Whole-genome sequencing analysis of *Glaesserella parasuis* isolates reveals extensive genomic variation and diverse antibiotic resistance determinants

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

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

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

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

Corresponding Author: He Yan
Email address: 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 the genetic variation, antibiotic resistance, and virulence mechanisms of *Glaesserella parasuis* (*G. parasuis*).

Methods: The whole-genome sequence of a ST328 isolate from diseased swine in China was then compared with 54 isolates from China sequenced in this study and 39 strains from China and other 8 countries sequenced by others in order to better understand the genetic variation, antibiotic resistance, and virulence mechanisms of *G. parasuis*.

Results: The ST328 genome contained a novel Tn*6678* transposon harboring unique resistance determinants. It also contained a small broad-host-range plasmid pYL1 carrying *aac(6')-Ie-aph(2'')-Ia* and bla_{ROB-1} that was transferred to *Staphylococcus aureus* RN4220 by electroporation and was highly stable under kanamycin selection. The diversity and wide variation of *G. parasuis* are mainly reflected in the accessory genomes (85.13–91.74%). Phylogenetic analysis revealed two major subgroups associated with country, serotypes, and MLST. Novel virulence factors (*gigP, malQ,* and *gmhA*) and drug resistance genes (*norA, bacA*, *ksgA,* and *bcr*) 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 to date provides valuable insight into the emergence and transmission of *G. parasuis* in the swine industry. The result suggests the importance of genomic variations, especially transposon-related and/or plasmid-related gene variations, in the evolution of *G. parasuis*.

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

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

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

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

[3] 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 (He Yan)

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 the genetic variation, antibiotic resistance, and virulence mechanisms of *Glaesserella parasuis* (*G. parasuis*).

Methods: The whole-genome sequence of a ST328 isolate from diseased swine in China was then compared with 54 isolates from China sequenced in this study and 39 strains from China and other 8 countries sequenced by others in order to better understand the genetic variation, antibiotic resistance, and virulence mechanisms of *G. parasuis*.

Results: The ST328 genome contained a novel Tn*6678* transposon harboring unique resistance determinants. It also contained a small broad-host-range plasmid pYL1 carrying *aac(6')-Ie-aph(2'')-Ia* and bla_{ROB-1} that was transferred to *Staphylococcus aureus* RN4220 by electroporation and was highly stable under kanamycin selection. The diversity and wide variation of *G. parasuis* are mainly reflected in the accessory genomes (85.13–91.74%). Phylogenetic analysis revealed two major subgroups associated with country, serotypes, and MLST. Novel virulence factors (*gigP, malQ,* and *gmhA*) and drug resistance genes (*norA, bacA*, *ksgA,* and *bcr*) 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 to date provides valuable insight into the emergence and transmission of *G. parasuis* in the swine industry. The result suggests the importance of genomic variations, especially transposon-related and/or plasmid-related gene variations, in the evolution of *G. parasuis*.

Introduction

Glaesserella parasuis, a gram-negative bacterium of the family *Pasteurellaceae*, 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 [1]. As China is one of the world's largest pork producers, with more than 463 million pigs accounting for approximately 50% of the global population [2], *G. parasuis* represents a significant threat to pig health and economic loss worldwide [3]. Penicillins and aminopenicillins are common treatments for *G. parasuis* infections [4], and clinical strains resistant to β-lactams and aminoglycosides have been reported in several countries [5-8].

Whole-genome sequencing (WGS) has revealed that the diversity and wide variation among *G. parasuis* strains are mainly reflected in their accessory genomes, which occupy a large fraction (~86%) of the gene content compared with the core genome [9]. The accessory genes are also associated with serovar-specific genes, including capsule genes and putative virulence factors [9]. A comparative study of five pathogenic *G. parasuis* strains of three serovars from diseased pigs in Japan, China, and the USA demonstrated high genomic diversity among the strains, which was partly attributed to a high number of mobile genetic elements such as transposons and bacteriophages; moreover, 4–6% of the analysed genomes comprised strain-specific accessory genes [10]. WGS analyses on 11 *G. parasuis* strains from Japan, USA, China, Germany, Sweden, and Switzerland identified several noteworthy differences in coding regions, including genes for outer membrane, metabolism, and pilin or adhesin-related genes [3]. Another study revealed that the toxin-antitoxin systems between phenotypically distinct *G. parasuis* strains from Japan and Sweden had genomic differences [11].

Antibiotic resistance in *G. parasuis* is mainly conferred by a combination of transferable antibiotic resistance genes (ARGs) and multiple target gene mutations. For instance, to date, two β-lactam resistance genes ($bla_{\text{ROB-1}}$ and bla_{TEM}) and an aminoglycoside-resistance gene (*aac (6')-Ib-cr*) have been identified in *G. parasuis* [7, 12, 13]. Moreover, the multi-resistant efflux pump AcrB may have an impact on the emergence of resistance in *G. parasuis*. Additionally, putative

compensatory mechanisms to overcome fitness impairment coinciding with the acquisition of resistance have been described for some drugs [2, 14].

However, the structural organization of genetic information, ARGs, potential pathogenicity, and evolutionary relationships among global *G. parasuis* strains remain unclear. Therefore, in this study, we sequenced a multidrug-resistant isolate from diseased swine in Dongguan, China, which was compared with 54 isolates from China sequenced by us and 39 strains from China and other 8 countries sequenced by others in order to improve our understanding and provide information for gaining better control to treat these infections.

Materials & Methods

Isolates

The multidrug-resistant *G. parasuis* isolate (HPS-1) of serotype 4 examined in this study 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. The isolate was resistant to β-lactams, aminoglycosides, macrolides, quinolones, lincomycin, and sulfonamides as determined by the disc agar diffusion method and the broth microdilution method (Table S1) [15].

The other 54 *G. parasuis* isolates were obtained from diseased pigs from more than 20 geographically dispersed farms in China between November 2007 and May 2017 (Table S2). Bacteria were identified by biochemical tests and 16S diagnostic PCR [6, 16]. All 55 *G. parasuis* isolates were characterised using serotyping and MLST as previously described [17, 18].

Genome sequencing, assembly, and bioinformatics analysis

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

Among the 55 isolates, one multidrug-resistant isolate (HPS-1) and one sensitive isolate (HPS-2) from diseased swine in Guangdong were randomly selected for WGS using the PacBio RSII (Pacific Biosciences, MenloPark, CA, USA) and Illumina MiSeq (Illumina, San Diego, CA, USA) platforms as previously described [19]. The genome assemblies of HPS-1 generated in this study were deposited in GenBank under accession number CP040243. The plasmid pYL1 and transposon Tn*6678* of HPS-1 were submitted to GenBank under accession number MK182379 and MK994978, respectively. Gene prediction and annotation were performed as previously described [19]. Genomic libraries of the other 53 genomes were prepared and sequenced as previously described [20]. WGS data were assembled *de novo*, and gene prediction and annotation were performed as previously described [21].

Phylogenetic and clustering analyses

Two phylogenetic trees were constructed to assess the relatedness of the 55 *G. parasuis* strains with 39 available, previously published, genome sequences (Table S2) using single-copy core orthologs and SNPs with the maximum-likelihood optimality criterion using PhyML v3.0; the SNP tree was virtually rooted using *Glaesserella* sp.15-184 as an outgroup [22]. The gene contents of all 94 isolates were compared using cd-hit (v 4.6.1) software, which utilizes protein sequences to effectively generate non-paralogous gene clusters (identity ≥ 0.8, $\geq 80\%$ the length of the longest cluster).

