Association between *iscR*-based phylogeny, serovars and potential virulence markers of *Haemophilus parasuis*

Junxing Li, Lihua Xu, Fei Su, Bin Yu and Xiufang Yuan

Zhejiang Academy of Agricultural Sciences, Institute of Animal Husbandry and Veterinary Medicine, Hangzhou, Zhejiang, China

ABSTRACT

Haemophilus parasuis is an economically important bacterial pathogen of swine. Extensive genetic and phenotypic heterogeneity among *H. parasuis* strains have been observed, which hinders the deciphering of the population structure and its association with clinical virulence. In this study, two highly divergent clades were defined according to iron–sulphur cluster regulator (*iscR*)-based phylogeny analysis of 148 isolates. Clear separation of serovars and potential virulence markers (PVMs) were observed between the two clades, which are indicative of independent evolution of the two lineages. Previously suggested virulence factors showed no correlation with clinical virulence, and were probably clade or serovar specific genes emerged during different stage of evolution. PVMs profiles varied widely among isolates in the same serovar. Higher strain diversity in respect of PVMs was found for isolates from multi-strain infected farms than those from single strain infected ones, which indicates that multi-strain infection in one farm may increase the frequency of gene transfer in *H. parasuis*. Systemic isolates were more frequently found in serovar 13 and serovar 12, while no correlation between clinical virulence and *iscR*-based phylogeny was observed. It shows that *iscR* is a reliable marker for studying population structure of *H. parasuis*, while other factors should be included to avoid the interference of gene exchange of *iscR* between isolates. The two lineages of *H. parasuis* may have undergone independent evolution, but show no difference in clinical virulence. Wide distribution of systemic isolates across the entire population poses new challenge for development of vaccine with better cross-protection. Our study provides new information for better deciphering the population structure of *H. parasuis*, which helps understanding the extreme diversity within this pathogenic bacterium.

Subjects Microbiology, Veterinary Medicine, Epidemiology
Keywords Serovar, *iscR*, Population structure, *Haemophilus parasuis*, Virulence

Submitted 4 January 2019
Accepted 9 April 2019
Published 14 May 2019

Corresponding author
Xiufang Yuan, yuanxf@zaas.ac.cn

Academic editor
Paul Tulkens

Additional Information and Declarations can be found on page 16

DOI **10.7717/peerj.6950**

INTRODUCTION

Haemophilus parasuis is a member of the family *Pasteurellaceae*, and is the causative agent of Glässer's disease in swine. *H. parasuis* could be a commensal organism of the upper respiratory tract of healthy pigs or pathogens causing systemic infection (*Biberstein, Gunnarsson & Hurvell, 1977*). Extensive heterogeneity among *H. parasuis* strains have been observed, and increasing efforts are made in deciphering the population structure and

its possible relevance with clinical outcome (*Oliveira, Blackall & Pijoan, 2003*; *Howell et al., 2014*; *Turni, Singh & Blackall, 2018*; *Zhao et al., 2018*).

Strain classification of *H. parasuis* has been widely studied serotypically, and genotypically, since differentiation of non-virulent strains from virulent strains is essential for diagnosis and control of the disease (*Kielstein & Rapp-Gabrielson, 1992*; *Turni & Blackall, 2005*; *Olvera, Calsamiglia & Aragon, 2006a*; *Moreno et al., 2016*).

Serotyping is the most commonly used subtyping method, and is traditionally considered to be associated with virulence, but increasing evidence indicated that serovar is a poor proxy for virulence (*Aragon et al., 2010*; *Brockmeier et al., 2014*; *Lawrence et al., 2014*). Sequencing studies of *hsp60* and *16s rRNA* gene revealed two clades of strains in *H. parasuis* and the presence of a separate lineage of virulent strains (*Olvera, Calsamiglia & Aragon, 2006a*; *Angen et al., 2007*). Population structure analysis at genome level revealed the separation of *H. parasuis* into two clades, which were further divided into five Bayesian analysis of population structure sub-clade, but no influence of disease association on population structure was observed (*Howell et al., 2014*). Hence, strains in certain sub-type tend to be virulent, but the relationship between sub-type and virulence phenotype is still not clear.

An increasing number of potential virulence markers (PVMs) have been identified for *H. parasuis*, of which some were already used for development of molecular methods for virulence prediction (*Sack & Baltes, 2009*; *Zhou et al., 2010*; *Wang et al., 2011*; *Yu et al., 2014*; *Galofre-Mila et al., 2017*; *Howell et al., 2017*). As clinical data on these PVMs are still limited, its clinical relevance needs to be further confirmed.

Iron–sulfur clusters are essential protein cofactors that participate in numerous biological processes in most organisms (*Santos, Pereira & Macedo-Ribeiro, 2015*). Iron–sulphur cluster regulator (*iscR*), a member of the Rrf2 family transcriptional regulators, is a global transcription regulator involving in various physiological processes during bacteria growth and stress responses. Growing evidence shows that *iscR* governs the proper regulation of virulence factors in many bacterial pathogens such as type III secretion system of *Yersinia pseudotuberculosis*, capsular polysaccharide biosynthesis of *Klebsiella pneumonia*, antigen I fimbriae of enterotoxigenic *Escherichia coli* and catalase A activity of *Pseudomonas aeruginosa* (*Haines et al., 2015*; *Miller & Auerbuch, 2015*; *Santos, Pereira & Macedo-Ribeiro, 2015*). *IscR* deficient mutant displays dramatic attenuation of virulence in several pathogens, and the important role of *iscR* in bacterial pathogenesis is becoming increasingly apparent (*Crack et al., 2012*; *Miller & Auerbuch, 2015*).

The *iscR* gene was found in reference strains of all 15 serovars, but little is known about the role of *iscR* in pathogenicity of *H. parasuis* (*Yue et al., 2009*; *Howell et al., 2013*). *IscR* is a highly conserved protein among Proteobacteria and particularly within a certain species, but apparent amino acid substitutions were observed among 15 reference strains of *H. parasuis* (*Howell et al., 2013*; *Mettert, Perna & Kiley, 2014*; *Santos, Pereira & Macedo-Ribeiro, 2015*). As *iscR* is a global regulator of virulence factors in many pathogenic bacteria, it is essential to evaluate the possible impact of the amino acid substitutions in *iscR* protein on population structure and clinical virulence of *H. parasuis*. In this study, population structure of *H. parasuis* was analyzed based on the phylogeny,

serovars and PVMs profiles of 148 isolates. Meanwhile, potential correlation among population structure and clinical phenotype was explored.

MATERIALS AND METHODS

Isolates and culture conditions

A total of 148 *H. parasuis* strains were used in this study, of which 145 strains were isolated from clinical samples sent to our diagnostic center from September 2009 to January 2017, and the other three strains were isolated from nasal cavity (Dataset S1). The isolates were originated from 125 independent pig farms in three provinces of eastern China. A sterile loop was used to inoculate TSA plates (supplemented with 2.5% bovine serum and 0.05% NAD) from samples of nasal cavity, lung, spleen, joint, brain, pericardium liquid, pleural effusion, and seroperitoneum. The plates were kept in a 37 °C incubator under humidified 5% CO_2 for 24–72 h until the colony appears. Unequal number of colonies (from four to seven) was picked based on the number of samples, isolation sites or morphology of the colony for each independent farm, and each selected colony was inoculated on another TSA plate for further purification. Then, the isolates were transferred from TSA plates to TSB broth (supplemented with 2.5% bovine serum and 0.05% NAD) for bacteria enrichment.

Serovar determination

Genome DNA of *H. parasuis* was extracted by digestion with proteinase K, and the isolates were identified by a species specific PCR test (*Angen et al., 2007*). Serovars of the isolates were determined by the two newly developed molecular serotyping methods, which were faster, more sensitive, and more specific than the traditional IHA (*Howell et al., 2015*; *Jia et al., 2017*). Using of the two molecular serotyping methods in combination allows the differentiation of all the 15 serovars. If all the picked colonies from one independent farm were determined to be the same serovar, then one of colonies was selected to be the representative. If more than one serovar were detected in a farm, then one representative for each different serovar was selected for further analysis. The designation of each clinical isolate includes information of province (ZJ for Zhejiang province), year (09 for year of 2009), sequence number (01 for the first isolates of the year) of isolation, and serovar (S1 for serovar 1 etc., Su for unknown serovar) of the isolate.

IscR sequencing and analysis

Primers for amplification of complete open reading frame of *iscR* gene were designed based on the *iscR* gene sequence of the strain SH0165 (GenBank accession number: CP001321) (*Xu et al., 2011*), and listed in Table 1. Amplicon of each isolate was purified and cloned into pMD 19 T vector, and send to TSINGKE Biological Technology Co., Ltd for gene sequencing. *IscR* gene for each isolate was sequenced twice in both the forward and reverse direction. *IscR* sequences of 15 reference strains (GenBank accession numbers: KC795297, KC795318, KC795335, KC795364, KC795350, KC795383, KC795404, KC795422, KC795440, KC795459, KC795480, KC795495, KC795512, KC795532, KC795549) of *H. parasuis* were originated from a previous study (*Howell et al., 2013*).

