



Genotyping and investigating capsular polysaccharide synthesis gene loci of non-serotypeable *Streptococcus suis* isolated from diseased pigs in Canada

Han Zheng, Xiaotong Qiu, David Roy, Mariela Segura, Pengchen Du, Jianguo Xu, Marcelo Gottschalk

► To cite this version:

Han Zheng, Xiaotong Qiu, David Roy, Mariela Segura, Pengchen Du, et al.. Genotyping and investigating capsular polysaccharide synthesis gene loci of non-serotypeable *Streptococcus suis* isolated from diseased pigs in Canada. *Veterinary Research*, 2018, 48 (1), pp.10. 10.1186/s13567-017-0417-6 . hal-01499742

HAL Id: hal-01499742

<https://hal.science/hal-01499742>

Submitted on 31 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Open Access



Genotyping and investigating capsular polysaccharide synthesis gene loci of non-serotypeable *Streptococcus suis* isolated from diseased pigs in Canada

Han Zheng^{1†}, Xiaotong Qiu^{1†}, David Roy², Mariela Segura², Pengchen Du³, Jianguo Xu¹ and Marcelo Gottschalk^{2*}

Abstract

Streptococcus suis (*S. suis*) is an important swine pathogen and an emerging zoonotic agent. Most clinical *S. suis* strains express capsular polysaccharides (CPS), which can be typed by antisera using the coagglutination test. In this study, 79 *S. suis* strains recovered from diseased pigs in Canada and which could not be typed using antisera were further characterized by capsular gene typing and sequencing. Four patterns of *cps* locus were observed: (1) fifteen strains were grouped into previously reported serotypes but presented several mutations in their *cps* loci, when compared to available data from reference strains; (2) seven strains presented a complete deletion of the *cps* locus, which would result in an inability to synthesize capsule; (3) forty-seven strains were classified in recently described novel *cps* loci (NCLs); and (4) ten strains carried novel NCLs not previously described. Different virulence gene profiles (based on the presence of *mrp*, *epf*, and/or *sly*) were observed in these non-serotypeable strains. This study provides further insight in understanding the genetic characteristics of *cps* loci in non-serotypeable *S. suis* strains recovered from diseased animals. When using a combination of the previously described 35 serotypes and the complete NCL system, the number of untypeable strains recovered from diseased animals in Canada would be significantly reduced.

Introduction

Streptococcus suis is recognized as one of the most important causes of bacterial disease in post-weaned piglets worldwide, generating important economic losses to the swine industry. In addition, it is an important emerging zoonotic agent [1–3]. Clinical strains of *S. suis* generally have a capsule (capsular polysaccharide or CPS), which is the basis of the serotyping traditionally used for epidemiological studies. Thirty-five serotypes of *S. suis* (serotype 1 through 34 and serotype 1/2) were identified in the 1980s and the 1990s [4–7]. More recently, serotypes 20, 22, 26, 32, 33 and 34 have been suggested as

belonging to a species different from *S. suis* [8, 9]. Strains isolated from diseased pigs primarily belong to serotype 2 in most countries, followed by serotypes 3, 4, 5, 7, 8 and 1/2 [10–12]. In some European countries, serotype 9 is also one of the most frequently recovered capsular types from diseased animals [12, 13]. Traditionally, *S. suis* is routinely serotyped by the coagglutination test using serotype-specific antisera. However, non-serotypeable *S. suis* strains are frequently reported in many studies [12, 14–18]. Given that strains not expressing the CPS cannot be serotyped using antisera, serotyping based on molecular techniques has been proposed. Since the *S. suis* CPS is synthesized by the Wzx/Wzy pathway in the CPS locus, *wzy* genes have been demonstrated to be serotype-specific [19]. Thus, high-throughput capsular gene typing systems based on serotype-specific *wzy* genes have become attractive alternatives/complement to the existing serological tests [18, 20, 21]. However, even with the

*Correspondence: marcelo.gottschalk@umontreal.ca

[†]Han Zheng and Xiaotong Qiu contributed equally to this work

² Faculty of Veterinary Medicine, Swine and Poultry Infectious Diseases Research Center, University of Montreal, Quebec, Canada

Full list of author information is available at the end of the article

use of multiplex PCR tests, non-serotypeable strains are still commonly isolated from both clinically healthy and diseased animals [18, 22, 23].

In recent years, 17 novel *cps* loci (NCLs) were identified from non-serotypeable *S. suis* and were designated as NCL1 to 16 and serotype Chz [22–24]. Meanwhile, an 18-plex Luminex assay was also developed to detect these 17 NCLs and nearly 60% of non-serotypeable strains from healthy pigs carried one of these NCLs [22]. However, little is known about the distribution and characteristics of the *cps* loci of potentially virulent non-serotypeable strains recovered from diseased animals.

In this study, the *cps* loci of 79 Canadian non-serotypeable *S. suis* strains (as determined by the coagglutination test) recovered from diseased pigs were studied using two capsular gene typing systems [20, 22] and the genetic characteristics of the NCLs were analyzed. To elucidate the non-serotypeable mechanisms of strains grouped into previously described serotypes, the study was extended to compare their *cps* sequence to that of corresponding reference strains. Furthermore, the prevalence of minimum core genome (MCG) sequence typing group and virulence gene profile were also investigated in all 79 strains.

Materials and methods

Bacterial strains and chromosomal DNA preparation

A total of 79 *S. suis* strains isolated from diseased pigs on non-related farms in Canada were used in this study (Additional file 1). All strains have been isolated from primary affected organs of clinically diseased pigs, including brain (meningitis; $n = 18$), heart (endocarditis; $n = 18$), multiple organs (septicemia; $n = 14$), pleura (polyserositis; $n = 9$), lungs (pneumonia; $n = 9$) and joints (arthritis; $n = 1$). For a very few isolates, the information was not available, but they were all recovered from diseased animals with a primary diagnosis of *S. suis* infection. All isolates were serotyped using the coagglutination test [25]. Chromosomal DNA was prepared from all strains as previously described [21]. The species identity of the 79 strains was determined to be *S. suis* by amplification of the 16S rRNA, *recN*, *gdh*, and *thrA* genes [20, 26–28].

Capsular gene typing

The *cps* locus type of the 79 strains was identified by the 32-plex and 18-plex Luminex assays previously reported [20, 22]. The subtypes of known NCLs were determined based on the arrangement of subtype-specific homology groups (HGs) and transposases [22, 23].