The search scope was limited to 65% identity to determine the phylogenetic profile of *bcr*. Multiple sequence alignment was performed with MUSCLE in MEGA 7.0, resulting in 92 candidate genes. The neighbour-joining phylogenetic tree was constructed in MEGA 7.0 with 1000 bootstrap replicates.

Comparison of antimicrobial resistance and virulence genes

A whole-genome Blast [23] 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). (see supplementary method for details)

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

Based on the results of the BLASTn search, the 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, which were aligned by MUSCLE in MEGA 7.0 and adjusted manually. The default parameter for both gap open and gap extend was -400. The phylogenetic tree was constructed using the neighbour-joining method in MEGA 7.0 with 1000 bootstrap replicates.

Electrotransformation and plasmid stability test

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 [24]. 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 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 $bla_{\text{ROB-1}}$ and *aac(6')-Ie-aph(2'')-Ia*.

Results and discussion

G. parasuis core and unique genes

Compilation of the 94 genomes covering all serovars and disease- and non-disease-causing backgrounds from nine geographic locations (Table S2) demonstrated expansion of the pan-genome, whereas the number of core genes remained relatively stable with the addition of new strains (Fig. 1A). This suggested the possibility of open pan-genomes experiencing frequent evolutionary changes through gene gains and losses or lateral gene transfers for efficient environmental adaptation. The size of the pan-genome was 5,243 genes, including 175 core genes shared among the 94 isolates (Fig. 1B). By contrast, 1,049 core genes were previously identified among 212 *G. parasuis* isolates from the UK (n = 121), Argentina (n = 1), Belgium (n = 1), China (n = 2), Denmark (n = 29), Germany (n = 6), Greece (n = 1), Italy (n = 2), Japan (n = 7), Spain (n = 21), Sweden (n = 4), Switzerland (n = 2), the USA (n = 6), and unknown origins (n=9) [9]. The diversity and wide variation were mainly reflected in the accessory genomes, occupying a large fraction (85.13–91.74%) of the *G. parasuis* gene content compared with the core genomes (Table 1). The number of unique genes ranged from 0 in 10 strains to 103 in strain D74, indicating that 0–4.6% of the genomes consisted of strain-specific accessory genes (Table 1). This further indicated the possibility of frequent horizontal gene transfer events in *G. parasuis*; thus, investigating the source of this variation is critical to understanding the adaptive potential of this species.

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

Phylogenetic analysis of *G. parasuis* isolates

The core gene sequence-based phylogeny for our isolates and reference isolates is shown in Fig. 2. This tree resolved two strongly structured lineages, lineages I and II, with each subgroup comprising multiple closely related isolates. This pattern was similar to a UPGMA-based dendrogram that exhibited two main clades, but differed from the population grouping predicted by MLST, which showed six main subgroups [18].

Phylogenetic lineages exhibited association with country, serotypes, and MLST types (Fig. 2). There was a certain degree of geographic structure, with a greater proportion of Chinese isolates in both clades, likely due to a larger proportion of Chinese isolates in the collection. This was analogous to the phylogenetic tree of the *G. parasuis* core genome built from concatenated core genes [9, 18]. Lineages I and II comprised eight and two countries, respectively. The fact that strains from different regions belonged to the same clade of the whole-genome phylogenetic tree suggested frequent migration of isolates between geographic regions. Serovars 5, 12, and 14 were identified predominantly in lineage I, while serovars 2 and 10 were mostly found in lineage II. For serovars 3, 8, 9, and 11, the numbers of isolates were too low to be able to restrict them to one clade. The remainder of the serovars were found in both clades. MLST analysis allocated the 39 isolates in GenBank to 13 unknown STs and 20 different STs, including six new STs, and the 55 isolates obtained in our study belonged to 49 different STs, including 39 new STs (Table S2).

Thus, the results of this study enrich the *G. parasuis* MLST databases and highlight the wide distribution of *G. parasuis* strains. Most strains of the same STs formed single clades, and there was no definitive association between STs and serotypes (Fig. 2), consistent with previous studies [18, 25]. 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

Variations in virulence and stress resistance genes were 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, the virulence factors *rfa* cluster, encoding enzymes for LPS core biosynthesis of *G. parasuis* [26], and *galU* and *galE,* resulting in impaired biofilm formation, were universally present in the *G. parasuis* isolates. The *rfaF* gene has been linked to serum resistance, adhesion, and invasion; *galU* plays a role in autoagglutination and biofilm formation; whereas *galE* appears to affect biofilm production indirectly in *G. parasuis* [27]. Serum resistance may play a role in the virulence of *G. parasuis* [28]. 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 [29, 30].

The sulfonamide resistance gene (*sul2*), β-lactam-resistant genes (*pbp1a*, *pbp3a*), and aminoglycoside resistance factors (*aph(3'')-Ib*) were universally present in the *G. parasuis* isolates (Fig. 3). The main ARGs associated with resistance in *G. parasuis*, including the β-lactam-resistant gene, tetracycline resistance genes, aminoglycoside resistance genes, fluoroquinolone resistance gene, chloramphenicol resistance genes, and two efflux pump genes mentioned above were discovered among these isolates (Fig. 3). The $bla_{\text{ROB-1}}$, *sul2*, *aph(3'')-Ib, tetB, aac(6')-Ie-aph(2'')-Ia, catIII,* and *floR* genes have previously been identified in *G. parasuis* [8]. The 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*, which was contained in the ISSag10 sequence of all three isolates (Fig. 3). The *lunC* gene was also identified in plasmid pHN61 of *G. parasuis* [31]. 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*, which have not previously been reported in *G. parasuis*.

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, two genes associated with the two-component signal transduction system, one efflux pump-associated gene, 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.

The clinical isolate *G. parasuis* HPS-1 also harbours two drug resistance genes (*cpxA* and *cpxR*) and an efflux pump gene (*bcr*) organized within a putative complete transposon. Association between the Cpx system and bacterial antimicrobial resistance has been reported in *Escherichia coli*, *Salmonella enterica, Klebsiella pneumoniae,* and *G. parasuis* [32-36]. CpxR

253 plays essential roles in mediating macrolide resistance (i.e., erythromycin) [36]. The Bcr/CflA
254 efflux system was identified as a group of antiporters that confer resistance to chloramphenicol,
255 florfenicol, and bicyclomycin by actively transporting these compounds out of the cell [37].

256 BLASTn searches for the *bcr* gene returned a large set of divergently related sequences
257 using default parameters. These sequences were annotated as bicyclomycin/multidrug efflux
258 system, Bcr/CflA family drug resistance efflux transporter, Bcr/CflA family multidrug efflux
259 MFS transporter or drug resistance transporter, and Bcr/CflA subfamily. Phylograms revealed
260 that the *bcr* gene in HPS-1 was most closely related to homologues identified in other members
261 of the Pasteurellaceae, particularly *G. parasuis*, *Actinobacillus indolicus*, *Bibersteinia trehalosi,*
262 *Actinobacillus* (*A. pleuropneumoniae, A. suis, A. equuli, A. lignieresii, A. indolicus*, and *A.*
263 *porcitonsillarum*), and *Mannheimia* (*M. haemolytica* and *M. varigena*), all of which are known
264 causative agents of upper respiratory tract infections (Fig. 6).

