Study of the population structure of *Haemophilus parasuis* by multilocus sequence typing

Alex Olvera, Marta Cerdà-Cuellar and Virginia Aragon

Correspondence

Virginia Aragon

virginia.aragon@cresa.uab.es

Centre de Recerca en Sanitat Animal (CRESA), Campus de Bellaterra-Universitat Autònoma de Barcelona, 08193-Bellaterra, Barcelona, Spain

Haemophilus parasuis is the aetiological agent of Glässer's disease in swine. In addition, this bacterium causes other clinical outcomes and can also be isolated from the upper respiratory tract of healthy pigs. Isolates of *H. parasuis* differ in phenotypic features (e.g. protein profiles, colony morphology or capsule production) and pathogenic capacity. Differences among strains have also been demonstrated at the genetic level. Several typing methods have been used to classify *H. parasuis* field strains, but they had resolution or implementation problems. To overcome these limitations, a multilocus sequence typing (MLST) system, using partial sequences of the house-keeping genes *mdh*, *6pgd*, *atpD*, *g3pd*, *frdB*, *infB* and *rpoB*, was developed. Eleven reference strains and 120 field strains were included in this study. The number of alleles per locus ranged from 14 to 41, *6pgd* being the locus with the highest diversity. The high genetic heterogeneity of this bacterium was confirmed with MLST, since the strains were divided into 109 sequence types, and only 13 small clonal complexes were detected by the Burst algorithm. Further analysis by unweighted-pair group method with arithmetic mean (UPGMA) identified six clusters. When the clinical background of the isolates was examined, one cluster was statistically associated with nasal isolation (putative non-virulent), while another cluster showed a significant association with isolation from clinical lesions (putative virulent). The remaining clusters did not show a statistical association with the clinical background of the isolates. Finally, although recombination among *H. parasuis* strains was detected, two divergent branches were found when a neighbour-joining tree was constructed with the concatenated sequences. Interestingly, one branch included almost all isolates of the putative virulent UPGMA cluster.

Received 27 June 2006

Revised 24 August 2006

Accepted 29 August 2006

INTRODUCTION

Haemophilus parasuis is a member of the family *Pasteurellaceae* and the causative agent of Glässer's disease in pigs, which is pathologically characterized by fibrinous to fibrinopurulent polyserositis and polyarthritis (Rapp-Gabrielson *et al.*, 2006). In addition to Glässer's disease, *H. parasuis* produces other clinical outcomes, such as pneumonia, and colonizes the upper respiratory tract of healthy animals (Rapp-Gabrielson *et al.*, 2006). Although it is commonly accepted that one strain is responsible for each clinical outbreak, diagnosis of *H. parasuis* infection is complicated by the fact that it is usual to detect several

strains in a farm and even within a single animal. Therefore, it is essential to determine the causative strain by its isolation from organs with the characteristic lesions of the disease.

Differences among strains in phenotypic and genotypic characteristics have been reported, although no clear association with virulence could be determined (Oliveira & Pijoan, 2004; Rapp-Gabrielson *et al.*, 2006). However, several studies have confirmed that different strains of *H. parasuis* have different pathogenic capacity (Kielstein & Rapp-Gabrielson, 1992; Nielsen, 1993; Rapp-Gabrielson *et al.*, 1992; Vahle *et al.*, 1995). Classically, strains of *H. parasuis* have been classified by serotyping, and although this method is useful for vaccine implementation, it is not discriminative enough for epidemiology. Moreover, a high percentage of strains are non-typable by serotyping (Oliveira & Pijoan, 2004; Rapp-Gabrielson *et al.*, 2006). Recently, different genotyping methods have been proposed to differentiate *H. parasuis* strains. The majority of them are fingerprinting methods, and, even though the reported techniques have a higher level of discrimination than serotyping, they present application problems, such as limited resolution (de la Puente-Redondo *et al.*, 2000,

Abbreviations: CC, clonal complex; d_N-d_S , mean difference between non-synonymous and synonymous codon changes; I_A , index of association; MLST, multilocus sequence typing; NJ, neighbour-joining; ST, sequence type; UPGMA, unweighted-pair group method with arithmetic mean.

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this work are DQ781411–DQ782327.

A supplementary table giving strain details is available with the online version of this paper.