Sequencing *cps* loci and bioinformatics analyses

Seventeen strains which could not be grouped using the 32- and 18-plex Luminex assays and 3 strains which

could not be grouped into known subtypes, as well as 15 strains which were grouped into reference serotypes, were sequenced by Illumina sequencing as previously described [23]. Each *cps* locus sequence was extracted from the draft genome sequence and was analyzed using the same bioinformatics methods described in previous studies [19, 22, 23]. The products of the *cps* genes were assigned to novel HGs if both of the global match regions and identity of the amino acid or nucleotide sequences were below 50% when compared to the 420 currently known HGs of the 35 reference serotypes and 17 NCLs. The novel HGs were assigned numerical values from HG421 onwards [19, 22, 23]. Novel HGs that were present in all strains of a given NCL were identified as NCL-specific HGs. The strains harboring the same *wzy* were clustered into the same NCL. The Artemis comparison tool (ACT) was used to visualize the data [29].

MCG typing and PCR assays for *mvp* (muramidase released protein), *sly* (suilysin) and *epf* (extracellular protein)

MCG sequence typing was performed using PCR amplification and DNA sequencing as previously described [30, 31]. The full-length *mvp* gene was amplified and sequenced using a previously described method [11]. Amplification of the *sly* and *epf* genes was performed according to methods previously described [32, 33].

Nucleotide sequence accession number

Sequences of *cps* loci obtained in this study were deposited in GenBank under the accession numbers KX870047–KX870056, KX870058–KX870064, KX870067–KX870072, and KX870074–KX870076. Reads of the sequenced strains were deposited in GenBank under accession number SRR5177663–SRR5177696 and SRR5177711. All accession numbers can also be found in Additional file 1.

Results

Serotyping of strains

The 79 strains used in the present study showed auto-agglutination, poly-agglutination or non-agglutination using the reference antisera and the coagglutination test and were thus considered as non-serotypeable. All strains were then typed using our previously developed capsular gene typing systems [20, 22]. Fifteen strains (18.9%) were grouped into reference serotypes while 47 (59.4%) were grouped into 17 known NCLs. The remaining 17 strains remained non-typeable (Additional file 1).

Of the 15 strains belonging to the previously described serotypes, serotype 2 or 1/2 ($n = 4$), which cannot be distinguished by capsular gene typing, was the most frequent, followed by serotypes 15 ($n = 3$), 11 ($n = 2$), and

30 ($n = 2$). Serotypes 5, 17, 27 and 29 only contained a single strain (Additional file 1).

Of the 47 strains which were assigned to previously known NCLs, NCL3 ($n = 18$) was the most prevalent, followed by NCL4 ($n = 8$), NCL7 ($n = 4$), NCL2 ($n = 3$), NCL12 ($n = 3$), and NCL13 ($n = 3$). In addition, one strain each of the NCL1, 5, 6, 10, 11, 14, 16, and Chz were also found (Additional file 1).

Identification of four new NCLs

The remaining 17 non-serotypeable strains mentioned above were sequenced by Illumina sequencing. The *cps* locus was absent from 7 strains. The *cps* loci of the remaining 10 strains were divided into four new NCLs which were named NCL17 to 20 based on their *wzy* gene sequences. NCL17 was the most prevalent ($n = 4$), followed by NCL18 ($n = 3$), NCL19 ($n = 2$), and NCL20 ($n = 1$) (Additional file 1). In addition, two types of patterns were found in the four new NCLs. NCL17 and NCL18 belonged to pattern I-a, while NCL19 and NCL20 belonged to pattern I-b [19].

The sizes of these NCLs ranged from 21.34 to 29.90 kb and the percentage of G+C content varied between 33.9 and 35.1%. Fifty-nine predicted coding sequences were designated *cps* HGs. Twenty-two HGs were also present in the *cps* loci of the reference strains of known serotypes and 17 known NCLs. An initial sugar transferase gene was located in the 5' region and was classified into three HGs: HG6 (NCL20), HG8 (NCL17), and HG295 (NCL18 and NCL19). The 5' regions of four NCLs were conserved, whereas the central and 3' regions of these were highly variable (Figure 1A).

Thirty-two HGs were NCL-specific. Each NCL contained 4–11 NCL-specific genes, with 4 HGs for NCL17, 7 HGs for NCL18, 11 HGs for NCL19, and 10 HGs for NCL20. Among these, 11 HGs encoded putative glycosyl transferases and two encoded acetyltransferases. As expected, all Wzy polymerases and Wzx flippases were NCL-specific (Additional file 2).

Determining the subtypes of NCLs

NCL2, NCL3, NCL7, and NCL11 strains were found to belong to a single subtype; NCL2-4, NCL3-1, NCL7-1, and NCL11-5, respectively. Genetic heterogeneity was not found within strains of NCL12, NCL15, and NCL17 to 20 (Additional file 2).

- i. NCL1: strain 1640373 could not be classified into any known NCL1 subtype and was sequenced by Illumina sequencing, named as NCL1-12. The replacement of HG293, HG294, and HG292 by the HG354 and HG355 was found in its three side regions (Figure 1B).

