Genomic and pathogenic investigations of *Streptococcus suis* serotype 7 population derived from a human patient and pigs

Pujun Lianga**, Mingliu Wangb**, Marcelo Gottschalkc*, Ana I. Velad*, April A. Estradae, Jianping Wangf, Pengcheng Duf, Ming Luog, Han Zhenga and Zongfu Wu h

**State Key Laboratory of Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, People’s Republic of China; aGuangxi Zhuang Autonomous Region Center for Disease Prevention and Control, Nanning, People’s Republic of China; bSwine and Poultry Infectious Diseases Research Center, Faculty of Veterinary Medicine, University of Montreal, Canada; cDepartamento de Sanidad Animal, Facultad de Veterinaria and Centro de Vigilancia Sanitaria Veterinaria (VISAVET), Universidad Complutense de Madrid, Madrid, Spain; dThe College of Veterinary Medicine, University of Minnesota, St. Paul, MN, USA; eBeijing Key Laboratory of Emerging Infectious Diseases, Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, People’s Republic of China; fYulin Center for Disease Prevention and Control, Yulin, People’s Republic of China; gOIE Reference Lab for Swine Streptococcus, MOE Joint International Research Laboratory of Animal Health and Food Safety, College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, People’s Republic of China

ABSTRACT

*Streptococcus suis* is one of the important emerging zoonotic pathogens. Serotype 2 is most prevalent in patients worldwide. In the present study, we first isolated one *S. suis* serotype 7 strain GX69 from the blood culture of a patient with septicemia complicated with pneumonia in China. In order to deepen the understanding of *S. suis* serotype 7 population characteristics, we investigated the phylogenetic structure, genomic features, and virulence of *S. suis* serotype 7 population, including 35 strains and 79 genomes. Significant diversities were revealed in *S. suis* serotype 7 population, which were clustered into 22 sequence types (STs), five minimum core genome (MCG) groups, and six lineages. Lineages 1, 3a, and 6 were mainly constituted by genomes from Asia. Genomes of Lineages 2, 3b, and 5a were mainly from Northern America. Most of genomes from Europe (41/48) were clustered into Lineage 5b. In addition to strain GX69, 13 of 21 *S. suis* serotype 7 representative strains were classified as virulent strains using the C57BL/6 mouse model. Virulence-associated genes preferentially present in highly pathogenic *S. suis* serotype 2 strains were not suitable as virulence indicators for *S. suis* serotype 7 strains. Integrative mobilizable elements were widespread and may play a critical role in disseminating antibiotic resistance genes of *S. suis* serotype 7 strains. Our study confirmed *S. suis* serotype 7 is a non-negligible pathotype and deepened the understanding of the population structure of *S. suis* serotype 7, which provided valuable information for the improved surveillance of this serotype.

ARTICLE HISTORY Received 6 September 2021; Revised 28 September 2021; Accepted 29 September 2021

KEYWORDS *Streptococcus suis* serotype 7; zoonotic pathogens; phylogeny; integrative mobilizable elements; virulence

Introduction

*Streptococcus suis* is an important emerging zoonotic pathogen responsible, among other infections, for septicemia, meningitis, endocarditis, and arthritis in humans [1]. To date, serotyping is an important routine diagnostic procedure and is widely used for subtyping *S. suis* strains. Among 29 confirmed serotypes (1–19, 21, 23–25, 27–31, and 1/2) and 28 novel *cps* types [2–5], serotype 2 is most frequently isolated from clinical cases in swine and humans worldwide [1,6]. Two outbreaks featured by high rates of streptococcal toxic-shock-like syndrome (STSLS) were caused by *S. suis* serotype 2 sequence type (ST)7 strains in China [7,8]. Recently, the prevalence of serotype 14 has also increased among sporadic human cases in China [9]. Serotype 9 has become the most prevalent serotype in diseased pigs in some European counties [1,10], and one human case of serotype 9 infection was reported [11]. Serotypes 4, 5, 16, 21, 24, and 31 have also been reported in human infections [1,12]. Serotype 7 is an important serotype frequently isolated from diseased pigs in European countries, North America, and Thailand [13–18], and it was also related to severe herd problems of meningitis and arthritis in nursery and grower pigs [14].

In the present study, an *S. suis* serotype 7 strain was isolated for the first time from the blood culture of a patient with septicemia complicated with pneumonia in China, suggesting that some serotype 7 strains may possess zoonotic potential. Except for limited epidemiologic studies [19] and *in vitro* survival assay in swine blood [20], little information is available for
the phylogeny, evolution, and pathogenicity of S. suis serotype 7 population. The present study included 35 strains and 79 genomes of strains from 1999 to 2019 originating from nine countries to represent the S. suis serotype 7 population. The phylogenetic relationship, dissemination mechanisms of antibiotic resistance (AR) genes, variation of cps arrangements, and virulence were investigated to elucidate the population structure, genomic features, evolution, and pathogenicity of S. suis serotype 7.

Materials and methods

Case description

On 22 July 2016, a 71-year-old female patient with a history of hypertension was admitted to the First People’s Hospital of Yulin in Yulin city because of repeating fever and chill (highest body temperature of 39.5°C), cough, and abdominal pain for five days. A computerized tomography scan image indicated inflammation of both lungs. The serum level of high-sensitivity C-reactive protein and total counts of white blood cells were 112.77 mg/L and 5.68 × 10^9/L, respectively. The neutrophil percentage was 72.7%. The patient’s blood pressure was 103/63 mm Hg. Meropenem, piperacillin/tazobactam, and levofloxacin were given as antibiotic therapy. The patient recovered and was discharged ten days later. A strain (named GX69) was isolated from the blood culture of the patient. The strain was confirmed as S. suis by amplifying S. suis-specific recN gene [21].

GX69 was first identified as serotype 7 by the agglutination test using the serum purchased from Statens Serum Institute, Copenhagen, Denmark, and further confirmed with a molecular serotyping method [22].

Bacterial strains, genomes, and sequencing

For comparison purposes, 35 strains and 79 genomes were used in this study (Table 1). Twenty-seven of them were from China (24 of them were sequenced in the present study), 23 from the United States of American (USA), 22 from the United Kingdom (UK), 16 from Canada (sequenced in the present study), 13 from Spain (sequenced in the present study), 9 from the Netherlands (3 of them were sequenced in the present study), 2 from France (sequenced in the present study), 1 from Germany (sequenced in the present study) and Denmark each. Genomes of unspecified origin were from Genbank database. All genomes were re-confirmed to belong to S. suis by analysing their full length of 16s rRNA sequences [23] and recN gene specific to S. suis [24]. In addition, these genomes harboured S. suis serotype 7 specific wzy gene. They were isolated from 1999 to 2019.

In the present study, the complete genome of strain GX69 was sequenced using PacBio Sequel platform and Illumina NovaSeq PE150, whereas the draft genomes were sequenced using Illumina NovaSeq PE150. Sequencing libraries were generated using the methods described previously [25]. The valid reads filtered low-quality reads were assembled into contigs and scaffolds with SOAPdenovo (release 1.04). Genes were predicted by using Glimmer 3.02, and gene orthologs were determined by using GO (Gene ontology) V20171011, KEGG (Kyoto Encyclopedia of Gene and Genomes) V20181107, and COG (Clusters of Othologous Database) V20171127.

Bioinformatics analysis

MLST and MCG typing

The multilocus sequence type (MLST) and the minimum core genome (MCG) group of the genomes were determined by using PubMLST (https://pubmlst.org/bigsdb?db=pubmlst_ssuis_seqdef&page=sequenceQuery), and a method previously described [26], respectively.

Phylogenetic analysis

Single-nucleotide polymorphisms (SNPs) were detected using Bowtie 2, and MUMmer v3.23 for sequencing reads and complete genomes, respectively, and the genome sequence of SC84 (accession No. FM252031) [27] was used as a reference. The mutational SNP sites were selected based on the method described in a previous study [26], and then the phylogenetic tree was constructed using the maximum likelihood method by FastTree v2.1.10. Streptococcus pneumoniae ATCC 700669 (accession No. NC_011990) was used as an outgroup to root the tree. The tree was presented using FigTree v1.4.0.

