Research Article

Yongliang Che*, Longbai Wang, Xuemin Wu, Rujing Chen, Chenyan Wang, Lunjiang Zhou*

Characterization of *Haemophilus parasuis* Serovar 2 CL120103, a Moderately Virulent Strain in China

https://doi.org/10.1515/biol-2018-0026
Received September 23, 2017; accepted February 27, 2018

Abstract: *Haemophilus parasuis* is an important bacterium affecting pigs, causing Glässer's disease. To further characterize this species, we determined the complete genomic sequence of *H. parasuis* CL120103, which was isolated from diseased pigs. The strain *H. parasuis* CL120103 was identified as serovar 2. The size of the largest scaffold is 2,326,318 bp and contains 145 large contigs, with the N50 contig being 20,573 bp in length. The complete genome of *H. parasuis* CL120103 is 2,305,354 bp in length with 39.97% GC content and contains 2227 protein-coding genes, 19 ribosomal rRNA operons and 60 tRNA genes. Sequence similarity of the genome of *H. parasuis* CL120103 to the previously sequenced genome of *H. parasuis* was up to 96% and query cover to 86%. Annotation of the genome of *H. parasuis* CL120103 identified a number of genes encoding potential virulence factors. These virulence factors are involved in metabolism, adhesion, secretion and LPS biosynthesis. These related genes pave the way to better understand mechanisms underlying metabolic capabilities. The comprehensive genetic and phylogenetic analysis shows that *H. parasuis* is closely related to *Actinobacillus pleuropneumoniae* and provides a foundation for future experimental confirmation of the virulence and pathogen-host interactions in *H. parasuis*.

Keywords: *Haemophilus parasuis*, genome sequencing, comprehensive genetic analysis, virulence-associated gene, LPS biosynthesis, virulence

***Corresponding authors: Yongliang Che, Lunjiang Zhou,** Institute of Animal Husbandry and Veterinary Medicine, Fujian Academy of Agricultural Sciences, Fuzhou, P. R. China, E-mail:cheyongliang1976@163.com, lunjiang@163.com
Longbai Wang, Xuemin Wu, Rujing Chen, Chenyan Wang, Institute of Animal Husbandry and Veterinary Medicine, Fujian Academy of Agricultural Sciences, Fuzhou, P. R. China

1 Introduction

Haemophilus parasuis (*H. parasuis*), is an important disease-causing agent in pigs. It can cause Glässer's disease and has shown multiple clinical manifestations including severe pant, pneumonia, pleurisy, peritonitis, polyserositis, arthritis, meningitis and septicemia [1-7]. In China, Glässer's disease can cause great economic loss ascribed to half to two-thirds of the fatal cases occurring in finishing pigs who were previously healthy [3, 6, 7]. Moreover, Glässer's disease outbreaks are seriously damaging to pigs, both on their own or when co-infections with other swine pathogens occur in pig-breeding companies, especially large-scale pig-raising enterprises [4, 8]. Previous studies have shown there are 15 serovars of *H. parasuis* with differences in virulence, including highly virulent serovars 1, 5, 10 and 12-14; virulent serovars 2, 4, 8 and 15; and avirulent serovars 3, 6, 7, 9 and 11 [9], and other studies have described the prevalent serovar for controlling infection [10-14]. However, vaccine immunity confers only limited cross-serovar protection [15]. Thus, further characterization to accurately identify serotypes is critical for epidemiological investigations or vaccine selection studies in *H. parasuis* infections. In this study, the field strain CL120103 of *H. parasuis* was identified through bacterial cultivation, morphological observation, PCR analysis of the 16S rRNA gene sequence and biochemical identification of traits. Furthermore, the genome of *H. parasuis* CL120103 was sequenced and compared with serovar 5, with a focus on the investigation of potential virulence factors, antibiotic-resistance genes and pathogen-host interactions in *H. parasuis*.

2 Materials and Methods

2.1 Bacterial strain isolation, identification and DNA purification

H. parasuis strain CL120103 was isolated from heart blood, lungs, ascites, articular fluid, and brain tissue

Open Access. © 2018 Yongliang Che et al., published by De Gruyter. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License.

samples of diseased pigs in Fujian province. The major clinical signs of the diseased pigs were observed as severe pant, pleurisy, peritonitis, arthritis, or meningitis. The bacterium was cultured on tryptic soy agar supplemented with 5 mg/ml nicotinamide adenine dinucleotide (NAD) at 37°C and inverted for 24-72 h before picking typical colonies as observed by Gram staining microscopy and pure culture. The strain was identified by determining the biological characteristics and by 16S rRNA gene sequencing according to the previous descriptions [16]. For amplification of the 16S rRNA gene, forward primer H1, 5′-GTGATGAGGAAGGGTGGTGT-3′ and reverse primer H2, 5′-GCTTCGTCACCCTCTGTA-3′ was used in this study. The reaction conditions were as follows: 5 min at 95°C, 35 cycles of 1 min at 94°C, 1 min at 58°C and 1 min at 72°C, followed by a final extension at 72°C for 10 min. Resulting reactions were subjected to agarose gel electrophoresis. Serotypes of *H. parasuis* strain CL120103 were identified by agar diffusion test via each type of standard serum of each separate strain according to the previous research [9].

