Comparative Genomics Analysis of Two Different Virulent Bovine Pasteurella multocida Isolates

Huihui Du,1,2 Rendong Fang,2 Tingting Pan,2 Tian Li,1 Nengzhang Li,2 Qiang He,1 Rui Wu,2 Yuanyi Peng,2 and Zeyang Zhou1

1The State Key Laboratory of Silkworm Genome Biology, Southwest University, Beibei, Chongqing 400716, China
2College of Animal Science and Technology, Southwest University, Beibei, Chongqing 400716, China

Correspondence should be addressed to Yuanyi Peng; ppy2002@sina.com and Zeyang Zhou; zyzhou@swu.edu.cn

Received 28 July 2016; Accepted 2 November 2016

Academic Editor: Sylvia Hagemann

Copyright © 2016 Huihui Du 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.

The Pasteurella multocida capsular type A isolates can cause pneumonia and bovine respiratory disease (BRD). In this study, comparative genomics analysis was carried out to identify the virulence genes in two different virulent P. multocida capsular type A isolates (high virulent PmCQ2 and low virulent PmCQ6). The draft genome sequence of PmCQ2 is 2.32 Mbp and contains 2,002 protein-coding genes, 9 insertion sequence (IS) elements, and 1 prophage region. The draft genome sequence of PmCQ6 is 2.29 Mbp and contains 1,970 protein-coding genes, 2 IS elements, and 3 prophage regions. The genome alignment analysis revealed that the genomes similarity between PmCQ2 and PmCQ6 is 99% with high colinearity. To identify the candidate genes responsible for virulence, the PmCQ2 and PmCQ6 were compared together with that of the published genomes of high virulent Pm36950 and PmHN06 and avirulent Pm3480 and Pm70 (capsular type F). Five genes and two insertion sequences are identified in high virulent strains but not in low virulent or avirulent strains. These results indicated that these genes or insertion sequences might be responsible for the virulence of P. multocida, providing prospective candidates for further studies on the pathogenesis and the host-pathogen interactions of P. multocida.

1. Introduction

Pasteurella multocida (P. multocida) is the etiologic agent of bovine pneumonia and hemorrhagic septicemia in cattle which has been estimated to cause huge economic losses. Five capsule types are routinely identified in P. multocida (A, B, D, E, and F) and each is generally associated with, but not completely restricted to, a specific host [1]. P. multocida has the typical characteristics of an opportunistic pathogen that is affected by various host and pathogen specific determinants and can survive in the oral cavity and upper respiratory tract of wild and domestic animals. In both, animals and humans, P. multocida is often associated with chronic as well as acute infections that can lead to significant morbidity (manifested as pasteurellosis, pneumonia, atrophic rhinitis, hemorrhagic septicemia and/or cellulitis, abscesses, and meningitis) and mortality, particularly in animals [2, 3]. Nevertheless, pasteurellosis is still a relatively uncommon cause of mortality in human, even though deaths due to pasteurellosis have increased in recent years in the United States [4, 5], and pasteurellosis in human is often due to bites or scratches by cats or dogs [6, 7].

The first complete genome sequence of P. multocida was Pm70, isolated from avian species in 2001 [8]. Since then, the complete or incomplete genomes of 57 P. multocida isolates have been sequenced, including at least ten complete genomes from the species in the NCBI database. All of the currently available P. multocida genomes are between 1.43 Mbp and 2.44 Mbp in length and comprise a single circular genome with a G+C content between 36.9% and 41%. The available data were used to identify a number of important similarities and differences between these strains and determine their virulence [9].

Several species-specific putative virulence factors, including the capsular and virulence-associated genes, have been proposed to play a key role in the interactions with the host [10, 11]. P. multocida possesses a number of virulence factors which include polysaccharide capsule, endotoxins or

Hindawi Publishing Corporation
International Journal of Genomics
Volume 2016, Article ID 4512493, 14 pages
http://dx.doi.org/10.1155/2016/4512493
lipopolysaccharide (LPS), outer membrane proteins (OMPs), fimbriae, exotoxins, multocidins or siderophores, extracellular enzymes, plasmids, and the virulence-associated genes (tbpA, phaA, toxA, hgbB, hgbA, Fur, tonB, exbB, hgbB, nanH, nanB, sodA, sodC, ompA, ompH, oma87, PlpB, fimA, hsf-1, hsf-2, tadD, and pifa) [1, 12, 13]. It is speculated that the virulence factors expressed by P. multocida are likely to play key roles in pathogenesis. Comparative genomics provides an effective source for better understanding the virulence of different isolated strains. In this study, genome sequencing and comparative genomics analysis were carried out to investigate the underlying virulence factors of the high virulent and low virulent bovine P. multocida capsular type A strains, PmCQ2 and PmCQ6, respectively.