Table 1 Primers used for the PCR assays for potential virulence markers and *iscR*.

Gene	Primer name	Primers(5′-3′)	Amplicon size (bp)	References
HAPS_0254	B14-F B14-R	ACACCTTATGCTTCCGCTAT ACGGTAACAGAACAAGAGCC	146	*Wang et al. (2011)*
HAPS_0254	D45-F D45-R	CTTCCGCTATCGCATTAA AACAGAACAAGAGCCTAAA	131	
nhaC	E30-F E30-R	GTCCAGGAAGCATAATACA TACAAGGTGGCGAGATAA	312	
fhuA	E35-F E35-R	TCTAAGCGATGGGATTGAGC GGTGGCGTAAGACGTGATT	461	
hhdA	MP_A1 MP_A2	GGTTCTAGTTCACAAACAGCCAATAC GATATTTACCCCTGCCTTCATTGTATC	964	*Sack & Baltes (2009)*
hhdB	MP_B1 MP_B2	ATCTTGCCCTGATTAGAGAGTAGGAGT GTGAATATAGCCCTTATCCAAATAGGC	557	
hsdR	hsdR-F hsdR-R	ATGACTATTGAAACAACGCCTATCA GAACTTGGTTAAAGGCTTCC	544	*Yu et al. (2014)*
hsdS*	hsdS-F hsdS-R	ATGATAGAAAGTCGTTTTATTGAAAAG GAAAACTTAATAATTTCTCTCTGTA	1,219	This study
ompP2	ompP2-F ompP2-R	ATGAAAAAAACACTAGTAGCA TTACCATAATACACGTAAACC	1,092	
HPS_21058	21058-F 21058-R	CCGAAAGCATAGATCCAAATGC CCACCTTGTTTACTTGCTTCTGC	590	*Howell et al. (2017)*
HPS_21059	21059-F 21059-R	CGTAGCATACGCACACCTAAAG GAAAGGGCAATAGATACATTTCGG	720	
vtaA	AV1-F V1-R NV1-R	AAATATTTAGAGTTATTTGGAGTC AATATACCTAGTAATACTAGACTTAAAAG CAGAATAAGCAAAATCAGC	190 222	*Galofre-Mila et al. (2017)*
iscR	iscR-F iscR-R	ATGAAATTAACTTCACGAGGACG TTAATGGTGATGATCGTGGCAG	447	This study

Note:
* Primers of the gene were newly designed in this study and were different from the original one.

Alignment of the *iscR* sequence was prepared with MEGA 7.0, Lasergene 7 or DNAMAN 9. Phylogenetic tree was decorated with an online tool iTOL v4.2.3 (https://itol.embl.de/).

Clinical backgrounds and detection of PVMs

Isolates from systemic sites such as spleen, brain, pericardium liquid, pleural effusion, seroperitoneum, and joint were considered to be systemic ones, and those from lungs and nasal cavities were considered to be respiratory ones and nasal ones, respectively.

To further confirm the population structure and its relevance with clinical phenotype, 12 PVMs from five pieces of literatures were selected. Primers of one of the selected PVMs (*hsdS*) were newly designed in this study (Table 1), and primers of other selected PVMs are the same with that in the literatures (*Sack & Baltes, 2009*; *Wang et al., 2011*; *Yu et al., 2014*; *Howell et al., 2017*; *Galofre-Mila et al., 2017*). The presence or absence of the selected PVMs in all the isolates were determined by PCR. Two or more strains from the same farm with any difference in respect of serovar, *iscR* clade, presence or absence of any PVMs were considered to be different isolates. The differences in average number of

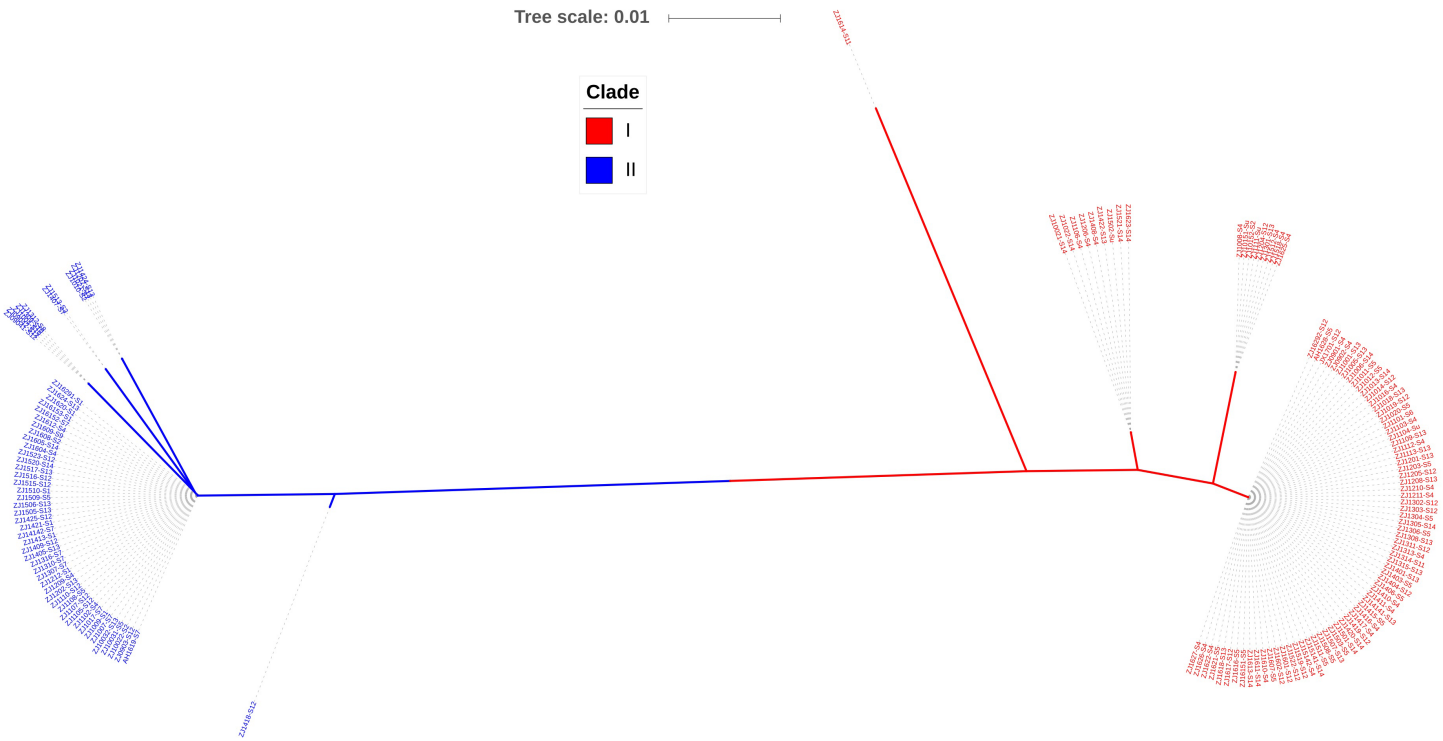

Figure 1 Phylogenetic tree of *iscR* proteins from 148 isolates. The proteins were highly conserved within both of the clade but deeply divergent between the two clades. Molecular Phylogeny was analyzed by Maximum Likelihood method using MEGA7 and decorated with an online tool iTOL v4.2.3 (https://itol.embl.de/).
Full-size DOI: 10.7717/peerj.6950/fig-1

PVMs between serovars were determined by One-way ANOVA with SPSS Statistics 17.0. Differences were considered statistically significant when $P < 0.05$.

Accession numbers

The nucleotide sequences of *iscR* for the 148 isolates were submitted to GenBank. The accession numbers for the individual CDSs are MH982282–MH982429.

RESULTS

Separation of clinical isolates into two deeply divergent groups by *iscR*-based phylogenetic analysis

Iron-sulfur cluster regulator gene was presented in all 148 isolates used in this study with the consistent open reading frame size of 447 bp. Phylogenetic sequence analysis of *iscR* amino acid sequences showed that the isolates were divided into two deeply divergent clades, which were designated as clade I and clade II (Fig. 1). Dominant sequence type (ST1 and ST6) in each clade were further aligned with *iscR* of other Gram-negative bacteria, and simultaneous amino acid substitutions at highly conserved residue 29, 68, 77 made clade I *H. parasuis* an evolutionarily distant lineage from this and other Gram-negative species (Fig. 2A). Amino acid sequence comparison revealed nine sequence types from the clinical isolates, and amino acid substitutions at residue 29, 68, 70–71, 77–78, 140 took place at 100% or nearly 100% between *iscR* of the two clades (Fig. 2B).