Bacterial genomic DNA was extracted and purified with the QIAamp DNA Mini Kit (Qiagen, Germany). The concentration of genomic DNA was measured using a Qubit 2.0 Fluorometer (Thermo Scientific, USA). Purity of the DNA samples (UV A260/A280) was assessed using a NanoDrop 2000 Spectrophotometer (Thermo Scientific, USA). The Quant-iT Picogreen dsDNA kit (Invitrogen, Shanghai, China), Nano-2000 (Thermo Scientific, Waltham, US), and gel electrophoresis were used to evaluate the quality and quantity of genomic DNA.

Ethical approval: The research related to animals use has been complied with all the relevant national regulations and institutional policies for the care and use of animals.

2.2 Genome sequencing, assembly and annotation of H. parasuis CL120103

The genome of strain CL120103 was sequenced using the PacBio RSII platform. A 20-kb DNA library was constructed according to the manufacturer's instructions and PacBio using single-molecule real-time sequencing technology (Pacific Biosciences, Gene Denovo, Guangzhou, China) [17, 18]. The rDNA sequences were publicly accessible and those which have high similarity to various serovars of *H. parasuis* by the BLASTn analysis were chosen [19]. The complete genome of *H. parasuis* CL120103 was added to the NCBI GenBank. The NCBI non-redundant (nr) database was applied to align the amino acid sequences by BLASTp [19].

2.3 Phylogenetic analysis of H. parasuis CL120103

The complete genomes of eight members of the genus *Haemophilus* and four closely related bacteria from other genera were used in the phylogenetic analysis. The accession no. for all species was shown in the front of the description (CP015099.1, CP007715.1, CP011226.1, CP001091.1, CP005384.1, CP006957.1, NC_011852.1, CP009237.1, CP007471.1, CP006955.1, CP000947.1, CL120103, and CP009158.1). The sequenced draft genome of *H. parasuis* serovar 5 strain 29755 (GenBank accession no. NZ_ABKM00000000) was downloaded from NCBI and aligned to the complete genome of *H. parasuis* strain SH03 by using the BLASTN (expected threshold of 1e^{-5} and minimum alignment length of 91%). Orthologous genes were identified by BLASTn suite [20]. A Bayesian phylogenetic tree was reconstructed in the software MEGA 6 [21]. For comparison within the species of *H. parasuis*, reciprocal BLAST was performed according to the previous description and numbers of orthologs shared between them were calculated by in-house Perl scripts [19].

2.4 COG and Pfam analysis of H. parasuis CL120103

The COG annotations were verified by comparing them to the annotations of the COG members in RefSeq databases [22]. The protein domain names in the Pfam database were used to predict protein-coding sequences and protein structure domains (ftp://ftp.sanger.ac.uk/pub/databases/Pfam/Tools/) [23, 24]. Alignment length over 90% of amino acid sequences and over 20% match identity were chosen and the description of the best hit was assigned as the annotation of the predicted gene. All annotated genes were then classified based on the COG database [25] and COG classes. COG-annotated genes and Pfam-annotation of *H. parasuis* strain CL120103 were compared to that of *H. parasuis* strain SH03.

2.5 Virulence factors of pathogenic bacteria (VFDB) analysis of H. parasuis CL120103

Virulence factors database (VFDB) of pathogenic bacteria was used to analyse *H. parasuis* CL120103 using chlamydia and mycoplasma pathogenic factors (http://www.mgc.ac.cn/VFs/main.htm). Also, the VFDB database was applied to align the amino acid sequences by BLASTp, and amino acid sequences with alignment length over 90%

and match over 20% were identified as predicted genes [19].

2.6 Virulence factors involved in pathway analysis of H. parasuis CL120103 and pathogenesis

All sequences of *H. parasuis* CL120103 were translated into amino acids and submitted to the KEGG database for pathway annotation (http://www.genome.jp/kaas-bin/kaas_main) [26]. All VFDB annotated and involved in the pathway were manually downloaded.

3 Results

3.1 Characterization and complete genome sequencing and assembly of *H. parasuis* CL120103

The isolated strain was cultured for 24h on an agar plate containing 0.005% NAD and 5% horse blood [27] and results showed that the bacterial colonies were translucent, moist, smooth and a single small colony was the size of the sample tip (Figure 1A). The strain did not grow on MacConkey agar plates and was Gram-negative (Figure 1B). The PCR results [16] showed that

