

## Research Article

# Comparative Genomics of *Actinobacillus pleuropneumoniae* Serotype 8 Reveals the Importance of Prophages in the Genetic Variability of the Species

Isabelle Gonçalves de Oliveira Prado <sup>1</sup>, Giarlã Cunha da Silva,<sup>1</sup> Josicelli Souza Crispim,<sup>1</sup> Pedro Marcus Pereira Vidigal <sup>2</sup>, Moysés Nascimento,<sup>3</sup> Mateus Ferreira Santana,<sup>1</sup> and Denise Mara Soares Bazzolli <sup>1</sup>

<sup>1</sup>Laboratório de Genética Molecular de Bactérias/Bioagro-Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa 36570-900, Brazil

<sup>2</sup>Núcleo de Análise de Biomoléculas (NuBioMol), Universidade Federal de Viçosa, Viçosa, Brazil

<sup>3</sup>Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, Brazil

Correspondence should be addressed to Denise Mara Soares Bazzolli; dbazzolli@ufv.br

Received 9 October 2019; Accepted 3 January 2020; Published 19 February 2020

Guest Editor: Tarja Sironen

Copyright © 2020 Isabelle Gonçalves de Oliveira Prado et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Actinobacillus pleuropneumoniae* is the etiologic agent of porcine pleuropneumonia. Currently, there are 18 different serotypes; the serotype 8 is the most widely distributed in the United States, Canada, United Kingdom, and southeastern Brazil. In this study, genomes of seven *A. pleuropneumoniae* serotype 8 clinical isolates were compared to the other genomes of twelve serotypes. The analyses of serotype 8 genomes resulted in a set of 2352 protein-coding sequences. Of these sequences, 76.6% are present in all serotypes, 18.5% are shared with some serotypes, and 4.9% were differential. This differential portion was characterized as a series of hypothetical and regulatory protein sequences: mobile element sequence. Synteny analysis demonstrated possible events of gene recombination and acquisition by horizontal gene transfer (HGT) in this species. A total of 30 sequences related to prophages were identified in the genomes. These sequences represented 0.3 to 3.5% of the genome of the strains analyzed, and 16 of them contained complete prophages. Similarity analysis between complete prophage sequences evidenced a possible HGT with species belonging to the family Pasteurellaceae. Thus, mobile genetic elements, such as prophages, are important components of the differential portion of the *A. pleuropneumoniae* genome and demonstrate a central role in the evolution of the species. This study represents the first study done to understand the genome of *A. pleuropneumoniae* serotype 8.

## 1. Introduction

Pork is an important source of animal protein and is currently one of the most commonly consumed meat products in the world [1]. However, the use of an intensive production system has frequently given rise to the occurrence of respiratory diseases, having a major impact on production, causing significant economic losses in pig farming [2, 3]. Porcine pleuropneumonia is one of the most important respiratory diseases in pigs and is caused by the bacterium *Actinobacillus pleuropneumoniae*; this species can be divided into two bio-

types according to their dependence on nicotinamide adenine dinucleotide (NAD) [4]. Currently, this species is classified into 18 serotypes based on the antigenic properties of capsule polysaccharides [5–7].

The pathogenesis of porcine pleuropneumonia is complex and involves different virulence factors produced by the bacterium [8–10]. Virulence is multifactorial and is related to a combination of factors such as toxins from the RTX family, composition and structure of capsule polysaccharides, outer membrane lipopolysaccharide (LPS), iron siderophores, biofilm formation, and adhesins [8]. In addition to the

TABLE 1: *A. pleuropneumoniae* genomes used in this study.

Strain/serotype	Genome size (pb)	CDS	% GC	Accession code (WGS)	Reference
4074/1	2318649	2135	41.2	CP029003.1	Xu et al., [16]
4226/2	2314083	2228	41.2	ADXN00000000.1	Zhan et al., 2010
JL03, 3	2242062	2115	41.2	CP000687.1	Xu et al., 2008
M62/4	2260565	2186	41.2	ADOF00000000.1	Xu et al., [16]
L20/5b	2274482	2168	41.3	CP000569.1	Foote et al., 2008
Femo/6	2302700	2300	41.0	ADOG00000000.1	Xu et al., [16]
AP76/7	2345435	2234	41.2	CP001091.1	Linke et al., 2008
MV460/8	2213381	2116	41.1	JSVG00000000.1	
MV518/8	2275540	2189	41.1	JSVZ00000000.1	
MV597/8	2219395	2115	41.1	JSVX00000000.1	Pereira et al., [23]
MV780/8	2274000	2179	41.1	JSVV00000000.1	
MV1022/8	2262828	2179	41.1	JSVF00000000.1	
MV5651/8	2264279	2179	41.1	JSVY00000000.1	
MIDG2331/8	2337633	2235	41.1	LN908249.1	Bossé et al., [22]
CVJ13261/9	2256417	2163	41.2	ADOI00000000.1	
D13039/10	2266276	2155	41.2	ADOD00000000.1	
56153/11	2257884	2154	41.2	ADOK00000000.1	Xu et al., [16]
1096/12	2185499	2082	41.2	ADOL00000000.1	
N273/13	2236660	2148	41.2	ADOM00000000.1	

abovementioned virulence factors, some *A. pleuropneumoniae* serotypes present natural competence, and therefore, the occurrence of natural transformation is common in this species and dissemination of resistance genes in Pasteurellaceae family members, as *A. pleuropneumoniae*, is common [11–13]. Finally, virulence is complex, and antimicrobial resistance genes that can be encoded by both the chromosome and plasmids are essential depending on the specific niche, as in the natural hosts and in specific conditions [14, 15].

Although comparative genomic studies with different genotypes of *A. pleuropneumoniae* serotypes were carried out [16], no information on *A. pleuropneumoniae* serotype 8 was provided so far. Over the years, *A. pleuropneumoniae* serotype 8 has been neglected in identification studies due to failures in serotyping techniques, and as a result, genomic studies involving this serotype are nonexistent. Although recent studies have showed a wide distribution of this serotype in several regions, such as the United Kingdom [17, 18], North America [19], and Brazil [20], just recently *A. pleuropneumoniae* serotype 8 genome sequence was available [21, 22].

A study carried out by our research group using the alternative host, *Galleria mellonella* larvae, detected different virulence patterns in clinical isolates of serotype 8 *A. pleuropneumoniae* [23]. Based on the results obtained from that study, six isolates with different phenotypic profiles were selected for genomic sequencing. Clinical isolates from *A. pleuropneumoniae* serotype 8 have virulence complexity [9], but no specific information on genotypic variability is available so far. Then, this study is the first to describe *A. pleuropneumoniae* serotype 8 genomic from comparative analysis between *A. pleuropneumoniae* serotype 8 genomes (clinical isolates: Brazilian origin [21], one of English origin [22])

and twelve genomes of different serotypes of *A. pleuropneumoniae* deposited in databases.

## 2. Materials and Methods

**2.1. Actinobacillus pleuropneumoniae Genome Sequences.** Nineteen genomes from different serotypes of *A. pleuropneumoniae* available at the GenBank database (<https://www.ncbi.nlm.nih.gov/genbank/>) were used in the present study. The genomes of *A. pleuropneumoniae* serotypes 1 (4074), 13 (JL03), 5 (L20), 7 (AP76), and 8 (MIDG2331) are closed. The genomes of the other serotypes are in contigs (Table 1).

