JOURNAL OF BACTERIOLOGY, Nov. 2010, p. 5625–5636 Vol. 192, No. 21
0021-9193/10/$12.00 doi:10.1128/JB.00535-10

Comparative Genomic Characterization of *Actinobacillus pleuropneumoniae*▽†

Zhuofei Xu,[1] Xiabing Chen,[1] Lu Li,[1] Tingting Li,[1] Shengyue Wang,[2] Huanchun Chen,[1] and Rui Zhou[1]*

Division of Animal Infectious Disease in the State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China,[1] and Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai 201203, China[2]

Received 10 May 2010/Accepted 28 July 2010

The Gram-negative bacterium *Actinobacillus pleuropneumoniae* is the etiologic agent of porcine contagious pleuropneumoniae, a lethal respiratory infectious disease causing great economic losses in the swine industry worldwide. In order to better interpret the genetic background of serotypic diversity, nine genomes of *A. pleuropneumoniae* reference strains of serovars 1, 2, 4, 6, 9, 10, 11, 12, and 13 were sequenced by using rapid high-throughput approach. Based on 12 genomes of corresponding serovar reference strains including three publicly available complete genomes (serovars 3, 5b, and 7) of this bacterium, we performed a comprehensive analysis of comparative genomics and first reported a global genomic characterization for this pathogen. Clustering of 26,012 predicted protein-coding genes showed that the pan genome of *A. pleuropneumoniae* consists of 3,303 gene clusters, which contain 1,709 core genome genes, 822 distributed genes, and 772 strain-specific genes. The genome components involved in the biogenesis of capsular polysaccharide and lipopolysaccharide O antigen relative to serovar diversity were compared, and their genetic diversity was depicted. Our findings shed more light on genomic features associated with serovar diversity of *A. pleuropneumoniae* and provide broader insight into both pathogenesis research and clinical/epidemiological application against the severe disease caused by this swine pathogen.

Actinobacillus pleuropneumoniae, a Gram-negative facultative anaerobic encapsulated coccobacillus, belongs to the *Actinobacillus* genus of the *Pasteurellaceae* family (19). *A. pleuropneumoniae* is a primary bacterial etiologic agent of porcine contagious pleuropneumonia, a severe respiratory disease leading to great economic losses to the global swine industry (7). The cases usually display pleuropneumonia and pulmonary lesions characterized by serious hemorrhage and necrosis. To date, several factors involved in the virulence of *A. pleuropneumoniae* have been described, including Apx exotoxins, capsular polysaccharides (CPS), lipopolysaccharides (LPS), outer membrane proteins, iron-acquisition proteins and adhesin factors (11, 19, 24). However, the genetic differences of pathogenesis remain poorly characterized and are worth interpreting from the perspective of comparative genomics for this bacterium.

Thus far, 15 serovars and two biotypes of *A. pleuropneumoniae* have been recognized, with great variations in virulence and interlocal distributions (6). The predominant serovar-specific antigens are composed of CPS, which could rigorously define serovars of *A. pleuropneumoniae* (6, 34). Antigenic differences in the LPS can further determine *A. pleuropneumoniae* subtypes within a same capsular serovar (13). The metabolic and virulent characteristics of this pathogen have been systematically described based on the prior knowledge and two complete genomes (18, 47), but the molecular basis and evolutionary mechanism of serotypic diversity are still not well explained due to the lack of sequence information. To investigate the associations of serovar diversity with the underlying genetic components, more serovar-related genomic islands involved in the biosynthesis of capsular and lipopolysaccharide antigens should be decoded at the pan-genome level of *A. pleuropneumoniae*. At present, through the next-generation of sequencing technique (454 GS FLX pyrosequencing platform), more and more bacterial species, subspecies or typical strains have been quickly sequenced, such as eight species in the *Yersinia* genus (9), 17 strains of *Streptococcus pneumoniae* (22), and 5 strains from different *Francisella tularensis* subspecies (8). Multiple genome sequences from different strains of a single species can offer comprehensive information to explore the relationship between genotypes and phenotypes and to further discover additional genetic markers for clinical purpose.

In the present study, we sequenced the *A. pleuropneumoniae* genomes of nine reference strains of serovars 1, 2, 4, 6, 9, 10, 11, 12, and 13. Together with three public complete genome sequences of *A. pleuropneumoniae* serovars 3, 5b, and 7, the analysis of comparative genomics was performed to report a global genomic characterization of this pathogenic bacterium. The acquisition and loss of genome compositions that contribute to virulence and serovar diversity were identified. The genetic loci involved in the biogenesis of capsule and O-specific polysaccharide were compared, and their vital roles in serotypic diversity were investigated.

MATERIALS AND METHODS

Bacterial strains. Nine reference strains from different *A. pleuropneumoniae* serovars sequenced in the present study (Table 1) were kindly provided by Pat

* Corresponding author. Mailing address: Division of Animal Infectious Disease, State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China. Phone: 86-27-87281878. Fax: 86-27-87282608. E-mail: rzhou@mail.hzau.edu.cn.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 27 August 2010.

TABLE 1. *A. pleuropneumoniae* strains and publicly available genomes used in this study

Virulence	Serovar	Strain	GenBank accession no.	Source or reference
High	1	4074		This study
Medium	2	S1536		This study
Low	3	JL03	CP000687	47
Medium	4	M62		This study
High	5b	L20	CP000569	18
Medium	6	Femφ		This study
Medium	7	Ap76	CP001091	
High	9	CVJ13261		This study
Rarely isolated	10	D13039		This study
High	11	56153		This study
Medium	12	1096		This study
Rarely isolated	13	N273		This study

Blackall of Australasian Pig Institute, Australia. All strains were grown overnight in tryptic soy broth (TSB) medium at 37°C shaking on a rotary shaker (200 rpm), supplemented with 10 μg of NAD/ml and 10% bovine serum. Total genomic DNA was extracted by using the DNeasy tissue kit (Qiagen).

Sequencing and assembly. Bacterial genomes were sequenced at the Chinese National Human Genome Center at Shanghai using the 454 GS FLX platform (Roche, Germany) (32). For each sample, a library containing fragments (500 to 800 bp) for sequencing was prepared from 4 μg of genomic DNA. On average, 59 contigs per genome were assembled by using the 454 Newbler *de novo* assembler (version 2.3).

Genome annotation. Nine newly sequenced draft genomes and three publicly available complete genomes of different *A. pleuropneumoniae* serovars were annotated by using an automated bacterial annotation pipeline DIYA (40) and custom Perl scripts. Using DIYA, the genomic contigs were tiled against a complete genome (from strain L20) by the promer script in the MUMmer package (28). The order of the matching contigs was thus inferred, and all contigs for each sample were concatenated into a pseudogenome with random non-matching contigs on the end. To identify the interrupt protein-coding genes at the terminals of a contig, a contig linker (NNNNNCATTCCATTCATTAATT AATTAATGAATGAATGNNNNN) that contains stop and start codons in all six reading frames was used (44). Each pseudogenome was then used for the identification of protein-coding sequences (CDSs), tRNAs, and rRNAs implemented by the programs Glimmer3 (37), tRNASCAN-SE v1.23 (30), and RNAmmer v1.2 (29), respectively. All protein-coding genes translated in all six reading frames were searched against the proteins from the UniRef50 database (which was updated in February 2010) (42) by using Blastx (10^{-10} cutoff E-value and 35% minimum identity) (2). Gene annotation was supplied with the cluster of orthologous groups (COG) database (43) by using rpsblast (10^{-10} cutoff E-value) (2). A combined data set including nine genomes sequenced in the present study and three reannotated intact genomes was generated for the following analysis of comparative genomics for *A. pleuropneumoniae*.

Accession numbers. The draft genome assemblies for the nine *A. pleuropneumoniae* strains were deposited in the GenBank database (http://www.ncbi.nih.gov/GenBank/index.html). The accession numbers for these sequences, as well as the accession numbers for three publicly available *A. pleuropneumoniae* complete genomes, were as follows with corresponding strains: ADOD00000000 (4074), ADOE00000000 (S1536), ADOF00000000 (M62), ADOG00000000 (Femφ), ADOI00000000 (CVJ13261), ADOJ00000000 (D13039), ADOK00000000 (56153), ADOL00000000 (1096), ADOM00000000 (N273), CP000569 (L20), CP000687 (JL03), and CP001091 (AP76).

Whole-genome alignment. Based on the blastn hits (minimum identity of 95% and a cutoff E-value of 10^{-5}) of a reference complete genome (from strain L20) searched against the other genomes of *A. pleuropneumoniae*, genome comparative circular maps were constructed by using CGview software (41). Multiple *A. pleuropneumoniae* genomes were aligned and visualized by using Mauve v2.3.1 with the default settings (12).