2003; del Rio *et al.*, 2006) or difficulty in comparing results from different laboratories (Rafiee *et al.*, 2000; Smart *et al.*, 1988). To improve the epidemiological study of *H. parasuis* strains, a single-locus sequence typing method was recently used by our group (Olvera *et al.*, 2006). The heterogeneity of *H. parasuis* field isolates was confirmed by *hsp60* partial sequencing, but although a virulent cluster was detected, the classification of the strains was not satisfactory. Moreover, the results of our study with *hsp60* and the 16S rRNA gene indicated a possible lateral gene transfer within *H. parasuis* strains and between *H. parasuis* and *Actinobacillus* spp. Thus, to achieve robustness against the effects of recombination and maintain an adequate resolution, we have developed a multilocus sequence typing (MLST) for *H. parasuis*.

MLST is based on the sequencing of 450–600 bp fragments of core genes and the assignment of allelic profiles, which leads to sequence types (ST). The advantages of MLST for local and global epidemiology have been extensively discussed elsewhere (Cooper & Feil, 2004; Enright & Spratt, 1999; Maiden *et al.*, 1998; Spratt, 1999). MLST has been successfully used for the determination of clonal complexes (CC) of several human and animal pathogens (Dingle *et al.*, 2001; Enright & Spratt, 1998; Enright *et al.*, 2001; Feavers *et al.*, 1999; Heym *et al.*, 2002; Homan *et al.*, 2002; King *et al.*, 2002; Kriz *et al.*, 2002; Lemee *et al.*, 2004; Nallapareddy *et al.*, 2002; Noller *et al.*, 2003; Shi *et al.*, 1998; van Loo *et al.*, 2002; Wang *et al.*, 2003), including *Haemophilus influenzae* (Meats *et al.*, 2003). Since the genomic sequence of *H. parasuis* is not available, we used primers designed for *H. influenzae* (Meats *et al.*, 2003), universal primers (Christensen *et al.*, 2004), or primers designed to areas of homology of the selected genes in other bacteria, including *Pasteurellaceae*. On the other hand, a major problem of current MLST databases is the poor representation of non-pathogenic isolates within species of clinical interest. In some cases, non-virulent strains can represent a significant part of the population, which is important in order to define population structures and to estimate population parameters (Perez-Losada *et al.*, 2006). For that reason, an effort was made to have a representative sample of the natural population of *H. parasuis* by also sampling asymptomatic carriers.

METHODS

Bacterial strains. This study included 120 *H. parasuis* field isolates and 11 reference strains. The online version of this paper (at <http://mic.sgmjournals.org>) contains a supplementary table with relevant data from all the strains. Field strains were isolated from lungs or systemic sites from pigs with clinical lesions (57 strains; referred to as clinical isolates throughout the text for clarity) and from the nasal cavity of piglets from farms with or without Glässer's disease (74 strains; referred to as nasal isolates throughout the text for clarity). To obtain nasal isolates, 19 farms, in four separate regions of Spain, were selected based on their health status. Eight to ten nasal swabs were taken from each farm and transported in Amies medium to the laboratory, where they were plated on chocolate agar to isolate colonies. Isolation and identification of *H. parasuis* were performed as previously described (Olvera *et al.*, 2006). Clinical strains were

isolated from characteristic lesions from diseased animals or kindly provided by the Department of Infectious Diseases of the Veterinary School of the Universitat Autònoma de Barcelona (Spain), by Dr E. Rodríguez Ferri (Universidad de León, Spain), by Dr Gustavo C. Zielinski (Instituto Nacional de Tecnología Agropecuaria-INTA, Argentina), Dr Øystein Angen (Danish Institute for Food and Veterinary Research, Denmark) and by Dr T. Blaha (Federal Institute for Health Protection of Consumers and Veterinary Medicine, Germany). All the strains were maintained in 20% glycerol-Brain Heart Infusion at -80°C and routinely cultured in chocolate agar plates at 37°C with 5% CO_2 for 24–48 h.

DNA extraction, primers and PCR conditions. Genomic DNA from each strain was extracted using a commercial kit (Nucleospin Blood, Macherey-Nagel) following the manufacturer's instructions.