**2.2. Determination of Protein-Coding Sequence Set in A. pleuropneumoniae Serotype 8.** The set of clusters of the coding DNA sequences (CDS) predicted for *A. pleuropneumoniae* serotype 8 was based on the seven genomes of clinical isolates taken from pig farms, six from Brazil [21] and one from the United Kingdom [22]. In this analysis, the CD-HIT v.4.6.1 program [24, 25] was used to consider an identity threshold of 0.85 to cluster the CDS. For the functional annotation of the *A. pleuropneumoniae* serotype 8 reference genome, five databases were used: COG [26], CDD [27], PFAM [28], SMART [29], and UNIPROT [30]. The similarity searches were carried out using the BLAST algorithm [31] considering an  $E$  value  $\leq 10^{-5}$ .

**2.3. Comparative Analysis of Predicted CDS of Serotype 8 and the Other 12 Different Serotypes of A. pleuropneumoniae.** In this analysis, 12 genomes of *A. pleuropneumoniae* different serotypes (1, 2, 3, 4, 5b, 6, 7, 9, 10, 11, 12, and 13) were used (Table 1). These sequences are deposited in the UNIPROT

database [30]. The comparative analysis was carried out using the BLAST algorithm [31], contrasting the genomes of the serotypes analyzed against the *A. pleuropneumoniae* serotype 8 reference assembled in this study.

**2.4. Analysis of *A. pleuropneumoniae* Orthologous Gene Groups.** From the predicted CDS of the 12 different serotypes and 7 serotype 8 *A. pleuropneumoniae* genomes, a database containing 28002 CDS corresponding to all serotypes of the species was assembled. Using the CD-HIT v.4.6.1 program [24, 25], with an identity threshold of 0.70 identity to cluster the sequences, an analysis was carried out to characterize the total set of CDS of the species. The CD-HIT was used for clustering the sequence and for reducing redundancy among them, to improve the results.

The groups of CDS identified by the CD-HIT were classified as core, shared, or differential. Additionally, the predicted protein sequences of *A. pleuropneumoniae* serotype 8 were individually compared to the predicted protein sequences of the other serotypes using the BLAST algorithm [31].

**2.5. Genome-Wide Analysis of Preferential Codon Usage and GC%.** The analysis of the preferential use of codons and GC content was carried out using the EMBOSS program [32] for the different serotype genomes of *A. pleuropneumoniae*. The use of each synonymous codon was determined by calculating the RSCU (Relative Synonymous Codon Usage). The RSCU value calculated for each codon was the parameter used to evaluate the codon selection type, with values = 1 characteristic of codons used with equal frequency; values > 1 were positive selection and <1 negative selection.

**2.6. Synteny Analysis.** The analyses were derived from the closed genomes of *A. pleuropneumoniae* serotypes 03 (JL03), 5b (L20), 7 (AP76), and 8 (MIDG2331) and from the six genomes of *A. pleuropneumoniae* serotype 8 isolates of Brazilian origin (MV460, MV518, MV597, MV780, MV1022, and MV5651). Multiple alignments of the sequences of the *A. pleuropneumoniae* genomes were derived from the Progressive Mauve v.2.3.1 software program [33].

**2.7. Analysis of Sequences Similar to Prophages.** Sequences similar to the prophages present in all the *A. pleuropneumoniae* genomes used in this study were obtained through the PHASTER program [34]. The prophage sequences were aligned by MAFFT [35], and the alignment was edited using the GBLOCKS program [36]. A dendrogram using the Neighbor-Joining genetic distance grouping method was generated by the MEGA 6 program [37] with a bootstrap containing 2000 replicates. Prophage complete sequences were compared using BLAST [31] against the GenBank databases to identify possible horizontal gene transfers between bacteria. For this, a coverage and identity above 70% and an *E* value less than  $10^{-5}$  were used as cutoff points. After editing, the alignment was obtained using the GBLOCKS program. Using the same complete sequences of prophages, an alignment was done with the Clustal Omega [38]. From the values of the identity matrix provided, a heat map was constructed with software R under version 3.5.1.

### 3. Results

**3.1. Genomic Analysis of *A. pleuropneumoniae* Serotype 8.** The total set of predicted CDS of *A. pleuropneumoniae* serotype 8 generated from the seven clinical isolates corresponded to 2352 sequences (Table 2). Of these, 1801 (76.6%) were considered core, 436 (18.5%) were shared with other serotypes, though not all, and 115 (4.9%) were predicted to be differential to *A. pleuropneumoniae* serotype 8 genomes. Among the 2352 CDS of *A. pleuropneumoniae* serotype 8, 1925 (81.8%) were categorized into the COG database. Among these were 1542 (80.1%) encode proteins with known functional categories (excluding “Unknown function” and “Prediction of general functions”).

From the distinction of the core, shared and differential regions of *A. pleuropneumoniae* serotype 8 CDS and clusters of ortholog groups were analyzed. Of the 1801 sequences comprising the core portion, 1685 were affiliated to the categories of the COG database. Among these, 1358 sequences (80.6%) represent known functional categories (Table 2). The majority of sequences characterized as core are related to amino acid metabolism and transport; ribosomal translation, structure, and biogenesis processes; biogenesis of the wall, membrane, and cell envelope; and production and conservation of energy, among other activities considered essential to the survival of the pathogen (Table 2).

As regards the shared portion, of the 436 sequences, 220 were affiliated with the COG database categories, of which 166 (75.5%) were known functional categories. Most sequences are related to the metabolism and transport of inorganic ions; biogenesis of the cellular envelope; replication, recombination, and DNA repair; and metabolism in general (Table 2; Supplementary Data Table 1).

The differential portion of the *A. pleuropneumoniae* serotype 8 genomes showed 115 CDS. Only 20 sequences were affiliated to the COG database categories, of which 18 (90.0%) were known functional categories (Table 2). These differential C are related to the regulatory processes and HGT mechanisms such as plasmids and prophages (Table 2; Supplementary Data Table 1). In this portion, CDS related to resistance to antibiotics such as tetracycline and florfenicol genes, transcriptional regulators such as LysR, DNA repair protein, transposon gamma-delta resolvase, transport proteins such as sodium and glutamate symmetric acetyltransferase, and prophage-related protein-coding sequences were reported (Table 2).

**3.2. The Pangenome of *A. pleuropneumoniae*.** From the 2984 clusters obtained from the total set of CDS of the species, the general characterization of the CDS with the distinction of the core, shared, and differential portions was carried out. Of the total, 1737 clusters were characterized as core region, present in the thirteen serotypes analyzed; 756 were clusters of CDS of shared proteins, and 491 clusters corresponded to CDS of differential proteins of each serotype. As regards the total genome of each serotype, the core portion averaged 82.5% (Table 3) showing conservation among the different serotypes.

TABLE 2: Coding sequences of *A. pleuropneumoniae* serotype 8.