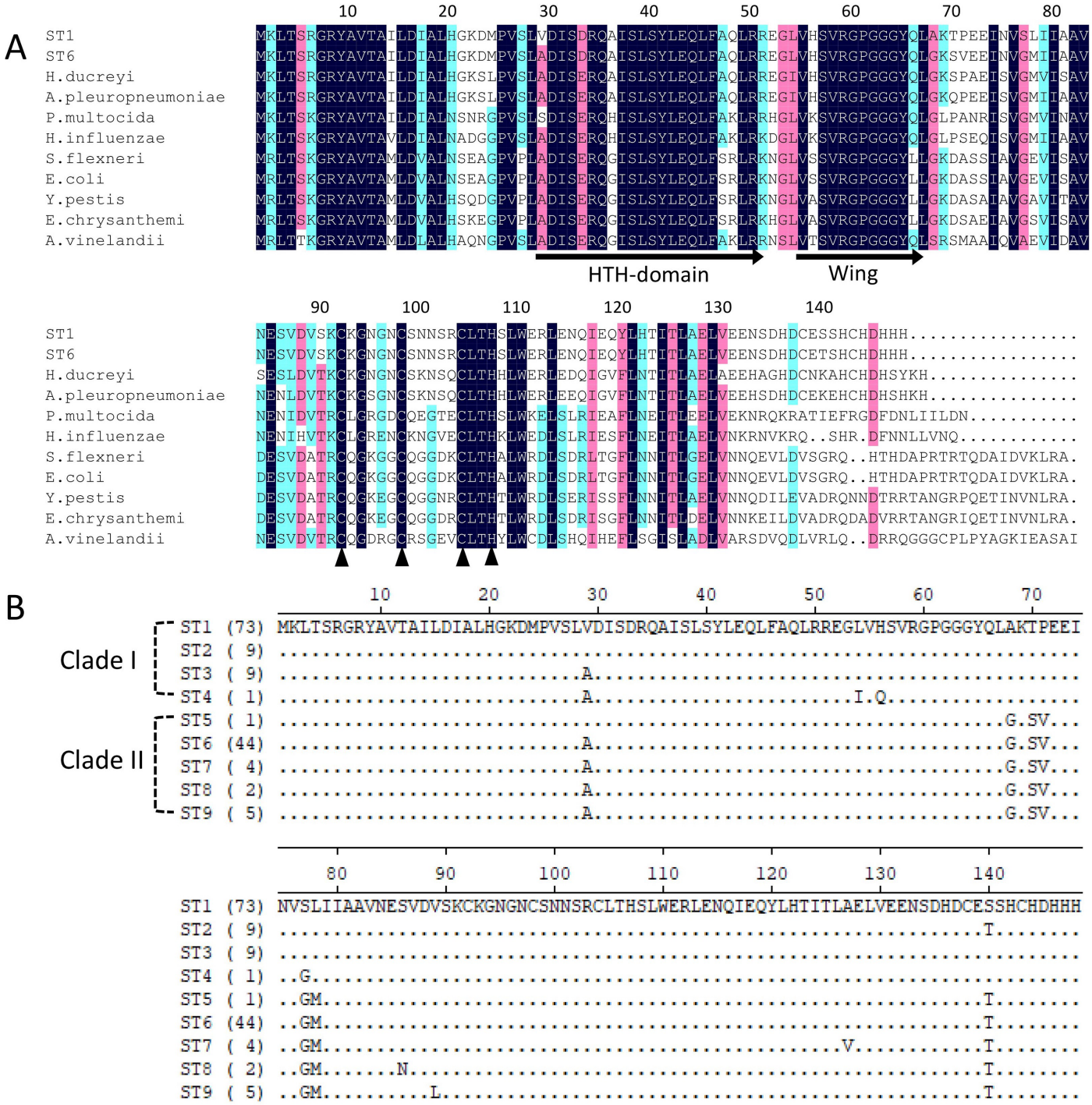

Figure 2 Alignment of *iscR* protein sequences of *H. parasuis*. (A) Conservation of *iscR* proteins from Gram-negative bacteria. *IscR* protein of the dominant sequence type in each clade (ST1 for clade I and ST6 for clade II) was aligned. Residue numbering refers to the *H. parasuis iscR* sequence. DNA-binding domain, encompassing the HTH motif and a wing, was indicated with arrows. Fe/S cluster ligands were indicated with triangles. Strictly conserved amino acids were highlighted in dark blue, and decreasing residue conservation was represented by a color gradient from red to green. Alignment was prepared with DNAMAN 9. (B) Representatives of the nine sequence types were aligned using Lasergene (MegAlign). Amino acid residue of the sequence type 1 (ST1) were used as marker, and residues that match ST1 exactly were hide (as "."). The number of isolates represented by each sequence type is shown in the brackets. The clades distribution of each sequence type were indicated.

Full-size DOI: 10.7717/peerj.6950/fig-2

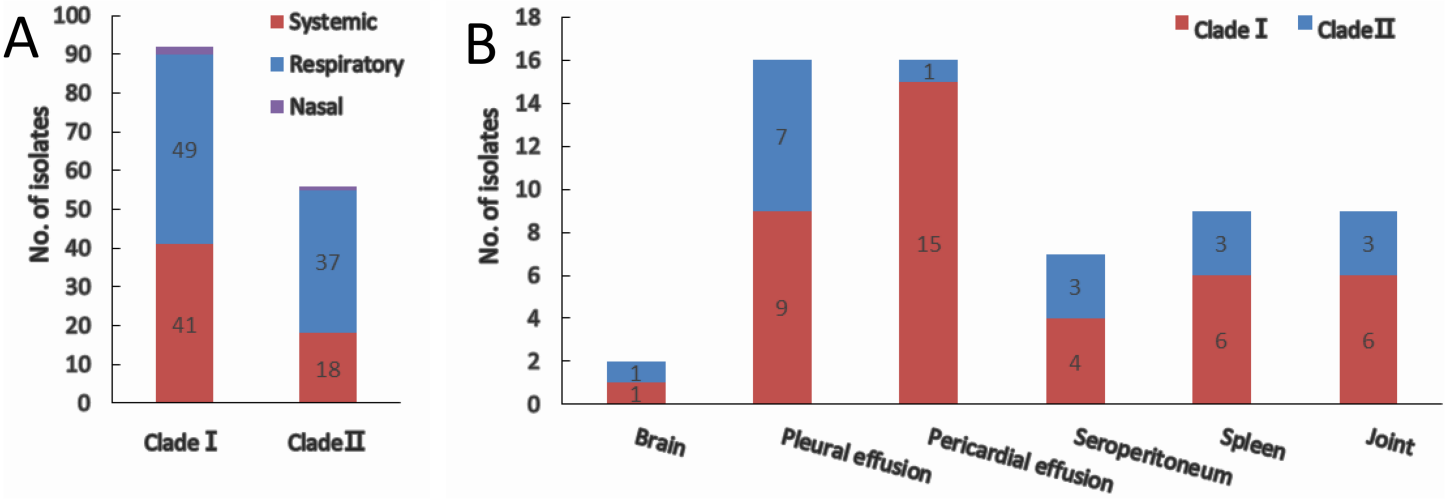

Figure 3 Clade distribution (A) and isolation site (B) of systemic isolates. The number of isolates are indicated with data tags in the histogram. (A) Distribution of nasal, respiratory, and systemic isolates in clade I and clade II. (B) Isolation site of systemic isolates in clade I and clade II.
Full-size DOI: 10.7717/peerj.6950/fig-3

Four *iscR* protein sequence types were identified from the 15 reference strains, of which three were identical to those from the clinical isolates in this study (Fig. S1).

Of all the isolates used in this study, clade I isolates were more prevalent (62.2%) than clade II isolates (37.8%). A higher proportion of isolates were systemic ones in clade I (44.6%) than that in clade II (30.4%) (Fig. 3A). No tissue tropism was seen in respect of *iscR*-based phylogeny, as isolates from all the systemic sites can be found in both clades (Fig. 3B). It is noteworthy that 93.8% of the isolates from pericardial effusion are belong to clade I (Fig. 3B).

Clear correlation between serovar and *iscR*-based phylogeny

The two molecular serotyping methods allow the assignment of 144 (97.3%) isolates to 13 serovars except serovar 3 and 15, and 4 isolates (2.7%) could not be assigned to any of the 15 serovar known so far. Serovar 4 (18.9%) and serovar 12 (18.9%) were the most prevalent followed by serovar 13 (16.2%) and serovar 5 (13.5%) (Fig. 4A). Isolates within most of the serovars tend to be distributed in the same clade. Serovars 4, 5, and 14 were dominated by clade I isolates, while serovars 1, 2 and 7 were dominated by clade II isolates (Fig. 4A). Although slightly higher proportion of serovar 12 and 13 isolates were found in clade I, the distribution of isolates in the two serovars was relatively equal in both clades. For the remainder of the serovars, the number of isolates was too low to be able to restrict them to one clade.

Systemic isolates were found in 10 of the 13 serovars identified. Of all the serovars with not less than five isolates, systemic isolates were more common in serovar 13 (54.2%) and serovar 12 (50%), and similar proportion of systemic isolates were found in serovar 1, 4, 5 ranging from 35% to 37.5%, which were followed by serovar 7 (20%) and serovar 14 (14.3%) (Fig. 4B). These data indicate that virulence varies among serovars, and serovars 13 and 12 are more strongly linked to systemic infections.