2. Materials and Methods

2.1. Bacterial Strains and Culture Conditions. Two P. multocida isolates (PmCQ2 and PmCQ6) have been previously isolated from the fatal pneumonia lungs of feedlot calves at Gaociazheng farms in Fengdu (Chongqing, China, longitude/latitude 107.70/29.89) from 2011 to 2012. Based on morphological characteristics, biochemical properties, and 16S rRNA gene sequence analysis, the bacteria were identified as P. multocida. Further analysis with PCR amplification of P. multocida species-specific gene Knt-1 and serotype-specific genes hyaD-hyaC, bcbD, dcbF, ecbj, and fcbD [16] indicated that the isolates were P. multocida capsular type A, named as PmCQ2 and PmCQ6, and the virulence of the two strains determined by LD_{50} in Kunming mice showed that PmCQ2 is a high virulent strain and PmCQ6 is a low virulent strain with 2.2 × 10^{7} CFU and 1.14 × 10^{6} CFU, respectively [17]. Isolated strains were maintained at −80°C in Martin Broth (MB) plus 10% glycerol. PmCQ2 and PmCQ6 were inoculated in 5 mL MB at 37°C overnight with shaking. The concentration was determined by viable cell counting on Martin agar plates at 37°C for 24 h.

2.2. Genome Sequencing and Annotation. Genomic DNAs of the two strains were isolated using the Qiagen DNA extraction kits. Genome sequencing was performed using an Illumina MiSeq platform. A total of 6,394,560 and 525,022,200 paired-end 100 bp reads of each genome were assembled into 7 and 32 contigs for strains PmCQ2 and PmCQ6, respectively. The sequences of PmCQ2 and PmCQ6 were assembled by SOAPdenovo [18]. Assemblies were submitted to NCBI for annotation using BLASTP (e-value < 1e−5) and comprehensive online resource for curating information. Among the potential virulent genes of P. multocida and gene annotation information, putative virulence genes for each strain were identified by PhiSpy [25]. Genomic islands (GIs) are clusters of genes in prokaryotic genomes of probable horizontal origin. GIs of P. multocida were predicted with IslandPick [26]. Insertion sequences (ISs) of P. multocida were identified by searching sequences against the IS Database (Table S1) that collect all ISs of bacteria and archaea. ISFinder [27] was implemented to launch BLAST with the e-value 1e−10 to search the database. Membrane proteins generally include transmembrane domains and were predicted by TMHMM Server 2.0 [28]. Signal peptide, transmembrane domain, GPI-anchor, and general subcellular localization were predicted with SignalP v3.0 [29], TMHMM Server 2.0, GPI-SOM [30], and PSORTb [31] to screen potential secretory proteins that contain signal peptide and no membrane localization signals. The virulence factor database (VFDB) is an integrated and comprehensive online resource for curating information about virulence factors of bacterial pathogens (Table S1). Based on homologous analysis, some virulent factors (ISs, GIs, VF, secretory proteins, and membrane proteins) were obtained in the sequenced strains. In combination with the potential virulent genes of P. multocida and gene annotation information, putative virulence genes for each strain were presented.

2.4. BLAST Score Ratio Analysis. Genes that were unique to each strain were also identified using BLASTN. The BLAST score ratio (BSR) method was used to compare peptide identities within three genomes (PmCQ2, PmCQ6, and Pm36950) using a measure of similarity based on the ratio of BLAST scores. The output of the BSR analysis enables global visualization of the degree of proteome similarity among genomes and enables the genomic synten (conserved gene order) between each genome pair to be assessed [23]. Pm36950 was used as a reference genome sequence. The BSR was calculated by dividing the query score by the reference score for each reference peptide. Following calculation of the BSRs, the four quadrants were derived from a BSR threshold value of 0.4, which was empirically determined to represent approximately 30% amino acid identity over approximately 30% of the peptide length and is a commonly used threshold for peptide similarity [24]. The four quadrants were determined for each of the query genomes and colored accordingly: yellow, unique to the reference, PmCQ2 < 0.4, and PmCQ6 < 0.4; red, common to all three, PmCQ2 ≥ 0.4, and PmCQ6 ≥ 0.4; Green, common between PmCQ2 and Pm36950, but absent in PmCQ6, PmCQ2 < 0.4, and PmCQ6 ≥ 0.4; Blue, common between PmCQ6 and Pm36950, but absent in PmCQ2, PmCQ2 ≥ 0.4, and PmCQ6 < 0.4.

2.5. Virulence Factors. Prophage-associated gene clusters were identified by PhiSpy [25]. Genomic islands (GIs) are clusters of genes in prokaryotic genomes of probable horizontal origin. GIs of P. multocida were predicted with IslandPick [26]. Insertion sequences (ISs) of P. multocida were identified by searching sequences against the IS Database (Table S1) that collect all ISs of bacteria and archaea. ISFinder [27] was implemented to launch BLAST with the e-value 1e−10 to search the database. Membrane proteins generally include transmembrane domains and were predicted by TMHMM Server 2.0 [28]. Signal peptide, transmembrane domain, GPI-anchor, and general subcellular localization were predicted with SignalP v3.0 [29], TMHMM Server 2.0, GPI-SOM [30], and PSORTb [31] to screen potential secretory proteins that contain signal peptide and no membrane localization signals. The virulence factor database (VFDB) is an integrated and comprehensive online resource for curating information about virulence factors of bacterial pathogens (Table S1). Based on homologous analysis, some virulent factors (ISs, GIs, VF, secretory proteins, and membrane proteins) were obtained in the sequenced strains. In combination with the potential virulent genes of P. multocida and gene annotation information, putative virulence genes for each strain were presented.
3. Results