COG	Description of COG classes	Core	Shared	Differential	Total
A	RNA modification and processing	1	0	0	1
C	Conversion and production of energy	117	9	0	126
D	Cycle control and cell division, chromosome partitioning	25	3	0	28
E	Amino acid metabolism and transport	158	10	1	169
F	Nucleotide metabolism and transport	60	2	0	62
G	Carbohydrate metabolism and transport	117	11	2	130
H	Coenzyme metabolism and transport	91	10	1	102
I	Lipid metabolism and transport	39	3	0	42
J	Translation, ribosomal structure, and biogenesis	153	7	2	162
K	Transcript	77	10	7	94
L	Replication, recombination, and repair	95	20	4	119
M	Biogenesis of cell wall, membrane, and envelope	127	21	0	148
N	Cellular motility	6	1	1	8
O	Posttranslational modification, protein turnover, and chaperones	96	6	0	102
P	Metabolism and transport of inorganic ions	104	24	0	128
Q	Biosynthesis of secondary metabolites, transport, and catabolism	8	5	0	13
T	Signal transduction mechanisms	31	1	0	32
U	Intracellular traffic, secretion, and vesicular transport	33	5	0	38
V	Defense mechanisms	20	18	0	38
R	Prediction of general functions	158	23	2	183
S	Unknown function	169	31	0	200
NC	Proteins not categorized on COG	116	216	95	427
	Total of affiliated proteins	1685	220	20	1925
	Total of serotype 8 proteins	1801	436	115	2352

TABLE 3: Characterization of protein groups in different serotypes of *A. pleuropneumoniae*.

Serotype	Total proteins	Core Proteins (%)	Shared Proteins (%)	Differential Proteins (%)
Serotype 1	2176	1765 (81.1)	404 (18.6)	7 (0.3)
Serotype 2	2064	1774 (86.0)	275 (13.3)	15 (0.7)
Serotype 3	2026	1756 (86.7)	260 (12.8)	10 (0.5)
Serotype 4	2219	1790 (80.7)	325 (14.7)	104 (4.7)
Serotype 5b	2004	1765 (88.1)	208 (10.4)	31 (1.6)
Serotype 6	2211	1768 (80.0)	384 (17.4)	59 (2.7)
Serotype 7	2113	1774 (84.0)	327 (15.5)	12 (0.6)
Serotype 8*	2352	1801 (76.6)	436 (18.5)	115 (4.9)
Serotype 9	2197	1779 (81.0)	416 (18.9)	2 (0.1)
Serotype 10	2170	1774 (82.0)	321 (14.8)	75 (3.5)
Serotype 11	2184	1767 (81.0)	414 (19.0)	3 (0.1)
Serotype 12	2081	1771 (85.1)	294 (14.1)	16 (0.8)
Serotype 13	2145	1760 (82.1)	381 (17.8)	4 (0.2)

\*Reference genome represents all the sequences encoding the seven clinical isolate genomes previously sequenced.

**3.3. Similarity Analysis between the Amino Acid Sequences Predicted for *A. pleuropneumoniae* Serotype 8 and Other Serotypes.** An alignment between predicted amino acid

sequences of *A. pleuropneumoniae* serotype 8 was created against all other *A. pleuropneumoniae* amino acid sequences used in this study generating clusters based on the pattern of similarity (Figure 1). Of the total 2352 amino acid sequences, 2196 (93.4%) had similarity patterns higher than 95%, thus revealing high serotype 8 sequence conservation in relation to the others. Based on the analysis of the BLAST results, three main groups of similarity related to high, medium, and low virulence standards were obtained. There was a greater sharing of the predicted CDS of serotype 8 with the serotype 6 sequences, followed by serotype 3 (Figure 1).

**3.4. Codon Preferential Usage.** As regards the codon analysis, a high standard of conservation was observed in the use of codons among all *A. pleuropneumoniae* serotypes investigated, which includes the clinical isolates of serotype 8 analyzed in this study. In Figure 2, we have represented the use of codons by *A. pleuropneumoniae*. Codons with higher RSCU values result in higher positive selection for their respective amino acids (Figure 2). We observed no significant differences in the proportions of the use of amino acids between the different isolates nor between the serotypes. The most commonly used amino acids were leucine (L: 10.6%), alanine (A: 8.7%), isoleucine (I: 6.8%), and valine (V: 6.8%), while cysteine (C: 1.0%) and tryptophan (W: 1.2%) were the most rarely used (Figure 2).

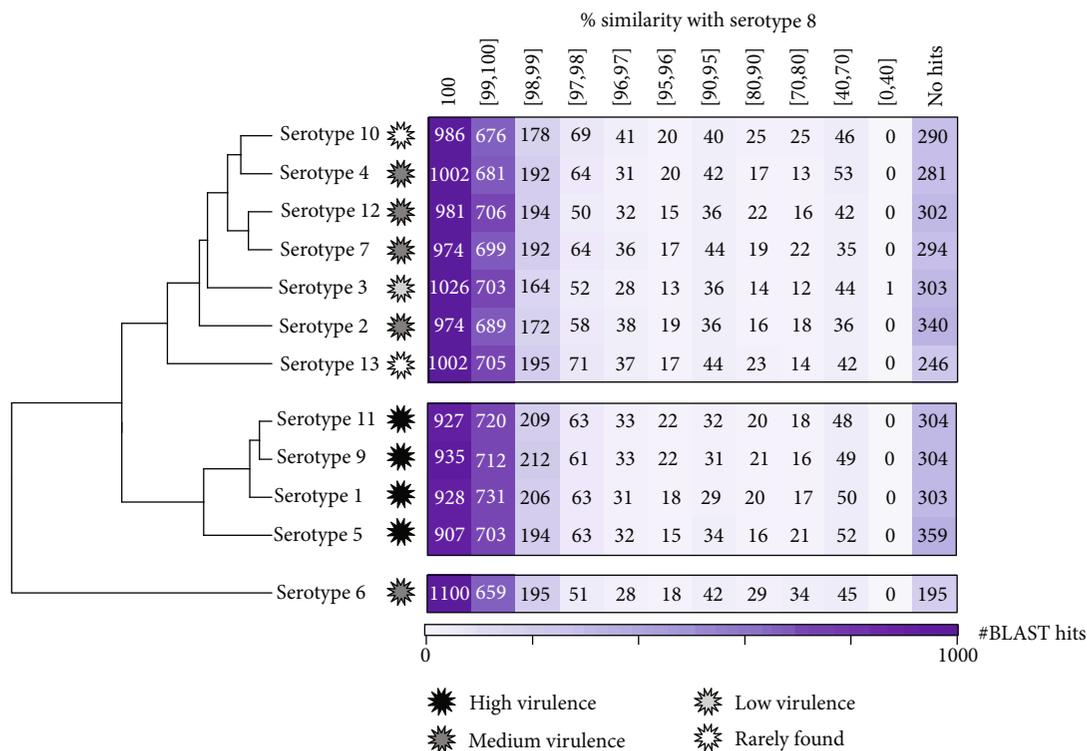


FIGURE 1: Analysis of similarity between predicted amino acid sequences of *A. pleuropneumoniae* serotype 8 and the other serotypes. The protein-coding sequences were clustered according to similarity in percentages to serotype 8. The serotypes were also characterized in relation to virulence.