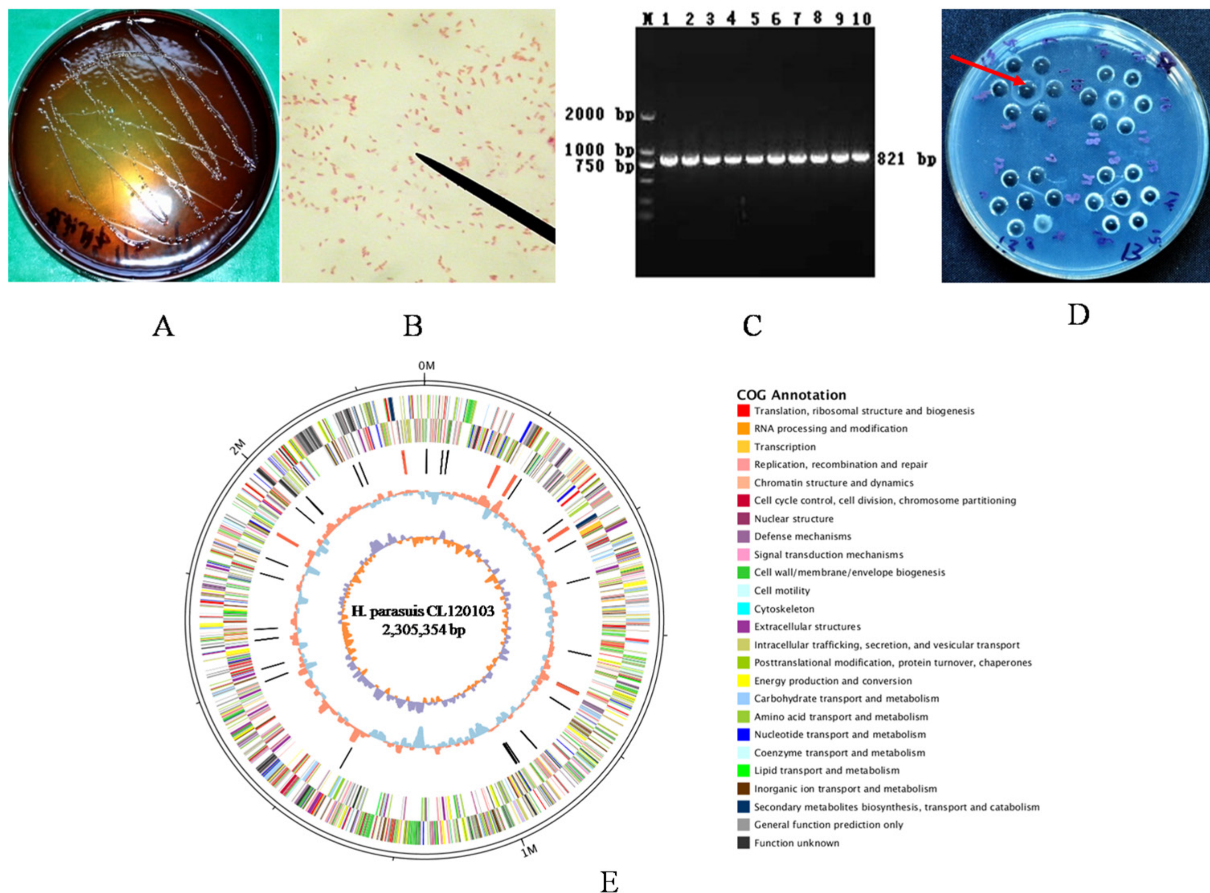

Figure 1. Identification, complete genome sequencing and assembly and circular representation of *H. parasuis* strain CL120103 genome. (A) Colony of *H. parasuis*; (B) Gram staining analysis of *H. parasuis*; (C) PCR appraisal of *H. parasuis*; (D) serotype classification of *H. parasuis* determined by agar gel immunodifusion (AGID; the red arrow means standard serotype 2); (E) circular representation of *H. parasuis* strain CL120103 genome. Circles range from the outermost circle to the innermost circle. The outer two circles show protein-coding genes on the forward and reverse strands in CL120103, colored according to COG categories. All genes are colored based on biological functions and different colors in the COG collection; The third circle shows the coordinates of BLAST hits of the *H. parasuis* CL120103 complete genome; Fourth circle, insertion sequence elements; Fifth circle, tRNA genes; Sixth circle, rRNA operons; Seventh circle, G+C content; Eighth circle, GC skew plot [(G2C)/(G+C)].

the 16S rRNA gene fragment length was 821bp, equal to the designed and expected fragment size (Figure 1C). The agar diffusion test showed that the isolated *H. parasuis* CL120103 was identified as serovar 2 (Figure 1D). The *Haemophilus parasuis* CL120103 genome was sequenced and its complete de novo assembly was achieved by way of overlap using Single Molecule Real Time [18]. A total of 150, 292 reads (481,933,423 bases) and 30,653 paired-end reads (429,578,421 bases) were generated by PacBio RS II sequencing, in which read quality was 99.796% and 99.143%, respectively. The size of the largest scaffold was 2,326,318 bp, which contained 145 large contigs and the N50 contig was 20,573 bp in length, suggesting that this raw assembly is highly continuous [7]. The complete genome of *H. parasuis* CL120103 was 2,305,354 bp in length with GC content of 39.97% (Figure 1E).

3.2 Genome annotation of *H. parasuis* CL120103

Based on the 2,227 predicted genes of *H. parasuis* CL120103, the 2,133 CDS were annotated by BLAST search from the NCBI non-redundant database (File S1). Sixty tRNA genes and 19 rRNA genes were found in the genome of *H. parasuis* CL120103. The same number of rRNA genes was found in the genome of strain CL120103, strain SH03, strain 19392 and strain SC1401, and the full annotation of repetitive sequences is attached as File S2a, S2b and S2c. General features of the whole genomes of *H. parasuis* strains are shown in Table 1.

