisingly, the eight outbreak isolates were clustered together with the Chinese strain (ATCC 35246), which caused outbreaks of high mortality in 1975 (Ma et al., 2011). In addition, three isolates from human cases with guinea pig exposure (NVSLVA-S19, NVSLVA-S2 and NVSLVA-S22) (Gruszynski et al., 2015) were also closely related to the outbreak isolates. In contrast, another swine isolate (AZ-45470), which did not cause a high mortality event, was distant to the outbreak isolates.

The MLST analysis on all 48 isolates revealed 24 previously described ST types, while 11 strains (22.91%) represented by 9 novel allelic profiles did not match to any STs available in the current PubMLST database of *S. zooepidemicus* as of December 2019 (Figure 2 and Table S3). Even though clusterings based on whole-genome core SNPs and MLST allelic profiles were overall similar, strains with the same MLST types were distributed both within the same and adjacent whole-genome core SNP clusters in the phylogenetic tree, indicating that the SNP-based genotyping was more discriminatory, as expected. The eight strains from Ohio and Tennessee outbreaks and two strains (CY and ATCC 35246) from China were grouped together in ST194, while the outbreak-unrelated Arizona swine isolate (AZ-45470) was clustered distantly in ST340. Three guinea pig isolates from human cases (NVSLVA-S19, NVSLVA-S2 and

Tree scale: 0.01



**FIGURE 1** Whole-genome sequence-based phylogenetic analysis was conducted using SNPs located in all tested *Streptococcus zooepidemicus* genome to generate a core SNP parsimony tree. The branches of the tree are proportional to the distance between the isolates. 11 colour strips demonstrate different hosts of total of 48 isolates

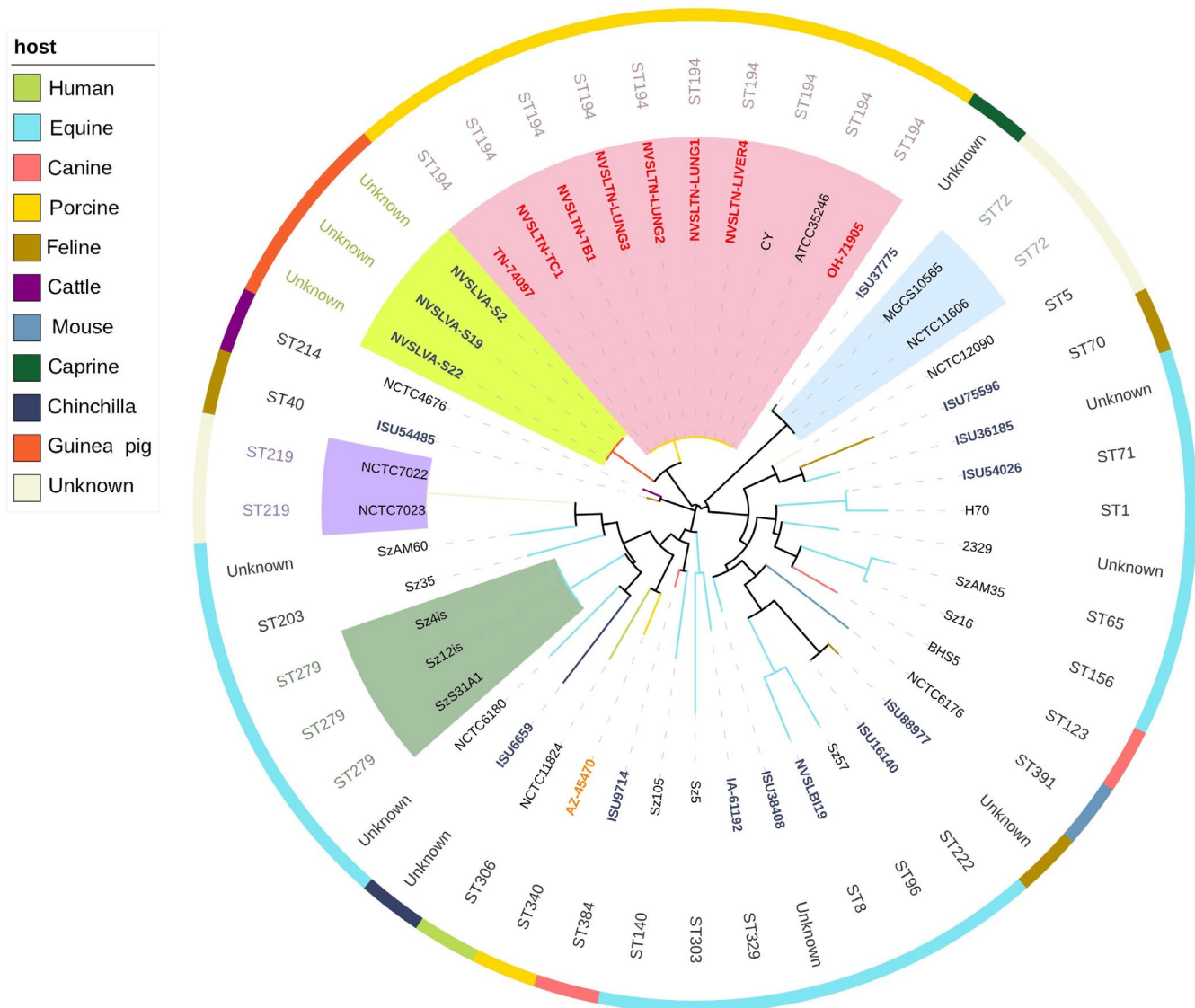
NVSLVA-S22) with the same allelic profile and an unassigned ST type, differing from ST194 at two allele sequences, were closely clustered with the eight outbreak strains from Ohio and Tennessee as well as the Chinese outbreak strain ATCC 35246 within the ST194 lineage.