Detection of S. suis virulence-associated genes, AR genes and AR genes associated with mobile genetic elements (MGEs)

Distributions of virulence-associated genes and regions of difference (RDs) preferentially present in highly pathogenic S. suis serotype 2 strains were investigated among S. suis serotype 7 genomes, consisting of genes mrp, sly, epF, sao, nadR, NisR, NisK, SalR, SalK, revS, ofS, RD6, RD12, RD14, RD21, RD29, RD40, RD53, and RD60 [28,29]. Genes having a global match region at <80% of the amino-acid sequence with an identity of <80% were determined to be absent.

AR genes were analysed by searching Comprehensive Antibiotic Resistance database (CARD) and Antibiotic Resistance genes database (ARDB). A resistance gene was only regarded as a homologue in tested strains if it showed at least 80% identity in amino-acid sequence across 80% of the length of the protein.
| Lineage | Name of Strain | MCG | Serotype | Sequence Type | cps subtype | Accession number | Host | Isolation source | Location | Year | AR genes |
|---------|----------------|-----|----------|---------------|-------------|------------------|------|------------------|----------|------|----------|
| Lineage 1 | 93.01B* | 1 | 7 | 1609 | 7-II | SRR123202835 | Diseased pig | Heart | Spain | 2001 | tet(O) |
| | YS12 | 1 | 7 | 17 | 7-b | SRR123202808 | Healthy pig | Nasopharynx swab | CN | 2012 | tet(O) |
| | WUSS415* | 1 | 7 | 1611 | 7-1 | SRR1232028954 | Healthy pig | Tonsil | CN | 2017.12 | ant(6)-la, erm(B), tet(M) |
| | WUSS417* | 1 | 7 | 1611 | 7-1 | SRR1232028955 | Healthy pig | Tonsil | CN | 2017.12 | ant(6)-la, erm(B), tet(M) |
| Lineage 2 | 21459559 | 2 | 7 | 1613 | 7-Ib | SRR1232028966 | Diseased pig | Brain | Canada | 2018 | tet(O), erm(B) |
| | 21505555 | 2 | 7 | 89 | 7-Ia | SRR1232028966 | Diseased pig | Brain | Canada | 2019 | tet(O), erm(B) |
| | 21208181 | 2 | 7 | 89 | 7-Ia | SRR1232028966 | Diseased pig | Liver | Canada | 2018 | tet(O), erm(B) |
| | 21082848 | 2 | 7 | 89 | 7-Ib | SRR1232028966 | Diseased pig | Joint | Canada | 2018 | tet(O), erm(B) |
| | 128.01B | 2 | 7 | 24 | 7-II | SRR1232028966 | Diseased pig | Brain | Spain | 2001 | tet(O), aph(3')-Ia, ant(6)-Ia, sat-4, erm(B) |
| | WUSS417* | 2 | 7 | 24 | 7-II | SRR1232028966 | Diseased pig | Brain | Spain | 1999 | tet(O), aph(3')-Ia, ant(6)-Ia, sat-4, erm(B) |
| | 21208181 | 3 | 7 | 24 | 7-II | SRR1232028966 | Diseased pig | Brain | Spain | 1999 | tet(O), aph(3')-Ia, ant(6)-Ia, sat-4, erm(B) |
| | 255B | 3 | 7 | 24 | 7-II | SRR1232028966 | Diseased pig | Brain | Spain | 1999 | tet(O), aph(3')-Ia, ant(6)-Ia, sat-4, erm(B) |
| Lineage 3a | Ssuis120 | 3 | 7 | 373 | 7-Ib | SRR1232028966 | Diseased pig | Meninges | USA | 2016 | tet(O) |
| | Ssuis95 | 3 | 7 | 373 | 7-Ia | SRR1232028966 | Diseased pig | Joint | USA | 2016 | tet(O) |
| | 2018WUSS020* | 3 | 7 | 373 | 7-Ib | SRR1232028966 | Diseased pig | Lung | USA | 2015 | tet(O), erm(B) |
| | 2019WUSS025* | 3 | 7 | 373 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | tet(O), erm(B) |
| | WUSS401* | 3 | 7 | 373 | 7-Ib | SRR1232028966 | Diseased pig | Brain | USA | 2015 | tet(O), erm(B) |
| | WUSS417* | 3 | 7 | 373 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | tet(O), erm(B) |
| Lineage 3b | Ssuis359 | 3 | 7 | 373 | 7-Ia | SRR1232028966 | Diseased pig | Lung | USA | 2017 | tet(O) |
| | Ssuis93 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | Canada | 2018 | – |
| | 21566696* | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2016 | tet(O) |
| | Ssuis51 | 3 | 7 | 980 | 7-lll | SRR1232028966 | Diseased pig | Lung | USA | 2015 | tet(O), erm(B) |
| | Ssuis118 | 3 | 7 | 979 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2016 | tet(O), erm(B) |
| | Ssuis98 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | erm(B), tet(O) |
| | Ssuis39 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | tet(O), erm(B) |
| | Ssuis40 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Meninges | USA | 2015 | tet(O), erm(B) |
| | 2130772* | 3 | 7 | 839 | 7-Ia | SRR1232028966 | Diseased pig | Brain | Canada | 2018 | tet(O), erm(B) |
| | Ssuis45 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Lung | USA | 2016 | – |
| | Ssuis109 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | erm(B) |
| | Ssuis303 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2016 | – |
| | Ssuis41 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | erm(B) |
| | Ssuis46 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Lung | USA | 2014 | tet(O) |
| | Ssuis324 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2017 | – |
| | Ssuis309 | 3 | 7 | 94 | 7-Ia | SRR1232028966 | Diseased pig | Brain | USA | 2015 | erm(B) |
| Lineage 4 | WUSS004* | 4 | 7 | 225 | 7-Ib | SRR1232028966 | Diseased pig | / | CN | 2016 | tet(O), erm(B) |
| | WUSS029* | 4 | 7 | 225 | 7-Ib | SRR1232028966 | Diseased pig | / | CN unknown | tet(O), erm(B) |
| | Ssuis8 | 4 | 7 | 225 | 7-Ib | SRR1232028966 | Diseased pig | Brain | USA | 2014 | aph(3')-Ila, ant(6)-la, sat-4, erm(B), tet(O) |
| | Ssuis11 | 4 | 7 | 225 | 7-Ib | SRR1232028966 | Diseased pig | Brain | USA | 2014 | ant(6)-Ia, ant(9)-la, aph(3')-Ila, sat-4, erm(B), tet(O) |
| | 2234124* | 4 | 7 | 1614 | 7-II | SRR1232028966 | Diseased pig | Brain | Canada | 2019 | erm(B) |
| | WUSS013* | 4 | 7 | 225 | 7-Ib | SRR1232028966 | Diseased pig | / | CN unknown | tet(O), erm(B) |
| | Ssuis100 | 4 | 7 | 225 | 7-Ib | SRR1232028966 | Diseased pig | Meninges | USA | 2015 | tet(O) |
| Accession | Isolate Code | Lineage | Country | Region | Tissue | Antimicrobial Resistance | Year | Notes |
|-----------|--------------|---------|---------|--------|--------|--------------------------|------|-------|
| SAMN18117671 | 7-ib | SAMN18117671 | Diseased pig | Heart | Canada | 2019 | tet(O), erm(B) |
| SAMN20087851 | 7-ia | SAMN20087851 | Healthy pig | Nasopharynx swab | CN | 2012 | tet(O), erm(B) |
| SAMA3136674 | Pig | SAME1316674 | Diseased pig | Lung | UK | 2010 | – |
| SAMN18117672 | 7-ia | SAMN18117672 | Diseased pig | Heart | Canada | 2019 | tet(O), erm(B) |
| SAMN18117673 | 7-ib | SAMN18117673 | Diseased pig | Brain | Canada | 2018 | tet(O), erm(B) |
| SAMA3233911 | Pig | SAME1316674 | Diseased pig | Lung | UK | 2014 | – |
| SAMA3136674 | Pig | SAME1316674 | Diseased pig | Lung | UK | 2014 | – |
| SAMA13166697 | Pig | SAME1316697 | Diseased pig | Lung | UK | 2014 | – |
| SAMA3595225 | Diseased pig | CSF | Netherlands | 2002 | erm(B), tet(O) |
| SAMA3595239 | Diseased pig | CSF | Netherlands | 2004 | erm(B), tet(O) |
| SAMA3595236 | Diseased pig | CSF | Netherlands | 2004 | dfrF, ant(6)-Ia, tet(O), cat-TC |
| SAMA13166897 | Pig | SAME13166897 | Diseased pig | Lung | UK | 2010 | – |
| SAMA1316697 | Pig | SAME1316697 | Diseased pig | Lung | UK | 2014 | – |
| SAMA31366768 | 7-ia | SAMN18117680 | Diseased pig | Lung | Canada | 2019 | tet(O), erm(B) |
| SAR9123095 | Pathogenic | Liver | USA | 2014 | erm(B) |
| SAMA3233998 | Diseased pig | Lung | UK | 2010 | ln(u(B), aph(3')-Ila, IscC, ant(9)-Ia) |
| SAMA3136681 | Diseased pig | Lung | USA | 2016 | tet(W), erm(B) |
| SAMA3136681 | Diseased pig | Lung | USA | 2016 | tet(W), erm(B) |
| SAMA3136681 | Diseased pig | Lung | USA | 2016 | tet(W), erm(B) |
| SAMA3136681 | Diseased pig | Lung | USA | 2016 | tet(W), erm(B) |
| SAMA3234014 | Diseased pig | Brain | UK | 2010 | erm(B), tet(O) |
| SAMA18117685 | Diseased pig | Spleen | Netherlands | 2018 | tet(O), erm(B) |
| SAMA17982947 | Healthy pig | Tonsil | CN | 2017 | tet(O), erm(B) |
| SAMA3233988 | Diseased pig | Lung | UK | 2010 | ln(u(B), aph(3')-Ila, IscC, ant(9)-Ia) |
| SAMA13166897 | Pig | SAME13166897 | Diseased pig | Lung | UK | 2010 | – |
| SAMA3595225 | Diseased pig | CSF | Netherlands | 2002 | erm(B), tet(O) |
| SAMA3595239 | Diseased pig | CSF | Netherlands | 2004 | erm(B), tet(O) |
| SAMA3595236 | Diseased pig | CSF | Netherlands | 2004 | dfrF, ant(6)-Ia, tet(O), cat-TC |
| SAMA18117682 | Diseased pig | Brain | Spain | 2019 | tet(O) |
| SAMA17982950 | Healthy pig | Tonsil | CN | 2017 | tet(O), erm(B) |
| SAMA17982947 | Healthy pig | Tonsil | CN | 2017 | tet(O), erm(B) |
| SAMA1316697 | Pig | SAME1316697 | Diseased pig | Lung | UK | 2014 | – |
| SAMA18117680 | Diseased pig | Lung | Canada | 2019 | tet(O), erm(B) |
| SAMA3595225 | Diseased pig | CSF | Netherlands | 2002 | erm(B), tet(O) |
| SAMA3595239 | Diseased pig | CSF | Netherlands | 2004 | erm(B), tet(O) |
| SAMA3595236 | Diseased pig | CSF | Netherlands | 2004 | dfrF, ant(6)-Ia, tet(O), cat-TC |
| SAMA13166897 | Pig | SAME13166897 | Diseased pig | Lung | UK | 2010 | – |
| SAMA3234014 | Diseased pig | Brain | UK | 2010 | erm(B), tet(O) |
| SAMA17982938 | Diseased pig | Heart | Spain | 1999 | – |
| SAMA3233974 | Diseased pig | CSF | Netherlands | 2006 | erm(B), tet(O) |
| SAMA3233929 | Diseased pig | Lymphatic gland | Spain | 1999 | – |
| SAMA3233938 | Diseased pig | Lung | UK | 2011 | dfr(M) |
| SAMA13166699 | Pig | SAME13166699 | Diseased pig | Lung | UK | 2014 | – |
| SAMA18117687 | Diseased pig | Brain | Spain | 2018 | tet(O), erm(B) |
| SAMA17982938 | Diseased pig | Heart | Spain | 1999 | – |
| SAMA3233974 | Diseased pig | CSF | Netherlands | 2006 | erm(B), tet(O) |
| SAMA18117689 | Diseased pig | Joint | Netherlands | 2019 | tet(O), erm(B) |
| SAMA3233926 | Healthy pig | Tonsil | UK | 2011 | dfr(M) |
| SAMA1316684 | Pig | SAME1316684 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA3233974 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA3233938 | Diseased pig | Lymphatic gland | Spain | 1999 | – |
| SAMA13166699 | Pig | SAME13166699 | Diseased pig | Lung | UK | 2014 | – |
| SAMA18117687 | Diseased pig | Brain | Spain | 2018 | tet(O), erm(B) |
| SAMA17982938 | Diseased pig | Heart | Spain | 1999 | – |
| SAMA3233974 | Diseased pig | CSF | Netherlands | 2006 | erm(B), tet(O) |
| SAMA18117689 | Diseased pig | Joint | Netherlands | 2019 | tet(O), erm(B) |
| SAMA3233926 | Healthy pig | Tonsil | UK | 2011 | dfr(M) |
| SAMA1316684 | Pig | SAME1316684 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA3233974 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA18117689 | Diseased pig | Heart | Canada | 2019 | tet(O), erm(B) |
| SAMA17982938 | Diseased pig | Heart | Spain | 1999 | – |
| SAMA3233926 | Diseased pig | CSF | Netherlands | 2006 | erm(B), tet(O) |
| SAMA1316684 | Pig | SAME1316684 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA3233974 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA18117689 | Diseased pig | Joint | Netherlands | 2019 | tet(O), erm(B) |
| SAMA3233926 | Healthy pig | Tonsil | UK | 2011 | dfr(M) |
| SAMA1316684 | Pig | SAME1316684 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA3233974 | Diseased pig | Brain | UK | 2014 | dfr(M) |
| SAMA18117689 | Diseased pig | Joint | Netherlands | 2019 | tet(O), erm(B) |
| SAMA3233926 | Healthy pig | Tonsil | UK | 2011 | dfr(M) |
The prophages and ICEs were predicted by PHAST (http://phast.wishartlab.com/) and ICEberg (https://db-mml.sjtu.edu.cn/ICEberg/), respectively. For the identification of integrative and conjugative elements (ICEs), signature proteins including integrase, relaxase, and VirB4 were typed using the database from a previous study [31]. Search strategies and the definitions of integrative mobilizable elements (IMEs) and cis-IMEs (CIMEs) were carried out according to the methods previously described [31,32].