3.3 Phylogenetic analysis of *H. parasuis* CL120103

Phylogenetic analysis showed that *H. parasuis* CL120103 shares the closest evolutionary origin to strain SH03, as expected (Figure 2). Interestingly, with numbers of orthologs (Table 2), *Actinobacillus pleuropneumoniae* serovar 7 strain AP76 had a similar evolutionary relationship to *H. parasuis* strain SC1401. *H. parasuis* CL120103 was illustrated by COG-annotated class distribution and the top COG classes are shown in Figure 3. The majority of the genes were involved in basic cellular functions, such as general function prediction

Table 1 General features of whole genomes of *H. parasuis* (CL120103)

Genbank accession No.	None	CP009158.1	CP011226.1	CP007715.1	CP009237.1	CP015099.1
Strain	CL120103	SH03	GU5	19392	KL0138	SC1401
Total length (bp)	2305354 bp	2265927 bp	1624874 bp	2431533 bp	2280878 bp	2277540 bp
Genes	2227	2148	1574	2264	2175	2180
CDS	2133	1979	1422	2117	1997	2098
Pseudo Genes	54	90	101	11	101	101
Ribosome RNA	19	19	0	19	20	19
Number of tRNA	60	59	50	62	56	59
Frameshifted Gene	35	78	69	7	78	0

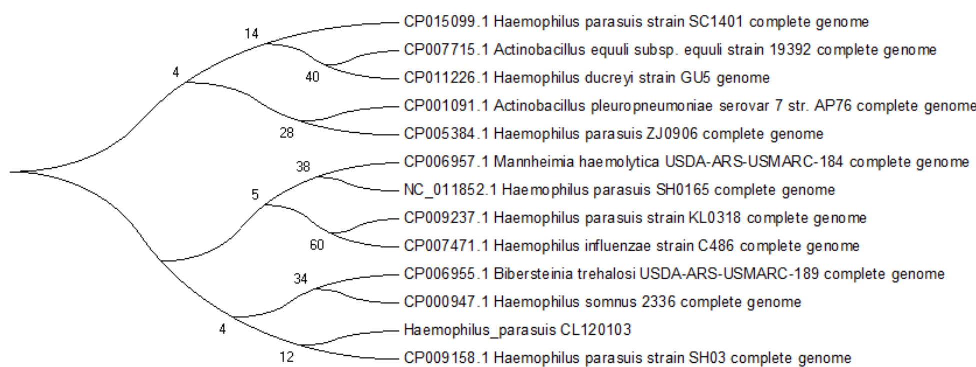

Figure 2. Bayesian phylogenetic tree of *H. parasuis* strain CL120103 and other closely related bacteria. Phylogenetic tree was reconstructed using MEGA 6 (Tamura et al., 2013).

Table 2. Orthologs of predicted CDSs of *H. parasuis* CL120103 compared with complete genomes of related organisms.

	Name	No. of orthologs	% of CDS in H. parasuis
1	Homologous to H. parasuis strain SH03	1462	68.5%
2	Homologous to H. somnus strain 2336	1236	57.9%
3	Homologous to B. trehalosi USDA-ARS-USMARC-189	1226	57.5%
4	Homologous to H. influenzae strain C486	1205	56.5%
5	Homologous to H. parasuis strain KL0318	1143	53.5%
6	Homologous to H. parasuis strain SH0165	1133	53.1%
7	Homologous to M. haemolytica USDA-ARS-USMARC-184	1097	51.4%
8	Homologous to H. parasuis strain ZJ0906	1065	49.9%
9	Homologous to A. pleuropneumoniae serovar 7 AP76	1061	49.7%
10	Homologous to H. ducreyi strain GU5	1054	49.4%
11	Homologous to A. equuli subsp. equuli strain 19392	1022	47.9%
12	Homologous to H. parasuis strain SC1401	1007	47.2%

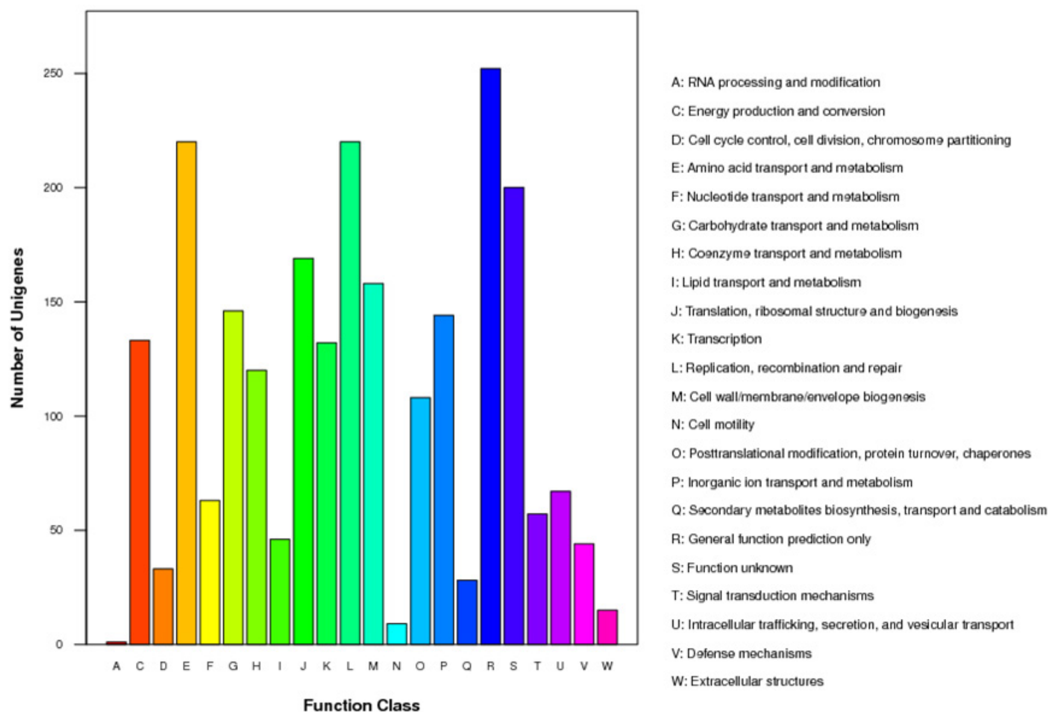

Figure 3. COG class distribution of *H. parasuis* CL120103 genome. The COG-annotated genes are grouped under their respective COG classes. Percentages of the top ten classes are labeled for easy reference and presentation of the gene functional-categories.

only, replication, recombination and repair, amino acid transport and metabolism, and energy production and conversion. However, 8.46% of the genes have unknown functions in the COG database. For the full COG functional annotation, refer to File S3.