3.5. *Synten Analysis of A. pleuropneumoniae*. In the alignment between the genotype-representative contigs of the six Brazilian clinical isolates of *A. pleuropneumoniae* serotype 8 and closed genomes of serotypes 3 (JL03), 5b (L20), 7 (AP76), and 8 (MIDG2331), we verified conservation in the genome structure (Figure 3). The genomes of *A. pleuropneumoniae* share practically the same blocks, denominated LCBs: "Selecting Locally Collinear Blocks." Although genome alignment conservation was found, it was possible to observe regions of acquisition/loss of genetic material and rearrangements (Figure 3). The presence of a differential and conserved block in the genomes of serotypes 5 (L20) and 7 (AP76) as well as in four of the clinical isolates of serotype 8 (MV518, MV780, MV1022, and MV5651) was observed (Figure 3). In this block composed of approximately 42.206 pb and with GC content of 40.4%, there are sequences encoding proteins related to integrase (WP\_005620278.1), pyrophosphatase (WP\_005620280.1), RdgC recombination protein (WP\_011848390.1), DNA methyltransferase (WP\_011848391.1), antirepressor (WP\_011848395.1, WP\_011848414.1), DNA methylase (WP\_011848397.1), endodeoxyribonuclease RuvA (WP\_011848399.1), terminase (WP\_043880767.1), peptidase (WP\_043877971.1), and various phage proteins (WP\_011848407.1, WP\_011848405.1, WP\_011848408.1, and WP\_011848412.1). In a second differential genomic segment, present only in serotype 7, a region of inversion and rearrangement of a block of approximately 59.649 bp and GC content of 40.9% was present (Figure 3). In this segment, we found sequences corresponding to genes encoding carboxylase enzymes (WP\_005617934.1, WP\_

005602033.1), oxidoreductases (WP\_005598646.1, WP\_005602054.1, and WP\_012478542.1), reductase (WP\_005598644.1), virulence factors involved in iron uptake (WP\_005602070.1), integrase (WP\_005617888.1), and transposases (WP\_005599960.1).

Minor variations were also observed between the different genomes (Figure 3). Among these differences, we found a small region of approximately 13000 pb present only in the seven genomes of *A. pleuropneumoniae* serotype 8 and the CDS found were for tRNA-glutamate ligase (WP\_005608501.1), tRNA-Ala (WP\_005612726.1), preproteins translocase (WP\_005612726.1), transcriptional regulator of the Rha family (WP\_039768145.1), antirepressor (WP\_058230489.1), propanediol utilization protein (WP\_039768152.1), host death prevention protein family (Phd) (WP\_005598318.1), YoeB toxin (WP\_005605064.1), tetracycl disaccharide kinase (WP\_005598320.1, WP\_005608502.1), and nine sequences encoding hypothetical proteins (WP\_039709488.1, WP\_039768147.1, WP\_039709486.1, WP\_039709484.1, WP\_039709483.1, WP\_052250595.1, WP\_014991324.1, WP\_039768150.1, and WP\_014991326.1). Another region of approximately 13180 pb was found in the genomes of *A. pleuropneumoniae* serotype 5 (L20), serotype 7 (AP76), and in three serotype 8 clinical isolates (MIDG2331, MV1022, and MV518). Sequences corresponding to genes present in this region were encoded for *flpC* operon (WP\_039709641.1), *flpB* operon (WP\_039709034.1, WP\_011848427.1), fimbriae protein (WP\_058230512.1, WP\_039709035.1, WP\_005611759.1, and WP\_011848428.1), ATP-dependent ATP-Rhi-RNA (WP\_039709036.1), ATP-

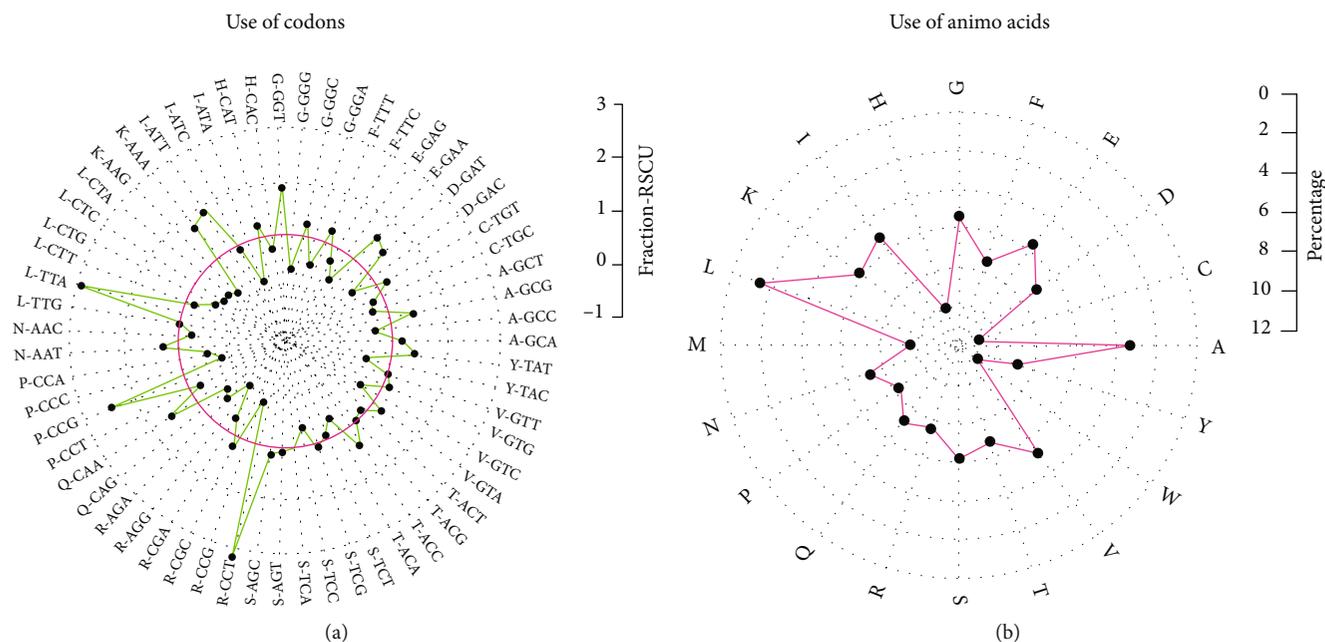


FIGURE 2: Use of codons and their respective amino acids of *A. pleuropneumoniae*. (a) The trend in the use of codons is represented in the circular map. Methionine, tryptophan, and stop codons were omitted. Synonymous codons for an amino acid used with equal frequency have RSCU = 1, indicated by the red circular line. (b) Percentage of the use of amino acids represented in the circular map. A: alanine; C: cysteine; D: aspartic acid; E: glutamic acid; F: phenylalanine; G: glycine; H: histidine; I: isoleucine; K: lysine; L: leucine; M: methionine; N: asparagine; P: proline; Q: glutamine; R: arginine; S: serine; T: threonine; V: valine; W: tryptophan; Y: tyrosine.

binding protein (WP\_005611764.1), iron-ABC transport substrate binding (WP\_005611765.1), and three hypothetical proteins (WP\_005611761.1, WP\_009875478.1, and WP\_005611761.1).

**3.6. Analysis of Prophage Sequences.** 30 sequences similar to the prophages were found in the 19 strains of *A. pleuropneumoniae* analyzed. From the total of sequences similar to prophages, 16 were classified by the PHASTER program as complete, 11 as incomplete, and 3 as questionable (Supplementary Data Table 2). Incomplete and questionable sequences were considered genomic regions containing sequences derived from phages. The regions containing prophage-related genes represent 0.3 to 3.5% of the genomes analyzed (Supplementary Data Table 2). The largest prophage identified had 48.1 kb and the smallest 22.4 kb, identified in the MV1022 and M62 strains, respectively (Supplementary Data Table 2). The GC content of the identified prophages varied between 39.3 and 44.6% (Supplementary Data Table 2).