### 3.2 | Comparative genomic analysis of *Streptococcus zooepidemicus*

To further characterize the *S. zooepidemicus* isolates associated with the high mortality outbreaks in Ohio and Tennessee (suggesting a hypervirulence trait), the genome gaps of three isolates were closed to obtain full-length genome sequences and perform

comparative genomic analysis. Considering the indistinguishability of seven porcine isolates from Tennessee, one representative isolate from Tennessee (TN-74097), the single swine outbreak isolate from Ohio (OH-71905) and the outbreak-unrelated swine isolate from Arizona (AZ-45470) were subjected to Oxford Nanopore Sequencing to generate long reads. Three full-length complete genome sequences (Accession Numbers: CP046040, CP046042 and CP046041) were obtained using hybrid assembly of both Illumina short reads and Nanopore long reads. The general genomic features of all three genomes are summarized in Table 1. Both *S. zooepidemicus* strains OH-71905 and TN-74097 had identical average GC content of 41.65% while contained the circular chromosome of 2,189,155 bp and 2,188,654 bp, respectively. Genome annotation

Tree scale: 0.01



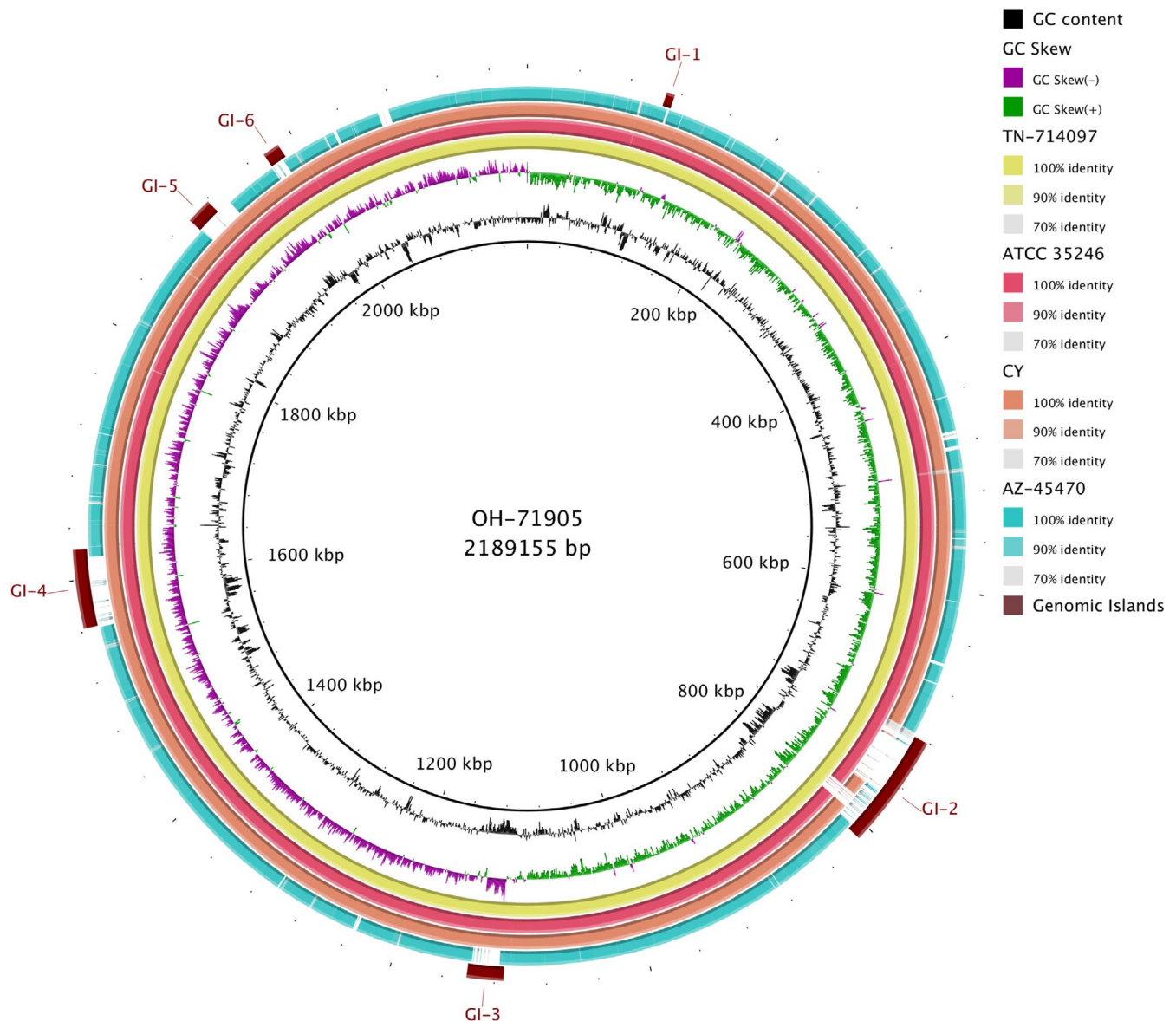
**FIGURE 2** MLST analysis of 48 *Streptococcus zooepidemicus* isolates. The outer colour ring represents different hosts. ST type of every isolate is listed at the inner circle. The neighbour-joining tree was based on a concatenated alignment of 7 housekeeping genes. Colour shades indicate different isolates with identical ST type