265 In *G. parasuis*, only the efflux pump AcrB, belonging to the RND family, has been
266 analysed to date. Efflux pump AcrB may play a role in multidrug resistance, and the *acrAB* gene
267 cluster could affect the efflux of macrolides in *G. parasuis* [14]. However, this is the first
268 description of the efflux pump Bcr/CflA in *G. parasuis*, belonging to the MFS. This efflux
269 encoded by *bcr* was shown to be involved in bicyclomycin (sulfonamide) resistance, and its
270 presence on a transposon indicated its potential transferability.

271 The neighbour-joining phylogenetic tree using 92 *bcr* genes selected from the BLASTn
272 searches clearly demonstrated two distinctive clades. The first clade contained *bcr* genes of
273 *Hemophilus influenzae,* which colonizes humans, and other bacteria that colonize animals.
274 Members of the second clade were divided into four apparent subclades, including *G. parasuis*,
275 *B. trehalosi*, *Actinobacillus* spp., and *Mannheimia* spp. Except for *G. parasuis*, the
276 chromosomally encoded Bcr/CflA from *G. parasuis* HPS-1 most closely clustered with that
277 found in *A. indolicus*. The phylogenetic tree indicated a divergent evolutionary pattern between
278 animal-origin *Pasteurellaceae* bacteria.

279 **General features and electrotransformation of the plasmid pYL1**

280 The plasmid pYL1 identified in HPS-1 contained seven ORFs with an average length of 912
281 bp, with one encoded protein of undetermined function (Fig. 4), and two antimicrobial resistance
282 genes, *bla*$_{ROB-1}$ and *aac(6')-Ie-aph(2'')-Ia*. The ROB-1 of plasmid pYL1 had a typical size of 305
283 bp, in line with functionally active members of the ROB-1 family from different plasmids in
284 *Pasteurellaceae* species. AAC(6')-Ie-APH(2')-Ia, the most important aminoglycoside-resistance
285 enzyme in gram-positive bacteria conferring resistance to almost all known aminoglycoside
286 antibiotics in clinical use, also had a typical size of 479 amino acids in this family [38]. Four
287 ORFs were identified to encode a 3′-truncated transposase protein ISApl1 (30 amino acids), a
288 Rep-like protein (444 amino acids) involved in plasmid replication, and two Mob proteins,
289 MobC (144 amino acids) and MobA (541 amino acids), associated with plasmid mobilization
290 (Fig. 4). To date, two β-lactam resistance genes (*bla*$_{ROB-1}$ and *bla*$_{TEM}$) have been reported in *G.*
291 *parasuis* [13]. A β-lactam resistance plasmid, pB1000, harbouring *bla*$_{ROB-1}$ was detected in *G.*
292 *parasuis* clinical strains isolated from Glässer's disease lesions [7]. Although *aac(6')-Ib-cr* is
293 considered the most prominent aminoglycoside-resistance gene in *G. parasuis* [7, 12], the
294 bifunctional aminoglycoside-resistance enzyme AAC(6')-Ie-APH(2')-Ia in plasmids is also
295 reported in GenBank for *G. parasuis* strains.

296 Except for resistance genes, pYL1 had the same backbone and genetic structure and showed
297 100% nucleotide identity to four plasmids, namely pFZ51, pFS39, pHN61, and pHB0503 (Table
298 S3) [31, 39, 40]. Further, the pYL1 segment of the resistant genes and its flanking region exhibited
299 ≥58% sequence homology to these four plasmids (Fig. 7), suggesting evolution among these
300 small plasmids.

301 The transposase gene of ISApl1 in pYL1 had an internal deletion of 659 bp, and the 3′and 5′
302 ends were intact. The truncated ISApl1 linked with *bla*$_{ROB-1}$ suggested that ISApl1 played a key
303 role in transposition of *bla*$_{ROB-1}$ by facilitating the horizontal transfer of β-lactam and
304 aminoglycoside resistance among *G. parasuis* isolates and it was also found in *A.*

porcitonsillarum or *G. parasuis* plasmids pFJS5863, pQY431, and pFS39. However, the function of ISApl1 requires further investigation.

Transformation of pYL1 into *S. aureus* RN4220 was achieved at a frequency of 10^{-9} cells per recipient cell by electroporation. Hence, 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_{ROB-1} and *aac(6')-Ie-aph(2'')-Ia* contributed to the penicillin resistance and aminoglycoside antibiotic resistance of *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, confirmed by PCR mapping.

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 potential 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 is necessary to gain a better understanding of the rapid development of antibiotic resistance associated with mobile genetic elements in *G. parasuis* based on WGS analysis.

Acknowledgements

We thank members of our laboratories for fruitful discussions.

References

[1] Brockmeier SL. Prior infection with *Bordetella bronchiseptica* increases nasal colonization by *Haemophilus parasuis* in swine. Vet Microbiol 2004; 99: 75-8. https://doi.org/10.1016/j.vetmic.2003.08.013.

[2] Zhou LJ, Ying GG, Liu S, Zhang RQ, Lai HJ, Chen ZF, et al. Excretion masses and environmental occurrence of antibiotics in typical swine and dairy cattle farms in China. Sci Total Environ 2013;444:183-95. https://doi.org/10.1016/j.scitotenv.2012.11.087.

[3] Brockmeier SL, Register KB, Kuehn JS, Nicholson TL, Loving CL, Bayles DO, et al. Virulence and draft genome sequence overview of multiple strains of the swine pathogen *Haemophilus parasuis*. Plos One 2014;9:e103787. https://doi.org/10.1371/journal.pone.0103787.

[4] Moleres J, Santos-Lopez A, Lazaro I, Labairu J, Prat C, Ardanuy C, et al. Novel bla_{ROB-1}-bearing plasmid conferring resistance to beta-lactams in *Haemophilus parasuis* isolates from healthy weaning pigs. Appl Environ Microbiol 2015;81:3255-67. https://doi.org/10.1128/AEM.03865-14.

[5] Wissing A, Nicolet J, Boerlin P. The current antimicrobial resistance situation in Swiss veterinary medicine. Schweiz Arch Tierheilkd 2001;143:503-10.

[6] de la Fuente AJ, Tucker AW, Navas J, Blanco M, Morris SJ, Gutierrez-Martin CB. Antimicrobial susceptibility patterns of *Haemophilus parasuis* from pigs in the United Kingdom and Spain. Vet Microbiol 2007;120:184-91. https://doi.org/10.1016/j.vetmic.2006.10.014.

[7] San MA, Escudero JA, Catalan A, Nieto S, Farelo F, Gibert M, et al. Beta-lactam resistance in *Haemophilus parasuis* is mediated by plasmid pB1000 bearing bla_{ROB-1}. Antimicrob Agents Chemother 2007;51:2260-4. https://doi.org/10.1128/AAC.00242-07.

354 [8] Zhao Y, Guo L, Li J, Huang X, Fang B. Characterization of antimicrobial resistance genes in
355 *Haemophilus parasuis* isolated from pigs in China. Peerj 2018;6:e4613.
356 https://doi.org/10.7717/peerj.4613.

357 [9] Howell KJ, Weinert LA, Chaudhuri RR, Luan SL, Peters SE, Corander J, et al. The use of
358 genome wide association methods to investigate pathogenicity, population structure and serovar
359 in *Haemophilus parasuis*. Bmc Genomics 2014;15:1179. https://doi.org/10.1186/1471-2164-15-
360 1179.

361 [10] Bello-Orti B, Aragon V, Pina-Pedrero S, Bensaid A. Genome comparison of three serovar 5
362 pathogenic strains of *Haemophilus parasuis*: insights into an evolving swine pathogen.
363 Microbiology+ 2014;160:1974-84. https://doi.org/10.1099/mic.0.079483-0.