The Neighbor-Joining method to cluster analysis between sequences similar to complete prophages allowed us to identify four different clusters (P1-P4) (Figure 4). Analysis of the complete prophages using the BLAST algorithm against sequences from the GenBank database showed that the sequences contained in the P3 cluster (prophage 2 (4074 strain), prophage 1 (4226 strain), prophage 1 (CVJ13261), and prophage 2 (56153 strain)), the M62 prophage 2, and the AP76 prophage 2 share over 70% identity and were found in genomes of *Haemophilus ducreyi* (AE017143.1), *Mannheimia haemolytica* (KP137440.1), and

*Actinobacillus suis* (CP009159.1), respectively (Table 4). For the other 13 complete prophages, no significant identity or sequence coverage was found in GenBank.

The heat map distribution showed a high identity among the prophage 1 from 518, 780, and 5651 (serotype 8) and femo (serotype 6) (Figure 5). This could also be observed among prophage 2 from 56153 (serotype 11) and 4074 (serotype 1) and prophage 1 from CVJ13261 (serotype 9) and 4226 (serotype 2). A considerable identity also was observed among prophage 1 from 518, 780, and 5651 (serotype 8), femo (serotype 6), and N273 (serotype 13). Similarly, it also was observed between N273 (serotype 13) and AP76 (serotype 7) and L20 (serotype 5) and AP76 (serotype 7).

## 4. Discussion

Analyses of GC content, codon usage, and amino acid use among the different *A. pleuropneumoniae* serotypes showed that they share a set of conserved CDS. The core portion of the genome that is well conserved among the serotypes also reinforces these results. Among the most commonly used amino acids are branched chain amino acids, such as leucine, isoleucine, and valine. These branched chain amino acids are required for the survival and virulence of *A. pleuropneumoniae* in swine, capable of synthesizing these amino acids critical for respiratory tract pathogens [39].

In the analyses of clusters from the set of CDS shared between serotypes, the pattern of clustering by similarity was compatible with the classification of serotypes into three virulence categories [16]: low, medium, and high virulence. We observed that serotype 8 shares a high number of

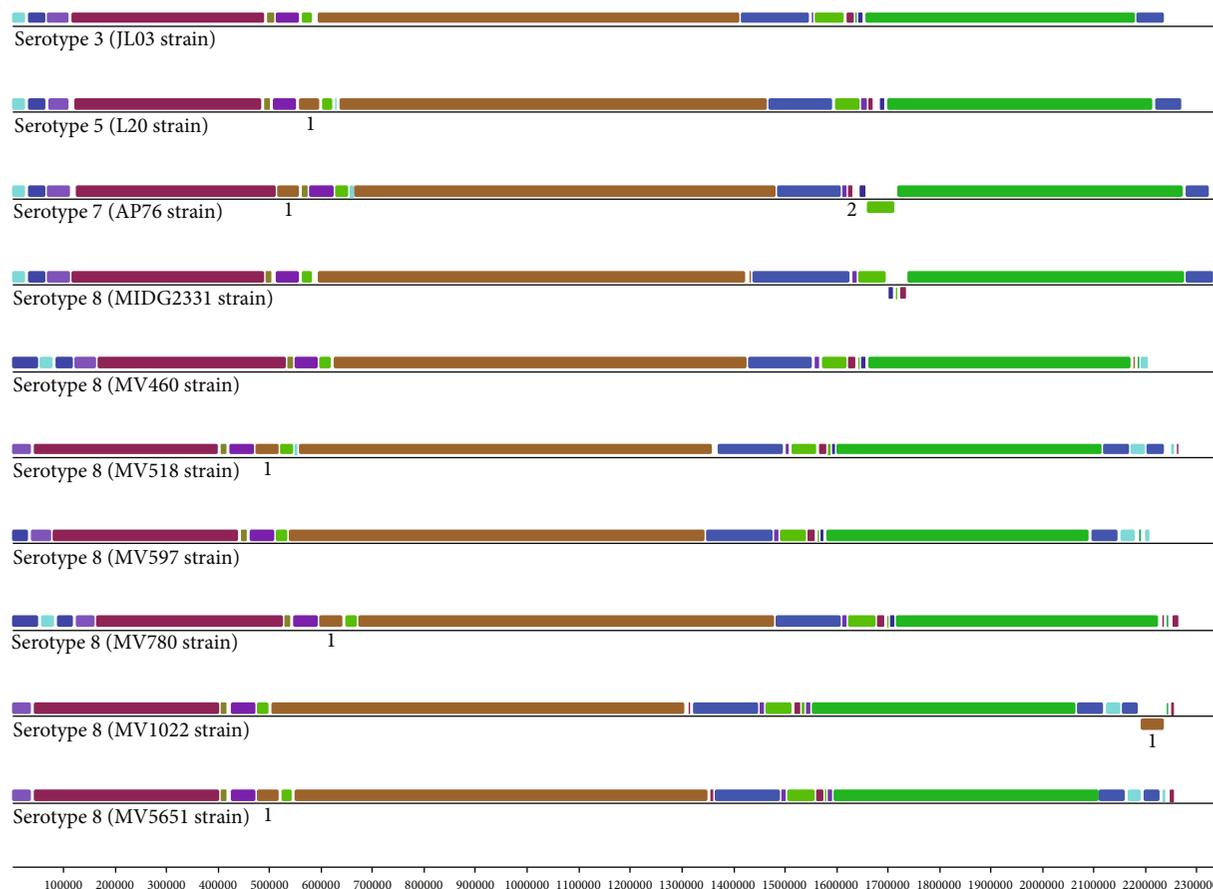


FIGURE 3: Synteny analysis of *A. pleuropneumoniae* genomes. The genomes represented correspond to serotype 3 (JL03 strain), serotype 5 (L20 strain), serotype 7 (AP76), and serotype 8 (MIDG2331, MV460, MV518, MV597, MV780, MV1022, and MV5651 strains). Horizontal bars represent the size of the genome (kb). The region identified in 1 represents the acquisition and loss of genomic information, and region 2 represents a recombination event.

protein-coding sequences with the serotypes characterized as having medium virulence, such as serotypes 2, 4, 6, 7, and 12. The characteristic of the serotypes considered as medium virulence category is associated with the persistence of the pathogen in the environment [8]. Additionally, a large sharing of CDS for serotype 8 proteins was observed in serotype 6, followed by serotype 3. As already reported in serotyping analyses, certain groups may cross-react and be mischaracterized. Serotypes 3, 6, and 8 of *A. pleuropneumoniae* in serotyping studies in North America constitute a single group, and discrimination of these three serotypes within this group is extremely difficult when using the antiserum technique [19].

In the COG analyses of the predicted amino acid sequences of *A. pleuropneumoniae* serotype 8, the core region is characterized by housekeeping genes. However, genes belonging to the core region may have differences in the level of DNA sequences. A number of genes classified as core like those encoding anaerobic glycerol-3-phosphate dehydrogenase subunit A (*glpA*), oxygen-independent coproporphyrinogen-III oxidase (*hemN*), heptosyltransferase family (*mutM*), tellurite resistance protein (*tehA*), sulfate transport system permease (*cysW*), thiazole biosynthesis protein (*thiH*), haloacid dehalogenase-like hydrolases (*had*) superfamily

(*cof*), nucleoside diphosphate sugar epimerase, and oligopeptide transporter testify to positive selection in *A. pleuropneumoniae* [40]. In general, these genes are involved in the transporting of nutrients and cellular metabolism that show that *A. pleuropneumoniae* has responded to different environmental pressures.