**TABLE 1** General genomic features of three swine *Streptococcus zooepidemicus* isolates OH-71905, TN-714097 and AZ-45470

Feature	OH-71905	TN-714097	AZ-45470
Length, bp	2,189,155	2,188,654	2,074,453
GC content	41.65%	41.65%	41.54%
Genes (total)	2,083	2084	1,958
CDSs (total)	1,994	1,995	1,869
Genes (coding)	1,919	1,917	1,822
rRNAs	6, 6, 6 (5S, 16S, 23S)	6, 6, 6 (5S, 16S, 23S)	6, 6, 6 (5S, 16S, 23S)
tRNAs	67	67	67
ncRNAs	4	4	4
Pseudogenes (total)	75	78	47

Abbreviations: CDS, coding sequence; ncRNA, non-coding RNA; rRNA, ribosomal RNA; tRNA, transfer RNA.





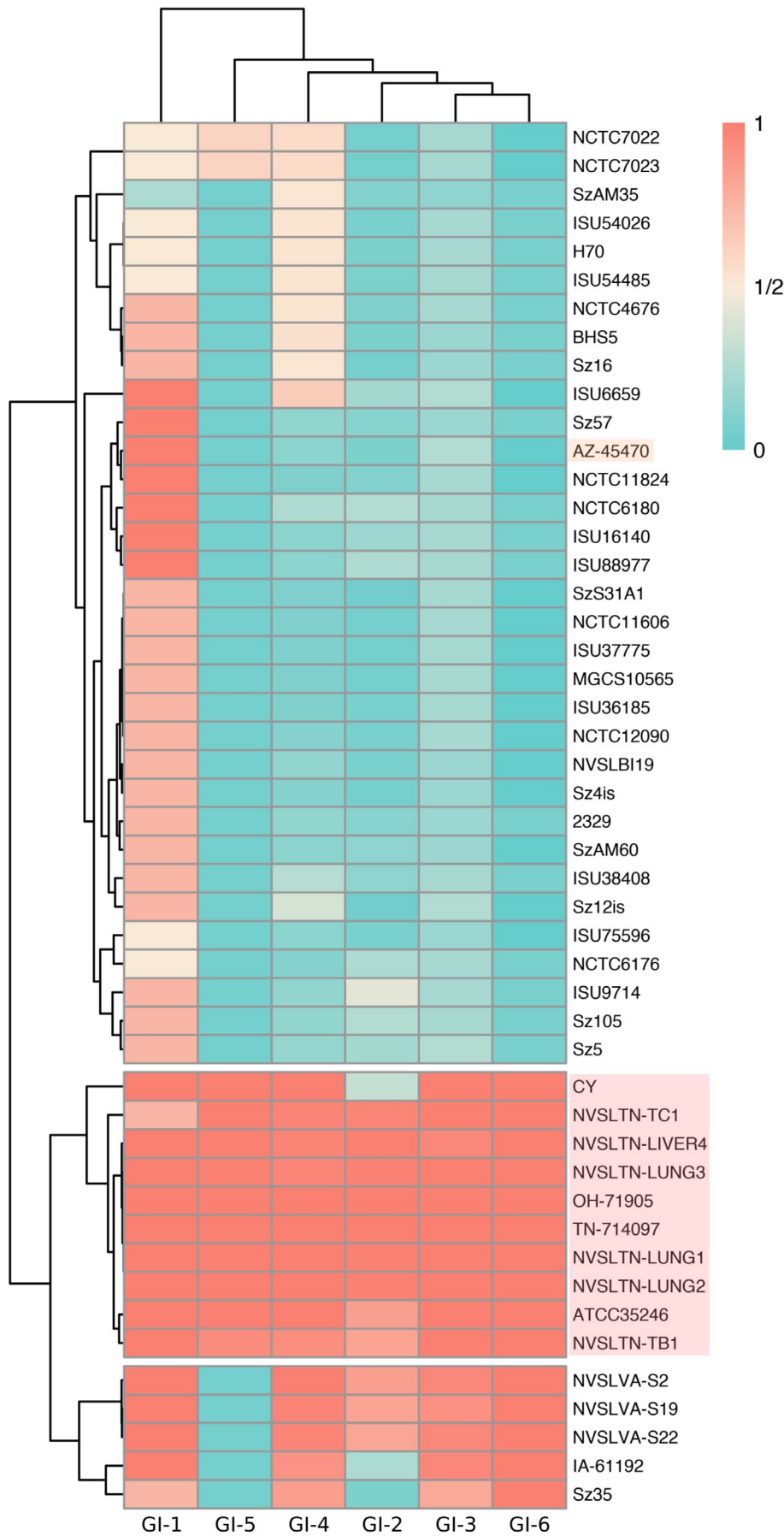
**FIGURE 3** Comparative genome analysis between *Streptococcus zooepidemicus* isolates from pigs. Rings from outside to inside: 1) predicted genomic islands of OH-71905 (brown); 2) AZ-45470 from Arizona (blue); 3) CY from China (orange); 4) ATCC 35246 from China (pink); 5) TN-714097 from Tennessee (yellow); 6) GC skew of OH-71905; 8) GC content of OH-71905; and 9) OH-71905 genome. The blank spaces in the rings represented matches with less than 70% identity to the reference genome

pipeline (PGAP) detected total of 89 RNA genes in both isolates OH-71905 and TN-74097 including 67 tRNA genes, 18 rRNA genes and 4 ncRNA genes. There were only differences detected in two coding and three pseudogenes between these two hypervirulent strains. Compared to two outbreak isolates from Ohio and Tennessee, the genome size of *S. zooepidemicus* strain AZ-45470 from an outbreak-unrelated case was noticeably smaller with 2,074,453 bp of length and an average GC content of 41.54%. Accordingly, 1,869 CDSs

were identified in the genome of AZ-45470, although it had the same numbers of tRNA, rRNA and ncRNA genes as with OH-71905 and TN-74097 genomes.

