Whole-genome sequence analyses of Glaesser*ella parasuis* isolates reveals extensive genomic variation and diverse antibiotic resistance determinants

Xiulin Wan [1], **Xinhui Li** [2], **Todd Osmundson** [3], **Chunling Li** [Corresp., 4], **He Yan** [Corresp. 1]

[1] School of Food Science and Engineering, South China University of Technology, Guangzhou, China

[2] Department of Microbiology, University of Wisconsin-La Crosse, La Crosse, United States

[3] Department of Biology, University of Wisconsin-La Crosse, La Crosse, United States

[4] Institute of Animal Health Guangdong Academy of Agricultural Sciences, Guangzhou, China

Corresponding Authors: Chunling Li, He Yan
Email address: lclclare@163.com, yanhe@scut.edu.cn

Background: *Glaesserella parasuis* (*G. parasuis*) is a respiratory pathogen of swine and the etiological agent of Glässer's disease. The structural organization of genetic information, antibiotic resistance genes, potential pathogenicity, and evolutionary relationships among global *G. parasuis* strains remain unclear. The aim of this study was to better understand patterns of genetic variation, antibiotic resistance factors, and virulence mechanisms of this pathogen.

Methods: The whole-genome sequence of a ST328 isolate from diseased swine in China was determined using Pacbio RS II and Illumina MiSeq platforms and compared with 54 isolates from China sequenced in this study and 39 strains from China and eight other countries sequenced by previously. Patterns of genetic variation, antibiotic resistance, and virulence mechanisms were investigated in relation to the phylogeny of the isolates. Electrotransformation experiments were performed to confirm the ability of pYL1 - a plasmid observed in ST328 - to confer antibiotic resistance.

Results: The ST328 genome contained a novel Tn*6678* transposon harbouring a unique resistance determinant. It also contained a small broad-host-range plasmid pYL1 carrying *aac(6')-Ie-aph(2'')-Ia* and bla_{ROB-1}; when transferred to *Staphylococcus aureus* RN4220 by electroporation, this plasmid was highly stable under kanamycin selection. Most (85.13–91.74%) of the genetic variation between *G. parasuis* isolates was observed in the accessory genomes. Phylogenetic analysis revealed two major subgroups distinguished by country of origin, serotype, and multilocus sequence type (MLST). Novel virulence factors (*gigP, malQ,* and *gmhA*) and drug resistance genes (*norA, bacA*, *ksgA,* and *bcr*) in *G. parasuis* were identified. Resistance determinants (*sul2, aph(3'')-Ib, norA, bacA*, *ksgA,* and *bcr*) were widespread across isolates, regardless of serovar, isolation source, or geographical location.

Conclusions: Our comparative genomic analysis of worldwide *G. parasuis* isolates provides valuable insight into the emergence and transmission of *G. parasuis* in the swine industry. The result suggests the importance of transposon-related and/or plasmid-related gene variation in the evolution of *G. parasuis*.

Whole-genome sequence analyses of Glaesser*ella parasuis* isolates reveals extensive genomic variation and diverse antibiotic resistance determinants

Xiulin Wan[1], Xinhui Li[2], Todd Osmundson[3], Chunling Li[4,*], He Yan[1,*]

[1] School of Food Science and Engineering, South China University of Technology, Guangzhou, Guangdong, China

[2] Department of Microbiology, University of Wisconsin-La Crosse, 1725 State Street, La Crosse, WI 54601, USA

[3] Department of Biology, University of Wisconsin-La Crosse, 1725 State Street, La Crosse, WI 54601, USA

[4] Institute of Animal Health Guangdong Academy of Agricultural Sciences, Guangzhou, Guangdong, China

*Corresponding author:

He Yan

381 Wushan Road, Tianhe District, Guangzhou, Guangdong, 510641, China

E-mail address: yanhe@scut.edu.cn

Chunling Li

21 Baishigang Street, Wushan Road, Tianhe District, Guangzhou, Guangdong, 510640, China

E-mail address: lclclare@163.com

[†]These authors have contributed equally to this work.

Abstract:

Background: *Glaesserella parasuis* (*G. parasuis*) is a respiratory pathogen of swine and the etiological agent of Glässer's disease. The structural organization of genetic information, antibiotic resistance genes, potential pathogenicity, and evolutionary relationships among global *G. parasuis* strains remain unclear. The aim of this study was to better understand patterns of genetic variation, antibiotic resistance factors, and virulence mechanisms of this pathogen.

Methods: The whole-genome sequence of a ST328 isolate from diseased swine in China was determined using Pacbio RS II and Illumina MiSeq platforms and compared with 54 isolates from China sequenced in this study and 39 strains from China and eigtht other countries sequenced by previously. Patterns of genetic variation, antibiotic resistance, and virulence mechanisms were investigated in relation to the phylogeny of the isolates. Electrotransformation experiments were performed to confirm the ability of pYL1 - a plasmid observed in ST328 - to confer antibiotic resistance.

Results: The ST328 genome contained a novel Tn*6678* transposon harbouring a unique resistance determinant. It also contained a small broad-host-range plasmid pYL1 carrying *aac(6')-Ie-aph(2'')-Ia* and bla_{ROB-1}; when transferred to *Staphylococcus aureus* RN4220 by electroporation, this plasmid was highly stable under kanamycin selection. Most (85.13–91.74%) of the genetic variation between *G. parasuis* isolates was observed in the accessory genomes. Phylogenetic analysis revealed two major subgroups distinguished by country of origin, serotype, and multilocus sequence type (MLST). Novel virulence factors (*gigP, malQ,* and *gmhA*) and drug resistance genes (*norA, bacA*, *ksgA,* and *bcr*) in *G. parasuis* were identified. Resistance determinants (*sul2, aph(3'')-Ib, norA, bacA*, *ksgA,* and *bcr*) were widespread across isolates, regardless of serovar, isolation source, or geographical location.

Conclusions: Our comparative genomic analysis of worldwide *G. parasuis* isolates provides valuable insight into the emergence and transmission of *G. parasuis* in the swine industry. The result suggests the importance of transposon-related and/or plasmid-related gene variations in the evolution of *G. parasuis*.

Introduction

Glaesserella parasuis, a gram-negative bacterium in the family *Pasteurellaceae* (*Dickerman et al., 2020*), is a respiratory pathogen that affects swine. It is the etiological agent of Glässer's disease, which can lead to pneumonia without signs of systemic disease (*Brockmeier*, *2004*). As China is one of the world's largest pork producers, with more than 463 million pigs accounting for approximately 50% of global population (*Zhou et al.*, *2013*), *G. parasuis* outbreaks in this country could pose a significant threat to pig health and economic loss worldwide (*Brockmeier et al.*, *2014*). Disease progression and severity are influenced by virulence and antibiotic resistance, both of which can result from evolutionary processes including mutation and horizontal gene transfer (*Deng et al., 2019*). Although antibiotic resistance may incur fitness costs in terms of virulence, the two phenomena may also act synergistically (*Geisinger and Isberg, 2017*).

Antimicrobial agents are widely used to prevent and control *G. parasuis* infection; however, overuse of antibiotics for non-therapeutic applications - including promoting growth in healthy individuals - has resulted in the evolution of antibiotic resistant *G. parasuis* in farming environments (*Zhao et al.*, *2018*). Antibiotic resistance in *G. parasuis* is mainly conferred by a combination of transferable antibiotic resistance genes (ARGs) and multiple target gene mutations. To date, two β-lactam resistance genes ($bla_{\text{ROB-1}}$ and bla_{TEM}), an aminoglycoside-resistance gene (*aac (6')-Ib-cr*) and a mutation in the six copies of the 23S rRNA gene, associated with macrolide resistance, have been reported in *G. parasuis* (*Doi and Arakawa*, *2007*; *San et al.*, *2007*; *By Guo et al.*, *2012*). *G. parasuis* strains often harbour multiple resistance genes and multi-drug resistance phenotypes, thus deterring clinical treatment.

PCR-based studies of *G. parasuis* strains have identified ARGs including *tetB*, *aph(3'')-Ib, aph(6)-Id*, *floR*, *sul1*, and *sul2* (*Wissing et al.*, *2001*; *San et al.*, *2007*; *Zhao et al.*, *2018*), and virulence factors including the haemolysin operon (*hhdBA*), iron acquisition genes (*cirA*, *tbpA/B* and *fhuA*), the restriction modification system *hsdS*, and genes involved in sialic acid utilization

(neuraminidase *nanH* and sialyltransferase genes *neuA*, *siaB* and *lsgB*) (*Martinez-Moliner et al.*, *2012*; *Costa-Hurtado and Aragon*, *2013*). Recently, whole-genome sequencing (WGS) has emerged as a powerful tool for predicting antibiotic resistance and pathogenic potential in *G. parasuis*. For instance, Li et al. (*2013*) reported two *G. parasuis* strains with potential resistance towards the antibiotics ciprofloxacin, trimethoprim, and penicillin, based on the presence of associated resistance genes; Nicholson et al. (*2018*) reported genomic differences in the toxin-antitoxin systems between phenotypically distinct *G. parasuis* strains from Japan and Sweden; and Bello-Orti et al. (*2014*) noted the role of mobile genetic elements and strain-specific accessory genes in fostering high genomic diversity between pathogenic strains of the same serovar from diseased pigs in Japan, China, and the USA.

Though significant effort has been focused on exploring ARGs, virulence factors and other genetic characteristics of various *G. parasuis* strains, the structural organization of genetic information, ARGs, potential pathogenicity determinants, and evolutionary relationships among global *G. parasuis* strains remain unclear. In this study, we sequenced a multidrug-resistant isolate from diseased swine in Dongguan, China, then compared this genome sequence with those of 54 isolates from China sequenced by us and 39 strains from China and eight other countries sequenced by other researchers in order to improve our understanding of genomic diversity in *G. parasuis* and provide information for gaining better control to treat these infections.