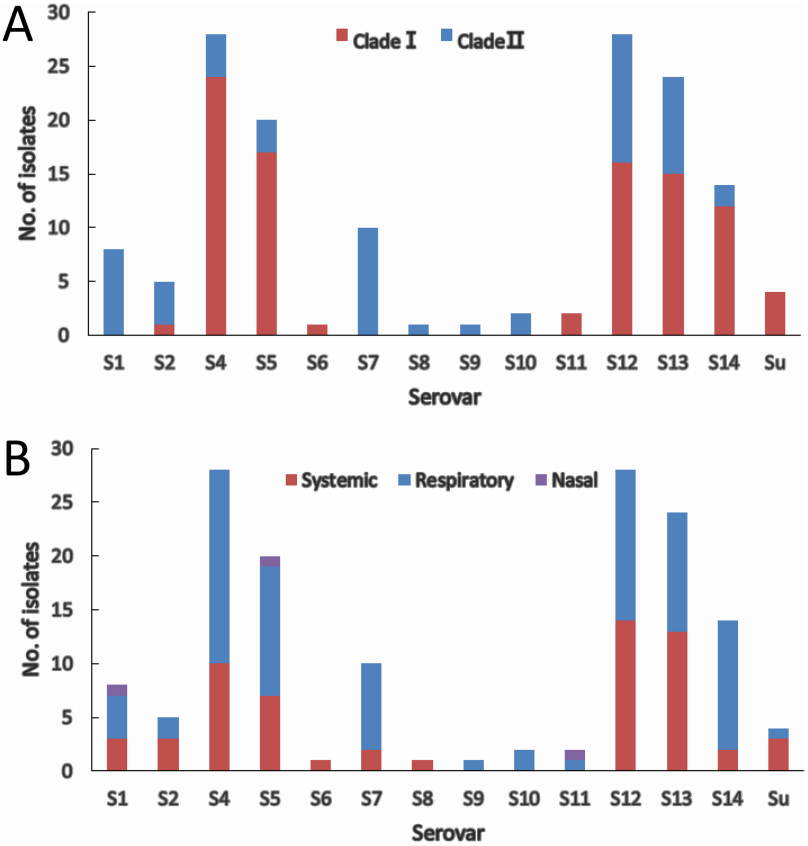

Figure 4 ***IscR*-based phylogenic (A) and clinical relevance (B) of isolates in each serovar.** (A) Clade distribution of isolates in each serovar. (B) Number of systemic, respiratory, and nasal isolates in each serovar. Full-size DOI: 10.7717/peerj.6950/fig-4

Presentation of PVMs showed obvious correlation with *iscR*-based phylogeny and serovar

To further confirm the population structure and its relevance with clinical phenotype, distributions of 12 previously identified PVMs were investigated in all the 148 isolates. OmpP2 and *vtaA* genes could be detected in 100% and 98.6% of the isolates, respectively. Only four isolates were identified to be avirulent by the *vtaA* method, of which one was isolated from joint of diseased pig. Hence, presence of *ompP2* and *vtaA* gene showed low resolution of the isolates, and the other 10 PVMs were selected in the following studies.

The positive rates of most of the PVMs were similar among serovars (4, 5, 12, 13, and 14) which were dominated by isolates in clade I and among serovars (1, 2, and 7) which were dominated by isolates in clade II. The positive rates were clearly different when compared between serovars of one group (serovars 4, 5, 12, 13, and 14) and the other group (serovars 1, 2, and 7). Most of the PVMs (8/10) were more prevalent in clade I than in clade II (Fig. 5). Of all the serovars with not less than 5 isolates, serovar 5 isolates carried the most PVMs with the average number of 7.35, while serovar 1 isolates carried the least with an average number of 1.5. Serovars dominated by clade I isolates were found with presence of remarkable higher average number of PVMs than those dominated by clade II

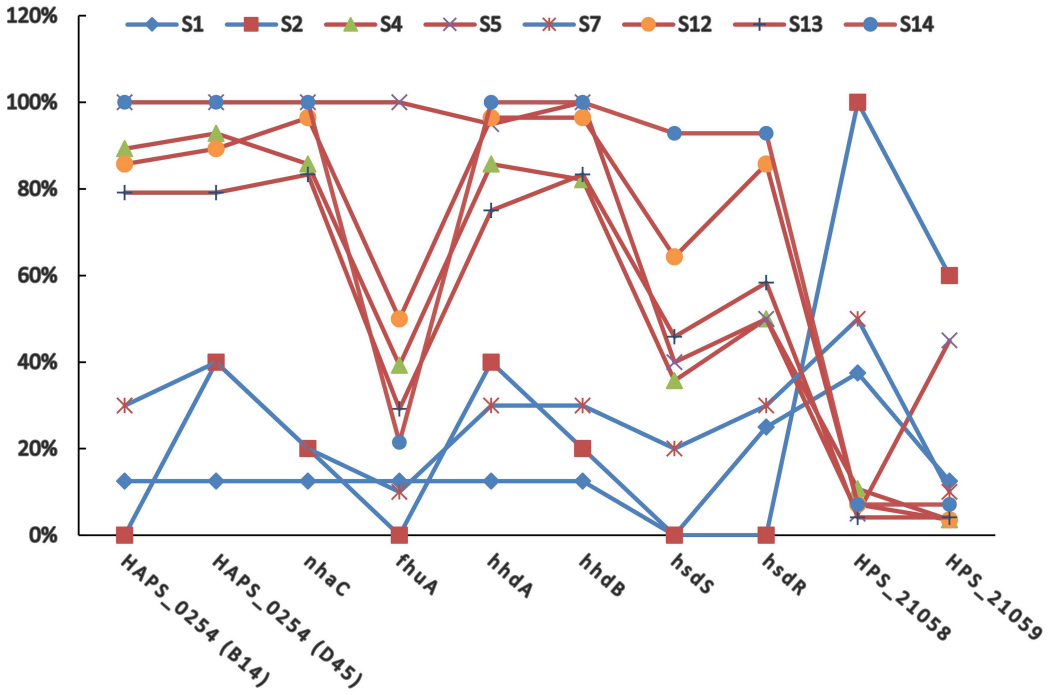

Figure 5 Distribution of potential virulence markers in each serovar. Percent of isolates with the presence of potential virulence markers in each serovar. Full-size DOI: 10.7717/peerj.6950/fig-5

isolates ($P < 0.05$) (Fig. 6). Interestingly, average number of PVMs in serovar 12 and 13 were similar to the serovars dominated by clade I isolates (Fig. 6), although the distribution of isolates of the two serovars was relatively equal in both clades. In this study, only one isolate was found with the presence of all the 10 selected PVMs (Fig. 7).

Higher strain diversity in isolates from multi-strain infected farms

A total of 34 PVMs profiles were identified based on the presence or absence of the 10 selected genes (Table 2). The most prevalent profile was shared by 39 isolates, followed by 22 isolates in profile 20. More than half of the profiles (58.8%) were represented by only one isolate, which suggests high gene heterogeneity in accessory genome among isolates. Of the six profiles with no less than five isolates, systemic isolates were most prevalent in profile 27 (4/5) and least prevalent in profile 29 (0/10). Similar proportion of systemic isolates (29–46%) were observed in other profiles (Table 2). Moreover, 10.2% of the systemic isolates do not carry any of the 10 selected PVMs (Table 2), and all of which were distributed in clade II with the exception of one serovar unknown isolate (Fig. 7). Neither presence of any of the selected PVMs nor the number of PVMs carried by the isolates showed correlation with clinical phenotype of systemic infection (Fig. 7).

Of all the 125 independent pig farms, multi-strain (two or three) infection was found in 16% of the pig farms and single strain infection was found in the other 84% pig farms. An average of every 4.4 isolates from the single strain infection farm produce one PVMs profile, while every 2.2 isolates from the multi-strain infection farm produce one PVMs profile (Table 3). Hence, multi-strain infection in one farm resulted in more diverse

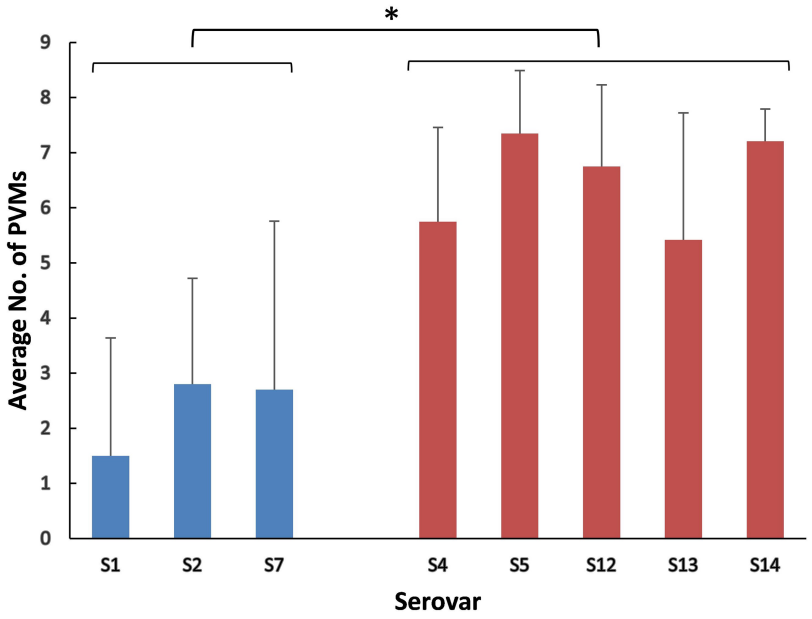

Figure 6 Average number of potential virulence markers (PVMs) carried by the isolates in each serovars. Data were shown as mean + S.D. A *P*-value was computed using the one-way ANOVA with SPSS Statistics 17.0. Variations were considered statistically significant when the *P*-value was ≤ 0.05, which is indicated with "*". Full-size DOI: 10.7717/peerj.6950/fig-6

PVMs profiles in those isolates, which indicates that frequent gene transfer among isolates took place in multi-strain infected farms.