Comparative genomic analysis was performed with OH-71905, TN-74097 and AZ-45470 genomes as well as the genome of ATCC 35246 and CY strains from China (Figure 3). Genome differences were visualized using BLAST Ring Image Generator (BRIG). As previously mentioned, the genome sequences of OH-71905 and

**FIGURE 4** Two-way clustering of GI prevalence among *Streptococcus zooepidemicus* isolates. A red-white-blue heat map was constructed based on the percentage of the CDSs in each predicted OH-71905 GIs presented in all *S. zooepidemicus* isolates. Clustering was performed to illustrate the similarities between the prevalence of the GIs examined and between the *S. zooepidemicus* isolates with regard to CDS proportion. Red shade indicates porcine isolates associated with high mortality, while orange shade indicates another porcine isolate AZ-45470 without high mortality



**FIGURE 5** Distribution of putative virulence genes among 48 *Streptococcus zooepidemicus* isolates. Right side of the vertical line shows *S. zooepidemicus* isolates clustered by the presence or absence of the putative virulence genes. Coloured boxes indicate the presence of 3 classified groups: 1, M-like protein-coding genes, yellow; 2, superantigen coding genes, green; and 3, other putative virulence genes, red; white boxes indicate the absence of virulence genes. Red shade indicates porcine isolates associated with high mortality, while orange shade indicates another porcine isolate AZ-45470 without high mortality

TN-74097 were 99.98% nt identical. Overall, OH-71905 showed an average nucleotide identity of 99.96% with ATCC 35246 within the 99.36% aligned sequence length, indicating that they are highly similar. Compared with the ATCC 35246 genome, 18 insertions and 6 deletions were found in the genomes of OH-71905, and the largest insertion was 12,494 bp of length from the position 786,295 bp to position 798,785 bp. In contrast, the genome sequence of AZ-45470 displayed an average nucleotide identity of 97.20% with those of OH-71905 and covering 91.12% of aligned sequence length. Additionally, 189 deletions and 186 insertions were identified in the genome sequence of AZ-4570 compared with those of OH-71905. The top six largest deletions ranged from 5,000 bp to 50,000 bp in length.

### 3.3 | Identification and distribution of genomic islands

A genomic island (GI) is part of a bacterial genome that has evidence of horizontal transmission origins. The GIs in OH-71905 and TN-74097 were predicted using IslandViewer 4. Six GIs with significantly different GC contents compared with the core genome were identified in both strains. The sizes of identified GIs varied greatly from 6 kb to 89 kb. Several GIs encode putative virulence genes and possibly contribute to the virulence of *S. zooepidemicus* (Table S4 and S5). GI-2 encodes a putative holing-like toxin and a putative type VI secretion system protein; GI-3 encodes a virulence-associated protein E; GI-4 encodes a putative toxin PezT, a nucleotidyltransferase which belongs to the AbiEii/AbiGii toxin family protein, and a putative VirB4 of type IV secretion system.

The distribution of the identified six GIs in all of the 45 *S. zooepidemicus* strains (Figure 4, Table S6) was assessed by examining the percentage of CDS present in six GIs. All six GIs were present (CDSs > 75%, with more than 75% of the CDS in a GI present in other genomes) in all eight outbreak isolates from Ohio and Tennessee. Five complete GIs could be detected in ATCC 35246 and CY strains of the Chinese outbreaks, and only partial CDS (90/106, 84.91%) of GI-2 was found in ATCC 35246. In contrast, only one complete GI (GI-1) was present in swine isolate AZ-45470, which originated from a swine case that was not associated with high swine mortality outbreaks. More than 82% of CDS in all GIs except for GI-2 were present in the three isolates from human cases with guinea pig exposure (NVSLVA-S19, NVSLVA-S2 and NVSLVA-S22). In addition, the occurrence of five GIs (GI-2, GI-3, GI-4, GI-5 and GI-6) in other *S. zooepidemicus* isolates from human and other animal species was extremely low, two of 35 (5.71%) isolates possessed GI-3, GI-4 and

GI-6 (CDS > 81%), while none of 35 isolates contained GI-2 or GI-5 (CDS < 61%).

### 3.4 | Detection and distribution of putative virulence genes

The presence of 15 putative virulence genes, which have been previously reported in *S. zooepidemicus* (Alber et al., 2005; Bergmann et al., 2019; Kittang et al., 2017; Ma et al., 2019; Rash et al., 2014), was examined in all 48 isolates included in this study (Figure 5, Table S7). The M-like protein gene *szP* was present in all *S. zooepidemicus* isolates in our study albeit with variable sequences. The second M-like protein *SzM* (AEJ26159.1) and the newly identified Fic domain-containing protein *BifA* were recently reported virulence factors of *S. zooepidemicus* (Ma et al., 2019). ATCC35246 type *szM* genes and *bifA* were encoded by all eight isolates from Ohio and Tennessee outbreaks, and another swine strain from China (CY), but absent from the outbreak-unrelated swine isolate AZ-45470. The distribution of these two genes in other strains isolated from human or other animal species was only 18.42% (7/38). In addition, a fimbrial subunit protein-coding gene *fszF* was also present in all eight outbreak isolates from Ohio and Tennessee and the ATCC 35246 and CY strains from Chinese outbreaks, but negative in the outbreak-unrelated swine isolate AZ-45470. Several superantigen genes including *szfF*, *szfL*, *szfM*, *szfN* and *szfP* have been previously identified in *S. zooepidemicus* (Alber et al., 2005; Paillot, Darby, et al., 2010; Paillot, Robinson, et al., 2010; Xu et al., 2019). These superantigen genes were all absent from all of the 11 swine isolates and infrequently present in other isolates from human and other animal species tested in our study (43.24%, 16/37).