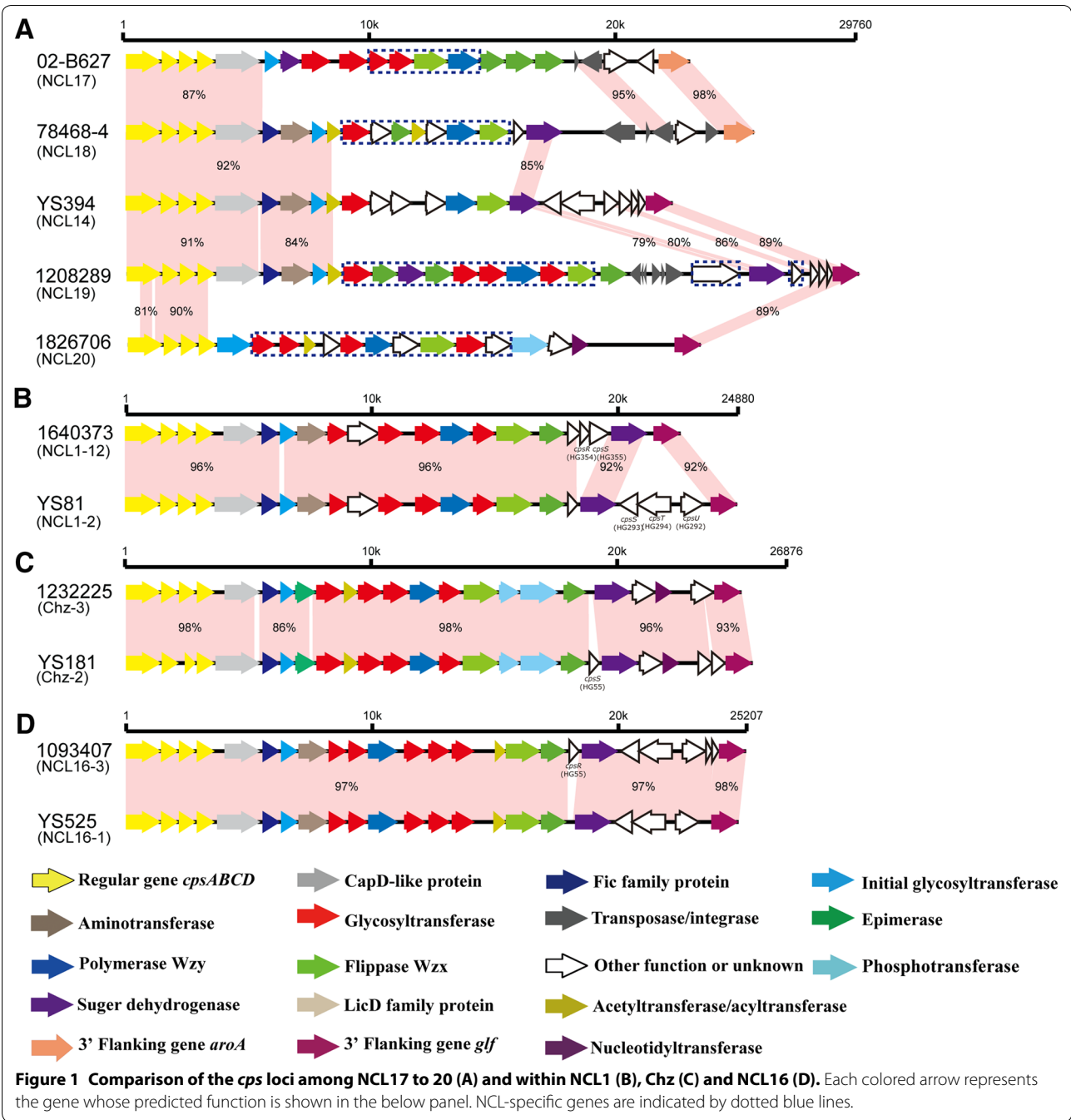
- ii. Chz: compared to the reference strain Chz-2, the deletion of HG55 was found in strain 1232225, named as Chz-3 (Figure 1C).

- iii. NCL16: compared to the reference strain YS525 (NCL16-1), the insertion of HG55 was found in strain 1093407, named NCL16-3 (Figure 1D).

Mutations in the *cps* loci of strains belonging to previously described serotypes

The 15 strains that were negative by coagglutination test but positive by multiplex Luminex assay for the reference serotypes were further analyzed. Comparing to the *cps* locus of the corresponding serotype reference strains, insertions and deletions were found in the serotype 5, 11, 15, 17 and 30 strains. The *cps* loci of four serotype 2 or 1/2 strains and one serotype 27 strain were intact and small-scale mutations were detected in these (Table 2).

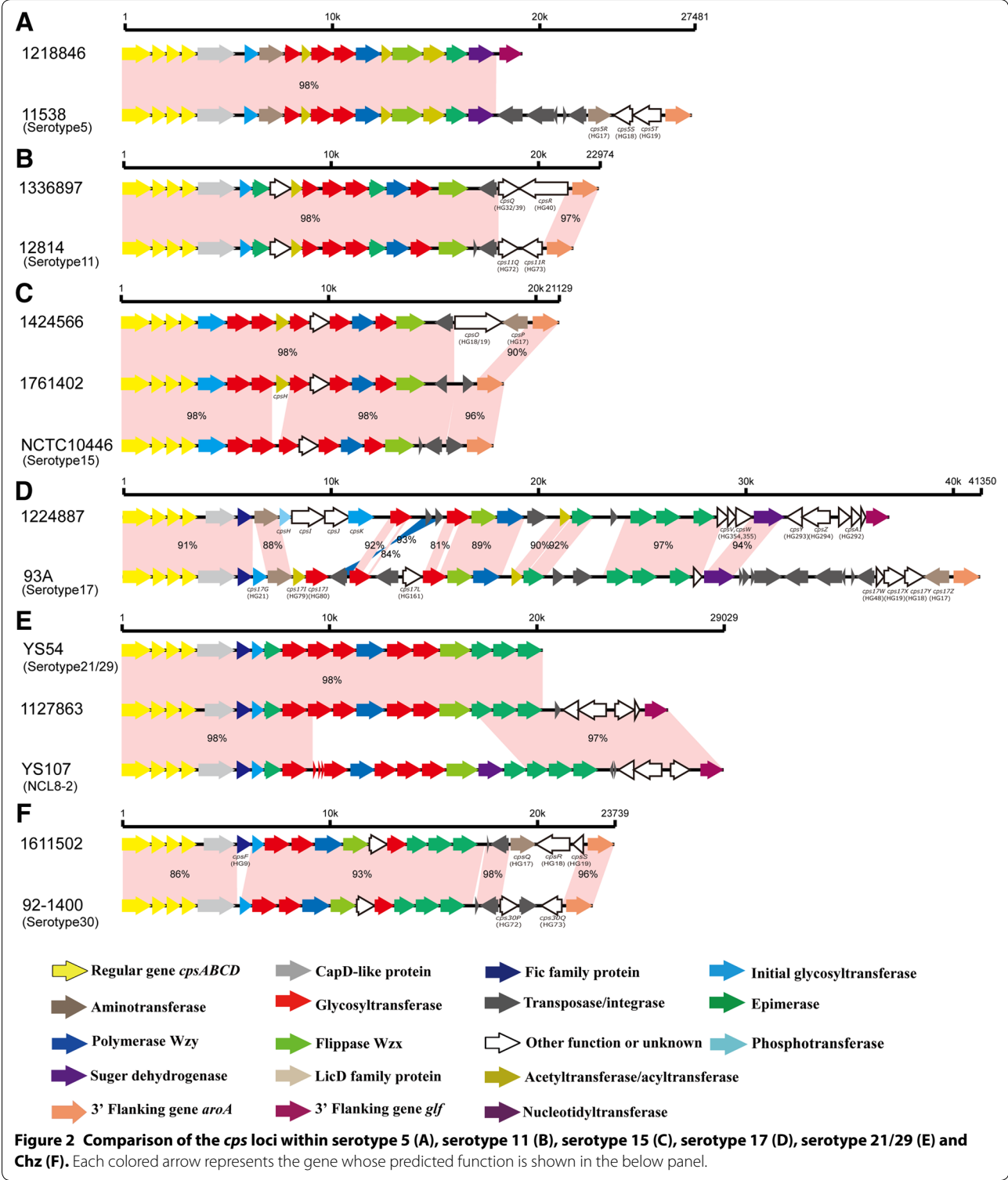
- i. Serotype 2 or 1/2: compared to the serotype 2 reference strain P1/7 (GenBank accession number BR001000), all four strains had a 33 bp insertion in *wxy* genes and four strains had single-nucleotide substitutions in *wzx* genes. The single-nucleotide substitutions in glycosyltransferase genes and a 27 bp deletion in the side-chain formation gene were also found in five strains (Table 2).
- ii. Serotype 5: compared to the serotype 5 reference strain 11538 (GenBank accession number BR001003), the deletions of HG17 to HG19 at the 3' end were found in strain 1218846 (Figure 2A).
- iii. Serotype 11: compared to the serotype 11 reference strain 12814 (GenBank accession number AB737819), HG72 and HG73 were replaced by HG32 and HG40 in strains 1336897 and 1336915. In addition, the nucleotide substitutions (TA→CC) of the termination codon of HG32 were found in the *cpsQ* gene of strains 1336897 and 1336915, which resulted in the chimeric HG32/HG39 gene (Figure 2B).
- iv. Serotype 15: two types of variations were found within this serotype. Strains 1424566 and 1449343 possessed identical *cps* sequences. A novel HG (*cpsH*, putative acetyltransferase) was inserted between HG33 and HG77, and the insertions of HG19 and HG17 at the 3' end were found in two strains (Table 1). Moreover, the transversion (T→G) was found at the site of the termination codon of HG19, which resulted in the chimeric HG18/HG19 gene in two strains. Compared to strains 1424566 and 1449343, HG18 and HG17 were replaced by a transposase in strain 1761402 (Figure 2C).
- v. Serotype 17: compared to the serotype 17 reference strain 93A (GenBank accession number