Primers were preliminarily tested with the 11 *H. parasuis* reference strains. All the primers of the *H. influenzae* MLST (Meats *et al.*, 2003) were tested, but finally only the primers for the malate dehydrogenase gene (*mdh*) were useful for *H. parasuis*. Previously published primers for *rpoB*, *atpD* and *infB* were also tested (Christensen *et al.*, 2004). The remaining primers were designed by homology with other bacterial genes (sequences from *Pasteurellaceae* were used when available). Target genes were those for the β chain of ATP synthase (*atpD*), 60 kDa heat-shock protein (*hsp60*), translation initiation factor IF-2 (*infB*), ribosomal protein β subunit (*rpoB*), superoxide dismutase A (*sodA*), phosphoglucosyltransferase (*pgm*), 6-phosphogluconate dehydrogenase (*6pgd*), glyceraldehyde-3-phosphate dehydrogenase (*g3pd*) and fumarate reductase B (*frdB*). Primers amplifying all reference strains without non-specific bands were selected. The seven selected were *atpD*, *infB*, *mdh*, *rpoB*, *6pgd*, *g3pd* and *frdB* (Table 1). All PCR amplifications were carried out in a final volume of 50 μl containing 1.5 U *Taq* polymerase and 200 μM dNTP. MgCl_2 and primer concentration were optimized in order to perform all the PCR reactions under the same cycling conditions (Table 1). Cycling conditions were 5 min at 95°C , 35 cycles of 1 min at 95°C , 30 s at 50°C and 30 s at 72°C , followed by a final step of 10 min at 72°C .

DNA sequencing and data analysis. After PCR, 5 μl aliquots of the reactions were visualized in a 2% agarose gel to confirm the absence of non-specific bands. Amplicons were then purified using Nucleofast 96 PCR kit (Macherey-Nagel) and 1 μl product was sequenced using the corresponding primers (Table 1), BigDye terminator v.3.1 kit and the ABI 3100 DNA sequencer (Applied Biosystems).

Fingerprinting II v.3.0 software (Bio-Rad) was used to edit, assemble and align the sequences and to carry out allele assignment. Congruence between loci was calculated by Pearson product-moment correlation coefficient comparing neighbour-joining (NJ) trees for each gene. Mean diversity for each locus was calculated as previously described (Blackall *et al.*, 1997). Afterwards, START (Jolley *et al.*, 2001) was used to perform Burst analysis (one strain was assigned to a CC when it shared five alleles with any other strain in the same group), cluster analysis [unweighted-pair group method with arithmetic mean (UPGMA) dendrogram using the matrix of pairwise differences] and recombination analysis [index of association (I_A) and square sum of the condensed fragment length statistic of Sawyer's test]. To examine the association of the clustering with the origin of the strains, the number of clinical isolates within a cluster was compared to the number of other isolates by chi-squared test (significance at $P < 0.001$) using SPSS 12.0 software. Finally, the partial sequences of the seven genes were concatenated using DAMBE (Xia & Xie, 2001), multiple alignments were constructed using BioEdit (Hall, 1998) and NJ trees using 10 000 bootstraps were constructed using MEGA3.1 (Kumar *et al.*, 2004). MEGA3.1 was also used to calculate the overall mean distances and the overall mean difference between non-synonymous and synonymous codon changes ($d_N - d_S$) for the seven genes independently.

Table 1. Primer sequences and relevant PCR conditions for partial amplification of the genes used in the MLST of *H. parasuis*

Gene	Primer sequences	[Primer] (μ M)	[MgCl ₂] (mM)	Reference
<i>atpD</i>	atpDF CAAGATGCAGTACCAAAAGTTTA	0.4	1.5	This work
	atpDR ACGACCTTCATCACGGAAT			
<i>infB</i>	infBF CCTGACTAYATTCGTAAAGC	0.5	2.0	Christensen <i>et al.</i> (2004)
	infBR ACGACCTTTATCGAGGTAAG			
<i>mdh</i>	mdh-up TCATTGTATGATATTGCCCC	0.4	4.5	www.mlst.net
	mdh-dn ACTTCTGTACCTGCATTTTG			
<i>rpoB</i>	rpoBF TCACAACTTTCICAATTTATG	0.4	3.0	This work
	rpoBR ACAGAAACCACTTGTTGCG			
<i>6pgd</i>	6pgdF TTATTACCGCACTTAGAAG	0.4	3.0	This work
	6pgdR CGTTGATCTTTGAATGAAGA			
<i>g3pd</i>	3gpdF GGTCAAGACATCGTTTCTAAC	0.4	1.5	This work
	3gpdR TCTAATACTTTGTTTGAGTAACC			
<i>frdB</i>	frdBf CATATCGTTGGTCTTGCCGT	0.4	1.5	This work
	frdBR TTGGCACTTTCGATCTTACCTT			

RESULTS

Variability of loci

Seven selected genes (*rpoB*, *6pgd*, *mdh*, *infB*, *frdB*, *g3pd* and *atpD*) were amplified and sequenced from the 131 *H. parasuis* isolates (GenBank accession nos DQ781411–DQ782327). The main parameters of the different loci are summarized in Table 2. The overall mean distance and d_N-d_S indicated differences in the variability of the loci and absence of strong positive or negative selection. In summary, we obtained a mean of 29 alleles per locus and an approximate potential resolution of 10^{10} STs. The 131 isolates studied were assigned to 109 STs, with a mean diversity per locus of 0.777. The most frequent ST was ST46, which was present in 3.8 % of the dataset.