**Analysis of cps gene cluster**

Each *cps* gene cluster was extracted from the genomes and compared with that of the serotype 7 reference strain 8074 (GenBank accession No. BR001004.1). The homology groups (HGs) of *cps* genes were assigned according to the nomenclature described in a previous study [33]. The sequence comparison of *cps* gene cluster was performed using blastN programme in BLAST with an e-value cutoff of e-10 and was visualized using an in-house Perl script (https://github.com/dupengcheng/BlastViewer).

**Infection experiments**

The virulence of strain GX69 from the patient and 21 additional representative strains based on the distribution in the phylogenetic tree were tested. For comparison, the highly pathogenic and well-characterized *S. suis* serotype 2 reference strain P1/7 (ST1) [34] was included [35,36]. C57BL/6 mice (6 weeks old, female) were injected intraperitoneally with 5 × 10⁷ CFU of *S. suis* strain in 1 mL PBS or 1 mL PBS only as a control group. The infection dose of each strain was confirmed by plating the serial dilutions of the suspension onto the Todd–Hewitt broth (THB, Oxoid Ltd, London, UK) agar before and after the infection. Each infected group contained ten mice, and the mock-infected group contained five mice. The mortality was recorded per six hours within 24 h post-infection and per 12 h from 24 h to 96 h post-infection. The experiment was performed independently at least twice for each strain. The mortality of each infected group was calculated via the Kaplan–Meier method.