Materials & Methods

Isolates

The multidrug-resistant *G. parasuis* isolate HPS-1 examined in this study belongs to serotype 4 and was originally isolated from the lungs of a pig suffering from Glässer's disease in a commercial pig farm in Dongguan city, Guangdong province, China, in 2017. Susceptibility to 19 antimicrobial agents was determined by the disc agar diffusion method and the broth

102 microdilution method (*Pruller et al.*, *2017*). The isolate was determined to be resistant to β-
103 lactams, aminoglycosides, macrolides, quinolones, lincomycin, and sulfonamides (Table S1).

104 The other 54 *G. parasuis* isolates were obtained from diseased pigs from more than 20
105 geographically dispersed farms in China between November 2007 and May 2017 (Table S2).
106 Bacteria species were identified by biochemical tests and 16S diagnostic PCR (*Oliveira et al.*,
107 *2001*; *de la Fuente et al.*, *2007*).. All 55 *G. parasuis* isolates were characterised using serotyping
108 and MLST as previously described (*Wang et al.*, *2016*; *Jia et al.*, *2017*).

109 **Genome sequencing, assembly, and bioinformatics analysis**

110 Isolates were cultured on tryptic soy agar or in tryptic soy broth (Oxoid, Hampshire, UK)
111 supplemented with 10 mg/mL nicotinamide adenine dinucleotide and 5% bovine serum at 37°C
112 in 5% CO_2 for 24 h. Total genomic DNA was extracted using the DNeasy DNA extraction kit
113 (Axygen, Union City, CA, USA).

114 Among the 55 isolates, one multidrug-resistant isolate (HPS-1) and one sensitive isolate
115 (HPS-2) from diseased swine in Guangdong were randomly selected for WGS using the PacBio
116 RSII (Pacific Biosciences, MenloPark, CA, USA) and Illumina MiSeq (Illumina, San Diego, CA,
117 USA) platforms as previously described (*Zheng et al.*, *2017*). The genome assemblies of HPS-1
118 generated in this study were deposited in GenBank under accession number CP040243. The
119 plasmid pYL1 and transposon Tn*6678* of HPS-1 were submitted to GenBank under accession
120 number MK182379 and MK994978, respectively. Genomic libraries of the other 53 genomes
121 were generated and sequenced using the Illumina HiSeq 4000 system (Illumina, San Diego, CA,
122 USA) as previously described (*Soge et al.*, *2016*). WGS data were assembled using SOAPdenovo
123 v1.05 software (assembly statistics available in supplementary materials, Table S3). Gene
124 prediction was performed using GeneMarkS (*Besemer et al.*, *2001*), and a whole-genome
125 BLAST (*Altschul et al.*, *1990*) searches (E-value ≤ $1e^{-5}$, minimal alignment length percentage ≥
126 80%) against 6 databases: Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of
127 Orthologous Groups (COG), NCBI non-redundant protein database (NR), Swiss-Prot, Gene

Ontology (GO), and TrEMBL.

Phylogenetic and clustering analyses

Two phylogenetic trees were constructed to assess the relatedness of the 55 *G. parasuis* strains and 39 previously published genome sequences using single-copy core orthologs and single nucleotide polymorphisms (SNPs) (Table S2). Phylogenetic inference was conducted using a maximum-likelihood optimality criterion as implemented in PhyML v3.0 (*Guindon et al.*, *2010*). The WAG amino acid substitution matrix was used for inference of the single-copy core ortholog tree, and the HKY85 nucleotide substitution model was used for inference of the SNP tree. The SNP tree was rooted using *Glaesserella* sp.15-184 as an outgroup. The gene contents of all 94 isolates were compared using CD-HIT (v 4.6.1) software to generate non-paralogous gene clusters (identity ≥ 0.8, $\geq 80\%$ the length of the longest cluster).

Comparison of antimicrobial resistance and virulence genes

A whole-genome BLAST search (E-value $\leq 1e^{-5}$, minimal alignment length percentage $\geq$ 80%) was performed against four databases for pathogenicity and drug resistance analysis: Pathogen Host Interactions (PHI), Virulence Factors of Pathogenic Bacteria (VFDB), Carbohydrate-Active enZYmes Database (CAZy), and Integrated Antibiotic Resistance Genes Database (IARDB).

Features of the novel Tn*6678* transposon in HPS-1

Based on the results of the BLASTn search, genomic characteristics were compared among four isolates that harboured a transposon Tn*6678*-like structure. BLASTn searches were performed to identify genes homologous to *bcr*, encoding the multidrug efflux system BCR/CflA, The homologuous sequences were aligned using MUSCLE algorithm in MEGA 7.0.26 (*Kumar et al.*, *2016*) and manually adjusted, yielding 92 candidate genes. The default parameter for gap opening and gap extension were used. The phylogenetic tree was generated using MEGA 7.0.26 software using the neighbour-joining method (*Kumar et al.*, *2016*) with the Kimura 2-parameter substitution model; branch support was assessed using 1000 bootstrap replicates.

Electrotransformation and plasmid stability test

Plasmid pYL1 harboring two antimicrobial resistance genes, bla_{ROB-1} and *aac(6')-Ie-aph(2'')-Ia*, which confer to β-lactams and aminoglycosides resistance. To determine the contributions of pYL1 to penicillin and aminoglycoside antibiotic resistance, electrotransformation experiments were performed using *Staphylococcus aureus* RN4220 as the recipient as previously described (*Wang et al., 2015*). Transformants were selected on brain-heart infusion (BHI) agar supplemented with kanamycin (25 μg/mL) for colony growth at 37°C for 16 h. Transformation efficiency was calculated based on the ratio of transformants to the total number of viable cells. The presence of the *aac(6')-Ie-aph(2'')-Ia* and bla_{ROB-1} genes in transformants was confirmed by PCR amplification followed by DNA sequence analysis. The primers for bla_{ROB-1} (494 bp) were 5′-CGCTTTGCTTATGCGTCCAC-3′ (forward) and 5′-ACTTTCCACGATGTTGGCGT-3′. The primers for *aac(6')-Ie-aph(2'')-Ia* (412 bp) were 5′-AGAGCCTTGGGAAGATGAAGTT-3′ (forward) and 5′-TGCCTTAACATTTGTGGCATT-3′ (reverse). The primers were designed using NCBI Primer-BLAST. The PCR conditions were as follows: initial denaturation at 95 °C for 5 min, 30 cycles of amplification (30 s at 95 °C, 30 s at 58 °C, and 90 s at 72 °C), followed by extension at 72 °C for 10 min. The PCR products were purified and sequenced by Majorbio Company (Shanghai, China). The MICs of *S. aureus* RN4220 and five transformants were determined by Etest (Liofilchems.r.l.) according to the manufacturer's instructions.

The stability of plasmids carrying *aac(6')-Ie-aph(2'')-Ia* and bla_{ROB-1} was determined by serial passages for 15 consecutive days at 1:1000 dilutions into fresh BHI, with or without antibiotic (kanamycin) pressure. Serially diluted cultures were spread on BHI agar plates with or without kanamycin (8 μg/mL), and the resistance retention rate was determined by randomly picking at least 50 colonies from the BHI plates, spotting them onto new BHI plates with kanamycin (8 μg/mL), and calculating the ratio of resistant to total colonies. Both the resistant and susceptible colonies from the plates were randomly picked and subjected to PCR for detection of bla_{ROB-1} and *aac(6')-Ie-aph(2'')-Ia*.

181 # Results

182 ## *G. parasuis* core and unique genes

183 Compilation of the 94 genomes covering all serovars and disease- and non-disease-causing
184 backgrounds from nine geographic locations (Table S2) demonstrated expansion of the pan-
185 genome, whereas the number of core genes remained relatively stable with the addition of new
186 strains (Fig. 1A). This result suggests the presence of an open pan-genome experiencing frequent
187 evolutionary changes through gene gains and losses or lateral gene transfer. The size of the pan-
188 genome was 5,243 genes, including ~3.34% core genes shared among the 94 isolates mainly
189 from China (Fig. 1B). Meanwhile, accessory genomes occupied a large fraction (85.13–91.74%)
190 of the *G. parasuis* gene content compared with the core genomes and the number of unique
191 genes ranged from 0 to 103 indicating that 0–4.6% of the genome consists of strain-specific
192 accessory genes (Table S4).

193 Clusters of Orthologous Groups classification indicated that core genes were significantly
194 enriched in defense mechanisms and inorganic ion transport and metabolism, whereas unique
195 genes were significantly enriched in unknown function, nucleotide transport and metabolism,
196 and carbohydrate transport and metabolism (Fig. 1C).

197 ## Phylogenetic analysis of *G. parasuis* isolates

198 A phylogenetic tree based on single-copy core genes of our isolates and reference isolates
199 resolved two well-supported lineages, lineages I and II, exhibiting association with country,
200 serotypes, and MLST types (Fig. 2). Lineages I and II comprised eight and two countries,
201 respectively. Serovars 5, 12, and 14 were identified predominantly in lineage I, while serovars 2
202 and 10 were mostly found in lineage II. For serovars 3, 8, 9, and 11, the numbers of isolates were
203 too low to draw conclusions about phylogenetic patterns. The remainder of the serovars were
204 found in both clades.

MLST analysis assigned the 39 isolates in GenBank to 20 different STs, including six new STs, with 13 isolates not determined. The 55 isolates obtained in our study belonged to 49 different STs, including 39 new STs (Table S2). Most strains of the same STs formed single clades (Fig. 2). The SNP-based tree with and without an outgroup (Fig. S1 and Fig. S2) was consistent with the phylogenetic analysis based on single-copy core orthologs. The number of whole-genome SNP differences among the 94 isolates ranged from 8,603 to 8,730.

Biological features of *G. parasuis* isolates

Variation in virulence and stress resistance genes was observed among *G. parasuis* lineages and subgroups (Fig. 3). All 94 *G. parasuis* isolates harboured more than five types of pathogenic factors. The virulence factors *gigP, malQ,* and *gmhA* were carried by all the tested *G. parasuis* isolates. Moreover, other virulence factors including the *rfa* cluster, encoding enzymes for lipopolysaccharide (LPS) core biosynthesis, and *galU* and *galE,* resulting in impaired biofilm formation, were universally present in the *G. parasuis* isolates.