In the core portion, we also found genes that, according to [41], have increased expression during the acute phase of natural infection of *A. pleuropneumoniae* in pigs. These genes were related, for example, to the assembly of curli fibers, important in the formation of biofilms [42]; to the maltose operon that may increase the competition capacity in some strains of pathogenic bacteria [43]; and to the *ula* operon involved with an ascorbate transport system under anaerobic conditions that can also be considered an important virulence factor for this species [41].

The accessory portion, comprising the shared portion and the differential, is characterized by genes that confer benefits to the microorganism under certain environmental conditions. The differential portion, as observed in *in silico* assays, has a strong relationship with HGT processes, containing sequence-encoding proteins common to plasmids and phages. This region can result in important adaptations, influencing the differentiated interaction of the pathogen

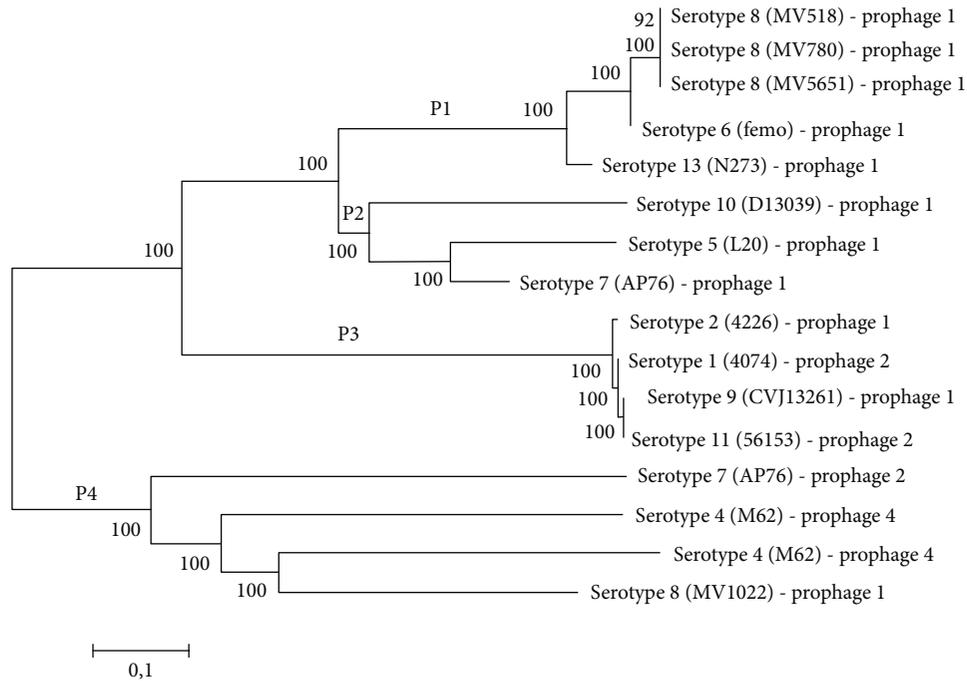


FIGURE 4: Phylogenetic relationships between sequences related to complete prophages found in *A. pleuropneumoniae*. The phylogenetic tree using the Neighbor-Joining method with 2000 bootstraps was generated by the MEGA 6 program after alignment by MAFFT and GBLOCKS. The scale is represented below, with 0.1 nucleotide substitutions per site.

TABLE 4: Identification of complete prophage sequences of *A. pleuropneumoniae*.

Strain/serotype	Prophage sequence identified	Organism/accession	ID %	Coverage	<i>E</i> value
4074/1	2	<i>Haemophilus ducreyi</i> /AE017143.1	97	85	0.0
4226/2	1	<i>Haemophilus ducreyi</i> /AE017143.1	98	95	0.0
M62/4	2	<i>Actinobacillus equuli</i> /CP007715.1	94	7	0.0
M62/4	4	<i>Mannheimia haemolytica</i> /KP137440.1	86	76	0.0
L20/5b	1	<i>Mannheimia</i> sp./CP006942.1	87	23	0.0
Femo/6	1	<i>Mannheimia haemolytica</i> /CP004753.1	89	25	0.0
AP76/7	1	<i>Mannheimia</i> sp./CP006942.1	89	26	0.0
	2	<i>Actinobacillus suis</i> /CP009159.1	89	86	0.0
MV518/8	1	<i>Mannheimia haemolytica</i> /CP004753.1	86	12	0.0
MV780/8	1	<i>Mannheimia haemolytica</i> /CP004753.1	86	14	0.0
MV1022/8	1	<i>Mannheimia haemolytica</i> /CP004753.1	86	12	0.0
MV5651/8	1	<i>Mannheimia haemolytica</i> /CP004753.1	86	15	0.0
CVJ13261/9	1	<i>Haemophilus ducreyi</i> /AE017143.1	97	89	0.0
D13039/10	1	<i>Mannheimia haemolytica</i> /CP004753.1	83	23	0.0
56153/11	2	<i>Haemophilus ducreyi</i> /AE017143.1	97	85	0.0
N273/13	1	<i>Mannheimia haemolytica</i> /CP004753.1	86	14	0.0

with the host, as well as having an important role in the differentiation of serotypes and mechanisms of virulence. As regards the differential portion, few differential C were affiliated with the COG categories. As this part of the genome has not been studied in a judicious way, we have a great network of sequences that codify proteins characterized as hypothetical, which are not categorized in the COG analysis. Among the sequences found in the differential portion, two sequences relating to the LysR family are present in the refer-

ence genome of *A. pleuropneumoniae* serotype 8. LysR is a family of transcriptional regulators that regulate a diverse set of genes, including those involved in virulence, metabolism, quorum sensing, and motility [44]. This regulator has also been related to processes of regulation of genes that code for urease in pathogenic bacteria [45]. In the differential portion of reference genome serotype 8, sequences encoding tetracycline, florfenicol, and sulfonamide resistance proteins were also found. In previous studies, the



identified high similarity and coverage with sequences present in the genomes of *A. suis*, *M. haemolytica*, and *H. ducreyi*, which may be related to HGT among species belonging to the family Pasteurellaceae. *A. suis* is commonly found in swine as tonsil commensal, but in the presence of unknown stimuli, it may invade the bloodstream, causing septicemia and sequelae, such as meningitis and arthritis, and even lead to the death of the host [52]. On the other hand, *M. haemolytica* is frequently involved in respiratory diseases in cattle [53] while *H. ducreyi* is a bacterium that causes soft chancre, a sexually transmitted disease in humans, and which has pigs as a model for studying the disease [54, 55]. Of the 6 sequences with high similarity and coverage identified in GenBank, only prophage 4 of the M62 strain had significant alignment correspondence with the phage sequence already described in the literature. This prophage was found in *M. haemolytica* and named vB\_MhM\_3927AP2 by the authors, being a phage belonging to the *Myoviridae* family [56]. The remaining 13 prophages have low identity and coverage in biological databases, suggesting that they may be phages unique to this species or not reported yet.

In conclusion, the genome of *A. pleuropneumoniae* serotype 8 is conserved in relation to the other serotypes, being more related to serotypes 3 and 6, which justifies the problems of serotyping to distinguish these three serotypes. We detected strong evidence of DNA sequence acquisition and recombination in the genomes of the different isolates/serotypes, and these differences were attributed to the presence of mobile genetic material, mainly prophages. In this study, we have identified 16 complete prophages, 6 of which may have suffered HGT among species belonging to the family Pasteurellaceae. However, the other prophages seem to be exclusive of *A. pleuropneumoniae* and not yet reported in the literature. Thus, prophages seem to play a key role in the restructuring of genomes and in the emergence of new strains of this pathogen.