364 [11] Nicholson TL, Brunelle BW, Bayles DO, Alt DP, Shore SM. Comparative genomic and
365 methylome analysis of non-virulent D74 and virulent Nagasaki *Haemophilus parasuis* isolates.
366 Plos One 2018;13:e205700. https://doi.org/10.1371/journal.pone.0205700.

367 [12] Doi Y, Arakawa Y. 16S ribosomal RNA methylation: emerging resistance mechanism
368 against aminoglycosides. Clin Infect Dis 2007;45:88-94. https://doi.org/10.1086/518605.

369 [13] By Guo LGL, Zhang JZJ, Xu CXC, Ren TRT, Zhang BZB, Chen JCJ, et al. Detection and
370 Characterization of β-Lactam Resistance in *Haemophilus parasuis* Strains from Pigs in South
371 China. J Integr Agr 2012;11:116-21.

372 [14] Feng S, Xu L, Xu C, Fan H, Liao M, Ren T. Role of *acrAB* in antibiotic resistance of
373 *Haemophilus parasuis* serovar 4. Vet J 2014;202:191-4.
374 https://doi.org/10.1016/j.tvjl.2014.05.045.

375 [15] Pruller S, Turni C, Blackall PJ, Beyerbach M, Klein G, Kreienbrock L, et al. Towards a
376 Standardized Method for Broth Microdilution Susceptibility Testing of *Haemophilus parasuis*. J
377 Clin Microbiol 2017;55:264-73. https://doi.org/10.1128/JCM.01403-16.

378 [16] Oliveira S, Galina L, Pijoan C. Development of a PCR test to diagnose *Haemophilus*

379 *parasuis* infections. J Vet Diagn Invest 2001;13:495-501.
380 https://doi.org/10.1177/104063870101300607.

381 [17] Jia AJA, Zhou RZR, Fan HFH, Yang KYK, Zhang JZJ, Xu YXY, et al. Development of
382 Serotype-Specific PCR Assays for Typing of *Haemophilus parasuis* Isolates Circulating in
383 Southern China. JOURNAL OF CLINICAL MICROBIOLOGY 2017;55:3249-57.

384 [18] Wang L, Ma L, Liu Y, Gao P, Li Y, Li X, et al. Multilocus sequence typing and virulence
385 analysis of *Haemophilus parasuis* strains isolated in five provinces of China. Infect Genet Evol
386 2016;44:228-33. https://doi.org/10.1016/j.meegid.2016.07.015.

387 [19] Zheng B, Yu X, Xu H, Guo L, Zhang J, Huang C, et al. Complete genome sequencing and
388 genomic characterization of two Escherichia coli strains co-producing MCR-1 and NDM-1 from
389 bloodstream infection. Sci Rep 2017;7:17885. https://doi.org/10.1038/s41598-017-18273-2.

390 [20] Soge OO, No D, Michael KE, Dankoff J, Lane J, Vogel K, et al. Transmission of MDR
391 MRSA between primates, their environment and personnel at a United States primate centre. J
392 Antimicrob Chemother 2016;71:2798-803. https://doi.org/10.1093/jac/dkw236.

393 [21] Larsen J, Andersen PS, Winstel V, Peschel A. *Staphylococcus aureus* CC395 harbours a
394 novel composite staphylococcal cassette chromosome mec element. J Antimicrob Chemother
395 2017;72:1002-5. https://doi.org/10.1093/jac/dkw544.

396 [22] Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms
397 and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML
398 3.0. Syst Biol 2010;59:307-21. https://doi.org/10.1093/sysbio/syq010.

399 [23] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J
400 Mol Biol 1990;215:403-10. https://doi.org/10.1016/S0022-2836(05)80360-2.

401 [24] Wang XM, Li XS, Wang YB, Wei FS, Zhang SM, Shang YH, et al. Characterization of a
402 multidrug resistance plasmid from *Enterococcus faecium* that harbours a mobilized *bcrABDR*
403 locus. J Antimicrob Chemother 2015;70:609-11. https://doi.org/10.1093/jac/dku416.

404 [25] Olvera A, Cerda-Cuellar M, Aragon V. Study of the population structure of *Haemophilus*
405 *parasuis* by multilocus sequence typing. Microbiology+ 2006;152:3683-90.
406 https://doi.org/10.1099/mic.0.29254-0.

407 [26] Zhang B, Xu C, Zhang L, Zhou S, Feng S, He Y, et al. Enhanced adherence to and invasion
408 of PUVEC and PK-15 cells due to the overexpression of RfaD, ThyA and Mip in the
409 DeltaompP2 mutant of *Haemophilus parasuis* SC096 strain. Vet Microbiol 2013;162:713-23.
410 https://doi.org/10.1016/j.vetmic.2012.09.021.

411 [27] Zou Y, Feng S, Xu C, Zhang B, Zhou S, Zhang L, et al. The role of galU and galE of
412 *Haemophilus parasuis* SC096 in serum resistance and biofilm formation. Vet Microbiol
413 2013;162:278-84. https://doi.org/10.1016/j.vetmic.2012.08.006.

414 [28] Cerda-Cuellar M, Aragon V. Serum-resistance in *Haemophilus parasuis* is associated with
415 systemic disease in swine. Vet J 2008;175:384-9. https://doi.org/10.1016/j.tvjl.2007.01.016.

416 [29] Amano H, Shibata M, Kajio N, Morozumi T. Pathogenicity of *Haemophilus parasuis*
417 serovars 4 and 5 in contact-exposed pigs. J Vet Med Sci 1996;58:559-61.
418 https://doi.org/10.1292/jvms.58.559.

419 [30] Brockmeier SL, Loving CL, Mullins MA, Register KB, Nicholson TL, Wiseman BS, et al.
420 Virulence, transmission, and heterologous protection of four isolates of *Haemophilus parasuis*.
421 Clin Vaccine Immunol 2013;20:1466-72. https://doi.org/10.1128/CVI.00168-13.

422 [31] Chen LCL, Cai XCX, Wang XWX, Zhou XZX, Wu DWD, Xu XXX, et al. Characterization
423 of plasmid-mediated lincosamide resistance in a field isolate of *Haemophilus parasuis*. J
424 Antimicrob Chemoth 2010;65:2256-8.

425 [32] Audrain B, Ferrieres L, Zairi A, Soubigou G, Dobson C, Coppee JY, et al. Induction of the
426 Cpx envelope stress pathway contributes to *Escherichia coli* tolerance to antimicrobial peptides.
427 Appl Environ Microbiol 2013;79:7770-9. https://doi.org/10.1128/AEM.02593-13.

428 [33] Hu WS, Chen HW, Zhang RY, Huang CY, Shen CF. The expression levels of outer

membrane proteins STM1530 and OmpD, which are influenced by the CpxAR and BaeSR two-component systems, play important roles in the ceftriaxone resistance of *Salmonella enterica* serovar Typhimurium. Antimicrob Agents Chemother 2011;55:3829-37. https://doi.org/10.1128/AAC.00216-11.

[34] Kurabayashi K, Hirakawa Y, Tanimoto K, Tomita H, Hirakawa H. Role of the CpxAR two-component signal transduction system in control of fosfomycin resistance and carbon substrate uptake. J Bacteriol 2014;196:248-56. https://doi.org/10.1128/JB.01151-13.

[35] Srinivasan VB, Vaidyanathan V, Mondal A, Rajamohan G. Role of the two component signal transduction system CpxAR in conferring cefepime and chloramphenicol resistance in *Klebsiella pneumoniae* NTUH-K2044. Plos One 2012;7:e33777. https://doi.org/10.1371/journal.pone.0033777.