Gene clustering and alignment. The CDSs were extracted from all annotated genomes of *A. pleuropneumoniae*. The small CDSs with length shorter than 40 amino acids were removed from the data set. The orthologous proteins were grouped by using the program Blastclust (1). An all-versus-all BLASTP of all proteomes was first performed to define the orthologous pairs satisfying the following criteria: a cutoff E-value of 10^{-6}, over 70% length coverage, and at least 70% identity. Pairs of sequences that have statistically significant matches were then clustered into the same group by using single-linkage clustering. Multiple protein sequence alignment for each cluster was performed with the program MUSCLE 3.6 (14). The corresponding nucleotide sequence alignment was produced based on the aligned amino acid sequences from each gene cluster using custom Perl scripts.

Phylogeny of *A. pleuropneumoniae* serovars. To infer the evolutionary relationships between 12 different serovar reference strains of *A. pleuropneumoniae*, two types of dendrograms were generated, respectively. According to a gene possession-based phylogenetic approach, we defined the genetic distance between a pair of genomes (i and k), to be $\Sigma_n | g_{n,i} - g_{n,k} |$, where $g_{n,i}$ is 1 if gene n is present in strain i and 0 otherwise (23). The distance matrix was then used to reconstruct species phylogeny with the unweighted pair group method with arithmetic mean (UPGMA) method implemented in the Phylip package (15). The second type of phylogenetic tree was reconstructed using sequence alignments of 1,287 single-copy core genes with nearly identical length and exactly one member in each of the 12 genomes. These gene alignments were concatenated into a large alignment of 1,211,061 nucleotides. A maximum-likelihood tree was built under the HKY85 substitution model with the estimated transition/transversion rate ratio (κ) and gamma distributed rate heterogeneity of four categories (Γ4) in PhyML (20).

RESULTS AND DISCUSSION

General features of sequenced genomes. In the present study, nine genomes from different *A. pleuropneumoniae* serovar reference strains, 4074 of serovar 1, S1536 of serovar 2, M62 of serovar 4, Femφ of serovar 6, CVJ13261 of serovar 9, D13039 of serovar 10, 56153 of serovar 11, 1096 of serovar 12, and N273 of serovar 13, were sequenced by using 454 GS FLX. The depth of each genomic sequencing is 14- to 26-fold, and the average number of assembled contigs for each genome is 59 with a range from 44 to 89 (see Table S1 in the supplemental material). All nine genome draft sequences have been submitted to GenBank.

Although these genomes were extracted from the reference strains of different *A. pleuropneumoniae* serovars that can be assigned to three levels of virulence (Table 1) (10), the overall genomic characteristics were quite similar (Table 2). All of the genomes were comprised of a circular chromosome with ~2.19 Mb ~2.33 Mb in length. The average GC content of each genome was 41%, which was consistent with that of an entire *A. pleuropneumoniae* chromosome (47). The median number of CDSs per strain was 2,174, the largest number was 2,223 for serovar 4 strain M62, and the least was 2,096 for serovar 12 strain 1096. Pairwise nucleotide alignments using blastn (>95% identity) revealed high sequence conservation between each draft genome and the *A. pleuropneumoniae* serovar 5b strain L20 complete genome (Fig. 1). The percentage of total length of matched sequences accounting for the L20 genome (2,274,482 bp) was ranged from ~90.8% (S1536 versus L20) to ~92.7% (56153 versus L20). Meanwhile, the global pairwise genomic alignment also showed several large genetic differences that may be relative to bacterial virulence and serotypic diversity. Notably, the genomic regions bearing aberrant GC content may represent the occurrence of horizontal gene transfer events in different serovar reference strain, such as biosynthetic loci of the LPS O antigen and capsule (Fig. 1). Detailed analyses of these featured genomic regions using the local multiple sequence alignments among subsets of genomes are described below.

Identification of gene clusters. A multi-fasta file with 26,012 CDSs of all 12 *A. pleuropneumoniae* genomes used for clustering was available in Text S1 in the supplemental material. The

TABLE 2. Summary of genome features and gene clusters of *A. pleuropneumoniae*

Serovar	Strain	Genome size (Mb)	G+C (%)	CDSs	Total gene clusters	Distributed clusters	Unique clusters	Noncore clusters (%)
1	4074	2.26	41.2	2,180	2,150	415	26	20.5
2	S1536	2.22	41.1	2,137	2,111	351	51	19.0
3	JL03	2.24	41.2	2,101	2,048	289	50	16.6
4	M62	2.27	41.2	2,223	2,193	317	167	22.1
5b	L20	2.27	41.3	2,137	2,076	289	78	17.7
6	Femφ	2.31	41.0	2,219	2,184	375	100	21.7
7	AP76	2.33	41.2	2,203	2,134	377	48	19.9
9	CVJ13261	2.26	41.2	2,204	2,178	432	37	21.5
10	D13039	2.27	41.2	2,168	2,137	321	107	20.0
11	56153	2.27	41.2	2,195	2,164	438	17	21.0
12	1096	2.19	41.2	2,096	2,072	302	61	17.5
13	N273	2.25	41.2	2,149	2,124	385	30	19.5

total number of *A. pleuropneumoniae* orthologous gene clusters (designated APO, hereafter), also including unique genes that were exclusive by only a single strain, was 3,303. Of these, 52% were identified to be core gene clusters that were shared by all strains and accounts for 79% of the total number of CDSs, 25% were dispensable gene clusters that were found to be possessed by at least two strains but not all, and the remaining 23% were unique genes, only accounting for ca. 3% of the

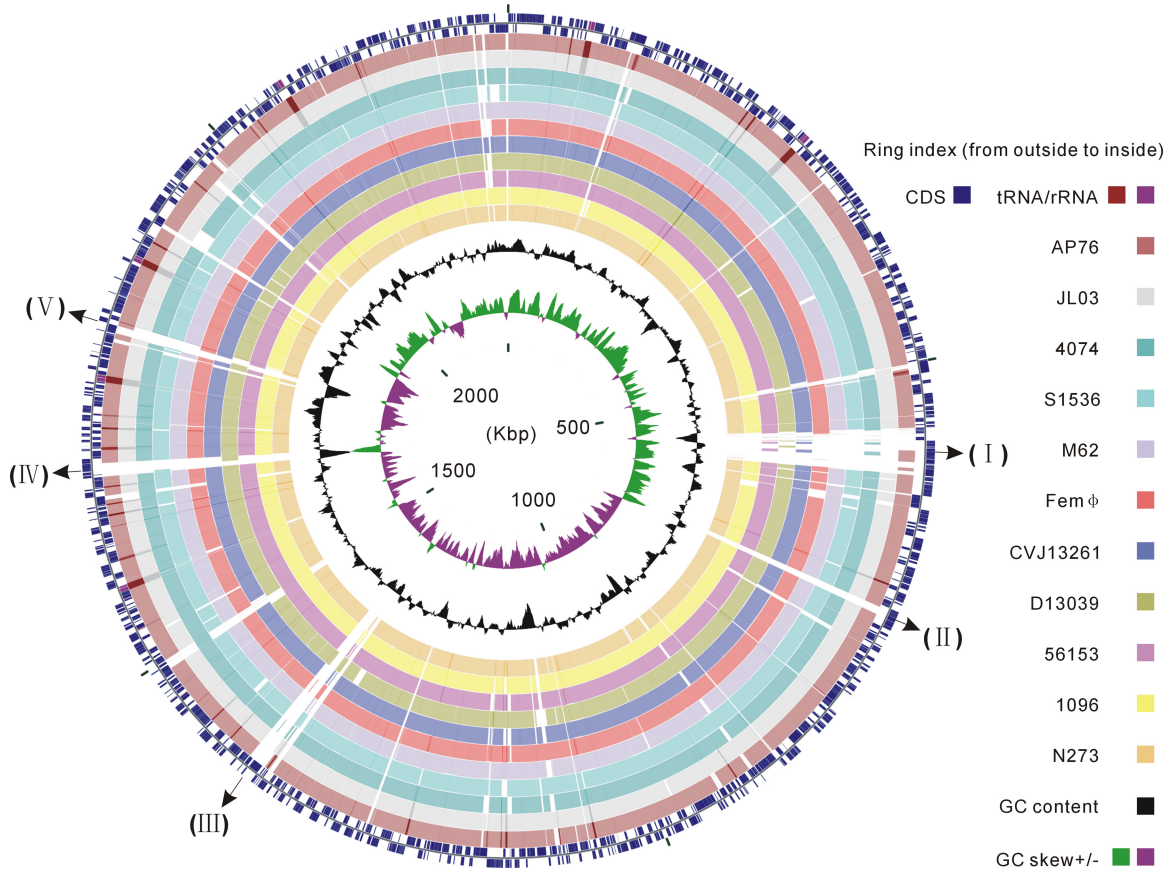

FIG. 1. Circular representation of sequence conservation between *A. pleuropneumoniae* serovar 5b strain L20 and 11 strains belonging to different serovars. Circles are numbered from 1 (outermost circle) to 16 (innermost circle). The outermost two circles show CDSs, rRNAs and tRNAs in the L20 genome of *A. pleuropneumoniae* serovar 5b. The next 11 circles show the coordinates of BLAST hits detected through blastn comparisons (minimum sequence identity of 95% and expected threshold of 10^{-5}) of the L20 reference genome against 11 *A. pleuropneumoniae* genomes, including two public complete genomes and nine contig sets of new genomes, and each circle is colored according to serovar reference strains: maroon for serovar 7 strain AP76, silver for serovar 3 strain JL03, teal for serovar 1 strain 4074, cyan for serovar 2 strain S1536, light purple for serovar 4 strain M62, red for serovar 6 strain Femφ, blue for serovar 9 strain CVJ13261, olive for serovar 10 strain D13039, fuchsia for serovar 11 strain 56153, yellow for serovar 12 strain 1096, and orange for serovar 13 strain N273. Overlapping hits appear as darker arcs. The innermost two circles show GC content and GC skew plot of the L20 genome. Several known serovar-specific genomic regions with low sequence identity were numbered as follows: I, the ~38-kb prophage region; II and III, the coding gene cluster involved in type I restriction-modification system; IV, the LPS O-antigen biosynthesis region; V, the CPS biosynthesis region.