3.4 Virulence factors of pathogenic bacteria (VFDB) analysis of *H. parasuis* CL120103

As shown in Table 3, virulence-factor annotated genes in strain CL120103 were identified, and a list of potential virulence factors was compiled, which included gene clusters such as the peptidoglycan-binding protein

LysM, 3-ketoacyl-ACP reductase, opacity-associated protein (*OapA*), opacity-associated A LysM-like domain protein, GlcNAc transferase, capsular polysaccharide biosynthesis protein, protein WbjB, and Fnl, the iron(III) ABC transporter ATP-binding protein, glycosyl transferase 2 protein, *FbpC*, ADP-L-glycero-D-manno-heptose-6-epimerase (*ADP-LgDmh6e*), UDP-glucose 4-epimerase (*UDP-g4e*), polysaccharide biosynthesis family protein

Table 3. Identification of the potential virulence factors involved in adhesion, secretion and lipopolysaccharide (LPS) biosynthesis in the *H. parasuis* CL120103 genome

Sequenced No.	Gene	Functional description
LPS biosynthesis		
*H.parasuis*_org000375	*GmhD*	ADP-L-glycero-D-manno-heptose-6-epimerase
H.parasuis _org001728	*GmhC*	heptose 1-phosphate adenyltransferase; bifunctional protein RfaE, domain I; bifunctional heptose 7-phosphate kinase/ heptose 1-phosphate adenyltransferase
Adhesion and secretion		
*H.parasuis*_orf000018 *H.parasuis*_orf000080 *H.parasuis*_orf001940 *H.parasuis*_orf001745 *H.parasuis*_orf001750	*aidA*	pertactin family Virulence factor, outer membrane autotransporter/type V secretion pathway, adhesin AidA; IgA1 protease, Type V secretory pathway, adhesin AidA/ outer membrane autotransporter barrel domain protein/
*H.parasuis*_orf000066	*Rlp*	rare lipoA family protein
*H.parasuis*_orf001352 *H.parasuis*_orf002200	*secA*	preprotein translocase subunit SecA
*H.parasuis*_orf000468	*secB*	preprotein translocase subunit SecB
*H.parasuis*_orf000285	*secE*	preprotein translocase subunit SecE
*H.parasuis*_orf000603	*secY*	preprotein translocase subunit SecY
*H.parasuis*_orf001367	*lolA*	outer membrane lipocarrier protein LolA
*H.parasuis*_orf000143 *H.parasuis*_orf001118	*lolB*	membrane protein/ outer membrane lipoprotein LolB
*H.parasuis*_orf000637	*IspA*	signal peptidase II/ lipoprotein signal peptidase
*H.parasuis*_orf000663	*ftsY*	cell division protein FtsY
*H.parasuis*_orf001041	*pulG*	Type II secretory pathway, pseudopilin PulG
*H.parasuis*_orf001355	*nlpE*	lipoprotein copper homeostasis and adhesion, NlpE
*H.parasuis*_orf001452	*yajC*	preprotein translocase subunit YajC
*H.parasuis*_orf001453	*sirA*	sirA-like family protein
*H.parasuis*_orf001539	*secG*	preprotein translocase subunit SecG
*H.parasuis*_orf001857	*yajC*	preprotein translocase YajC subunit
*H.parasuis*_orf001858	*SecD*	preprotein translocase subunit SecD
*H.parasuis*_orf001859	*SecF*	protein-export membrane protein SecF
*H.parasuis*_orf001928	*pilB/hofB*	Tfp pilus assembly pathway, ATPase PilB/protein transporter HofB
*H.parasuis*_orf000937	*pilW*	type IV pilus biogenesis/stability protein PilW
*H.parasuis*_orf001929	*fimB*	fimbrial protein/type II secretion system protein F
*H.parasuis*_orf001930	*pilD*	type IV leader peptidase family protein/Tfp pilus assembly pathway, fimbrial leader peptidase PilD
*H.parasuis*_orf002180	*yidC*	inner-membrane protein insertion factor/membrane protein insertase/ protein translocase component YidC
*H.parasuis*_orf002181	*tatA*	protein translocase TatA
*H.parasuis*_orf002182	*tatB*	preprotein translocase subunit TatB/twin arginine-targeting protein translocase TatB
*H.parasuis*_orf002183	*tatC*	preprotein translocase subunit TatC