## 4 | DISCUSSION

Although *S. zooepidemicus* is considered as an opportunistic pathogen in a large variety of host species including cats, rodents, mink, monkeys and seals, the majority of cases were reported from domestic animals such as horses, dogs, ruminants and pigs (Fulde & Valentin-Weigand, 2013). *S. zooepidemicus* is a commensal organism in horses, but may act as an opportunistic pathogen causing abscesses, neonatal septicaemia and endometritis (Timoney, 2004). Several studies revealed that *S. zooepidemicus* was the major causative agent of purulent infections in horses and foals leading to severe respiratory diseases (Laus et al., 2007; Newton, Laxton, Wood, & Chanter, 2008). *S. zooepidemicus* has also been reported as a causative agent of mastitis in ruminants





inducing severe and deep tissue infections (Sharp et al., 1995). In addition, emerging infections of *S. zooepidemicus* in dogs frequently occur as outbreaks causing severe and life-threatening diseases such as haemorrhagic pneumonia and septicaemia (Pesavento, Hurley, Bannasch, Artiushin, & Timoney, 2008; Priestnall & Erles, 2011). In the swine industry, *S. zooepidemicus* has been a significant pathogen in some Asian countries with outbreaks reported in China (1975) where more than 300,000 pigs died (Ma et al., 2011) and in Indonesia nearly twenty years later (1994) (Soedarmanto et al., 1996). However, high mortality events due to *S. zooepidemicus* in swine have not previously been reported in the United States. Our genomic epidemiological analysis revealed that all eight *S. zooepidemicus* isolates from Ohio and Tennessee outbreaks were indistinguishable, yet clearly different from those isolated from horses, dogs, ruminants and most other host species regardless of European or North American origin. These eight outbreak isolates were also significantly divergent from another outbreak-unrelated swine strain AZ-45470 from Arizona. What is especially concerning is that these eight outbreak isolates were clustered together with the ATCC 35246, the strain that caused the high mortality outbreak in China (Ma et al., 2011), and showed high similarity in their contents of genomic islands and virulence genes.

*S. zooepidemicus* is a known zoonotic pathogen, and nearly 30 human cases of meningitis, septicaemia, pneumonia and glomerulonephritis have been documented (Abbott et al., 2010; Beres et al., 2008; Bordes-Benitez et al., 2006; Kuusi et al., 2006; Mincses, Brown, & Veldkamp, 2011; Sesso & Pinto, 2005). These human cases are usually associated with ingestion of animal products including unpasteurized milk and cheese or contact with companion animals such as horses, dogs and guinea pigs (Fulde & Valentin-Weigand, 2013). Transmission from pigs to human has, thus far, never been reported. Our results showed that the eight outbreak isolates of *S. zooepidemicus* in this study were classified as ST194, an ST type recorded in the *S. zooepidemicus* database from two human blood isolates recovered during 2001. Based on the whole-genome phylogeny, the eight isolates were also closely clustered with three isolates from human patients who had severe clinical illness with guinea pig exposure (Gruszynski et al., 2015), and moreover, their genomic islands and virulence genes were very similar. These results highlight significant public health concerns with these recent US outbreak isolates. Swine producers, veterinarians and other personnel who may directly or indirectly have contact with pigs should be aware of the potential of this organism to cause serious disease and death, and *S. zooepidemicus* infection should be considered if they have purulent wounds or systemic symptoms of infection.

Our results also suggest the contributions of two previously identified virulence genes ATCC 35246 type *szM* and *bifA* to the pathogenicity of *S. zooepidemicus*. Both genes were present in all eight *S. zooepidemicus* isolates and the ATCC 3546 strain, all of which have been involved in high mortalities. Conversely, both genes were absent from another outbreak-unrelated swine isolate AZ-45470 and were also extremely rare in strains from other animal species. It is possible that these two virulence genes are specifically associated with the

hypervirulence of *S. zooepidemicus* to pigs, and further study of these specific genes is warranted. Superantigens are potent toxins, which may disrupt both innate and adaptive immune responses and trigger non-specific T-cell proliferation and overzealous inflammatory production in the host (Fraser, Arcus, Kong, Baker, & Proft, 2000). Several superantigen genes were shown to be significantly associated with non-strangles lymph node abscessation in the horse and probably contribute to virulence (Alber et al., 2005). However, in our study, all 11 swine isolates were negative for any superantigen gene, suggesting that they might not be necessary for the virulence of *S. zooepidemicus* in pigs. In addition, many deletions were identified from the genome of the outbreak-unrelated swine isolate AZ-45470 compared with those of eight *S. zooepidemicus* outbreak isolates and the ATCC 3546 strain. Among them, six large deletions were predicted genomic islands that were absent from this strain and several of these islands encoded putative virulence genes including putative toxin genes and type IV and type VI secretion system proteins. Four out of six PAIs were previously founded in the ATCC 35246 strain and partial deletion of one PAI (from SeseC\_01869 to SeseC\_01898 of the SeseCisland\_3) indeed led to the virulence attenuation (Ma et al., 2013). It will be very interesting to determine whether other genes and other islands would contribute to the virulence of *S. zooepidemicus* in pigs.