TABLE 3. Summary of gene clusters for 12 *A. pleuropneumoniae* strains

Gene class	No. of orthologous clusters	No. of CDSs
Core	1,709	20,597
Distributed	822	4,411
Unique	772	772
Excluded by size		232
Total	3,303	26,012

total CDSs (Table 3). A gene clustering table that contained the gene content of 12 *A. pleuropneumoniae* genomes was summarized (see Table S2 in the supplemental material) and the relative identifiers of CDSs were listed (see Table S3 in the supplemental material). Each genome contained some strain-specific protein coding genes, with a range of 167 for strain M62 of serovar 4 to 17 for strain 56153 of serovar 11 (Table 2). Among 1,594 noncore clusters, including distributed and unique genes, serovar 3 strain JL03 had the lowest percentage (16.6%), and serovar 4 strain M62 had the highest percentage (22.1%) (Table 2). The pairwise comparison of gene content demonstrated that the average number of genes associated with the gain or loss between any two strains was 429, with a standard deviation of 97. The maximum and minimum numbers of genic differences, 611 and 96, were identified in strain pairs M62 (serovar 4)/CVJ13261 (serovar 9) and CVJ13261 (serovar 9)/56153 (serovar 11), respectively.

To some extent, the proteins encoded by the core genes present in all 12 genomes should participate in the fundamental metabolic activities of *A. pleuropneumoniae* and be essential for the growth and survival of this bacterium. The distribution of cellular functions of these core proteins indicated that protein-coding genes involved in translation were assigned into the largest category (8.54%) (Fig. 2; see also Table S4 in the supplemental material). As expected, there was no core protein involved in cell motility, which was coincident with the common phenotype of nonmotile. It was worth noting that a flagellin gene *fliC* reported previously in *A. pleuropneumoniae* (33) was absent in the genomic sequences of the 12 reference strains. The set of noncore proteins had relatively more elements involved in surface polysaccharide biogenesis and the bacterial pathogenic process compared to that of core proteins. These distributed or unique proteins that may play a potential role in differentiating serovars and virulence were assigned to the function categories of defense mechanism, replication and recombination, and the type I and III restriction-modification system, including diverse transposases, recombinases, integrases, and DNA helicases (see Table S5 in the supplemental material). Twenty-four genes encoding autotransporter adhesins involved in extracellular structures were found to be distributed or unique in 12 strains. In addition, ~5.1% of the unique genes and ~5.2% of distributed genes had annotations associated with phage, prophage, or bacteriophage; whereas very few phage protein coding genes of 0.2% was present in the core genes. Among 2,531 core and distributed orthologous gene clusters (Table 3), ca. 21.4% (542) were annotated as hypothetical or uncharacterized proteins (see Table S5 in the supplemental material), suggesting that a significant percentage of even the bacterial housekeeping genes remain unknown in *A. pleuropneumoniae*.

The differences of gene components between the high- and low-virulence strains may provide insight for identifying novel candidate virulence factors. The 54 distributed genes that were shared by the high-virulence strains from serovars

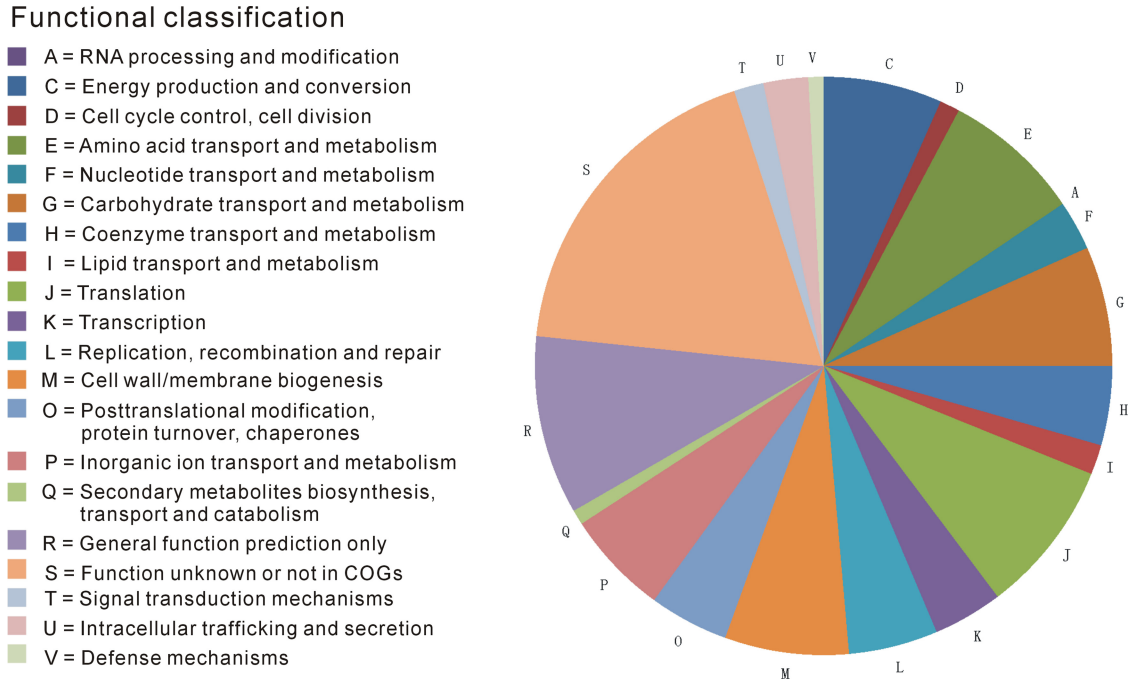

FIG. 2. Distribution of cellular function categories of core orthologous protein clusters.

TABLE 4. Gene clusters shared by the highly virulent serovars in *A. pleuropneumoniae*