(*pbf*), and lacto-N-neotetraose biosynthesis glycosyl transferase (*LgtA*) (File S4). These genes are involved in bacterial metabolism, adhesion, secretion and surface lipopolysaccharide (LPS) biosynthesis (File S5). Despite the increased complexity of virulence in the mature bacterium, this work has shown that the list of virulence factors found in the genome of *H. parasuis* CL120103 was compared with the other serovar strains SH03, KL0138 and SC1401. A list of genes coding for enzymes involved in the pentose phosphate pathway (PPP) was identified, including heptose 1-phosphate adenyltransferase, *RfaE I*, *prsA*, *rbsK*, *gntK*, *rpe-3*, *rpiA*, *talB*, *pgi*, *fba*, *pfkA*, *gnd*, *devB*, *zwf*, *tktA*, *deoC*, and *fbp* (Figure 4A, File S6). In this study, we also identified the sedoheptulose-7-phosphate isomerase (*GmhA*, *H. parasuis*_orf001260 and *H. parasuis*_orf001739), which could increase the sedoheptulose-7-phosphate content to synthesize LPS (Figure 4B and File S1). As shown in Figure 4C, ATP-dependent Clp protease subunit B (*ClpB*) is an important component in the complex cellular process of CtrA degradation entry into S phase. Furthermore, the identified genes were also involved in energy metabolism, glucose metabolism, DNA damage repair, autophagy and protection of proteins by chaperones according to the results of the KEGG analysis (File S7). The chaperones included *GroEL* (molecular chaperone belonging to heat-shock protein system), *MutS* (DNA mismatch repair) and lysine-tRNA ligase, and might increase the virulence of *H. parasuis* CL120103, which should be identified in future research.

4 Discussion

In the present study, genomic analysis showed that the genome sequence is 2,305,354 bp in length with GC content of 39.97%, and contains 145 large contigs and the N50 contig. Sequence similarity to the SH0165 has been identified and results showed that the sequence similarity was up to 96% and query cover to 86% [6]. In the *H. parasuis* CL120103 genome, 60 tRNA genes and 19 rRNA genes were found and associated with prophage and elements targeted to tRNA-Arg-4 (File S2b). According to previous studies, phage elements as a strategy of horizontal gene transfer might play a potentially important role in genetic diversity and virulence variations in many bacteria [10, 28, 29]. However, the inheritance differences and virulence were closely related to the evolution of the bacteria. In the present study, evolutionary relationships showed that *H. parasuis* CL120103 displays an evolutionary relationship to *A. pleuropneumoniae* serovar 7 strain AP76, which was in accordance with previous findings for *H. parasuis* by Xu's group [6]. Similarly, *A. pleuropneumoniae* is an important etiological agent in pig disease, pleuropneumonia, which

Figure 4. Virulence factors of pathogenic bacteria (VFDB) analysis of *H. parasuis* strain CL120103 genome. (A) VFDB analysis involved in the pentose phosphate pathway; (B) VFDB analysis involved in lipopolysaccharides biosynthesis; (C) VFDB analysis involved in *H. parasuis* maturation. In the figure, red color box shows the identified gene based on the *H. parasuis* strain CL120103 genome.

is transmitted via direct contact or airborne route. This characteristic of *H. parasuis* and *A. pleuropneumoniae* may be partially explained by their common habitat in the upper respiratory tract of pigs.

With regard to *H. parasuis*, the facultative anaerobe possessed metabolic pathways of both fermentation and respiration for energy generation, and carbon source utilization was important to produce energy [5, 30, 31]. In this study, virulence factors genes coding for adhesins or invasins may be located on transmissible genetic elements such as transposons [32], outer membrane protein P5 and outer membrane protein A, etc. [11, 33, 34]. However, beyond that, we have identified one of the two sugar transport systems in the *H. parasuis* CL120103 genome encoding ATP-binding cassette (ABC) transport complexes involved in the utilization of sugars [35]. ABC transport complexes comprise the largest protein transporter super-family in all organisms. This family of genes codes for different proteins that transport molecules such as amino acids, proteins, ions, sugars, cholesterol, peptides, metabolites and toxins across extra- and intracellular membranes [35-39]. A previous study indicated that all ABC transporters contain two domains, the nucleotide-binding domain (NBD) and the transmembrane domain (TMDs) [40, 41]. The two domains are roughly divided into two functional areas, which appear to specialize in handling various tasks, for example, the NBD catalyzes ATP hydrolysis and the TMDs translocates substances to use energy through the membrane by conformational changes [42]. These confirmed genotypes support the previously observed biochemical patterns of carbon source utilization in *H. parasuis* [5, 30]. As with previous studies, fructose as an alternative start point to glycolysis was controlled by *Rbsk2*, which can catalyze the phosphorylation of fructose to fructose-6-phosphate [43]. Furthermore, based on PPP, sedoheptulose was produced and used in the crucial components of the biosynthesis process of the LPS, lipooligosaccharides (LOS), capsules, O-antigens, and glycan moieties of bacterial cell surface (S-layer) glycoproteins [44]. In this study, we also identified the sedoheptulose-7-phosphate isomerase (*GmhA*, H. parasuis_orf001260 and _orf001739), which is the first biosynthesis step of the L,D-heptose component of the LPS and responsible for catalyzing isomerization of the D-sedoheptulose 7-phosphate into D-glycero-α, β-D-manno-heptose-7-phosphate, and leading to generation of GDP-D-glycero-α-D-*manno*-heptose and ADP-L-glycero-β-D-*manno*-heptose [44]. Thus, we speculated that the PPP plays a key role in the virulence of *H. parasuis*.