In summary, we performed genomic epidemiological and comparative genomic analyses with *S. zooepidemicus* isolates associated with high swine mortality in the United States. Our findings provide significant and timely insights for a better understanding of the epidemiology and virulence of *S. zooepidemicus* isolates associated with highly unexpected and severe outbreaks that occurred very recently in the US swine population. In addition, identification of specific virulence genes and genomic islands may lead to the development of novel molecular diagnostic tools, and provide the basis for future investigation of virulence mechanisms and control measures.

Ms. Chen is a graduate student in the College of Veterinary Medicine at Iowa State University. Her primary research interests include molecular epidemiology and bacterial pathogenesis.

## ACKNOWLEDGEMENTS

This study was partially supported by the Swine Health Information Center (SHIC#19-236). We thank Ying Zheng and Dr. Huigang Shen for excellent technical assistance.

## CONFLICT OF INTEREST

The authors declare no conflicts of interests.

## ETHICS STATEMENT

Ethical statement is not applicable since samples have been gathered from animals.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are at NCBI under BioProject Accession Numbers PRJNA588803 and PRJNA591128 and available in the supplementary material of this article.

## ORCID

Panchan Sitthicharoenchai  <https://orcid.org/0000-0002-4412-7121>  
 Ganwu Li  <https://orcid.org/0000-0002-6370-9573>