Cluster	Distributed serotypes	Annotation of predicted protein product
APO_1721	1, 2, 5b, 6, 7, 9, 10, 11, 12, 13, 4	SNF2-related protein
APO_1723	1, 2, 7, 12, 13, 6, 5b, 9, 11, 4, 10	Truncated transferrin-binding protein 1
APO_1726	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Homoserine dehydrogenase
APO_1729	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	GTP pyrophosphokinase
APO_1730	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Elongation factor G
APO_1746	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Exodeoxyribonuclease 7 large subunit
APO_1764	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Malate dehydrogenase (oxaloacetate-decarboxylating) [NADP(+)], phosphate acetyltransferase
APO_1767	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Molybdopterin biosynthesis protein moeA
APO_1779	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Glycerophosphoryl diester phosphodiesterase
APO_1797	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Uncharacterized periplasmic iron-binding protein
APO_1806	1, 2, 4, 6, 7, 9, 10, 11, 12, 13, 5b	DacC protein
APO_1811	1, 2, 4, 5b, 6, 7, 10, 11, 12, 13, 9	Carbonic anhydrase
APO_1812	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	tRNA [adenine-*N*(6)-]-methyltransferase
APO_1817	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	FKBP-type peptidyl-prolyl *cis-trans* isomerase 1
APO_1825	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	50S ribosomal protein L10
APO_1852	1, 2, 4, 5b, 6, 7, 9, 10, 11, 12, 13	Hypothetical protein
APO_1866	1, 7, 9, 11, 13, 2, 4, 5b, 12, 6	Acetyltransferase
APO_1874	1, 2, 5b, 6, 7, 9, 10, 11, 13, 12	Threonine efflux protein
APO_1878	1, 2, 4, 5b, 7, 9, 10, 11, 12, 13	Uncharacterized protein
APO_1879	1, 2, 5b, 6, 7, 9, 10, 11, 12, 13	Endoribonuclease L-PSP
APO_1881	1, 2, 4, 5b, 6, 7, 9, 10, 11, 13	Putative uncharacterized protein
APO_1895	1, 9, 11, 4, 5b, 7, 10, 12, 13	ATP-dependent RNA helicase RhlB
APO_1896	1, 4, 5b, 7, 9, 11, 12, 13, 10	Flp operon protein C, RcpC
APO_1897	1, 2, 4, 5b, 7, 9, 11, 12, 13	Probable ABC transporter permease protein
APO_1901	7, 12, 1, 2, 5b, 6, 9, 11, 13	Putative uncharacterized protein
APO_1905	1, 2, 5b, 6, 7, 9, 11, 12, 13	Putative uncharacterized protein
APO_1916	1, 5b, 9, 11, 12, 6, 7, 13	Hemolysin activation/secretion protein
APO_1920	1, 2, 5b, 7, 9, 11, 12, 13	Transposase IS200-like protein
APO_1927	1, 2, 4, 5b, 9, 10, 11, 12	Hypothetical protein
APO_1961	10, 5b, 7, 1, 9, 11	Phage terminase large subunit
APO_1967	10, 5b, 7, 1, 9, 11	Bacteriophage capsid protein
APO_1968	1, 9, 11, 12, 5b, 10	Undecaprenyl-phosphate galactose phosphotransferase, WbaP
APO_1973	12, 10, 1, 5b, 9, 11	Hypothetical protein
APO_1987	2, 5b, 6, 1, 9, 11	Outer membrane receptor proteins, mostly Fe transport
APO_1988	10, 1, 5b, 7, 9, 11	Uncharacterized protein
APO_2004	1, 5b, 9, 10, 11	RTX-I toxin determinant A from serotypes 5/10, ApxIA
APO_2006	7, 1, 9, 11, 5b	ATP-dependent Clp protease proteolytic subunit
APO_2010	1, 5b, 9, 10, 11	Putative uncharacterized protein
APO_2012	1, 5b, 6, 9, 11	Predicted phage integrase family protein
APO_2016	1, 9, 11, 5b, 10	ABC transporter ATP binding subunit, Wzm
APO_2018	1, 5b, 9, 10, 11	Possible bacteriophage antirepressor
APO_2024	1, 5b, 7, 9, 11	Putative phage tail component
APO_2025	1, 5b, 9, 11, 10	Possible DNA methylase
APO_2026	5b, 10, 1, 9, 11	Hemolysin-activating lysine-acyltransferase HlyC, ApxIC
APO_2027	5b, 1, 9, 11, 12	Putative uncharacterized protein
APO_2030	5b, 10, 1, 9, 11	Putative uncharacterized protein
APO_2031	1, 5b, 7, 9, 11	Possible bacteriophage tail protein
APO_2033	1, 5b, 9, 10, 11	Hypothetical protein
APO_2038	1, 5b, 9, 10, 11	Hypothetical protein
APO_2039	10, 1, 5b, 9, 11	Hypothetical protein
APO_2040	1, 9, 11, 5b, 10	Hypothetical protein
APO_2043	1, 2, 5b, 9, 11	Hypothetical protein
APO_2085	1, 5b, 9, 11	Site-specific DNA methylase
APO_2140	5b, 9, 11, 1	Hypothetical protein

1, 5b, 9, and 11 but absent in the low-virulence strain JL03 of serovar 3 are summarized in Table 4. As expected, genes (APO_2026 and APO_2004) encoding a toxin activator ApxIC and a structural toxin ApxIA, respectively, were exclusively present in the reference strains of serovars 1, 5b, 9, 10, and 11, all of which secrete the strongly hemolytic and cytotoxic ApxI (19, 38). Compared to *apxIC* and *apxIA*, the genes *apxIIIC* (APO_2098) and *apxIIIA* (APO_2049) involved in expression of the nonhemolytic but strongly cytotoxic ApxIII were present only in the reference strains of serovars 2, 3, 4, and 6. The genetic compositions of *apx* genes in newly sequenced genomes conformed to the *apx* gene patterns of corresponding serovar reference strain previously reported (4, 25). Of 54 distributed genes, 17 were annotated as hypothetical proteins, and their roles in bacterial virulence need to be further investigated.

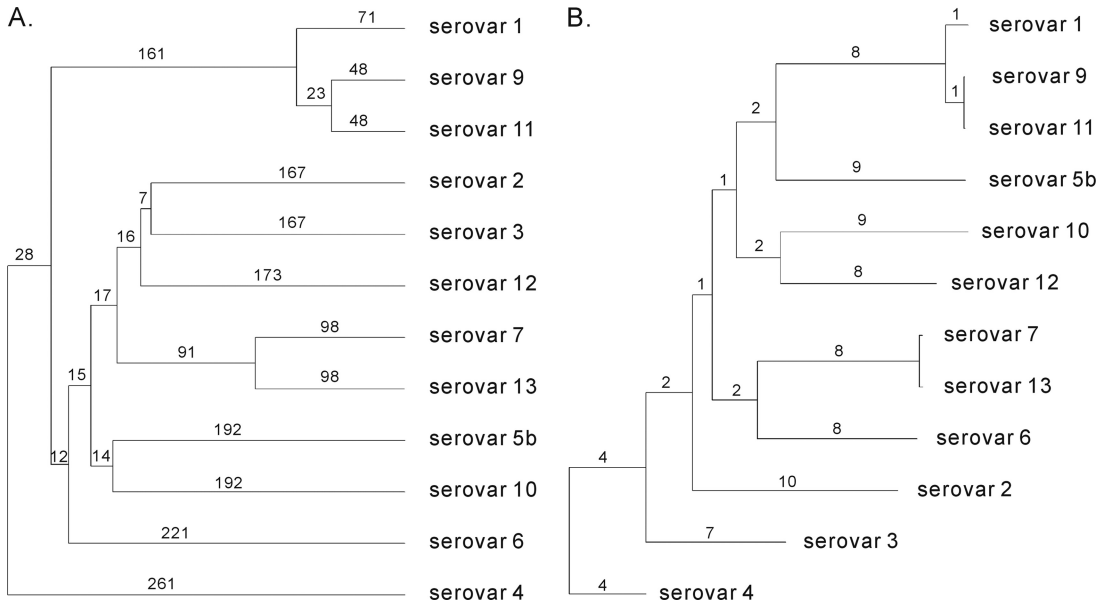

FIG. 3. *A. pleuropneumoniae* whole-genome phylogeny. (A) Dendrogram showing the phylogenetic relationship based on differences in genetic gain or loss of noncore genes among the 12 strains of diverse *A. pleuropneumoniae* serovars. The numbers on the branch represent the number of genic differences that occurs from the previous bifurcation node. (B) Maximum-likelihood tree estimated from a data set of 1,287 concatenated, conserved genic sequences in 12 *A. pleuropneumoniae* genomes. The numbers on the branch stand for the number of nucleotide substitutions per kilobase that occur prior to the next level of separation.

Phylogenetic relationships among serovar reference strains. To understand the phylogenetic relationships among the 12 serovars of *A. pleuropneumoniae*, we used two approaches based on noncore genic differences and the concatenated sequences of 1,287 single-copy core genes among all serovar reference strains, respectively. Figure 3A demonstrated phylogenetic differentiation among *A. pleuropneumoniae* serovars. The differentiation was represented by the total numbers of gene loss and gain between any two genomes of reference strains. Except for *A. pleuropneumoniae* strain L20 of serovar 5b, which originated in the United States, the other highly virulent strains from serovars 1, 9, and 11 belonged to a common clade and differed from each other by fewer than 157 genes, hinting that the strains from the three serovars probably derived from a recent common ancestor. As expected, serovar 9 strain CVJ13261 and serovar 11 strain D13039, which were both isolated from the same geographical location in the Netherlands, had a more close relationship (27). Notably, strains AP76 of serovar 7 and N273 of serovar 13 isolated in Canada and Hungary, respectively, were the second closely related pair and had the genic differentiation bearing 196 genes with gain or loss between them (3). Figure 3B showed a maximum-likelihood tree estimated by the large sequence alignment of 1,287 single-copy core genes. The partial topology types of the genic differences and multi-sequence alignments based trees were similar. Three *A. pleuropneumoniae* reference strains of serovars 1, 9, and 11 were also grouped into an individual clade, as well as serovars 7 and 13.

Whole-genome alignment. A global multiple genome alignment is shown in Fig. S1 in the supplemental material, and it demonstrates that *A. pleuropneumoniae* chromosomes had highly colinear arrangements without largely internal rearrangements among all genomes from 12 serovars. Comparative analysis between two complete genomes of strains JL03 and L20 has shown that the serotypic diversity of *A. pleuropneumoniae* is likely to associate with several serovar-specific genomic regions, which encode the gene clusters involved in the biosynthesis of CPS and LPS O antigen (47).