The main ARGs associated with resistance in *G. parasuis*, including the β-lactam-resistant gene $bla_{\text{ROB-1}}$, tetracycline resistance genes *tetB*, aminoglycoside resistance genes *aph(3'')-Ib* and *aac(6')-Ie-aph(2'')-Ia,* fluoroquinolone resistance gene *norA,* chloramphenicol resistance genes *catIII* and *floR*, sulfonamide resistance gene *sul2* were discovered (Fig. 3). Among all of these genes, the genes *sul2* and *aph(3'')-Ib,* and β-lactam-resistant genes *pbp1a* and *pbp3a* were universally present in the *G. parasuis* isolates (Fig. 3). Three different serotype isolates (H82, H92, and H313) obtained from different sites in different years that clustered closely in one branch all harboured the lincosamide antibiotic resistance factor *lunC* (Fig. 3). Moreover, 91.5% of the isolates had *bcr*, 90.42% of the isolates had *bacA*, 100% of the isolates had *ksgA*, but five isolates had *norA.*

Genomic features of *G. parasuis* HPS-1

Following sequencing and assembly, a 2,326,414-bp chromosome with an average G+C content of 40.03%, and a 7,777-bp small plasmid sequence (pYL1) with an average G+C content

of 33.32% were identified in strain HPS-1 (Supplementary Fig. S3 and Fig. 4). HPS-1 exhibited a novel ST (ST328) with undescribed MLST alleles or previously unreported allelic combinations. This ST328 genome harbored resistance genes against several types of antibiotics, including sulfonamides (*sul2*), aminoglycosides (*aph(3'')-Ib*, *aac(6')-Ie-aph(2'')-Ia*), and β-lactam (bla_{ROB-1}) (Table S1). Further, this genome contained efflux pump-related genes that confer resistance to sulfonamides (*bcr*) and multidrug resistance (*acrB*).

We also identified a novel transposon in the ST328 isolate, designated Tn*6678* in the Tn Number Registry (https://transposon.lstmed.ac.uk/). This transposon harbours two 966-bp IS110 family transposases at both ends, two toxin genes *pilT* and *phd*, two genes associated with the two-component signal transduction system *cpxA* and *cpxR*, one efflux pump-associated gene *bcr*, and four genes encoding hypothetical proteins with unknown function (Fig. 5). Genome analysis revealed that Tn*6678* was inserted between the molybdopterin molybdotransferase MoeA encoded by *moeA* and 3-isopropylmalate dehydratase large subunit encoded by *leuC*. A LacI family transcriptional regulator and a bifunctional tRNA (5-methylaminomethyl-2-thiouridine)(34)-methyltransferase MnmD/FAD-dependent 5-carboxymethylaminomethyl-2-thiouridine (34) oxidoreductase MnmC flanked the transposon to the right and left, respectively.

Through BLASTN searches, highly conserved homologous sequences to Tn*6678* (>97% nucleotide sequence similarity) were identified in four *G. parasuis* strains [29755 (GenBank accession number CP021644, USA), SH0165 (CP001321, China), ZJ0906 (CP005384, China), and str. Nagasaki (NZ_APBT00000000, Japan)]. The only differences in these five chromosomes were in the transposases, but transposon Tn*6678* had two complete inverted repeats of IS110 transposases flanked by 32-bp inverted repeats of ISNme5 at both ends (Fig. 5), suggesting mobility potential. The *bcr*-containing Tn*6678* also contained an antibiotic resistance gene cassette, suggesting its potential to transfer antibiotic resistance genes.

BLASTn searches for the *bcr* gene returned a large set of divergently related sequences using default parameters. These sequences were annotated as bicyclomycin/multidrug efflux system, Bcr/CflA family drug resistance efflux transporter, Bcr/CflA family multidrug efflux

major facilitator superfamily (MFS) transporter or drug resistance transporter, and Bcr/CflA subfamily. Phylograms revealed that the *bcr* gene in HPS-1 was most closely related to homologs identified in other members of the Pasteurellaceae, particularly *G. parasuis*, *Actinobacillus indolicus*, *Bibersteinia trehalosi, Actinobacillus* (*A. pleuropneumoniae, A. suis, A. equuli, A. lignieresii, A. indolicus*, and *A. porcitonsillarum*), and *Mannheimia* (*M. haemolytica* and *M. varigena*), all of which are known causative agents of upper respiratory tract infections (Fig. 6).

The neighbour-joining phylogenetic tree using 92 *bcr* genes selected from the BLASTn searches clearly demonstrated two distinctive clades. The first clade contained *bcr* genes of *Hemophilus influenzae,* which colonizes humans, and other *Haemophilus* species that colonize non-human animals. Members of the second clade were divided into four apparent subclades, including *G. parasuis*, *B. trehalosi*, *Actinobacillus* spp., and *Mannheimia* spp. Except for *G. parasuis*, the chromosomally encoded Bcr/CflA from *G. parasuis* HPS-1 most closely clustered with that found in *A. indolicus*. The phylogenetic tree indicated a divergent evolutionary pattern between animal-origin *Pasteurellaceae* bacteria. The *bcr* gene tree is consistent with the organismal phylogeny, suggesting that horizontal gene transfer does not play an important role in the evolution of *bcr*-mediated resistance.

General features and electrotransformation of the plasmid pYL1

The plasmid pYL1 identified in HPS-1 contained seven ORFs with an average length of 912 bp, with one encoded protein of undetermined function (Fig. 4), and two antimicrobial resistance genes, $bla_{\text{ROB-1}}$ and *aac(6')-Ie-aph(2'')-Ia*. Four ORFs were identified to encode a 3′-truncated transposase protein ISApl1 (30 amino acids), a Rep-like protein (444 amino acids) involved in plasmid replication, and two Mob proteins, MobC (144 amino acids) and MobA (541 amino acids), associated with plasmid mobilization (Fig. 4). Except for resistance genes, pYL1 had the same backbone and genetic structure and showed 100% nucleotide identity to four previously-identified plasmids, pFZ51, pFS39, pHN61, and pHB0503 (Table S5) (*Kang et al.*, *2009*; *Chen*

et al., *2010*; *Yang et al.*, *2013*). In contrast, the resistance genes and flanking regions in pYL1 exhibited as little as 58% sequence identity to the other four plasmids (Fig. 7).

Transformation of pYL1 into *S. aureus* RN4220 was achieved at a frequency of 10^{-9} cells per recipient cell by electroporation, confirming that pYL1 is a mobilizable plasmid with active mobilization genes. The transformants had increased MICs for oxacillin, gentamicin, amikacin, kanamycin, and streptomycin as compared with those of the parental strain (0.047 to > 256 mg/L, 0.094 to 1.5 mg/L, 0.38 to 16 mg/L, 0.38 to 32 mg/L, and < 0.25 to 32 mg/L, respectively). This finding indicated that plasmid pYL1 carrying $bla_{\text{ROB-1}}$ and *aac(6')-Ie-aph(2'')-Ia* contributed to the penicillin resistance and aminoglycoside antibiotic resistance in *S. aureus* RN4220 transformants. Furthermore, the plasmid showed low stability in *S. aureus* without antibiotic pressure, as only 52.5%, 30.48%, and 2.68% of transformants maintained the kanamycin resistance after five, six, and seven subcultures, respectively. However, the plasmid can be conserved in *S. aureus* cultured with kanamycin, as 100% of the colonies remained resistant to kanamycin after 10 subcultures, as confirmed by PCR mapping.

Discussion

In the current study, we observed an open pan-genome. Similar result that the size of pan-genome was 7,431 genes including 1,049 core genes has been reported (*Howell et al.*, *2014*). This suggested that the *G. parasuis* pan-genome is vast, and unique genes can be continuously be identified upon sequencing more *G. parasuis* genomes. However, the isolates in this study with ~3.34% core genes, primarily isolated from China, displayed further diversity and higher variability than isolates with only ~14% core genes, primarily obtained from the UK (*Howell et al.*, *2014*). Besides, we identified 54 new STs enriching the *G. parasuis* MLST databases and highlight the wide distribution of *G. parasuis* strains. Although most strains of the same STs formed single clades, there was no definitive association between ST and serotype (Fig. 2), consistent with previous studies (*Olvera et al.*, *2006*; *Wang et al.*, *2016*).

The pattern of the phylogenetic tree based on single-copy core genes was different from the population grouping predicted via MLST, which showed six main subgroups (*Wang et al.*, *2016*). Both phylogenetic lineages contain both Asian and North American isolates, in agreement with previous phylogenetic analyses (*Howell et al.*, *2014*; *Wang et al.*, *2016; Dickerman et al.,2020*) and supporting the hypothesis of frequent migration of isolates between geographic regions.

Five types of pathogenic factors *gigP, malQ, gmhA, rfa* and *gal* cluster were universally carried by *G. parasuis* isolates in this study. The *rfaF* gene has been linked to serum resistance, adhesion, and invasion (*Zhang et al.*, *2013*); *galU* plays a role in autoagglutination and biofilm formation, and *galE* appears to affect biofilm production indirectly in *G. parasuis* (*Zou et al.*, *2013*). Serum resistance may play a role in the virulence of *G. parasuis* (*Cerda-Cuellar and Aragon*, *2008*). However, *lsgB*, previously associated with *G. parasuis* virulence potential, was predominant in six isolates (29755 and HPS9 from the USA, Nagasaki from Japan, and KL0318, SH0104, and SH0165 from China), in line with potentially virulent strains isolated from the nasal cavities of healthy pigs (*Amano et al.*, *1996*; *Brockmeier et al.*, *2013*).