S. suis serotype 7 strains initiating lethal infection with a mortality ≥80% at 96 h post-infection were classified as virulent strains.

**Investigation of antimicrobial susceptibility profiles**

To determine whether the AR genes in genomes conferred the predicted resistance to the corresponding bacteria, we used the MIC-test strip (Liofilchem, Roseto degli Abruzzi, Italy) to assess the antimicrobial susceptibility of strains carrying AR genes. The
following antibiotics were tested: clindamycin (0.016–
56 μg/mL), erythromycin (0.016–256 μg/mL), azithro-
mycin (0.016–256 μg/mL), tetracycline (0.016–256 μg/
ml), gentamicin (0.016–256 μg/mL), kanamycin (0.016–256 μg/mL), and streptomycin (0.064–
1024 μg/mL). For tetracycline, azithromycin, erythro-
mycin, and clindamycin, breakpoints were used as re-
commended by the Clinical and Laboratory Standard
Institute (CLSI) guidelines 2019 (M100-S29) for Strept-
occus spp. Viridans group. No breakpoint values of
streptomycin, kanamycin, and gentamicin were avail-
able for Streptococcus. Their breakpoints were taken
from a previous study [37].

**Statistics**

The survival curves of different infected groups were
compared using Gehan–Breslow–Wilcoxon test. For
the test, a p-value < .05 was considered to be signifi-
cant.

**Nucleotide sequence accession numbers**

The sequences of the genomes sequenced in the study
were deposited in the GenBank under accession num-
bers listed in Table 1.

**Ethical approval**

This study and the application of the animal experi-
ments with code 2020-024 were reviewed and
approved by the ethics committee of the National
Institute for Communicable Disease Control and
Prevention, Chinese Center for Disease Control and
Prevention.

**Results**

**MLST and MCG typing**

Among 114 genomes, 22 different STs were identi-
ﬁed, revealing high heterogeneity of S. suis serotype 7
population. ST29 (n = 47) was most prevalent, fol-
lowed by ST373 (n = 16), ST94 (n = 12), ST225 (n =
6), ST24 (n = 4), ST32 (n = 4), ST971 (n = 4), ST89
(n = 3), ST907 (n = 2), ST1610 (n = 2), ST1611 (n =
2), and ST1612 (n = 2). The remaining ST17, ST34,
ST839, ST854, ST973, ST979, ST980, ST1609,
ST1613, and ST1614 only contained one strain each.
The strain GX69 from the patient was ST373 which
was prevalent in China, whereas ST29 and ST94
were predominant in Europe and North America,
respectively (Table 1).

The 114 genomes were clustered into five MCG
groups, including MCG groups 1, 2, 3, 4, and 7–2.
MCG group 4 was predominant and included ST29
strains. It is noteworthy that genomes of MCG group 4
were composed of 10 STs and 69 genomes widely
distributed in all nine countries. Five STs and 31 gen-
omes were classiﬁed into MCG group 3, including
the strain GX69 from the patient. Eight and four genomes
were classiﬁed into MCG groups 2 and 1, respectively.
Both of them contained 3 STs. Two ST1612 genomes
were classiﬁed into MCG group 7–2 (Figure 1).

Based on the distribution of mutational SNPs in
core genomes, 114 genomes were clustered into six
lineages. Each MCG group consisted of one lineage,
except for MCG group 4. Both Lineages 4 and 5
were composed of MCG group 4, whereas contained
14 and 55 genomes, respectively. Lineage 3 was
divided into Lineages 3a and 3b. Interestingly, gen-
omes of Lineage 3a and Lineage 3b were mainly
from China and USA, respectively. Compared with
Lineage 3a, composed of ST373 genomes, Lineage 3b
was mainly composed of ST94 genomes. Lineages 5a
and 5b were mainly composed of ST971 and ST29
genomes, respectively (Figure 1).

The difference in virulence among S. suis serotype 7
strains.

In order to evaluate the virulence level of S. suis ser-
type 7 population, we compared the survival level of
C57BL/6 mice infected with S. suis highly pathogenic
serotype 2 strain P1/7, strain GX69, and additional
21 serotype 7 representative strains. Most mice
infected with S. suis serotype 7 strains showed obvious
septic signs during the infection, such as rough hair
coat, swollen eyes, weakness, and shivering. The
apparent diversity in the survival curves of mice
infected with S. suis serotype 7 strains were observed.
A signiﬁcant difference was observed in survival
curves between mice infected with strains P1/7 and
GX69 (p = .0002), which attributed to the di-

ference in mortality at the early phase of the infection. Mice
infected with P1/7 had a 20% survival rate at 12 h
post-infection, while mice infected with strain GX69
had a 65% survival rate at the same time point. Nota-

ably, the survival levels of mice infected with strain
GX69 dramatically decreased after 12 h post-infection.
Its survival rate decreased to 10% at 24 h post-infect-
gen and was identical to that of strain P1/7 (Figure S1A).
Thus, strain GX69 possessed the capacity to initiate lethal infection in C57BL/6 mice and was classiﬁed as a virulent strain.

(1) Among additional 21 serotype 7 representative
strains, the mortalities of mice infected with
eight S. suis serotype 7 strains at 96 h post-infect-
ion were less than 50%. These strains were
classiﬁed as lowly virulent strains. Interestingly,
half of them were isolated from diseased pigs.

(1) None of the mice infected with strains
128.01B, 173B, WUSS316, WUSS302, and
8074 died within the infection period (Table
2), even though strains 128.01B, 173B, and
8074 were isolated from diseased pigs.
Figure 1. A maximum-likelihood phylogenetic tree of 114 S. suis serotype 7 genomes. The phylogenetic tree was constructed based on mutational SNPs differences across the whole core genome. The S. pneumoniae ATCC 700669 was used as an outgroup to root the tree. The strains were coloured based on the isolation regions, grey for Europe, orange-yellow for North America, and orange-red for China. The scale is given as the number of substitutions per variable site.
Table 2. The value of mortality and statistical comparison in the survival assay.

| Lineage | Strains | 6h  | 12h  | 18h  | 24h  | 36h  | 48h  | 60h  | 72h  | 96h  | p value⁴ |
|---------|---------|-----|------|------|------|------|------|------|------|------|----------|
|         |         | 0±0 | 50±7 | 75±4 | 100±0| 100±0| 100±0| 100±0| 100±0| 100±0|         |
| lineage1| YS12    | 0±0 | 10±7 | 85±4 | 90±7 | 90±7 | 90±7 | 90±7 | 90±7 | 90±7 | V        |
|         | WUSS415 | 0±0 | 0±0  | 45±32| 90±0 | 90±0 | 95±4 | 95±4 | 95±4 | 100±0| V        |
|         | 93.01B  | 0±0 | 0±0  | 70±5 | 80±0 | 80±0 | 80±0 | 80±0 | 80±0 | 80±0 | V        |
| lineages2| 126.01B | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L        |
|         | 128.01B | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L        |
|         | 173B    | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
| lineages3a| GX69   | 0±0 | 35±14| 65±7 | 90±7 | 90±7 | 90±7 | 90±7 | 90±7 | 90±7 | V<.0002  |
|         | WUSS316 | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | WUSS382 | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | 2018WUSS020 | 0±0 | 23±3 | 57±7 | 70±5 | 80±0 | 80±0 | 80±0 | 80±0 | 80±0 | L<.0001  |
|         | 2019WUSS020 | 0±0 | 30±7 | 70±7 | 80±0 | 80±0 | 80±0 | 80±0 | 80±0 | 80±0 | L<.0001  |
| lineages4| YS63    | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | WUSS004 | 0±0 | 23±11| 37±12| 70±8 | 90±5 | 93±5 | 93±5 | 93±5 | 93±5 | V<.0001  |
|         | WUSS005 | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | WUSS029 | 0±0 | 10±7 | 15±4 | 45±18| 75±11| 95±4 | 95±4 | 95±4 | 95±4 | V<.0001  |
|         | WUSS032 | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | WUSS413 | 0±0 | 20±5 | 47±7 | 83±10| 97±3 | 100±0| 100±0| 100±0| 100±0| V<.0001  |
|         | 149B    | 0±0 | 37±14| 50±17| 93±3 | 100±0| 100±0| 100±0| 100±0| 100±0| V<.0001  |
|         | 150B    | 0±0 | 0±0  | 75±3 | 90±0 | 100±0| 100±0| 100±0| 100±0| 100±0| V<.0001  |
|         | 8074    | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | 151B    | 0±0 | 33±12| 60±14| 100±0| 100±0| 100±0| 100±0| 100±0| 100±0| V<.0001  |
| lineages6| 2018WUSS100 | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L<.0001  |
|         | P1/7    | 60±7| 80±7 | 90±0 | 90±0 | 90±0 | 90±0 | 90±0 | 90±0 | 90±0 | H        |
|         | Control | 0±0 | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | 0±0  | L        |