Potential gene exchange at the capsule locus

To study the possible recombination between strains at the capsule locus, all isolates were clustered based on pattern similarity of PVMs. Two distinct clusters were identified, and were designated as cluster A and cluster B (Fig. 8). Most of the isolates (92.1%) of serovar 4, 5, 12, 13, and 14 harbored more than five PVMs, and were distributed in cluster B. Most of the isolates (78.3%) of serovar 1, 2, and 7 only have less than two PVMs, and were listed in cluster A (Fig. 8). Coincidently, clear separation of isolates from those serovars (4, 5, 12, 13, and 14) and other serovars (1, 2, and 7) was observed in *iscR*-based phylogeny analysis (Fig. 4A). It indicates that these two clusters were possibly originated from two independently evolutionary lineages, which may have branched off from each other for a long evolutionary period.

Recombination event may take place when two or more strains co-exist in one farm, especially when in the same infected organ or commensal site. A small portion of isolates (21.7%) of serovar 1, 2, and 7 such as ZJ1510-S1 and ZJ1310-S7 were divided into cluster B. Similarly, a small portion of strains (7.9%) of serovar 4, 5, 12, 13, 14 were listed in cluster A, for instance, ZJ1604-S4, ZJ1207-S13, ZJ1405-S13, and ZJ1424-S13 etc., (Fig.8). All these isolates harbored PVMs characteristic that do not match the majority of isolates of the same serovar, but consistent with PVMs characteristic of serovars which were mainly distributed in the other cluster. It indicates that these isolates may have undergone gene exchange between one serovar group (1, 2 or 7) and the other serovar

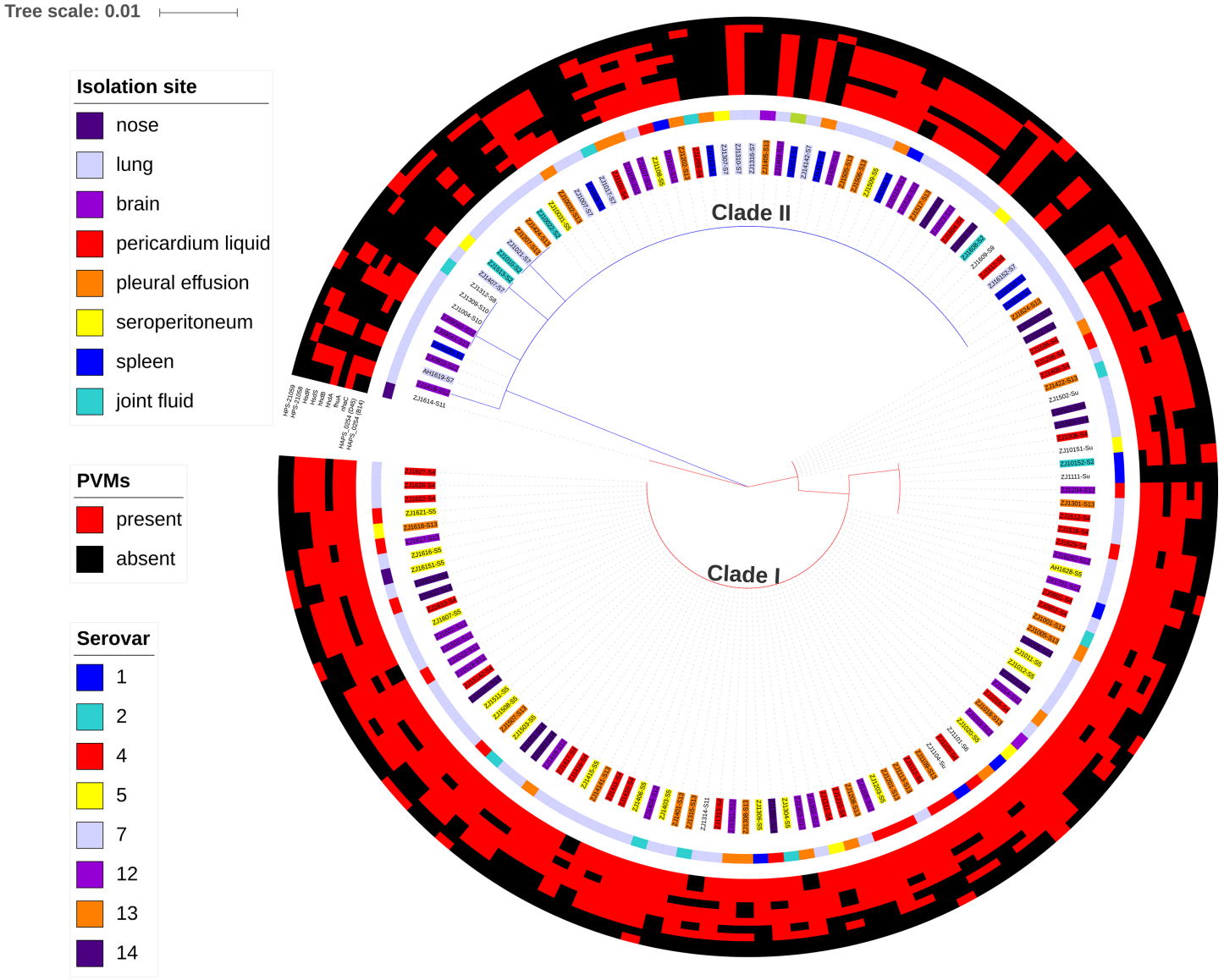

Figure 7 Maximum likelihood phylogenetic tree with the 148 isolates of *H. parasuis*, based on iscR protein sequence alignment. The tree was constructed with MEGA7, and decorated using ItoL (https://itol.embl.de/). Clade colors indicate the divergence of clade I (red) and clade II (blue). Inner circle represents lineage for each strains, and background color of the text label at each node indicates the serovar of the strain. The middle circle shows the isolation site of the strain. The outer circle shows the presence (red) or absence (green) of the 10 selected potential virulence markers (PVMs) in each isolates.
Full-size DOI: 10.7717/peerj.6950/fig-7

group (4, 5, 12, 13, and 14) at or around the site of the capsule locus. Hence, isolates in the same serovar can vary greatly in terms of PVMs, which could lead to differences in the pathogenicity and immunogenicity within one serovar. It would be more difficult to identify in this study when exchange of capsule locus take place between serovars with similar PVMs characteristics.

DISCUSSION

Phylogeny and population structure of *H. parasuis* were widely studied, and *H. parasuis* strains could be divided into two groups in a variety of ways (*Oliveira & Pijoan, 2004*;

Table 2 Potential virulence markers (PVMs) profiles and systemic isolates distribution in each profile.