## Data Availability

The data that support the results of this study are available in databases described in the manuscript and from the corresponding authors upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

We are grateful to the Núcleo de Análise de Biomoléculas of the Universidade Federal de Viçosa for providing facilities for the implementation of the experiments and Tomas Gomes Reis Veloso for helping with additional analysis. This work was supported by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG; APQ-02732-15), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/PROEX), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## Supplementary Materials

Supplementary Data Table 1: differential protein sequences found in *A. pleuropneumoniae* serotype 8. Supplementary Data Table 2: putative prophage sequences identified in the *A. pleuropneumoniae* genomes analyzed. (*Supplementary Materials*)

## References

- [1] OECD, *Meat consumption (indicator)*, 2019, October 2019.
- [2] B. Clark, L. A. Panzone, G. B. Stewart et al., “Consumer attitudes towards production diseases in intensive production systems,” *PLoS One*, vol. 14, no. 1, article e0210432, 2019.
- [3] K. D. Stärk, “Epidemiological investigation of the influence of environmental risk factors on respiratory diseases in swine—a literature review,” *The Veterinary Journal*, vol. 159, no. 1, pp. 37–56, 2000.
- [4] S. Pohl, H. U. Bertschinger, W. Frederiksen, and W. Mannheim, “Transfer of *Haemophilus pleuropneumoniae* and the *Pasteurella haemolytica*-like organism causing porcine necrotic pleuropneumonia to the genus *Actinobacillus* (*Actinobacillus pleuropneumoniae* comb. nov.) on the basis of phenotypic and deoxyribonucleic acid relatedness,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 33, no. 3, pp. 510–514, 1983.
- [5] R. Sárközi, L. Makrai, and L. Fodor, “Identification of a proposed new serovar of *Actinobacillus pleuropneumoniae*: serovar 16,” *Acta Veterinaria Hungarica*, vol. 63, no. 4, pp. 444–450, 2015.
- [6] J. T. Bossé, Y. Li, R. Fernandez Crespo et al., “Comparative sequence analysis of the capsular polysaccharide loci of *Actinobacillus pleuropneumoniae* serovars 1–18, and development of two multiplex PCRs for comprehensive capsule typing,” *Veterinary Microbiology*, vol. 220, pp. 83–89, 2018.
- [7] J. T. Bossé, Y. Li, R. Sárközi et al., “Proposal of serovars 17 and 18 of *Actinobacillus pleuropneumoniae* based on serological and genotypic analysis,” *Veterinary Microbiology*, vol. 217, pp. 1–6, 2018.
- [8] K. Chiers, T. De Waele, F. Pasmans, R. Ducatelle, and F. Haesebrouck, “Virulence factors of *Actinobacillus pleuropneumoniae* involved in colonization, persistence and induction of lesions in its porcine host,” *Veterinary Research*, vol. 41, no. 5, p. 65, 2010.
- [9] M. F. Pereira, C. C. Rossi, L. E. Seide, S. Martins Filho, C. M. Dolinski, and D. M. S. Bazzolli, “Antimicrobial resistance, biofilm formation and virulence reveal *Actinobacillus pleuropneumoniae* strains' pathogenicity complexity,” *Research in Veterinary Science*, vol. 118, pp. 498–501, 2018.
- [10] E. L. Sassu, J. T. Bossé, T. J. Tobias, M. Gottschalk, P. R. Langford, and I. Hennig-Pauka, “Update on *Actinobacillus pleuropneumoniae* — knowledge, gaps and challenges,” *Transboundary and Emerging Diseases*, vol. 65, pp. 72–90, 2018.
- [11] J. T. Bossé, S. Sinha, T. Schippers, J. S. Kroll, R. J. Redfield, and P. R. Langford, “Natural competence in strains of *Actinobacillus pleuropneumoniae*,” *FEMS Microbiology Letters*, vol. 298, no. 1, pp. 124–130, 2009.
- [12] J. T. Bossé, D. M. Soares-Bazzolli, Y. Li et al., “The generation of successive unmarked mutations and chromosomal insertion of heterologous genes in *Actinobacillus pleuropneumoniae*