Identification of clusters

After Burst analysis, STs were grouped in 13 lineages (or CCs) and 69 singletons (63 %). Only one predicted founder, ST21 for CC3, could be defined (for details see supplementary Table S1, available with the online version of this paper). CC1 contained 9 STs (14 isolates), CC2 contained 6

STs (7 isolates), CC3 contained 4 STs (5 isolates), CC4 contained 3 STs (4 isolates) and the remaining CCs contained 2 STs (between 6 and 2 isolates). The CCs showed an association with the geographical origin of the strains, since almost all the CCs included isolates from the same country. Only CC7 included isolates from different countries, Spain and Germany. When the clinical origin of the strains was examined, CCs 1, 2, 4, 5, 6, 8 and 13 included only nasal isolates, CC3 was formed by five systemic isolates, CC10 consisted of two lung isolates, and CCs 7, 9, 11 and 12 included mainly clinical isolates from different or unknown sites.

When a UPGMA dendrogram was built, six monophyletic clusters of related genotypes were defined (Fig. 1). Notably, cluster A was mainly formed by nasal isolates from asymptomatic carriers (Table 3). Statistical analysis by chi-squared test showed a significantly ($P < 0.001$) higher number of nasal isolates in this cluster. Clusters B, C, D and E did not show any significant percentage of nasal or clinical isolates (Table 3). Finally, cluster F included a high percentage of systemic isolates (Table 3) and showed a significantly ($P < 0.001$) higher number of clinical isolates

Table 2. Main parameters of the selected loci

Locus	Fragment size (bp)	No. of alleles	Mean distance	$d_N-d_S^*$
<i>rpoB</i>	470	31	0.031 ± 0.004	-0.085 ± 0.013
<i>6pgd</i>	599	41	0.054 ± 0.006	$+0.055 \pm 0.008$
<i>mdh</i>	537	24	0.025 ± 0.004	-0.075 ± 0.009
<i>infB</i>	501	36	0.008 ± 0.002	$+0.008 \pm 0.002$
<i>frdB</i>	553	33	0.014 ± 0.002	-0.043 ± 0.007
<i>g3pd</i>	564	14	0.002 ± 0.001	-0.005 ± 0.002
<i>atpD</i>	582	23	0.010 ± 0.002	$+0.013 \pm 0.004$

*Difference between non-synonymous and synonymous codon changes.

when compared using the chi-squared test. It is noteworthy that cluster F also included virulent reference strains Nagasaki, 1A-84-22113 and 5D-84-15995 (Rapp-Gabrielson *et al.*, 2006). Indeed, some nasal isolates (e.g. CA38-4 and CC6-7) included in cluster F came from farms affected by Glässer's disease. When the alleles were compared, strains in cluster F had 54 alleles that were not present in other clusters (although 36 of them were only present in a single isolate) and 22 alleles that were shared with other clusters.

Analysis of concatenated sequences

To complete the analysis, the sequences were concatenated and an unrooted NJ tree with 10 000 bootstraps was built with the 3806 bp resulting sequences (Fig. 2). Strains were divided into two main branches strongly supported by high bootstrap values (> 95 %). There was little structure inside the two main branches, although many of the bootstrap values were above 50 %. The first branch (branch 1, Fig. 2) included 65 nasal isolates out of a total of 101 (64 %) and non-virulent reference strains C5, D74, 174 and SW114. On the other hand, it also included virulent reference strain SW124. The second branch (branch 2, Fig. 2) included 83 % clinical isolates and virulent reference strains Nagasaki, 5D-84-15995 and 1A-84-22113 (Rapp-Gabrielson *et al.*, 2006). This branch contained almost all the strains included in cluster F of Fig. 1.