[36] Cao Q, Feng F, Wang H, Xu X, Chen H, Cai X, et al. *Haemophilus parasuis* CpxRA two-component system confers bacterial tolerance to environmental stresses and macrolide resistance. Microbiol Res 2018;206:177-85. https://doi.org/10.1016/j.micres.2017.10.010.

[37] Marklevitz J, Harris LK. Prediction driven functional annotation of hypothetical proteins in the major facilitator superfamily of S. aureus NCTC 8325. Bioinformation 2016; 12:254-62. https://doi.org/10.6026/97320630012254.

[38] Rouch DA, Byrne ME, Kong YC, Skurray RA. The *aacA-aphD* gentamicin and kanamycin resistance determinant of Tn*4001* from *Staphylococcus aureus*: expression and nucleotide sequence analysis. J Gen Microbiol 1987;133:3039-52. https://doi.org/10.1099/00221287-133-11-3039.

[39] Yang SS, Sun J, Liao XP, Liu BT, Li LL, Li L, et al. Co-location of the *erm*(T) gene and $bla_{\text{ROB-1}}$ gene on a small plasmid in *Haemophilus parasuis* of pig origin. J Antimicrob Chemother 2013;68:1930-2. https://doi.org/10.1093/jac/dkt112.

[40] Kang M, Zhou R, Liu L, Langford PR, Chen H. Analysis of an *Actinobacillus*

454 *pleuropneumoniae* multi-resistance plasmid, pHB0503. Plasmid 2009;61:135-9.
455 https://doi.org/10.1016/j.plasmid.2008.11.001.

Table 1(on next page)

Orthologous clusters in the *G. parasuis* pan-genome

1 **Table 1 Orthologous clusters in the *G. parasuis* pan-genome**

Strain	total	not_core	%Non-core	Unique	%Unique
12939	1974	1780	90.17%	24	1.20%
131	2050	1854	90.44%	19	0.90%
16	1936	1739	89.82%	1	0.10%
174	1433	1252	87.37%	46	3.20%
29755	2182	1991	91.25%	3	0.10%
84-15995	2206	2017	91.43%	33	1.50%
84-17975	2061	1869	90.68%	1	0.00%
CCUG3712	2176	1979	90.95%	11	0.50%
CL120103	1962	1765	89.96%	1	0.10%
D74	2218	2024	91.25%	103	4.60%
F9	2318	2121	91.50%	51	2.20%
gx033	2113	1923	91.01%	10	0.50%
H12	1955	1763	90.18%	3	0.20%
H35	2045	1854	90.66%	0	0.00%
H36	2034	1843	90.61%	0	0.00%
H38	2153	1955	90.80%	0	0.00%
H39	2157	1959	90.82%	0	0.00%

H465	1540	1350	87.66%	32	2.10%
H47	1997	1806	90.44%	8	0.40%
Hp100-13	2251	2044	90.80%	8	0.40%
HPS10	2081	1888	90.73%	14	0.70%
HPS11	1912	1718	89.85%	14	0.70%
HPS4	1973	1777	90.07%	0	0.00%
HPS6	1932	1737	89.91%	3	0.20%
HPS9	2060	1870	90.78%	10	0.50%
K3	2195	1997	90.98%	3	0.10%
KL0318	1997	1807	90.49%	1	0.10%
MN-H	1859	1667	89.67%	46	2.50%
Nagasaki	2260	2070	91.59%	18	0.80%
SC1401	1997	1802	90.24%	1	0.10%
SH0104	1993	1802	90.42%	1	0.10%
SH0165	2031	1841	90.65%	18	0.90%
SH03	1979	1790	90.45%	5	0.30%
ST4-1	2041	1851	90.69%	5	0.20%
ST4-2	2059	1870	90.82%	8	0.40%
SW114	1947	1759	90.34%	73	3.70%

SW140	1484	1300	87.60%	41	2.80%
YT	1925	1729	89.82%	0	0.00%
ZJ0906	2211	2021	91.41%	42	1.90%
H100	2143	1951	91.04%	7	0.30%
H105	2088	1896	90.80%	5	0.20%
H106	2121	1925	90.76%	3	0.10%
H110	2163	1966	90.89%	3	0.10%
H112	2124	1928	90.77%	2	0.10%
H115	2086	1894	90.80%	2	0.10%
H134	2150	1954	90.88%	0	0.00%
H140	2063	1871	90.69%	2	0.10%
H143	2103	1912	90.92%	2	0.10%
H157	2357	2159	91.60%	20	0.80%
H159	2225	2027	91.10%	17	0.80%
H160	2350	2156	91.74%	23	1.00%
H164	2074	1882	90.74%	6	0.30%
H178	2370	2169	91.52%	45	1.90%
H190	2095	1903	90.84%	4	0.20%
H191	2128	1936	90.98%	2	0.10%

H197	2210	2013	91.09%	18	0.80%
H199	2121	1925	90.76%	11	0.50%
H19	2209	1987	89.95%	6	0.30%
H201	2166	1972	91.04%	8	0.40%
H222	2056	1863	90.61%	11	0.50%
H223	2321	2120	91.34%	17	0.70%
H233	2207	2010	91.07%	15	0.70%
H257	2139	1948	91.07%	8	0.40%
H259	2118	1926	90.93%	2	0.10%
H25	1231	1048	85.13%	16	1.30%
H263	2049	1857	90.63%	2	0.10%
H26	2036	1845	90.62%	17	0.80%
H275	2133	1942	91.05%	6	0.30%
H27	2046	1853	90.57%	5	0.20%
H285	2044	1851	90.56%	5	0.20%
H292	2109	1917	90.90%	19	0.90%
H299	2174	1983	91.21%	11	0.50%
H312	2180	1986	91.10%	2	0.10%
H313	2193	2001	91.24%	5	0.20%

H33	2166	1975	91.18%	2	0.10%
H40	2124	1932	90.96%	6	0.30%
H43	2162	1946	90.01%	3	0.10%
H45	2015	1824	90.52%	9	0.40%
H46	2079	1887	90.76%	3	0.10%
H49	2256	2060	91.31%	34	1.50%
H52	2092	1899	90.77%	7	0.30%
H60	2029	1838	90.59%	6	0.30%
H61	2270	2079	91.59%	16	0.70%
H64	2047	1851	90.43%	3	0.10%
H68	2044	1852	90.61%	5	0.20%
H74	2187	1995	91.22%	7	0.30%
H78	2072	1879	90.69%	0	0.00%
H80	2079	1886	90.72%	3	0.10%
H82	2248	2056	91.46%	2	0.10%
H87	2140	1944	90.84%	0	0.00%
H90	2138	1942	90.83%	0	0.00%
H92	2227	2035	91.38%	3	0.10%
HPS-1	2288	2097	91.65%	27	1.20%