Furthermore, the identified gene was involved in the biosynthesis process of surface lipopolysaccharide, which is generally the initiation step on the bacterial surface for bacterial infection or bacterial adhesion on cells or substrates. Previous studies also reported that LPS was an important functional component of the Gram-negative bacterial outer membrane that can mediate bacterial adhesion on substrates/cells [45, 46]. As previously reported, LPSs includes three covalently linked biochemical moieties: the core oligosaccharide, the O-polysaccharide and the lipid A [6, 11, 33, 46]. The O-polysaccharide plays a vital role in bacterial adherence, invasion and immune evasion [46]. Furthermore, a gene cluster pilABCDW coding for type IV leader peptidase/fimbrial family protein has been identified in a number of Gram-negative pathogens of the genera *Haemophilus*, *Vibrio*, *Actinobacillus* and others[47]. Also identified in the genome of *H. parasuis* CL120103 was the type IV fimbrial genes encoding the major structural unit *pilB* (*H. parasuis*_orf001928) and biogenesis/stability protein *pilD* (*H. parasuis*_orf001930) and *pilDW* (*H. parasuis*_orf001930) for mediating bacterial adherence.

In conclusion, understanding the function of the complete genome of *H. parasuis* strain CL120103 will facilitate the development of safe and effective vaccines via approaches focused on genomic analysis to prevent and control swine disease. In particular, our work demonstrated the crucial function of virulence factors in metabolism, adhesion, secretion and LPS synthesis. These findings underscore the significance of VFDB as a target for therapeutics. A putative communication with the pentose phosphate pathway and virulence involved in LPS synthesis is proposed and requires further experimental confirmation in *H. parasuis*.

Acknowledgments: This work was supported by the provincial research institutes of special public-funded projects on epidemiology and diagnosis of *Haemophilus parasuis* (No. 2010R1025-2), the science and technology innovation team of Fujian Academy of Agricultural Sciences project on the development and application of *Haemophilus parasuis* vaccine (No.CXTD-1-1306) and the Fujian province modern agricultural industry technology system of pig.

Conflict of interest: Authors state no conflict of interest.

References

[1] Smart NL, Miniats OP, Rosendal S, Friendship RM. Glasser's disease and prevalence of subclinical infection with Haemophilus parasuis in swine in southern Ontario. The

Canadian veterinary journal = La revue veterinaire canadienne 1989; 30(4): 339-43.
[2] Aragon V, Cerda-Cuellar M, Fraile L, *et al.* Correlation between clinico-pathological outcome and typing of Haemophilus parasuis field strains. Veterinary microbiology 2010; 142(3-4): 387-93.
[3] Oliveira S,Pijoan C. Haemophilus parasuis: new trends on diagnosis, epidemiology and control. Veterinary microbiology 2004; 99(1): 1-12.
[4] Palzer A, Haedke K, Heinritzi K, *et al.* Associations among Haemophilus parasuis, Mycoplasma hyorhinis, and porcine reproductive and respiratory syndrome virus infections in pigs with polyserositis. The Canadian veterinary journal = La revue veterinaire canadienne 2015; 56(3): 285-7.
[5] Norskov-Lauritsen N, Bruun B,Kilian M. Multilocus sequence phylogenetic study of the genus Haemophilus with description of Haemophilus pittmaniae sp. nov. International journal of systematic and evolutionary microbiology 2005; 55(Pt 1): 449-56.
[6] Xu Z, Yue M, Zhou R, *et al.* Genomic characterization of Haemophilus parasuis SH0165, a highly virulent strain of serovar 5 prevalent in China. PloS one 2011; 6(5): e19631.
[7] Li Y, Kwok AH, Jiang J, *et al.* Complete genome analysis of a Haemophilus parasuis serovar 12 strain from China. PloS one 2013; 8(9): e68350.
[8] Yu J, Wu J, Zhang Y, *et al.* Concurrent highly pathogenic porcine reproductive and respiratory syndrome virus infection accelerates Haemophilus parasuis infection in conventional pigs. Veterinary microbiology 2012; 158(3-4): 316-21.
[9] Kielstein P, Rapp-Gabrielson VJ. Designation of 15 serovars of Haemophilus parasuis on the basis of immunodiffusion using heat-stable antigen extracts. Journal of clinical microbiology 1992; 30(4): 862-5.
[10] Hu M, Zhang Y, Xie F, *et al.* Protection of piglets by a Haemophilus parasuis ghost vaccine against homologous challenge. Clinical and vaccine immunology : CVI 2013; 20(6): 795-802.
[11] Brockmeier SL, Loving CL, Mullins MA, *et al.* Virulence, transmission, and heterologous protection of four isolates of Haemophilus parasuis. Clinical and vaccine immunology : CVI 2013; 20(9): 1466-72.
[12] Olvera A, Pina S, Perez-Simo M, *et al.* Immunogenicity and protection against Haemophilus parasuis infection after vaccination with recombinant virulence associated trimeric autotransporters (VtaA). Vaccine 2011; 29(15): 2797-802.
[13] Bak H, Riising HJ. Protection of vaccinated pigs against experimental infections with homologous and heterologous Haemophilus parasuis. The Veterinary record 2002; 151(17): 502-5.
[14] Takahashi K, Naga S, Yagihashi T, *et al.* A cross-protection experiment in pigs vaccinated with Haemophilus parasuis serovars 2 and 5 bacterins, and evaluation of a bivalent vaccine under laboratory and field conditions. The Journal of veterinary medical science 2001; 63(5): 487-91.
[15] Miniats OP, Smart NL, Rosendal S. Cross protection among Haemophilus parasuis strains in immunized gnotobiotic pigs. Canadian journal of veterinary research = Revue canadienne de recherche veterinaire 1991; 55(1): 37-41.
[16] Cai X, Chen H, Blackall PJ, *et al.* Serological characterization of Haemophilus parasuis isolates from China. Veterinary microbiology 2005; 111(3-4): 231-6.
[17] Berlin K, Koren S, Chin CS, *et al.* Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. Nature biotechnology 2015; 33(6): 623-30.
[18] Chin CS, Alexander DH, Marks P, *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature methods 2013; 10(6): 563-9.
[19] Altschul SF, Madden TL, Schaffer AA, *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research 1997; 25(17): 3389-402.
[20] Kent WJ. BLAT--the BLAST-like alignment tool. Genome research 2002; 12(4): 656-64.
[21] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular biology and evolution 2013; 30(12): 2725-9.
[22] Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic acids research 2015; 43(7): 3872.
[23] Finn RD, Coggill P, Eberhardt RY, *et al.* The Pfam protein families database: towards a more sustainable future. Nucleic acids research 2016; 44(D1): D279-85.
[24] Finn RD, Bateman A, Clements J, *et al.* Pfam: the protein families database. Nucleic acids research 2014; 42(Database issue): D222-30.
[25] Galperin MY, Makarova KS, Wolf YI, Koonin EV. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic acids research 2015; 43(Database issue): D261-9.
[26] Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research 2012; 40(Database issue): D109-14.
[27] Biberstein EL, White DC. A proposal for the establishment of two new Haemophilus species. Journal of medical microbiology 1969; 2(1): 75-8.
[28] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic acids research 1997; 25(5): 955-64.
[29] Lagesen K, Hallin P, Rodland EA, *et al.* RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic acids research 2007; 35(9): 3100-8.
[30] Frickmann H, Podbielski A, Essig A, Schwarz NG, Zautner AE. Difficulties in species identification within the genus Haemophilus - A pilot study addressing a significant problem for routine diagnostics. European journal of microbiology & immunology 2014; 4(2): 99-105.
[31] Janovska D, Vymola F. [Taxonomy of the genus Haemophilus]. Ceskoslovenska epidemiologie, mikrobiologie, imunologie 1987; 36(2): 104-9.
[32] Hacker J, Blum-Oehler G, Muhldorfer I, Tschape H. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Molecular microbiology 1997; 23(6): 1089-97.
[33] Olvera A, Pina S, Perez-Simo M, Oliveira S, Bensaid A. Virulence-associated trimeric autotransporters of Haemophilus parasuis are antigenic proteins expressed in vivo. Veterinary research 2010; 41(3): 26.