3.1. Overview of the *P. multocida* PmCQ2 and PmCQ6 Genomes. The genome sequences of both PmCQ2 and PmCQ6 strains were successively sequenced by Illumina MiSeq platform. Using Pm36950 as a reference strain, PmCQ2 genome is 2.32 Mbp in size with 39.12% G+C content, containing 2,000 predicted coding regions, 4 rRNAs operons, and 49 tRNAs. PmCQ6 genome is 2.29 Mbp in size with 40.09% G+C content, containing 1,969 predicted coding regions, 1 rRNA operon, and 43 tRNAs. The single circular genome maps of the two *P. multocida* genomes were shown in Figure 1. There are no obvious species-specific features of the coding density, and the G+C content is highly conserved. Compared with some other *P. multocida* strains carrying multiple plasmids that may either be cryptic or carry antibiotic resistance genes, both PmCQ2 and PmCQ6 genomes do not contain any plasmids. Taken together, there are only slightly differences in genome sizes, predicted gene numbers, and G+C contents between PmCQ2 and PmCQ6.

3.2. COG Classification. The predicted protein sequences were annotated to various COG categories. Some differences in protein numbers among COG categories of PmCQ2 and PmCQ6 were identified (including those listed as protein numbers for PmCQ2 and PmCQ6, resp.): “energy production and conversion” (109 and 111), “amino acid transport and metabolism” (158 and 156), “nucleotide transport and metabolism” (60 and 57), “carbohydrate transport and metabolism” (165 and 166), “coenzyme transport and metabolism” (89 and 86), “translation, ribosomal structure, and biogenesis” (132 and 129), “transcription” (81 and 79), “replication, recombination, and repair” (111 and 100), “cell wall/membrane/envelope biogenesis” (145 and 158), “inorganic ion transport and metabolism” (121 and 120), “general function prediction only” (183 and 181), “function unknown” (158 and 157), “signal transduction mechanisms” (42 and 44), and “intracellular trafficking, secretion, and vesicular transport” (38 and 40) (Figure 2).

3.3. Global Alignment Analysis. The colinearity analysis at the nucleotide level provides information on sequence insertion or deletion [32]. By aligning the genome at the nucleotide level, there was no significant differences among the large segments between high virulent PmCQ2 and low virulent PmCQ6, and the two strains revealed high colinearity with Pm36950 (Figures 3(a)–3(c)). Direct comparison of the complete nucleotide sequences using BLAST revealed the similarity between PmCQ2 and Pm36950, PmCQ6 and Pm36950, and PmCQ2 and PmCQ6 is 90%, 90%, and 99%, respectively. PmCQ2 and PmCQ6 showed higher homology as indicated by matched CDS (Figure 3(d)). By BSR analysis, the protein sequences shared a high degree of synteny among PmCQ2, PmCQ6, and Pm36950, using Pm36950 as a reference strain (Figure 4). However, some unique proteins were identified, PmCQ2 and PmCQ6 (BLAST score ratio is less than 0.4). There are 32 unique proteins in PmCQ2 genome (including transposase IS200, elongation factor Tu-A-1/2, SrfC, IsrR, TolA, and peptidase B) and only two unique proteins found in PmCQ6 genome (*Pasteurella* filamentous hemagglutinin protein and mercuric transport protein MerT). The relative chromosomal locations of the unique proteins (red thick marks) of PmCQ2 and PmCQ6 were shown in Figure 5.

Using a Venn diagram of three bovine *P. multocida* strains, the majority of homologous gene groups and unique gene groups were identified. The unique gene groups were significantly different among three strains, containing 37, 29, and 245 gene groups in PmCQ2, PmCQ6, and Pm36950, respectively (Figure 5).

3.4. Virulence Factors. The pathogenicity of *P. multocida* is associated with different virulence factors. The major virulence factors identified in *P. multocida* are capsule proteins, lipopolysaccharides, membrane proteins, and secreted proteins. Here, together with genome sequences of PmCQ2 and PmCQ6, published genome sequences of high virulent strains (Pm36950 and PmHN06) and avirulent strains (Pm3480 and Pm70) from NCBI were selected for comparative genomics analysis (Table 1). Comparing the PmCQ2 and PmCQ6 genomes with the complete genome sequences of Pm36950 (G+CA_000234743.1), PmHN06 (G+CA_000255915.1), Pm3480 (G+CA_000259545), and Pm70 (G+CA_000006825.1) using BLAST, a number of virulence-associated genes were identified that were absent or present in all of the comparison strains (Table 2).

A number of genes or gene clusters have been implicated as important for virulence of *P. multocida* [9]. Some of these genes encoding putative virulence factors are universally present in all six *P. multocida* genomes, including genes encoding prophage, genomic islands, insertion sequences, virulence factor, secretory proteins, and outer membrane proteins.