Genes involved in CPS biosynthesis. The genetic organization of the CPS biosynthesis and export locus is shown in Fig. 4, demonstrating that the genes *cpxDCBA* involved in CPS export are present and highly conserved in all of the serovar reference strains, whereas genes of the *cps* cluster involved in the capsule biosynthesis exhibited high genetic diversity in different sets of serovars. According to the results generated by Blastclust, 33 orthologous genes encoding CPS biosynthetic enzymes were identified in the reference strains of 12 *A. pleuropneumoniae* serovars (Table 5), 24 of which encode strain- or serovar-specific enzymes that are probably responsible for the dissimilarity of the CPS chemical structures. However, previous studies have pointed out that *A. pleuropneumoniae* serovars 1 to 13 can be divided into three groups according to differences of their chemical compositions and the structures of the capsule: type I of CPS consisted of teichoic acid polymers joined by phosphate diester bonds is present in serovars 2, 3, 6, 7, 8, 9, 11, and 13; type II consisted of oligosaccharide polymers joined through phosphate bonds includes serovars 1, 4, and 12; and type III solely containing repeats of oligosaccharide units includes serovars 5a, 5b, and 10 (26, 34).

The genetic organization of the *cps* biosynthetic loci of 12 serovar reference strains provided molecular evidence to further support the grouping of *A. pleuropneumoniae* serovars described above. First, the proteins encoded by *cps2ABC* were present only in serovars 2, 3, 6, 7, 9, 11, and 13 of type I. Cps2A and Cps2B encode teichoic acid glycerol transferase and glyc-

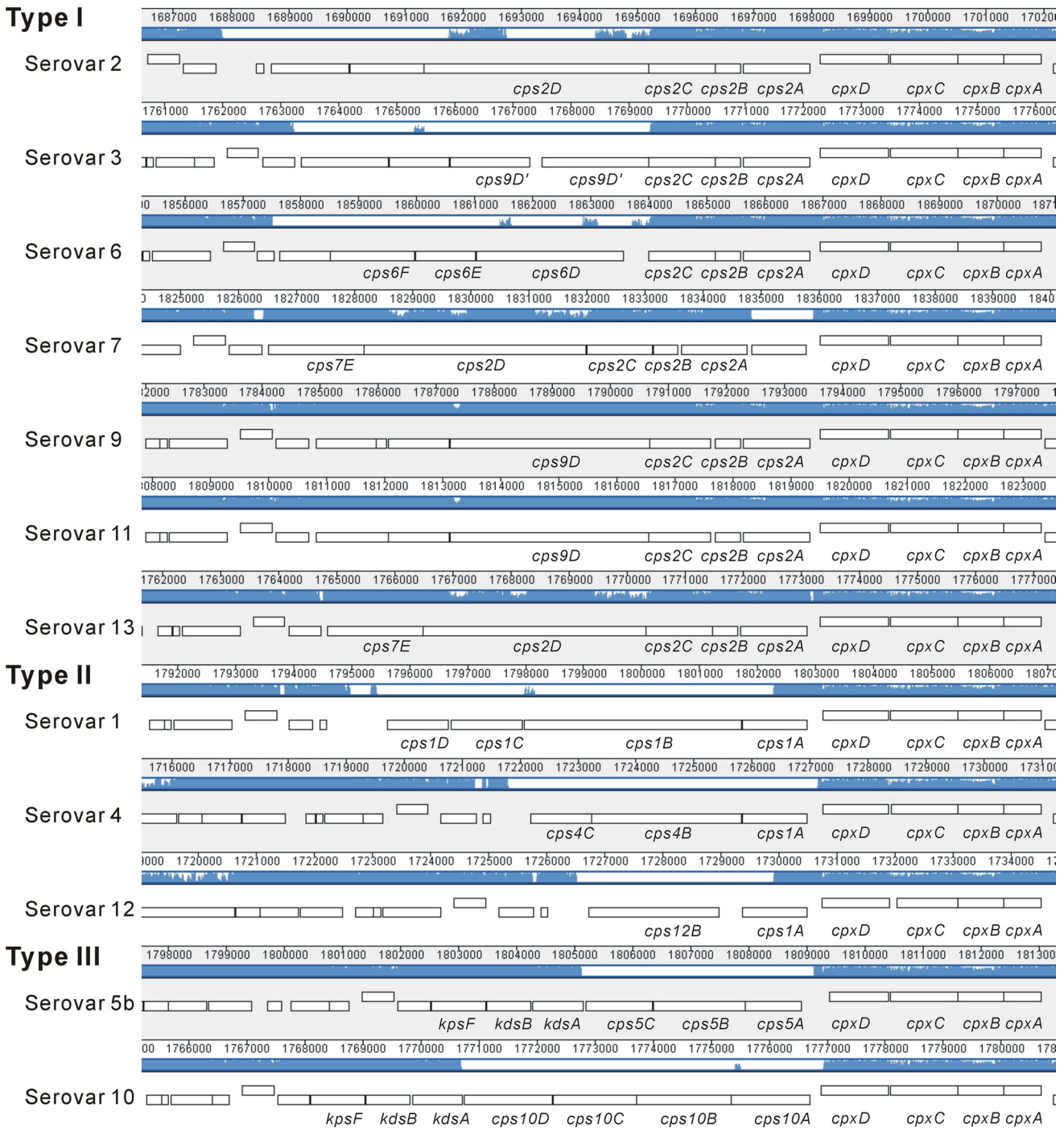

FIG. 4. Schematic comparison of the genetic organizations of the CPS biosynthesis and export gene clusters in the reference strains of 12 *A. pleuropneumoniae* serovars. Three types of CPS have been defined in *A. pleuropneumoniae* as follows: type 1, serovars 2, 3, 6, 7, 9, 11, and 13 are composed of teichoic acid polymers linked by phosphate diester bonds; type II, serovars 1, 4, and 12 are composed of oligosaccharide polymers linked by phosphate bonds; and type III, serovars 5b and 10 are composed of repeating oligosaccharide units (26, 34).

erol-3-phosphate cytidylyltransferase, respectively, which are required for the sequential transfer of glycerol phosphate units (17). Teichoic acid synthases encoded by the genes *cps2D* (APO_2170), *cps6D* (APO_2555), and *cps9D* (APO_2331) were also identified in the serovars of type I. Although these synthases share low sequence identity, they all contain two conserved domains *Glyphos_transf* (PF04464) and *Glycos_transf_1* (PF00534). It is worth mentioning that the teichoic acid biosynthetic enzyme encoded by *cps9D* (APO_2331) was present only in strains CVJ13261 and 56153 of serovars 9 and 11 but became a pseudogene in strain JL03 of serovar 3. Second, Cps1A encodes a capsular polysaccharide phosphotrans-

TABLE 5. Genes encoding enzymes in the capsular polysaccharide biosynthesis locus of *A. pleuropneumoniae*

ClusterID	Name	Putative function	CDS no.
APO_1934	*cps2C*	Cps2C	appser2_15930 appser3_16160 appser6_17380 appser7_16930 appser13_16580 appser9_17130 appser11_17250
APO_1935	*cps2A*	Cps2A	appser2_15950 appser3_16180 appser6_17400 appser7_16950 appser9_17150 appser11_17270 appser13_16600
APO_1947	*cps2B*	Glycerol-3-phosphate cytidylyltransferase	appser2_15940 appser3_16170 appser6_17390 appser7_16940 appser9_17140 appser11_17260 appser13_16590
APO_2170	*cps2D*	Cps2D	appser2_15920 appser7_16920 appser13_16570
APO_2202	*cps1A*	Capsular polysaccharide phosphotransferase	appser12_16150 appser1_17000 appser4_16240
APO_2331	*cps9D*	Teichoic acid biosynthesis protein	appser9_17120 appser11_17240
APO_2349	*cps7E*	CapZD protein	appser7_16910 appser13_16560
APO_2361		Putative uncharacterized protein	appser11_17220 appser9_17090
APO_2380		Putative uncharacterized protein	appser9_17110 appser11_17230
APO_2543	*cps1B*	Glycosyl transferase family protein	appser1_16990
APO_2554	*cps4B*	Glycosyl transferase family protein	appser4_16230
APO_2555	*cps6D*	Cps6D	appser6_17370
APO_2562	*cps12B*	Cps12B	appser12_16140
APO_2582	*cps9D'*	Teichoic acid biosynthesis protein	appser3_16150
APO_2601	*cps10B*	Acetyltransferase (Isoleucine patch superfamily) protein	appser10_16450
APO_2602	*cps5B*	Region 2 capsular polysaccharide biosynthesis protein	appser5b_16510
APO_2607	*cps10D*	Putative uncharacterized protein	appser10_16430
APO_2609		Putative uncharacterized protein	appser3_16120
APO_2611	*cps6F*	Glycosyl transferase group 1	appser6_17350
APO_2615	*cps10C*	Cap29eA protein	appser10_16440
APO_2625	*cps9D'*	Putative uncharacterized protein	appser3_16140
APO_2628	cps10A	GDP-mannose pyrophosphorylase	appser10_16460
APO_2632		Putative uncharacterized protein	appser2_15900
APO_2644		Putative uncharacterized protein	appser2_15910
APO_2651	*cps1C*	Putative uncharacterized protein	appser1_16980
APO_2663	*cps5C*	Region 2 capsular polysaccharide biosynthesis protein	appser5b_16500
APO_2684	*cps1D*	Acetyltransferase (Isoleucine patch superfamily) protein	appser1_16970
APO_2685	*cps4C*	Possible lipopolysaccharide biosynthesis glycosyltransferase	appser4_16220
APO_2686		Putative uncharacterized protein	appser3_16130
APO_2688	*cps6E*	Cps6E (fragment)	appser6_17360
APO_2702	*cps5A*	Putative glycosyltransferase	appser5b_16520
APO_2724		Hypothetical protein	appser6_17340
APO_3193		Hypothetical protein	appser9_17100