No. of PVMs	Serial number of the profiles	No. of systemic/ total isolates in the profile	HAPS_0254 (B14)	HAPS_0254 (D45)	nhaC	fhuA	hhdA	hhdB	hsdS	hsdR	HPS-21058	HPS-21059
0	1	6/13	−	−	−	−	−	−	−	−	−	−
1	2	2/7	−	−	−	−	−	−	−	−	+	−
	3	1/1	−	+	−	−	−	−	−	−	−	−
2	4	0/1	−	−	−	−	−	−	−	+	+	−
	5	2/3	−	−	−	−	−	−	−	−	+	+
	6	0/1	−	−	−	−	−	−	+	+	−	−
	7	0/1	−	−	+	−	+	−	−	−	−	−
3	8	0/1	−	+	−	−	+	−	−	−	+	−
	9	0/1	−	−	−	−	−	−	−	+	+	+
	10	0/1	+	+	−	−	−	−	−	+	−	−
	11	0/1	+	+	−	−	+	−	−	−	−	−
4	12	0/1	−	−	+	−	+	+	−	+	−	−
	13	2/2	+	+	+	−	−	+	−	−	−	−
	14	1/1	−	+	−	−	+	−	−	+	+	−
5	15	0/1	−	+	+	−	+	+	−	−	+	−
	16	1/1	+	+	+	−	+	−	−	+	−	−
	17	1/1	+	+	+	−	+	+	−	−	−	−
	18	2/2	+	+	+	−	−	+	−	+	−	−
	19	1/3	−	−	+	−	+	+	+	+	−	−
6	20	9/22	+	+	+	+	+	+	−	−	−	−
	21	1/1	−	−	+	+	+	+	+	+	−	−
	22	3/4	+	+	+	−	+	+	−	+	−	−
	23	1/1	+	+	+	+	−	+	−	+	−	−
	24	1/1	−	+	+	−	+	+	−	−	+	+
	25	0/1	+	+	+	−	+	−	+	+	−	−
	26	0/1	+	+	−	−	+	+	+	+	−	−
7	27	4/5	+	+	+	+	+	+	−	+	−	−
	28	13/39	+	+	+	−	+	+	+	+	−	−
	29	0/10	+	+	+	+	+	+	−	−	−	+
8	30	1/1	+	+	+	+	+	+	−	+	+	−
	31	5/13	+	+	+	+	+	+	+	+	−	−
9	32	1/3	+	+	+	+	+	+	+	+	+	−
	33	0/2	+	+	+	+	+	+	+	+	−	+
10	34	1/1	+	+	+	+	+	+	+	+	+	+
Total	34	59/148										

Olvera, Cerda-Cuellar & Aragon, 2006b; *Turni & Blackall, 2010*; *Howell et al., 2014*; *Moreno et al., 2016*; *Wang et al., 2016*). Results of current study on correlation between population structure and virulence were contradictory. *H. parasuis* strains in some sub-groups tend to be virulent (*Ruiz et al., 2001*; *Olvera, Calsamiglia & Aragon, 2006a*; *Olvera, Cerda-Cuellar & Aragon, 2006b*; *Wang et al., 2016*), while no association between

Table 3 Potential virulence markers (PVMs) profiles in single and multi-strain infected farms.

Infection mode	No. of farms (%)	No. of isolates	No. of PVMs profiles
Single strain	105 (84)	105	24
Multi-strain	20 (16)	43	20

population structure and clinical virulence was found in others (*Zehr, Lavrov & Tabatabai, 2012*; *Boerlin et al., 2013*; *Howell et al., 2014*). No strong correlation between clinical phenotype and *iscR*-based phylogeny or PVMs profiles was observed in this study.

The relationship between serotype and other typing methods has been evaluated (*Turni & Blackall, 2010*; *Ma et al., 2016*; *Zhao et al., 2018*), but correlation were only been identified in a few reports (*Boerlin et al., 2013*; *Howell et al., 2014*). In this study, isolates of serovar 4, 5, and 14 were found predominantly distributed in clade I, and those of serovar 12 and 13 were mostly found in clade I, while those of serovar 1, 2, and 7 were predominantly in clade II. It is similar to the results that separation of serovars (serovar 5, 12, 13 predominantly in clade I and serovar 7 in clade II) was found between the two clades in a previous study (*Howell et al., 2014*). Our study was also consistent with another report, in which serovar 5, 13, 14 were found only in cluster I and serovar 2 was only found in cluster II (*Boerlin et al., 2013*).

Systemic isolates were most frequently seen in serovar 13 in this study. In a previous study, serovar 13 were not the most prevalent one in respect of total isolation frequency, but was the most prevalent one with regard to the frequency of systemic isolates (*Cai et al., 2005*). Serovar 5 and 4 were determined to be more strongly linked to disease in another study (*Howell et al., 2014*). Mega data on correlation between serovar and virulence or clinical phenotype is still too limited to understand the relationship. Moreover, virulence of isolates varies greatly within the same serovar (*Lawrence et al., 2014*; *Yu et al., 2014*), and the reason for the variation is still to be determined. In this study, some isolates carry completely different PVMs compared to most of the counterparts within the same serovar, which indicates that those isolates may vary extensively in the accessory genome. The difference in the accessory genome may be responsible to the variation of virulence within one serovar. Genes in the accessory genome had been linked to differences in clinical phenotype (*Howell et al., 2014*). The vast genetic heterogeneity among isolates may pose a challenge to vaccine development (*Turni, Singh & Blackall, 2018*).

In this study, selected PVMs were similarly distributed in serovars of which the isolates were predominantly distributed in the same clade, but huge differences were found when compared in between the clades. Hence, similarity of PVMs distribution in the accessory genome further suggest the close relationship among serovars dominated by isolates in the same clade. Moreover, phylogenetic analysis of gene content of the capsule loci support the close relationship among serovar 1, 2, and 7 (*Howell et al., 2013*), which is consistent with our finding that isolates of those serovars were all found in the same clade with exception of one isolate of sevovar 2. The coincidence of i*scR*-based phylogeny, separation of serovars and similarity of PVMs distribution in the accessory genome strongly suggests the independent evolution of the two branches of *H. parasuis*.

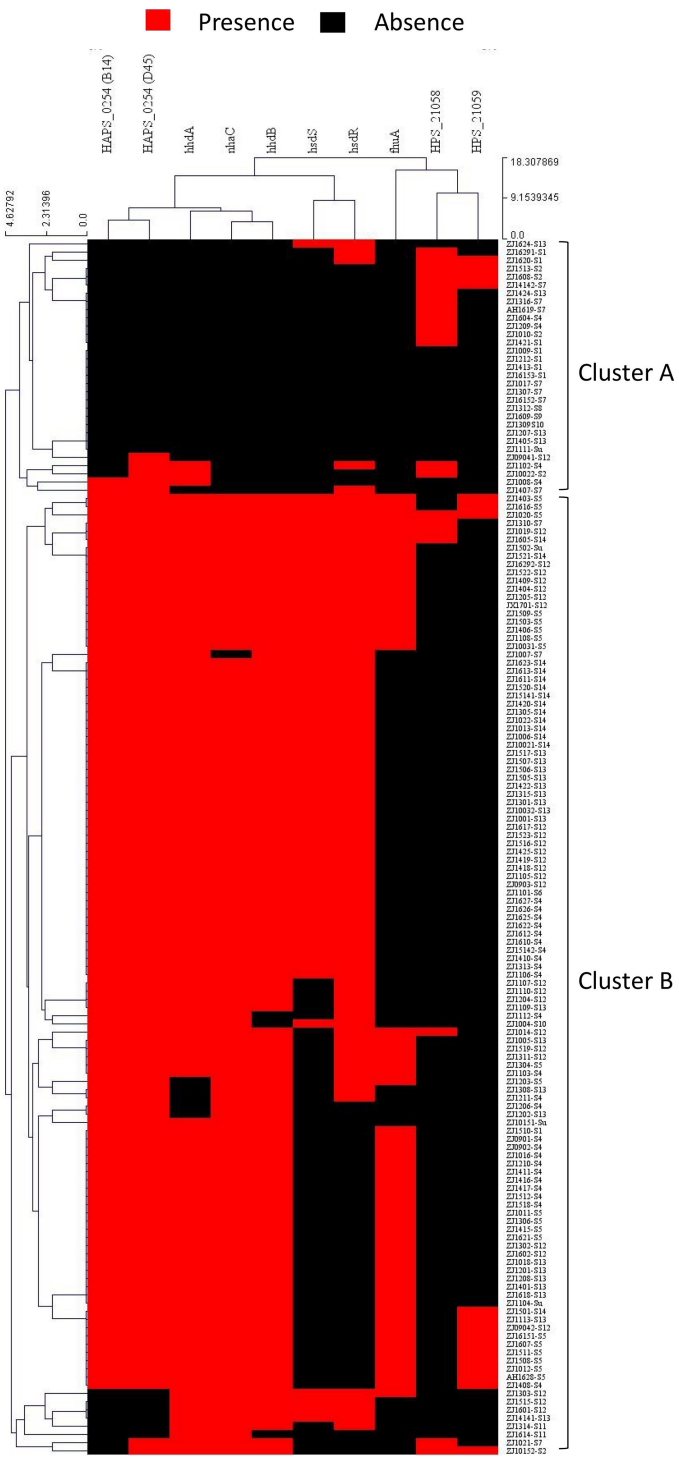

Figure 8 Hierarchical clustering of *H. parasuis* isolates based on the pattern similarity of potential virulence factors. Isolates were ordered along the columns based on the similarity in the presence (red) and absence (black) pattern of 10 potential virulence factors (HAPS_0254 (B14), HAPS_0254 (D45), *nhaC*, *fhuA*, *hhdA*, *hhdB*, *hsdS*, *hsdR*, HPS_21058, and HPS_21059). The similarity matrix was calculated using the Euclidean distance. Cluster analysis was performed by the Hierarchical clustering (HCL) through the cluster analysis procedure of MeV (Multiple Experiment Viewer) v4.9.0 package.