By comparing the high virulent strains (PmCQ2, Pm36950, and PmHN06) with low virulent strain (PmCQ6) and avirulent strains (Pm3480 and Pm70), unique genes which were correlated with virulence and only presented in high virulent strains were identified. For instance,
Figure 1: Circular genome maps of PmCQ2 (a) and PmCQ6 (b) from inside to outside indicate the following: Circle 1, G+C skew; yellow green, G+C skew > 0; purple, G+C skew < 0; Circle 2, G+C content (median represents the above average content, the outer circle is greater than the average content, and the inner circle is less than the average content); Circle 3, rRNA genes distribution represented in scaffold sequence; Circle 4, tRNA gene distribution represented in scaffold sequence; Circle 5, open reading frame (ORF) distribution, plus strand; and Circle 6, multiple scaffold exhibition.
insertion sequence (transposase IS200) only existed in three high virulent strains, suggesting that IS200 elements are not conserved sequences and do not spread among all *P. multocida* strains. IS605 and secreted protein PmCQ2-2g0088 (ModB) and nonspecific tight adherence protein D PmCQ2-3g0367 were presented only in PmCQ2 genome (Table 2).

In addition, genomic islands (GIs) are clusters of genes in prokaryotic genomes and are probable horizontal origin. GIs of Pm70, Pm3480, Pm36950, and PmHN06 were predicted with IslandPick. Homology analysis of these GIs with the draft genomes of PmCQ2 and PmCQ6 was carried out using ORTHOMCL1.4 (BLAST *p* value 1e−5, percent identity cutoff 60%, and percent match cutoff 60%). The result showed that transcriptional regulator *PmCQ2_7g0006* and hypothetical proteins *PmCQ2_5g0013* and *PmCQ2_5g0025* are present in high virulent strains (PmCQ2 and PmHN06) but absent in low virulent strain PmCQ6 and the avirulent strains (Pm70 and Pm3480).

Taken together, comparative genomics analysis supplies essential information for understanding the virulence of different capsular type (A, D, and F) and different host origin (bovine, avian, and swine) strains. Five unique genes and two insertion sequences were identified only in high virulent strains, providing candidate virulence factors for further studies on the pathogenesis of different *P. multocida* strains (Table 3).

### 4. Discussion

Moreover, comparative genomic analysis allows the identification of core genes and/or disease-specific factors. The first complete *P. multocida* genome was sequenced from strain Pm70 in 2001, from which 104 putative virulence-associated genes were identified [8]; this facilitated new approaches for studying the pathogenesis of *P. multocida*. Until now, the complete and incomplete genomes of 57 *P. multocida* have been sequenced in NCBI database. In this study, two bovine *P. multocida* capsular type A genomes (high virulent PmCQ2 and low virulent PmCQ6) were sequenced. Comparative genomics analysis was performed among PmCQ2, PmCQ6, and four other *P. multocida* genomes (Pm36950, PmHN06, Pm3480, and Pm70) from NCBI. Some virulence genes were identified among different virulent strains; five genes and two insertion sequences were only identified in high virulent strains, which might be responsible for the virulence differences among high virulent, low virulent, and avirulent strains.

The genome sequences of high virulent PmCQ2 and the low virulent PmCQ6 have high similarity, but the virulence of two strains is significantly different. It could be speculated that the unique genes may play a key role in virulence. Compared with PmCQ6, the five genes and two insertion sequences are predicted virulence-associated genes in PmCQ2 and other high virulent strains. Further studies to construct mutant strains targeting these genes would be of great importance to prove their contributions to virulence. Besides, PmCQ2 has more than 30 other unique genes that might also orchestrate the virulence differences of PmCQ2 and PmCQ6. These genes include recombinase, phage-related genes, phage N-6-adenine-methyltransferase, phage terminase, and prophage integrase.

Based on homology analysis, prophage-associated genes, GIs, ISs, secretory proteins, and membrane proteins were
screened for different virulence-associated genes among different virulent strains. Insertion sequences usually only carry genes of transposon sequences for the transposition in bacteria and can also induce a variety of genomic rearrangements; they also play an important role in bacterial host specificity and virulence [33, 34]. Transposase IS200 was found in three high virulent isolated strains encoding the 7 genes (PmCQ2.1g0197, PmCQ2.1g0267, PmCQ2.1g0316, PmCQ2.1g0378, PmCQ2.2g0113, PmCQ2.4g0323, and PmCQ2.4g0359), but IS200 was not present in the low virulent strains (PmCQ6, Pm70) or the avirulent strain (Pm3480). The IS200 elements may adapt to different hosts in closely related genera but stochastic loss can appear in some low virulent or avirulent strains. According to previous reports, IS200-related transposons may have already existed in remote stages of bacterial evolution, such as Salmonellae, and IS200-based methods have been described for the identification of certain Salmonella serovars [35]. The function and host range of transposase IS200 in P. multocida still need to be further studied.