aThe mortality represented as mean ± SED (calculated via the Kaplan–Meier method) at different post-infection time points was present.
bH indicates highly virulent, V indicates virulent, and L indicates lowly virulent.
cThe survival curves of different infected groups were compared using Gehan–Breslow–Wilcoxon test.

Emerging Microbes & Infections
(2) The survival mice infected with strains 
126.01B \( (p < .0001) \), WUSS382 \( (p < .0001) \), and 2018WUSS100 \( (p < .0001) \) were significantly higher than that of mice infected with strain GX69 (Table 2). The strain 126.01B was isolated from diseased pig.

(2) The survival levels of mice infected with remaining 13 S. suis serotype 7 representative strains were significantly higher than that of mice infected with strains PI/7. However, the mortalities of mice infected with these strains at 96 h post-infection reached or exceeded 80% (Table 2). These strains were classified as virulent strains. It is noteworthy that seven of them were isolated from healthy pigs.

(1) The survival curves of mice infected with YS63 \( (p < .0001) \), WUSS013 \( (p < .0001) \), and WUSS029 \( (p = .0021) \) were significantly different from that of mice infected strain GX69, since mice infected with three strains mainly died after 24 h post-infection (Table 2 and Figure S1B). Two strains YS63 and WUSS029 were isolated from healthy pigs.

(2) The survival curves of mice infected with 10 strains YS12 \( (p = .2817) \), WUSS415 \( (p = .7557) \), 93.01B \( (p = .0592) \), 2018WUSS020 \( (p = .2705) \), 2019WUSS020 \( (p = .8214) \), WUSS004 \( (p = .0838) \), WUSS413 \( (p = .2015) \), 149B \( (p = .7383) \), 150B \( (p = .3338) \), and 151B \( (p = .9834) \) were similar to that of mice infected strain GX69 (Table 2 and Figure S1C). Therefore, they were classified as virulent strains. Among them, strains YS12, WUSS415, 2018WUSS020, 2019WUSS020, and WUSS413 were isolated from healthy pigs.

Interestingly, all three tested strains from Lineage 1 were classified as virulent strains, while all strains from Lineages 2 and 6 belonged to lowly virulent strains. On the contrary, Lineages 3a, 4, and 5b contained both virulent and lowly virulent strains.

**Distribution of putative S. suis virulence-related genes**

Only three genomes (YS12, WUSS415, and WUSS417) of Lineage 1 were positive for *epf*. Sixteen genomes from Lineages 4 and 6 were *mrp* gene negative. Most of the genomes (98/114, 86.0%) contained putative full-length *mrp* gene copies. Based on the variation in the central portion of the gene, *mrp* was grouped into three subtypes EU, NA1, and NA2 [38]. Subtype NA2 \( (n = 12) \) was only present in genomes of Lineages 1 and 2. Subtype EU \( (n = 6) \) was only found in genomes of Lineage 5a. All genomes of Lineage 3 and 5b harboured subtype NA1 \( (n = 31) \). Compared with those of Lineage 3, *mrp* gene of Lineage 5b \( (n = 49) \) did not contain an intact open reading frame because of a frameshift mutation in 2.1 kb, which may result in the truncated MRP expression. *sly* gene was only present in genomes of Lineages 1, 2, 3, and 4. A premature stop codon was present in *sly* gene of ST373 genomes of Lineage 3 (Figure 1).

Genes *nadR*, *nisR*, *nisK*, *salR*, and *salK* were absent from all serotype 7 genomes. Genes *revS* and *ofs* were only present in genomes of Lineage 1. *sao* gene was widely distributed in serotype 7 genomes, except for genomes of Lineage 6. RD6 was present in strains WUSS415, WUSS417, and YS12, while the remaining RDs tested in the study were absent from all serotype 7 genomes.

**The distribution of AR genes in S. suis serotype 7 genomes**

Thirteen genomes did not harbour any AR genes. Totally, 216 AR genes were present in the remaining 101 genomes. The AR genes belonged to six categories tetracycline, macrolides/lincosamides/streptogramin (MLS), lincosamide, aminoglycosides, trimethoprim, and chloramphenicol (Table 1).

**The tetracyclines resistance genes**

Ninety-three genomes carried tetracycline-resistant genes. Three types of tetracycline-resistant genes were found among them, consisting of *tet* (O), *tet* (M), and *tet* (W). *tet*(O) was the prevalent tetracycline-resistant gene and was present in 72 genomes. Eighteen genomes carried *tet*(M) gene, mainly from Lineage 5b (14/18). *tet*(W) gene was present in three genomes.

**The MLS and lincosamide resistance genes**

Three types of genes were found, consisting of genes *erm* (B), *lsaC*, and *lnu*B. The MLS resistance gene *erm* (B), encoding rRNA adenine N-6-methyltransferase, was prevalent and present in 72 genomes. The lincosamides-streptogramin A resistance gene *lsaC* was found in three genomes of Lineage 6, which simultaneously harboured lincosamide resistance gene *lnu*B.

**The aminoglycosides resistance genes**

Fifteen genomes carried aminoglycosides resistance genes, including streptomycin resistance gene *ant*6ia encoding aminoglycoside O-nucleotidyltransferase \( (n = 13) \), kanamycin resistance gene *aph*(3′)-IIIa encoding aminoglycoside O-phosphotransferase \( (n = 9) \), spectinomycin resistance gene *ant*9ia encoding aminoglycoside 3′-phosphotransferase \( (n = 4) \), and gentamicin and kanamycin resistance gene *aac*(6′)-Ie-aph(2″)-1a encoding aminoglycoside acetyltransferase \( (n = 2) \).
The trimethoprim and chloramphenicol resistance genes

Eleven genomes harboured trimethoprim resistance gene \( dfrF \) encoding dihydrofolate reductase. It is noteworthy that all genomes carried \( dfrF \) gene from Lineage 6. One genome of Lineage 6 harboured chloramphenicol resistance gene \( cat-TC \) encoding chloramphenicol acetyltransferase.

Antimicrobial susceptibility profiles of available strains

To investigate whether AR genes conferred resistance to host strains, we tested the antimicrobial susceptibility of available strains, including 25 Chinese and five Spanish strains, carrying genes responsible for resistance to tetracycline, erythromycin, clindamycin, streptomycin, kanamycin, or gentamycin. Thirty strains harboured tetracycline resistance genes were all resistant to tetracycline, with a MIC value between 12 and 128 \( \mu \)g/mL. Concomitant resistance to erythromycin and clindamycin was found in all strains \((n = 25)\) carrying \( erm(B) \) gene due to the overlapping ribosomal binding sites of the two antibiotics. MIC values for both antibiotics were between 128 and \( >256 \) \( \mu \)g/mL. A high level of kanamycin (MICs > 256 \( \mu \)g/mL) carrying both gentamycin and clindamycin was found in all strains \((n = 34)\).

Resistance to aminoglycosides, trimethoprim, and MLS, and aminoglycosides were present in these ICEs (Figure 2).