Detection of recombination

In addition, recombination in *H. parasuis* was evaluated due to its impact on phylogenetic reconstructions and its value as an indicator of a clonal population structure. The I_A for the whole database was 0.752 and for clusters A, B, C, D, E and F was 0.726, 0.111, 0.412, 0.641, 1.646 and 0.647, respectively. To study the recombination within the gene fragments, Sawyer's test (10 000 trials) was performed. Only the *6pgd* gene showed a significant ($P < 0.0001$) non-random distribution of synonymous polymorphic sites. Individual NJ trees for the seven fragments showed congruences by Pearson product-moment correlation between 51.4 and 0, indicating little agreement among trees. Interestingly, isolate GN-254 showed an unusually divergent *rpoB* sequence, which placed this isolate in a separated branch in the corresponding NJ tree (data not shown). This divergent sequence showed a range of identity with the sequences of the remaining isolates of 0.86–0.84, which may indicate a possible horizontal gene transfer. A BLASTN search (<http://www.ncbi.nlm.nih.gov/blast/>) with this sequence reported a best hit, with 97 % identity, to *Actinobacillus porcini* strain CCUG 38924.

Fig. 1. UPGMA dendrogram constructed with the pairwise mean differences in allelic profile of the 131 *H. parasuis* isolates. Isolates from clinical lesions are highlighted by a black background and the different clusters by letters.

Table 3. Percentages of clinical and nasal isolates in the different MLST clusters

The site of isolation or disease status is stated when known.

MLST cluster*	Clinical isolates			Nasal isolates			Total clinical	Total nasal
	Pulmonary	Systemic	ND†	Healthy farm	Diseased farm	ND†		
A (33)	3	0	3	76	9	9	6	94
B (6)	33	0	33	33	0	0	67	33
C (28)	21	18	21	32	4	4	60	40
D (15)	14	26	0	27	20	13	40	60
E (20)	20	0	15	35	30	0	35	65
F (29)	21	31	31	3	7	7	83	17

*Clusters A to F are defined in Fig. 1. The number of strains within each cluster is indicated in parentheses.

†No clinical data available.

DISCUSSION

The objective of this study was to gain an insight into the population structure of *H. parasuis* using a reliable typing method and to analyse the association between the genotype and the virulence of the strains. For that reason, a MLST scheme was developed and an effort was made to compile a collection of clinical isolates and putative non-virulent strains from the nasal cavity. The lack of information on the genome sequence of *H. parasuis* has limited the development of this MLST scheme, since only partial sequences of three conserved genes (*atpD*, *infB* and *rpoB*) were accessible at the time. As a result, the locations of the seven loci in the genome and the adjacent sequences are not known. Despite these limitations, this MLST scheme provided enough resolution power to unambiguously characterize *H. parasuis* and describe two lineages with different putative virulence.

As expected, the mean diversity per locus in the MLST scheme was higher (0.777) than the mean diversity reported for multilocus enzyme electrophoresis (0.405) (Blackall *et al.*, 1997). In agreement with previous reports (Oliveira *et al.*, 2003; Olvera *et al.*, 2006; Rafiee *et al.*, 2000; Smart *et al.*, 1988), our results confirmed that several strains (between one and five) can circulate in a farm. In addition, different strains can be isolated from the same animal [e.g. IQ7N7 (ST56) and IQ7N8 (ST84)] and even from the same systemic lesion [e.g. RU15-4P (ST51) and RU15-5P (ST75)]. Contrary to what is commonly accepted, the latter results indicate that more than one strain can be involved in a clinical outbreak. Nevertheless, some clones seem to have a wider distribution, since some STs could be detected in different farms (e.g. ST44, ST97, ST56 and ST34).

In this study, we also confirmed the high heterogeneity of *H. parasuis*. Accordingly, MLST analysis did not detect any predominant ST (the highest frequency of a ST is 3.8 %) and singletons were very frequent, even when a relaxed CC definition (five common alleles instead of six) was used in the Burst analysis. Although a certain geographical association of strains was found, our sample could be biased by a

more intense sampling in Spain and this association should be studied further.

Taking into account the I_A values, the sign of recombination in *6pgd* and the lack of congruence between individual gene trees, it seems that recombination events have significant incidence in *H. parasuis*. Indeed, this bacterium seems to have no clonal framework. Furthermore, the shared alleles among groups indicate a recent exchange of alleles, a phenomenon that has also been reported for *H. influenzae* (Meats *et al.*, 2003). Suitably, the facts that *H. parasuis* is naturally transformable (Bigas *et al.*, 2005; Lancashire *et al.*, 2005) and that more than one strain can colonize an individual create suitable conditions for frequent recombination.