The *bla*$_{\text{ROB-1}}$, *sul2*, *aph(3'')-Ib, tetB, tetD, aac(6')-Ie-aph(2'')-Ia, catIII,* and *floR* genes have previously been identified in *G. parasuis* (*Zhao et al.*, *2018*). In the current study, we identified all of genes mentioned above. This is the first report of genes *tetA*, *tetH* and *tetR* genes in *G. parasuis* isolates and needs further study. Tetracycline resistance genes are often associated with conjugative and mobile genetic elements enabling horizontal transfer (*Lancashire et al., 2005*; *Zhao et al.*, *2018*). Moreover, this is the first report describing the presence of the *bcr*, *bacA*, *ksgA* and *norA* genes in *G. parasuis*, to the best of our knowledge. All of these benefits from the application of whole genome sequencing method. Three isolates clustered closely in one branch all harboured *lunC* gene, contained in the ISSag10 sequence of all three isolates. The *lunC* gene was only identified in plasmid pHN61 of *G. parasuis* (*Chen et al.*, *2010*). The results suggested that the resistance of these three strains to lincomycin may be mediated by the plasmid carrying *lunC* gene.

This is also the first report describing the transoson Tn*6678* containing toxin genes *pilT* and *phd*, drug resistance genes *cpxA* and *cpxR*, and an efflux pump gene *bcr*. Association between the Cpx system and bacterial antimicrobial resistance has been reported in *Escherichia coli*, *Salmonella enterica, Klebsiella pneumoniae,* and *G. parasuis* (*Hu et al.*, *2011*; *Srinivasan et al.*, *2012*; *Audrain et al.*, *2013*; *Kurabayashi et al.*, *2014*; *Cao et al.*, *2018*). CpxR plays essential roles in mediating macrolide (i.e., erythromycin) resistance (*Cao et al.*, *2018*). The Bcr/CflA efflux system was identified as a group of antiporters that confer resistance to chloramphenicol, florfenicol, and bicyclomycin by actively transporting these compounds out of the cell (*Marklevitz and Harris*, *2016*). The transposon Tn*6678* had two complete inverted repeats of IS110 transposases flanked by 32-bp inverted repeats of ISNme5 at both ends suggesting mobility potential and its potential to transfer antibiotic resistance genes. In *G. parasuis*, only the efflux pump AcrB, belonging to the resistance-nodulation division (RND) family, has been analysed to date. Efflux pump AcrB may play a role in multidrug resistance, and the *acrAB* gene cluster could affect the efflux of macrolides in *G. parasuis* (*Feng et al.*, *2014*). However, this is the first description of the efflux pump Bcr/CflA in *G. parasuis*, belonging to the MFS. This efflux pump, encoded by *bcr*, harbored on a transposon indicated its potential transferability.

To date, two β-lactam resistance genes ($bla_{\text{ROB-1}}$ and bla_{TEM}) have been reported in *G. parasuis* (*By Guo et al.*, *2012*). A β-lactam resistance plasmid, pB1000, harbouring $bla_{\text{ROB-1}}$ was previously detected in *G. parasuis* clinical strains isolated from Glässer's disease lesions (*San et al.*, *2007*). The plasmid pYL1 harboured two antimicrobial resistance genes, $bla_{\text{ROB-1}}$ and *aac(6')-Ie-aph(2'')-Ia*. The ROB-1 of plasmid pYL1 had a typical size of 305 bp, in line with functionally active members of the ROB-1 family from different plasmids in *Pasteurellaceae* species. AAC(6')-Ie-APH(2')-Ia, the most important aminoglycoside-resistance enzyme in gram-positive bacteria conferring resistance to almost all known aminoglycoside antibiotics in clinical use, also had a typical size of 479 amino acids in this family (*Rouch et al.*, *1987*). Although *aac(6')-Ib-cr* is considered the most prominent aminoglycoside-resistance gene in *G. parasuis* (*Doi and Arakawa*, *2007*; *San et al.*, *2007*), the bifunctional aminoglycoside-resistance enzyme

AAC(6')-Ie-APH(2')-Ia in plasmids is also reported in GenBank for *G. parasuis* strains. Comparing with other four previously-identified plasmids which have similar structure with pYL1 suggested more rapid evolution among the resistance-associated components of these small plasmids. The transposase gene of ISApl1 in pYL1 had an internal deletion of 659 bp, but intact 3′ and 5′ ends. The truncated ISApl1 linked with bla_{ROB-1} suggested that ISApl1 played a key role in transposition of bla_{ROB-1}, facilitating the horizontal transfer of β-lactam and aminoglycoside resistance among *G. parasuis* isolates. These results are consistent with a previous study presenting evidence for spread of β-lactam resistance (*Yang et al.*, *2013*). A similar occurrence was also identified in *A. porcitonsillarum* or *G. parasuis* plasmids pFJS5863, pQY431, and pFS39, suggesting a more widespread role and highlighting that the function of ISApl1 requires further investigation.

Conclusions

In summary, our results shed new light on the importance of genomic variations, especially transposon-related and/or plasmid-related gene variations, in the evolution of *G. parasuis*. This comparative analysis identified potentially novel virulence factors (*gigP*, *malQ*, and *gmhA*) and drug resistance genes (*norA*, *bacA*, *ksgA*, and *bcr*) in *G. parasuis*. Resistance determinants (*sul2*, *aph(3'')-Ib*, *norA*, *bacA*, *ksgA*, and *bcr*) were widespread across isolates, regardless of serovar, isolation source, or geographical location. Future research focused on a larger sample of *G. parasuis* isolates worldwide will further increase understanding of the rapid development of antibiotic resistance associated with mobile genetic elements in this important animal pathogen.

Acknowledgements

We thank members of our laboratories for fruitful discussions.

Funding

This work was supported by the National Key Basic Research Program (Grant No. 2016YFD0500606), Guangdong Provincial Key Research and Development Plan Project (Grant No. 2019B020212002), and the Central Universities constructs the world fist-class university (discipline) and Characteristic Development Guidance Special Fund (Grant No. K5174960).

References

Altschul, SF, Gish, W, Miller, W, Myers, EW and Lipman, DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215: 403-410 DOI: 10.1016/S0022-2836(05)80360-2.

Amano, H, Shibata, M, Kajio, N and Morozumi, T. 1996. Pathogenicity of *Haemophilus parasuis* serovars 4 and 5 in contact-exposed pigs. *Journal of Veterinary Medical Science*, 58: 559-561 DOI: 10.1292/jvms.58.559.

Audrain, B, Ferrieres, L, Zairi, A, Soubigou, G, Dobson, C, Coppee, JY, Beloin, C and Ghigo, JM. 2013. Induction of the Cpx envelope stress pathway contributes to *Escherichia coli* tolerance to antimicrobial peptides. *Appl Environ Microbiol*, 79: 7770-7779 DOI: 10.1128/AEM.02593-13.

Bello-Orti, B, Aragon, V, Pina-Pedrero, S and Bensaid, A. 2014. Genome comparison of three serovar 5 pathogenic strains of *Haemophilus parasuis*: insights into an evolving swine pathogen. *Microbiology*, 160: 1974-1984 DOI: 10.1099/mic.0.079483-0.

Besemer, J, Lomsadze, A and Borodovsky, M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Research*, 29: 2607-2618 DOI: 10.1093/nar/29.12.2607.

Brockmeier, SL. 2004. Prior infection with *Bordetella bronchiseptica* increases nasal colonization by *Haemophilus parasuis* in swine. *Veterinary Microbiology*, 99: 75-78 DOI: 10.1016/j.vetmic.2003.08.013.

Brockmeier, SL, Loving, CL, Mullins, MA, Register, KB, Nicholson, TL, Wiseman, BS, Baker, RB and Kehrli, MJ. 2013. Virulence, transmission, and heterologous protection of four isolates

410 of *Haemophilus parasuis*. *Clinical and Vaccine Immunology*, 20: 1466-1472 DOI:
411 10.1128/CVI.00168-13.

412 Brockmeier, SL, Register, KB, Kuehn, JS, Nicholson, TL, Loving, CL, Bayles, DO, Shore, SM
413 and Phillips, GJ. 2014. Virulence and draft genome sequence overview of multiple strains of the
414 swine pathogen *Haemophilus parasuis*. *PLoSOne*, 9: e103787 DOI:
415 10.1371/journal.pone.0103787.

416 Guo, LGL, Zhang, JZJ, Xu, CXC, Ren, TRT, Zhang, BZB, Chen, JCJ and Liao, MLM. 2012.
417 Detection and characterization of β-lactam resistance in *Haemophilus parasuis* strains from pigs
418 in South China. *Journal of Integrative Agriculture*, 11: 116-121 DOI.

419 Cao, Q, Feng, F, Wang, H, Xu, X, Chen, H, Cai, X and Wang, X. 2018. *Haemophilus parasuis*
420 CpxRA two-component system confers bacterial tolerance to environmental stresses and
421 macrolide resistance. *Microbiological Research*, 206: 177-185 DOI:
422 10.1016/j.micres.2017.10.010.

423 Cerda-Cuellar, M and Aragon, V. 2008. Serum-resistance in *Haemophilus parasuis* is associated
424 with systemic disease in swine. *Veterinary Journal*, 175: 384-389 DOI:
425 10.1016/j.tvjl.2007.01.016.

426 Chen, LCL, Cai, XCX, Wang, XWX, Zhou, XZX, Wu, DWD, Xu, XXX and Chen, HCH. 2010.
427 Characterization of plasmid-mediated lincosamide resistance in a field isolate of *Haemophilus*
428 *parasuis*. *Journal of Antimicrobial Chemotherapy*, 65: 2256-2258 DOI.

429 Costa-Hurtado, M and Aragon, V. 2013. Advances in the quest for virulence factors of
430 *Haemophilus parasuis*. *Veterinary Journal*, 198: 571-576 DOI: 10.1016/j.tvjl.2013.08.027.

431 Deng, Y, Xu, H, Su, Y, Liu, S, Xu, L, Guo, Z, Wu, J, Cheng, C, and Feng, J. 2019. Horizontal
432 gene transfer contributes to virulence and antibiotic resistance of *Vibrio harveyi* 345 based on
433 complete genome sequence analysis. *BMC Genomics*, 20: 761 DOI: 10.1186/s12864-019-6137-8.

434 de la Fuente, AJ, Tucker, AW, Navas, J, Blanco, M, Morris, SJ and Gutierrez-Martin, CB. 2007.