AB737824), two deletions (HG21 and HG161) and two insertions (HG354 and HG355) were found. Furthermore, HG79 and HG80 were replaced by a putative phosphotransferase, a putative hypothetical protein, a putative biotin carboxylase, and a putative glycosyltransferase (initial sugar transferase), which were not assigned to any previous homology group (Table 1). Moreover, the replacement of HG48, HG17, HG18, and HG19 by

HG293, HG294, and HG292 was also found (Figure 2D).

- vi. Serotype 27: compared to the serotype 27 reference strain 89–5259 (GenBank accession number AB737831), the single-nucleotide substitutions and small-scale indels in glycosyltransferase genes, *wzx* gene, and side-chain formation gene were found (Table 2).



vii. Serotype 29: compared to strain YS54 agglutinated with both serotypes 21 and 29 antisera (GenBank accession number KC537387), the insertions of a transposase gene, HG293, HG294, and HG292

on the 3'-side were found in strain 1127863 (Figure 2E).
viii. Serotype 30: compared to the serotype 30 reference strain 92-1400 (GenBank accession number

Table 1 Information of the novel HGs inserted *cps* loci of strains belonging to reference serotypes

Strain ID	<i>cps</i> locus type	Gene name	Predicted products	Similar protein, species (GenBank accession number)	Coverage/identity (%)
1424566, 1449343, 1761402	Serotype 15	<i>cpsH</i>	Acetyltransferase	Maltose O-acetyltransferase, <i>Lactobacillus reuteri</i> (CUR43586.1)	96/68
1224887	Serotype 17	<i>cpsH</i>	UDP-phosphate galactose phosphotransferase	UDP-phosphate galactose phosphotransferase, <i>Sphaerochaeta pleomorpha</i> (WP_014270310.1)	99/57
		<i>cpsI</i>	Hypothetical protein	Biotin carboxylase, <i>Sphaerochaeta pleomorpha</i> (WP_014270309.1)	68/48
		<i>cpsJ</i>	Biotin carboxylase	Biotin carboxylase, <i>Ruminococcus</i> sp. (CBL19829.1)	99/60
		<i>cpsK</i>	Glycosyltransferase	Glycosyltransferase family 1 protein, <i>Gallibacterium anatis</i> (WP_065231950.1)	92/58

AB737834), insertion of HG9 was found in strains 1611502 and 1839679. Moreover, HG72, transposase gene, and HG73 were replaced by HG17, HG18, and HG19 (Figure 2F).

Variations of chromosomal loci

In a previous study, the chromosomal loci of *cps* gene clusters of reference serotype 5 and 17 strains were classified into pattern I-a [19]. In the present study, strains 1218846 (serotype 5) and 1224887 (serotype 17) were classified into pattern I-b (Figures 2A and D, respectively).

MCG typing

The majority of the 79 strains were clustered in the MCG group 6 (44.3%, 35/79 strains), followed by ungroupable (24%, 19/79 strains), and group 7 (20.2%, 16/79 strains). MCG groups 4, 3, 2, and 1 also contained four, three, one, and one strains, respectively (Additional file 1).