The genes \( SNF2 \) and \( PPI \) encoding a putative adenine-specific DNA methylase and a putative peptidyl-prolyl isomerase, respectively, are two specific insertion hot spots for integrating IMEs or CIMEs [32]. In the present study, 34 and 22 IMEs were integrated into the PPI and \( SNF2 \) genes, respectively. All three CIMEs were integrated into \( SNF2 \) gene (Figure 2). All IMEs or CIMEs harboured an 11-bp inverted repeat 5′-TTTTGCAGGACA-3′ in their flanking region. Interestingly, 25 IMEs and two CIMEs were integrated into ICEs and all AR genes in the ICEs were carried by these integrated IMEs and CIMEs. The remaining 31 IMEs and one CIME were located in non-ICE regions. Thirty-two \( tet(O) \) and 23 \( ermB \) genes were present in these IMEs and CIMEs. The integrases of all IMEs were identical and belonged to serine integrase type 3, regardless of their integration site. Meanwhile, the relaxases of the IMEs belonged to the MobV superfamily. Based on their integrase and relaxase types, all IMEs belonged to IME_Class_6. It is noteworthy that all AR genes responsible for resistance to aminoglycosides, trimethoprim, and chloramphenicol were not present in the above MGEs.

AR genes associated with MGEs

To investigate the mechanism to disseminate AR genes, the MGEs harbouring AR genes in \( S. suis \) serotype 7 genomes were predicted. Among 114 genomes, 27 ICEs, 56 IMEs, and three CIMEs (absent of the integrase and relaxase genes) with a complete sequence were detected. These ICEs were distributed in Lineages 2, 4, and 5, whereas IMEs were distributed in Lineages 2, 3, 4, and 5. Totally, 111 of 216 AR genes were present in these MGEs (Figure 2).

Twenty-five ICEs were inserted into \( rplL \) locus. All of them harboured a 15-bp \( att \) sequence 5′-TTATT-TAAGAGTAAC-3′ in the flanking region. ICESSuWUSS029 and ICESuWUSS004 were integrated into downstream of \( rum \) gene. Both of them harboured the 14-bp \( att \) sequences 5′-CACGTGGAGTGCAGT-3′ and 5′-CATGTTGAAGTTGT-3′ in the 5′ and 3′ flanking regions, respectively. All ICEs were classified as Tn5225 group and harboured intact signature proteins VirB4, integrase, and canonical relaxase of the MobP family. Fifty-six AR genes resistant to tetracycline,
In the present study, we first reported a *S. suis* serotype 7 strain (GX69) isolated from a patient with septicemia complicated with pneumonia. The strain GX69 was ST373 and belonged to MCG group 3, whereas ST1 and ST7 are predominant in *S. suis* strains from patients [39], mainly clustered into MCG group 1 [26]. The genotype of *S. suis* classical virulence markers in strain GX69 was $mrp^{NA1}$ $sly^+$ $epf^+$, whereas the prevalent genotype of corresponding virulence markers in human strains was $mrp^{EU}$ $sly^+$ $epf^+$ or $mrp^{NA2}$ $sly^+$ $epf^+$ [9,40]. It is noteworthy that a premature stop codon was present in *sly* gene of strain GX69 and may result in the truncated SLY expression. To evaluate the virulence of strain GX69, the survival level was compared with that of highly pathogenic *S. suis* serotype 2 strain P1/7. Significant differences at the early phase of infection and the similarity at the middle phase of infection were observed between the two strains in the mouse infection model. Our result confirmed that GX69 was a virulent strain and possessed the capacity to initiate lethal infection, even though virulence-associated genes and RDs preferentially present in highly pathogenic *S. suis* serotype 2 strains were almost absent in strain GX69. We proposed that *S. suis* serotype 7 may be considered as a potential zoonotic pathotype, and further investigation of *S. suis* serotype 7 population is needed to improve the prevention and control strategies.

In the present study, *S. suis* serotype 7 population composed of 35 strains and 79 genomes of strains from 1999 to 2019 in nine countries was investigated. Twenty-two STs and five MCG groups were identified among *S. suis* serotype 7 genomes clustered into six lineages based on the distribution of mutational SNPs in the core genomes. Interestingly, since most predominant ST29, ST373, and ST94 were prevalent in respective regions, it suggests that the evolution of *S. suis* serotype 7 population was relevant to the geographical distribution. The evolutionary affinity between ST373 and ST94 was revealed in that they belonged to MCG group 3 and were clustered into Lineage 3. It is noteworthy that significant heterogeneity was observed within ST373 strains, which were clustered into three clades. Similar heterogeneity was previously reported in phylogenetic analysis of ST1 [40], ST7 [9], and ST25 [41].

Based on the results of the survival assay using the C57BL/6 mouse model, the strain GX69 and over 60% additional representative strains tested (13/21) were classified as virulent strains. Among 13 virulent strains, the survival curves of ten *S. suis* serotype 7 representative strains were similar to that of strain GX69. In a recent study, 82.6% *S. suis* serotype 7 strains from North America were pathogenic based
on the clinical information and site of isolation [20].

In the present study, seven strains isolated from healthy pigs were classified as virulent strains, and two of them belong to ST373, the same ST of strain GX69. Therefore, we proposed that the public health threat of S. suis serotype 7, especially those virulent

**Figure 3.** The schematic comparison of the cps gene cluster subtype la to that of lb (A), II (B), and III (C). Each colored arrow represents the gene whose predicted function is shown in the blow panel. HG17, HG18, HG19, HG72, and HG73 genes are indicated. The aroA gene is located on the 3’ side of each locus. Regions of over 70% identity were marked by blue shading.
ST373 strains, should not be ignored. Coincidentally, healthy pigs were reported to be a reservoir of strains with high virulence potential in humans [39,42]. Moreover, four strains isolated from diseased pigs were classified as low virulent. A correlation between the virulence level of strains and their origin (diseased or healthy pigs) could not be observed in the present study. A similar result was also reported in our previous study [43]. It should be noted that the presence of clinical signs in pigs may also depend on co-infection with some viral and bacterial pathogens [42]. Thus, the public health significance of strains may not be accurately evaluated only based on the clinical information of their host.

In the present study, three classical virulence markers _mrp_, _sly_, and _epf_ were not critical virulence indicators of the _S. suis_ serotype 7 strains. However, a significant correlation of genotypes and variations of three genes and their distribution in lineages was observed, suggesting that these genes correlate with the evolution of _S. suis_ serotype 7 population rather than virulence. Most virulence-associated genes preferentially present in highly pathogenic _S. suis_ serotype 2 strains were absent from all serotype 7 genomes. Previous studies also reported that these virulence markers studied in _S. suis_ serotype 2 strains were not suitable as virulence markers for _S. suis_ non-serotype 2 strains [3,12,44]. Thus, _S. suis_ serotype 7 virulent strains may utilize a different pathogenesis strategy. Because of the high diversity of virulence levels within _S. suis_ serotype 7 population, further studies are necessary to identify reliable virulence indicators of _S. suis_ serotype 7 strains. Using multiple animal models to accurately pathotype ST373 strains combined with comparative genomic analysis of ST373 strains with different virulence levels may be feasible.

Six categories of AR genes are present in _S. suis_ serotype 7 genomes. The predominant categories were tetracycline and MLS resistance genes. High rates of resistance to tetracycline, macrolide, lincosamide, and erythromycin have been reported in both human and pig isolates of _S. suis_ in the last 20 years [45–49]. Tetracycline, lincosamide, and macrolide are used extensively for therapy and metaphylaxis in the swine industry [32,50,51], contributing to the emergence and spread of associated resistance. The most prevalent tetracycline resistance gene was _tet_ (O). This is different from what was previously reported for serotype 2 strains, which have been shown to mainly carry _tet_(M) and _tet_(W) [45,48].