435 Antimicrobial susceptibility patterns of *Haemophilus parasuis* from pigs in the United Kingdom
436 and Spain. *Veterinary Microbiology*, 120: 184-191 DOI: 10.1016/j.vetmic.2006.10.014.

437 Dickerman, A, Bandara, AB, and Inzana, TJ. 2020. Phylogenomic analysis of *Haemophilus*
438 *parasuis* and proposed reclassification to *Glaesserella parasuis*, gen. nov., comb. nov.
439 *International Journal of Systematic and Evolutionary Microbiology*, 70: 180-186. DOI:
440 10.1099/ijsem.0.003730.

441 Doi, Y and Arakawa, Y. 2007. 16S ribosomal RNA methylation: emerging resistance mechanism
442 against aminoglycosides. *Clinical Infectious Diseases*, 45: 88-94 DOI: 10.1086/518605.

443 Feng, S, Xu, L, Xu, C, Fan, H, Liao, M and Ren, T. 2014. Role of *acrAB* in antibiotic resistance
444 of *Haemophilus parasuis* serovar 4. *Veterinary Journal*, 202: 191-194 DOI:
445 10.1016/j.tvjl.2014.05.045.

446 Geisinger, E, and Isberg, RR. 2017. Interplay between antibiotic resistance and virulence during
447 disease promoted by multidrug-resistant bacteria. *Journal of Infectious Diseases*, 215: S9-S17
448 DOI: 10.1093/infdis/jiw402.

449 Guindon, S, Dufayard, JF, Lefort, V, Anisimova, M, Hordijk, W and Gascuel, O. 2010. New
450 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
451 performance of PhyML 3.0. *Systematic Biology*, 59: 307-321 DOI: 10.1093/sysbio/syq010.

452 Howell, KJ, Weinert, LA, Chaudhuri, RR, Luan, SL, Peters, SE, Corander, J, Harris, D, Angen,
453 O, Aragon, V, Bensaid, A, Williamson, SM, Parkhill, J, Langford, PR, Rycroft, AN, Wren, BW,
454 Holden, MT, Tucker, AW and Maskell, DJ. 2014. The use of genome wide association methods
455 to investigate pathogenicity, population structure and serovar in *Haemophilus parasuis*. *BMC*
456 *Genomics*, 15: 1179 DOI: 10.1186/1471-2164-15-1179.

457 Hu, WS, Chen, HW, Zhang, RY, Huang, CY and Shen, CF. 2011. The expression levels of outer
458 membrane proteins STM1530 and OmpD, which are influenced by the CpxAR and BaeSR two-
459 component systems, play important roles in the ceftriaxone resistance of *Salmonella enterica*

460 serovar Typhimurium. *Antimicrob Agents Chemother*, 55: 3829-3837 DOI:
461 10.1128/AAC.00216-11.

462 Jia, AJA, Zhou, RZR, Fan, HFH, Yang, KYK, Zhang, JZJ, Xu, YXY, Wang, GWG and Liao,
463 MLM. 2017. Development of serotype-specific PCR assays for typing of *Haemophilus parasuis*
464 isolates circulating in southern China. *Journal of Clinical Microbiology*, 55: 3249-3257 DOI:
465 10.1128/JCM.00688-17.

466 Kang, M, Zhou, R, Liu, L, Langford, PR and Chen, H. 2009. Analysis of an *Actinobacillus*
467 *pleuropneumoniae* multi-resistance plasmid, pHB0503. *Plasmid*, 61: 135-139 DOI:
468 10.1016/j.plasmid.2008.11.001.

469 Kumar, S, Stecher, G and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis
470 version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33: 1870-1874 DOI:
471 10.1093/molbev/msw054.

472 Kurabayashi, K, Hirakawa, Y, Tanimoto, K, Tomita, H and Hirakawa, H. 2014. Role of the
473 CpxAR two-component signal transduction system in control of fosfomycin resistance and
474 carbon substrate uptake. *Journal of Bacteriology*, 196: 248-256 DOI: 10.1128/JB.01151-13.

475 Lancashire JF, Terry TD, Blackall PJ, Jennings MP. 2005. Plasmid-encoded TetB tetracycline
476 resistance in *Haemophilus parasuis*. *Antimicrob Agents Chemother*, 49: 1927-1931 DOI:
477 10.1128/AAC.49.5.1927-1931.2005

478 Li, Y, Kwok, AH, Jiang, J, Zou, Y, Zheng, F, Chen, P, Hou, C, Leung, FC and Jiang, P. 2013.
479 Complete genome analysis of a *Haemophilus parasuis* serovar 12 strain from China. *PLoS One*,
480 8: e68350 DOI: 10.1371/journal.pone.0068350.

481 Marklevitz, J and Harris, LK. 2016. Prediction driven functional annotation of hypothetical
482 proteins in the major facilitator superfamily of *S. aureus* NCTC 8325. *Bioinformation*, 12: 254-
483 262 DOI: 10.6026/97320630012254.

484 Martinez-Moliner, V, Soler-Llorens, P, Moleres, J, Garmendia, J and Aragon, V. 2012.

485 Distribution of genes involved in sialic acid utilization in strains of *Haemophilus parasuis*.
486 *Microbiology*, 158: 2117-2124 DOI: 10.1099/mic.0.056994-0.

487 Nicholson, TL, Brunelle, BW, Bayles, DO, Alt, DP and Shore, SM. 2018. Comparative genomic
488 and methylome analysis of non-virulent D74 and virulent Nagasaki *Haemophilus parasuis*
489 isolates. *PLoSOne*, 13: e205700 DOI: 10.1371/journal.pone.0205700.

490 Oliveira, S, Galina, L and Pijoan, C. 2001. Development of a PCR test to diagnose *Haemophilus*
491 *parasuis* infections. *Journal of Veterinary Diagnostic Investigation*, 13: 495-501 DOI:
492 10.1177/104063870101300607.

493 Olvera, A, Cerda-Cuellar, M and Aragon, V. 2006. Study of the population structure of
494 *Haemophilus parasuis* by multilocus sequence typing. *Microbiology*, 152: 3683-3690 DOI:
495 10.1099/mic.0.29254-0.

496 Pruller, S, Turni, C, Blackall, PJ, Beyerbach, M, Klein, G, Kreienbrock, L, Strutzberg-Minder, K,
497 Kaspar, H, Meemken, D and Kehrenberg, C. 2017. Towards a standardized method for broth
498 microdilution susceptibility testing of *Haemophilus parasuis*. *Journal of Clinical Microbiology*,
499 55: 264-273 DOI: 10.1128/JCM.01403-16.

500 Rouch, DA, Byrne, ME, Kong, YC and Skurray, RA. 1987. The *aacA-aphD* gentamicin and
501 kanamycin resistance determinant of Tn4001 from *Staphylococcus aureus*: expression and
502 nucleotide sequence analysis. *J Gen Microbiol*, 133: 3039-3052 DOI: 10.1099/00221287-133-
503 11-3039.

504 San, MA, Escudero, JA, Catalan, A, Nieto, S, Farelo, F, Gibert, M, Moreno, MA, Dominguez, L
505 and Gonzalez-Zorn, B. 2007. Beta-lactam resistance in *Haemophilus parasuis* is mediated by
506 plasmid pB1000 bearing $bla_{\text{ROB-1}}$. *Antimicrob Agents Chemother*, 51: 2260-2264 DOI:
507 10.1128/AAC.00242-07.

508 Soge, OO, No, D, Michael, KE, Dankoff, J, Lane, J, Vogel, K, Smedley, J and Roberts, MC.
509 2016. Transmission of MDR MRSA between primates, their environment and personnel at a

510 United States primate centre. *J Antimicrob Chemother*, 71: 2798-2803 DOI: 10.1093/jac/dkw236.

511 Srinivasan, VB, Vaidyanathan, V, Mondal, A and Rajamohan, G. 2012. Role of the two
512 component signal transduction system CpxAR in conferring cefepime and chloramphenicol
513 resistance in *Klebsiella pneumoniae* NTUH-K2044. *PLoSOne*, 7: e33777 DOI:
514 10.1371/journal.pone.0033777.

515 Wang, L, Ma, L, Liu, Y, Gao, P, Li, Y, Li, X and Liu, Y. 2016. Multilocus sequence typing and
516 virulence analysis of *Haemophilus parasuis* strains isolated in five provinces of China. *Infection*
517 *Genetics and Evolution*, 44: 228-233 DOI: 10.1016/j.meegid.2016.07.015.

518 Wang, XM, Li, XS, Wang, YB, Wei, FS, Zhang, SM, Shang, YH and Du XD. 2015.
519 Characterization of a multidrug resistance plasmid from *Enterococcus faecium* that harbours a
520 mobilized *bcrABDR* locus. *J Antimicrob Chemother*, 70: 609-611 DOI: 10.1093/jac/dku416.

521 Wissing, A, Nicolet, J and Boerlin, P. 2001. The current antimicrobial resistance situation in
522 Swiss veterinary medicine. *Schweiz Arch Tierheilkd*, 143: 503-510.

523 Yang, SS, Sun, J, Liao, XP, Liu, BT, Li, LL, Li, L, Fang, LX, Huang, T and Liu, YH. 2013. Co-
524 location of the *erm*(T) gene and $bla_{\text{ROB-1}}$ gene on a small plasmid in *Haemophilus parasuis* of pig
525 origin. *J Antimicrob Chemother*, 68: 1930-1932 DOI: 10.1093/jac/dkt112.

526 Zhang, B, Xu, C, Zhang, L, Zhou, S, Feng, S, He, Y and Liao, M. 2013. Enhanced adherence to
527 and invasion of PUVEC and PK-15 cells due to the overexpression of RfaD, ThyA and Mip in
528 the DeltaompP2 mutant of *Haemophilus parasuis* SC096 strain. *Veterinary Microbiology*, 162:
529 713-723 DOI: 10.1016/j.vetmic.2012.09.021.