Identifying genotypes of *mrp*, *epf*, and *sly*

Twelve strains were *mrp* positive. Frameshift mutations at 2740 bp from the reported initiator ATG codon were present in the *mrp* gene of strain 1114193, which resulted in premature stop codons. Eleven other strains contained intact full-length *mrp* gene copies and may express MRP. Based on the *mrp* subtypes reported in North America (NA) [11], the sequences of *mrp* were grouped into one of three subtypes, EU (European, $n = 3$), NA1 ($n = 7$), or NA2 ($n = 1$). Only twelve strains contained the *sly* gene and 6 strains were positive for *epf*. There were eight genotypes of *mrp*, *epf*, and *sly*, primarily based on *mrp* variation: most of the strains in this study ($n = 62$) were *mrp*⁻*sly*⁻*epf*⁻, followed by *mrp*^{NA1}*sly*⁻*epf*⁻ ($n = 4$), *mrp*^{NA1}*sly*⁺*epf*⁻ ($n = 4$), and *mrp*^{EU}*sly*⁺*epf*⁺ ($n = 3$). It is noteworthy that the latter strains were serotype 2 or 1/2,

serotype 15 and serotype 30, as revealed by the 32-plex Luminex assay. In addition, *mrp*⁻*sly*⁺*epf*⁺ ($n = 2$), *mrp*⁻*sly*⁺*epf*⁻ ($n = 2$), *mrp*^{NA2}*sly*⁺*epf*⁻ ($n = 1$), and *mrp*⁻*sly*⁻*epf*⁺ ($n = 1$) genotypes were also found (Additional file 1).

Discussion

In addition to the traditional 35 serotypes originally described for *S. suis*, 17 NCLs have recently been reported in non-serotypeable *S. suis* strains isolated from healthy animals using high-throughput typing systems and online bioinformatics [22, 23]. However, the genetic characteristics of *cps* loci in potentially virulent non-serotypeable *S. suis* strains recovered from diseased animals are still scarce.

In the present study, the *cps* loci of 79 Canadian non-serotypeable strains isolated from the internal organs of diseased pigs were analyzed. Non-serotypeable strains are frequently isolated from diseased animals in this country [34]. Based on previous gene typing and sequencing results [22, 35], the non-serotypeable phenotype may be attributed to one of three causes: (1) strains belonging to previously described serotypes harboring mutated *cps* loci causing loss of capsule expression or antigenic variation; (2) strains without *cps* locus completely losing their ability to synthesize capsule; or (3) strains with not-previously described NCL referring to novel serotypes.

In this study, 15 non-serotypeable strains could be grouped into reference serotypes by the 32-plex Luminex assay. To elucidate the lack of positive identification by the coagglutination test, we further sequenced and compared their *cps* loci to those of corresponding reference strains. Previous studies showed that replacements and large indels, as well as small-scale mutations of *cps* genes, caused phenotypical changes in agglutination tests [21,

Table 2 Mutations in glycosyltransferase, side-chain formation, wzy and wzx genes of serotype 2 or 1/2 representative strains and a serotype 27 strain

Strain	Affected gene(s)	Types of mutations	Affected nucleotide(s) [Affected amino acid]
1827702	<i>cps2E</i>	Missense	A61G [Thr21Ala]
	<i>cps2I</i>	Insertion	IS element: 33 bp
	<i>cps2N</i>	Deletion	27 bp
1090772	<i>cps2E</i>	Missense	A61G [Thr21Ala]
	<i>cps2F</i>	Missense	A149G [Asp50Gly]
		Missense	A1047T [Leu349Phe]
	<i>cps2I</i>	Insertion	IS element: 33 bp
	<i>cps2N</i>	Deletion	27 bp
1160406	<i>cps2O</i>	Missense	C859A [Arg287Ser]
	<i>cps27E</i>	Missense	A506G [Asp169Gly]
		Missense	T508A [Ser170Thr]
		Missense	A513T [Glu171Asp]
		Missense	A522T [Lys174Asn]
		Missense	A524T A525T [Lys175Ile]
		Missense	A541C C543G [Ile181Leu]
		Missense	A553G [Ile185Val]
		Missense	G617T T618G [Ser206Met]
		Missense	A623T [Tyr208Phe]
		Missense	T633G [Leu211Val]
		Missense	C640T A642C [Leu214Phe]
		Missense	A651T [Glu217Asp]
		Missense	C692T [Ser231Leu]
		Missense	G706A A708G [Ala236Thr]
		Missense	G874A A876T [Val292Ile]
		Missense	T905C [Val302Ala]
		Missense	A922C A924G [Lys308Gln]
		Missense	A941C G942A [Lys314Thr]
		Missense	A967G [Ile323Val]
		Missense	A997C G999C [Met333Leu]
		Missense	G1000A C1001G T1002C [Ala334Ser]
		Missense	A1174C A1175G [Lys392Arg]
		Missense	G1186A G1187T T1188G [Gly396Met]
		Missense	A1272C [Glu424Asp]
		Missense	A1279C A1280G [Lys427Arg]
		Missense	G1288T T1290A [Val430Leu]
		Missense	A1292C [Glu431Ala]
		Missense	G1324A A1326T [Val442Ile]
		Missense	A1357T T1359G [Ile453Leu]
		Missense	A1360T A1361T [Lys454Leu]
	<i>cps27F</i>	Deletion	27 bp
		Missense	A662T [His221Leu]
	<i>cps27G</i>	Insertion	IS element: 21 bp
	<i>cps27I</i>	Missense	G483C [Trp161Cys]
		Missense	C513A [Asp171Glu]
		Missense	A611G [Gln204Arg]
	<i>cps27L</i>	Deletion	57 bp
	<i>cps27M</i>	Missense	C1373T C1374T [Ala458Val]
		Missense	C1375T [Leu459Phe]

36–38]. We also found similar mutations in the *cps* loci of strains tested.

HG17, HG18, HG19, HG32, HG39, and HG40, which were present in the *cps* loci of the reference strains belonging to serotypes 1, 2, 4, 5, 7, 14, 17, 18, 19, 23 and 1/2, were detected in the *cps* loci of strains in the present study belonging to serotypes 11, 15, and 30. Moreover, chimeric HG18/HG19 and HG32/HG39 genes were found in serotype 11 and 15 strains. It is noteworthy that HG292, HG293, HG294, HG354, and HG355, only present in the *cps* loci of NCLs, were also detected in the *cps* locus of strain 1224889, typed herein as serotype 17. In addition, some genes which were never before assigned to any HG were found to be inserted in the *cps* loci of strains belonging to serotypes 17 and 15. The sequence differences between strains NCL8-2 and 1127863 were mainly caused by the replacement of 8 NCL-specific HGs in the center of NCL8-2 and by 6 HGs in 1127863. The replacement and insertion activities may indicate recombination events or horizontal gene transfer between the *cps* loci of *S. suis* strains, probably leading to antigenic variations that would be beneficial to *S. suis* in the course of infection or through immunity evasion.