[34] Zhou M, Guo Y, Zhao J, *et al.* Identification and characterization of novel immunogenic outer membrane proteins of Haemophilus parasuis serovar 5. Vaccine 2009; 27(38): 5271-7.
[35] Broehan G, Kroeger T, Lorenzen M, Merzendorfer H. Functional analysis of the ATP-binding cassette (ABC) transporter gene family of Tribolium castaneum. BMC genomics 2013; 14: 6.
[36] Benadiba M, Maor Y. Importance of ABC Transporters in Drug Development. Current pharmaceutical design 2016; 22(38): 5817-29.
[37] Videira M, Reis RL, Brito MA. Deconstructing breast cancer cell biology and the mechanisms of multidrug resistance. Biochimica et biophysica acta 2014; 1846(2): 312-25.
[38] Wu F, Shao ZY, Zhai BJ, Zhao CL, Shen DM. Ultrasound reverses multidrug resistance in human cancer cells by altering gene expression of ABC transporter proteins and Bax protein. Ultrasound in medicine & biology 2011; 37(1): 151-9.
[39] Auner V, Sehouli J, Oskay-Oezcelik G, *et al.* ABC transporter gene expression in benign and malignant ovarian tissue. Gynecologic oncology 2010; 117(2): 198-201.
[40] ter Beek J, Guskov A, Slotboom DJ. Structural diversity of ABC transporters. The Journal of general physiology 2014; 143(4): 419-35.
[41] Wilkens S. Structure and mechanism of ABC transporters. F1000prime reports 2015; 7: 14.
[42] Silva R, Vilas-Boas V, Carmo H, *et al.* Modulation of P-glycoprotein efflux pump: induction and activation as a therapeutic strategy. Pharmacology & therapeutics 2015; 149: 1-123.
[43] Jahreis K, Bentler L, Bockmann J, *et al.* Adaptation of sucrose metabolism in the Escherichia coli wild-type strain EC3132. Journal of bacteriology 2002; 184(19): 5307-16.
[44] Valvano MA, Messner P, Kosma P. Novel pathways for biosynthesis of nucleotide-activated glycero-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides. Microbiology 2002; 148(Pt 7): 1979-89.
[45] Abu-Lail NI, Camesano TA. Role of lipopolysaccharides in the adhesion, retention, and transport of Escherichia coli JM109. Environmental science & technology 2003; 37(10): 2173-83.
[46] Strauss J, Burnham NA, Camesano TA. Atomic force microscopy study of the role of LPS O-antigen on adhesion of E. coli. Journal of molecular recognition : JMR 2009; 22(5): 347-55.
[47] Mattick JS. Type IV pili and twitching motility. Annual review of microbiology 2002; 56: 289-314.

Supplemental Material: The online version of this article (DOI: 10.1515/biol-2018-0026) offers supplementary material.