530 Zhao, Y, Guo, L, Li, J, Huang, X and Fang, B. 2018. Characterization of antimicrobial resistance
531 genes in *Haemophilus parasuis* isolated from pigs in China. *PeerJ*, 6: e4613 DOI:
532 10.7717/peerj.4613.

533 Zheng, B, Yu, X, Xu, H, Guo, L, Zhang, J, Huang, C, Shen, P, Jiang, X, Xiao, Y and Li, L. 2017.
534 Complete genome sequencing and genomic characterization of two *Escherichia coli* strains co-

producing MCR-1 and NDM-1 from bloodstream infection. *Sci Rep*, 7: 17885 DOI: 10.1038/s41598-017-18273-2.

Zhou, LJ, Ying, GG, Liu, S, Zhang, RQ, Lai, HJ, Chen, ZF and Pan, CG. 2013. Excretion masses and environmental occurrence of antibiotics in typical swine and dairy cattle farms in China. *Science of the Total Environment*, 444: 183-195 DOI: 10.1016/j.scitotenv.2012.11.087.

Zou, Y, Feng, S, Xu, C, Zhang, B, Zhou, S, Zhang, L, He, X, Li, J, Yang, Z and Liao, M. 2013. The role of *galU* and *galE* of *Haemophilus parasuis* SC096 in serum resistance and biofilm formation. *Veterinary Microbiology*, 162: 278-284 DOI: 10.1016/j.vetmic.2012.08.006.

Figure 1

Analysis of the core and pan-genome of *G. parasuis* isolates.

(A) Core and pan-genomic calculations in *G. parasuis* isolates. Each green point represents the number of genes conserved between genomes. All of the points are plotted as a function of the strain number(x). The deduced pan-genome size: $P(x)= 2483.54x^{0.18}-461.72$. The height of the curve continues to increase because the pan-genome of *G. parasuis* is open. **(B) Genes missing or present in *G. parasuis* isolates.** The heat map illustrates the distribution of core and accessory genes across the *G. parasuis* strains. The columns represent *G. parasuis* isolates. The rows represent genes. The red and black regions represent the presence or absence of genes in a particular genome, respectively. The black regions indicate features missing in that strain but present in one or more of the other *G. parasuis* strains. **(C) The distribution of all, core, and specific genes according to the COG classification.** The y-axis indicates the percentage of genes in various COG categories. A: RNAprocessing and modification. C: Energy production and conversion. D: Cell cycle control, cell division, chromosome partitioning. E:Amino acid transport and metabolism. F: Nucleotide transport and metabolism. G: Carbohydrate transport and metabolism. H: Coenzyme transport and metabolism. I: Lipid transport and metabolism. J: Translation, ribosomal structure and biogenesis. K: Transcription. L: Replication, recombination and repair. M: Cell wall/membrane/envelope biogenesis. N: Cell motility. O: Posttranslational modification, protein turnover, chaperones. P: Inorganic ion transport and metabolism. Q: Secondary metabolites biosynthesis, transport and catabolism. R: General function prediction only. S: Function unknown. T: Signal transduction mechanisms. U: Intracellular trafficking, secretion, and vesicular transport. V: Defense mechanisms.

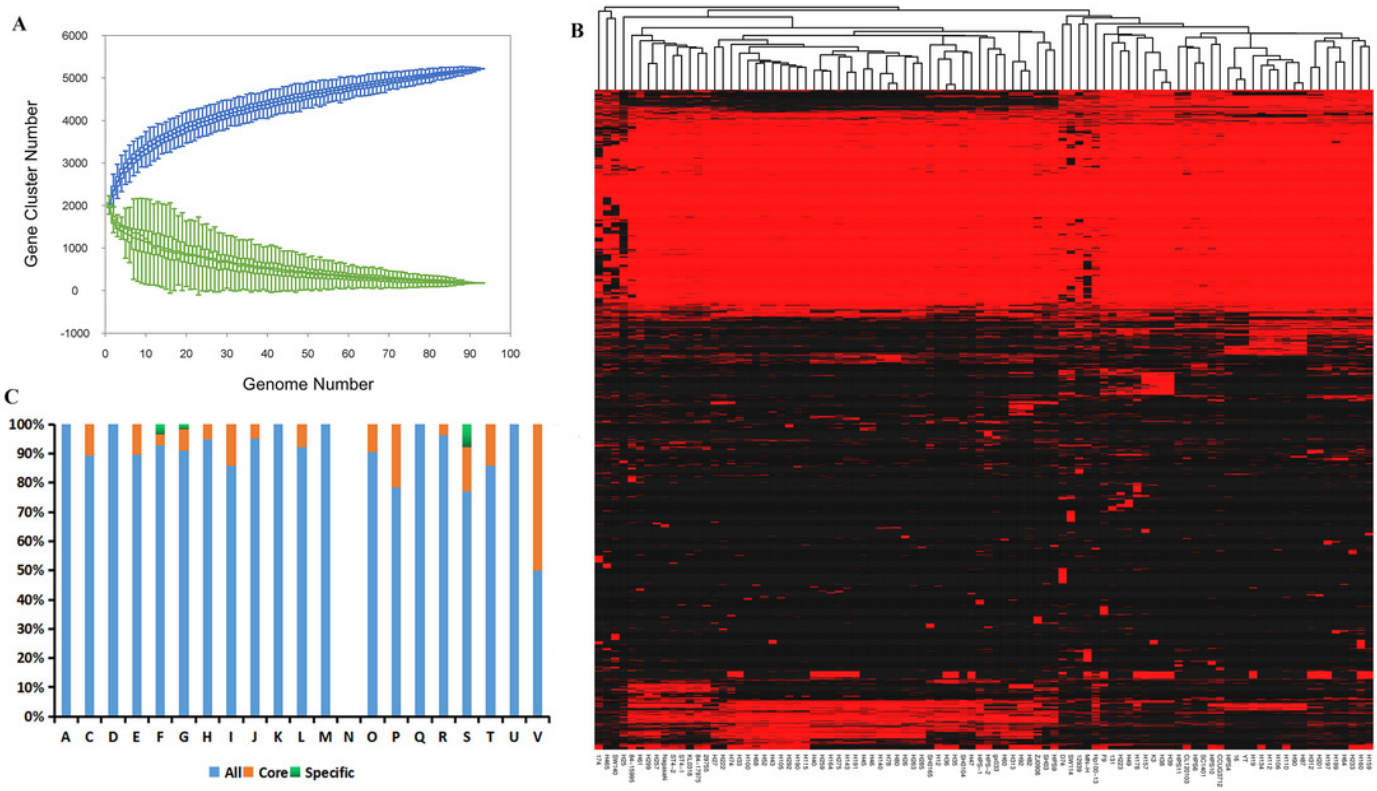
A
Gene Cluster Number
6000
5000
4000
3000
2000
1000
0
-1000
0 10 20 30 40 50 60 70 80 90 100
Genome Number
B
C
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
A C D E F G H I J K L M N O P Q R S T U V
All Core Specific

Figure 2

Maximum-likelihood phylogeny of 94 *Glaesserella parasuis* isolates based on 93 single-copy core genes.

The tree was constructed with MEGA 7 with 1,000 bootstrap replicates. The annotation rings surrounding the tree, from outside to inside, depict (1) geographic region, (2) year of sample collection, (3) site of sample, and (4) serotype. The different colors of the branches represent lineages, lineage Ⅰ in pink and lineage Ⅱ in green.

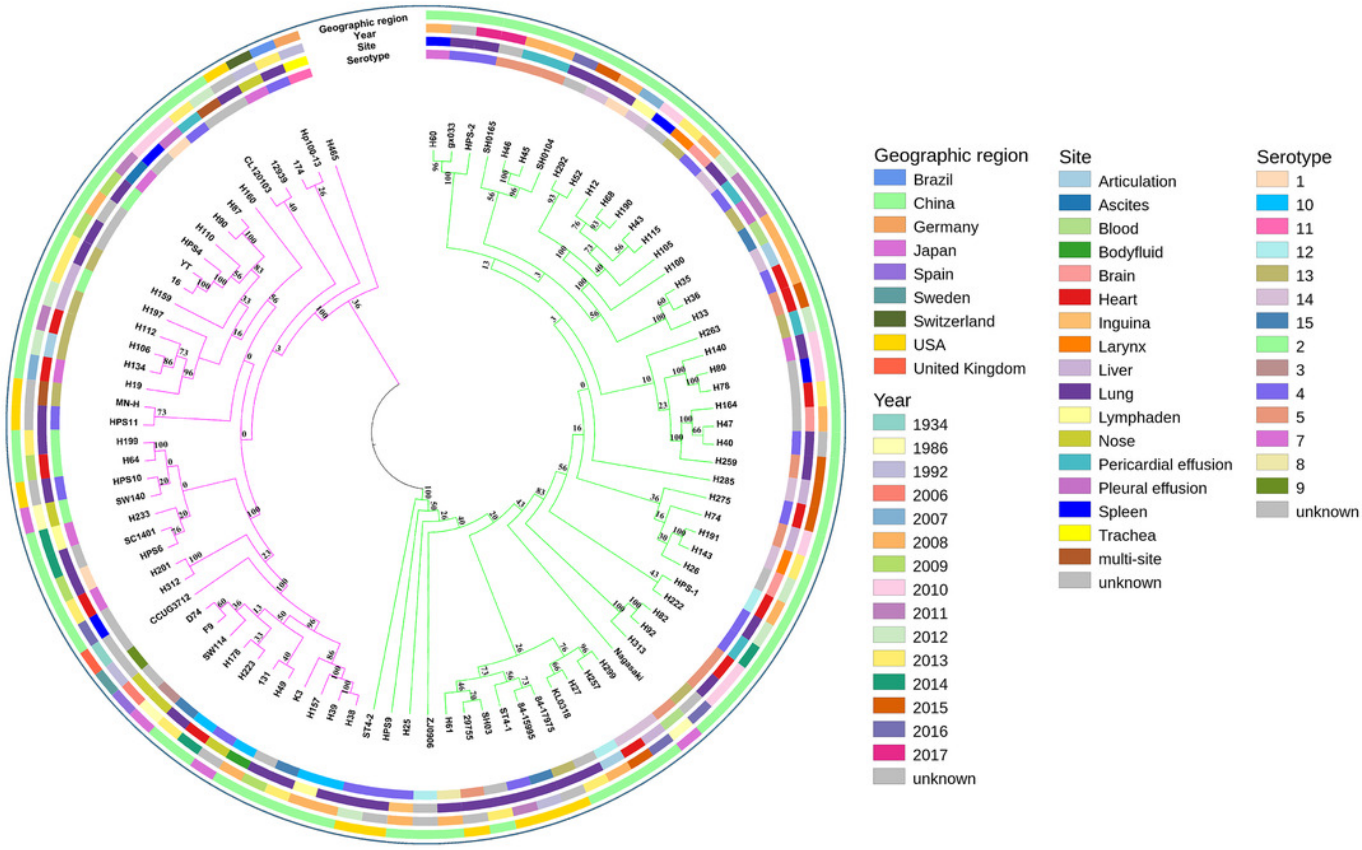

Figure 3

Virulence and resistance profiles across the phylogeny of the 94 *G. parasuis* isolates.