Full-size DOI: 10.7717/peerj.6950/fig-8

The phylogenetic and amino acid substitution analysis of *iscR* inferred that clade II was the first to branch off from its ancestor, and clade I was possibly originated from clade II at the early stage of evolution. Long term of independent evolution may have resulted in a huge difference in the accessory genome. Pronounced difference between the accessory genome of the two clades had been observed in a previous study (*Howell et al., 2014*).

Serovar 12 and 13 were found in relatively equal proportions in both clades in this study, but were identified predominantly in one clade in other literature (*Howell et al., 2014*). Most of the serovar 12 or 13 isolates in clade II were found harboring typical PVMs characteristics of the counterparts in clade I, which indicates that those isolates were probably originated from clade I and transformed into clade II isolates through homologous recombination at or around the site of *iscR* gene. Since *iscR* was the target gene for population structure analysis in this study, isolates that came across gene exchange at this site of the chromosome may be shifted from one clade to another. It is noteworthy that gene exchange among isolates within the same clade would be difficult to spot due to the similar genetic background. Higher rates of homologous recombination for the clinical isolates compared with the non-clinical isolates have been observed in *H. parasuis* (*Vos & Didelot, 2008*). More complex of PVMs profile were observed in isolates from multi-strain infected farm, indicating more frequent gene exchange in those isolates. Both higher prevalence and higher proportion of systemic isolates in serovar 12 and 13 increase the opportunity for them to contact with other isolates through co-infection, which may be part of the reason that leads to the higher rates of gene exchange in those serovars. Hence, extensive recombination within the species makes it harder to decipher the population structure of *H. parasuis*.

No isolate from healthy pigs was included for comparison in this study, and the absence of commensals was a possible reason for the low resolution of *vtaA* and *ompP2* in differentiation of the isolates although these two genes were also found in all strains including both commensals and systemic isolates in other studies (*Mullins et al., 2009*; *Howell et al., 2014*). The common pathogens that co-infected with the *H. parasuis* isolates in our diagnostic center are porcine reproductive and respiratory syndrome virus, porcine circovirus type 2, *Streptocccus suis*, *Bordetella bronchiseptica* etc. Recent research showed that many previously suggested virulence factors were not appropriate markers of virulence (*Howell et al., 2014*; *Turni, Singh & Blackall, 2018*). In this study, a total of six systemic isolates harbored none of the PVMs, co-infection with other viral or bacterial pathogens were found for only three of the isolates. No correlation was found between the 10 selected PVMs and clinical virulence of the isolates. Those PVMs were probably serovar or clade specific genes emerged during evolution. To a certain extent, those PVMs may contribute to virulence, but they are not the genes that enable the organism to be virulent.

Of the 20 farms of multi-strain infection, systemic disease caused by multi-strains (ZJ10151-Su and ZJ10152-S2) was found in one farm, which stress the need for full assessment of the causative strains before treatment. Of the 86 lung isolates, a total of

24 isolates were isolated from lungs with presence of fibrinous inflammation, and the others were isolated from lungs with barely pneumonia. No difference was observed between the lung isolates from the two different clinical cases in respect of serovar distribution, PVMs profile and *iscR* phylogeny. Moreover, some of the pathogens such as *S. suis* and *Mycoplasma hyorhinis* could also cause lesions similar to Glässer's disease, and we could not exclude the co-infection of all of these possible agents especially for *M. hyorhinis* in this study. Hence, virulence factors involved in invasion are still to be determined for the agent.

CONCLUSIONS

Iron-sulfur cluster regulator is a reliable marker for studying population structure of *H. parasuis*, while other factors should be included to avoid the interference of gene exchange. The two lineages of *H. parasuis* may have undergone independent evolution, but shows no difference in clinical virulence. Wide distribution of systemic isolates across the entire population poses new challenge for development of vaccine with better cross-protection. Our study provide new information for better deciphering the population structure of *H. parasuis*, which helps understanding the extreme diversity within this pathogenic bacterium.

ACKNOWLEDGEMENTS

We thank the members of the Swine Disease diagnostic Center of the institute of the animal husbandry and veterinary science for their underlying contributions.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was financially supported by the National Natural Science Foundation of China (31201934), the Zhejiang Provincial Key Research and Development programs (2019C02052-2), the Projects of Zhejiang Provincial Science and Technology Plans (2016C02054-7) and the Young Talent Training Project of Zhejiang Academy of Agricultural Sciences (10102000617CF0101G/003/003). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Natural Science Foundation of China: 31201934.
Zhejiang Provincial Key Research and Development programs: 2019C02052-2.
Zhejiang Provincial Science and Technology Plans: 2016C02054-7.
Zhejiang Academy of Agricultural Sciences: 10102000617CF0101G/003/003.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Junxing Li conceived and designed the experiments, performed the experiments, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Lihua Xu performed the experiments, prepared figures and/or tables.
- Fei Su analyzed the data.
- Bin Yu analyzed the data.
- Xiufang Yuan conceived and designed the experiments, contributed reagents/materials/analysis tools, approved the final draft.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The nucleotide sequences of iscR for the 148 isolates are available in GenBank.

The accession numbers for the individual CDSs are MH982282–MH982429.

Data Availability

The following information was supplied regarding data availability:

The raw measurements are available in Dataset S1.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.6950#supplemental-information.

REFERENCES

Angen O, Oliveira S, Ahrens P, Svensmark B, Leser TD. 2007. Development of an improved species specific PCR test for detection of *Haemophilus parasuis*. *Veterinary Microbiology* **119(2–4)**:266–276 DOI 10.1016/j.vetmic.2006.10.008.

Aragon V, Cerdà-Cuéllar M, Fraile L, Mombarg M, Nofrarías M, Olvera A, Sibila M, Solanes D, Segalés J. 2010. Correlation between clinico-pathological outcome and typing of *Haemophilus parasuis* field strains. *Veterinary Microbiology* **142(3–4)**:387–393 DOI 10.1016/j.vetmic.2009.10.025.

Biberstein EL, Gunnarsson A, Hurvell B. 1977. Cultural and biochemical criteria for the identification of haemophilus spp from swine. *American Journal of Veterinary Research* **38**:7–11.

Boerlin P, Poljak Z, Gallant J, Chalmers G, Nicholson V, Soltes GA, MacInnes JI. 2013. Genetic diversity of *Haemophilus parasuis* from sick and healthy pigs. *Veterinary Microbiology* **167(3–4)**:459–467 DOI 10.1016/j.vetmic.2013.07.028.

Brockmeier SL, Register KB, Kuehn JS, Nicholson TL, Loving CL, Bayles DO, Shore SM, Phillips GJ. 2014. Virulence and draft genome sequence overview of multiple strains of the swine pathogen *Haemophilus parasuis*. *PLOS ONE* **9(8)**:e103787 DOI 10.1371/journal.pone.0103787.

Cai X, Chen H, Blackall PJ, Yin Z, Wang L, Liu Z, Jin M. 2005. Serological characterization of *Haemophilus parasuis* isolates from China. *Veterinary Microbiology* **111(3–4)**:231–236 DOI 10.1016/j.vetmic.2005.07.007.

Crack JC, Green J, Hutchings MI, Thomson AJ, Le Brun NE. 2012. Bacterial iron-sulfur regulatory proteins as biological sensor-switches. *Antioxidants & Redox Signaling* **17(9)**:1215–1231 DOI 10.1089/ars.2012.4511.

Galofre-Mila N, Correa-Fiz F, Lacouture S, Gottschalk M, Strutzberg-Minder K, Bensaid A, Pina-Pedrero S, Aragon V. 2017. A robust PCR for the differentiation of potential virulent strains of *Haemophilus parasuis*. *BMC Veterinary Research* **13(1)**:124 DOI 10.1186/s12917-017-1041-4.

Haines S, Arnaud-Barbe N, Poncet D, Reverchon S, Wawrzyniak J, Nasser W, Renauld-Mongenie G. 2015. IscR Regulates Synthesis of Colonization Factor Antigen I Fimbriae in Response to Iron Starvation in Enterotoxigenic *Escherichia coli*. *Journal of Bacteriology* **197(18)**:2896–2907 DOI 10.1128/JB.00214-15.

Howell KJ, Peters SE, Wang J, Hernandez-Garcia J, Weinert LA, Luan S-L, Chaudhuri RR, Angen O, Aragon V, Williamson SM, Parkhill J, Langford PR, Rycroft AN, Wren BW, Maskell DJ, Tucker AW. 2015. Development of a multiplex PCR assay for rapid molecular serotyping of *Haemophilus parasuis*. *Journal of Clinical Microbiology* **53(12)**:3812–3821 DOI 10.1128/JCM.01991-15.

Howell KJ, Weinert LA, Chaudhuri RR, Luan S-L, Peters SE, Corander J, Harris D, Angen O, Aragon V, Bensaid A, Williamson SM, Parkhill J, Langford PR, Rycroft AN, Wren BW, Holden MT, Tucker AW, Maskell DJ. 2014. The use of genome wide association methods to investigate pathogenicity, population structure and serovar in *Haemophilus parasuis*. *BMC Genomics* **15(1)**:1179 DOI 10.1186/1471-2164-15-1179.