In some bacteria, disease is caused by specific clones, which spread out, causing outbreaks. Those genotypes are favoured by selection and expand, thus creating an 'epidemic' population structure (Smith *et al.*, 2000). In *H. parasuis*, few CCs were identified by Burst analysis, and although a link between some of them and clinical (putative virulent) isolates was found, no dominant CC associated with systemic infection could be demonstrated. However, it is possible that, when more clinical isolates are tested, CCs with a worldwide distribution and linked to disease onset could be found.

When a UPGMA dendrogram was constructed, the 131 isolates were divided into six monophyletic clusters. Interestingly, cluster A (Fig. 1) was clearly associated with nasal isolation, and it is probably formed by non-virulent strains. Cluster B showed a tendency to include mainly pulmonary isolates, which were also present, although at a lower percentage, in cluster E. Finally, systemic isolates were primarily found in cluster F, although they were also included in clusters C and D. Our results, together with the clinical background of the strains, suggest that *H. parasuis* comprises strains with three levels or capacities of virulence: first, non-virulent strains, belonging to the biota of the upper respiratory tract; second, pulmonary strains with the pathogenic capacity to produce bronchopneumonia but not invasive disease; and third, systemic strains with the

Fig. 2. Unrooted NJ tree constructed with the concatenated sequences using 10 000 bootstraps for the 131 *H. parasuis* isolates. Bootstrap values (>50%) are indicated in the nodes. Isolates from clinical lesions are highlighted by a black background and the two main branches are indicated by numbers.

capacity to produce Glässer's disease. Unfortunately, the putative virulence of some isolates was difficult to establish, since, even when systemic infection is observed, lung tissue is frequently used for diagnosis of *H. parasuis* infection. Nevertheless, these results should be confirmed by experimental animal infections in order to determine the real virulence of the strains.

Finally, the NJ tree indicated the existence of two divergent branches within *H. parasuis*. Other algorithms were not used due to excessive computation time. However, there are several studies indicating that all methods tend to perform well when they are provided with enough data (Nei & Kumar, 2000). The division of *H. parasuis* isolates into two branches was strongly supported by high bootstrap values. One branch (branch 2, Fig. 2) showed an association with pathogenic isolates. This branch included all strains of UPGMA cluster F, with only one exception, and it could be indicative of a divergent lineage with increased virulence. The existence of this highly virulent cluster was already suggested, although less obviously, by our study of *H. parasuis* strains by partial sequencing of *hsp60* (Olvera *et al.*, 2006). On the other hand, branch 1 (Fig. 2) did not show an association with disease. The majority of strains in this group were of nasal origin, and only some of the isolates seem to be potentially virulent. Even though branch 2 appears to be very consistent, it has to be taken into account that recombination has a major impact on phylogenetic reconstructions and these results have to be interpreted carefully.

In conclusion, *H. parasuis* strains were classified by MLST and two clusters were statistically associated with nasal and clinical isolation, respectively. After NJ analysis, the isolates in the disease-associated cluster were found to be clearly divergent from the remaining *H. parasuis* isolates, forming a different lineage.

ACKNOWLEDGEMENTS

We thank Núria Galofré for technical support. This work was supported by grant AGL2004-07349 from the Ministerio de Educación y Ciencia of Spain. Fellowship support for A. O. from CReSA is also acknowledged.

REFERENCES

- Bigas, A., Garrido, M. E., de Rozas, A. M., Badiola, I., Barbe, J. & Llagostera, M. (2005). Development of a genetic manipulation system for *Haemophilus parasuis*. *Vet Microbiol* **105**, 223–228.
- Blackall, P. J., Trott, D. J., Rapp-Gabrielson, V. & Hampson, D. J. (1997). Analysis of *Haemophilus parasuis* by multilocus enzyme electrophoresis. *Vet Microbiol* **56**, 125–134.