Previous studies have shown that MGEs play a significant role in the horizontal transfer of AR genes in _S. suis_ [31,32]. Twenty-seven ICEs carrying AR genes were found in _S. suis_ serotype 7 genomes, although intact prophages carrying AR genes were not detected. Conversely, AR genes in _S. suis_ serotype 31 population were majorly present in prophages [12]. In the present study, two types of DNA cargo of IMEs and CIMEs with AR genes were integrated into genes SNF2 or _PPI_ of all ICEs. Similar IMEs were also integrated into SNF2 gene of ICESsuZJ20091101-1 (KX077882.1), ICESsuLP081102 (KX077885.1), ICESsuJH1301 (KX077887.1) [31], and ICEsSuD9 [52]. A similar CIME was also inserted into the same integration site of ICESsuBS6 [53]. Likely, the exchange, acquisition, and deletion of the IME/CIME module may contribute to the evolution of ICEs. In the present study, all AR genes in ICEs were carried by these IMEs and CIMEs. Moreover, IMEs and CIMEs carrying AR genes were also present in non-ICE regions of additional 32 _S. suis_ serotype 7 genomes.

IMEs were reported to be more widespread than ICEs in _S. suis_ [52]. In this work, IMEs were also found to be highly prevalent in _S. suis_ serotype 7 genomes. Over 50% AR genes identified in serotype 7 genomes were present in IMEs. IMEs mainly carried tetracycline, erythromycin, and lincosamide resistance genes. We propose that IMEs may play a critical role in the horizontal transfer of these AR genes in _S. suis_ serotype 7. Interestingly, the proportion of genome carrying IMEs was higher in Lineages 2, 3a, 4, and 5a. Our data indicated that the transmission patterns of AR genes might be related to the evolution of serotype 7 population. CIMEs are decayed IMEs, which are _cis_-mobilizable elements without integration and relaxase genes but with _attL_ and _attR_ sites. CIMEs carrying _tet_(O) and _ermB_ genes were found in two ICEs and non-ICE region of one additional genome. Further study is needed to investigate the role of CIMEs in the transmission of AR genes.

Finally, different organizations of _cps_ loci were observed among _S. suis_ serotype 7 population. These differences can be attributed to the variable presence of HG17, HG18, and HG19. The function of HG17 was related to aminotransferase, while the functions of both HG18 and HG19 were ATP-binding proteins. HG17, HG18, and HG19 existed widely in _S. suis_ serotypes 4, 5, 17, 18, 19, and 23 reference strains. It is noteworthy that HG17, HG18, and HG19 were also inverted in _cps_ gene clusters of _S. suis_ serotypes 17 and 23 reference strains. Among 23 _cps_ gene clusters of _S. suis_ serotype 7 genomes, HG17, HG18 and HG19 were replaced by HG72 and HG73. The functions of HG72 and HG73 were related to carboxyvinyltransferase and unknown, respectively. HG72 and HG73 also existed in _cps_ gene clusters of _S. suis_ serotypes 11 and 30 reference strains. Based on the agglutination test results, the capsular antigenic phenotype was not affected by the variable presence of HG17, HG18, and HG19. Therefore, these HGs may not be involved in the forming of serotype 7-specific epitopes. The subtype _cps_7-I was most predominant among _S. suis_ serotype 7 population and dispersed in different lineages and geographical regions. On the other hand, subtype _cps_7-II was majorly present in
strains from Europe (such as UK and Spain), while it was absent in strains from North America. Different cps subtypes may enhance the fitness of corresponding host strains in specific environments.

In conclusion, our data confirmed *S. suis* serotype 7 is a non-negligible pathotype and deepened the understanding of *S. suis* serotype 7 population. Geographically dependent characteristics were revealed in the evolution of *S. suis* serotype 7 population. Our study provided valuable information for the improved surveillance of *S. suis* provided valuable information for the improved surveillance of *S. suis* serotype 7 strains. Further studies are needed to identify the virulence indicators to predict the public health significance of *S. suis* serotype 7 strains.

**Acknowledgements**

HZ and ZW designed the project; HZ drafted the manuscript; ZW, MG and AV reviewed the manuscript; PL, and JW carried out the experiments and generated the data; PL, AE, WM, MG, AV, PD, ML, and HZ analysed the data. All authors have read and approved the final version of the manuscript.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by the National Science and Technology Major Project from the Ministry of Health of the People’s Republic of China [grant number 2017ZX10303405-002] and the National Natural Science Foundation of China [grant number 81572044].

**ORCID**

Zongfu Wu  http://orcid.org/0000-0002-6048-6829

**References**

[1] Goyette-Desjardins G, Auger JP, Xu J, et al. *Streptococcus suis*, an important pig pathogen and emerging zoonotic agent—an update on the worldwide distribution based on serotyping and sequence typing. Emerg Microb Infect. 2014 Jun;3(6):e45.

[2] Bojarska A, Janas K, Pejsak Z, et al. Diversity of serotypes and new cps loci variants among *Streptococcus suis* isolates from pigs in Poland and Belarus. Vet Microbiol. 2020;240:108534.

[3] Huang J, Liu X, Chen H, et al. Identification of six novel capsular polysaccharide loci (NCL) from *Streptococcus suis* multidrug resistant non-typeable strains and the pathogenic characteristic of strains carrying new NCLs. Transbound Emerg Dis. 2019;66 (2):995–1003.

[4] Zheng H, Qiu X, Roy D, et al. Genotyping and investigating capsular polysaccharide synthesis gene loci of non-serotypeable *Streptococcus suis* isolated from diseased pigs in Canada. Vet Res. 2017;48(1):10.

[5] Qiu X, Bai X, Lan R, et al. Novel capsular polysaccharide loci and new diagnostic tools for high-throughput capsular gene typing in *Streptococcus suis*. Appl Environ Microbiol. 2016;82(24):7102–7112.

[6] Huong VT, Ha N, Huy NT, et al. Epidemiology, clinical manifestations, and outcomes of *Streptococcus suis* infection in humans. Emerg Infect Dis. 2014;20 (7):1105–1114.

[7] Yu H, Jing H, Chen Z, et al. Human *Streptococcus suis* outbreak, Sichuan, China. Emerg Infect Dis. 2006;12 (6):914–920.

[8] Ye C, Zhu X, Jing H, et al. *Streptococcus suis* sequence type 7 outbreak, Sichuan, China. Emerg Infect Dis. 2006;12(8):1203–1208.

[9] Wang M, Du P, Wang J, et al. Genomic Epidemiology of *Streptococcus suis* sequence type 7 sporadic infections in the Guangxi Zhuang autonomous region of China. Pathogens. 2019;8(4).

[10] Schultsz C, Jansen E, Keijzers W, et al. Differences in the population structure of invasive *Streptococcus suis* strains isolated from pigs and from humans in The Netherlands. PLoS One. 2012;7(5):e33854.

[11] Kersdin A, Hatrongjit R, Gottschalk M, et al. Emergence of *Streptococcus suis* serotype 9 infection in humans. J Microbiol Immunol Infect. 2017;50(4):545–546.

[12] Wang X, Sun J, Bian C, et al. The population structure, antimicrobial resistance, and pathogenicity of *Streptococcus suis* cps31. Vet Microbiol. 2021;259:109149.

[13] Prüfer TL, Rohde J, Verspohl J, et al. Molecular typing of *Streptococcus suis* strains isolated from diseased and healthy pigs between 1996–2016. PLoS One. 2019;14 (1):e0210801.

[14] Unterweger C, Baums CG, Hocher M, et al. [Clinical situation, diagnosis and prevention of a *Streptococcus suis* serotype 7 problem on a farm]. Berl Munch Tierarztl Wochenschr. 2014 May-Jun;127(5–6):194–201.

[15] Nutrivong T, Angkititrakul S, Jiwakanon N, et al. Identification of major *Streptococcus suis* serotypes 2, 7, 8 and 9 isolated from pigs and humans in upper northeastern Thailand. Southeast Asian J Trop Med Public Health. 2014 Sep;45(5):1173–1181.

[16] Tian Y, Aarestrup FM, Lu CP. Characterization of *Streptococcus suis* serotype 7 isolates from diseased pigs in Denmark. Vet Microbiol. 2004;103(1–2):55–62.

[17] Tarradas C, Perea A, Vela AI, et al. Distribution of serotypes of *Streptococcus suis* isolated from diseased pigs in Spain. Vet Rec. 2004;154(21):665–666.