## REFERENCES

- Abbott, Y., Acke, E., Khan, S., Muldoon, E. G., Markey, B. K., Pinilla, M., ... Waller, A. (2010). Zoonotic transmission of *Streptococcus equi* subsp. *zooepidemicus* from a dog to a handler. *Journal of Medical Microbiology*, 59(Pt 1), 120–123. <https://doi.org/10.1099/jmm.0.012930-0>
- Alber, J., El-Sayed, A., Estoepangestie, S., Lammler, C., & Zschock, M. (2005). Dissemination of the superantigen encoding genes *seeL*, *seeM*, *seeL* and *seeM* in *Streptococcus equi* subsp. *equi* and *Streptococcus equi* subsp. *zooepidemicus*. *Veterinary Microbiology*, 109(1–2), 135–141. <https://doi.org/10.1016/j.vetmic.2005.05.001>
- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics*, 12, 402. <https://doi.org/10.1186/1471-2164-12-402>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Beres, S. B., Sesso, R., Pinto, S. W., Hoe, N. P., Porcella, S. F., Deleo, F. R., & Musser, J. M. (2008). Genome sequence of a Lancefield group C *Streptococcus zooepidemicus* strain causing epidemic nephritis: New information about an old disease. *PLoS ONE*, 3(8), e3026. <https://doi.org/10.1371/journal.pone.0003026>
- Bergmann, R., Jentsch, M. C., Uhlig, A., Muller, U., van der Linden, M., Rasmussen, M., ... Baums, C. G. (2019). Prominent binding of human and equine fibrinogen to *Streptococcus equi* subsp. *zooepidemicus* is mediated by specific *SzM*-types and is a distinct phenotype of zoonotic isolates. *Infection and Immunity*, 88(1), e00559-19. <https://doi.org/10.1128/IAI.00559-19>
- Bertelli, Claire, Laird, Matthew R., Williams, Kelly P., Lau, Britney Y., Hoad, Gemma, Winsor, Geoffrey L., & Brinkman, Fiona S. L. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research*, 45(W1), W30–W35. <https://doi.org/10.1093/nar/gkx343>
- Blum, S., Elad, D., Zukin, N., Lysnyansky, I., Weisblith, L., Perl, S., ... David, D. (2010). Outbreak of *Streptococcus equi* subsp. *zooepidemicus* infections in cats. *Veterinary Microbiology*, 144(1–2), 236–239. <https://doi.org/10.1016/j.vetmic.2009.12.040>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bordes-Benítez, A., Sánchez-Oñoro, M., Suárez-Bordón, P., García-Rojas, A. J., Saéz-Nieto, J. A., González-García, A., ... Bolaños-Rivero, M. (2006). Outbreak of *Streptococcus equi* subsp. *zooepidemicus* infections on the island of Gran Canaria associated with the consumption of inadequately pasteurized cheese. *European Journal of Clinical Microbiology and Infectious Diseases*, 25(4), 242–246. <https://doi.org/10.1007/s10096-006-0119-x>
- Byun, J. W., Yoon, S. S., Woo, G. H., Jung, B. Y., & Joo, Y. S. (2009). An outbreak of fatal hemorrhagic pneumonia caused by *Streptococcus equi* subsp. *zooepidemicus* in shelter dogs. *Journal of Veterinary Science*, 10(3), 269–271. <https://doi.org/10.4142/jvs.2009.10.3.269>
- Fan, H., Wang, Y., Tang, F., & Lu, C. (2008). Determination of the mimic epitope of the M-like protein adhesin in swine *Streptococcus equi* subsp. *zooepidemicus*. *BMC Microbiology*, 8, 170. <https://doi.org/10.1186/1471-2180-8-170>
- Fraser, J., Arcus, V., Kong, P., Baker, E., & Proft, T. (2000). Superantigens - powerful modifiers of the immune system. *Mol Med Today*, 6(3), 125–132. [https://doi.org/10.1016/s1357-4310\(99\)01657-3](https://doi.org/10.1016/s1357-4310(99)01657-3)
- Fulde, M., & Valentin-Weigand, P. (2013). Epidemiology and pathogenicity of zoonotic streptococci. *Current Topics in Microbiology and Immunology*, 368, 49–81. [https://doi.org/10.1007/82\\_2012\\_277](https://doi.org/10.1007/82_2012_277)
- Gardner, S. N., Slezak, T., & Hall, B. G. (2015). kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics*, 31(17), 2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>
- Gruszynski, K., Young, A., Levine, S. J., Garvin, J. P., Brown, S., Turner, L., ... Beall, B. (2015). *Streptococcus equi* subsp. *zooepidemicus* infections associated with guinea pigs. *Emerging Infectious Diseases*, 21(1), 156–158. <https://doi.org/10.3201/eid2101.140640>
- Jolley, K. A., Chan, M. S., & Maiden, M. C. (2004). mlstdbNet - distributed multi-locus sequence typing (MLST) databases. *BMC Bioinformatics*, 5, 86. <https://doi.org/10.1186/1471-2105-5-86>
- Kelly, C., Bugg, M., Robinson, C., Mitchell, Z., Davis-Poynter, N., Newton, J. R., ... Waller, A. S. (2006). Sequence variation of the *SeM* gene of *Streptococcus equi* allows discrimination of the source of strangles outbreaks. *Journal of Clinical Microbiology*, 44(2), 480–486. <https://doi.org/10.1128/JCM.44.2.480-486.2006>
- Kittang, B. R., Pettersen, V. K., Oppegard, O., Skutlaberg, D. H., Dale, H., Wiker, H. G., & Skrede, S. (2017). Zoonotic necrotizing myositis caused by *Streptococcus equi* subsp. *zooepidemicus* in a farmer. *BMC Infectious Diseases*, 17(1), 147. <https://doi.org/10.1186/s12879-017-2262-7>
- Kuusi, M., Lahti, E., Virolainen, A., Hatakka, M., Vuento, R., Rantala, L., ... Huotari, K. (2006). An outbreak of *Streptococcus equi* subspecies *zooepidemicus* associated with consumption of fresh goat cheese. *BMC Infectious Diseases*, 6, 36. <https://doi.org/10.1186/1471-2334-6-36>
- Laus, F., Prezioso, S., Spaterna, A., Beribe, F., Tesei, B., & Cuteri, V. (2007). Clinical and epidemiological investigation of chronic upper respiratory diseases caused by beta-haemolytic *Streptococci* in horses. *Comparative Immunology, Microbiology and Infectious Diseases*, 30(4), 247–260. <https://doi.org/10.1016/j.cimid.2007.02.003>
- Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, 47(W1), W256–W259. <https://doi.org/10.1093/nar/gkz239>
- Ma, Z., Geng, J., Yi, L. I., Xu, B., Jia, R., Li, Y., ... Hu, S. (2013). Insight into the specific virulence related genes and toxin-antitoxin virulent pathogenicity islands in swine streptococcosis pathogen *Streptococcus equi* ssp. *zooepidemicus* strain ATCC35246. *BMC Genomics*, 14, 377. <https://doi.org/10.1186/1471-2164-14-377>
- Ma, Z., Geng, J., Zhang, H., Yu, H., Yi, L., Lei, M., ... Hu, S. (2011). Complete genome sequence of *Streptococcus equi* subsp. *zooepidemicus* strain ATCC 35246. *Journal of Bacteriology*, 193(19), 5583–5584. <https://doi.org/10.1128/JB.05700-11>
- Ma, Z., Peng, J., Yu, D., Park, J. S., Lin, H., Xu, B., ... Waldor, M. K. (2019). A streptococcal Fic domain-containing protein disrupts blood-brain barrier integrity by activating moesin in endothelial cells. *PLoS Pathogens*, 15(5), e1007737. <https://doi.org/10.1371/journal.ppat.1007737>
- Ma, Z., Zhang, H., Zheng, J., Li, Y., Yi, L., Fan, H., & Lu, C. (2012). Interaction between M-like protein and macrophage thioredoxin facilitates anti-phagocytosis for *Streptococcus equi* ssp. *zooepidemicus*. *PLoS ONE*, 7(2), e32099. <https://doi.org/10.1371/journal.pone.0032099>
- Metzgar, D., & Zampolli, A. (2011). The M protein of group A *Streptococcus* is a key virulence factor and a clinically relevant strain identification marker. *Virulence*, 2(5), 402–412. <https://doi.org/10.4161/viru.2.5.16342>
- Mikheenko, A., Prijbelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG.