HPS-2	2336	2145	91.82%	37	1.60%

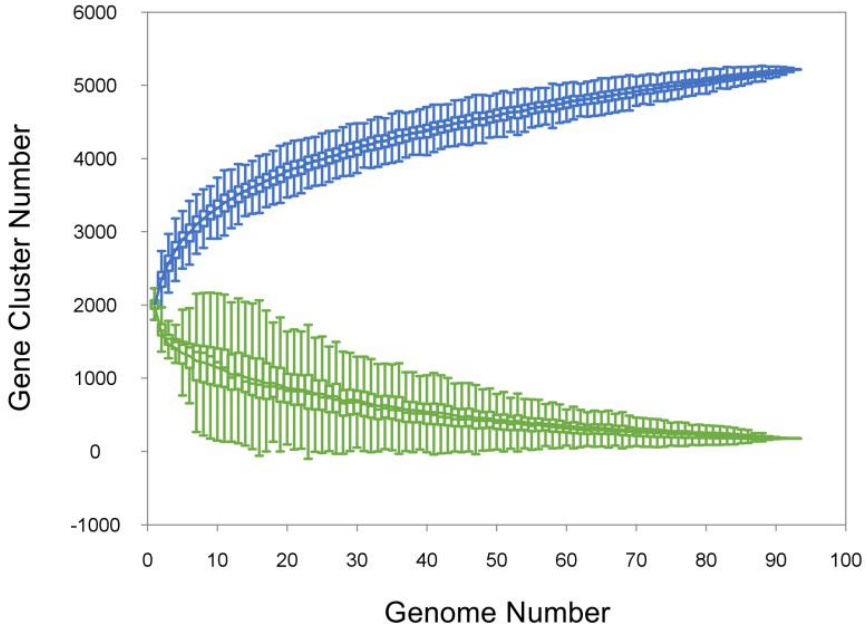
2

Figure 1(on next page)

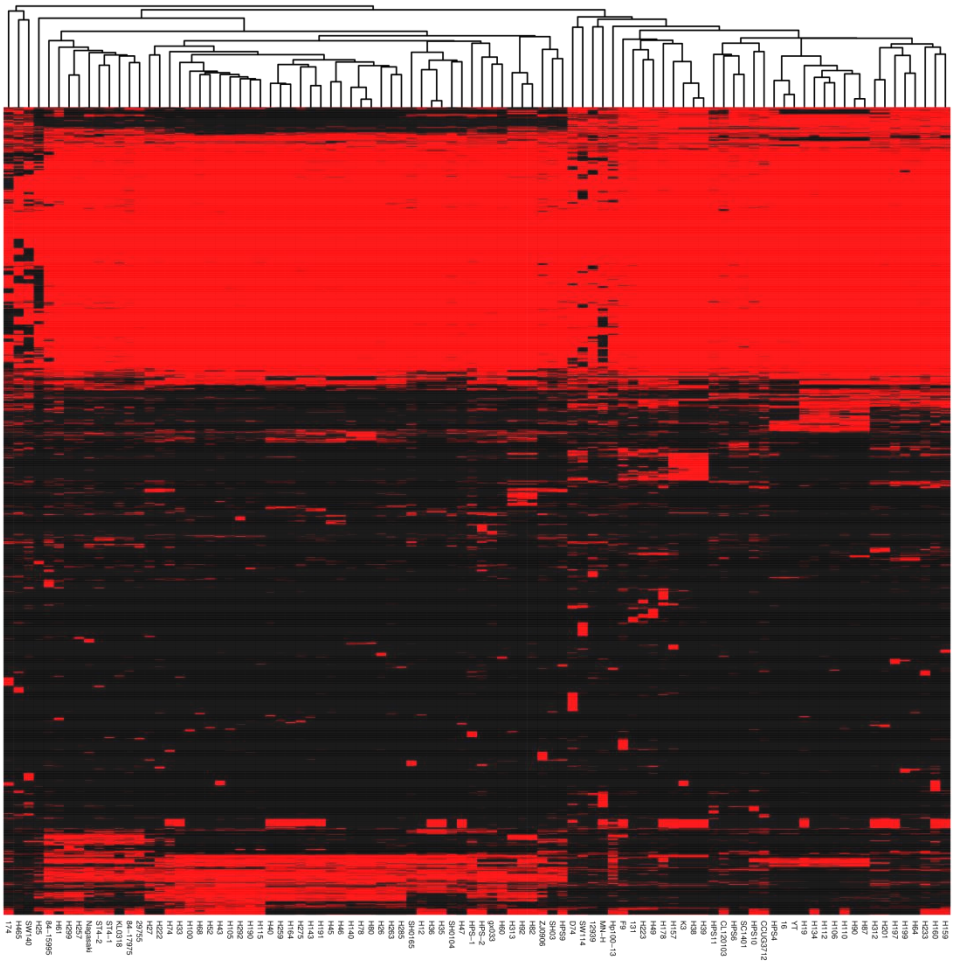
Analysis of the core and pan-genome of *G. parasuis* isolates.

(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)= 2483.54x0.18 -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. J: Translation, ribosomal structure and biogenesis. A: RNAprocessing and modification. K: Transcription. L: Replication, recombination and repair. B: Chromatin structure and dynamics. D: Cell cycle control, cell division, chromosome partitioning. Y: Nuclear structure. V: Defense mechanisms. T: Signal transduction mechanisms. M: Cell wall/membrane/envelope biogenesis. N: Cell motility. Z: Cytoskeleton. W: Extracellular structures. U: Intracellular trafficking, secretion,and vesicular transport. O: Posttranslational modification, protein turnover, chaperones. C: Energy production and conversion. G: Carbohydrate transport and metabolism. E:Amino acid transport and metabolism. F: Nucleotide transport and metabolism. H: Coenzyme transport and metabolism. I: Lipid transport and metabolism. P: Inorganic ion transport and metabolism. Q: Secondary metabolites biosynthesis, transport and catabolism. R: General function prediction only. S: Function unknown.

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

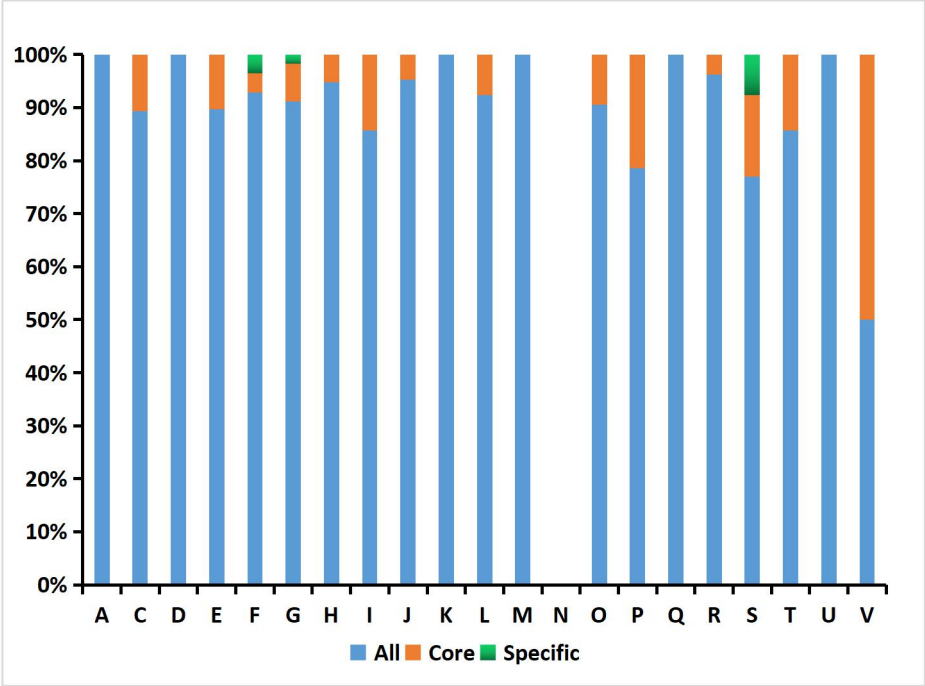


Figure 1. Analysis of the core and pan-genome of *G. parasuis* isolates. (A) Core and pan-genomic calculations in *G. parasuis* isolates. Each green point represents the number of genes conserved between genomes. All of the points are plotted as a function of the strain number(x). The deduced pan-genome 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. J: Translation, ribosomal structure and biogenesis. A: RNAprocessing and modification. K: Transcription. L: Replication, recombination and repair. B: Chromatin structure and dynamics. D: Cell cycle control, cell division, chromosome partitioning. Y: Nuclear structure. V: Defense mechanisms. T: Signal transduction mechanisms. M: Cell wall/membrane/envelope biogenesis. N: Cell motility. Z: Cytoskeleton. W: Extracellular structures. U: Intracellular trafficking, secretion,