Howell KJ, Weinert LA, Luan S-L, Peters SE, Chaudhuri RR, Harris D, Angen O, Aragon V, Parkhill J, Langford PR, Rycroft AN, Wren BW, Tucker AW, Maskell DJ, on behalf of the BRaDP1T Consortium. 2013. Gene content and diversity of the loci encoding biosynthesis of capsular polysaccharides of the 15 serovar reference strains of *Haemophilus parasuis*. *Journal of Bacteriology* **195(18)**:4264–4273 DOI 10.1128/JB.00471-13.

Howell KJ, Weinert LA, Peters SE, Wang J, Hernandez-Garcia J, Chaudhuri RR, Luan S-L, Angen O, Aragon V, Williamson SM, Langford PR, Rycroft AN, Wren BW, Maskell DJ, Tucker AW. 2017. Pathotyping" multiplex PCR assay for *Haemophilus parasuis*: a tool for prediction of virulence. *Journal of Clinical Microbiology* **55(9)**:2617–2628 DOI 10.1128/JCM.02464-16.

Jia A, Zhou R, Fan H, Yang K, Zhang J, Xu Y, Wang G, Liao M. 2017. Development of Serotype-Specific PCR assays for typing of *Haemophilus parasuis* isolates circulating in Southern China. *Journal of Clinical Microbiology* **55(11)**:3249–3257 DOI 10.1128/JCM.00688-17.

Kielstein P, Rapp-Gabrielson VJ. 1992. Designation of 15 serovars of *Haemophilus parasuis* on the basis of immunodiffusion using heat-stable antigen extracts. *Journal of Clinical Microbiology* **30**:862–865.

Lawrence PK, Wiener BL, Kolander-Bremer T, Bey RF, Stine DL, Kittichotirat W, Bumgarner RE. 2014. Genome-wide association studies of virulent and avirulent *Haemophilus parasuis* serotype 4 strains. *Genome Announcements* **2(5)**:e00884-14 DOI 10.1128/genomeA.00884-14.

Ma L, Wang L, Chu Y, Li X, Cui Y, Chen S, Zhou J, Li C, Lu Z, Liu J, Liu Y. 2016. Characterization of Chinese *Haemophilus parasuis* isolates by traditional serotyping and molecular serotyping methods. *PLOS ONE* **11(12)**:e0168903 DOI 10.1371/journal.pone.0168903.

Mettert EL, Perna NT, Kiley PJ. 2014. Sensing the cellular Fe-S cluster demand: A structural, functional, and phylogenetic overview of *Escherichia coli* IscR. In: Rouault T, ed. *Iron-Sulfur Clusters in Chemistry and Biology*. Berlin: Walter de Gruyter GmbH, 326–345.

Miller HK, Auerbuch V. 2015. Bacterial iron-sulfur cluster sensors in mammalian pathogens. *Metallomics* **7(6)**:943–956 DOI 10.1039/C5MT00012B.

Moreno LZ, Silva GFR, Gomes VTM, Matajira CEC, Silva APS, Mesquita RE, Lotto NP, Ferreira TSP, Christ APG, Sato MIZ, Gherpelli Y, Dottori M, Bonilauri P, Luppi A, Moreno AM. 2016. Application of protein profiling of virulent *Haemophilus parasuis* by MALDI-TOF mass spectrometry. *Journal of Infection in Developing Countries* **10(6)**:678–681 DOI 10.3855/jidc.7787.

Mullins MA, Register KB, Bayles DO, Loving CL, Nicholson TL, Brockmeier SL, Dyer DW, Phillips GJ. 2009. Characterization and comparative analysis of the genes encoding *Haemophilus parasuis* outer membrane proteins P2 and P5. *Journal of Bacteriology* **191(19)**:5988–6002 DOI 10.1128/JB.00469-09.

Oliveira S, Blackall PJ, Pijoan C. 2003. Characterization of the diversity of *Haemophilus parasuis* field isolates by use of serotyping and genotyping. *American Journal of Veterinary Research* **64(4)**:435–442 DOI 10.2460/ajvr.2003.64.435.

Oliveira S, Pijoan C. 2004. Computer-based analysis of *Haemophilus parasuis* protein fingerprints. *Canadian Journal of Veterinary Research* **68(1)**:71–75.

Olvera A, Calsamiglia M, Aragon V. 2006a. Genotypic diversity of *Haemophilus parasuis* field strains. *Applied and Environmental Microbiology* **72(6)**:3984–3992 DOI 10.1128/AEM.02834-05.

Olvera A, Cerda-Cuellar M, Aragon V. 2006b. Study of the population structure of *Haemophilus parasuis* by multilocus sequence typing. *Microbiology* **152(12)**:3683–3690 DOI 10.1099/mic.0.29254-0.

Ruiz A, Oliveira S, Torremorell M, Pijoan C. 2001. Outer membrane proteins and DNA profiles in strains of *Haemophilus parasuis* recovered from systemic and respiratory sites. *Journal of Clinical Microbiology* **39(5)**:1757–1762 DOI 10.1128/JCM.39.5.1757-1762.2001.

Sack M, Baltes N. 2009. Identification of novel potential virulence-associated factors in *Haemophilus parasuis*. *Veterinary Microbiology* **136(3–4)**:382–386 DOI 10.1016/j.vetmic.2008.11.008.

Santos JA, Pereira PJ, Macedo-Ribeiro S. 2015. What a difference a cluster makes: the multifaceted roles of IscR in gene regulation and DNA recognition. *Biochimica et Biophysica Acta (BBA)—Proteins and Proteomics* **1854(9)**:1101–1112 DOI 10.1016/j.bbapap.2015.01.010.

Turni C, Blackall PJ. 2005. Comparison of the indirect haemagglutination and gel diffusion test for serotyping *Haemophilus parasuis*. *Veterinary Microbiology* **106(1–2)**:145–151 DOI 10.1016/j.vetmic.2004.12.019.

Turni C, Blackall PJ. 2010. Serovar profiling of *Haemophilus parasuis* on Australian farms by sampling live pigs. *Australian Veterinary Journal* **88(7)**:255–259 DOI 10.1111/j.1751-0813.2010.00592.x.

Turni C, Singh R, Blackall PJ. 2018. Virulence-associated gene profiling, DNA fingerprinting and multilocus sequence typing of *Haemophilus parasuis* isolates in Australia. *Australian Veterinary Journal* **96(6)**:196–202 DOI 10.1111/avj.12705.

Vos M, Didelot X. 2008. A comparison of homologous recombination rates in bacteria and archaea. *ISME Journal* **3(2)**:199–208 DOI 10.1038/ismej.2008.93.

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

Wang X, Xu X, Zhang S, Guo F, Cai X, Chen H. 2011. Identification and analysis of potential virulence-associated genes in *Haemophilus parasuis* based on genomic subtraction. *Microbial Pathogenesis* **51(4)**:291–296 DOI 10.1016/j.micpath.2011.06.007.

Xu Z, Yue M, Zhou R, Jin Q, Fan Y, Bei W, Chen H. 2011. Genomic characterization of *Haemophilus parasuis* SH0165, a highly virulent strain of serovar 5 prevalent in China. *PLOS ONE* **6(5)**:e19631 DOI 10.1371/journal.pone.0019631.

Yu J, Wu J, Zhang Y, Du Y, Peng J, Chen L, Sun W, Cong X, Xu S, Shi J, Li J, Huang B, Zhu X, Wang J. 2014. Identification of putative virulence-associated genes among *Haemophilus parasuis* strains and the virulence difference of different serovars. *Microbial Pathogenesis* **77**:17–23 DOI 10.1016/j.micpath.2014.10.001.

Yue M, Yang F, Yang J, Bei W, Cai X, Chen L, Dong J, Zhou R, Jin M, Jin Q, Chen H. 2009. Complete genome sequence of *Haemophilus parasuis* SH0165. *Journal of Bacteriology* **191(4)**:1359–1360 DOI 10.1128/JB.01682-08.

Zehr ES, Lavrov DV, Tabatabai LB. 2012. Comparison of *Haemophilus parasuis* reference strains and field isolates by using random amplified polymorphic DNA and protein profiles. *BMC Microbiology* **12(1)**:108 DOI 10.1186/1471-2180-12-108.

Zhao Y, Wang Q, Li J, Lin X, Huang X, Fang B. 2018. Epidemiology of *Haemophilus parasuis* isolates from pigs in China using serotyping, antimicrobial susceptibility, biofilm formation and ERIC-PCR genotyping. *PeerJ* **6(5)**:e5040 DOI 10.7717/peerj.5040.

Zhou H, Yang B, Xu F, Chen X, Wang J, Blackall PJ, Zhang P, Xia Y, Zhang J, Ma R. 2010. Identification of putative virulence-associated genes of *Haemophilus parasuis* through suppression subtractive hybridization. *Veterinary Microbiology* **144(3–4)**:377–383 DOI 10.1016/j.vetmic.2010.01.023.