Cluster analysis based on single-copy core orthologs. Pattern of gene presence (colored line) or absence (white).

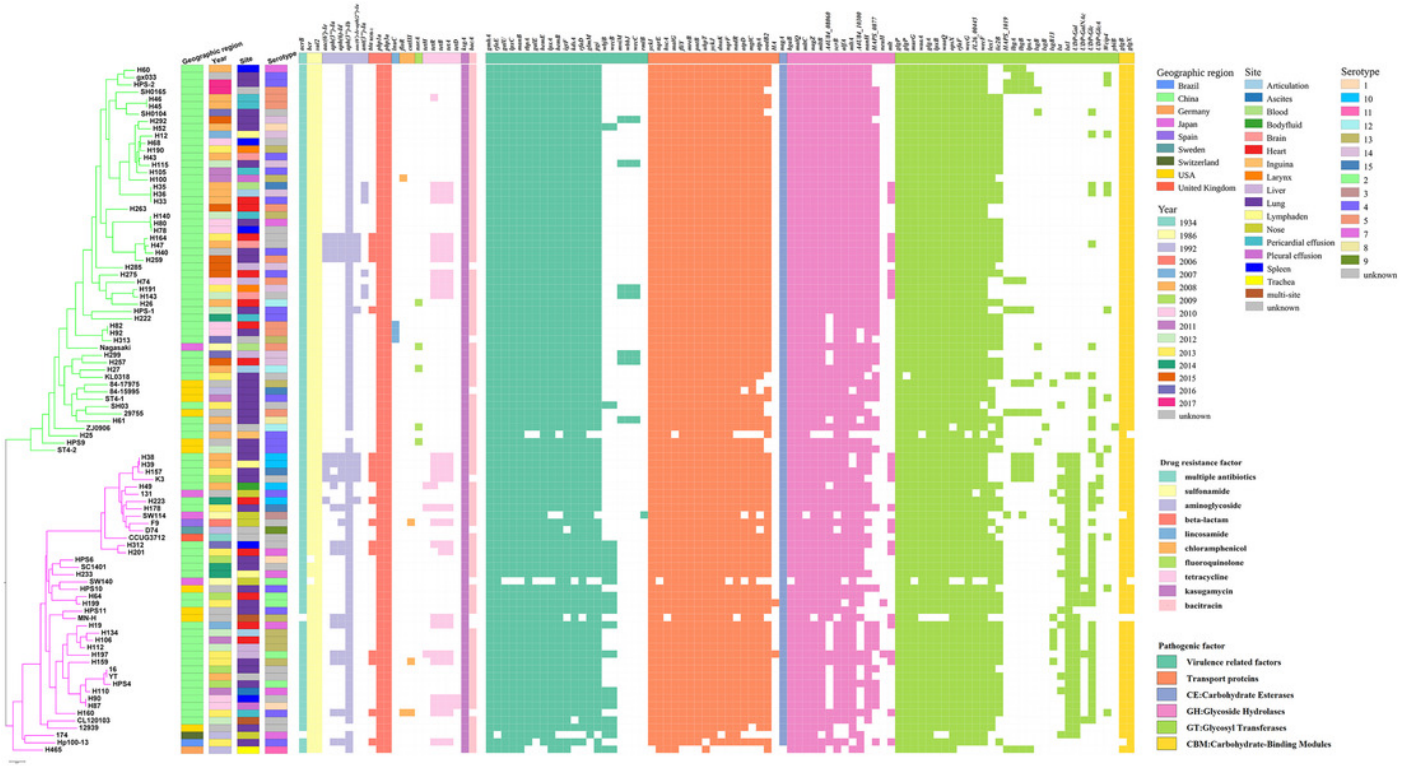

Figure 4

Schematic map of plasmid pYL1.

The circles show, from outside to inside: first and second, putative open reading frames, the positions and orientations of the genes; third, G+C content (deviation from the average); and, fourth, G+C skew (green, +; purple, −).

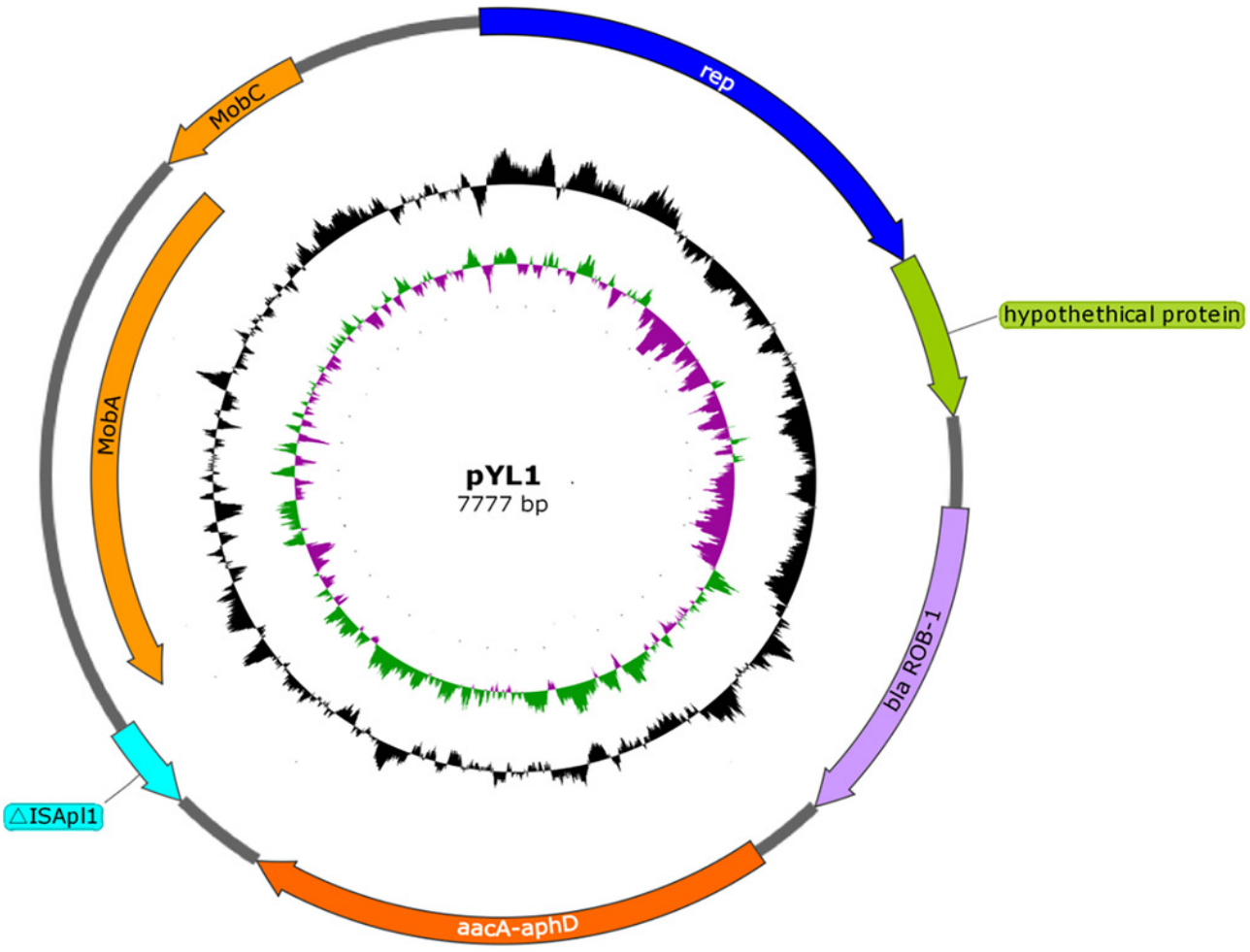

Figure 5

Organization of the *G. parasuis* HPS-1 Tn*6678* transposon and comparison with the similar structure.

ORFs are shown as arrows, indicating the transcription direction, and the colors of the arrows represent different fragments. Gene color code: transposase, purple; toxin genes (*pilT* and *phd*), yellow; resistance genes(*cpxA, cpxR* and *bcr)*, blue; proteins with other or unknown functions, gray. Homologous gene clusters in different isolates are shaded in gray (>97%).