and vesicular transport. O: Posttranslational modification, protein turnover, chaperones. C: Energy production and conversion. G: Carbohydrate transport and metabolism. E:Amino acid transport and metabolism. F: Nucleotide transport and metabolism. H: Coenzyme transport and metabolism. I: Lipid transport and metabolism. P: Inorganic ion transport and metabolism. Q: Secondary metabolites biosynthesis, transport and catabolism. R: General function prediction only. S: Function unknown.

Figure 2

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

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 represents lineage.

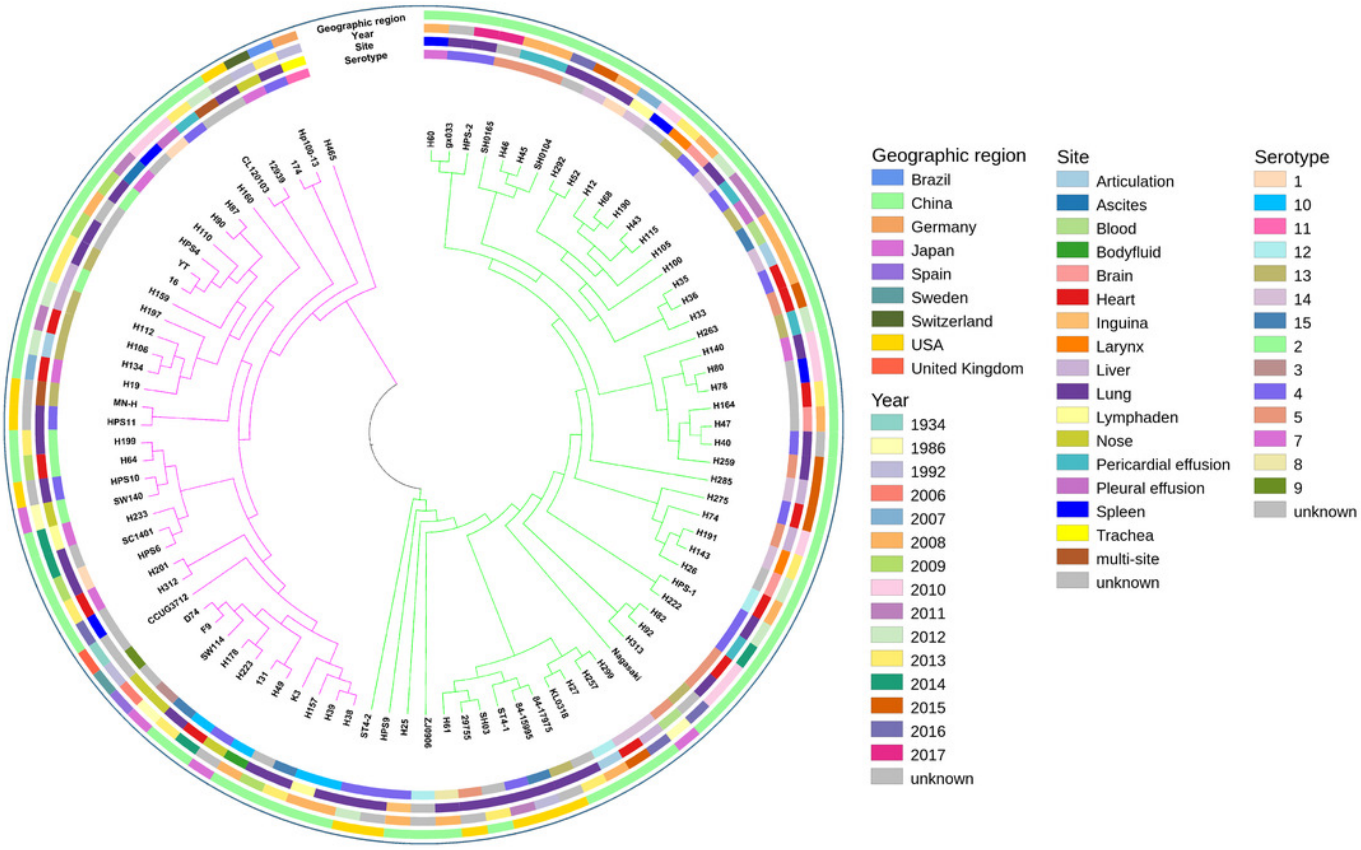


Figure 3

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).

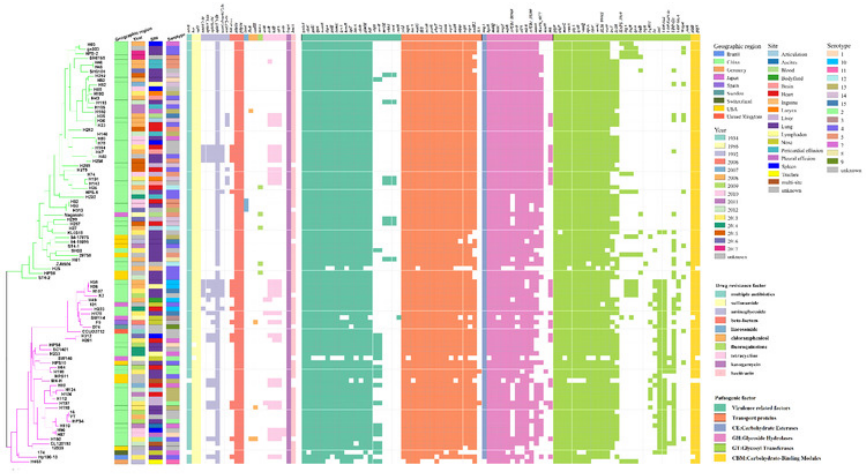


Figure 3. Virulence and resistance profiles across the phylogeny of the 94 *G. parasuis* isolates. Cluster analysis based on single-copy core orthologs. Pattern of gene presence (colored line) or absence (white).

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, −).

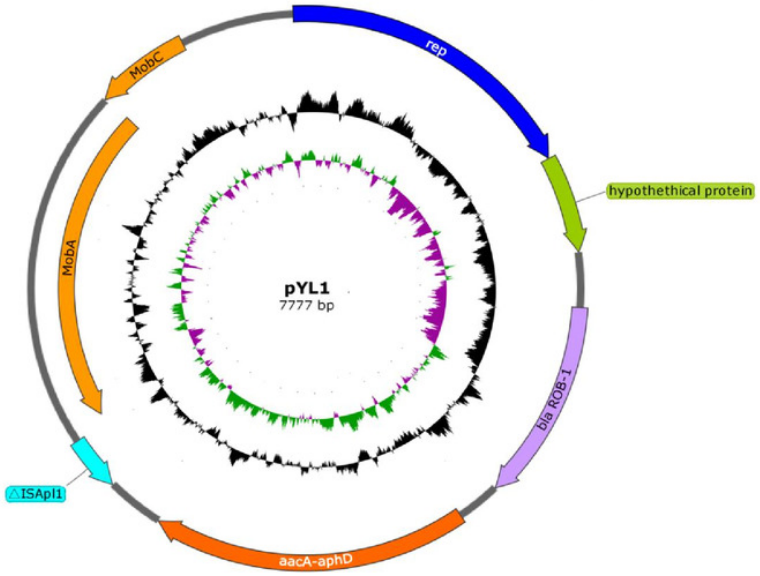


Figure 4. Schematic map of plasmid pYL1. The circles show, from outside to inside: first and second, putative open reading frames, the positions and orientations of the genes; third, G+C content (deviation from the average); and, fourth, G+C skew (green, +; purple, −).