ferase and is 73% similar to the LcbA protein of *Neisseria meningitidis*, which may help pathogens to evade the host innate immune system (21, 39). This enzyme shared by serovars 1, 4, and 12 of type II may be involved in the chemical linkage of phosphate in the linear CPS backbone. Third, a KdsA homolog required for the synthesis of monosaccharide dOclA that is a structural component of the *A. pleuropneumoniae* serovars 5 and 10 CPSs (13, 46), was identified just in their reference strains L20 and D13039, respectively. It has been reported that the CPSs of *A. pleuropneumoniae* are the immunodominant antigens bearing greater serological specificities than the O antigens (35). Concordantly, the genetic organizations of the *cps* biosynthetic loci were found to be distinct from each other in all 12 serovar reference strains, potentially leading to differences in their CPS structures.

Genes involved in LPS biosynthesis. LPS, the major adhesin of *A. pleuropneumoniae* involved in adherence to porcine respiratory tract cells and mucus, plays an important role in virulence (5, 36). The structures of LPS O antigen (or O polysaccharide) were reported to be chemically identical or similar in different sets of *A. pleuropneumoniae* serovars: serovars 3, 6, 8, and 15; serovars 1, 9, and 11; and serovars 4, 7, and 13 (31, 35). In the present study, we identified 19 core genes involved in the synthetic pathways of lipid A and core oligosaccharide, the majority of which were dispersed throughout the chromosome and highly conserved among different organisms within the family *Pasteurellaceae* (see Table S6 in the supplemental material). On the other hand, a cluster of genes coding for enzymes that catalyze the biosynthesis of O antigen were identified between the conserved genes *erpA* (APO_1556) and *rpsU* (APO_0575), and these genes were transcribed in the same orientation in all *A. pleuropneumoniae* serovar reference strains. A total of 52 orthologous genes distributed in the 12 O-antigen chains were identified, 20 of which were strain- or serovar-specific genes probably involved in the structural diversity of O polysaccharide (Table 6). The genetic organizations of the LPS O-antigen biosynthesis are shown in Fig. 5. Three sets of serovars (serovars 1, 9, and 11; serovars 7 and 13; and serovars 3 and 6) were observed to have identical gene components; this finding was consistent with the characterization of the O-antigen structures described above. Two different O-antigen biosynthetic pathways were identified for the first time in *A. pleuropneumoniae*. The Wzy/Wzx-dependent pathway of O-antigen biosynthesis was possessed by the serovars 2, 3, 4, 6, 7, and 13 (group I), whereas the ABC-2 transporter-dependent pathway was shared by the serovars 1, 5b, 9, 10, 11, and 12 (group II) (45). The genes *wzm* and *wzt* encoding ABC-2 transporters that were the integral membrane subunit and the ATP-binding subunit, respectively, were identified within the O-

TABLE 6. Genes encoding enzymes in the lipopolysaccharide O-antigen biosynthesis locus of *A. pleuropneumoniae*

Cluster	Name	Putative function	CDS no.
APO_0575	*rmlB*	dTDP-glucose 4,6-dehydratase	appser3_15020 appser5b_15380 appser6_16250 appser10_15330 appser12_15050 appser2_14830 appser4_15060 appser7_15190 appser13_15410 appser1_15770 appser9_15920 appser11_16040
APO_1940	*rmlA*	Glucose-1-phosphate thymidylyltransferase	appser1_15780 appser2_14840 appser4_15070 appser7_15200 appser9_15930 appser11_16050 appser13_15420
APO_1941	*rmlD*	dTDP-4-dehydrorhamnose reductase	appser1_15790 appser2_14850 appser4_15080 appser7_15210 appser9_15940 appser11_16060 appser13_15430
APO_1942	*rmlC*	dTDP-4-dehydrorhamnose 3,5-epimerase	appser1_15800 appser2_14860 appser4_15090 appser7_15220 appser9_15950 appser11_16070 appser13_15440
APO_1968	*wbaP*	Undecaprenyl-phosphate galactose phosphotransferase, WbaP	appser1_15880 appser9_16030 appser11_16150 appser12_15040 appser5b_15370 appser10_15320
APO_1972	*wzz*	Wzz-like protein	appser2_14740 appser3_14920 appser4_15050 appser6_16150 appser7_15180 appser13_15400
APO_2016	*wzt*	ABC transporter ATP binding subunit	appser1_15820 appser9_15970 appser11_16090 appser5b_15300 appser10_15260
APO_2172		Putative glycosyltransferase	appser1_15830 appser9_15980 appser11_16100
APO_2187	*wbaP*	Undecaprenyl-phosphate galactose phosphotransferase, WbaP	appser4_15150 appser7_15280 appser13_15500
APO_2191		Serotype b-specific antigen synthesis gene cluster	appser1_15840 appser11_16110 appser9_15990
APO_2192	*wzx*	Putative O-antigen transporter	appser4_15100 appser7_15230 appser13_15450
APO_2194	*wzy*	Putative membrane protein	appser7_15260 appser13_15480 appser4_15130
APO_2198	*glf1*	Probable UDP-galactopyranose mutase	appser12_15010 appser10_15280 appser5b_15330
APO_2209		Glycosyl transferase, family 2	appser1_15850 appser9_16000 appser11_16120
APO_2212		Glycosyl transferase, family 2	appser5b_15350 appser10_15300 appser12_15020
APO_2213	*rfbN*	O antigen biosynthesis rhamnosyltransferase *rfbN*	appser1_15870 appser9_16020 appser11_16140
APO_2215		Glycosyl transferase family 2	appser4_15140 appser7_15270 appser13_15490
APO_2216	*rfbF*	Putative rhamnosyl transferase	appser1_15860 appser9_16010 appser11_16130
APO_2228	*wzm*	ABC transporter integral membrane subunit	appser1_15810 appser9_15960 appser11_16080
APO_2230	*epsF*	Exopolysaccharide biosynthesis protein	appser5b_15360 appser10_15310 appser12_15030
APO_2352	*wzx*	Flippase Wzx	appser3_14980 appser6_16210
APO_2357	*wzy*	Oligosaccharide repeat unit polymerase	appser3_14970 appser6_16200
APO_2359		Undecaprenyl-phosphate galactose phosphotransferase	appser3_15000 appser6_16230
APO_2373	*glf*	UDP-galactopyranose mutase	appser3_14990 appser6_16220
APO_2376		Polysacchride biosynthesis protein	appser6_16240 appser3_15010
APO_2378		Glycosyltransferase	appser3_14950 appser6_16180
APO_2379		Glycosyltransferase	appser3_14960 appser6_16190
APO_2383		Putative glycosyltransferase	appser7_15240 appser13_15460
APO_2398		Glycosyltransferase	appser7_15250 appser13_15470
APO_2402	*rfaG*	CpsF protein	appser3_14930 appser6_16160
APO_2403	*wzm*	ABC transporter integral membrane subunit	appser5b_15290 appser10_15250
APO_2430		Putative uncharacterized protein	appser5b_15340 appser10_15290
APO_2585		Putative uncharacterized protein	appser10_15270
APO_2587		Putative glycosyltransferase	appser12_15000
APO_2613	*wzx*	Flippase Wzx	appser2_14800
APO_2631		Putative uncharacterized protein	appser12_14990
APO_2634		Undecaprenyl-phosphate galactosephosphotransferase	appser2_14820
APO_2668	*wciN*	WciN beta, glycosyltransferase	appser2_14750
APO_2674		VI polysaccharide biosynthesis protein	appser6_16170
APO_2678		VI polysaccharide biosynthesis protein	appser3_14940
APO_2696		Putative glycosyltransferase	appser4_15110
APO_2699		Glycosyl transferase, group 2 family protein	appser2_14770
APO_2703		Glycosyl transferase family 2	appser12_14980
APO_2711	*wzy*	Eps4N	appser2_14790
APO_2725	*cpsP*	CpsP, glycosyltransferase	appser2_14760
APO_2755	*wzm*	ABC-2 type transporter family protein	appser12_14960
APO_2759		Putative uncharacterized protein	appser5b_15310
APO_2791		Chloramphenicol acetyltransferase	appser2_14810
APO_2796	*wzt*	O-antigen export system ATP-binding protein RfbB	appser12_14970
APO_2810	*cpsM*	Eps7I	appser2_14780
APO_2980		Putative uncharacterized protein	appser5b_15320
APO_3164		Glycosyltransferase	appser4_15120

antigen chains of group II. Wzm proteins in the corresponding *A. pleuropneumoniae* reference strains shared low amino acid sequence identity (58%) and were all predicted to have six transmembrane domains, like the Wzm homologue in *Escherichia coli* (16). We deduced that the diversified gene composition of O-antigen chains should also play a role in the serotypic diversity of *A. pleuropneumoniae*.