- Christensen, H., Kuhnert, P., Olsen, J. E. & Bisgaard, M. (2004). Comparative phylogenies of the housekeeping genes *atpD*, *infB* and *rpoB* and the 16S rRNA gene within the Pasteurellaceae. *Int J Syst Evol Microbiol* **54**, 1601–1609.
- Cooper, J. E. & Feil, E. J. (2004). Multilocus sequence typing – what is resolved? *Trends Microbiol* **12**, 373–377.
- de la Puente-Redondo, V. A., del Blanco, N. G., Gutierrez-Martin, C. B., Mendez, J. N. & Rodriguez Ferri, E. F. (2000). Detection and subtyping of *Actinobacillus pleuropneumoniae* strains by PCR-RFLP analysis of the *tbpA* and *tbpB* genes. *Res Microbiol* **151**, 669–681.
- de la Puente Redondo, V. A., Navas Mendez, J., Garcia del Blanco, N., Ladron Boronat, N., Gutierrez Martin, C. B. & Rodriguez Ferri, E. F. (2003). Typing of *Haemophilus parasuis* strains by PCR-RFLP analysis of the *tbpA* gene. *Vet Microbiol* **92**, 253–262.
- del Rio, M. L., Martin, C. B., Navas, J., Gutierrez-Muniz, B., Rodriguez-Barbosa, J. I. & Rodriguez Ferri, E. F. (2006). *aroA* gene PCR-RFLP diversity patterns in *Haemophilus parasuis* and *Actinobacillus* species. *Res Vet Sci* **80**, 55–61.
- Dingle, K. E., Colles, F. M., Wareing, D. R. & 7 other authors (2001). Multilocus sequence typing system for *Campylobacter jejuni*. *J Clin Microbiol* **39**, 14–23.
- Enright, M. C. & Spratt, B. G. (1998). A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**, 3049–3060.
- Enright, M. C. & Spratt, B. G. (1999). Multilocus sequence typing. *Trends Microbiol* **7**, 482–487.
- Enright, M. C., Spratt, B. G., Kalia, A., Cross, J. H. & Bessen, D. E. (2001). Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between emm type and clone. *Infect Immun* **69**, 2416–2427.
- Feavers, I. M., Gray, S. J., Urwin, R., Russell, J. E., Bygraves, J. A., Kaczmarek, E. B. & Maiden, M. C. (1999). Multilocus sequence typing and antigen gene sequencing in the investigation of a meningococcal disease outbreak. *J Clin Microbiol* **37**, 3883–3887.
- Hall, T. (1998). BIOEDIT – Biological Sequence Alignment Editor for Windows. North Carolina, USA: Carolina State University.
- Heym, B., Le Moal, M., Armand-Lefevre, L. & Nicolas-Chanoine, M. H. (2002). Multilocus sequence typing (MLST) shows that the 'Iberian' clone of methicillin-resistant *Staphylococcus aureus* has spread to France and acquired reduced susceptibility to teicoplanin. *J Antimicrob Chemother* **50**, 323–329.
- Homan, W. L., Tribe, D., Poznanski, S., Li, M., Hogg, G., Spalburg, E., Van Embden, J. D. & Willems, R. J. (2002). Multilocus sequence typing scheme for *Enterococcus faecium*. *J Clin Microbiol* **40**, 1963–1971.
- Jolley, K. A., Feil, E. J., Chan, M. S. & Maiden, M. C. (2001). Sequence type analysis and recombinational tests (START). *Bioinformatics* **17**, 1230–1231.
- Kielstein, P. & Rapp-Gabrielson, V. J. (1992). Designation of 15 serovars of *Haemophilus parasuis* on the basis of immunodiffusion using heat-stable antigen extracts. *J Clin Microbiol* **30**, 862–865.
- King, S. J., Leigh, J. A., Heath, P. J., Luque, I., Tarradas, C., Dowson, C. G. & Whatmore, A. M. (2002). Development of a multilocus sequence typing scheme for the pig pathogen *Streptococcus suis*: identification of virulent clones and potential capsular serotype exchange. *J Clin Microbiol* **40**, 3671–3680.
- Kriz, P., Kalmusova, J. & Felsberg, J. (2002). Multilocus sequence typing of *Neisseria meningitidis* directly from cerebrospinal fluid. *Epidemiol Infect* **128**, 157–160.
- Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.
- Lancashire, J. F., Terry, T. D., Blackall, P. J. & Jennings, M. P. (2005). Plasmid-encoded Tet B tetracycline resistance in *Haemophilus parasuis*. *Antimicrob Agents Chemother* **49**, 1927–1931.
- Lemee, L., Dhalluin, A., Pestel-Caron, M., Lemeland, J. F. & Pons, J. L. (2004). Multilocus sequence typing analysis of human and animal *Clostridium difficile* isolates of various toxigenic types. *J Clin Microbiol* **42**, 2609–2617.