PmCQ2.2g0088 has been suggested to encode a subfamily of ATP-binding cassette (ABC) transporters that have a possible role in remodeling the cell envelope and entry of the pathogen into nonphagocytic cells [36]. Bacterial ABC transporters are essential for the uptake of nutrients, including rare elements such as molybdenum [37]. ABC transporters are integral membrane proteins that actively transport molecules across cell membranes [38], and these three proteins are coded by modA, modB, and modC genes, respectively. The ModA, ModB, and ModC proteins are very similar in various
Table 2: The difference of virulence-associated genes in some or all comparison genomes using BLAST.

| Database | PmCQ2 | PmCQ6 | Pm36950 | PmHN06 | Pm3480 | Pm70 | Annotation |
|----------|-------|-------|---------|--------|--------|------|------------|
| PmCQ2   | Pmu_00900 | PMCN06_0843 |
| PmCQ2   | Pmu_13570 |
| PmCQ2   | Pmu_13960 |
| PmCQ2   | Pmu_16290 |
| PmCQ2   | Pmu_17660 |
| PmCQ2   | Pmu_18340 |
| PmCQ2   | Pmu_20570 |
| PmCQ2   | Pmu_21310 |
| PmCQ2   | Pmu_24490 |
| PmCQ2   | Pmu_25590 |
| PmCQ2   | Pmu_27000 |
| PmCQ2   | Pmu_27001 |
| PmCQ2   | Pmu_27002 |

Phages-associated genes

| Database | PmCQ2 | PmCQ6 | Pm36950 | PmHN06 | Pm3480 | Pm70 | Annotation |
|----------|-------|-------|---------|--------|--------|------|------------|
| PmCQ2   | NT08PM_0122 | PM01098 |
| PmCQ2   | NT08PM_0197 | PM0109 |
| PmCQ2   | NT08PM_0930 | PM0385 |
| PmCQ2   | NT08PM_1618 | PM1778 |
| PmCQ2   | NT08PM_0285 |
| PmCQ2   | NT08PM_0286 |
| PmCQ2   | NT08PM_0288 |
| PmCQ2   | NT08PM_0294 |
| PmCQ2   | NT08PM_0295 |
| PmCQ2   | NT08PM_0298 |
| PmCQ2   | PMCN06_2102 |
| PmCQ2   | PMCN06_2103 |
| PmCQ2   | PMCN06_2105 |
| PmCQ2   | PMCN06_2102 |
| PmCQ2   | PMCN06_2103 |
| PmCQ2   | PMCN06_2105 |
| PmCQ2   | Tail assembly protein I |
| PmCQ2   | Host specificity protein, putative, partial |
| PmCQ2   | Tail protein |
| PmCQ2   | Tail assembly protein |
| Database | PmCQ2 | PmCQ6 | Pm36950 | PmHN06 | Pm3480 | Pm70 | Annotation |
|----------|-------|-------|---------|--------|--------|------|------------|
| Genomic islands |       |       |         |        |        |      |            |
| PmCQ2, 1g0097 |       |       |         |        |        |      | phage terminase, large subunit, pbsx family, putative |
| PmCQ2, 1g0267 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 1g0316 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 1g0378 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0113 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 4g0323 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 4g0359 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0231 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 6g0025 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 5g0021 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 5g0039 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0228 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0229 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0230 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0233 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0234 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 2g0235 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 7g0006 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 5g0013 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 5g0025 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 17g0009 |       |       |         |        |        |      | transcriptional regulator |
| PmCQ2, 17g0011 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 5g0001 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, 17g0004 |       |       |         |        |        |      | hypothetical protein |
| PmCQ2, C4143g0001 |       |       |         |        |        |      | hypothetical protein |