In summary, comparative genomic analysis using genome

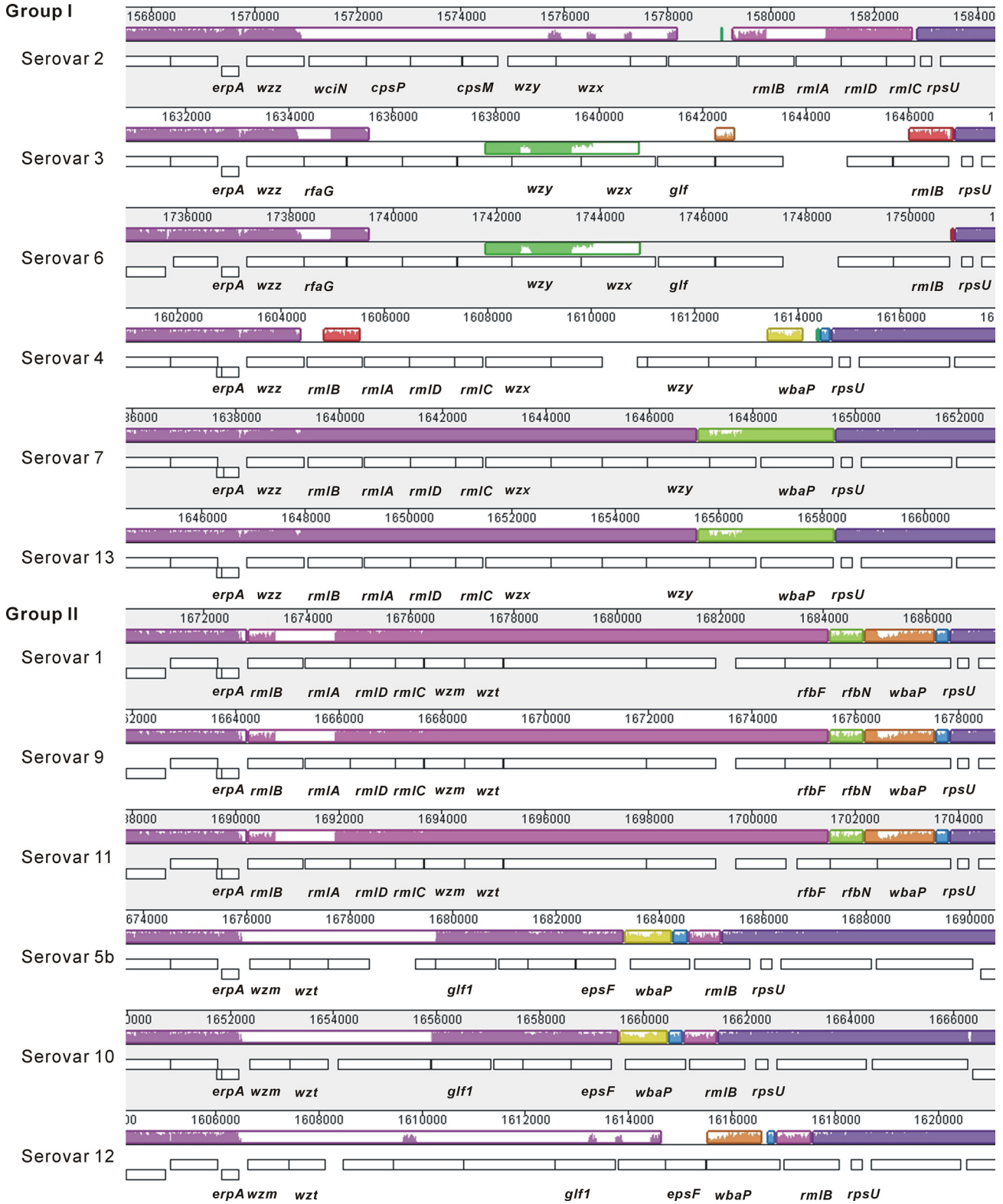

FIG. 5. Schematic comparison of the genetic organizations of the LPS O-antigen biosynthesis gene clusters in the reference strains of 12 *A. pleuropneumoniae* serovars. *A. pleuropneumoniae* serovars can be divided into two groups based on two different mechanisms for the assembly and translocation of O antigen: group I, the Wzy/Wzx-dependent pathway is present in serovars 2, 3, 4, 6, 7, and 13; and group II, the ABC-2 transporter-dependent pathway is present in serovars 1, 5b, 9, 10, 11, and 12.

sequences originated from 12 serovars showed that the pan genome of *A. pleuropneumoniae* consists of 3,303 gene clusters, which contain 1,709 core genome genes, 822 distributed genes, and 772 strain-specific unique genes. The genetic diversity of strain (serovar)-specific genomic islands related to the biogenesis of capsule and lipopolysaccharide O antigen should offer powerful molecular evidence explaining the mechanisms of the serotypic diversity of *A. pleuropneumoniae*. We believe that these findings will provide crucial clues for the development of genomic typing of *A. pleuropneumoniae* and new-style universal vaccines against the severe swine disease caused by this pathogen.

ACKNOWLEDGMENTS

We thank Huajun Zheng and members in the Sequencing Division of the Chinese National Human Genome Center in Shanghai, China, for the genome sequencing and analysis work.

This study was supported by grants from the National Basic Research Program of China (973 Program; 2006CB504402), the National Natural Science Foundation of China (30771599, 30901075), the National Scientific and Technical Supporting Program of China (2006BAD06A01), and the Program for Changjiang Scholars and Innovative Research Team in University (IRT0726).