Comparing to the *cps* loci of their corresponding reference strains, only small-scale mutations were observed in four strains typed as serotype 2 or 1/2 by the 32-plex Luminex. Previous study revealed that all serotype 2 and all serotype 14 strains had a G nucleotide at position 483 of the *cpsK* gene, while all serotype 1 and all serotype 1/2 strains (including 13 serotype 1/2 strains recovered in Canada) contained either a C or T at that nucleotide position [39]. In present study, all four strains had a G nucleotide at position 483 of the *cpsK* gene. We postulated that they were most probably non-encapsulated serotype 2 strains. A previous study reported that single-nucleotide substitutions and frameshift mutations in two glycosyltransferase genes (*cps2E* and *cps2F*) were the main causes of capsule loss in serotype 2 strains. Moreover, mutations in the genes involved in side-chain formation (*cps2J* and *cps2N*), *wzy* (*cps2I*), and *wzx* (*cps2O*) also appeared to be lethal to serotype 2 strains [36]. It may be hypothesized that the missense mutations and small scale indels found in these genes in strains of the present study also had a deleterious effect on the capsular expression. Indeed, high hydrophobic indexes have been obtained with these strains (unpublished data), which strongly suggest lack of capsule expression [15, 40]. Although non-encapsulated *S. suis* strains had originally been considered to be avirulent, they are frequently isolated from cases of endocarditis; as such, non-encapsulation could be, under certain circumstances, beneficial for *S. suis* in the course of such infections [36, 41]. In some cases, non-encapsulated strains resulting from small point mutations may switch

to a capsulated phenotype in vivo [42]. Interestingly, small-scale mutations or clear deletions of *cps* loci were also found in an additional eight strains, which are also probably non-encapsulated. Finally, strains without *cps* locus completely losing their ability to synthesize capsule were also found in this study. It is possible that these strains are not able to reverse the encapsulated phenotype. The biological and pathological significance of these non-encapsulated strains need to be further evaluated. Although never described, it is not impossible that some strains lose the capsular phenotype after in vitro culture.

In this study, 60% of non-serotypeable strains carried one of the recently described NCLs. The most common NCLs were 3, 4, 7, and 17, whereas in a previous study with strains recovered from healthy pigs in China, the most common NCLs were 1, 2, 3, and 7 [22]. Differences may be due to the geographical origin of strains (Canada vs. China) and/or their virulence potential (strains from diseased or clinically healthy animals). Since many strains of NCL3 have been identified in this study, further research on its virulence potential should be performed. In addition, and similarly to a previous study [22], high diversity within the same NCL was observed. The *S. suis* species is composed of phenotypically and genetically diverse strains. Host specificity and ecological environment may contribute to this diversity. The *cps* loci could provide important information regarding the ecology of strains. The differences in dominant NCLs between clinical strains from Canada and field strains from China and the emergence of novel NCLs or subtypes in clinical strains from Canada are expected.

In this study, new NCLs (CNL17–20), distributed in 10 strains, are reported for the first time. These NCLs possess completely different Wzy and transferases from those of the previously reported serotypes and NCLs, which in turn may express unique oligosaccharide structures and antigen identities. It is noteworthy that, taking into consideration all NCLs, more than 70% of non-typeable strains could now be typed. The use of the complete serotyping and NCL typing system would considerably reduce the number of non-typeable strains recovered from diseased animals in Canada.

The presence of some genes, such as *mrp*, *epf*, and *sly*, has been associated with virulence [43, 44]. Three distinct *mrp* genotypes have been reported so far and NA1 was the dominant genotype in *S. suis* strains recovered from diseased pigs in the USA [11]. In the present study, 11 strains possessed an intact *mrp* gene and NA1 genotype whereas three strains harbored the EU genotype. One of latter strains was typed as being a serotype 2 or 1/2 by the 32-plex Luminex but, as mentioned above, it is probably a real serotype 2 as shown by the presence of a G nucleotide at position 483 of the *cpsK* gene. The fact that most

mrp⁺, *epf*⁺, and *sly*⁺ Eurasian serotype 2 strains belong to the clonal complex 1 [12] also indicated the strain is most probably a non-encapsulated serotype 2 strain with an Eurasian profile [11] that might have been introduced to North America through the importation of animals. In fact, it has been reported that up to 5% of serotype 2 strains recovered in the United States are ST1 and probably originated from Europe [11]. Although the most prevalent virulence gene profile was *mrp*⁺*sly*⁺*epf*⁺, 17 strains contained at least one of these three genes. The relevance of these virulence markers in strains of serotypes different from serotype 2 is still controversial.

The most prevalent MCG groups amongst the strains harboring NCLs were the groups 6 and 7, which had been described as being the most ancient groups in the *S. suis* population [30]. This indicates that their *cps* loci have existed for a long time and play important roles in the serotype diversity of *S. suis* population. The most prevalent MCG groups amongst the strains harboring mutated *cps* loci of previously described serotypes and the strains losing their *cps* loci were MCG ungroupable. These strains possibly had a more significant recombination history that prevented them from being reliably assigned; meanwhile these recombination events may facilitate the mutations and loss of their *cps* loci.

In conclusion, this study provides further insight in understanding the *cps* diversity of *S. suis* and may contribute to future epidemiological studies that will allow characterization of potentially virulent and previously non-serotypeable strains isolated from diseased animals. Use of the 35 serotype-based system complemented with the NCL typing system would significantly reduce the number of untypeable strains recovered from diseased animals in Canada. Further studies with *S. suis* strains isolated in other countries are needed.

Additional files

Additional file 1. Strains of *S. suis* used in the present study. List and identification of strains used in the present study.

Additional file 2. Predicted product and homology group in each NCL. Predicted product of coding genes and homology group in each characterized NCL.

Abbreviations

S. suis: *Streptococcus suis*; CPS: capsular polysaccharides; NCL: novel *cps* loci; MCG: minimum core genome; *mrp*: muramidase released protein; *sly*: sulysin; *epf*: extracellular protein factor.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

HZ, MS, JX, and MG conceived and designed this study. HZ, XQ, and DR contributed to the majority of laboratory experiments. HZ, XQ and PD interpreted

the data. XQ prepared the manuscript. HZ, MS, and MG reviewed and critically revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by Grant (81572044) from the National Natural Science Foundation of China, "the Fundamental Research Funds for the Central Universities" from the Key Laboratory of Molecular Microbiology and Technology, Ministry of Education TEDA Institute of Biological Sciences and Biotechnology, Nankai University, and by Grants 125684 from the Canadian Institutes of Health Research (China-Canada Joint Health Research Initiative) and 154280 from the Natural Sciences and Engineering Research Council of Canada.