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%).

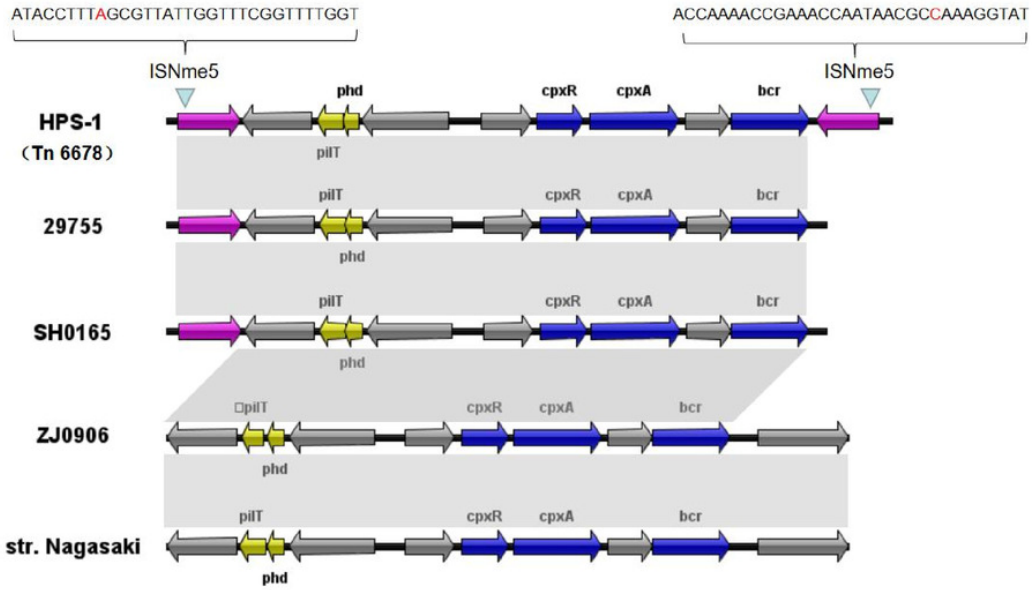


Figure 5. Organization of the *G. parasuis* HPS-1 Tn*6678* transposon and comparison with the similar structure. ORFs are shown as arrows, indicating the transcription direction, and the colors of the arrows represent different fragments. Gene color code: transposase, purple; toxin genes (*pilT* and *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 trees were constructed using MEGA version 7.0 with 1,000 bootstrap replicates. The *G. parasuis* HPS-1 is indicated by a solid circle.

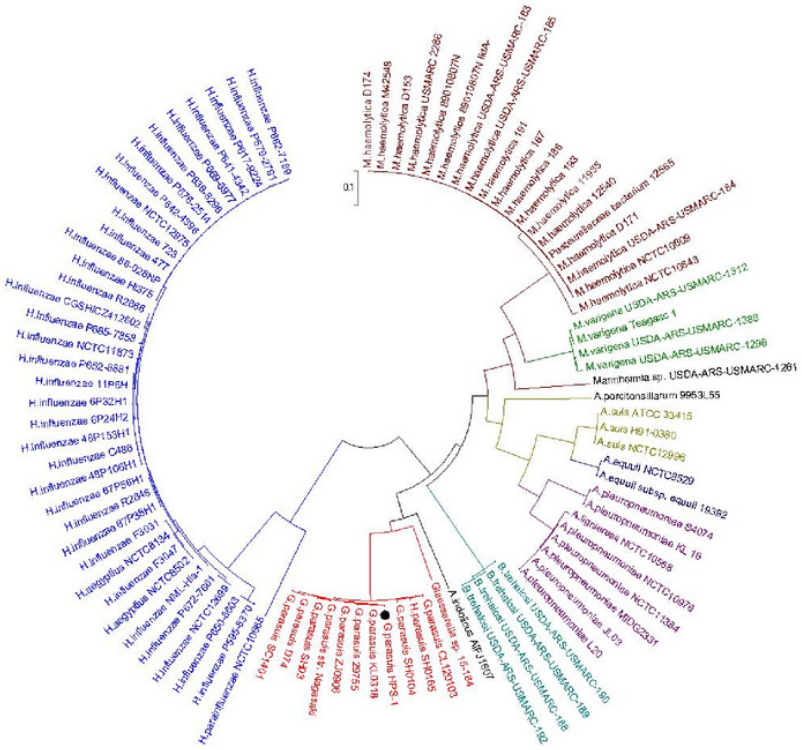


Figure 6. Neighbor-joining phylogenetic tree based on *bcr* gene sequences obtained from the current study and downloaded from NCBI. The trees were constructed using MEGA version 7.0 with 1,000 bootstrap replicates. The *G. parasuis* HPS-1 is indicated by a solid circle.

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 98% nucleotide sequence identity are shaded grey.

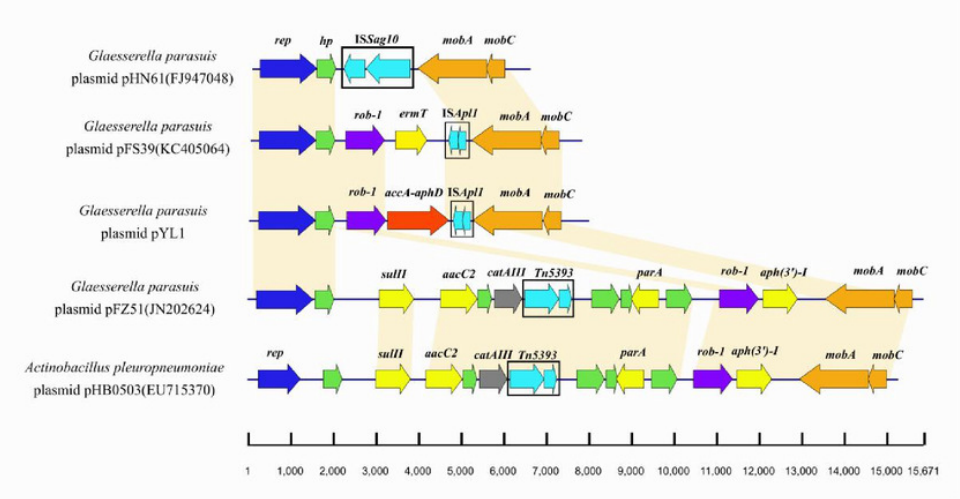


Figure 7. Comparison of the genetic structures of pHN61, pFS39, pYL1, pFZ51 and pHB0503. The accession numbers and origins of these plasmids are displayed on the left side. Arrows represent putative open reading frames, the positions and orientations of the genes. Blue arrows indicate Rep-like protein involved in plasmid replication. Green arrows indicate hypothetical protein. Regions with 98% nucleotide sequence identity are shaded grey.