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **Anderson, C., A. A. Potter, and G.-F. Gerlach.** 1991. Isolation and molecular characterization of spontaneously occurring cytolysin-negative mutants of *Actinobacillus pleuropneumoniae* serotype 7. Infect. Immun. **59:**4110–4116.
4. **Beck, M., J. F. van den Bosch, I. M. Jongenelen, P. L. Loeffen, R. Nielsen, J. Nicolet, and J. Frey.** 1994. RTX toxin genotypes and phenotypes in *Actinobacillus pleuropneumoniae* field strains. J. Clin. Microbiol. **32:**2749–2754.
5. **Bélanger, M., D. Dubreuil, J. Harel, C. Girard, and M. Jacques.** 1990. Role of lipopolysaccharides in adherence of *Actinobacillus pleuropneumoniae* to porcine tracheal rings. Infect. Immun. **58:**3523–3530.
6. **Blackall, P. J., H. L. Klaasen, H. van den Bosch, P. Kuhnert, and J. Frey.** 2002. Proposal of a new serovar of *Actinobacillus pleuropneumoniae*: serovar 15. Vet. Microbiol. **84:**47–52.
7. **Bossé, J. T., H. Janson, B. J. Sheehan, A. J. Beddek, A. N. Rycroft, J. S. Kroll, and P. R. Langford.** 2002. *Actinobacillus pleuropneumoniae*: pathobiology and pathogenesis of infection. Microbes Infect. **4:**225–235.
8. **Champion, M. D., Q. Zeng, E. B. Nix, F. E. Nano, P. Keim, C. D. Kodira, M. Borowsky, S. Young, M. Koehrsen, R. Engels, M. Pearson, C. Howarth, L. Larson, J. White, L. Alvarado, M. Forsman, S. W. Bearden, A. Sjöstedt, R. Titball, S. L. Michell, B. Birren, and J. Galagan.** 2009. Comparative genomic characterization of *Francisella tularensis* strains belonging to low and high virulence subspecies. PLoS Pathog. **5:**e1000459.
9. **Chen, P. E., C. Cook, A. C. Stewart, N. Nagarajan, D. D. Sommer, M. Pop, B. Thomason, M. P. Thomason, S. Lentz, N. Nolan, S. Sozhamannan, A. Sulakvelidze, A. Mateczun, L. Du, M. E. Zwick, and T. D. Read.** 2010. Genomic characterization of the *Yersinia* genus. Genome Biol. **11:**R1.
10. **Christensen, H., and M. Bisgaard.** 2004. Revised definition of *Actinobacillus sensu stricto* isolated from animals. A review with special emphasis on diagnosis. Vet. Microbiol. **99:**13–30.
11. **Chung, J. W., C. Ng-Thow-Hing, L. I. Budman, B. F. Gibbs, J. H. Nash, M. Jacques, and J. W. Coulton.** 2007. Outer membrane proteome of *Actinobacillus pleuropneumoniae*: LC-MS/MS analyses validate in silico predictions. Proteomics **7:**1854–1865.
12. **Darling, A. C., B. Mau, F. R. Blattner, and N. T. Perna.** 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. **14:**1394–1403.
13. **Dubreuil, J. D., M. Jacques, K. R. Mittal, and M. Gottschalk.** 2000. *Actinobacillus pleuropneumoniae* surface polysaccharides: their role in diagnosis and immunogenicity. Anim. Health Res. Rev. **1:**73–93.
14. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792–1797.
15. **Felsenstein, J.** 1989. PHYLIP: Phylogeny Inference Package (version 3.2). Cladistics **5:**164–166.
16. **Feng, L., S. N. Senchenkova, J. Yang, A. S. Shashkov, J. Tao, H. Guo, J. Cheng, Y. Ren, Y. A. Knirel, P. R. Reeves, and L. Wang.** 2004. Synthesis of the heteropolysaccharide O antigen of *Escherichia coli* O52 requires an ABC transporter: structural and genetic evidence. J. Bacteriol. **186:**4510–4519.
17. **Fitzgerald, S. N., and T. J. Foster.** 2000. Molecular analysis of the *tagF* gene, encoding CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase of *Staphylococcus epidermidis* ATCC 14990. J. Bacteriol. **182:**1046–1052.
18. **Foote, S. J., J. T. Bossé, A. B. Bouevitch, P. R. Langford, N. M. Young, and J. H. Nash.** 2008. The complete genome sequence of *Actinobacillus pleuropneumoniae* L20 (serovar 5b). J. Bacteriol. **190:**1495–1496.
19. **Frey, J.** 1995. Virulence in *Actinobacillus pleuropneumoniae* and RTX toxins. Trends Microbiol. **3:**257–261.
20. **Guindon, S., and O. Gascuel.** 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52:**696–704.
21. **Hammerschmidt, S., C. Birkholz, U. Zähringer, B. D. Robertson, J. van Putten, O. Ebeling, and M. Frosch.** 1994. Contribution of genes from the capsule gene complex (*cps*) to lipooligosaccharide biosynthesis and serum resistance in *Neisseria meningitidis*. Mol. Microbiol. **11:**885–896.
22. **Hiller, N. L., B. Janto, J. S. Hogg, R. Boissy, S. Yu, E. Powell, R. Keefe, N. E. Ehrlich, K. Shen, J. Hayes, K. Barbadora, W. Klimke, D. Dernovoy, T. Tatusova, J. Parkhill, S. D. Bentley, J. C. Post, G. D. Ehrlich, and F. Z. Hu.** 2007. Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome. J. Bacteriol. **189:**8186–8195.
23. **Hogg, J. S., F. Z. Hu, B. Janto, R. Boissy, J. Hayes, R. Keefe, J. C. Post, and G. D. Ehrlich.** 2007. Characterization and modeling of the *Haemophilus influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains. Genome Biol. **8:**R103.
24. **Jacques, M.** 2004. Surface polysaccharides and iron-uptake systems of *Actinobacillus pleuropneumoniae*. Can. J. Vet. Res. **68:**81–85.
25. **Jansen, R., J. Briaire, A. B. van Geel, E. M. Kamp, A. L. Gielkens, and M. A. Smits.** 1994. Genetic map of the *Actinobacillus pleuropneumoniae* RTX-toxin (Apx) operons: characterization of the ApxIII operons. Infect. Immun. **62:**4411–4418.
26. **Jessing, S. G., P. Ahrens, T. J. Inzana, and Ø. Angen.** 2008. The genetic organization of the capsule biosynthesis region of *Actinobacillus pleuropneumoniae* serotypes 1, 6, 7, and 12. Vet. Microbiol. **129:**350–359.
27. **Kokotovic, B., and Ø. Angen.** 2007. Genetic diversity of *Actinobacillus pleuropneumoniae* assessed by amplified fragment length polymorphism analysis. J. Clin. Microbiol. **45:**3921–3929.
28. **Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, and S. L. Salzberg.** 2004. Versatile and open software for comparing large genomes. Genome Biol. **5:**R12.
29. **Lagesen, K., P. Hallin, E. A. Rødland, H. H. Staerfeldt, T. Rognes, and D. W. Ussery.** 2007. RNAmmer: consistent and rapid annotation of rRNA genes. Nucleic Acids Res. **35:**3100–3108.
30. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
31. **MacLean, L. L., M. B. Perry, and E. Vinogradov.** 2004. Characterization of the antigenic lipopolysaccharide O chain and the capsular polysaccharide produced by *Actinobacillus pleuropneumoniae* serotype 13. Infect. Immun. **72:**5925–5930.
32. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, and J. M. Rothberg.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
33. **Negrete-Abascal, E., M. E. Reyes, R. M. García, S. Vaca, J. A. Girón, O. García, E. Zenteno, and M. De La Garza.** 2003. Flagella and motility in *Actinobacillus pleuropneumoniae*. J. Bacteriol. **185:**664–668.
34. **Perry, M. B., E. Altman, J.-R. Brisson, L. M. Beynon, and J. C. Richards.** 1990. Structural characteristics of the antigenic capsular polysaccharides and lipopolysaccharides involved in the serological classification of *Actinobacillus pleuropneumoniae* strains. Serodiagn. Immuno. Infect. Dis. **4:**299–308.
35. **Perry, M. B., L. L. MacLean, and E. Vinogradov.** 2005. Structural characterization of the antigenic capsular polysaccharide and lipopolysaccharide O-chain produced by *Actinobacillus pleuropneumoniae* serotype 15. Biochem. Cell Biol. **83:**61–69.
36. **Rioux, S., C. Bégin, J. D. Dubreuil, and M. Jacques.** 1997. Isolation and characterization of LPS mutants of *Actinobacillus pleuropneumoniae* serotype 1. Curr. Microbiol. **35:**139–144.
37. **Salzberg, S. L., A. L. Delcher, S. Kasif, and O. White.** 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Res. **26:**544–548.
38. **Schaller, A., S. P. Djordjevic, G. J. Eamens, W. A. Forbes, R. Kuhn, P. Kuhnert, M. Gottschalk, J. Nicolet, and J. Frey.** 2001. Identification and

detection of *Actinobacillus pleuropneumoniae* by PCR based on the gene *apxIVA*. Vet. Microbiol. **79:**47–62.
39. **Sperisen, P., C. D. Schmid, P. Bucher, and O. Zilian.** 2005. Stealth proteins: *in silico* identification of a novel protein family rendering bacterial pathogens invisible to host immune defense. PLoS Comput. Biol. **1:**e63.
40. **Stewart, A. C., B. Osborne, and T. D. Read.** 2009. DIYA: a bacterial annotation pipeline for any genomics lab. Bioinformatics **25:**962–963.
41. **Stothard, P., and D. S. Wishart.** 2005. Circular genome visualization and exploration using CGView. Bioinformatics **21:**537–539.
42. **Suzek, B. E., H. Huang, P. McGarvey, R. Mazumder, and C. H. Wu.** 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics **23:**1282–1288.
43. **Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:**33–36.
44. **Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy, T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C. R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J. Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K. J. O'Connor, S. Smith, T. R. Utterback, O. White, C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R. Wessels, R. Rappuoli, and C. M. Fraser.** 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome." Proc. Natl. Acad. Sci. U. S. A. **102:**13950–13955.
45. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front. Biosci. **8:**s452–471.
46. **Ward, C. K., M. L. Lawrence, H. P. Veit, and T. J. Inzana.** 1998. Cloning and mutagenesis of a serotype-specific DNA region involved in encapsulation and virulence of *Actinobacillus pleuropneumoniae* serotype 5a: concomitant expression of serotype 5a and 1 capsular polysaccharides in recombinant *A. pleuropneumoniae* serotype 1. Infect. Immun. **66:**3326–3336.
47. **Xu, Z., Y. Zhou, L. Li, R. Zhou, S. Xiao, Y. Wan, S. Zhang, K. Wang, W. Li, L. Li, H. Jin, M. Kang, B. Dalai, T. Li, L. Liu, Y. Cheng, L. Zhang, T. Xu, H. Zheng, S. Pu, B. Wang, W. Gu, X. L. Zhang, G. F. Zhu, S. Wang, G. P. Zhao, and H. Chen.** 2008. Genome biology of *Actinobacillus pleuropneumoniae* JL03, an isolate of serovar 3 prevalent in China. PLoS One **3:**e1450.