**Table 2: Continued.**
| Database   | PmCQ2       | PmCQ6       | Pm36950 | PmHN06 | Pm3480 | Pm70 | Annotation                                      |
|------------|-------------|-------------|---------|--------|--------|------|------------------------------------------------|
|            | PmCQ6_17g0007 | PMCN06,2072 |         |        |        |      | Hypothetical protein                            |
|            | PmCQ6_17g0008 | PMCN06,2073 |         |        |        |      | Hypothetical protein                            |
|            | PmCQ6_17g0010 | PMCN06,2075 |         |        |        |      | Hypothetical protein                            |
|            | PmCQ6_9g0009  | PMCN06,2102 |         |        |        |      | Hypothetical protein                            |
|            | PmCQ6_9g0008  | PMCN06,2103 |         |        |        |      | Hypothetical protein                            |
|            | PmCQ6_9g0006  | PMCN06,2105 |         |        |        |      | Hypothetical protein                            |
| ISs        | PmCQ2_1g0197  | PMu_00900   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_1g0267  | PMu_01900   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_1g0316  | PMu_13570   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_1g0354  | PMu_13960   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_2g0013  | PMu_16290   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_4g0323  | PMu_17660   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_4g0359  | PMu_19340   | PMCN06,0843 |        |        |      | IS200 transposase                               |
|            | PmCQ2_12g0001 | PMu_19340   | PMCN06,0843 |        |        |      | Putative transposase for insertion sequence IS1162 |
| VFDB       | PmCQ2_4g0316  | PMu_13500   | PMCN06,1329 |        | NT08PM,1414 | PM1994 | UDP-3-O-[3-hydroxymyristoyl]; UDP-3-O-acylglucosamine N-acyltransferase |
|            | PmCQ2_4g0631  | PMu_21120   | PMCN06,1329 |        | NT08PM,2001 | PM1666 | Noncanonical purine NTP pyrophosphatase, RdgB/ham1 family |
|            | PmCQ2_4g0241  | PMu_12710   | PMCN06,1257 |        | NT08PM,1342 | PM0051 | Iron-binding protein FbpA                        |
|            | PmCQ2_3g0252  | PMu_08050   | PMCN06,0796 |        | NT08PM,0537 | PM0734 | Periplasmic serine protease do/hhoA-like protein |
|            | PmCQ2_2g0162  | PMu_01460   | PMCN06,0215 |        | NT08PM,0212 | PM0105 | Hypothetical protein PM0105                      |
|            | PmCQ2_3g0101  | PMu_15140   | PMCN06,1551 |        | NT08PM,1574 | PM0105 | Putative virulence effector, SrfC                 |
|            | PmCQ2_1g0553  | PMu_15880   | PMCN06,1607 |        | NT08PM,1650 | PM0105 | Elongation factor Tu, partial                    |
|            | PmCQ2_1g0157  | PMu_20230   | PMCN06,2025 |        | NT08PM,2100 | PM1746 | Nonspecific tight adherence protein D, partial    |
| Database | PmCQ2 | PmCQ6 | Pm36950 | PmHN06 | Pm3480 | Pm70 | Annotation |
|----------|-------|-------|---------|--------|--------|------|------------|
| PmCQ2   |       |       |         |        |        |      |            |
| PmCQ2.1g0033 |       |       |         |        |        |      | Hypothetical protein, uncharacterized lipoprotein |
| PmCQ2.3g0252 |       |       |         |        |        |      | Periplasmic serine protease do/hhoA-like protein |
| PmCQ2.2g0088 |       |       |         |        |        |      | ModB, partial |
| PmCQ2.1g0666 |       |       |         |        |        |      | Penicillin-binding protein IA |
| PmCQ2.4g0293 |       |       |         |        |        |      | Bicyclomycin resistance protein-1 |
| Membrane proteins |       |       |         |        |        |      | Mercuric transport protein MerT |
| PmCQ6.1g0027 |       |       |         |        |        |      | Glycoside transport protein |
| PmCQ6.C4143g0001 |       |       |         |        |        |      | Hypothetical protein |
| PmCQ6.17g0010 |       |       |         |        |        |      | Tail assembly protein I |
| PmCQ6.9g0003 |       |       |         |        |        |      | C4-dicarboxylate ABC transporter permease |
| PmCQ6.14g0071 |       |       |         |        |        |      | Hypothetical protein, partial |

Table 2: Continued.
Table 3: The distribution of predicted virulence factors among different *P. multocida* strains.

| Virulence   | Strains   | IS 200 | PmCQ2_5g0025 | PmCQ2_7g0006 | PmCQ2_5g0013 | IS 605 | PmCQ2_2g0088 | PmCQ2_3g0367 | Capsular type | Host |
|------------|-----------|--------|---------------|---------------|---------------|--------|---------------|---------------|---------------|------|
| Highly virulent | PmCQ22 | +      | +             | +             | +             | +      | +             | +             | +             | A    |
| Highly virulent | Pm36950 | +      | +             | +             | −             | −      | −             | −             | −             | A    |
| Highly virulent | PmHN06 | +      | −             | −             | −             | −      | −             | −             | −             | D    |
| Lowly virulent | PmCQ6  | −      | −             | −             | −             | −      | −             | −             | −             | A    |
| Avirulent   | Pm3480   | −      | −             | −             | −             | −      | −             | −             | −             | A    |
| Avirulent   | Pm70     | −      | −             | −             | −             | −      | −             | −             | −             | F    |

+ stands for the gene present in certain strain; − stands for the gene absent in certain strain.
organisms (Escherichia coli, Haemophilus influenzae, Azotobacter vinelandii, and Rhodobacter capsulatus) [39]. In this study, PmCQ2_g0088 (ModB) is only present in virulent PmCQ2 but absent in PmCQ6. PmCQ2_g0088 contains a signal peptide and a SBP_bac_ID structural domain. The SBP-box gene family is specific to plants and encodes a class of zinc-finger-containing transcription factors with a broad range of functions [40]. However, the function of the ModB protein family has not been clearly established; PmCQ2_g0088 might affect the virulence of strain and needs to be further studied as a candidate virulence factor.

The present study revealed that P. multocida strains carry different virulence genes which may indicate variation in the pathogenicity. It could be speculated that the specific genes of different strains play the most important role for the difference of pathogenicity. By extensive genomics and proteomics analysis, the intensive study on virulence genes provides deeper understanding of host specificity and pathogenesis and also provides insights into the host-microbe interactions and the immunologic mechanism, contributing to the development of novel vaccines.

**Additional Points**

**Availability of Data and Materials.** The genome sequences of PmCQ2 and PmCQ6 have been deposited at GenBank under the accession numbers of LIUN00000000 and LIUO00000000, respectively.

**Competing Interests**

The authors declare no conflict of interests.

**Authors’ Contributions**

Huihui Du and Rendong Fang contributed equally to the present study. Yuanyi Peng and Zeyang Zhou conceived and designed the experiments. Huihui Du and Tingting Pan performed the experiments. Huihui Du and Rendong Fang wrote the paper. Nengzhang Li, Qiang He, Tian Li, and Rui Wu contributed to data analysis. Rendong Fang, Yuanyi Peng, and Zeyang Zhou supervised the project.