Author details

¹ State Key Laboratory for Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Changping, Beijing, China. ² Faculty of Veterinary Medicine, Swine and Poultry Infectious Diseases Research Center, University of Montreal, Quebec, Canada. ³ Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing Key Laboratory of Emerging Infectious Diseases, Beijing, People's Republic of China.

Received: 20 November 2016 Accepted: 20 January 2017

Published online: 20 February 2017

References

- Gottschalk M, Segura M, Xu J (2007) *Streptococcus suis* infections in humans: the Chinese experience and the situation in North America. *Anim Health Res Rev* 8:29–45
- Ye C, Zhu X, Jing H, Du H, Segura M, Zheng H, Kan B, Wang L, Bai X, Zhou Y, Cui Z, Zhang S, Jin D, Sun N, Luo X, Zhang J, Gong Z, Wang X, Wang L, Sun H, Li Z, Sun Q, Liu H, Dong B, Ke C, Yuan H, Wang H, Tian K, Wang Y, Gottschalk M, Xu J (2006) *Streptococcus suis* sequence type 7 outbreak, Sichuan, China. *Emerg Infect Dis* 12:1203–1208
- Yu H, Jing H, Chen Z, Zheng H, Zhu X, Wang H, Wang S, Liu L, Zu R, Luo L, Xiang N, Liu H, Liu X, Shu Y, Lee SS, Chuang SK, Wang Y, Xu J, Yang W (2006) Human *Streptococcus suis* outbreak, Sichuan, China. *Emerg Infect Dis* 12:914–920
- Gottschalk M, Higgins R, Jacques M, Beaudoin M, Henrichsen J (1991) Characterization of six new capsular types (23 through 28) of *Streptococcus suis*. *J Clin Microbiol* 29:2590–2594
- Gottschalk M, Higgins R, Jacques M, Mittal KR, Henrichsen J (1989) Description of 14 new capsular types of *Streptococcus suis*. *J Clin Microbiol* 27:2633–2636
- Higgins R, Gottschalk M, Boudreau M, Lebrun A, Henrichsen J (1995) Description of six new capsular types (29–34) of *Streptococcus suis*. *J Vet Diagn Invest* 7:405–406
- Perch B, Pedersen KB, Henrichsen J (1983) Serology of capsulated streptococci pathogenic for pigs: six new serotypes of *Streptococcus suis*. *J Clin Microbiol* 17:993–996
- le Tien HT, Nishibori T, Nishitani Y, Nomoto R, Osawa R (2013) Reappraisal of the taxonomy of *Streptococcus suis* serotypes 20, 22, 26, and 33 based on DNA-DNA homology and *sodA* and *recN* phylogenies. *Vet Microbiol* 162:842–849
- Hill JE, Gottschalk M, Brousseau R, Harel J, Hemmingsen SM, Goh SH (2005) Biochemical analysis, *cpn60* and 16S rDNA sequence data indicate that *Streptococcus suis* serotypes 32 and 34, isolated from pigs, are *Streptococcus orisratti*. *Vet Microbiol* 107:63–69
- Wei Z, Li R, Zhang A, He H, Hua Y, Xia J, Cai X, Chen H, Jin M (2009) Characterization of *Streptococcus suis* isolates from the diseased pigs in China between 2003 and 2007. *Vet Microbiol* 137:196–201
- Fittipaldi N, Fuller TE, Teel JF, Wilson TL, Wolfram TJ, Lowery DE, Gottschalk M (2009) Serotype distribution and production of muramidase-released protein, extracellular factor and sulysin by field strains of *Streptococcus suis* isolated in the United States. *Vet Microbiol* 139:310–317
- Goyette-Desjardins G, Auger JP, Xu J, Segura M, Gottschalk M (2014) *Streptococcus suis*, an important pig pathogen and emerging zoonotic