- Bioinformatics, 34(13), i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>
- Mincses, L. R., Brown, P. J., & Veldkamp, P. J. (2011). Human meningitis from *Streptococcus equi* subsp. *zooepidemicus* acquired as zoonoses. *Epidemiology and Infection*, 139(3), 406–410. <https://doi.org/10.1017/S0950268810001184>
- Newton, J. R., Laxton, R., Wood, J. L., & Chanter, N. (2008). Molecular epidemiology of *Streptococcus zooepidemicus* infection in naturally occurring equine respiratory disease. *The Veterinary Journal*, 175(3), 338–345. <https://doi.org/10.1016/j.tvjl.2007.02.018>
- Paillot, R., Darby, A. C., Robinson, C., Wright, N. L., Steward, K. F., Anderson, E., ... Waller, A. S. (2010). Identification of three novel superantigen-encoding genes in *Streptococcus equi* subsp. *zooepidemicus*, *szf*, *szn*, and *szeP*. *Infection and Immunity*, 78(11), 4817–4827. <https://doi.org/10.1128/IAI.00751-10>
- Paillot, R., Robinson, C., Steward, K., Wright, N., Jourdan, T., Butcher, N., ... Waller, A. S. (2010). Contribution of each of four Superantigens to *Streptococcus equi*-induced mitogenicity, gamma interferon synthesis, and immunity. *Infection and Immunity*, 78(4), 1728–1739. <https://doi.org/10.1128/IAI.01079-09>
- Pelkonen, S., Lindahl, S. B., Suomala, P., Karhukorpi, J., Vuorinen, S., Koivula, I., ... Tuuminen, T. (2013). Transmission of *Streptococcus equi* subspecies *zooepidemicus* infection from horses to humans. *Emerging Infectious Diseases*, 19(7), 1041–1048. <https://doi.org/10.3201/eid1907.121365>
- Pesavento, P. A., Hurley, K. F., Bannasch, M. J., Artiushin, S., & Timoney, J. F. (2008). A clonal outbreak of acute fatal hemorrhagic pneumonia in intensively housed (shelter) dogs caused by *Streptococcus equi* subsp. *zooepidemicus*. *Veterinary Pathology*, 45(1), 51–53. <https://doi.org/10.1354/vp.45-1-51>
- Priestnall, S., & Erles, K. (2011). *Streptococcus zooepidemicus*: An emerging canine pathogen. *The Veterinary Journal*, 188(2), 142–148. <https://doi.org/10.1016/j.tvjl.2010.04.028>
- Rash, N. L., Robinson, C., DeSouza, N., Nair, S., Hodgson, H., Steward, K., ... Paillot, R. (2014). Prevalence and disease associations of superantigens *szf*, *szn* and *szeP* in the *S. zooepidemicus* population and possible functional redundancy of *szf*. *Research in Veterinary Science*, 97(3), 481–487. <https://doi.org/10.1016/j.rvsc.2014.09.001>
- Sesso, R., & Pinto, S. W. (2005). Five-year follow-up of patients with epidemic glomerulonephritis due to *Streptococcus zooepidemicus*. *Nephrology, Dialysis, Transplantation*, 20(9), 1808–1812. <https://doi.org/10.1093/ndt/gfh904>
- Sharp, M. W., Prince, M. J., & Gibbens, J. (1995). *S. zooepidemicus* infection and bovine mastitis. *The Veterinary Record*, 137(5), 128. <https://doi.org/10.1136/vr.137.5.128-b>
- Soedarmanto, I., Pasaribu, F. H., Wibawan, I. W., & Lammler, C. (1996). Identification and molecular characterization of serological group C streptococci isolated from diseased pigs and monkeys in Indonesia. *Journal of Clinical Microbiology*, 34(9), 2201–2204. <https://doi.org/10.1128/JCM.34.9.2201-2204.1996>
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., ... Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research*, 44(14), 6614–6624. <https://doi.org/10.1093/nar/gkw569>
- Timoney, J. F. (2004). The pathogenic equine streptococci. *Veterinary Research*, 35(4), 397–409. <https://doi.org/10.1051/vetres:2004025>
- Torres, R., Santos, T. Z., Bernardes, A. F. L., Soares, P. A., Soares, A. C. C., & Dias, R. S. (2018). Outbreak of Glomerulonephritis Caused by *Streptococcus zooepidemicus* SzPHV5 Type in Monte Santo de Minas, Minas Gerais, Brazil. *Journal of Clinical Microbiology*, 56(10), e00845-18. <https://doi.org/10.1128/JCM.00845-18>
- Velineni, S., & Timoney, J. F. (2013). Characterization and protective immunogenicity of the SzM protein of *Streptococcus zooepidemicus* NC78 from a clonal outbreak of equine respiratory disease. *Clinical and Vaccine Immunology*, 20(8), 1181–1188. <https://doi.org/10.1128/CVI.00069-13>
- Walker, J. A., & Timoney, J. F. (1998). Molecular basis of variation in protective SzP proteins of *Streptococcus zooepidemicus*. *American Journal of Veterinary Research*, 59(9), 1129–1133.
- Webb, K., Jolley, K. A., Mitchell, Z., Robinson, C., Newton, J. R., Maiden, M. C. J., & Waller, A. (2008). Development of an unambiguous and discriminatory multilocus sequence typing scheme for the *Streptococcus zooepidemicus* group. *Microbiology*, 154(Pt 10), 3016–3024. <https://doi.org/10.1099/mic.0.2008/018911-0>
- Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Computational Biology*, 13(6), e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Xu, B., Zhang, P., Zhou, H., Sun, Y., Tang, J., & Fan, H. (2019). Identification of novel genes associated with anti-phagocytic functions in *Streptococcus equi* subsp. *zooepidemicus*. *Veterinary Microbiology*, 233, 28–38. <https://doi.org/10.1016/j.vetmic.2019.04.023>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Chen X, Resende-De-Macedo N, Sitthicharoenchai P, et al. Genetic characterization of *Streptococcus equi* subspecies *zooepidemicus* associated with high swine mortality in the United States. *Transbound Emerg Dis*. 2020;00:1–12. <https://doi.org/10.1111/tbed.13645>