**Acknowledgments**

This work was supported by the earmarked fund for China Agriculture Research System (Beef/Yak Cattle, CARS-38), Chongqing Science & Technology Commission (cstc2015shmszx8022, cstc2015jcyjBX0108, cstc2015shmszx80010, and cstc2015jcyjA80021), the National Natural Science Foundation of China (31400762), and the Fundamental Research Funds for the Central Universities (XDJK2015B002).

**References**

[1] C. Ewers, A. Lübke-Becker, A. Bethe, S. Kießling, M. Filter, and L. H. Wieler, “Virulence genotype of Pasteurella multocida strains isolated from different hosts with various disease status,” Veterinary Microbiology, vol. 114, no. 3-4, pp. 304–317, 2006.

[2] W. A. Hill and J. P. Brown, “Zoonoses of rabbits and rodents,” Veterinary Clinics of North America—Exotic Animal Practice, vol. 14, no. 3, pp. 519–531, 2011.

[3] M. J. Souza, “Bacterial and parasitic zoonoses of exotic pets,” Veterinary Clinics of North America—Exotic Animal Practice, vol. 12, no. 3, pp. 401–415, 2009.
[4] A. C. Adler, C. Cester, and R. B. Brown, "Septic shock from Pasteurella multocida following a cat bite: case report and review of literature," *Connecticut Medicine*, vol. 75, no. 10, pp. 603–605, 2011.

[5] J. Heydemann, J. S. Heydemann, and S. Antony, "Acute infection of a total knee arthroplasty caused by Pasteurella multocida: a case report and a comprehensive review of the literature in the last 10 years," *International Journal of Infectious Diseases*, vol. 14, no. 3, pp. e242–e245, 2010.

[6] C. Dendle and D. Looke, "Review article: animal bites: an update for management with a focus on infections," *Emergency Medicine Australasia*, vol. 20, no. 6, pp. 458–467, 2008.

[7] B. A. Wilson and M. Ho, "Pasteurella multocida: from zoonosis to cellular microbiology," *Clinical Microbiology Reviews*, vol. 26, no. 3, pp. 631–655, 2013.

[8] B. J. May, Q. Zhang, L. L. Li, M. L. Paustian, T. S. Whittam, and V. Kapur, "Complete genomic sequence of Pasteurella multocida, Pm70," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 6, pp. 3460–3465, 2001.

[9] J. D. Boyce, T. Seemann, B. Adler, and M. Harper, "Pathogenomics of Pasteurella multocida," *Current Topics in Microbiology and Immunology*, vol. 361, pp. 23–38, 2012.

[10] F. Dziva, A. P. Muhairwa, M. Bisgaard, and H. Christensen, "Diagnostic and typing options for investigating diseases associated with Pasteurella multocida," *Veterinary Microbiology*, vol. 128, no. 1-2, pp. 1–22, 2008.

[11] S. Verma, M. Sharma, S. Katooch et al., "Profiling of virulence associated genes of Pasteurella multocida isolated from cattle," *Veterinary Research Communications*, vol. 37, no. 1, pp. 83–89, 2013.

[12] M. Harper, J. D. Boyce, and B. Adler, "Pasteurella multocida pathogenesis: 125 years after Pasteur," *FEMS Microbiology Letters*, vol. 265, no. 1, pp. 1–10, 2006.

[13] S. Katooch, M. Sharma, R. D. Patil, S. Kumar, and S. Verma, "In vitro and in vivo pathogenicity studies of Pasteurella multocida strains harbouring different ompA," *Veterinary Research Communications*, vol. 38, no. 3, pp. 183–191, 2014.

[14] G. B. Michael, K. Kadlec, M. T. Sweeney et al., "ICEPmnl, an integrative conjugative element (ICE) of Pasteurella multocida: structure and transfer," *Journal of Antimicrobial Chemotherapy*, vol. 67, no. 1, pp. 91–100, 2012.

[15] W. Liu, M. Yang, Z. Xu et al., "Complete genome sequence of Pasteurella multocida HN06, a toxigenic strain of serogroup D," *Journal of Bacteriology*, vol. 194, no. 12, pp. 3292–3293, 2012.

[16] K. M. Townsend, J. D. Boyce, J. Y. Chung, A. J. Frost, and B. Adler, "Genetic organization of Pasteurella multocida cap loci and development of a multiplex capsular PCR typing system," *Journal of Clinical Microbiology*, vol. 39, no. 3, pp. 924–929, 2001.

[17] B. F. Yang, N. Z. Li, L. X. Zou et al., "Identification and partial biological characteristics of six serotype A Pasteurella multocida from beef cattle," *Chinese Journal of Preventive Veterinary Medicine*, vol. 36, no. 6, pp. 487–489, 2014.

[18] R. Li, H. Zhu, J. Ruan et al., "De novo assembly of human genomes with massively parallel short read sequencing," *Genome Research*, vol. 20, no. 2, pp. 265–272, 2010.