- using natural transformation,” *PLoS One*, vol. 9, no. 11, article e111252, 2014.
- [13] G. C. da Silva, C. C. Rossi, M. F. Santana, P. R. Langford, J. T. Bossé, and D. M. S. Bazzolli, “p518, a small *floR* plasmid from a South American isolate of *Actinobacillus pleuropneumoniae*,” *Veterinary Microbiology*, vol. 204, pp. 129–132, 2017.
- [14] A. Beceiro, M. Tomas, and G. Bou, “Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world?,” *Clinical Microbiology Reviews*, vol. 26, no. 2, pp. 185–230, 2013.
- [15] M. Schroeder, B. D. Brooks, and A. E. Brooks, “The complex relationship between virulence and antibiotic resistance,” *Genes*, vol. 8, no. 1, p. 39, 2017.
- [16] Z. Xu, X. Chen, L. Li et al., “Comparative genomic characterization of *Actinobacillus pleuropneumoniae*,” *Journal of Bacteriology*, vol. 192, no. 21, pp. 5625–5636, 2010.
- [17] C. O’Neill, S. C. P. Jones, J. T. Bosse et al., “Prevalence of *Actinobacillus pleuropneumoniae* serovars in England and Wales,” *Veterinary Record*, vol. 167, no. 17, pp. 661–662, 2010.
- [18] Y. Li, J. T. Bossé, S. M. Williamson et al., “*Actinobacillus pleuropneumoniae* serovar 8 predominates in England and Wales,” *Veterinary Record*, vol. 179, pp. 276–277, 2016.
- [19] M. Gottschalk and S. Lacouture, “*Actinobacillus pleuropneumoniae* serotypes 3, 6, 8 and 15 isolated from diseased pigs in North America,” *Veterinary Record*, vol. 174, no. 18, p. 452, 2014.
- [20] C. C. Rossi, A. M. Vicente, W. V. Guimarães, E. F. Araújo, M. V. De Queiroz, and D. M. S. Bazzolli, “Face to face with *Actinobacillus pleuropneumoniae*: landscape of the distribution of clinical isolates in Southeastern Brazil,” *African Journal of Microbiology Research*, vol. 7, no. 23, pp. 2916–2924, 2013.
- [21] M. F. Pereira, C. C. Rossi, F. M. de Carvalho et al., “Draft genome sequences of six *Actinobacillus pleuropneumoniae* serotype 8 Brazilian clinical isolates: insight into new applications,” *Genome Announcements*, vol. 3, no. 2, article e01585-14, 2015.
- [22] J. T. Bossé, R. R. Chaudhuri, Y. Li et al., “Complete genome sequence of MIDG2331, a genetically tractable serovar 8 clinical isolate of *Actinobacillus pleuropneumoniae*,” *Genome Announcements*, vol. 4, no. 1, article e01667-15, 2016.
- [23] M. F. Pereira, C. C. Rossi, M. Vieira de Queiroz et al., “*Galleria mellonella* is an effective model to study *Actinobacillus pleuropneumoniae* infection,” *Microbiology*, vol. 161, no. 2, pp. 387–400, 2015.
- [24] W. Li and A. Godzik, “Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences,” *Bioinformatics*, vol. 22, no. 13, pp. 1658–1659, 2006.
- [25] L. Fu, B. Niu, Z. Zhu, S. Wu, and W. Li, “CD-HIT: accelerated for clustering the next-generation sequencing data,” *Bioinformatics*, vol. 28, no. 23, pp. 3150–3152, 2012.
- [26] R. L. Tatusov, M. Y. Galperin, D. A. Natale, and E. V. Koonin, “The COG database: a tool for genome-scale analysis of protein functions and evolution,” *Nucleic Acids Research*, vol. 28, no. 1, pp. 33–36, 2000.
- [27] A. Marchler-Bauer, S. Lu, J. B. Anderson et al., “CDD: a conserved domain database for the functional annotation of proteins,” *Nucleic Acids Research*, vol. 39, no. Database, pp. D225–D229, 2011.
- [28] R. D. Finn, J. Tate, J. Mistry et al., “The Pfam protein families database,” *Nucleic Acids Research*, vol. 36, pp. D281–D288, 2008.
- [29] I. Letunic, R. R. Copley, B. Pils, S. Pinkert, J. Schultz, and P. Bork, “SMART 5: domains in the context of genomes and networks,” *Nucleic Acids Research*, vol. 34, no. 90001, pp. D257–D260, 2006.
- [30] C. H. Wu, R. Apweiler, A. Bairoch et al., “The universal protein resource (UniProt): an expanding universe of protein information,” *Nucleic Acids Research*, vol. 34, no. 90001, pp. D187–D191, 2006.
- [31] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic local alignment search tool,” *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.
- [32] P. Rice, I. Longden, and A. Bleasby, “EMBOSS: the European Molecular Biology Open Software Suite,” *Trends in Genetics*, vol. 16, no. 6, pp. 276–277, 2000.
- [33] A. C. E. Darling, B. Mau, F. R. Blattner, and N. T. Perna, “Mauve: multiple alignment of conserved genomic sequence with rearrangements,” *Genome Research*, vol. 14, no. 7, pp. 1394–1403, 2004.
- [34] D. Arndt, J. R. Grant, A. Marcu et al., “PHASTER: a better, faster version of the PHAST phage search tool,” *Nucleic Acids Research*, vol. 44, no. W1, pp. W16–W21, 2016.
- [35] K. Katoh and D. M. Standley, “MAFFT multiple sequence alignment software version 7: improvements in performance and usability,” *Molecular Biology and Evolution*, vol. 30, no. 4, pp. 772–780, 2013.
- [36] J. Castresana, “Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis,” *Molecular Biology and Evolution*, vol. 17, no. 4, pp. 540–552, 2000.
- [37] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, “MEGA6: molecular evolutionary genetics analysis version 6.0,” *Molecular Biology and Evolution*, vol. 30, no. 12, pp. 2725–2729, 2013.
- [38] F. Sievers and D. G. Higgins, “Clustal Omega, accurate alignment of very large numbers of sequences,” *Methods in Molecular Biology*, vol. 1079, pp. 105–116, 2014.
- [39] S. Subashchandrabose, R. M. LeVeque, T. K. Wagner, R. N. Kirkwood, M. Kiupel, and M. H. Mulks, “Branched-chain amino acids are required for the survival and virulence of *Actinobacillus pleuropneumoniae* in swine,” *Infection and Immunity*, vol. 77, no. 11, pp. 4925–4933, 2009.
- [40] Z. Xu, H. Chen, and R. Zhou, “Genome-wide evidence for positive selection and recombination in *Actinobacillus pleuropneumoniae*,” *BMC Evolutionary Biology*, vol. 11, no. 1, p. 203, 2011.
- [41] V. Deslandes, M. Denicourt, C. Girard, J. Harel, J. H. E. Nash, and M. Jacques, “Transcriptional profiling of *Actinobacillus pleuropneumoniae* during the acute phase of a natural infection in pigs,” *BMC Genomics*, vol. 11, no. 1, p. 98, 2010.
- [42] M. M. Barnhart and M. R. Chapman, “Curli biogenesis and function,” *Annual Review of Microbiology*, vol. 60, no. 1, pp. 131–147, 2006.
- [43] S. A. Jones, M. Jorgensen, F. Z. Chowdhury et al., “Glycogen and maltose utilization by *Escherichia coli* O157: H7 in the mouse intestine,” *Infection and Immunity*, vol. 76, no. 6, pp. 2531–2540, 2008.
- [44] S. E. Maddocks and P. C. F. Oyston, “Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins,” *Microbiology*, vol. 154, no. 12, pp. 3609–3623, 2008.

- [45] J. T. Bossé, H. D. Gilmour, and J. I. MacInnes, "Novel genes affecting urease activity in *Actinobacillus pleuropneumoniae*," *Journal of Bacteriology*, vol. 183, no. 4, pp. 1242–1247, 2001.
- [46] J. T. Bossé, Y. Li, T. G. Atherton et al., "Characterisation of a mobilisable plasmid conferring florfenicol and chloramphenicol resistance in *Actinobacillus pleuropneumoniae*," *Veterinary Microbiology*, vol. 178, no. 3-4, pp. 279–282, 2015.
- [47] Y. Li, G. C. da Silva, Y. Li et al., "Evidence of illegitimate recombination between two Pasteurellaceae plasmids resulting in a novel multi-resistance replicon, pM3362MDR, in *Actinobacillus pleuropneumoniae*," *Frontiers in Microbiology*, vol. 9, p. 2489, 2018.
- [48] P. Siguier, E. Gourbeyre, and M. Chandler, "Bacterial insertion sequences: their genomic impact and diversity," *FEMS Microbiology Reviews*, vol. 38, no. 5, pp. 865–891, 2014.
- [49] M. Chandler and J. Mahillon, "Insertion sequences revisited," in *Mobile DNA II*, pp. 305–366, American Society of Microbiology, 2002.
- [50] T. Ooka, Y. Ogura, M. Asadulghani et al., "Inference of the impact of insertion sequence (IS) elements on bacterial genome diversification through analysis of small-size structural polymorphisms in *Escherichia coli* O157 genomes," *Genome Research*, vol. 19, no. 10, pp. 1809–1816, 2009.
- [51] L. C. Fortier and O. Sekulovic, "Importance of prophages to evolution and virulence of bacterial pathogens," *Virulence*, vol. 4, no. 5, pp. 354–365, 2013.
- [52] A. R. Bujold and J. I. MacInnes, "Identification of putative adhesins of *Actinobacillus suis* and their homologues in other members of the family Pasteurellaceae," *BMC Research Notes*, vol. 8, no. 1, p. 675, 2015.
- [53] J. Gioia, X. Qin, H. Jiang et al., "The genome sequence of *Mannheimia haemolytica* A1: insights into virulence, natural competence, and Pasteurellaceae phylogeny," *Journal of Bacteriology*, vol. 188, no. 20, pp. 7257–7266, 2006.
- [54] M. M. Hobbs, L. R. San Mateo, P. E. Orndorff, G. Almond, and T. H. Kawula, "Swine model of *Haemophilus ducreyi* infection," *Infection and Immunity*, vol. 63, no. 8, pp. 3094–3100, 1995.
- [55] W. G. Fusco, N. R. Choudhary, P. A. Routh et al., "The *Haemophilus ducreyi* trimeric autotransporter adhesin DsrA protects against an experimental infection in the swine model of chancroid," *Vaccine*, vol. 32, no. 30, pp. 3752–3758, 2014.
- [56] Y. D. Niu, S. R. Cook, J. Wang et al., "Comparative analysis of multiple inducible phages from *Mannheimia haemolytica*," *BMC Microbiology*, vol. 15, no. 1, p. 175, 2015.