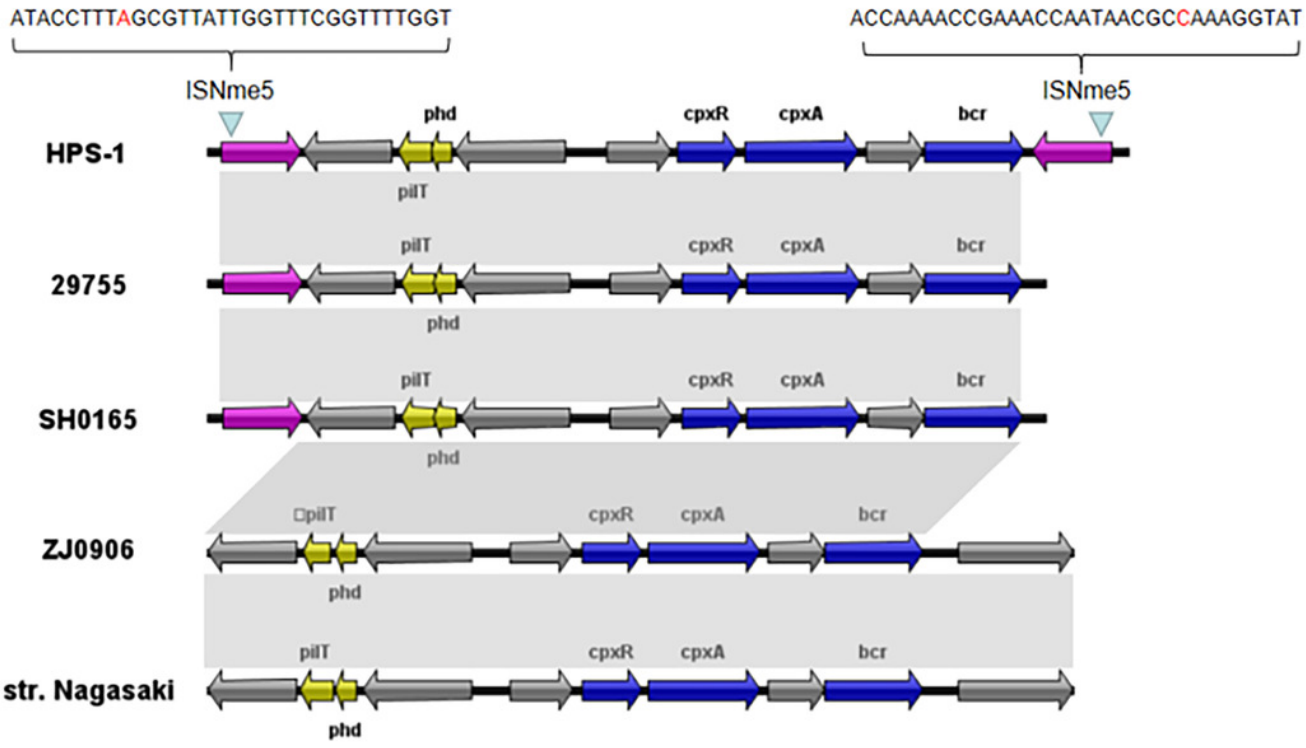

Figure 6

Neighbor-joining phylogenetic tree based on *bcr* gene sequences obtained from the current study and downloaded from NCBI.

The tree was constructed using MEGA 7 with 1,000 bootstrap replicates. The different colors of the branches represent lineages. The *G. parasuis* HPS-1 is indicated by a solid circle.

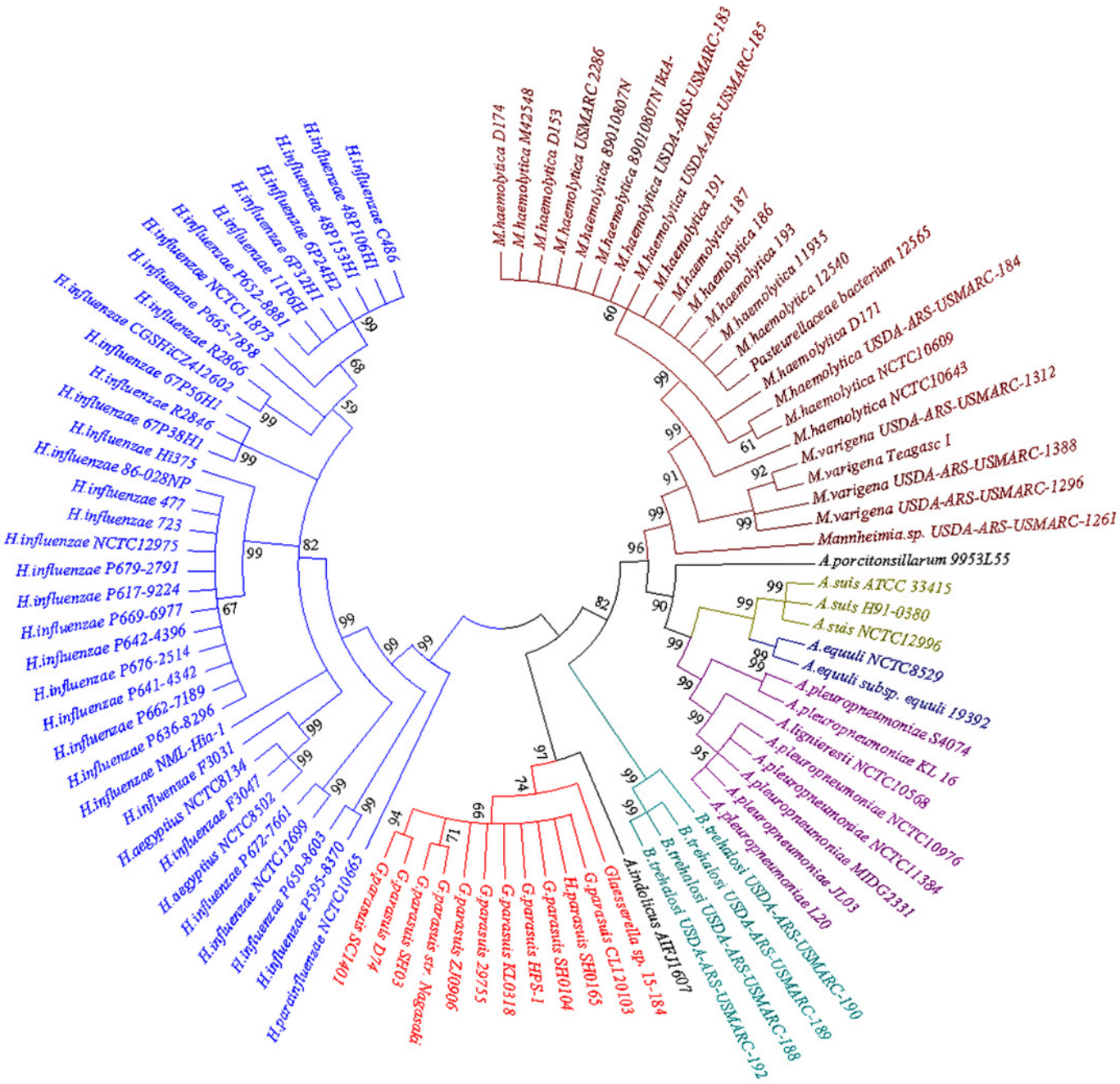
M.haemolytica D174
M.haemolytica M42548
M.haemolytica D153
M.haemolytica USMARC 2286
M.haemolytica 89010807N
M.haemolytica 89010807N ilvA-
M.haemolytica USDA-ARS-USMARC-183
M.haemolytica USDA-ARS-USMARC-185
M.haemolytica 191
M.haemolytica 187
M.haemolytica 186
M.haemolytica 193
M.haemolytica 11935
M.haemolytica 12540
Pasteurellaceae bacterium 12565
M.haemolytica D171
M.haemolytica USDA-ARS-USMARC-184
M.haemolytica NCTC10609
M.haemolytica NCTC10643
M.varigena USDA-ARS-USMARC-1312
M.varigena Teagasc 1
M.varigena USDA-ARS-USMARC-1388
M.varigena USDA-ARS-USMARC-1296
Mannheimia.sp. USDA-ARS-USMARC-1261
A.porcitonsillarum 9953L55
A.suis ATCC 33415
A.suis H91-0380
A.suis NCTC12996
A.equuli NCTC8529
A.equuli subsp. equuli 19392
A.pleuropneumoniae S4074
A.lignieresii NCTC10568
A.pleuropneumoniae KL 16
A.pleuropneumoniae NCTC10976
A.pleuropneumoniae MIDG2331
A.pleuropneumoniae JL03
A.pleuropneumoniae L20
B.trehalosi USDA-ARS-USMARC-189
B.trehalosi USDA-ARS-USMARC-190
B.trehalosi USDA-ARS-USMARC-188
B.trehalosi USDA-ARS-USMARC-192
A.indolicus AIR71607
Glaesserella sp. 15-184
G.parasuis CL120103
H.parasuis SH0165
H.parasuis SH0104
G.parasuis HPS-1
G.parasuis KL0318
G.parasuis 29755
G.parasuis 29090
G.parasuis str. Nagasaki
G.parasuis SH03
G.parasuis D74
G.parasuis SC1401
H.parainfluenzae NCTC10665
H.influenzae P595-8370
H.influenzae P650-8603
H.influenzae NCTC12699
H.influenzae P672-7661
H.aegyptius NCTC8502
H.influenzae F3047
H.aegyptius NCTC8134
H.influenzae F3031
H.influenzae NML-Hia-1
H.influenzae P636-8296
H.influenzae P662-7189
H.influenzae P641-4342
H.influenzae P676-2514
H.influenzae P642-4396
H.influenzae P669-6977
H.influenzae P617-9224
H.influenzae P679-2791
H.influenzae NCTC12975
H.influenzae 723
H.influenzae 477
H.influenzae 86-028NP
H.influenzae HI375
H.influenzae 67P38H1
H.influenzae R2846
H.influenzae 67P56H1
H.influenzae CGSHiCZ412602
H.influenzae R2866
H.influenzae P665-7858
H.influenzae NCTC11873
H.influenzae P652-8881
H.influenzae 11P6H
H.influenzae 6P32H1
H.influenzae 6P24H2
H.influenzae 48P153H1
H.influenzae 48P106H1
H.influenzae C486

Figure 7

Comparison of the genetic structures of pHN61, pFS39, pYL1, pFZ51 and pHB0503.

The accession numbers and origins of these plasmids are displayed on the left side. Arrows represent putative open reading frames, the positions and orientations of the genes. Blue arrows indicate Rep-like protein involved in plasmid replication. Green arrows indicate hypothetical protein. Regions with more than 98% nucleotide sequence identity are shaded yellow.