- Maiden, M. C., Bygraves, J. A., Feil, E. & 10 other authors (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* **95**, 3140–3145.
- Meats, E., Feil, E. J., Stringer, S., Cody, A. J., Goldstein, R., Kroll, J. S., Popovic, T. & Spratt, B. G. (2003). Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. *J Clin Microbiol* **41**, 1623–1636.
- Nallapareddy, S. R., Duh, R. W., Singh, K. V. & Murray, B. E. (2002). Molecular typing of selected *Enterococcus faecalis* isolates: pilot study using multilocus sequence typing and pulsed-field gel electrophoresis. *J Clin Microbiol* **40**, 868–876.
- Nei, M. & Kumar, S. (editors) (2000). *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.
- Nielsen, R. (1993). Pathogenicity and immunity studies of *Haemophilus parasuis* serotypes. *Acta Vet Scand* **34**, 193–198.
- Noller, A. C., McEllistrem, M. C., Stine, O. C., Morris, J. G., Jr, Boxrud, D. J., Dixon, B. & Harrison, L. H. (2003). Multilocus sequence typing reveals a lack of diversity among *Escherichia coli* O157:H7 isolates that are distinct by pulsed-field gel electrophoresis. *J Clin Microbiol* **41**, 675–679.
- Oliveira, S. & Pijoan, C. (2004). *Haemophilus parasuis*: new trends in diagnosis, epidemiology and control. *Vet Microbiol* **99**, 1–12.
- Oliveira, S., Blackall, P. J. & Pijoan, C. (2003). Characterization of the diversity of *Haemophilus parasuis* field isolates by use of serotyping and genotyping. *Am J Vet Res* **64**, 435–442.
- Olvera, A., Calsamiglia, M. & Aragon, V. (2006). Genotypic diversity of *Haemophilus parasuis* field strains. *Appl Environ Microbiol* **72**, 3984–3992.
- Perez-Losada, M., Browne, E. B., Madsen, A., Wirth, T., Viscidi, R. P. & Crandall, K. A. (2006). Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. *Infect Genet Evol* **6**, 97–112.
- Rafiee, M., Bara, M., Stephens, C. P. & Blackall, P. J. (2000). Application of ERIC-PCR for the comparison of isolates of *Haemophilus parasuis*. *Aust Vet J* **78**, 846–849.
- Rapp-Gabrielson, V. J., Gabrielson, D. A. & Schamber, G. J. (1992). Comparative virulence of *Haemophilus parasuis* serovars 1 to 7 in guinea pigs. *Am J Vet Res* **53**, 987–994.
- Rapp-Gabrielson, V., Oliveira, S. & Pijoan, C. (2006). *Haemophilus parasuis*. In *Diseases of Swine*, pp. 681–690. Edited by B. E. Straw, J. J. Zimmerman, S. D'Allaire & D. J. Taylor. Ames, IA: Blackwell Publishing.
- Shi, Z. Y., Enright, M. C., Wilkinson, P., Griffiths, D. & Spratt, B. G. (1998). Identification of three major clones of multiply antibiotic-resistant *Streptococcus pneumoniae* in Taiwanese hospitals by multilocus sequence typing. *J Clin Microbiol* **36**, 3514–3519.
- Smart, N. L., Miniats, O. P. & MacInnes, J. I. (1988). Analysis of *Haemophilus parasuis* isolates from southern Ontario swine by restriction endonuclease fingerprinting. *Can J Vet Res* **52**, 319–324.

Smith, J. M., Feil, E. J. & Smith, N. H. (2000). Population structure and evolutionary dynamics of pathogenic bacteria. *Bioessays* **22**, 1115–1122.

Spratt, B. G. (1999). Multilocus sequence typing: molecular typing of bacterial pathogens in an era of rapid DNA sequencing and the internet. *Curr Opin Microbiol* **2**, 312–316.

Vahle, J. L., Haynes, J. S. & Andrews, J. J. (1995). Experimental reproduction of *Haemophilus parasuis* infection in swine: clinical, bacteriological, and morphologic findings. *J Vet Diagn Invest* **7**, 476–480.

van Loo, I. H., Heuvelman, K. J., King, A. J. & Mooi, F. R. (2002). Multilocus sequence typing of *Bordetella pertussis* based on surface protein genes. *J Clin Microbiol* **40**, 1994–2001.

Wang, X. M., Noble, L., Kreiswirth, B. N., Eisner, W., McClements, W., Jansen, K. U. & Anderson, A. S. (2003). Evaluation of a multilocus sequence typing system for *Staphylococcus epidermidis*. *J Med Microbiol* **52**, 989–998.

Xia, X. & Xie, Z. (2001). DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* **92**, 371–373.