- agent—an update on the worldwide distribution based on serotyping and sequence typing. *Emerg Microbes Infect* 3:e45
13. Wisselink HJ, Smith HE, Stockhofe-Zurwieden N, Peperkamp K, Vecht U (2000) Distribution of capsular types and production of muramidase-released protein (MRP) and extracellular factor (EF) of *Streptococcus suis* strains isolated from diseased pigs in seven European countries. *Vet Microbiol* 74:237–248
 14. Sanchez del Rey V, Fernandez-Garayzabal JF, Briones V, Iriso A, Dominguez L, Gottschalk M, Vela AI (2013) Genetic analysis of *Streptococcus suis* isolates from wild rabbits. *Vet Microbiol* 165:483–486
 15. Gottschalk M, Lacouture S, Bonifait L, Roy D, Fittipaldi N, Grenier D (2013) Characterization of *Streptococcus suis* isolates recovered between 2008 and 2011 from diseased pigs in Quebec, Canada. *Vet Microbiol* 162:819–825
 16. Marois C, Le Devendec L, Gottschalk M, Kobisch M (2007) Detection and molecular typing of *Streptococcus suis* in tonsils from live pigs in France. *Can J Vet Res* 71:14–22
 17. Han DU, Choi C, Ham HJ, Jung JH, Cho WS, Kim J, Higgins R, Chae C (2001) Prevalence, capsular type and antimicrobial susceptibility of *Streptococcus suis* isolated from slaughter pigs in Korea. *Can J Vet Res* 65:151–155
 18. Okura M, Lachance C, Osaki M, Sekizaki T, Maruyama F, Nozawa T, Nakagawa I, Hamada S, Rossignol C, Gottschalk M, Takamatsu D (2014) Development of a two-step multiplex PCR assay for typing of capsular polysaccharide synthesis gene clusters of *Streptococcus suis*. *J Clin Microbiol* 52:1714–1719
 19. Okura M, Takamatsu D, Maruyama F, Nozawa T, Nakagawa I, Osaki M, Sekizaki T, Gottschalk M, Kumagai Y, Hamada S (2013) Genetic analysis of capsular polysaccharide synthesis gene clusters from all serotypes of *Streptococcus suis*: potential mechanisms for generation of capsular variation. *Appl Environ Microbiol* 79:2796–2806
 20. Bai X, Liu Z, Ji S, Gottschalk M, Zheng H, Xu J (2015) Simultaneous detection of 33 *Streptococcus suis* serotypes using the luminex xTAG(R) assay. *J Microbiol Methods* 117:95–99
 21. Liu Z, Zheng H, Gottschalk M, Bai X, Lan R, Ji S, Liu H, Xu J (2013) Development of multiplex PCR assays for the identification of the 33 serotypes of *Streptococcus suis*. *PLoS One* 8:e72070
 22. Qiu X, Bai X, Lan R, Zheng H, Xu J (2016) Novel capsular polysaccharide loci and new diagnostic tools for high-throughput capsular gene typing in *Streptococcus suis*. *Appl Environ Microbiol* 82:7102–7112
 23. Zheng H, Ji S, Liu Z, Lan R, Huang Y, Bai X, Gottschalk M, Xu J (2015) Eight novel capsular polysaccharide synthesis gene loci identified in nontypeable *Streptococcus suis* isolates. *Appl Environ Microbiol* 81:4111–4119
 24. Pan Z, Ma J, Dong W, Song W, Wang K, Lu C, Yao H (2015) Novel variant serotype of *Streptococcus suis* isolated from piglets with meningitis. *Appl Environ Microbiol* 81:976–985
 25. Gottschalk M, Higgins R, Boudreau M (1993) Use of polyvalent coagglutination reagents for serotyping of *Streptococcus suis*. *J Clin Microbiol* 31:2192–2194
 26. Ishida S, Ie Tien HT, Osawa R, Tohya M, Nomoto R, Kawamura Y, Takahashi T, Kikuchi N, Kikuchi K, Sekizaki T (2014) Development of an appropriate PCR system for the reclassification of *Streptococcus suis*. *J Microbiol Methods* 107:66–70
 27. Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. *J Microbiol Methods* 55:541–555
 28. Okwumabua O, O'Connor M, Shull E (2003) A polymerase chain reaction (PCR) assay specific for *Streptococcus suis* based on the gene encoding the glutamate dehydrogenase. *FEMS Microbiol Lett* 218:79–84
 29. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J (2005) ACT: the Artemis comparison tool. *Bioinformatics* 21:3422–3423
 30. Chen C, Zhang W, Zheng H, Lan R, Wang H, Du P, Bai X, Ji S, Meng Q, Jin D, Liu K, Jing H, Ye C, Gao GF, Wang L, Gottschalk M, Xu J (2013) Minimum core genome sequence typing of bacterial pathogens: a unified approach for clinical and public health microbiology. *J Clin Microbiol* 51:2582–2591
 31. Zheng H, Ji S, Lan R, Liu Z, Bai X, Zhang W, Gottschalk M, Xu J (2014) Population analysis of *Streptococcus suis* isolates from slaughtered swine by use of minimum core genome sequence typing. *J Clin Microbiol* 52:3568–3572
 32. Staats JJ, Plattner BL, Stewart GC, Changappa MM (1999) Presence of the *Streptococcus suis* suilysin gene and expression of MRP and EF correlates with high virulence in *Streptococcus suis* type 2 isolates. *Vet Microbiol* 70:201–211
 33. Wisselink HJ, Reek FH, Vecht U, Stockhofe-Zurwieden N, Smits MA, Smith HE (1999) Detection of virulent strains of *Streptococcus suis* type 2 and highly virulent strains of *Streptococcus suis* type 1 in tonsillar specimens of pigs by PCR. *Vet Microbiol* 67:143–157
 34. Gottschalk M, Lacouture S (2015) Canada: distribution of *Streptococcus suis* (from 2012 to 2014) and *Actinobacillus pleuropneumoniae* (from 2011 to 2014) serotypes isolated from diseased pigs. *Can Vet J* 56:1093–1094
 35. Salter SJ, Hinds J, Gould KA, Lamberts L, Hanage WP, Antonio M, Turner P, Hermans PW, Bootsma HJ, O'Brien KL, Bentley SD (2012) Variation at the capsule locus, cps, of mistyped and non-typable *Streptococcus pneumoniae* isolates. *Microbiology* 158:1560–1569
 36. Lakkitjaroen N, Takamatsu D, Okura M, Sato M, Osaki M, Sekizaki T (2014) Capsule loss or death: the position of mutations among capsule genes sways the destiny of *Streptococcus suis*. *FEMS Microbiol Lett* 354:46–54
 37. Morona JK, Morona R, Paton JC (1999) Comparative genetics of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* types belonging to serogroup 19. *J Bacteriol* 181:5355–5364
 38. Yun KW, Cho EY, Choi EH, Lee HJ (2014) Capsular polysaccharide gene diversity of pneumococcal serotypes 6A, 6B, 6C, and 6D. *Int J Med Microbiol* 304:1109–1117
 39. Athey TB, Teatero S, Lacouture S, Takamatsu D, Gottschalk M, Fittipaldi N (2016) Determining *Streptococcus suis* serotype from short-read whole-genome sequencing data. *BMC Microbiol* 16:162
 40. Bonifait L, Gottschalk M, Grenier D (2010) Cell surface characteristics of nontypeable isolates of *Streptococcus suis*. *FEMS Microbiol Lett* 311:160–166
 41. Lakkitjaroen N, Takamatsu D, Okura M, Sato M, Osaki M, Sekizaki T (2011) Loss of capsule among *Streptococcus suis* isolates from porcine endocarditis and its biological significance. *J Med Microbiol* 60:1669–1676
 42. Auger JP, Meekhanon N, Okura M, Osaki M, Gottschalk M, Sekizaki T, Takamatsu D (2016) *Streptococcus suis* serotype 2 capsule in vivo. *Emerg Infect Dis* 22:1793–1796
 43. Takeuchi D, Akeda Y, Nakayama T, Kerdin A, Sano Y, Kanda T, Hamada S, Dejsirilert S, Oishi K (2014) The contribution of suilysin to the pathogenesis of *Streptococcus suis* meningitis. *J Infect Dis* 209:1509–1519
 44. Vecht U, Wisselink HJ, Jellema ML, Smith HE (1991) Identification of two proteins associated with virulence of *Streptococcus suis* type 2. *Infect Immun* 59:3156–3162