[18] MacLennan M, Foster G, Dick K, et al. *Streptococcus suis* serotypes 7, 8 and 14 from diseased pigs in Scotland. Vet Rec. 1996;139(17):423–424.

[19] Rieckmann K, Seydel A, Szewczyk K, et al. *Streptococcus suis* cps7: an emerging virulent sequence type (ST29) shows a distinct, IgM-determined pattern of bacterial survival in blood of piglets during the early adaptive immune response after weaning. Vet Res. 2018;49(1):48.

[20] Estrada AA, Gottschalk M, Rossov S, et al. Serotype and genotype (Multilocus sequence type) of *Streptococcus suis* isolates from the United States serve as predictors of pathotype. J Clin Microbiol. 2019;57(9).

[21] Ishida S, Tien le HT, Osawa R, et al. Development of an appropriate PCR system for the reclassification of *Streptococcus suis*. J Microbiol Meth. 2014;107:66–70.

[22] Bai X, Liu Z, Ji S, et al. Simultaneous detection of 33 *Streptococcus suis* serotypes using the luminex xTAG (R) assay. J Microbiol Meth. 2015;117:95–99.
23. Chatellier S, Harel J, Zhang Y, et al. Phylogenetic diversity of Streptococcus suis strains of various serotypes as revealed by 16S rRNA gene sequence comparison. Int J Syst Bacteriol. 1998;48(Pt 2):581–589.

24. Tien LHT, Nishibori T, Nishitani Y, et al. Reappraisal of the taxonomy of Streptococcus suis serotypes 20, 22, 26, and 33 based on DNA-DNA homology and sodA and recN phylogenies. Vet Microbiol. 2013;162(2-4):842–849.

25. Wang J, Yi X, Liang P, et al. Investigation of the genomic and pathogenic features of the potentially zoonotic Streptococcus parasitisi. Pathogens. 2021 Jul 2;10(7):1–18.

26. Chen C, Zhang W, Zheng H, et al. Minimum core genome sequence typing of bacterial pathogens: a unified approach for clinical and public health microbiology. J Clin Microbiol. 2013;51(8):2582–2591.

27. Ye C, Zheng H, Zhang J, et al. Clinical, experimental, and genomic differences between intermediate pathogenic, highly pathogenic, and epidemic Streptococcus suis. J Infect Dis. 2009;199(1):97–107.

28. Zheng H, Du P, Qiu X, et al. Genomic comparisons of Streptococcus suis serotype 9 strains recovered from diseased pigs in Spain and Canada. Vet Res. 2018;49(1):1.

29. Zheng X, Zheng H, Lan R, et al. Identification of genes and genomic islands correlated with high pathogenicity in Streptococcus suis using whole genome tiling microarrays. PLoS One. 2011;6(3):e17987.

30. Hu Y, Yang X, Qin J, et al. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. Nature Commun. 2013;4:2151.

31. Huang J, Ma J, Shang K, et al. Evolution and diversity of the antimicrobial resistance associated mobiles in Streptococcus suis: A probable mobile genetic elements reservoir for other streptococci. Front Cell Infect Microbiol. 2016;6(118):1–14.

32. Libante V, Nombre Y, Coluzzi C, et al. Chromosomal conjugal and mobilizable elements in Streptococcus suis: major actors in the spreading of antimicrobial resistance and bacteriocin synthesis genes. Pathogens. 2020;9(1):1–23.

33. Okura M, Takamatsu D, Maruyama F, et al. Genetic analysis of capsular polysaccharide synthesis gene clusters from all serotypes of Streptococcus suis: potential mechanisms for generation of capsular variation. Appl Environ Microbiol. 2013;79(8):2796–2806.

34. Holden MT, Hauser H, Sanders M, et al. Rapid evolution of virulence and drug resistance in the emerging zoonotic pathogen Streptococcus suis. PLoS One. 2009;4(7):e6072.

35. Lachance C, Gottschalk M, Gerber PP, et al. Exacerbated type II interferon response drives hyper-virulence and toxic shock by an emergent epidemic strain of Streptococcus suis. Infect Immun. 2013;81(6):1928–1939.

36. Lachance C, Segura M, Gerber PP, et al. Toll-like receptor 2-independent host innate immune response against an epidemic strain of Streptococcus suis that causes a toxic shock-like syndrome in humans. PLoS One. 2013;8(5):e65031.

37. Marie J, Morvan H, Berthelot-Herault F, et al. Antimicrobial susceptibility of Streptococcus suis isolated from swine in France and from humans in different countries between 1996 and 2000. J Antimicrob Chemother. 2002;50(2):201–209.

38. Fittipaldi N, Fuller TB, Teel JE, et al. Serotype distribution and production of muramidase-released protein, extracellular factor and suilysin by field strains of Streptococcus suis isolated in the United States. Vet Microbiol. 2009;139(3-4):310–317.

39. Dong X, Chao Y, Zhou Y, et al. The global emergence of a novel Streptococcus suis clade associated with human infections. EMBO Mol Med. 2021;13(7):e13810.

40. Callejo R, Zheng H, Du P, et al. Streptococcus suis serotype 2 strains isolated in Argentina (South America) are different from those recovered in North America and present a higher risk for humans. JMM Case Rep. 2016;3(5):e005066.

41. Atthey TB, Teatero S, Takamatsu D, et al. Population structure and antimicrobial resistance profiles of Streptococcus suis serotype 2 sequence type 25 strains. PLoS One. 2016;11(3):e0150908.

42. Obradovic MR, Segura M, Segales J, et al. Review of the speculative role of co-infections in Streptococcus suis-associated diseases in pigs. Vet Res. 2021;52(1):49.

43. Zheng H, Lan R, Zheng X, et al. Comparative genomic hybridization identifies virulence differences in Streptococcus suis. PLoS One. 2014;9(2):e87866.

44. Dong W, Zhu Y, Ma Y, et al. Multilocus sequence typing and virulence genotyping of Streptococcus suis serotype 9 isolates revealed high genetic and virulence diversity. FEMS Microbiol Lett. 2017;364(22):1–8.

45. Palmieri C, Varaldo PE, Facinelli B. Streptococcus suis, an emerging drug-resistant animal and human pathogen. Front Microbiol. 2011;2:235.

46. Princivalle MS, Palmieri C, Magi G, et al. Genetic diversity of Streptococcus suis clinical isolates from pigs and humans in Italy (2003-2007). Euro Surveill. 2009;14(33):1–7.

47. Zhang C, Ning Y, Zhang Z, et al. In vitro antimicrobial susceptibility of Streptococcus suis strains isolated from clinically healthy sows in China. Vet Microbiol. 2008;131(3-4):386–392.

48. Ye C, Bai X, Zhang J, et al. Spread of Streptococcus suis sequence type 7, China. Emerg Infect Dis. 2008;14(5):787–791.

49. Wisselink HJ, Veldman KT, Van den Eede C, et al. Quantitative susceptibility of Streptococcus suis strains isolated from diseased pigs in seven European countries to antimicrobial agents licensed in veterinary medicine. Vet Microbiol. 2006;113(1-2):73–82.

50. Xu Z, Xie J, Peters BM, et al. Longitudinal surveillance on antibiogram of important gram-positive pathogens in southern China, 2001 to 2015. Microb Pathog. 2017;103:80–86.

51. Seitz M, Valentín-Weigand P, Willenborg J. Use of antibiotics and antimicrobial resistance in veterinary medicine as exemplified by the swine pathogen Streptococcus suis. Curr Top Microbiol Immunol. 2016;398:103–121.

52. Huang K, Song Y, Zhang Q, et al. Characterisation of a novel integrative and conjugative element ICESsD9 carrying erm(B) and tet(O) resistance determinants in Streptococcus suis, and the distribution of ICESsD9-like elements in clinical isolates. J Global Antimicrob Resis. 2016;7:13–18.

53. Huang J, Chen L, Li D, et al. Emergence of a vanG-carrying and multidrug resistant ICE in zoonotic pathogen Streptococcus suis. Vet Microbiol. 2018;222:109–113.