[19] K. Lagesen, P. Hallin, E. A. Rodland, H.-H. Stærfeldt, T. Rognes, and D. W. Usery, "rRNAmmer: consistent and rapid annotation of ribosomal RNA genes," *Nucleic Acids Research*, vol. 35, no. 9, pp. 3100–3108, 2007.

[20] T. M. Lowe and S. R. Eddy, "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence," *Nucleic Acids Research*, vol. 25, no. 5, pp. 955–964, 1997.

[21] C. Frech, C. Choo, and N. Chen, "FeatureStack: Perl module for comparative visualization of gene features," *Bioinformatics*, vol. 28, no. 23, pp. 3137–3138, 2012.

[22] S. Kurtz, A. Phillippy, A. L. Delcher et al., "Versatile and open software for comparing large genomes," *Genome Biology*, vol. 5, no. 2, p. R12, 2004.

[23] D. A. Rasko, G. S. A. Myers, and J. Ravel, "Visualization of comparative genomic analyses by BLAST score ratio," *BMC Bioinformatics*, vol. 6, article 2, 2005.

[24] D. A. Rasko, J. Ravel, O. A. Okstad et al., "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pX01," *Nucleic Acids Research*, vol. 32, no. 3, pp. 977–988, 2004.

[25] S. Akhter, R. K. Aziz, and R. A. Edwards, "PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies," *Nucleic Acids Research*, vol. 40, no. 16, article e126, 2012.

[26] M. G. L. Langille, W. W. L. Hsiao, and F. S. L. Brinkman, "Evaluation of genomic island predictors using a comparative genomics approach," *BMC Bioinformatics*, vol. 9, article 329, 2008.

[27] P. Siguer, J. P. Brancatelli, E. Neve, and M. Chandler, "ISfinder: the reference centre for bacterial insertion sequences," *Nucleic Acids Research*, vol. 34, pp. D32–D36, 2006.

[28] Y. Chen, P. Yu, J. Luo, and Y. Jiang, "Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT," *Mammalian Genome*, vol. 14, no. 12, pp. 859–865, 2003.

[29] J. D. Bentdtsen, H. Nielsen, G. Von Heijne, and S. Brunak, "Improved prediction of signal peptides: signalP 3.0," *Journal of Molecular Biology*, vol. 340, no. 4, pp. 783–795, 2004.

[30] N. Fankhauser and P. Mäser, "Identification of GPI anchor attachment signals by a Kohonen self-organizing map," *Bioinformatics*, vol. 21, no. 9, pp. 1846–1852, 2005.

[31] M. C. C. de Lencastre, N. Fankhauser, and P. Mäser, "A new method for predicting GPI anchor attachment signals using a Kohonen self-organizing map," *Bioinformatics*, vol. 2, no. 1, pp. 15–20, 2005.

[32] Y. Chen, P. Yu, J. Luo, and Y. Jiang, "Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT," *Mammalian Genome*, vol. 14, no. 12, pp. 859–865, 2003.

[33] B. Xiao, Y. F. Sun, B. Lian, and T. M. Chen, "Complete genome sequence and comparative genome analysis of the Pasteobacillus massiliensis N01, Microbial Pathogenesis*, vol. 93, pp. 194–203, 2016.

[34] J. W. Choi, S. S. Yim, M. J. Kim, and K. J. Jeong, "Enhanced production of recombinant proteins with Corynebacterium glutamicum by deletion of insertion sequences (IS elements)," *Microbial Cell Factories*, vol. 14, no. 1, article 207, 2015.

[35] T. Ooka, Y. Ogura, M. Asadulhghani 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.

[36] C. R. Beuzón, D. Chessa, and J. Casadesús, "IS200: an old and still bacterial transposon," *International Microbiology*, vol. 7, no. 1, pp. 3–12, 2004.

[37] T. Sekizuka, M. Kawai, and K. Nakasone et al., "Complete genome sequence and comparative genomic analysis of Mycobacterium massiliense JCM 15300 in the Mycobacterium abcessus group reveals a conserved genomic island MmG1-1 related to putative lipid metabolism," *PLoS ONE*, vol. 9, no. 12, Article ID e114848, 2014.
[37] K. Hollenstein, D. C. Frei, and K. P. Locher, “Structure of an
ABC transporter in complex with its binding protein,” *Nature*,
vol. 446, no. 7132, pp. 213–216, 2007.

[38] A. Moussatova, C. Kandt, M. L. O’Mara, and D. P. Tieleman,
“ATP-binding cassette transporters in *Escherichia coli*,” *Biochim-
ica et Biophysica Acta (BBA)—Biomembranes*, vol. 1778, no. 9,
pp. 1757–1771, 2008.

[39] A. M. Grunden and K. T. Shanmugam, “Molybdate transport
and regulation in bacteria,” *Archives of Microbiology*, vol. 168,
no. 5, pp. 345–354, 1997.

[40] Y. Ma, J. W. Guo, R. Bade, Z. H. Men, and A. Hasi, “Genome-
wide identification and phylogenetic analysis of the SBP-box
gene family in melons,” *Genetics and Molecular Research*, vol.
13, no. 4, pp. 8794–8806, 2014.